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

Ruoying Gong
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.Sc. (Chemistry)
GRADE / DEGREE

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

Synthesis Towards O-Linked AFGO Analogue with Lactose as Its Carbohydrate Component
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

Tony Durst

André Beauchemin

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

Synthesis towards *O*-Linked AFGP Analogues with Lactose as Its Carbohydrate Component

By

Ruoying Gong

Thesis submitted to the

Faculty of Graduate and Postdoctoral Studies

University of Ottawa

In partial fulfillment of the requirements for the M.Sc degree

In the Ottawa-Carleton Chemistry Institute

Your file Votre référence
ISBN: 978-0-494-61177-7
Our file Notre référence
ISBN: 978-0-494-61177-7

Abstract

Antifreeze glycoproteins (AFGPs) are biological antifreezes which are found in the blood serum of arctic fish to help them living in the subzero temperature environments. AFGPs have a common structure composed with a carbohydrate and a peptide moiety by α *O*-linkage and can inhibit ice recrystallization which is useful in medical and commercial applications. The α *O*-linkage is very sensitive to acidic, basic and enzymatic conditions, therefore design and synthesizing AFGP analogues are more attractive. Thermal hysteresis (TH) and IRI activities of AFGPs have no relationship with one another and antifreeze analogues having good IRI and no TH will be good candidates for cryopreservation.

Recent reports suggested that hydration numbers and hydration indices of carbohydrates are important factors for its antifreeze activity. Antifreezes having saccharides with high hydration number and hydration index show good IRI activities which are in accordance with antifreeze activities of the saccharides.

My research goal was to synthesize **53** which is a truncated version of compound **52** (originally synthesized by Nishimura group) and test its IRI activity. Lactose has a high hydration number and hydration index in disaccharides and shows a good recrystallization inhibition. If compound **53** has good IRI activity, we can then conclude

that the IRI activities of AFGP analogues which have disaccharides as their carbohydrate moiety are in accordance with the recrystallization inhibition activities of their carbohydrate moiety themselves.

The compound **53** was synthesized through a convergent method. In its carbohydrate component synthesis, the glycosyl donor, benzyl ether protected lactose fluoride **55** was prepared starting from commercially available lactose. The synthesis of peptide component **56** was started from commercially available *N*-Fmoc-*O*-*t*-butyl-L-threonine **61**. Coupling reaction between **55** and **56** and then deprotecting the allyl ether to obtain the building blocks **54** and **65**. The glycopeptides **68** and **72** were obtained from standard SPPS. Following cleaved **68** and **72** from the solid support, the final products should be obtained from removing their benzyl groups. Unfortunately, the debenzylation could not be achieved, and the final polymer **69** and **73** were not obtained.

ACKNOWLEDGEMENTS

At the time I am finishing my graduate study, I would like to thank for many people who help me a lot to make my dream come true. First, I would like to thank my supervisor, Professor Robert Ben for his support, encouragement and advice during my two years study. His incredible ideas, broad knowledge and his enthusiasm for research deeply impress me and inspirit me working hard.

I would like to thank all of my colleagues for their help of my seminar. Every practice and every review, all of them give me a lot of suggestions, from concepts to contents, from English grammar to typo of slides, etc. Especially I want to thank Liz and Jenn who proofread my slides, thank Roger, Sandra and Wendy who gave me a lot of encouragement for my public presentation

I would highly like to thank Roger and Jenn for their help in my research. They help me with a lot of my research. Taline, thank you for proofreading my thesis and help me with my English grammar and the contents of it. I know it took a lot of spare time from you especially when you have a really tight schedule. I am also thankful to Dr. Glenn and Ms. Cheryl for NMR support.

I also like to thank Mathieu, Michael, John, Jackie, and Taz for your help in my whole graduate study. And I have to thank everybody in my lab for your kindly friendship in the baby shower and birthday party.

At the end, I would like to thank my family and my baby Ellison. Their love, support and encouragement are the motivation to let me complete my study.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF SCHEMES

LIST OF TABLES

ABBREVIATIONS

AFGP	antifreeze glycoprotein
AFP	antifreeze protein
Bn	Benzyl
Boc	*tert*-Butyloxycabonyl
Bz	Benzoyl
Cbz	benzyloxylcarbonyl
CD	Circular dichroism
d	doublet
DCC	Dicyclohexylcarbodiimide
DIPEA	Diisopropylethylamine
DIS	Dynamic ice-shaping
DMAP	4-Dimethylamino pyridine
DMF	*N,N*-Dimethyl formamide
DMSO	Dimethyl sulfoxide
DPPA	Diphenylphosphorylazide
Dt	Doublet of triplet
Fmoc	9-Fluorenylmethoxycarbonyl
GalNAc	*N*-Acetyl-galactosamine
HBTU	2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium
IRI	Ice recrystallization inhibiton
kDa	kiloDalton
M	multiplet

MALDI	Matrix-assisted laser desorption ionization
Me	Methyl
MS	Mass spetctrometry
NIS	*N*-Iodosuccinimide
PBS	Phosphate-buffered saline
Pd/C	Palladium on carbon
Pfp	Pentaflourophenyl
Ppm	Parts per million
PVP	Polyvinylpyrrolidone
SAR	Structure-activity relationship
t	triplet
TBDMS	*tert*-Butyldimethylsilyl
Tf	trifluoromethanesulfonyl
TFA	Trifluoroacetic acid
TH	Thermal hysteresis
TMSOTf	Trimethylsilyltrifluoromethanesulphonate

1 Biological Antifreezes

1.1 General Introduction

Antifreeze glycoprotein (AFGP) is a class of biological antifreezes found in organisms and can prevent damage from exposure to subzero temperatures. These compounds have special conformation of their carbohydrate and peptide moieties which is effective to inhibit the growth and recrystallization of ice crystal in the organism tissue. This characteristic property is highly desirable for medical and commercial applications.

1.2 Biological Antifreezes

The first natural antifreeze compound, antifreeze glycoprotein (AFGP), was isolated from Antarctic marine fishes by Devries and Wohlschlag in 1969.[1] From that time, many different antifreeze proteins were discovered in different species of organisms, such as polar fishes, amphibians, plants, bacteria[2] and insects.[3] The most studied biological antifreeze proteins are those isolated from Teleost fish which lives in the Antarctic and Arctic Oceans.

Normally, organisms cannot live in subzero temperatures as they sustain extensive cell damage due to ice crystals formed *in vivo*. However, some polar fish can still be alive where the water temperature is around -1.9°C in the polar oceans.[4] Those fish can lower

their blood freezing point to -2.0 to -2.1°C.[5] Sholander and co-workers first discovered the special ability of blood serum from bullhead fish (*Pagahenia borchgrevinki*) and Antarctic cold icefish (*Dissostichus mawsoni*) in 1953. [6, 4]

A depression of the freezing point of a solution can be achieved by colligative property of a solution which depends only on the number of molecules in a given volume of solvent but not on the mass of the substance. However, the ability of Antarctic fish to lower their blood serum freezing point is not only due to the colligative substances (such as sodium chloride, urea, and amino acids) in their blood serum since the colligative substances are in very low concentrations in their blood serum and can only contribute 0.86°C freezing point decreasing. There is still a 1°C difference between the freezing point of the seawater and their blood which is contributed by the antifreeze glycoproteins (AFGPs).[1, 7] The AFGPs can interact with ice crystals and inhibit their growth in blood serum by changing the shape of growing ice crystals.

Figure 1.1 shows the comparison of freezing points of those different solutions, such as AFGP plus 0.05 M NaCl, galactose, AFGP, NaCl, and a normal protein, chicken lysozyme.[8] It is shown that a low concentration of AFGP is slightly better than NaCl in decreasing freezing point of water. As the AFGP concentration is increased, it reaches a threshold. However, in comparison with colligative substances, AFGP is 200 to 500 times more effective than NaCl in lowering the freezing point of intramolecular fluid of cells on a molar basis.[2] In contrast, the galactose and lysozyme are not very effective in decreasing the freezing point. However, the addition of NaCl to an AFGP solution shows

a large increase in the ability to depress the freezing point. This implies the colligative substances can enhance the effect of AFGP.

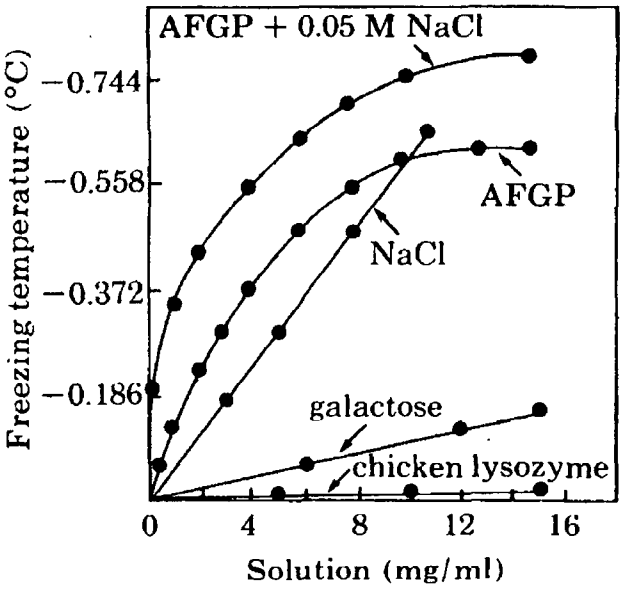


Figure 1.1 Comparing the different freezing temperatures of AFGP plus sodium chloride with AFGP, sodium chloride, galactose and chicken lysozyme[8]

1.3 Classification of Biological Antifreezes

Biological antifreezes exist in different species of organisms such as polar fishes, amphibians, plants, bacteria and insects.[3] Even though different antifreezes are found in different locations *in vitro* and have different structures, all of them have similar

properties and can be classified into two main groups; antifreeze proteins (AFPs) and antifreeze glycoproteins (AFGPs) based on their structural difference.[9]

1.3.1 Antifreeze Proteins (AFPs)

There are four different classes of antifreeze proteins based on their structural identities. Each type of the AFPs is found in different species and different organism tissues.[10]

1.3.1.1 Type I AFP

Type I AFP was first isolated by Duman, DeVries[11], Hew and Yip[12] from the winter flounder (*Pseudopleuronectus americanus*). It is about 3.3 – 4.5 kDa in size and contains about 60% of alanine which forms α – helical structure (Figure 1.2).

Figure 1.2 Type I AFP shows α - helix structure due to alanine-rich in composition [2]

1.3.1.2 Type II AFP

Type II AFPs are typically found in sea raven (*Hemitripterus americanus*), smelt (*Osmerus mordax*) and herring (*Clupea harengus harengus*) [2] which are cysteine-rich and only have about 14% of alanine. The disulfide bond is very typical in type II AFPs due to their cysteine-rich structures. These AFPs have the globular structure which is a mixture of α – helix, β – sheet and random coil. They are typically 11 – 24 kDa in size due to the increase of cysteine component (Figure 1.3). [13,14]

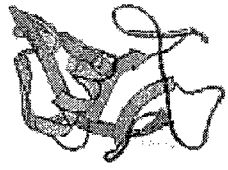


Figure 1.3 Type II AFP shows a mixed α-helix, with α – helix, β – sheet and random coil structure [2]

1.3.1.3 Type III AFP

Type III AFP is found in the ocean pout (*Macrozoarces americanus*). It contains at least 8 components which fall into two distinct groups based on their different behaviour.

It is around 6.5 kDa in size and has a flat-faced globular structure which contains 8 β strands (Figure 1.4).[2]

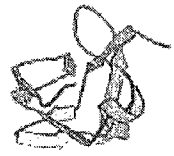


Figure 1.4 Type III AFP shows flat-faced globular structure [2]

1.3.1.4 Type IV AFP

Type IV AFP was isolated from the longhorn sculpin *(Myoxocephalus octodecimspinosis)*[15]. It is a highly α – helical structure with a heavy molecular weight about 12 kDa (Figure 1.5).

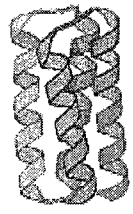


Figure 1.5 Type IV AFP shows α-helical bundle structure [2]

1.3.2 Antifreeze Glycoproteins (AFGPs)

The AFGPs are the other group of biological antifreezes which are mainly found in the blood serum of Antarctic notothenioid and Arctic cod.[16] Comparing with AFPs, AFGPs show very little structural variation among their members. The first AFGP structure was reported by DeVries in 1960.[1] The general structure of AFGPs is composed of an L-alanine-L-alanine-L-theronine tripeptide repeating unit in which the β-hydroxyl of each threonine is attached to a β-D-galactopyranosyl-(1,3)-2-acetamido-2-deoxy-α-D-galactopyranose disaccharide unit (Figure 1.6). The tripeptide repeating unit can be repeated from 4 – 55 times in different subtypes of AFGPs.[17]

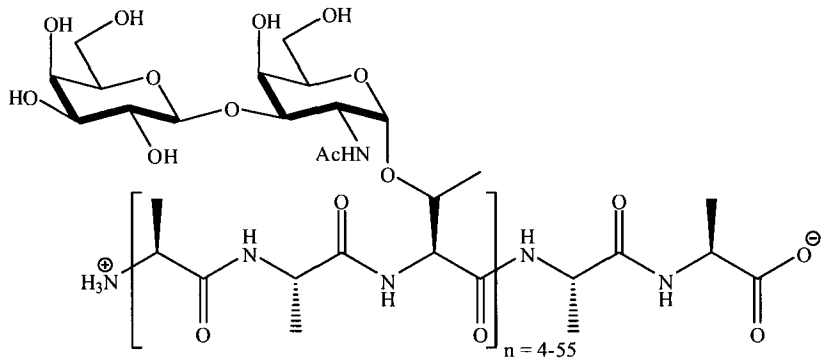


Figure 1.6 General Structure of AFGPs with repeat units n= 4 – 55 [17]

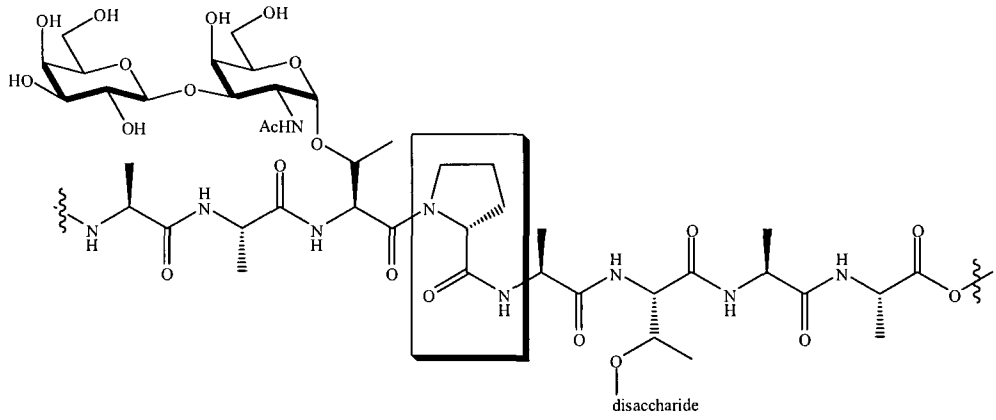
Based on the molecular mass difference of AFGPs due to variable number of the tripeptide repeating units, AFGPs can be divided into eight subtypes, AFGP 1 - 8. Their molecular weights vary from 33.7 kDa to 2.6 kDa, in which, AFGP 1 has the greatest molecular weight (33.7 kDa) and AFGP 8 has the smallest molecular weight (2.6 kDa) as shown in Table 1.1.[7]

Table 1.1 Classification of AFGPs based on different molecular mass

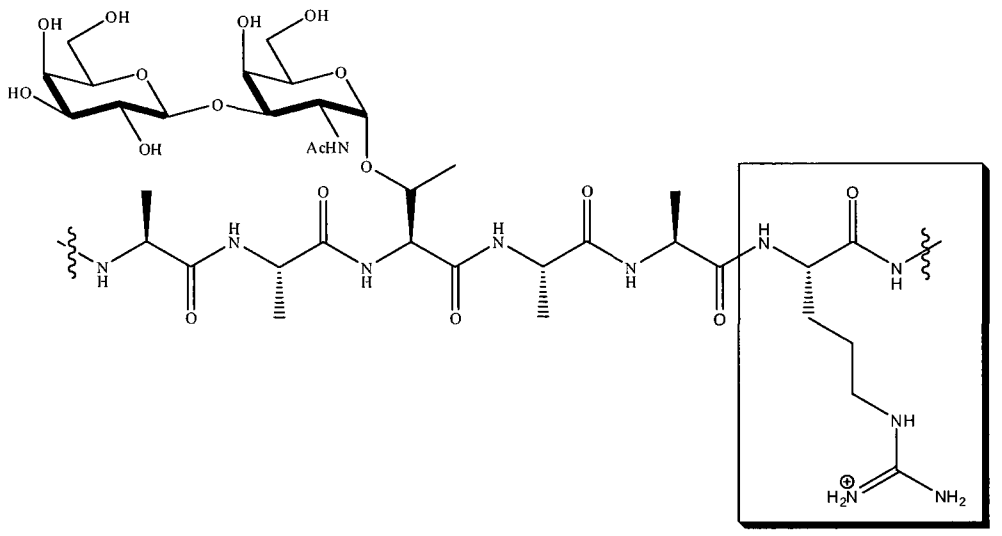
AFGP Subtype	Molecular Weight
AFGP 1,2	33.7 – 28.8 kDa
AFGP 3,4,5	21.5 – 10.5 kDa
AFGP 6,7	5 – 3.8 kDa
AFGP 8	2.6 kDa

It is found that the heavier AFGPs (i.e. AFGP 1-5) are more effective in decreasing the freezing point.[18] However, among the different subtypes of AFGPs, the AFGP 8 is the most studied since it contains only 4 tripeptide repeating units and it is easier to be synthesized than the other AFGP subtypes.

Both the molecular weights and structures are variable among AFGP subtypes by making changes in the tripeptide unit. It has been reported that the first L-alanine can be substituted by L-proline in some lower molecular weight AFGP subtypes such as AFGP 6 – 8 (Figure 1.7a); and also, the L-threonine can be replaced by L-arginine in some polar fish AFGPs such as Arctic cod AFGP (Figure 1.7b).[2]



(a)



(b)

Figure 1.7 Structure variable from typical AFGPs; (a) L-alanine substituted to be L-proline; (b) L-threonine substituted to be L-arginine [2]

Some research suggested that AFGPs predominately adopt random coil conformation with α – helix and β – sheet structures in solution at low temperature. [19, 20] The study of vacuum ultraviolet circular dichroism (CD) proposed that the AFGPs structure should be three-fold left handed helix similar to poly--L-proline type II of helical conformation,[21 ,22] and these disaccharide moieties are on the same side of this three-fold left handed helical structure which construct a hydrophilic face. In addition, the methyl groups on the

alanine residue and acetylmethyl groups in the GalNAc residues clustered into a hydrophobic face. [23] This amphiphatic structure of AFGPs is very similar to the amphiphatic single α – helix structure formed by alanines in type I AFP.[24] However, the X-ray crystal structure of an AFGP has not been obtained yet.

1.4 Properties of Antifreezes

1.4.1 Thermal Hysteresis (TH)

Normally, the ice starts to form when the temperature of a solution is below its melting point which is also its freezing point. However, by adding antifreezes into a solution, its freezing point will drop to below its original freezing point but the melting point will remain unchanged. The temperature gap between the melting point and freezing point is called thermal hysteresis (TH) [8] which is illustrated in detail in Figure 1.8. In this figure both the melting point and freezing point of water are 0°C. Upon addition of colligative substances such as NaCl into it, both the melting point and freezing point are dropped below 0°C but they are still equal to each other. However, water containing the biological antifreeze only decreases its freezing point but keeps the melting point similar to pure water and the temperature difference is TH.

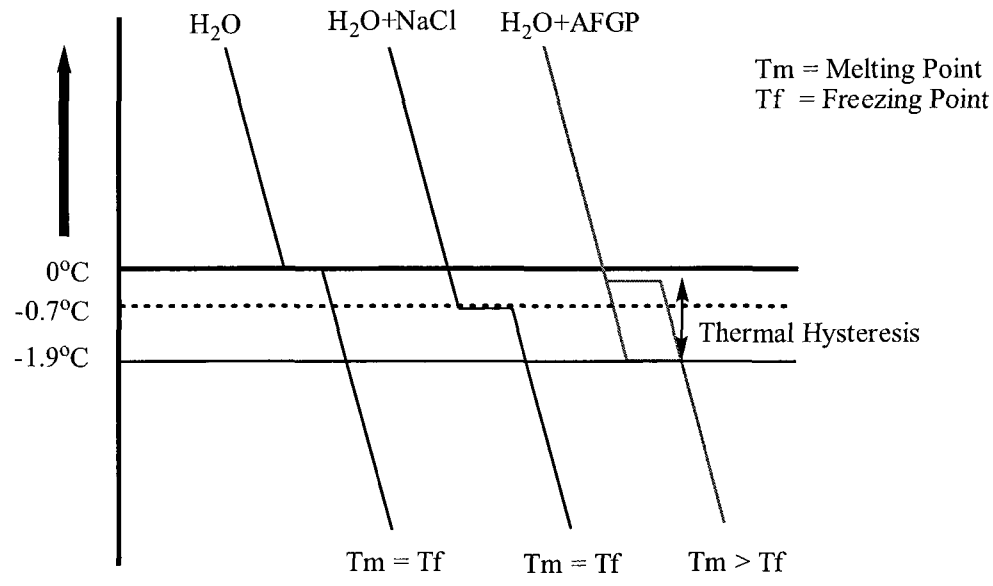


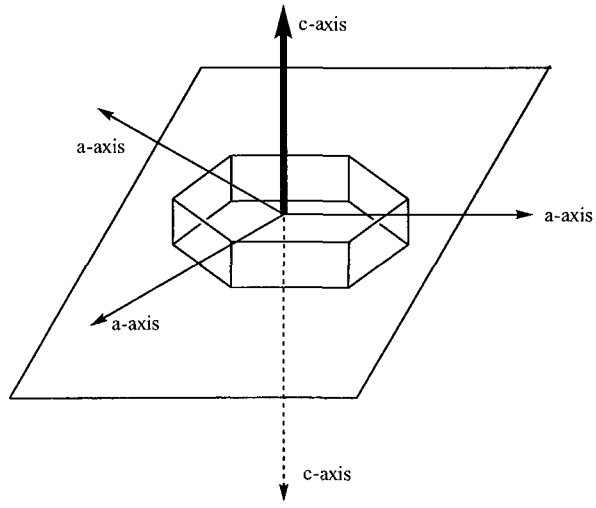
Figure 1.8 Illustration of thermal hysteresis

TH is a characteristic of biological antifreezes.[2] The TH gap varies among different types of biological antifreeze. For instance, the average TH of AFPs in the blood serum of Teleost fish is about 0.7 to 1.5 °C, terrestrial arthropod hem lymph AFP has a wide average TH gap which is approximately 3 to 6 °C, but the AFP-producing plants only have a very narrow TH gap about 0.2 to 0.6 °C.[25] Irrespective of the size of the TH gap, the biological antifreezes can help those organisms to survive in subzero environments by preventing ice growth when their ambient temperatures are within the TH gap.

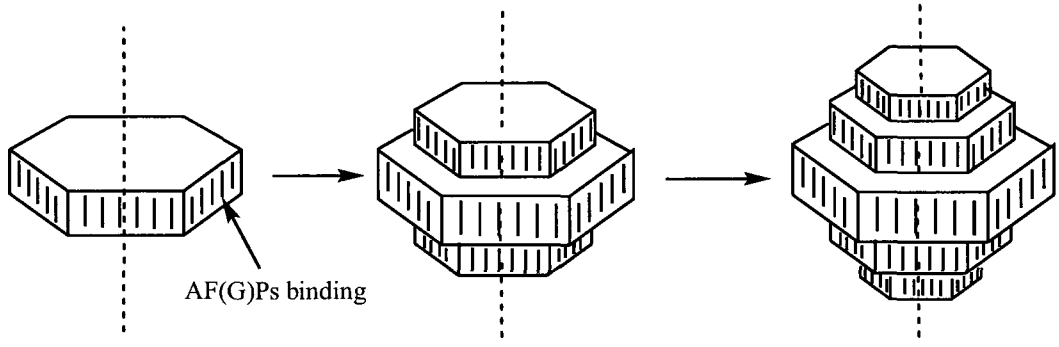
1.4.2 Dynamic Ice Shaping (DIS)

The TH activity of AFGPs is always accompanied with a modulation of ice crystal morphology to form different shapes. Normally, when small ice crystals are held for a prolonged period of time at a temperature lower than the melting point, the ice crystals will grow in size and change shape. However, when antifreezes were added into the solution, the ice growth process is inhibited on specific faces and resulted in different shapes because of the binding of antifreezes onto those specific faces.[26]

The whole process of DIS is shown in Figure 1.9. A typical ice crystal is a hexagonal structure below 0°C (Figure 1.9a). In pure water, ice will grow along the a-axes in the shape of elongated plates. [27] However, adding the biological antifreeze blocks the ice crystal growth along the a-axes is blocked, there is only limited ice crystal growth along c-axis which is perpendicular to the a-axes, finally it forms a hexagonal bipyrimidal structure (Figure 1.9b).[28]



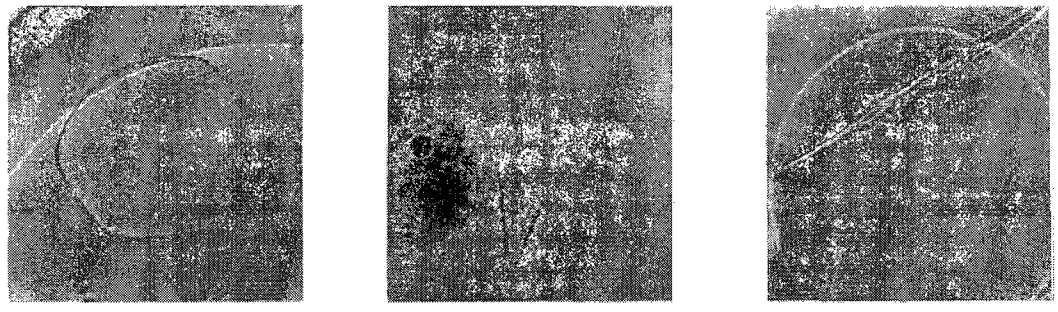
(a)



(b)

Figure 1.9 Qualitative illustration of dynamic ice shaping (a) Normal ice crystal grows along the a-axes to form hexagonal plate; (b) Binding with biological antifreezes, the ice crystal grows along the c-axis and forms hexagonal bipyrimidal structure. [27, 28]

Figure 1.10 shows the different morphologies of ice crystal structures. The hexagonal plate (Figure 1.10a) grew without adding biological antifreezes; the hexagonal bipyrimidal ice crystal (Figure 1.10b) was formed due to the basal planes (a-axes) of the hexagonal ice crystal binding with biological antifreezes which inhibited the ice crystal expanded along the basal planes within the TH gap; when the temperature was below the TH gap, the ice crystal grew rapidly along c-axis and formed needle-shape crystal (Figure 1.10c) which would damage the cell membranes.[29]



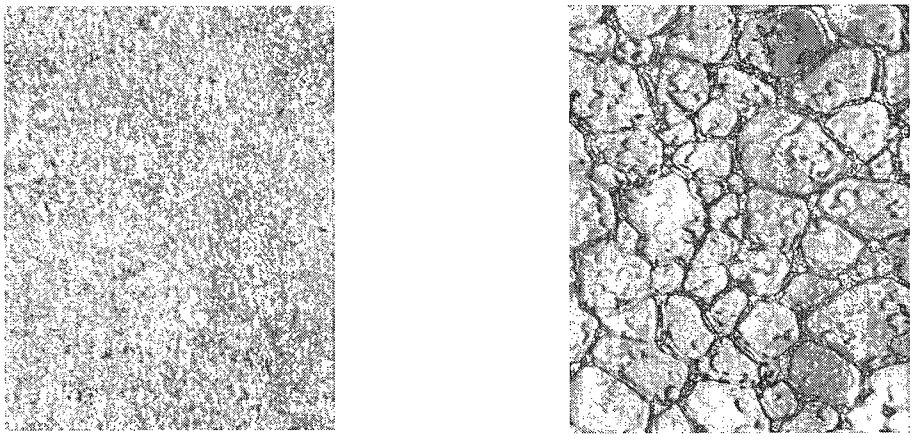
(a) (b) (c)

Figure 1.10 Ice crystal growth patterns in the absence and presence of AFPs (a) Regular hexagonal ice crystal grown in the absence of AFPs; (b) Hexagonal bipyrimidal ice crystal grown in the presence of AFPs within TH temperature gap; (c) Needle-shaped ice crystal grown in AFP solution below TH temperature gap.[29]

1.4.3 Ice Recrystallization Inhibition (IRI)

Biological antifreezes have another characteristic which is the ability to inhibit ice recrystallization. When the temperature is within the TH temperature gap, small ice crystals with hexagonal bipyrimidal shape formed in the organism blood and body fluid due to the DIS of biological antifreeze. After the small ice crystals are formed, the biological antifreezes can prevent them grow larger to form large ice crystals which finally cause tissue damages. This IRI ability can help those organisms such as Teleost fish survive in subzero temperature. [30]

Smaller ice grains tend to be absorbed by larger ice grains resulting in larger ice crystals when the solution containing small ice crystals is held at temperatures slightly below the freezing point. This is a spontaneous process which is driven by the overall reduction of ice crystal boundary energy. Since ice grain boundaries contain many broken and strained bonds which make the free energy of the whole system very high, migration of these boundaries will lower the free energy of the whole system. [31, 32] It has been suggested that antifreeze molecules can bind the ice crystals at the ice-water interface and prevent ice crystal boundary migration and eventually inhibit ice recrystallization (Figure 1.11).[33, 34] The minimum concentration of biological antifreeze with effective IRI activity can be as low as 10^{-12} M.[35]



(a) (b)

Figure 1.11 Comparing sizes of ice crystals in a solution of (a) AFGP8 (b) PBS as a control at -6°C (Pictures are at same magnification) [33, 34]

It has been reported that TH and IRI activities of biological antifreezes have no relationship between each other. [36] IRI can be effective at very low concentration of biological antifreezes which is 100 to 500 times less than the concentration of antifreezes required for TH activity.[37]

1.5 Mechanism of Antifreeze Activity

The most acceptable hypothesis of antifreezes was given by Raymond and DeVries in 1977 who proposed that the mechanism of antifreezes at the macroscopic level is an adsorption-inhibition process.[33, 3] Biological antifreeze molecules bind onto the surfaces

of the growing ice and additional water molecules cannot attach to those ice surfaces. Therefore, ice crystals only can grow between those antifreeze molecules attached places which result in very large curved ice surfaces (Figure 1.12). However, adding more water molecules to these convex ice surfaces is not thermodynamically favourable; ice growth therefore stops unless the temperature is further depressed. This results in a localized (high curvature of ice growth front) freezing point depression, and it is also known as Kelvin effect. [38, 1]

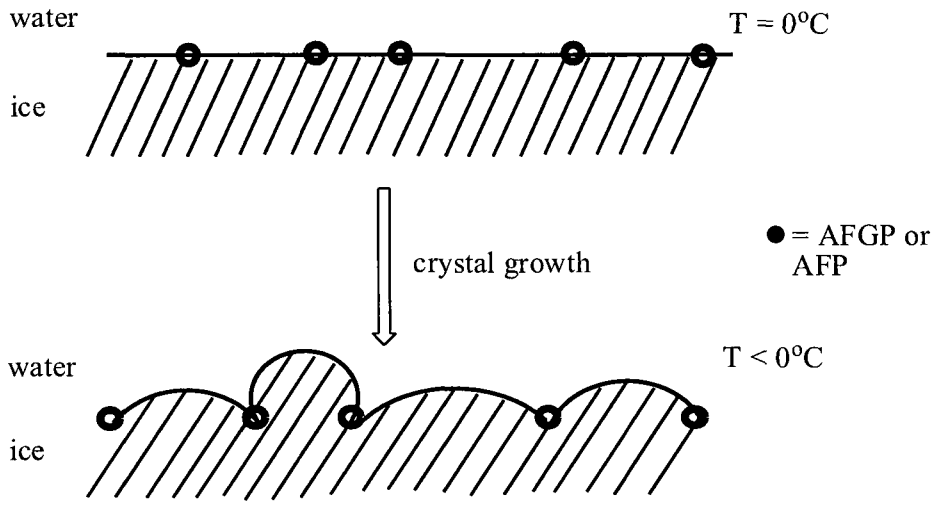


Figure 1.12 Adsorption-inhibition mechanism of biological antifreezes [33]

It was assumed the biological antifreezes were adsorbed onto the surface of ice via multiple numbers of hydroxyl groups in antifreezes and water molecules in the ice crystals, and the possibility that all hydrogen bonds are released simultaneously is very low. [39] Therefore, it was hypothesized by Knight, Devries [40] and Wilson [41] that the

binding between the ice crystal surfaces and the biological antifreezes is thought to be irreversible. However, there was another assumption about a reversible hydrogen bonding model between antifreezes and ice crystals depending on different antifreeze concentrations given by Burcham.[42] In which, the hydrogen bonding is reversible when antifreeze concentration is low due to large freedom and the hydrogen bonding is irreversible when antifreeze concentration is high due to tight packing.

All of above assumptions are based on that the hydrogen bonding is a hydrophilic effect between the antifreeze molecules and water molecules. However, whether the bonding is a hydrophilic effect or a hydrophobic effect is still being debated.

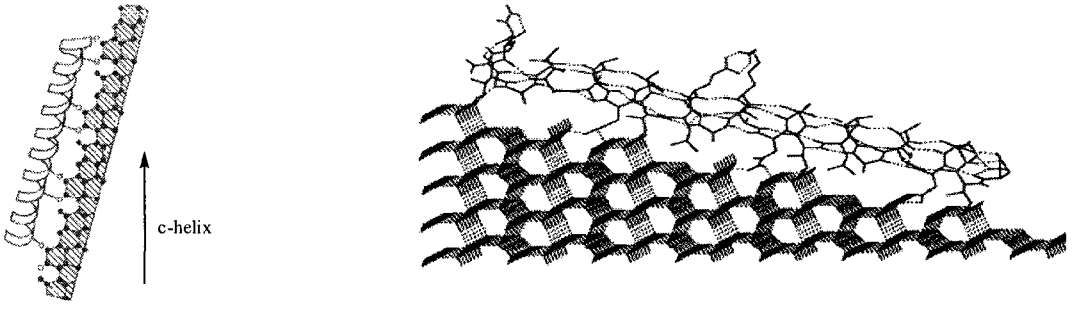
1.6 Structure - Activity Relationships

The assumptions about the adsorption-inhibition process of biological antifreeze activity are based on the interaction between the antifreeze and ice crystal is hydrogen bonding, a hydrophilic effect. However, there are some studies showing that the hydrophobic effects are still important for antifreeze activity.

1.6.1 Structure – Activity Relationship Study of AFP

In the biological antifreeze family, type I AFP has a near perfect single α-helix structure and studied very well. Some research showed that the number of hydrogen

bonding donor, the polar side chains of threonine residues of the antifreeze, and the hydrogen bonding acceptor, the water molecules of the ice crystal surface are matched very well.[43] The repeat spacing in the ice along the alignment direction is 16.7 Å which is very close to the distance between those polar residues of threonines, 16.5 Å, along the helix. The rotational freedom of the threonine side chains and the helix flexibility both readily compensate for the 0.2 Å distance difference.[40]



(a) (b)

Figure 1.13 Hydrogen bonding model of type I AFP. (a) Bonding view along c-helix. Filled circles represent water molecules in the ice lattice. (b) Line model of an AFP molecule bound to the prism face of ice, showing the hydrogen bonding pattern. [43]

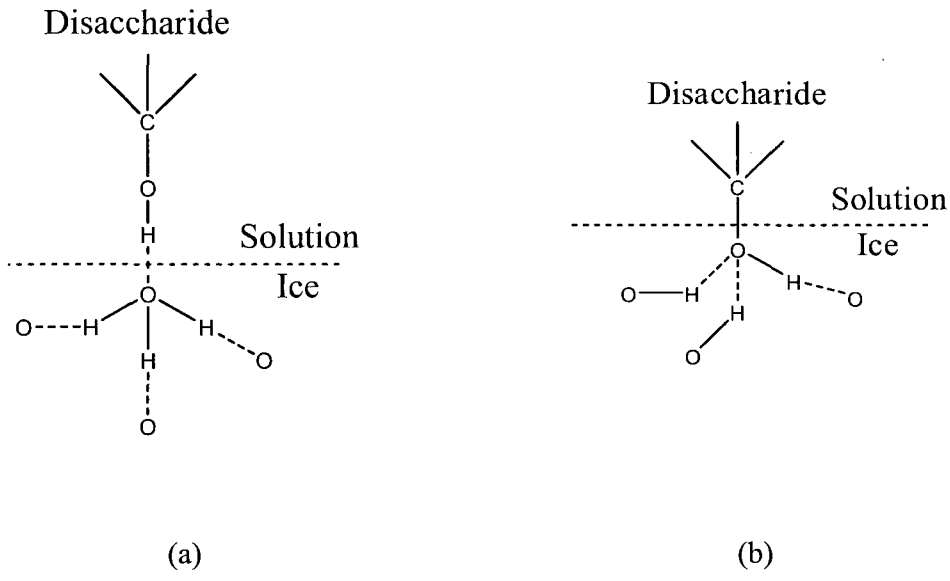
Figure 1.13 shows that the polar groups of the amino acid side chains of type I AFP within the hydrogen bonding range are embedded into the ice lattice and bind to the pyramidal plane of prism faces on the ice lattice.[43]

But there are some different assumptions, in which the researchers said that the mechanism of antifreeze activity is due to a hydrophobic effect. When threonine was mutated to be serine in type I AFP which also can form similar hydrogen bond with the water molecule of ice surface, the antifreeze activity was almost lost; but when mutated threonine to be valine in which one of the two methyl groups has similar spacing structure as threonine, the antifreeze activity still remained. This study showed that the β-methyl groups on those threonine residues are more important than the hydroxyl groups for the interaction between type I AFP and ice crystal surface and it is a hydrophobic effect between the β-methyl groups and the ice crystals. [44,45]

1.6.2 Structure – Activity Relationship Study of AFGP

The hydrogen bonding between AFGP and ice surface was also been studied. The periodicity of the polyproline helix of the AFGP 8 is 9.31 Å and the repeating space in the ice along the α-helix is 4.519 Å, [39] therefore, twice of the spacing distance of the water molecules in the ice surface are close to the periodicity of the AFGP 8. Therefore, it needs two hydroxyl groups in every disaccharide to form enough hydrogen bonds with water molecules. However, the number of the polar side chain of AFGP 8 theronine is not enough to tightly hold the ice crystals by bonding with water molecules on ice surface. If one hydroxyl group of AFGP 8 forms one hydrogen bond, the total will be eight hydrogen bonds formed by a single AFGP 8 molecule since there are four repeating

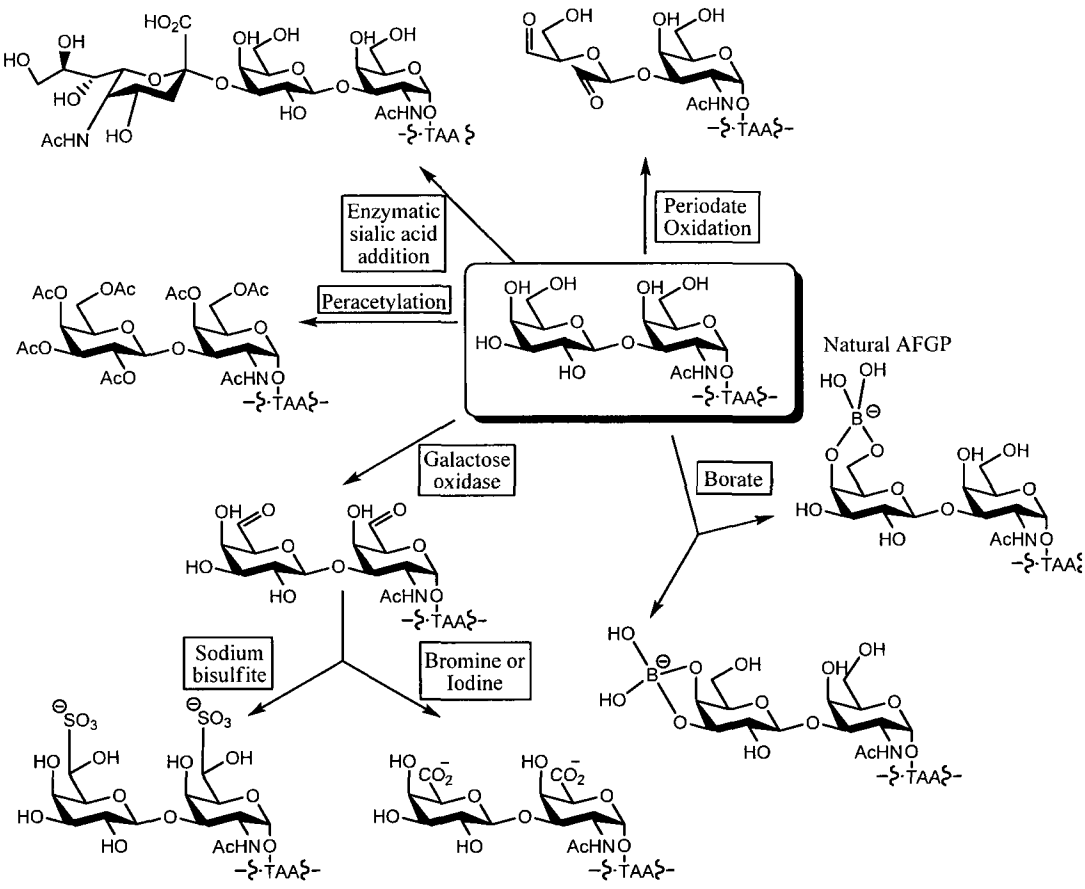
units (Figure 1.14a). This theory was revised to another model in which the hydroxyl groups of the disaccharide are incorporated into the ice lattice (Figure 1.14b). [39]



(a) (b)

Figure 1.14 Hydrogen bonding models between AFGP 8 and ice crystal surface (a) AFGP 8 is at outside of ice surface. (b) AFGP 8 sticks into the ice surface. [39]

Similar results were obtained on AFGP study by different groups. The Feeney [46] and Nishimura[23] group separately synthesized a group of AFGP analogues with mutated disaccharide moieties and tested for antifreeze activity, from which they both found that the hydroxyl groups with specific orientations are important for antifreeze activity (Scheme 1.1).

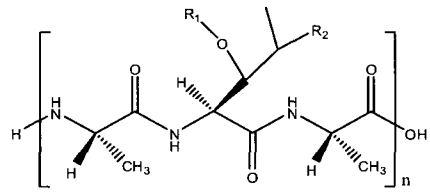


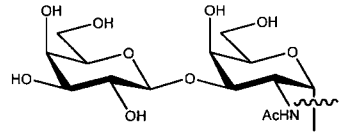
Scheme 1.1 Different AFGP analogues synthesized by modification of natural AFGP

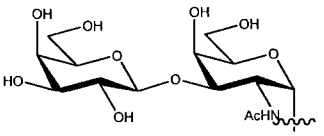
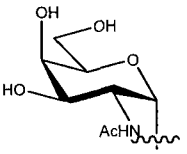
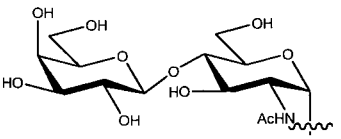
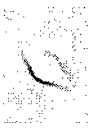
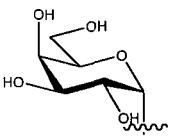
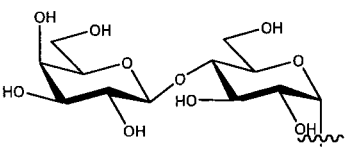
Except the importance of certain hydroxyl group of the disaccharide moiety for antifreeze activity, the hydrophobic effect reported is also important. The research done by the Nishimura group showed that the antifreeze activity is also decided by some groups other than the hydroxyls. In their research, when threonine was replaced by alanine to form alanine-serine-alanine (ASA) tripeptide, this antifreeze lost its TH

activity; secondly, replacing different saccharide moieties which also contain C2 NHAc group as the original dissacharide, only GalNAc has full antifreeze ability to form bipyramidal crystals and shows a little bit lower TH. This illustrates that the terminal galactose residue is not necessary for antifreeze activity but may enhance the antifreeze-ice binding. LacNAc also shows inhibition of ice growth and a lower TH which means carbohydrate hydroxyl group is important for antifreeze activity. Gal or Lac analogue can produce hexagonal-like crystals but does not show any TH, which means that C2 NHAc of the reducing end sugar residues is a key functional group required for the antifreeze activity.[23]

Table 1.2 Structure-activity relationship of AFGP analogues. [23]



	R_1	R_2	Ice Crystal Morphology
AFGP		CH_3	

ASA		H	
GalNAc		CH_3	
LacNAc		CH_3	
Gal		CH_3	
Lac		CH_3	

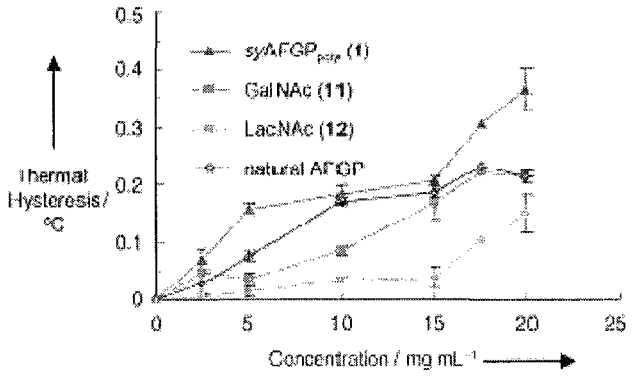


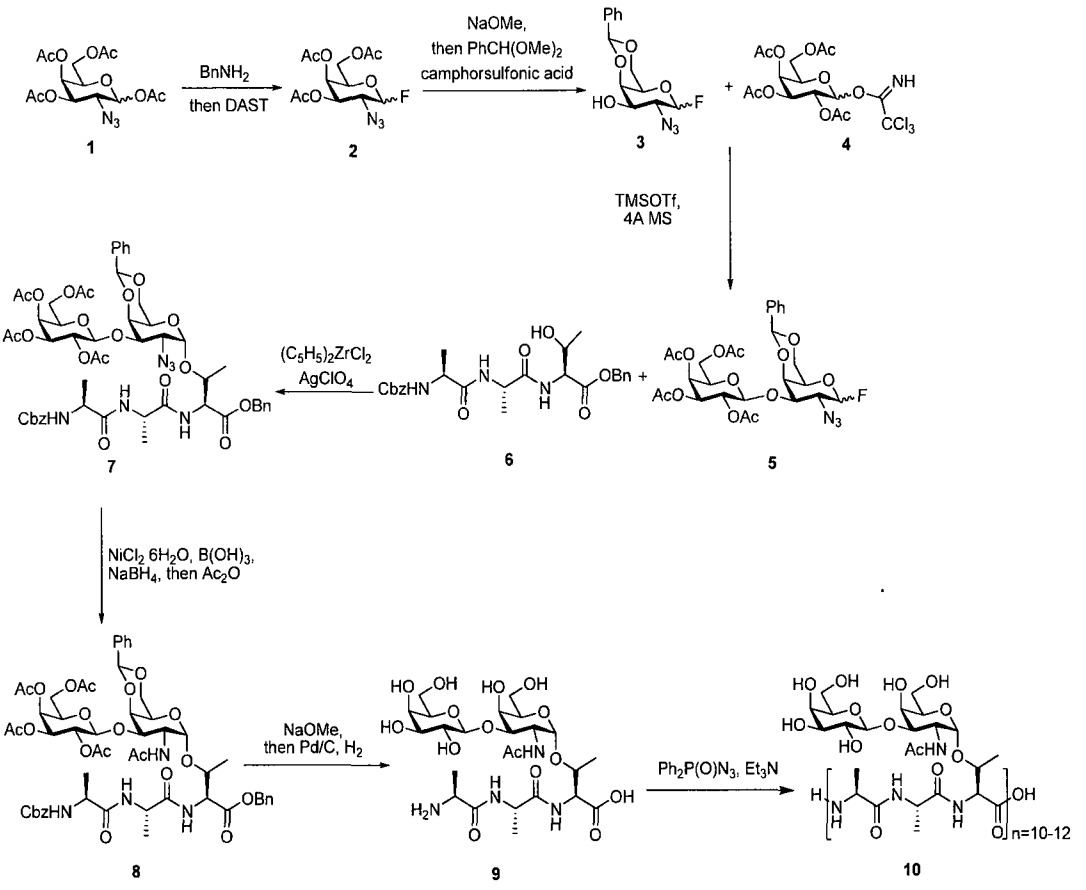
Figure 1.15 TH activity of AFGP analogues [23]

1.7 Synthesis of AFGPs and AFGP Analogues

Most of the AFGP researches were done by using the biological antifreeze isolated and purified from Teleost fish. There are very tiny quantities of pure AFGPs can be isolated from large amount of fishes which make it very difficult to obtain enough AFGPs for research and commercial applications. Synthesis of AFGPs is still a challenging work since the AFGPs containing carbohydrate and polypeptide moieties which are hard to be incorporated together. Some synthesis methods were developed by different groups.

1.7.1 Synthesis of Native AFGP

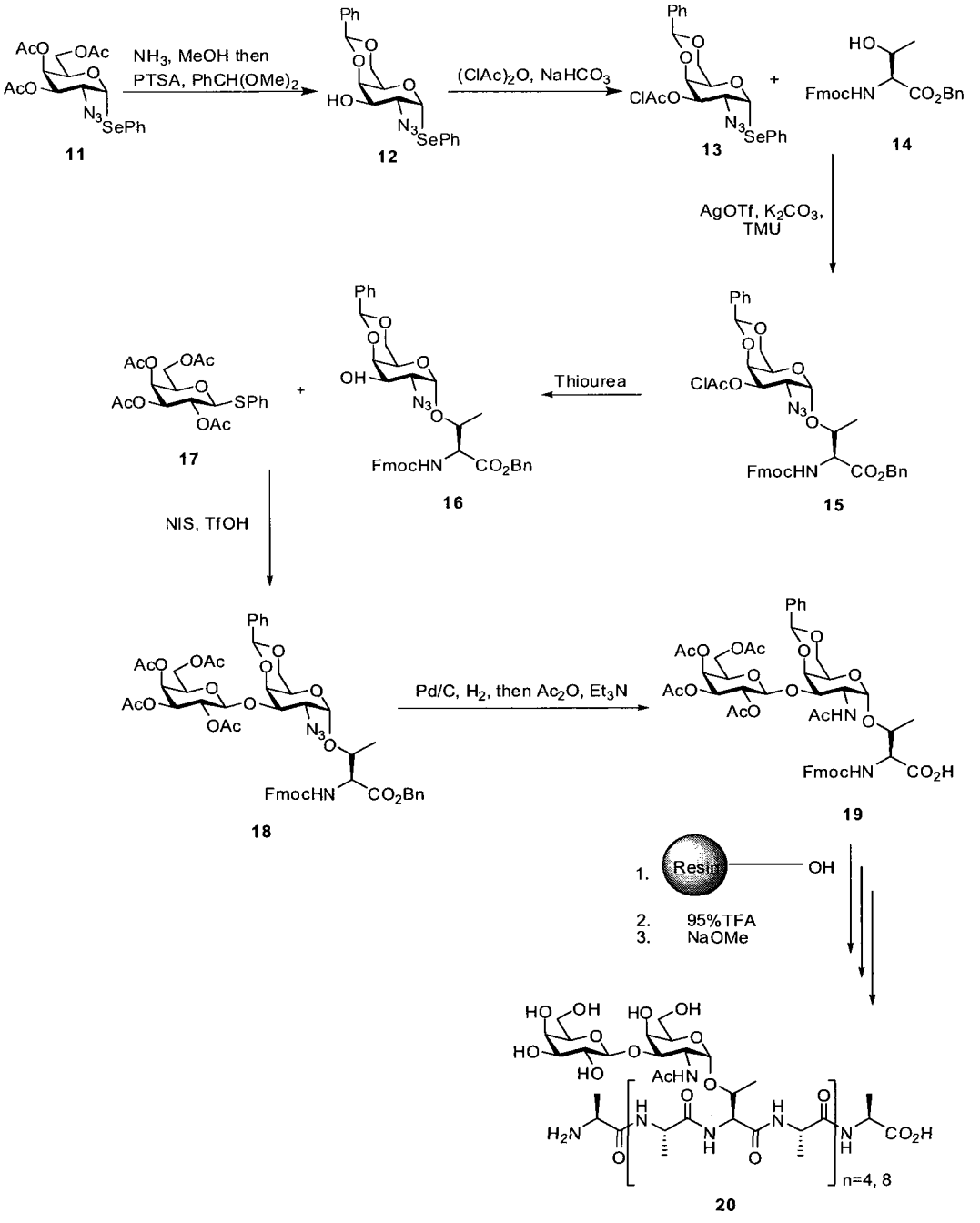
The first total synthesis of natural AFGPs was reported by the Nishimura group in 1996.[47] In this method (Scheme 1.2), a glycosyl acceptor, properly protected glycosyl fluoride **3** was modified from commercially available 1-azido-1-deoxy-β-D-galactopyranoside tetraacetate **1** in two steps. Following coupling with a glycosyl donor, properly protected galactosyl trichloroacetimidate **4** using trimethylsilyl trifluoromethane sulfonate (TMSOTf) as the glycosylation promoter, the dissacharide **5** was obtained in 81% yield. And then **5** was coupled with an alanine-alanine-threonine tripeptide **6** to obtain a single α-glycosidic linked product **7** using cyclopentadienyl zirconium dichloride and silver perchlorate as coupling promoters. The C2 azide group of **7** was converted into the acetamido group by a reduction reaction with nickel chloride in sodium borohydride followed with acetylation to obtain **8.** Following fully removing all protect groups by sodium methoxide and hydrogen gas with catalytic amount of palladium on charcoal led to **9**. Polymerization of **9** to afford mixture of polypeptides **10** as AFGP analogues with molecular weight about 6.7 kDa was accomplished with catalyst diphenylphosphoryl azide (DPPA) and triethyl amine.



Scheme 1.2 Nishimura's natural AFGPs synthesis

Another natural AFGP synthetic route was published by Chen and coworkers who employed the standard solid phase synthesis prototol (Scheme 1.3).[48] Coupling reaction between glycosyl selenide **13** and Fmoc (fluorenylmethyloxycarbonyl) group and benzyl group protected threonine **14** using silver triflate as a promoter yielded **15**. C3 chloroacetyl group was removed to obtain **16** using thiourea. Conjugating **16** with

thioglycoside **17** in the presence of *N*-iodosuccinimide (NIS) and TfOH yielded the building block, disaccharide glycosylated threonine **18** in 93% yield. Removing the benzyl group of **18** afforded a free carboxylic acid group in **19**. The polymer **20** was obtained by standard solid phase synthesis process using HMP resin.



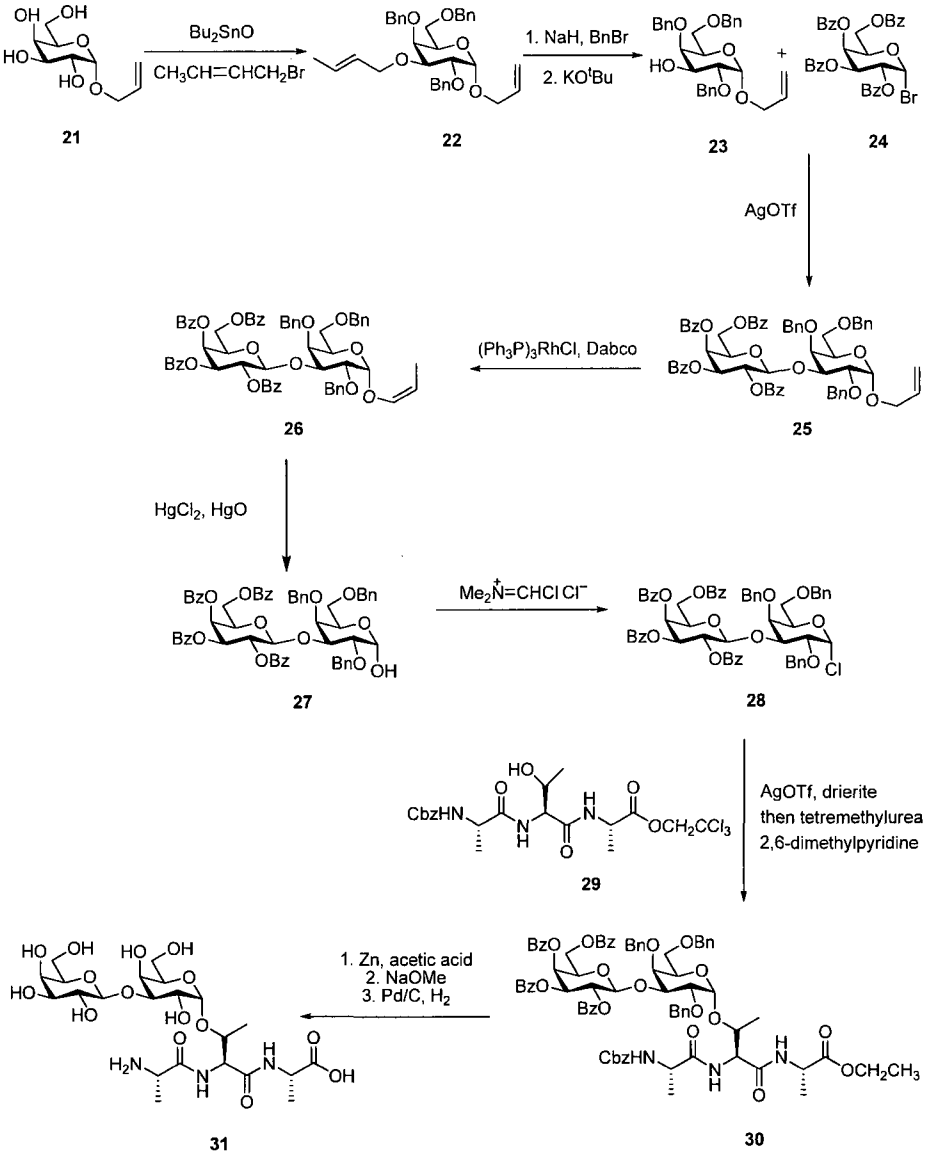
11 12 13 14 15 16 17 18 19 20

Scheme 1.3 Chen's natural AFGPs synthetic route

1.7.2 *O*-linked AFGP Analogue Synthesis

There were some papers describing the preparation of AFGP analogues which had either one or more structural variations to the native AFGPs and shared a common structural feature of an *O*-linkage between saccharide and peptide moieties.

The Anderson group reported an *O*-linked AFGP analogue synthetic route via a convergent method (Scheme 1.4).[49] The glycosyl donor, benzyl protected galactose bromide **24** coupled with the glycosyl acceptor, benzyl protected and allyl protected galactoside **23** to obtain disaccharide **25** using silver triflate as promoter. The anomeric allyl protecting group of **25** was converted to 1-propenyl group in **26** and then removed to hydroxyl to afford **27** using mercury chloride and mercury oxide as reagents, and then used Vilsmeier reagent to convert hydroxyl group to glycosyl α-chloride **28**. Directly coupled a carbobenzyloxy (Cbz) protected alanine-threonine-alanine tripeptide **29** with the disaccharide got the protected glycosylated tripeptide **30**. The final AFGP analogue **31** was obtained by removing all protecting groups from the glycosylated tripeptide **30**.

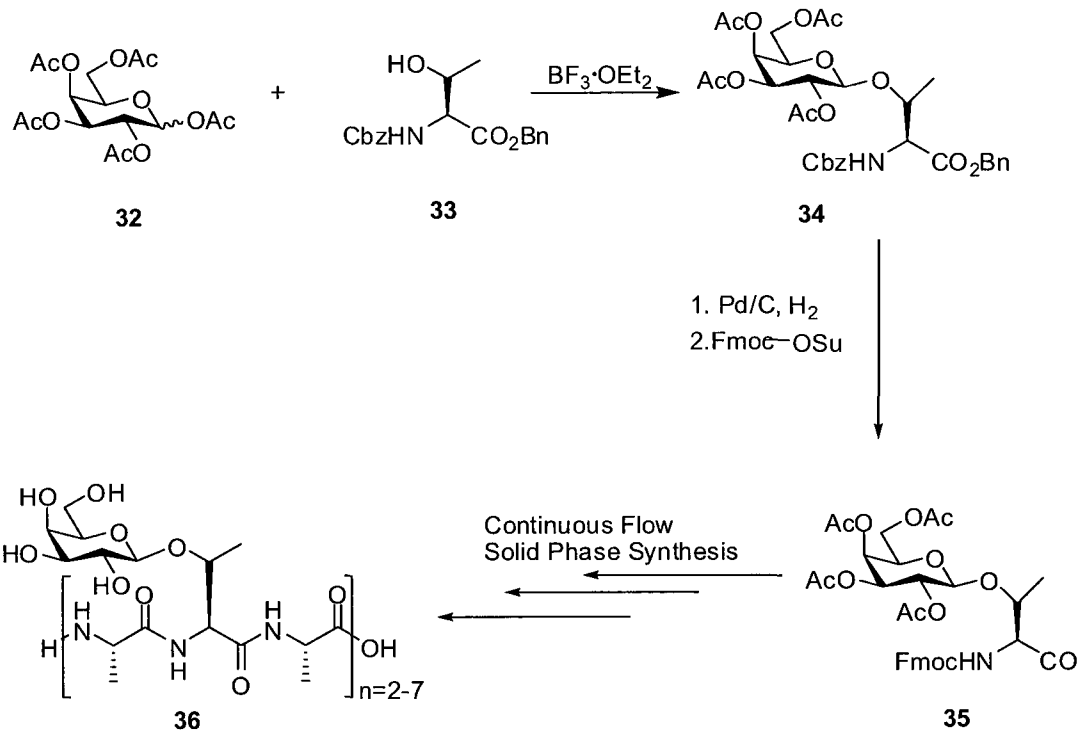


Scheme 1.4 Anderson' *O*-linkage AFGP analogue synthesis

The convergent method was not used frequently in *O*-linked AFGP analogue preparation because of the low solubility of the peptides for glycosylation [50] and the low

reactivity of the hydroxyl group of threonine residue in the tripeptide [51]. It was shown that glycosylation of single amino acids to form building blocks and then stepwise assembly of the glycopeptides is more reliable.

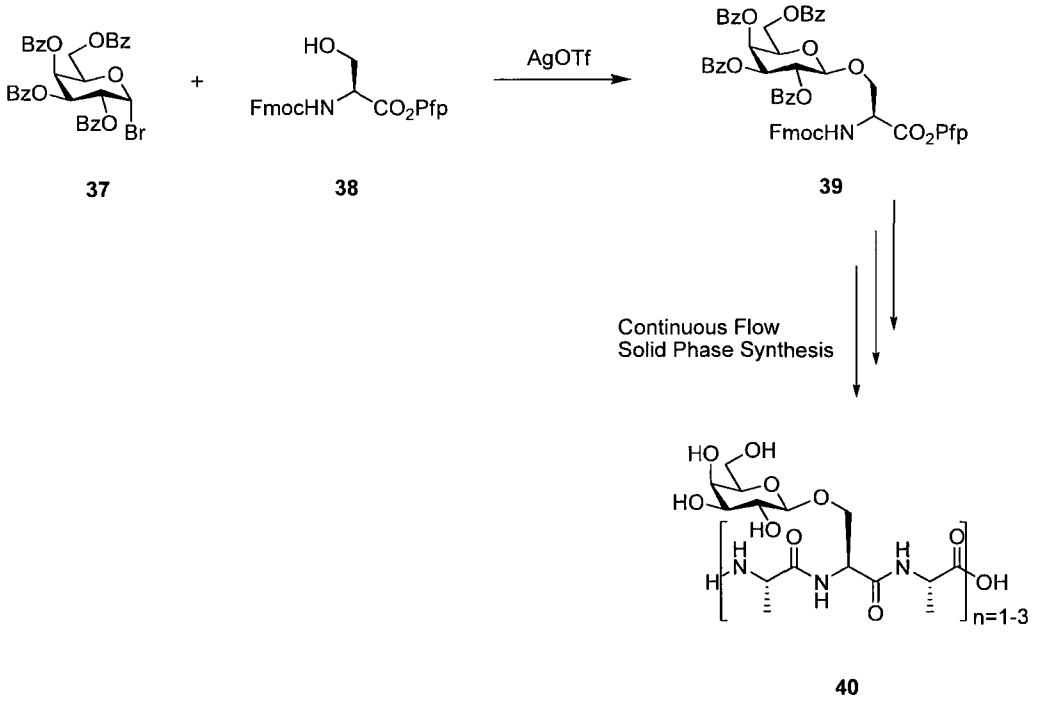
The Filira group reported the first AFGP analogue synthesis via solid phase peptide synthesis using Fmoc-chemistry.[52] In this synthesis (Scheme 1.5), α-*O*-disaccharide was replaced by a β *O*-galactose to mimic the terminal non-reducing galactose residue in native AFGPs. Galactose pentaacetate **32** coupled with hydroxyl group of protected threonine **33** generating glycosylated threonine **34**, and then modified the protecting group Cbz to be Fmoc group to obtain **35**. Following standard solid phase peptide synthesis generated polymer AFGPs analogue **36**.



Scheme 1.5 Filira's solid phase synthesis of AFGP analogues

Meldal and Jensen reported a similar synthetic approach with a pentafluorophenyl (Pfp) group protected serine **38** replacing threonine residue (Scheme 1.6).[50] In this synthesis, Pfp ester was used not only as a protecting group for the carboxylic acid of serine, but also an activating group for the peptide bond forming reaction in solid phase peptide synthesis. The building block **39** was obtained by the glycosylation reaction between **37** and **38** using silver triflate as promoter. Polypeptide **40** was prepared with 1

to 3 serine-alanine-alanine tripeptide repeating units with a continuous flow solid phase synthesis protocol.

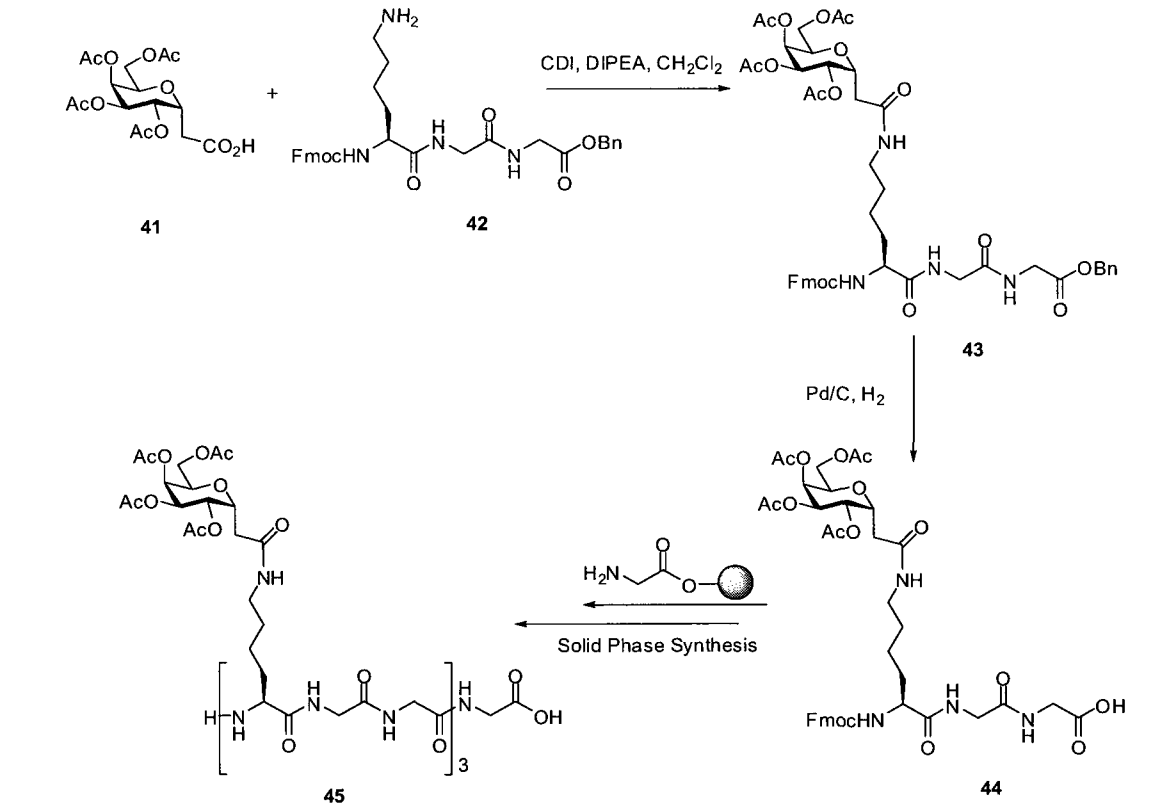


Scheme 1.6 Meldal's solid phase synthesis of AFGP analogue

1.7.3 *C*-linked AFGP Analogues Synthesis

The Ben group did research on AFGP analogue with α *C*-linkage instead of *O*-linkage. It showed that *C*-linkage is more stable than the *O*-linkage (Scheme 1.7).[53] In their approach, they used a series of monosaccharide to replace the disaccharide. Also, they used a lysine to replace threonine because the amide bond structure is similar to the

guanidinyl group in arginine which sometimes is used to replace threonine in native AFGPs. The monosaccharide **41** was coupled with lysine-glycine-glycine tripeptide **42** using CDI as activator to afford **43**. Removing the benzyl group lead a free carboxylic acid on building block **44**. The AFGP analogue **45** was prepared from **44** via solid phase peptide synthesis.



Scheme 1.7 Ben's AFGP analogue solid phase synthesis

1.8 Applications

The antifreeze properties of biological antifreezes make them have potential applications in different areas.

1.8.1 Biological Antifreezes in Medicine

Long term storage of human organ cells, tissues is always a challenging work. Due to very limited preservation techniques, the storage is restricted to only 12-20 hours depending on the organs. As temperature decreased, both extracellular and intracellular ice crystals can be formed [54] and cause cell membrane disruption though osmotic imbalance. In some cryopreservation protocols, addition of glycerol or DMSO can increase intracellular osmosis but they cannot block extra cellular ice formation.[55] Non-penetrating polymers such as polyvinylpyrrolidone (PVP) can be used in extracellular cryopreservation to prevent cell membrane disruption. [56] However, those cryopreservation techniques all have limits and have low recovery ability from frozen state to normal state.

Some research showed that antifreezes can be used as cryopreservation agents. Rubinsky and co-workers have reported that biological antifreezes can protect mammalian cells at hypothermic and cryogenic temperatures.[57] For example, treating pig oocytes with 40μg/mL of AFGPs, oocyte cells showed improvement of cell integrity; and liver cells treated with type III AFP can be stored at hypothermic temperatures.

Cryosurgery is a minimally invasive surgical technique which uses freezing to destroy undesired tissue by inserting a cooled probe into it and producing local freezing. Cryosurgery has advantages such as low discomfort, quicker recovery. Studies *in vivo* showed that the property of forming needle-shape ice crystals when temperature is dropped below the TH gap makes the biological antifreezes attractive in cryosurgery. [58] Rubinsky and co-workers have found that when concentrations are higher than 5 mg/mL, biological antifreezes can cause frozen cell destruction.

1.8.2 Biological Antifreezes in Food Storage

Ice crystals may grow in frozen foods due to recrystallization. When large ice crystals formed intracellular in food tissues, it damages cell membranes and causes increased drip during thawing which results in a lower quality frozen food. The ability of recrystallization inhibition is very important for frozen food texture, especially for foods such as ice cream which are eaten while frozen and need higher texture quality.[59] Payne and co-workers studied the effect of biological antifreezes on frozen meat and got some positive results.[60] They observed two groups of bovine muscle, one group was frozen without adding AFGP which was used as control; another group was frozen after adding 0.1 mg/mL AFGP. Examining those two samples using scanning electron microscope (SEM) under low magnification (50X) revealed macroscopic difference between those two groups; the treated group had much smaller intracellular space between the cells

comparing with the control. They concluded that low concentrated AFGP can help to get higher quality food during frozen process.

---

[1] DeVries, A. L., Komatsu, S. K., Feeney, R. E. *J. Biol. Chem.* **1970,** 245, 2901–2908.

[2] Harding, M. M., Anderberg, P. I., Haymet, A. D. *J. Eur. J. Biochem.* **2003,** 270, 1381-1392.

[3] Ben, R.N. *Chem. Biochem.* **2001,** 2, 161-166.

[4] Scholander, P.F., Dam, L.V., Kanwisher, J.W., Hammel, H.T., Gordon, M.S. *J. Cell. Comp. Physiol.* **1957,** 49, 5.

[5] Feeney, R.E., Burcham, T.S., Yeh, Y *Ann. Rev. Biophys, Biophys. Chem.* **1986.** 15, 59-78.

[6] Scholander, P.F., Flagg, W., Walter, V., Irving, L. *Physiol. Zool.* **1953,** 26, 67.

[7] DeVries, A.L., Vandenheede, J., Feeney, R.E. *J. Biol. Chem.* **1971,** 246, 305-308.

[8] Yeh, Y., Feeney, R. *Chem. Rev.* **1996,** 96, 601-618.

[9] (a) Y. Yeh, R. E. Feeney, *Chem. Rev.* **1996,** 96, 601- 617.

(b) P. L. Davies, B. D.Sykes, *Curr. Opin. Struct. Biol.* **1997,** 7, 828 – 834.

(c) K. V. Ewart, Q. Lin, C. L.Hew, *Cell. Mol. Life Sci.* **1999,** 55, 271 - 283.

[10] Cheng, C.H. *Curr. Opin. Genet. Dev.* **1998,** 8, 715-720.

[11] Duman, J.G. and DeVries, A.L. *Nature* **1974,** 247, 237-238.

[12] Hew, C.L., Yip, C. C. *Biochem. Biophys. Res. Commun.* **1976,** 71, 845-850.

[13] Sönnichsen, F.D., Sykes,B.D., Davies, P.L. *Protein Sci.* **1995,** 4,460-471.

[14] Sonnichsen, F. D., Sykes, B. D., Chao, H., Davies, P. L. *Science* **1993**, *259*, 1154-1157.

[15] Deng, G. J., Andrews, D. W., Laursen, R. A. *FEBS Lett.* **1997**, *402*, 17-20.

[16] Fletcher, G.L., Hew, C. L., Devries, P.L. *Annu. Rev. Physiol.* **2001,** 63, 359-390.

[17] Kelli, B., Jeffry, D.M., Andrzej, W. *J. Mol. Recognit.* **2000,** 13, 101-113.

[18] Wu, Y. L., Banoub, J., Goddard, S. V., Kao, M. H., Fletcher, G. L. *Comp. Biochem. Physiol. B*, **2001,** 128, 265-273.

[19] Franks, F., Morris, E. R., *Biochim. Biophys. Acta.* **1978,** *540,* 346 – 356.

[20] Bouvet, V. R., Lorello, G. R., Ben, R. N, *Biomacromolecules,* **2006,** 7, 565-571.

[21] Bush, C. A., Ralapati, S., Matson, G. M., Yamasaki, R. B., Osuga, D. T., Yeh, Y., Feeney, R. E., *Arch. Biochem. Biophys.* **1984,** 232, 624 - 631;

[22] Bush, C. A., Feeney, R. E., Osuga, D. T., Ralapati, S., Yeh, Y., *Int. J. Pept. Protein Res.* **1981**, *17*, 125-129

[23] Tachibana, Y., Fletcher, G. L., Fujitani, N., Tsuda, S., Monde, K., Nishimura, S. *Angew. Chem. Int. Ed.* **2004**, *43*, 856 –862

[24] Lin, Y., Duman, J. G., DeVries, A. L. *Biochem. Biophys. Res. Commun.* **1972**, *46*, 87-92.

[25] Dumas, J. G. *Annu. Rev. Physiol.* **2001,** 63, 327 - 357

[26] Yang, D. S. C., Sax, M., Chakarbartty, A., Hew, C. L. *Nature,* **1988,** 333, 232-237

[27] Raymond, J.A., Wilson, P. W., DeVries, A.L. *Proc. Natl Acad. Sci. USA,* **1989,** 86, 881–885.

[28] Hew, C. L., Yang, S. C. D. *Eur. J. Biochem.* **1992,** 203, 33-42.

[29] Davies, P. L., Hew, C. L. *FASEB. J.* **1990,** 4, 2460-2468.

[30] Kuiper, M. J., Davies, P. L., Walker, V. K. *Biophys. J.* **2001,** 81, 3560-3565.

[31] Knight, C. A., Wen, D., Laursen, R. A., *Cryobiology,* **1995,** 32, 23-34.

[32] Knight, C. L. *Nature* **2000,** 406, 249-251.

[33] Raymond, J. A., DeVries, A. L. *Proc. Natl. Acad. Sci. USA.* **1977.** 74, 2589-2593.

[34] Knight, C. A., DeVries, A. L. *Science* **1989,** 245, 505-507.

[35] Knight, C. A., DeVries, A. L., Oolman, L. D. *Nature* **1984,** 308, 295-296.

[36] Horwath, K. L., Easton, C. M., Poggioli Jr., G. J., Myers, K., Schnorr, L. *Eur. J. Entomol.* **1996**, *93*, 419-433.

[37] Kuiper, M. J., Davies, P. L., Walker, V. K. *Biophys. J.* **2001,** 81, 3560-3565.

[38] Bouvet, V., Ben, R., *Cell Biochemistry and Biophysics,* **2003,** 39, 133.

[39] Knight, C. A., Driggers, E., DeVries, A. L., *Biophys. J.* **1993**, *64*, 252-259

[40] Knight, C. A., Cheng, C. C., DeVries, A. L., *Biophys. J.* **1991,** 59, 409-418.

[41] Wilson, P. *Cryoletters,* **1993,** 14, 31-36.

[42] Burcham, T. S., Osuga, D. T., Yeh, Y., Feeney, R. E. *J. Biol. Chem.* **1986,** 6390-6397.

[43] Wen, D., Laursen, R. A. *Biophys. J.* **1992,** 63, 1659-1662.

[44] Chao, H., Houston, M. E. Jr., Hodges, R. S., Kay, C. M., Sykes, B. D., Loewen, M. C., Davies, P. L., Sonnichsen, F. D. *Biochemistry* **1997,** 36, 14652-14660.

[45] Haymet, A. D. J., Ward, L. G., Harding, M. *J. Am. Chem. Soc.* **1999,** 121, 941-948.

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

[47] Tsuda, T., Nishimura, S. *Chem. Commun.* **1996,** 24, 2779-2780.

[48] Tseng, P. H., Jiang, W. T., Chang, M. Y., Chen, S. T. *Chem. Eur. J.* **2001,** 7, 585-590.

[49] Anisuzzaman, A. K. M., Anderson, L., Navia, J. L. *Carbohydr. Res.* **1988,** 174, 265-278.

[50] Meldal, M., Jensen, K. J. *J. Chem. Soc. Chem. Commun.* **1990**, 483-485.

[51] Maeji, N. J., Inoue, Y., ChÙjÔ, R. *Carbohydr. Res.* **1986**, *146*, 174-176.

[52] Rao, B. N. N., Bush, C. A. *Biopolymers,* **1987,** 26, 1227-1244.

[53] Eniade, A., Ben, R. N. *Biomacromolecules* **2001,** 2, 557-561.

[54] Costanzo, J. P., Lee, R. E. Jr., DeVries, A. L., Wang, T., Layne, J. R. Jr. *FASEB. J.* **1995,** 9, 351-358.

[55] Baust, J. M., Van Buskirk, R., Baust, J. C. *In Vitro Cell Dev. Biol. Ani.* **2000,** 36, 262-270.

[56] Meryman, H. T., Hornblower, M. *Cryobiology,* **1972,** 9, 262-267.

[57] Rubinsky, B., Arav, A., Fletcher, G. L. *Biochem. Biophys. Res. Commun.* **1991,** 180, 566-571.

[58] Pham, L., Dahiya, R., Rubinsky, B. *Cryobiology,* **1999,** 38, 169-175.

[59] Griffith, M., Ewart, K. V., *Biotechnol. Adv.* **1995,** 13, 375-402.

[60] Payne, S. R., Sandford, D., Harris, A., Young, O. A. *Meat Sci.* **1994,** 37, 429-438.

2 Research Goals and Objectives

2.1 Previous synthesized AFGP 8 Analogues

Glycoproteins are widely contained in biological systems and their carbohydrate moiety is very important in various cell functions, such as cell growth, cell-cell adhesion, fertilization, viral replication, and inflammation. [1] Therefore, the synthesis of glycoproteins is of great interest due to its intrinsic properties.

Glycoproteins have a common structure; they consist of a carbohydrate and a peptide moiety which are brought together by either an *O*-glycosidic or *N*-glycosidic linkage. The most common glycoproteins are *O*-linked glycoproteins made of L-threonine or L-serine as the peptide moiety while the most common *N*-linked glycoproteins have L-asparagine as the peptide moiety (Figure 2.1).

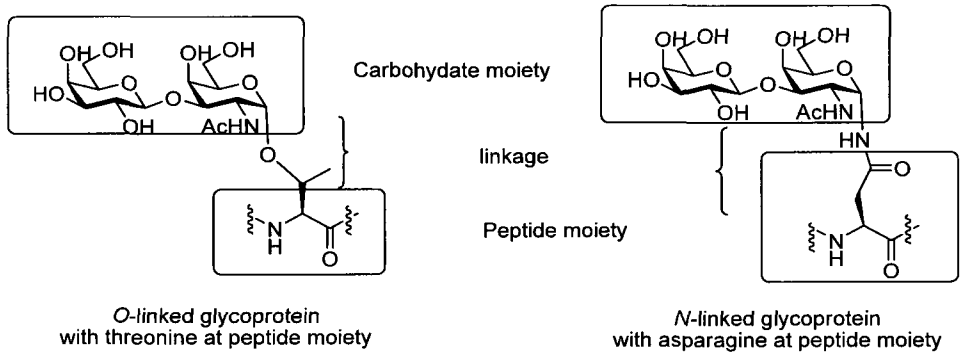
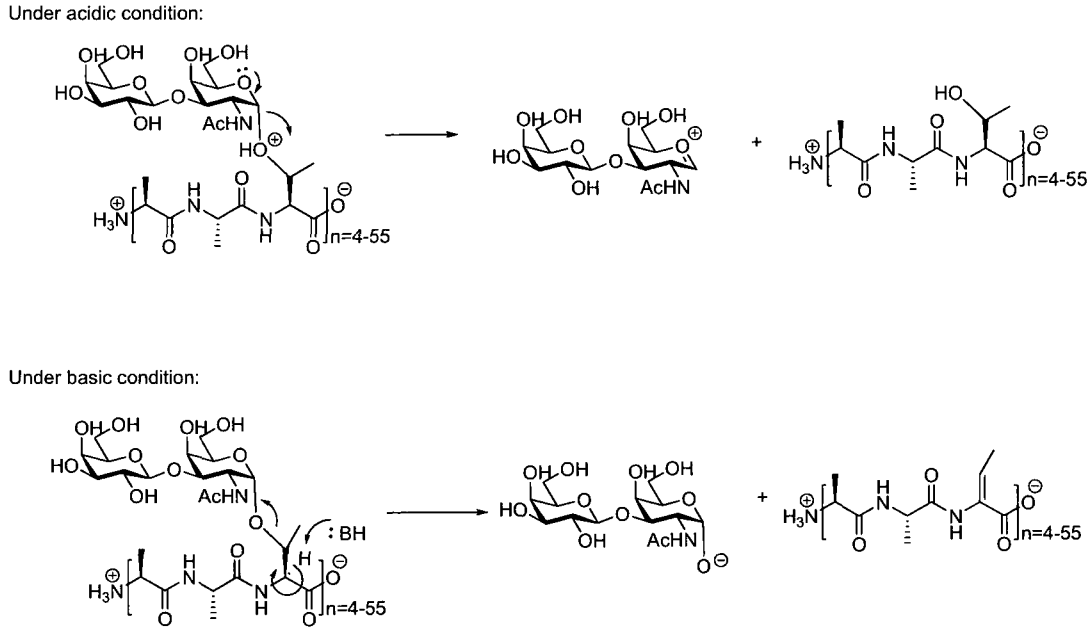


Figure 2.1 Structure of common glycopeptides

AFGP 8 is an example of *O*-linked glycoprotein found in the plasma of Teleost fish. As previously mentioned, AFPG 8 has the ability to inhibit ice formation and recrystallization in fish body fluid, allowing the fish to survive and live in subzero temperatures. These interesting properties have been extensively studied by scientists due to their biological, medical and commercial relevance. Although these glycoproteins have desirable properties and applications, they are also problematic; they are difficult to isolate and purify, making the process very time consuming and costly. Therefore, to obtain these glycoproteins, synthesis of these compounds and their analogues is another feasible solution for chemists and biologists. Another problem associated with native AFGPs is that the α *O*-linkage between the carbohydrate and the peptide moieties is very unstable under acidic, basic and enzymatic conditions.[2] Shown in Scheme 2.1 is the α *O*-linked glycosidic bond that can be easily hydrolyzed under strong acidic condition to form an oxocarbenium or under strong basic conditions to cause a β-elimination of the carbohydrate from the peptide moiety.

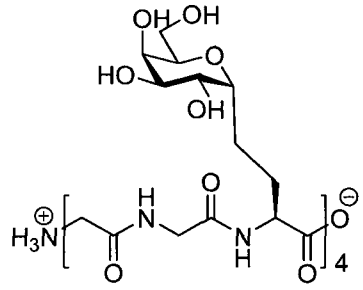


Scheme 2.1 Sensitivity of AFGP 8 under acidic and basic conditions

The instability of *O*-linked glycoproteins has therefore made the synthesis of native AFGP 8 very difficult. A solution to this problem is to synthesize AFGP 8 analogues that mimic the native AFGP 8 structures. The most common AFGP 8 analogues include *C*-linked and *S*-linked glycoproeins, which have been synthesized by different groups and shown improved stability under acidic, basic and enzymatic conditions compared to their native form.[3] The *S*-linked AFGP analogues have similar chemical properties as *O*-linked but are more stable, while *C*-linked analogues differ from *O*-linked in their relative polarity and electronic effect. However, it is important to make note of these similarities: research done by Kishi and co-workers has shown that *C*-linked and *O*-linked analogues

adopt similar conformation in solution, irrespective to the changes in bond length, bond angle and dihedral angle in *C*-linked analogues caused by substitution of C-O bond with a methylene group,[4] and also the *C*-linked analogues have showed similar antifreeze activity as their native form. Glaudemans studied the binding of di-, tri- and tetrasaccharides to three monoclonal antigalactan immunoglobulins, and compared the results with corresponding *C*-linked oligosaccharide analogues to understand the hydrogen bond interactions involved in this process. Both *O*- and *C*-linked oligosaccharides showed nearly identical affinities to the same proteins and similar ligand-induced tryptophanyl fluorescence changes, which strongly implied the near identity in binding for *C*- and *O*-oligosaccharides.[5]

The Ben group has taken interest in AFGP 8 and has thus synthesized various analogues that have modifications at either the carbohydrate or peptide moiety. Furthermore, these analogues have been tested for their thermal hysteresis (TH), dynamic ice shaping (DIS) and ice recrystallization inhibition (IRI) properties. Of these analogues *C*-serine analogue **46** (Figure 2.2) showed the most potent IRI activity and exhibited no TH. This analogue was synthesized based on several structural modifications and simplification of the native system. Firstly, the Gal-GalNAc disaccharide unit was replaced by a single saccharide being galactose. Secondly, the C2 NHAc was replaced by a regular hydroxyl group. Thirdly, the *O*-linkage was replaced by a more stable *C*-linkage. Lastly, the L-threonine-alanine-alanine tripeptide repeating unit was replaced by L-serine-glycine-glycine.



46

Figure 2.2 Structure of the most potent AFGP 8 analogue C-serine.

Subsequent to the synthesis of *C*-serine analogue **46**, two other *C*-linked analogues with serine-like peptide moieties were synthesized to investigate the effect of side chain length. Influenced by *C*-serine analogue **46**, the linkage between the carbohydrate and peptide moieties of analogue **47** was extended by one methylene group while analogue **48** was extended by two (Figure 2.3). The IRI results of **46** and these two *C*-serine-like AFGP 8 analogues are shown in Figure 2.4. From the data presented in this graph, analogues **47** and **48** did not exhibit improved IRI activity.

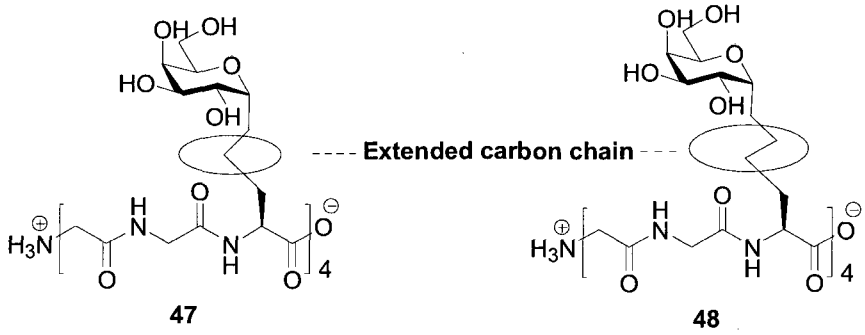


Figure 2.3 Structure of C-serine AFGP 8 Analogue

Previous studies showed that the IRI activity of *C*-linked AFGP analogues was proportional to the carbohydrate concentration [6] and the rigidity of the *C*-linked analogue linkage could directly affect the orientation of the carbohydrate moiety against the peptide backbone. A rigid linkage could limit the extent of free rotation and stabilize the carbohydrate moiety against the polypeptide backbone; so as to increase vicinal carbohydrate concentration in solution thus increases the IRI activity. However, the linkages in analogue **47** and **48** were too flexible to keep carbohydrate moieties close to polypeptide backbones, consequently, the vicinal carbohydrate concentrations were decreased; therefore, these two analogues showed very limited IRI activities.

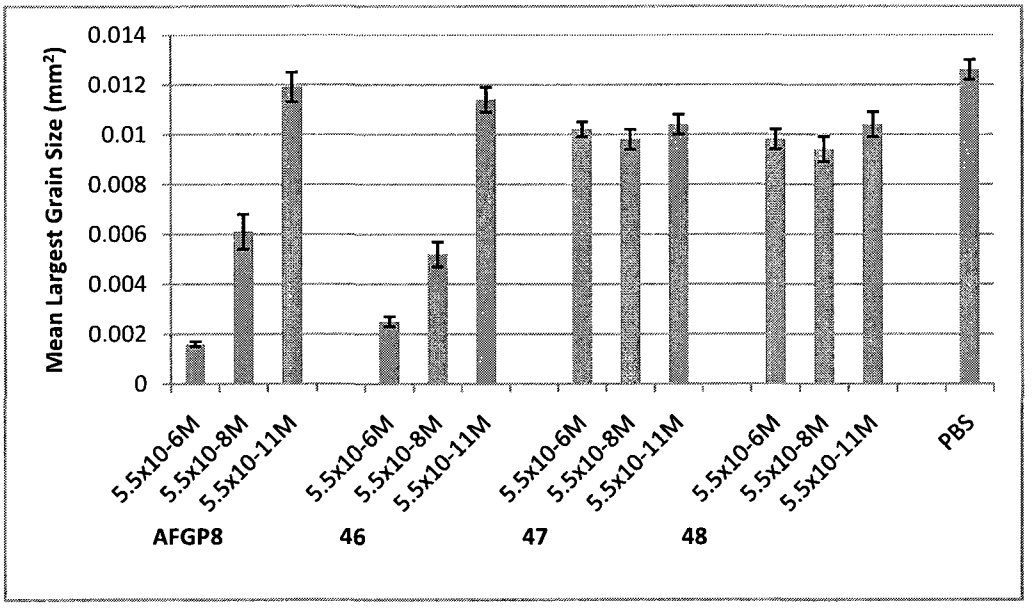
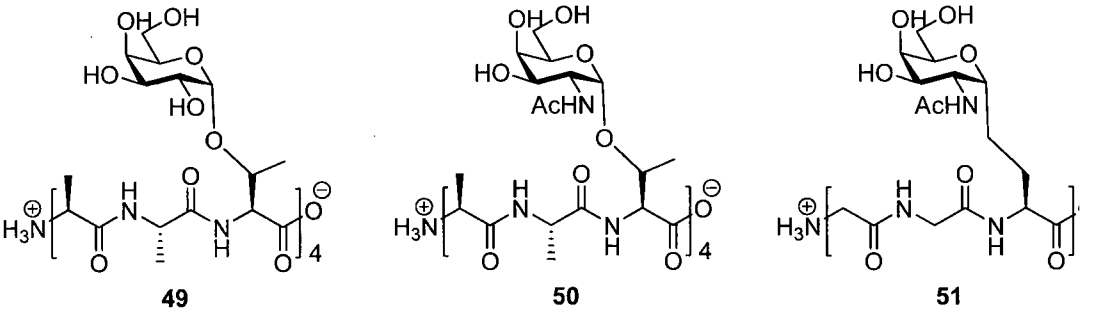


Figure 2.4 Ice recrystallization inhibition of different *C*-serine analogues.[7]

We then wanted to study the C2 NHAc group and its influence on IRI activities, thus analogues **49** and **50,** which contained C2 hydroxyl group and NHAc group respectively, were synthesized. In addition, a *C*-serine analogue **51** which contains C2 NHAc group was synthesized used to compare with the *C*-serine analogue **46** containing C2 hydroxyl group (Figure 2.5). The IRI results of each analogue are shown in Figure 2.6.



49 50 51

Figure 2.5 Structures of O-linked AFGP 8 analogues containing C2 hydroxyl or NHAc group and *C*-serine analogue containing C2 NHAc group

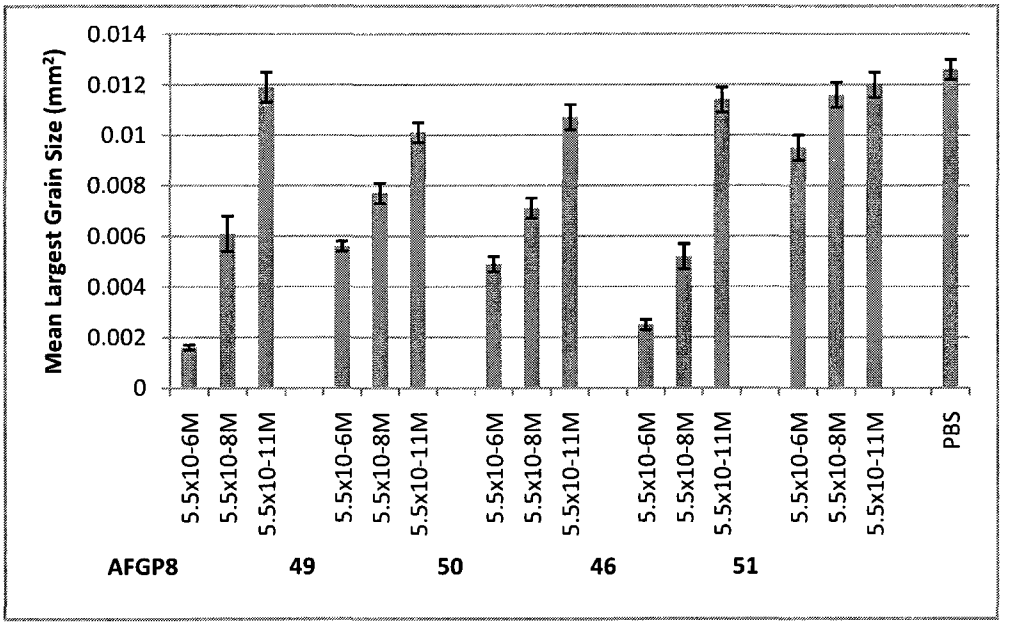


Figure 2.6 Results of ice recrystallization inhibition activity of AFGP 8 analogues with or without C2 NHAc groups.[5,8]

Figure 2.6 illustrates IRI results of various AFGP 8 analogues which have either a C2 hydroxyl (**46** and **49**) or NHAc group (**50** and **51**) in their saccharide moieties and are either *O*-linked (**49** and **50**) or *C*-linked (**46** and **51**). At concentration of 5.5×10^{-6} M and 5.5×10^{-8} M, *O*-linked analogue **50** which contains a C2 NHAc group showed an increased activity by 12% comparing with the *O*-linked analogue **49** that has a C2 hydroxyl group. However, the *C*-linked analogues **46** and **51** showed opposite results. *C*-serine analogue **51** which has a C2 NHAc group showed about only 31% IRI activity comparing with the *C*-serine analogue **46** which has a C2 hydroxyl. Thus, it can conclude from above data that the C2 NHAc group plays different roles in both *O*-linked and *C*-linked analogues. Moreover, because *C*-serine GalNAc analogue **51** only displayed DIS activity but *C*-serine galatose analogue **46** only showed IRI activity, it could be confirmed that the DIS, TH and IRI activities do not imply a similar interactions with the ice lattice therefore it is possible to synthesize AFGP 8 analogues possessing one activity or another.[6]

2.2 Research Goals and Objectives

Biological antifreezes have three particular properties, IRI, DIS and TH as mentioned previously. When the temperature of a solution is lower than its freezing point, small ice crystals formed in solution can grow into larger ones and this is known as the process of ice recrystallization.

DIS, another antifreeze property, is a direct result of a solute binding at the interface of the ice lattice and the quasi-liquid layer (QLL), a semi-ordered layer between ice and bulk water surrounding the carbohydrates. When the ambient temperature is within the TH gap, small ice crystals having hexagonal bipyrimidal shape are formed within the organism's blood and body fluid. This occurs because the biological antifreezes block the a-axes of ice crystals allowing the ice to grow only along the c-axis, giving it its unique shape.[9] However, when the temperature is below the TH gap, ice crystals grow rapidly along the c-axis giving it a needle-like shape, which will eventually cause damage to cell membranes.[10]

It has been reported that TH and IRI are independent activities.[11] The focus of the Ben lab is to synthesize compounds used in cryopreservation, which have enhanced IRI activity with no TH because studies showed that the major cause of decreased cellular viability upon cryopreservation is cellular damage caused by ice recrystallization rather than TH.[12]

Recent reports suggested that the ability of inhibiting ice recrystallization of simple mono- and disaccharides is based on complimentarity of hydroxyl groups with the ice lattice and proposed that the optimal distance between hydroxyl groups is 4.2 – 4.5 Å.[13] My colleagues have recently tested the TH activities of different mono- and disaccharides and suggested that the carbohydrate interacts with the QLL rather than the ice lattice directly.[14] Analysis for TH activity suggests that the specific distance between hydroxyl groups is not the only factor responsible for the antifreeze activity. Hydration number, a

parameter describing the number of water molecules surrounding the carbohydrate molecule, is also important.[15] In general, a larger hydration number implies that there are more water molecules surrounding the carbohydrate and forming more hydrogen-bonds between them, which cause greater disorder in the surrounding bulk water (Figure 2.7).[16] Therefore, this will hinder the addition of bulk water molecules into the QLL because it will cost very high energy to break the existing hydrogen-bonds between the water molecules which results in smaller ice crystals. Thus, carbohydrate has larger hydration number shows better IRI activity.

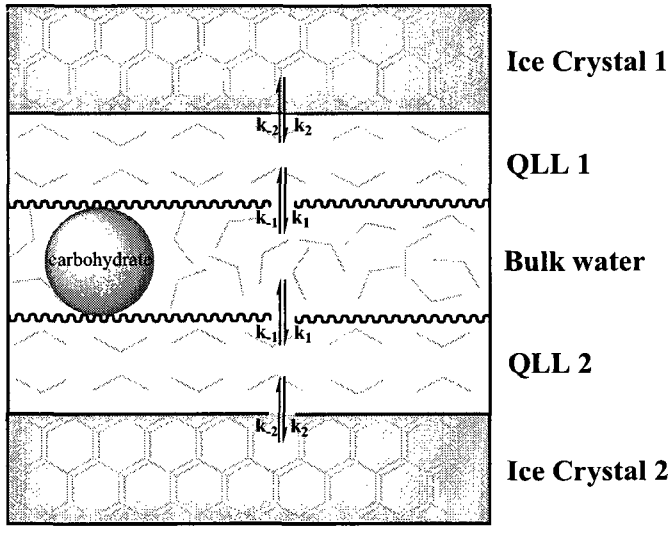


Figure 2.7 Proposed relationship between carbohydrate, surrounding bulk water and QLL

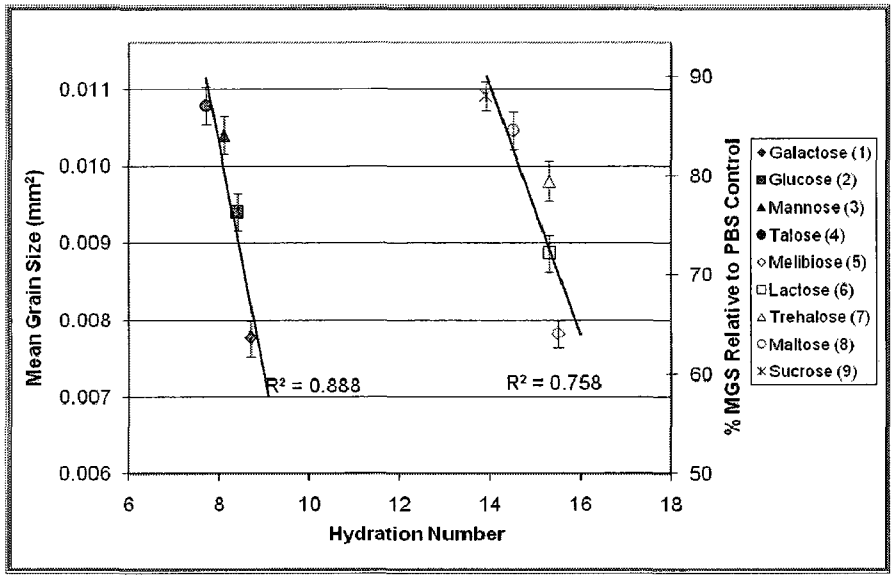


Figure 2.8 Relationship between the hydration numbers of carbohydrates and respective IRI activities at 0.022 M in PBS solution.[13]

Figure 2.8 shows a strong linear relationship between the hydration numbers of mono- and disaccharides and their respective IRI activities at 0.022 M in PBS solution.[13] However, the general increase in hydration number for disaccharides relative to monosaccharides does not correlate to an expected increase in IRI activity. The Ben lab has shown[14] this is due to the fact that volume of the solvent has been neglected. Another parameter, hydration index is used to correct this difference. It describes the number of tightly bound water molecules per molar volume of carbohydrate which is obtained by dividing hydration numbers by partial molar volumes. [17] Figure 2.9 shows the linear relationship between the hydration index of mono- and disaccharides and their respective

IRI activities which means that not only the absolute number of water molecules surrounding a solute is important for its IRI activity, but also is the number of water molecules per unit volume of solute. A higher concentration of water molecules around a solute will increase disorder of the surrounding bulk water[18] leading to increased IRI activity.

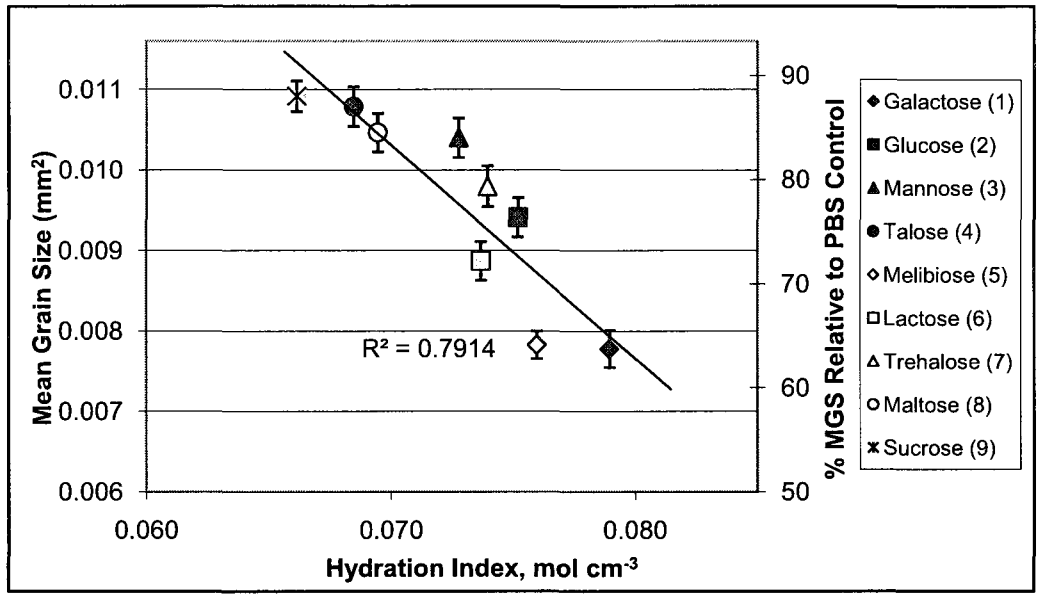
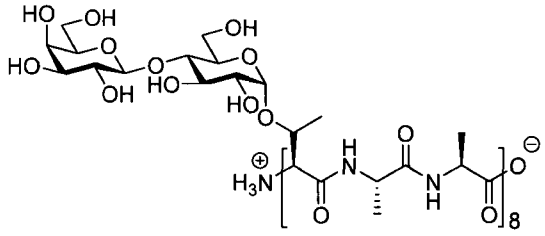


Figure 2.9 Relationship between the hydration index of mono- and disaccharides and their respective IRI activity.[13]

It is known that the *O*-linked native AFGP 8 which has a Gal-GalNAc as its carbohydrate moiety showed potent IRI activity. However, its TH activity does not make it a suitable choice for cryopreservation. The *C*-serine analogue **46** has also been tested

and shown to be the most potent AFGP analogue with no TH activity, which makes it a good candidate for cryopreservation. Both of the two glycopeptides have galactose as their carbohydrate component, and their IRI activities are in accordance with the result that galactose has a large hydration number and hydration index and shows a good recrystallization inhibition activity by itself.



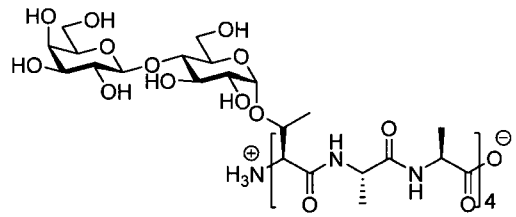
52

Figure 2.10 Structure of compound 52

According to Nishimura's research, compound **52** (Figure 2.10) which contains lactose as its carbohydrate component and 8 L-alanine-L-alanine-L-threonine repeating units as its peptide component showed no TH activity but was able to produce hexagonal-like crystals. [19] However, its IRI activity is still unknown as it had not been tested. Based on the data in Figure 2.8 and 2.9, lactose has a large hydration number and large hydration index among the other tested disaccharides and has good recrystallization inhibition activity comparable to galactose. According to the good IRI activities of native AFGP 8 and C-serine analogue **46** are in accordance with the ice recrystallization inhibition activity of their carbohydrate component galactose, therefore, we hypothesize

that if compound **52** possesses potent IRI activity which was in accordance with the activity of lactose and then it would be a suitable candidate for cryopreservation applications.

The focus of my research is to synthesize an antifreeze analogue **53** and test its IRI activity. This compound has same overall structure as compound **52** but only contains 4 L-alanine-L-alanine-L-threonine repeating units (Figure 2.11). It is also my goal to compare its IRI activity with native AFGP 8 which has an axial hydroxyl group at C4 and has a 1, 3 *O*-linkage between the two sugar units. Based on the result of compound **53** IRI activity along with the good ice recrystallization inhibition activity of lactose, then we can conclude if the IRI activity of AFGP analogue containing disaccharide as its carbohydrate component is in accordance with the hydration number and hydration index of the disaccharide.

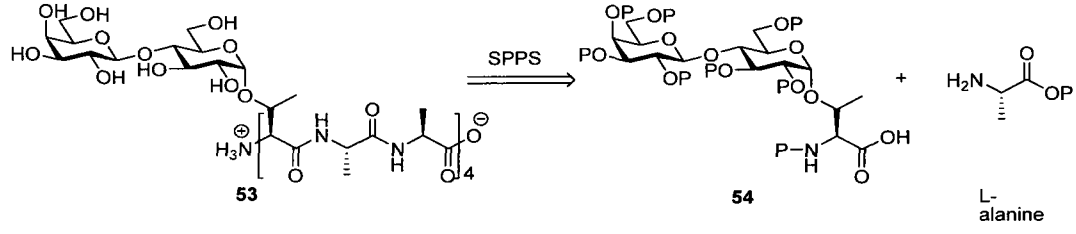


53

Figure 2.11 Target AFGP 8 analogue 53.

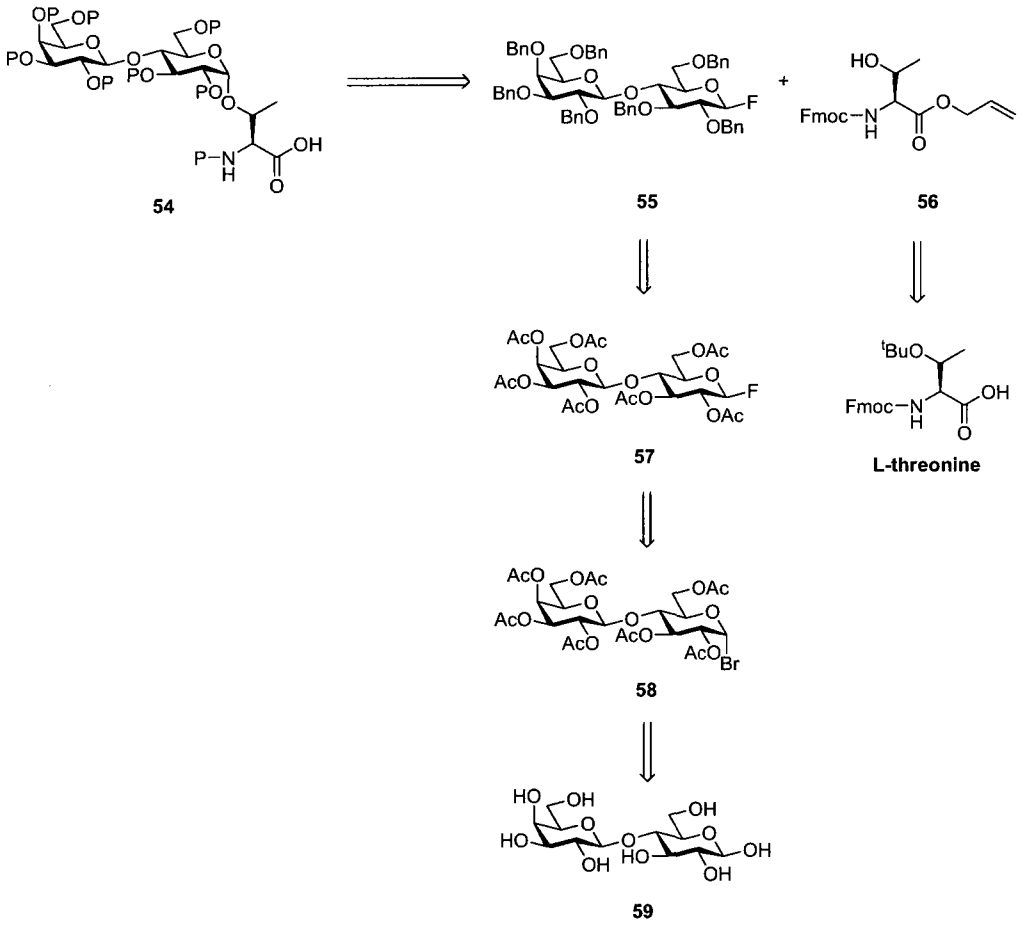
2.2.1 Retrosynthesis of AFGP 8 analogue

In order to build compound **53**, it is important to first synthesize building block **54,** which is made up of lactose and L-threonine to give the carbohydrate and peptide moieties, respectively. After synthesizing the building block, it can then be coupled with commercially available L-alanines to assemble the final glycoprotein by using solid phase protein synthesis (SPPS) (Scheme 2.2).



Scheme 2.2 Disconnection of the target AFGP 8 analogue

Building block **54** was synthesized from commercially available lactose and L-threonine. The retrosynthetic analysis is shown in Scheme 2.3.



Scheme 2.3 Retrosynthesis of lactose - threonine building block 54

In order to couple threonine at the α anomeric position of the lactose **55** to form the building block **54**, it is important to have a leaving group such as a fluorine at the β anomeric position of lactose **57** which would be synthesized from α bromo lactose acetate **58** by an S_N2 reaction. In order to perform α coupling reaction between threonine and β lactose fluoride, the acetate protecting groups of those hydroxyl groups on β lactose

fluoride **57** have to be replaced by benzyl groups to give lactose fluoride **55**. Changing the protecting groups is important because acetate groups have the neighbouring group participating effect which give rise to only β coupling products. The α bromo lactose acetate **58** would be synthesized from commercially available lactose **59**.

2.3 Summary

Due to the difficult synthesis of the native AFGP 8, a lot of research on synthesizing AFGP 8 analogues had been done. In those synthesized AFGP 8 analogues, C-serine analogue has the most potent antifreeze activity. For *O*-linked AFGP 8 analogue, C2 NHAc group plays an important rule.

Antifreeze activity of carbohydrates appears to be related to its absolute hydration number and hydration index. In tested disaccharides, lactose has a large hydration number and hydration index and good IRI activity. Therefore, if incorporated the lactose into the peptide unit to form a glycopeptide **53** based on the synthetic analysis shown in Scheme 2.2 and Scheme 2.3, this glycopeptide might have good IRI activity and be a good candidate of AFGP 8 analogue used in cryopreservation since this glycopeptide was tested with no TH activity by Nishimura group.

[1] Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683-720.

[2] Boons, G. J., Hale, K. J. *Organic Synthesis with Carbohydrates*, Blackwell Science: Malden, MA. **2000**, page 156-160.

[3] Marcaurelle, L. A., Bertozzi, C. R., *Chem. Eur. J.* **1999**, *5*, 1384-1390.

[4] Wei, A., Boy, K. M., Kishi, Y. *J. Am. Chem. Soc.* **1995**, *117*, 9432-9436.

[5] Wang, J., Kováč, P., Sinaÿ, P., Glaudemans, C. P. J. *Carbohydr. Res.* **1998**, *308*, 191-193.

[6] Murphy, A. V. "Preparation of structurally diverse C-linked antifreeze antifreeze glycoprotein analogues and assessment for antifreeze protein-specific activity" PhD Thesis, Binghamton University (S. U. N. Y.), 2004.

[7] Liu, S., Ph. D. Thesis, University of Ottawa, **2006.**

[8] Bouvet, V., Ph. D. Thesis, University of Ottawa, **2005.**

[9] Hew, C. L., Yang, S. C. D. *Eur. J. Biochem.* **1992,** 203, 33-42.

[10] Davies, P. L., Hew, C. L. *FASEB. J.* **1990,** 4, 2460-2468.

[11] Horwath, K. L., Easton, C. M., Poggioli Jr., G. J., Myers, K., Schnorr, L. *Eur. J. Entomol.* **1996**, *93*, 419-433.

[12] (a) Wang, T., Zhu, Q., Yang, X., Jr., DeVries, A. L., *Cryobiology,* **1994,** 31, 185 – 192.

(b) Mazur, P. C., *Science,* **1970,** 168, 939 – 949.

(c) Mazur, P. C. *Am. J. Physiol. –Cell Ph.* **1984,** 247, C125 –C142.

(d) Rubinsky. B. *Heart Fail. Rev.* **2003,** 8, 277 – 284.

[13] Baruch, E., Belostotskii, A. M., Mastai, Y., *J. Mol. Struct.* **2008,** 874, 170 – 177.

[14] Tam, R.Y., Ferreira, S. S., Czechura, P., Chaytor, J. L., Ben, R. N., *J. Am. Chem.Soc.* **2008,** 130, 17494-17501.

[15] Madura, J. D., Baran, K.; Wierzbicki, A., *J. Mol. Recognit.* **2000,** 13, 101 – 113.

[16] (a) Leinala, E. K.; Davies, P. L.; Jia, Z. *Structure* **2002**, *10*, 619-627.

(b)Baardsnes, J.;Davies, P.L. *Biochim. Biophys. Acta*, **2002**, *1601*, 49-54.

[17] (a) Høiland, H., Holvik, H. *J. Solution Chem.* **1978**, *7,* 587-596.

(b) Shahidi, F.; Farrell,P.G.; Edward, J. T. *J. Solution Chem.* **1976**, *5*, 807-816.

[18] (a) Leinala, E. K.; Davies, P. L.; Jia, Z. *Structure* **2002**, *10*, 619-627.

(b) Baardsnes, J.;Davies, P.L. *Biochim. Biophys. Acta*, **2002**, *1601*, 49-54.

[19] Tachibana, Y., Fletcher, G. L., Fujitani, N., Tsuda, S., Monde, K., Nishinura, S. I., *Angew. Chem. Int. Ed.* **2004,** 43, 856 – 862.

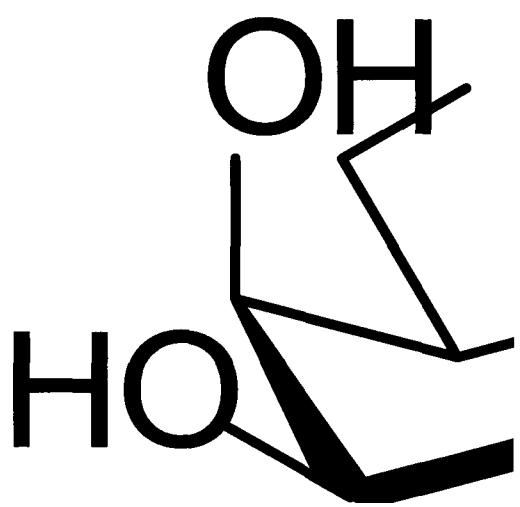
3 Synthesis of Lactose – Threonine *O*-linked AFGP Analogues

As shown in Scheme 3.1, the target glycopeptides **53** was prepared by SPPS between building block **54** and commercially available L-alanine. The requisite building block **54** was assembled by coupling a fluorinated disaccharide **55** to a fully protected threonine such as **56**.

Scheme 3.1 Retrosynthetic analysis of target analogue 53

3.1 Preparation of the Carbohydrate Component

The synthetic route for the carbohydrate component **55** is shown in Scheme 3.2.



Scheme 3.2 Synthesis of the Carbohydrate Component 55

The synthesis began by refluxing commercially available lactose **59** (95% of β-lactose) with acetic anhydride and sodium acetate for 2 hours to fully protect the

hydroxyls with acetate groups to afford **60**. After work up, peracetylated lactose **60** was recrystallized from cold ether as white powder in 90% yield.[1] The bromo-lactose derivative **58** was then synthesized by stirring peracetylated lactose **60** in 33% HBr/AcOH at 0 °C for 2 hours and purified by work up and recrystallization from cold ether in almost quantitative yield.[2]

During the above bromination to produce disaccharide **58**, the bromide anion is normally added to the oxocarbenium in the α configuration at the anomeric position due to the anomeric effect.[3] This effect was first proposed by Dr. Lemieux to address the tendency of strong electron withdrawing groups to prefer the α position at the anomeric centre and there are two explanations for its nature. The first is a dipole-dipole effect which is not commonly accepted by most scientists. The dipole moments of a peracetylated sugar with a bromine at the α or β position are shown in Figure 3.1. In Figure 3.1a, the dipole moments of the oxygen and bromine atoms are in opposite direction for the α anomer. Therefore the net dipole moment of the whole system is decreased. In Figure 3.1b, both of the dipole moments of oxygen and bromine point in the same direction causing the overall molecular dipole moment to increase. Comparing the two overall dipole moments, the α anomer is more stable than β anomer because it has a lower dipole moment. As a result, bromine is more likely to be in the α position.

Figure 3.1 Dipole moments of α and β anomer. (A) α anomer dipole moment. (B) β anomer dipole moment.

The more accepted explanation is a stereo-electronic effect. As seen in Figure 3.2a, the bromine is in the α position and the anti-bonding orbital of carbon-bromine (C-Br) bond is anti-periplanar to one of the lone pair orbitals of the oxygen atom. This can be easily seen in the Newman projection of Figure 3.2b. Thus, the lone pair of electrons on oxygen can donate electron density into the C-Br anti-bonding orbital to stabilize it. On the other hand, the β anomer does not have a C-Br anti-bonding orbital in proper orientation where it can be stabilized by the lone pair of electrons of oxygen. Therefore, the compound with the bromine atom in the α position is more stable than the β anomer and the bromo-lactose derivative is more likely to adopt the α configuration.

Figure 3.2 The Stereo-electronic Effect. Anti-periplanar relationship between oxygen lone pair of electrons and C-Br anti-bonding orbital.

In order to facilitate the formation of a glycoconjugate with a β-linkage at the anomeric position (Scheme 3.1) we decided to install a fluorine atom at the anomeric position (Scheme 3.2). Consequently, the bromine atom of glycosyl donor **58** was converted to fluorine by treating it with silver fluoride in acetonitrile to furnish glycosyl fluoride **57** with the desired β-configuration at the anomeric center in 65% yield.[4]

The use of acetates for a coupling reaction is quite problematic due to the neighbouring group participation effect. The presence of an ester at position C2 often results in product with a β configuration. In order to generate the desired α product, the C2 ester has to be replaced by non-participating protecting groups. For this reason, we have chosen to replace the acetates for benzyls. The acetates of compound **57** were removed using diluted MeONa solution (pH = 8) in methanol at 0 °C. The reaction mixture was then neutralized using Amberlite-120 beads, filtered, and concentrated under reduced pressure. The residue was immediately benzylated using BnBr and NaH in DMF

under nitrogen for 18 hours to obtain the perbenzylated lactose fluoride **55** in an average 63% yield.

3.2 Preparation of Peptide component

The synthetic route for the peptide component is shown in Scheme 3.3. Commercially available *N*-Fmoc-*O*-*tert*-butyl-L-threonine **61** was coupled with allyl alcohol using HBTU to afford orthogonally protected ester **62** in an average 62% yield. The *t*-butyl group in **62** was then removed by treating with 50% TFA in water to furnish the secondary alcohol in **56** in an average 67%.

Scheme 3.3 Synthesis of the peptide component.

3.3 Preparation of the Building Block for Solid Phase Peptide Synthesis

With both of the carbohydrate and peptide components in hand, a coupling reaction was performed to prepare the building block used for SPPS of the *O*-linked glycopeptides (Scheme 3.4).

Scheme 3.4 Synthetic route of Building Block for Solid Phase Peptide Synthesis

Threonine allyl ester **56** and perbenzylated lactose fluoride **55** were coupled in anhydrous DCM using tin chloride and silver perchlorate as promoters. The reaction was then stirred at -20 °C and allowed to warm to ambient temperature over 24 hours to afford the coupled product in a 2:1 mixture of α (**63**) and β products (**64**) in 50% overall yield.[5] Finally, $Pd(PPh_3)_4$ and *N*-methylaniline were used to remove the allyl groups of **63** and **64** to give building blocks **54** and **65** in about 70% yield.[6]

3.4 SPPS of Lactose *O*-linked AFGP8 Analogue

Solid phase peptide synthesis was first developed by Robert Bruce Merrifield in 1959[7] and has proven to be revolutionary for the synthesis of peptides and small proteins. There are two main parts to this technique. The first part is to elongate the peptide chain with protected amino acid derivatives on a solid support and the second is to cleave the synthesized peptide chain from the solid support to obtain the final peptide.

In the peptide chain elongation part, SPPS proceeds in a *C*-terminus to *N*-terminus fashion. The general principle is the repeated cycle of coupling-deprotection. The free *N*-terminal amine site of elongated peptide chain is always coupled to a carboxylic acid of a single amino acid which also contains a protected *N*-terminal amine. Following the coupling step, the protecting group of *N*-terminal amine is removed and the coupling-deprotection cycle is repeated until the desired peptide sequence is reached.

There are two approaches in SPPS, one is to use Fmoc-protecting groups (base labile) and another is to use *t*-Boc protecting groups (acid labile). For each method, there are specific types of resins and cleavage techniques. Fmoc is a base labile protecting group used for SPPS of acid sensitive substrates while *t*-Boc is an acid labile protecting group used for SPPS of base sensitive substrates.

The primary advantage of SPPS is its high yield and simple purification. Since the synthesized product is attached on solid supports, all of the purification procedures only involve washing the beads with regular solvents such as DMF, DCM or MeOH which is much simpler compared to solution phase synthesis which needs work up or flash chromatography. After washing the beads several times, the beads are ready for the next synthetic step. As a result of this easy purification technique, the yield of the peptides synthesized is typically greater than those prepared by solution phase synthesis.

Commercially available Wang resin preloaded with Fmoc-alanine was used as solid support for the lactose *O*-linked AFGP 8 analogue synthesis. As shown in Figure 3.3, the resin is made of cross-linked polystyrene matrix and attached with a *p*-benzyloxy benzylalcohol linker. The linker is preloaded with an Fmoc-alanine through an acid-labile ester bond formed between the hydroxyl group of the linker and the C-terminal carboxyl acid of Fmoc-alanine. This ester bond can be easily cleaved at end of SPPS using mild acid conditions. The lactose *O*-linked AFGP8 analogue **53** SPPS uses Fmoc-chemistry, which means that the Fmoc group used to protect the *N*-terminal amine of alanine is removed with 20% piperidine in DMF and every single amino acid added to

the elongated steps all has a Fmoc protected N-terminal amine site and a free carboxylic acid site.

Figure 3.3 General structure of Fmoc-alanine-Wang resin.

The general SPPS protocol of *O*-linked lactose AFGP analogues **69** and **73** is shown in Scheme 3.5. The very first step involved the swelling of the resin by soaking it in anhydrous DMF for 1 hour. The next step was to remove the Fmoc group of alanine to generate a free amine site which would couple with the *C*-terminal carboxylic acid of an individual Fmoc-alanine in the next elongation step.

In the elongation step, 3 equivalents of Fmoc-alanine-OH, 1.1 equivalents HCTU and diisopropylethylamine (DIPEA) were added into the synthesizer to perform the coupling reaction between the C-terminal carboxylic acid of Fmoc-alanine-OH and the free *N*-terminal amine on the resin by shaking the synthesizer for 2 hours. After repeating this process once, the *N*-terminal Fmoc group was then removed and another Fmoc-alanine-OH coupling reaction was performed and repeated once.

After removing the Fmoc group at *N*-terminus of the elongated peptide chain, the next substrate would be attached to the terminal alanine was the building block **54** or **65**. Therefore, 1.5 equivalents of building block, 1.1 equivalents HCTU and 7.5 equivalents of DIPEA were added into the synthesizer and shaken for 30 hours. At this point, two alanines and one building block have been lineally attached onto the resin to form the corresponding short peptides **66** and **70**.

Following the removal of the Fmoc group of the building block on the short peptide synthesized, the whole process was repeated by adding another two alanines and one building block, according to the above procedure for three more times to generate corresponding oligomer **67** (α) or **71** (β) attached on the resins.

Scheme 3.5 Solid phase synthesis of *O*-linked AFGP8 analogues 69 and 73.

As seen in Scheme 3.5, polymers **67** or **71** are cleaved from the resin using 30% TFA in DCM. The crude compounds are then precipitated from cold ether to afford **68** or **72** as white powder.

3.5 Debenzylation to Afford the Final Polymer

After obtaining the α and β oligomers **68** and **72**, the next step is to remove all benzyl ether groups on the carbohydrate moiety to generate the final polymers **69** or **73**.

Benzyl ether is widely used as a protecting group for hydroxyls in carbohydrate and nucleoside chemistry. Its deprotection is typically performed using a palladium catalyzed hydrogenolysis which generates the free hydroxyl group and toluene. This deprotection can also be achieved using strong acids such as HBr,[8] but this method is not applicable to acid-sensitive substrates. Alternatively, benzyl ether groups may be removed by oxidation to the benzoate ester which allows a subsequent hydrolysis under basic conditions like MeONa, [9] or by using a Lewis acid like BCl_3[10] or $BF_3 \bullet Et_2O$. [11] There are some substituted benzyl ethers which can use more specific and high yielding deprotection methods such as DDQ oxidation.[12] However, since the synthesized *O*-linked glycopeptides are both acid and base sensitive, [13] these conditions may be too harsh for glyconjugates **68** and **72**. Hence, the best choice is to employ hydrogenolysis under neutral conditions. For the palladium catalyzed hydrogenation, the common

catalyst used is 5% or 10% activated palladium on charcoal and the reaction is commonly carried out in a protic solvent such as methanol or ethanol.

It has been reported that debenzylation reactions are often carried out in alcoholic solvents or in acetic acid and the catalyst activity is usually diminished when using other solvents. Moreover, using acids can promote debenzylation and using water can also significantly increase the activity of the catalyst.[14] Therefore, the hydrogenolysis of the benzyl ether group is solvent dependent as illustrated in Table 3.1.[15]

Table 3.1 Solvent effect on the hydrogenolysis of benzyl ether (1.1 bar H_2, 50°C)

Solvent	Relative rate
MeOH	2.5
EtOH	3.5
propanol	7
hexanol	12.5
octanol	16
Acetic acid	17

hexane	3
toluene	1

It is shown in Table 3.1 that all of the protic solvents have enhanced reaction rates for the hydrogenolysis of benzyl ether compared to aprotic solvents such as toluene and hexane. Therefore, solvents which can provide hydrogen bonding greatly influence reaction rates compared to aprotic solvent.

While alcohol based solvents increase reaction rates, acetic acid may also be used to effectively remove benzyl groups. As seen in Table 3.1, the deprotection rate for acetic acid is much larger than the alcohols. Furthermore, the pKa of acetic acid is 4.75, and that of most aliphatic alcohols are about 16, which makes the proton of acetic acid much more accessible for protonating the benzyl ether. As a result, acids promote debenzylation and accelerate the hydrogen transfer step.

It is also seen in Table 3.1 that longer chained alcohols, such as hexanol and octanol, have greater reactivity compared to simpler ones (methanol and ethanol). The pKa differences among those alcohols are very small; therefore, the solvent polarity is more likely a factor contributing to accelerate hydrogenolysis rates. Furthermore, long chain alcohols are less polar, and compounds having benzyl ether groups also tend to become

non-polar. Therefore, it may be assumed that solvents having a longer hydrocarbon chain may dissolve benzyl compounds better thus enhancing the deprotection reactivity.

Synthesized glycopeptides **68** and **72** have a very non-polar carbohydrate component due to the seven benzyl ether groups making them insoluble in common protic solvents (methanol and ethanol) used for hydrogenolysis. Therefore, in order to dissolve those glycopeptides, a 1:1 THF/ethanol mixture was initially used as the solvent system to increase the solubility.

To deprotect compound **68** and **72**, there were four trials set up using different solvent systems, different metal catalysts or different hydrogen pressure conditions. However, none of them would yield the desired products **69** and **73**.

The first hydrogenation trial was set up under normal pressure overnight, using 10% Pd/C as catalyst and 1:1 THF/ethanol as solvent. The reaction residue was collected and analyzed by MALDI and the molecule ion peaks referred only to starting material indicating that the hydrogenolysis did not take place under the conditions stated above.

In contrast to the first trial, the second one was set up under 600 psi [16] in order to increase the concentration of hydrogen in the reaction mixture, which would in turn increase the formation of the metal catalytic complex of the proposed mechanism in Figure 3.4. Similarly, the reaction residue was collected and analyzed by MALDI. Unfortunately, there were no molecular ion peaks to confirm the presence of starting material, fully or partially debenzylated product, making the second trial unsuccessful.

Interestingly, precipitate had formed when re-dissolving the reaction residue in THF which suggested the presence of a polar product; however, it could not be identified as decomposed starting material, fully or partially debenzylated product from the molecular ion peaks.

As mentioned, acetic acid can accelerate the debenzylation process since it can protonate the benzyl ether and promote the hydrogen transfer process. However, since *O*-linked glycopeptides are very acid labile, there was a big risk in using the acid to debenzylate since the carbohydrate component might be cleaved from the peptide backbone to form an oxocarbenium. [17] While it has been reported that these glycoconjugates are acid sensitive, Nishimura's research indicates otherwise. His analogue which had a similar structure to *O*-linked **68** but differed only at the tripeptide unit was dissolved in a mixture of 3:1:1 DMF/AcOH/H_2O. While this solvent system is acidic, it proved to be a good choice as the reaction was successful.[18] Therefore, in my third trial, I tried a similar solvent system (3:1:1 THF/AcOH/H_2O). THF was used in place of DMF as compound **68** could not be dissolved in DMF.

Two experiments were set up in this trial. The first experiment was done under normal pressure, and the second experiment was set up under 500 psi pressure; both experiments were using 10% Pd/C as catalyst. Reaction residues of both experiments were collected and analyzed by MALDI. However, there were no molecular ion peaks to indicate the starting material or a fully or partially debenzylated product. This suggests that the starting material had decomposed and that deprotection was unsuccessful.

Since the first three trials all failed, the next option was to change the metal catalyst from 10% Pd/C to 5% Rh/Al_2O_3 since rhodium is much more reactive than palladium.[19] Four experiments were conducted in this trial. Two were done in THF/ethanol while the other two were done in a $THF/AcOH/H_2O$ solvent system. For each reaction pair a different pressure was used. One reaction was done at normal pressure while the other was carried out under 500 psi condition. Based on MALDI results, the molecular ion peaks of those four experiments indicated that there was no starting material or full or partial debenzylation products. This again would suggest that the starting material had decomposed and the deprotection was not a success.

Due to time restriction and lack of starting material, there was no chance to set up other trials with different solvents and different pressure systems. Therefore, the final polymer **69** and **73** were not obtained. .

It is known that rate and selectivity of debenzylation reactions depend on the nature of the benzyl group, the basicity of substrate or product, steric and electronic effects, the type and amount of catalyst, solvents, modifiers and reducing agent. [20] Therefore there are possible reasons why the hydrogenolysis failed. The first may be the large molecular weight (4882.69 g/mol) of the synthesized glycopeptides **68** and **72** and the seven non-polar benzyl groups on their carbohydrate moiety which made these compounds insoluble in protic solvents such as methanol and ethanol. Therefore, to improve solubility, a mixture of THF/ethanol was used. According to the proposed mechanism (Figure 3.4), the protic solvent is very important in hydrogenation since it allows the formation of an

active complex between the compound and the solvent to accelerate the hydrogen transfer step. Thus, using a mixture of protic and aprotic solvents system may likely hinder the formation of an active catalytic complex and decrease the reaction conversion rate.

The second reason may be due to the state of metal catalyst in solution. In the solvents with low polarity, the optimal Pd dispersions are much lower than those observed in polar and protic solvents.[21] Therefore, the conversion rate will be decreased in the mixed THF/ethanol solvent system due to decreased polarity compared with using ethanol as solvent.

It was also suggested that the significant difference in the compound structure sensitivity could be attributed to the change in the steric effect of the Pd metal surface on the adsorption of the catalyst–substrate adduct which depends on the substrate structure and on the most stable catalyst conformation in different solvents.[22] Therefore, the compound structure and its spacial conformation in solution will contribute a lot to the debenzylation rate. Glycoprotein **68** and **72** contain both carbohydrate and peptide components and the seven benzyl ether groups on the carbohydrate moiety make them more soluble in aprotic solvent, whereas the metal catalyst needs protic solvent to give optimal dispersions. It is hard to find a solvent to meet both requirements, therefore, the debenzylation reaction was hard to achieve using a mixture of THF/ethanol and we may speculate that synthesized polymer **68** and **72** could not be debenzylated or partially debenzylated for this reason.

There are two possibilities to modify this whole synthetic route and to avoid debenzylation at the end of the synthesis. The first is to make change of compound **55** and the second is to make change of compound **63** and **64**. However, since the desired product of coupling reaction between **55** and L-threonine is an α *O*-linkage, it is very hard to choose an appropriate protecting group to replace an ether for compound **55**. Esters (i.e. acetates and benzoyls) are common protecting groups used for carbohydrate chemistry; however, a neighbouring group participation effect may take place and affect the product conformation. Furthermore, there is a potential risk during cleavage as they require strong basic conditions. A more appropriate route to obtain the final polymer would be to deprotect the benzyl groups before building the polymer, as the deprotection of benzyl groups in the monomeric building block structure **63** and **64** is more facile. However, since the *O*-linkage is sensitive to acidic and basic conditions,[23] the protecting group that would be used to replace the benzyl ether must be able to be removed by neutral conditions. Due to these limitations, only allyl ether and TBDMS ether are good candidates to replace the benzyl ethers (Scheme 3.6). The allyl ether can be removed under oxidative condition using NMO, $NaIO_4$ and OsO_4 at near neutral condition, in which the allyl group undergoes hydroxylation and subsequent periodate scission of the vicinal diol;[24] the TBDMS group can be removed by $CuCl_2 \bullet H_2O$ under catalytic conditions.[25]

Scheme 3.6 Possible cleavage methods for allyl ether and TBDMS groups

[1] Hudson, C. S.; Johnson, J. M., *J. Am. Chem. Soc.* **1915,** 37, 1270 - 1275.

[2] Wang, R., Steensma, D. H., Takaoka, Y., Yun, J. W., Kajimoto, T., Wong, C. H., *Bioorganic & Medicinal Chem.* **1997,** 5, 661 – 672.

[3] Lemieux, R. U., *Pure Appl. Chem.* **1971,** 25, 527 - 548

[4] Armand, S., Drouillard, S., Schulein, M., Henrissat, B., Driguez, H., *J. Biol. Chem.* **1997,** 272, 2709 – 2713.

[5] Tachibana, Y., Fletcher, G. L., Fujitani, N., Tsuda, S., Monde, K., Nishinura, S. I., *Angew. Chem. Int. Ed.* **2004,** 43, 856 - 862

[6] Yang, Y., Ficht, S., Brik, A., Wong, C., *J. Am. Chem. Soc.*, **2007,** 129, 7690 – 7701.

[7] Merrifield, R. B., *J. Am. Chem. Soc.,* **1963,** 85, 2149–2154.

[8] Boovanahalli, S. K., Kim, D. W., Chi, D. Y., *J. Org. Chem.,* **2004**, *69*, 3340-3344.

[9] Angibeaud, P., Defaye, J., Gadelle, A., Utille, J.P., *Synthesis,* **1985,** 1123 – 1125.

[10] Williams, D. R., Brown, D. L., Benbow, J. W., *J. Am. Chem. Soc.,* **1989,** 111, 1923 – 1925.

[11] Daly, S.M., Armstrong, R. W., *Tetrahedron Lett.,* **1989,** 30, 5713 – 5716.

[12] Rahim, M. A., Matsumura, S., Toshima,, K., *Tetrahedron Lett.*, **2005**, *46*, 7307-7309.

[13] Boons, G. J., Hale, K. J. *Organic Synthesis with Carbohydrates*, Blackwell Science: Malden, MA. **2000**, 156-160.

[14] Blaser, H.-U., Indolese, A., Schnyder, A., Steiner, H., Studer, M., *J. Mol. Cata. A: Chemical,* **2001**, 173, 3–18.

[15] Hawker, S.,Bhatti, M. A., Griffin, K. G., *Chim. Oggi,.* **1992,** 10, 49.

[16] Pantulu, A. J., Achaya, K. T., *J. Am. Oil Chem. Soc.* **1964**, 41, 511 – 514.

[17] Boons, G. J., Hale, K. J. *Organic Synthesis with Carbohydrates*, Blackwell Science: Malden, MA. **2000**, 156-160.

[18] Tachibana, Y., Fletcher, G. L., Fujitani, N., Tsuda, S., Monde, K., Nishinura, S. I., *Angew. Chem. Int. Ed.* **2004,** 43, 856 – 862

[19] Yao, H.- C., Emmett, P. H., *J. Am. Chem. Soc.,* **1959,** 81, 4125–4132.

[20] Blaser, H.-U., Indolese, A., Schnyder, A., Steiner, H., Studer, M., *J. Mol. Cata. A: Chemical,* **2001,**173, 3–18.

[21] Nitta, Y., Kubota, T., Okamoto, Y., *J. Mol. Cata A: Chemical,* **2004,** 212, 155 – 159.

[22] Blaser, H.-U., Indolese, A., Schnyder, A., Steiner, H., Studer, M., *J. Mol. Cata. A: Chemical,* **2001,**173, 3–18.

[23] Boons, G. J., Hale, K. J. *Organic Synthesis with Carbohydrates*, Blackwell Science: Malden, MA. **2000**, 156-160.

[24] Kitov, P. I., Bundle, D. R., *Org. Lett.,* **2001**, *3*, 2835-2838.

[25] Tan, Z. P., Wang, L., Wang, J. B., *Chin. Chem. Lett.* **2000,** 11, 753 - 756.

4 Experimental: Materials and Methods

General experimental information

All anhydrous reactions were performed in flame-dried round bottom flasks under argon or nitrogen atmosphere. All air sensitive reagents and anhydrous solvents were transferred with oven-dried syringes or through cannulation. All analytical thin layer chromatography (TLC) was performed with 0.2mm pre-coated silica gel plates 60 F_{254} (E. Merck). All flash chromatography was performed with Silicycle Ultrapure SiliaFlash®F60 (230-400 mesh).

The reactions performed under dry conditions were using distilled solvents unless aqueous systems. Dry dichloromethane (DCM) and dry acetonitrile were distilled over calcium hydride. *N*, *N*-dimethylformamide (DMF) was dried over activated molecular sieves and stored under argon atmosphere. Diisopropylethylamine (DIPEA) was distilled from calcium hydride. All solution phase reactions were performed under normal atmosphere unless stated specify.

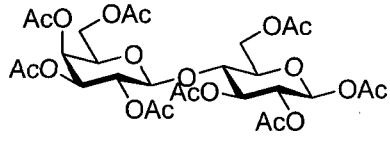
Solution phase reactions were monitored by thin layer chromatography (TLC) and visualized by illumination with a short-wavelength UV light (254nm) or stained with ceric ammonium molybdate (CAM) or Potassium permanganate ($KMnO_4$) stain. Purification was done by flash chromatography. All reverse phase HPLC purifications were done by using Varian (model 330) HPLC system equipped with a variable wavelength UV detector.

^{1}H and ^{13}C NMR experiments were performed by using Bruker Avance 400 MHz, Bruker Avance 500 MHz or INOVA 500 MHz spectrometers. Deuterated chloroform or deuterated methanol was used as NMR experiment solvents, unless otherwise stated. Chemical shifts were reported in ppm and were corrected using NMR solvent peak as a reference. ^{1}H NMR signal patterns are abbreviated as follows: s - singlet, d - doublet, t - triplet, q - quartet, m - multiplet, dd - doublet of doublet, td - triplet of doublet. Mass spectra were reported by using Electrospray Ionization (ESI) and Matrix Assisted Laser Desorption Ionization (MALDI-TOF) mass spectrometer.

Solid phase peptide syntheses (SPPS) were performed using Fmoc-Ala-Wang resin (Nova Biochem) in solid phase peptide synthesis machine (Chem Tech).

4.1 Preparation of the Carbohydrate Components

Preparation of 1,2,3,6-Tetra-*O*--acetyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-D-glucopyranose (60)



This compound was prepared according to the previously published procedure. [1] Sodium acetate (1.3 g, 0.016 mol) was added into acetic anhydride solution (50 mL) and refluxed at 100°C for 10 minutes, and then added β-lactose (5.13 g, 0.015 mol) into reaction mixture and continuous refluxing the reaction mixture at 115°C for about 2

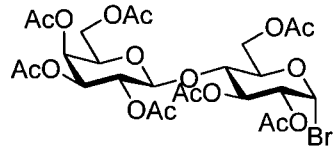
hours. Diluted the reaction mixture with cold water, and then extracted with CH_2Cl_2 (3X), washed aqueous phase with CH_2Cl_2. Combined organic phases combined and washed with water, sat. $NaHCO_3$ and brine, dried over $MgSO_4$ and concentrated. Crystallized from ether to afford white powder and the average yield was 90%. The ratio of β and α form is 10 to 1. NMRs of β anomer are listed below.

^{1}H NMR (400 MHz, $CDCl_3$) δ ppm 5.67 (d, J = 8.27 Hz, 1H), 5.34 (dd, J = 3.38, 0.78 Hz, 1H), 5.24 (t, J = 9.14, 9.14 Hz, 1H), 5.14-5.01 (m, 2H), 4.94 (dd, J = 10.41, 3.45 Hz, 1H), 4.46 (td, J = 10.95, 3.15, 3.15 Hz, 2H), 4.10 (m, 3H), 3.86 (m, 2H), 3.75 (ddd, J = 9.96, 4.79, 2.00 Hz, 1H, H4), 2.15 (3H), 2.12 (3H), 2.09 (s, 3H), 2.06 (3H), 2.04 (3H), 2.03(3H), 2.02(3H), 1.96 (s, 3H)

^{13}C NMR (101 MHz, $CDCl_3$) δ ppm 170.3, 170.3, 170.1, 170.0, 169.6, 169.5, 169.0, 168.8, 100.9, 91.5, 75.7, 73.5, 72.6, 70.9, 70.7, 70.4, 69.0, 66.6, 61.7, 60.8, 20.7, 20.7, 20.6, 20.5

MS (ESI): Calcd for $C_{28}H_{38}O_{19}$ *m/z* ($M+Na^+$) 701.59, found 701.5.

Preparation of 2,3,6-Tri-*O*-acetyl-4-O-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-α-D-glucopyranosyl bromide (58)



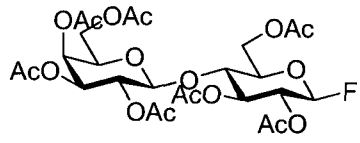
This compound was prepared according to the previously published procedure. [2] Peracetate lactose (2.0 g, 2.95 mmol) was dissolved into 3 mL of CH_2Cl_2 and then added 20 mL hydrogen bromide solution (33 wt. % in acetic acid) and cooled the reaction mixture to 0°C. The reaction was stirred at 0°C for 30 minutes and then allowed to warm up to room temperature slowly and stirred for another 2 hours. Following added reaction mixture into large amount of ice cold water, extracted the aqueous layer with CH_2Cl_2 (3X). Organic phases were combined, washed with water, $NaHCO_3$, and brine, dried over $MgSO_4$ and concentrated. Recrystallized from cold ethyl ether to afford white powder in almost quantitative yield. This product was used for farther reactions without further purification.

^{1}H NMR (400 MHz, $CDCl_3$) δ ppm 6.44 (d, J = 4.00 Hz, 1H) 5.67 (d, J = 8.27 Hz, 1H), 5.34 (dd, J = 3.38, 0.78 Hz, 1H), 5.24 (t, J = 9.14, 9.14 Hz, 1H), 5.14-5.01 (m, 1H), 4.94 (dd, J = 10.41, 3.45 Hz, 1H), 4.46 (td, J = 10.95, 3.15, 3.15 Hz, 2H), 4.10 (m, 4H), 3.81 (m, 2H), 2.15 (s, 3H), 2.12 (s, 3H), 2.09 (s, 3H), 2.06 (s, 3H), 2.04 (s, 3H), 2.03(s, 3H), 2.02(s, 3H), 1.96 (s, 3H)

^{13}C NMR (101 MHz, $CDCl_3$) δ ppm 170.3, 170.3, 170.1, 170.0, 169.6, 169.5, 169.0, 168.8, 100.9, 91.5, 75.7, 73.5, 72.6, 70.9, 70.7, 70.4, 69.0, 66.6, 61.7, 60.8, 20.7, 20.7, 20.6, 20.5

MS (ESI): Calcd for $C_{26}H_{35}BrO_{17}$ *m/z* ($M+NH_4^+$) 992, found 992.6

Preparation of 2,3,6-Tri-*O*-acetyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl fluoride (57)



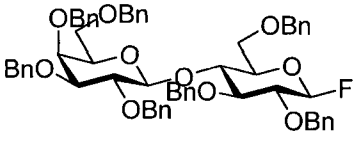
This compound was prepared according to the previously published procedure. [3] Flame dried the round bottom flask and the hepta-*O*-acetyllactosyl bromide (1.4 g, 0.002 mol) was dissolved in anhydrous acetonitrile (20 mL) under nitrogen atmosphere. Silver fluoride (0.52 g, 0.004 mol) was added into reaction mixture and stirred for overnight at room temperature under nitrogen atmosphere. The reaction mixture was concentrated under reduced pressure. The residue was re-dissolved in 20 mL ethyl acetate and the precipitate was filtered off by celite. The hepta-*O*-acetyllactosyl fluoride product was purified by flash chromatography on silica gel (ethyl acetate / hexane, 1:1 v/v) and recrystallized from cold ethyl ether in about 65% yield.

^{1}H NMR (400 MHz, *$CDCl_3$*) δ ppm 5.27 (dd, *J* = 52.73, 5.59 Hz, 1H), 5.22 (d, *J* = 3.08 Hz, 1H), 5.07 (t, *J* = 7.64, 7.64 Hz, 1H), 4.98 (dd, *J* = 10.37, 7.91 Hz, 1H), 4.87 (td, *J* = 6.97, 4.64, 4.64 Hz, 2H), 4.46 (d, *J* = 7.88 Hz, 1H), 4.42 (dd, *J* = 11.98, 1.80 Hz, 1H), 3.98 (m, 3H), 3.85 (td, *J* = 13.49, 8.04, 8.04 Hz, 2H), 3.76 (m, 1H), 2.02 (s, 3H), 2.00 (s, 3H), 1.97 (s, 3H), 1.96 (s, 3H), 1.93 (s, 3H), 1.92 (s, 3H), 1.84 (s, 3H)

^{13}C NMR (400 MHz, *$CDCl_3$*) δ ppm 170.0, 169.8, 169.7, 169.2, 169.0, 168.8, 106.5, 104.4, 100.8, 77.2, 74.9, 72.1, 71.8, 70.9, 68.7, 66.4, 61.5, 60.6, 20.4, 20.4, 20.3, 20.2, 20.2

MS (ESI): Calcd for $C_{26}H_{35}FO_{17}$ *m/z* ($M+NH_4^+$) 656.19, found 656.4.

Preparation of 2,3,6-Tri-*O*-benzyl-4-*O*-(2,3,4,6-tetra-*O*-benzyl-β-D-galactopyranosyl)-β-D-glucopyranosyl fluoride (55)



To a flask of β-lactose fluoride (1.21 g, 1.39 mmol) added 30ml NaOMe solution (pH =8.0), and stirred at 0°C for about 30 minutes, reaction was modified by TLC (ethyl acetate/hexane, 1:1, v/v) until completed. Then the reaction mixture was neutralized with the Amberlite120 [H^+] resin, filtered, concentrated under reduced pressure. The residue was directly benzylated without further characterization in an average 63% yield.

Dissolved the concentrated deprotected residue in DMF (50 mL) at 0°C, added NaH (0.669 g, 27.9 mmol) and stirred reaction mixture for 30 minutes under nitrogen atmosphere. Benzyl bromide (4.87 mL, 40.9 mmol) was slowly added into reaction mixture and stirred for about 10 minutes, and then the reaction mixture was allowed to slowly warm up to room temperature and stirred overnight under nitrogen. Reaction mixture was diluted with large amount of water, extracted with $CHCl_3$ (3X), and the organic phases were combined, washed with water (3X), 10% HCl, saturated $NaHCO_3$, brine, dried over $MgSO_4$ and concentrated. The residue was purified by flash column chromatography on silica to afford the product in an average 63% yield (Hexane : ethyl ether, 7:3 v/v). [4]

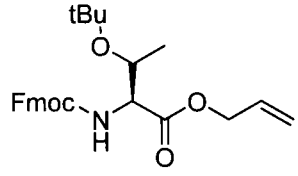
^{1}H NMR (400MHz, *$CDCl_3$*) δ ppm 7.52 – 7.29 (m, 35H) 5.40 (dd, *J* = 53.29, 6.42 Hz, 1H), 5.20 (m, 2H), 4.91 (m, 7H), 4.72 (m, 2H), 4.55 (m, 3H), 4.43 (d, *J* = 11.82 Hz, 1H), 4.24 (t, *J* = 9.22 Hz, 1H), 4.08 (d, *J* = 2.76 Hz, 1H), 4.01 (dd, *J* = 10.94, 3.62 Hz, 1H), 3.94 (dd, *J* = 9.64, 7.75 Hz, 1H)3.84 (dd, *J* = 10.87, 1.71 Hz, 1H), 3.69 (m, 4H), 3.55 (m, 3H)

^{13}C NMR (400MHz, *$CDCl_3$*) δ ppm 138.9, 138.7, 138.6, 138.336, 137.9, 137.9, 137.8, 128.2, 128.2, 128.1, 128.0, 127.9, 127.7, 127.7, 127.6, 127.6, 127.4, 127.3, 127.1, 110.6, 108.4, 102.5, 82.3, 81.5, 81.4, 80.8, 80.6, 79.8, 77.2, 75.6, 75.2, 74.9, 74.8, 74.6, 74.4, 73.3, 73.1, 73.0, 72.4, 71.9, 68.0, 67.6

MS (ESI): Calcd for $C_{61}H_{63}FO_{10}$ *m/z* ($M+NH_4^+$) 992.47, found 992.5

4.2 Preparation of the Peptide Components

Preparation of (2*S*,3*R*)-allyl 2-(((9H-fluoren-9-yl)methoxy)carbonylamino)-3-tert-butoxybutanoate (62)



Fmoc-threonine (240 mg, 0.6 mmol) was dissolved in 10 mL dry CH_2Cl_2 under argon, HBTU was added (532.3 mg, 0.72 mmol) into reaction mixture and stirred for 5 minutes under argon. Following added allyl alcohol (0.21 mL, 3 mmol) into the reaction mixture

and stirred for another 15 minutes, and then added DIPEA (0.2 mL, 1.2 mmol) into the reaction mixture and allow stirred for 4 hours under argon.

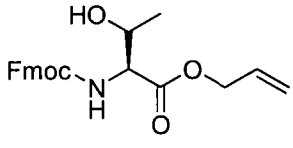
Diluted reaction mixture with 30ml CH_2Cl_2, extracted with saturated NH_4Cl solution (3X), and washed aqueous phase with CH_2Cl_2. Combined with organic phases and washed with water, brine, dried over $MgSO_4$, concentrated. The Fmoc-threonine allyl ester was purified by flash chromatography on silica gel (hexane: ethyl acetate = 4:1 v/v) to afford white foam, the average yield was 64%.

^{1}H NMR (400 MHz, *$CDCl_3$*) δ ppm 7.78 (d, *J* = 7.45 Hz, 2H), 7.66 (t, *J* = 7.57, 7.57 Hz, 2H), 7.38 (td, *J* = 30.49, 7.43, 7.43 Hz, 4H), 5.95 (tdd, *J* = 16.32, 10.51, 5.84, 5.84 Hz, 1H), 5.66 (d, *J* = 9.66 Hz, 1H), 5.38 (dd, *J* = 17.17, 1.36 Hz, 1H), 5.28 (dd, *J* = 10.38, 1.02 Hz, 1H), 4.64 (ddd, *J* = 52.46, 13.16, 5.83 Hz, 1H), 4.35 (m., *5*H), 1.28 (d, *J* = 5.45Hz, 3H), 1.17 (s, 9H)

^{13}C NMR (400 MHz, *$CDCl_3$*) δ ppm 170.7, 156.7, 143.9, 143.7, 141.2, 131.5, 127.6, 127.0 126.9, 125.1, 125.0, 119.8, 118.8, 74.0, 67.3, 67.1, 65.9, 59.8, 47.1, 28.3, 28.2, 20.8

MS (ESI): Calcd for $C_{26}H_{31}NO_5$ *m/z* ($M+Na^+$) 460.5, found 460.5

Preparation of (2*S*,3*R*)-allyl 2-(((9H-fluoren-9-yl)methoxy)carbonylamino)-3-hydroxybutanoate (56)



Dissolved 150 mg Fmoc-threonine allyl ester into TFA/water solution (10 mL, 1:1 v/v) and stirred overnight. Extracted reaction mixture with 20 mL ethyl acetate (3X) and combined organic phases. And then organic phase was washed with water, sat. $NaHCO_3$, brine, dried over $MgSO_4$ and concentrated under reduced pressure. The product was purified by flash chromatography on silica gel (hexane: ethyl acetate = 3:1 v/v) to afford of a white foam.

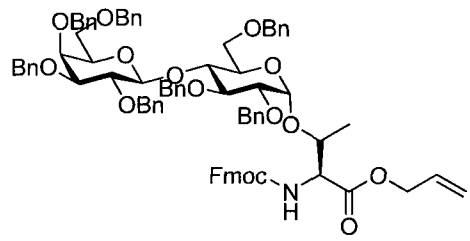
^{1}H NMR (400 MHz, *$CDCl_3$*) δ ppm 7.74 (d, *J* = 7.57 Hz, 2H), 7.60 (dd, *J* = 7.04, 3.71 Hz, 2H), 7.33 (td, *J* = 36.55, 7.47, 7.47 Hz, 4H), 5.87 (m, 2H), 5.32 (d, *J* = 17.20 Hz, 1H), 5.23 (dd, *J* = 10.43, 1.13 Hz, 1H), 4.65 (d, *J* = 5.60 Hz, 2H), 4.40 (m, 4H), 4.22 (t, *J* = 7.06, 7.06 Hz, 1H), 2.81 (s, 1H), 1.25 (d, *J* = 6.16 Hz, 3H)

^{13}C NMR (400 MHz, *$CDCl_3$*) δ ppm 170.9, 156.8, 143.8, 141.2, 131.3, 127.6, 127.0, 125.0, 119.9, 118.8, 67.9, 67.2, 66.1, 59.2, 47.0, 19.8,

MS (ESI): Calcd for $C_{22}H_{23}NO_5$ *m/z* ($M+Na^+$) 427.42, found 427.2

4.3 General Protocol for the Preparation of Building Blocks

Preparation of 2,3,6-Tri-*O*-benzyl-4-*O*-(2,3,4,6-tetra-*O*-benzyl-β-D-galactopyranosyl)-β-D-glucopyranosyl-L-N-(9-Fluorenylmethoxycarbonyloxy)-threonine allyl ester (63)



This compound was prepared according to the previously published procedure. [5] Flame dried two round bottom flasks. A mixture of $SnCl_2$ (134mg, 0.71 mmol), $AgClO_4$ (147.2 mg, 0.71 mmol), Fmoc and allyl protected threonine (406.2 mg, 1.065 mmol) and powdered molecular sievers 4Å (800 mg) were stirred in 20ml dry CH_2Cl_2 at room temperature for 10 minutes , and then cooled to -20°C and stirred for another 10 minutes under argon atmosphere. Glucopyranosyl fluoride (303.8 mg, 0.35 mmol) was dissolved in 5mL dry CH_2Cl_2 under argon atmosphere in another flask and cannulated into the reaction mixture and stirred for about 30 minutes at -20°C. The reaction mixture was allowed to slowly warm up to room temperature and stirred for 24 hours.

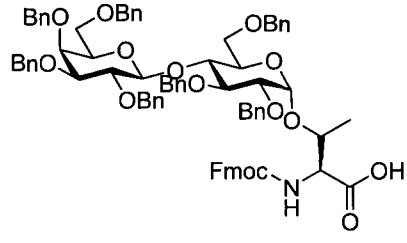
Reaction mixture was diluted with $CHCl_3$, filtered through celite then extracted with water(3X), the organic phases were combined and washed with sat. $NaHCO_3$, brine, dried over $MgSO_4$ and concentrated under reduced pressure. The crude residue was purified by flash column chromatography (petroleum ether:ethyl acetate = 2:1 v/v) to give product as a 2:1 mixutre of α:β isomers. Both anomers were separated by ingredient flash column chromatography (from petroleum ether:ethyl acetate = 7:2 to petroleum ether:ethyl acetate 2:1 v/v). The overall yield of both anomers is about 50%. NMRs of α anomer are listed below.

^{1}H NMR (500 MHz, *$CDCl_3$*) δ ppm 7.73 (d, J = 7.47 Hz, 2H), 7.62 (t, J = 6.92, 6.92 Hz, 2H), 7.29 (m., 39H), 6.32 (d, *J* = 7.64 Hz, 1H), 5.88 (ddd, *J* = 16.47, 11.35, 5.87 Hz, 1H), 5.27 (ddd, *J* = 13.81, 11.22, 1.04 Hz, 2H) 5.13 (d, *J* = 10.66 Hz, 1H), 5.00 (d, *J* = 11.48 Hz, 1H), 4.76 (m, 7H), 4.56 (m, *6*H), 4.46 (dd, *J* = 10.63, 7.16 Hz, 1H), 4.35 (m 6H), 4.24 (m, *2H*), 3.90 (m, *4*H), 3.78 (m, *2*H), 3.49 (m, *3*H), 3.36 (m, *3*H), 1.36 (d, *J* = 6.23 Hz, 3H)

^{13}C NMR (500 MHz, *$CDCl_3$*) δ ppm 170.4, 156.9, 143.9, 143.8, 141.3, 139.4, 139.1, 138.6, 138.2, 138.0, 131.6, 128.4, 128.3, 128.2, 128.0, 127.9, 127.9, 127.8, 127.7, 127.6, 127.5, 127.4, 127.1, 125.2, 125.1, 120.0, 118.8, 102.9, 98.3, 82.5, 78.4, 75.4, 75.3, 74.7, 73.6, 73.6, 73.4, 73.2, 73.1, 72.5, 71.0, 68.2, 67.2, 66.2, 59.3, 47.2, 29.9, 29.8

MS (ESI): Calcd for $C_{83}H_{85}NO_{15}$ *m/z* ($M+K^+$) 1375.66, found 1374.5

Preparation of (2*S*,3*R*)-2-(((9H-fluoren-9-yl)methoxy)carbonylamino)-3-((2*S*,3*R*,4*S*,5*R*,6*R*)-3,4-bis(benzyloxy)-6-(benzyloxymethyl)-5-((2*S*,3*R*,4*S*,5*S*,6*R*)-3,4,5-tris(benzyloxy)-6-(benzyloxymethyl)tetrahydro-2H-pyran-2-yloxy)tetrahydro-2H-pyran-2-yloxy)butanoic acid (54)



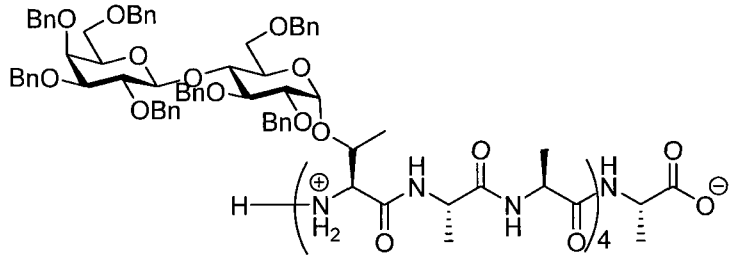
This compound was prepared according to the previously published procedure.[6] Coupling product (500 mg, 0.37 mmol) was dissolved into 10 mL THF, *N*-methylaniline (0.4 μLl, 3.7 mmol) and $Pd(PPh_3)_4$ (40 mg, 0.037 mmol) were added into reaction mixture and stirred for 45 minutes under argon atmosphere, then the reaction mixture was concentrated under reduced pressure. The deprotected product was purified by flash column chromatography (hexane: ethyl acetate = 2: 1 v/v followed by 5% MeOH in CH_2Cl_2).

^{1}H NMR (500 MHz, *$CDCl_3$*) δ ppm 7.71 (d, *J* = 7.27 Hz, 2H), 7.58 (t, *J* = 6.83, 6.83, 2H), 7.25 (m, 39H), 5.12 (d, *J* = 10.67 Hz, 1H), 4.97 (d, *J* = 11.24 Hz, 2H), 4.73 (m, 7H), 4.52 (dd, *J* = 22.65, 11.66 Hz, 2H), 4.30 (m, 9H), 3.83 (m, 7H), 3.47 (m, 3H), 3.33 (dd, *J* = 5.08, 4.23 Hz, 3H), 1.22 (d, *J* = 5.82 Hz, 3H)

^{13}C NMR (300 MHz, *$CDCl_3$*) δ ppm 143.8, 143.7, 141.2, 141.193, 139.1, 139.0, 138.7, 138.5, 138.1, 137.9, 137.0, 128.3, 128.1, 127.8, 127.6, 127.4, 125.0, 119.9, 102.8, 97.6, 96.6, 96.3, 82.4, 80.2, 79.9, 75.3, 75.2, 74.6, 74.4, 73.6, 73.4, 73.1, 73.1, 72.5, 71.1, 68.2, 67.6, 67.1, 47.1

MS (ESI): Calcd for $C_{80}H_{81}NO_{15}$ *m/z* ($M+K^+$) 1335.59, found 1343.5

4.4 Solid Phase Synthesis of AFGP Analogues



68

Both polypeptides are prepared by linear solid phase synthesis starting from Wang resin preloaded with Fmoc-alanine (0.1 g, 0.06 mmol active site) by using Chem Tech solid phase peptide synthesizer. The resin was swollen in dry DMF (3 mL) in 1 hour, and then deprotected the Fmoc group by adding 20% piperidine in DMF solution and reaction was shaken for 30 minutes followed by washing the resin with DMF for 3 times. Previously dissolved Fmoc-alanine-OH (3 equiv.) and HCTU (1.1 equiv.) in DMF were added into solid phase synthesizer, following added DIPEA (7.5 equiv.) in DMF into the synthesizer and shaken the synthesizer for 2 hours. Repeated this coupling reaction one more time, and then deprotected the Fmoc group on the alanine residue by 20% piperidine in DMF solution. Then repeated the same coupling reaction by adding same equivalents of Fmoc-alanine-OH, HCTU, DIPEA into the synthesizer, and the same coupling reaction was repeated two more times to install one more alanine residue onto the synthesized sequence. The terminal fmoc group was removed by 20% piperidine in DMF solution. Except of HCTU, DIPEA, the following coupling reaction was done with the synthesized glycopeptides building block instead of the commercially available Fmoc-alanine-OH, reaction mixture was shaken for 30 hours. After removing the

terminal Fmoc group of the building block, the whole coupling cycle was finished. The whole coupling process was repeated three more times to afford the desired glycopeptides attached on the resin. The last step was washed the resin with 5ml dry CH_2Cl_2 for 3 minutes and repeated the wash process for 3 times.

The resin was removed from the synthesizer to a vial and treated with 30% TFA in CH_2Cl_2 solution for 1 hrs. The solution was filtered and the resin was washed with 30% TFA in CH_2Cl_2 and CH_2Cl_2. All filtrate was collected and concentrated under reduced pressure. Cold Et_2O was added into the residue to afford white precipitate and the precipitate was dried under vacuum.

^{1}H NMR (400 MHz, *CDCl$_3$, MeOD*) δ ppm 7.20 (m, 33H), 5.11-5.01 (m, 2H), 4.93-4.83 (m, 1H), 4.77-4.51 (m, 6H), 4.50-4.36 (m, 3H), 4.37-4.10 (m, 6H), 4.08-3.93 (m, 2H), 3.91-3.58 (m, 6H), 3.55-3.19 (m, 6H), 1.37-1.15 (m, 4H), 1.14-0.99 (m, 3H), 0.98-0.76 (m, 3H)

^{13}C NMR (300 MHz, *CDCl$_3$, MeOD*) δ ppm 159.8, 154.0, 152.1, 151.0, 145.1, 143.5, 138.9, 137.9, 134.1, 131.6, 128.8, 127.9, 127.0, 126.9, 124.2, 120.0, 117.0, 115.4, 104.2, 104.1, 103.0, 102.8, 99.9, 93.9, 82.8, 82.7, 82.4, 81.7, 81.7, 79.7, 78.0, 77.8, 77.0, 76.5, 76.2, 75.8, 75.6, 74.7, 74.6, 73.5, 73.4, 73.4, 73.3, 73.1, 73.0, 72.9, 72.7, 72.7, 72.6, 72.5, 68.0, 65.0, 49.3, 49.2, 49.1, 39.9, 34.8, 34.1, 24.6, 22.9, 17.1, 17.0, 16.0, 15.9

MS (MALDI): Calcd for $C_{287}H_{323}N_{13}O_{58}$ *m/z* ($M+NH_3$) 4896.27, found 4895.9

---

[1] Hudson, C. S.; Johnson, J. M., *J. Am. Chem. Soc.* **1915,** *37,* 1270 - 1275.

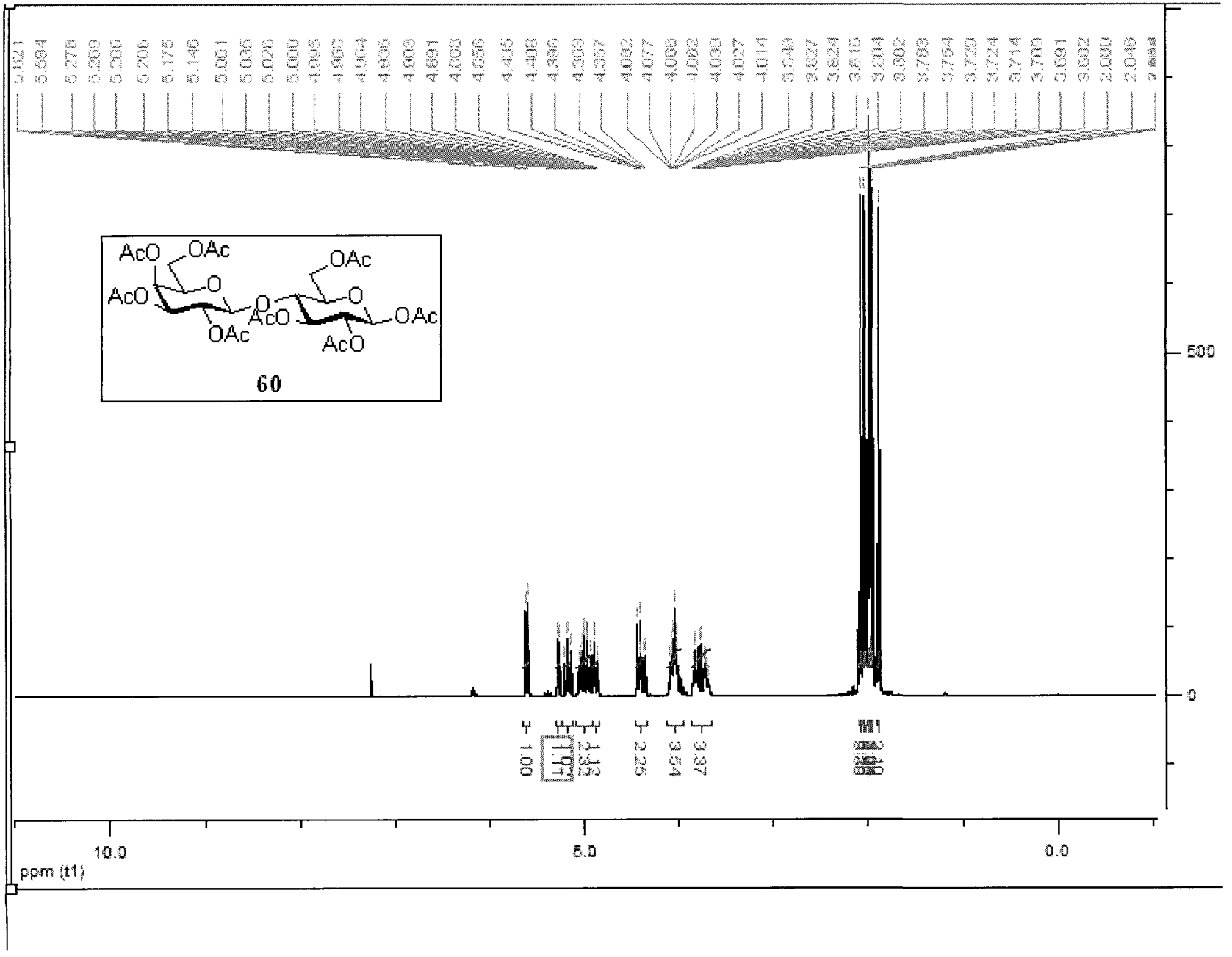
[2] Wang, R., Steensma, D. H., Takaoka, Y., Yun, J. W., Kajimoto, T., Wong, C. H., *Bioorganic & Medicinal Chem.* **1997,** *5*, 661 – 672.

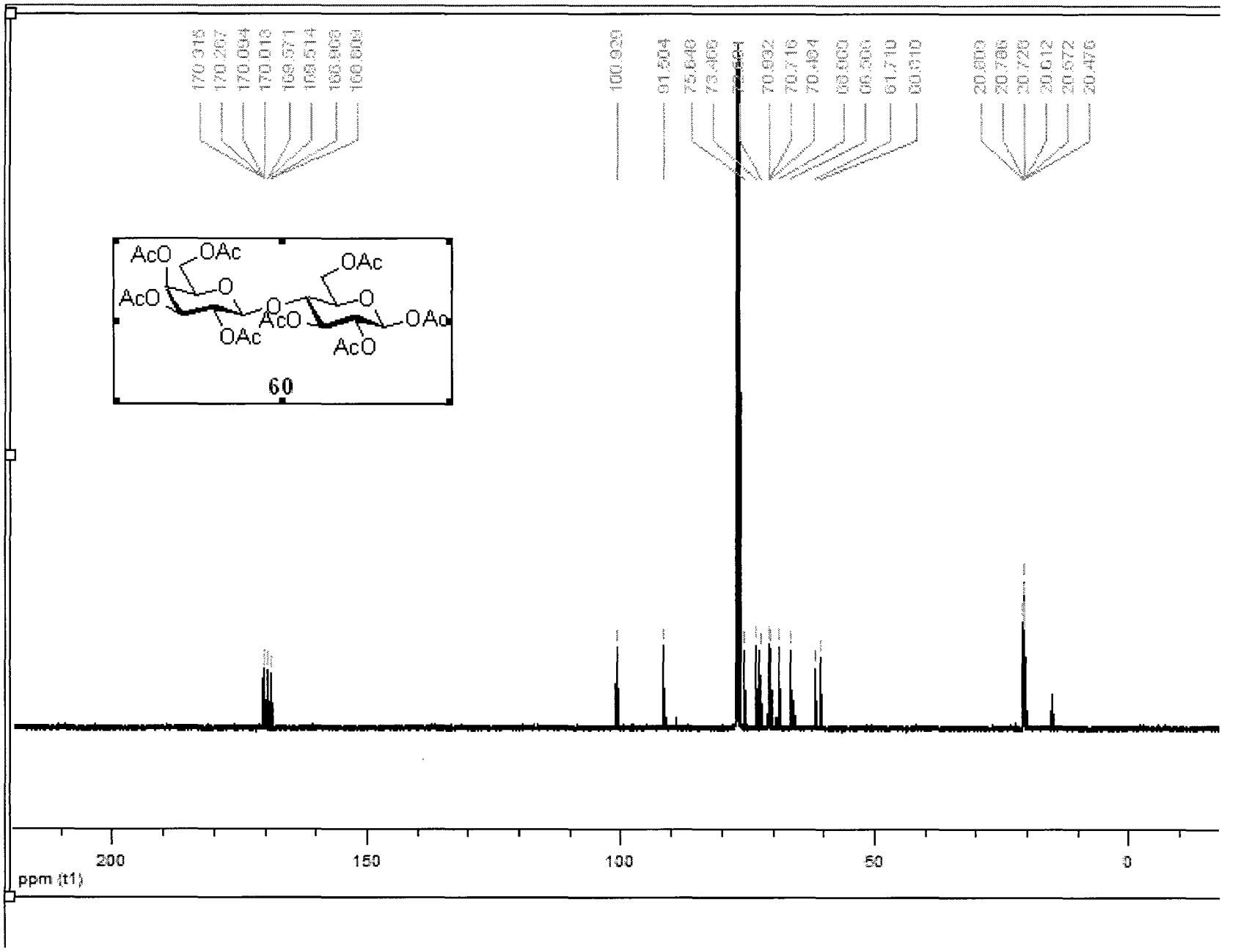
[3] Armand, S., Drouillard, S., Schulein, M., Henrissat, B., Driguez, H., *J. Biol. Chem.* **1997,** *272,* 2709 – 2713.

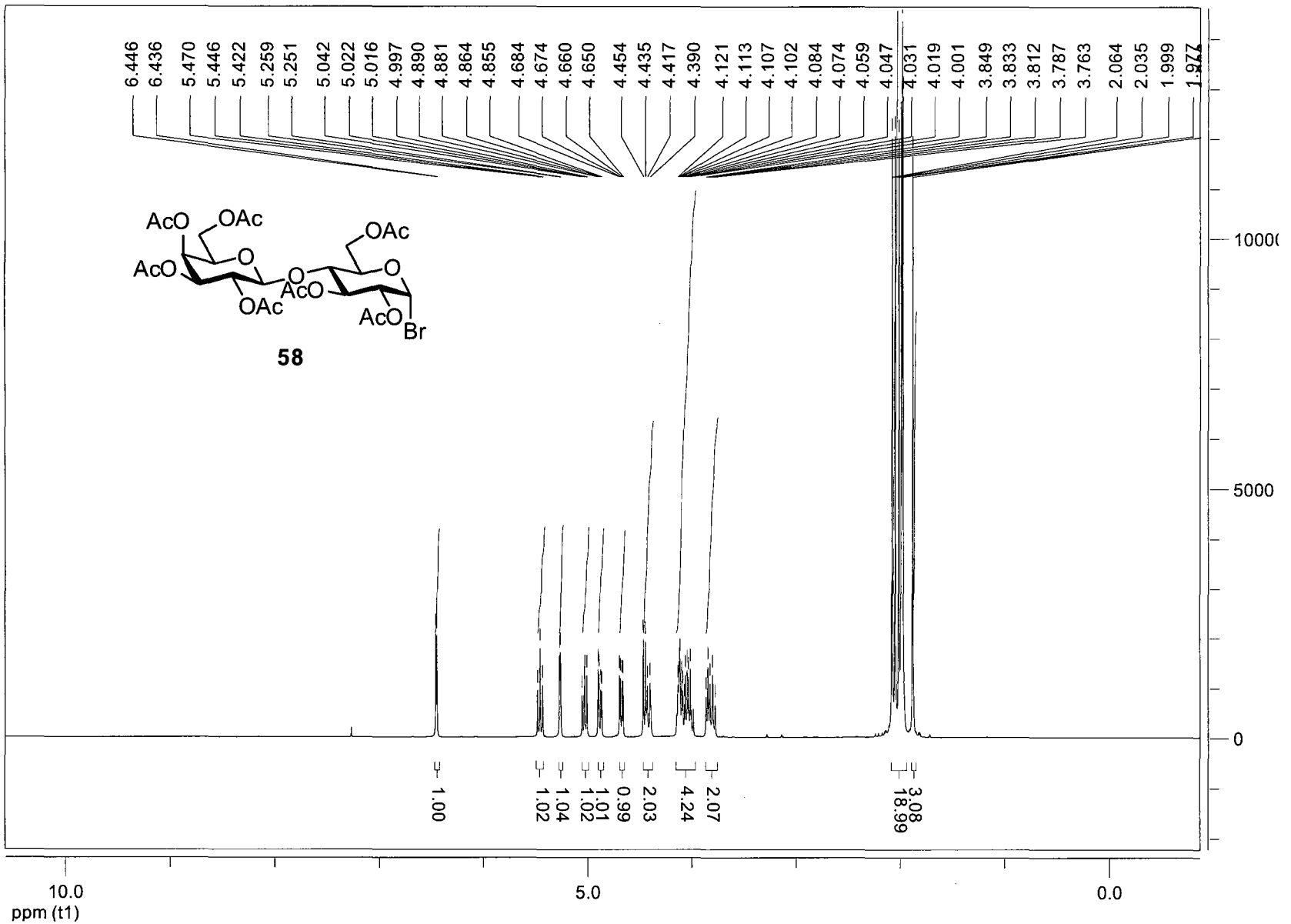
[4] Armand, S., Drouillard, S., Schulein, M., Henrissat, B., Driguez, H., *J. Biol. Chem.* **1997,** *272,* 2709 - 2713

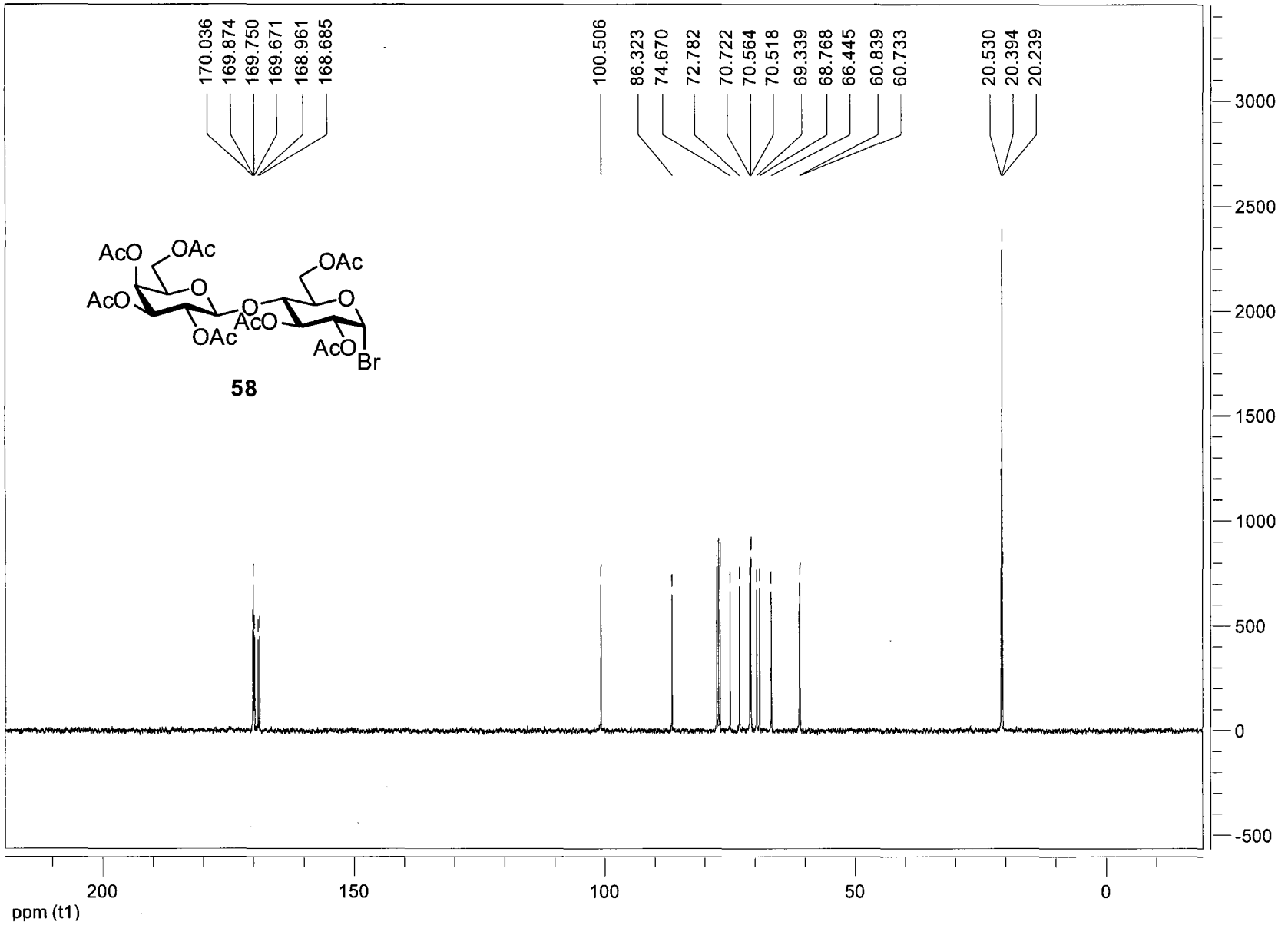
[5] Tachibana, Y., Fletcher, G. L., Fujitani, N., Tsuda, S., Monde, K., Nishinura, S. I., *Angew. Chem. Int. Ed.* **2004,** *43,* 856 - 862

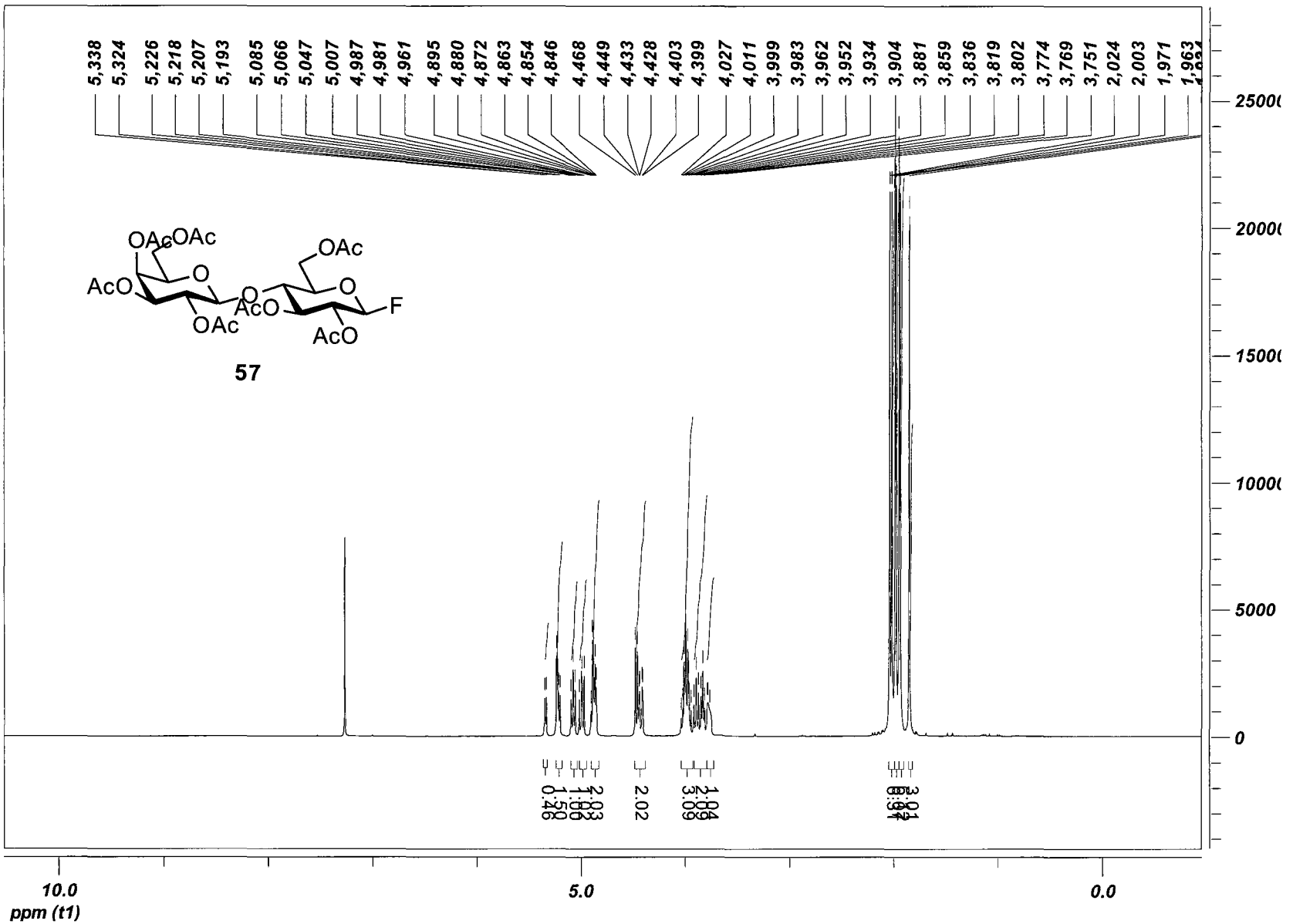
[6] Yang, Y., Ficht, S., Brik, A., Wong, C., *JACS*, **2007,** 129, 7690 - 7701

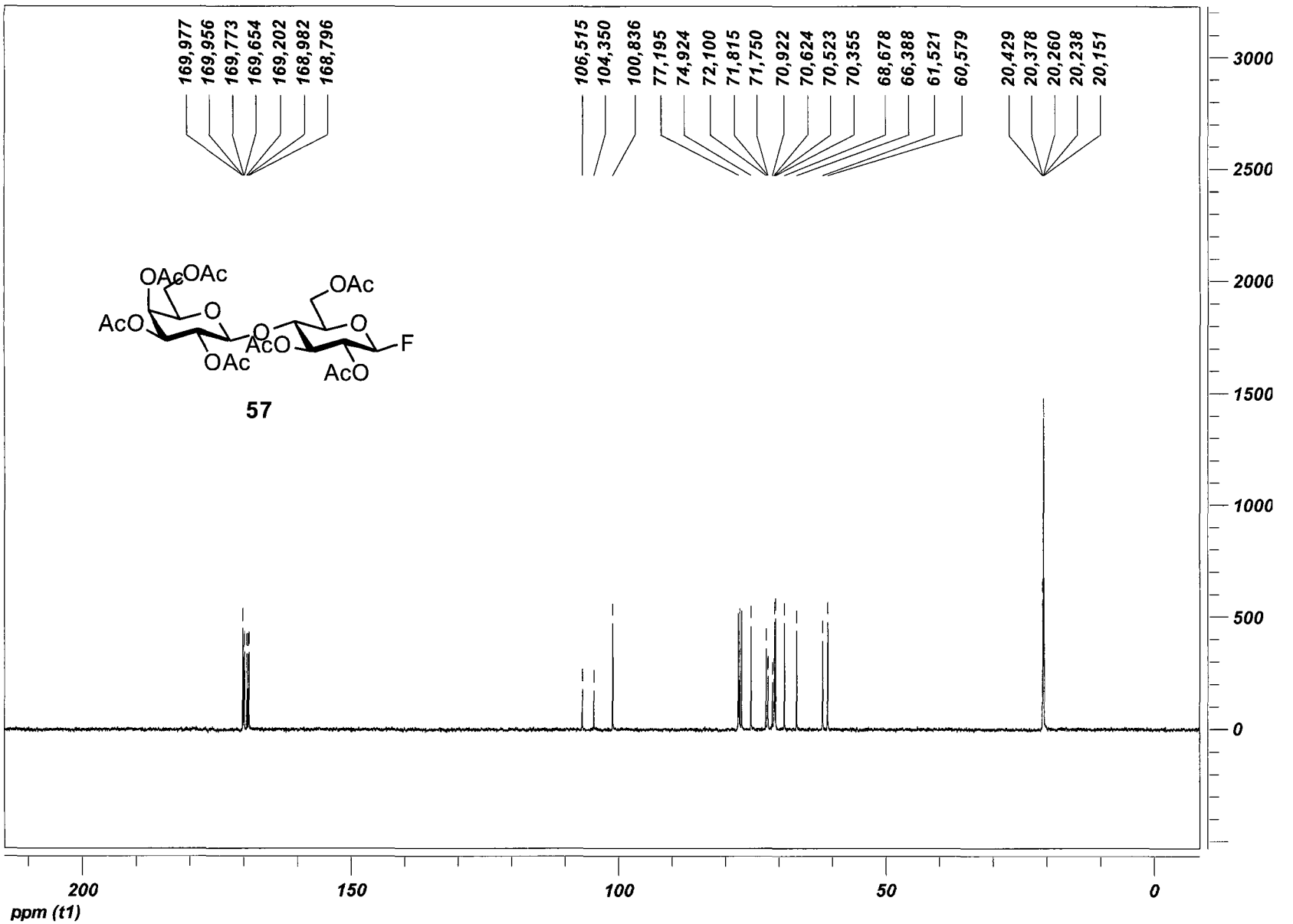

60
500
0
10.0
5.0
0.0
ppm (t1)


170.315
170.267
170.094
170.018
169.971
169.514
169.005
168.889
100.926
91.504
75.640
73.405
70.932
70.716
70.494
68.460
68.206
61.710
60.810
20.680
20.786
20.728
20.612
20.572
20.476
AcO
OAc
AcO
O
OAc
O
AcO
O
AcO
OAc
OAc
60
200
150
100
50
0
ppm (t1)

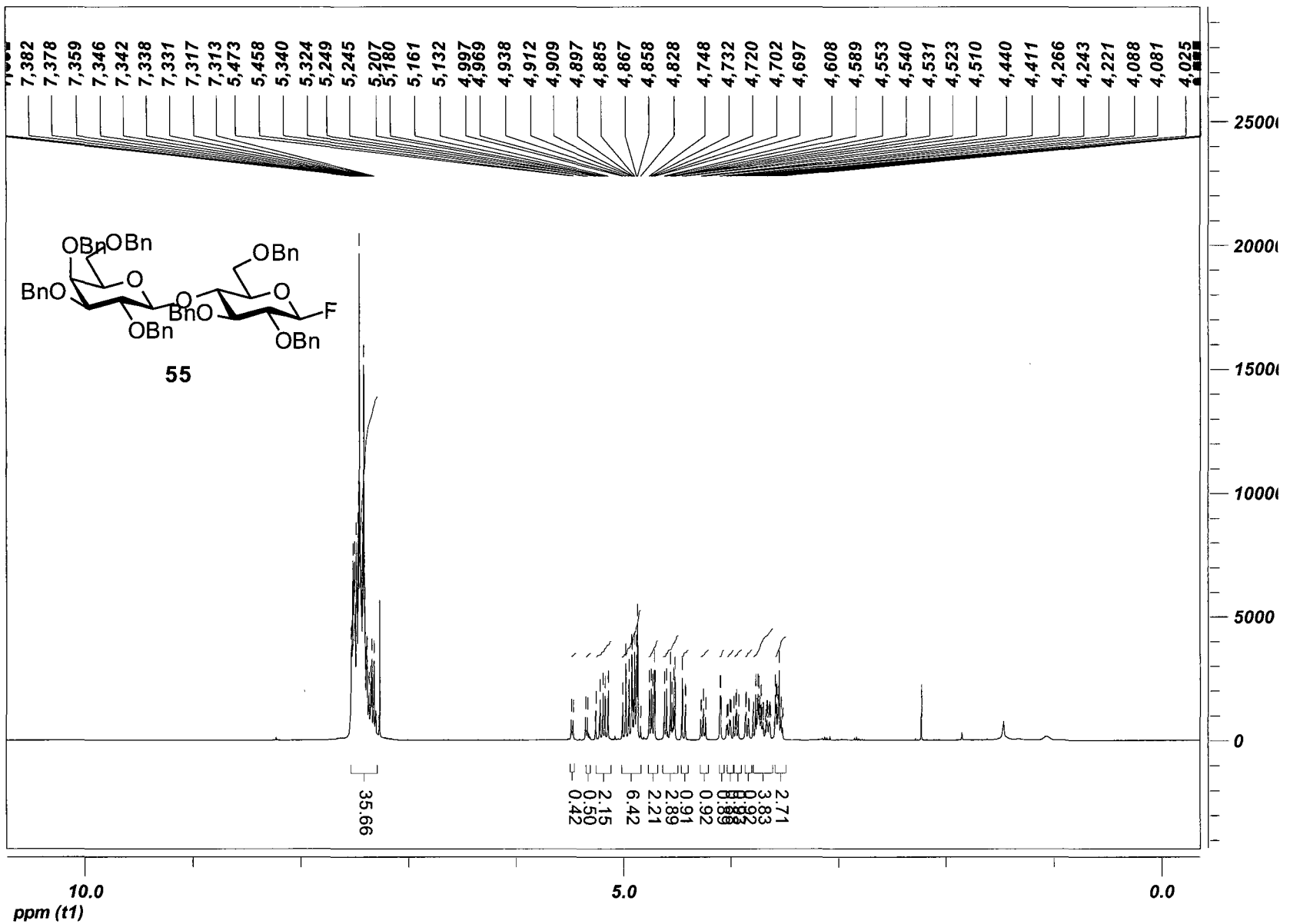

6.446
6.436
5.470
5.446
5.422
5.259
5.251
5.042
5.022
5.016
4.997
4.890
4.881
4.864
4.855
4.684
4.674
4.660
4.650
4.454
4.435
4.417
4.390
4.121
4.113
4.107
4.102
4.084
4.074
4.059
4.047
4.031
4.019
4.001
3.849
3.833
3.812
3.787
3.763
2.064
2.035
1.999
1.977
AcO
OAc
AcO
O
OAc
O
AcO
OAc
O
AcO
Br
58
1000
5000
0
1.00
1.02
1.04
1.02
1.01
0.99
2.03
4.24
2.07
3.08
18.99
10.0
ppm (t1)
5.0
0.0

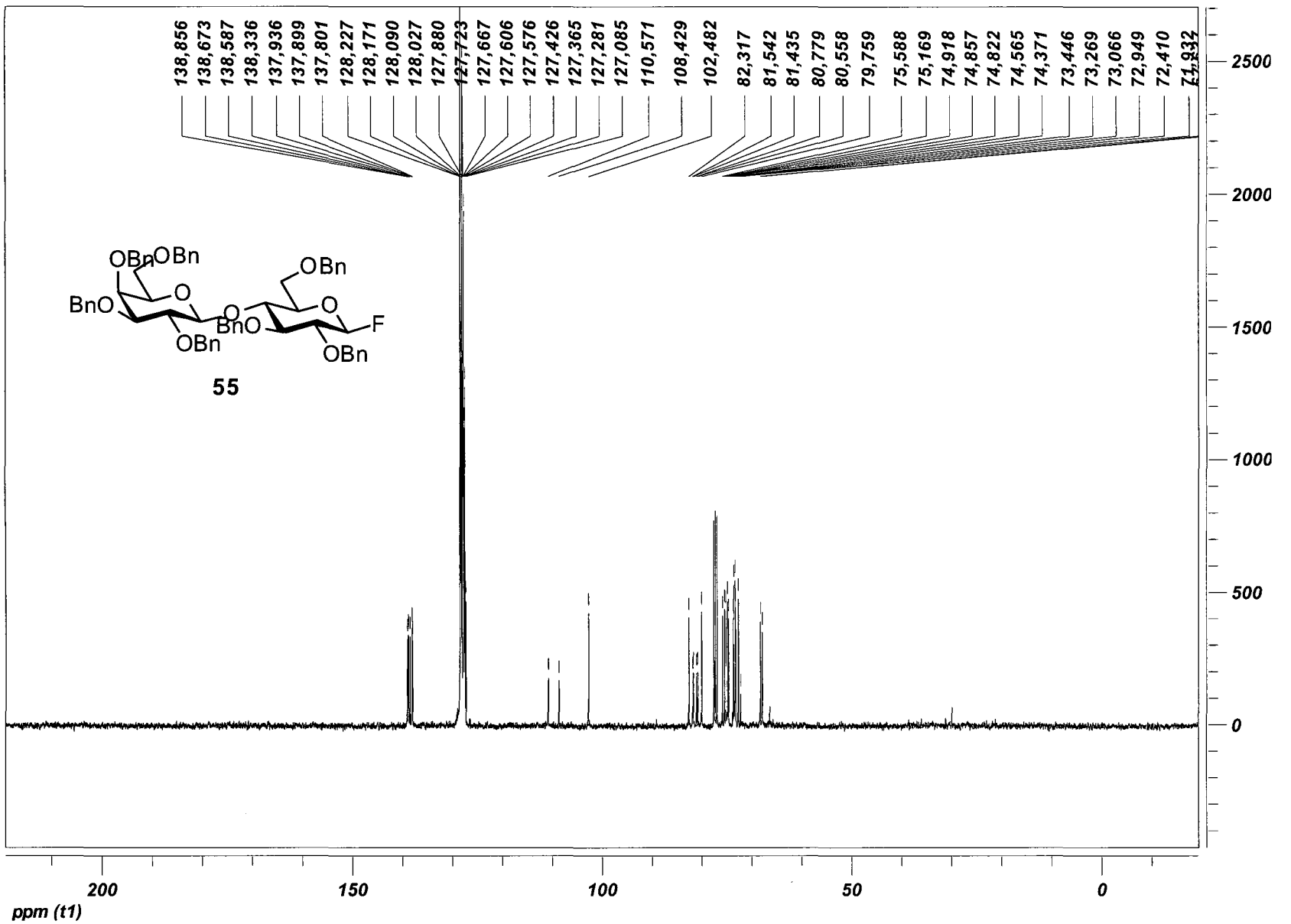

170.036
169.874
169.750
169.671
168.961
168.685
100.506
86.323
74.670
72.782
70.722
70.564
70.518
69.339
68.768
66.445
60.839
60.733
20.530
20.394
20.239
AcO
OAc
AcO
O
OAc
O
AcO
OAc
AcO
Br
58
3000
2500
2000
1500
1000
500
0
-500
200
150
100
50
0
ppm (t1)

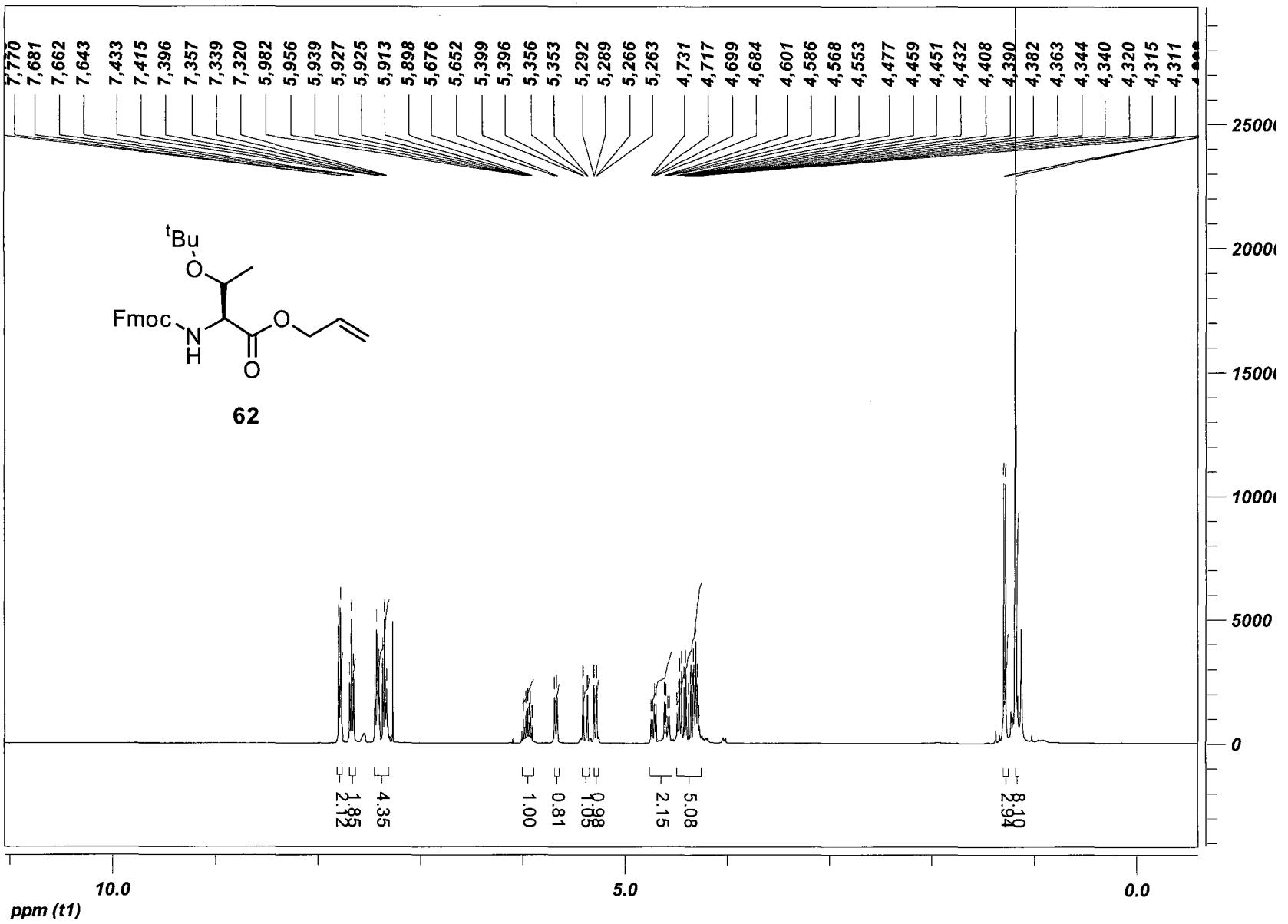

5,338
5,324
5,226
5,218
5,207
5,193
5,085
5,066
5,047
5,007
4,987
4,981
4,961
4,895
4,880
4,872
4,863
4,854
4,846
4,468
4,449
4,433
4,428
4,403
4,399
4,027
4,011
3,999
3,983
3,962
3,952
3,934
3,904
3,881
3,859
3,836
3,819
3,802
3,774
3,769
3,751
2,024
2,003
1,971
1,963
OAc
OAc
OAc
AcO
O
O
AcO
OAc
OAc
O
F
AcO
57
25000
20000
15000
10000
5000
0
0.46
1.50
1.06
2.83
2.02
3.09
2.09
1.04
10.0
ppm (t1)
5.0
0.0

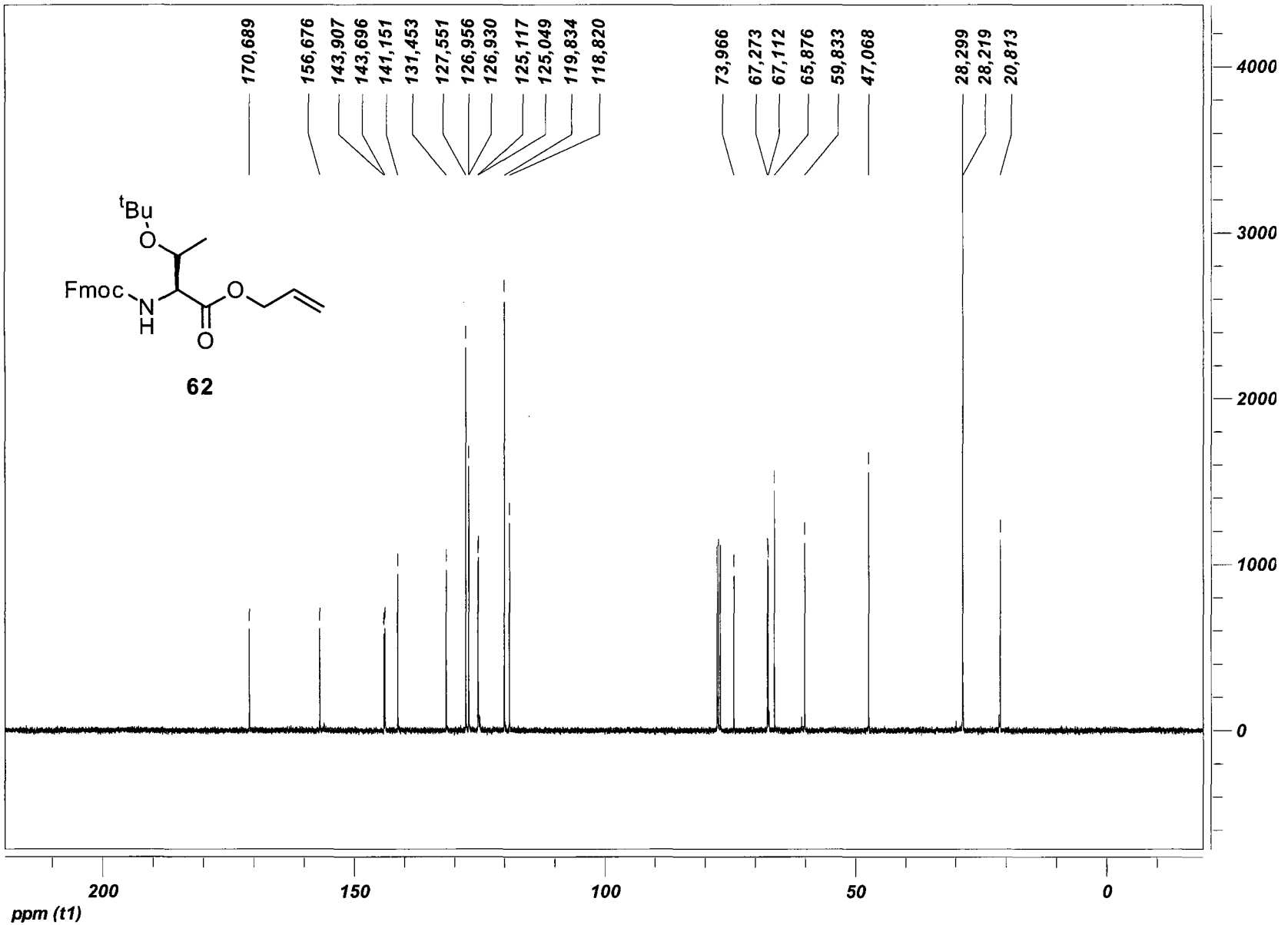

169,977
169,956
169,773
169,654
169,202
168,982
168,796
106,515
104,350
100,836
77,195
74,924
72,100
71,815
71,750
70,922
70,624
70,523
70,355
68,678
66,388
61,521
60,579
20,429
20,378
20,260
20,238
20,151
OAc
OAc
OAc
AcO
O
O
AcO
OAc
O
F
AcO
3000
2500
2000
1500
1000
500
0
200
150
100
50
0
ppm (t1)

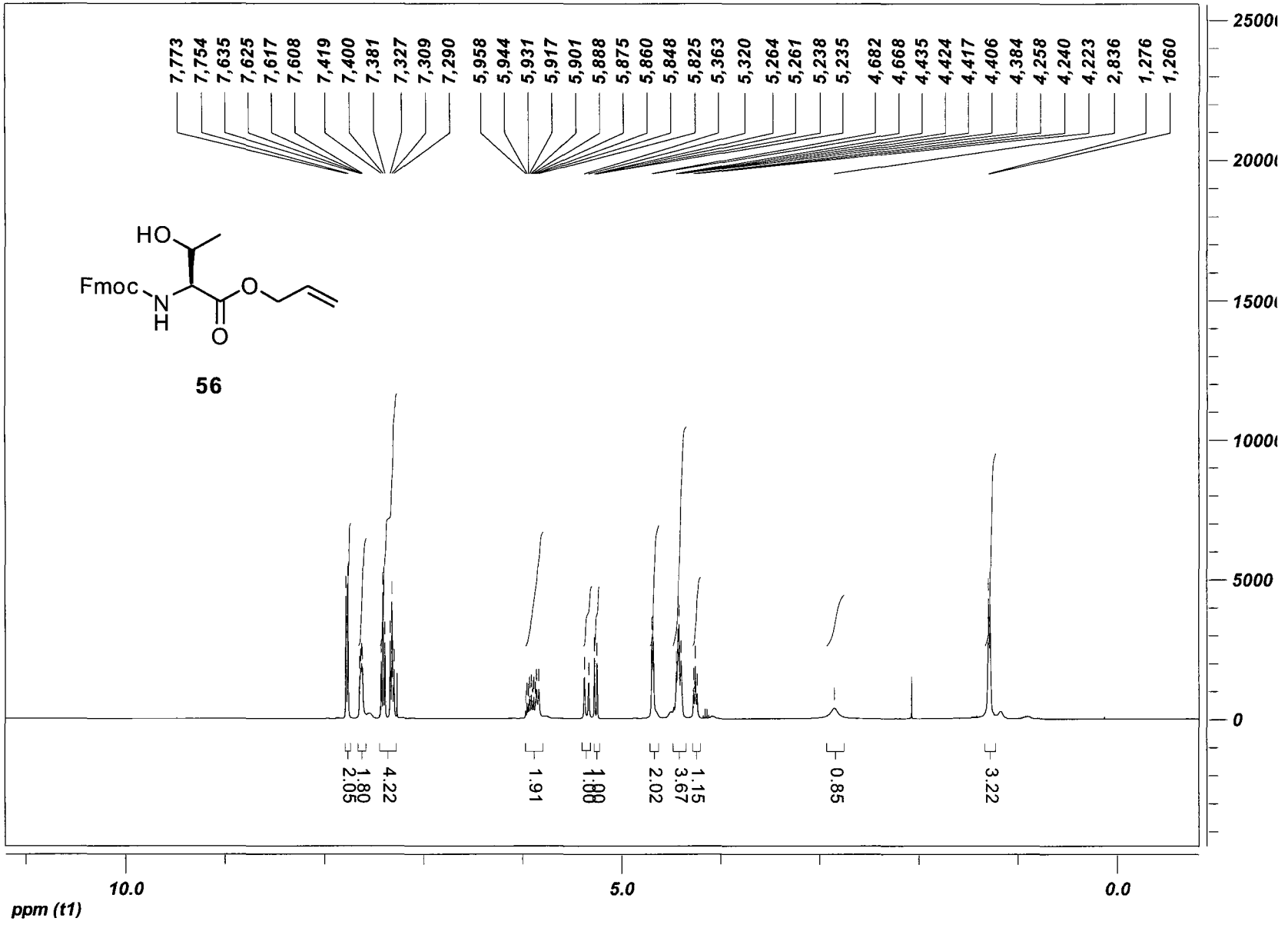
57


7,382
7,378
7,359
7,346
7,342
7,338
7,331
7,317
7,313
5,473
5,458
5,340
5,324
5,249
5,245
5,207
5,180
5,161
5,132
4,997
4,969
4,938
4,912
4,909
4,897
4,885
4,867
4,858
4,828
4,748
4,732
4,720
4,702
4,697
4,608
4,589
4,553
4,540
4,531
4,523
4,510
4,440
4,411
4,266
4,243
4,221
4,088
4,081
4,025
OBn
OBn
OBn
BnO
O
O
BnO
O
F
OBn
OBn
55
35.66
0.42
2.15
0.50
6.42
2.21
2.89
0.91
0.92
3.83
2.71
25000
20000
15000
10000
5000
0
10.0
5.0
0.0
ppm (t1)

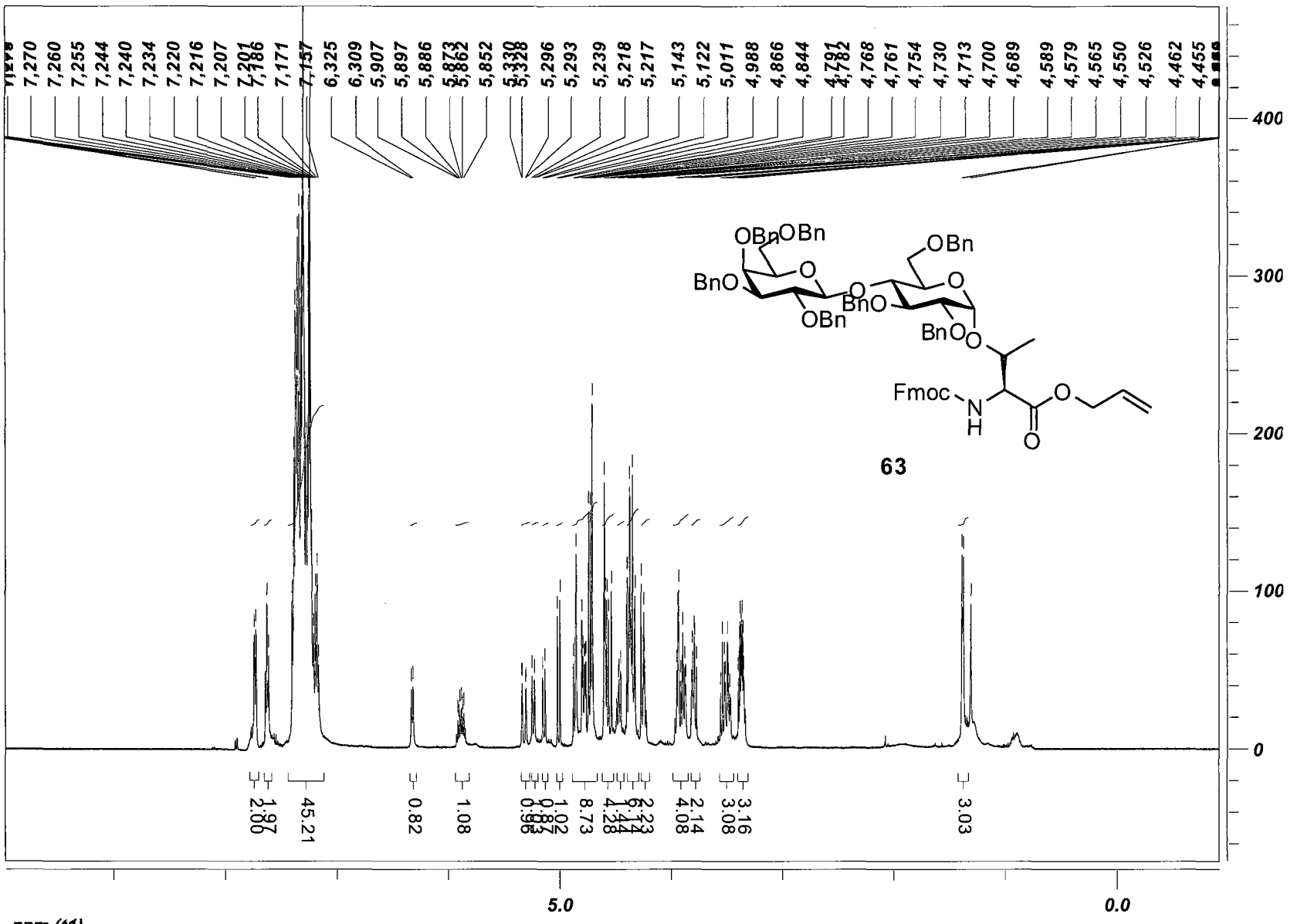

138,856
138,673
138,587
138,336
137,936
137,899
137,801
128,227
128,171
128,090
128,027
127,880
127,723
127,667
127,606
127,576
127,426
127,365
127,281
127,085
110,571
108,429
102,482
82,317
81,542
81,435
80,779
80,558
79,759
75,588
75,169
74,918
74,857
74,822
74,565
74,371
73,446
73,269
73,066
72,949
72,410
71,932
OBn
OBn
OBn
BnO
O
O
BnO
O
F
OBn
OBn
55
2500
2000
1500
1000
500
0
200
150
100
50
0
ppm (t1)

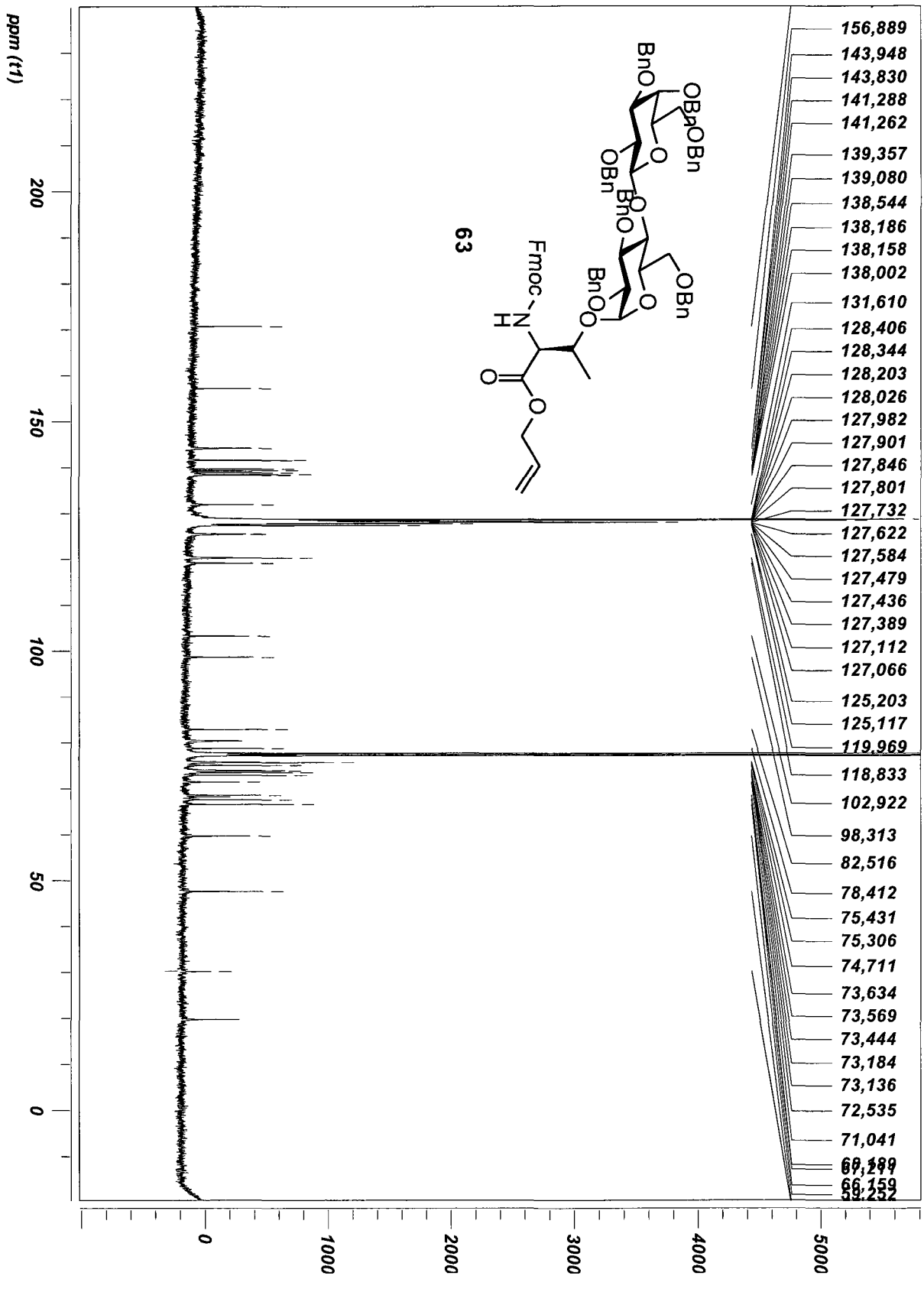

7,770
7,681
7,662
7,643
7,433
7,415
7,396
7,357
7,339
7,320
5,982
5,956
5,939
5,927
5,925
5,913
5,898
5,676
5,652
5,399
5,396
5,356
5,353
5,292
5,289
5,266
5,263
4,731
4,717
4,699
4,684
4,601
4,586
4,568
4,553
4,477
4,459
4,451
4,432
4,408
4,390
4,382
4,363
4,344
4,340
4,320
4,315
4,311
tBu
O
Fmoc
N
H
O
O
O
62
2.12
1.85
4.35
1.00
0.81
1.05
2.15
5.08
2.94
8.10
25000
20000
15000
10000
5000
0
10.0
5.0
0.0
ppm (t1)

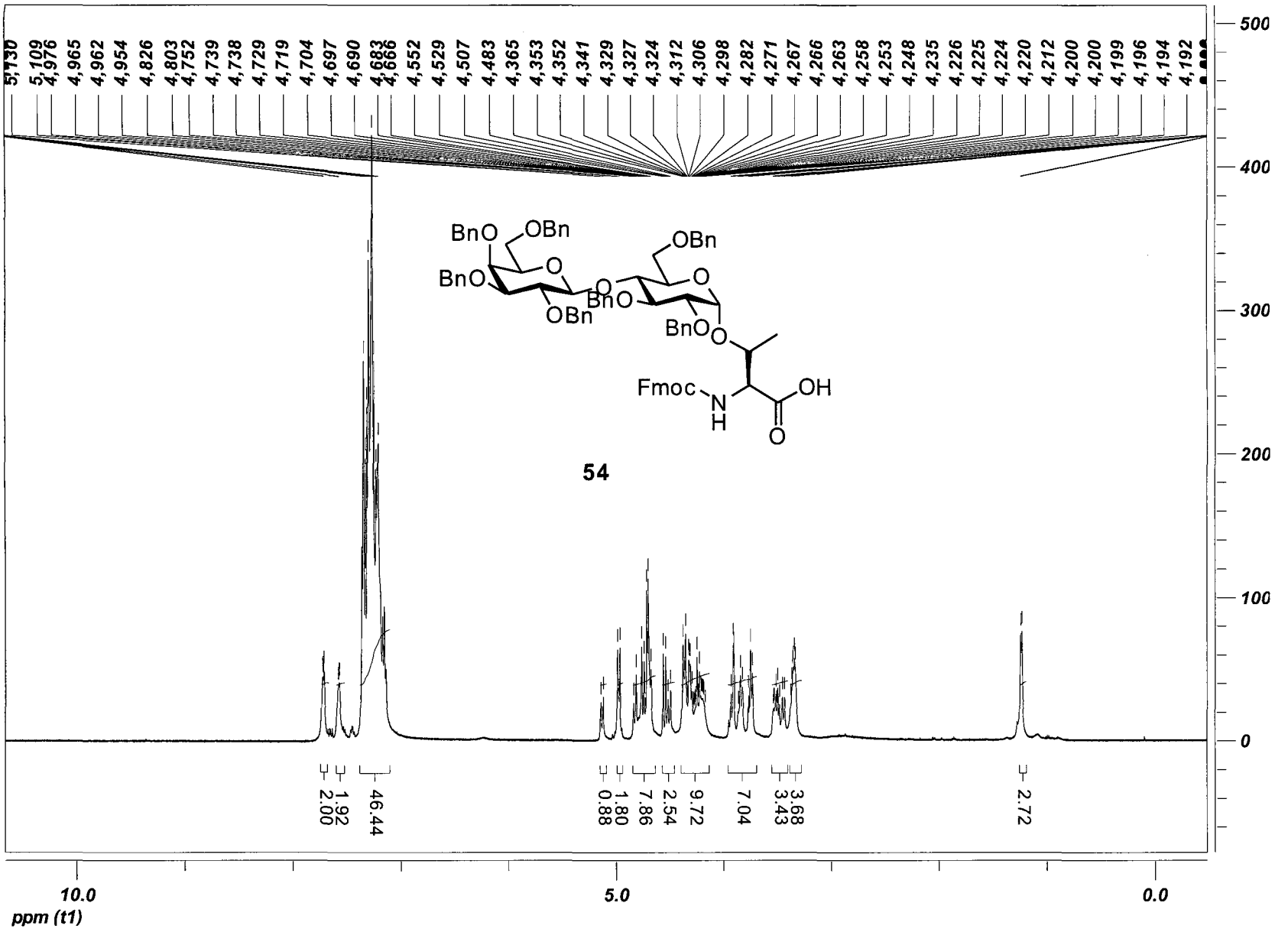

170,689
156,676
143,907
143,696
141,151
131,453
127,551
126,956
126,930
125,117
125,049
119,834
118,820
73,966
67,273
67,112
65,876
59,833
47,068
28,299
28,219
20,813
tBu
O
Fmoc
N
H
O
O
O
62
4000
3000
2000
1000
0
200
150
100
50
0
ppm (t1)

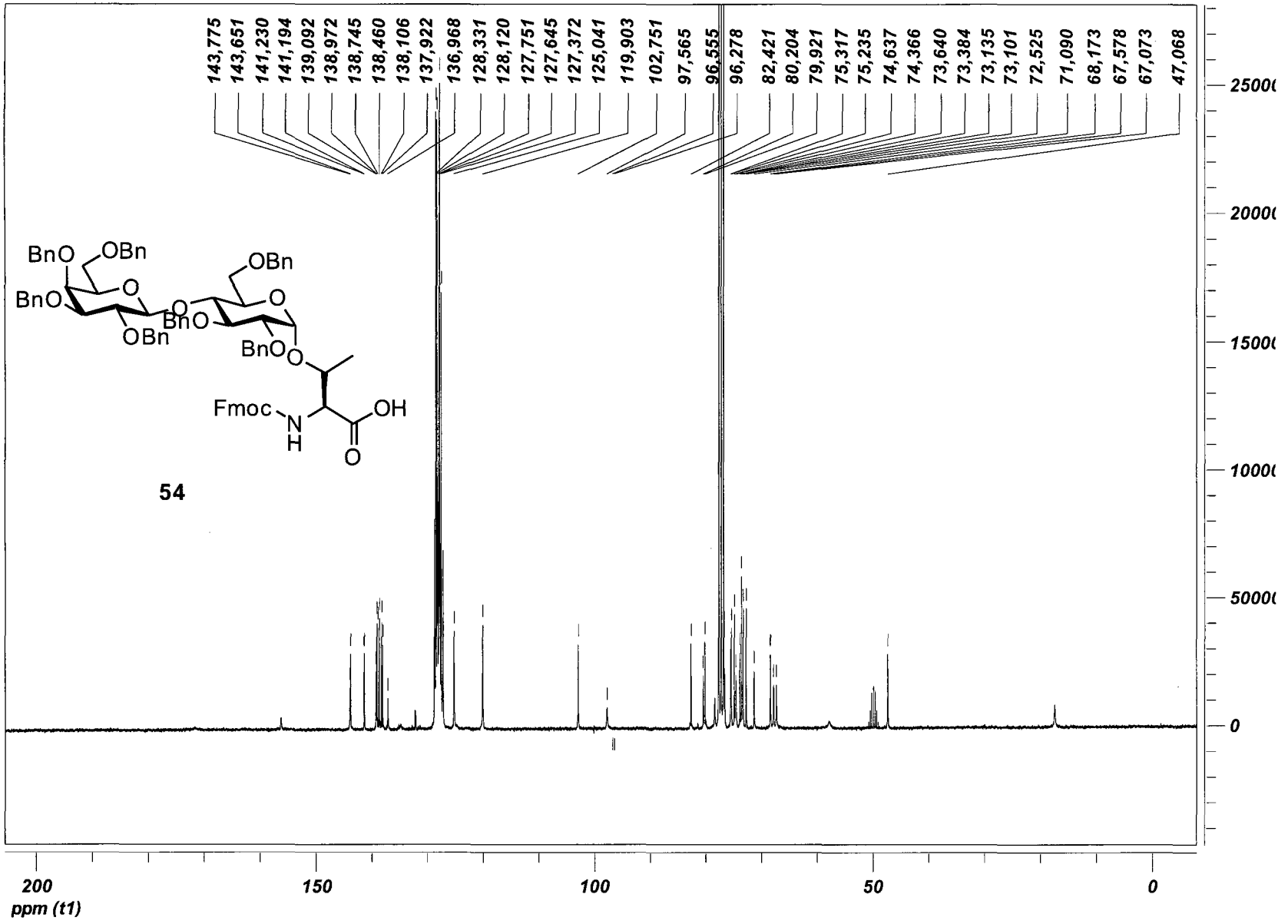

7,773
7,754
7,635
7,625
7,617
7,608
7,419
7,400
7,381
7,327
7,309
7,290
5,958
5,944
5,931
5,917
5,901
5,888
5,875
5,860
5,848
5,825
5,363
5,320
5,264
5,261
5,238
5,235
4,682
4,668
4,435
4,424
4,417
4,406
4,384
4,258
4,240
4,223
2,836
1,276
1,260
HO
Fmoc
N
H
O
O
O
56
2.05
1.80
4.22
1.91
1.00
1.00
2.02
3.67
1.15
0.85
3.22
25000
20000
15000
10000
5000
0
10.0
5.0
0.0
ppm (t1)

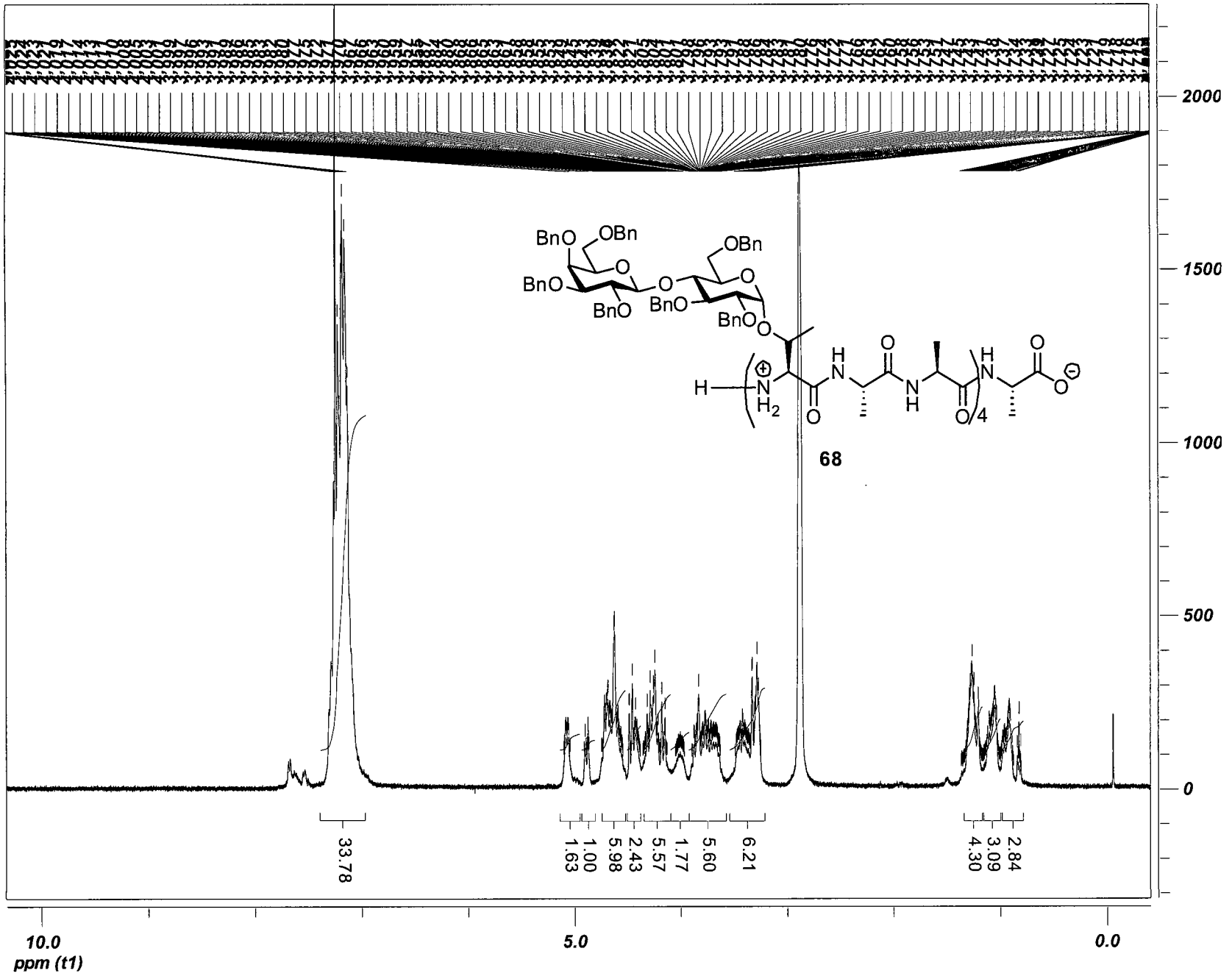
ppm (t1)
200
150
100
50
0
0
500
1000
1500
2000
2500
Fmoc
N
H
HO
O
O
56
170,851
156,752
143,755
143,576
141,190
131,324
127,633
126,992
125,024
119,893
118,836
67,888
67,166
66,130
59,168
47,026
19,844

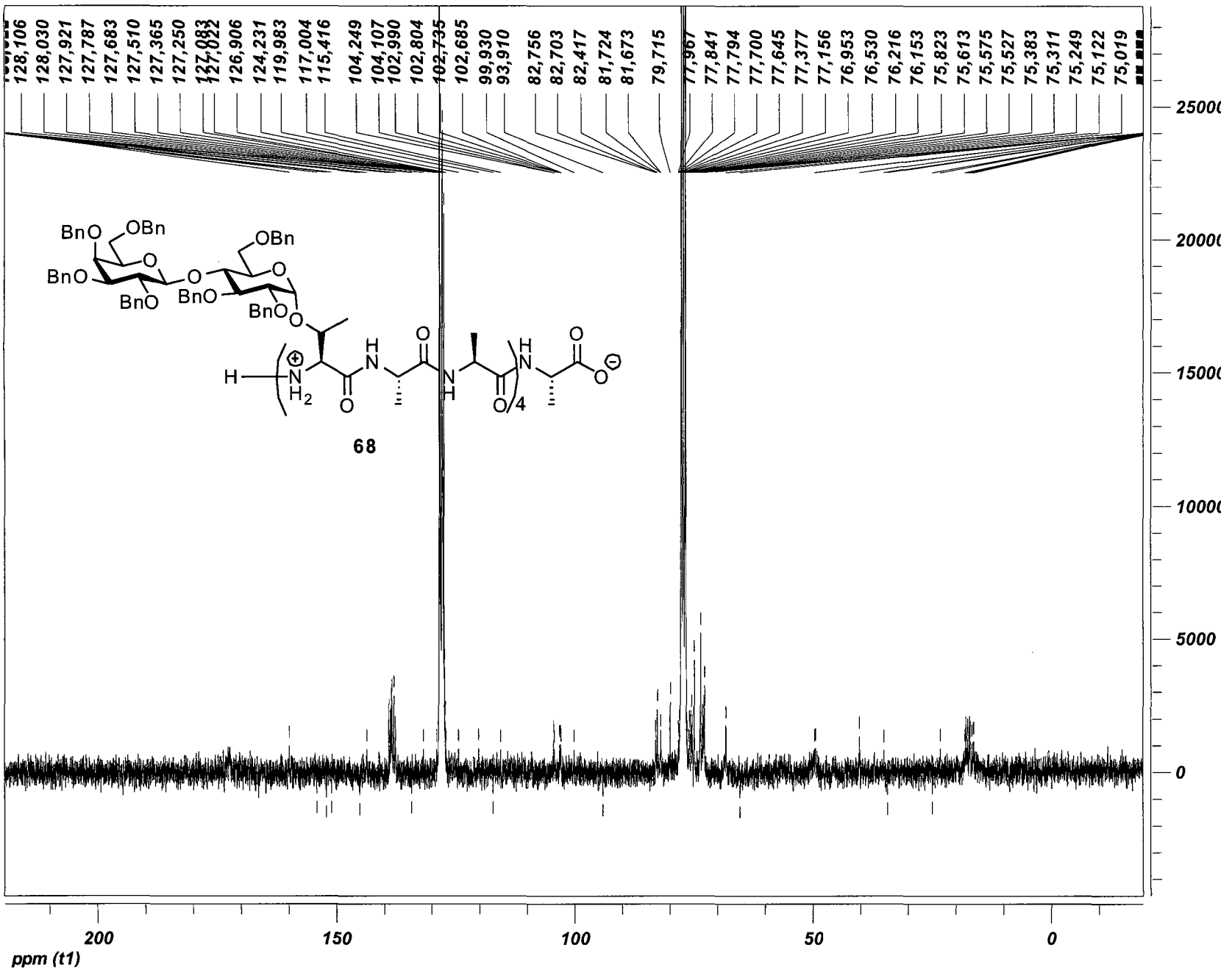

7,270
7,260
7,255
7,244
7,240
7,234
7,220
7,216
7,207
7,201
7,186
7,171
7,157
6,325
6,309
5,907
5,897
5,886
5,873
5,862
5,852
5,330
5,328
5,296
5,293
5,239
5,218
5,217
5,143
5,122
5,011
4,988
4,866
4,844
4,791
4,782
4,768
4,761
4,754
4,730
4,713
4,700
4,689
4,589
4,579
4,565
4,550
4,526
4,462
4,455
OBn
OBn
OBn
BnO
O
O
OBn
BnO
OBn
BnO
O
Fmoc
N
H
O
O
63
2.00
1.97
45.21
0.82
1.08
0.96
0.87
1.02
8.73
4.28
2.23
2.14
4.08
3.08
3.16
3.03
400
300
200
100
0
5.0
0.0
ppm (t1)


ppm (t1)
0
50
100
150
200
63
Fmoc
N
H
O
O
O
BnO
OBn
OBn
OBn
OBn
OBn
156,889
143,948
143,830
141,288
141,262
139,357
139,080
138,544
138,186
138,158
138,002
131,610
128,406
128,344
128,203
128,026
127,982
127,901
127,846
127,801
127,732
127,622
127,584
127,479
127,436
127,389
127,112
127,066
125,203
125,117
119,969
118,833
102,922
98,313
82,516
78,412
75,431
75,306
74,711
73,634
73,569
73,444
73,184
73,136
72,535
71,041
66,159
0
1000
2000
3000
4000
5000


5,130
5,109
4,976
4,965
4,962
4,954
4,826
4,803
4,752
4,739
4,738
4,729
4,719
4,704
4,697
4,690
4,683
4,666
4,552
4,529
4,507
4,483
4,365
4,353
4,352
4,341
4,329
4,327
4,324
4,312
4,306
4,298
4,282
4,271
4,267
4,266
4,263
4,258
4,253
4,248
4,235
4,226
4,225
4,224
4,220
4,212
4,200
4,200
4,199
4,196
4,194
4,192
BnO
OBn
BnO
O
OBn
O
BnO
OBn
O
BnO
O
Fmoc
N
H
OH
O
54
2.00
1.92
46.44
0.88
1.80
7.86
2.54
9.72
7.04
3.43
3.68
2.72
500
400
300
200
100
0
10.0
5.0
0.0
ppm (t1)


143,775
143,651
141,230
141,194
139,092
138,972
138,745
138,460
138,106
137,922
136,968
128,331
128,120
127,751
127,645
127,372
125,041
119,903
102,751
97,565
96,555
96,278
82,421
80,204
79,921
75,317
75,235
74,637
74,366
73,640
73,384
73,135
73,101
72,525
71,090
68,173
67,578
67,073
47,068
BnO
OBn
OBn
BnO
O
O
BnO
OBn
BnO
O
Fmoc
N
H
OH
O
54
25000
20000
15000
10000
5000
0
200
150
100
50
0
ppm (t1)


BnO
OBn
OBn
BnO
BnO
BnO
BnO
O
H
N
H2
O
4
68
2000
1500
1000
500
0
33.78
1.63
1.00
5.98
2.43
5.57
1.77
5.60
6.21
4.30
3.09
2.84
10.0
ppm (t1)
5.0
0.0


128,106
128,030
127,921
127,787
127,683
127,510
127,365
127,250
126,906
124,231
119,983
117,004
115,416
104,249
104,107
102,990
102,804
102,735
102,685
99,930
93,910
82,756
82,703
82,417
81,724
81,673
79,715
77,967
77,841
77,794
77,700
77,645
77,377
77,156
76,953
76,530
76,216
76,153
75,823
75,613
75,575
75,527
75,383
75,311
75,249
75,122
75,019
BnO
OBn
O
BnO
BnO
BnO
O
BnO
OBn
O
BnO
O
H
N
H2
O
H
N
O
H
N
O
O/4
O
O
200
150
100
50
0
ppm (t1)

68