

*Synthesis of Fused Tricyclic Compounds from Glycols &
Click Chemistry Inspired Imaging of Microbes*

Mr. Ashish Tripathi

Dr. Srinivas Hotha
(Research Guide)

**DIVISION OF ORGANIC CHEMISTRY
NATIONAL CHEMICAL LABORATORY
PUNE – 411008 (INDIA)**

[May 2008]

**SYNTHESIS OF FUSED TRICYCLIC COMPOUNDS FROM GLYCAL
& CLICK CHEMISTRY INSPIRED IMAGING OF MICROBES**

**A THESIS
SUBMITTED TO
UNIVERSITY OF PUNE
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
(IN CHEMISTRY)**

**BY
Mr. ASHISH TRIPATHI**

**DIVISION OF ORGANIC CHEMISTRY
NATIONAL CHEMICAL LABORATORY
PUNE – 411008 (INDIA)**

May 2008

Dedicated
To
My Parents, Brother & Sister-in-law

CERTIFICATE

This is to certify that the research work presented in thesis entitled “*Synthesis of Fused Tricyclic Compounds from Glycals & Click Chemistry inspired Imaging of Microbes*” has been carried out under my supervision at National Chemical Laboratory, Pune and is a bonafide work of **Mr. Ashish Tripathi**. This work is original and has not been submitted for any other degree or diploma of this or any other University.

Pune-411008

May 2008

(Dr. Srinivas Hotha)

Research Guide

DECLARATION

The research work embodied in this thesis has been carried out at National Chemical Laboratory, Pune under the supervision of **Dr. Srinivas Hotha**, Organic Chemistry Division, National Chemical Laboratory, Pune - 411 008. This work is original and has not been submitted in part or full, for any degree or diploma of this or any other university.

Organic Chemistry Division
National Chemical Laboratory
Pune-411008
May 2008

(Ashish Tripathi)

Acknowledgements

Science is fun doing and pleasure giving. Above all we get paid reasonably well for doing what we like most. It is a pleasant feeling for me to have this opportunity to express my gratitude for all of them who have been accompanied and supported throughout the time I spent working for my doctoral degree.

First and foremost, I would take this chance to express my unreserved thanks to my guide and mentor Dr. S. Hotha for his excellent guidance, continuous encouragement, and generous support during the every stage of my Ph.D. The confidence he had in me, willingness to share new ideas, enthusiasm to initiate novel projects and determination to drive them to completion helped me in a real sense to shape my research career. I do sincerely acknowledge the freedom rendered by him in the laboratory for the independent thinking, planning and execution of research. Although this eulogy is insufficient, I preserve an everlasting gratitude for him.

It gives me immense pleasure to thank Prof. Gopala Krishna Aradhyam, IIT Madras for providing me an opportunity to work under his guidance and use the facilities in his laboratory. My work on fluorescent imaging of microorganisms would not have been complete without his help, advice and co-operation.

Thanks are due to Dr. Sayam Sen Gupta and Dr. Mahesh Kulkarni for permitting me to work in their respective laboratories. Both of them have been extremely co-operative and have always provided me with valuable suggestions which have helped me to improve my research.

It would be very inappropriate of me not to make mention Mr. I. Sivakumar, Dr. C. V. Ramana and Dr. H.V. Thulasiram who were always ready for their timely help whenever required.

During my tenure in NCL, I learnt that a journey is easier when we travel together. I would like to thank my labmates Sushil, Girish, Sudhir, Suresh, Rao, Ashif, Suneel, Ram, Sandesh, Mahesh, Abhijeet and Shivaji for their kind help, invaluable discussions which we shared and maintaining a lively environment in the laboratory during the course of my work. My friends at IIT Madras, Bindu, Jeba, Harsha, Sai, Lavanya and Vijai were forever willing me to teach me the basics of biochemistry, without their assistance my work at IIT Madras would not have been successful.

Special thanks are due to my beloved friends Sameer, Divya, Sreedhar, James, Atul, Patwa, Susheem, Noor, Dillu, Prasanna, Ashwani, Kannan, Ramanujan, Srikant, Mohsin, JP,

Saurabh, Ganya, Rishi, Somesh for their unconditional support and continuous encouragement during my stay in NCL.

I wish to thank my fellow colleagues in NCL, Raman, Roopa, Mahesh, Satyendra, Pushpesh, Dr. Umashankar, Nagendra, Kamendra, Abhishek, Nishant, Arshad, Baagh, Dharmendra, Ankush, Dr. Khirud, Lakshi, Bhalchandra, Swaroop, Rajendra, Nagraj, Sachin, Pankaj, Sanjay, Chetan, Sarvesh, Abhilash, Arun, Dr. Manish, Dr. Arif, Pinak, Amol, Satish, Sashi, Ambrish, Bhuvan, Manje, Rakesh, Sunil and Sudharshan for their cheerful company and making my life in NCL very lively and enjoyable.

My thesis would not have been complete without the timely help from the spectroscopy group especially Dr. Rajmohan, Mr. Sathe and Mrs. Phalgune from NMR facility.

I am grateful to Council of Scientific and Industrial Research, Government of India, for awarding the junior and senior research fellowships and Dr. S. Sivaram, Director, National Chemical Laboratory to carry out my research works, extending all infrastructure facilities.

Finally, it has been a difficult task to capture and express my feelings for my family members. What I am and intend to be in the future is because of the good will and unstinted support of my parents, elder brother, Abhishek and my bhabhi, Suchitra, without knowing much what I am doing exactly, just wishing me all the time with no expectations. No words are enough to acknowledge them for their patience and sacrifice which were always remain a source of inspiration and will remain throughout my life. My success now and always will be dedicated to them.

Ashish

Contents

	<i>Page Number</i>
General Remarks	i
Abbreviations	ii
Abstract	iv
Chapter 1: Diversity Oriented Synthesis of Fused Tricyclic Compounds from Glycals	
Introduction	1
Present work	12
Experimental Section	27
Spectra Charts	38
References	72
Chapter 2: Click Chemistry Inspired Imaging of Microorganisms	
Introduction	74
Present work	82
Experimental Section	97
Spectra Charts	108
References	127
Chapter 3: Photocleavable Linkers for Bioconjugation of Proteins	
Introduction	129
Present work	137
Experimental Section	146
Spectra Charts	153
References	164
List of Publications	166

General Remarks

- ^1H NMR spectra were recorded on AV-200 MHz, AV-400 MHz, and DRX-500 MHz spectrometer using tetramethylsilane (TMS) as an internal standard. Chemical shifts have been expressed in ppm units downfield from TMS.
- ^{13}C NMR spectra were recorded on AV-50 MHz, AV-100 MHz, and DRX-125 MHz spectrometer.
- EI Mass spectra were recorded on Finnigan MAT-1020 spectrometer at 70 eV using a direct inlet system.
- Infrared spectra were scanned on Shimadzu IR 470 and Perkin-Elmer 683 or 1310 spectrometers with sodium chloride optics and are measured in cm^{-1} .
- UV/Visible spectra were recorded on Perkin Elmer Lambda 35 spectrophotometer.
- Fluorescent Graphs were scanned with JASCO FP-6500 fluorescence spectrophotometer.
- Fluorescent Images were analyzed under Leica DM5000B fluorescent microscope.
- Optical rotations were measured with a JASCO DIP 370 digital polarimeter.
- All reactions are monitored by Thin Layer chromatography (TLC) carried out on 0.25 mm E-Merck silica gel plates (60F-254) with UV light, I_2 , and anisaldehyde in ethanol as developing agents.
- All reactions were carried out under nitrogen or argon atmosphere with dry, freshly distilled solvents under anhydrous conditions unless otherwise specified. Yields refer to chromatographically and spectroscopically homogeneous materials unless otherwise stated.
- All evaporations were carried out under reduced pressure on Büchi rotary evaporator below 40 °C unless otherwise specified.
- Silica gel (60–120), (100-200), and (230-400) mesh were used for column chromatography.
- Scheme, Figure and Compound numbers in abstract and individual chapters are different.

Abbreviations

Ac	Acetyl/Acetate
CAN	Acetonitrile
Bz	Benzoyl
Bn	Benzyl
BSA	Bovine Serum Albumin
COSY	Correlation Spectroscopy
DIC	<i>N,N</i> -Diisopropylcarbodiimide
DIPEA	<i>N,N</i> -Diisopropylethyl amine
DMAP	Dimethylaminopyridine
DME	Dimethoxyethane
DMF	Dimethylformamide
DMSO	Dimethylsulphoxide
g	Gram
h	hour
Hz	Hertz
<i>J</i>	Coupling constant
M	Molar
mL	Milliliter
mol	Mole
mmol	Millimole
MsCl	Methanesulphonyl chloride
Ms	Methanesulphonyl
Me	Methyl
MOM	Methoxy- <i>O</i> -methyl
NOESY	Nuclear Overhauser Enhancement Spectroscopy
r.t.	Room temperature
TBDMSCl	<i>tert</i> -Butyldimethylsilyl chloride
TBDPSCl	<i>tert</i> -Butyldiphenylsilyl chloride
THF	Tetrahydrofuran
Tr	Trityl

Abstract

Thesis Organization

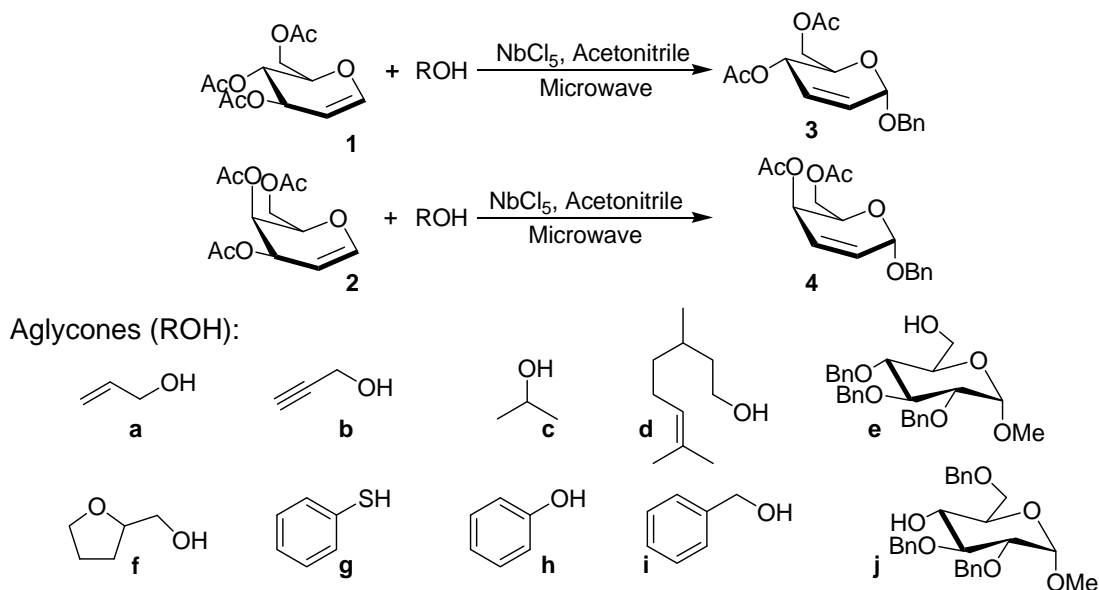
The research work has been divided into three parts each comprising of one chapter. The first chapter elaborates the utility of *Diversity Oriented Synthesis* in the organic chemistry with the aim to develop novel natural product like small compounds while in the second and third chapter describe the work influenced by chemical biology which is a fast emerging field of interfacial science. Taking advantage of the click chemistry, novel probes were synthesized which were employed for the imaging of micro-organisms and bioconjugation of the proteins. A short description of each chapter is provided here under.

1. Diversity Oriented Synthesis of Fused Tricyclic Compounds from Glycals

Cell permeable small molecules can be used to perturb protein function reversibly with temporal and quantitative control in biological systems. One of the goals of chemical genomics is to discover small molecules that affect biological processes through perturbation of protein function. High-throughput screening of chemically diverse libraries provides unprecedented opportunity for rapid identification of small molecules with better physiological effects. The chances of finding a hit molecule depends grossly on the collection of compounds present in the chemical library and ideally the library should be of highest level of structural, skeletal and stereochemical diversity. Diversity oriented synthesis is a newly proposed algorithm that enables synthesis of variety of scaffolds simultaneously in a predictable manner exploiting concepts from parallel synthesis. Our interest in the utilization of carbohydrate scaffolds for the development of DOS pathways prompted us to investigate synthesis of fused tricyclic cyclopentenones. We have utilized complexity generating reactions comprising the Pauson-Khand reaction, Ferrier reaction and Michael addition on glycals to achieve tricyclic enones in a highly diastereoselective manner.

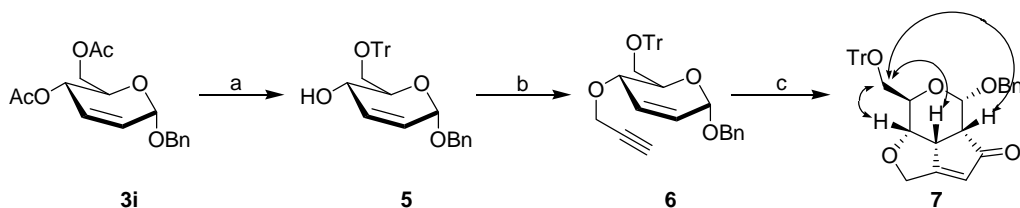
We choose to use the readily accessible unsaturated sugars as our starting material as the whole system could be subjected to Ferrier Reaction to yield 2,3-unsaturated glycosides. The double bond of the then glycoside can thus be subjected to various reactions e.g. asymmetric dihydroxylation, epoxidation etc to achieve structural complexity and diversity. The S_N2' attack of alcohols or allylic rearrangement of per-*O*-acetylated glycals was discovered by Ferrier using $BF_3 \cdot Et_2O$ as the Lewis acid catalyst is known as the

Ferrier reaction. The reaction is also known to occur in the presence of other Lewis Acid e.g. InCl_3 , SnCl_4 , BiCl_3 , FeCl_3 , $\text{Sc}(\text{OTf})_3$, ZnCl_2 etc. However, all the methods have some drawbacks namely form varying yields to low stereoselectivity and long reaction times. We



Scheme 1. NbCl_5 mediated microwave assisted synthesis of various 2,3 unsaturated glycosides demonstrated on per-*O*-acetylated glucal and galactal

then found out that NbCl_5 was effecting the formation of 2,3-unsaturated glycosides from per-*O*-acetylated glycols under microwave conditions. As a test reaction the 3,4,6-tri-*O*-acetyl D-glucal was treated with a catalytic amount of NbCl_5 in the presence of benzyl alcohol to yield benzyl 4,6-di-*O*-acetyl-2,3-dideoxy-D-erythro-hex-2-eno- α -D-glucopyranoside. The methodology was then extended to per-*O*-acetylated glucal and galactal using ten aglycones resulting in the formation of various mono- and disaccharides (**3a-3j** and **4a-4j**) as shown in Scheme 1. All the reactions were complete in less than a minute and gave only α isomer and good yields.



Reagent and Conditions: (a) (i) NaOMe , MeOH , r, 95%; (ii) Ph_3CCl , Pyridine , CH_2Cl_2 , 90%; (b) Propargyl bromide, NaH , DMF , $\text{nBu}_4\text{N}^+\text{I}^-$, 0 - rt; 3 hr, 91%; (c) $\text{Co}_2(\text{CO})_8$, CH_2Cl_2 , rt, 1 hr then acetonitrile-dimethoxyethane (4:1), 70 C, 4 hr, 92%.

Scheme 2. Synthesis of fused tricyclic cyclopentenones from 2,3 unsaturated glycosides and NOE interactions of the fused tricyclic enone

Going ahead and creating the next level of diversity on the unsaturated glycoside we thought that the 2,3-double bond would serve as an excellent platform for the 2+2+1 Pauson Khand reaction of the enyne. Subsequently the obtained Ferrier product, **3i** was deacetylated under the Zemplen conditions (NaOMe, MeOH). The C-4 hydroxyl of **5** was converted as propargyl ether (NaH in DMF, Propargyl bromide), before which the primary hydroxyl group was blocked with trityl chloride (CH₂Cl₂, Et₃N, Trityl Chloride).

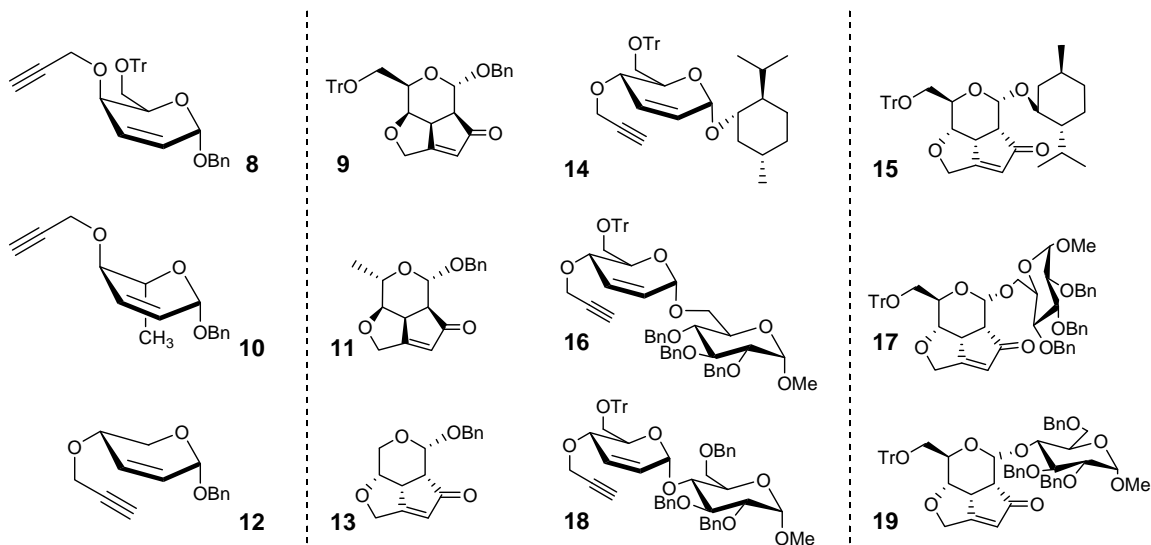
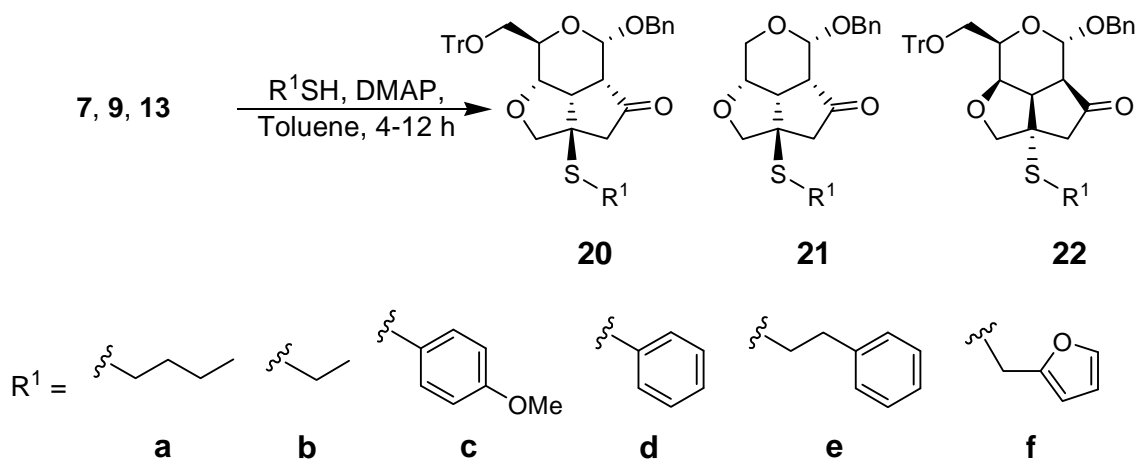


Figure 1. Synthesis of natural product like fused tricyclic enones

Having synthesized the required enyne, **6** we performed the Pauson-Khand reaction on the **6** employing Co₂(CO)₈, DCM under N₂ atmosphere to obtain the cobalt-enyne complex as thick red oil after passing through a pad of silica gel. Subsequently, the cobalt-enyne complex was cleaved by the NMO in DCM to yield fused tricyclic cyclopentenone (**7**) in 70% yield over the two steps. However, the cleavage of the cobalt complex affected by NMO was very slow and also sensitive to the amount of moisture present in the reaction mixture. The other protocols tried were wherein we used cyclohexylamine and acetonitrile instead of NMO did not give the desired results but instead led to the decomposition of the complex. Meanwhile, we found out that heating the cobalt-enyne complex in acetonitrile-dimethoxyethane (4:1) effected the cleavage of the complex and gave the desired tricyclic enone, **7** (Scheme 2). We observed no acetylenic proton in the ¹H NMR spectrum of the enone, however the olefinic proton was present at the δ 6.00 ppm as a singlet showing the presence of the α, β-unsaturated system in the molecule. The carbonyl resonances were present in the ¹³C NMR at the δ 206.7 ppm further confirming the success of the reaction. The other resonance signals e.g.

corresponding to the presence of the benzylic moiety remained unchanged confirming the configuration of the compound. The reaction was found to occur in a highly diastereoselective fashion leading to the formation of single diastereomer although three new chiral centers were formed simultaneously (confirmed by their NOESY experiments, Scheme 2). The chirality of the C-4 center has a very important role in configuration of the resulting tricyclic enones as the Co-enyne complex will be below the plane of the paper and thus resulting in the unidirectional fusion. Having developed the synthetic strategy for the formation of natural product like fused tricyclic cyclopentenones, we then successfully demonstrated the reaction on the other glycals (*per-O*-acetylated galactal, rhamninal and xylal) and using various *mono*- and *di*-saccharides, the results are elaborated in Figure 1. It is pertinent to mention that the compound **9** possesses the complementary stereochemistry at the new chiral centre formed when compared to **7** as there is an inversion at C-4 in the enyne **8**. The generality of the Pauson-Khand reaction can be gauged by the fact that all the compounds were synthesized in excellent yields and were chirally pure. The fused tricyclic library was characterized by ^1H , ^{13}C NMR spectra and the elemental analysis and the results found were in conformity with the assigned structure.



Scheme 3. Michael Addition of Thiol onto the Tricyclic Enones

The presence of α,β -unsaturated system in the enones would enable to function as excellent Michael acceptors. So we chose thiols due to the tremendous ability of sulphur nucleophiles to undergo 1,4-addition. In a typical experiment the enone **7** was taken in toluene and stirred along with the thiol in the presence of catalytic amount of DMAP. We observed the formation of a single diastereomer in case of all the enones due to bowl-like configuration of the fused tricyclic moiety. The incoming sulphur nucleophile attacked from the least sterically

demanding side so as to furnish a single diastereomer in excellent yields. A total of six thiols and three tricyclic enones were chosen for synthesizing the fused tricyclic library. The results are presented in Scheme 3.

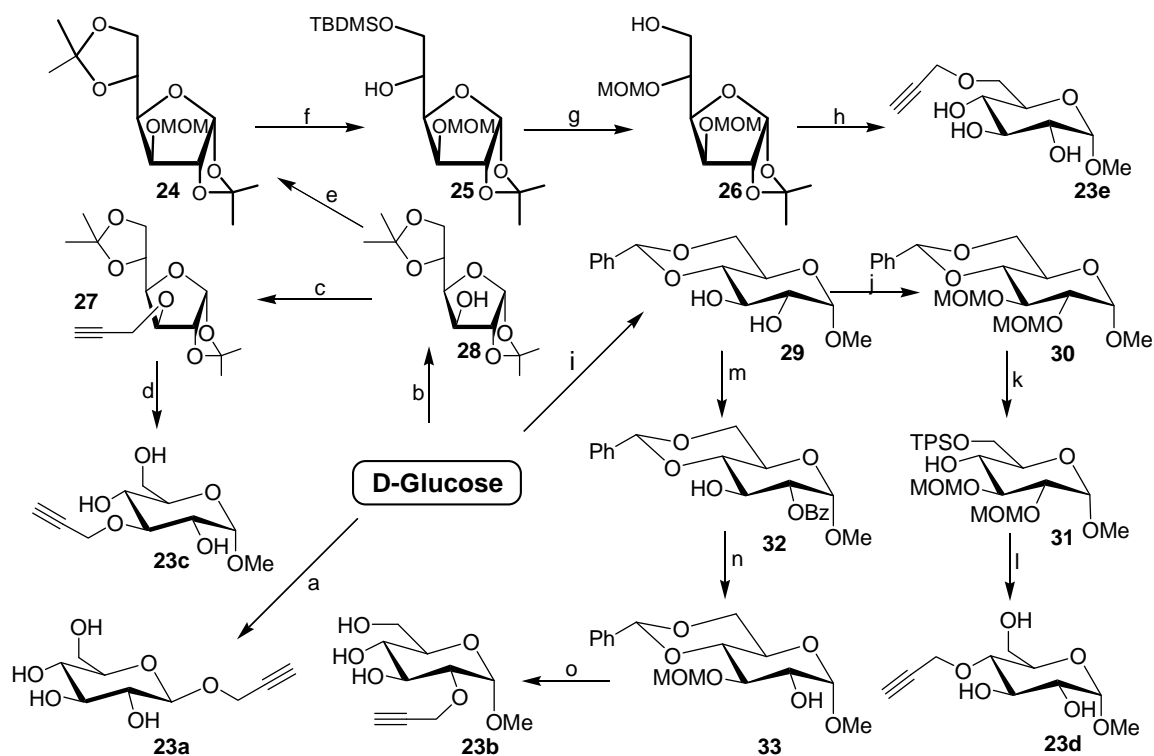
In conclusion, we have developed a practical protocol for the synthesis of natural product-like tricyclic compounds and a total of 25 tricyclic compounds were synthesized as a pilot library, using the complexity generating reactions like Ferrier, Pauson-Khand and the Michael reactions. A little peep at the library reveals the unique mixture of diversity and complexity among the members of the library. While the glucal and galactal derived scaffolds demonstrate the complementary stereochemistry, on the other hand, the rhamnal and xylal derived compounds provide the structural complexity to the carbohydrate library.

2. Click Chemistry Inspired Imaging of Microorganisms

The ability to visualize, track and quantify molecules and events in living cells with high spatial and temporal resolution is essential for understanding biological systems. There are very few techniques known in literature for probing the cellular pathways. Although fluorescent tags have long been used in cell biology, the recent advances like confocal microscopy fuelled it much further. In spite of all this, the major disadvantage is that fluorescent tags available commercially are bulkier in nature and when attached changes the structure and thus it tends to change the function of the biological molecules. Here in this effort we showed the utility of a coumarinyl based pro-fluorophore, which can be easily be tethered to any scaffold and can be triggered to become fluorescent through a Cu(I)-catalyzed 1,3 dipolar cycloaddition reaction with an incorporated alkyne functionality. Glucose is one of the most abundant molecule and is the major source of the energy inside the cells. It is pertinent to mention that bacteria use glucose as a sole C-source for their survival. So, we decided to track the most active carbon of glucose. For this we decided to selectively block the hydroxyl functionality of glucose sequentially with propargyl group and provide it to the microbial cell as the carbon source. We can thus monitor the cellular uptake by addition of the coumarinyl azide into the cell-lysate and take the fluorescence spectra of the cell-lysate.

All the five positional isomers of glucose (**23a-23e**) with propargyl appendages were synthesized *via* the intermediates **24-33** following the protocol shown in Scheme 4, and made to react with coumarinyl azide to yield the corresponding glucoconjugates (**34a-34e**, Scheme 5). It was heartwarming to notice that all the triazoles were fluorescent. Now having acquired

the desired water-soluble propargyl modified glucose derivatives we decided check the feasibility of the molecules to image the microorganisms. The *E.coli* cells (BL21) were

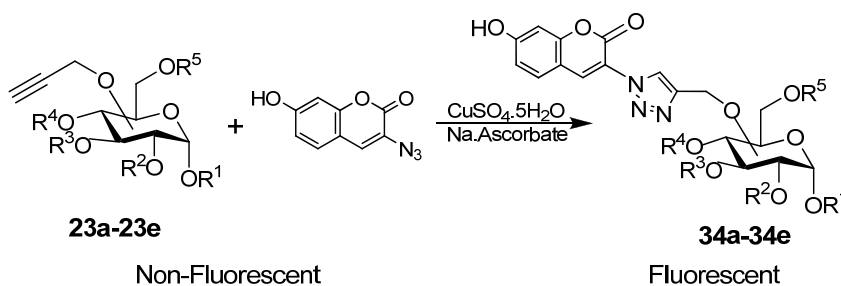


Reagents and conditions: (a) Propargyl Alcohol, Dioxane-HCl, reflux, 6 h, 75% (b) Acetone, CuSO_4 , 24h, 80% (c) Propargyl bromide, NaH, DMF, $\text{nBu}_4\text{N}^+\text{I}^-$, 0-rt, 3h, 94% (d) MeOH-HCl, reflux, 12h, 87% (e) MOMCl, DIPEA, CH_2Cl_2 , 12h, 95% (f) (i) p-toluene sulfonic acid, Methanol, 5h, 85% (ii) TBDMSCl, Triethylamine, CH_2Cl_2 , 4h, 95% (g) (i) MOMCl, DIPEA, CH_2Cl_2 , 12h, 90% (ii) p-toluene sulfonic acid, Methanol, 1h, 94% (h) MeOH-HCl, reflux, 12h, 85% (i) PhCHO, ZnCl_2 , 24h, 80% (j) MOMCl, DIPEA, CH_2Cl_2 , 12h, 97% (k) (i) p-toluene sulfonic acid, Methanol, 2h, 94% (ii) TBDPSCI, Triethylamine, CH_2Cl_2 , 4h, 92% (l) (i) Propargyl bromide, NaH, DMF, $\text{nBu}_4\text{N}^+\text{I}^-$, 0-rt, 3h, 96% (ii) MeOH-HCl, reflux, 4h, 88% (m) Dibutyltin oxide, Dioxane, Benzoyl Chloride, 6h, 85% (n) (i) MOMCl, DIPEA, CH_2Cl_2 , 12h, 97% (ii) NaOMe, MeOH, rt, 30 min, 97% (o) (i) Propargyl bromide, NaH, DMF, $\text{nBu}_4\text{N}^+\text{I}^-$, 0-rt, 3h, 96% (ii) MeOH-HCl, reflux, 2h, 92%

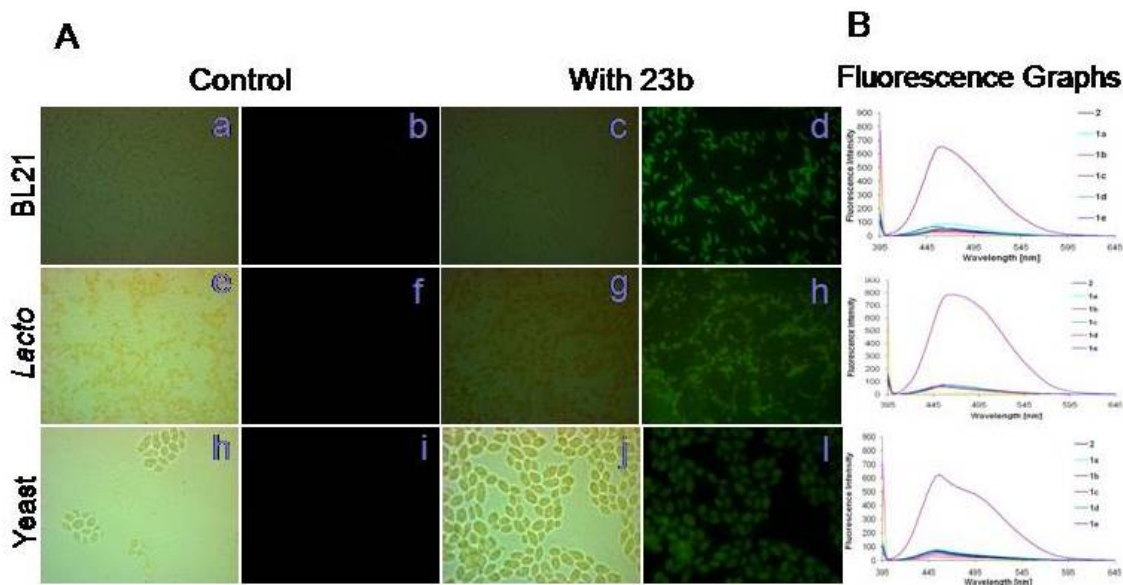
Scheme 4. Modified glucose derivative

allowed to grow in the presence of modified glucose for 12 h. The cells were then subsequently washed and lysed with sonification. The fluorescent spectra of the cell lysate were examined after their reaction with the coumarinyl azide. Analysis of the fluorescent spectra illustrated that only compound **23b** propargyl ether at C-2 position was observed to have been incorporated into the cell wall more as compared to other propargyl ethers (Figure 2B). The feasibility of exploiting coumarin based probe for visualizing bacterial cells through fluorescence microscope was then investigated using C-2 only. The cells were fixed on cover slips and click reaction was carried out. *E.coli*. cells incubated in the presence of **23b** showed strong fluorescence. The images of the bacterial results are shown respectively in Figure 3A.

Furthermore, the experiments were also performed on *Lactobacillus lacti* and yeast cells (*Sacchromyces* wild type). These microorganisms also demonstrated the similar results in accordance with that shown by *E. coli*. The C-2 propargyl compound, **23b**, was found to be more potent as compared to the other propargyl derivatives for incorporation into the cell wall of the micro-organisms. It may be noted that the control for all the micro-organisms did not exhibited any fluorescence. The control for all the experiments did not have any modified glucose derivative added to itself but it was made to undergo the same procedures as the experimental samples. This shows that modified glucose **23b** is being incorporated into microbes and after the conjugation with coumarin azide results in the fluorescence.



Scheme 5. Formation of Fluorescent Triazole



(A) Fluorescent images of BL21, *Lactobacillus lacti* and Yeast cells; (B) Fluorescence spectra of the BL21, *Lactobacillus lacti* and Yeast cells. Indicated cells were inoculated in normal complex media containing either glucose (control) or propargyl ether (with 23b) was allowed to grow overnight at 37°C. Cells were treated with coumarin azide in presence of CuSO_4 , sodium ascorbate after which the images and spectra were recorded.

Figure 2. Fluorescent Microscopic and Spectroscopic Studies

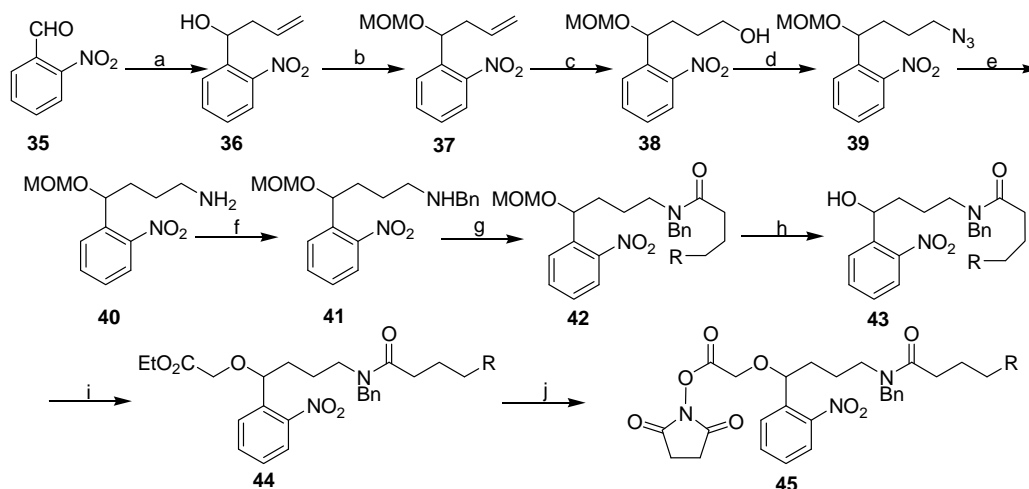
In summary, the utility of profluorophoric coumarine-based azide for fluorescent labelling and imaging of microorganism culture was demonstrated for the first time for bacterial and yeast cultures. The current protocol of fluorescent labelling *via* the click chemistry facilitates monitoring of glycan expression and the related processes which are poorly accessible. Additionally, this protocol also scores over the reported fluorescence probes with the size of the probe being smaller and the not interfering with the cellular processes.

3. Photocleavable Linkers for Bioconjugation of Proteins

Proteins are a large molecules comprising of various amino acids arranged from one end to another. Various different class of proteins are present inside the cell namely – hydrolases, proteases, kinases, nucleases, toxins, cell- matrix proteins, receptors, serum proteins, galactosidases and antibodies and are involved in various important biological functions inside the body. Manipulation of protein structure and thus its activity leads to unravel critical details of protein function at molecular and cellular level. These can be statistically modified utilizing the several active side chains of amino acid residues in them (e.g. Lysine, Arginine, Cysteine etc.) free amino group and sulfhydryl group can selectively be targeted to introduce the desired moieties to tap the cellular mechanisms. Post translational modifications affect the protein structure and thus its stability, activity and interactions with other molecules. Combination of chemical and biological techniques to chemoselectively modify proteins has proved to be an excellent resource for monitoring their function at molecular level. Photosensitive protecting groups act as valuable tools for investigating the biological phenomenon. Uses of photolabile groups have long been known for synthetic purposes and they have come a long way since their first application in caging of ATP molecule by Hoffman *et. al.* Photodeprotection is known to occur in high quantum yields at shorter wavelengths (> 300 nm) with no harmful side products. These can be used for the formation of stable covalent bonds on the reactive side chains of the proteins which can be easily cleaved using the light. Photocleavable groups can be used as conditional trigger to observe the various cellular mechanisms which in turn are controlled by the proteins. Various other functions of protein such as folding and unfolding, interactions with other bio-molecules etc. can also be monitored on a time dependent manner.

Herein we designed a photocleavable linker which can be easily introduced on the available free amino groups of the proteins. Further the designed linker will have o-nitro benzyl group on one end while on the other end it has alkyne or the azide moiety. The

conjugation of alkyne and azide moieties under the click conditions results in the formation of the 1,2,3-triazoles. The reaction is already known to be high yielding and atom economic. With the application of click chemistry we aim to investigate the fluorescence tagging, cross linking of proteins and the targeted delivery of drugs. Following our envisioned path, we started with our synthesis of the linkers. Starting with *o*-nitro benzaldehyde, **35** which was allylated employing allyltributyl tin and the resulting hydroxyl group, **36** was blocked with the MOMCl to obtain the MOM ether, **37**. For lengthening of the side chain, the hydroboration of the double bond was executed followed by conversion of hydroxyl moiety to amine, **40** using Staudinger reaction. The amide bond was formed utilizing six carbon chain bearing the alkyne or azide functionality at the other end. For the installation of the synthesized linker onto the protein we removed the MOM ether from hydroxyl group of **42** and installed the carboxylic acid functionality by simple alkylation using bromoethyl acetate. *N*-hydroxy succinamide was prepared, **45** with the hypothesis in the mind that it can be utilized to form an amide bond with the free amines present on the protein surface. The schematic representation of the protocol along with the reaction conditions are represented in the Scheme 6.



Reagents and Conditions: (a) Allyltributyl Tin, ZnCl₂, ACN/Water (9:1), 3h, 95% (b) MOMCl, DIPEA, CH₂Cl₂, 12h, 98% (c) BMS, NaOH, H₂O₂, THF, 2h, 80% (d) (i) MsCl, CH₂Cl₂, 30min (ii) NaN₃, DMF, 60°C, 4h, 90 % over two steps; (e) PPh₃, THF, Water, 4h (f) (i) PhCHO, MgSO₄, THF, 3h (ii) NaBH₄, THF, 1h, 70% over three steps (g) R'CO₂H, DIC, DMAP (cat.), DCM, 4h, 90% (h) HCl(cat.), Methanol reflux, 4h, 95% (i) NaH, THF, Bromoethylacetate, *n*-Bu₄N⁺I⁻, 0-rt, 3h, 70% (j) (i) NaOH, THF/Water, 1h, 80% (ii) *N*-hydroxy succinamide, DIC ,DMAP(cat.),CH₂Cl₂, 2h, 85%. R = CH₂N₃, ≡, R' = 6-azido-1-hexanoic acid, 5-hexynoic acid

Scheme 6: *Synthesis of photocleavable linkers*

The azide and the alkyne appendages at one end of the linker would serve as a handle for the triazole formation while the *N*-hydroxy group at the other end will facilitate in the bioconjugation onto the proteins. The triazole formation can be accomplished employing

another protein to form a homo- or hetro-dimeric conjugate of the protein or if the small molecules are engaged then it would help unravel the involvement of protein inside the various metabolic processes inside the cell. Having completed the synthesis of the linkers, we moved on to investigate the utility of the linkers for the bioconjugation. We choose the readily available Bovine Serum Albumin (BSA) for our experiments for its numerous biological applications including ELISA (Enzyme-Linked Immunosorbent Assay), blots etc. Additionally, it is also used as a nutrient in cell and microbial culture. One of the main advantages for preferring BSA over the others was due to its stability as compared to the other proteins.

BSA is known to have several free amines free in its tertiary structure. We choose to covalently link only a few of them in our experiments, firstly it will be impossible to react all the free amines and secondly, it will be tiresome job to cleave all the linked amines after the click reaction. The stock solution of 4 mg/mL BSA was prepared in 0.1 M phosphate buffer (pH 7.2) and mixed with NHS-linkers taken in DMSO. The resulting mixture was allowed to rotate in the rotary for 2 h. After the completion of the bioconjugation, the protein was dialysed overnight employing the cellulose membrane (12,000 kDa cut off). The samples were analysed by MALDI-TOF. Delightfully, both the azide and alkyne linkers were found to link onto the protein. Surprisingly, if the low concentration of the linkers were employed for the reaction then MALDI-TOF analysis showed that only one molecule of the linker is covalently attached to the protein.

To conclude we have successfully synthesized the photocleavable linkers having the azide and alkyne appendages and also exploited them to covalently link onto BSA which can be further employed for linking small molecules as drugs or fluorescence tags to the proteins. On the other hand, the linkers can also be exploited for protein-protein bioconjugation to form homo- and hetro-dimeric conjugates to assist in the studies regarding the behaviour of proteins in promiscuity with the other proteins.

Note: Compound numbers in abstract are different from those in the thesis.

Chapter 1
Diversity Oriented Synthesis of Fused Tricyclic Compounds from Glycals

Chapter 1: Introduction

Nature provides a plethora of structurally diverse and biologically active compounds. Natural products have a profound effect on both chemical biology and drug discovery. These natural products have been attractive probes for cell-cell signalling pathways, protein-protein interactions, and provide suitable cures for various disorders, disabilities and syndromes.¹ Cell-permeable and selective small molecules that are used to perturb protein function rapidly, reversibly and conditionally with temporal and quantitative control in any biological function helps in providing a better insight of the biological pathways.² The rich structural diversity and complexity of the natural products have prompted synthetic chemists to produce them in laboratory, often with medicinal applications in mind. Nature produces an astonishing array of structurally complex and diverse molecules as the secondary metabolites. Many drugs in clinical use are either natural products or natural product derivatives. For example over the last 20 years, 5% of the 1031 chemical entities approved as drugs were natural products, and another 23% were natural product-derived.³

The advent of the twentieth century saw the isolation of a large number of natural products from the plant extracts and with the progression of characterization techniques, UV; IR; and more important NMR and MS, there have been great ease in the identification of their structures.⁴ Even with the sophistication of purification and chromatographic methods, the separation and characterization of natural product extracts still remains a challenging task. Unfortunately, there are serious disadvantages of working with the extracts. Firstly, the nature does not synthesize secondary metabolites in pure form waiting to be isolated; therefore, the extracts comprise a collection of huge number of compounds with complementary structure and properties. This leaves us with the problem of isolation, identification of the active component(s). Secondly, there is always a limited supply of the active component(s) from the extract, usually in few milligrams from the huge plant extracts. Thirdly, the active component may be so complex structurally, such as Taxol that making analogues to characterize or optimize its activity remains a difficult exercise. Fourthly, the chemical space encompassed by the natural products is a limited space which may not be the only region useful for discovering physical or biological properties of interest.

The increasing entrepreneurship of synthetic organic chemists to in the art of total synthesis has helped to reduce the limited supply of natural products, even with the materials possessing very complex structures. In principle, the total synthesis is the complete chemical

synthesis of complex natural product starting with the much simpler starting materials without the aid of any biological processes. The target molecule can be a secondary metabolite from a plant; bacteria, or a medicinally active ingredient or organic compounds of theoretical interest in chemistry or biology.

The earliest demonstration of organic total synthesis was Friedrich Wöhler's synthesis of urea in 1828, which showed that organic molecules can be produced from inorganic precursors. The first commercialized total synthesis was Gustaf Komppa's synthesis and industrial production of camphor in 1903.⁵ Early efforts focussed on building chemicals which were extracted from plants or isolated from microorganisms. Today, the total synthesis remains a playground for synthetic chemists to develop new reactions, methodologies and highlight the sophistication of modern synthetic organic chemistry.^{5b} Sometimes, it inspires the development of novel mechanisms, catalysts or techniques. From a chemist's perspective, the art of total synthesis demands an accurate sense of chemical intuition and encyclopaedic knowledge of chemical reactions. Some of the classical examples of total synthesis include that of Cholesterol, Cortisone, Chlorophyll, Colchicine, Taxol, Vancomycin etc.^{5b, 5c} It is impertinent to mention E. J. Corey for his contribution in advancing the art and science of total synthesis by introducing retrosynthetic analysis.⁶

Nature produces molecules which can be simple, such as Serotonin and Histamine, or structurally complex, like Vancomycin, Taxol etc. These accommodate a variety of functional groups and are usually supplemented with the problem of instability, for most of these are stable inside the biological systems which provide an amazing stability and also enhance their activity. Target Oriented Synthesis (TOS) starts with the powerful planning algorithm, retrosynthetic analysis and subjecting simple and easily available synthons. In majority of the cases, the well trained eye of the professional is required in order to identify the basic fragment(s) present in the complex target molecule which can be derived from the suitable precursor. In the elaboration of the synthetic strategy, one should never forget that even the well established procedures may fail when applied in a specific structural context and thus an otherwise chemically sound synthetic plan may prove to be unworkable. Hence, the most risky synthetic operations should be shifted to the earliest possible step of the entire synthetic scheme so as to avoid the possible loss of time. A number of other criteria must also be considered when making a final selection between the options that emerge for the total synthesis of any given natural product. Among the most important are the length of the scheme (the fewer steps the better); availability and price of the starting materials, catalysts, the complexity of the equipments needed and last but not the least an in-depth knowledge of a rich

arsenal of available synthetic methods and a clear understanding of the ultimate goal of the whole endeavour. In fact a synthetic plan designed for the laboratory may appear nearly ideal but may be totally unacceptable for the industrial applications considering the cost measures, the amount of toxic waste being produced or simply the number of steps being involved.

It is heart-warming to see how the synthetic chemists solve the problems of synthetic chemistry using the rich arsenal of complementary methods. Woodward's synthesis of steroids beautifully illustrates the value of a carefully thought out general plan for a synthesis.³ His syntheses were most often described as having an element of art in them, and since then, the synthetic chemists have always looked for elegance as well as utility in their synthetic plan. Thus in the process of target oriented synthesis we come across the invention of new general reactions and reagents for organic synthesis, the design and execution of various efficient multi-step synthesis of complex organic molecules. However, inspite of the tremendous amount of effort, meticulous planning and money being poured into the total synthesis, it has failed to provide any suitable answer to the needs of pharmaceutical companies for regular and surplus supply of drug molecules. With the demand of pharmaceutical drugs shooting up to astronomical heights with the population boom, it has prompted the scientists to venture into more novel fields of science rather than banking on mother nature to provide us with the cure to our problems.

Chemical Space:

Chemical space is the space spanned by all possible (i.e. energetically stable) stoichiometric combinations of electrons and atomic nuclei and topologies (isomers) in molecules and also compounds in general (Figure 1).⁷ Chemical reactions allow us to move in chemical space. The mapping between chemical space and molecular properties is often not unique, meaning that there is usually more than one molecule which exhibits the same property which is being explored when carrying out material design and drug discovery.

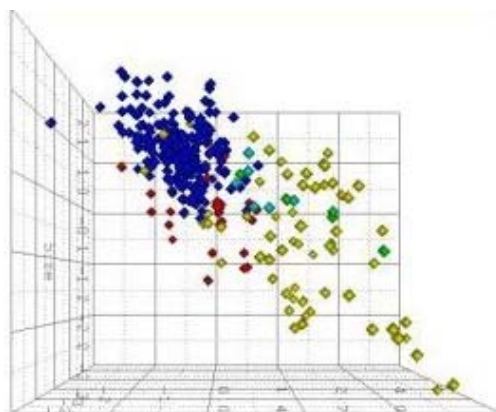


Figure 1: *Three dimensional view of Chemical Space.*

The number of drug like molecules in the chemical space has been estimated to be 10^{62} . By comparison, there are approximately 10^{51} atoms on earth. Therefore, it is impossible to synthesize every drug-like molecule and chemists must hence be selective. This becomes evident when knowing that only 27,000,000 molecules have been registered (and thus been synthesized) so far. Systematic exploration of chemical space is possible by creating *in silico* databases of virtual molecules.⁸ By representing these compounds as a series of chemical descriptors (molecular weight, lipophilicity, dipole moment, etc.) it is possible to plot them in chemical space.

Not all biologically active compounds encountered in the chemical space have the desired physicochemical properties to be a drug. A biologically active compound may be too lipophilic to be orally absorbed, too polar to be able to cross the gastrointestinal wall or may have too vulnerable chemical functionality that can be attacked by metabolizing systems in the body, and therefore not remain intact long enough to have a fruitful *in vivo* effect. The determination of characteristics of compounds that are more likely to be safe, orally bio-available medicines has led to the concept of 'drug-likeness'. Compounds that are drug-like have the potential to be developed into orally administered drugs, which are generally favoured owing to their use by the patients. In this regard, the Christopher Lipinski has postulated the Rule of Five which states that, in general, an orally active drug has no more than one violation of the following criteria: (i) Not more than 5 hydrogen bond donors (nitrogen or oxygen atoms with one or more hydrogen atoms); (ii) Not more than 10 hydrogen bond acceptors (nitrogen or oxygen atoms); (iii) A molecular weight under 500 gm/mol; (iv) A partition coefficient $\log P$ less than 5. Note that all the numbers are multiples of five and thus the postulation became popular as Lipinski's Rule of Five.⁹

With the introduction of various combinatorial techniques and high-throughput screening assays, there is an increasing demand for the small molecules within the chemical space which are more probable to be drug-like. Surprisingly, after the billions of years of evolution, nature still thrives on small molecules for signalling, protection and other essential functions. Simple life forms can function on a few hundred small molecules. In order to reap the benefits of small molecule approach for chemical genetics, advances must be made in finding the selective small molecules that bind to the protein target quickly, cheaply and with adequate selectivity. If chemical space is huge and we cannot synthesize everything, then we should decide on what to make and how to make it.

Many research groups have calculated where currently available drugs are located in chemical space and it had been noted that they cluster together. There are two schools of

thought on how to use this information. One approach is to look for new molecular entities based on ‘privileged’ core structures (pioneered by Waldmann group),¹⁰ which are known to be commonly bioactive. Another approach is to look for new molecular entities in uncharted regions of chemical space, by synthesizing new core structures (introduced by Schreiber group)¹¹. Diversity Oriented Synthesis (DOS) finds application in both of them. Generating libraries around new or unexplored templates with the aim of generating structurally, skeletally and functionally diverse compounds has become more common.¹²

Diversity Oriented Synthesis:

The search for new molecular entities on the drug discovery and chemical genetics programs requires the screening of high-quality collections of structurally diverse small molecules. The design and synthesis of such collections remains a major challenge to synthetic chemists. As mentioned earlier that the natural products do not occupy all regions of chemical space that are relevant to discovering bioactive compounds. Therefore, a question posed is, are there more productive, uncharted areas of chemical space that should be investigated to discover new molecular entities? In the early 1990’s chemist turned to combinatorial chemistry as a technique to efficiently synthesize large number of compounds by appending building blocks onto a core structure in search of finding novel active compounds. Despite, resulting in a large number of compounds being synthesized, this methodology was not successful as initially expected. The failure of the approach to discover a broad range of activities was due to the lack of structural diversity being attained. Any structural diversity of the molecules was only supplied by the building blocks and starting scaffold, while the resulting molecular framework was the same in every case. In order to achieve the highest level of structural diversity the following factors must be varied: (i) the building blocks; (ii) the stereochemistry; (iii) the functional groups; (iv) the molecular framework. Today chemists are investigating ways to synthesize libraries of compounds with high degree of structural diversity. Efficiently enriching chemical space in this way has been termed as *Diversity-Oriented Synthesis* (DOS), which concentrates on the synthesis of combinatorial libraries of structurally diverse (and complex) small molecules for biological screening.

The approach to *DOS* is in contrast to target-oriented synthesis (TOS), which aims to synthesize a single target, or traditional combinatorial chemistry, which generates structurally similar target structure (Figure 2). In target oriented synthesis we follow a convergent pathway to achieve the synthesis of any single natural product while in combinatorial chemistry we subject a set of similar compounds to a same set of reactions to compose a library which is more or less structurally similar. Complementary to these practices, the synthetic pathways in

DOS are branched and divergent, and the planning strategy extends from simple and similar compounds to more complex and diverse compounds both in terms of structure and activity. Retrosynthetic analysis traditionally focuses on the existence of a defined target structure. In *DOS* there is no single target structure and, therefore, retrosynthesis analysis cannot be used directly; instead a forward synthetic analysis algorithm is required which involves the transformation of a collection of substrates into a group of products by

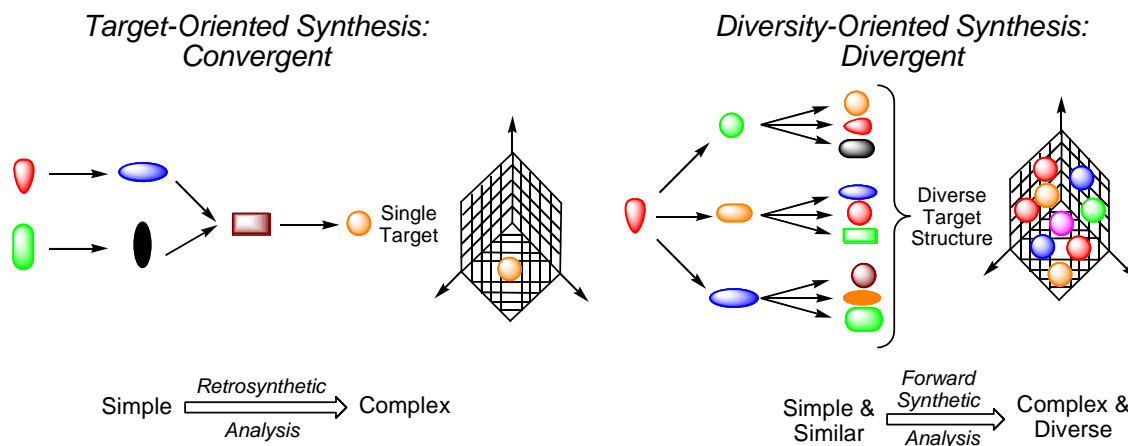


Figure 2: *Target-Oriented Synthesis and Diversity-Oriented Synthesis*

performing a number of chemical reactions together in forward direction. The inherent chemical reactivity common to all the substrates remains a key element for the implementation of forward synthetic analysis. Although, retrosynthetic analysis cannot be applied for the *DOS*, but its foundations have regularly been practised there to develop a complementary strategy to facilitate planning and implementation of synthetic strategies for the development of diverse libraries. Also complementary to *TOS*, the *DOS* does not involve a huge number of steps, the structural diversity and complexity is planned to achieve over as few steps as possible. With the help of various complexity generating reaction, reagent variations, substrate differentiation and stereochemical modifications *DOS* libraries aims to introduce rich diversity to enable them to target a wide range of proteins and thus inhibit the biological processes. However, the degree of diversity obtained within libraries is subjective: some libraries are reliant on a large number of appendages while others focus on synthesizing different skeletal structures. *DOS* libraries may seem to be inherited from a natural product scaffold or can be found out in an entire new region within the chemical space. Both of these categories are discussed henceforth.

Libraries inspired from the natural products:

The functional group diversity and architectural platforms engineered into natural products during their biosynthesis continue to attract several synthetic and medicinal chemists in their strategies for the creation of biologically active mimics, and provide selective ligands for cellular targets. Moreover, the natural products have also been effective in teaching us about chemical functionality that is compatible with most of the biological systems so they remain an invaluable tool for deciphering the logic for biosynthesis and as a platform for the development of frontline drugs. Individual natural products are proving to be valuable, biologically validated starting points for the library design. With the underlying structural frameworks selected for binding to certain protein domains, structural diversity arising from variations in appended functional groups can provide selectivity for related targets. These are thus used as a major source of innovation for therapeutic drugs for infectious diseases. However, the core scaffold of a natural product can also provide a biologically validated framework which can be altered with diverse functional groups to optimize its activity. Library design strategies have been divided into three broad categories: (i) libraries based in core scaffold of an individual product; (ii) libraries based in specific structural motifs; (iii) libraries that mimic the structural characteristic of natural products.

(i) Libraries based on core scaffold of an individual product: This approach has primarily been used to enhance the inherent activity of the parent natural product. Libraries designed on a privileged core of a natural product, known to inhibit a specific protein target, offer an excellent platform to provide an array of compounds which will have more chances to provide the lead molecule. Compounds in libraries that are based on core structures known to exhibit biological activity will have a higher intrinsic ability to bind to targets than those compounds in a library not based on natural products. A further extension suggests that if the parent core is a fortunate motif, defined as being able to target a specific class of proteins, then the offshoot library will be extremely useful for targeting multiple classes of protein having different folds. In a recent example, Schreiber and co-workers synthesized a library of 3520 spirooxindoles based on spirotryprostatin B, a mammalian cell-cycle inhibitor produced by *Aspergillus fumigatus*.¹³ Using a yeast cell-based screen, they identified 19 enhancers of latrunculin B, an actin polymerization inhibitor that induces growth arrest, and concurrently identified preliminary structure-activity relationship. In contrast, the consecutive library may not bind to the same target but can also display a wide range of activity against multiple protein targets having different folds. An important early demonstration in this approach was by Shair and co-workers discovery of secramine, a novel secretory pathway inhibitor from a library of

2527 compounds based on enantiomeric scaffold of galanthamine, a known acetylcholinesterase inhibitor.¹⁴ Notably, galanthamine had no effect upon the secretory pathway at upto 100 μ M (Figure 3).

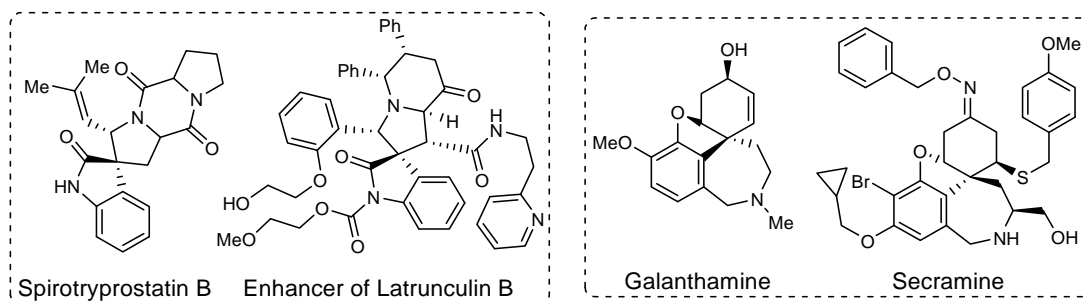


Figure 3. Some of the natural product derived active Pharmacophore molecules obtained from the DOS libraries synthesized by Shair group

Individual natural products are proving to be valuable starting materials for the library design. These provide biologically active framework, which when exploited to various chemical reactions and decorated with the various functional groups have been one of the most successful ways of drug design. However, the foremost limitation, in this approach is that the fixed scaffold is most likely suited to address a narrow range of biological targets with related activity.

(ii) *Libraries based on specific sub-structural motifs of natural products:* This strategy exploits the use of natural products sub-structures which are distributed across in most of the natural products. Inherent activities of the drugs are due to the presence of these core scaffolds. Ornamenting them with various functional groups provides an increased chance for the multiple target inhibition.

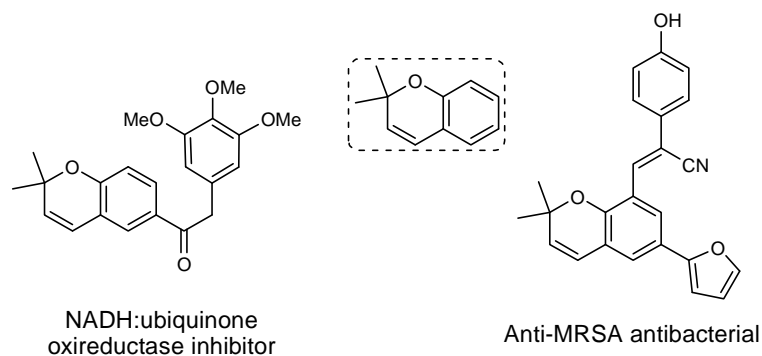


Figure 4. Nicolaou and group discovered two of the active drugs originated from 2,2-dimethylbenzopyran skeleton

One of the earliest discoveries in this regard was the approach of Nicolaou and co-workers' synthesis of a collection of 10,000-membered compounds derived from the skeleton

of 2,2-dimethylbenzopyrans, a privileged sub-structural motif present in a large number of natural products, with a wide range of activities.¹⁵ Their efforts yielded in identification and optimization inhibitors of NADH:ubiquinone oxidoreductase, a target known to be inhibited by 2,2-dimethylbenzopyran-containing natural products. Further, screening the library, for antibacterials they identified several cyanostilbenes with low micro molar activity against methicillin-resistant *Staphylococcus aureus* (MRSA) strains (Figure 4). In spite of holding the immense potential for the drug design there remains an important question which always remains with an element of doubt, whether the resulting library will preserve the parent activity of the core composition. However, the discoveries made forward in this regard are encouraging and suggest that such libraries will be of constant interest for the drug development.

(iii) *Libraries that mimic the structural characteristic of natural products:* This line of approach makes use of the various structural features of natural products, such as functional and stereochemical diverse feature around a variety of molecular frameworks. The main drawback to this approach had been the fact that it does not bear any direct connection to any natural scaffolds. However, in the same line of thinking, there remains an unexplored mine for the discovery of truly novel druggable targets.

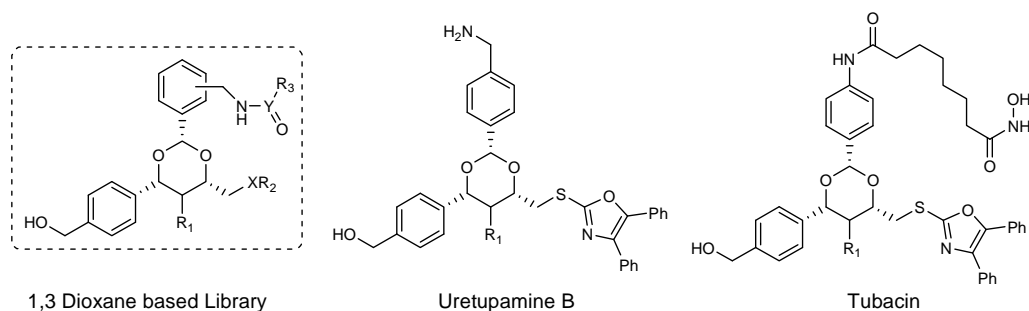


Figure 5. Active drugs synthesized by Schreiber group possessing the 1,3-dioxane framework

Schreiber and co-workers' 1,3-dioxane based libraries are a unique example in this endeavour. Uretupamine B, a novel inhibitor of the yeast nutrient responsive signalling protein Ure2p, was an early product of the first generation library of 1890 compounds.¹³ Importantly, Uretupamine B provided function selective Ure2p inhibition that could not be achieved with conventional genetic knockouts or RNA interference, allowing the dissection of two separate Ure2p-mediated signalling pathways. They further expanded this library to 18,000 members through the incorporation of additional building blocks and by introduction of stereochemical diversity. General screening analysis revealed that the resulting library comprised of over 2000 known pharmacologically active compounds (Figure 5).

This final approach of the DOS of deriving the libraries from the natural products has yielded significant number of the pharmacologically active molecules. The exploitation of skeletal diversity coupled with the stereochemical features of a variety of the natural product core scaffolds which are classified across a vast region of chemical space provides us with an unexplored gold mine to reveal novel drugs. The region of the chemical space occupied by these molecules neighbours that to the natural products, thus shows us an important pathway for discovering the novel biologically molecules which can be used for the further challenging targets.

Each of the above three strategies balance the degree of connection to natural-product structure space against the accessibility of structural diversity that is likely required to address multiple different biological targets. This approach has primarily been used to enhance the original activity of the parent natural product. However, the complexity of the natural products coupled with the limited supply can limit the structural modification and thus library design. But still the advances made forward in understanding the natural product biosynthesis, have helped us in the identification of the active motifs. The libraries designed around the natural product scaffolds lead us into an uncharted area of the chemical space to yield numerous small molecules which hold the potential of giving us a better insight of the understanding of the biological processes.

Novel libraries which do not bear any resemblance to the natural products:

Synthetic chemists get inspired after considering the vast amount of chemical space being under-utilised and also being aware of the fact that most of the natural products or their derived libraries are clustered together or occupy adjacent space to venture into new territories of chemical space in search for novel drug-like molecules. Generating libraries around new or underexploited templates with the aim of synthesizing structurally, skeletally and functionally diverse and novel scaffolds has become more common in the recent years. The idea is to exploit number of various diverse and unique skeletons and synthesize libraries exploiting the various complexity generating reactions, the resulting libraries although do not have any likeliness to natural products but are still able to bear the fruitful results. It is significant to mention the library synthesis by Mitchell and Shaw group which prepared a series of oxazoline compounds which were either spirocyclic or fused lactams (Figure 6).¹⁷ The researchers produced five core structures overall and found several compounds comprising this core which promote the growth of yeast, while others were cytotoxic to HeLa cells in a dose dependent manner.

However, together with the excitement of exploring new frontiers of chemical space is also comes the fear of failure. The key to the success of these libraries lies behind the careful planning and execution of the synthetic plan. If the core scaffold around which the whole library has been generated is biologically inactive or toxic then the whole effort can prove to be a futile exercise. Quite lately, a lot of effort has been forward in this direction and the chemists have exploited both linear as well as the convergent pathways for the generation of libraries with high degree of structural as well as stereochemical diversity.

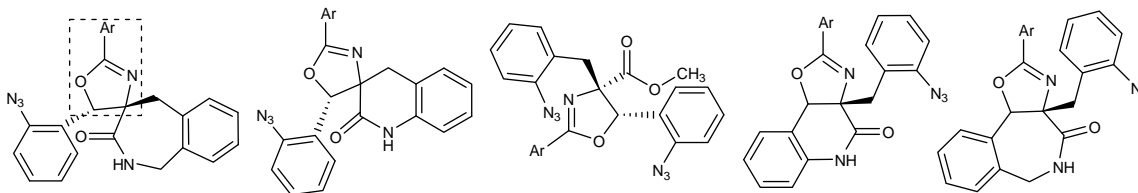


Figure 6. Small molecule inhibitors discovered by Mitchell and Shaw group using DOS

The ability of high-throughput screening methods in the rapid screening of huge number of molecules for various biological assays has made *DOS* a favourite sport amongst most researchers involved in the drug discovery. *DOS* has scored over the traditional target-oriented synthesis and lately introduced combinatorial chemistry with its uniqueness for the rapid synthesis of drug-like molecules. The success for *DOS* lies in the fact that its techniques had evolved from both of its precursors which were the traditional methods of drug design. The clever exploitation of forward synthetic analysis (the basic concepts of which finds its roots in retrosynthetic plans) and the combinatorial techniques to generate libraries with high degree of complexity has been the success of this technique. However, the real challenge for the future is to maintain a continuous supply of the libraries with high degree of structural diversity. This would enable the exploration of the previously uncharted regions of chemical space in an effort for the search of novel pharmacologically active compounds.

Chapter 1: Present Work

Carbohydrates are the most abundant of the four major classes of biomolecules, which also include proteins, lipids and nucleic acids, found in the living organisms which are the least exploited for the Diversity Oriented Synthesis (DOS). They along with their derivatives play a direct role in the smooth functioning of cellular mechanisms like immune system, fertilization, pathogenesis, blood clotting etc. and are also involved as the receptors for cell-cell recognition and pathogen defence. In most of the cases, these have a decisive role in the biological processes rather than being present only as a source of energy. To add to all this, sugars are also involved in providing a structural framework for the cells and in turn to the living organisms.

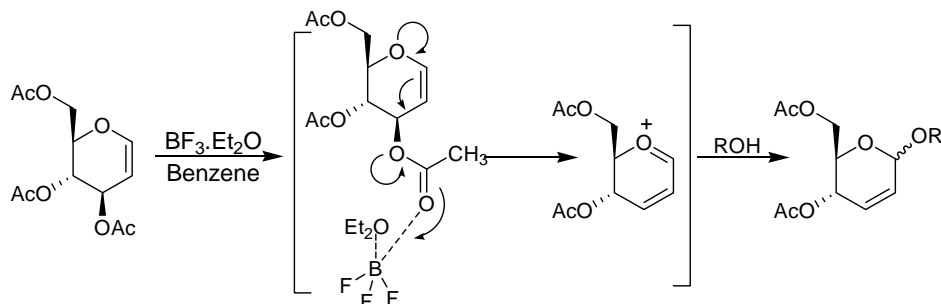
Renewed interest in the biology and chemistry of carbohydrates over the past two decades has dramatically accelerated the development of carbohydrate based therapeutics.¹⁷ Modern medicine already employs a range of bioactive carbohydrates, including heparin-based anticoagulants, polysaccharide vaccines, aminoglycoside antibiotics, azasugars for diabetes and with further developments, they also promise a cure for influenza and cancer. Besides, carbohydrates are also being used to modify the pharmacological profile of other drugs, including recombinant proteins. Some hybrid molecules often display superior efficacy and reduced side effects when compared with the parent drug. The use of such biomolecules for drug and gene delivery is actively being pursued. The development of many carbohydrate based pharmaceuticals has been facilitated by the gains made in the field of glycobiology, which focuses on defining the biological significance of carbohydrate interactions.¹⁸ The physical chemistry of carbohydrate-protein interactions strongly influences the pharmacology of many carbohydrate drugs. As relatively polar substances, they are unable to rapidly cross the lipid membranes and as a result they often act on the cell surfaces or in extracellular domain. And it will not be surprising to know that these are also distributed among a wide variety of natural products originating from different classes and having diverse applications. These are widely found in a number of the antibiotics produced by the microorganisms.

Notably carbohydrates also possess many of the favourable physical attributes, including their ready availability, chiral purity and structural diversity, even simple oligosaccharides have huge number of structural isomers owing to the poly-functionality of the constituent monomer.⁸ Although much of the efforts are being poured into the glycobiology, carbohydrates are not been utilized to their potential in *DOS*. The inherent chirality and the

poly-functionality of the carbohydrates can be easily manipulated to introduce the high degree of complexity in the resulting libraries, which help us in expanding our boundaries in the chemical space. The starting template in the *DOS* occupies a predominant role in dictating the resulting skeletal and stereochemical diversities. As evident by the Lipinski's Rule of Five and from various records generated from the data-mining, large number of biologically active natural products is oxygen-rich compared with corresponding congeners thus far synthesized from the combinatorial libraries. In view of the above facts, we became interested in initiating a diversity oriented synthesis using glycols as the starting template to enable oxygen-rich stereochemically pure scaffolds. The unsaturated bond as well the poly-hydroxyl functionality coupled with the inherent chirality of the glycols can be easily exploited to generate structurally diverse libraries in short steps. We planned to introduce a three level of diversity into the library using the range of complexity generating reactions like Ferrier, Pauson-Khand and Michael Addition reactions, which are discussed in detail hereunder.

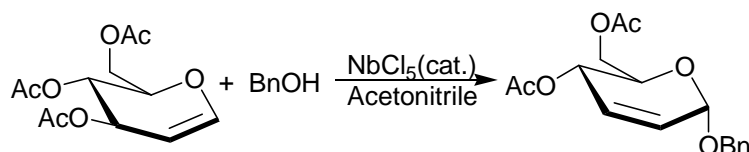
First Level of Diversity: *Ferrier Reaction*

One of the important reactions to produce diversity in glycol chemistry is the Ferrier reaction that gives access to 2,3-unsaturated glycosides. The S_N2' attack of alcohols or allylic rearrangement of per-*O*-acetylated glycols was discovered by R. J. Ferrier using $BF_3 \cdot Et_2O$ as the Lewis acid catalyst and is known as the Ferrier reaction (Scheme 1).¹⁹ The reaction has evolved many folds since its discovery in 1964, a variety of Lewis acids have been employed to effect this transformation include $InCl_3$,^{20a} Montmorillonite K10,^{20b} $SnCl_4$,^{20c} $BiCl_3$,^{20d} $FeCl_3$,^{20e} $Sc(OTf)_3$,^{20f} $ZnCl_4$,^{20g} $LiBF_4$,^{20h} $Dy(OTf)_3$,²⁰ⁱ and $ZrCl_4$.^{20j} In addition to these Lewis acids, oxidizing agents such as 2,3-dichloro-5,6-dicyano-*p*-benzoquinone (DDQ),^{21a} *N*-iodosuccinimide (NIS),^{21b} iodine, iodonium dicollidine perchlorate (IDCP),^{21c} ceric ammonium nitrate^{21d} and $HClO_4$ impregnated on silica gel^{21e} produce the desired 2,3-unsaturated glycopyranosides. However, all these methods have their own drawbacks in generality, low yields, low stereoselectivity and functional group compatibility, and practicality due to the harsh reaction conditions, acidic medium, high temperature, strong oxidizing conditions,



Scheme 1. Mechanism of the Ferrier Reaction.

longer reaction time and stoichiometric use of the costly reagents being employed. In view of the above and in our efforts to utilize glycal chemistry to achieve diverse scaffolds, we focussed on developing a practical procedure for synthesizing 2,3-dideoxy glycopyranosides. Recently niobium(V)chloride has emerged as a Lewis acid for a variety of the reactions.^{21g} In particular NbCl₅ has many advantages compared with other Lewis acids such as ease of handling, moisture stability, shelf life and economic viability. So we decided to take advantage of the Lewis acidity of the NbCl₅ for effecting the Ferrier reaction. To check our hypothesis, per-*O*-acetylated glucal in the presence of benzyl alcohol was treated with one mole percent of NbCl₅ in acetonitrile under microwave irradiation in an open vessel for 2 minutes (Scheme 2).²²



Scheme 2. NbCl₅ promoted Ferrier reaction of per-*O*-acetylated glucal

Not only is the desired product formed with 97% yield but also only α -anomer was observed. The use of microwave reaction conditions with carbohydrate is not yet a mature subject, albeit various successful efforts in conventional organic synthesis exist. The use of microwaves in the Ferrier reaction was explored using Montmorillonite K10^{20a} and InCl₃.^{20b} The microwave induced Montmorillonite K10 mediated Ferrier rearrangement of per-*O*-acetylated glucal resulted in the formation of 2,3-unsaturated glucosides; the corresponding InCl₃ reaction gave α -anomeric selectivity, however, 30 mol% of InCl₃ was required. Complementary to the other Lewis acids, only catalytic amounts of NbCl₅ effected the reaction with excellent yields and stereoselectivity. It is also pertinent to mention that the corresponding reaction under room temperature or reflux conditions took almost one hour for completion as evident by the consumption of the starting material, thereby making this reaction a microwave oven-induced reaction enhancement (MORE) technique.

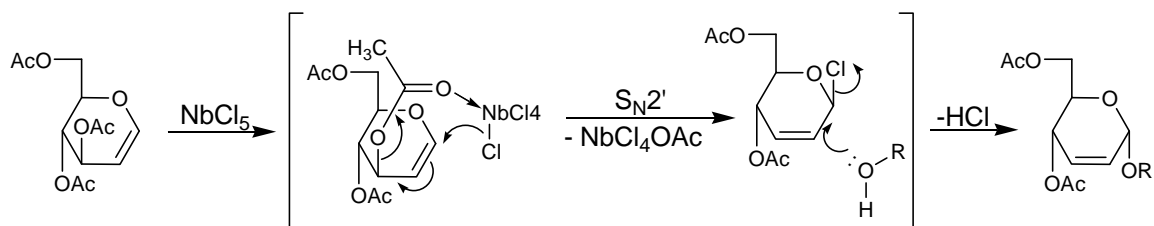


Figure 7. Mechanism of the NbCl₅ promoted Ferrier reaction

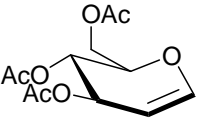
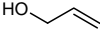
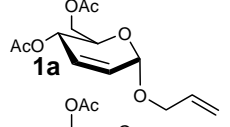
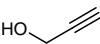
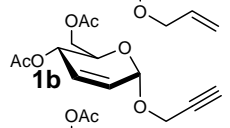
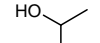
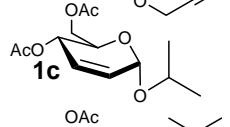
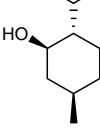
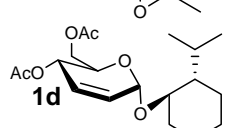
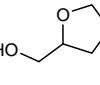
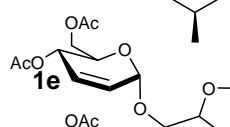
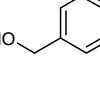
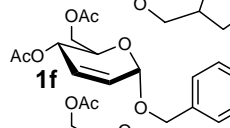
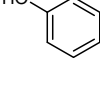
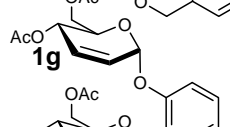
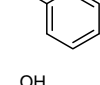
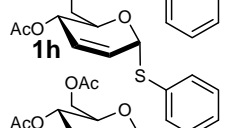
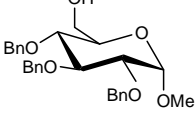
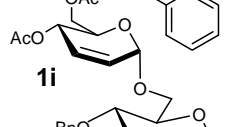
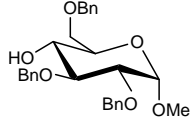
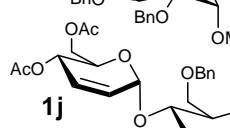
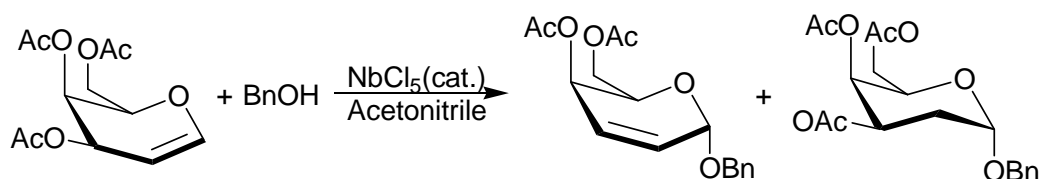
Entry	Substrate	Alcohol	Product	Time(min)	Yield(%)
1			 1a	2.0	95
2			 1b	2.5	92
3			 1c	2.0	80
4			 1d	1.5	85
5			 1e	2.0	93
6			 1f	2.0	97
7			 1g	3.0	84
8			 1h	2.0	87
9			 1i	3.0	90
10			 1j	3.0	87

Table 1. *NbCl₅* promoted 2,3-unsaturated Ferrier rearranged products from glucal

A recent report by Sinou et. al. provided the plausible mechanistic details of the NbCl_5 catalyzed Ferrier reaction.²³ Niobium complexes with the carbonyl oxygen atom of the *O*-acetyl group at C-3. Then one of the chlorine atom of NbCl_5 could attack at C-1 via a 8-membered transition state to furnish 4,6-di-*O*-acetyl-2,3-dideoxy- β -D-*erythro*-hex-2-enopyranosyl chloride by a $\text{S}_{\text{N}}2'$ mechanism. This instantaneously reacts with the alcohol to form the Ferrier product. The exclusive formation of the α -anomer was justified with the explanation that the displacement of the chlorine atom with that of the alcohol is an instantaneous reaction which takes place before the anomerization of the intermediate halide (Figure 7). The generality of this reaction with NbCl_5 was shown by exposing per-*O*-

acetylated glucal to various alcohols including allylic, benzylic, aliphatic, phenolic and monosaccharide donors to the corresponding glycosides of mono- and disaccharides. As it is evident from the results revealed in Table 1, all the donors reacted with great ease in shorter time to give α -anomer as the sole product of the reaction in excellent yields. All the glycoside products were confirmed by comparing with the reported values.^{20,21}

Encouraged by our success, we decided to test our protocol on 3,4,6-tri-*O*-acetyl galactal. Treatment of per-*O*-acetylated galactal in the presence of an alcohol with NbCl₅(cat.) in acetonitrile and irradiation with microwaves for less than 3 min resulted in the formation of the Ferrier product 4,6-di-*O*-acetyl-2,3-dideoxy-D-*threo*-hex-2-eno- α -D-galactopyranoside along with a minor amount of 2-deoxy- α -D-*lyxo*-hexopyranoside (Scheme 3). It is important to mention that not so many synthetic procedures are available in the literature for the Ferrier reaction involving 3,4,6-tri-*O*-acetyl galactal and many report the formation of 2-deoxy



Scheme 3. NbCl₅ catalyzed Ferrier reaction of per-*O*-acetylated galactal

compound as the frustrating side product in considerable yield. In order to check the versatility of this reaction we exposed a variety of aglycones to per-*O*-acetylated galactal (Table 2). As is evident from the results shown, all of the reactions were completed within no time and enhanced amount of Ferrier product was obtained in contrast to the earlier published results. The ratio of the Ferrier product versus the 2-deoxy compound was found to be 4:1 after the isolation of the respective compounds by conventional silica gel column chromatography. ¹H and ¹³C NMR spectra of all the various saccharides synthesized were in conformity with those of reported values.^{20,21}

Successful in our efforts of exploiting the chirality and structural complexity of the carbohydrates, we were able to synthesize a library of 20 members. All the constituents/members of the library are chirally pure and structurally diverse even though they are derived from the same skeleton. The derivatives of 2,3-unsaturated glycosides constitute the structural motifs of several antibiotics and these are also found to reduce the plasma cholesterol and triglyceride levels significantly in mice.

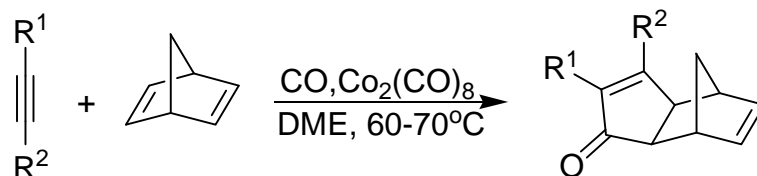
Entry	Substrate	Alcohol	Product	Time(min)	Yield(%)
1			 2a	1.5	79
2			 2b	2.5	73
3			 2c	3.0	72
4			 2d	2.0	69
5			 2e	2.0	74
6			 2f	1.0	80
7			 2g	2.0	76
8			 2h	2.5	73
9			 2i	3.0	74
10			 2j	3.0	68

Table 2. *NbCl₅* catalysed 2,3-unsaturated Ferrier rearranged products from galactal.

Second Level of Diversity: Pauson-Khand Reaction

The 2,3-unsaturated bond of the Ferrier product leaves enough scope to explore the chemical space around the glycols for the further extension of diversity in the library. The double bond in the pyran ring can be subjected to various complexity generating reactions like, stereoselective dihydroxylation, hydrogenation, epoxidation and amino hydroxylation. However the above mentioned reactions have been exploited for the *DOS* rather frequently but the utility of the Pauson-Khand reaction (PKR) has started recently. The Pauson-Khand

reaction was an unexpected finding that was discovered in 1973 by P. L. Pauson and his student I. U. Khand while attempting to synthesize new organometallic compounds (arene-cobalt complexes).²⁴ PKR incorporates three components in a formal [2+2+1] cycloaddition process- an alkyne, an alkene and carbon monoxide molecule with the aid of cobalt carbonyl complexes to form a α,β -cyclopentenone (Scheme 4).



Scheme 4. *General Pauson-Khand Reaction*

The reaction at that time has serious practical drawbacks, like, use of stoichiometric quantities of toxic and costly $\text{Co}_2(\text{CO})_8$, harsh reaction conditions, high pressure use of poisonous carbon monoxide and low functional group compatibility. Discovered using $\text{Co}_2(\text{CO})_8$, this transformation can now be promoted by complexes of titanium, zirconium, molybdenum, tungsten, iron, ruthenium, rhodium, nickel and iridium.^{25a} In the recent years, large amount of literature has been generated related to PKR variants, ranging from photochemical to thermal, stoichiometric to catalytic, achiral to asymmetric and metallic to heterobimetallic reactions.^{25b} As the reaction continues to be thoroughly investigated, the scope of this reaction broadened, catalysis has been achieved, milder versions have been developed and its application in organic synthesis has been realised. One of the major highlights in the advancement of PKR has been the introduction of promoters, the minimal use of cobalt catalyst and with the reaction being successfully carried out in carbon monoxide free atmosphere.^{25c} PKR has now emerged as a powerful method for the synthesis of cyclopentenone derivatives and one important factor which has contributed to its widespread use is the multiple bond formation in a single step with high degree of predictivity of the regio- and stereochemical outcome of the substrates, both in intermolecular and intramolecular versions. The schematic representation of the reaction mechanism is shown below (Figure 8).

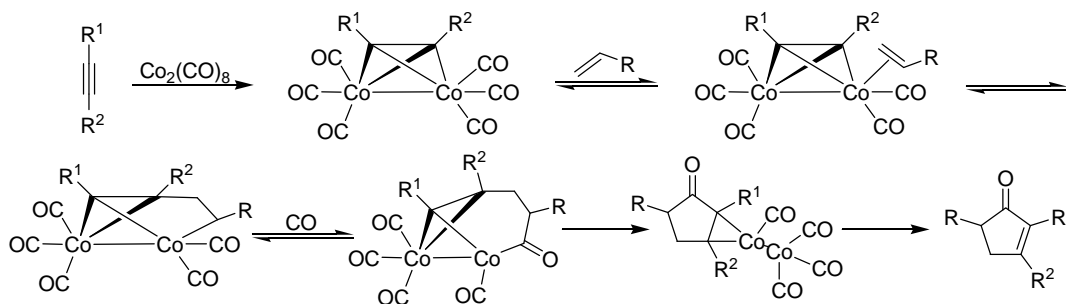
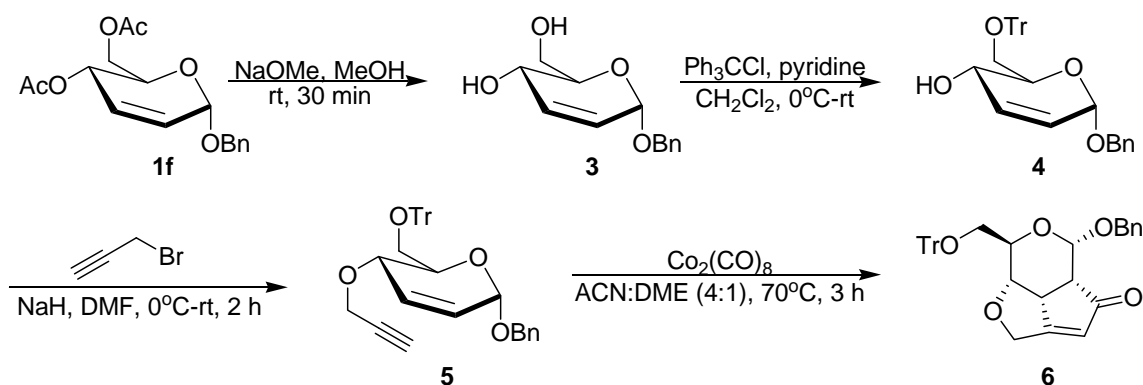


Figure 8. *Mechanism of the Pauson Khand Reaction*

Although the synthetic utility of this reaction has been well documented, its application on the carbohydrate templates for the generation of *DOS* libraries has not been exploited to the potential. Recently, Schreiber group treated suitably substituted per-*O*-acetyl glucal with various propargyl alcohols in the presence of $\text{BF}_3 \cdot \text{Et}_2\text{O}$ to obtain 2,3-unsaturated enynes, which were then subjected to PKR reaction to obtain a library of tricyclic compounds.²⁶ However, we envisaged that our α -glycosides, (Ferrier products) obtained using a variety of aglycones, with the functional group manipulation should enable the installation of propargyl moiety at the *C*-4 position so that the Pauson-Khand reaction can be carried out stereoselectively to obtain fused tricyclic enones.

To begin our investigation, benzyl 4,6-di-*O*-acetyl-2,3-dideoxy-D-*erythro*-hex-2-eno- α -D-glucopyranoside (**1f**) was deacetylated under Zemplen conditions (NaOMe , MeOH) to afford the diol **3** and the primary hydroxyl group was selectively blocked as a trityl ether using pyridine, trityl chloride in dichloromethane as the solvent to obtain the compound, **4**. The remaining allylic hydroxyl group of **4** was converted to propargyl ether employing propargyl bromide, NaH in DMF to enyne **5** (Scheme 5). In the ^1H NMR spectrum of **5**, characteristic acetylenic methine proton was observed at δ 2.24 ppm as a triplet ($J = 2.4$ Hz), benzylic methylene group was identified at δ 4.78 ppm ($J = 2.4$ Hz) as an AB type quartet (ABq), and olefinic resonances were noticed at δ 5.81 (td, $J = 2.1, 10.2$ Hz) and δ 6.07 (d, $J = 10.3$ Hz) ppm with the rest of the resonances in accordance with the assigned structure. In the ^{13}C NMR spectrum of **5**, resonances corresponding to the anomeric and acetylenic carbons were noticed at δ 93.3 and δ 74.4 ppm, respectively.



Scheme 5. Schematic representation of synthesis of enone **6**

Having the required enyne **5** in hand, we dwelt upon carrying out Pauson-Khand reaction; accordingly, the enyne **5** was treated with stoichiometric quantities of $\text{Co}_2(\text{CO})_8$ for 1 hour under N_2 atmosphere in anhydrous CH_2Cl_2 and was passed through a pad of silica gel to obtain the $\text{Co}_2(\text{CO})_6$ -alkyne complex as thick red oil. At this stage, cleavage of the cobalt

complex and carbonyl insertion became an arduous task. Initially we used *N*-methylmorpholine *N*-oxide (NMO) in CH₂Cl₂ to effect cleavage of the cobalt complex, and we obtained **6** in 70% yield. However, to our surprise, the reaction was sluggish (3 days) and sensitive to the quality of NMO; hence, we sought alternate protocols. Heating the Co₂(CO)₆-alkyne complex in the presence of other cleavage reagents such as cyclohexylamine or acetonitrile did not result in the formation of the **6** but led to the decomposition of the cobalt complex. Treatment of the Co₂(CO)₆-alkyne complex with H₂O₂ furnished the desired tricyclic enone **6** in 30% yield. Use of dimethoxyethane and acetonitrile at 70 °C as the Co₂(CO)₆-alkyne complex cleavage cocktail afforded the enone **6** in 92% yield (Scheme 5).

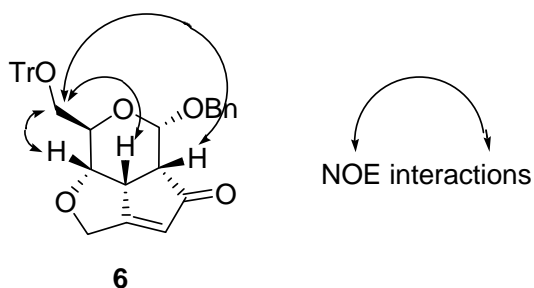
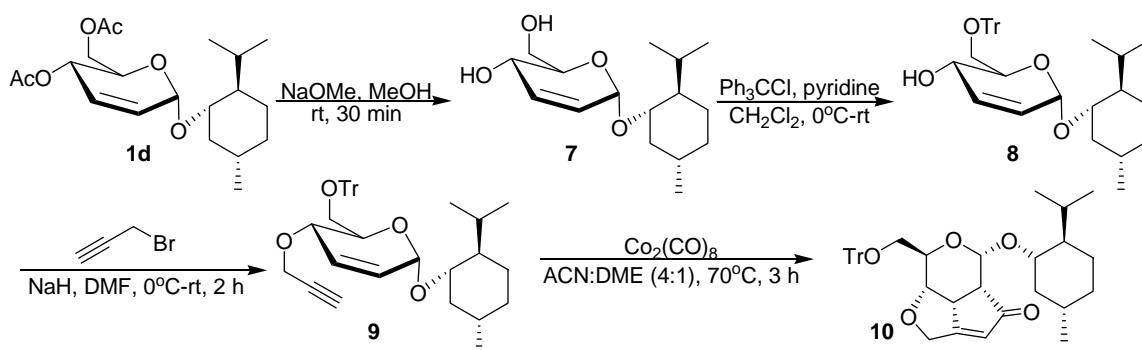
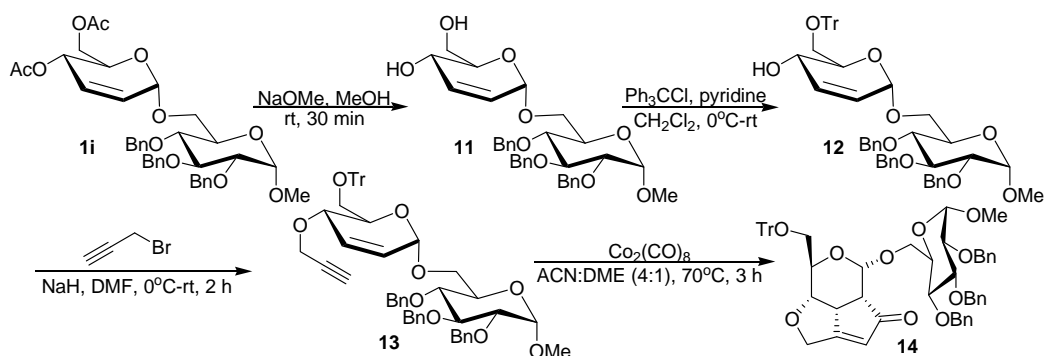


Figure 9. NOE correlations of enone **6**.

The ¹H NMR spectrum of **6** showed no resonances corresponding to the enyne, and new resonances characteristic of olefinic proton associated with the enone were observed at δ 6.00 ppm as a singlet. Resonances due to the benzylic -CH₂ group were noticed at δ 4.67 ppm as an ABq (*J* = 12.2 Hz), and that of anomeric proton were evident at δ 5.22 ppm as a doublet (*J* = 7.8 Hz). It is pertinent to mention that the Pauson-Khand annulation occurred in a highly diastereoselective fashion; the NOESY spectrum showed NOE cross-peaks between H-2, H-3, H-4, and H-6 (Figure 9). Configuration at the C-4 position plays a predominant role in dictating the chirality of the Pauson-Khand annulations as the Co₂(CO)₆-alkyne will be below the plane. As a consequence, carbonyl insertion can take place from below only; hence, the observation of a single diastereomer **6** with the simultaneous formation of three chiral centres.

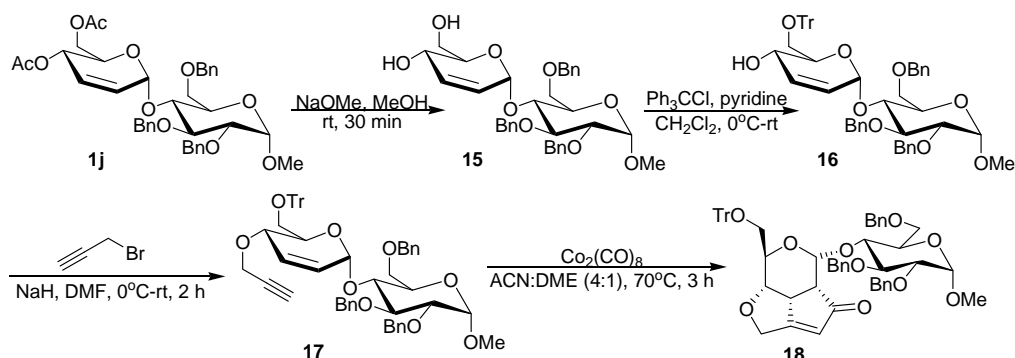


Scheme 6. Schematic representation for the synthesis of enone **10**.

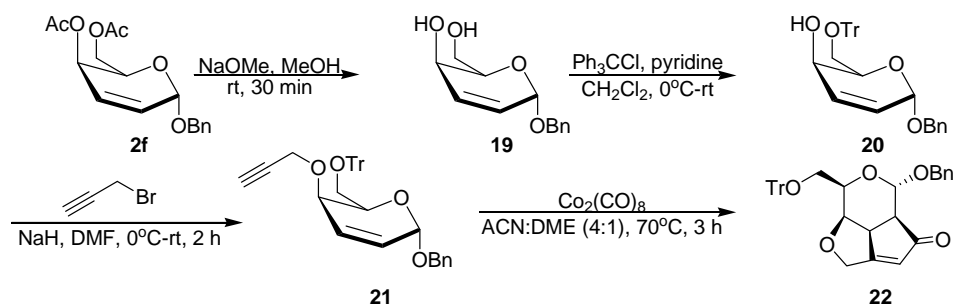


Scheme 7. Schematic representation for the synthesis of enone **14**

Excited with our success and in progression with the introduction of more structural complexity and diversity, we subjected the Ferrier products **1d**, **1i** and **1j** to the same set of reactions to get enones **10**, **14** and **18** via intermediates **7-9**, **11-13** and **15-17** as shown in Scheme 6, 7 and 8 respectively. The practicality of the whole procedure was established, as both of them resulted in the formation of the single diastereomers. The products were confirmed by their ^1H and ^{13}C NMR spectra. Thrilled after having identified a practical route for the synthesis of fused tricyclic enones with the successful formation of the three glucal derived cyclopentenones, we decided to check the generality of the reaction on the other glycals namely, galactal, xylal and rhamnal. This will add another dimension to our library and would offer the much needed stereochemical diversity to our *DOS* library. Not only the resulting library will be uniquely diverse but would also explore the new boundaries of the chemical space. Persistent in our synthetic endeavours, we then subjected benzyl 4,6-di-*O*-acetyl-2,3-dideoxy-*D*-erythro-hex-2-eno- α -*D*-galactopyranoside (**2f**), the Ferrier product, to tritylation followed by alkylation of the resulting allylic alcohol with propargyl bromide, NaH in DMF to yield compound **21**. In the ^1H NMR spectrum of the galactal derived enyne **21**, the distinguishable acetylenic proton was observed at δ 2.28 Hz as the triplet ($J = 2.28$ Hz), the olefinic protons were located at δ 6.03 (dd, $J = 2.79, 9.99$ Hz) and δ 6.21 (td, $J = 0.77, 5.19, 9.90$ Hz). The ^{13}C NMR spectrum of enyne **21**, resonances corresponding to anomeric and



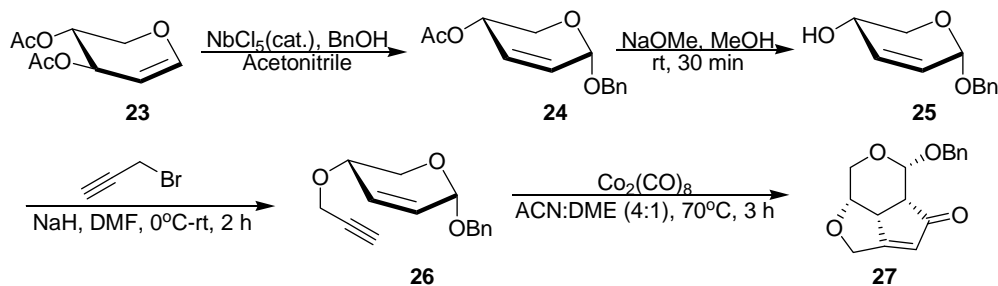
Scheme 8. Schematic representation for the synthesis of enone **18**.



Scheme 9. Schematic representation for the synthesis of enone **22**

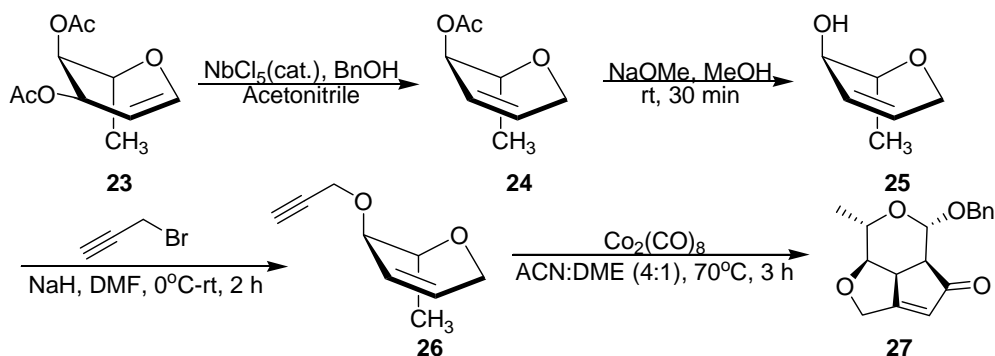
acetylenic carbon were noticed at δ 92.9 and δ 74.5 ppm respectively. Rest of the spectrum was in accordance with the structure. Exposing enyne **21** to Pauson-Khand reaction conditions, we were thrilled to have the enone **22** in 90% yield and as a single diastereomer (Scheme 9). The product was confirmed by the presence of the distinguishable resonances of olefinic proton associated with the enone at δ 5.93 as a singlet and with the disappearance of the respective acetylenic peaks in the ^1H NMR spectrum. The anomeric proton was observed at δ 97.5 ppm in the ^{13}C NMR spectrum. In accordance with the earlier observation, the formation of **22** was envisioned to occur in a highly diastereoselective fashion. It is heart warming to mention that contrary to glucal derived enyne **23**, the insertion of the olefin into the cobalt-enyne complex occurs from above the plane due to the configuration at C-4 atom and results in the formation of a single diastereomer which is having complementary stereochemistry to the compound **24**.

Continuing in our attempts to provide more structural complexity to our *DOS* and improve the diversity of the carbohydrate library, we envisaged the formation of the rhamnal and xylal based enones. Starting with the per-*O*-acetylated xylal, we performed the Ferrier reaction employing benzyl alcohol as the aglycone in the presence of one mole percent of NbCl_5 in acetonitrile under the microwave conditions for one min. The corresponding monosaccharide, **24**, was subjected to same set of reactions as cited above for the sister template to yield the enyne, **26**. On the examination of ^1H NMR spectrum of **26**, the distinctive triplet of the methine proton was observed at δ 2.45 ($J = 2.28$ Hz) along with the olefinic protons at δ 6.01 (dd, $J = 2.72, 10.23$ Hz) and δ 6.14 (td, $J = 1.12, 4.98$ Hz).



Scheme 10. Schematic representation for the synthesis of enone **27**

The ^{13}C NMR spectrum showed anomeric and alkyne carbon at δ 92.2 and 74.7 ppm respectively. The synthesis of the PKR product of **26** was carried out by heating the cobalt-alkyne complex in acetonitrile-dimethoxyethane at 70°C for 3 h to affect the cleavage of the complex and result in the enone **27** (Scheme 10). The tricyclic enone, **27** was confirmed by the disappearance of the acetylenic proton in the ^1H NMR spectrum and the appearance of a singlet at δ 6.01, the olefinic peak characteristic of the enone. The ^{13}C NMR spectrum also confirmed the product with the emergence of the carbonyl peak at δ 207.3 ppm. All the other peaks were in agreement with the assigned structure.



Scheme 11. Schematic representation for the synthesis of enone **27**.

To add another feather into our carbohydrate template based library, we performed the same string of reactions as mentioned for per-*O*-acetylated xylal on per-*O*-acetylated rhamnal to acquire the final natural product like enone **32** (Scheme 11). All the substrates were confirmed through NMR and elemental analysis. After the successful completion of the synthesis of six natural product-like tricyclic enones, we performed an initial survey of the substrate scope. We could successfully develop a general protocol for the synthesis of various mono- and disaccharides derived from a variety of glycals such as per-*O*-acetylated glucal, galactal, xylal, and rhamnal adopting our optimized aforementioned conditions. Not only are the six compounds structurally complex but also offer elegant stereochemical diversity.

Third Level of Diversity: *Michael Addition*

The Michael addition of nucleophiles to electron deficient alkenes is one of the most important reactions in organic chemistry for constructing carbon-sulfur bond. Among various nucleophilic additions, the reaction of thiols to form a carbon-sulfur bond constitutes a key reaction in the biosynthesis as well as in the synthesis of biologically active compounds such as calcium antagonist diltiazem. The α,β -unsaturation of the products from Pauson-Khand products offers enough possibility for the diversification for exposing them to the Michael addition using thiols. In continuation of our efforts for offering more complexity to our

carbohydrate library, we chose six mono-thiols and three tricyclic enones synthesized *vide supra* for our studies. Diversification through thiolates was chosen because of their propensity toward nucleophilic addition to enones under mild conditions coupled with their ready availability and thiol-substituted compounds were found to be more potent than corresponding enones.

Accordingly, we performed the thiolate nucleophilic conjugate addition to enones and found that *N,N*-dimethylaminopyridine (DMAP) catalyzes thiolate addition very efficiently. In a typical experiment, a solution containing ethane thiol was stirred with enone **6** in the presence of a catalytic quantity of DMAP in toluene for 12 h to afford thiol incorporated tricyclic ketone **32a**. ^1H NMR spectrum of **32a** confirmed the product with the disappearance of the enone proton around δ 6.06 Hz and the emergence of the desired signals of thiol around δ 1.22 (t, 3H, $J = 7.31$ Hz). The ^{13}C NMR spectrum of **32a** revealed anomeric carbon at δ 93.8 ppm and all the other resonances were in conformity with the assigned structure. It is pertinent to mention that the conjugate addition favoured a single diastereomer that can be rationalized based on the *bowl*-like conformation of enone; hence, the thiolate addition can take place from the least sterically demanding phase only. Consecutively, the Michael addition was performed with the other thiols to yield substrates **32b-32f** as shown in Figure 10.

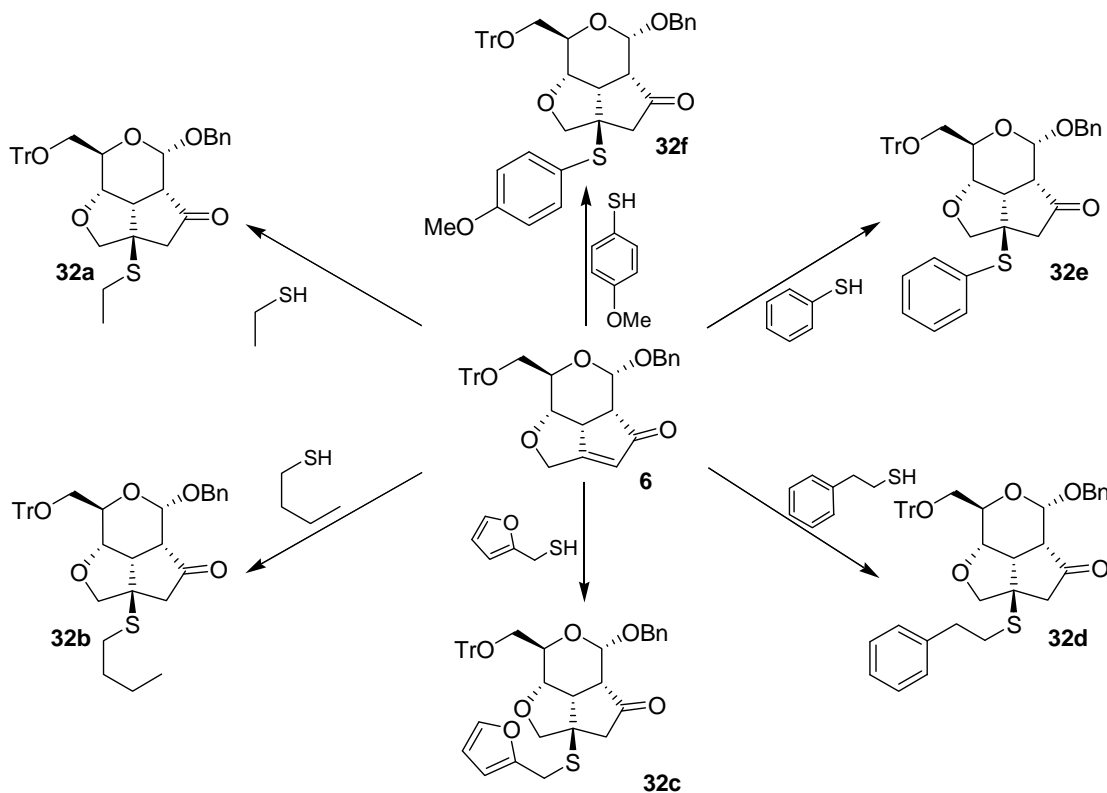


Figure 10. Michael addition onto enone **6**

With the established protocol and successful formation of the six Michael adducts from enone **6**, we tried the same reaction on the xylal-derived enone **27**. As a general synthetic procedure, the enone, **27** was taken in toluene and allowed to stir in the presence of ethane thiol to afford the product **33a** as a single isomer only. The stereochemistry of the resulting Michael adduct was defined taking a cue from the spatial arrangement of **33a**, as the fundamental skeleton is identical to both of them. Hereafter, the other thiols were promptly were reacted with the enone **27** to afford the Michael products **33b-33f** (Figure 11).

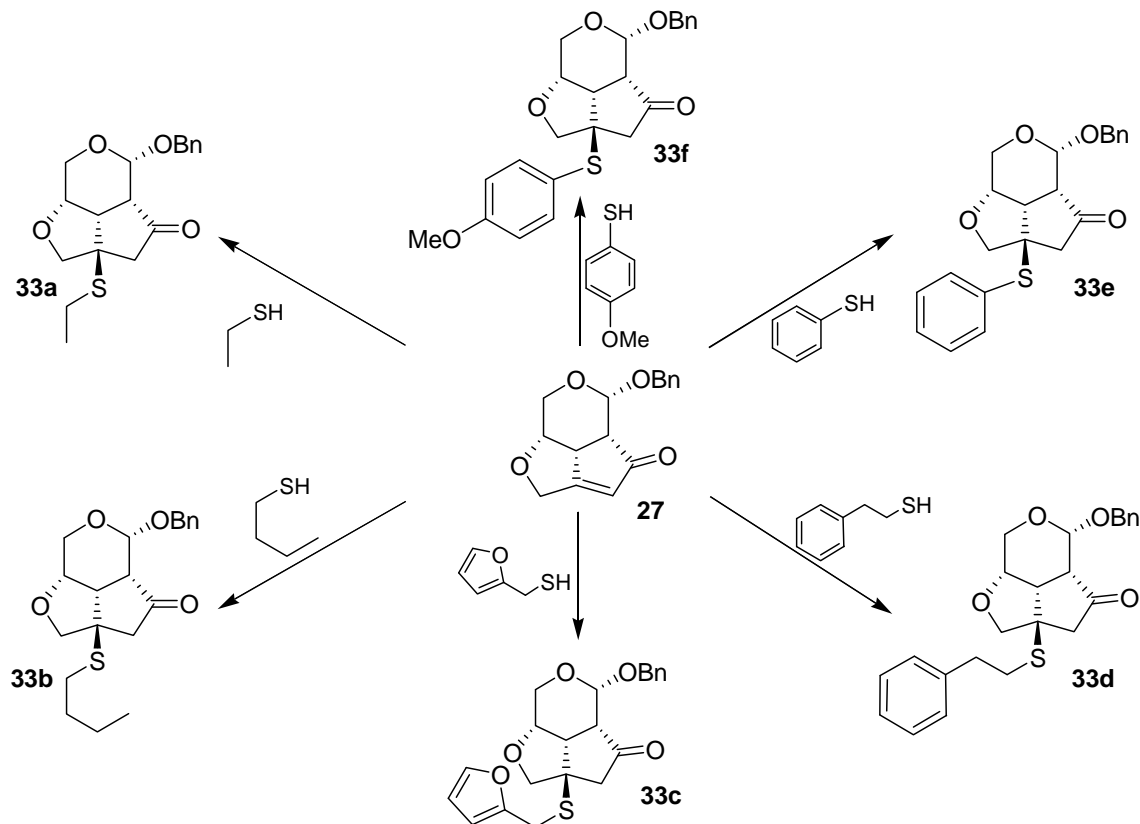


Figure 11. Michael addition onto enone **27**.

Delighted with the synthesis of the Michael adducts for glucal derived library, we undertook synthesis of the same Michael products with the galactal derived enone to have more stereochemical flexibility in our library. We envisaged that the conjugate addition onto enone **22** would afford the complementary stereochemistry to compound **6**. This could be rationally explained on the grounds that both them are having complementary structures to each other, bowl-like structures, and the addition takes place from the least hindered side to afford single isomer only. In the usual experiment, the enone **22** was dissolved in toluene and was stirred with ethane thiol for 12 hr. After the completion of the reaction, inspection of the ^1H NMR spectrum of **34a** revealed the characteristic resonances of thiol at δ 1.24 (t, 3H, $J = 7.48$ Hz)

and the ^{13}C NMR spectrum also confirmed the product with all the carbons at the expected positions along with the anomeric carbon at δ 95.5 ppm. Consecutively, the remaining thiol adducts were synthesized following the same protocol to yield the compounds **34b-34f** (Figure 12).

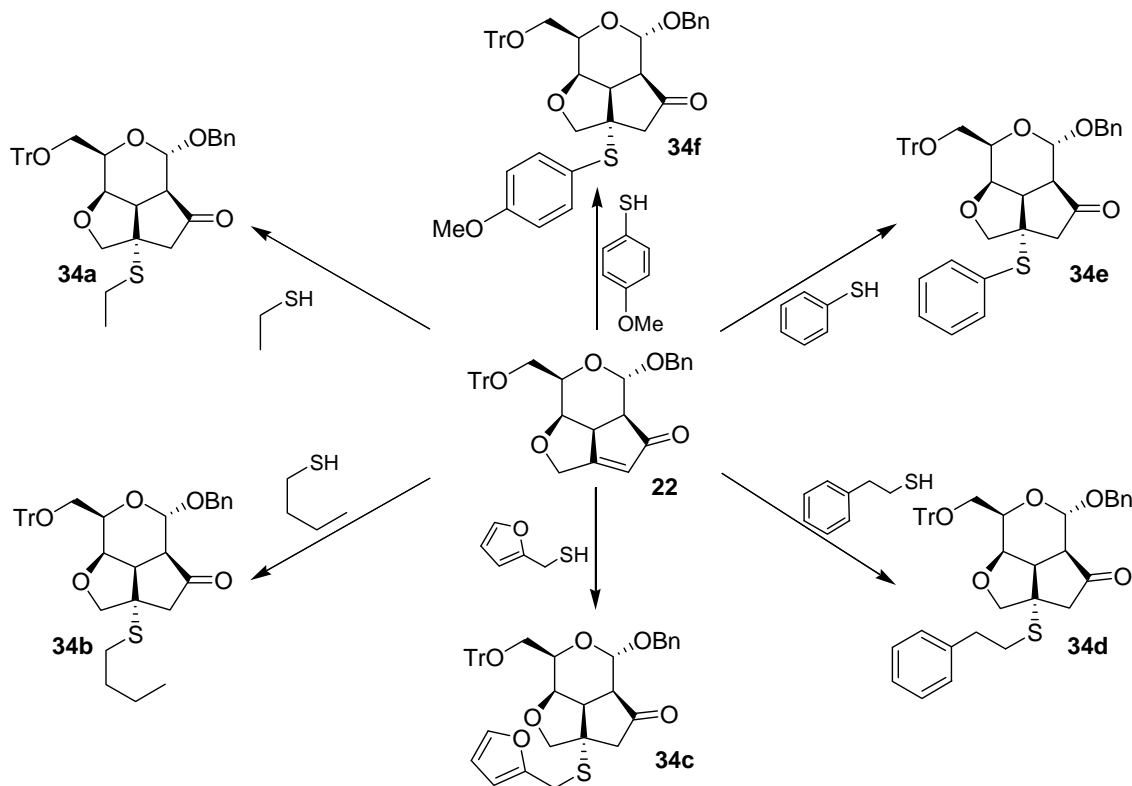


Figure 12. Michael addition onto enone **22**

In conclusion, we have developed a practical protocol for the synthesis of natural product-like tricyclic compounds and a total of 25 tricyclic compounds were synthesized as a pilot library, using the Ferrier, Pauson-Khand reactions and the conjugate thiolate addition to enones were performed under mild reaction conditions to afford a thiol-substituted library.²⁷ It is pertinent to mention that the complexity generating reactions (viz., the Ferrier, Pauson-Khand, and Michael addition reactions) were highly diastereoselective thereby enabling chirally pure, oxygen-rich, tricyclic derivatives from easily accessible glycals. A little examination of the pilot library beautifully illustrates the fact that the collection of the compounds possesses a high level of structural complexity and diversity. On one hand, the complementary spatial arrangement of the glucal and galactal provided the stereochemical flexibility on the other hand the structural dissimilarity of the xylal and rhamnol provides an advantage of structural variation as well.

Chapter 1: Experimental Section

General experimental procedures for the Ferrier reaction:

To a mixture of per-*O*-acetyl glycal (1 mmol) and alcohol (1.5 mmol) in acetonitrile (5 mL) was added NbCl₅ (0.01 mmol). The resulting solution was exposed to microwave at ambient temperature for the specified time. The reaction mixture was then diluted with water and extracted two times with ethyl acetate. The combined organic extracts were dried over anhydrous sodium sulfate and concentrated *in vacuo*. The product was purified by silica gel column chromatography using ethyl acetate and light petroleum (60-80°C) as the mobile phase.

General experimental procedures for deacetylation of 2,3-Unsaturated *O*-glycosides:

NaOMe (3 mmol) was added to the solution of 2,3-unsaturated *O*-glycoside in anhydrous methanol (5 mL). The resulting solution was stirred at room temperature until the reaction showed completion by TLC (takes about 30 min). The solvent was removed *in vacuo*, and the reaction mixture was diluted with water and extracted three times with ethyl acetate. The combined organic extracts were dried over anhydrous sodium sulfate and concentrated *in vacuo* to yield the diol, which was used without further purification in the next step.

General experimental procedures for *O*-Tritylation of 2,3-Unsaturated *O*-glycosides:

The diol (1 mmol) prepared *vide supra* was dissolved in anhydrous pyridine (10 mL), and the solution was cooled to 0°C. Trityl chloride (1.5 mmol) was added to the above solution, and the reaction mixture was brought to room temperature and allowed to stir for 24 h. The reaction mixture was poured into water and was extracted three times with ethyl acetate. The combined organic extracts were washed with brine and dried over sodium sulphate. The solvent was removed under reduced pressure to yield the corresponding *O*-trityl derivative, which was purified by silica gel column chromatography using ethyl acetate and light petroleum (60-80°C).

General experimental procedures for *O*-propargylation:

To an ice-cooled solution of the trityl derivative of the glycal (1 mmol) in anhydrous DMF (5 mL) was added sodium hydride (1.5 equiv, 60% oil suspension) and stirred for 1 h at room temperature. Propargyl bromide (1.5 equiv) was introduced dropwise to the mixture at 0 °C and stirred at room temperature for 1 h. The resulting suspension was quenched with saturated ammonium chloride and extracted three times with ethyl acetate. The combined organic extracts were washed with brine solution and concentrated under reduced pressure to give the

crude *O*-propargyl derivative that was purified by silica gel column chromatography using ethyl acetate and light petroleum (60-80°C).

General experimental procedures for Pauson-Khand Reaction:

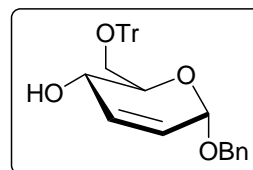
To an ice-cooled solution of alkyne in anhydrous dichloromethane was added $\text{Co}_2(\text{CO})_8$ (1.2 equiv) and stirred at room temperature for 1 h. The reaction mixture was adsorbed on silica gel and was eluted with ethyl acetate and light petroleum to yield the cobalt-ene-yne complex as thick red oil. Subsequently, the Co-alkyne complex was redissolved in anhydrous 1,2-dimethoxy ethane (5 mL) and anhydrous acetonitrile (20 mL) and was refluxed until the colour of the solution changed from thick red to greyish black. The solution was filtered through a pad of silica gel and concentrated to give light yellow oil, which was purified by silica gel column chromatography using ethyl acetate and light petroleum (60-80°C).

General experimental procedures for thiolate additions on enones:

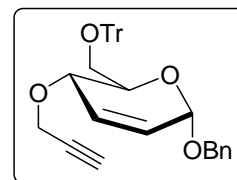
The tricyclic enone (1 equiv) was dissolved in anhydrous toluene (10 mL/mmol), and the thiol (2 equiv) was added to it. The resulting solution was stirred in the presence of catalytic amount of DMAP until the reaction showed completion by TLC (typically takes about 12-24 h). Thereafter the solvent was removed in vacuo, and the adduct was purified by silica gel column chromatography using ethyl acetate and light petroleum (60-80°C).

Compound Characterization Data:

Compound 4: $[\alpha]_D$ (CHCl_3 , c 0.9) = + 50.6. IR (cm^{-1}) = 3455. ^1H NMR (200 MHz, CDCl_3): δ 3.29-3.46 (m, 2H), 3.83-3.94 (m, 1H), 4.02-4.12 (m, 1H), 4.72 (ABq, 2H, $J = 11.77$ Hz), 5.05 (s, 1H), 5.76 (td, 1H, 2.38, 10.17 Hz), 5.92 (td, 1H, 1.56, 10.16 Hz), 7.23-7.35 (m, 14H), 7.44-7.50 (m, 6H). ^{13}C NMR (50 MHz, CDCl_3): δ 61.8, 63.4, 70.0, 71.5, 93.4, 125.5, 127.7-129.6, 137.5, 133.6, 142.5. Anal. Calcd. for $\text{C}_{32}\text{H}_{30}\text{O}_3$: C, 83.09; H, 6.54. Found C, 82.95; H, 6.64.



Compound 5: $[\alpha]_D$ (CHCl_3 , c 0.7) = +33.6°. IR (cm^{-1}) = 3303. ^1H NMR (200 MHz, CDCl_3): δ 2.24 (t, 1H, $J = 2.39$ Hz), 3.19 (dd, 1H, $J = 5.44, 10.31$ Hz), 3.37 (dd, 1H, $J = 1.34, 10.23$ Hz), 4.01 (m, 3H), 4.23 (dd, 1H, $J = 1.18, 9.46$ Hz), 4.78 (ABq, 2H, $J = 11.80$ Hz), 5.15 (d, 1H, $J = 1.62$ Hz), 5.81 (td, 1H, $J = 2.14, 10.22$ Hz), 6.07 (d, 1H, $J = 10.28$ Hz), 7.15-7.42 (m, 14H), 7.47-7.54 (m, 6H). ^{13}C NMR (50 MHz, CDCl_3): δ 56.6, 63.1, 69.4, 69.7, 70.1, 74.4, 79.7, 86.4,

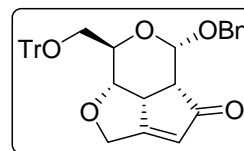


93.3, 126.7-128.8, 131.0, 137.9, 144.1. Anal. Calcd. for C₃₅H₃₂O₄: C, 81.37; H, 6.24. Found: C, 80.96; H, 6.21.

Compound 6: [α]_D (CHCl₃, *c* 0.9) = +177.9°. IR (cm⁻¹) = 1716, 1647.

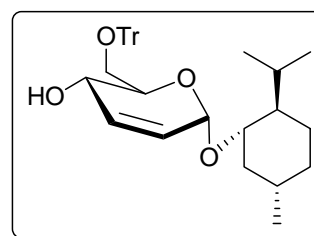
¹H NMR (500 MHz, CDCl₃): δ 3.21-3.33 (m, 4H), 3.59 (td, 1H, *J* = 1.98, 7.34 Hz), 3.95 (t, 1H, *J* = 9.13 Hz), 4.45 (m, 2H), 4.67 (ABq, 2H, *J* = 12.18 Hz), 5.22 (d, 1H, *J* = 7.76 Hz), 6.00 (s, 1H), 7.12-7.31 (m, 14H), 7.33-7.45 (m, 6H).

¹³C NMR (50 MHz, CDCl₃): δ 45.5, 1.5, 64.0, 65.1, 65.7, 68.9, 71.2, 86.8, 94.3, 125.2, 126.8-128.7, 137.3, 143.9, 178.36, 206.7. Anal. Calcd. for C₃₆H₃₂O₅: C, 79.39; H, 5.92. Found: C, 79.89; H, 6.20. MALDI-TOF: mol wt calcd 544.64; found 567.17 (M⁺ + 23 for Na).



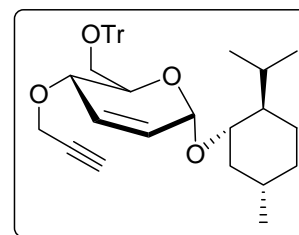
Compound 8: [α]_D (CHCl₃, *c* 0.7) = -48.35°. IR (cm⁻¹) = 3460.

¹H NMR (500 MHz, CDCl₃): δ 0.64-1.01 (m, 11H), 1.12-1.31 (m, 3H), 1.58-1.65 (m, 2H), 2.05-2.12 (m, 2H), 2.61 (d, 1H, *J* = 4.68 Hz), 3.25-3.53 (m, 3H), 3.85-3.95 (m, 1H), 4.08-4.15 (m, 1H), 5.01 (s, 1H), 5.73 (td, 1H, *J* = 2.28, 10.20 Hz), 5.88 (m, 1H), 7.17-7.34 (m, 9H), 7.43-7.49 (m, 6H). ¹³C NMR (50 MHz, CDCl₃): δ 16.2, 21.1, 22.3, 23.2, 25.5, 31.7, 34.3, 43.5, 48.8, 65.5, 66.6, 69.4, 80.3, 87.4, 95.8, 126.0-128.5, 132.6, 143.5. Anal. calcd. for C₃₅H₄₂O₃: C, 82.31; H, 8.29. Found: C, 82.39; H, 8.32.



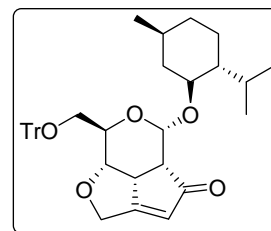
Compound 9: [α]_D (CHCl₃, *c* 1.6) = +8.76°. IR (cm⁻¹): 3216. ¹H

NMR (200 MHz, CDCl₃): δ 0.72, 0.75, 0.79, 0.83, 0.89, 0.93 (6s, 9H), 0.98-1.11 (m, 3H), 1.19-1.70 (m, 4H), 2.05-2.35 (m, 3H), 3.23 (dd, 1H, *J* = 4.04, 10.23 Hz), 3.38-3.58 (m, 2H), 3.85-4.08 (m, 3H), 4.24 (td, 1H, *J* = 0.9, 9.23 Hz), 5.14 (s, 1H), 5.73-6.10 (m, 2H), 7.16-7.35 (m, 9H), 7.40-7.56 (m, 6H). ¹³C NMR (50 MHz, CDCl₃): δ 16.2, 21.1, 22.2, 23.2, 25.6, 31.8, 34.4, 43.3, 48.8, 56.5, 63.1, 69.0, 70.5, 74.3, 79.8, 80.5, 86.4, 96.1, 126.8-128.8, 130.5, 143.9, 144.0. Anal. Calcd. for C₃₈H₄₄O₄: C, 80.82; H, 7.85. Found: C, 80.49; H, 7.89.



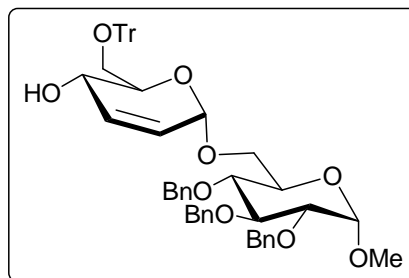
Compound 10: [α]_D (CHCl₃, *c* 1.0) = +147.5°. IR (cm⁻¹): 1749,

1454. ¹H NMR (200 MHz, CDCl₃): δ 0.60, 0.63 (2s, 3H), 0.76, 0.79, 0.86, 0.92 (4s, 6 H), 0.83-0.97 (m, 3H), 1.08-1.22 (m, 2H), 1.58 (m, 3H), 2.11 (m, 2 H), 3.24-4.45 (m, 5H), 4.27 (dd, 1 H, *J* = 8.40, 9.61 Hz), 4.56 (bs, 2 H), 5.26 (d, 1H, *J* = 7.59 Hz), 5.99 (d, 1 H, *J* = 1.51 Hz), 7.10-7.35 (m, 9H), 7.42-7.53 (m, 6H). ¹³C NMR (50 MHz, CDCl₃): δ 15.8, 21.4, 22.2, 22.6, 24.7, 31.6, 34.2, 42.3, 45.7, 49.0, 52.2, 63.8, 65.2, 65.4, 71.3, 80.1, 86.8, 96.3, 125.0,



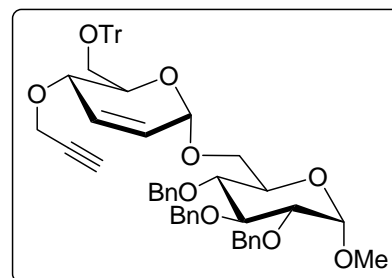
126.9-128.9, 144.0, 178.1, 206.7. Anal. Calcd. for $C_{39}H_{44}O_5$: C, 79.02; H, 7.48. Found: C, 79.38; H, 6.96.

Compound 12: $[\alpha]_D$ ($CHCl_3$, c 1.26) = -2.46° . IR (cm^{-1}) = 3450. 1H NMR (200 MHz, $CDCl_3$): δ 2.34 (d, 1H, J = 4.69 Hz), 3.18-3.34 (m, 2H), 3.36 (s, 3H), 3.49-3.61 (m, 2H), 3.69-3.85 (m, 3H), 3.93-4.05 (m, 2H), 4.12-4.20 (m, 1H), 4.54-4.69 (m, 3H), 4.76-4.83 (m, 3H), 4.98 (d, 1H, J = 10.83 Hz), 5.04 (m, 1H), 5.75 (td, 1H, J = 2.25,



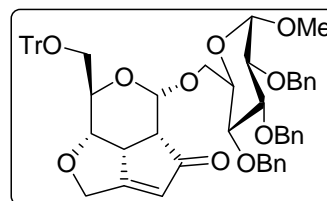
10.25 Hz), 5.90 (td, 1H, J = 1.25, 10.34 Hz), 7.14-7.40 (m, 30H). ^{13}C NMR (50 MHz, $CDCl_3$): δ 55.1, 64.7, 66.0, 66.8, 70.1, 73.3, 75.0, 75.6, 77.8, 80.0, 82.0, 87.2, 94.5, 98.0, 125.9, 127.1, 127.6-128.6, 132.8, 138.2, 138.8, 143.5. Anal. Calcd. for $C_{53}H_{54}O_9$: C, 76.30; H, 6.76. Found: C, 76.45; H, 6.84.

Compound 13: $[\alpha]_D$ ($CHCl_3$, c 1.06) = $+31.2^\circ$. IR (cm^{-1}) = 3205. 1H NMR (200 MHz, $CDCl_3$): δ 2.25 (t, 1H, J = 2.41 Hz), 2.98 (dd, 1H, J = 3.92, 10.24 Hz), 3.25-3.42 (m, 4H), 3.44-3.68 (m, 3H), 3.70-3.85 (m, 3H), 3.91-4.20 (m, 4H), 4.31-4.42 (m, 1H), 4.50-5.05 (m, 7H), 5.75-5.90 (m, 1H), 6.07 (m, 1H), 7.05-7.55 (m, 30H). ^{13}C NMR (50 MHz,



$CDCl_3$): δ 55.1, 56.8, 62.5, 66.3, 66.7, 69.6, 70.0, 70.3, 73.3, 74.3, 75.0, 75.6, 79.9, 82.0, 86.3, 94.9, 98.1, 126.5-128.8, 131.1, 138.1, 138.9, 144.0. Anal. Calcd. for $C_{56}H_{56}O_9$: C, 77.04; H, 6.47. Found: C, 77.45; H, 6.64.

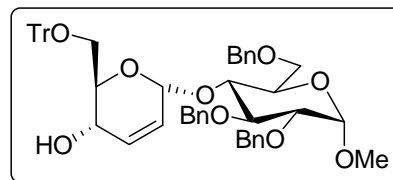
Compound 14: $[\alpha]_D$ ($CHCl_3$, c 1.0) = $+85.0^\circ$. IR (cm^{-1}) = 1720, 1649. 1H NMR (200 MHz, $CDCl_3$): δ 3.25-3.49 (m, 8H), 3.51-3.82 (m, 4H), 3.85-4.13 (m, 3H), 4.75-5.15 (m, 9H), 5.31 (d, 1H, J = 7.33 Hz), 5.97 (s, 1H), 7.14-7.38 (m, 24H), 7.39-7.40 (m, 6H). ^{13}C NMR (50 MHz, $CDCl_3$): δ 29.7, 45.5, 51.5, 54.8, 64.0,



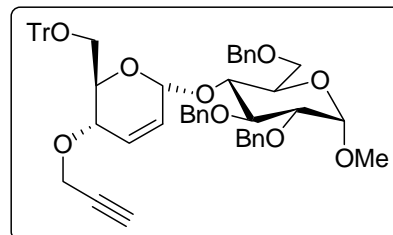
65.2, 65.5, 66.0, 70.2, 71.2, 73.4, 75.0, 75.6, 78.3, 80.1, 82.0, 86.8, 95.5, 97.8, 125.0, 127.0 - 128.8, 138.2, 138.3, 138.9, 143.9, 178.3, 206.3. Anal. Calcd. for $C_{57}H_{56}O_{10}$: C, 75.98; H, 6.26. Found: C, 75.48; H, 5.90.

Compound 16: $[\alpha]_D$ ($CHCl_3$, c 1.98) = $+13.82^\circ$. IR (cm^{-1}) = 3450. 1H NMR (200 MHz, $CDCl_3$): δ 2.45 (d, 1H, J = 3.42 Hz), 3.16 (m, 1H), 3.32 (m, 1H), 3.39 (s, 3H), 3.47-3.68 (m, 5H), 3.79 (m, 1H), 3.94 (t, 1H, J = 9.12 Hz), 4.14-4.37 (m, 3H), 4.58-4.77 (m, 4H), 5.02 (d, 1H, J = 11.14 Hz), 5.39 (s, 1H), 5.47 (m, 1H), 5.86 (d, 1H, J = 10.23 Hz), 7.13-7.40 (m, 30H).

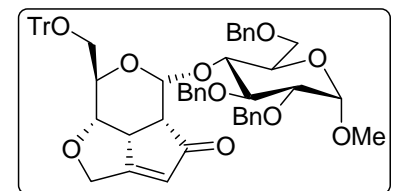
^{13}C NMR (50 MHz, CDCl_3): δ 55.1, 64.6, 69.2, 69.5, 69.8, 72.9, 73.2, 75.6, 76.0, 80.0, 81.9, 87.3, 95.5, 97.7, 125.5, 127.2-128.6, 132.8, 137.9, 138.2, 138.2, 138.5, 143.3. Anal. Calcd. for $\text{C}_{57}\text{H}_{54}\text{O}_{10}$: C, 75.98; H, 6.26. Found: C, 76.08; H, 6.44.



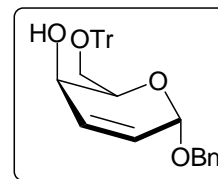
Compound 17: $[\alpha]_{\text{D}}$ (CHCl_3 , c 1.08) = $+35.34^\circ$. IR (cm^{-1}) = 3306. ^1H NMR (200 MHz, CDCl_3): δ 2.29 (t, 1H, J = 2.40 Hz), 3.02 (dd, 1H, J = 2.82, 10.22 Hz), 3.27 (d, 1H, J = 9.84 Hz), 3.39 (s, 3H), 3.44-3.71 (m, 5H), 3.73-4.05 (m, 4H), 4.17 (d, 1H, J = 12.26 Hz), 4.32-4.44 (m, 2H), 4.57-4.7 (m, 4H), 5.04 (d, 1H, J = 11.11 Hz), 5.45-5.6 (m, 2H), 6.04 (d, 1H, J = 9.82 Hz), 7.07-7.40 (m, 24H), 7.45-7.5 (m, 6H). ^{13}C NMR (50 MHz, CDCl_3): δ 55.1, 56.8, 61.9, 69.6, 70.0, 72.8, 73.2, 74.4, 75.6, 80.1, 81.9, 86.3, 96.0, 97.6, 126.1, 126.9-128.8, 131.3, 138.0, 138.1, 138.6, 143.9. Anal. Calcd. for $\text{C}_{56}\text{H}_{56}\text{O}_9$: C, 77.04; H, 6.47. Found: C, 77.29; H, 6.25.



Compound 18: $[\alpha]_{\text{D}}$ (CHCl_3 , c 0.9) = $+113.88^\circ$. IR (cm^{-1}) = 1720. ^1H NMR (200 MHz, CDCl_3): δ 3.08-3.35 (m, 6H), 3.38 (s, 3H), 3.41-3.51 (m, 2H), 3.58-3.72 (m, 2H), 3.98 (dd, 1H, J = 8.44, 9.22 Hz), 4.10 (m, 1H), 4.17 (s, 1H), 4.28 (t, 1H, J = 8.59 Hz), 4.39-4.55 (m, 2H), 4.56 (d, 1H, J = 3.28 Hz), 4.67 (ABq, 2H, J = 12.13 Hz), 5.04 (ABq, 2H, J = 10.85 Hz), 5.81 (s, 1H), 5.85 (s, 1H), 7.07-7.36 (m, 22H), 7.37-7.49 (m, 8H). ^{13}C NMR (50 MHz, CDCl_3): δ 45.5, 51.1, 55.1, 62.9, 64.8, 65.5, 69.3, 69.5, 70.6, 72.9, 73.2, 74.8, 80.3, 81.9, 86.7, 94.6, 97.7, 125.0, 126.9-128.9, 138.0, 138.2, 139.1, 143.9, 177.9, 206.2. Anal. Calcd. for $\text{C}_{57}\text{H}_{56}\text{O}_{10}$: C, 75.98; H, 6.26. Found: C, 75.88; H, 5.99.

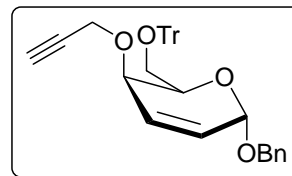


Compound 20: $[\alpha]_{\text{D}}$ (CHCl_3 , c 2.84) = -52.99° . IR (cm^{-1}) = 3455. ^1H NMR (200 MHz, CDCl_3): δ 3.25 (dd, 1H, J = 5.07, 9.95 Hz), 3.51 (dd, 1H, J = 6.92, 9.94 Hz), 3.84 (m, 1H), 4.3 (m, 1H), 4.78 (ABq, 2H, J = 11.56 Hz), 5.14 (d, 1H, J = 2.92 Hz), 5.93 (dd, 1H, J = 3.05, 10.03 Hz), 6.14 (dd, 1H, J = 5.55, 10.12 Hz), 7.22-7.38 (m, 14H), 7.47-7.52 (m, 6H). ^{13}C NMR (50 MHz, CDCl_3): δ 62.2, 63.6, 69.5, 70.1, 86.8, 93.2, 127.0, 127.7-129.6, 137.7, 143.9. Anal. Calcd. for $\text{C}_{32}\text{H}_{30}\text{O}_3$: C, 83.09; H, 6.54. Found: C, 83.15; H, 6.49.



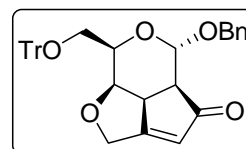
Compound 21: $[\alpha]_{\text{D}}$ (CHCl_3 , c 2.82) = -80.4° . IR (cm^{-1}) = 3301, 1491, 1449. ^1H NMR (200 MHz, CDCl_3): δ 2.28 (t, 1H, J = 2.28 Hz), 3.27 (dd, 1H, J = 5.47, 9.66 Hz), 3.54 (dd, 1H, J = 7.10, 9.78 Hz), 3.85 (dd, 1H, J = 2.54, 5.34 Hz), 4.09 (dd, 2H, J = 1.52, 2.28 Hz), 4.35 (m,

1H), 4.73 (ABq, 2H, $J = 11.63$ Hz), 5.15 (d, 1H, $J = 2.98$ Hz), 6.02 (dd, 1H, $J = 2.79, 9.99$ Hz), 6.21 (ddd, 1H, $J = 0.77, 5.19, 9.90$ Hz), 7.15-7.42 (m, 14H), 7.45-7.55 (m, 6H). ^{13}C NMR (50 MHz, CDCl_3): δ 56.2, 63.1, 67.0, 69.4, 69.8, 74.5, 79.8, 86.6, 92.9, 126.7-128.7, 129.9, 137.6, 143.9. Anal. Calcd. for $\text{C}_{35}\text{H}_{32}\text{O}_4$: C, 81.37; H, 6.24. Found: C, 80.65; H, 6.58.



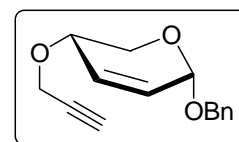
Compound 22: $[\alpha]_{\text{D}} (\text{CHCl}_3, c 0.8) = -130.9^\circ$. IR (cm^{-1}) = 1738, 1713.

^1H NMR (200 MHz, CDCl_3): δ 2.79 (d, 1H, $J = 6.39$ Hz), 3.25 (dd, 1H, $J = 2.29, 10.53$ Hz), 3.33 (m, 1H), 3.43-3.64 (m, 1H), 4.09 (dd, 1H, $J = 3.24, 9.62$ Hz), 4.18 (dd, 1H, $J = 4.10, 7.79$ Hz), 4.31 (ABq, 2H, $J = 14.75$ Hz), 4.74 (ABq, 2H, $J = 11.45$ Hz), 4.03 (d, 1H, $J = 11.85$ Hz), 5.93 (s, 1H), 7.12-7.38 (m, 14H), 7.41-7.48 (m, 6H). ^{13}C NMR (50 MHz, CDCl_3): δ 44.1, 51.7, 63.0, 65.8, 68.0, 69.3, 71.0, 86.9, 97.5, 124.0, 126.9-128.8, 137.0, 144.0, 183.6, 209.4. Anal. Calcd. for $\text{C}_{36}\text{H}_{32}\text{O}_5$: C, 79.39; H, 5.92. Found: C, 79.61; H, 5.58.



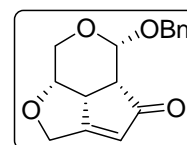
Compound 26: $[\alpha]_{\text{D}} (\text{CHCl}_3, c 1.0) = +69.4^\circ$. IR (cm^{-1}) = 3306. ^1H

NMR (200 MHz, CDCl_3): δ 2.45 (t, 1H, $J = 2.28$ Hz), 3.87 (m, 1H), 3.96 (td, 1H, $J = 1.10, 12.81$ Hz), 4.12 (dd, 1H, $J = 2.69, 4.12$ Hz), 4.26 (d, 2H, $J = 2.41$ Hz), 4.68 (ABq, 2H, $J = 11.77$ Hz), 5.09 (d, 1H, $J = 2.35$ Hz), 6.01 (dd, 1H, $J = 2.72, 10.23$ Hz), 6.14 (td, 1H, $J = 1.12, 4.98$ Hz), 7.25-7.38 (m, 5H). ^{13}C NMR (50 MHz, CDCl_3): δ 55.5, 61.1, 66.6, 69.7, 74.7, 79.7, 92.2, 126.2, 127.7, 128.1, 128.4, 130.1, 137.6. Anal. Calcd. for $\text{C}_{15}\text{H}_{16}\text{O}_3$: C, 73.75; H, 6.60. Found: C, 73.69; H, 6.61.



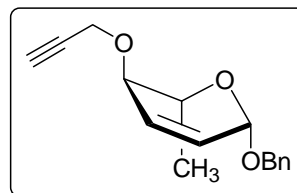
Compound 27: $[\alpha]_{\text{D}} (\text{CHCl}_3, c 1.5) = +52.2^\circ$. IR (cm^{-1}) = 1738, 1713. ^1H

NMR (200 MHz, CDCl_3): δ 2.92 (d, 1H, $J = 6.78$ Hz), 3.51 (m, 2H), 4.03 (dd, 1H, $J = 5.30, 13.09$ Hz), 4.38 (m, 1H), 4.58 (ABq, 2H, $J = 11.50$ Hz), 4.67 (t, 2H, $J = 14.29$ Hz), 4.84 (s, 1H), 6.10 (s, 1H), 7.28-7.40 (m, 5H). ^{13}C NMR (50 MHz, CDCl_3): δ 43.8, 51.5, 62.2, 65.2, 70.0, 70.1, 96.8, 124.6, 128.0, 128.5, 136.8, 183.5, 208.7. Anal. Calcd. for $\text{C}_{16}\text{H}_{16}\text{O}_4$: C, 70.57; H, 5.92. Found: C, 70.74; H, 6.29.



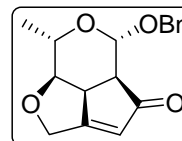
Compound 31: $[\alpha]_{\text{D}} (\text{CHCl}_3, c 1.7) = -96.4^\circ$. IR (cm^{-1}) = 3288. ^1H

NMR (200 MHz, CDCl_3): δ 1.27 (d, 3H, $J = 5.78$ Hz), 2.43 (t, 1H, $J = 2.39$ Hz), 3.85 (m, 2H), 4.23 (d, 2H, $J = 2.32$ Hz), 4.68 (ABq, 2H, $J = 12.00$ Hz), 5.03 (d, 1H, $J = 2.61$ Hz), 5.79 (ddd, 1H, $J = 1.76, 2.65,$

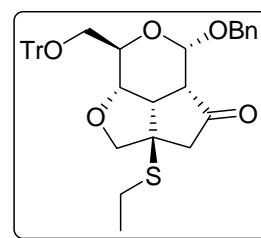


10.25 Hz), 6.07 (td, 1H, $J = 1.19, 10.37$ Hz), 7.23-7.40 (m, 5H). ^{13}C NMR (75 MHz, CDCl_3): δ 18.4, 56.7, 66.1, 70.3, 74.9, 76.5, 80.2, 94.1, 127.4-128.8, 130.65, 138.62. Anal. Calcd. for $\text{C}_{16}\text{H}_{18}\text{O}_3$: C, 74.39; H, 7.02. Found: C, 73.94; H, 6.58.

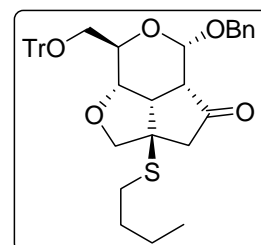
Compound 32: $[\alpha]_{\text{D}}$ (CHCl_3 , c 1.2) = -268.2° . IR (cm^{-1}) = 1714, 1722. ^1H NMR (200 MHz, CDCl_3): δ 1.24 (d, 3H, $J = 6.13$ Hz), 3.27-3.61 (m, 3H), 3.93 (t, 1H, $J = 9.10$ Hz), 4.50 (ABq, 2H, $J = 12.41$ Hz), 4.62 (d, 2H, $J = 1.57$ Hz), 5.16 (d, 1H, $J = 7.54$ Hz), 6.05 (d, 1H, $J = 2.04$ Hz), 7.22-7.40 (m, 5H). ^{13}C NMR (50 MHz, CDCl_3): δ 18.1, 46.0, 51.9, 61.9, 85.6, 69.7, 78.0, 95.1, 125.3, 125.8-128.2, 137.6, 179.1, 207.3. Anal. Calcd. for $\text{C}_{17}\text{H}_{18}\text{O}_4$: C, 71.31; H, 6.34. Found: C, 71.65; H, 6.71.



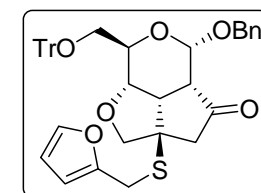
Compound 32a: $[\alpha]_{\text{D}}$ (CH_2Cl_2 , c 0.3) = $+59.2^\circ$. IR (cm^{-1}): 1747. ^1H NMR (200 MHz, CDCl_3): δ 1.22 (t, 3 H, $J = 7.31$ Hz), 2.48-2.64 (m, 3H), 2.70-2.83 (m, 2H), 3.08 (m, 1H), 3.32 (m, 2H), 3.70 (m, 1H), 3.86-4.02 (m, 3H), 4.70 (ABq, 2H, $J = 12.56$ Hz), 5.30 (d, 1H, $J = 7.44$ Hz), 7.20 (m, 14H), 7.42-7.53 (m, 6H). ^{13}C NMR (75 MHz, CDCl_3): δ 14.6, 24.1, 29.7, 46.2, 48.6, 50.1, 52.2, 65.5, 69.0, 69.3, 73.2, 86.8, 93.9, 96.6, 127.0-128.8, 137.2, 144.0, 211.8. Anal. Calcd for $\text{C}_{38}\text{H}_{38}\text{O}_5\text{S}$: C, 75.22; H, 6.31; S, 5.28. Found: C, 75.59; H, 6.25; S, 5.34.



Compound 32b: $[\alpha]_{\text{D}}$ (CHCl_3 , c 1.16) = $+67.51^\circ$. IR (cm^{-1}): 1747. ^1H NMR (200 MHz, CDCl_3): δ 0.80 (t, 3H, $J = 7.04$ Hz), 1.25-1.53 (m, 4H), 2.45-2.90 (m, 5H), 3.02-3.13 (m, 1H), 3.32 (m, 2H), 3.73-3.99 (m, 4 H), 4.72 (ABq, 2H, $J = 11.99$ Hz), 5.30 (d, 1H, $J = 7.50$ Hz), 7.15-7.36 (m, 14H), 7.41 - 7.52 (m, 6H). ^{13}C NMR (50 MHz, CDCl_3): δ 13.6, 22.1, 29.7, 31.5, 46.0, 48.5, 50.0, 55.0, 65.4, 68.8, 69.1, 73.0, 77.3, 86.7, 93.8, 127.0-128.9, 137.1, 143.9, 212.0. Anal. Calcd for $\text{C}_{40}\text{H}_{42}\text{O}_5\text{S}$: C, 75.68; H, 6.67; S, 5.05. Found: C, 75.01; H, 6.34; S, 5.23.



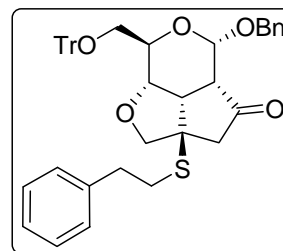
Compound 32c: $[\alpha]_{\text{D}}$ (CHCl_3 , c 0.9) = $+53.85^\circ$. IR (cm^{-1}): 1748. ^1H NMR (200 MHz, CDCl_3): δ 1.16 (dd, 1 H, $J = 1.16, 17.82$ Hz), 2.65-2.82 (m, 2H), 3.05 (ddd, 1H, $J = 1.38, 7.81, 10.82$ Hz), 3.25-3.39 (m, 2H), 3.67 (ABq, 2H, $J = 9.58$ Hz), 3.77 (m, 2H), 3.90-3.98 (m, 2H), 4.70 (ABq, 2H, $J = 12.22$ Hz), 5.28 (d, 1H, $J = 7.54$ Hz), 6.14 (d, 1H, $J = 3.12$ Hz), 6.28 (dd, 1H, $J = 2.01, 3.17$ Hz), 7.15-7.39 (m, 15H), 7.40-7.52 (m, 6H). ^{13}C NMR (50 MHz, CDCl_3): δ 27.1, 46.0, 48.4, 49.4, 55.3, 65.4, 68.8, 69.2, 72.7, 77.1, 86.7, 91.0, 93.8, 107.9, 110.8, 127.0-



128.7, 137.1, 142.1 143.9, 150.7, 211.9. Anal. Calcd. for C₄₁H₃₈O₆S: C, 74.75; H, 5.81; S, 4.87. Found: C, 74.89; H, 5.86; S, 4.78.

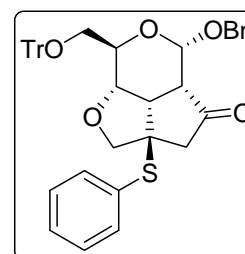
Compound 32d: [α]_D (CHCl₃, *c* 1.16) = +76.38°. IR (cm⁻¹): 1747.

¹H NMR (200 MHz, CDCl₃): δ 2.56 (dd, 1H, *J* = 1.01, 17.50 Hz), 2.68-2.91 (m, 6H), 3.06 (ddd, 1H, *J* = 1.50, 7.77, 10.78 Hz), 3.22-3.37 (m, 2H), 3.81 (ABq, 2H, *J* = 9.40 Hz), 3.94 (m, 2H), 4.71 (ABq, 2H, *J* = 12.27 Hz), 5.28 (d, 1H, *J* = 7.56 Hz), 7.10-7.39 (m,



19H), 7.41-7.50 (m, 6H). ¹³C NMR (50 MHz, CDCl₃): δ 31.7, 35.9, 46.1, 48.6, 50.1, 55.4, 65.5, 68.8, 69.3, 73.2, 77.4, 86.8, 93.8, 126.6-128.9, 137.2, 139.9, 143.9, 212.0. Anal. Calcd. for C₄₄H₄₂O₅S: C, 77.39; H, 6.20; S, 4.70. Found: C, 77.12; H, 6.55; S, 4.94.

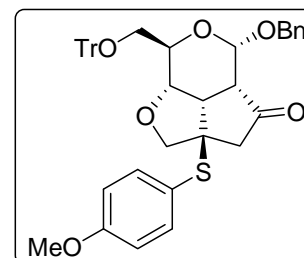
Compound 32e: [α]_D(CHCl₃, *c* 0.44) = +84.14°. IR (cm⁻¹): 1747. ¹H NMR (200 MHz, CDCl₃): δ 2.62 (m, 2H), 2.85 (m, 1H), 3.06 (m, 1H), 3.28 (m, 2H), 3.83-3.95 (m, 4H), 4.70 (ABq, 2H, *J* = 12.36 Hz), 5.30 (d, 1H, *J* = 7.75 Hz), 7.20-7.36 (m, 18H), 7.41-7.49 (m, 7H). ¹³C NMR (75 MHz, CDCl₃): δ 29.7, 45.4, 48.5, 50.0, 58.1, 65.5, 69.3, 73.4,



77.0, 86.8, 94.0, 127.0-129.3, 131.5, 135.9, 137.2, 144.0, 211.4. Anal. Calcd for C₄₂H₃₈O₅S: C, 77.04; H, 5.85; S, 4.90. Found: C, 77.11; H, 5.99; S, 5.01.

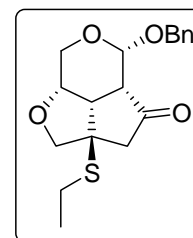
Compound 32f: [α]_D (CHCl₃, *c* 0.98) = +84.92°. IR (cm⁻¹): 1747.

¹H NMR (200 MHz, CDCl₃): δ 2.54-2.74 (m, 2H), 2.83 (dd, 1H, *J* = 8.72, 10.89 Hz), 3.05 (ddd, 1H, *J* = 1.39, 7.58, 10.72 Hz), 3.21 (dd, 1H, *J* = 5.93, 9.92 Hz), 3.35 (dd, 1H, *J* = 2.57, 9.98 Hz), 3.73-3.95 (m, 4H), 3.76 (s, 3H), 4.69 (ABq, 2H, *J* = 12.36 Hz), 5.29 (d, 1H, *J* = 7.49 Hz), 6.75-6.81 (m, 2H), 7.20-7.48 (m, 22H). ¹³C



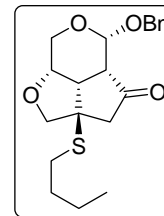
NMR (50 MHz, CDCl₃): δ 45.0, 48.6, 50.0, 55.4, 58.2, 65.5, 69.3, 73.6, 76.8, 86.7, 94.0, 114.8, 116.1, 121.7, 127.1-128.6, 137.2, 138.3, 144.0, 160.9, 212.1. Anal. Calcd. for C₄₃H₄₀O₆S: C, 75.41; H, 5.89; S, 4.68. Found: C, 75.21; H, 5.65; S, 4.49.

Compound 33a: [α]_D (CH₂Cl₂, *c* 3.5) = -95.9°. IR (cm⁻¹): 1749. ¹H NMR (200 MHz, CDCl₃): δ 7.48 (t, 3H, *J* = 7.48 Hz), 2.55-2.78 (m, 4H), 2.85 (dd, 1H, *J* = 1.77, 18.77 Hz), 3.00 (dd, 1H, *J* = 7.27, 10.70 Hz), 3.70-4.13 (m, 5H), 4.60 (ABq, 2H, *J* = 11.60 Hz), 5.26 (s, 1H), 7.30 (m, 5H). ¹³C NMR (50 MHz, CDCl₃): δ 14.7, 23.8, 44.1, 48.4, 51.2, 53.7, 58.2, 69.4, 73.3,

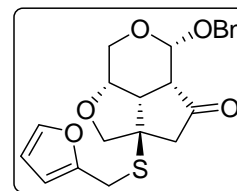


80.3, 94.8, 128.0, 128.5, 137.1, 211.3. Anal. Calcd. for C₁₈H₂₂O₄S: C, 64.65; H, 6.63; S, 9.59. Found: C, 64.72; H, 6.58; S, 9.37.

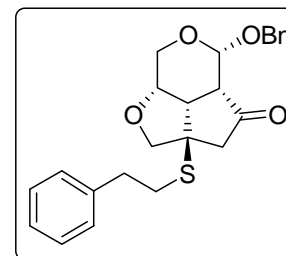
Compound 33b: $[\alpha]_D$ (CHCl₃, *c* 1.5) = -97.7°. IR (cm⁻¹): 1749. ¹H NMR (200 MHz, CDCl₃): δ 0.92 (t, 3H, *J* = 6.80 Hz), 1.32-1.65 (m, 4H), 2.54-2.76 (m, 4H), 2.84 (dd, 1H, *J* = 1.74, 18.68 Hz), 3.01 (dd, 1H, *J* = 7.29, 10.77 Hz), 3.70-4.13 (m, 5H), 5.01 (ABq, 2H, *J* = 11.50 Hz), 5.26 (s, 1H), 7.35 (m, 5H). ¹³C NMR (50 MHz, CDCl₃): δ 13.6, 22.1, 29.5, 31.7, 44.1, 48.5, 51.3, 53.7, 58.2, 69.5, 73.4, 80.3, 94.9, 128.0, 128.5, 137.1, 211.4. Anal. Calcd. for C₂₀H₂₆O₄S: C, 66.27; H, 7.23; S, 8.85. Found: C, 66.03; H, 7.31; S, 8.98



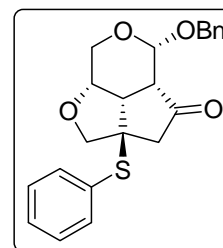
Compound 33c: $[\alpha]_D$ (CHCl₃, *c* 0.82) = +24.14°. IR (cm⁻¹): 1746. ¹H NMR (200 MHz, CDCl₃): δ 2.61-2.85 (m, 4H), 3.56 (d, 1H, *J* = 9.17 Hz), 3.70-3.98 (m, 5H), 4.14 (dd, 1H, *J* = 1.87, 12.86 Hz), 4.33 (d, 1H, *J* = 4.33 Hz), 4.71 (ABq, 2H, *J* = 12.45 Hz), 6.31 (dd, 1H, *J* = 1.98, 3.21 Hz); 6.13 (ddd, 1H, *J* = 0.54, 1.08, 3.16 Hz), 7.25-7.38 (m, 6H). ¹³C NMR (50 MHz, CDCl₃): δ 26.8, 47.0, 47.3, 50.0, 54.7, 65.3, 69.8, 72.2, 77.8, 97.1, 107.9, 110.8, 127.7-128.4, 136.9, 142.2, 150.8, 210.4. Anal. Calcd. for C₂₁H₂₂O₅S: C, 65.27; H, 5.74; S, 8.30. Found: C, 65.43; H, 5.97; S, 8.99.



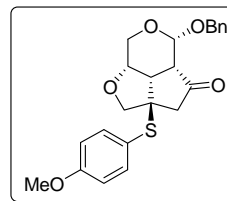
Compound 33d: $[\alpha]_D$ (CHCl₃, *c* 0.9) = +68.23°. IR (cm⁻¹): 1747. ¹H NMR (200 MHz, CDCl₃): δ 2.64 (ddd, 1H, *J* = 1.01, 2.23, 17.4 Hz), 2.73-2.91 (m, 7H), 3.69 (d, 1H, *J* = 9.09 Hz), 3.79 (dd, 1H, *J* = 4.83, 12.88 Hz), 3.92-4.03 (m, 2H), 4.14 (dd, 1H, *J* = 2.03, 13.03 Hz), 4.51 (d, 1H, *J* = 4.51 Hz), 4.71 (ABq, 2H, *J* = 12.31 Hz), 7.13-7.38 (m, 10H). ¹³C NMR (50 MHz, CDCl₃): δ 31.3, 35.9, 46.9, 47.4, 50.6, 54.8, 65.4, 69.7, 72.5, 77.8, 97.1, 126.6-128.4, 136.9, 139.7, 210.5. Anal. Calcd for C₂₄H₂₆O₄S: C, 70.22; H, 6.38; S, 7.81. Found: C, 69.87; H, 6.15; S, 7.46.



Compound 33e: $[\alpha]_D$ (CH₂Cl₂, *c* 1.2) = -97.29°. IR (cm⁻¹): 1749. ¹H NMR (200 MHz, CDCl₃): δ 2.52 (d, 1H, *J* = 18.93 Hz), 2.63 (ddd, 1H, *J* = 0.75, 1.75, 10.86 Hz), 2.79 (dd, 1H, *J* = 1.88, 18.93 Hz), 3.10 (dd, 1H, *J* = 10.73, 7.33 Hz), 3.68-4.09 (m, 5H), 4.59 (ABq, 2H, *J* = 11.62 Hz), 5.22 (s, 1H), 7.30-7.42 (m, 8H), 7.48-7.56 (m, 2H). ¹³C NMR (50 MHz, CDCl₃): δ 43.9, 48.5, 50.7, 56.8, 58.3, 69.5, 73.3, 79.8, 94.9, 128.1-129.5, 131.0, 136.0, 137.1, 211.4. Anal. Calcd. for C₂₂H₂₂O₄S: C, 69.09; H, 5.80; S, 8.38. Found: C, 68.58; H, 6.01; S, 8.48.

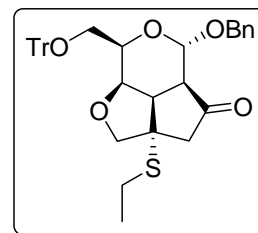


Compound 33f: $[\alpha]_D$ (CHCl₃, *c* 0.8) = +64.4°. IR (cm⁻¹): 1749. ¹H NMR (200 MHz, CDCl₃): δ 2.56-2.75 (m, 3H), 2.88 (dd, 1H, *J* = 8.65, 10.14 Hz), 3.72-3.95 (m, 5H), 3.81 (s, 3H), 4.66 (d, 1H, *J* = 5.06 Hz), 4.71 (ABq, 2H, *J* = 12.18 Hz), 6.84, 6.88 (2s, 2H), 7.30 (m, 5H), 7.38, 7.42



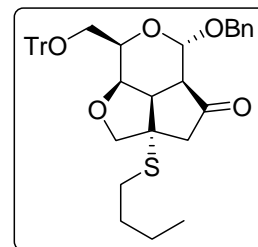
(2s, 2H). ¹³C NMR (50 MHz, CDCl₃): δ 46.1, 47.5, 50.3, 55.3, 57.7, 65.7, 69.7, 72.7, 77.3, 97.3, 114.8, 121.2, 127.7-128.4, 137.0, 138.1, 160.8, 210.5. Anal. Calcd. for C₂₃H₂₄O₅S: C, 66.97; H, 5.86; S, 7.77. Found: C, 66.73; H, 5.99; S, 7.94.

Compound 34a: $[\alpha]_D$ (CHCl₃, *c* 0.48) = +10.04°. IR (cm⁻¹): 1747. ¹H NMR (200 MHz, CDCl₃): δ 1.24 (t, 3H, *J* = 7.42 Hz), 2.50-2.89 (m, 5H), 3.04 (dd, 1H, *J* = 6.95, 10.97 Hz), 3.27 (dd, 1H, *J* = 4.03, 9.99 Hz), 3.45 (dd, 1H, *J* = 7.57, 9.97 Hz), 3.70 (dd, 1H, *J* = 2.14, 6.84 Hz), 3.77 (ABq, 2H, *J* = 9.10 Hz), 4.24 (m, 1H), 4.74 (ABq, 2H, *J* =



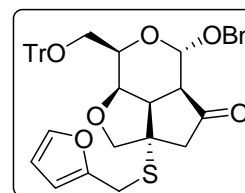
11.47 Hz), 5.36 (s, 1H), 7.10-7.52 (m, 20H). ¹³C NMR (75 MHz, CDCl₃): δ 14.7, 23.9, 45.9, 48.3, 51.3, 53.2, 64.0, 66.3, 69.0, 74.2, 80.9, 86.7, 95.5, 127.0-128.7, 137.2, 144.0, 211.1. Anal. Calcd. for C₃₈H₃₈O₅S: C, 75.22; H, 6.31; S, 5.28. Found: C, 75.45; H, 6.13; S, 5.34.

Compound 34b: $[\alpha]_D$ (CHCl₃, *c* 1.4) = +10.96°. IR (cm⁻¹): 1743. ¹H NMR (200 MHz, CDCl₃): δ 0.91 (m, 3H), 1.32-1.63 (m, 4H), 2.50-2.85 (m, 5H), 3.03 (dd, 1H, *J* = 6.94, 11.02 Hz), 3.26 (dd, 1H, *J* = 3.94, 9.87 Hz), 3.50 (dd, 1H, *J* = 7.58, 9.85 Hz), 3.70 (dd, 1H, *J* = 2.02, 6.96 Hz), 3.76 (ABq, 2H, *J* = 9.04 Hz), 4.25 (m, 1H), 4.74



(ABq, 2H, *J* = 11.49 Hz), 5.36 (s, 1H), 7.11-7.52 (m, 20H). ¹³C NMR (50 MHz, CDCl₃): δ 13.6, 22.1, 29.5, 31.7, 45.9, 48.3, 51.4, 53.2, 64.1, 66.3, 69.0, 74.2, 80.8, 86.8, 95.5, 127.0-128.8, 137.3, 144.0, 211.1. Anal. Calcd. for C₄₀H₄₂O₅S: C, 75.68; H, 6.67; S, 5.05. Found: C, 75.34; H, 6.76; S, 5.33.

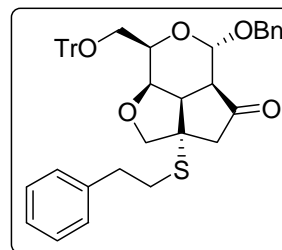
Compound 34c: $[\alpha]_D$ (CHCl₃, *c* 0.82) = +24.14°. IR (cm⁻¹): 1747. ¹H NMR (200 MHz, CDCl₃): δ 2.49 (d, 1H, *J* = 18.69 Hz), 2.67 (dd, 1H, *J* = 1.10, 11.14 Hz), 2.79 (dd, 1H, *J* = 1.52, 18.66 Hz), 2.97 (dd, 1H, *J* = 6.85, 10.98 Hz), 3.24 (dd, 1H, *J* = 4.16, 9.98 Hz), 3.48 (dd, 1H, *J* =



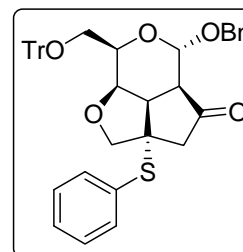
7.69, 9.09 Hz), 3.61 (ABq, 2H, *J* = 9.26 Hz), 3.66 (dd, 1H, *J* = 2.15, 6.86 Hz), 3.80 (m, 2H), 4.21 (m, 1H), 4.67 (ABq, 2H, *J* = 11.71 Hz), 5.35 (s, 1H), 6.16 (dd, 1H, *J* = 0.63, 3.27 Hz), 6.31 (dd, 1H, *J* = 1.89, 3.16 Hz), 7.15-7.48 (m, 21H). ¹³C NMR (50 MHz, CDCl₃): δ 27.1, 46.0, 48.1, 50.9, 53.2, 64.1, 66.3, 69.0, 73.9, 81.1, 86.8, 95.5, 108.0, 111.0, 127.0-128.8, 137.3,

142.3, 144.1, 150.9, 210.9. Anal. Calcd. for C₄₁H₃₈O₆S: C, 74.75; H, 5.81; S, 4.87. Found: C, 74.39; H, 5.45; S, 4.60.

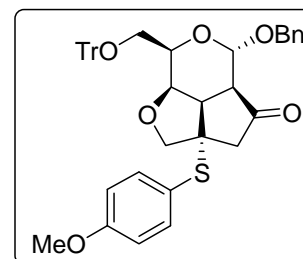
Compound 34d: [α]_D (CHCl₃, *c* 0.9) = +68.23°. IR (cm⁻¹): 1747. ¹H NMR (200 MHz, CDCl₃): δ 2.46-2.74 (m, 3H), 2.82 (m, 4H), 3.00 (dd, 1H, *J* = 6.96, 11.01 Hz), 3.25 (dd, 1H, *J* = 4.06, 9.97 Hz), 3.53 (dd, 1H, *J* = 7.58, 9.96 Hz), 3.67 (dd, 1H, *J* = 2.31, 6.74 Hz), 3.71 (ABq, 2H, *J* = 9.12 Hz), 4.22 (m, 1H), 4.73 (ABq, 2H, *J* = 11.56 Hz), 5.35 (s, 1H), 7.11 - 7.45 (m, 25H). ¹³C NMR (50 MHz, CDCl₃): δ 31.4, 35.8, 45.7, 48.2, 51.3, 53.3, 64.0, 66.2, 69.0, 74.2, 80.6, 86.7, 95.4, 126.7-128.7, 137.2, 139.7, 144.0, 210.8. Anal. Calcd. for C₄₄H₄₂O₅S: C, 77.39; H, 6.20; S, 4.70. Found: C, 77.89; H, 5.99; S, 4.49.



Compound 34e: [α]_D (CHCl₃, *c* 1.0) = +14.37°. IR (cm⁻¹): 1747. ¹H NMR (200 MHz, CDCl₃): δ 2.43 (d, 1H, *J* = 18.79 Hz), 2.61 (dd, 1H, *J* = 1.52, 10.89 Hz), 2.75 (dd, 1H, *J* = 1.51, 18.71 Hz), 3.09-3.27 (m, 2H), 3.48 (dd, 1H, *J* = 7.83, 9.96 Hz), 3.61 (dd, 1H, *J* = 2.03, 9.23 Hz), 3.76 (ABq, 2H, *J* = 9.23 Hz), 4.22 (m, 1H), 4.74 (ABq, 2H, *J* = 11.51 Hz), 5.33 (s, 1H), 7.15-7.53 (m, 25H). ¹³C NMR (75 MHz, CDCl₃): δ 45.6, 48.2, 50.7, 56.3, 64.0, 66.2, 69.0, 74.1, 80.1, 86.7, 95.5, 126.9-129.4, 131.0, 136.0, 137.2, 144.0, 211.0. Anal. Calcd. for C₄₂H₃₈O₅S: C, 77.04; H, 5.85; S, 4.90. Found: C, 77.57; H, 5.95; S, 4.99.

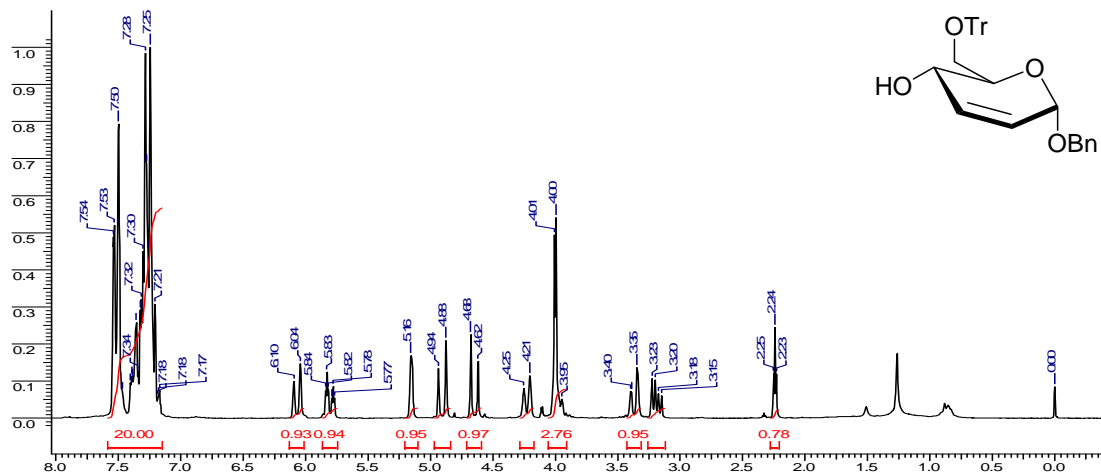


Compound 34f: [α]_D (CHCl₃, *c* 0.78) = +19.46°. IR (cm⁻¹): 1747. ¹H NMR (200 MHz, CDCl₃): δ 2.42 (d, 1H, *J* = 18.83 Hz), 2.58 (dd, 1H, *J* = 1.56, 11.00 Hz), 2.73 (dd, 1H, *J* = 1.49, 18.69 Hz), 3.06-3.25 (m, 2H), 3.48 (dd, 1H, *J* = 7.95, 9.97 Hz), 3.57 (dd, 1H, *J* = 1.90, 6.95 Hz), 3.71 (ABq, 2H, *J* = 9.26 Hz), 3.80 (s, 3H), 4.21 (m, 1H), 4.69 (ABq, 2H, *J* = 11.53 Hz), 5.32 (s, 1H), 6.84, 6.88 (2s, 2H), 7.17-7.51 (m, 22H). ¹³C NMR (50 MHz, CDCl₃): δ 45.5, 48.3, 50.6, 55.4, 56.5, 64.1, 66.3, 69.1, 74.3, 79.8, 86.8, 95.6, 114.9, 121.5, 127.0-128.8, 137.3, 138.2, 144.1, 160.9, 211.3. Anal. Calcd. for C₄₃H₄₀O₆S: C, 75.41; H, 5.89; S, 4.68. Found: C, 75.88; H, 5.98; S, 4.94.

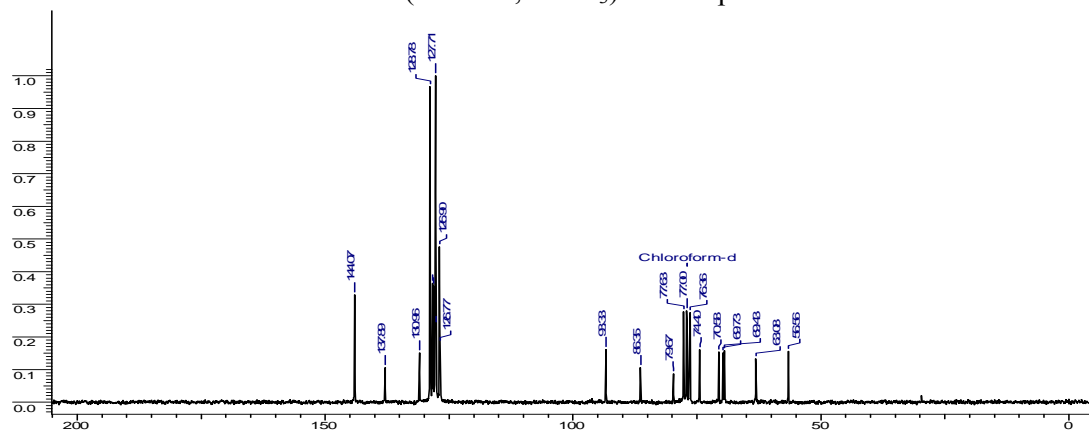


Chapter 1: Spectral Charts

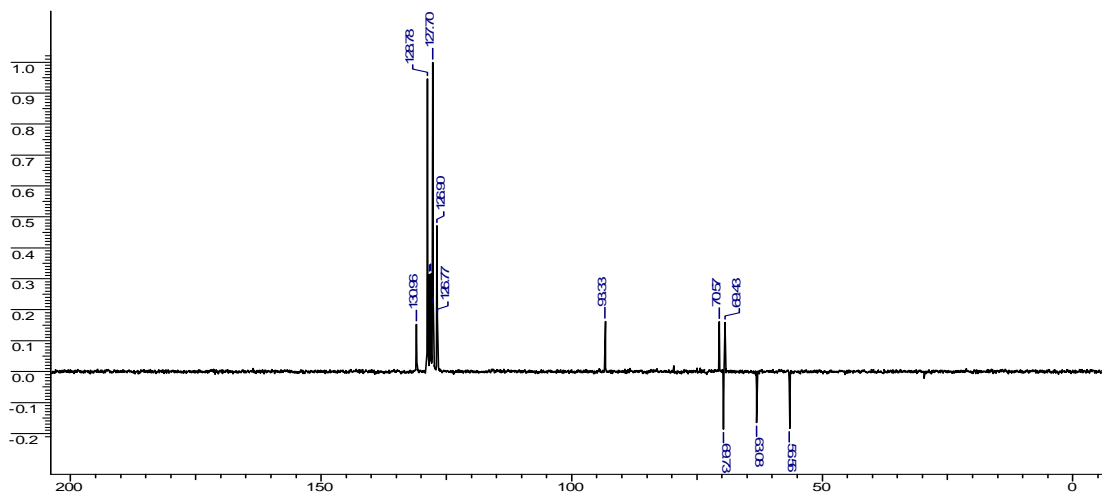
^1H NMR (200 MHz, CDCl_3) of Compound 5



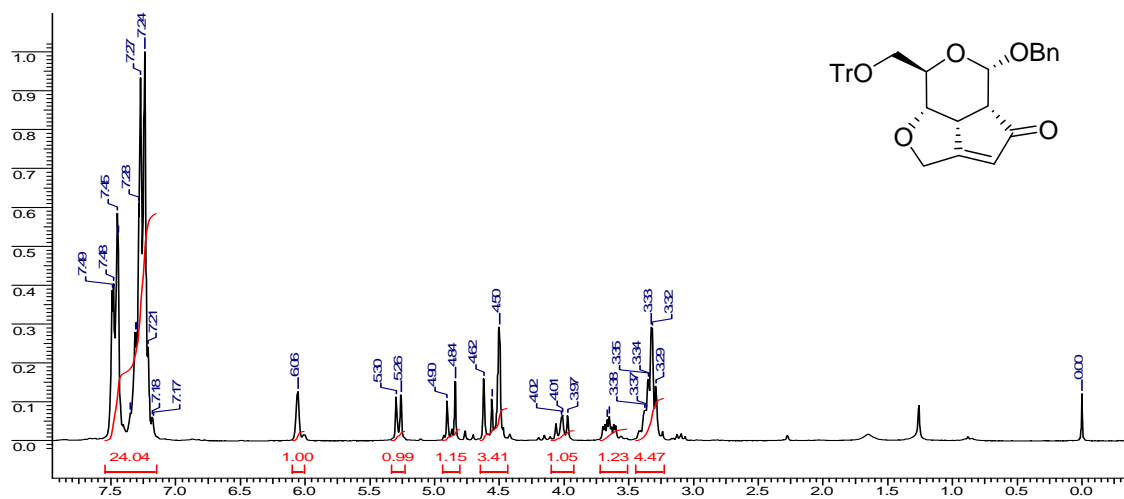
^{13}C NMR (50 MHz, CDCl_3) of Compound 5



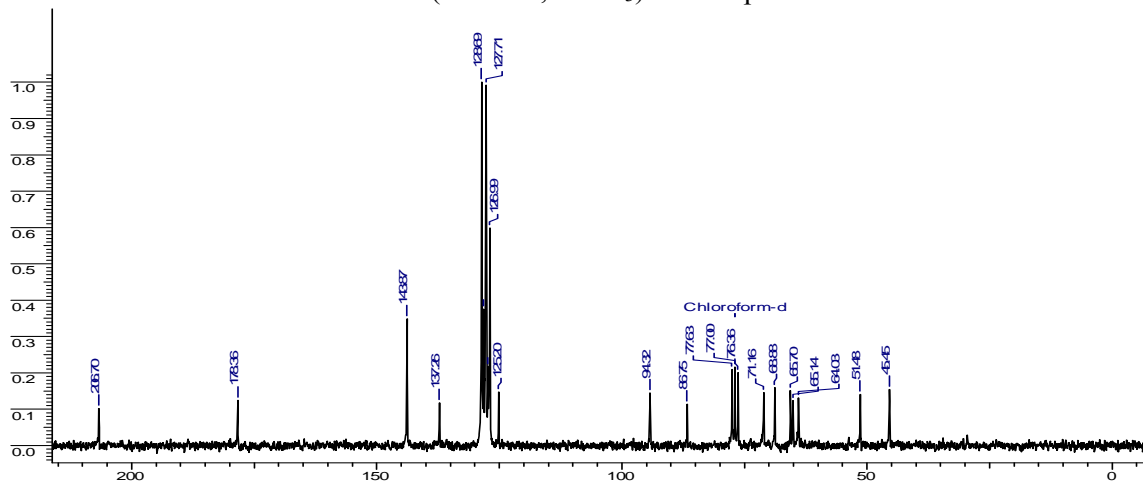
DEPT NMR (50 MHz, CDCl_3) of compound 5



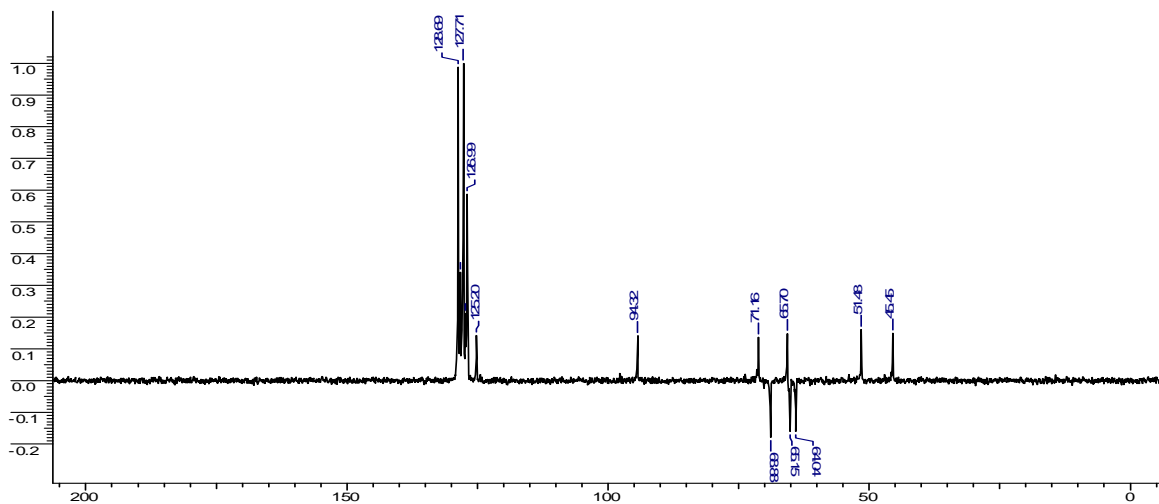
¹H NMR (200 MHz, CDCl₃) of Compound 6



¹³C NMR (50 MHz, CDCl₃) of Compound 6



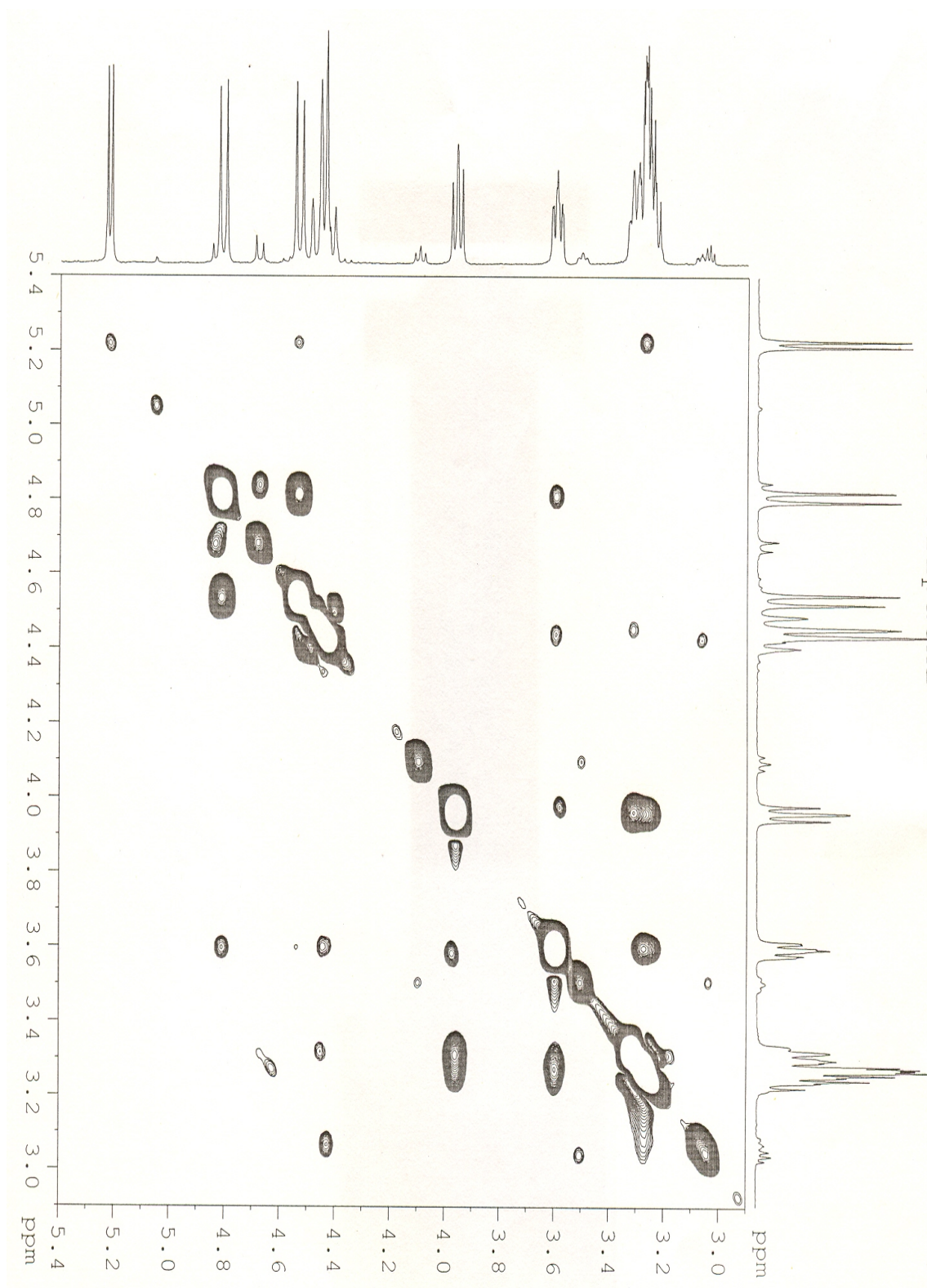
DEPT NMR (50 MHz, CDCl₃) of Compound 6



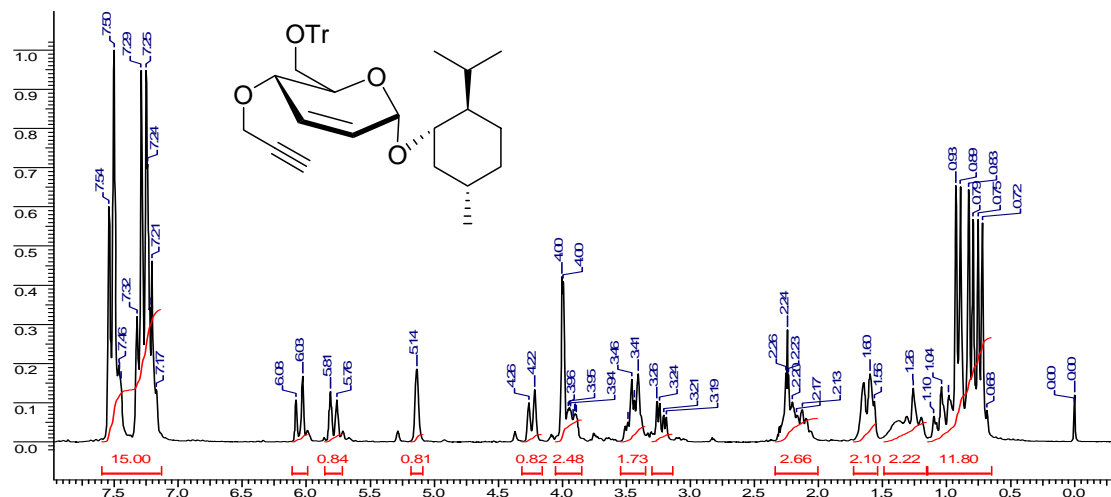
COSY (500 MHz, CDCl₃) Spectrum of Compound 6



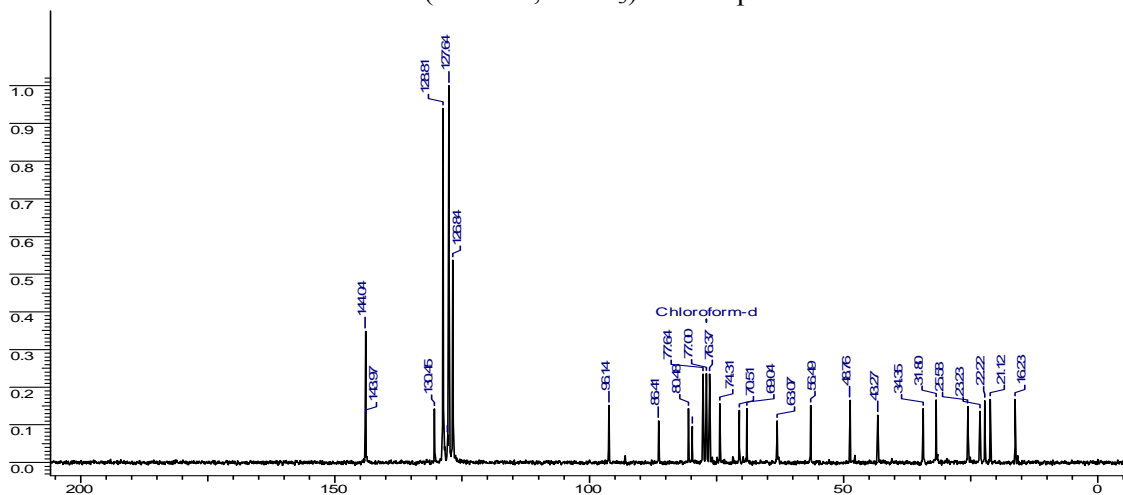
NOESY (500 MHz, CDCl₃) Spectrum of Compound 6



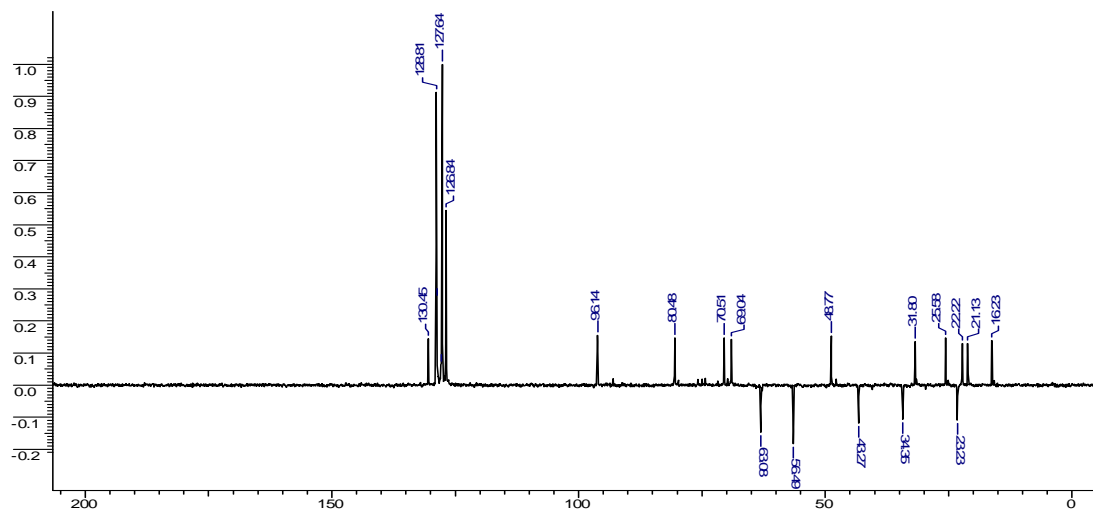
¹H NMR (200 MHz, CDCl₃) of compound **9**



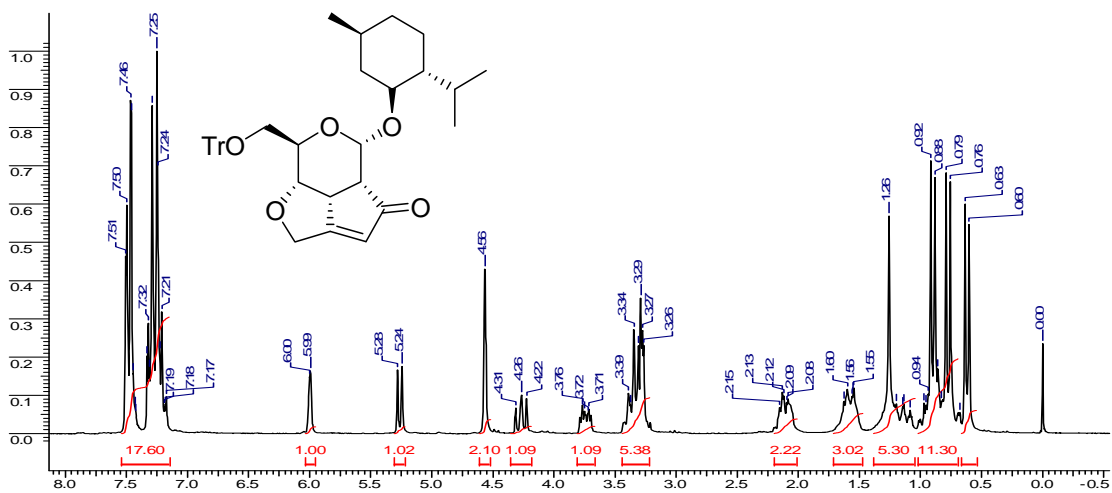
¹³C NMR (50 MHz, CDCl₃) of Compound **9**



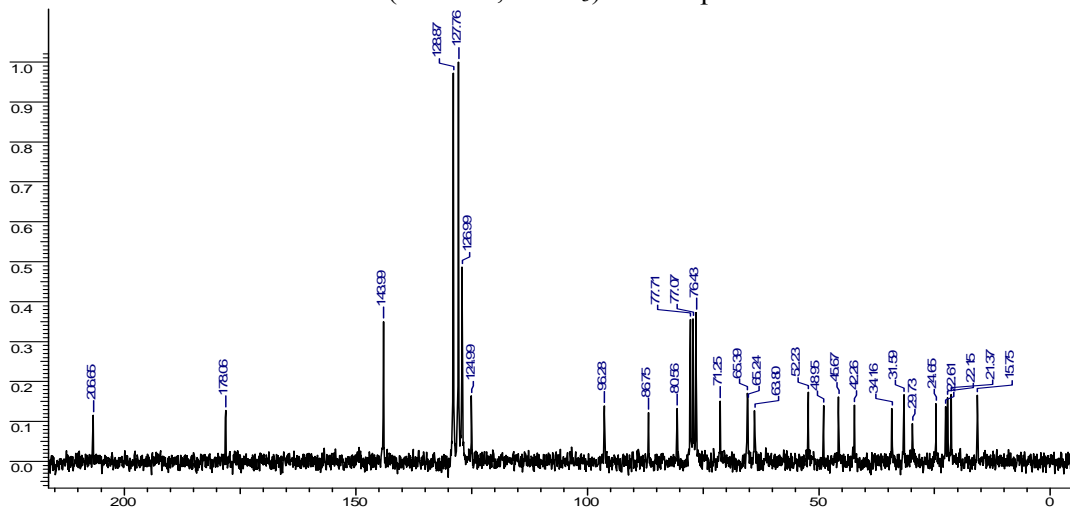
DEPT NMR (50 MHz, CDCl₃) of Compound **9**



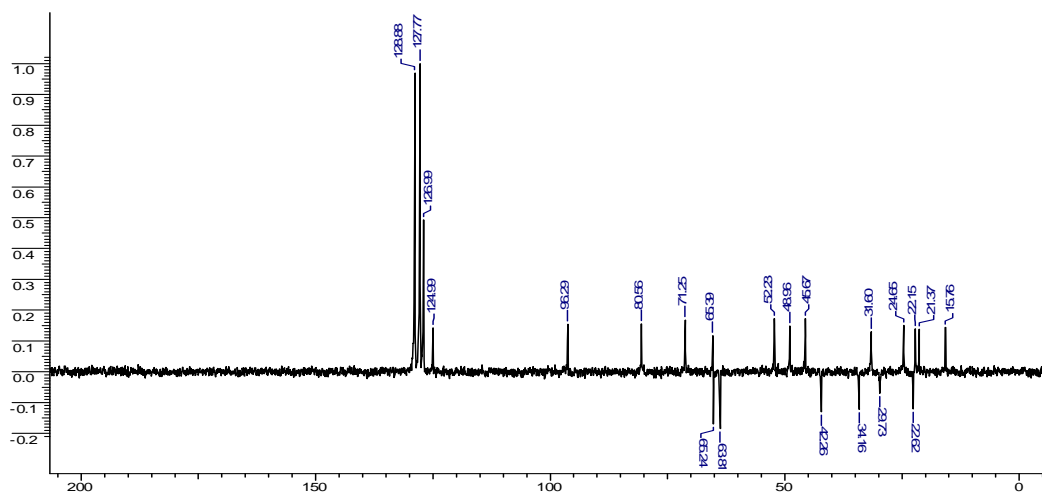
¹H NMR (200 MHz, CDCl₃) of Compound 10



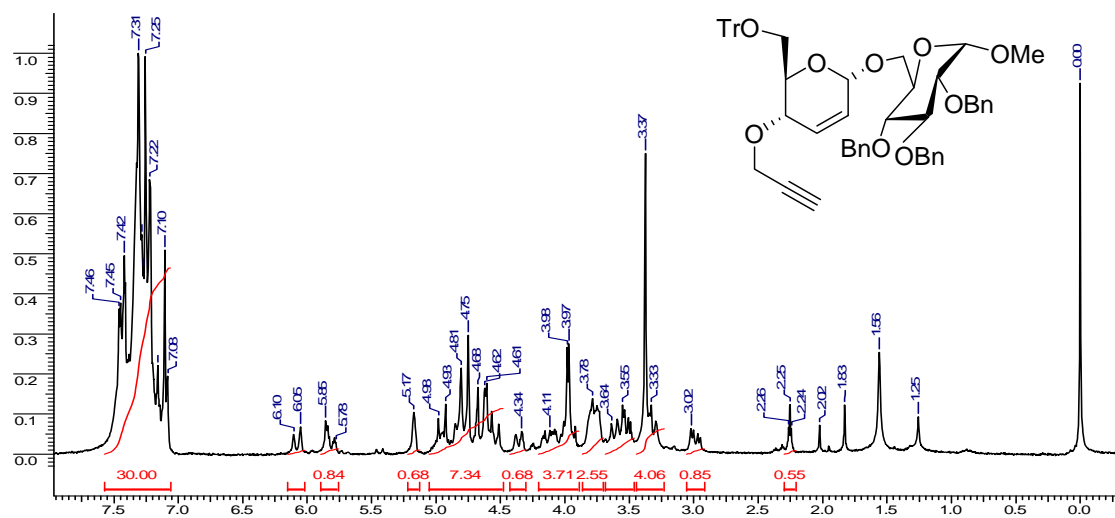
¹³C NMR (50 MHz, CDCl₃) of Compound 10



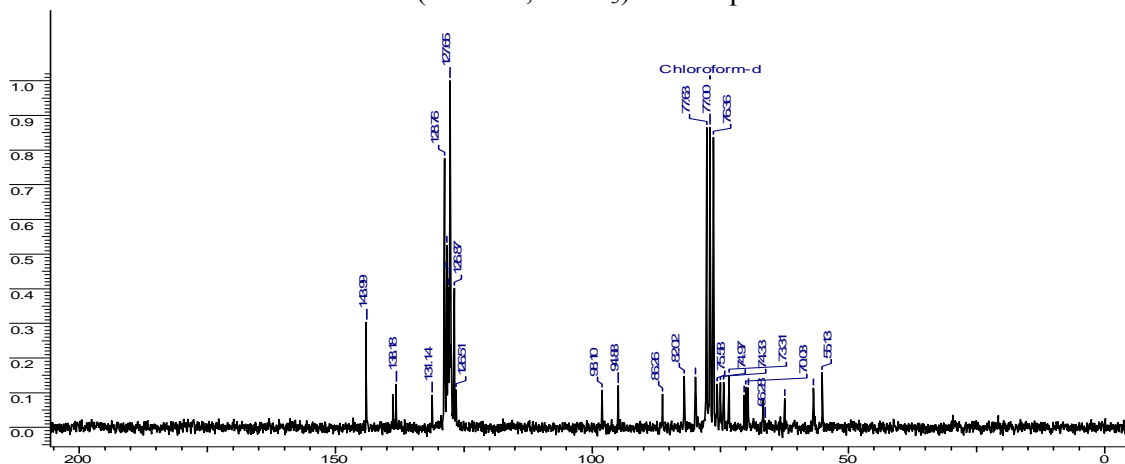
DEPT NMR (50 MHz, CDCl₃) of Compound 10



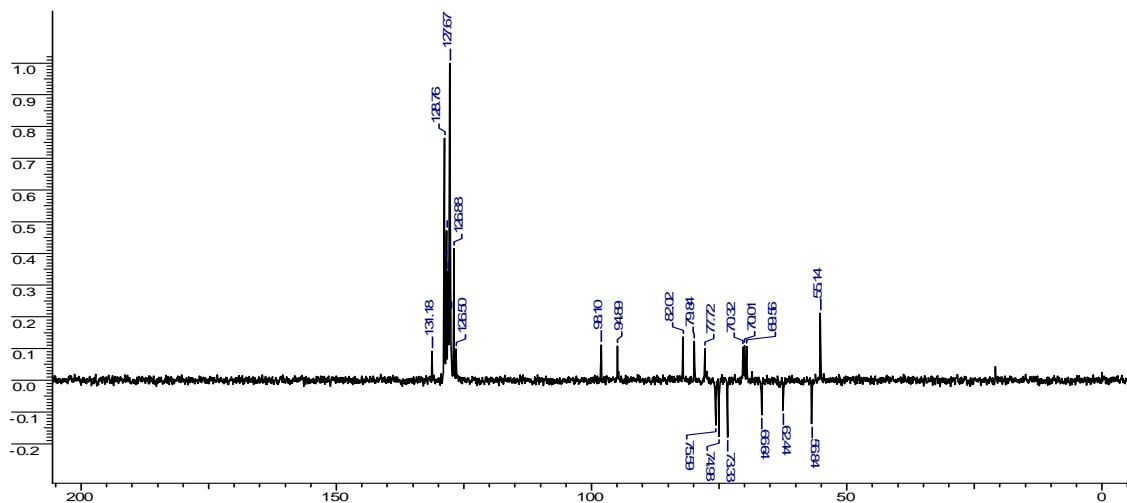
¹H NMR (200 MHz, CDCl₃) of Compound 13



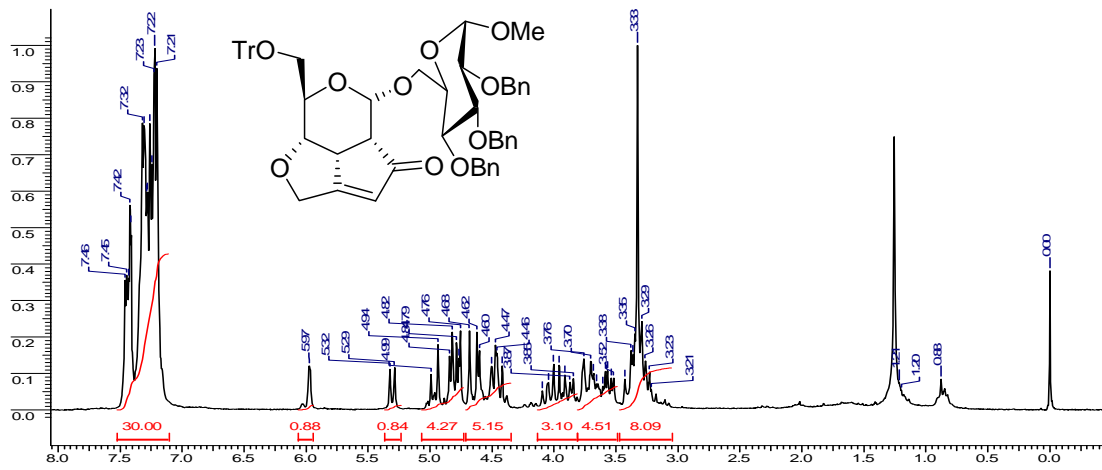
¹³C NMR (50 MHz, CDCl₃) of Compound 13



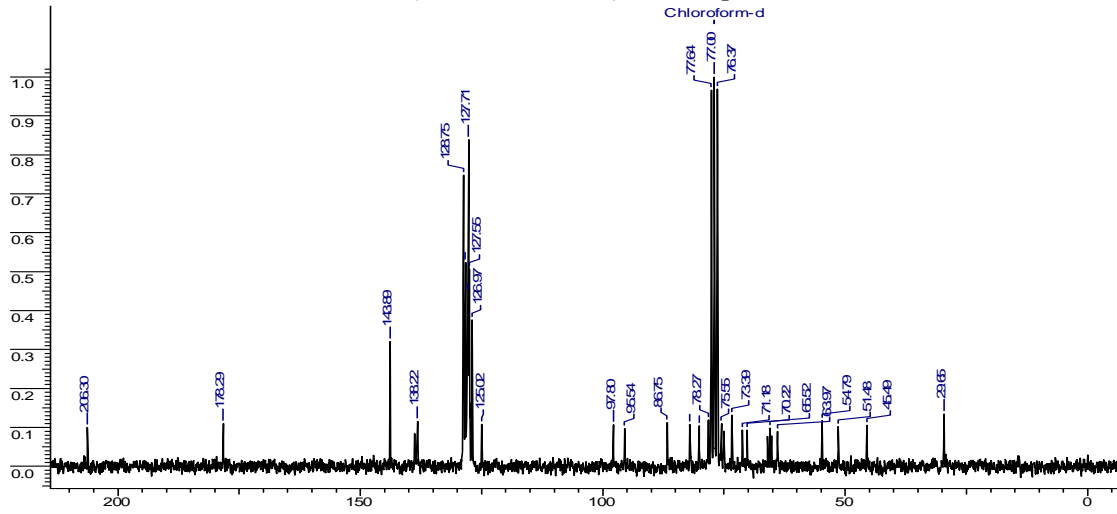
DEPT NMR (50 MHz, CDCl₃) of Compound 13



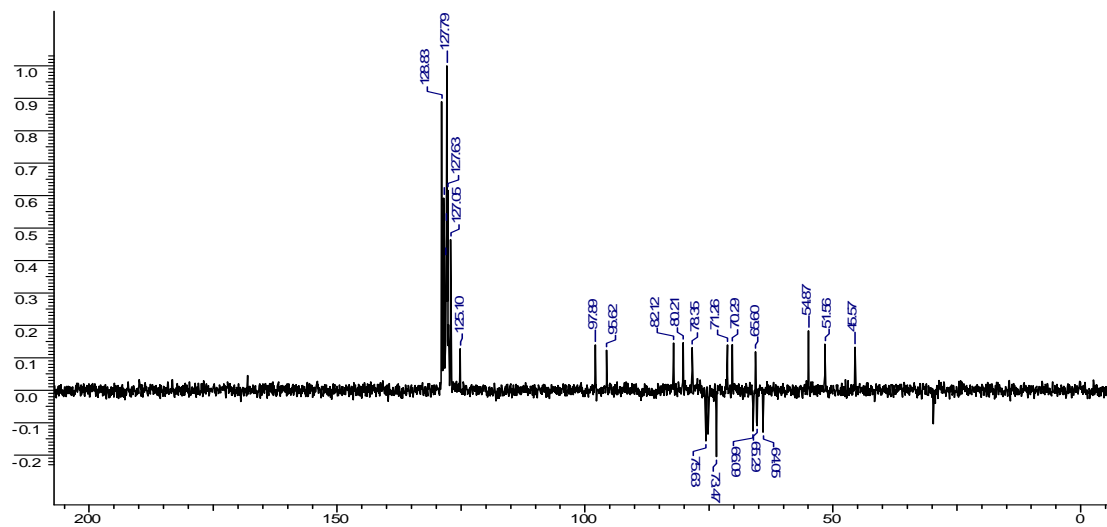
¹H NMR (200 MHz, CDCl₃) of Compound 14



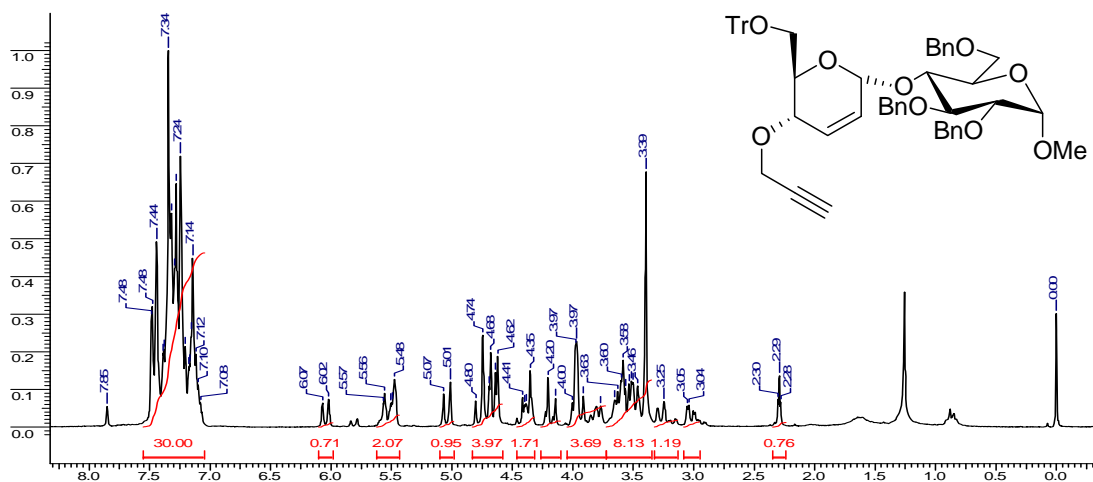
¹³C NMR (50 MHz, CDCl₃) of Compound 14



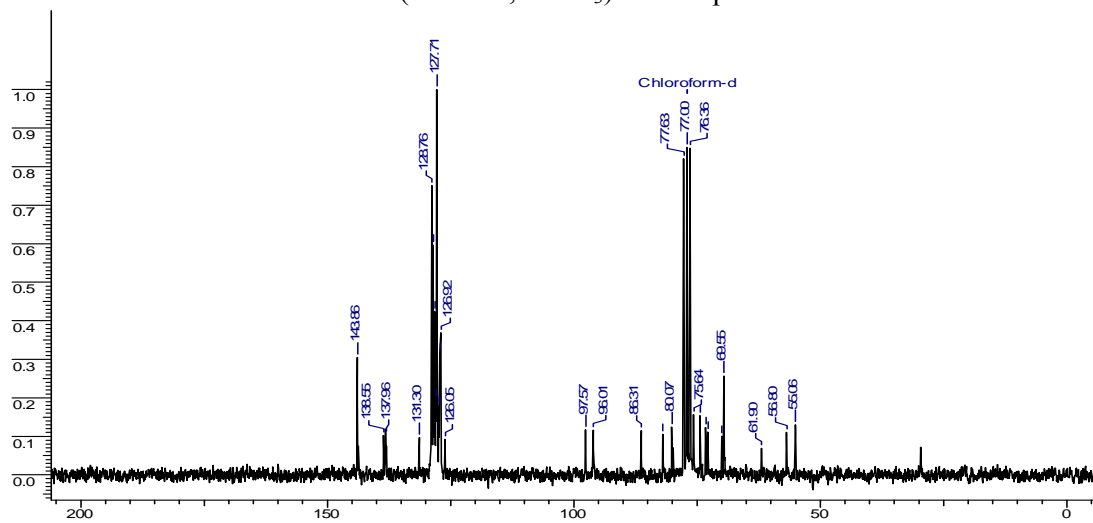
DEPT NMR (50 MHz, CDCl₃) of Compound 14



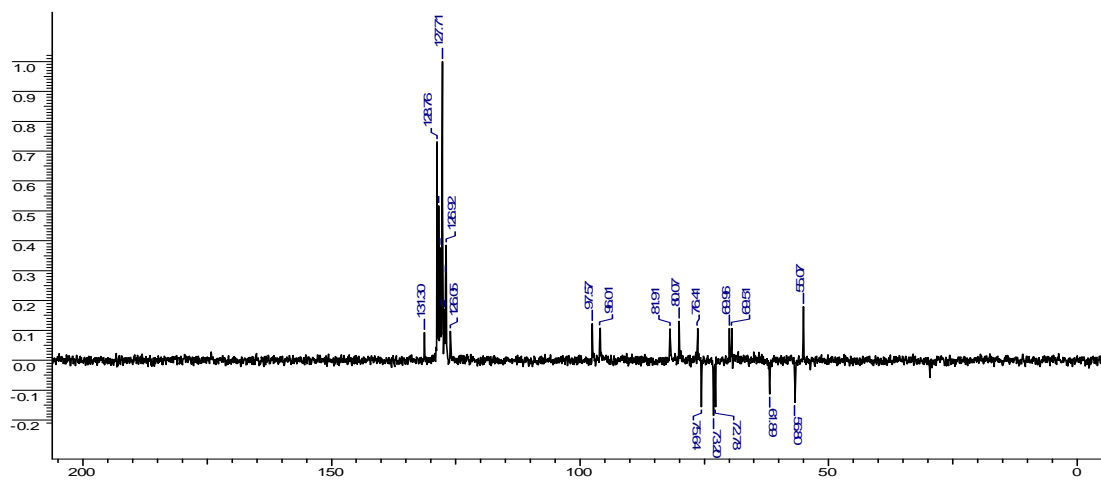
¹H NMR (200 MHz, CDCl₃) of Compound 17



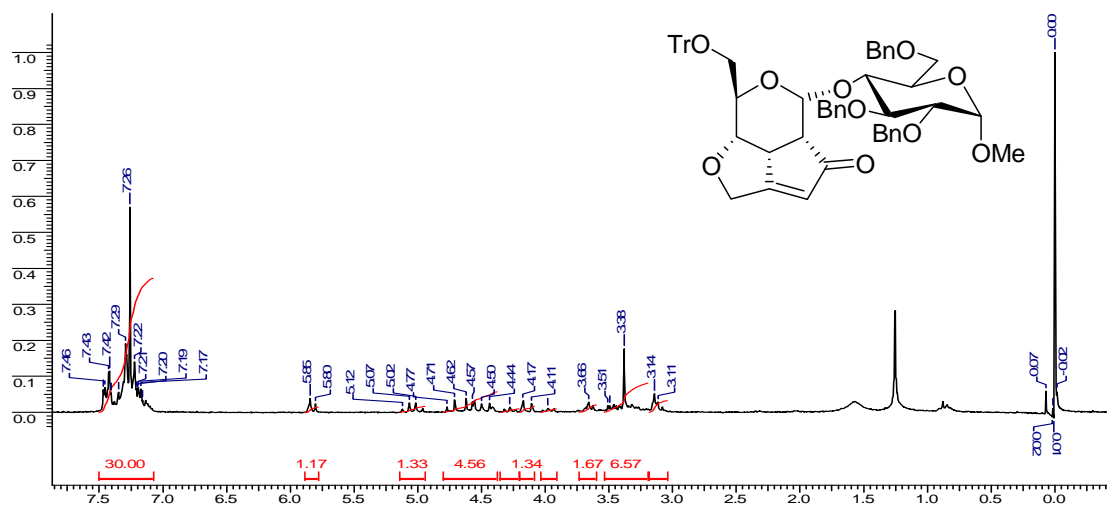
¹³C NMR (50 MHz, CDCl₃) of Compound 17



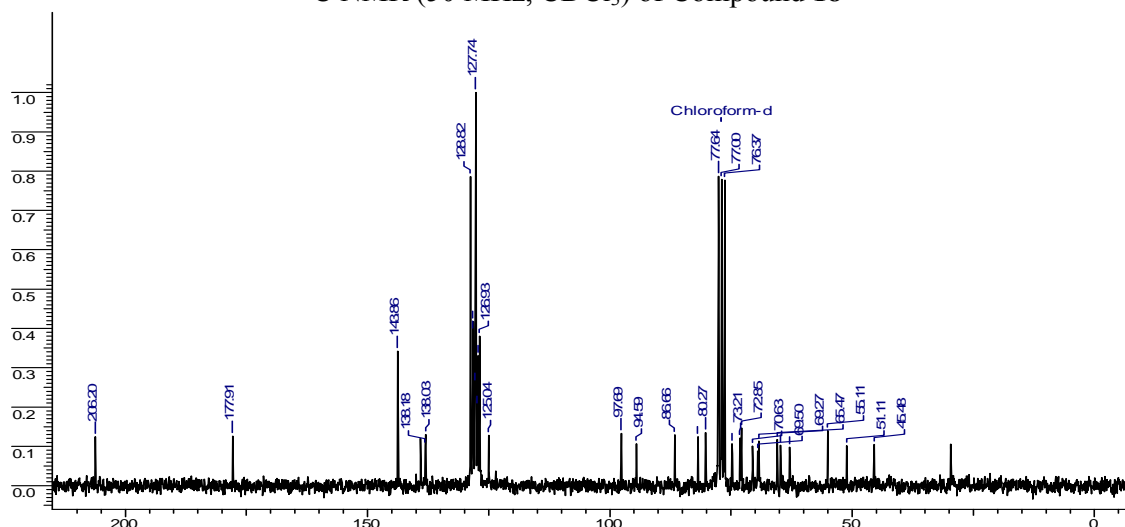
DEPT NMR (50 MHz, CDCl₃) of Compound 17



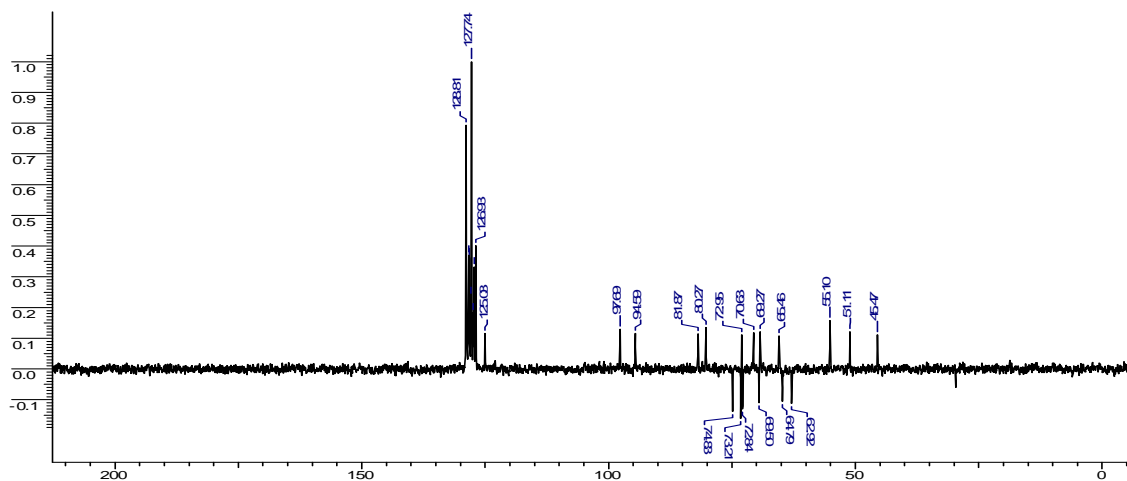
^1H NMR (200 MHz, CDCl_3) of Compound **18**



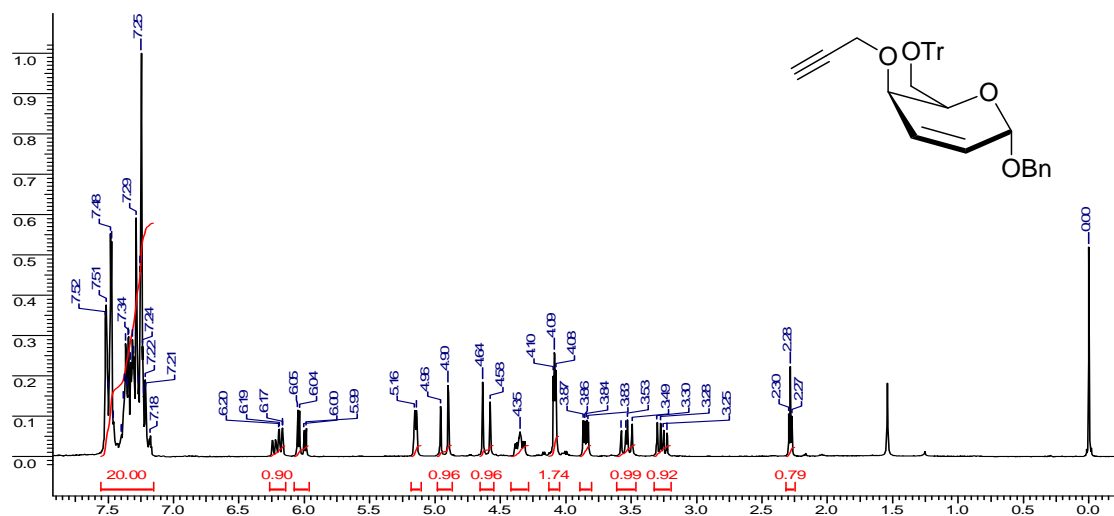
^{13}C NMR (50 MHz, CDCl_3) of Compound **18**



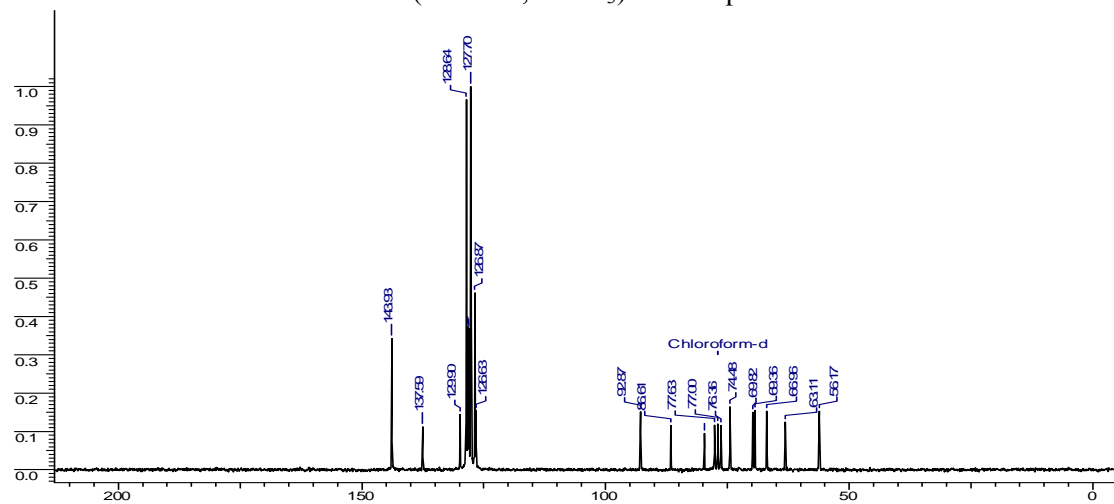
DEPT NMR (50 MHz, CDCl_3) of Compound **18**



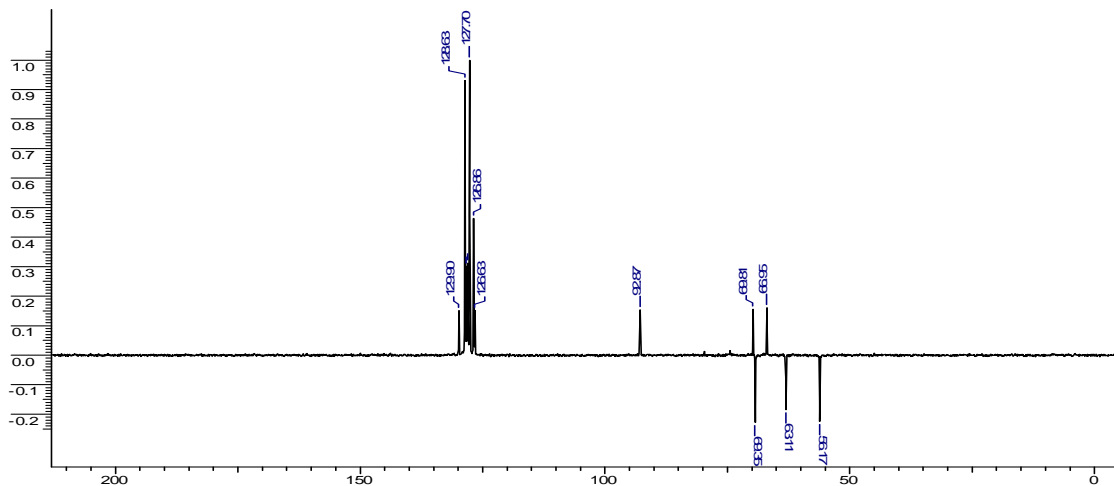
¹H NMR (200 MHz, CDCl₃) of Compound **21**



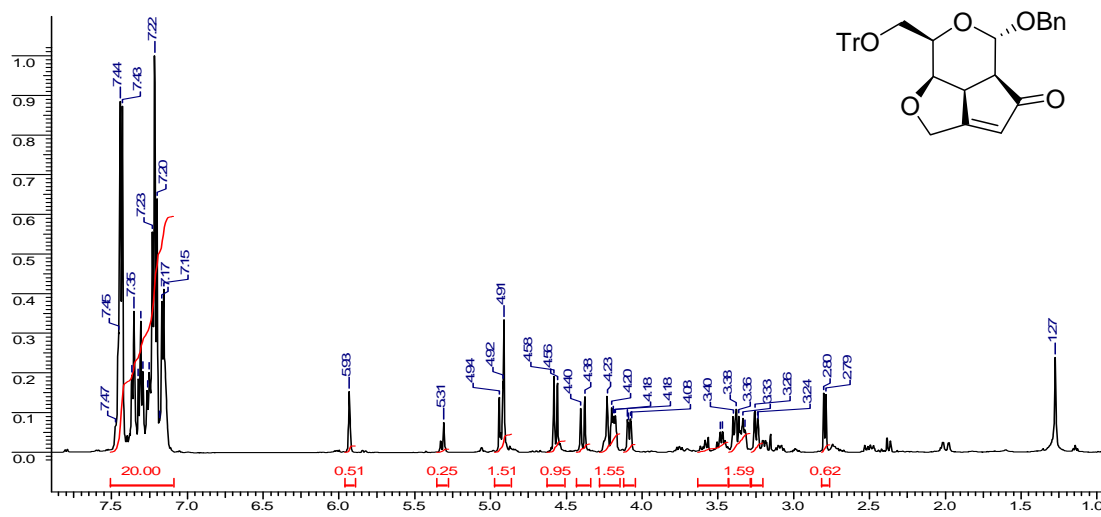
¹³C NMR (50 MHz, CDCl₃) of Compound **21**



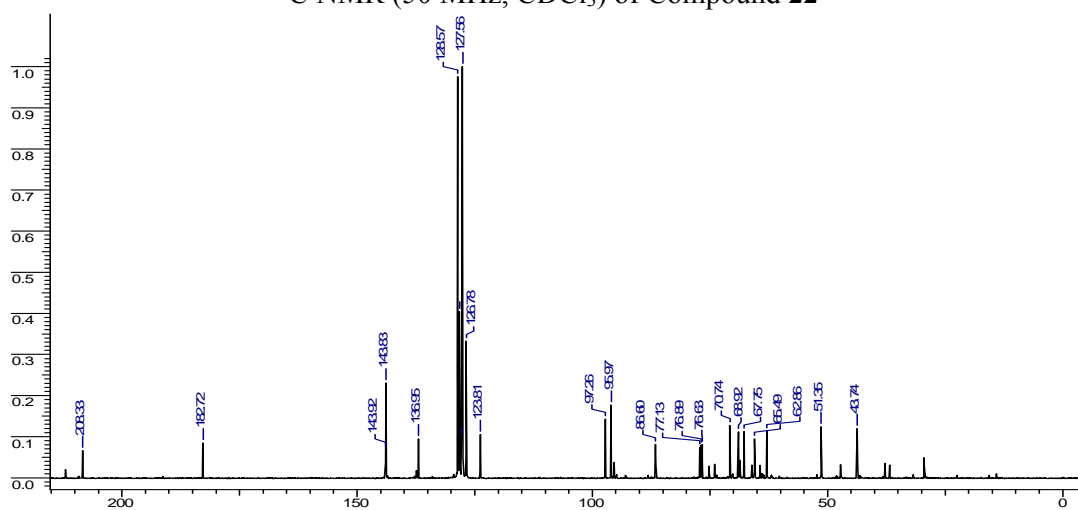
DEPT NMR (50 MHz, CDCl₃) of Compound **21**



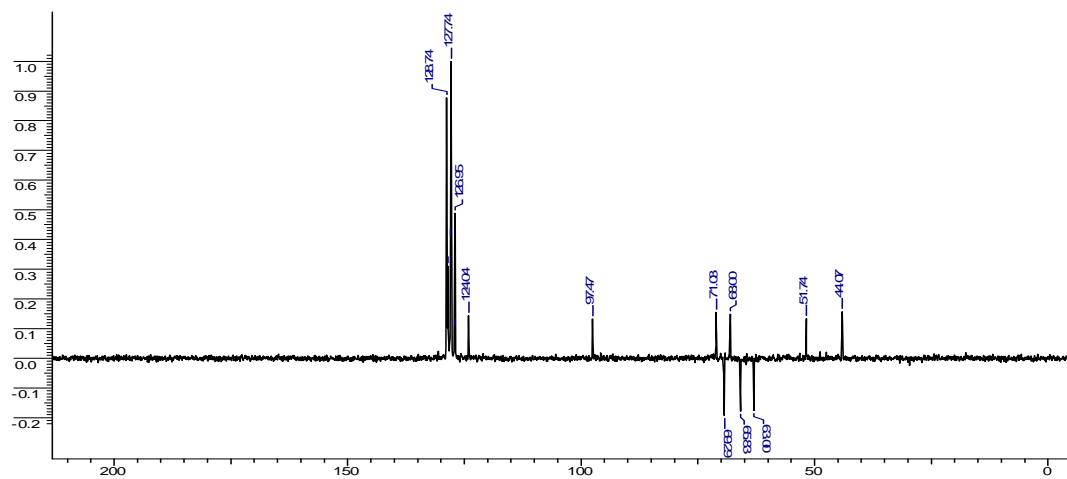
¹H NMR (200 MHz, CDCl₃) of Compound 22



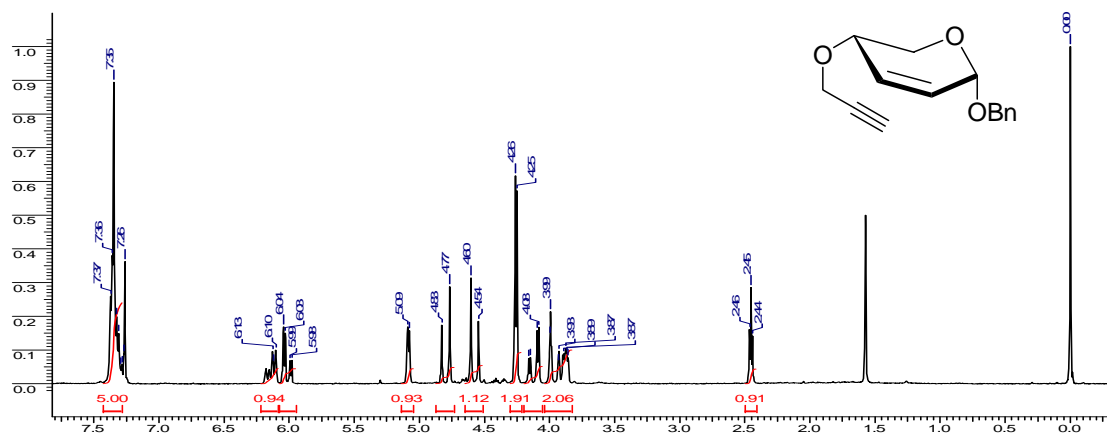
¹³C NMR (50 MHz, CDCl₃) of Compound 22



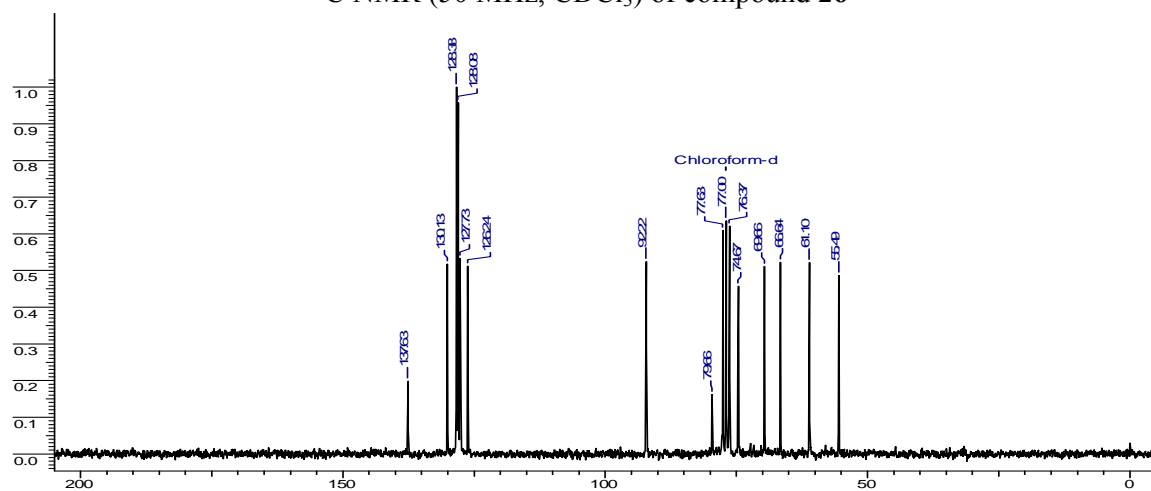
DEPT NMR (50 MHz, CDCl₃) of Compound 22



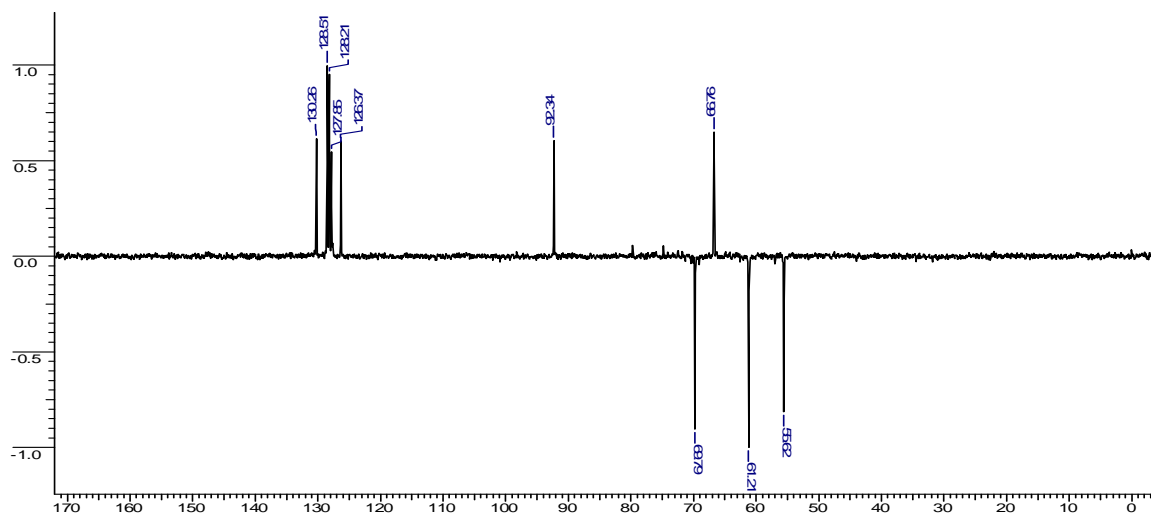
¹H NMR (200 MHz, CDCl₃) of Compound 26



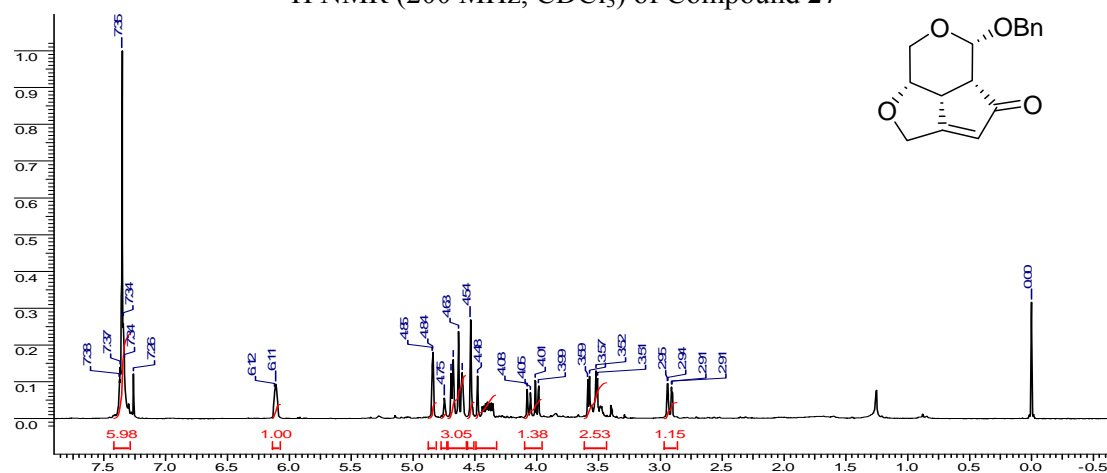
¹³C NMR (50 MHz, CDCl₃) of compound 26



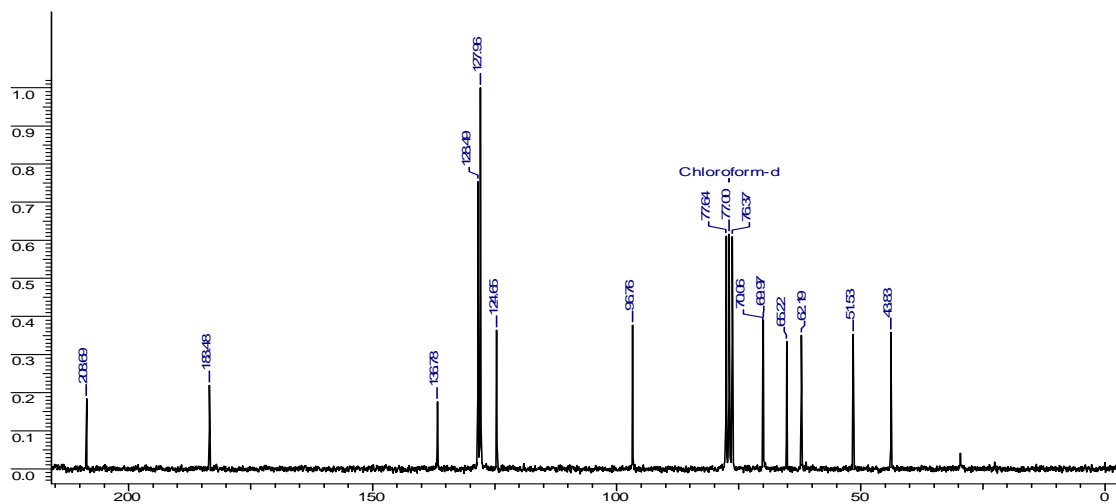
DEPT NMR (50 MHz, CDCl₃) of Compound 26



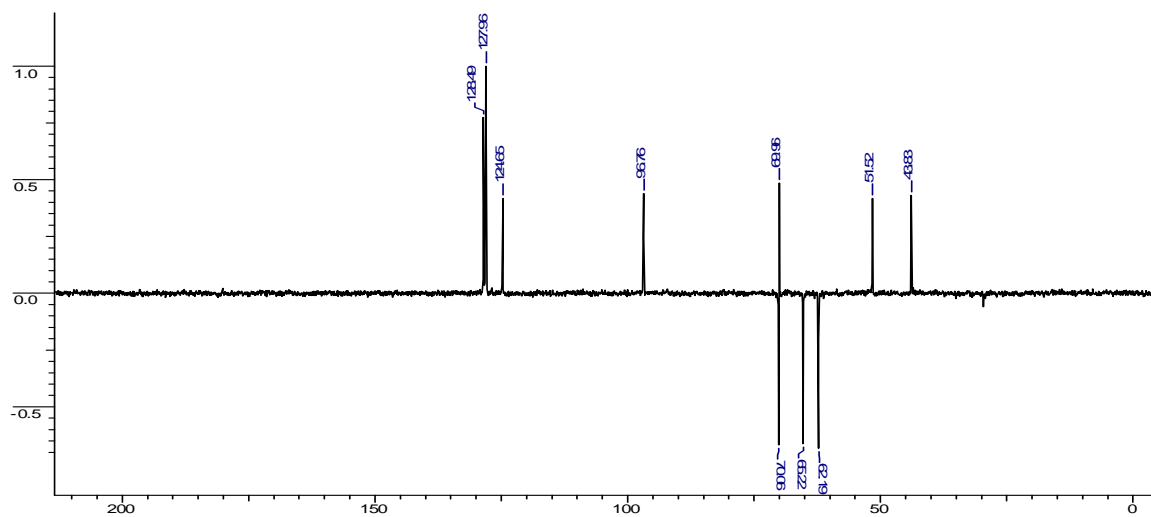
^1H NMR (200 MHz, CDCl_3) of Compound 27



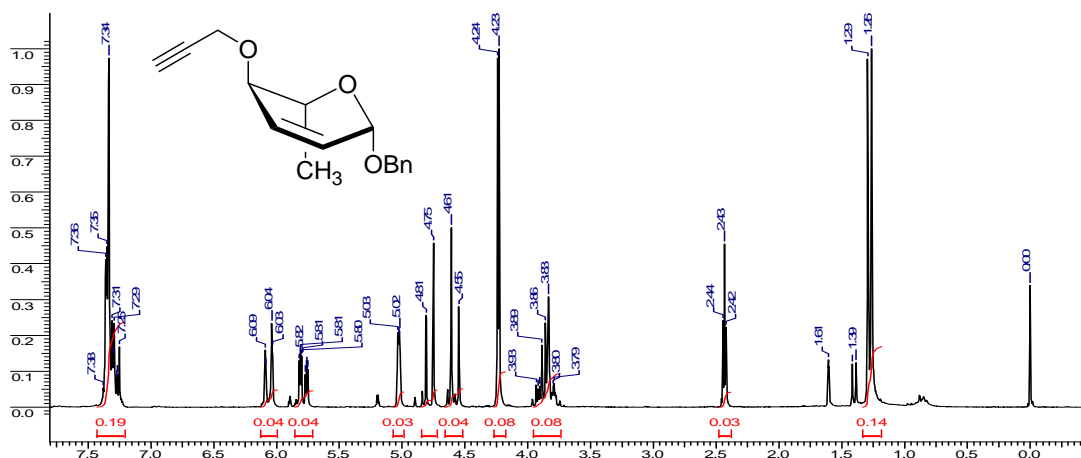
^{13}C NMR (50 MHz, CDCl_3) of compound 27



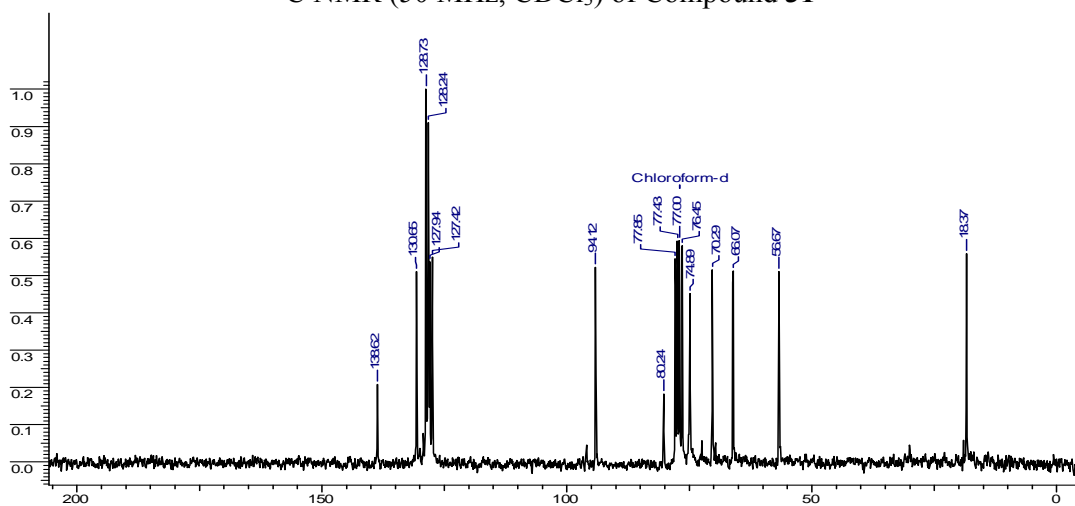
DEPT NMR (50 MHz, CDCl_3) of Compound 27



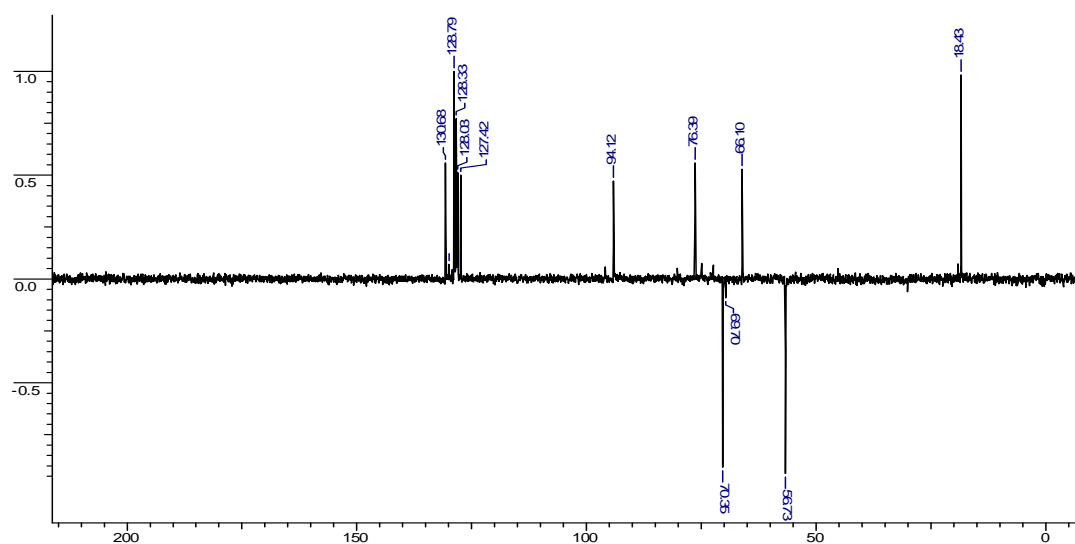
¹H NMR (200 MHz, CDCl₃) of Compound 31



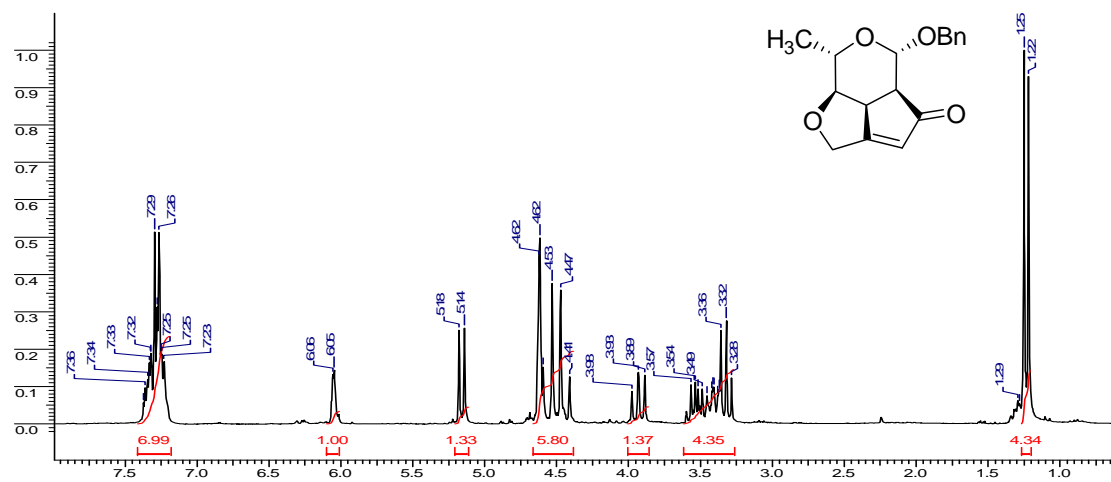
¹³C NMR (50 MHz, CDCl₃) of Compound 31



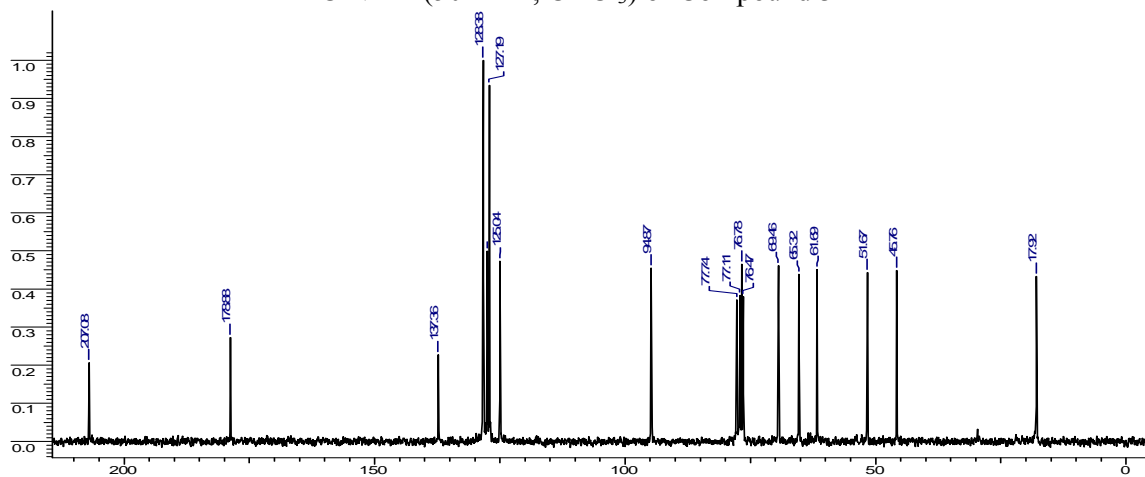
DEPT NMR (50 MHz, CDCl₃) of Compound 31



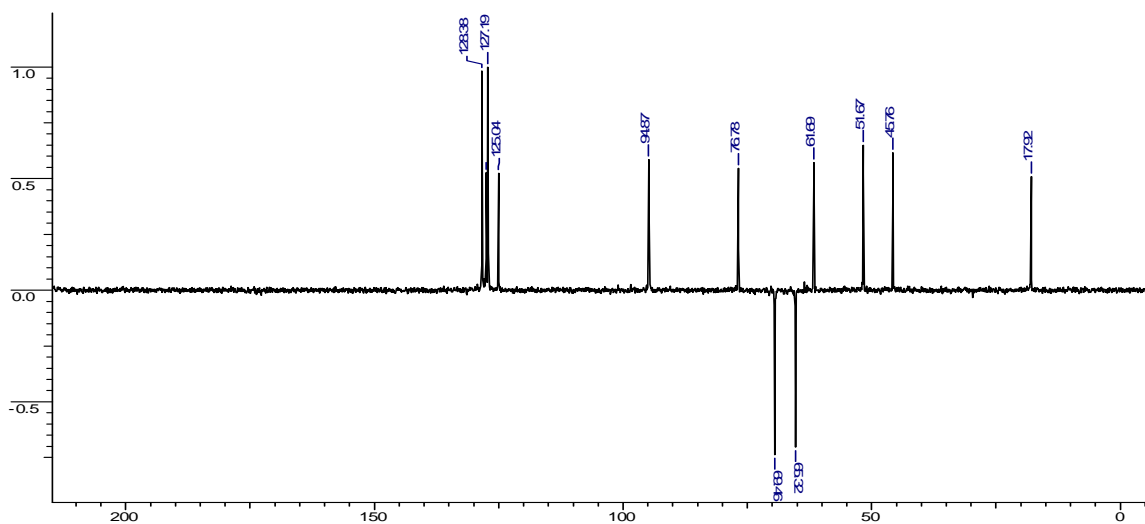
¹H NMR (200 MHz, CDCl₃) of Compound **32**



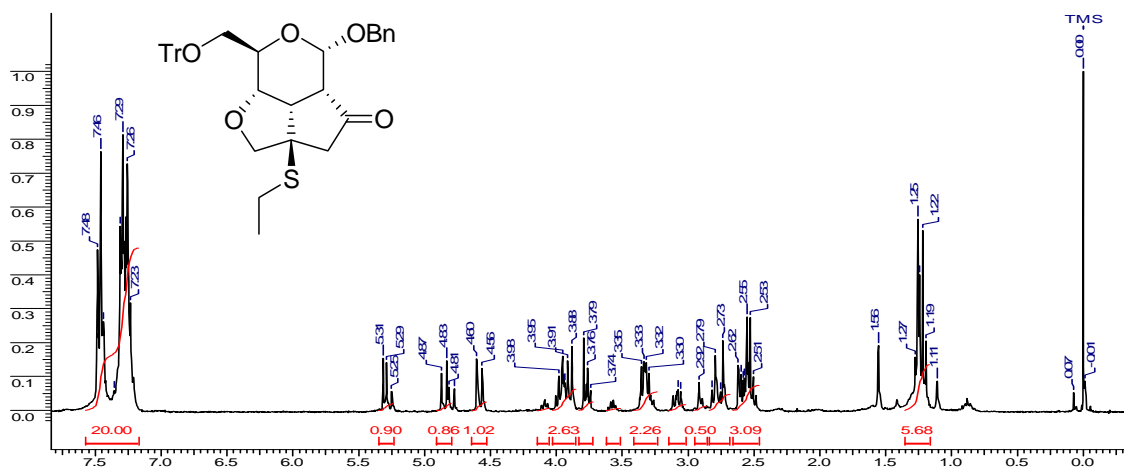
¹³C NMR (50 MHz, CDCl₃) of Compound **32**



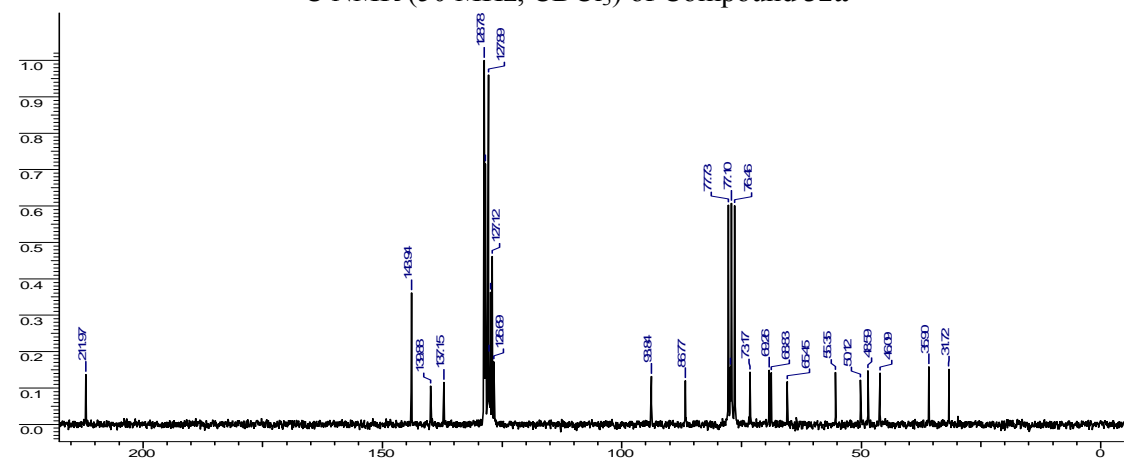
DEPT NMR (50 MHz, CDCl₃) of Compound **32**



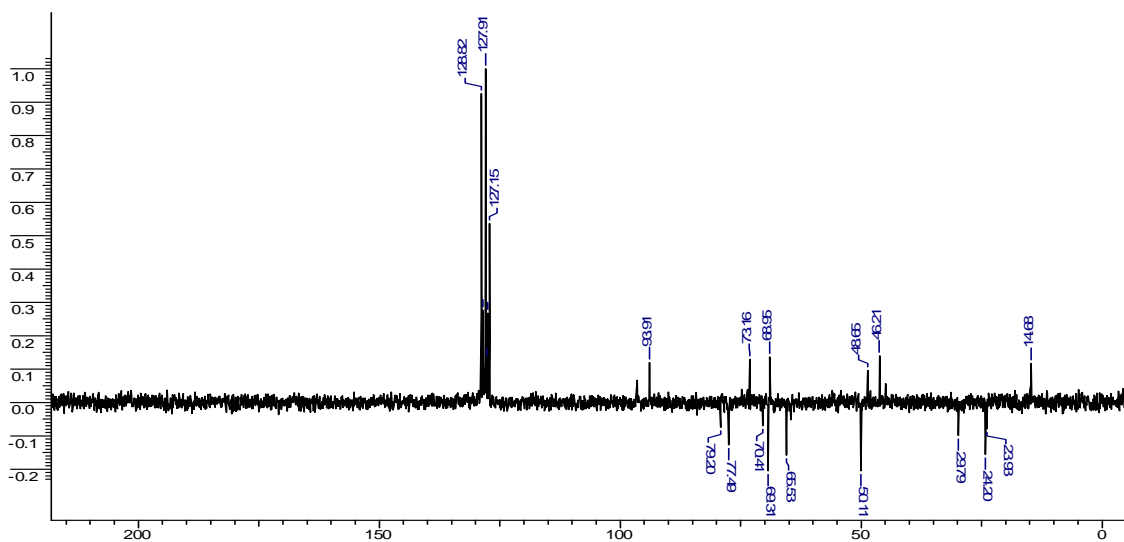
¹H NMR (200 MHz, CDCl₃) of Compound **32a**



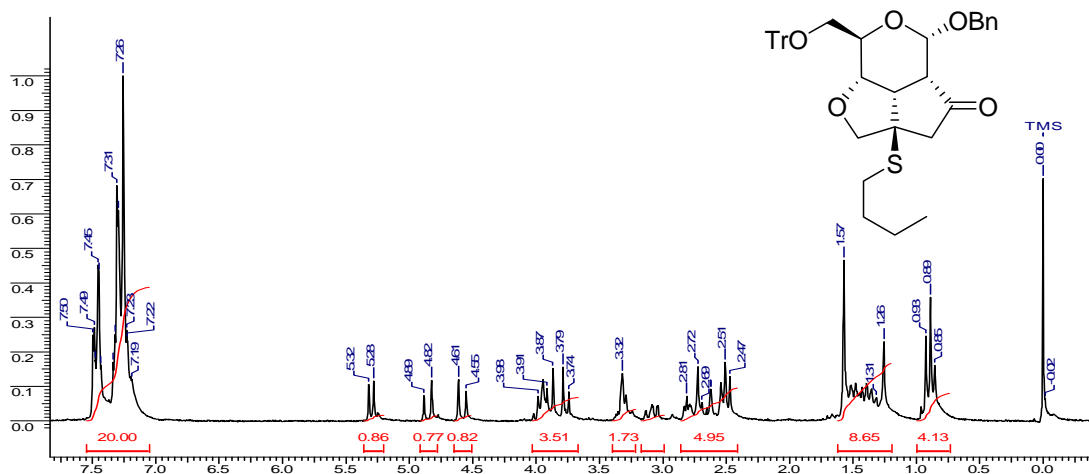
¹³C NMR (50 MHz, CDCl₃) of Compound **32a**



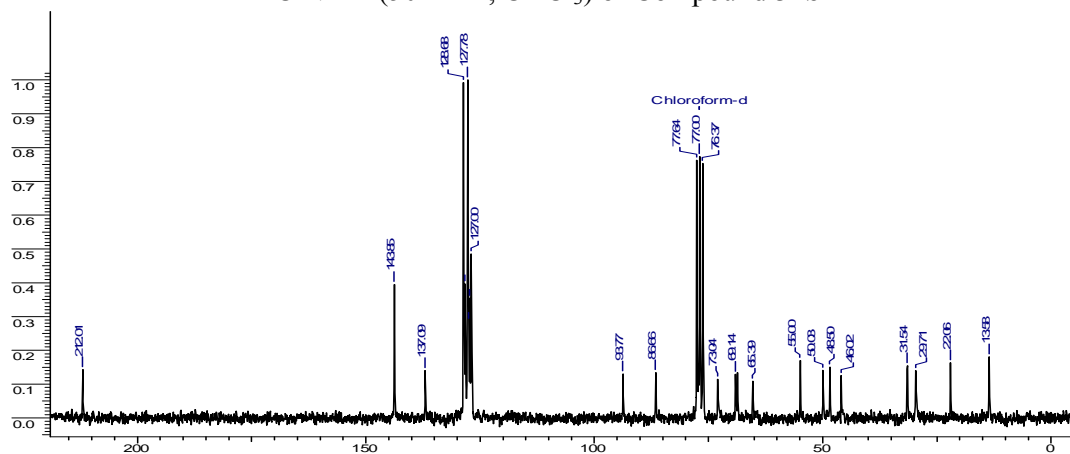
DEPT NMR (50 MHz, CDCl₃) of Compound **32a**



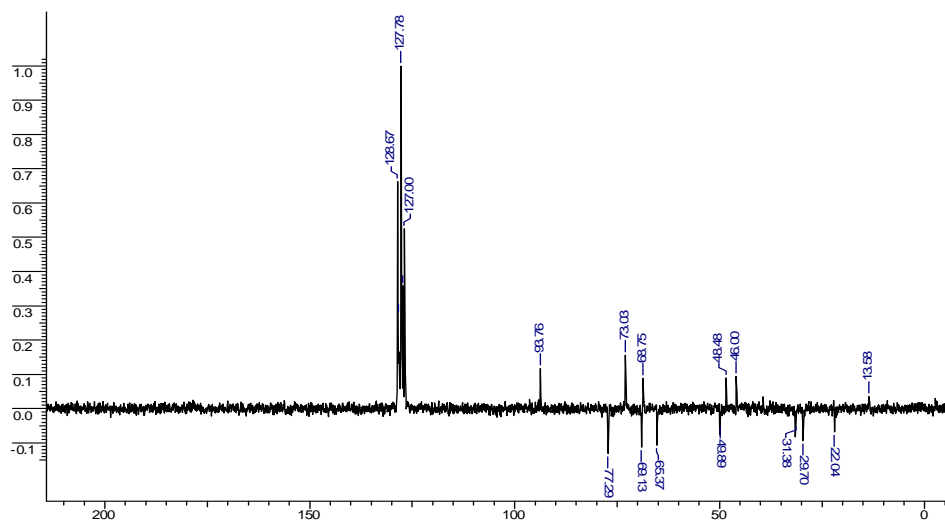
^1H NMR (200 MHz, CDCl_3) of Compound **32b**



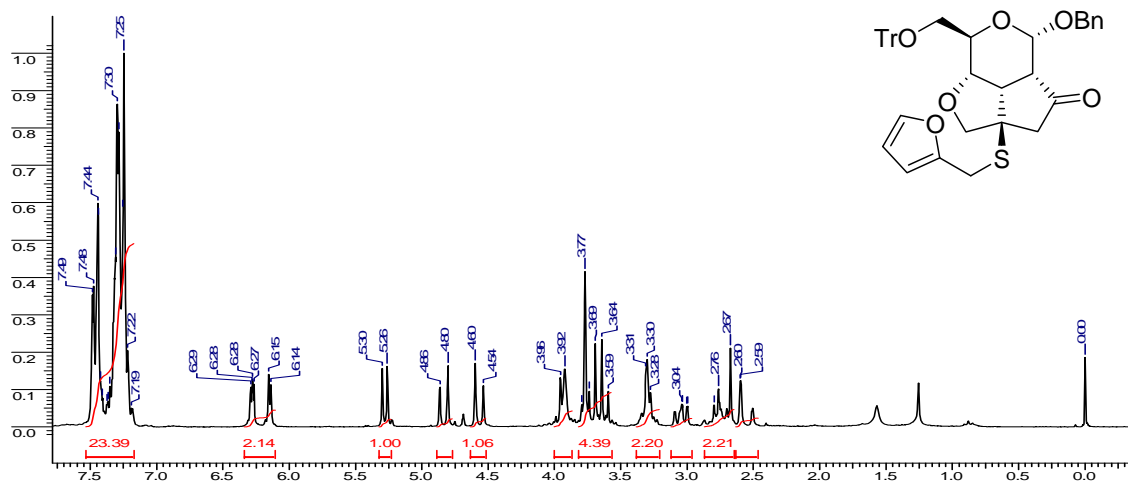
^{13}C NMR (50 MHz, CDCl_3) of Compound **32b**



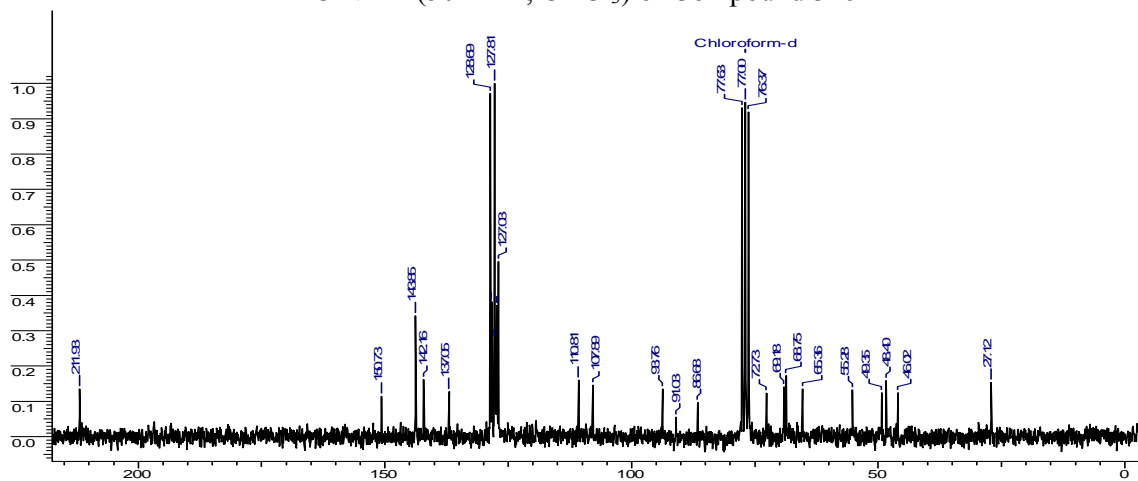
DEPT NMR (50 MHz, CDCl_3) of Compound **32b**



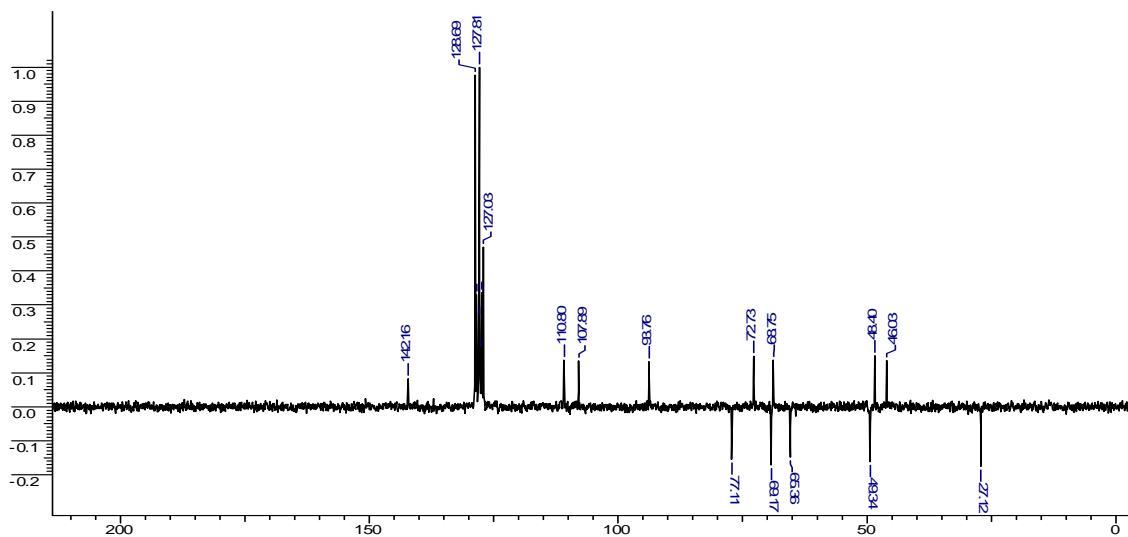
^1H NMR (200 MHz, CDCl_3) of Compound **32c**



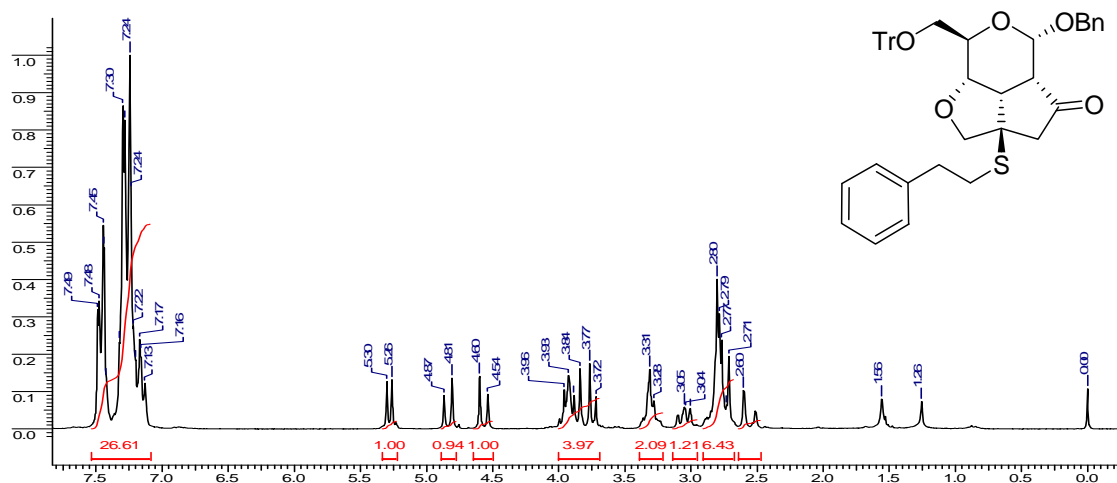
^{13}C NMR (50 MHz, CDCl_3) of Compound **32c**



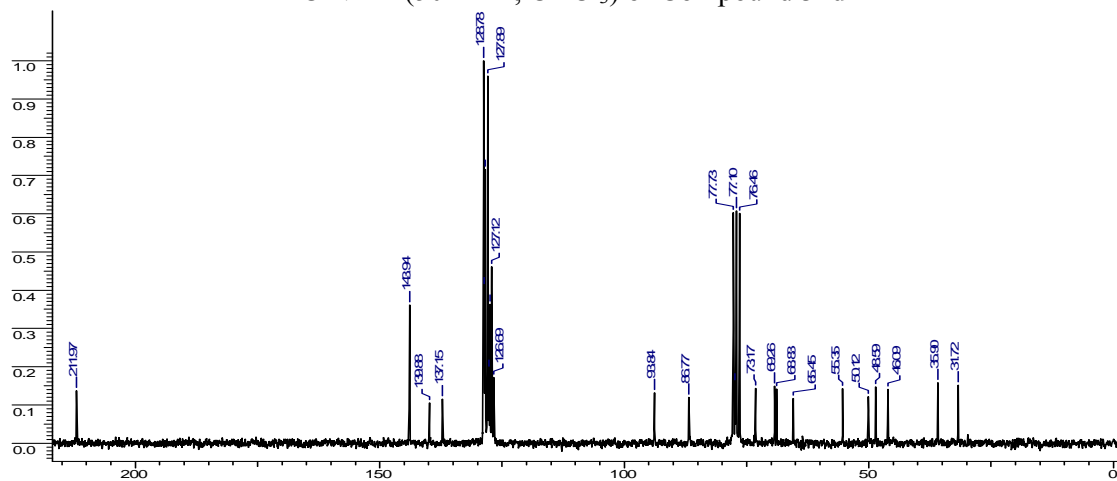
DEPT NMR (50 MHz, CDCl_3) of Compound **32c**



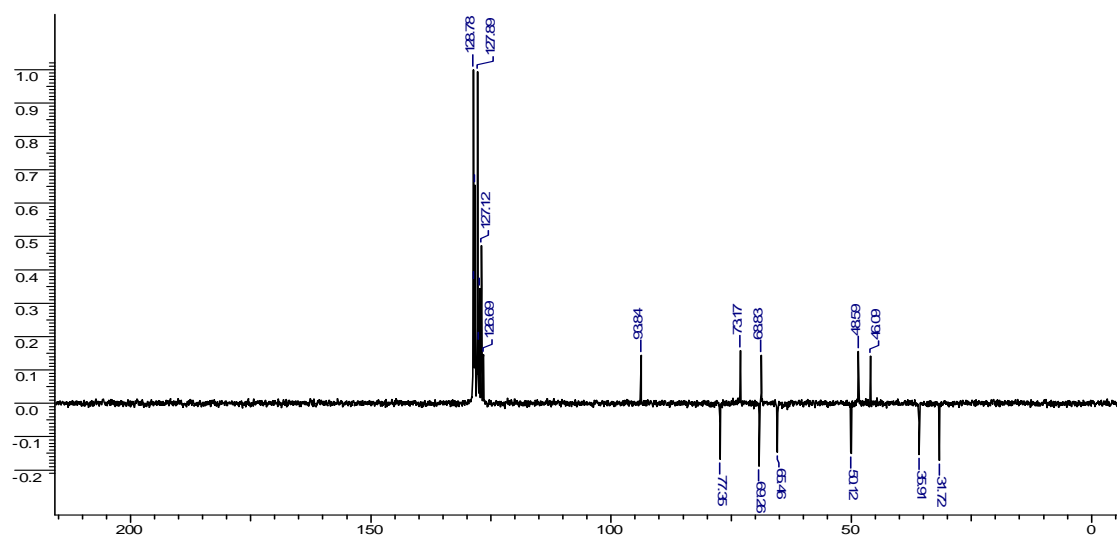
¹H NMR (200 MHz, CDCl₃) of Compound **32d**



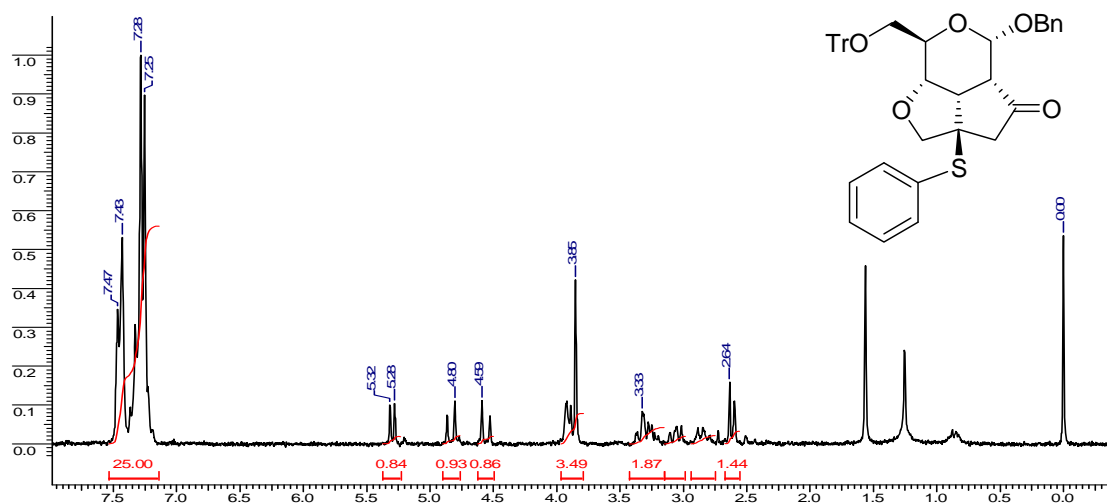
¹³C NMR (50 MHz, CDCl₃) of Compound **32d**



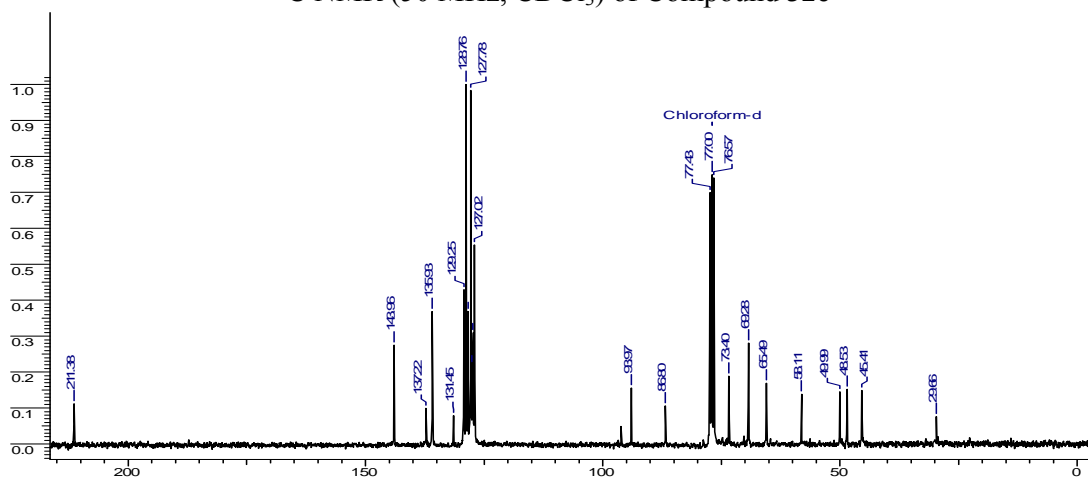
DEPT NMR (50 MHz, CDCl₃) of Compound **32d**



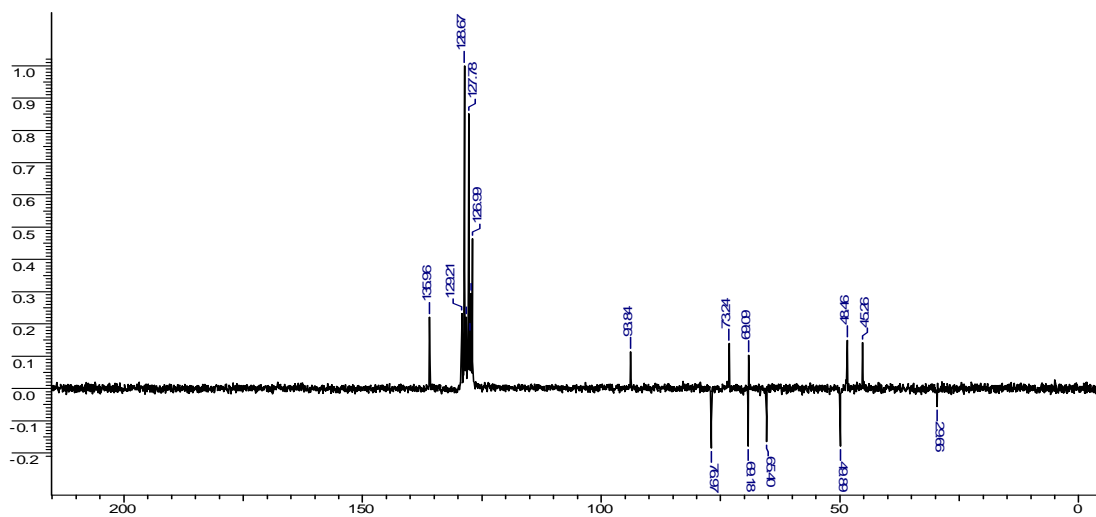
^1H NMR (200 MHz, CDCl_3) of Compound **32e**



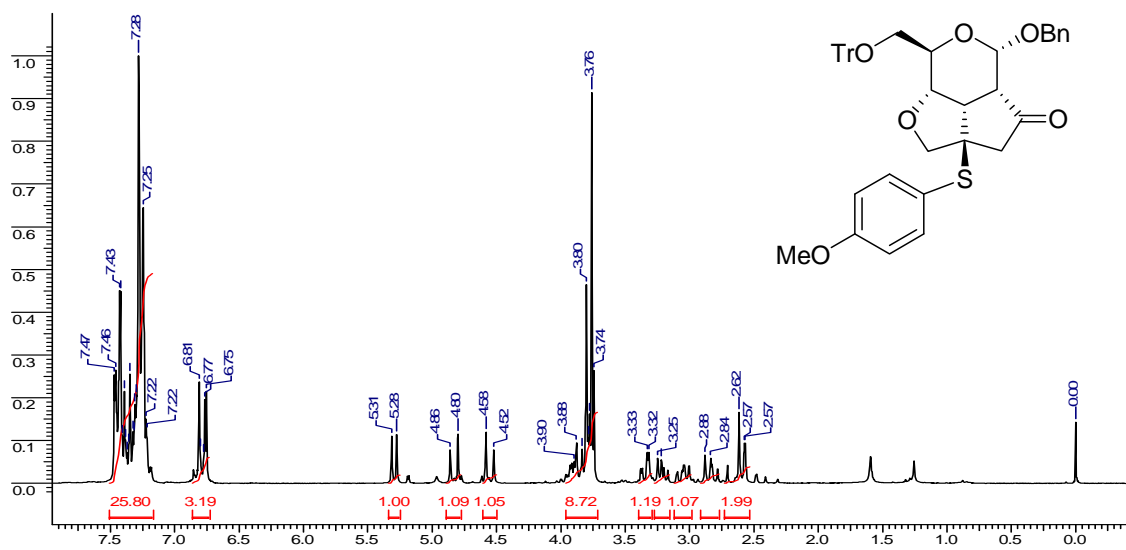
^{13}C NMR (50 MHz, CDCl_3) of Compound **32e**



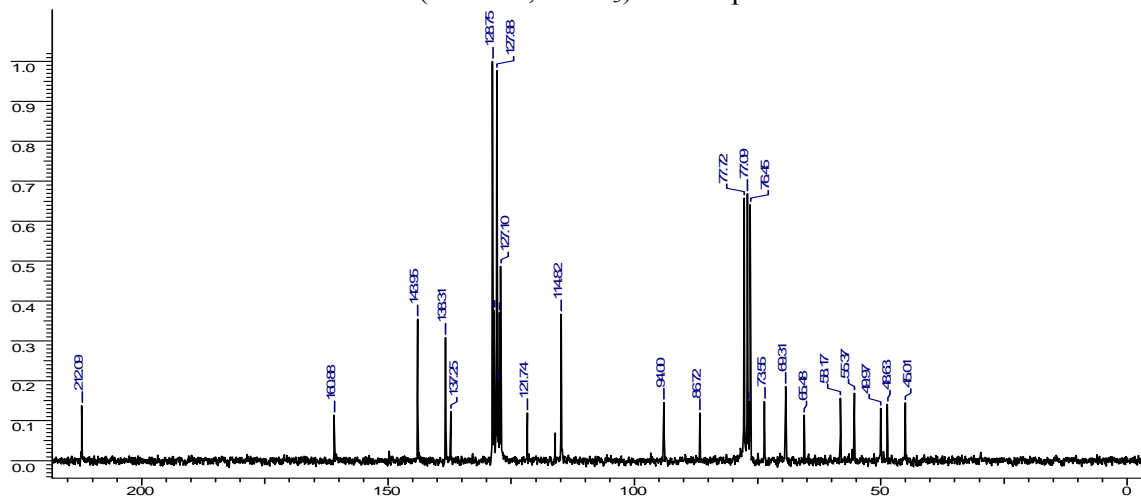
DEPT NMR (50 MHz, CDCl_3) of Compound **32e**



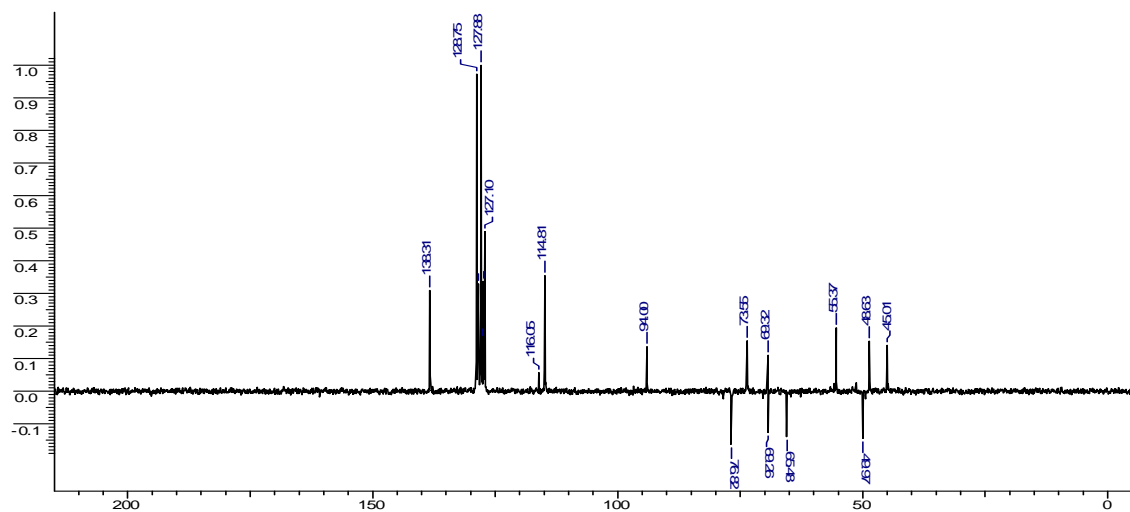
^1H NMR (200 MHz, CDCl_3) of Compound **32f**



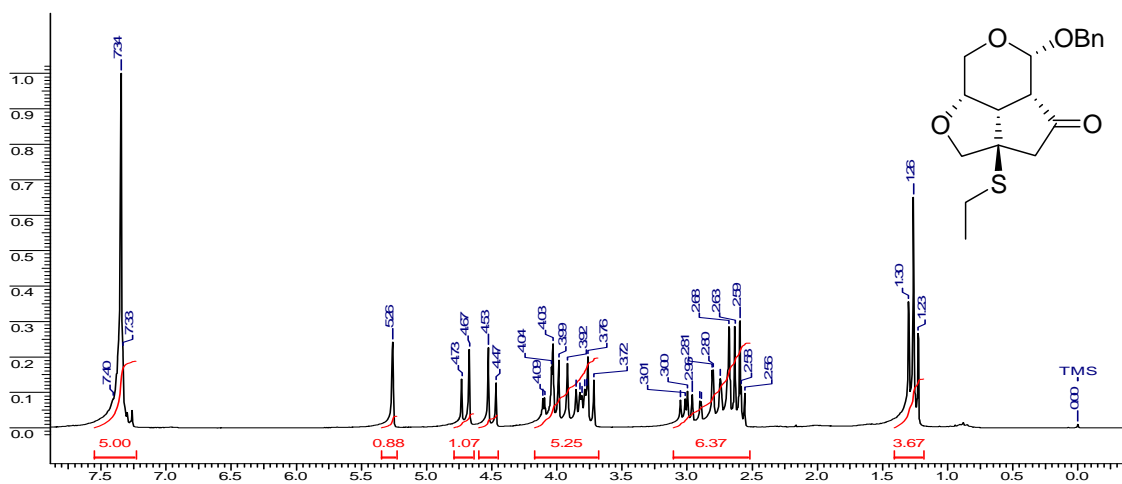
^{13}C NMR (50 MHz, CDCl_3) of Compound **32f**



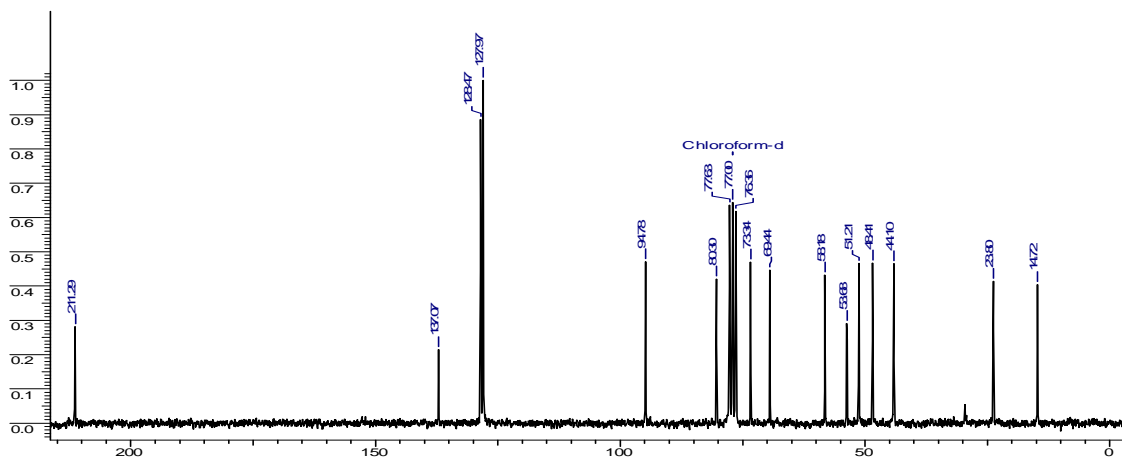
DEPT NMR (50 MHz, CDCl_3) of Compound **32f**



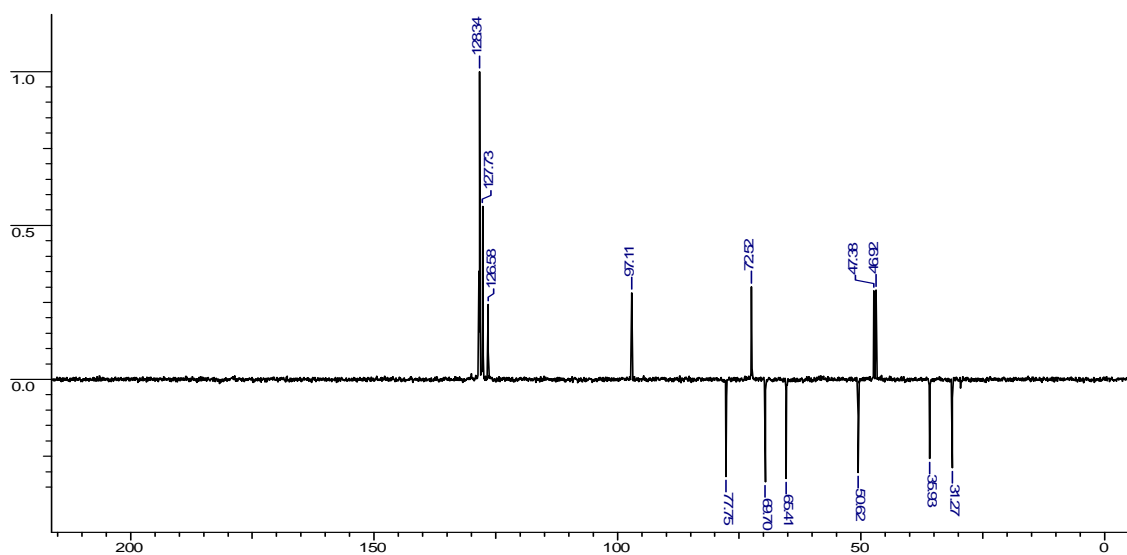
^1H NMR (200 MHz, CDCl_3) of Compound **33a**



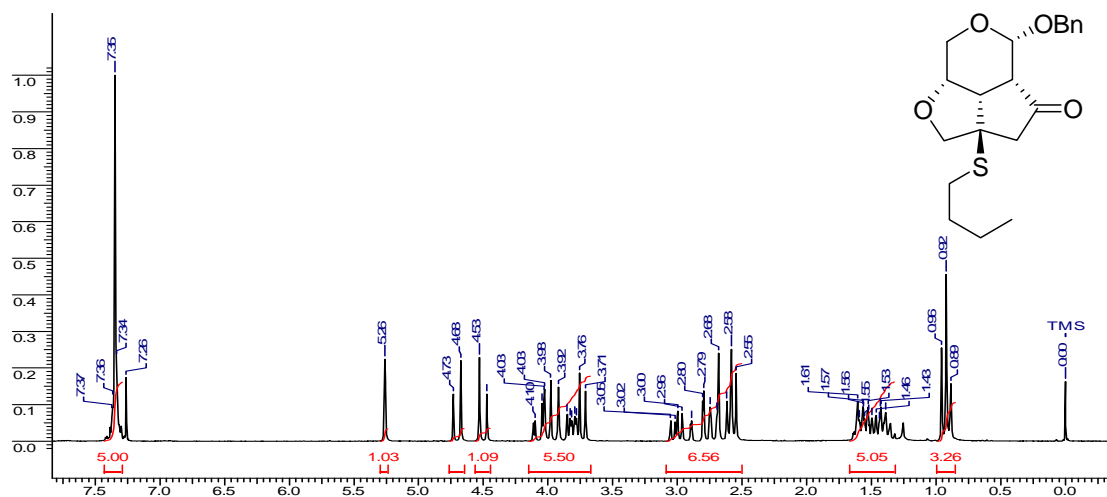
^{13}C NMR (50 MHz, CDCl_3) of Compound **33a**



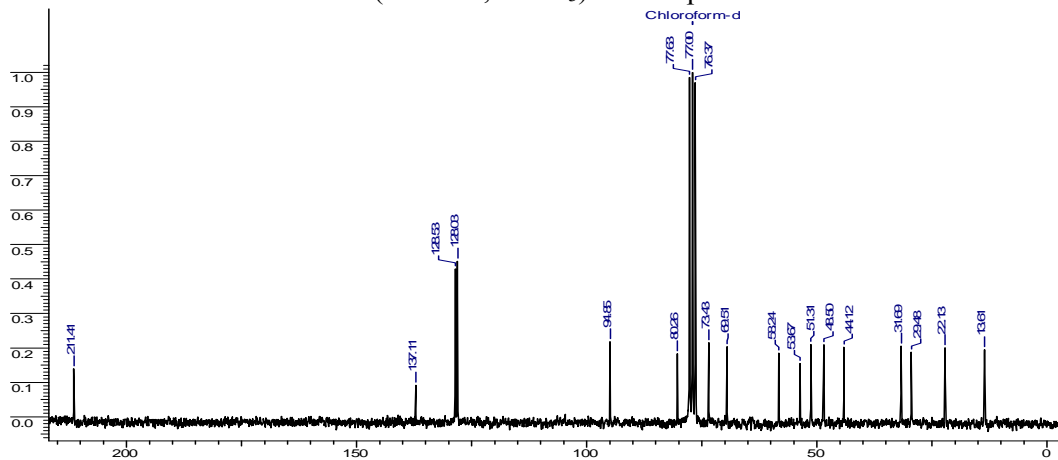
DEPT NMR (50 MHz, CDCl_3) of Compound **33a**



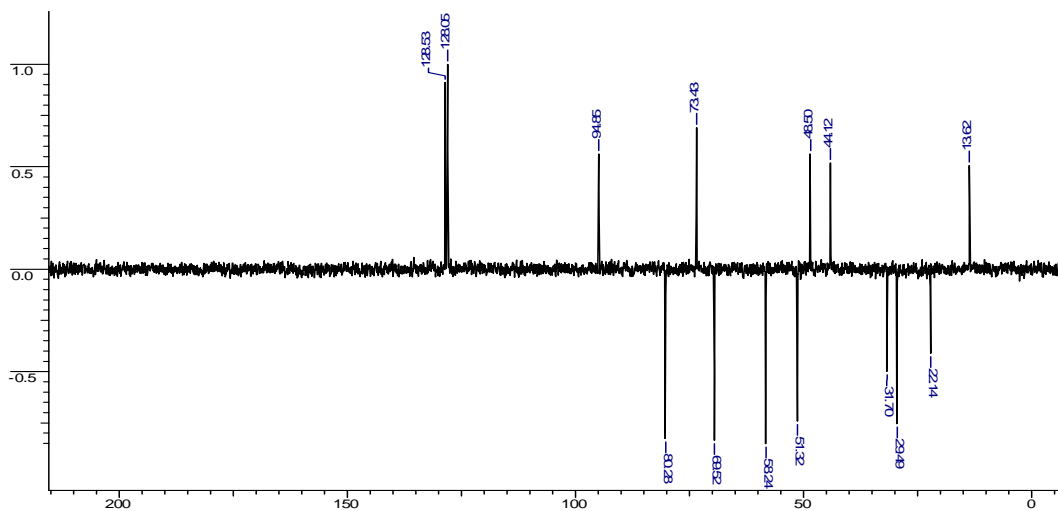
^1H NMR (200 MHz, CDCl_3) of Compound **33b**



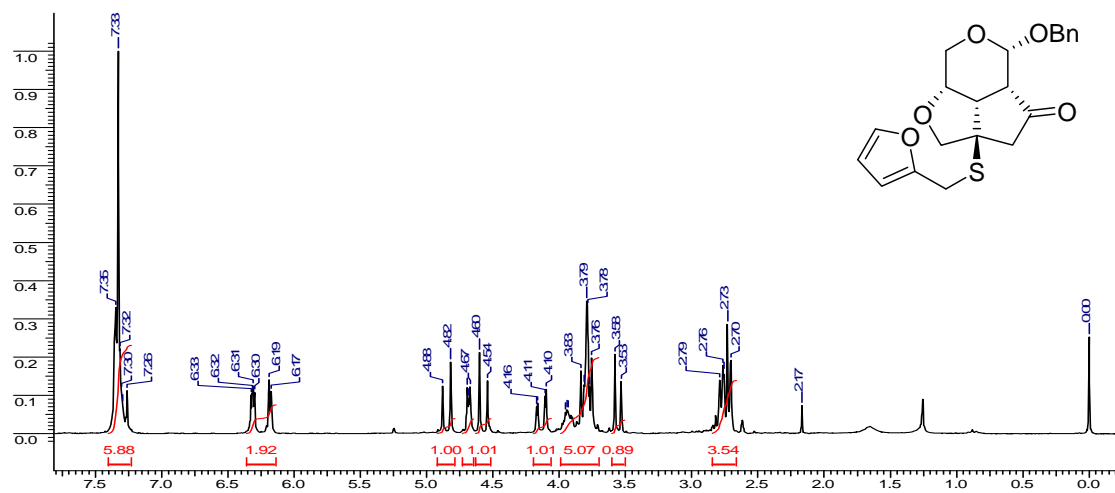
^{13}C NMR (50 MHz, CDCl_3) of Compound **33b**



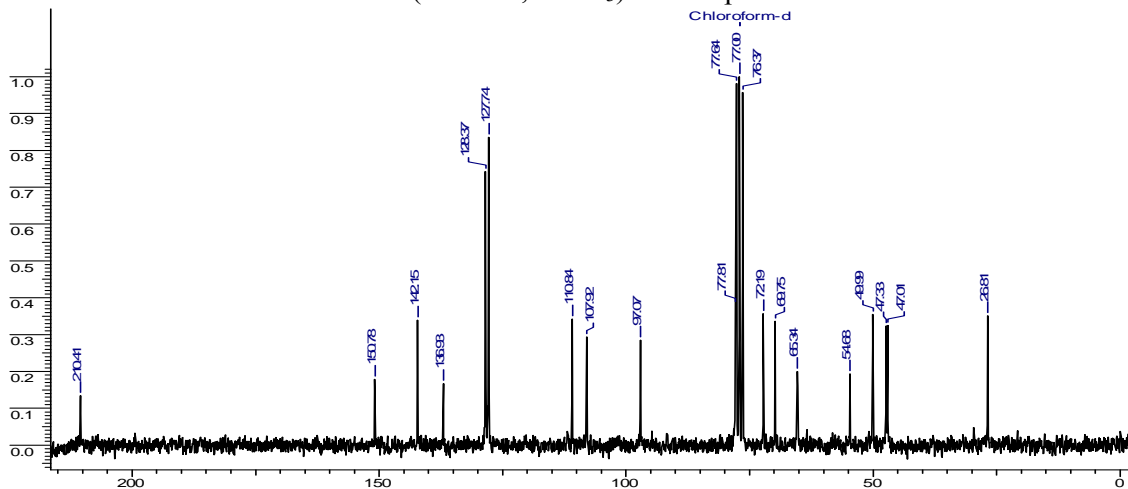
DEPT NMR (50 MHz, CDCl_3) of Compound **33b**



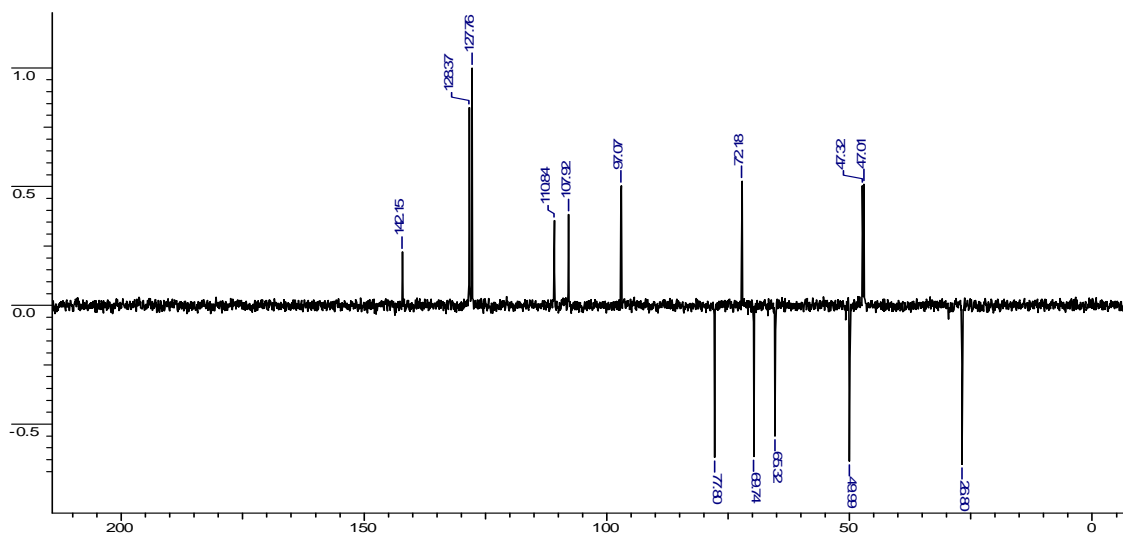
¹H NMR (200 MHz, CDCl₃) of Compound **33c**



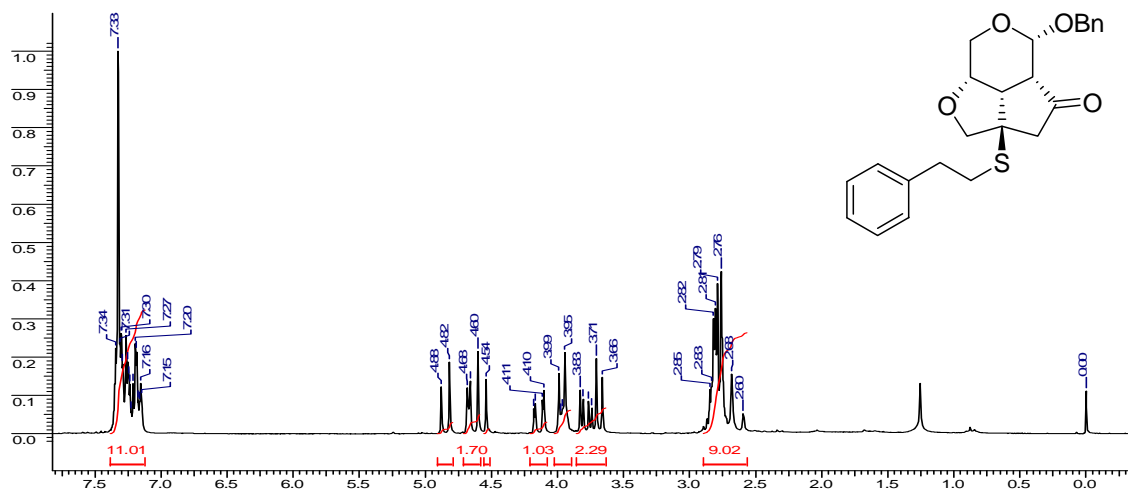
¹³C NMR (50 MHz, CDCl₃) of Compound **33c**



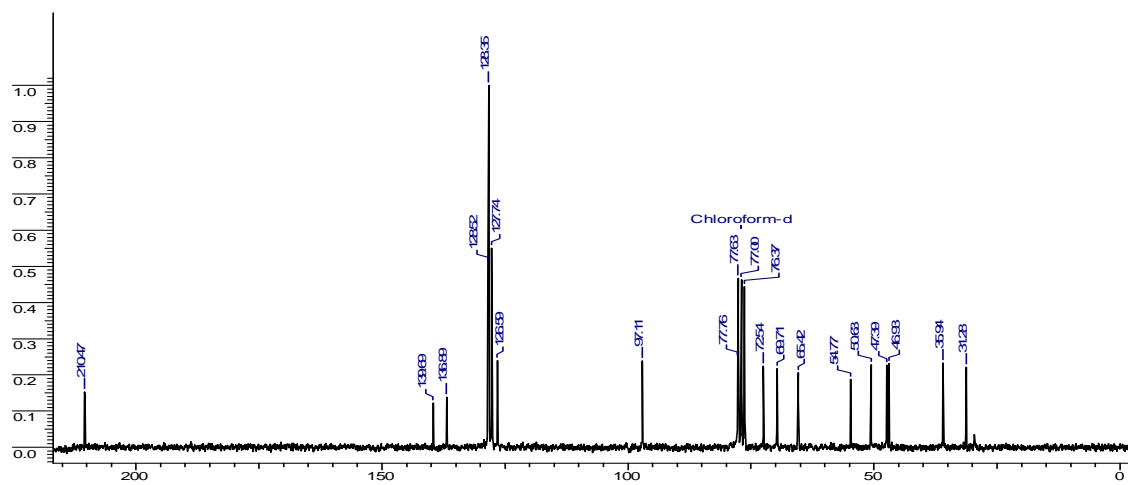
DEPT NMR (50 MHz, CDCl₃) of Compound **33c**



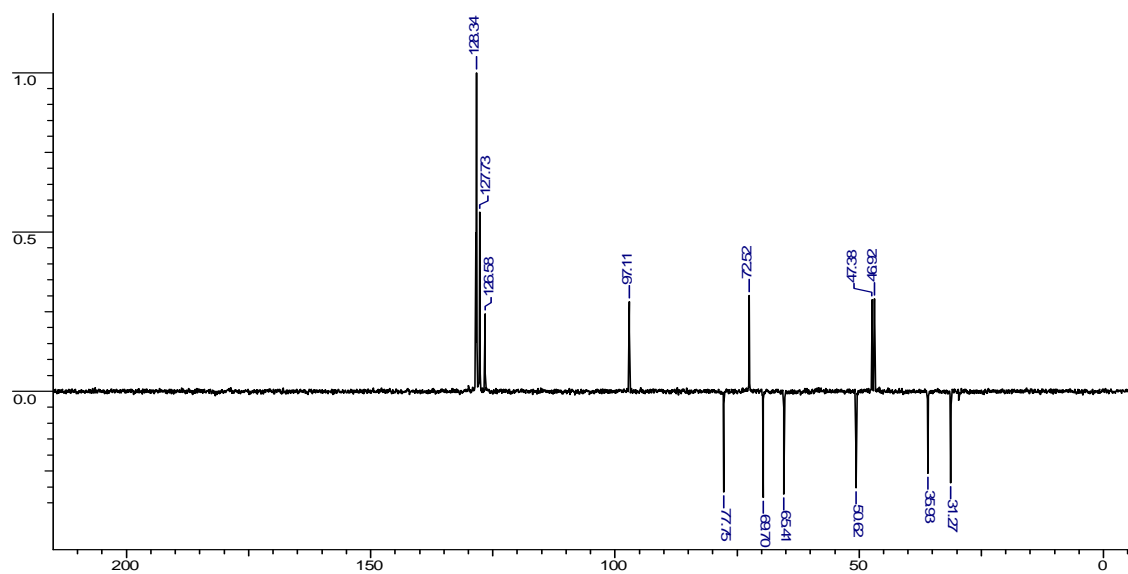
^1H NMR (200 MHz, CDCl_3) of Compound **33d**



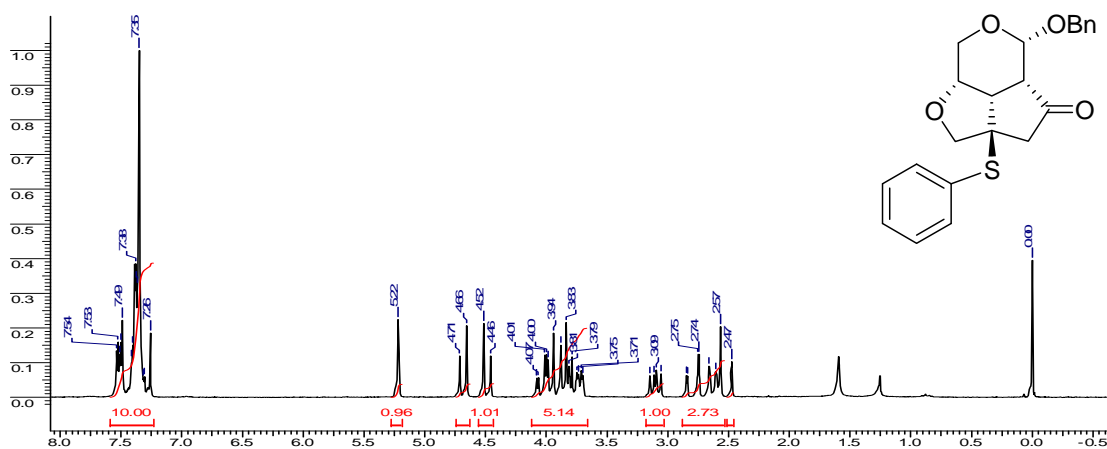
^{13}C NMR (50 MHz, CDCl_3) of Compound **33d**



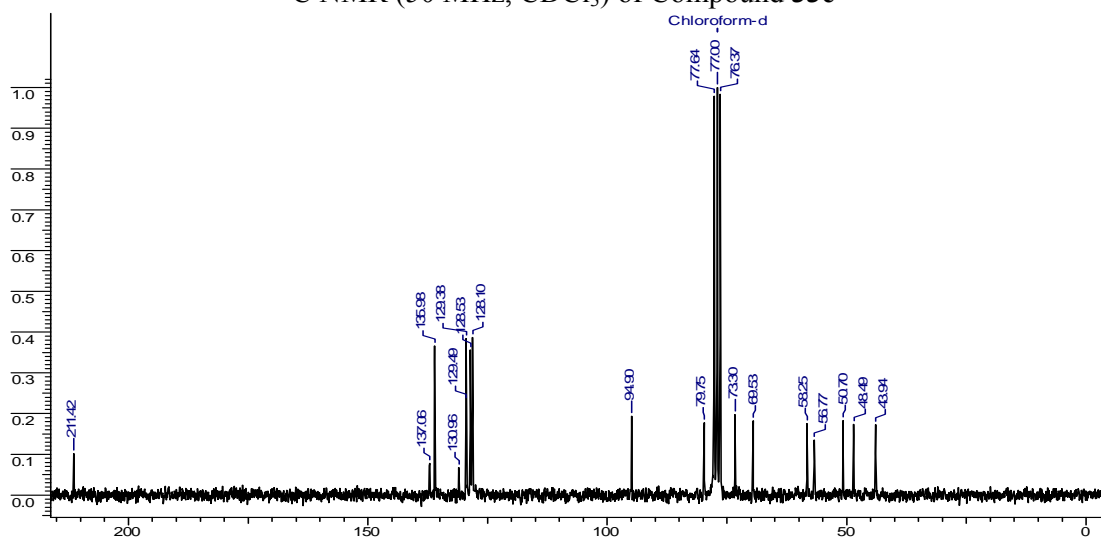
DEPT NMR (50 MHz, CDCl_3) of Compound **33d**



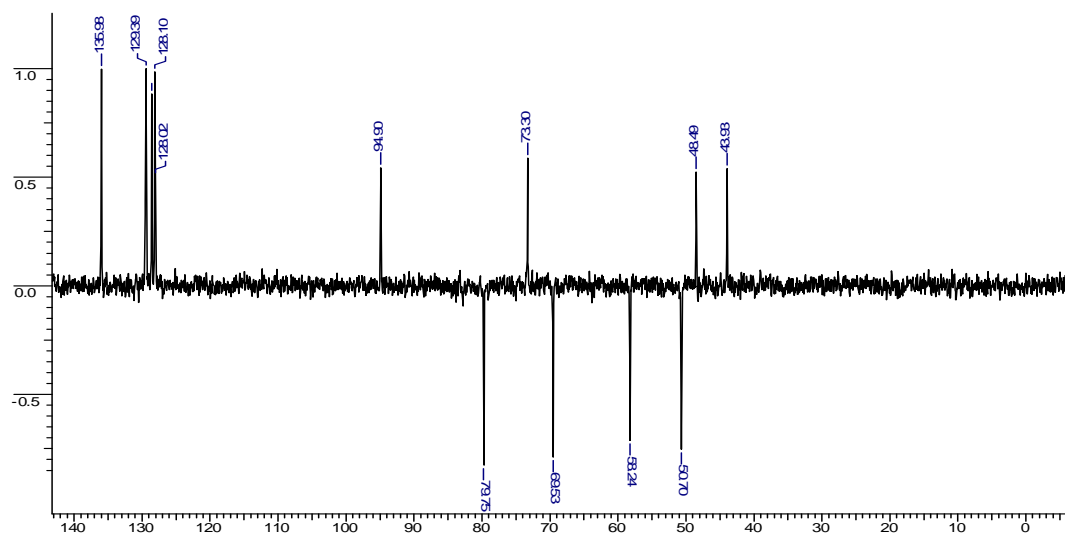
^1H NMR (200 MHz, CDCl_3) of Compound **33e**



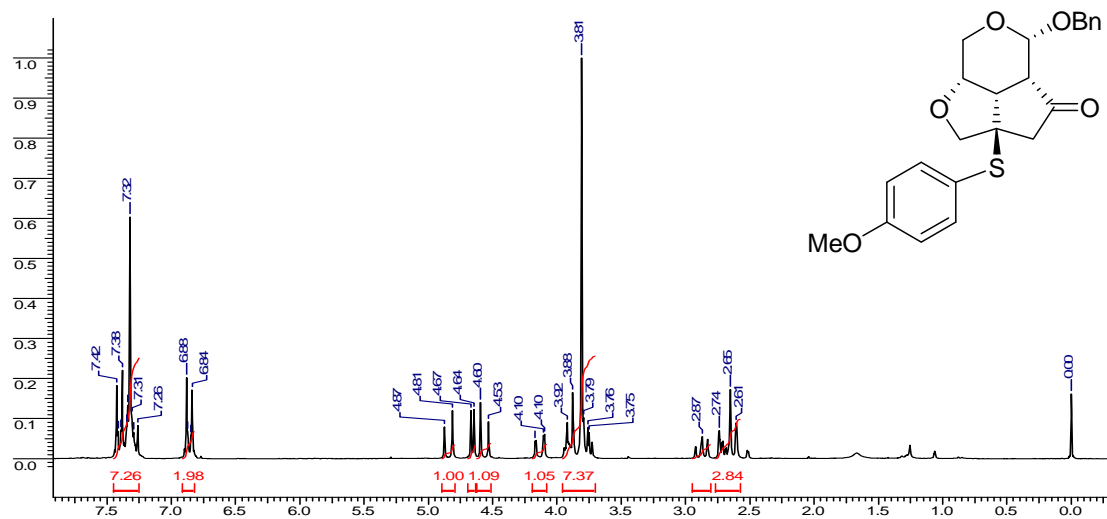
^{13}C NMR (50 MHz, CDCl_3) of Compound **33e**



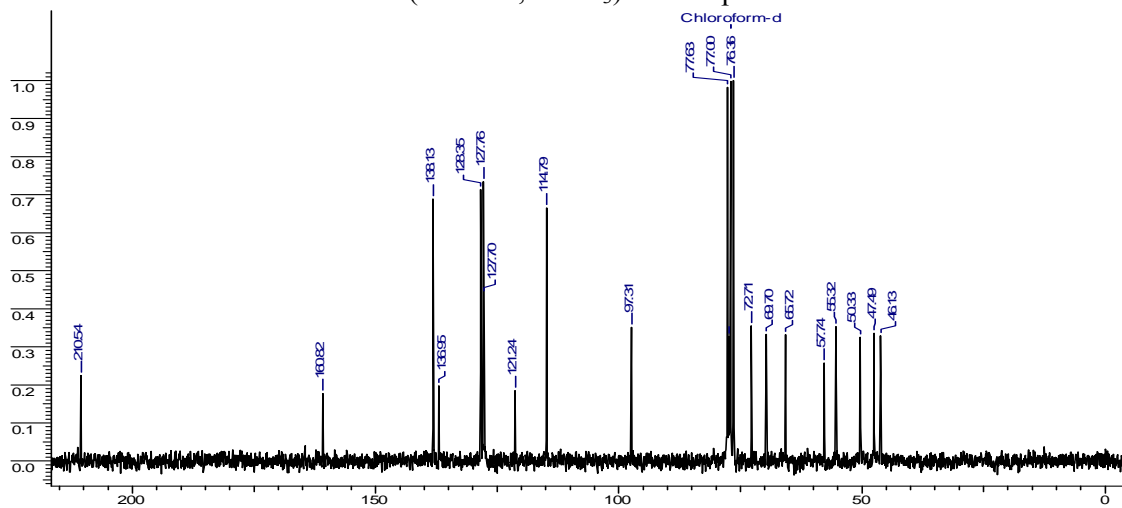
DEPT NMR (50 MHz, CDCl_3) of Compound **33e**



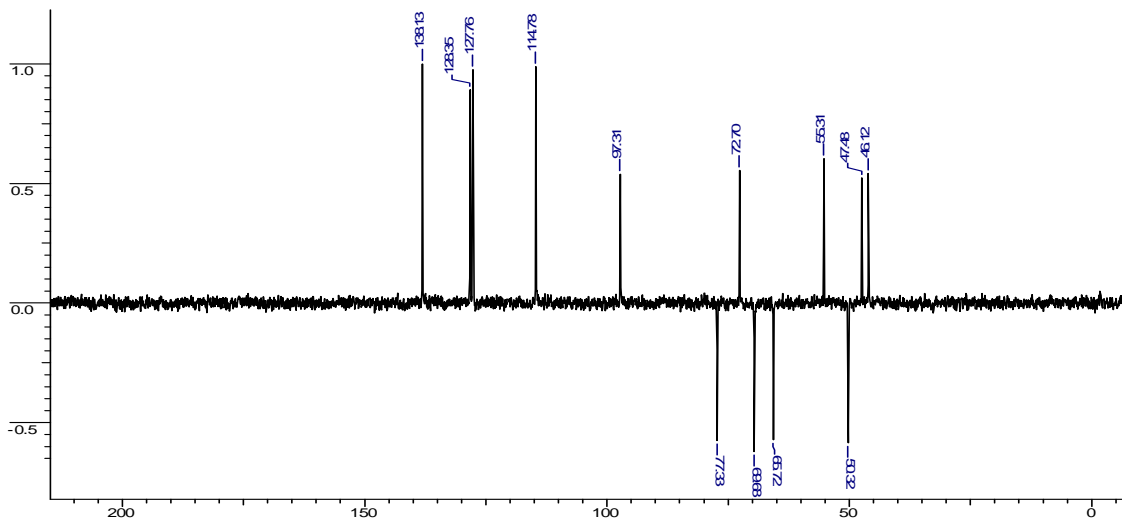
^1H NMR (200 MHz, CDCl_3) of Compound **33f**



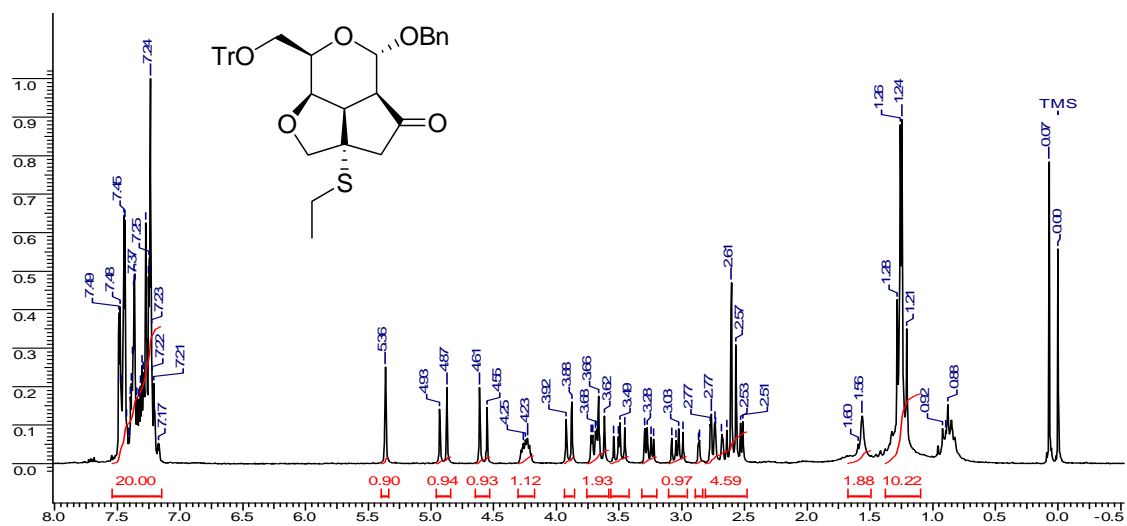
^{13}C NMR (50 MHz, CDCl_3) of Compound **33f**



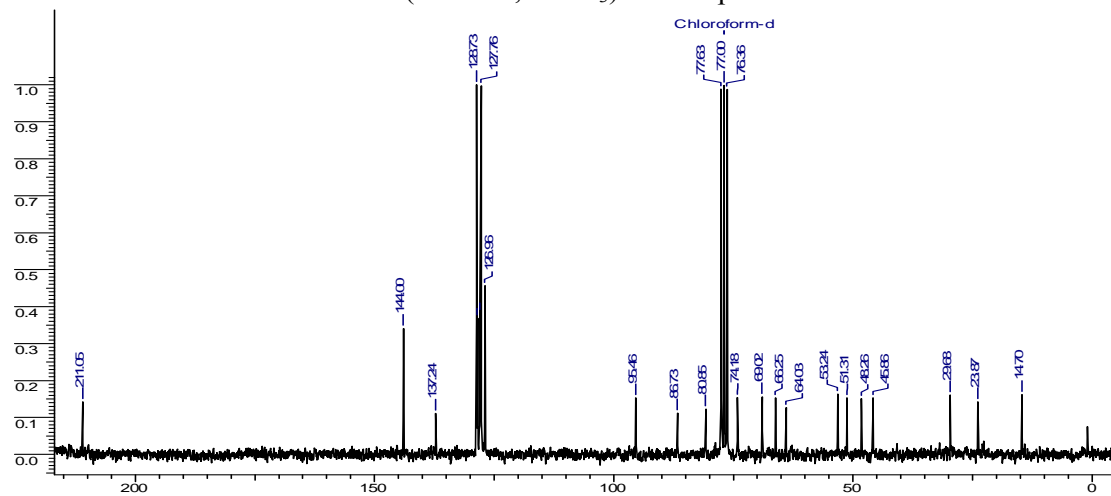
DEPT NMR (50 MHz, CDCl_3) of Compound **33f**



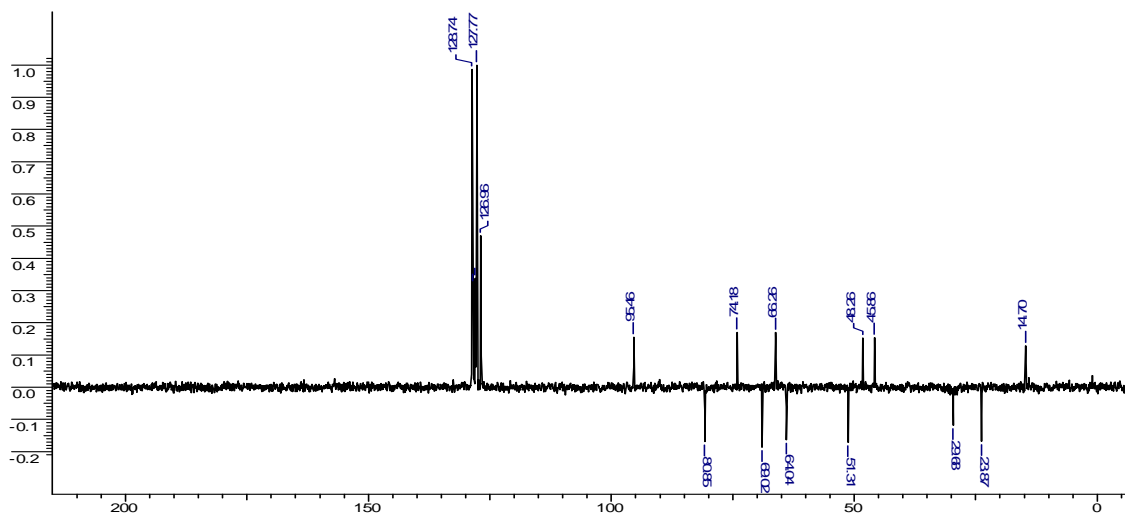
¹H NMR (200 MHz, CDCl₃) of Compound **34a**



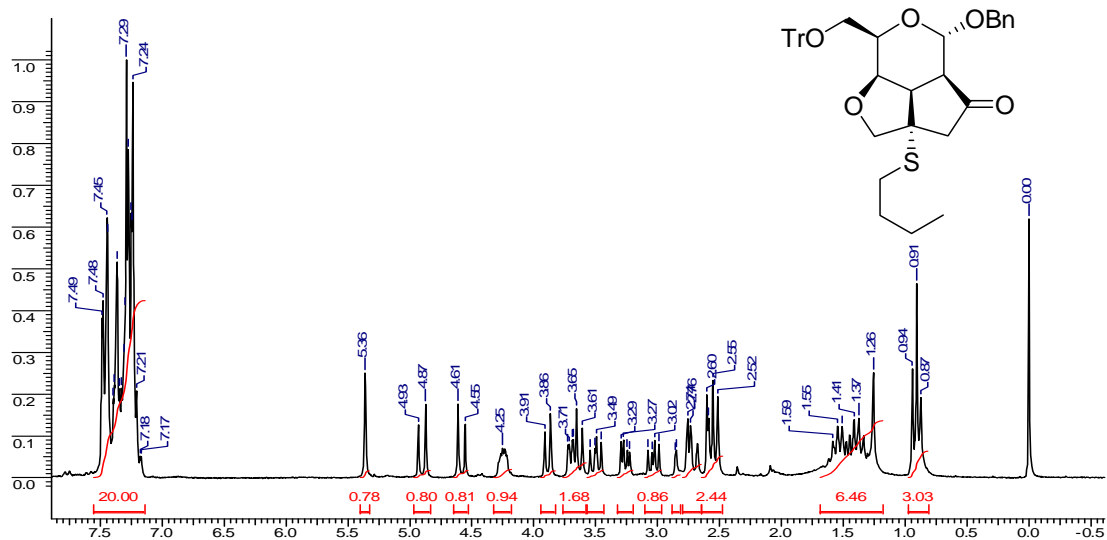
¹³C NMR (50 MHz, CDCl₃) of Compound **34a**



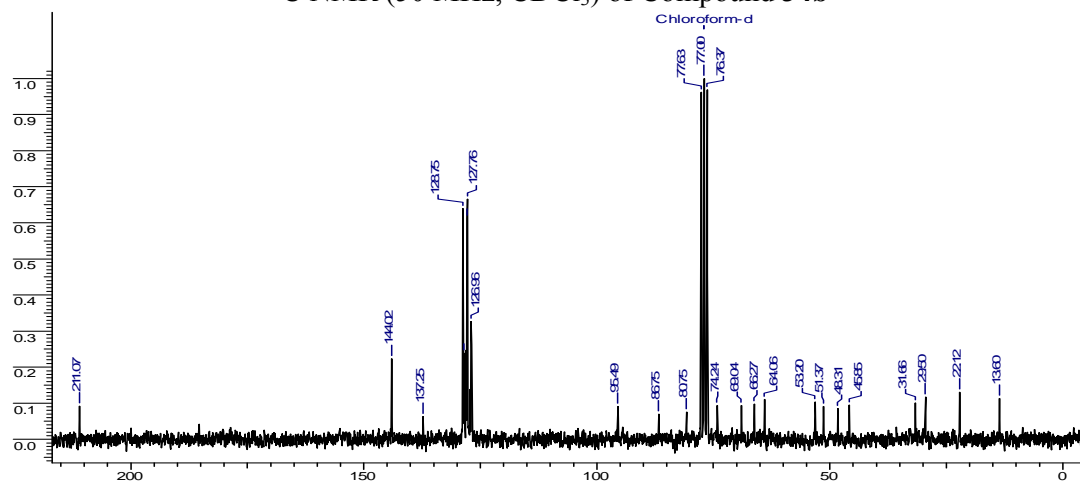
DEPT NMR (50 MHz, CDCl₃) of Compound **34a**



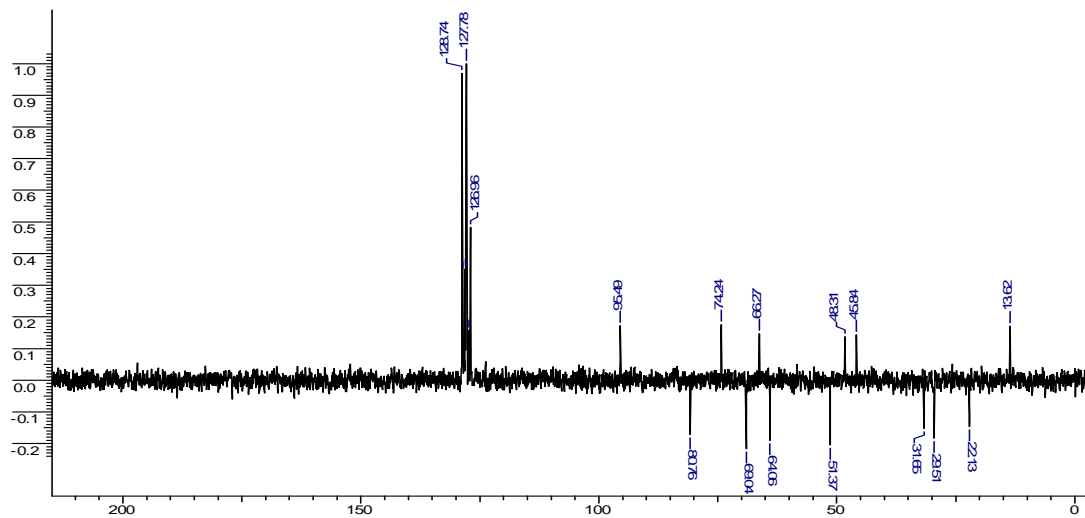
¹H NMR (200 MHz, CDCl₃) of Compound **34b**



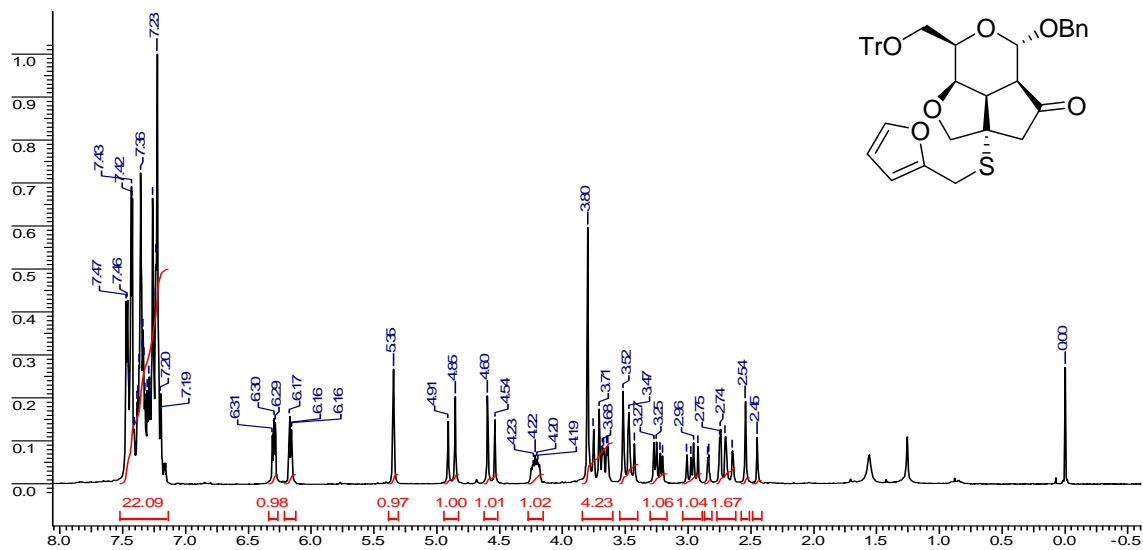
¹³C NMR (50 MHz, CDCl₃) of Compound **34b**



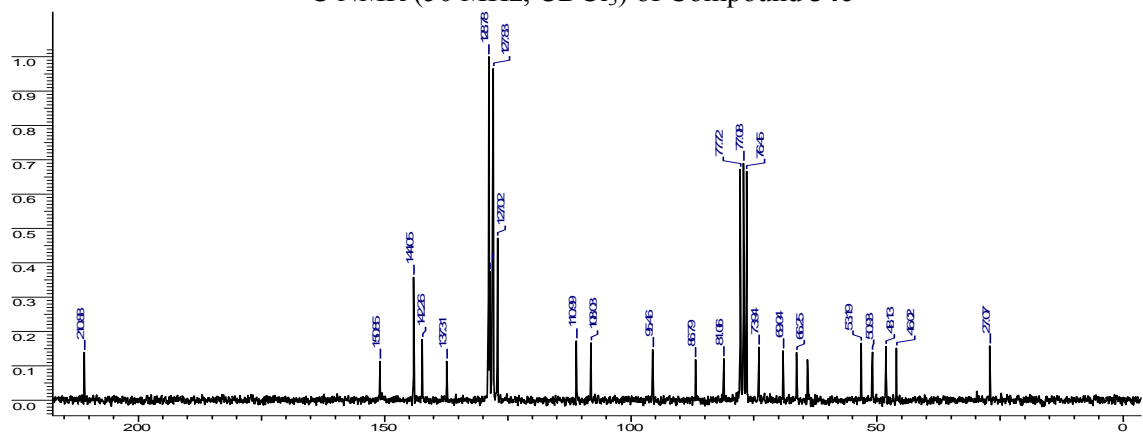
DEPT NMR (50 MHz, CDCl₃) of Compound **34b**



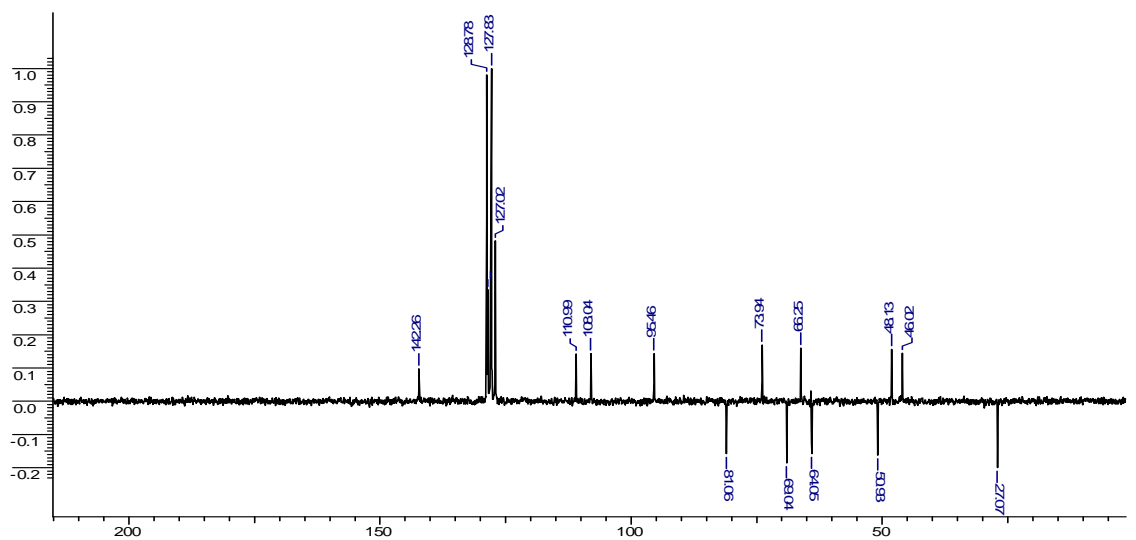
^1H NMR (200 MHz, CDCl_3) of Compound **34c**



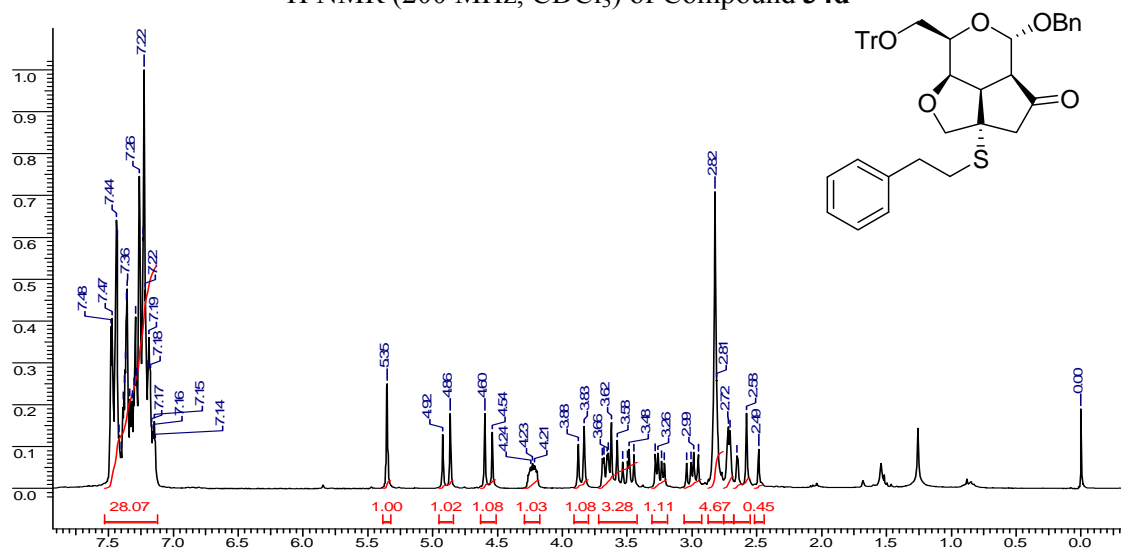
^{13}C NMR (50 MHz, CDCl_3) of Compound **34c**



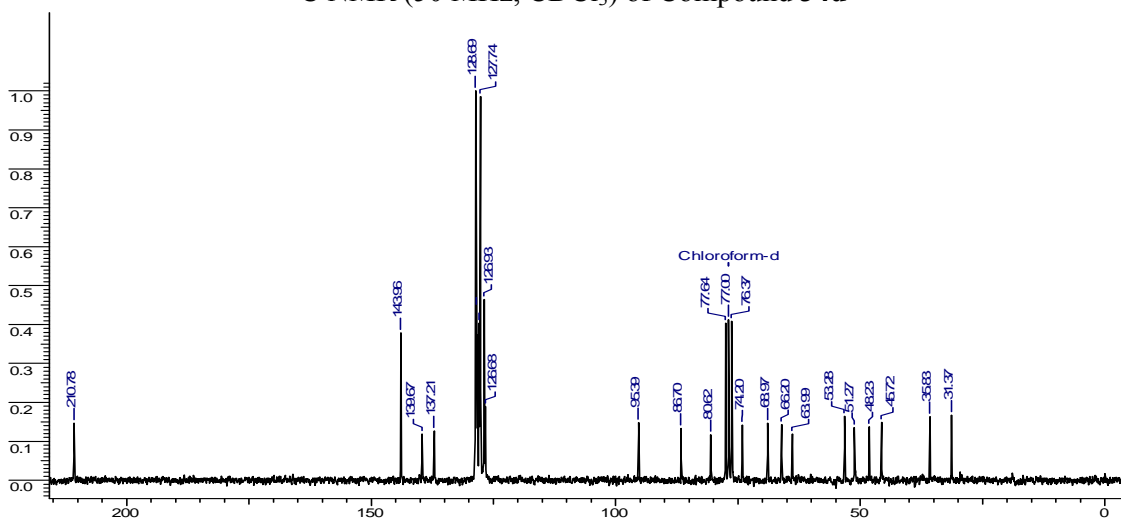
DEPT NMR (50 MHz, CDCl_3) of Compound **34c**



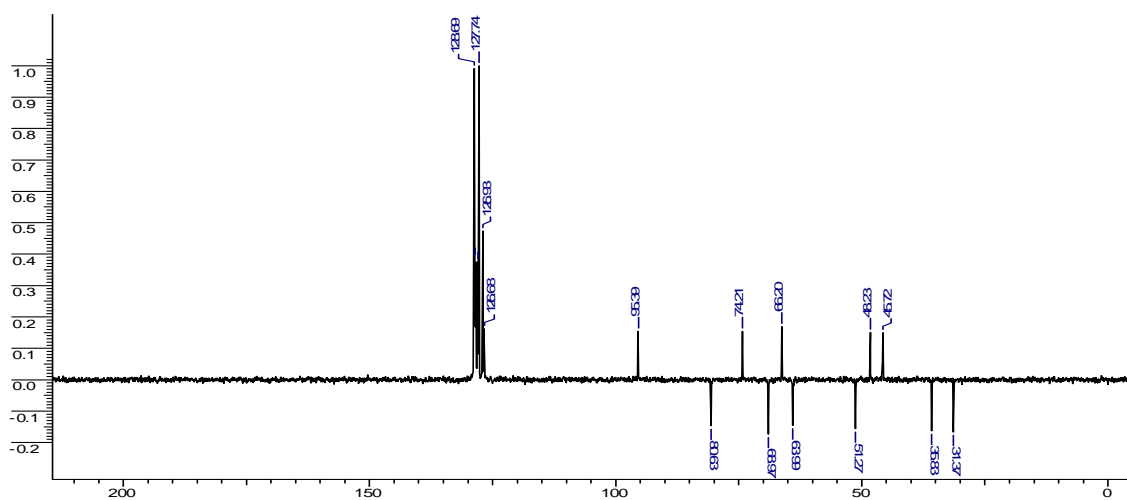
¹H NMR (200 MHz, CDCl₃) of Compound **34d**



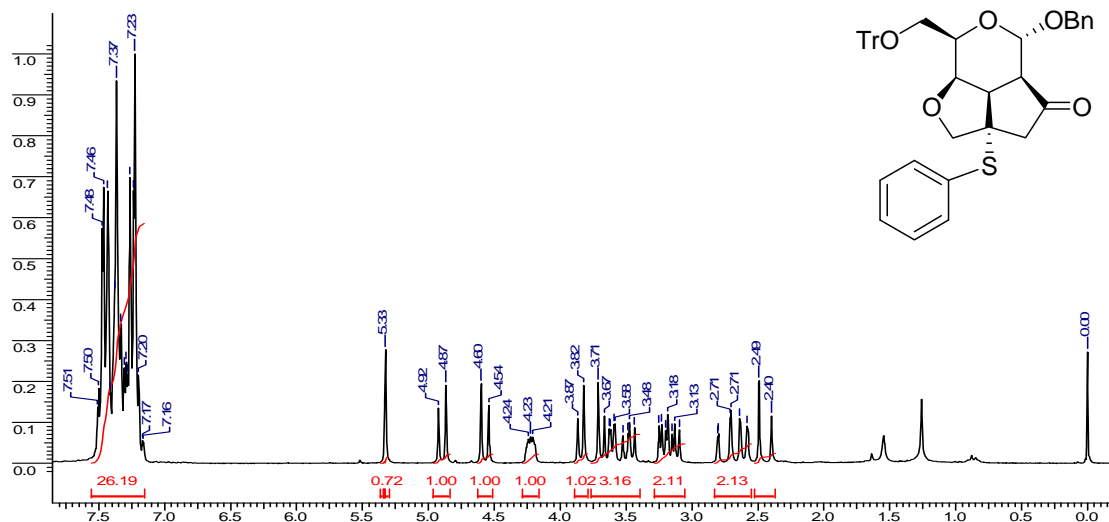
¹³C NMR (50 MHz, CDCl₃) of Compound **34d**



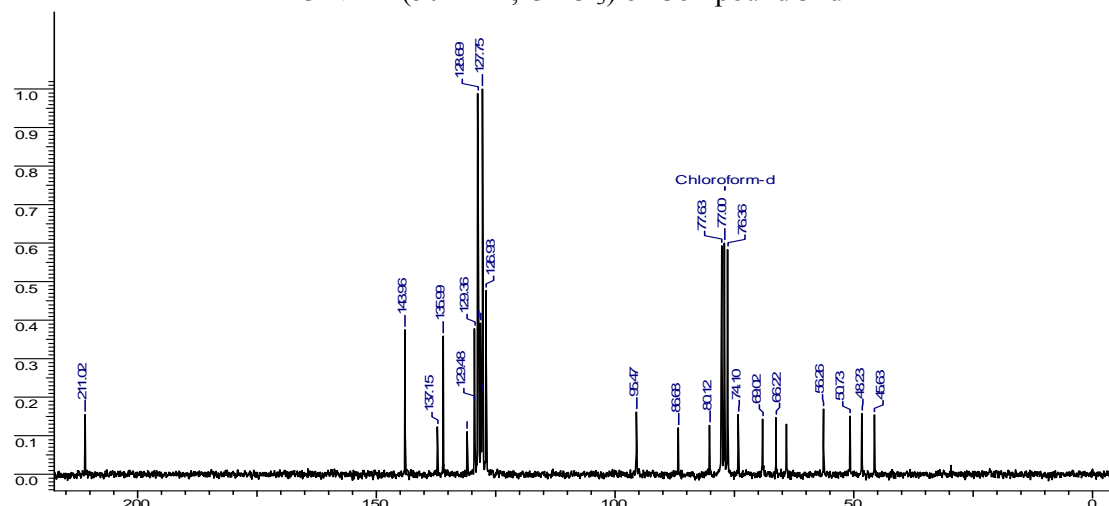
DEPT NMR (50 MHz, CDCl₃) of Compound **34d**



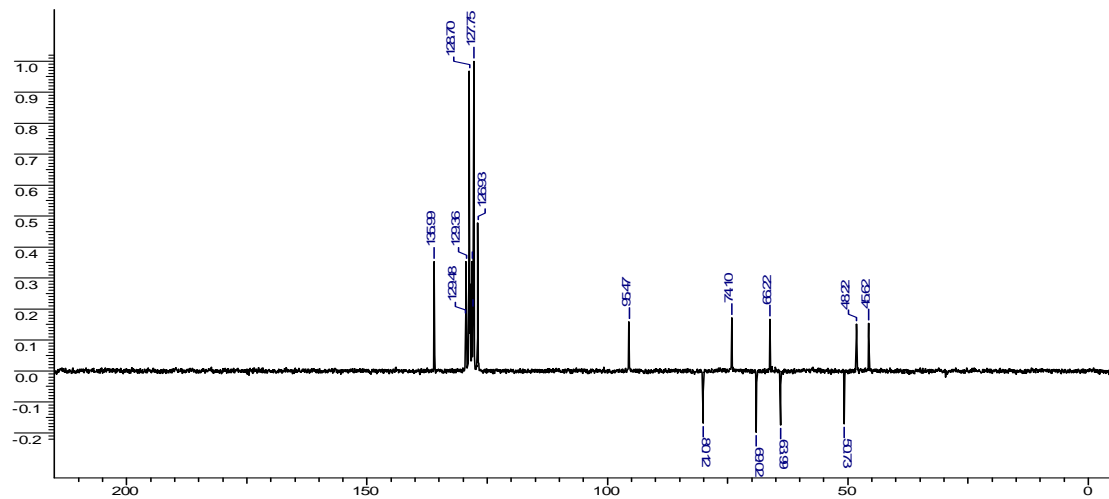
^1H NMR (200 MHz, CDCl_3) of Compound **34e**



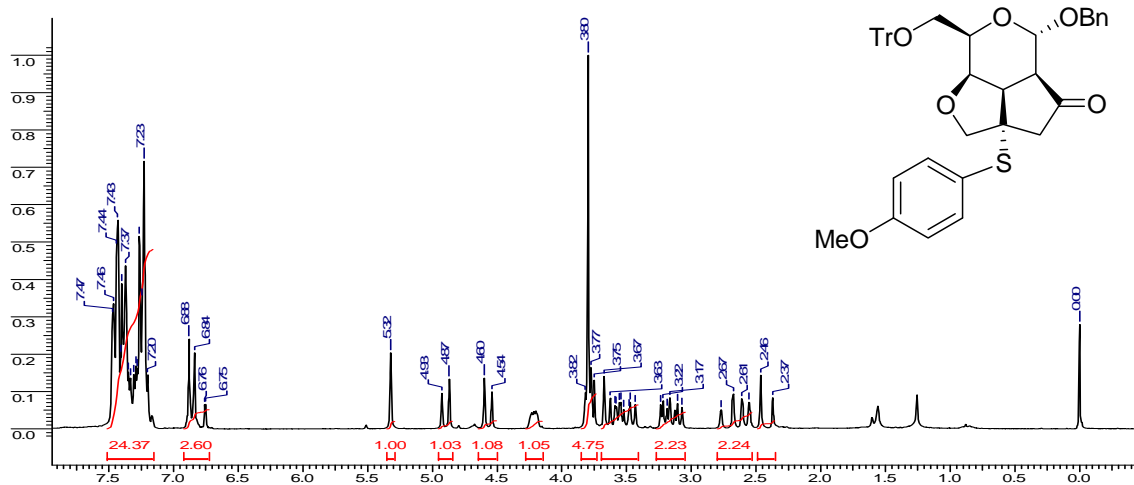
^{13}C NMR (50 MHz, CDCl_3) of Compound **34d**



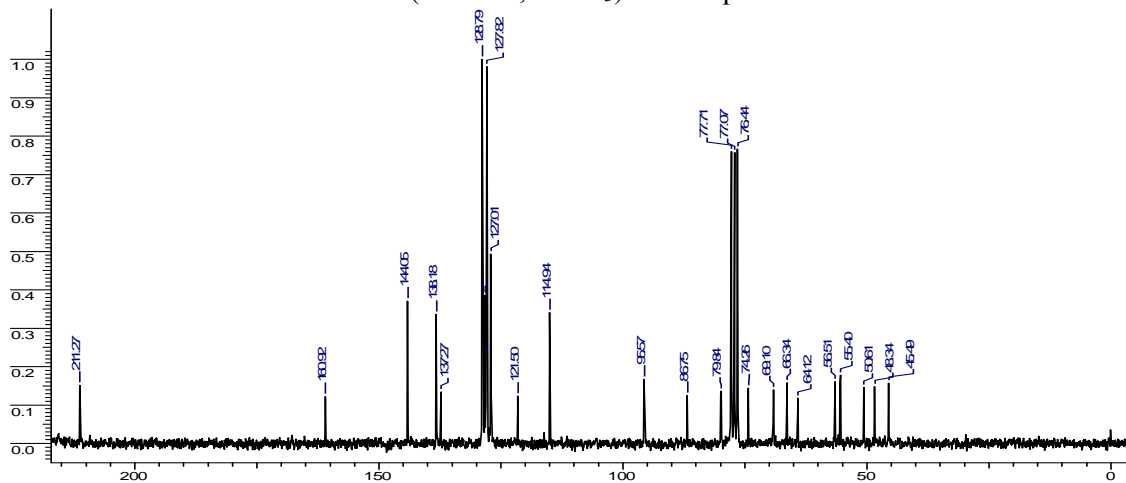
DEPT NMR (50 MHz, CDCl_3) of Compound **34d**



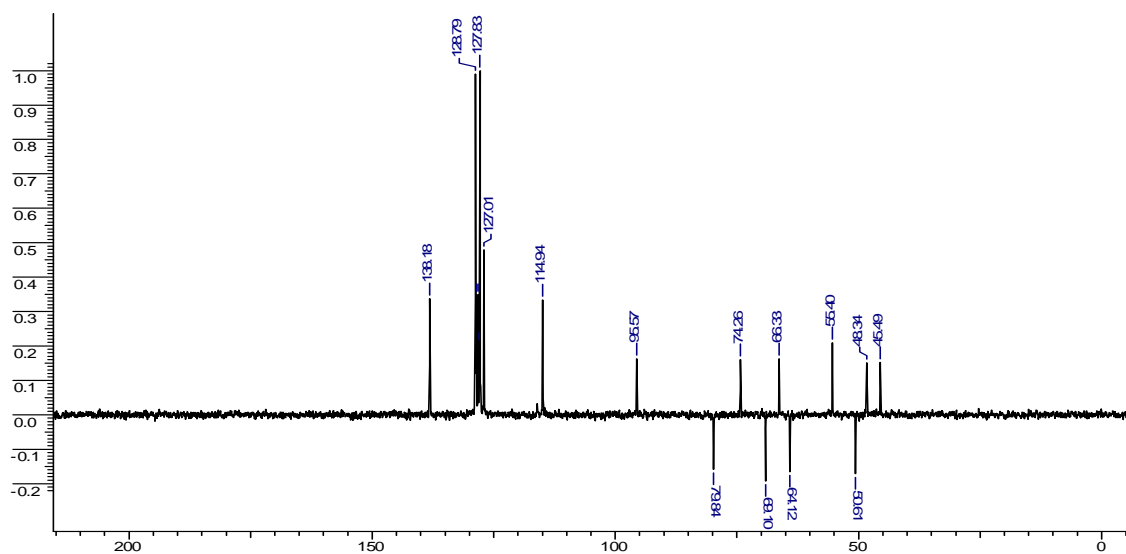
^1H NMR (200 MHz, CDCl_3) of Compound **34e**



^{13}C NMR (50 MHz, CDCl_3) of Compound **34e**



DEPT NMR (50 MHz, CDCl_3) of Compound **34e**



Chapter 1: References

1. (a) Schreiber, S. L. *Chem. Eng. News* **2003**, *81*, 51-61; (b) Spring, D. R. *Chem. Soc. Rev.* **2005**, *34*, 472-482.
2. Lee, Y. K.; Puong, K. Y. *Br. J. Nutr. Med.* **2002**, *88*, S101-S108.
3. Bohacek, R. S.; McMartin, C.; Guida, W. C. *Med. Res. Rev.* **1996**, *16*, 3-50.
4. Clardy, J.; Walsh, C. *Nature* **2004**, *432*, 829-837.
5. (a) Nichlaou, K. C.; Vourloumis, D.; Winssinger, N.; Baran, P. S. *Angew. Chem. Int. Ed.* **2000**, *39(1)*, 44-122; (b) Nicolaou, K. C.; Sorenson, E. J. *Classics in Total Synthesis* **1996**; (c) Nicolaou, K. C.; Snyder, S. A. *Classics in Total Synthesis II* **2003**.
6. Corey, E. J. *Chem. Soc. Rev.* **1988**, *17*, 111-133.
7. (a) Thomas, G. L.; Wyatt, E. E.; Spring, D. R. *Curr. Opin. Drug Disc. & Develop.* **2006**, *9(6)*, 700-712; (b) Kirkpatrick, P.; Ellis, C. *Nature* **2004**, *432*, 823-865; (c) Shang, S.; Tan, D. S. *Curr. Opin. Chem. Biol.* **2005**, *9*, 1-11.
8. Fink, T.; Reymond, J. L. *J. Chem. Inf. Model* **2007**, *47(2)*, 342.
9. Lipinski, C. A. et. al. *Adv. Drug Del. Rev.* **2001**, *46*, 3-26.
10. Breinbauer, R.; Vetter, I. R.; Waldmann, H. *Angew. Chem., Int. Ed.* **2002**, *41*, 2878-2890.
11. Schreiber, S. L. *Science* **2000**, *287*, 1964-1969.
12. (a) Schreiber, S. L. *Science* **2000**, *287*, 1964-1969; (b) Kwon, O.; Park, S. B.; Schreiber, S. L. *J. Am. Chem. Soc.* **2002**, *124*, 13402-13404; (c) Wu, C.-Y.; Chang, C.-F.; Chen, J. S.-y.; Wong, C.-H.; Lin, C.-H. *Angew. Chem. Int. Ed.* **2003**, *42*, 4661-4664. (d) Micalizio, G. C.; Schreiber, S. L. *Angew. Chem., Int. Ed.* **2002**, *41*, 3272-3276; (e) Tan, D. S.; M. A.; Shair, M. D.; Schreiber, S. L. *J. Am. Chem. Soc.* **1998**, *120*, 8565-8566; (f) Burke, M. D.; Schreiber, S. L. *Angew. Chem., Int. Ed.* **2004**, *43*, 46-58.
13. Neumann, C. S.; Nagayama, S.; Peristein, E. O.; Schreiber, S. L. *J. Am. Chem. Soc.* **2004**, *126*, 16077-16086.
14. Pelish, H. E.; Westwood, N. J.; Shair, M. D. *J. Am. Chem. Soc.* **2001**, *123*, 6740-6741.
15. Nicolaou, K. C.; Pfefferkorn, J. A.; Mithcell, H. J. *J. Am. Chem. Soc.* **2000**, *122*, 9968-9976.

16. Thomas, G. L.; Wyatt, E. E.; Spring, D. R. *Curr. Opin. Drug Disc. & Develop.* **2006**, *9*(6), 700-712.
17. Williams, N. R.; Wander, J. D. *The Carbohydrates in Chemistry and Biochemistry*; Academic Press: New York **1980**, 761.
18. Varki, A.; Cummings, R.; Marth, J. *Essentials of Glycobiology* **1999**.
19. (a) Ferrier, R. J.; Prasad, N. *J. Chem. Soc. (C)* **1969**, 570-574.
20. (a) Babu, B. S.; Balasubramanian, K. K. *Tetrahedron Lett.* **2000**, *41*, 1271-1274; (b) Shanmugasundaram, B.; Bose, A. K.; Balasubramanian, K. K. *Tetrahedron Lett.* **2002**, *43*, 6795-6798; (c) Grynkiewicz, G.; Priebe, W.; Zamojski, A. *Carbohydr. Res.* **1979**, *68*, 33-41; (d) Swami, N. R.; Venkateswarlu, A. *Synthesis* **2002**, 598-600; (e) Masson, C.; Soto, J.; Bessodes, M. *Synlett* **2000**, 1281-1282; (f) Yadav, J. S.; Reddy, B. V. S.; Murthy, C. V. S. R.; Kumar, G. M. *Synlett* **2000**, 1450-1451; (g) Bettadaih, B. K.; Srinivas, P. *Tetrahedron Lett.* **2003**, *44*, 7257-7259; (h) Babu, B. S.; Balasubramanian, K. K. *Synth. Commun.* **1998**, *29*, 4299-4305; (i) Yadav, J. S.; Reddy, B. V. S.; Reddy, J. S. S. *J. Chem. Soc., Perkin Trans. 1* **2002**, 2390-2394; (j) Smitha, G.; Reddy, C. S. *Synthesis* **2004**, 834-836.
21. (a) Toshima, K.; Ishizuka, T.; Matsuo, G.; Nakata, M.; Konoshita, M. *J. Chem. Soc., Chem. Commun.* **1993**, 704-706; (b) Fraser-Reid, B.; Madsen, R. *J. Org. Chem.* **1995**, *60*, 3851-3858; (c) Koreeda, M.; Houston, T. A.; Shull, B. K.; Klemke, E.; Tuinman, R. J. *Synlett* **1995**, 90-92; (d) Lo'pez, J. C.; Fraser-Reid, B. *J. Chem. Soc., Chem. Commun.* **1992**, 94-95; (e) Yadav, J. S.; Reddy, B. V. S.; Pandey, S. K. *New J. Chem.* **2001**, *25*, 538-540; (f) Agarwal, A.; Rani, S.; Vankar, Y. D. *J. Org. Chem.* **2004**, *69*, 6137-6140; (g) Suzuki, K.; Hashimoto, T.; Maeta, H.; Matsumoto, T. *Synlett* **1992**, 125-128.
22. Hotha, S.; Tripathi, A. *Tetrahedron Lett.* **2005**, *46*, 4555-4558.
23. Sinou, D. *Hetrocycles* **2006**, *68*, 2607-2613.
24. Khand, I. U.; Knox, G. R.; Pauson, P. L. *J. Chem. Soc. Perkin Trans* **1973**, *1*, 977.
25. (a) Gibson, S. E.; Stevenazzi, A. *Angew. Chem. Int. Ed.* **2003**, *42*, 1800-1810. (b) Krafft, M. E.; Bonaga, L. V. R. *Tetrahedron* **2004**, *60*, 9795-9833; (c) Jeong, N.; Chung, Y. K.; Yoo, S. E. *Synlett* **1991**, 204.
26. Kubota, H.; Lim, J.; Depew, K. M.; Schreiber, S. L. *Chem. & Biol.* **2002**, *9*, 265-276.
27. Hotha, S.; Tripathi, A. *J. Comb. Chem.* **2005**, *7*, 968-976.

Chapter 2
'Click' Chemistry Inspired Imaging of Microorganisms

Chapter 2: Introduction

Carbohydrates have long been underappreciated by the scientific community, and many researchers approach the complex structures and elaborate nomenclature with apprehension. Over the past three decades, complex carbohydrates have been widely recognized as more than just an energy source and are involved in broad range of biological processes varying from protein folding to oligomerization and stability, to the immune response and host pathogen interactions.¹⁻³ Recently, glycoconjugates, carbohydrates covalently linked with other chemical species have been recognised as an important class of the compounds in biology and consist of many different categories such as glycoproteins, glycopeptides, peptidoglycans, glycolipids and lipopolysaccharides. Glycoconjugates also play important role in developmental processes, as revealed by the pathology of human diseases caused by abnormal glycosylation and genetic studies in model organisms.⁴ These are involved in cell-cell interactions, including cell-cell recognition and cell-matrix interactions.^{5, 6} Although fewer scientists work with carbohydrates than with other biopolymers, advances in this field have been prolific. Chemists and biochemists have developed new methods to rapidly synthesize oligosaccharides, enabling them to generate complex polysaccharides and analogues from natural products that have increased activity *in vivo*. Biologists have explored the physiological role of various sugars, discovering that many have essential roles in all of the major organ systems and are involved in the several diseases.



Figure 1. Diagrammatic representation of central dogma of molecular biology

The central paradigm of modern molecular biology is that biological information flows from DNA to RNA to protein as first enunciated by Francis Crick (Figure 1).⁷ This powerful concept lies not only in its template-driven precision, but also in the ability to manipulate any one class of molecules based in knowledge of another, and in the patterns of sequence homology and relatedness that predict function and reveal evolutionary relationships. With the completion of genomic sequences of human and several other commonly studied model organisms, even more spectacular gains in the understanding of biological systems are anticipated.

Infact, creating a cell wall requires two major classes of molecule: lipids and carbohydrates.^{8b, 8c} Carbohydrates are indispensable to life on Earth. These biomolecules in their simplest life forms serve as a primary energy source for sustaining life. For the most part, however, carbohydrates exist not as simple sugars but as complex molecular conjugates, or *glycans*. Glycans come in many shapes and sizes, from linear chains (polysaccharides) to highly branched molecules bristling with antennae-like arms. And although proteins and nucleic acids such as DNA have traditionally attracted far more scientific attention, glycans are also key to life. They are ubiquitous in nature, forming the intricate sugar coat that surrounds the cells of virtually every organism and occupying the spaces between these cells.

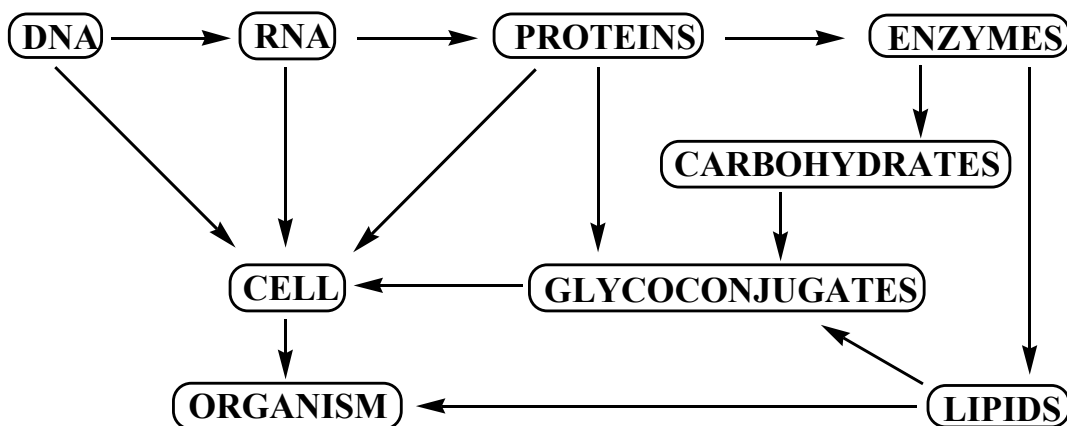


Figure 2. *Extended paradigm of the molecular biology*

The structural role of carbohydrates becomes particularly important in constructing complex multi-cellular organs and organisms, which requires interactions of cell with one another and with the surrounding matrix.⁸ Indeed, all cells and many macromolecules in nature carry a dense and complex array of covalently attached sugar chains (called oligosaccharides or glycans). Since most glycans are in the outer surface of cellular and secreted macromolecules, they are in a position to modulate or mediate a wide variety of events in cell-cell and cell-matrix interactions crucial to the development and function of a complex multicellular organism. They are also in a position to mediate interactions between organism (e.g. between host and parasite). In addition, simple, highly dynamic protein-bound glycans are abundant in the nucleus and cytoplasm, where they appear to serve as regulatory switches (Figure 2).

During the last century, the chemistry, biochemistry and biology of carbohydrates were prominent matters of interest. However, during the initial phase of the modern revolution in molecular biology, studies of glycans lagged far behind those major class of molecules. This was in large part due to their inherent structural complexity, the difficulty in easily determining

their sequence, and the fact that their biosynthesis can not be easily predicted from the DNA template. The development of a variety of new technologies for exploring the structures of these glycan chains has opened a new frontier of molecular biology which has been called glycobiology. This word was first coined by Rademacher, Parekh and Dwek to recognize the coming together of the traditional disciplines of carbohydrate chemistry and biochemistry with modern understanding of the cellular and molecular biology of glycans.⁹ Emerging from its roots in classical carbohydrate chemistry and biochemistry, glycobiology has become a vibrant, expanding and important extension of molecular biology. Glycobiology combines the expertise of synthetic, analytical and carbohydrate biochemistry, as well as molecular and cellular biology, to unravel the structural complexity, chemistry, biosynthesis, and biological functions of the sugar-bearing biomolecules.

As the surface of almost every living cell is decorated with the a layer of complex carbohydrates, many of the advances in complex carbohydrate synthesis revolve around methods to form the glycosidic bond, because this is the primary means by which monosaccharide building blocks are assembled into more complex oligosaccharide structures. The complexity of the glycosylation pathways that manufacture and constantly remodel surface sugars is truly astounding. Glycosylation can have a profound influence on the function of a variety of eukaryotic cells. It can be described as a complex process that requires a major commitment of cellular resources; it employs an estimated 2–3% of the genes in humans and many high-energy intermediates.¹⁰ The first step in glycosylation is the import of dietary sugars, such as glucose, into a cell. This is followed by a series of phosphorylation, epimerization, and acetylation reactions that diversify these sugars and convert them into high-energy nucleotide sugar donors. These compounds serve as the “building blocks” for the assembly of complex carbohydrates in the endoplasmic reticulum and Golgi apparatus. The newly synthesized carbohydrates, which are generally attached to proteins or lipids, are then transported to the cell surface where they contribute to the interaction between a cell and its environment by playing important structural and signalling roles. In particular, it can affect signal transduction and cell–cell communication properties and thus shape critical cell decisions, including the regulation of differentiation and apoptosis. Regulation of glycosylation has multiple layers of complexity, both structural and functional, which make its experimental and theoretical analysis difficult to perform and interpret.¹¹

Carbohydrate chemists and biochemists tend to be controlling personalities and aren't satisfied with the way sugar groups are attached naturally to glycosylated proteins and natural products. Around the world researchers are occupied most days devising new ways to tame

fractious glycosylated biomolecules into shape. Glycosylation of protein or a natural product can be achieved by the bioengineering of bacteria, yeast, plant cells, insect cells, and mammalian cells. Apart from this, it can also be achieved by manipulating protein or natural product glycosylation pathways in genetically engineered organisms. Compounds also can be glycosylated by exploiting the liberality with which sugar-attaching enzymes (glycosyl transferases) tolerate a wide range of different substrates. And one can glycosylate proteins and natural products through a variety of chemical approaches: with enzymes, protein ligation, metabolic bioengineering, or total synthesis.

More than 50% of human proteins are glycoproteins, which tend to be heterogeneously glycosylated.¹² That means the structures of attached sugars and the sites to which they're attached are variable. A wide range of bioactive natural products are glycosylated as well, including anthracycline antitumor antibiotics, avermectin antiparasitics, enediyne antibiotics, macrolides such as erythromycin, and glycopeptide natural products like vancomycin. Some natural products are also heterogeneously glycosylated, although generally to a lesser degree than glycoproteins. Whether glycoproteins and glycosylated natural products are heterogeneously glycosylated or not, researchers have felt the need to better control which sugars go where on these molecules. The idea is to improve on Mother Nature because sugars are often essential to the functions of biomolecules, and the ability to modify and control which ones are attached and exactly how they're attached can make it possible to tailor biomolecular structure, folding, stability, immunogenicity, uptake, distribution, target recognition, and other properties and functions of these compounds. For example, glycosylated biomolecules can be turned into better drugs by manipulating their sugars. Careful glycosylation of erythropoietin (EPO), an approved glycoprotein drug for anemia and cancer, is critical for its activity and longevity of action, and its properties have been improved by modifying its glycosylation through genetic engineering. Novel research methodologies provided by glycobiology can help to address many outstanding issues and integrate glycosylation with other metabolic and cell regulation processes. Furthermore, glycosylation often lies at the interface between metabolism and cell signalling, thereby making analysis even more challenging and interesting. Recently, researchers around the world are actively pursuing the cell surface remodelling exploiting glycosylation pathways through chemical means. As the cell surface govern many of the vital biological events, perturbation of any of these metabolic processes will be guided by the glycosylation pathways.

One of the recent advances in probing the glycosylation pathways was pioneered by C. Bertozzi and her group when they synthesized the unnatural sialic acids as tool for the

penetrating metabolic process of the cells.^{13a} Sialic acid is a generic term for the *N*- or *O*-substituted derivatives of neuraminic acid, a nine-carbon monosaccharide. It is also the name for the most common member of this group, *N*-acetylneuraminic acid. The amino group bears either an acetyl or a glycolyl group. The hydroxyl substituents may vary considerably: acetyl, lactyl, methyl, sulphate, and phosphate groups have been found. Sialic acids are the most abundantly distributed terminal components of oligosaccharides on mammalian glycoproteins and glycolipids. Sialic acids are biosynthesized from the six carbon precursor *N*-acetylmannosamine.^{13b-13e} Bertozzi group exploited the tolerance of enzymes, which take part in sialic acid biosynthesis, to unnatural mannosamine derivatives; ManNAc was substituted with *N*-levulinoyl having the ketone functionality at the position normally occupied by the *N*-acetyl group in the natural substrate. The ketone group on the surface was useful when covalently ligated with complementary functional group hydrazide (Figure 3).

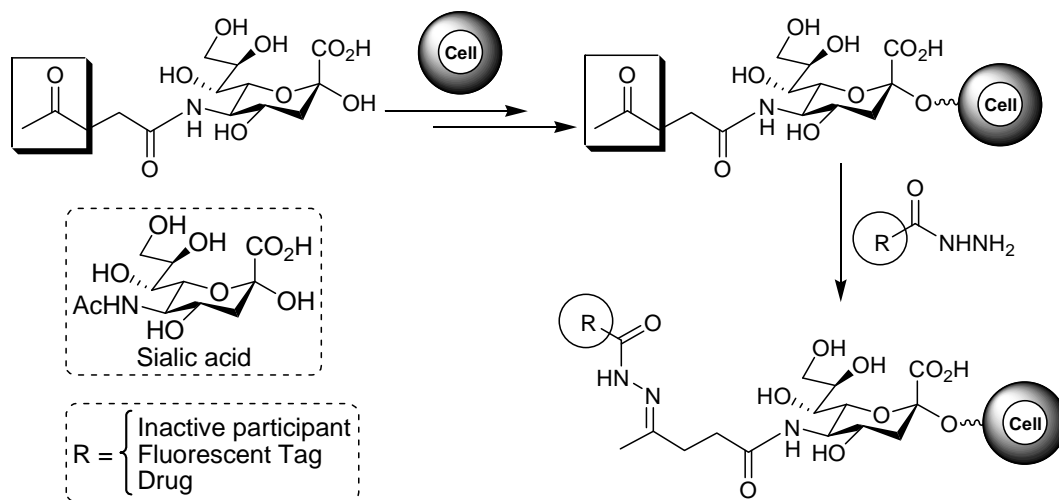


Figure 3. Biosynthetic incorporation of ketone groups of unnatural mannosamine derivatives onto cell surface

They selected three human cell lines for their studies- Jurkat (T cell derived), HL-60 (neutrophil derived), and HeLa (cervical epithelial carcinoma). After performing a series of experiments, they were able to successfully demonstrate the incorporation of ManLev on the cell surface exploiting the glycosylation pathway of the human cell lines. The presence of the ketone group on the cell surface was determined by the chemoselective ligation of a hydrazide based probe, biotinamidocaproyl hydrazide. The cells were then analysed by flow cytometry employing fluorescein isothiocyanate which shows a strong binding affinity for biotin. Researchers were also successful in their attempt to demonstrate the inhibition of the ketone group display on the cell surface *via* the use of Tunicamycin, known to inhibit the *N*-linked protein glycosylation.

Taking advantage of the research work carried out by Bertozzi group, K. J. Yarema also exploited the sialic acid metabolism pathway to decorate the cell surfaces with thiols.¹⁴ They synthesized thiol substituted unnatural mannosamine derivative, *N*-thioglyconueraminic acid, and incorporated them on Jurkat (human T lymphoma) cells exploiting the permissivity of sialic acid biosynthetic pathway for non-natural metabolic intermediates (Figure 4). Cell surface thiols were analysed by the labelling of the cell with (+)-biotinyl-3-maleimidopropionamidyl-3,6-dioxaoctanediamine followed by fluorescein-conjugated avidin staining and flow cytometry.

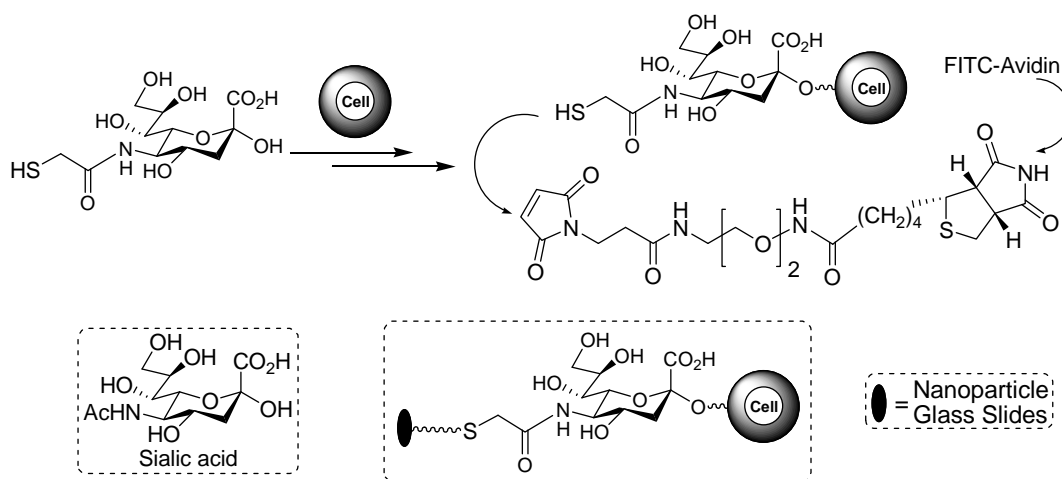


Figure 4. Cell surface display of thiol-bearing sialic acids

Yarema group was also successful in their attempts to covalently attach the mammalian cell lines onto the glass slides opening an all new frontier for the tissue engineering by creating custom designed binding interfaces. The ability of the unnatural thiol mannosamine derivative to alter the adhesive properties of glycocalyx enabled them to demonstrate the concept of cell adhesion on the glass slides coated with bovine serum albumin (BSA) functionalised with maleimide. Michael addition of cell surface thiols to the maleimide moiety of the BSA facilitated the adherence of the cell lines onto the glass slides. Further taking advantage of the affinity of the thiols for the gold nanoparticles, researchers also accomplished the adhesion of nanoparticles to the cell surface thiols. In the longer term, where the linkage of the cells to the surface of biomaterials, synthetic scaffolds and matrix materials remains a formidable challenge, the effort in this regard will go in a long for the advancement of interfacial science.

Researchers have also targeted numerous other glycosylation pathways other than related to sialic acids, fucosylation pathways (C. Wong group) being one of them. Fucose is a hexose deoxy sugar, generally found on *N*-linked glycans on the mammalian, insects and plant cell surfaces. Two structural features that distinguish fucose from other six-carbon sugars

found abundantly in mammals is the lack of hydroxyl functionality on C-6 and the L-configuration. It is most often the terminal sugar in glycans that participate in important biological events that are involved in various physiological and pathological processes of the body. Wong *et.al.* demonstrated the visualization of the incorporation of an unnatural fucose analog into glycoproteins *via* the fucose salvage pathway (Figure 5).¹⁵

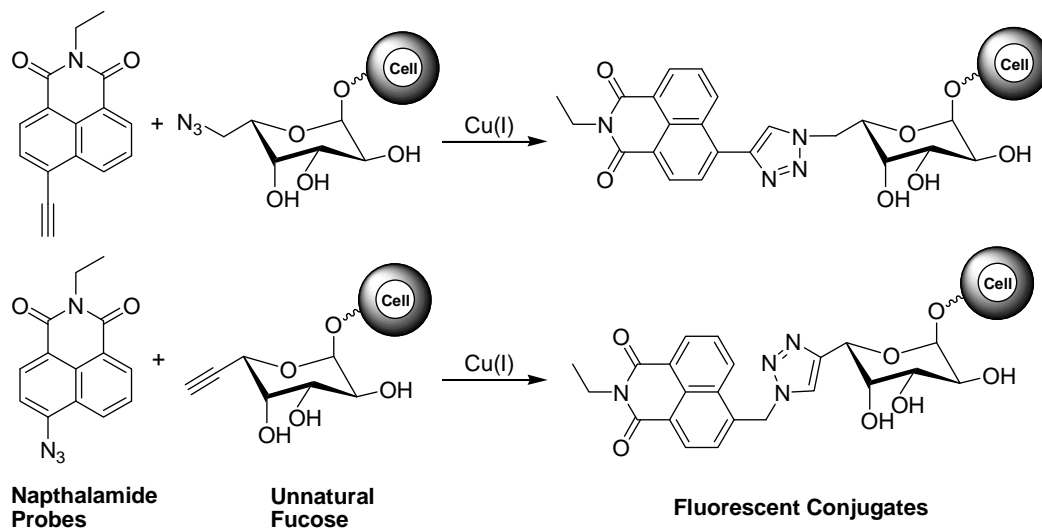


Figure 5. Acetylene and azido probes for fucosylation pathways

The unnatural fucose analogs were designed so as to exploit the tolerance of the fucosylating enzymes to a range of substituents at position 6 on fucose. Likewise, they synthesized the azido/alkynyl functionalised fucose and allowed them to enter the cells. The modified fucose derivatives were supplied to the cells in acetylated form facilitating incorporation into glycoproteins through the salvage biosynthetic pathway and monitored them by employing the fluorescent probes with the application of the click chemistry. The use of a click activated probe is highly practical in biological systems as a fluorescent signal is obtained only after a highly selective, bioorthogonal ligation takes place. The researchers were successful in their attempt to probe the fucosylating pathway using fluorogenic probes activated by click reaction allowing them to permit imaging of fucosylated glycoconjugates at cell surface and inside the cells.

Apart from the mammalian cells, scientists are also engaged in discovering novel probes for other microorganisms, like bacteria etc., in an attempt to display the cell wall with unnatural substrates with a promise to provide novel vaccines against infectious diseases caused by bacteria. In this regard, Nishimura *et.al.* reported a versatile chemical approach for displaying the target compounds on the bacterial cell wall.¹⁶⁻¹⁸ Bacterial cell wall differs from that of all the organisms by the presence of peptidoglycans (poly-*N*-acetylglucosamine and *N*-

acetylmuramic acid), which is located immediately outside the cytoplasmic membrane. The bacterial cell wall consists of repeating units of *N*-acetylglucosamine and *N*-acetylmuramic acid linked with a pentapeptide.¹⁹

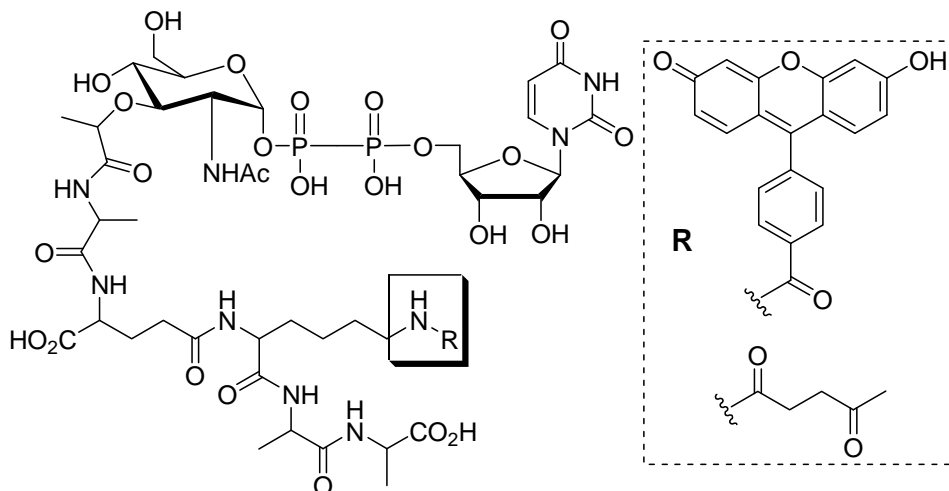


Figure 6. Unnatural UDP-MurNAc pentapeptide synthesized by Nishimura for bacterial cell wall engineering

Nishimura group cleverly designed the unnatural probe for the *Lactobacillus* bacterial cell wall comprising of a pentapeptide attached to the *N*-acetylglucosamine along with Uracil diphosphate moiety. The pentapeptide has the handle along which a fluorescent or a ketone probe can be tethered which can later be tracked down by a hydrazide based fluorescent probe. After examining a number of combinations of peptide, they arrived to a conclusion that presence of the pentapeptide was vital for the activity of the unnatural disaccharide. After a series of experiments, they were successful in demonstrating modification of the bacterial cell wall through the artificial precursors which were displayed across the membrane.

As the global scientific efforts in the field of glycobiology have grown, the success of these endeavours has prompted parallel efforts towards the next layer of biological complexity, namely post-translational modifications, with an emphasis on glycosylation. The complex glycosylation pathways of a cell do not exist in isolation; instead they are intimately intertwined with other critical metabolic and regulatory networks of the cells, illustrate the level of complexity involved to unravel the mechanistic details of glycosylation pathways.²⁰ However, the tremendous benefits which are seemingly guaranteed have stimulated an explosion of interest towards glycobiology and have been successful in luring lot of researchers to put up a spirited effort to come with the challenge.

Chapter 2: Present Work

Strep throat, cholera, pneumonia, whooping cough etc. these diseases, and more, are often the only things bacteria get credit for causing in human and animals. Bacteria are the oldest, the simplest and the most numerous forms of life on earth. Although we are never able to notice these tiny forms of life with the naked eye, these are present almost everywhere. These were first observed by Antonie van Leeuwenhoek in 1676, using a single-lens microscope of his own design.²¹⁻²³ Much later, the name *bacterium* was coined by Christian Gottfried Ehrenberg in 1828, which is derived from the Greek word meaning *small mass*. Afterwards, Robert Koch put forward his “*Germ Theory of Disease*”, for which he was awarded the Nobel Prize in 1908, laying down some benchmarks to test if any organism is the cause of a disease.

Bacteria are unicellular microorganisms, possessing a wide range of shapes from spheres to rods to spirals. These are prokaryotes and unlike eukaryotes do not contain a nucleus and rarely harbour membrane bound organelles. Bacterial cell wall is made of peptidoglycan, unlike the cell wall of plants and fungi which are made up of cellulose and chitin respectively (Figure 7).¹⁹ Broadly speaking, there are two different types of cell wall in bacteria, called Gram-positive and Gram-negative. The name originates from the reaction of cells to the Gram stain, a test long employed for the classification of bacterial species.²⁴ Gram-positive possesses a thick cell wall containing many layers of peptidoglycans and lipoteichoic acids external to cytoplasmic membrane. In contrast, Gram-negative have relatively thin cell wall consisting of a few layer of peptidoglycans surrounded by a second lipid membrane containing lipopolysaccharides and lipoproteins. Most bacteria have Gram-negative cell wall. This dissimilarity in structure is apparent in behavioral differences towards antibiotic susceptibility; for instance, vancomycin can kill only Gram-positive bacteria and is ineffective against Gram-negative pathogens. Peptidoglycan is a polymer of alternating *N*-acetylmuramic acid and *N*-acetylglucosamine.²⁵ Long strands of this alternating polymer may be linked by amino acids. Gram-positive cells have a much more highly cross-linked peptidoglycan structure than Gram-negative cells. Teichoic acid is a polymer of glycerol or ribitol *via* phosphodiester bonds. They can be either covalently bonded to *N*-acetylmuramic acid of the peptidoglycan layer, linked to the plasma membrane lipids found in cytoplasmic membrane or linked to the terminal D-alanine in the tetrapeptide cross-links between molecules of *N*-acetylmuramic acid. The combined unit s comprised of teichoic acid and lipids are referred to as lipoteichoic acid.

Teichoic acids are negatively charged and therefore contribute to the negative charge of the Gram-positive cell wall.

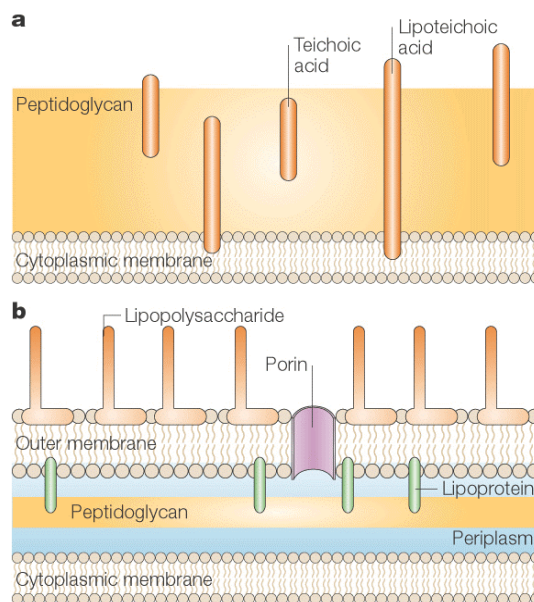


Figure 7. The overall cell wall structure of (a) Gram-positive bacteria; (b) Gram-negative bacteria

Understanding glycomics in general and surface exposed carbohydrates of many organisms in particular will play a vital role in enhancing our knowledge of human health ranging from congenital disorders to infectious diseases.²⁶ Advances in glycobiology have established the significance of oligosaccharides and glycoconjugates, often present on the cell surfaces.²⁷ Carbohydrate engineering by exogenously added monosaccharide supplements is a tool to probe various disease processes which would eventually facilitate discovery of highly specific small molecule inhibitors.²⁸ One of the major hindrances to the development of the field is the lack of proper glycan oriented imaging techniques.²⁹ Labeling *via* sialic acid and fucose metabolic pathway in mammalian cells by exogenously supplied monosaccharides, using a Staudinger ligation,^{28a} a ketone-aminoxyl/hydrazide ligation^{13, 28b} and a Michael addition¹⁴ have been extensively studied. However, this strategy will not be appropriate for those cells (e.g. many bacteria, yeast – mostly causing infectious diseases) which do not display appropriate sugars on their cell surfaces. Nishimura *et.al.* did study the cell surface display of fluorescein attached UDP-MurNAc which allows imaging of the bacterial cells.¹⁶ Nishimura's fluorescent conjugates are accessed through a multistep process and the total glycosyl character of the chemical probe was reduced to the minimum due to the presence of the appendages that would facilitate selective incorporation. As is evident from the Figure 8, that researchers have exploited various appendages on carbohydrate motifs for probing the glycosylation pathways utilizing fluorescence tags. But all the available methods have their

own drawbacks ranging from their specificity for only mammalian cells and bacterial cells or large background noise during real-time fluorescence imaging of cells. All this and more demands a universal approach applicable to a wide range of glycosylation pathways.

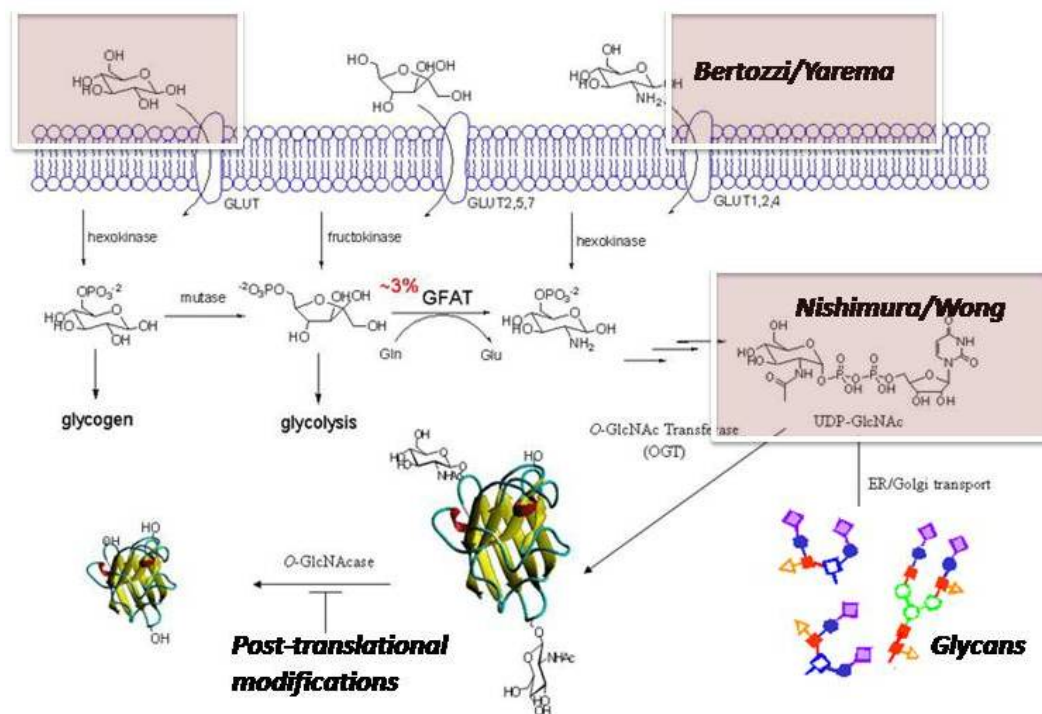


Figure 8. Techniques known so far for probing the glycosylation pathways of the cell

Fluorescent probes are one of the cornerstones of real-time imaging of live cells and a powerful tool for cell biologists. They provide high sensitivity and great versatility while minimally perturbing the cell under investigation. For an ideal imaging applications, the fluorescent label is a biologically inert participant that is used merely as a visible marker. The application of fluorescent probes will continue to expand and provide exciting new insights into the biology of living cells. However, most often employed fluorescent tags are bulky molecules with high molecular weight which no longer behave as an inert partner and more commonly the excess pre-labeled reagents are generally hard to remove from the experiments, which prohibit the application of multistep labeling procedures.³⁰ Recently, Wang *et. al.* reported 3-azidocoumarins as excellent profluorophores for the bioconjugation experiments (Figure 9).³¹ With the clever exploitation of the click conditions, the 3-azidocoumarins were made to react with alkynes to form triazoles which were fluorescent. Furthermore, both the precursors being non-fluorescent also rules out the possibility of the unbound reagent creating background fluorescence in our experiments. The mild condition under which the fluorophore

was generated non-fluorescent precursors provides an excellent platform to capitalize it for the real-time imaging of microorganisms.

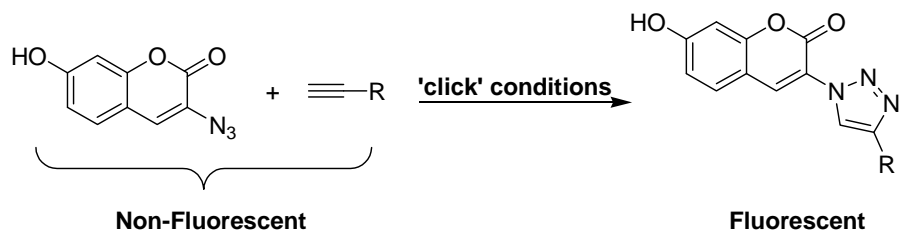


Figure 9. Schematic illustration of formation of fluorescent triazole

One of the significant procedures in probing glycosylation pathways involves modification of cell wall components chemically or enzymatically and thereby attaching a probe on the cell wall surface which can serve as a handle for fluorescence imaging or selective delivery of the drugs. Continuing in our efforts for the study and development of glycoconjugates, we hypothesized that an azidocoumarin based profluorophore would be immensely significant if it can be triggered to become fluorescent *via* Cu(I)-catalyzed 1,3 dipolar cycloaddition or 'click' reaction with an incorporated alkyne functionality by the carbohydrate engineering of bacterial cells. Thus we dwelt upon exploiting a simple appendage (a propargyl moiety) onto a monosaccharide which can be activated to become fluorescent *via* the conjugation with the coumarin azide.

Click Chemistry:

“Strategy for the rapid and efficient assembly of molecules with diverse functionality... enabled by a few nearly perfect reactions, it guarantees reliable synthesis of the desired product in high yield and purity...” – K. B. Sharpless.

Click chemistry is a chemical philosophy introduced by Sharpless in 2001 and describes chemistry tailored to generate substances quickly and reliably by joining small units together.³² This is inspired by the fact that nature also generates substances by joining small modular units. Click chemistry can be summarized neatly in one sentence: *all searches must be restricted to molecules that are easy to make.* A set of stringent criteria that a process must meet to be useful in the context of click chemistry has been defined by Sharpless *et al.*, as reactions that are wide in scope, stereospecific, easy to perform, and uses only readily available reagents to produce the best yields and highest rates.^{32b} More ideally, the reaction work-up and purification should involve the use of benign solvents and should avoid chromatography. Although meeting the requirements of a click reaction is a tall order, several processes have been identified which step up to the mark: nucleophilic ring opening reactions; non-aldol

carbonyl chemistry; additions to carbon–carbon multiple bonds; and cycloaddition reactions shown respectively in Figure 10.

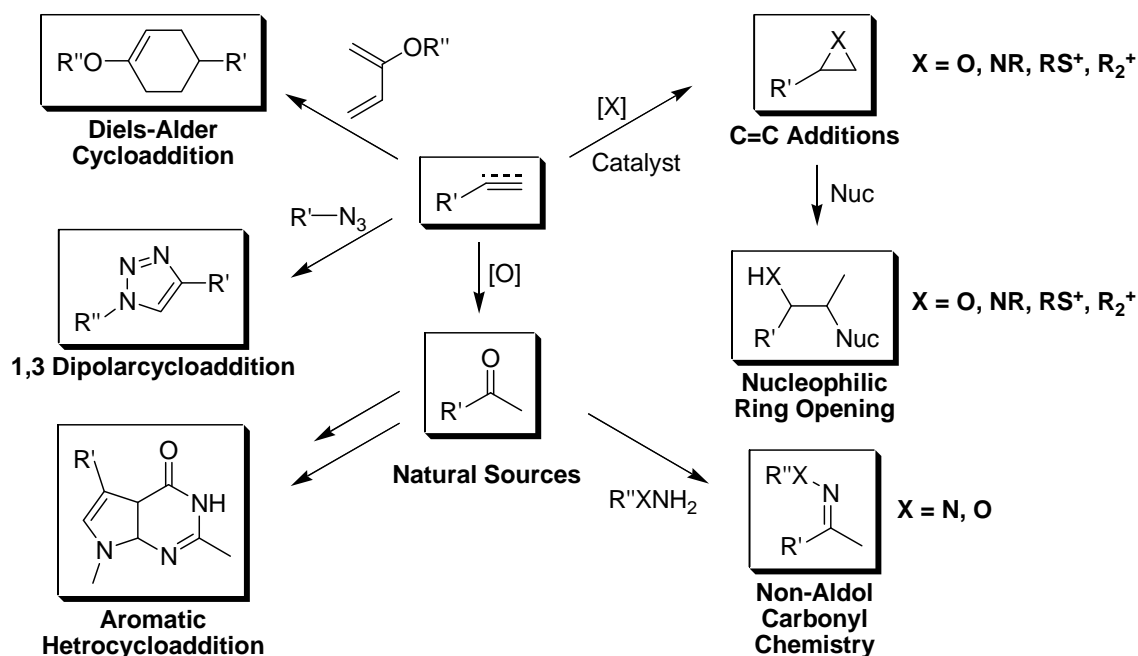


Figure 10. Schematic illustration of different types of Click reactions

Among the carefully selected reactions, as depicted in Figure 10, Cu(I)-catalyzed variant of the Huisgen 1,3-dipolar cycloaddition of azides and alkynes to afford 1,2,3-triazoles become the gold standard of click chemistry due to its reliability, specificity and biocompatibility. Rolf Huisgen was the first one to understand the scope of this reaction but it was K. Barry Sharpless who referred to this cycloaddition as the “cream of the crop” of click reactions.^{33, 32b} Further interest in this reaction stems from the interesting biological activity of 1,2,3-triazoles. These heterocycles function as rigid linking units that can mimic the atom placement and electronic properties of a peptide bond without the same susceptibility to hydrolytic cleavage. Since the foundations of click reactions were laid, there has been an explosive growth in publications describing a wealth of applications of this practical and sensible chemical approach.

A notable variant of the Huisgen 1,3-dipolar cycloaddition is the copper(I) mediated conjugation of organic azide and terminal alkynes to afford 1,4-regioisomers of 1,2,3-triazoles as the sole products. Nowadays, this reaction is better termed the Copper(I)-catalyzed Azide-Alkyne Cycloaddition (CuAAC) (Figure 11). While the reaction can be performed using commercial sources of copper(I) such as cuprous bromide or iodide, the reaction gave the best yields when a mixture of copper(II) salt (e.g. copper(II) sulfate) and a reducing agent (e.g.

sodium ascorbate) is employed to generate Cu(I) *in situ*. For improving the outcome of click reaction in biological systems, most often the ligands are required to stabilize Cu(I) which is

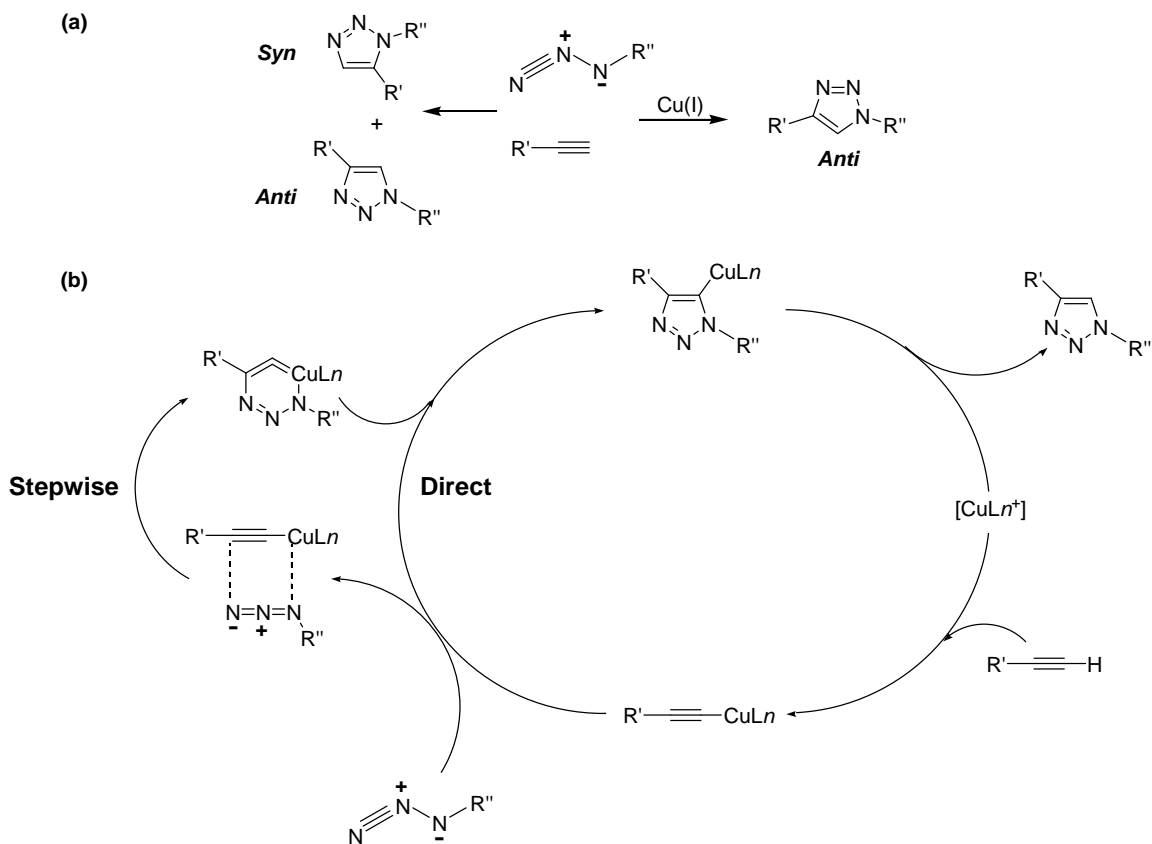


Figure 11. Schematic illustration of (a) Huisgen 1,3 dipolar cycloaddition; (b) Mechanistic pathway for the Cu(I) catalyzed azide-alkyne cycloaddition

otherwise unstable in aqueous solvents. True to the merits of click reaction, the reaction can be performed in a wide range of solvents and mixtures of water and a variety of (partially) miscible organic solvents including alcohols, DMSO, DMF, *t*BuOH and acetone work well.

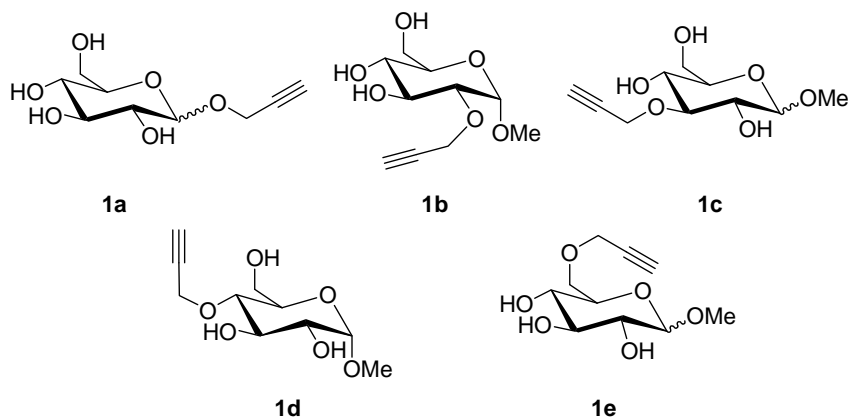


Figure 12. Positional isomers of propargylated glucose

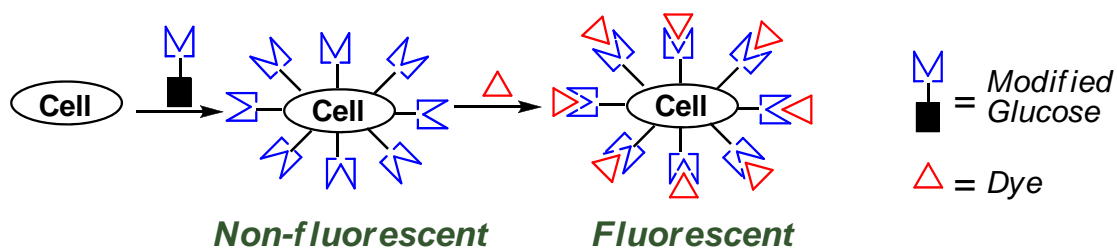
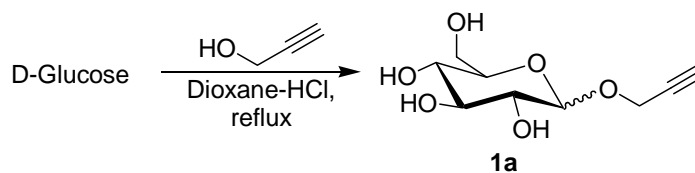


Figure 13. *Cartoon representation of the overall strategy*

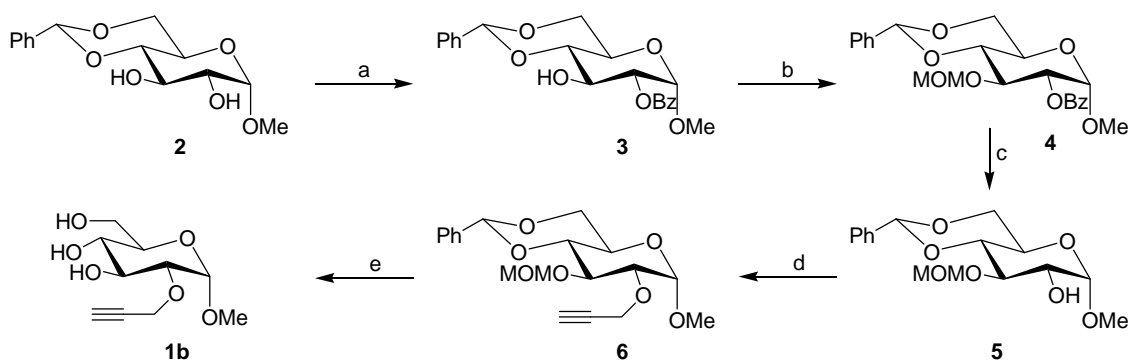
Moving ahead with our intensions of providing a more generalist and universal approach to probe the glycosylation pathways, we decided to choose the pyranose form of glucose for our experiments. Glucose, a monosaccharide, is an important carbohydrate in biological studies for its involvement in most of the metabolic processes inside the cells. It is pertinent to mention that bacteria uses glucose as a sole C-source for their survival. Following this, we hypothesized that it would be possible to probe the bacterial cell wall, if modified monosaccharide is provided to the bacteria as a carbon source. This would enable us to monitor the cellular uptake by the addition of azidocoumarin into the cell lysate which is further analysed for fluorescence (Figure 13). We envisioned synthesizing all the positional isomers of glucose (pyranose form) by selectively blocking the hydroxyl functionality sequentially with propargyl group (Figure 12). To begin our investigation, we started with the synthesis of the compounds which are discussed hereunder in detail.



Scheme 1. *Synthesis of 1a*

Starting with the synthesis of **1a**, D-Glucose was refluxed with propargyl alcohol under acidic conditions provided by dioxane-HCl to afford the propargylated glucose derivative, **1a** as an anomeric mixture (Scheme 1). The compound was confirmed by its ^1H and ^{13}C NMR spectra in which all the resonances were in conformity with the reported values.³⁴

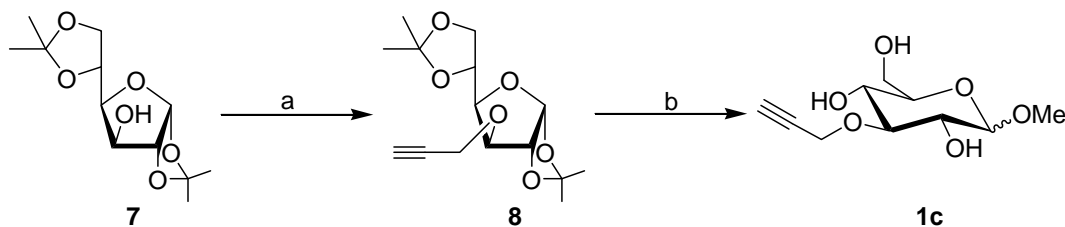
Commercially available methyl 4,6-*O*-benzylidene- α -D-glucopyranoside, **2** was converted to methyl 4,6-*O*-benzylidene-2-*O*-benzoyl- α -D-glucopyranoside, **3** reproducing the reported literature protocols. Further, the hydroxyl functionality of compound **3** was blocked employing the MOMCl in DIPEA using CH_2Cl_2 to yield the MOM ether, **4**. Successive deprotection of the benzoyl moiety of **4** under Zemplen conditions methanol resulted in



Reagents and conditions: (a) reference 35; (b) MOMCl, DIPEA, CH₂Cl₂, 12 hr, 97%; (c) NaOMe, MeOH, rt, 30 min, 96%; (d) Propargyl bromide, NaH, DMF, *n*Bu₄N⁺I⁻, 0°C-rt, 3 hr, 94%; (e) MeOH-HCl, reflux, 2 hr, 87 %

Scheme 2. Schematic illustration of the synthesis of **1b**

compound **5** which gave us an access to the free hydroxyl functionality at C-2 that enabled us to appropriately position the propargyl group at C-2. This was achieved by alkylation of the free alcohol **5** employing propargyl bromide in NaH using DMF as a solvent to result in propargyl ether **6**. Global deprotection of the alkylated MOM ether, **6** via refluxing in methanolic-HCl afforded the required C-2 isomer, **1b** as a free flowing brown solid (Scheme 2). The ¹H NMR spectrum of the propargylated glucose revealed resonances at δ 2.81 (t, 1H, *J* = 2.4 Hz) indicating the presence of alkyne functionality and a singlet at δ 3.31 (3H) of the methyl group positioned at anomeric position. Compound **1b** was further confirmed by the ¹³C NMR spectrum which indicated the presence of a single isomer with the resonances due to the anomeric carbon at δ 96.5 ppm.



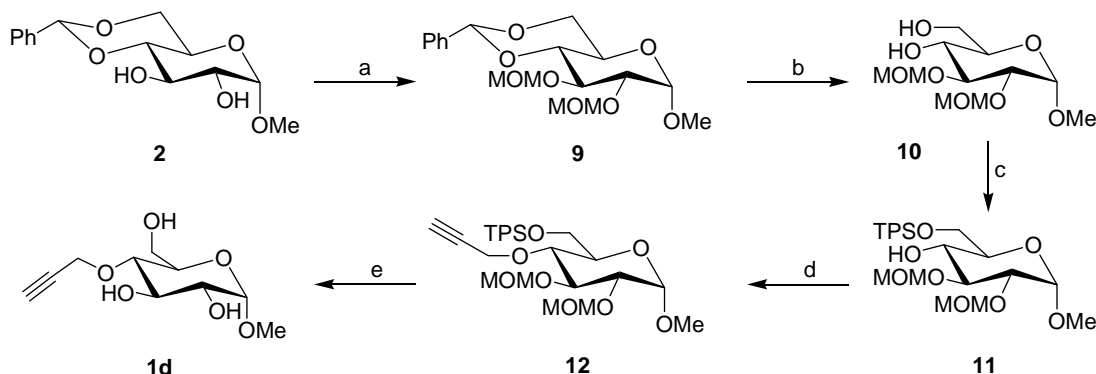
Reagents and conditions: (a) Propargyl bromide, NaH, DMF, *n*Bu₄N⁺I⁻, 0 °C-rt, 3 h, 90% (b) MeOH-HCl, reflux, 2 h, 80%

Scheme 3. Schematic illustration of the synthesis of **1c**

Successful in our efforts towards the synthesis of the two modified glucose derivatives, we continued with the synthesis of **1c**. Free hydroxyl functionality of the commercially available, 1,2:5,6-Di-*O*-isopropylidene glucofuranose, **7** was alkylated with the propargyl bromide employing NaH in DMF to yield, propargyl ether **8**. The furanose form of compound, **8** was respectively transformed to the compound **1c** as an anomeric mixture by refluxing in methanolic-HCl (Scheme 3). The ¹H NMR spectrum of compound **1c** established the existence of propargyl ether at δ 2.76 (t, 1H, *J* = 2.76 Hz) and due to the presence of the anomeric

mixture the ^{13}C NMR spectrum indicated two anomeric carbon at δ 98.8 and 102.4 ppm respectively.

Commercially available methyl 4,6-*O*-benzylidene- α -D-glucopyranoside, **2** was transformed to di-MOM ether, **9** using the conditions as indicated in Scheme 4. The benzylidene protecting group of **9** was removed employing acidic conditions provided by PTSA in methanol to result in the diol, **10**. The primary hydroxyl functionality of **10** was selectively blocked as silyl ether exploiting the steric hindrance at C-4 carbon. The resulting alcohol, **11** was then successfully converted to propargyl ether, **12** employing the basic environment. Respective removal of all the protecting groups from **12** gave the desired modified monosaccharide, **1d** as a free flowing white solid (Scheme 4). In the ^1H NMR spectrum of **1d**, resonances at δ 2.87 (t, 1H, $J = 2.37$ Hz), 3.35 (s, 3H) ppm indicated the presence of the alkyne functionality and methyl glucoside respectively, while ^{13}C spectrum pointed towards the occurrence of single isomer with the anomeric carbon resonances at δ 98.6 ppm.

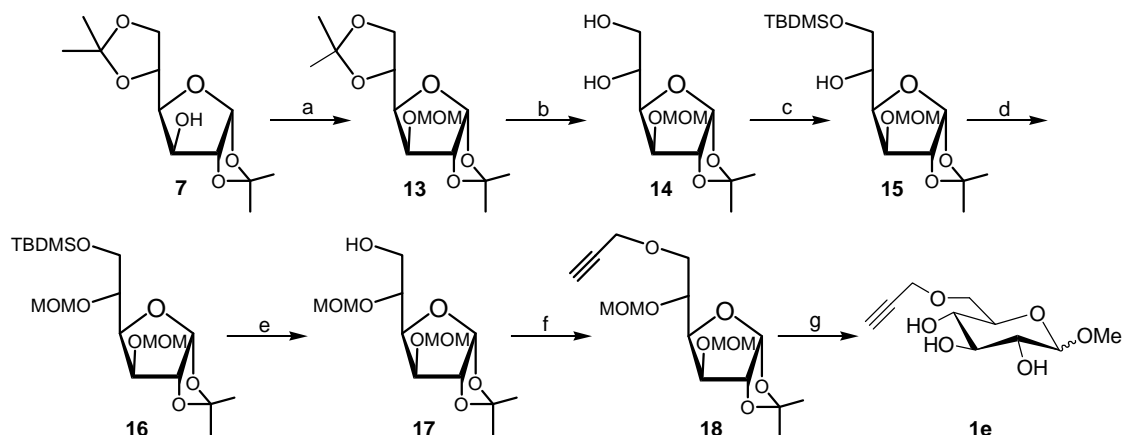


Reagents and conditions : (a) MOMCl, DIPEA, CH_2Cl_2 , 12h, 90%; (b) PTSA (cat.), CH_3OH , 2 h, 95%; (c) TBDPSCI, Triethylamine, CH_2Cl_2 , 4 hr, 97%; (d) Propargyl bromide, NaH, DMF, $n\text{Bu}_4\text{N}^+\text{I}^-$, 0 $^\circ\text{C}$ -rt, 3 h, 92%; (e) CH_3OH -HCl, 2 h, 85%.

Scheme 4. Schematic elaboration of the synthesis of **1d**

Once again 1,2:5,6-Di-*O*-isopropylidene glucofuranose, **7** was employed as the starting material for the synthesis of **1e**. The 3-OH group of **7** was converted to MOM ether after which the primary 5,6-isopropylidene protecting group was carefully deprotected under acidic environment provided by PTSA in methanol to afford the diol **14**. The primary alcoholic group was then protected as the silyl ether employing TBDMSCl, triethylamine in CH_2Cl_2 to afford the secondary alcohol **15** which was subsequently converted to MOM ether **16**. For the installation of propargyl group at C-6 position, the silyl moiety was removed under acidic conditions to yield the primary alcohol, **17**, which was then propargylated using propargyl bromide and NaH in DMF to afford the modified monosaccharide, **18**. Refluxing the

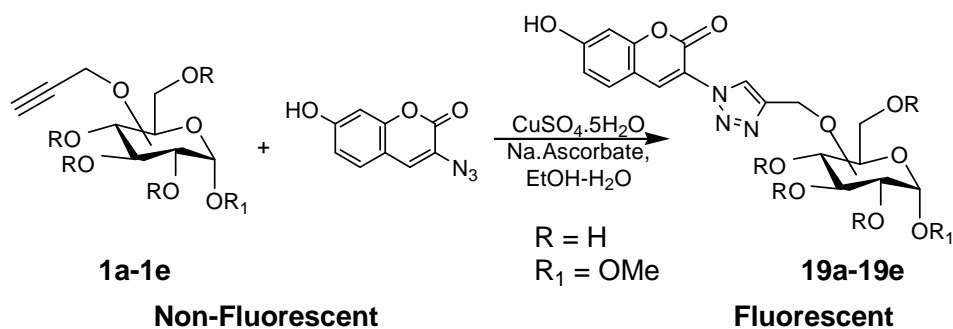
compound **18** in methanolic-HCl gave the much needed monosaccharide derivative **1e** (Scheme 5). In ^1H NMR spectrum of **1e**, due to the acetylenic group resonances were observed at δ 2.51 (t, 1H, $J=2.24$ Hz) ppm while the two anomeric carbons were observed in ^{13}C NMR spectrum at δ 99.6 and 103.5 ppm respectively suggesting the presence of the compound, **1e** as the anomeric mixture.



Reagents and Conditions : (a) MOMCl, DIPEA, CH_2Cl_2 , 6 h, 95%; (b) *p*-toluene sulfonic acid, CH_3OH , 5 h, 85%; (c) TBDMSCl, Triethylamine, CH_2Cl_2 , 4 h, 95%; (d) MOMCl, DIPEA, CH_2Cl_2 , 6 h, 94%; (e) *p*-toluene sulfonic acid, CH_3OH , 5 h, 85%; (f) Propargyl bromide, NaH, DMF, $n\text{Bu}_4\text{N}^+\text{I}^-$, 0 °C-rt, 3 h, 85%; (g) Methanol-HCl, 2 h, 70%.

Scheme 5. Schematic demonstration of the synthesis of **1e**

It is pertinent to mention here that although the compounds **1a**, **1c** and **1e** were obtained as the anomeric mixtures, we hypothesized that they would still be useful for the imaging of the bacterial cell wall as either or both the isomers will be able to penetrate the cell wall. Also, the glycan synthesis starts with the formation of UDP in which case the cell has to hydrolyse the glycoside. Having completed the synthesis of the modified monosaccharides **1a-1e**, we decided to test practicability of their reaction with the 3-azidocoumarin (Scheme 6), it was heartwarming to notice that all the resulting glycoconjugates triazoles **19a-19e** were fluorescent showing intensity maxima at 473 nm when excited with 365 nm light, as analysed by their fluorescent spectra (Figure 14).



Scheme 6. Synthesis of glycoconjugates triazoles

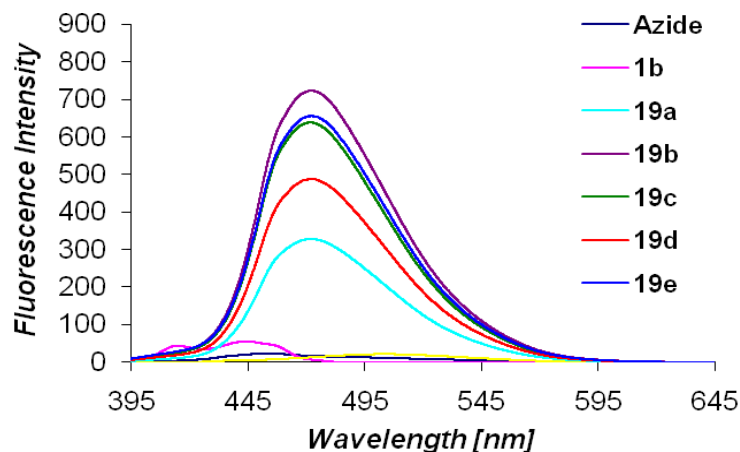


Figure 14. Fluorescence spectra of the glycoconjugate triazoles

Now having acquired all the desired water-soluble positional isomers of propargyl modified glucose derivatives, we decided to check the feasibility of the molecules to image the microorganisms. *Escherichia coli* BL21 cells were allowed to grow in the presence of propargyl ethers **1a-1e** and harvested after 12 h. Gratifyingly *E. coli* cells tolerated the propargyl substitution in glucose and grew well, and subsequently the cells were pelleted *via* centrifugation and washed two times with PBS buffer (pH 7.1) so as to remove the excess propargyl derivative. All the cells were suspended again in PBS buffer (pH 7.8) 2X and respectively subjected to sonification so as to obtain the cell lysate which was treated with coumarinyl azide to examine the fluorescence labeling of the modified cellular surfaces of *E. coli* BL21 cells *via* 1,3-dipolar cycloaddition or ‘click’ reaction promoted by copper sulphate and sodium ascorbate. Surprisingly, the cellular uptake of the propargyl substitution at C-2 position, **1b** was found to be the most preferred by the bacteria among all the isomers as deduced by the fluorescence spectra (Figure 15); hence the rest of the experiments were performed with **1b** only.

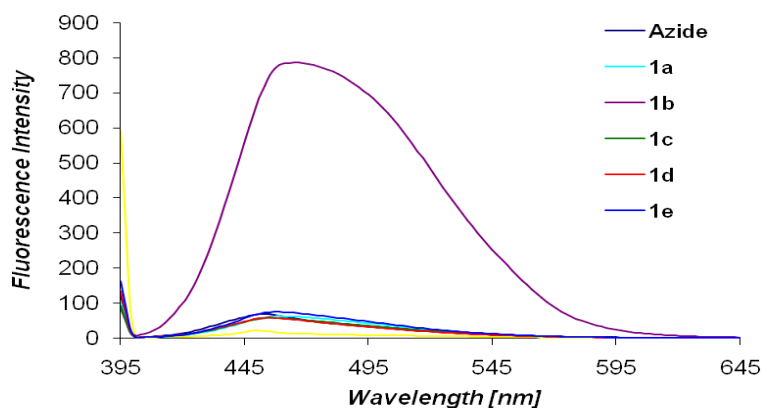
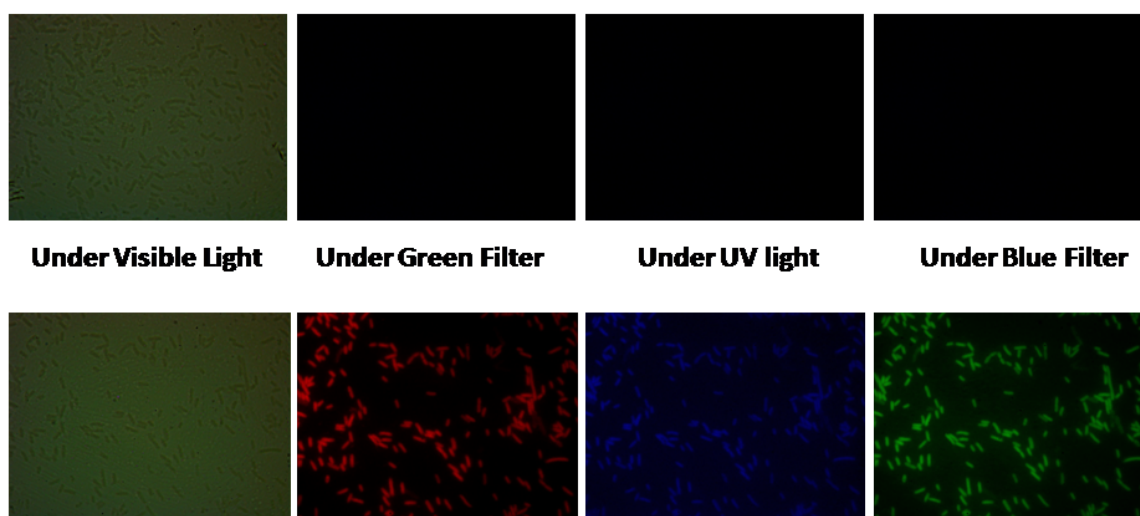


Figure 15. Fluorescence spectra of the BL21 cells

The feasibility of exploiting coumarin-based probe for visualizing bacterial cells through fluorescence imaging was then investigated. The *E. coli* BL21 cells were allowed to grow in the presence of glucose or propargyl ether **1b** for 12 h, and processed employing the same procedure as followed for examining the fluorescence spectrum except that the cells were not lysed but were subsequently fixed on cover slips and treated with the coumarinyl azide, copper sulphate and sodium ascorbate to effect the ‘click’ reaction. *E. coli* BL21 cells incubated in the presence of **1b** showed bright fluorescent images when compared; on the other hand control experiments carried out without **1b** showed no fluorescence (Figure 16).

Control Cells



Cells grown in the presence of **1b**

Figure 16. Fluorescence images of the BL21 cells

Furthermore, we thought of extending the overall strategy to *Lactobaccilus* and yeast cells as well. It may be mentioned here that while the *Escherichia coli* is a Gram positive bacteria, *Lactobaccilus lacti* is the Gram negative bacteria. We started with the *Lactobaccilus* cells which were allowed to grow in the presence of **1b** for 24 h after which they were processed as delineated above for *E. coli* cells. On the examination of the fluorescence spectra of the cell lysate after treatment with the click reagents, **1b** seemed to most preferred among all the monosaccharide derivatives supplied to the bacteria while the control cells which were devoid of the modified derivative showed no fluorescence absorption, Figure 17. Hereafter, we continued in our intentions towards the imaging experiments employing only **1b**. For this purpose, the *Lactobaccilus lacti* cell cultures grown in the presence of **1b** were cleansed of the excess sugar derivatives examined under the fluorescence microscope after performing the

click reaction. Delightfully the experimental cultures exhibited brilliant fluorescence as compared to the control cells (Figure 18).

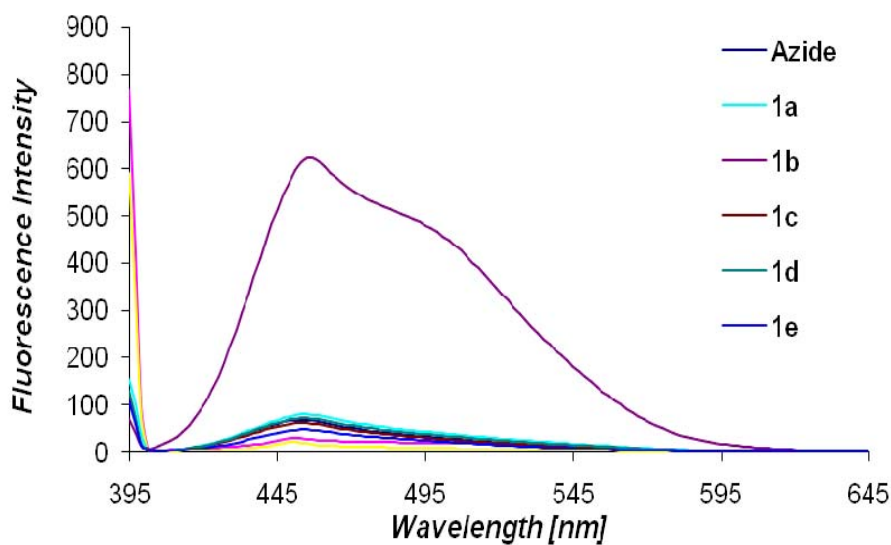
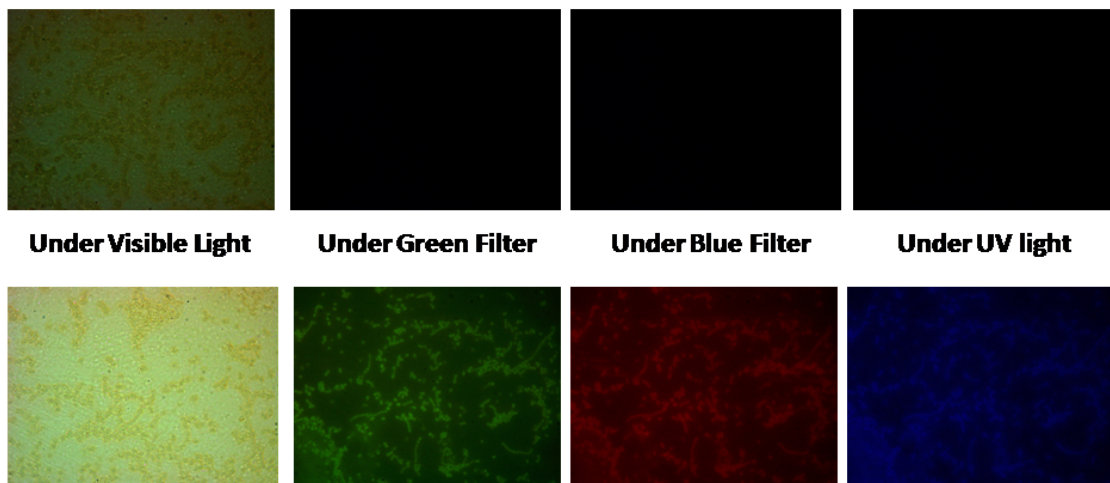


Figure 17. Fluorescence spectra of the *Lactobacillus lacti* cells

Control Cells



Cells grown in the presence of 1b

Figure 18. Fluorescence images of the *Lactobacillus lacti* cells

Still not satisfied after having been awarded with successful results in lieu of our efforts, we attempted to widen the scope of our experiments beyond the bacterial cultures. In this regard, *Sacchromyces* (wild type) cells were selected for our endeavors. The yeast cells were grown overnight with the propargyl derivatives **1a-1e**, and isolated by centrifugation.

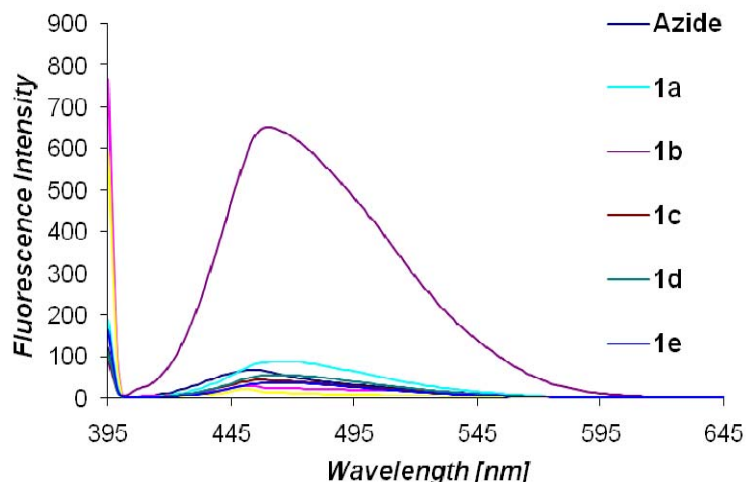
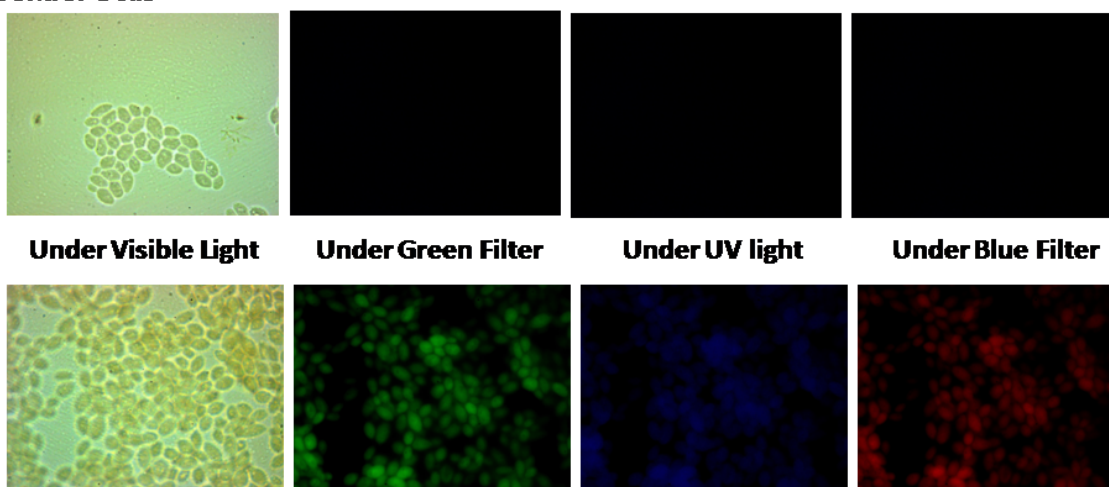


Figure 19. Fluorescence spectra of the *Sacchromyces* cells

Control Cells



Cells grown in the presence of 1b

Figure 20. Fluorescence images of the *Sacchromyces* cells

The cells were washed two times with phosphate buffer (pH 7.8) 2X so as to get rid of the excess sugars sticking to the cells. Subsequently, all the cell cultures were individually lysed by sonification and ‘click’ reaction was executed in the presence of azidocoumarin, copper sulphate and sodium ascorbate. On the inspection of the fluorescence spectra of the cell lysates, still **1b** appeared to be the favorable sugar for the yeast cells among all the modified positional isomers of glucose. The results of the fluorescence spectra are displayed in Figure 19. The cells were incubated overnight along with **1b** and processed accordingly to prepare the slides to examine under the microscope and satisfyingly the outcome of the imaging experiments as displayed in Figure 20 came out as anticipated with cells influenced by **1b** demonstrating intensely fluorescent images as compared to the control

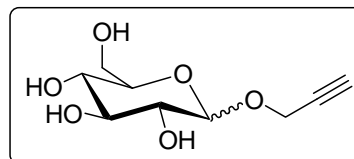
experiments which were fed with the normal glucose instead of the modified propargylated sugar **1b**.

In conclusion, carbohydrate engineering by exogenously added monosaccharide supplements is a technique of immense importance for studying various glycan specific diseases. The utility of profluorophoric coumarine-based azide for fluorescent labelling and imaging of bacterial, (both Gram positive as well as Gram negative cells), and yeast cultures was demonstrated for the first time. In this endeavour, we showed that a propargyl group of monosaccharide can be incorporated into the bacterial and yeast cultures which was probed using azido coumarinyl profluorophore under 1,3-dipolar cycloaddition or 'click' chemistry conditions. Often, it is difficult to identify the role of bacterial glycans as they exist in micro and heterogeneous forms. The current protocol of fluorescent labelling *via* the click chemistry facilitates monitoring of glycan expression and the related processes which are otherwise poorly accessible.

Chapter 2: Experimental Section

Synthesis of propargyl α/β -D-glucopyranoside:

D-Glucose (180mg, 1mmol) was suspended in 3 mL of dioxane-HCl and heated with propargyl alcohol (0.445mL, 5 mmol) at 100°C for 5 h. After the completion of the reaction, as evident by the

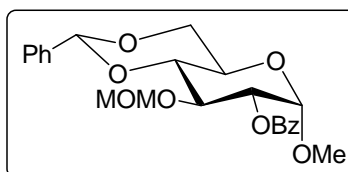


TLC (CH_2Cl_2 -MeOH; 5:1), the reaction was neutralised with triethylamine and the solvent was then concentrated *in vacuo* to give a black syrup. The syrup was subjected to column chromatography, excess of propargyl alcohol was eluted using ethyl acetate followed by the elution of the product **1a** (174 mg, 80%) as an anomeric mixture. The NMR resonances peaks were found to be in conformity with the literature values reported for the same compound.³⁴

Synthesis of methyl 4,6-O-benzylidene-2-O-benzoyl-3-O-

(methoxy-O-methyl)- α -D-glucopyranoside:

To an ice cooled stirred solution of methyl 4,6-O-benzylidene-2-O-benzoyl- α -D-glucopyranoside (386 mg, 1 mmol) **3** and diisopropyl amine (

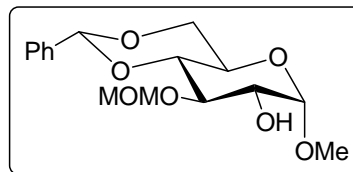


0.347 mL, 2 mmol) in dichloromethane (7 mL) was added MOMCl (0.150 mL, 2 mol) over a period of 5 min. The reaction was allowed to stir at room temperature for 12 h. TLC examination of the reaction indicated the formation of the compound as a single compound (Ethyl Acetate-Petroleum Benzine 1:3). The reaction was diluted with water and extracted with dichloromethane (2X); the combined extracts were dried over anhydrous sodium sulphate. The solvent was evaporated under reduced pressure to give the crude MOM ether. It was purified over silica gel column chromatography (Ethyl Acetate-Petroleum Benzine; 1:3) to give Compound **4** (418 mg, 97% yield). $[\alpha]_D$ (CHCl_3 , c 1.65) = +120.1°. ^1H NMR (200 MHz, CDCl_3): δ 3.32 (s, 3H), 3.39 (s, 3H), 3.66-3.95 (m, 3H), 4.29-4.43 (m, 2 H), 4.79 (ABq, 2 H, J = 6.86 Hz), 5.04-5.15(m, 2 H), 5.59 (s, 1H), 7.33-7.64 (m, 8H), 8.10-8.15 (m, 2H); ^{13}C NMR (50 MHz, CDCl_3): δ 55.4, 55.8, 62.3, 69.0, 72.8, 73.3, 81.5, 97.1, 97.8, 101.6, 126.1, 126.1, 128.2, 128.2, 128.4, 128.4, 129.0, 129.5, 129.9, 129.9, 133.3, 137.2, 165.9; Anal Calcd. for $\text{C}_{23}\text{H}_{26}\text{O}_8$: C, 64.18; H, 6.09. Found: C, 63.75; H, 6.48.

Synthesis of methyl 4,6-O-benzylidene-3-O-(methoxy-O-methyl)- α -D-glucopyranoside:

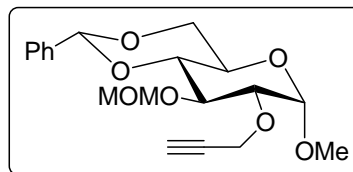
NaOMe (108 mg, 2 mmol) was added to the solution of methyl 4,6-O-benzylidene-2-O-benzoyl 3-O-(methoxy-O-methyl)- α -D-glucopyranoside, **4** (430 mg, 1 mmol) in anhydrous methanol (5

mL). The resulting solution was stirred at ambient temperature until the completion of the reaction as evident by the TLC (Ethyl acetate-Petroleum benzine; 1:3). The reaction was diluted with water and extracted with ethyl acetate. The



combined extracts were dried over anhydrous sodium sulphate and concentrated *in vacuo*. The alcohol **5** (312 mg, 96%) was obtained after the purification over silica gel chromatography. $[\alpha]_D$ (CHCl₃, *c* 0.92) = +140.5°; ¹H NMR (200 MHz, CDCl₃): δ 3.19 (s, 1H), 3.42 (s, 3H), 3.44 (s, 3H), 3.49-3.95 (m, 5H), 4.26-4.31 (m, 1H), 4.83 (s, 3H), 5.53 (s, 1H), 7.32-7.54 (m, 5H); ¹³C NMR (50 MHz, CDCl₃): δ 55.3, 55.9, 62.4, 69.0, 72.2, 77.8, 80.5, 97.4, 100.0, 101.5, 126.1, 126.1, 128.2, 128.2, 129.0, 137.2. Anal. Calcd. for C₁₆H₂₂O₇: C, 58.89; H, 6.79. Found C, 59.45; H, 7.21

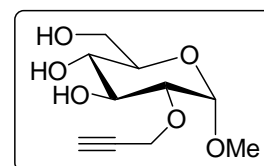
Synthesis of methyl 4,6-O-benzylidene-2-O-propargyl-3-O-(methoxy-O-methyl)-α-D-glucopyranoside: To an ice-cooled solution of methyl 4,6-O-benzylidene-3-O-(methoxy-O-methyl)-α-D-glucopyranoside, **5** (326 mg, 1 mmol) in



anhydrous DMF (5mL) was added sodium hydride (60 mg, 1.5 mmol, 60% oil suspension) and stirred for 1 h at room temperature. Propargyl bromide (0.132 mL, 1.5 mmol) was introduced drop-wise into the mixture at 0°C and stirred at room temperature for 1 h. The resulting suspension was quenched with saturated ammonium chloride and extracted three times with diethyl ether. The extracts were combined and dried over anhydrous sodium sulphate. The solvent was removed under vacuum and the resulting residue was subjected to silica gel column chromatography and the compound was eluted with Ethyl acetate-Petroleum benzine (1:4) to give compound **6** (342 mg, 94 %). $[\alpha]_D$ (CHCl₃, *c* 0.90) = + 38.0°. ¹H NMR (200 MHz, CDCl₃): δ 2.47 (t, 1H, *J*=2.38), 3.36 (s, 3H), 3.44 (s, 3H), 3.52-3.85 (m, 5H), 4.09-4.30 (m, 2H), 4.38 (d, 2H, *J* = 2.38 Hz), 4.78-4.94 (m, 2H), 5.53 (s, 1H), 7.33-7.50 (m, 5H). ¹³C NMR (50 MHz, CDCl₃): δ 55.1, 55.7, 58.8, 62.2, 68.9, 74.1, 74.9, 78.6, 79.6, 81.5, 97.2, 98.8, 101.4, 126.0, 126.0, 128.1, 128.1, 128.1, 137.2. Anal Calcd. for C₁₉H₂₄O₇ : C, 62.63; H, 6.64. Found : C, 62.13; H, 6.17.

Synthesis of methyl 2-O-propargyl-α-D-glucopyranoside:

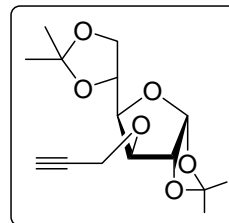
Compound **6** (364 mg, 1 mmol) prepared *vide supra* was refluxed in methanolic-HCl (10 mL) for 2 h. After the completion of the reaction, the solution was neutralised with triethylamine, the solvent was



removed under pressure and was transferred onto a silica gel column and the compound **1b** (201 mg, 87%) was eluted with dichloromethane-methane (9:1) and obtained as a white solid. $[\alpha]_D$ (H₂O, *c* 2.0) = +124.0°. ¹H NMR (200 MHz, CD₃OD): δ 2.81 (t, 1H, *J* = 2.4 Hz), 3.31 (s, 3H), 3.36-3.80 (m, 9H), 4.25 (d, 2H, *J* = 2.27 Hz), 4.98 (d, 1H, *J* = 3.53 Hz). ¹³C NMR (50 MHz, Acetone-*d*₆): δ 54.2, 57.5, 60.0, 69.1, 70.9, 71.7, 75.7, 78.2, 79.2, 96.5. Anal. Calcd. for C₁₀H₁₆O₆: C, 51.72; H, 6.94. Found: C, 51.99; H, 7.32.

Synthesis of 1,2:5,6-di-*O*-isopropylidene-3-*O*-propargyl- α -D-glucofuranose:

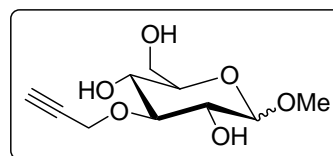
1,2:5,6-Di-*O*-isopropylidene-2-*O*-propargyl glucofuranose, **7** (261 mg, 1 mmol) was dissolved in anhydrous DMF (5mL), cooled to 0°C and was added sodium hydride (60 mg, 1.5 mmol, 60% oil suspension). The resulting reaction mixture was stirred at room



temperature for 1h. Propargyl bromide (0.132 mL, 1.5 mmol) was introduced drop-wise at 0°C and stirred at room temperature for 1 h. The reaction was quenched by the addition of methanol and extracted with diethyl ether (3X). The solvent was concentrated *in vacuo* to yield the propargyl furanoside derivative **8** (268 mg, 90%). $[\alpha]_D$ (CHCl₃, *c* 0.60) = -12.1°; ¹H NMR (200 MHz, CDCl₃): δ 1.32 (s, 3H), 1.35 (s, 3H), 1.43 (s, 3H), 1.50 (s, 3H), 2.49 (t, 1H, *J* = 2.40 Hz), 3.96-4.15 (m, 4H), 4.23-4.28 (m, 3H), 4.63 (s, 1H), 5.88 (s, 1H); ¹³C NMR (50 MHz, CDCl₃): δ 25.3, 26.2, 26.7, 26.7, 58.0, 67.1, 72.4, 74.9, 79.2, 80.9, 81.5, 82.8, 105.1, 108.9, 111.8; Anal. Calcd. for C₁₅H₂₂O₆: C, 60.39; H, 7.43. Found: C, 60.61; H, 7.65.

Synthesis of methyl 3-*O*-propargyl- α/β -D-glucofuranoside:

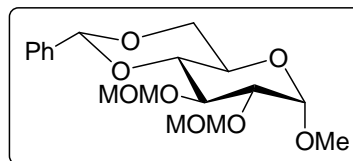
1,2:5,6-Di-*O*-isopropylidene-3-*O*-propargyl glucofuranose, **8** (298 mg, 1 mmol) was dissolved in methanolic-HCl (7 ml) and refluxed for 2 h. Crude **1c** was purified by silica gel column



chromatography (CH₃OH:CH₂Cl₂-1:10) to yield **1c** (185 mg, 80%) as the anomeric mixture which was analyzed as such for the characterization purpose. ¹H NMR (200 MHz, CD₃OD): δ 2.76 (t, 1H, *J* = 2.76 Hz), 3.28 (s, 3H), 3.33-3.75 (m, 9H), 4.36 (d, 2H, *J* = 2.37 Hz), 4.65 (m, 1H); ¹³C NMR (50 MHz, CDCl₃): δ 54.6, 55.4, 57.2, 57.4, 59.4, 60.0, 62.7, 68.7, 68.9, 70.4, 71.1, 74.2, 75.4, 75.7, 76.8, 78.9, 79.4, 80.5, 81.6, 82.6, 98.8, 102.4, 108.8. Anal. Calcd. for C₁₀H₁₆O₆: C, 51.72; H, 6.94. Found: C, 51.85; H, 6.83.

Synthesis of methyl 4,6-*O*-benzylidene-2,3-di-*O*-(methoxy-*O*-methyl)- α -D-glucofuranoside: To a solution of methyl 4,6-*O*-benzylidene- α -D-glucofuranose, **2** (282 mg, 1

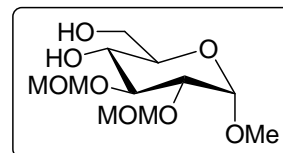
mmol) in CH₂Cl₂ (7 mL), DIPEA (0.695 mL, 4 mmol) was added and the resulting mixture was cooled in an ice bath. MOMCl (0.301 mL, 4 mmol) was added drop-wise over a period of 10 min, the reaction mixture was stirred at room



temperature for 12 h, was diluted with water and extracted with CH₂Cl₂ (3X). The combined organic extracts were dried over anhydrous sodium sulphate and the solvent was removed under reduced pressure and crude syrup was purified using silica gel column chromatography with ethyl acetate-petroleum benzene as the mobile phase to obtain the respective di-MOM ether **9** (333 mg, 90%). $[\alpha]_D$ (CHCl₃, *c*1.05) = +9.05°; ¹H NMR (200 MHz, CDCl₃): δ 3.35 (s, 3H), 3.43 (s, 3H), 3.53-3.85 (m, 7H), 4.09-4.27 (m, 2H), 4.73-4.83 (s, 5H), 5.53 (s, 1H), 7.34-7.47 (m, 5H); ¹³C NMR (50 MHz, CDCl₃): δ 55.2, 55.5, 55.7, 62.2, 69.0, 73.9, 78.4, 81.5, 97.3, 97.8, 99.8, 101.4, 126.0, 126.0, 128.1, 128.1, 128.9, 137.2; Anal. Calcd. for C₁₈H₂₆O₈: C, 58.37; H, 7.08. Found : C, 57.82; H, 7.52.

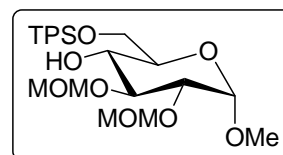
Synthesis of methyl 2,3-di-O-(methoxy-O-methyl)-α-D-glucopyranoside:

Compound **9** (370 mg, 1 mmol) was dissolved in methanol (5 mL) and deprotection of the benzylidene was effected by the addition of catalytic amount of p-toluene sulfonic acid. After usual work-up, the diol **10** was obtained which was directly used in the next step without further characterization.



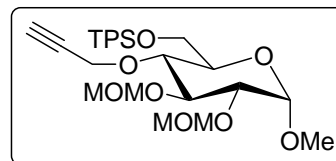
Synthesis of methyl 6-O-t-butyl diphenyl silyl-2,3-di-O-(methoxy-O-methyl)-α-D-glucopyranoside:

The diol **10** (282 mg, 1 mmol) was dissolved in CH₂Cl₂ (5 mL) and was added triethylamine (0.208 mL, 1.5 mmol) and cooled to 0°C. TBDPSCI (0.256 mL, 1mmol) was added portion-wise and the reaction mixture was stirred at room temperature until the completion of the reaction as indicated on the TLC (Ethyl acetate-Petroleum Benzene, 1:3). The reaction mixture was diluted with water and extracted with CH₂Cl₂ (3X). The combined extracts were dried over anhydrous sodium sulphate, the solvent was removed *in vacuo* and the title compound was purified by column chromatography to give compound **11** (505 mg, 97%). $[\alpha]_D$ (CHCl₃, *c* 1.22) = - 13.8°; ¹H NMR (200 MHz, CDCl₃): δ 1.06 (s, 9H), 3.41 (s, 3H), 3.43 (s, 3H), 3.44 (s, 3H), 3.49-3.74 (m, 4H), 3.83-4.02 (m, 3H), 4.69-4.86 (m, 5H), 7.35 – 7.42 (m, 6H), 7.69- 7.75 (m, 4H); ¹³C NMR(50 MHz, CDCl₃): δ 19.3, 26.8, 26.8, 26.8, 54.8, 55.4, 55.7, 63.7, 69.8, 77.1, 71.6, 83.9, 97.5, 98.4, 98.7, 127.6, 127.6, 127.6, 127.6, 129.5, 129.5, 133.5, 133.5, 135.6,



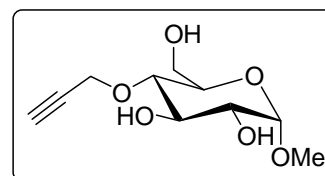
135.6, 135.6, 135.6; Anal Calcd. for C₂₇H₄₀O₈Si : C, 62.28; H, 7.74; Si, 5.39. Found: C, 62.55; H, 7.95; Si, 5.82.

Synthesis of methyl 6-*O*-*t*-butyl diphenyl silyl-2,3-di-*O*-(methoxy-*O*-methyl)-4-*O*-propargyl- α -D-glucopyranoside:



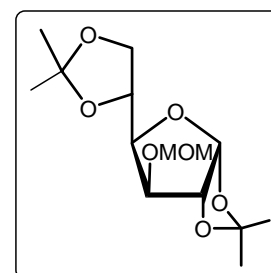
Preparative procedure is same as delineated above for compound **6**. $[\alpha]_D$ (CHCl₃, *c* 0.82) = +27.7°; ¹H NMR (200 MHz, CDCl₃): δ 1.06 (s, 9H), 2.32 (t, 1H, *J* = 2.25 Hz), 3.37 (s, 3H), 3.43 (s, 3H), 3.46 (s, 3H), 3.49 – 3.68 (m, 3H), 3.90- 4.0 (m, 3 H), 4.36–4.40 (m, 2H), 4.70-4.88 (m, 5H), 7.37–7.42 (m, 6H), 7.70-7.75 (m, 4H); ¹³C NMR (50 MHz, CDCl₃): δ 19.3, 26.8, 26.8, 26.8, 54.8, 55.5, 56.2, 56.2, 59.8, 62.9, 70.9, 74.3, 77.4, 78.1, 79.2, 79.8, 97.7, 98.3, 98.8, 127.5, 127.5, 127.6, 127.6, 129.6, 129.6, 133.3, 133.7, 135.6, 135.6, 135.8, 135.8; Anal Calcd. for C₃₀H₄₂O₈Si : C, 64.49; H, 7.58. Found : C, 65.12; H, 7.21.

Synthesis of methyl 4-*O*-propargyl- α -D-glucopyranoside:



Preparative protocol is same as described for compound **1b**. $[\alpha]_D$ (H₂O, *c* 2.0) = +170.6°; ¹H NMR (200 MHz, CD₃OD): δ 2.87 (t, 1H, *J* = 2.37 Hz), 3.35 (s, 3H), 3.39 (d, 2H, *J* = 9.69 Hz), 3.48-3.89 (m, 6H), 4.38 (d, 2H, *J* = 1.74 Hz), 4.74 (d, 2H, *J* = 3.72 Hz); ¹³C NMR (50 MHz, Acetone-d₆): δ 54.5, 59.2, 60.0, 70.0, 70.7, 72.6, 75.6, 76.9, 79.1, 98.6; Anal. Calcd. for C₁₀H₁₆O₆: C, 51.72; H, 6.94. Found: C, 51.44; H, 6.58.

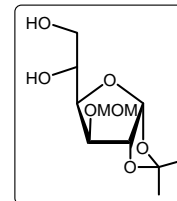
Synthesis of 1,2:5,6-Di-*O*-isopropylidene-3-*O*-(methoxy-*O*-methyl)- α -D-glucofuranose:



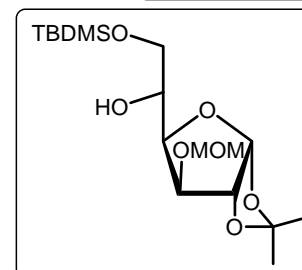
Preparative procedure is as delineated above for compound **4**. $[\alpha]_D$ (CHCl₃, *c* 0.84) = -12.0°; ¹H NMR (200 MHz, CDCl₃): δ 1.32 (s, 3H), 1.34 (s, 3H), 1.42 (s, 3H), 1.50 (s, 3H), 3.41 (s, 3H), 3.96 -4.33 (m, 4H), 4.56 (d, 1H, *J* = 3.52 Hz), 4.73 (d, 2H, *J* = 1.87 Hz), 5.89 (d, 1H, *J* = 3.53 Hz); ¹³C NMR (50 MHz, CDCl₃): δ 25.3, 26.2, 26.8, 30.9, 55.8, 67.5, 72.3, 79.1, 81.0, 83.3, 96.0, 105.2, 109.1, 111.9. Anal Calcd. for C₁₄H₂₄O₇ : C, 55.25; H, 7.95. Found: C, 55.75; H, 8.02.

Synthesis of 1,2-*O*-isopropylidene-3-*O*-(methoxy-*O*-methyl)- α -D-glucofuranose: To a solution of 1,2:5,6-Di-*O*-isopropylidene-3-*O*-(methoxy-*O*-methyl)- α -D-glucofuranose, **13** (500 mg, 1mmol) in methanol (5 mL) was added a catalytic amount of PTSA, the resulting solution was stirred at room temperature until the completion of the reaction as indicated by the

TLC (Ethyl Acetate; Petroleum Benzine 1:1). The solvent was then removed under reduced pressure and the diol thus obtained was carried further as such for the next step.



Synthesis of 1,2-*O*-isopropylidene-3-*O*-(methoxy-*O*-methyl)-6-*O*-*t*-butyl dimethylsilyl- α -D-glucofuranose: To a solution of the diol **14** (305 mg, 1 mmol) in dichloromethane (5 mL) was added triethylamine (0.208 mL, 1.5 mmol) and TBDMSCl (150 mg, 1 mmol) portion-wise at 0°C. The resulting reaction mixture was stirred at room temperature for 4 h. The reaction mixture was

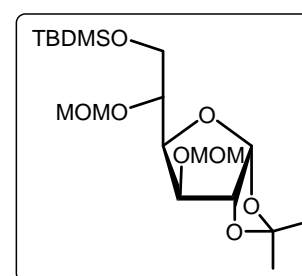


diluted with water and extracted with CH₂Cl₂ (3x50 mL). The combined extracts were dried on anhydrous sodium sulphate and purified on silica gel to obtain compound **15** (398 mg, 95 %). $[\alpha]_D$ (CHCl₃, *c* 0.90) = -14.17°; ¹H NMR (200 MHz, CDCl₃): δ 0.09 (s, 6H), 0.91 (s, 9H), 1.32 (s, 3H), 1.48 (s, 3H), 2.74 (b, 1H), 3.43 (s, 3H), 3.71-3.89 (m, 3H), 4.09-4.14 (m, 1H), 4.26 (d, 1H, *J* = 2.87 Hz), 4.56 (d, 1H, *J* = 3.70 Hz), 4.76 (dd, 2 H, *J* = 6.57, 12.37 Hz), 5.89 (d, 1H, *J* = 3.65 Hz); ¹³C NMR (50 MHz, CDCl₃): δ -5.5, -5.5, 18.3, 25.8, 25.8, 25.8, 26.3, 26.7, 55.8, 64.3, 68.2, 79.2, 79.9, 83.2, 96.5, 105.1, 111.7; Anal Calcd. for C₁₇H₃₄O₇Si: C, 53.94; H, 9.05; Si, 7.42. Found : C, 54.32; H, 8.89; Si, 7.82.

Synthesis of 1,2-*O*-isopropylidene-3,5-di-*O*-(methoxy-*O*-methyl)-6-*O*-*t*-butyl dimethylsilyl- α -D-glucofuranose:

Preparative procedure is same as delineated above for compound **4**.

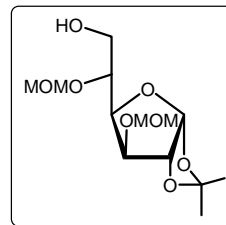
$[\alpha]_D$ (CHCl₃, *c* 0.72) = +7.3°; ¹H NMR (200 MHz, CDCl₃): δ 0.07 (s, 6H), 0.09 (s, 9H), 1.31 (s, 3H), 1.48 (s, 3H), 3.40 (s, 3H), 3.41 (s, 3H), 3.73-3.89 (m, 2H), 3.97- 4.04 (m, 1H), 4.13 (d, 1H, *J* =



2.74 Hz), 4.25-4.31 (m, 1H), 4.65 (d, 1H, *J* = 3.51 Hz), 4.70- 4.82 (m, 4 H), 5.88 (d, 1 H, *J* = 3.66 Hz); ¹³C NMR (50 MHz, CDCl₃): δ - 5.5, - 5.4, 18.3, 25.9, 25.9, 25.9, 26.3, 26.7, 55.7, 56.0, 63.7, 75.9, 78.1, 81.3, 83.0, 97.0, 97.3, 104.8, 111.7; Anal Calcd. for C₁₉H₃₈O₈Si : C, 54.00; H, 9.06; Si, 6.65. Found : C, 54.49; H, 8.58; Si, 6.35.

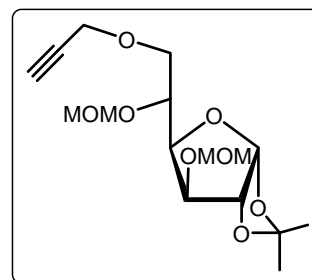
Synthesis of 1,2-*O*-isopropylidene-3,5-di-*O*-(methoxy-*O*-methyl)- α -D-glucofuranose: To a solution of the silyl ether **16** (464 mg, 1 mmol) was added catalytic amount of PTSA in methanol (5 mL) and stirred at room temperature for 5 h. The solvent was then evaporated *in vacuo* to yield a syrup which was purified by the silica gel column chromatography to achieve

the primary alcohol **17** (296 mg, 85%). $[\alpha]_D$ (CHCl₃, *c* 1.06) = +43.8°; ¹H NMR (200 MHz, CDCl₃): δ 1.32 (s, 3H), 1.49 (s, 3H), 2.99 (b, 1H), 3.40 (s, 3H), 3.44 (s, 3H), 3.64–3.73 (m, 1H), 3.83–3.95 (m, 2H), 4.16–4.21 (m, 2H), 4.63 (d, 1H, *J* = 3.68 Hz), 4.72 (s, 4H), 5.89 (d, 1H, *J* = 3.64 Hz); ¹³C NMR (50 MHz, CDCl₃): δ 26.2, 26.7, 55.8, 56.0, 63.8, 78.7, 79.1, 80.8, 82.7, 96.9, 97.3, 104.9, 111.8; Anal Calcd. for C₁₃H₂₄O₈: C, 50.64; H, 7.85. Found: C, 50.94; H, 7.67.



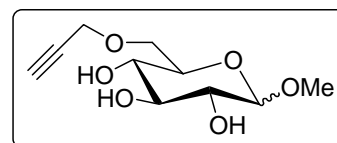
Synthesis of 1,2-*O*-isopropylidene-3,5-di-*O*-(methoxy-*O*-methyl)-6-*O*-propargyl- α -D-glucofuranose:

Preparative procedure is same as delineated above for compound **8**. $[\alpha]_D$ (CHCl₃, *c* 0.42) = +17.4°; ¹H NMR (200 MHz, CDCl₃): δ 1.32 (s, 3H), 1.49 (s, 3H), 2.41 (t, 1H, *J* = 2.38 Hz), 3.40 (s, 3H), 3.41 (s, 3H), 3.63–3.71 (m, 1H), 3.96–4.02 (m, 2H), 4.14 (d, 1H, *J* = 2.90 Hz), 4.19–4.24 (m, 2H), 4.29–4.34 (m, 1H), 4.66 (d, 1H, *J* = 3.78 Hz), 4.73 (d, 2H, *J* = 1.63 Hz), 4.75 (s, 2H), 5.88 (d, 1H, *J* = 3.55 Hz); ¹³C NMR (50 MHz, CDCl₃): δ 26.4, 26.8, 55.9, 56.0, 58.6, 70.3, 74.0, 74.4, 78.5, 79.7, 81.2, 83.0, 96.9, 97.3, 104.8, 111.9; Anal Calcd. for C₁₆H₂₆O₈: C, 55.48; H, 7.57. Found: C, 55.96; H, 7.37.

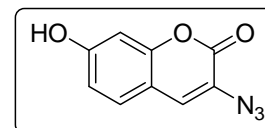


Synthesis of methyl 6-*O*-propargyl- α/β -D-glucopyranoside:

Preparative procedure is same as delineated above for compound **1b**. The compound was obtained as the anomeric mixture and characterization data for the anomeric mixture is provided. ¹H NMR (200 MHz, CDCl₃): δ 2.51 (t, 1H, *J* = 2.29 Hz), 3.43 (s, 3H), 3.50 (m, 3H), 3.67–3.86 (m, 4H), 4.21–4.25 (m, 3H), 4.61 (b, 1H), 4.77 (d, 1H, *J* = 3.68 Hz); ¹³C NMR (50 MHz, CDCl₃): δ 55.3, 57.1, 58.7, 68.3, 70.1, 70.5, 71.8, 73.3, 74.2, 75.0, 75.1, 79.6, 79.6, 99.6, 103.5; Anal Calcd. for C₁₀H₁₆O₆: C, 51.72; H, 6.94. Found: C, 51.83; H, 7.08.



Preparation of 3-Azido Coumarin: A mixture of 2,4 dihydroxy benzaldehyde (2.76 gm, 20 mmol), *N*-acetylglycine (2.34 gm, 20 mmol), anhydrous sodium acetate (4.91 gm, 60 mmol) was refluxed

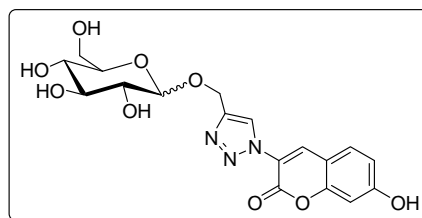


for 4 h. The resulting suspension was poured into ice to give a yellow precipitate which was subsequently filtered and washed with ice-water (2X). Further, it was refluxed in a solution of conc. HCl and ethanol (2:1, 30 mL) for 1 h and then diluted with ice-water (40 mL). The

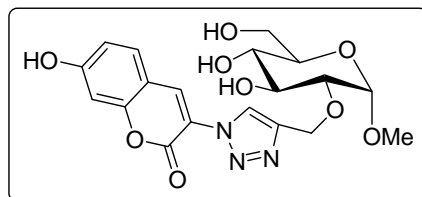
solution was cooled in an ice bath and sodium nitrite (2.75 gm, 40 mmol.) was added in portions and after which the reaction was allowed to stir for 15 min. at 0°C and then sodium azide (3.9 gm, 60 mmol) was added to the stirred solution maintaining the temperature of the reaction. After stirring for another 15 min., the resulting precipitate was filtered off and washed with water, dried and purified over silica gel chromatography to afford a brown solid (2.2 gm, 54%). The NMR characterization data was found to be in conformity with the reported values.^{31a}

General procedure for the synthesis of Triazoles (19a-19e): To an ethanol-water (5 mL; 1:1) solution of modified sugars **1a** (1 mol) was added coumarinyl azide (1 mol) and stirred at room temperature. Freshly prepared solution of CuSO₄·5H₂O (0.1 mol %) and sodium ascorbate (0.5 mol%) was added and the reaction mixture was stirred at room temperature for 30 min after which the TLC examination of the reaction mixture revealed formation of the fluorescent triazoles. We were able to characterize these compounds by ¹H NMR spectra only and not by ¹³C NMR spectral studies as we observed that these compounds have greater propensity to form gels at higher concentrations. These compounds were directly analysed for their UV (Figure 21) and fluorescence spectra.

Compound 19a: ¹H NMR (200 MHz, CD₃OD): δ 3.18-3.46 (m, 6H), 3.62-3.85 (m, 5H), 4.75 (m, 2H), 4.97 (m, 1H), 6.80 (d, 1H, *J* = 1.89 Hz), 6.89 (dd, 1H, *J* = 1.90, 8.57 Hz), 7.64 (d, 1H, *J* = 8.61 Hz), 8.49 (s, 1H), 8.63 (s, 1H). Anal. Calcd. for C₁₈H₁₉N₃O₉: C, 51.31; H, 4.55; N, 9.97. Found: C, 51.25; H, 4.63; N, 10.09.

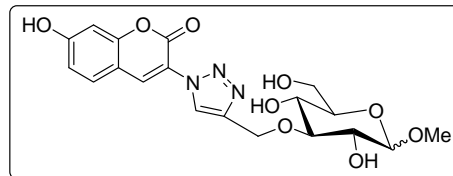


Compound 19b: ¹H NMR (200 MHz, CD₃OD): δ 3.29-3.41 (m, 7H), 3.44-3.55 (m, 2H), 3.63- 3.86 (m, 5H), 4.83 (m, 2H), 6.78 (d, 1H, *J* = 2.13 Hz), 6.87 (dd, 1H, *J* = 2.28, 8.46 Hz), 7.61 (d, 1H, *J* = 8.59 Hz), 8.48 (s, 1H), 8.59 (s, 1H). Anal. Calcd. for C₁₉H₂₁N₃O₉: C, 52.41; H, 4.86; N, 9.65. Found C, 52.33; H, 4.96; N, 9.73.

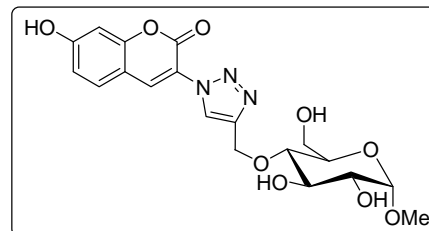


Compound 19c: ¹H NMR (200 MHz, CD₃OD): δ 3.29-3.35 (m, 3H), 3.42 (s, 3H), 3.53-4.24 (m, 7H), 4.68-4.78 (m, 1H), 5.03 (m, 2H), 6.77 (d, 1H, *J* = 2.14 Hz), 6.86 (dd, 1H, *J* = 2.27,

8.56 Hz), 7.61 (d, 1H, $J = 8.61$ Hz), 8.46 (s, 1H), 8.58 (s, 1H). Anal. Calcd. for $C_{19}H_{21}N_3O_9$: C, 52.41; H, 4.86; N, 9.65. Found C, 52.33; H, 4.96; N, 9.73.



Compound 19d: 1H NMR (200 MHz, CD_3OD): δ 3.29-3.32 (m, 2H), 3.35 (s, 2H), 3.39-3.47 (m, 6H), 3.51-3.86 (m, 5H), 4.67 (m, 1H), 5.07 (m, 1H), 6.81 (d, 1H, $J = 2.15$ Hz), 6.86 (dd, 1H, $J = 2.17, 8.60$ Hz), 7.63 (d, 1H, $J = 8.60$ Hz), 8.49 (s, 1H), 8.57 (s, 1H). Anal. Calcd. for $C_{19}H_{21}N_3O_9$: C, 52.41; H, 4.86; N, 9.65. Found C, 52.31; H, 4.94; N, 9.71.



Compound 19e: 1H NMR (200 MHz, CD_3OD): δ 3.29-3.43 (m, 7H), 3.51 (s, 1H), 3.57-3.87 (m, 4H), 4.04-4.20 (m, 1H), 4.67 (d, 1H, $J = 3.68$ Hz), 4.75 (m, 2H), 6.78 (d, 1H, $J = 2.15$ Hz), 6.87 (dd, 1H, $J = 2.15, 8.60$ Hz), 7.62 (d, 1H, $J = 8.60$ Hz), 8.47 (s, 1H), 8.57 (s, 1H). Anal. Calcd. for $C_{19}H_{21}N_3O_9$: C, 52.41; H, 4.86; N, 9.65. Found C, 52.61; H, 4.93; N, 9.51.

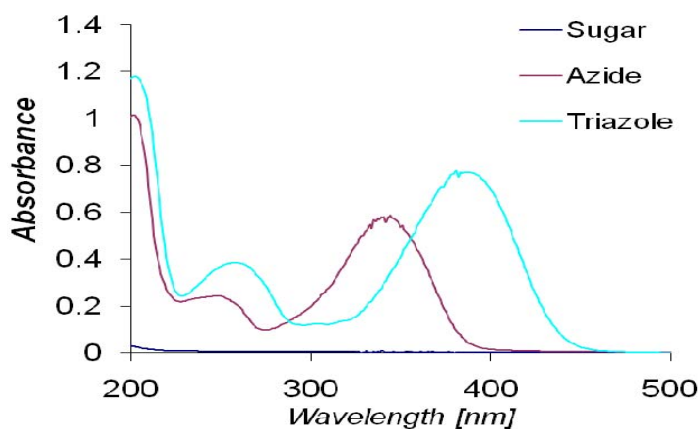
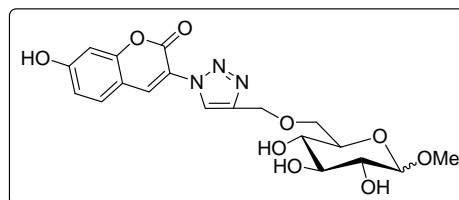


Figure 21. UV absorption spectra of sugar **1b**, triazole **19b** and azidocoumarin

Preparation of the stock solutions:

Sugar Stock Solution (1a-1e): Individual solutions of compound **1a-1e** were prepared in 27mg/ml in dist. water and then sterilized *via* autoclave.

3-azidocoumarin stock: 5 mg of the coumarinyl azide was dissolved in 2 ml of ethanol and used as such for the experiments.

100 mmol and 100 mmol stock solutions of copper sulphate and sodium ascorbate were prepared.

Composition of LB and Mineral Medium:

Luria Broth: Tryptone (1 gm), Yeast extract (0.5 gm), NaCl (1gm) were dissolved in 100 ml of dist. water and sterilised *via* autoclave.

Mineral Medium: Glucose (5 gm), Na₂HPO₄ (6 gm), KH₂PO₄ (3 gm), NH₄Cl (1 gm), NaCl (0.5 gm), MgSO₄ (0.12 gm), CaCl₂ (0.01 gm) were dissolved in 1000ml dist. H₂O and used after sterilisation.

General procedure for cell culturing:

Two sets of cultures were separately incubated simultaneously each time for all the cell variants employed. Firstly, the control which was devoid of the modified sugars and secondly, the cell cultures which were grown in the presence of modified sugar derivatives **1a-1e**, 40 µL/mL of the propargylated monosaccharides were added individually to all the tubes. However, after the cells have attained their maximum growth, both the sets were treated likewise separately. The general procedure for the cell culturing is given hereunder for all the cells used.

Escherichia coli: BL21 cells were inoculated into the 5 ml of LB medium. The cell culture was then incubated at 37°C and shaken at 180 rpm for 12 h. Cells were then isolated by centrifugation at 8000 rpm for 5 min and subsequently were washed with PBS buffer pH 7.8 (2x) and resuspended in 1 mL of PBS buffer (pH 7.8). The cells were subsequently cooled in ice-bath before being sonicated for 2 min (pulse 2 sec). The cell lysate was isolated by centrifugation of the resulting mixture at 8000 rpm for 5 min. The cell lysate was used as such for UV and Fluorescence spectroscopic analysis.

Lactobacillus lacti: *Lactobacillus* cells were inoculated into the 5 mL mineral medium. The culture was then incubated at 30°C without shaking for 24 h. Cells were then isolated by centrifugation at 8000 rpm for 5 min and subsequently were washed with PBS buffer pH 7.8 (2X) and resuspended in 1 mL of PBS buffer (pH 7.8). The cells were cooled in ice-bath and sonicated for 2 min (pulse 2 sec). The cell lysate was isolated by centrifugation of the resulting mixture at 8000 rpm for 5 min. The cell lysate was used as such for UV and Fluorescence spectroscopic analysis.

Sacchromyces (Wild Type): *Sacchromyces* cells were inoculated into the 5 mL LB medium. The medium was shaken at 37°C for 12 hr at 180 rpm and cells were harvested by centrifugation at 8000 rpm for 5 min. The cells were resuspended in 1ml PBS buffer pH 7.8

before they were washed two times with 2 mL of PBS buffer pH 7.8 and subsequently sonicated in an ice-bath to yield the cell lysate after centrifugation at 8000 rpm for 5 min which was used as such for UV and Fluorescence spectroscopic analysis.

Recording of UV and Fluorescence Spectra:

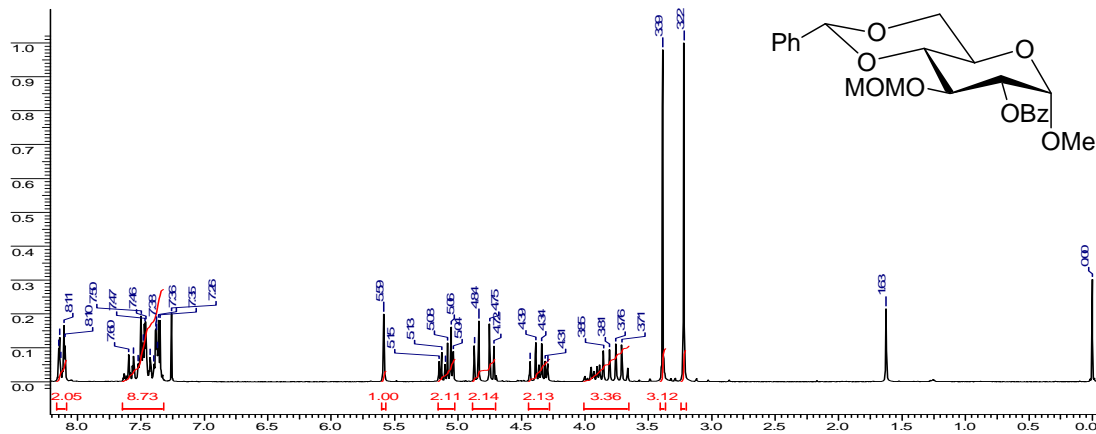
Cell lysates were treated with 5 µl of the above prepared azide stock solution and 5 µl of CuSO₄ and 10 µl of sodium ascorbate. The resulting solution was kept at room temperature for 30 min and then utilized for UV and Florescence studies.

Preparation of Slides:

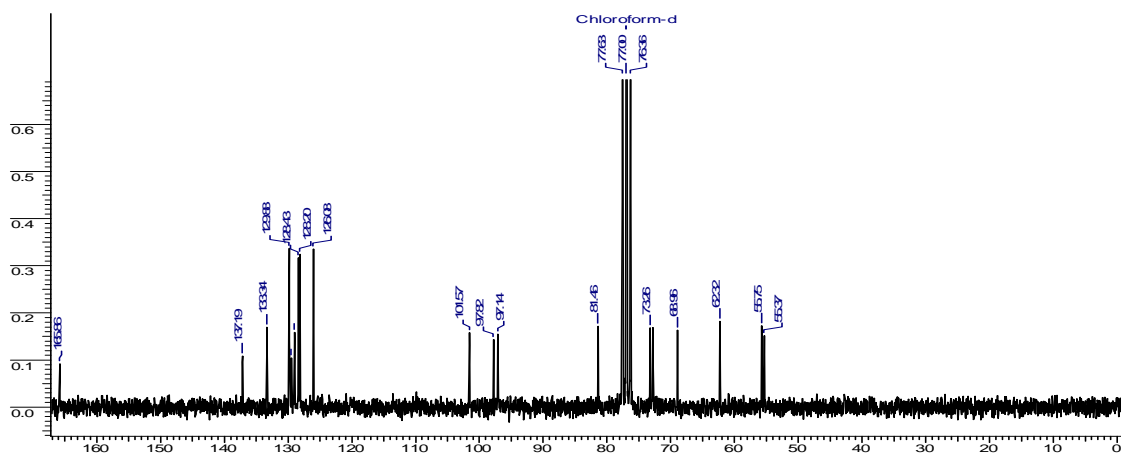
In a typical experiment, growing cells ($OD_{600} < 1.0$) were washed with 1 mL of PBS buffer (pH 7.8) and treated with 2 µl of azide stock solution and 1 µl and 2 µl of CuSO₄ and sodium ascorbate stock solutions. Cells were then incubated for 30 min at rt and then a drop of the uniformly suspended cells was loaded onto a glass slide and kept at room temperature for air drying, after which the cells were analysed under the fluorecence microscope and the images recorded.

Chapter 2: Spectral Charts

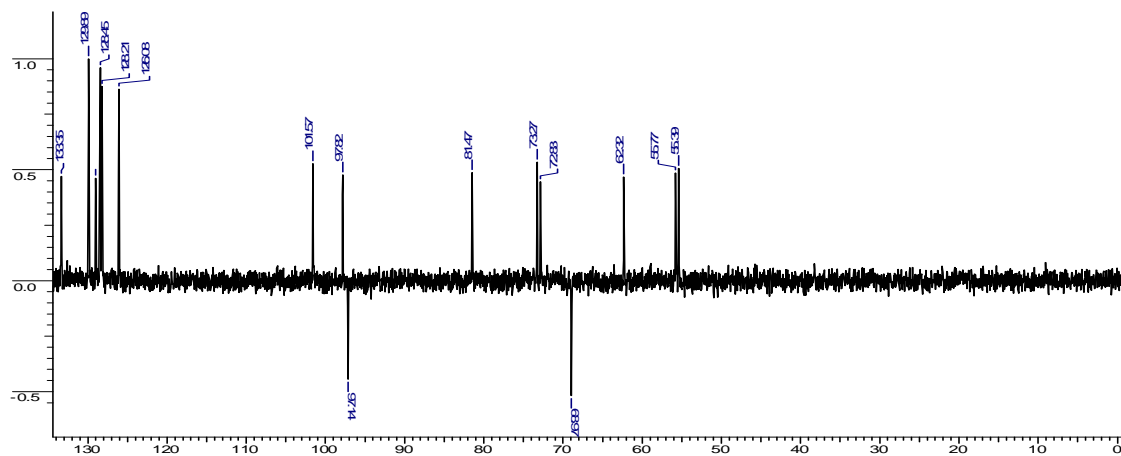
^1H NMR (200 MHz, CDCl_3) of Compound 4



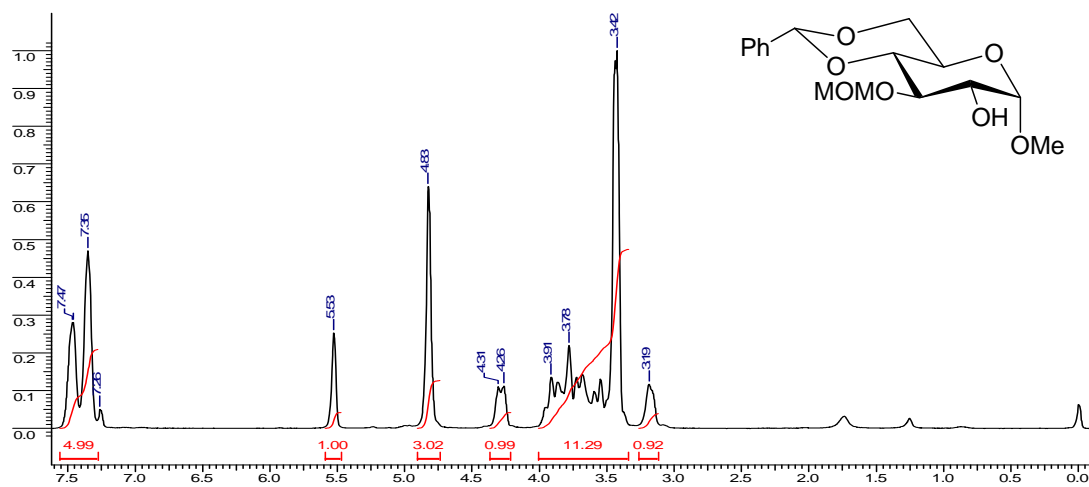
^{13}C NMR (75 MHz, CDCl_3) of Compound 4



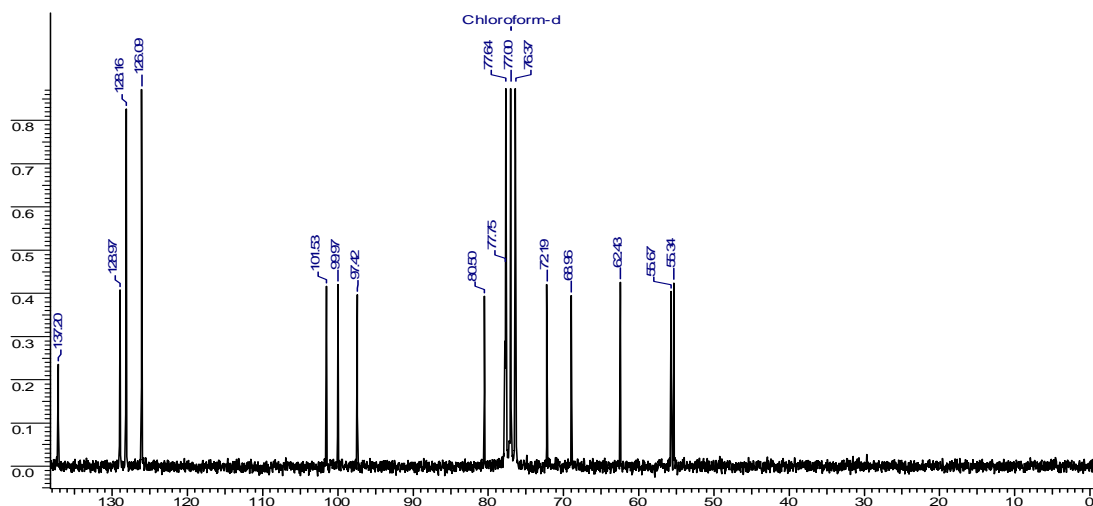
DEPT NMR (75 MHz, CDCl_3) of Compound 4



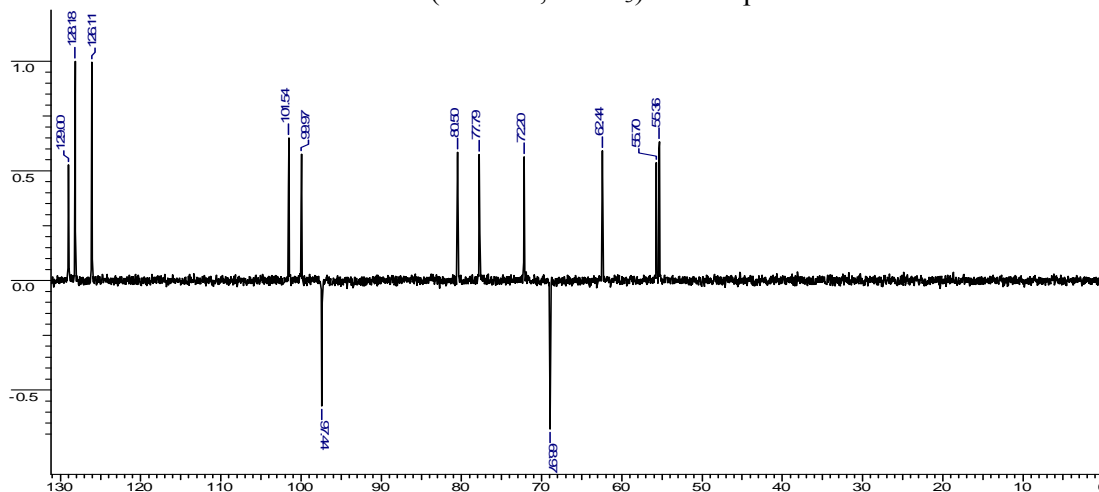
¹H NMR (200 MHz, CDCl₃) of Compound 5



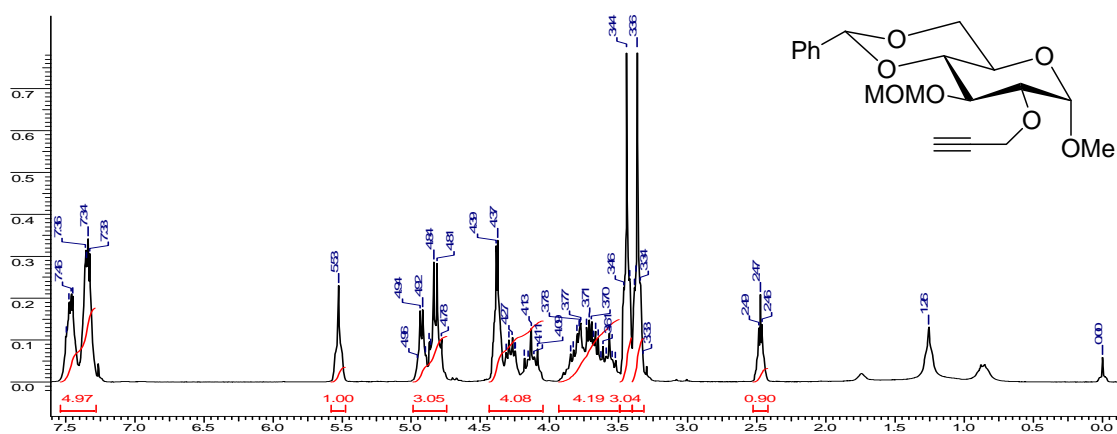
¹³C NMR (75 MHz, CDCl₃) of Compound 5



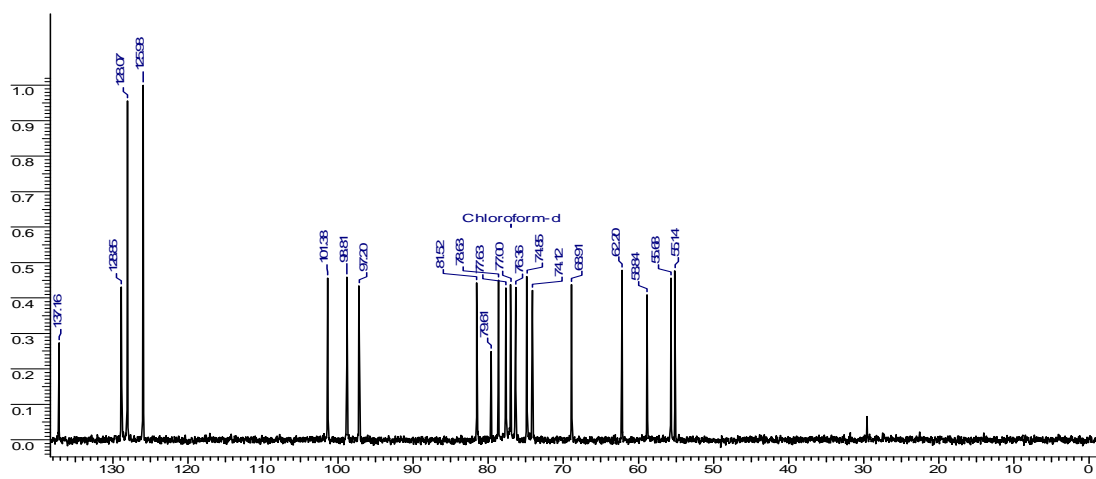
DEPT NMR (75 MHz, CDCl₃) of Compound 5



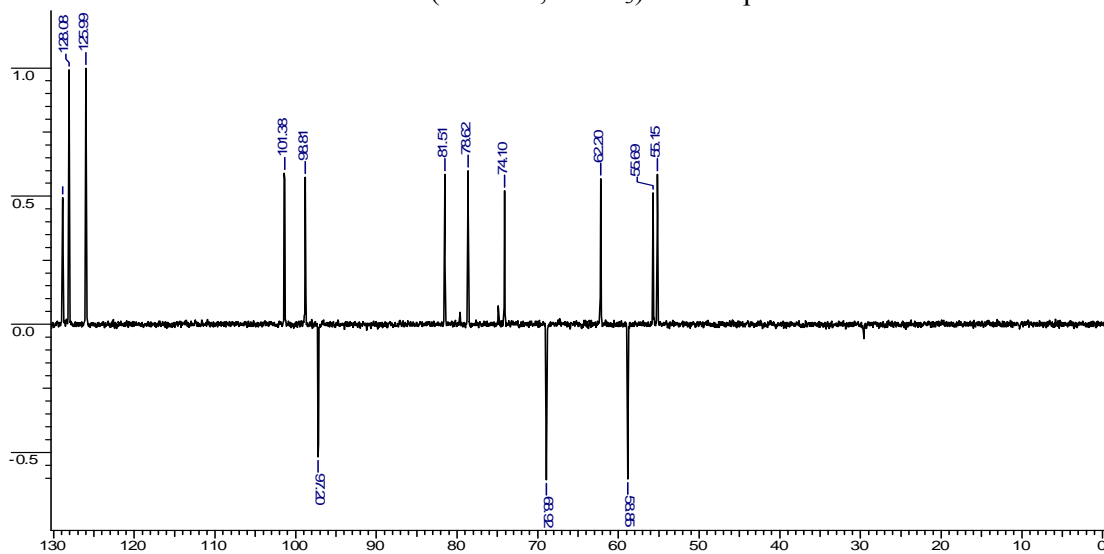
^1H NMR (200 MHz, CDCl_3) of Compound 6



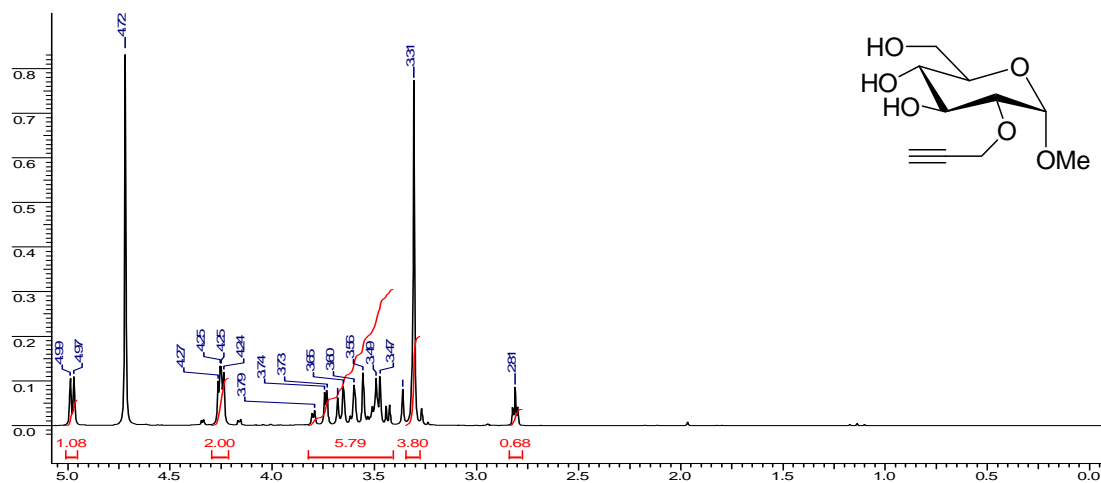
^{13}C NMR (75 MHz, CDCl_3) of Compound 6



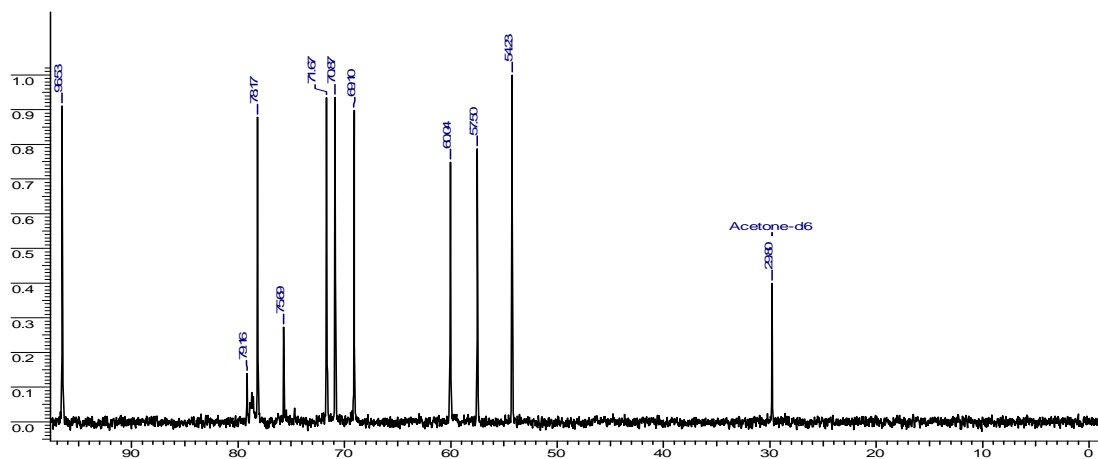
DEPT NMR (75 MHz, CDCl_3) of Compound 6



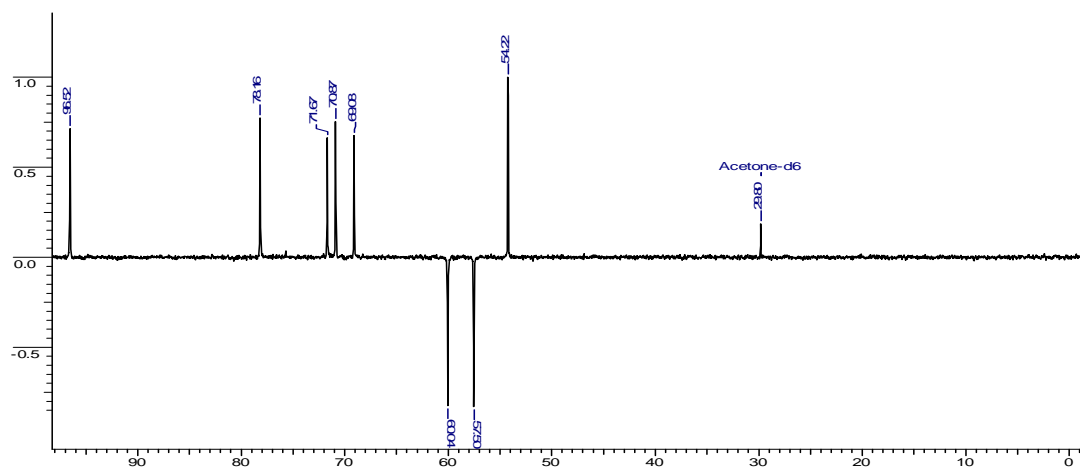
^1H NMR (200 MHz, CD_3OD) of Compound **1b**



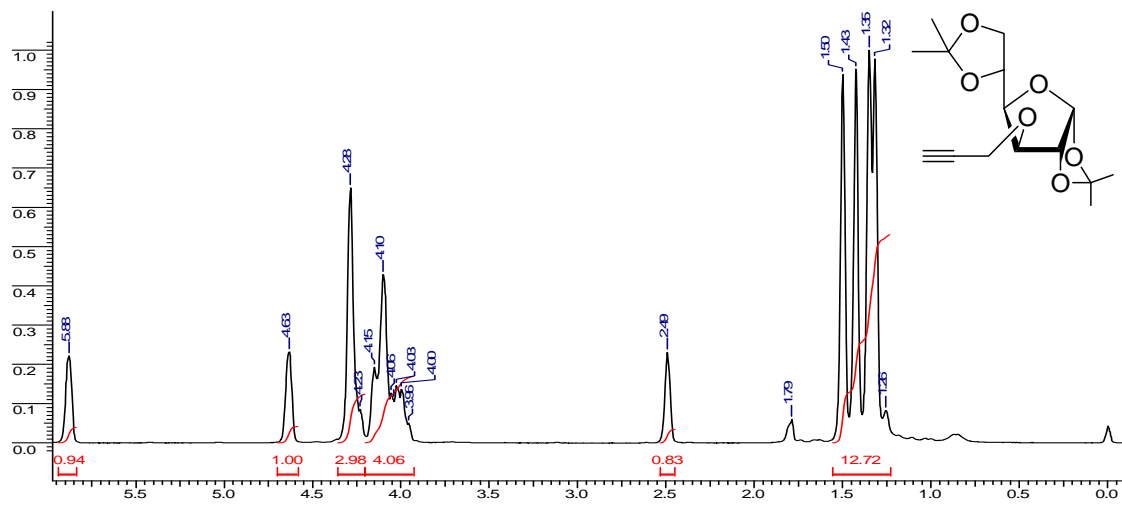
^{13}C NMR (75 MHz, CD_3OD) of Compound **1b**



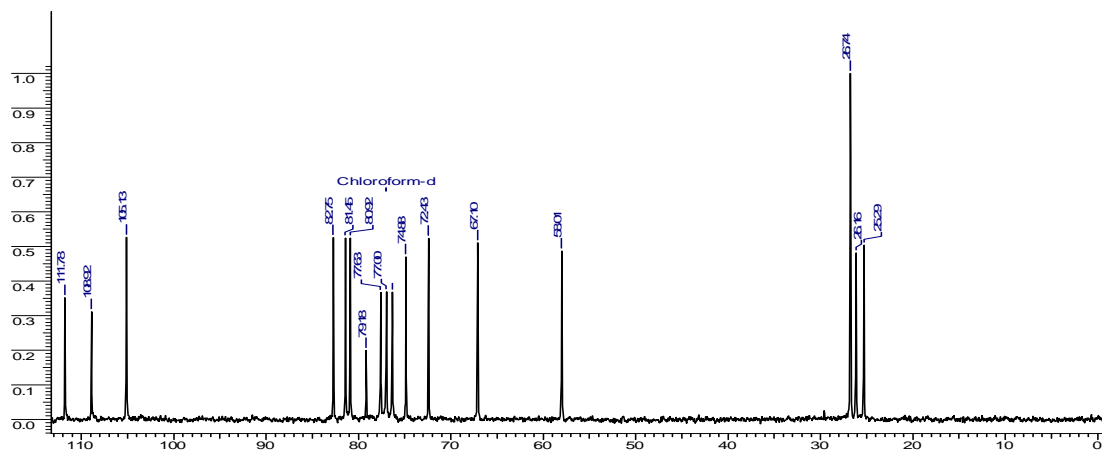
DEPT NMR (75 MHz, CD_3OD) of Compound **1b**



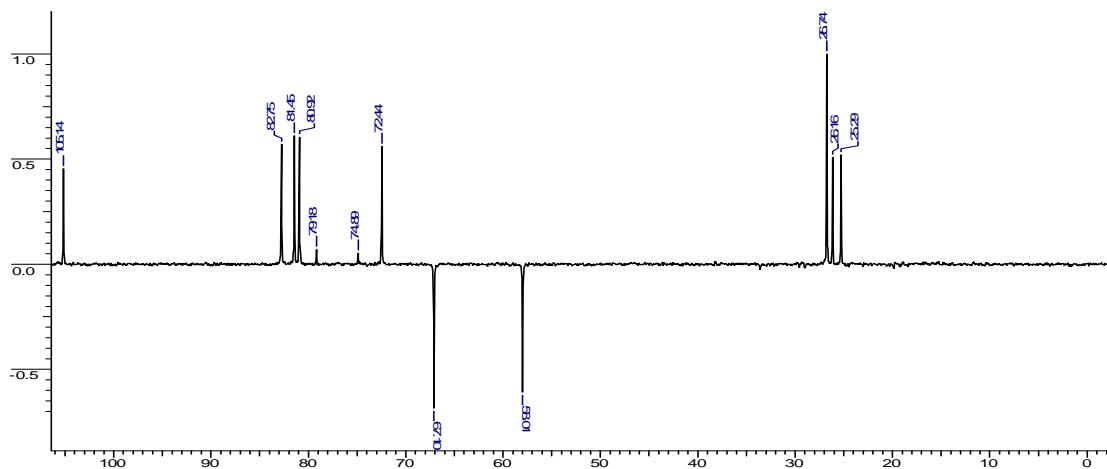
^1H NMR (200 MHz, CDCl_3) of Compound **8**



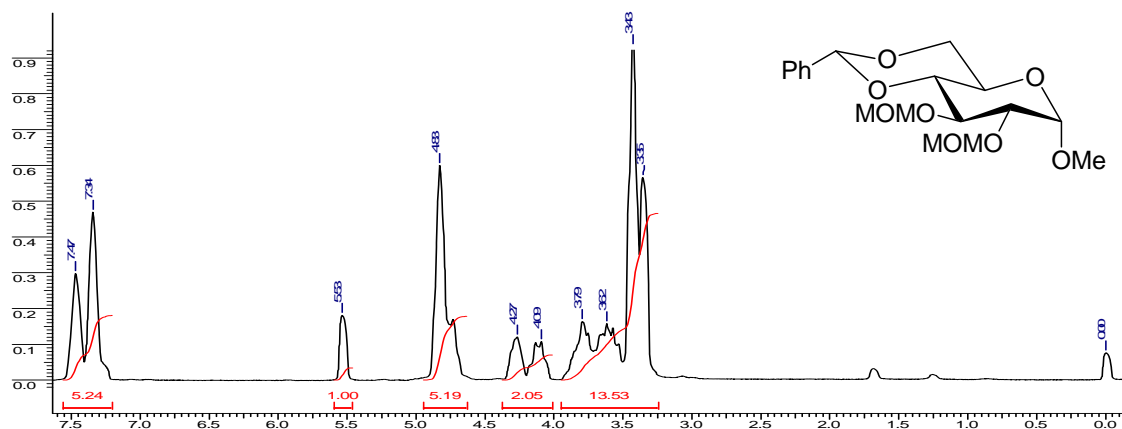
^1H NMR (200 MHz, CDCl_3) of Compound **8**



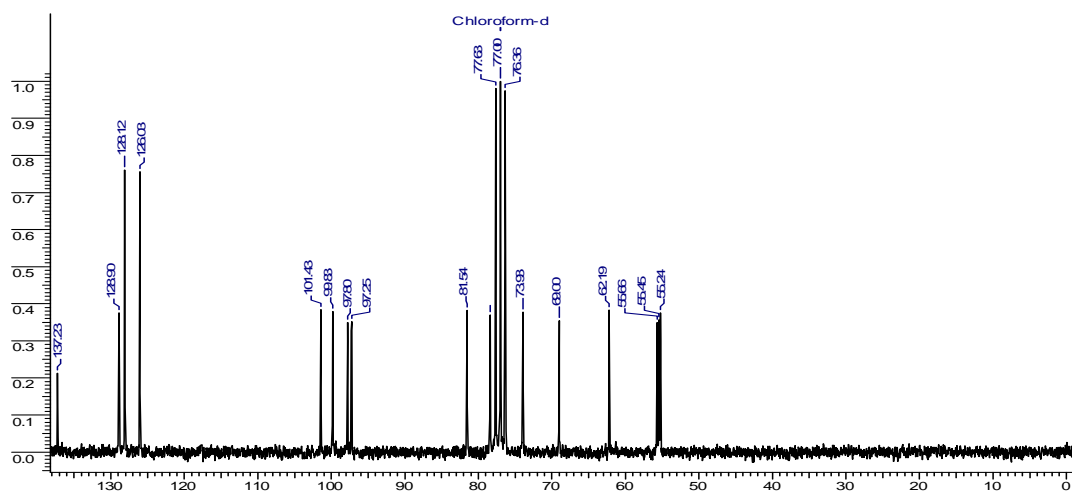
DEPT NMR (75 MHz, CDCl_3) of Compound **8**



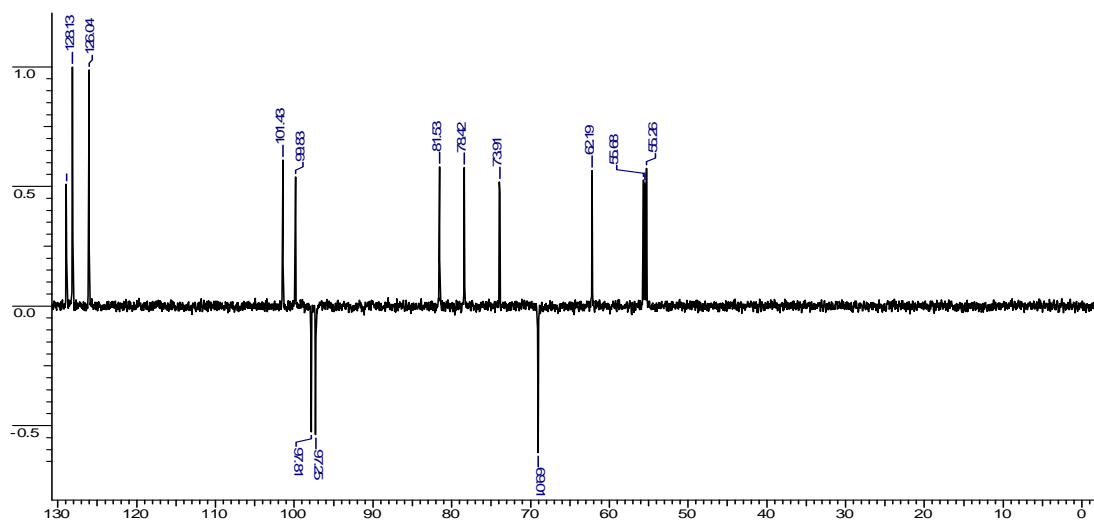
^1H NMR (200 MHz, CDCl_3) of Compound 9



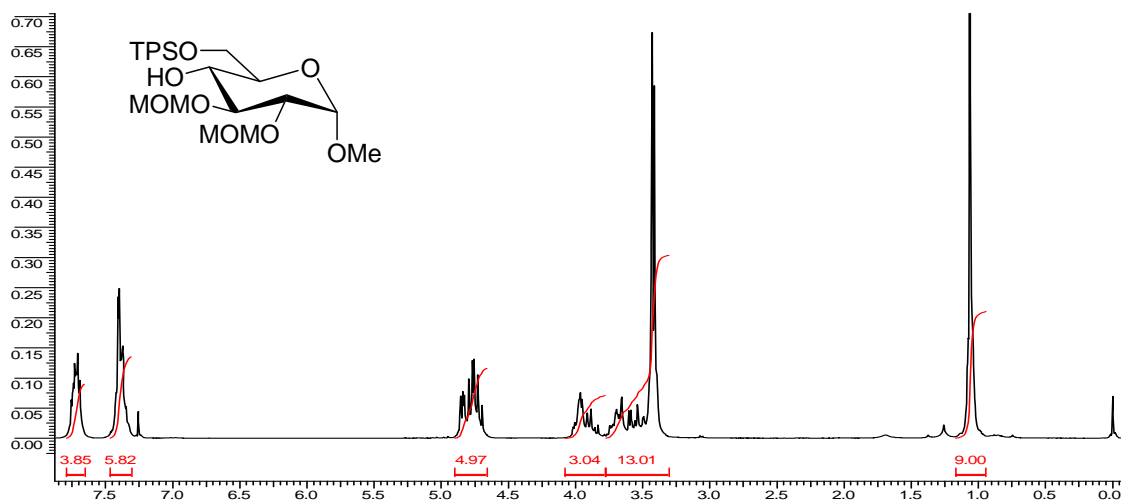
^{13}C NMR (75 MHz, CDCl_3) of Compound 9



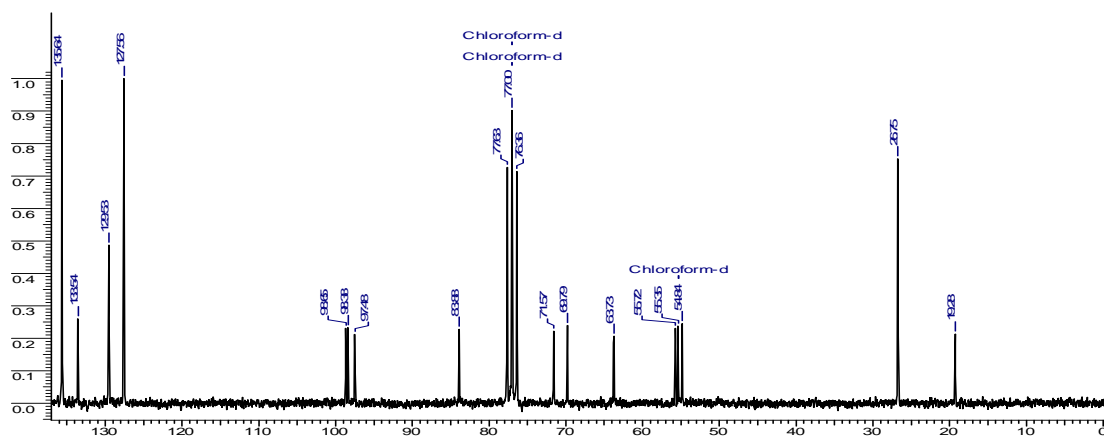
DEPT NMR (75 MHz, CDCl_3) of Compound 9



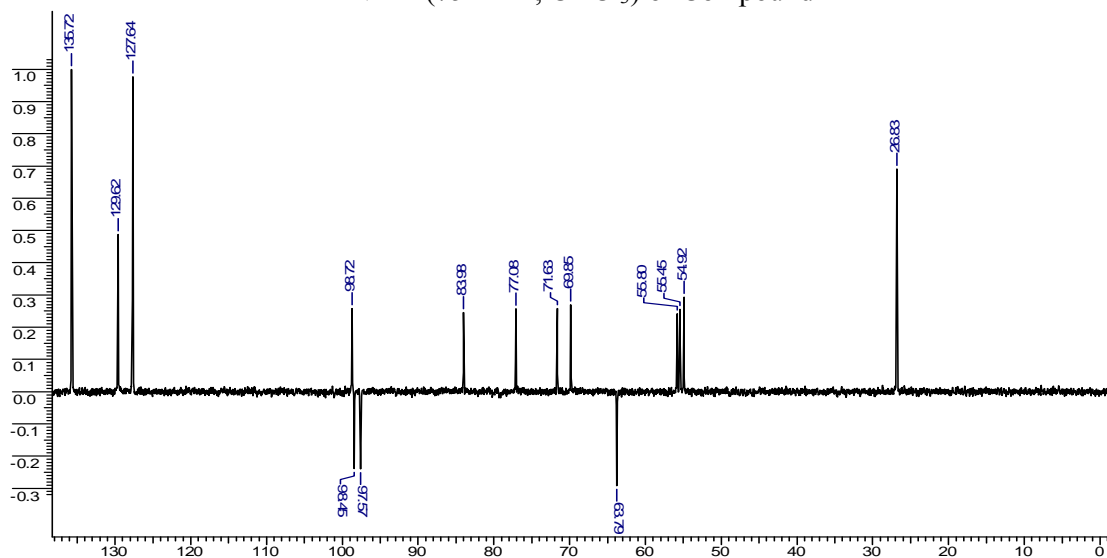
¹H NMR (200 MHz, CDCl₃) of Compound 11



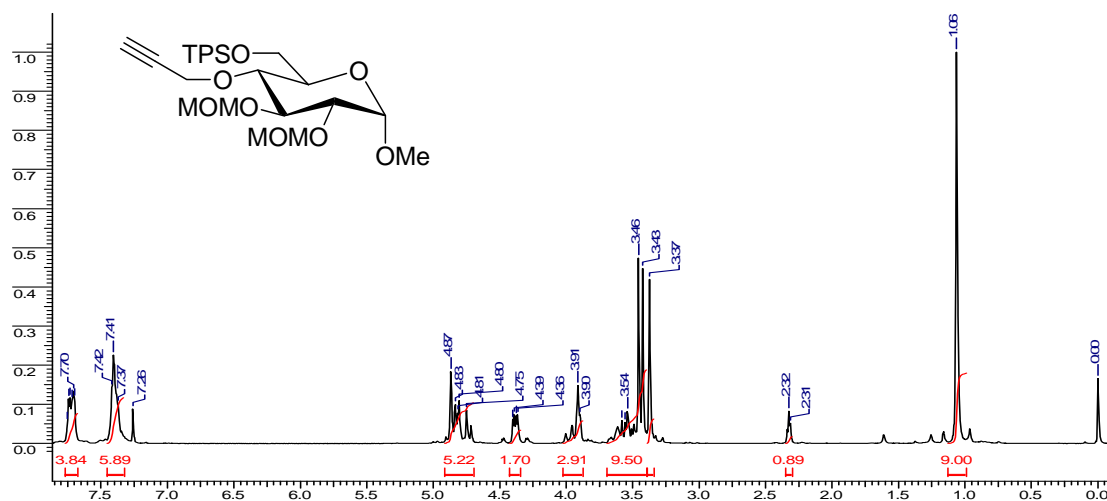
¹³C NMR (75 MHz, CDCl₃) of Compound 11



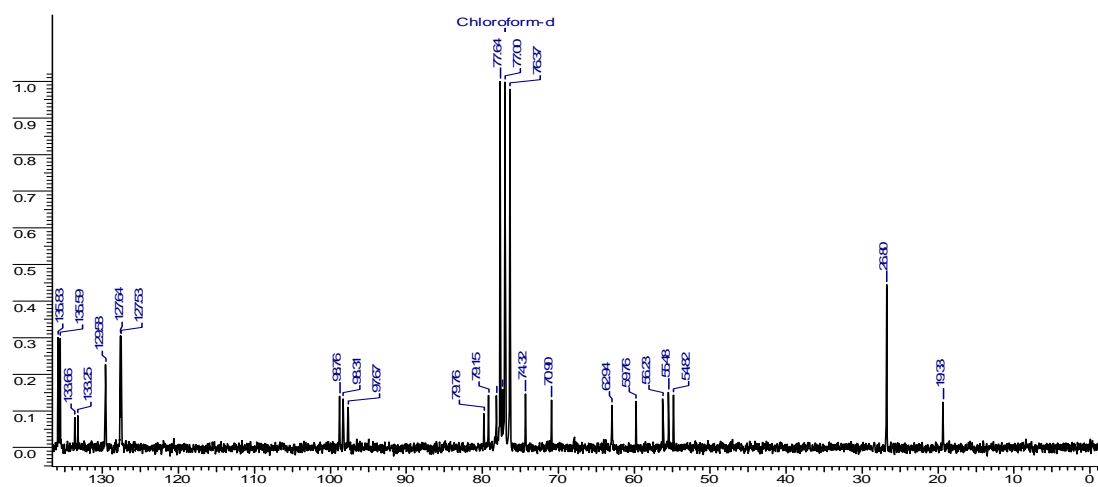
DEPT NMR (75 MHz, CDCl₃) of Compound 11



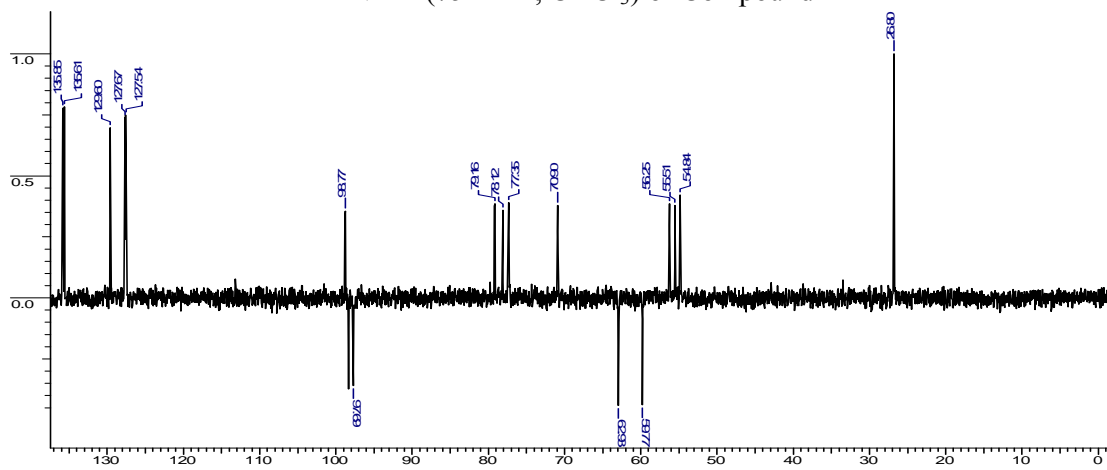
¹H NMR (200 MHz, CDCl₃) of Compound 12



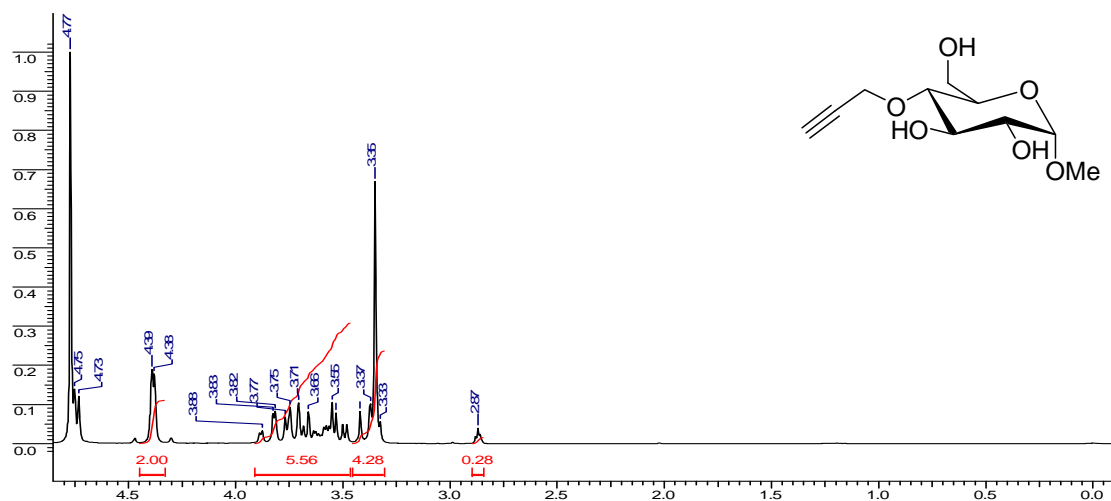
¹³C NMR (75 MHz, CDCl₃) of Compound 12



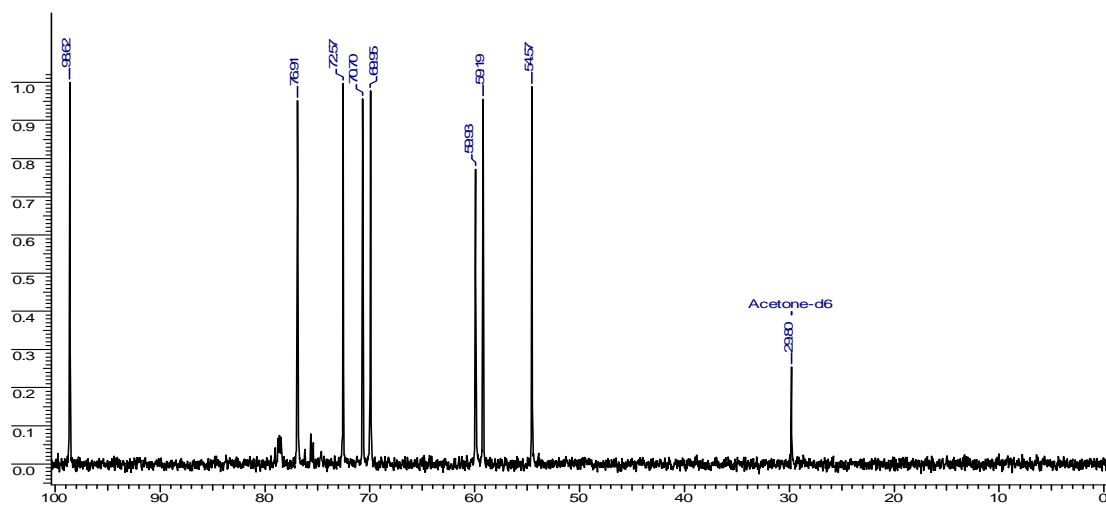
DEPT NMR (75 MHz, CDCl₃) of Compound 12



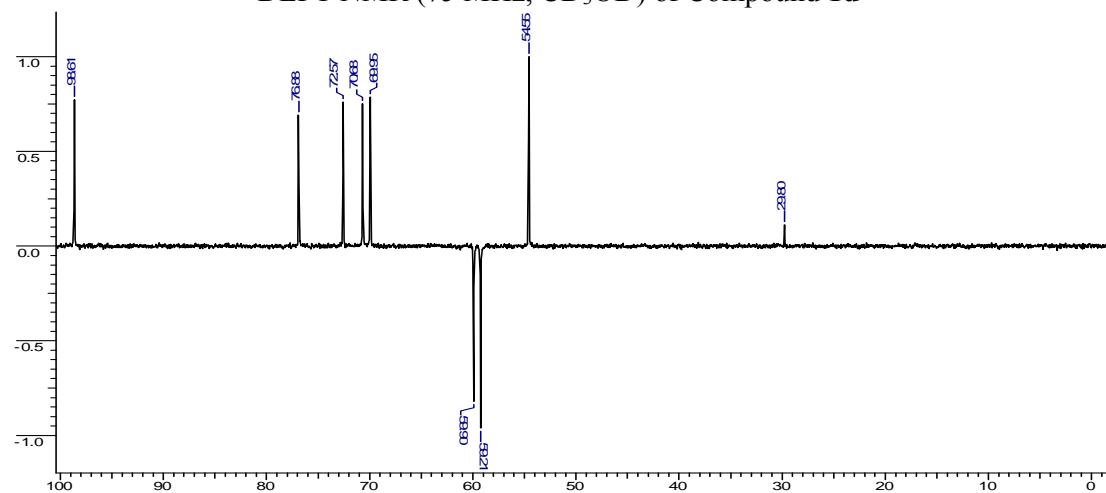
¹H NMR (200 MHz, CD₃OD) of Compound **1d**



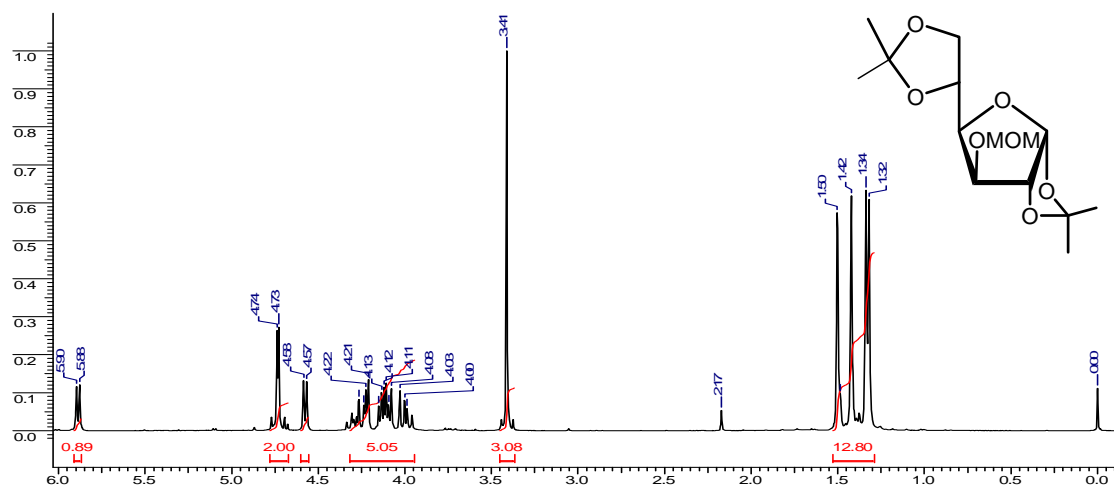
¹³C NMR (75 MHz, CD₃OD) of Compound **1d**



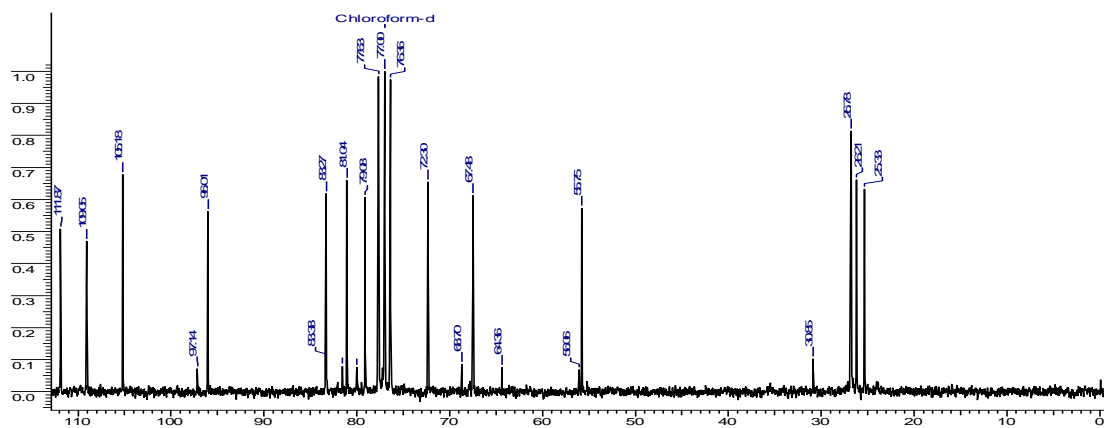
DEPT NMR (75 MHz, CD₃OD) of Compound **1d**



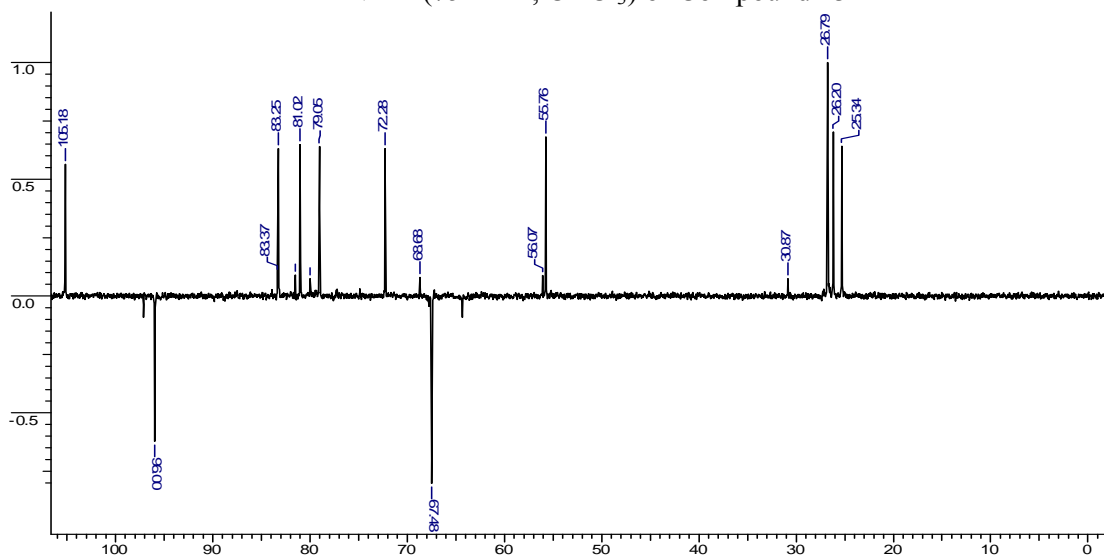
¹H NMR (200 MHz, CDCl₃) of Compound 13



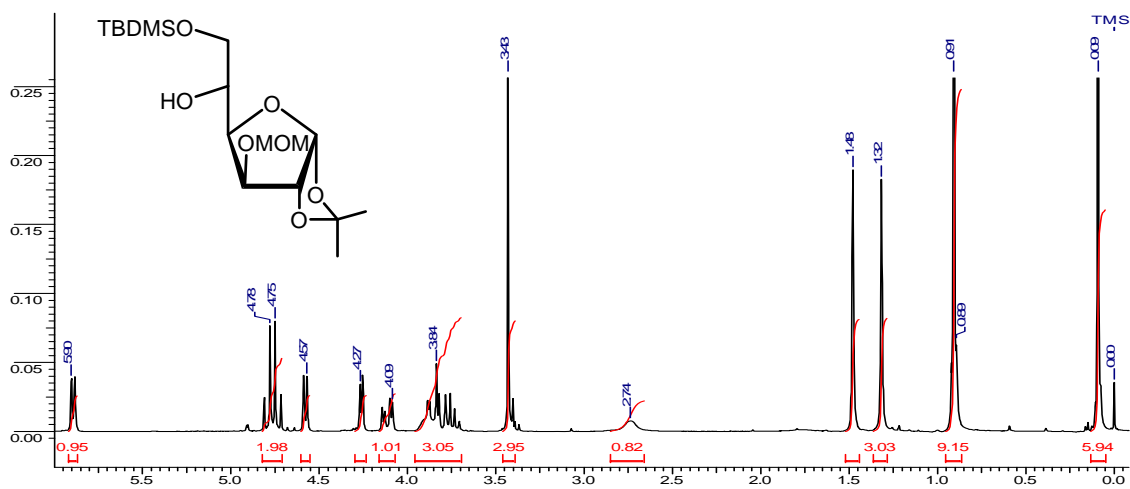
¹³C NMR (75 MHz, CDCl₃) of Compound 13



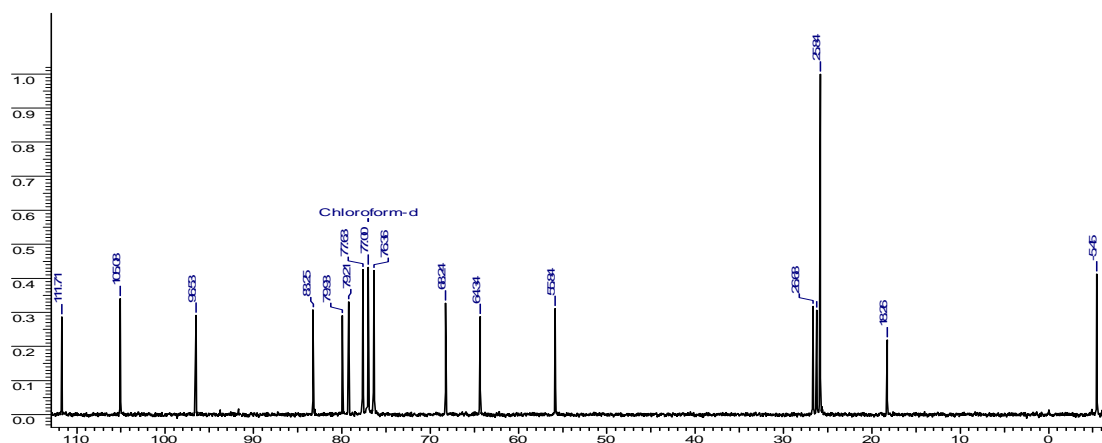
DEPT NMR (75 MHz, CDCl₃) of Compound 13



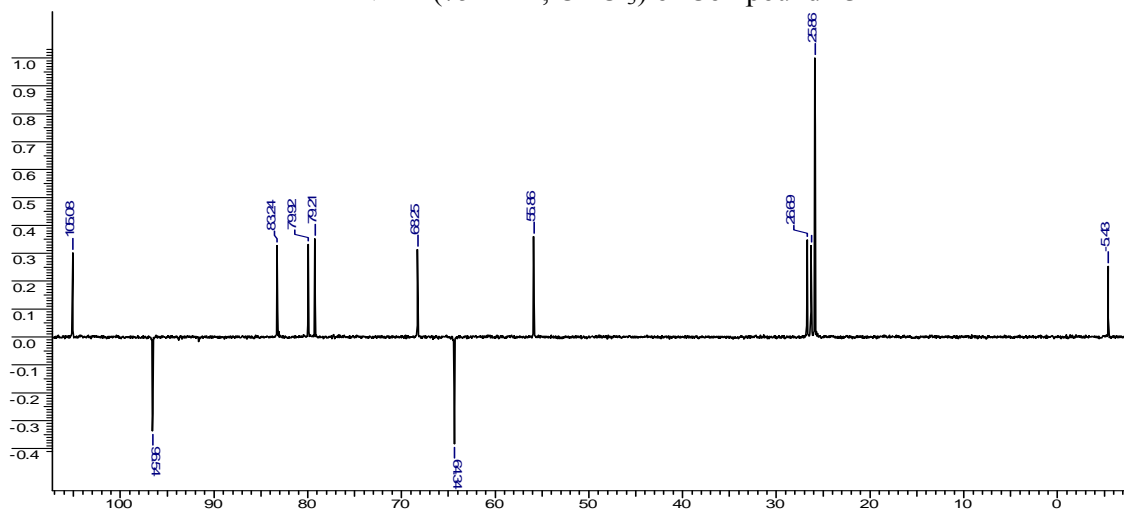
^1H NMR (200 MHz, CDCl_3) of Compound 15



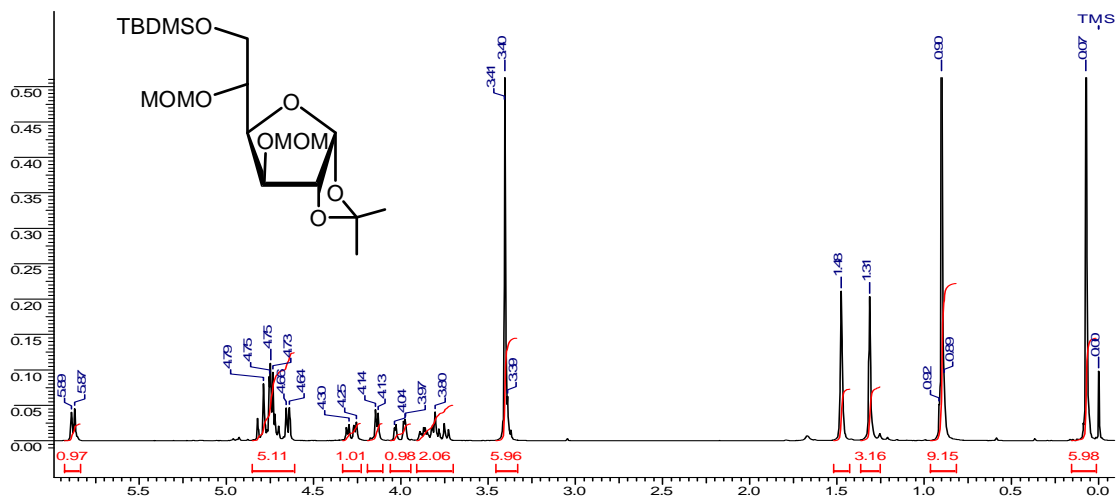
^{13}C NMR (75 MHz, CDCl_3) of Compound 15



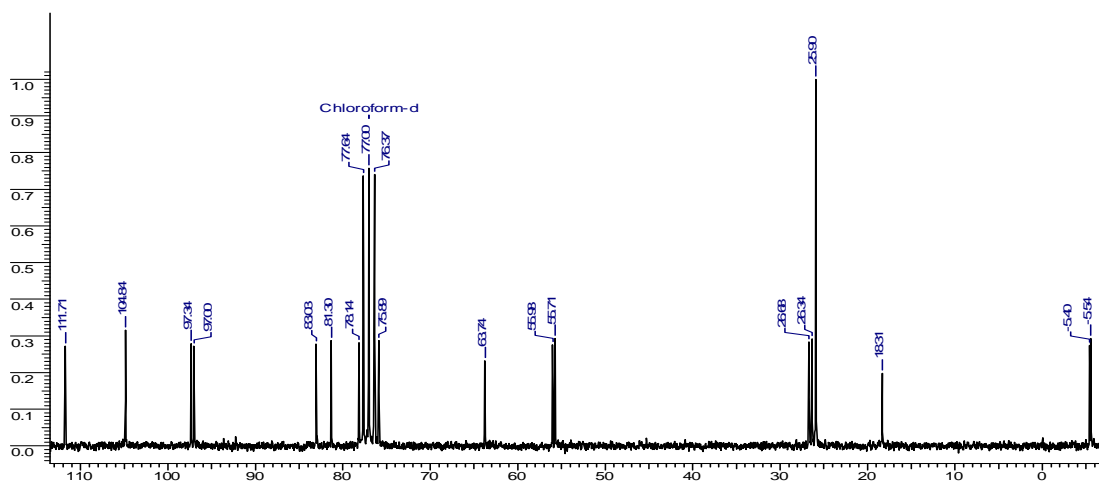
DEPT NMR (75 MHz, CDCl_3) of Compound 15



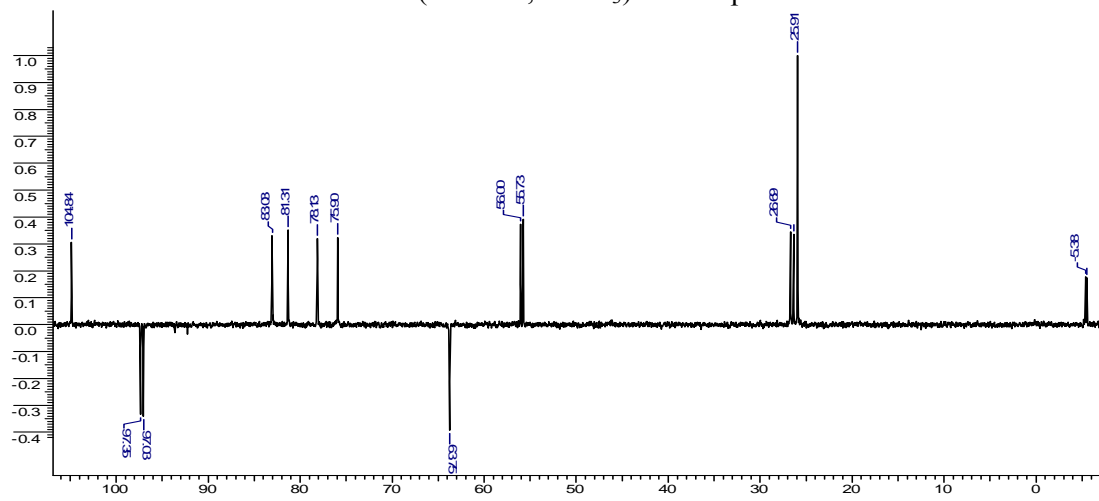
¹H NMR (200 MHz, CDCl₃) of Compound 16



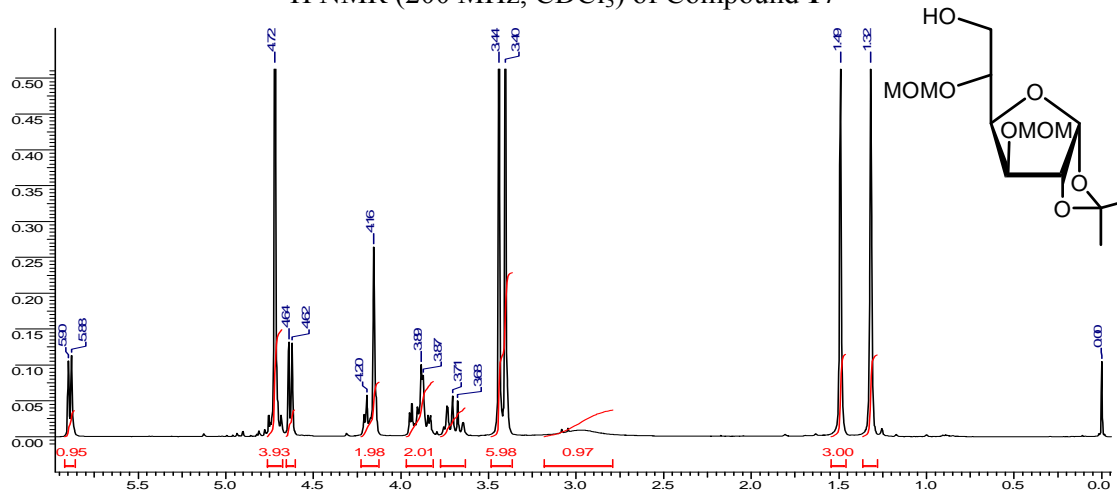
¹³C NMR (75 MHz, CDCl₃) of Compound 16



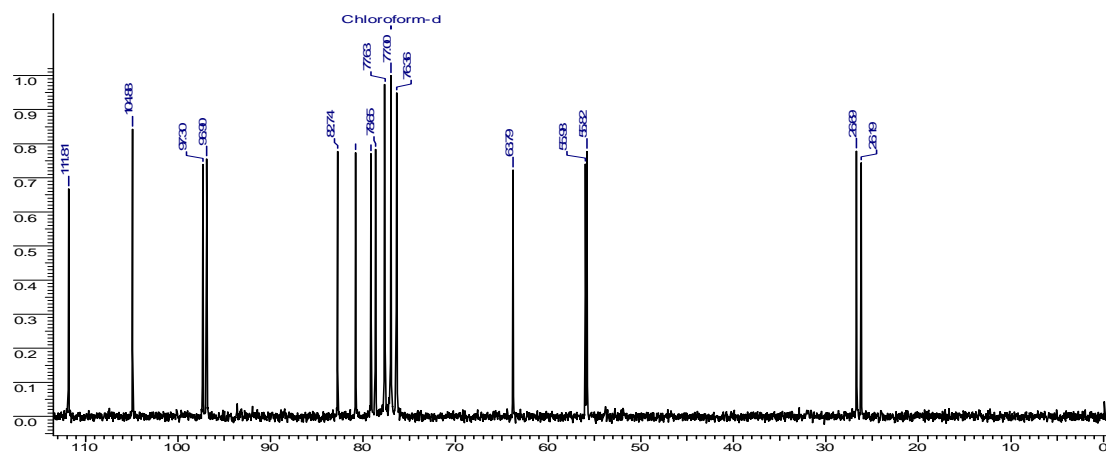
DEPT NMR (75 MHz, CDCl₃) of Compound 16



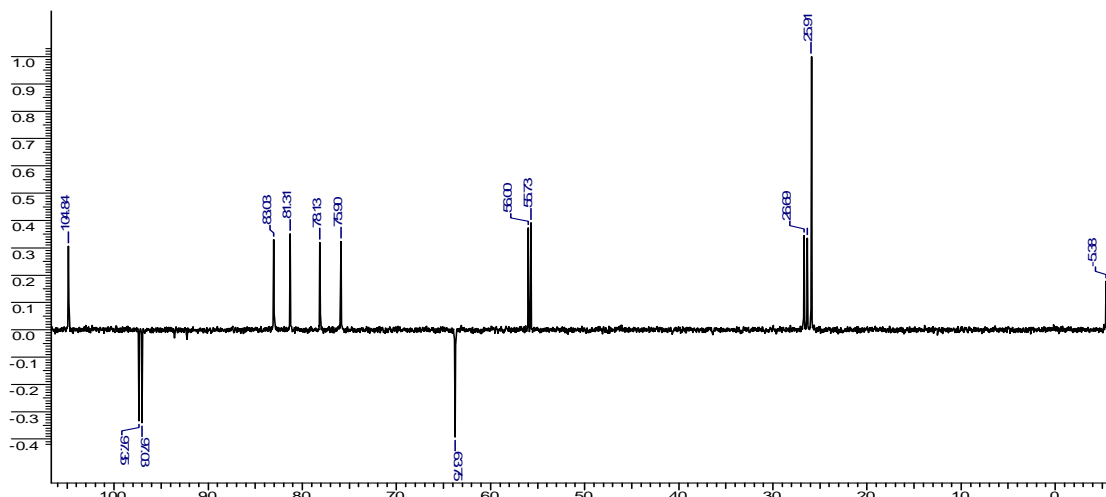
^1H NMR (200 MHz, CDCl_3) of Compound 17



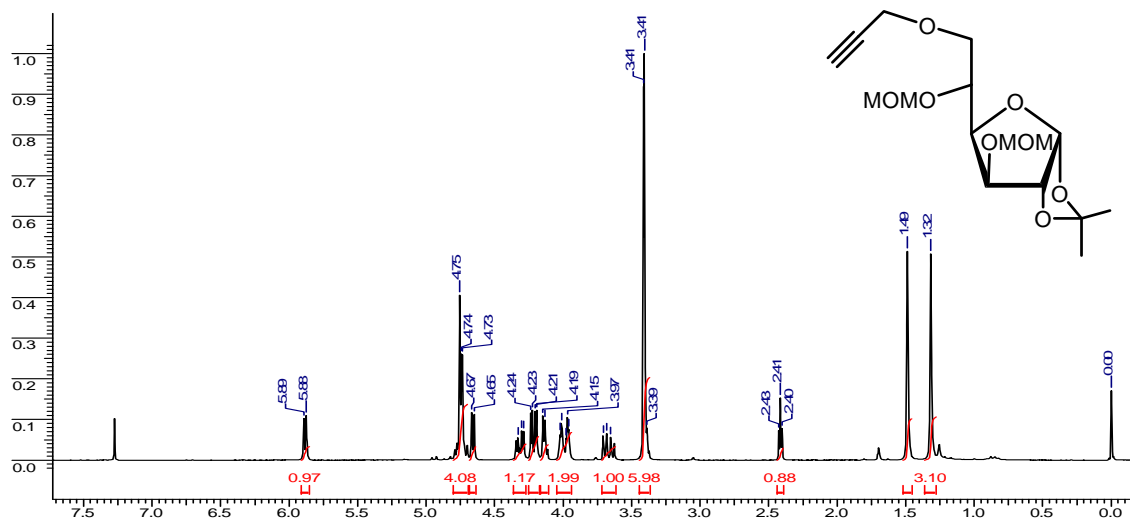
^{13}C NMR (75 MHz, CDCl_3) of Compound 17



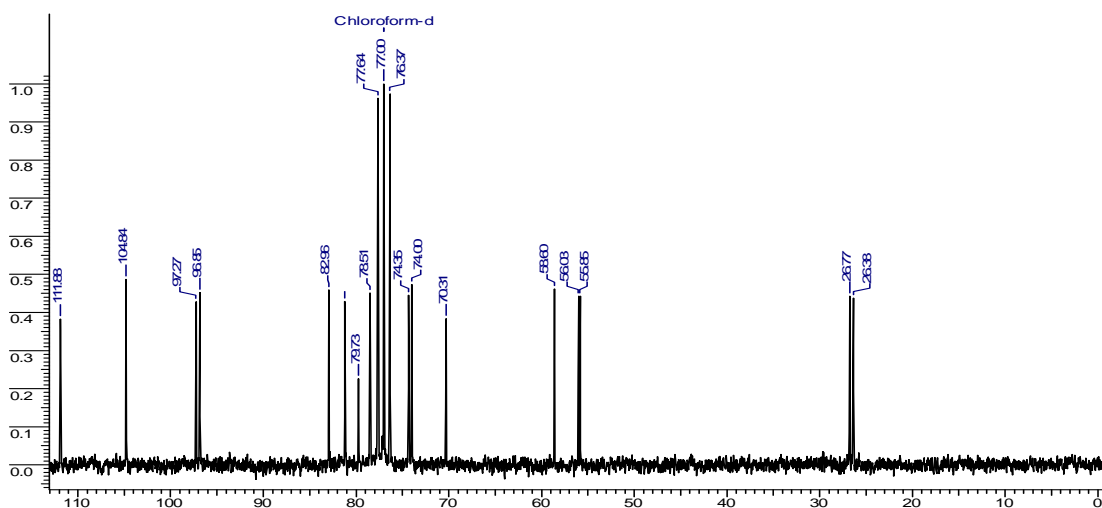
DEPT NMR (75 MHz, CDCl_3) of Compound 17



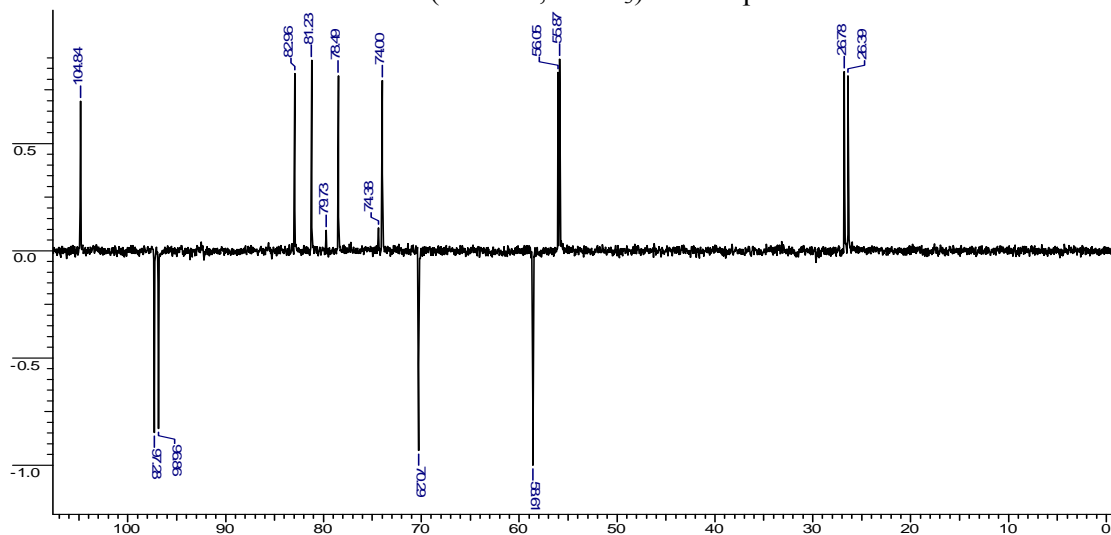
¹H NMR (200 MHz, CDCl₃) of Compound 18



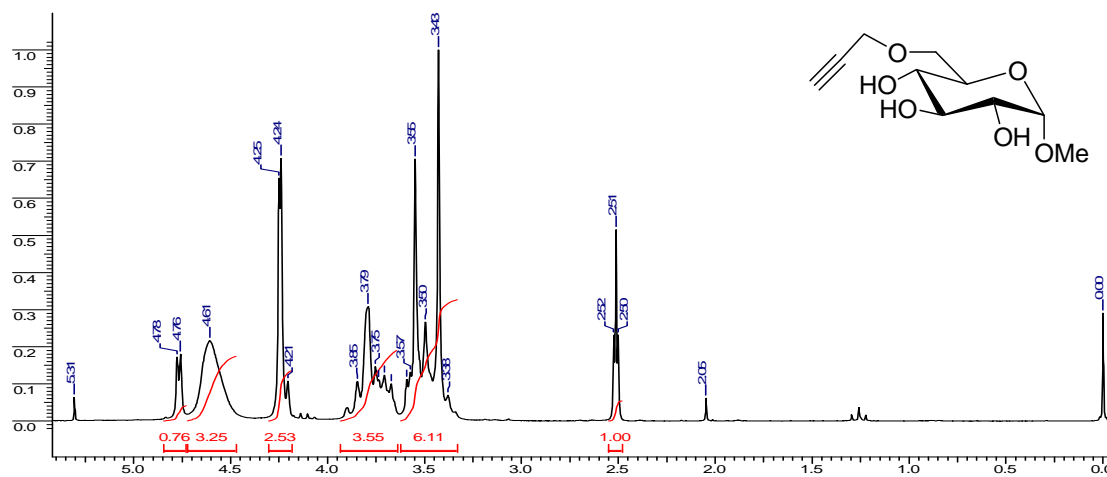
¹³C NMR (75 MHz, CDCl₃) of Compound 18



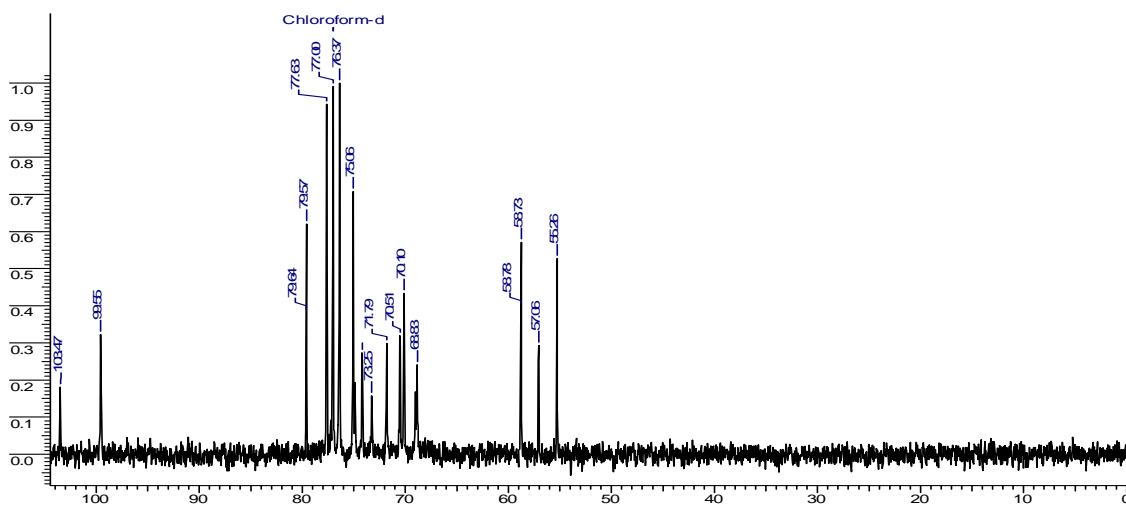
DEPT NMR (75 MHz, CDCl₃) of Compound 18



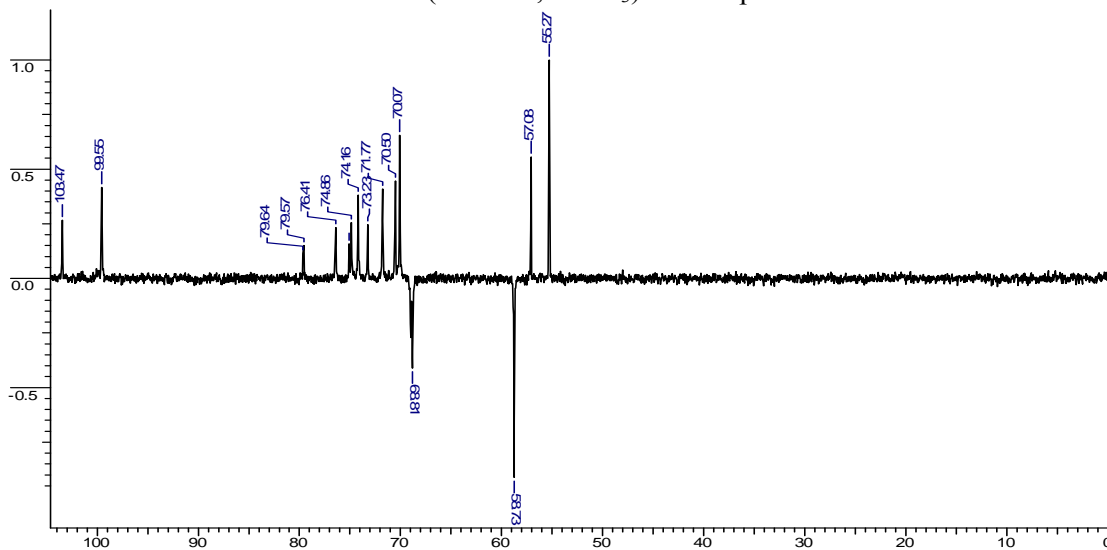
^1H NMR (200 MHz, CDCl_3) of Compound **1e**



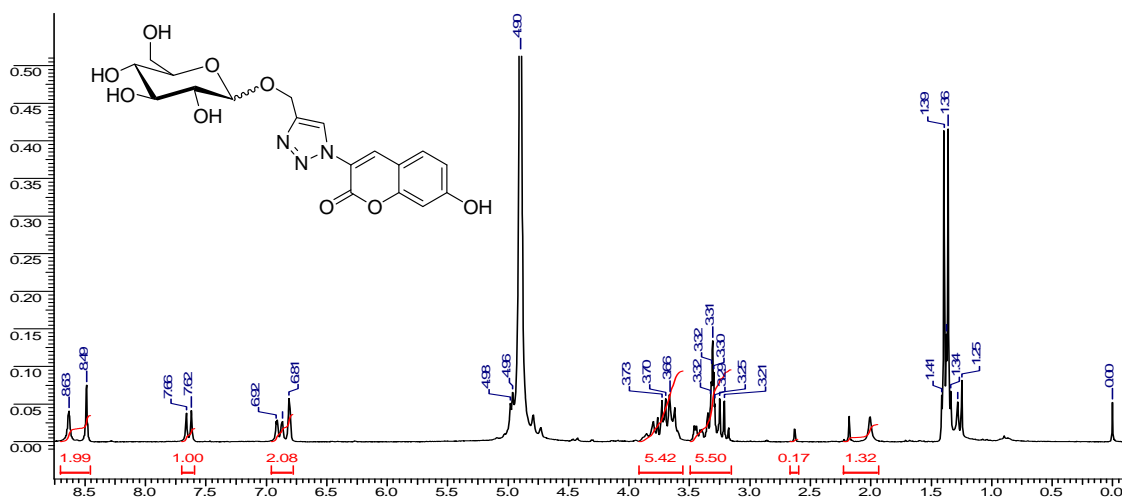
^{13}C NMR (75 MHz, CDCl_3) of Compound **1e**



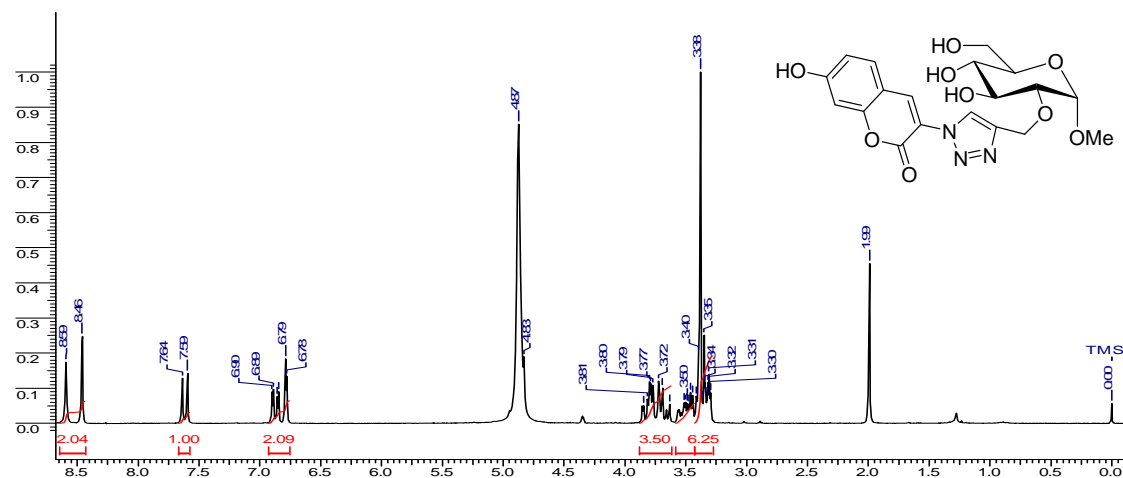
DEPT NMR (75 MHz, CDCl_3) of Compound **1e**



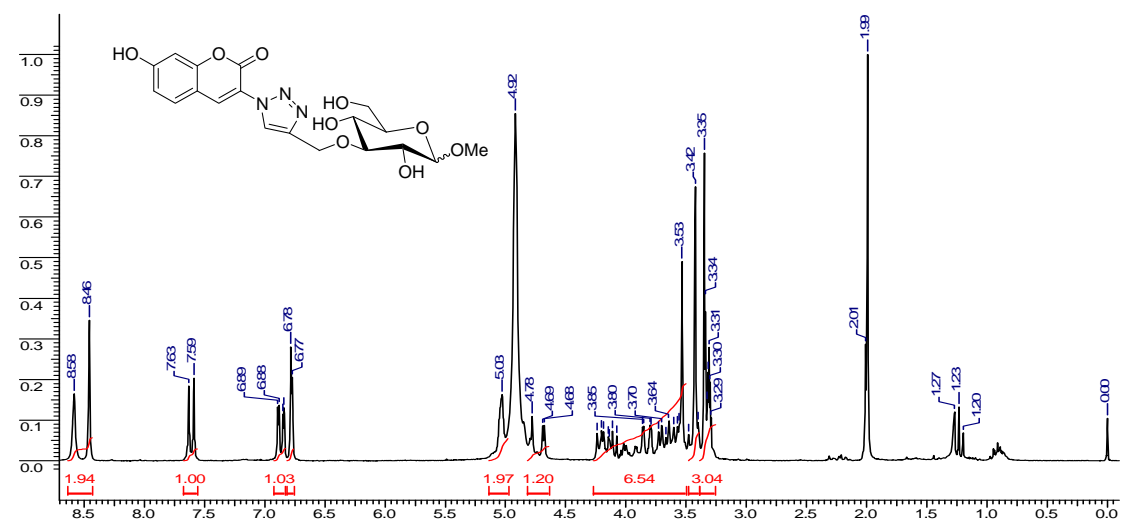
¹H NMR (200 MHz, CDCl₃) of Compound 19a



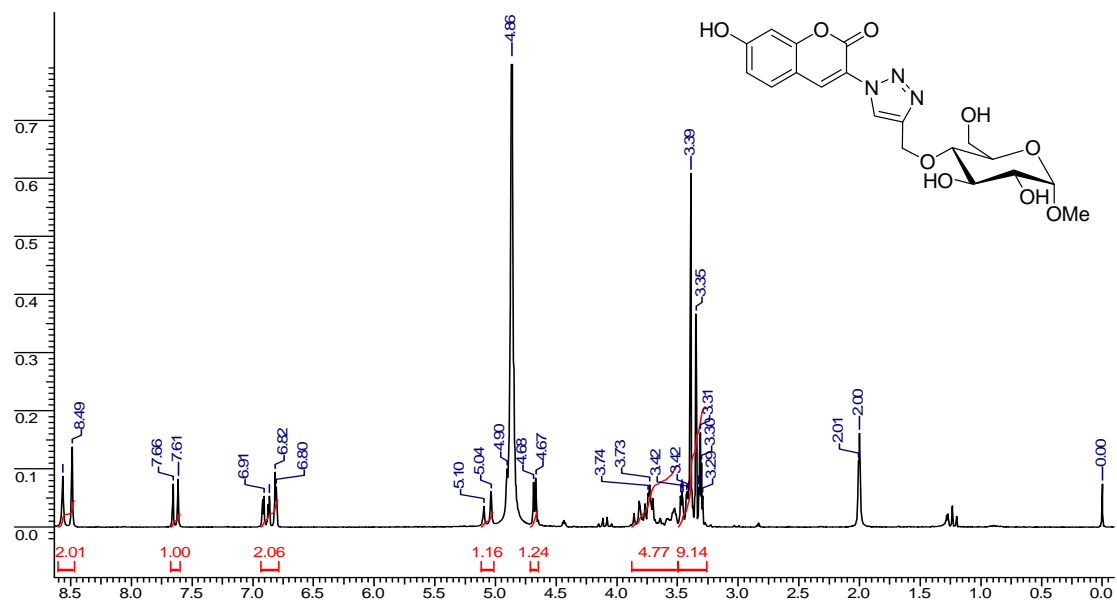
¹H NMR (200 MHz, CDCl₃) of Compound 19b



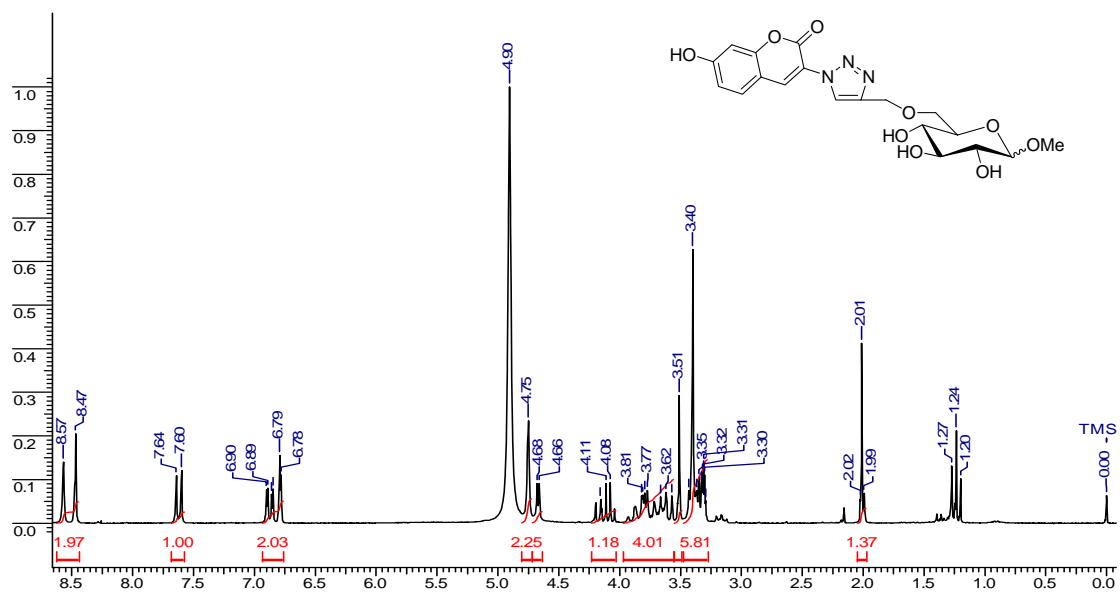
¹H NMR (200 MHz, CDCl₃) of Compound 19c



¹H NMR (200 MHz, CDCl₃) of Compound **19d**



¹H NMR (200 MHz, CDCl₃) of Compound **19e**



Chapter 2: References

1. Friedman, R. J.; An, Y. H. *Handbook of Bacterial Adhesion: Principles, Methods, and Applications*.
2. Sharon, N.; Ofek, I.; *Glycoconjugate J.* **2001**, *17*, 659-664.
3. Lee, Y. K.; Puong, K. Y. *Br. J. Nutr. Med.* **2002**, *88*, S101-S108.
4. Isolauri, E.; Joensuu, J.; Vesikari, T. *Vaccine* **1995**, *13*, 310-312.
5. Hakamori, S. I. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 225-232.
6. Sorice, M.; Longo, A.; Pavan, A.; *Glycoconjugate J.* **2004**, *20*, 63-70.
7. Crick, F. H. C. *Symp. Soc. Exp. Biol. XII* **1958**, 139-163.
8. (a) Rees, D. A. *Biochem. J.* **1972**, *126*, 257-273; (b) Holzl, G., Dormann, P. *Prog. Lipid Res.* **2007**, *46*, 225-243; (c) Wang, X. *Curr. Opin. Plant Biol.* **2004**, *7*(3), 329-336.
9. Rademacher, T. W.; Parekh, R. B.; Dwek, R. A. *Annu. Rev. Biochem.* **1988**, *57*, 785-838.
10. Bertozzi, C. R.; Kiessling, L. L. *Science* **2001**, *291*, 2357-2364.
11. Zachara, N. E.; Hart, G. W. *Chem. Rev.* **2002**, *102*, 431-438.
12. Aubert, A. K.; Desai, N. N.; Neuberger, A. C.; Michael, J. *Arch. Biochem. Biophys.* **1976**, *175*(2), 410-418.
13. (a) Mahal, L. K.; Yarema, K. J.; Bertozzi, C. R. *Science* **1997**, *276*, 1125-1128; (b) Varki, A. *FASEB*, **1991**, *5*, 226. (c) Fltz, W.; Wong, C.-H. *J. Org. Chem.* **1994**, *59*, 8279; (d) Shames, S. L. et al. *Glycobiology*, **1991**, *1*, 187; (e) Lin, C.-H.; Sugai, T.; Halcomb, R. L.; Ichikawa Y, Wong, C.-H. *J. Am. Chem. Soc.* **1992**, *114*, 10138.
14. Sampathkumar, S. G.; Li, A. V.; Jones, M. B.; Sun, Z.; Yarema, K. J. *Nat. Chem. Biol.* **2006**, *2*, 149-152.
15. Sawa, M.; Hsu, T. L.; Itoh, T.; Sugiyama, M.; Hanson, S. R.; Vogt, P. K.; Wong, C.-H. *Proc. Nat. Acad. Sci.* **2006**, *103*, 12371-1276.
16. Sadamoto, R.; Niikura, K.; Sears, P. S.; Liu, H.; Wong, C.-H.; Suksomcheep, A.; Tomita, F.; Monde, K.; Nishimura, S. I. *J. Am. Chem. Soc.* **2002**, *124*, 9018-9019.
17. Sadamoto, R.; Niikura, K.; Monde, K.; Nishimura, S.-I. *Methods Enzymol.* **2003**, *362*, 273-286.
18. Sadamoto, R.; Niikura, K.; Ueda, T.; Mode, K.; Fukuhar, N.; Nishimura, S.-I. *J. Am. Chem. Soc.* **2004**, *126*, 3755-3761.

19. Koch, A. *Clin. Microbiol. Rev.* **2003**, *16*(4), 673-687.
20. Hanover, J. A. *FASEB* **2001**, *15*, 1865-1876.
21. van Leeuwenhoek, A. *Philosophical Transactions*, **1684**, *14*, 568-574.
22. van Leeuwenhoek, A. *Philosophical Transactions* **1700**, *22*, 509-518.
23. van Leeuwenhoek, A. *Philosophical Transactions* **1702**, *23*, 1304-1311.
24. Gram, H. C. *Fortschr. Med.* **1884**, *2*, 185-189.
25. Madigan, M.; Martinko, J. *Brock Biology of Microorganisms*.
26. (a) Sharon, N. *Biochem. Biophys. Acta* **2006**, *1760*, 527-537. (b) Sharon, N. *FEBS Lett.* **1987**, *217*, 145-157. (c) Yarema, K. J.; Bertozzi, C. R. *Curr. Opin. Chem. Biol.* **1998**, *2*, 49-61.
27. (a) Rudd, P. M.; Elliott, T.; Cresswell, P.; Wilson, I. A.; Dwek, R. A. *Science* **2001**, *291*, 2370-2376. (b) McAuliffe, J. C.; Hindsgaul, O. *Front. Mol. Biol.* **2000**, *30*, 249. (c) Varki, A. *Glycobiology* **1993**, *3*, 97-130. (d) Campbell C. T.; Yarema, K. J. *Genome Biol.* **2005**, *6*, 236.1-236.7. (e) Murrell, M.P., Yarema, K.J., Levchenko, A. *ChemBiochem.* **2004**, *5*, 1334-1347.
28. (a) Saxon, E.; Bertozzi, C. R. *Science* **2000**, *287*, 2007-2010. (b) Tai, H. C.; Kidekel, N.; Ficarro, S. B.; Peters, E. C.; Hsieh-Wilson, L. C. *J. Am. Chem. Soc.* **2004**, *126*, 10500-10501; (c) Mahal, L. K.; Bertozzi, C. R. *Chem. Biol.* **1997**, *4*, 415-422. (d) Prescher, J. A.; Bertozzi, C. R. *Cell* **2006**, *126*, 851-854. (e) Dube, D. H.; Bertozzi, C. R. *Curr. Opin. Chem. Biol.* **2003**, *7*, 616-625. (f) Campbell, C.T., Sampathkumar, S.-G., Yarema, K.J. *Mol. Biosys.* **2007**, *3*, 187-194. (g) Sadamoto, R. *Trends Glycosci. Glycotechnol.* **2005**, *17*, 97-105.
29. (a) Zhang, J.; Campbell, R. E.; Ting, A. Y.; Tsien, R. Y. *Nat. Rev. Mol. Cell. Biol.* **2002**, *3*, 906-918. (b) Laughlin, S. T.; Agard, N. J.; Baskin, J. M.; Carrico, I. S.; Chang, P. V.; Ganguli, A. S.; Hangauer, M. J.; Lo, A.; Prescher, J. A.; Bertozzi, C. R. *Methods Enzymol.* **2006**, *415*, 230-250. (c) Chang, P. V.; Prescher, J. A.; Hangauer, M. J.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2007**, *129*, 8400-8401. (d) Link, A. J.; Tirrell, D. A. *J. Am. Chem. Soc.* **2003**, *125*, 11164-11165.
30. Zhang, J.; Campbell, R. E.; Ting, A. Y.; Tsien, R. Y. *Nature* **2002**, *3*, 906-918.
31. (a) Sivakumar, K.; Xie, F.; Cash, B. M.; Long, S.; Barnhill, H. N.; Wang, Q. *Org. Lett.* **2004**, *6*, 4603-4606. (b) Zhou, Z.; Fahrni, C. J. *J. Am. Chem. Soc.* **2004**, *126*, 8862-8863.
32. (a) Huisgen, R. In *1,3-Dipolar Cycloaddition Chemistry*, Padwa, A., Ed., Wiley: New York, **1984**, Chapter 1, pp 1-176. (b) Kolb, H. C.; Finn, M. G.; Sharpless, K.

- B. *Angew. Chem., Int. Ed.* **2001**, *40*, 2004-2021. (c) Lewis, W. G.; Green, L. G.; Grynszpan, F.; Radic, Z.; Carlier, P. R.; Taylor, P.; Finn, M. G.; Sharpless, K. B. *Angew. Chem., Int. Ed.* **2002**, *41*, 1053-1057. (d) Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem., Int. Ed.* **2002**, *41*, 2596-2599. (e) (a) TornØe, C.; Christensen, C.; Meldal, M. *J. Org. Chem.* **2002**, *67*, 3057-3064. (f) Buchardt, J.; SchiØdt, C. B.; Krog-Jensen, C.; Delaissé, J.-M.; Foged, N. T.; Meldal, M. *J. Comb. Chem.* **2000**, *2*, 634-638. (g) Wang, Q.; Chan, T. r.; Hilgraf, R.; Fokin, V. V.; Sharpless, K. B.; Finn, M. G. *J. Am. Chem. Soc.* **2003**, *125*, 3192-3193. (h) Speers, A. E.; Cravatt, B. F. *Chem. Biol.* **2004**, *11*, 535-545.
33. Huisgen, R. "Centaury Lecture – 1,3-dipolar cycloadditions" – *Proceedings of the Chemical Society of London* **1961**, 357.
34. Roy, B.; Mukhopadhyay, B. *Tetrahedron Lett.* **2007**, *48*, 3783-3787.
35. Munavu, R. M.; Szmant, H. H. *J. Org. Chem.* **1976**, *41*, 1832-1836.

Chapter 3
Photocleavable Linkers for Bioconjugation of Proteins

Chapter 3: Introduction

Living organisms are composed of lifeless molecules. Yet living organisms possess extraordinary attributes not exhibited by random collection of molecules.¹ First is their degree of complexity and organisation in which the thousands of different molecules make up a cell's intricate internal structures. In contrast, inanimate matter-clay, sand, rocks etc. usually consists of mixtures of relatively simple chemical compounds. Secondly, living organisms extract, transform, and use energy from their environment usually in the form of chemical nutrients or sunlight enabling them to build and maintain their complicated structure and support various cellular metabolisms. Complementarily, inanimate matter exploits energy in an unsystematic manner which tends to decay towards a more disordered state to come to the equilibrium with its surroundings. The third attribute if the living organisms is the capability for precise self-replication and self-assembly, a property that is quintessence of the living state. A single bacterium placed in a sterile nutrient medium can give rise to billion identical "daughter" cells in 24 hours.² Each of the cells contain thousands of different molecules, some extremely complicated; yet each bacterium is a faithful copy of the original, its construction directed entirely from the information contained within the genetic material of the original cell. Although the ability to self-replicate has no true analog in the nonliving world, there is an instructive analogy in the growth of the crystals in saturated solutions. Crystallization produce more material identical in lattice structure to the original "seed" crystal but these are less complex and their structure is static as compared to the dynamic for of life in living organisms.

Each component of living organisms has a specific function. This is true not only of macroscopic structures, such as leaves and stems or heart and lungs, but also true for the variety of bio-molecules present inside the cell. The interplay among the chemical components of a living organism is dynamic; changes in one component cause coordinating or compensating changes in another, with the whole ensemble displaying a character beyond that is its individual constituents. The collection of molecules carries out a program, the end result of which is reproduction of the program and self-perpetuation of that collection of molecules; in short, life. The fundamental building blocks of living cells are biomolecules which provide the bedrock of life and now provide a new research frontier for the interfacial science. All life forms can be considered to be self-organized systems assembled from these building blocks. Even though one can subdivide life into substructures like cell organelles, the entire cell, tissues, multicellular organisms, and even colonies, biomolecules make their presence felt

across the entire hierarchy of the biological order. So naturally the quest for a theory of the living organisms starts with the fundamental building blocks – biomolecules.

The biomolecules around which the dynamic form of life revolves can be broadly classified into three categories - nucleic acids, protein and carbohydrates, with the first two comprising of most important ones sharing the genetic information between them.³ It is amazing to notice that most of the molecular constituents of living systems are composed of carbon atoms covalently joined with other carbon atoms and with hydrogen, oxygen, or nitrogen. It is the structural and spatial arrangement of the atoms inside the molecule which enables them to interact with their environment in a particular manner is the dictating factor in deciding the ultimate behaviour of the molecule and thus its characteristic function. The relationship between the biomolecules and the fundamental equation of life is beautifully summed up in the central dogma of biology which holds the secret for life (Figure 1).³

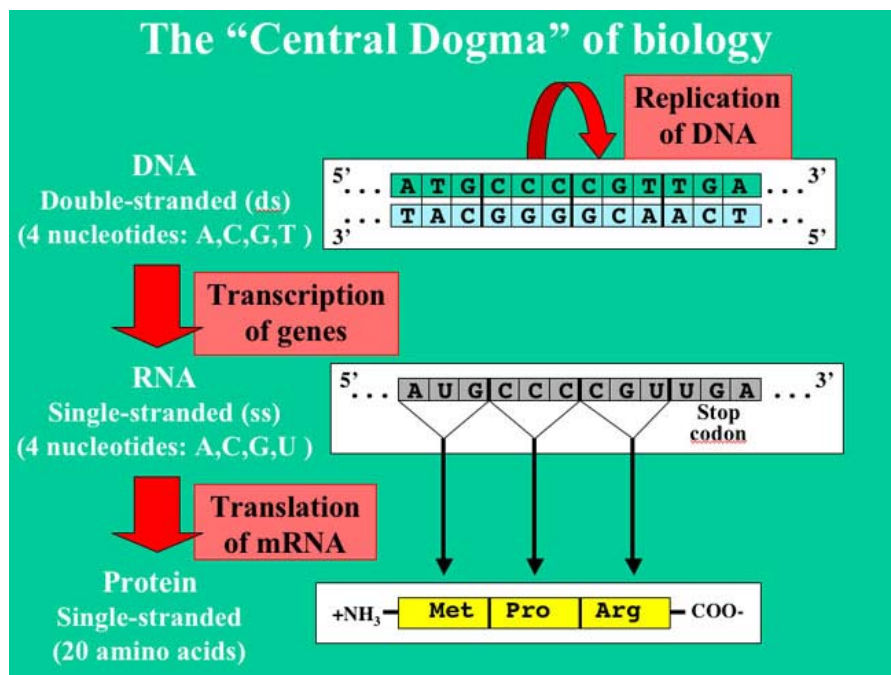


Figure 1. *Central Dogma: The secret for life*

The classical flow of coded genetic information between the different biopolymers is magnificently illustrated in the central dogma which deals with the detailed residue-by-residue transfer of sequential information in a manner that it cannot be transferred back from protein to either protein or nucleic acid. The central dogma was first enunciated by Francis Crick in 1958.^{3a} There are three class of the biopolymers: DNA and RNA (both nucleic acids), and protein, involved in the central dogma equation. Thus there can be $3 \times 3 = 9$ conceivable direct transfers of information that occur between these. Transcription is the process by which the information contained in a section of DNA is transferred to a newly assembled piece of

messenger RNA (mRNA). It is facilitated by RNA polymerase and transcription factors.^{3b, 3e} Eventually, this mature mRNA finds its way to a ribosome, where it delivers the genetic code for the protein assembly, a process commonly known as Translation.^{3b} In prokaryotic cells, which have no nuclear compartment, the process of transcription and translation may be linked together.⁴ In eukaryotic cells, the site of transcription (the cell nucleus) is usually separated from the site of translation (the cytoplasm), so the mRNA must be transported out of the nucleus into the cytoplasm, where it can be bound by ribosomes. The mRNA is read by the ribosome as triplet codons, usually beginning with an AUG, or initiator methionine codon downstream of the ribosome binding site.⁵ Complexes of initiation factors and elongation factors bring aminoacylated transfer RNAs (tRNAs) into the ribosome-mRNA complex, matching the codon in the mRNA to the anti-codon in the tRNA, thereby adding the correct amino acid in the sequence encoding the gene. As the amino acids are linked into the growing peptide chain, they begin folding into the correct conformation. This folding continues until the nascent polypeptide chains are released from the ribosome as a mature protein. However, ultimately the whole story boils down to the type of the protein translated from the mRNA to carry out the different metabolic functions inside the cells.

Proteins are the most abundant biological macromolecules, occurring in all parts of the cells. They also occur in great variety; thousands of different kinds, ranging in size from relatively small peptides to huge polymers with molecular weights in millions, may be found in a single cell. Moreover, proteins exhibit enormous diversity of biological function and are the most important final products of the information pathways (Central Dogma). Proteins are the molecular instruments through which the genetic information is expressed.³ Relatively simple monomeric subunits provide the key to the structure of the thousands different kinds of the proteins. All proteins, whether from the most ancient lines of bacteria or found among the complex forms of life are constructed from the same ubiquitous set of amino acids, covalently linked characteristic linear sequence to make up a polypeptide chain which is nothing but the primary structure of the protein. Because each of the amino acid has a side chain with distinctive chemical properties, this group of 20 precursor molecules may be regarded as the alphabet in which the language of protein structure is written. What is the most remarkable thing is that cells can produce proteins with strikingly different properties and activities by joining the same 20 amino acids in many different combinations and sequences. From these building blocks different organisms can make such widely diverse products as enzymes, hormones, antibodies, transporters, muscles, milk proteins and a myriad of other substances having distinct biological activities.⁶

For large macromolecules such as proteins, the tasks of describing and understanding structure are approached at several levels of complexity, arranged in a kind of conceptual hierarchy. Four levels of protein structure are commonly defined.⁷ A description of all covalent bonds (mainly peptides and disulfide bonds) linking amino acids residues in a polypeptide chain is its primary structure. The most important element of the primary structure is the sequence of the amino acids. Secondary structure refers to the particularly stable arrangements of amino acid residues giving rise to recurring structural patterns. In biochemistry and structural biology, secondary structure is the general term for local segments of biopolymers (proteins and nucleic acids) which is formally defined by the weak interactions arising within the molecule. By building models of peptides using known information about bond lengths and angles, the first elements of secondary structure, the alpha helix and beta-sheet, were first suggested by Linus Pauling and co-workers in 1951.⁸ Both the alpha helix and beta-sheet represent a way of saturating all the hydrogen bond donors and acceptors in the peptide backbone. Although the other elements of the secondary structures have been discovered, like various forms of loop and helices, but alpha helix and beta-sheet are the most predominant secondary structures found in protein.

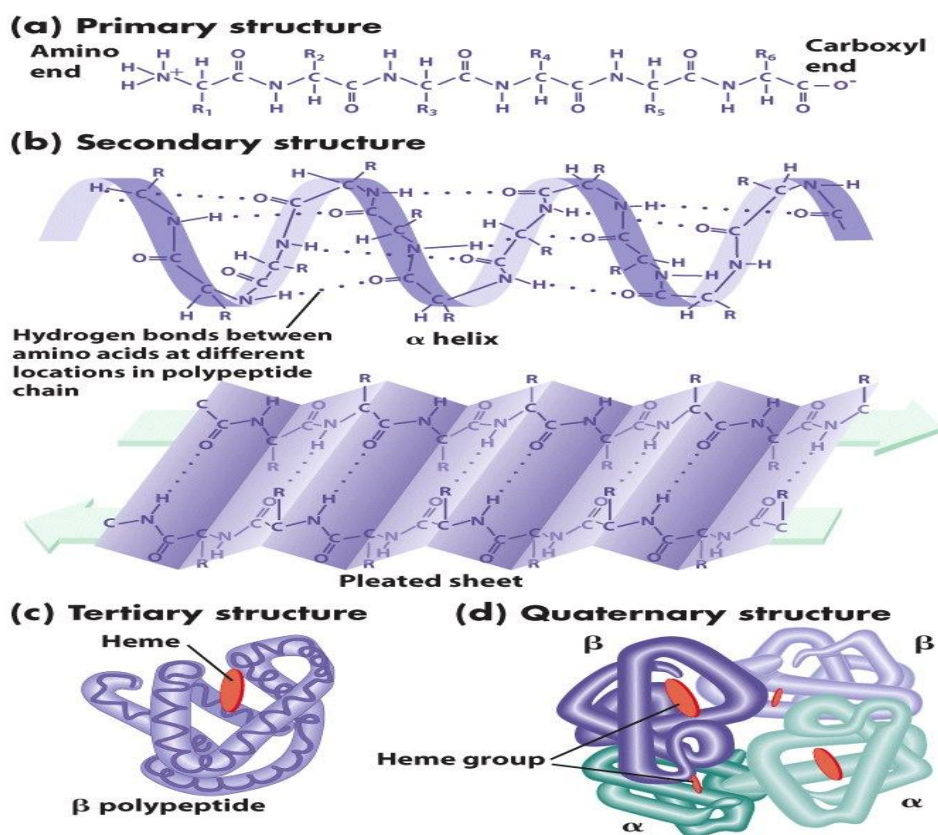


Figure 2. Diagrammatic illustration of the various structures of proteins

The elements of the secondary structure are usually folded into a compact shape using a variety of loops and turns giving rise to the tertiary structure of the proteins.⁹ It formally describes all aspects of the three-dimensional folding of the peptide backbone. The formation of the tertiary structure is usually driven by the burial of hydrophobic residues, but other interactions such as hydrogen bonding, ionic interactions and disulfide bonds can also stabilize the tertiary structure to encompass all the non-covalent interactions which are not covered in the secondary structure, and is what defines the overall fold of the protein. It is considered indispensable for the function of the protein and is usually prone to any influence (e.g. certain chemicals, heat etc.) that disrupts the weak interactions, a process which is commonly termed as denaturation of the protein. Quaternary structure of the protein involves the spatial relationship of multiple polypeptide chain (subunits) that are stably associated.⁹ Not all proteins are found to have quaternary structure. The individual subunits may not be covalently connected, but are stabilized by the same range of interactions as the tertiary structure. All the structures of the protein are diagrammatically depicted in Figure 2.

As is evident from the central dogma, proteins are an important class of molecules considering the variety of the cellular reactions they take part in. In the recent times, a wealth of methods had been developed to manipulate the protein structure and thus its activity to unravel the critical details of protein function at the molecular and cellular level. Notably, the use of genetic tools such as site-directed mutagenesis, knock-out and knock-in techniques, and the use of encoded reporters such as fluorescent proteins have facilitated the discovery of many protein functions within their molecular networks.¹⁰ One of the next formidable challenges in biology is directed towards understanding the molecular details of how post-translational modifications (PTMs) affect protein function, stability, cellular localization, three-dimensional conformation, and interaction with other molecules. Many of these biological investigations of protein function require access to protein molecules that are hard to obtain with purely biological methods such as site-directed mutagenesis and recombinant protein expression.¹¹ The combination of chemical and biological techniques to chemoselectively modify proteins, however has proved to be an excellent resource for studying protein function on a molecular level. A few of the prominent techniques related to post-translational modification of proteins can broadly be classified into the following categories: (a) PTMs involving addition of functional groups; (b) PTMs involving addition of small peptides or other proteins; (c) PTMs inducing the changes in the chemical nature of amino acids present on the peptide backbone; (d) PTMs involving structural changes in the tertiary structure of protein.

Based on early work by Wieland et. al., a synthetic method known as native chemical ligation (NCL) has evolved to be a powerful approach in peptide chemistry.¹² NCL allows the construction of large polypeptides through the condensation of two or more smaller peptides with C-terminal thioesters and N-terminal cysteines, in the presence of an added thiol catalyst. In a free reversible first step, a transthioesterification occurs to yield a thioesters-linked intermediate; this intermediate rearranges irreversibly under the usual reaction conditions to form a native amide bond at the ligation site. A short diagrammatic description is provided in Figure 3a. NCL of unprotected peptide segments was developed by Dawson *et. al.* but Wieland's work led to the 'active ester' method for making protected peptide segments in conventional solution synthesis in organic solvents.¹³ An elegant display of NCL was demonstrated by C. H. Wong group when they successfully synthesized *N*-linked glycopeptides using sugar assisted ligation (SAL) strategy from cysteines free peptides.¹⁴ This approach is based on SAL involving a peptide thioesters and a glycopeptide in which *N*-acetylglucosamine linked to serine side chain is modified at C-2 acetamide with a thiol handle to mimic the cysteines function. They also reported the similar rates of ligation for both *N*- and *O*-linked glycopeptides coupled with the wide tolerance for a variety of amino acids. The detailed diagrammatic representation is provided in Figure 3b.

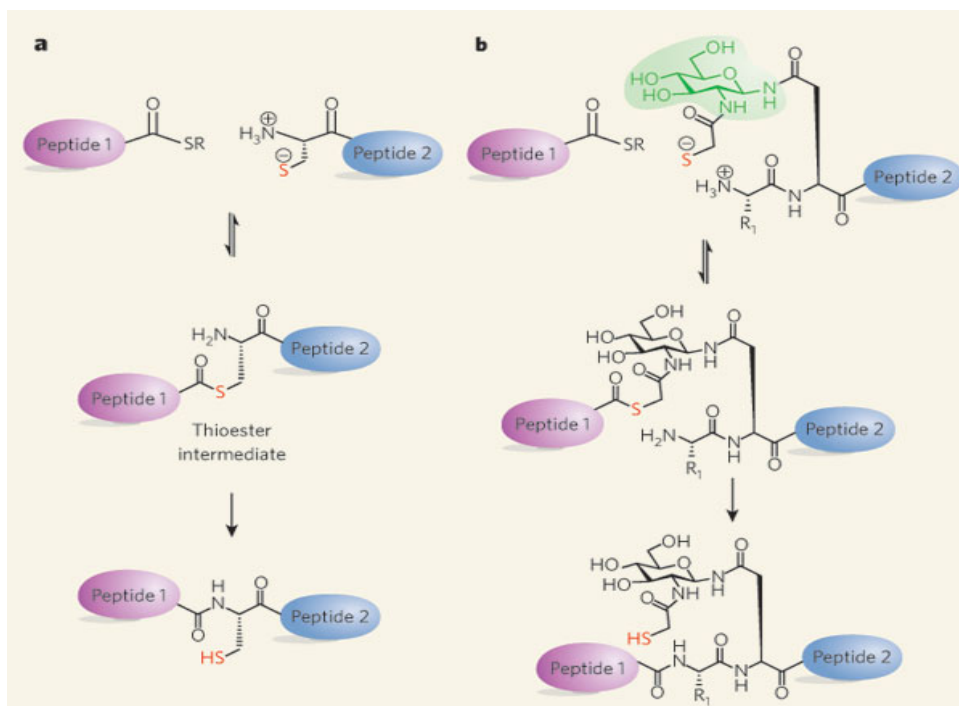


Figure 3. (a) Native Chemical Ligation: A well established method for preparing peptides, *R* is typically a phenyl ring; (b) Wong's modification of the NCL to prepare the glycopeptides, in which sugars are attached to peptide chains, *R*₁ represents the amino acid side chain

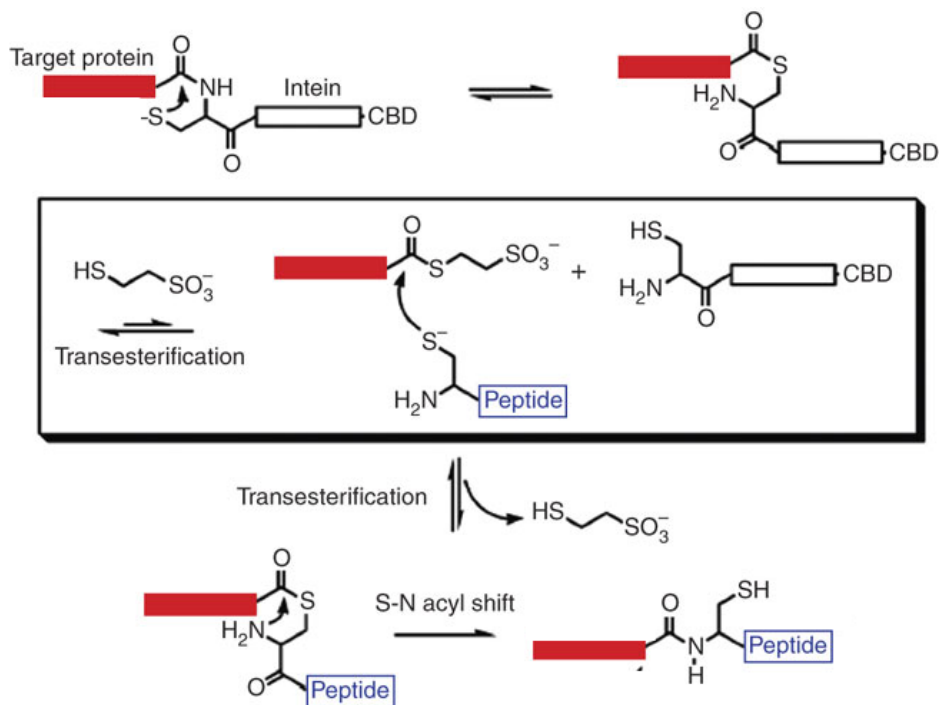


Figure 4. Schematic representation of Expressed Protein Ligation. CBD is the Chitin Binding Domain

The semi-synthetic version of NCL, known as Expressed Protein Ligation (EPL), permits the *in vitro* ligation of a chemically synthesized C-terminal segment of a protein to a recombinant N-terminal segment fused through its C terminus to an intein protein splicing element.¹⁵ EPL, also named as intein-mediated protein ligation, facilitates the site-specific protein labelling with a broad range of physical probes such as fluorophores, PTMs, stable isotopes, and unnatural amino acids. The growing popularity of EPL can be estimated by the increasing application of this technique to a wide range of protein-engineering problems, for e.g. generation of cyclic peptides and proteins, segmental isotopic labelling, site-specific protein modification, generation of cytotoxic proteins etc. Inteins are internal segments of precursor proteins that catalyze their *ipso* excision, in an intramolecular process called protein slicing, with the concurrent ligation of to flanking external regions (N- and C- inteins) through a native peptide bond, to finally yield a protein.¹⁶ The schematic illustration is shown in Figure 4 for the EPL. Note that the addition of the external as an additive is the key to the process since it generated reactive α -thioester derivative of the protein *in situ*. This underlines the extension of ligation techniques to the emerging area of chemical genetics providing innovative tools to probe protein function through perturbation of protein activity or protein-protein interactions in a rapid tunable manner that is necessary to track protein function within the dynamic and variable nature of biological processes.

The only disadvantage of NCL and EPL is the necessity of cysteines residue or a homologue at the ligation site. The occurrence of this amino acid in globular proteins is very low and the insertion of additional cycteine residues can alter the protein structure and function by the formation of disulfide bridges. Several approaches have been developed to circumvent this limitation. A few of the promising and most sought after techniques include NCL with Cys-mimetics,¹⁷ Staudinger ligation on proteins,¹⁸ expressed enzymatic ligation,¹⁹ maleimide and succinamide assisted protein conjugation¹⁹ etc. Complementary to the development of techniques for protein-protein ligation, researchers are also involved in the probing the behaviour of proteins in vicinity of abiotic surfaces like nanotubes/nanoparticles, glass slides, transition metal catalysts etc.

However, the mild and selective chemical modification of complex biomolecules such as proteins under physiological conditions remains a huge challenge. There are so many factors governing the stability of the proteins, like temperature and chemical sensitivity, denaturation and precipitation etc. which make the protein handling fairly difficult to handle. However, the power of protein ligation to investigation protein function is evident and the technology is constantly improving, aiming towards minimally invasive methods that allow analysis of target proteins within the ensemble of all the other proteins *in vivo*.

Chapter 3: Present Work

The continuing success of genome sequencing efforts has laid the foundation for understanding the molecular basis of life in its many forms. However, this task involves not only the characterization of genes but the products of gene expression as well. As primary mediators of most physiological and pathological processes, proteins may be viewed as the next major challenge, especially considering the difficulties involved in developing and applying methods for probing protein interactions as compared to DNA or RNA molecules.²⁰ For example, unlike oligonucleotides, proteins exhibit a diverse array of chemical and biochemical properties, are not amenable to molecular amplification, and do not possess predefined complementary binding partners.

Despite the many technical challenges that accompany the analysis of proteins, the need for global strategies to characterize the expression and function of these biomolecules is clear, especially given the multitude of post-transcriptional and post-translational processes that regulate the activity of proteins in cells and tissues.²¹ In most cases, the initial polypeptide-translation product undergoes some type of modification before it assumes its functional role in a living system. These changes are broadly termed as “post-translational modifications” and encompass a wide variety of reversible and irreversible chemical reactions. Specific delivery of the protein to specific subcellular or extracellular components is often achieved with leader or signal sequences, which may be proteolytically cleaved. At their cellular destinations, proteins carry out their many functions, and their activities are controlled by post-translational modifications. Most proteins function in collaboration with other proteins found within the perimeter of the cells. Sensing extracellular signals is a matter of receptor to adaptor interactions and the shape of the cell is maintained by an intricate network of structural protein interactions. The interactions between the proteins are central importance for virtually every process in a living cell, probing these interactions is a way to get a broader view of how they work cooperatively in a cell.³ Information about these interactions improves our understanding of diseases and can provide the basis for new therapeutic approaches.

The temporal and spatial regulation of protein function is central to biological processes. Hence, the ability to artificially trigger molecular events in a biologically relevant context is useful for the study of living organism.²² By chemically modifying a protein with groups that can respond to external inputs (such as reagents, light etc.), will enable us to reveal the complex mechanism through which proteins operate. The ability to alter the protein

structure and function by introducing unnatural amino acids has great potential to enhance our understanding of the proteins and thus generate new tool for biochemical research. The delicate structure of the proteins being vulnerable to the chemicals coupled with the thermal stability makes task of biochemists all the even more complicated. In spite of these short comings, innumerable methods have been reported in the literature for the construction of protein-protein conjugates *via* the chemical modifications of this complex biomolecules.^{18, 19, 20}

Most of the reactions employed for the protein derivatization involve the formation of a covalent bond with the reactive functional groups (of amino acids) present on the polypeptide backbone. Efficiency in this case depends entirely on the number of these reactive functional residues present and the ease in the accessibility of these groups. Indeed, Schultz and co-workers have demonstrated an elegant method for the selective modifications of the protein surfaces *via* the genetic incorporation of azide- or acetylene-containing unnatural amino acids in *E. coli* and yeast proteins.²³ These unnatural amino acids were then modified by Huisgen [3 + 2] cycloaddition reaction with the complementary derivatives respectively enabling them to bioconjugate the two proteins residues. They had earlier showed the incorporation of genetically encoded amino acids in bacteria and yeast cells possessing the ketone residues in amino acids. This methodology allowed them to selectively label the proteins with a host of reagents, including fluorophores and cytotoxic molecules.

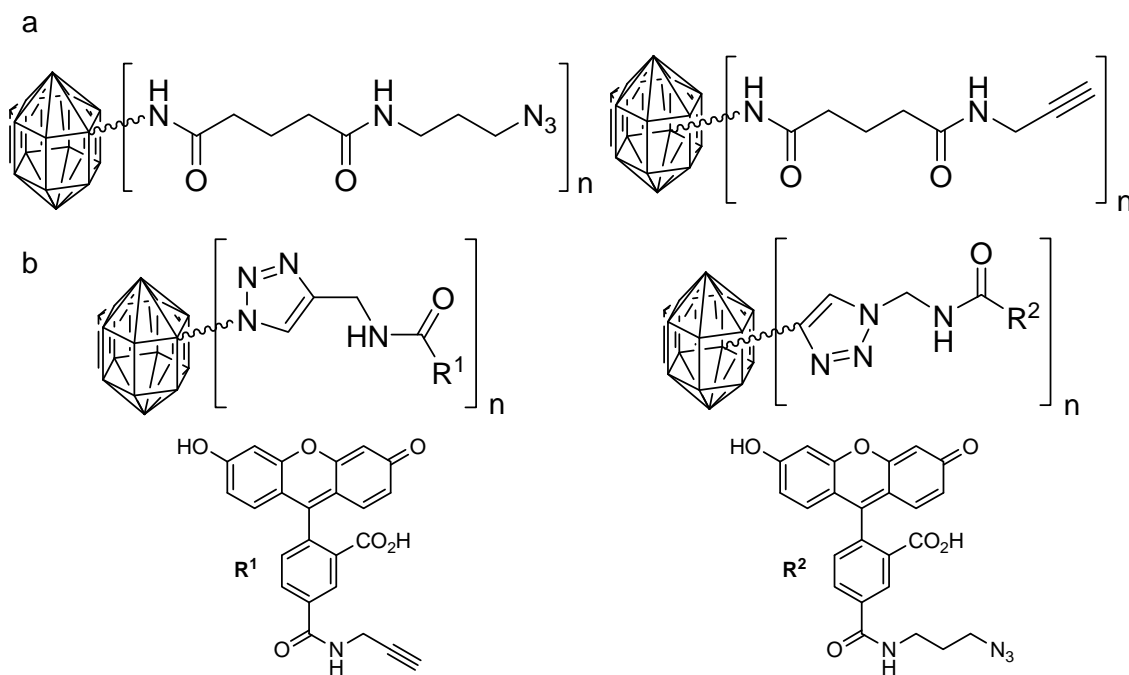


Figure 5. (a) Modified viral capsules using click ligands; (b) Viral capsules after click reaction. Modified Fluorescein dyes employed for click reaction

The exploitation of click chemistry protocols on the proteins is a relatively growing field of biochemistry and developments in this area are rapid. Earlier Finn group has decorated the surface of cowpea mosaic virus (genetic material encapsulated by a protein coat) with azide and alkyne linkers taking advantage of the reactivity of lysine and cysteine residues.²⁴ As is evident from the Figure 5, the researchers have efficiently modified the virus surface and further prepared the conjugates by reacting with them with the fluorescein derivatives through the Cu(I)-catalyzed azide-alkyne cycloaddition.

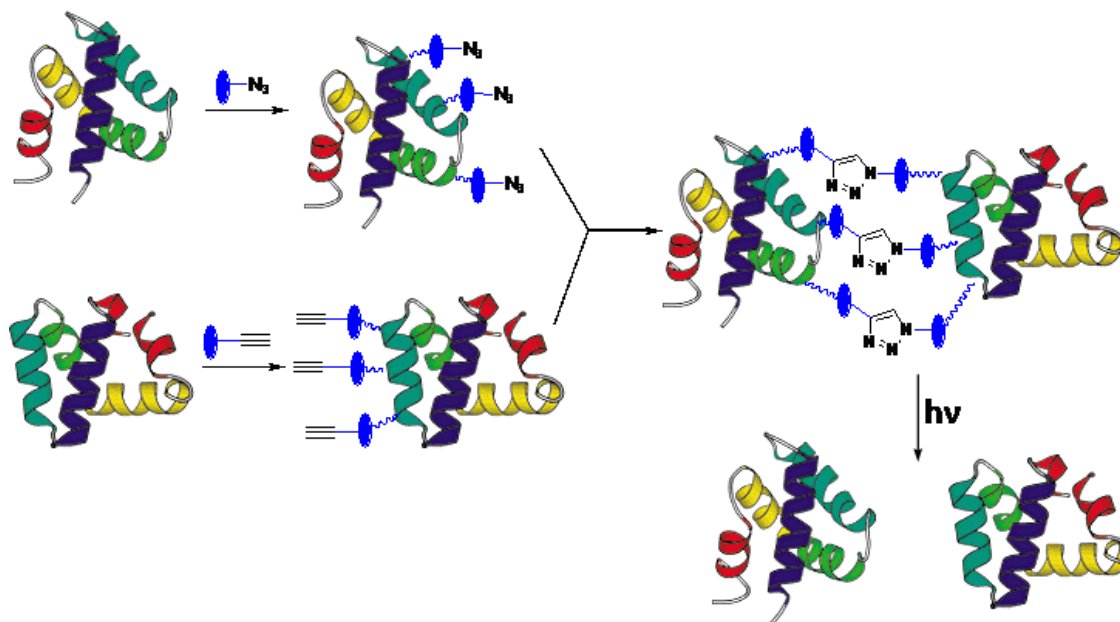


Figure 6. Hypothetical representation of the scheme for the bioconjugation of the proteins

Although powerful, each of the currently existing techniques to introduce unnatural molecules into proteins has associated with it certain synthetic or practical limitations that have limited their widespread application. There is a considerable demand for chemical reactions that can selectively modify under physiological conditions. Herein we felt the need for the development of the chemically synthesized linkers that can be covalently appended onto the proteins and removed thereafter as per the requirement. Inspired by the recent developments in the click chemistry, we hypothesized the structure of a bifunctional linker such that they will possess the click appendages (alkyne or azide functionality) along with a photocleavable moiety engrossed in it. Once installed on the protein, it will enable us to bioconjugate the two proteins by clicking them together, and would allow us a unique advantage to again separate the two proteins by cleaving the photochemical groups under light induced conditions. A short hypothetical scheme is displayed in Figure 6. Although, there are a few cleavable linkers reported earlier also for the protein-protein bioconjugation but those have taken advantage of

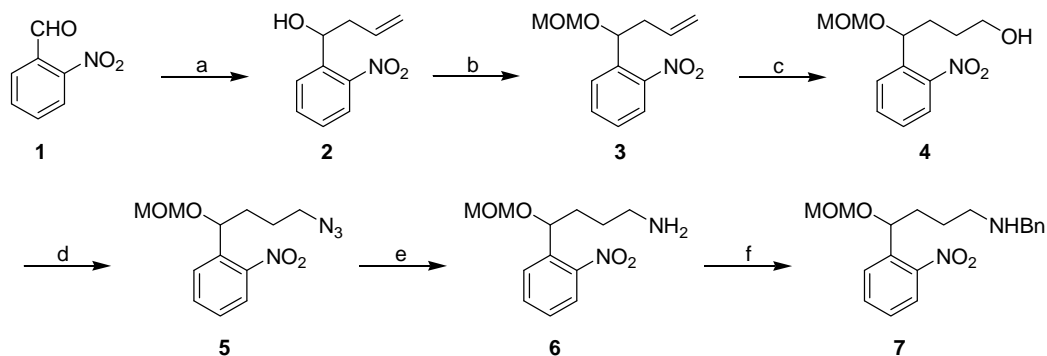
the disulfide bridges on the proteins but these are prone to disulfide exchange and show premature cleavage in cellular systems and reducing buffers.²⁵ The photocleavable group being inert to the biological mechanisms does not interfere in the cellular mechanisms and also can be triggered to expulsion from the biological systems on the exposure to light.

Photoactivable ligands are becoming increasingly common as remotely controllable tools in drug-discovery process.²⁶ These probes enable researchers to identify the target of drugs, to determine the affinity and selectivity of the drug-target interaction, and to identify the binding site on the target. In many cases, the cells/biomolecules do not react to light unless these are highly specialized cells, such as the photoreceptors of the eye (Rhodopsin) etc. Although, the photolabile groups were consistently engaged for the synthetic purposes but their utility for the biological systems was first demonstrated by Hoffman et.al. with the caging of adenosine triphosphate (ATP) derivative.²⁷ Since then photolabile groups are consistently being used for a wide range of biological experiments. Nowadays, a variety of photoactive functionalities, such as *o*-nitrobenzyl derivatives, azobenzenes, are available to suit the requirement of the designed experiments. A biosynthetic approach to site-specific incorporation of unnatural amino acids, such as *o*-nitrobenzyl serine, was shown by Schultz et. al. by incorporating the synthesized unnatural photolabile amino acid into the peptide sequences of the thermostable DNA polymerase.²⁸ One of the more practical application of the photocleavable groups in drug development is in the field of photodynamic therapy, which is currently under development for the cancer therapy. The DNA-alkylating site of a prodrug is blocked by a photocleavable group to give non-cytotoxic prodrugs.²⁹ Local irradiation releases the cytotoxic agent and inhibits cell growth.

In spite of all the positive developments which have taken place in this regards, we noticed that the simultaneous use of “click chemistry” along with the photolabile groups have not been exploited. As is evident from the proposed design of the linkers proposed in Figure 6, it would allow us to synthesize the protein-protein hetro- or homo-dimeric conjugates and would also permit us to again split the dimeric complex by photochemical detachment of the photosensitive groups on the linkers. Unlike, the previous linkers which offered the reversible conjugation of the proteins employing the disulfide bridges on the peptide backbone, the clickable photolabile linkers will not be prone to the reducing agents or unnecessary cleavage under the physiological conditions.

Having, the virtual design of our linkers, we decided to make use of *o*-nitrobenzyl functionality as the photosensitive group. The wide popularity of this reagent as a protecting

group for the synthetic purposes together with the efficient cleavage of this group in response to short wavelength made this an attractive proposition to start with. Likewise we decided to



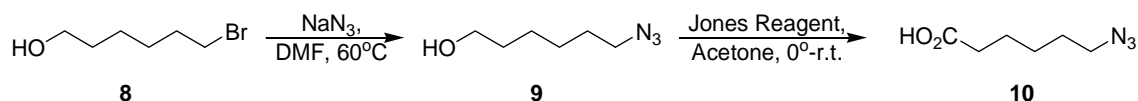
Reagents and Conditions: (a) Allyltributyl Tin, ZnCl₂, Acetonitrile/Water (9:1), 3h, 95%; (b) MOMCl, DIPEA, CH₂Cl₂, 12h, 98%; (c) BH₃.DMS, NaOH, H₂O₂, THF, 2h, 80%; (d) (i) MsCl, CH₂Cl₂, 30 min; (ii) NaN₃, DMF, 60°C, 4h, 90 % over two steps; (e) PPh₃, THF, Water, 4h; (f) (i) PhCHO, MgSO₄, THF, 3h; (ii) NaBH₄, THF, 1h, 70% over three steps.

Scheme 1. Synthetic protocol of the preliminary steps involved in the synthesis of photocleavable linkers

start our synthetic endeavour with the commercially available *o*-nitro benzaldehyde, **1** which was allylated employing allyltributyl tin in the presence of a Lewis acid, zinc chloride in a cocktail of the acetonitrile-water to yield the secondary alcohol, **2**. The alcohol was consecutively converted to MOM ether, **3** so as to curb its reactivity, utilizing MOMCl together with diisopropyl amine in CH₂Cl₂. For the extension of the side chain of the linker we decided to introduce the hydroxyl functionality into the double bond of the MOM ether, this was promptly achieved by the hydroboration of the alkene, **3** by exposing it to borane dimethyl sulfide complex in dry tetrahydrofuran to produce the primary alcohol, **4**. The alcohol was converted to azide *via* the nucleophilic displacement of the intermediate mesyl derivative of the primary alcohol using sodium azide in *N,N*-dimethylformamide under thermal conditions to result in the formation of primary azide **5**. ¹H NMR analysis revealed the presence of resonances at δ 3.32 (singlet, 3H) which can be attributed to the presence of a methyl group in the MOM protecting group while the methylene carbon was noted in ¹³C NMR at δ 95.2 ppm. Remaining resonances were found to be in perfect unanimity with the assigned structure. The azide, **5** was reduced to the primary amine, **6** *via* the Staudinger reaction,³⁰ by means of triphenyl phosphine in tetrahydrofuran-water system. The amine, **6** was converted to benzyl derivative, **7** by in situ reduction of the benzyl imine, formed by the reaction with benzaldehyde, with sodium borohydride. ¹H NMR spectrum examination of the compound **7** confirmed the presence of MOM group with the resonances at δ 3.33 (s, 3H) ppm amounting to the presence of methyl group in the molecule. The other characteristic feature observed in the spectrum was the increase in the number of the protons in the aromatic region which can be

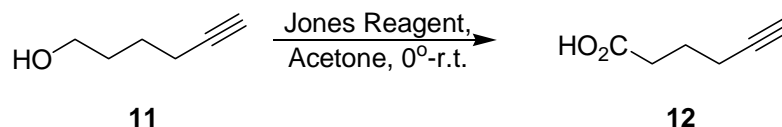
attributed to the existence of the benzyl functionality present in the secondary amine, **7**. Moreover, the presence of methylene carbon at δ 94.3 ppm further confirmed the assigned structure of the compound **7**. The synthetic steps involved are displayed in Scheme 1 along with the reagents and conditions employed.

Meanwhile, the functional side chain for the linkers was prepared. Commercially available 6-bromo hexanol, **8** was heated with sodium azide in DMF to result in the azido alcohol, **9**. The alcohol was subsequently converted to the corresponding 6-azido-1-hexynoic acid, **10** under the oxidizing conditions provided by the Jones reagent. The schematic display of the reagents and conditions involved are shown in Scheme 2.



Scheme 2. Synthetic scheme for the azido side chain

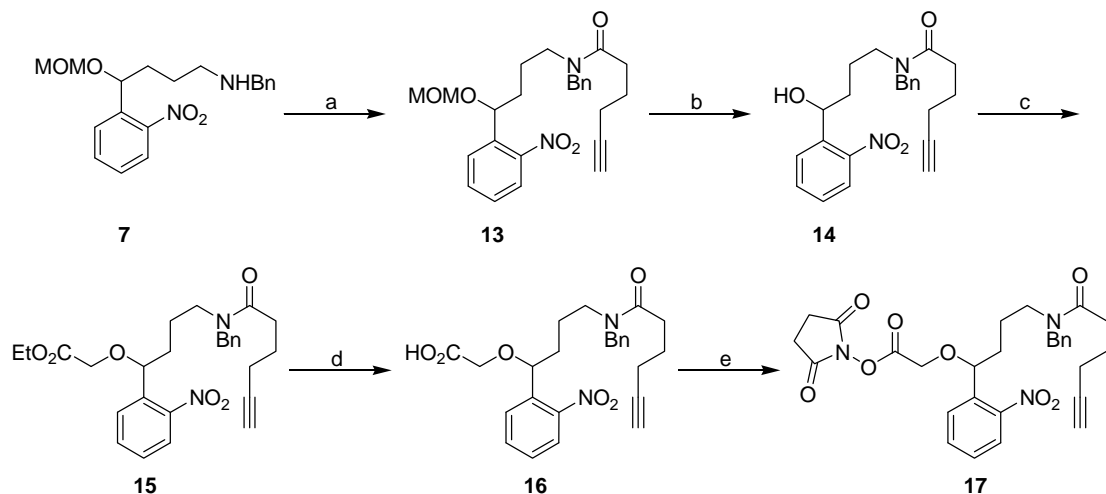
Likewise, the alkyne functionalized side chain was prepared employing the commercially available 6-hexynol, **11** which was thereafter converted to 5-hexynoic acid, **12** by oxidation of alcohol with Jones reagent. The reaction conditions are displayed in Scheme 3. The compound was characterized by ^1H and ^{13}C NMR spectra and the resonances were found to be in accordance with the reported values.³¹



Scheme 3. Synthetic scheme for the alkyne side chain

The acid functionality at one of the side chains, **10** and **12** would enable us to conjugate them onto the amine, **7** thereby extending the alkyl chain. The extension of the side chain is necessary because when appended onto the proteins, the azide/alkyne part of the alkyl chain should protrude out from the bulky structure of the protein so as to be freely available for modification. In continuation with our synthetic efforts, the alkyne derived hetro-bifunctional linker, **17** was synthesized. The free secondary amine, **7** previously prepared was joined with the alkyne side chain with the coupling conditions offered by *N,N*-diisopropyl carbodimide in CH_2Cl_2 to grant us the amide, **13**. Now having completed one part of the linker which seemed suitable for the click modification on the proteins, we continued in our task to introduce the proper functionalities on the linker which would enable us to chemically attach on the protein surface. For the purpose, the MOM ether on **13** was cleaved under acidic environment (HCl, Methanol) to afford the free secondary alcohol, **14** which was alkylated with bromoethylacetate facilitated by sodium hydride in DMF to afford the ester, **15**. Hydrolysis of the ester, **15** with

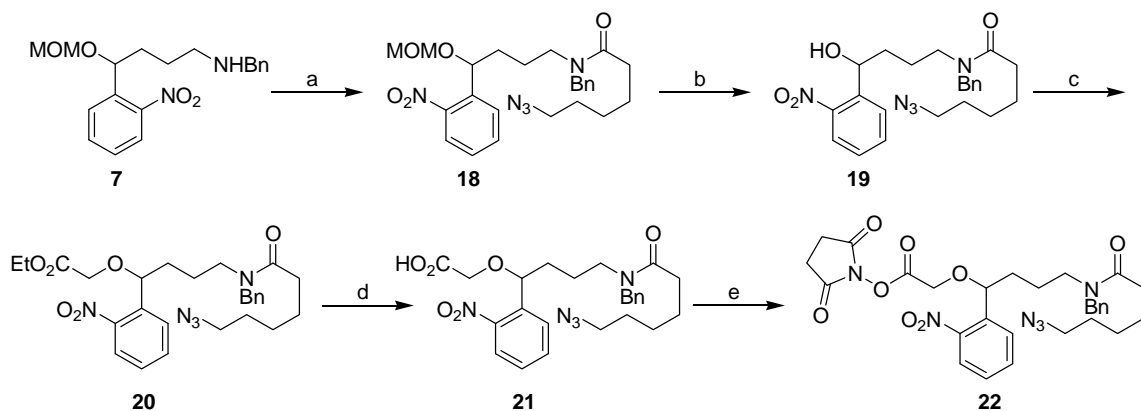
basic reagent (NaOH) in THF/water cocktail gave the free acid, **16**. Promptly the *N*-hydroxy succinamide ester of the free acid, **16** was prepared by using the usual coupling reagents, *N,N*-



Reagents and Conditions: (a) **12**, DIC, DMAP (cat.), DCM, 4h, 90%; (b) HCl(cat.), Methanol reflux, 4h, 95%; (c) NaH, THF, Bromoethylacetate, $n\text{-Bu}_4\text{N}^+\text{I}^-$, 0-rt, 3h, 70%; (d) NaOH, THF/Water, 1h, 80%; (e) *N*-hydroxy succinimide, DIC, DMAP(cat.), CH_2Cl_2 , 2h, 85%.

Scheme 4. Schematic representation of the synthesis of the photocleavable alkyne functionalized linker

diisopropyl carbodimide using CH_2Cl_2 as the solvent to provide the active ester, **17** which seemed ready for the bio-conjugation experiments. The synthetic protocol along with reagents and conditions are depicted in Scheme 4.



Reagents and Conditions: (a) **10**, DIC, DMAP (cat.), DCM, 4h, 90%; (b) HCl(cat.), Methanol reflux, 4h, 95%; (c) NaH, THF, Bromoethylacetate, $n\text{-Bu}_4\text{N}^+\text{I}^-$, 0-rt, 3h, 70%; (d) NaOH, THF/Water, 1h, 80%; (e) *N*-hydroxy succinimide, DIC, DMAP(cat.), CH_2Cl_2 , 2h, 85%.

Scheme 5. Schematic representation of the synthesis of the photocleavable azido functionalized linker.

Delighted with the preparation of linker, **17**, we persisted in our efforts towards the synthesis of the azido functionalized linker, **22**. Our synthetic endeavour was again commenced with the amine, **7**, which was coupled with the previously synthesized azido side

chain, **10** with the coupling conditions afforded by *N, N*-diisopropyl amine in CH₂Cl₂ to provide the amide, **18**. Following the same synthetic protocol as standardized in the Scheme 4, the secondary alcohol, **19** was obtained after the deprotection of the MOM ether in **18** in acidic environment employing HCl in methanol. Likewise we then progressed with the alkylation of the hydroxyl functionality with bromoethylacetate in sodium hydride using DMF as the solvent to result in the product, **20**. It was subsequently hydrolysed to afford the free acid, **21** and converted to the active succinamide ester, **22** following the same protocol as standardized previously (Scheme 5).

We moved on to investigate the utility of the linkers for the bioconjugation, with both the functionalised clickable linkers in hand. For the purpose readily available Bovine Serum Albumin (BSA) was chosen for our experiments due to its numerous biological applications including ELISA (Enzyme-Linked Immunosorbent Assay), used as additive for hybridoma cell cultures etc. Additionally, it is also used as a nutrient in cell and microbial culture. One of the main advantages for preferring BSA over the others was due to its stability as compared to the other proteins.

BSA is known to have several free amines in its tertiary structure. We choose to covalently link only a few of them in our experiments, firstly it will be impossible to react all the free amines and secondly, it will be tiresome job to cleave all the linked amines after the click reaction. Proceeding with our conjugation experiments, BSA was taken in 0.1 M phosphate buffer (pH 7.2) and allowed to stir with the NHS-functionalized linker, **17** for 2 h. Assuming the completion of the experiment, the functionalized protein was dialysed overnight employing the cellulose membrane (12, 000kD cut off) so as to remove the excess linker present in the protein mixture. The covalently modified BSA sample was then analyzed with MALDI-TOF. It was heartwarming to observe that there was an increase in mass of the protein sample. The natural BSA showed a mass of 66,682 kDa while the alkyne modified BSA revealed a mass of 68,084 kDa clearly indicating the presence of functionalized linker on the peptide backbone.

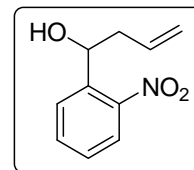
Delighted in having succeeded with initial bioconjugation experiments, we then proceeded to check the feasibility of the azido functionalized linker as well. Following the same protocol, the azido linker, **22** was engaged instead of **17**. Gratifyingly, upon the analysis with MALDI-TOF, the BSA modified with azido linker displayed a mass of 67,970 kDa as compared to the normal mass of BSA, 69,568 kDa, signalling towards the presence of azido linker on the protein surfaces.

We further decided to check the generality of the synthesized linkers, **17** and **22**, for the other proteins. For the purpose, Insulin was chosen as another protein to be engaged in our bioconjugation experiments. Insulin is an animal hormone whose presence in body makes liver and muscle cells to regulate the blood glucose levels. As compared to BSA, it is a small protein with molecular weight of around 5800 kDa. The utility of the linkers was probed by employing both the photocleavable linkers, **17** and **22**, for biological studies. Following the standardized protocol, the Insulin was made to undergo through the same conditions as established for BSA. It was pleasant to notice the modification of the Insulin samples as was evident through the MALDI-TOF analysis of protein samples. The azido- and alkyne-modified Insulin samples revealed the mass of 6238 kDa and 6187 kDa in their respective mass analysis as compared to the normal mass of the Insulin (5800 kDa).

To conclude, we have successfully synthesized the two photocleavable linkers having the azide and alkyne appendages and also exploited them to covalently link onto BSA and Insulin. Furthermore, these linkers can be employed for linking small molecules, such as drugs or fluorescence tags on to the proteins. On the other hand, the linkers can also be exploited for protein-protein bioconjugation to form homo- and hetro-dimeric conjugates to assist in the studies regarding the behaviour of proteins in promiscuity of the other proteins.

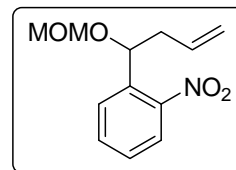
Chapter 3: Experimental Section

Synthesis of 1-(2-nitrophenyl)but-3-en-1-ol: Commercially available *o*-nitro benzaldehyde, **1** (151 mg, 1 mmol) was dissolved in a cocktail of 10 mL acetonitrile:water (9:1), to this was added allyltributyl tin (306 μ L, 1 mmol) along with zinc chloride (136 mg, 1 mmol) was allowed to stir at



ambient temperature for 3 h. After the completion of the reaction, as indicated by the consumption of the starting material on TLC examination, the reaction was subsequently filtered through a pad of celite so as to remove insoluble impurities. The filtrate was diluted with water and extracted with CH_2Cl_2 (3X), the combined extracts were dried over anhydrous sodium sulphate and consequently the solvent was removed *in vacuo* to yield crude compound **2** which was purified *via* silica gel chromatography (Ethyl Acetate-Petroleum Benzene; 1:3) to afford the pure allylated product **2** in 95% yield (182 mg). ^1H NMR (200 MHz, CDCl_3): δ 2.34-2.5 (m, 2H), 2.64-2.77 (m, 1H), 5.14-5.17 (m, 1H), 5.21-5.25 (m, 1H), 5.31 (dd, 1H, $J = 3.79, 8.39$ Hz), 5.79-6.00 (m, 1H), 7.38-7.47 (m, 1H), 7.60-7.69 (m, 1H), 7.82 (dd, 1H, $J = 1.49, 7.84$ Hz), 7.92 (dd, 1H, $J = 1.29, 8.10$ Hz). ^{13}C NMR (50 MHz, CDCl_3): δ 42.8, 68.4, 119.0, 124.4, 128.1, 128.1, 133.4, 134.0, 139.2, 147.8. Anal Calcd. for $\text{C}_{10}\text{H}_{11}\text{NO}_3$: C, 62.17; H, 5.74; N, 7.25. Found C, 62.22; H, 5.64; N, 7.33.

Synthesis of 1-(1-(methoxy-*O*-methyl)-but-3-enyl-2-nitrobenzene: To an ice cooled solution of **2** (192 mg, 1 mmol) and diisopropylethyl amine (347 μ L, 2 mmol) in dichloromethane (5 mL) was added MOMCl (146 μ L, 2 mmol) over a period of 5 min. The reaction was stirred at room

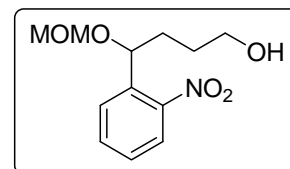


temperature for 12 h. TLC examination of the reaction indicated the formation of the compound as a single compound, after which it was diluted with water and extracted with dichloromethane (2X); the combined extracts were evaporated under reduced pressure before which they were dried over anhydrous sodium sulphate, to provide the crude MOM ether, **3**. It was subsequently purified over silica gel column chromatography (Ethyl Acetate-Petroleum Ether; 1:9) to afford the pure compound **3** in 98% yield (231 mg). ^1H NMR (200 MHz, CDCl_3): δ 2.46-2.65 (m, 2H), 3.31 (s, 3H), 4.52 (ABq, 2H, $J = 14.74$ Hz), 5.07-5.17 (m, 2H), 5.27-5.33 (m, 1H), 5.82-6.03 (m, 1H), 7.38-7.47 (m, 1H), 7.60-7.68 (m, 1H), 7.74-7.79 (m, 1H), 7.93 (dd, 1H, $J = 1.27, 8.22$ Hz). ^{13}C NMR (50 MHz, CDCl_3): δ 41.9, 55.8, 73.2, 95.2, 117.7,

124.3, 128.1, 128.5, 133.2, 134.1, 138.0, 148.3. Anal Calcd. for C₁₂H₁₅NO₄: C, 60.75; H, 6.37; N, 5.90. Found C, 60.82; H, 6.27; N, 5.81.

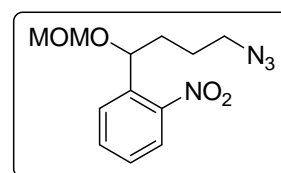
Synthesis of 4-(methoxy-*O*-methyl)-4-(2-nitrophenyl)-butan-1-

ol: The alkene, **3** (236 mg, 1 mmol) was dissolved in anhydrous tetrahydrofuran (5 mL) and was introduced borane dimethyl sulfide (104 μ L, 1.1 mmol) in tetrahydrofuran (1 mL) at 0°C under argon atmosphere. After 4 h of stirring at room temperature, the reaction was cooled to 0°C and 6 mL solution of sodium hydroxide (80 mg, 2 mmol) in ethanol-water (2:1) was added, this was soon followed by the dropwise addition of H₂O₂ (3 mL) such that the addition does not increase the internal temperature of the reaction mixture. Thereafter, the whole of the reaction mixture was diluted with water and extracted with ethyl acetate (2X), all the extracts were pooled together and dried over anhydrous sodium sulphate. Removal of the solvent under the reduced pressure to provide the crude primary alcohol which was purified by means of silica gel chromatography (Ethyl Acetate-Petroleum Benzine; 1:1) to give the primary alcohol, **4** in 80% yield (203 mg). ¹H NMR (200 MHz, CDCl₃): δ 1.79-1.88 (m, 5H), 3.32 (s, 3H), 3.72 (t, 2H, *J* = 5.97 Hz), 4.51 (ABq, 2H, *J* = 14.47 Hz), 5.20-5.26 (m, 1H), 7.38-7.46 (m, 1H), 7.60-7.68 (m, 1H), 7.75 (dd, 1H, *J* = 1.69, 7.80 Hz), 7.92 (dd, 1H, *J* = 1.26, 8.16 Hz). ¹³C NMR (50 MHz, CDCl₃): δ 29.1, 34.2, 55.9, 62.3, 73.4, 95.2, 124.3, 128.1, 128.3, 133.3, 138.6, 148.3. Anal Calcd. for C₁₂H₁₇NO₅: C, 56.46; H, 6.71; N, 5.49. Found C, 56.54; H, 6.82; N, 5.57.



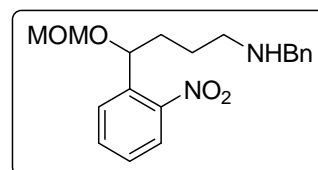
Synthesis of 1-(4-azido-1-(methoxy-*O*-methyl)-butyl)-2-

nitrobenzene: The primary alcohol, **4** (254 mg, 1 mmol) was taken in anhydrous dichloromethane (5 mL) cooled to 0°C and was added dropwise mesyl chloride (116 μ L, 1.5 mmol) was carried out and then allowed to stir at room temperature for 30 min. The solvent was removed under the reduced pressure and then the residue was redissolved in anhydrous DMF (7 mL) and sodium azide (140 mg, 2 mmol) was added. After heating the reaction at 60°C for 6 h, the reaction was diluted with water and compound was obtained by extracting the aqueous layer with ethyl acetate (3X). After the usual processing and purification the primary azide, **5** was obtained in 90% over the two steps. ¹H NMR (200 MHz, CDCl₃): δ 1.76-1.99 (m, 4H), 3.33 (s, 3H), 3.36-3.39 (m, 2H), 4.50 (ABq, 2H, *J* = 21.47 Hz), 5.18-5.23 (m, 1H), 7.39-7.47 (m, 1H), 7.61-7.69 (m, 1H), 7.75 (dd, 1H, *J* = 1.65, 7.82 Hz), 7.94 (dd, 1H, *J* = 1.21, 8.15 Hz). ¹³C NMR (50



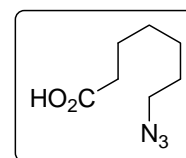
MHz, CDCl₃): δ 25.6, 34.9, 51.1, 57.0, 73.1, 95.2, 124.4, 128.3, 128.4, 133.4, 138.4, 148.2. Anal Calcd. for C₁₂H₁₆N₄O₄: C, 51.42; H, 5.75; N, 19.99. Found C, 51.54; H, 5.66, N, 19.89).

Synthesis of the *N*-benzyl-4-(methoxy-*O*-methyl)-4-(2-nitrophenyl)-butan-1-amine: The primary azide, **5** (1 mmol, 279 mg) was taken in tetrahydrofuran (4 mL), to this was added triphenyl phosphine (262 mg, 1 mmol) in a portion wise manner



followed by the addition of water (27 μ L, 1.5 mmol). As the reaction proceeded, evolution of nitrogen bubbles from the reaction solvent was observed. After the completion of the reaction, as indicated by the TLC examination, the solvent was removed under reduced pressure. The crude amine was redissolved in anhydrous methanol (5 mL) along with benzaldehyde (101 μ L, 1 mmol), the imine formation allowed to take place by stirring it for 4-5 h. It is useful to add a little amount of anhydrous magnesium sulphate into the reaction solvent so as to remove the traces of water present. After 4 h, the reaction vessel was cooled to 0°C and sodium borohydride (37 mg, 1 mmol) was added so as to effect the reduction of the imine formed. The formation of the secondary amine was observed on the TLC examination. Removal of the solvent under reduced pressure yielded the crude amine which was purified by silica gel column chromatography to afford the pure nitro compound, **7** (240 mg, 70%) over the three steps. ¹H NMR (200 MHz, CDCl₃): δ 1.75-1.89 (m, 4H), 2.66-2.73 (m, 1H), 3.33 (s, 3H), 3.40-3.55 (m, 2H), 3.68-3.79 (m, 2H), 4.59 (ABq, 2H, $J = 21.57$ Hz), 5.20-5.26 (m, 1H), 7.23-7.43 (m, 6H), 7.61 (td, 1H, $J = 1.08, 7.13$ Hz), 7.72 (dd, 1H, $J = 1.56, 7.76$ Hz), 7.91 (dd, 1H, $J = 1.16, 8.22$ Hz). ¹³C NMR (50 MHz, CDCl₃): δ 26.4, 35.5, 48.9, 53.9, 59.0, 67.5, 71.6, 73.6, 94.3, 124.3, 127.0, 128.0, 128.2, 128.4, 128.4, 133.3, 138.8, 140.0, 148.3. Anal. Calcd. for C₁₉H₂₄N₂O₄: C, 66.26; H, 7.02; N, 8.13. Found: C, 66.34; H, 6.96; N, 8.22.

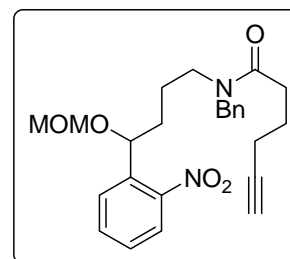
Synthesis of the 6-azido-1-hexanoic acid: Commercially available 6-bromo hex-1-ol (181 mg, 1 mmol) was heated at 60°C with sodium azide (140 mg, 2 mmol) in anhydrous DMF (5 mL). After 6 h, the reaction mixture was diluted with water and extracted with diethyl ether (3X), the combined



extracts upon the dehydration over anhydrous sodium sulphate gave the crude azido alcohol. This was purified over silica gel chromatography to yield the desired compound **9** (140 mg, 98%). The azido alcohol, **9** (143 mg, 1mmol) was dissolved in acetone (5 mL) and further cooled in an ice bath after which the drop wise addition of the Jones reagent (1.2 mL, 1.5 mmol) was carried out. The completion of the reaction was found to occur after 2 h. The

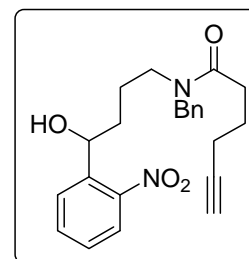
reaction mixture was diluted with water and the consequently extracted with diethyl ether (3X) to yield the crude azido acid. Further the purification of the crude acid afforded the pure side chain, **10** (125 mg, 80%). Whole spectrum features matched with those of the reported values.³¹

Synthesis of *N*-benzyl-*N*-(4-(methoxy-*O*-methyl)-4-(2-nitrophenyl)-butyl)-hex-5-ynamine: The free amine, **7** (343 mg, 1mmol), was dissolved in anhydrous CH₂Cl₂ (5 mL) along with 5-hexynoic acid (112 mg, 1 mmol) prepared *vide supra* and cooled to 0°C in an ice-bath. *N,N*-diisopropyl carbodimide (311 μL, 1.2 mmol) was introduced into the reaction mixture so as to effect the



coupling of the two molecules, the reaction mixture was allowed to stir at room temperature until the completion. After the completion of the reaction, the solvent was removed under the reduced pressure and the crude residue was purified by silica gel column chromatography (Ethyl Acetate-Petroleum Benzine; 2:3) to furnish desired amide **13** (393 mg, 90%). ¹H NMR (200 MHz, CDCl₃): δ 1.71-1.99 (m, 8H), 2.19-2.35 (m, 2H), 2.43-2.59 (m, 2H), 3.27 (s, 3H), 3.43-3.44 (m, 1H), 4.47 (ABq, 2H, *J* = 20.97 Hz), 4.59 (d, 2H, *J* = 9.73 Hz), 5.13 (m, 1H), 7.15-7.48 (m, 6H), 7.55-7.73 (m, 2H), 7.88-7.97 (m, 1H). ¹³C NMR (50 MHz, CDCl₃): δ 17.8, 23.9, 25.3, 31.6, 35.1, 45.7, 46.8, 48.2, 50.8, 55.9, 69.1, 73.5, 83.8, 95.2, 124.5, 126.2-128.5, 133.5, 136.9, 137.7, 138.3, 148.2, 172.7. Anal Calcd. for C₂₅H₃₀N₂O₅: C, 68.47; H, 6.90; N, 6.39. Found C, 68.58; H, 6.97; N, 6.43.

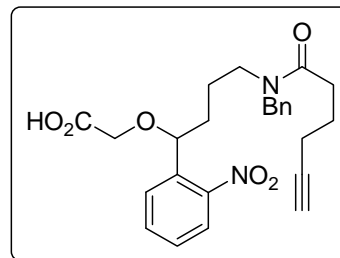
Synthesis of *N*-benzyl-*N*-(4-hydroxy-4-(2-nitrophenyl) butyl) hex-5-ynamide: The amide, **13** (437 mg, 1 mmol) was refluxed in methanol (5 mL) in the presence of HCl (cat.) until the completion of the reaction as evident by TLC examination. Thereafter, the reaction was diluted with water and extracted with ethyl acetate (3X) which were



pooled together and dried over anhydrous sodium sulphate. Solvent removal under reduced pressure and subsequent purification over the silica gel column yielded the free alcohol, **14** (372 mg, 95%). ¹H NMR (200 MHz, CDCl₃): δ 1.70-1.97 (m, 6H), 2.20-2.32 (m, 2H), 2.45-2.57 (m, 2H), 2.85 (d, 1H, *J* = 3.88 Hz), 3.21-3.31 (m, 1H), 3.66-3.80 (m, 1H), 4.10-4.14 (m, 1H), 4.56-4.61 (m, 2H), 5.21-5.31 (m, 1H), 7.15-7.46 (m, 6H), 7.58-7.68 (m, 1H), 7.77-7.94 (m, 2H). ¹³C NMR (50 MHz, CDCl₃): δ 17.8, 23.8, 24.6, 31.6, 35.0, 45.7,

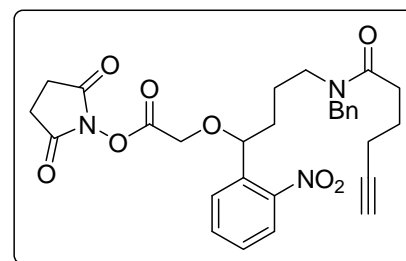
48.2, 51.1, 69.1, 83.6, 124.2, 126.2-128.9, 133.5, 136.5, 140.3, 140.9, 147.4, 173.2. Anal Calcd. for $C_{23}H_{26}N_2O_4$: C, 70.03; H, 6.64; N, 7.10. Found C, 70.12; H, 6.75; N, 7.19.

Synthesis of compound 16: To an ice-cooled solution of alcohol, **14** (392 mg, 1mmol) in anhydrous DMF (5 mL) was added sodium hydride (44 mg, 1.1 mmol, 60% oil suspension) and stirred for 1 h at room temperature. Bromoethylacetate (220 μ L, 2 mmol) was introduced dropwise at 0°C and stirred at room temperature for 1 h. The resulting suspension was



quenched with saturated ammonium chloride and extracted with diethyl ether (3X). The combined extracts were dried over anhydrous sodium sulphate and the solvent was removed *in vacuo* and the resulting residue was directly used for the next step for hydrolysis of the ethyl ester. For the purpose, the alkylated product, **15** (478 mg, 1 mmol) was dissolved in a mixture of THF/water (1:1, 5 mL) was added sodium hydroxide solution (1 M, 2 mL) and allowed to stir until the completion of the reaction, 5 h. After the completion of the reaction, the reaction mixture was diluted with water and extracted with ethyl acetate (3X) to offer the crude acid as a syrup after the evaporation of the solvent under reduced pressure. Purification achieved *via* the silica gel column chromatography to afford the pure acid, **16** (384 mg, 75% over the two steps). 1H NMR (200 MHz, $CDCl_3$): δ 0.83-0.88 (m, 1H), 1.14-1.30 (m, 3H), 1.79-1.95 (m, 7H), 2.22-2.25 (m, 2H), 2.48-2.61 (m, 2H), 3.27-3.48 (m, 1H), 3.37-4.14 (m, 2H), 4.58-4.63 (m, 2H), 5.03-5.06 (m, 1H), 6.03 (b, 1H), 7.15-7.50 (m, 6H), 7.60-7.87 (m, 2H), 7.91-7.99 (m, 1H). ^{13}C NMR (50 MHz, $CDCl_3$): δ 14.1, 17.8, 23.2, 23.8, 25.0, 31.3, 34.8, 42.5, 45.7, 46.9, 48.2, 50.9, 60.4, 66.1, 69.0, 83.8, 124.7, 126.2, 127.9-128.9, 133.8, 136.6, 137.6, 148.5, 172.6, 173.2. Anal. Calcd for $C_{29}H_{31}N_3O_8$: C, 63.38; H, 5.69; N, 7.65. Found C, 63.47; H, 5.76; N, 7.73.

Synthesis of the compound 17: The free acid, **16** (513 mg, 1 mmol) was stirred along with *N*-hydroxy succinamide (230 mg, 1.2 mmol) in anhydrous CH_2Cl_2 (5 mL). *N,N*-diisopropyl carbodimide (186 μ L, 1.2 mmol) was introduced into the reaction mixture at 0°C and the resulting mixture was made to stir at room temperature for



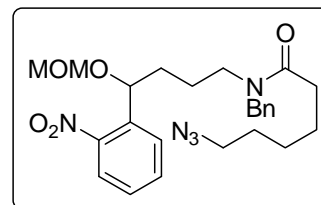
2 h. After the completion of the reaction, any attempts to further purify the active ester were foiled by the degradation of the compound inside the silica gel column. Hereafter, the reaction

was then cooled at 0°C so as to precipitate the insoluble unwanted side product (urea) and the resulting upper layer was carefully decanted to obtain the alkyne functionalized linker which was found to be pure enough to be utilized for the biological experiments.

Synthesis of 6-azido-*N*-benzyl-*N*-(4-(methoxy-*O*-methyl)-4-(2-

nitrophenyl) butyl) hexanamide: The synthetic protocol was same as followed for the compound **13**. ¹H NMR (200 MHz, CDCl₃): δ 1.52-1.78 (m, 10H), 2.28-2.45 (m, 2H), 3.19-3.32 (m, 6H), 3.44 (m, 1H), 4.40-4.62 (m, 4H), 5.13 (m, 1H), 7.14-7.48

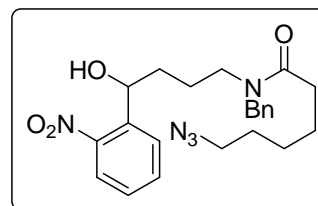
(m, 6H), 7.57-7.74 (m, 2H), 7.88-7.97 (m, 1H). ¹³C NMR (50 MHz, CDCl₃): δ 23.2- 26.5, 28.9, 32.8, 35.1, 42.0, 45.7, 46.7, 48.1, 51.1, 55.8, 73.5, 95.2, 124.2, 124.4, 126.0, 127.8-128.8, 133.3, 136.9, 137.7, 138.2, 148.1, 172.7. Anal Calcd. for C₂₅H₃₃N₅O₅: C, 62.10; H, 6.88; N, 14.48. Found C, 62.19; H, 6.75; N, 14.56.



Synthesis of 6-azido-*N*-benzyl-*N*-(4-hydroxy-4-(2-

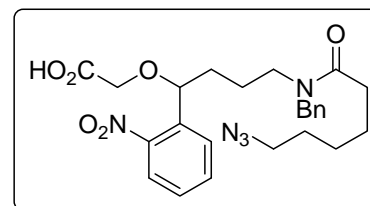
nitrophenyl)-butyl) hexanamide: Synthetic procedure same as followed for the synthesis of the compound **14**. ¹H NMR (200 MHz, CDCl₃): δ 1.18-1.81 (m, 10H), 2.29-2.42 (m, 2H), 3.19-3.31 (m, 3H), 3.66-3.76 (m, 1H), 4.10-4.13 (m, 1H), 4.54 (d, 1H,

$J = 1.90$ Hz), 4.58 (d, 1H, $J = 3.93$ Hz), 5.26 (t, 1H, $J = 8.94$ Hz), 7.13-7.46 (m, 6H), 7.58-7.68 (m, 1H), 7.77-7.94 (m, 2H). ¹³C NMR (50 MHz, CDCl₃): δ 24.8, 26.4, 28.7, 32.9, 33.1, 35.0, 45.8, 46.8, 51.1, 51.3, 68.9, 69.3, 124.3, 126.1-129.0, 133.5, 136.7, 141.0, 147.5, 173.8. Anal Calcd. for C₂₃H₂₉N₅O₄: C, 62.85; H, 6.65; N, 15.93. Found C, 62.93; H, 6.76; N, 15.86.



Synthesis of 21: Synthetic procedure same as followed for the compound **16**. ¹H NMR (200 MHz, CDCl₃): δ 1.22-1.77 (m, 10H), 2.29-2.47 (m, 2H), 3.16-3.30 (m, 3H), 3.58-4.18 (m, 3H), 4.54-4.62 (m, 2H), 5.04 (m, 1H), 6.68 (b, 1H), 7.12-7.45 (m, 6H), 7.59-7.72 (m, 2H), 7.90-7.98 (m, 1H).

¹³C NMR (50 MHz, CDCl₃): δ 23.9, 24.7, 25.0, 26.3, 28.5, 32.7, 33.0, 35.0, 46.0, 46.9, 48.2, 51.1, 51.2, 124.5, 126.1, 127.9-128.9, 133.9, 136.5, 137.6, 148.4, 173.4. Anal. Calcd. for C₂₅H₃₁N₅O₆: C, 60.35; H, 6.28; N, 14.08. Found: C, 60.46; H, 6.35; N, 14.17.



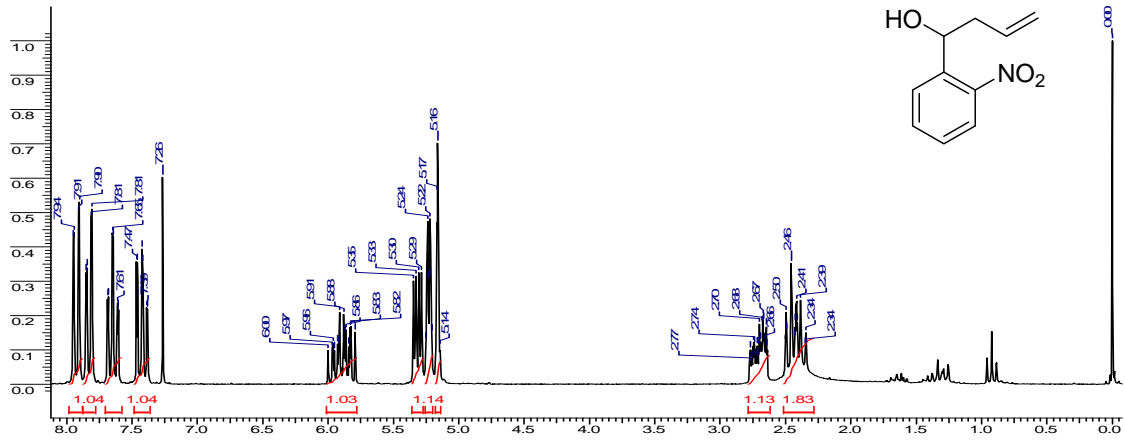
Synthesis of 22: Synthetic procedure same as followed for the compound **17**.

Stock solution of BSA/Insulin: A 4 mg/mL stock solution of the protein was prepared.

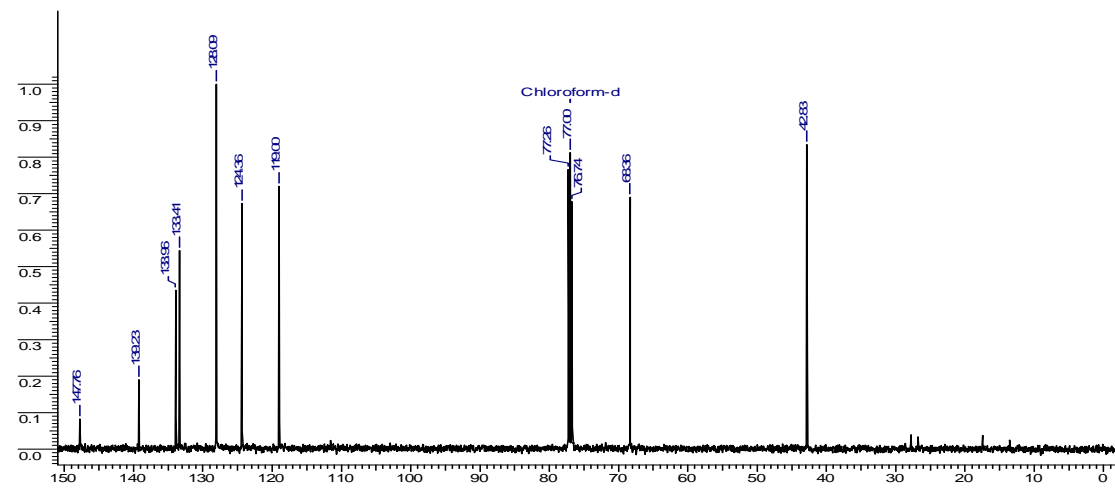
General procedure for the bioconjugation of the linkers on the protein: 1 mg (10 mmol) of the functionalized linker was dissolved in DMSO (200 μ L) this was added to the stock solution of the protein (800 μ L), which was allowed to rotate on a rotary shaker for 2 hr. The total concentration of the DMSO in the reaction was maintained as such that it doesn't exceed more than 20 percent of the total volume at any time. The protein samples were then dialysed overnight in 0.1 M tris-HCl buffer (pH 8.1), with the buffer being replaced three times at regular intervals. Thereafter, the protein samples were poured into an eppendroff and directly utilized for MALDI-TOF analysis.

Chapter 3: Spectral Charts

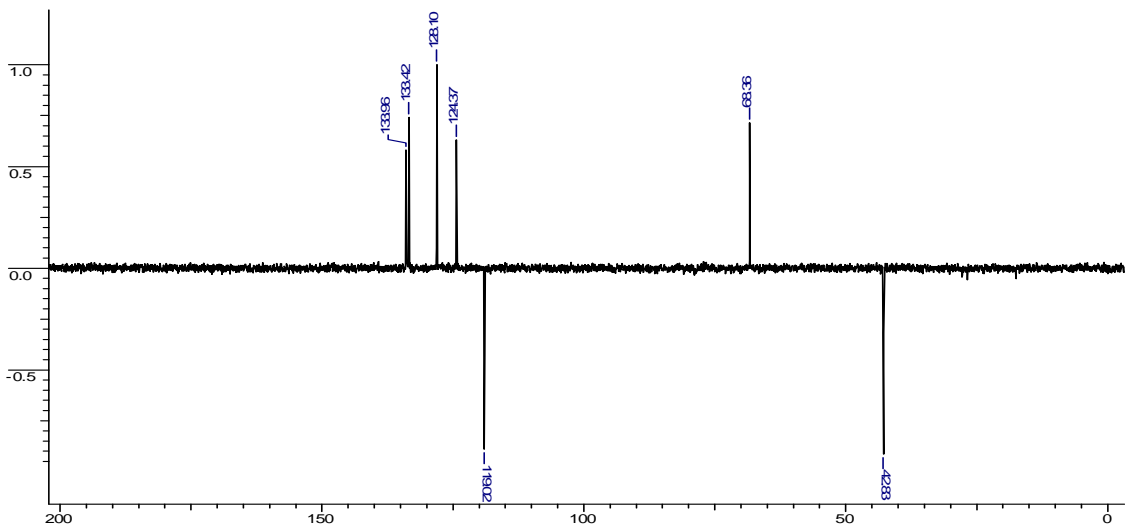
^1H NMR (200 MHz, CDCl_3) of Compound 2



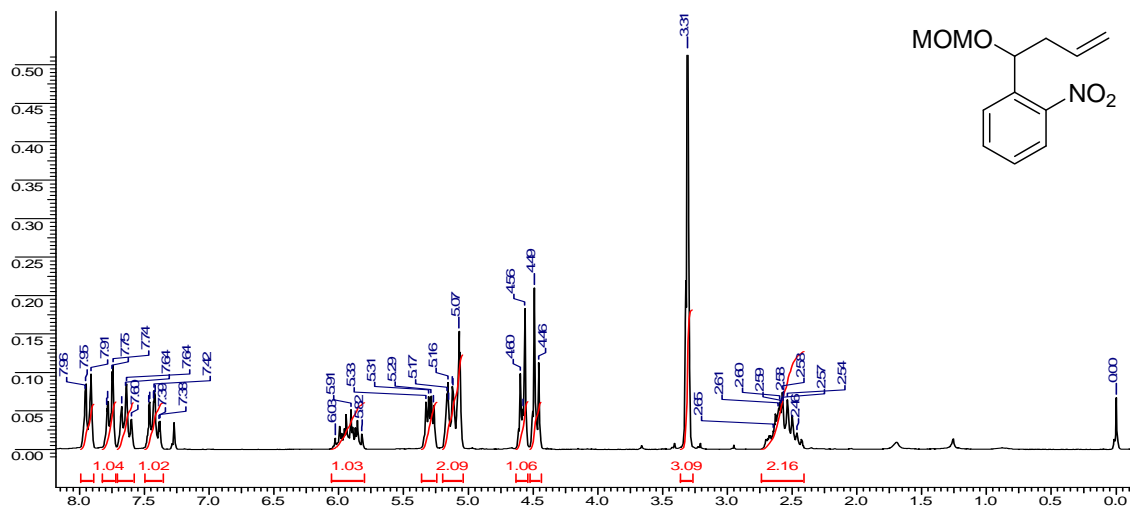
^{13}C NMR (50 MHz, CDCl_3) of Compound 2



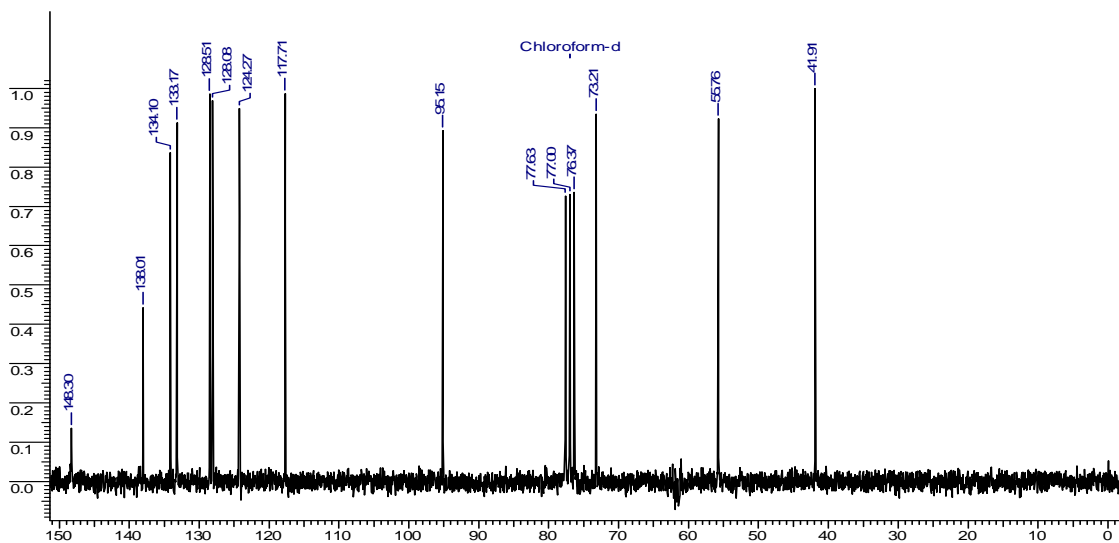
DEPT NMR (50 MHz, CDCl_3) of Compound 2



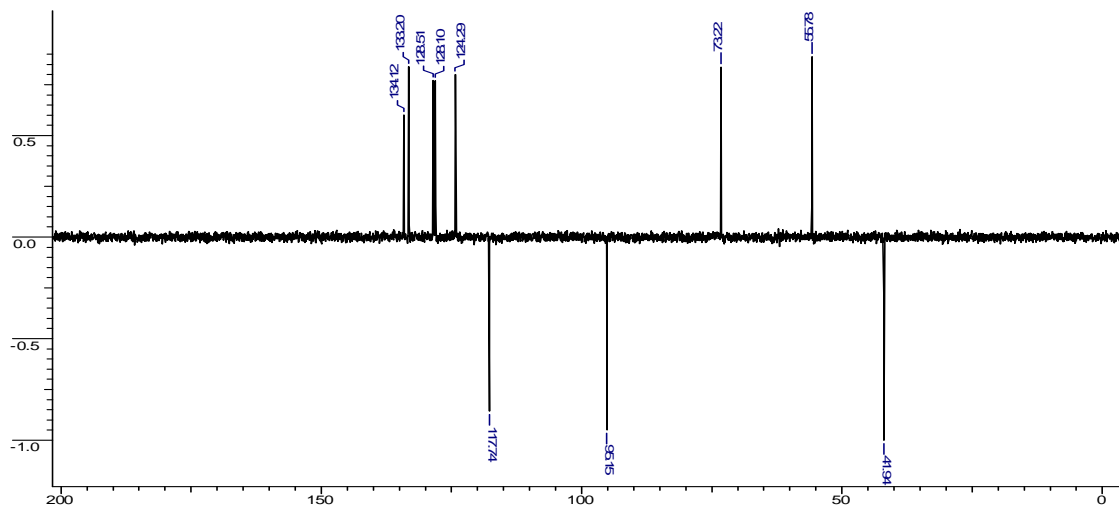
^1H NMR (200 MHz, CDCl_3) of Compound 3



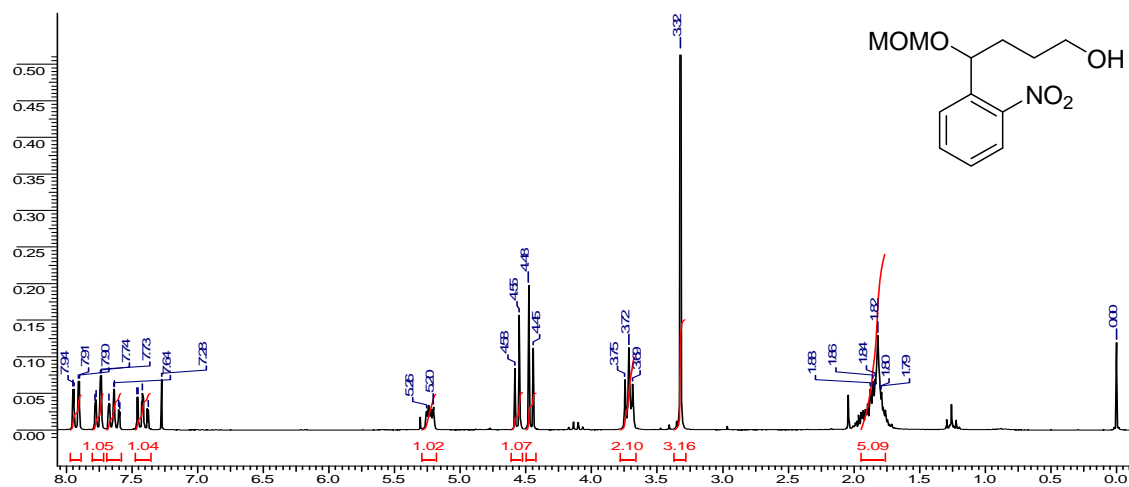
^{13}C NMR (50 MHz, CDCl_3) of compound 3



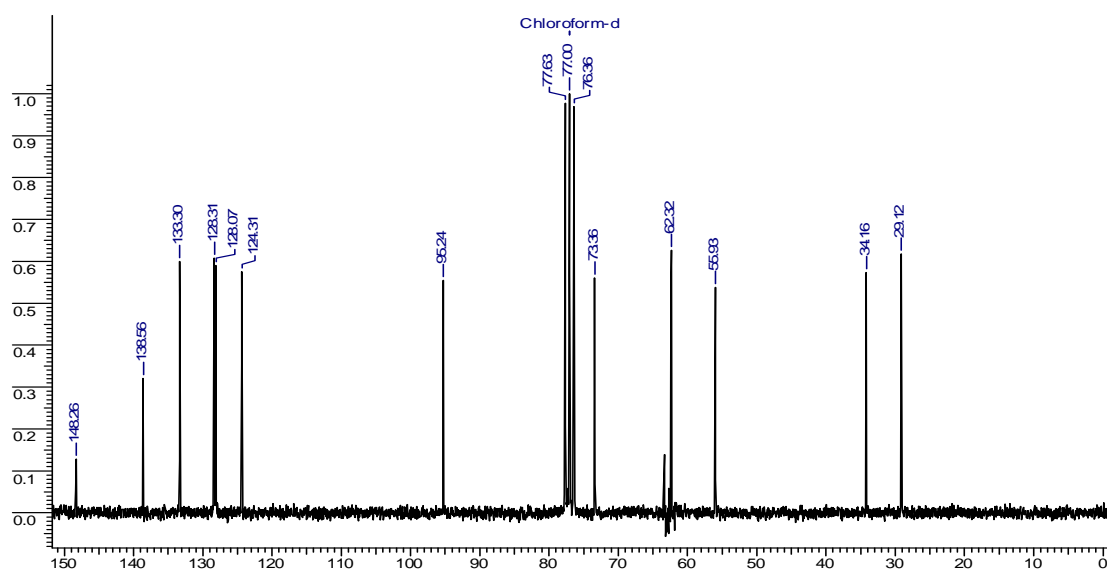
DEPT NMR (50 MHz, CDCl_3) of Compound 3



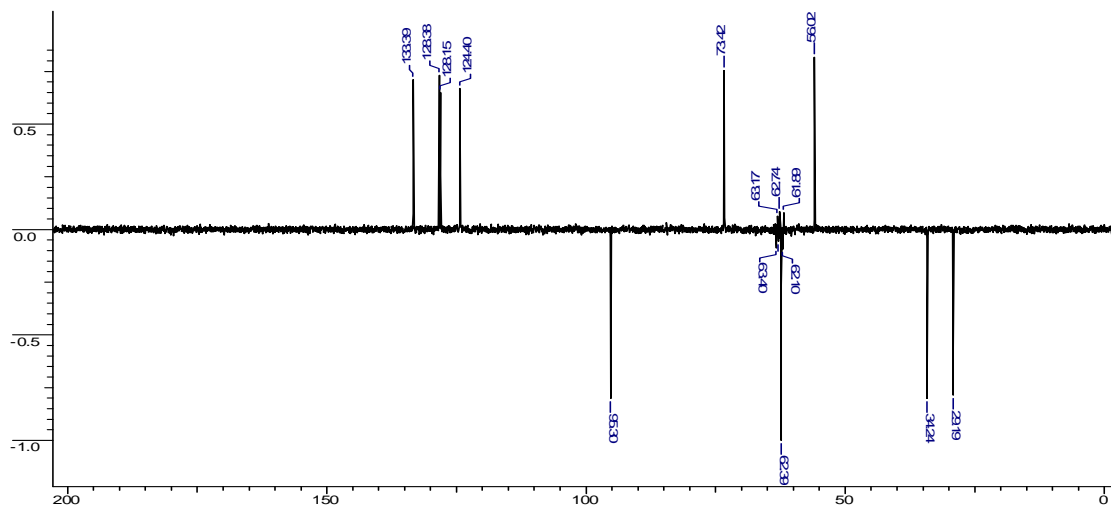
¹H NMR (200 MHz, CDCl₃) of Compound 4



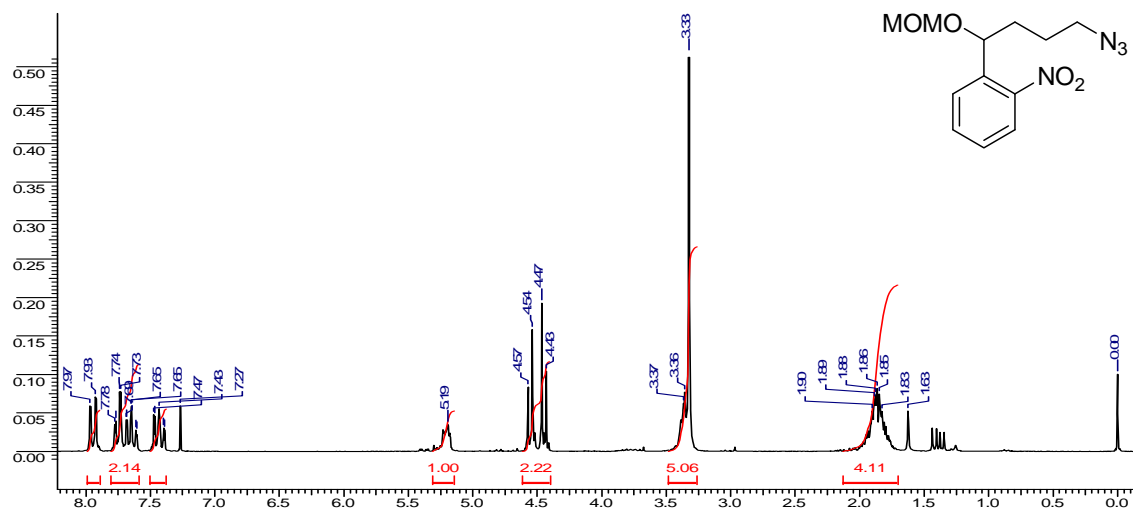
¹³C NMR (50 MHz, CDCl₃) of Compound 4



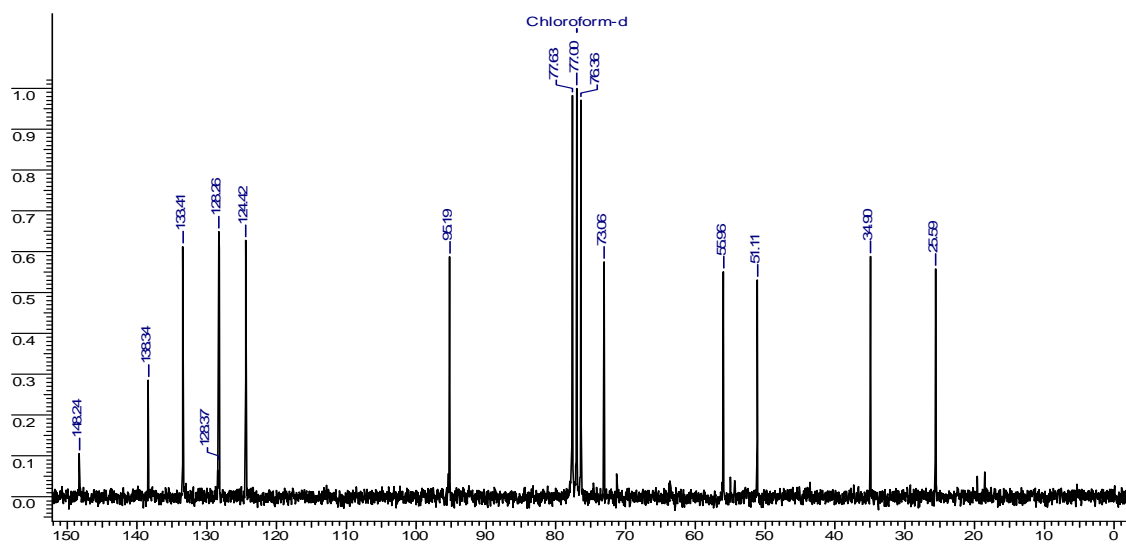
DEPT NMR (50 MHz, CDCl₃) of Compound 4



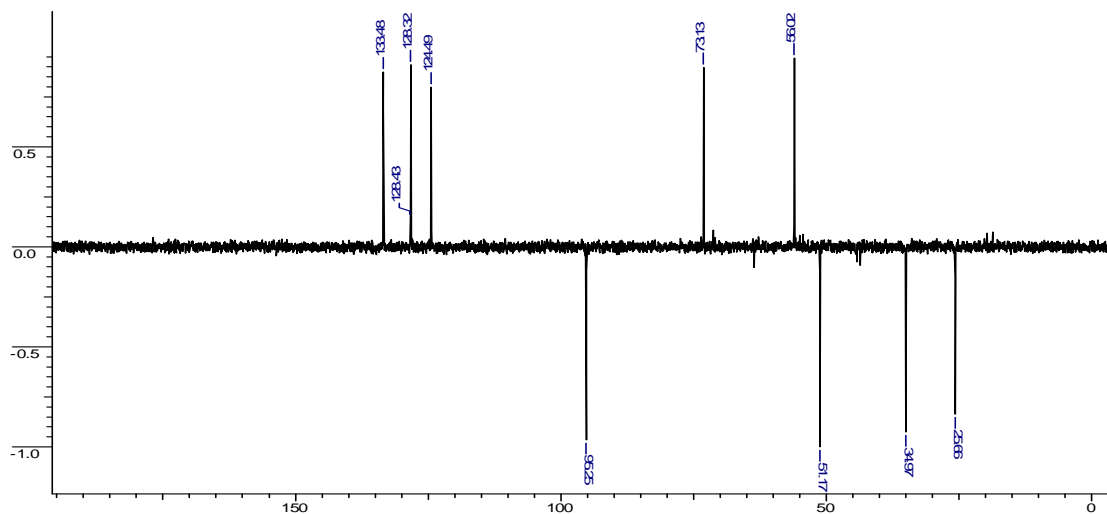
¹H NMR (200 MHz, CDCl₃) of Compound 5



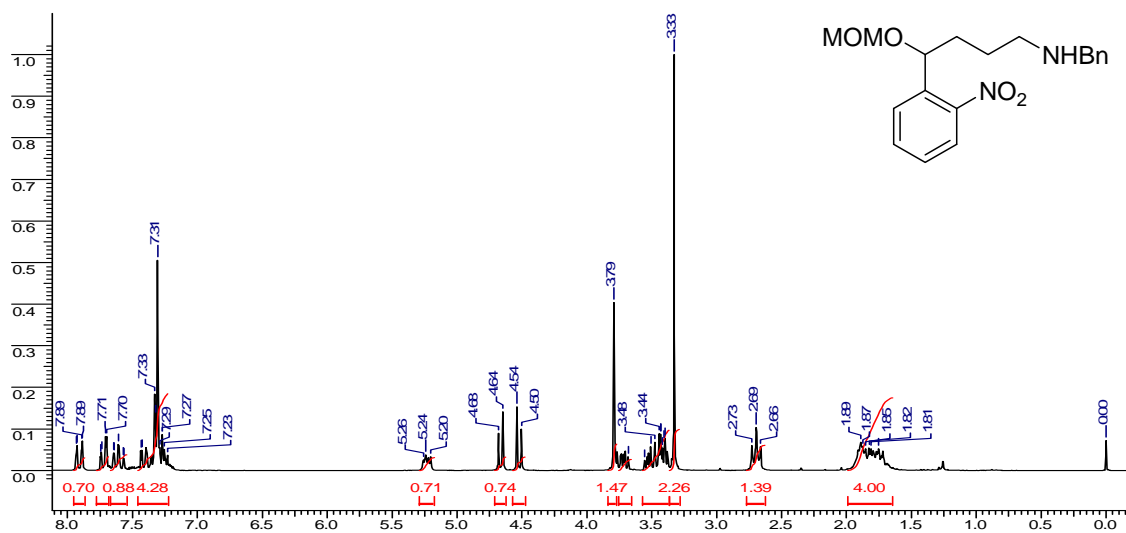
¹³C NMR (50 MHz, CDCl₃) of Compound 5



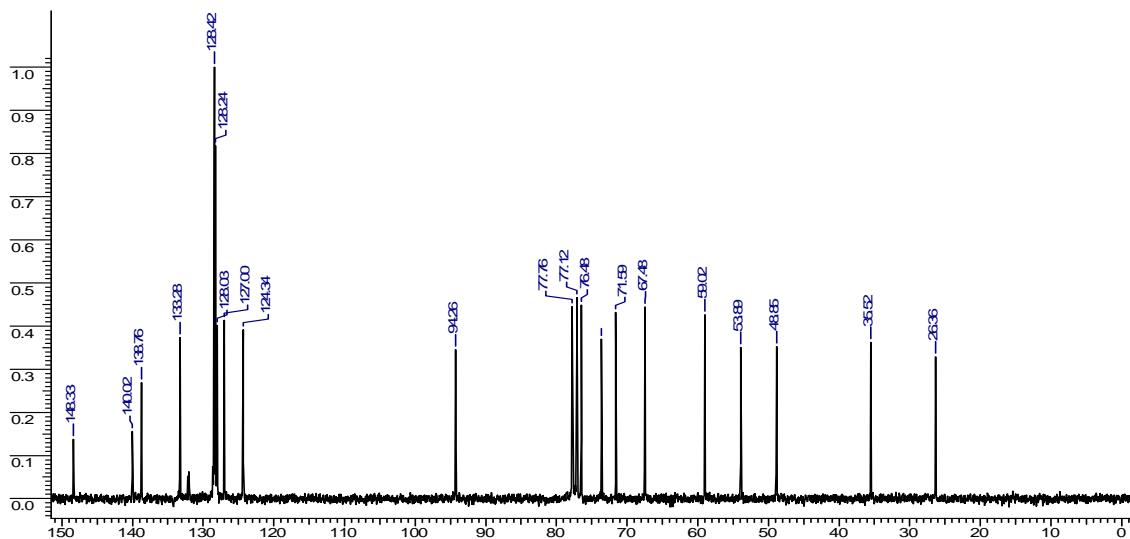
DEPT NMR (50 MHz, CDCl₃) of Compound 5



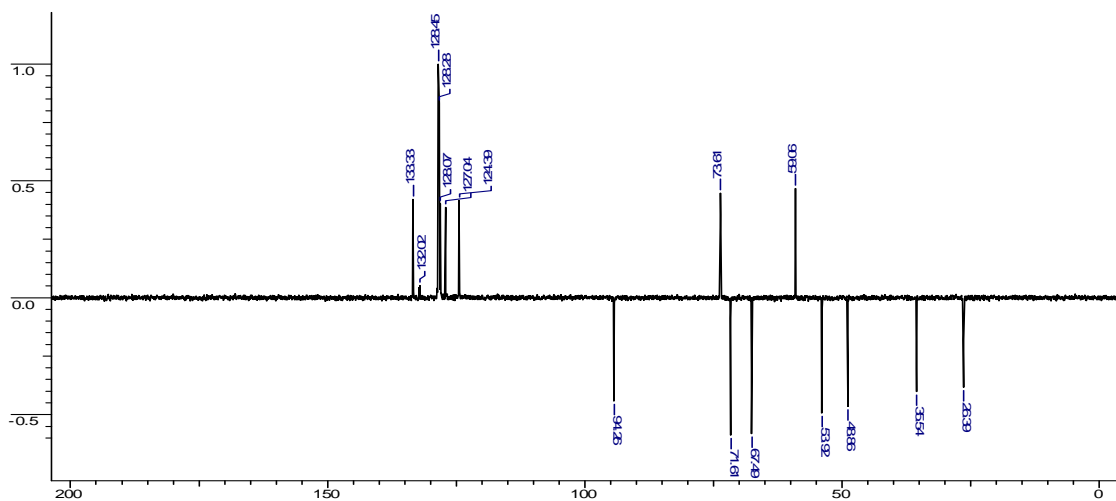
¹H NMR (200 MHz, CDCl₃) of Compound 7



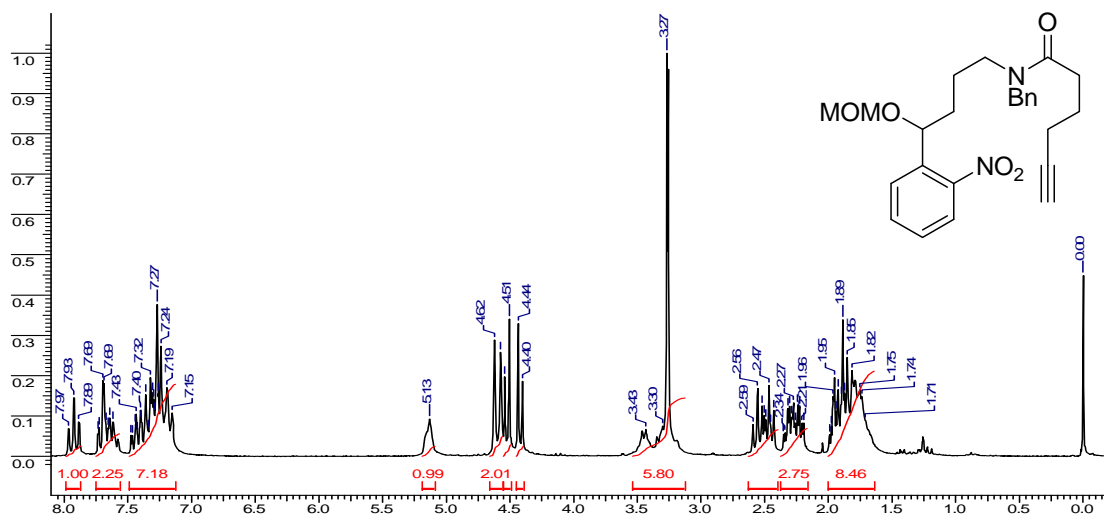
¹³C NMR (50 MHz, CDCl₃) of Compound 7



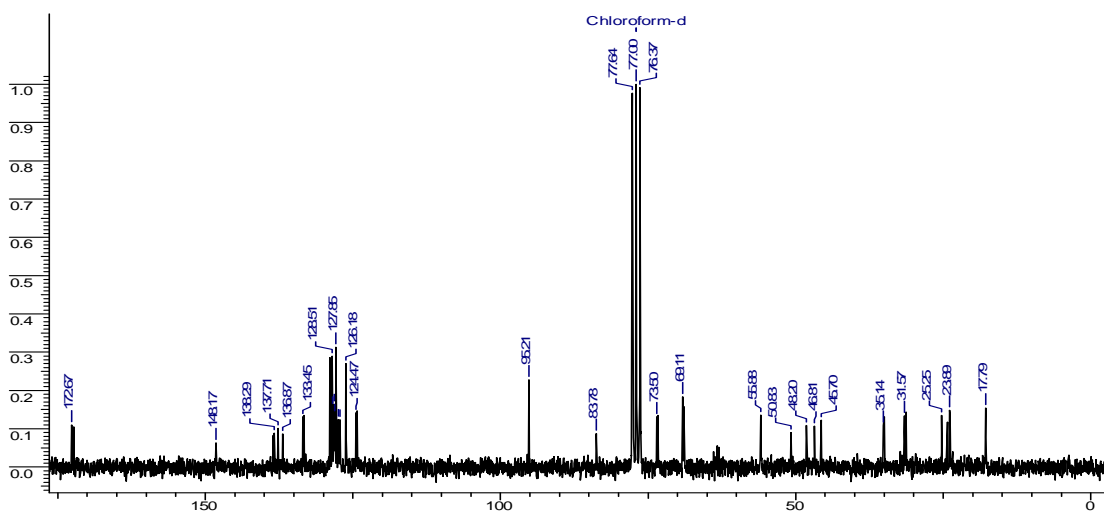
DEPT NMR (50 MHz, CDCl₃) of Compound 7



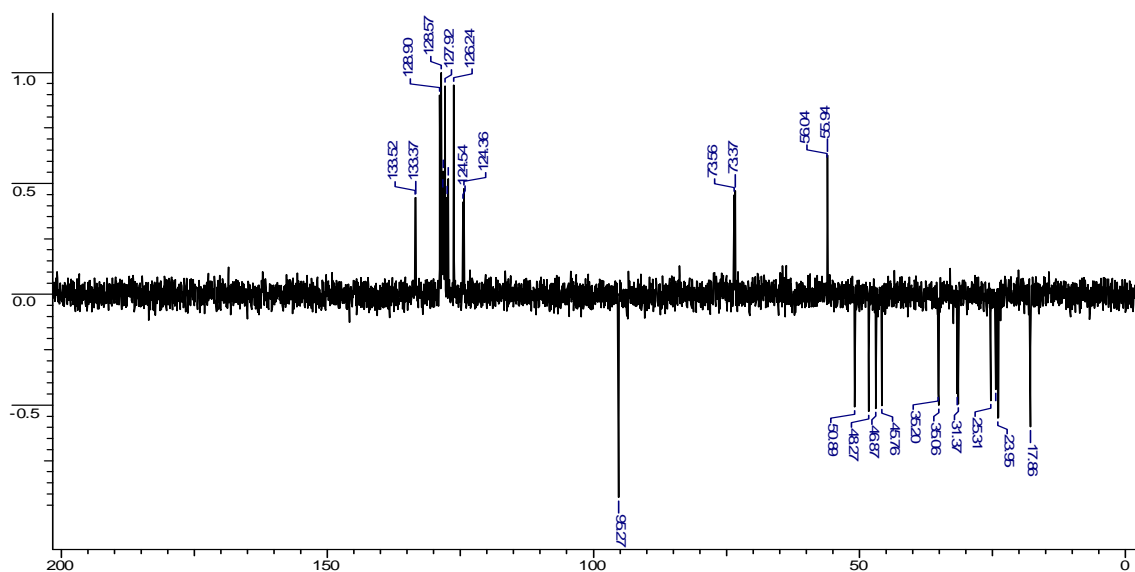
^1H NMR (200 MHz, CDCl_3) of Compound 13



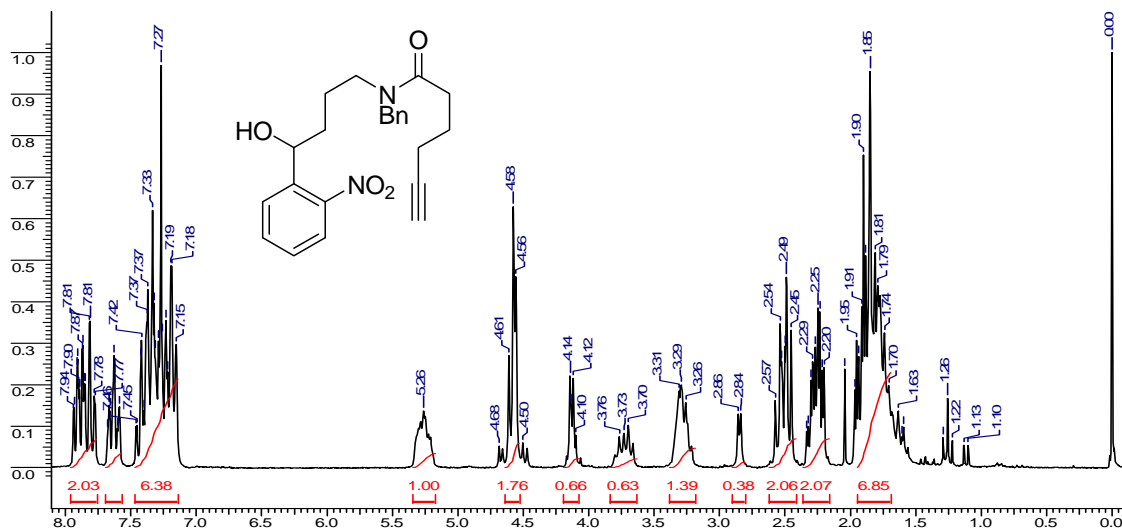
^{13}C NMR (50 MHz, CDCl_3) of Compound 13



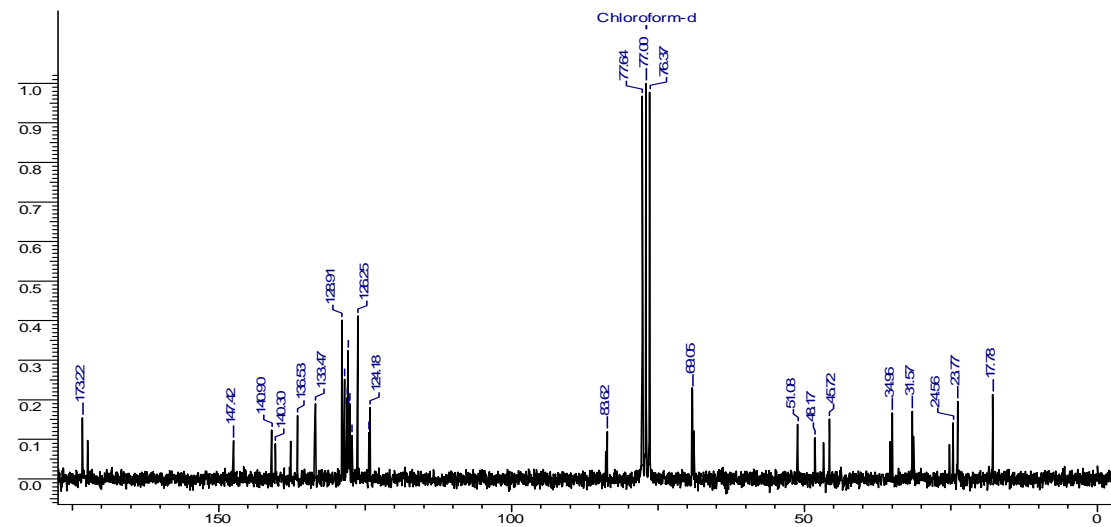
DEPT NMR (50 MHz, CDCl_3) of Compound 13



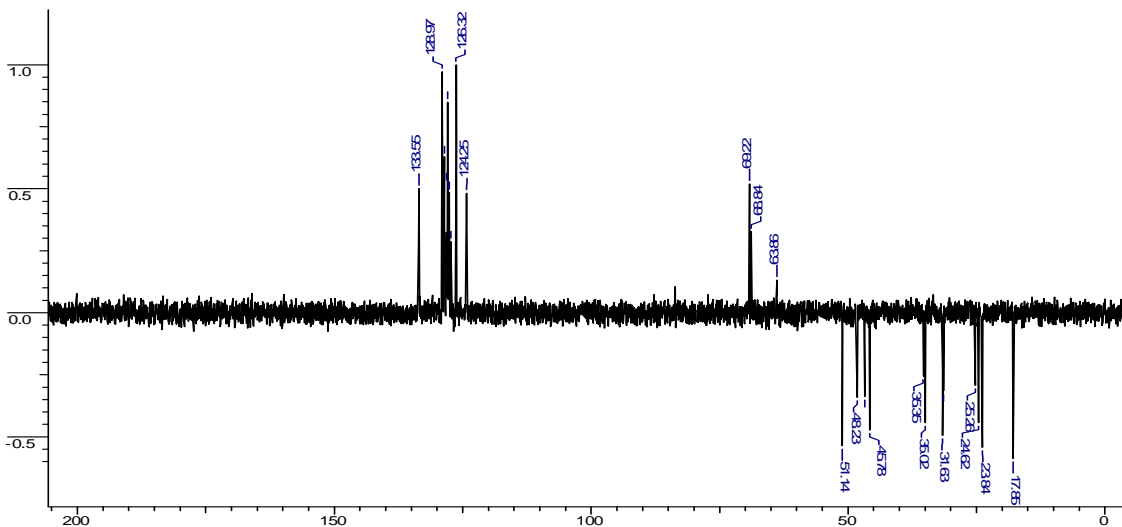
¹H NMR (200 MHz, CDCl₃) of Compound 14



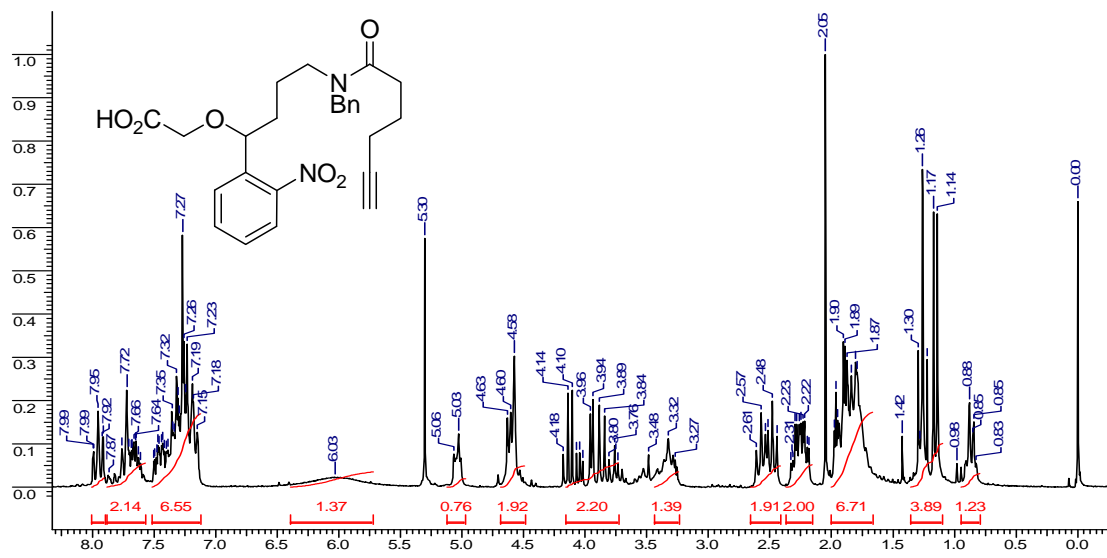
¹³C NMR (50 MHz, CDCl₃) of Compound 14



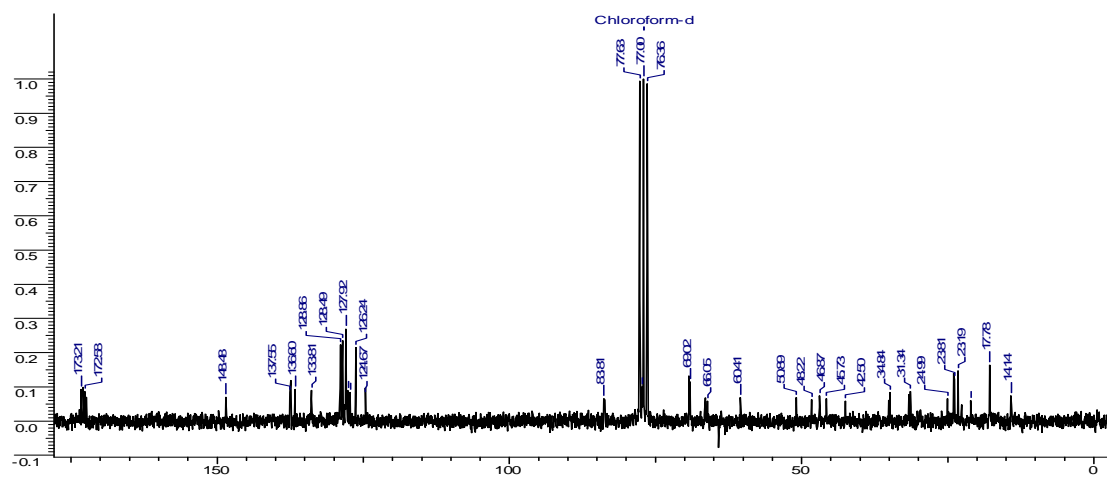
DEPT NMR (50 MHz, CDCl₃) of Compound 14



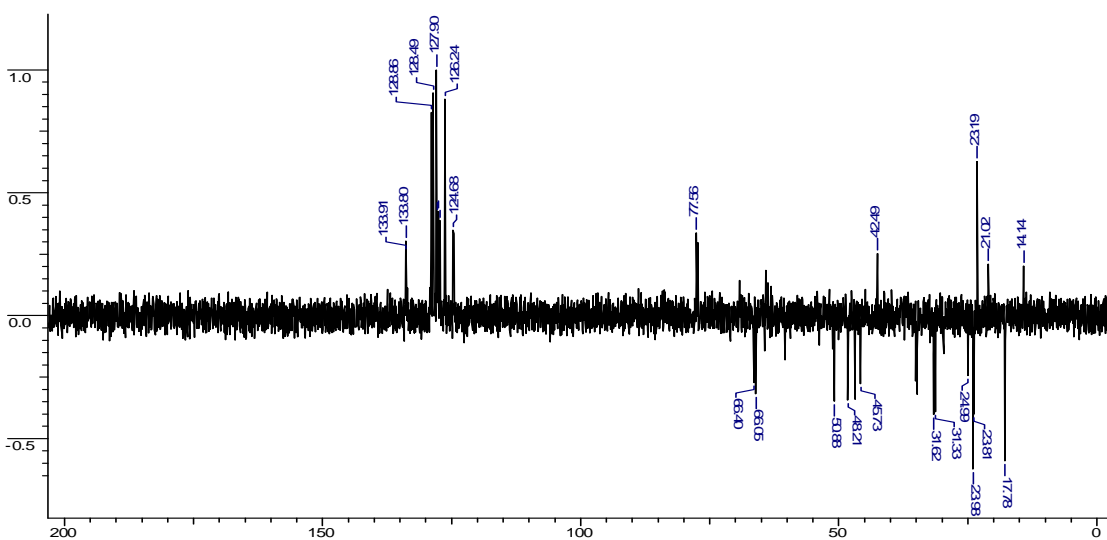
¹H NMR (200 MHz, CDCl₃) of Compound 16



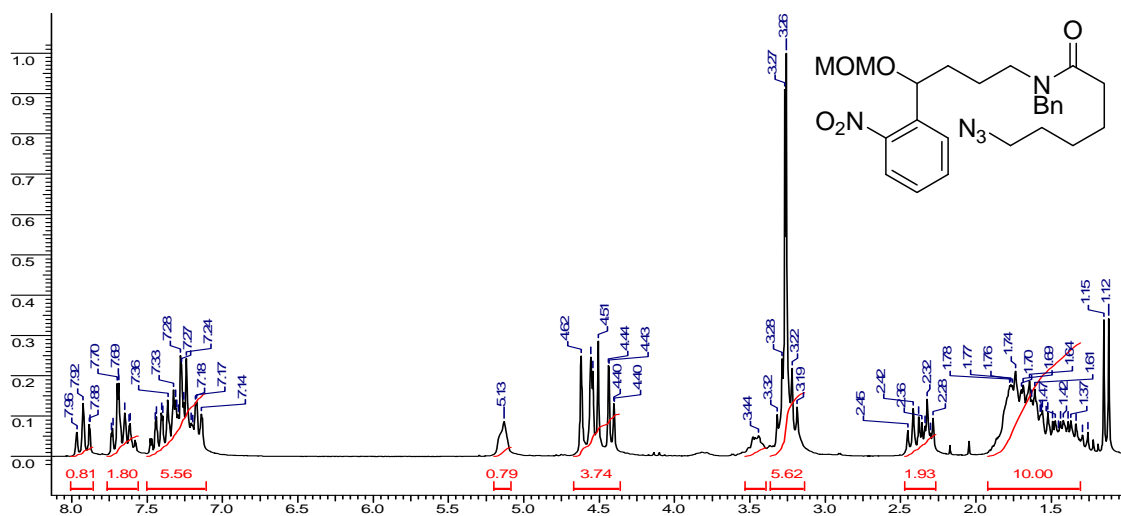
¹³C NMR (50 MHz, CDCl₃) of Compound 16



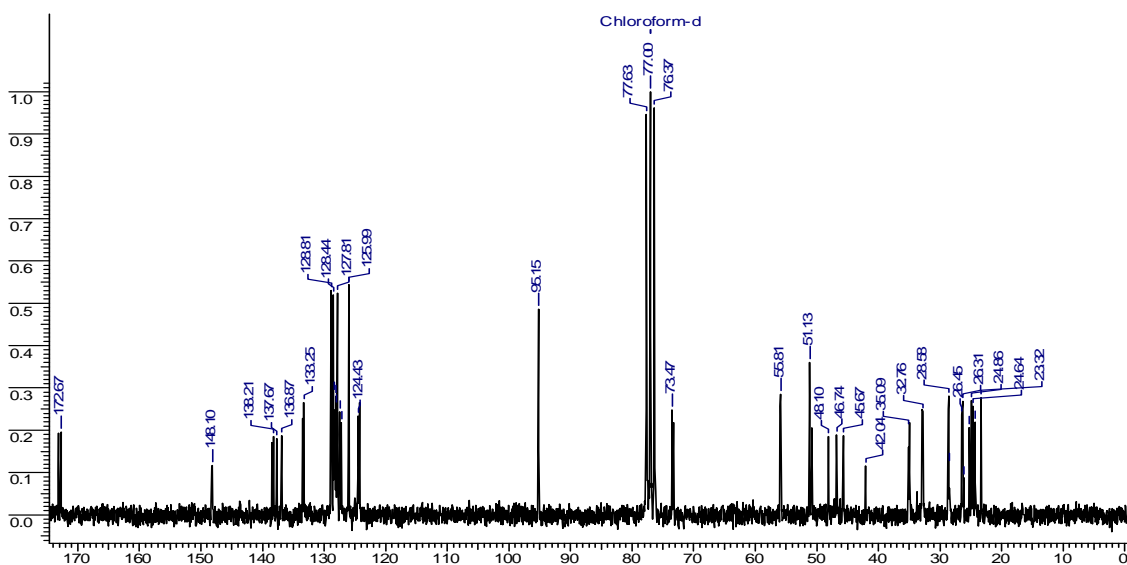
DEPT NMR (50 MHz, CDCl₃) of Compound 16



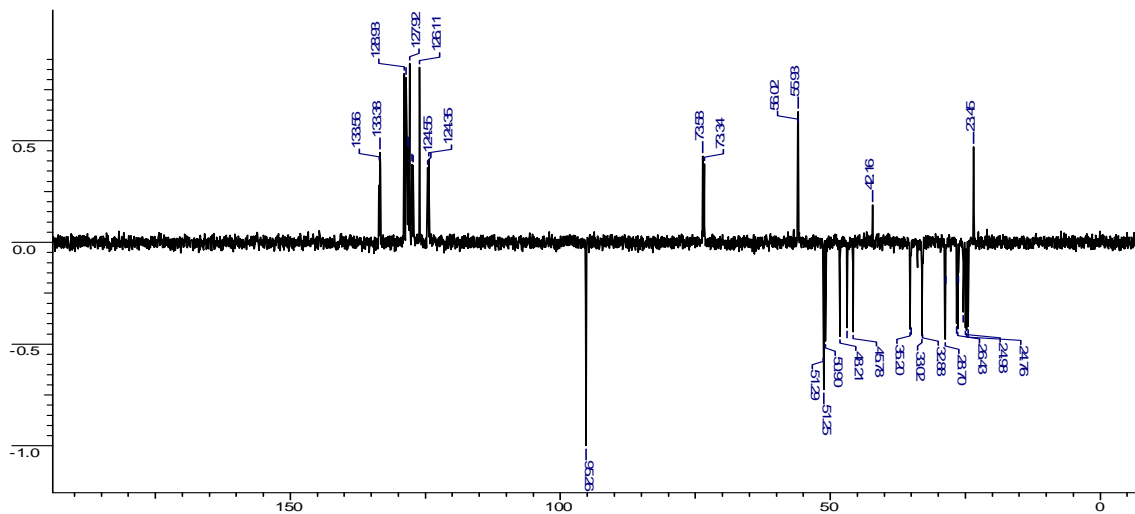
¹H NMR (200 MHz, CDCl₃) of Compound 18



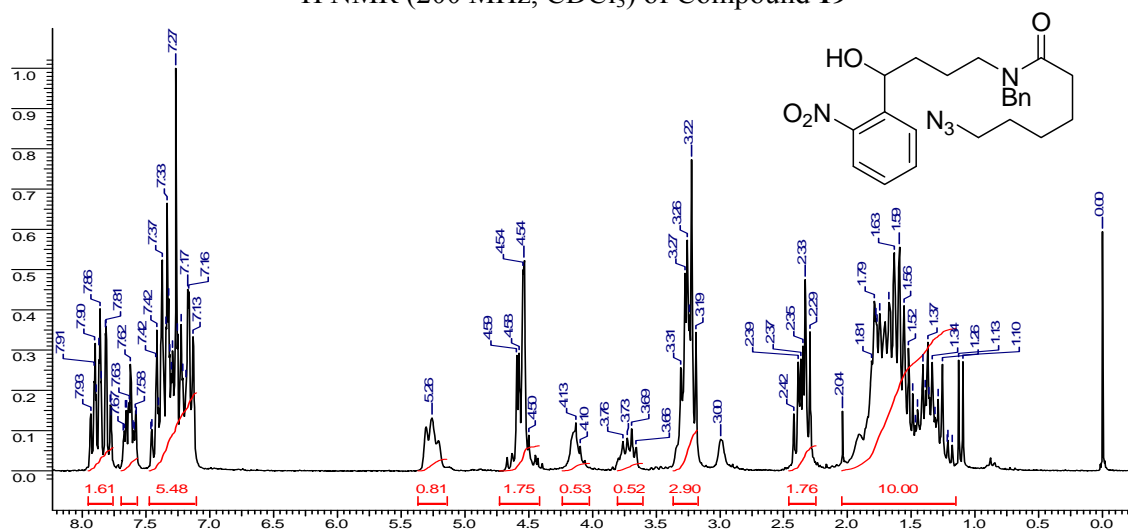
¹³C NMR (50 MHz, CDCl₃) of Compound 18



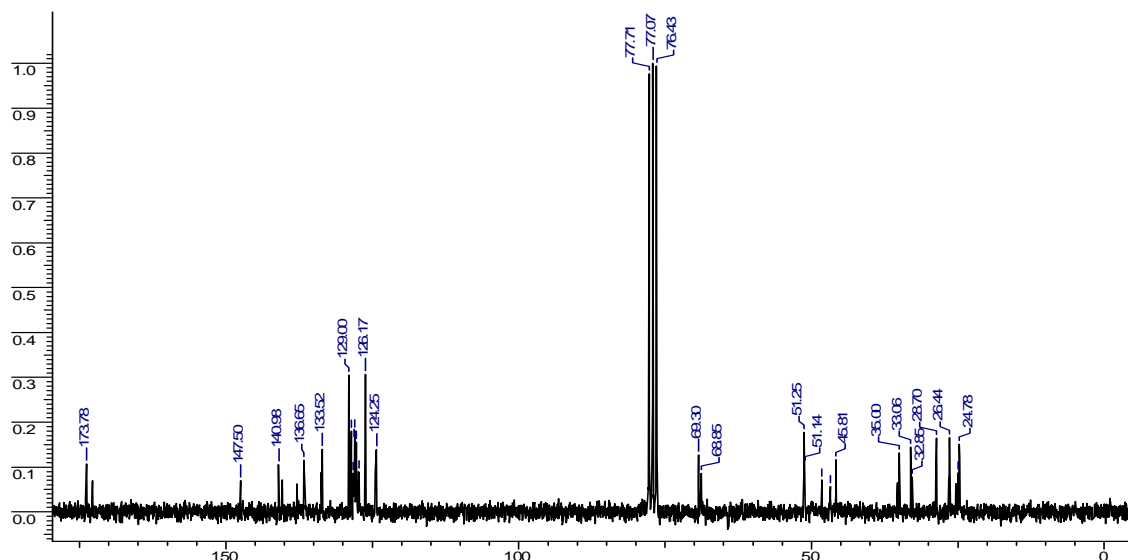
DEPT NMR (50 MHz, CDCl₃) of Compound 18



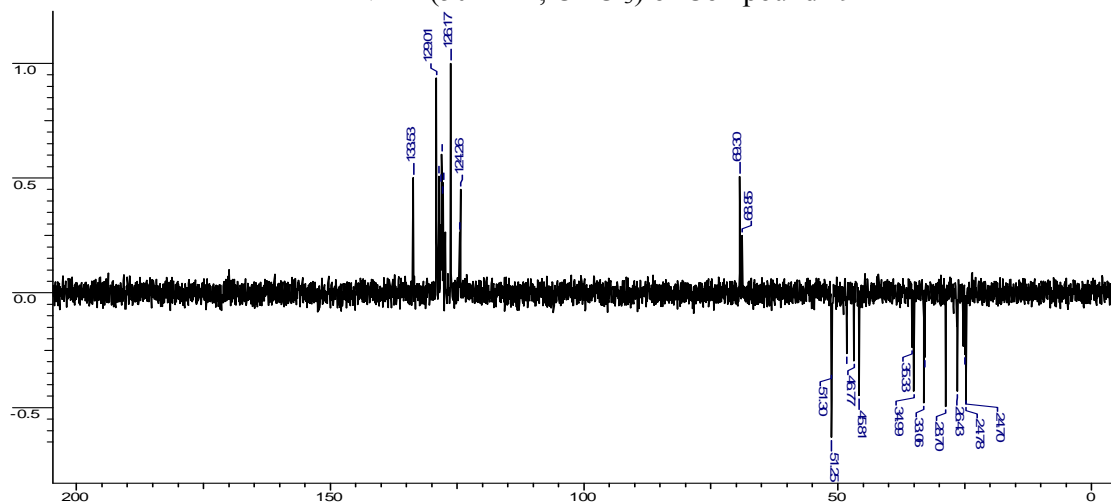
^1H NMR (200 MHz, CDCl_3) of Compound 19



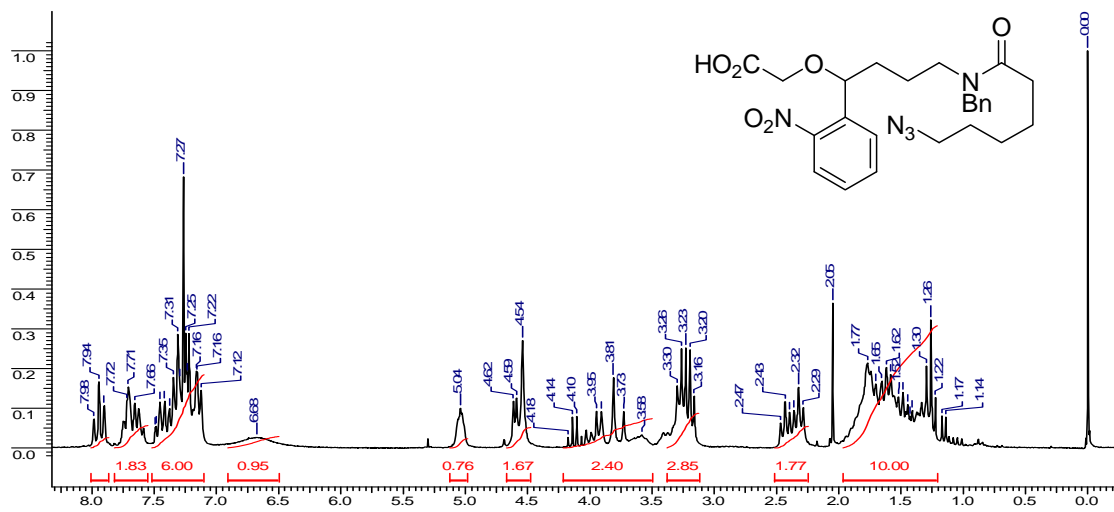
^{13}C NMR (50 MHz, CDCl_3) of Compound 19



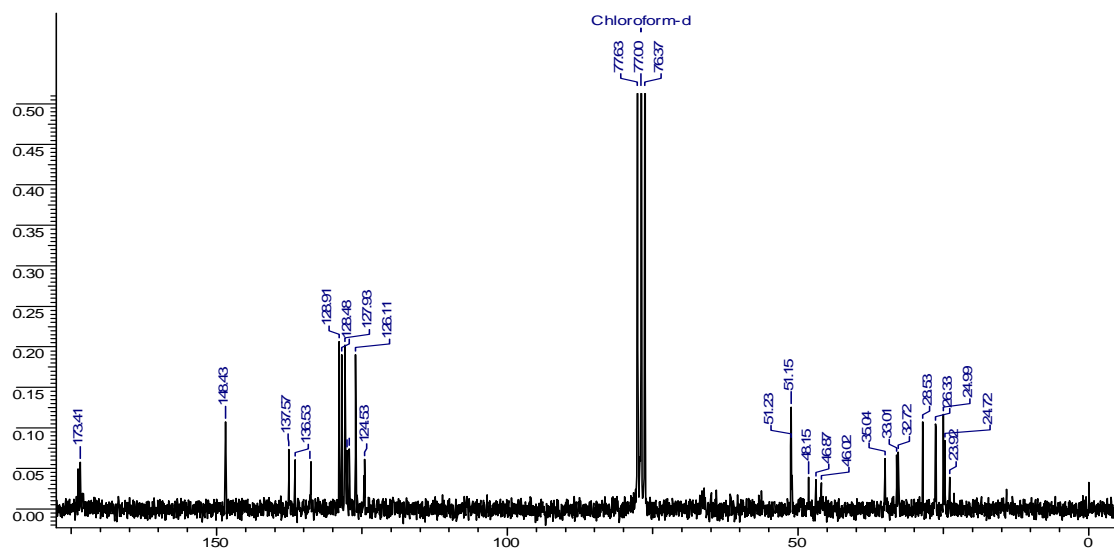
DEPT NMR (50 MHz, CDCl_3) of Compound 19



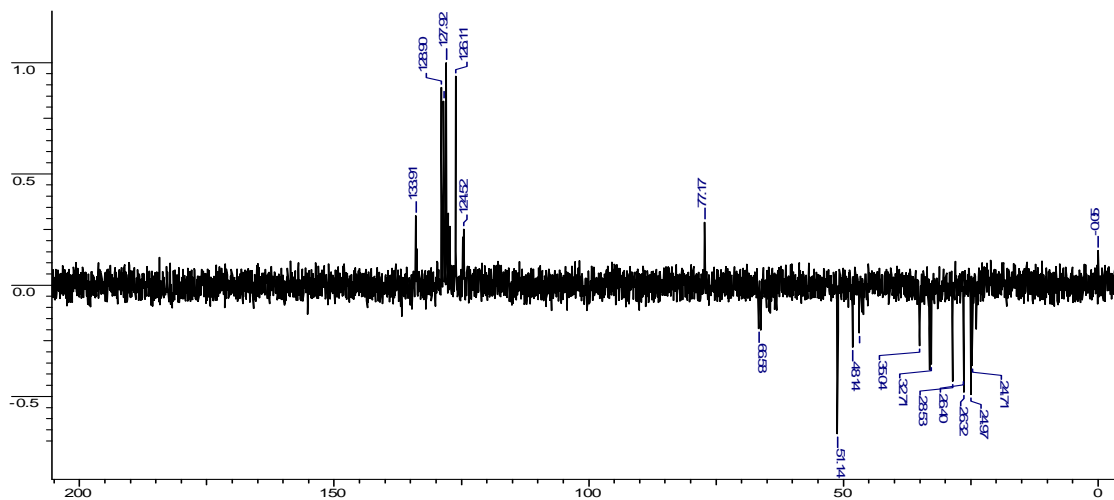
¹H NMR (200 MHz, CDCl₃) of Compound 21



¹³C NMR (50 MHz, CDCl₃) of Compound 21



DEPT NMR (50 MHz, CDCl₃) of Compound 21



Chapter 3: References

1. (a) Abdallah, D. J.; Weiss, R. G. *Adv. Mater.* **2000**, *12*, 1237-1247; (b) Terech, P.; Weiss, R. G. *Chem. Rev.* **1997**, *97*, 3133-3159.
2. Eagon, R. J. *Bacteriol.* **1995**, *83*, 736-737.
3. (a) Crick, F. H. C. *Symp. Soc. Exp. Biol XII* **1958**, 139-163; (b) Crick, F. H. C. *Nature* **1970**, *227*, 561-563; (c) Uzawa, T.; Yamagishi, A.; Oshima, T. *The Journal of Biochem.* **2002**, *131*, 849-853; (d) Werner, E. *FEBS Lett.* **2005**, *579*, 1779-1782; (e) Ptashne, M.; Gann, A. *Nature* **1997**, *386*, 569-577.
4. Thanbichler, M.; Wang, S.; Shapiro, L. *J. Cell Biochem.* **2005**, *96*(3), 506-521.
5. Knight, R. D.; Freeland, S. J.; Landweber, L. F. *Trends in Biochem. Sci.* **1999**, *24*(6), 241-247.
6. (a) Dwek, D. *Chem. Rev.* **1996**, *96*, 683-720; (b) Hakamori, S.; Zhang, Y. *Chem. Biol.* **1997**, *4*, 97-104.
7. Branden, C.; Tooze, J. *Introduction to Protein Structure 2nd ed.* (1999).
8. Pauling, L.; Corey, R. B.; Branson, H. R. *Proc. Nat. Acad. Sci.* **1951**, *37*, 205-211.
9. (a) Back, J. W.; de Jong, L.; Muijisers, A. O.; de Koster, C. G. *J. Mol. Biol.* **2003**, *331*, 303-313.; (b) Young, M. M.; Tang, N.; Hempel, J. C.; Oshiro, C. M.; Taylor, E. W.; Kuntz, I. D.; Gibson B. W.; Dollinger, G. *Proc. Nat. Acad. Sci.* **2000**, *97*, 5802-5806; (c) Doyle, D. A.; Chait, B. T.; Mackinnon, R. *Science* **1998**, *280*, 69-77.
10. (a) Noren, C. J.; Anthony-Cahill, S. J.; Griffith, M. C.; Schultz, P. G. *Science* **1989**, *244*, 182-188; (b) Wallace, C. J. A. *Curr. Opin. Biotechnol.* **1995**, *6*, 403-410; (c) Witte, K.; Sears, P.; Wong, C. H. *J. Am. Chem. Soc.* **1997**, *119*, 2114-2118.
11. Rauh, D.; Waldmann, H. *Angew. Chem. Int. Ed.* **2007**, *46*, 826-829.
12. Wieland, T.; Bockelmann, E.; Bauer, L.; Lau H. U. *Justus Leibigs Ann. Chem.* **1958**, 583, 129.
13. (a) Dawson, P. E.; Muir, T.; Clark-Lewis, I.; Kent, S. B. *Science* **1994**, 776-779; (b) Dawson, P. E.; Kent, S. B. *Annu. Rev. Biochem.* **2000**, *69*, 923-960.
14. Ashraf, B.; Ficht, S.; Yang, Y. Y.; Bennett, C. S.; Wong, C. H. *J. Am. Chem. Soc.* **2006**, *128*, 15026-15033.
15. David, R.; Richter, M. P. O.; Beck-Sickinger, A. G. *Eur. J. Biochem.* **2004**, *271*, 663-677.
16. Perler, F. B. *Nucleic Acids Res.* **2000**, *28*, 344-345.

17. Canne, L. E.; Bark, S. J.; Kent, S. B. H. *J. Am. Chem. Soc.* **1996**, *118*, 5891-5896.
18. Nilsson, B. L.; Kiessling, L. L.; Raines, R. T. *Org. Lett.* **2000**, *2*, 1939-1941.
19. Bordusa, F. *Chem. Rev.* **2002**, *102*, 4817-4868.
20. Gavin, A. C. *Nature* **2006**, *440*, 631-636.
21. Kobe, B.; Kemp, B. E. *Nature* **1999**, *402*, 373-376.
22. Cambridge, S. B.; Davis, R. L.; Minden, J. S. *Science* **1997**, *277*, 825.
23. Schultz, P. G. et.al. *J. Am. Chem. Soc.* **2003**, *125*, 11782-11783.
24. Wang, Q.; Chan, T. R.; Hilgraf, R.; Fokin, V. V.; Sharpless, K. B.; Finn, M. G. *J. Am. Chem. Soc.* **2003**, *125*, 3192-3193.
25. Verhelst, S. H. L.; Fonovic, M.; Bogoyo, M. *Angew. Chem. Int. Ed.* **2007**, *46*, 1284-1286.
26. Brunner, J. *Annu. Rev. Biochem.* **1993**, *62*, 483-514.
27. Kaplan, J. H.; Hoffman, J. F. *Biochemistry* **1978**, *17*, 1929-1935.
28. Cornish, V.; Schultz, P. G. *Curr. Opin. Stru. Biol.* **1998**, *4*, 601-607.
29. Wei, Y. et. al. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 2419-2422.
30. Staudinger, H.; Meyer, J. *Helv. Chim. Acta.* **1919**, *2*, 635.
31. Grandjean, C.; Boutonnier, A.; Guerreiro, C.; Fournier, J.-M.; Mulard, L. A. *J. Org. Chem.* **2005**, *70*, 7123.

List of Publications

1. “Niobium(V) chloride catalyzed microwave assisted synthesis of 2,3-unsaturated *O*-glycosides by the Ferrier reaction”: Srinivas Hotha and **Ashish Tripathi**, *Tetrahedron Lett.* **2005**, *46*, 4555-4558.
2. “Diversity Oriented Synthesis of Tricyclic Compounds from Glycals Using the Ferrier and the Pauson-Khand Reaction”: Srinivas Hotha and **Ashish Tripathi**, *Journal of Combinatorial Chemistry*, **2005**, *7*, 968-976.
3. “Click Chemistry Guided Carbohydrate Mediated Imaging of Bacteria and Yeast”: **Ashish Tripathi**, Gopala Krishna Aradhyam and Srinivas Hotha. (Manuscript under preparation).
