

DESIGN, SYNTHESIS, SELF-ASSEMBLY, AND
DIS-ASSEMBLY STUDIES OF MONODISPERSE
PROTEIN-DENDRON CONJUGATES

A Thesis

Submitted in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

by

PAVANKUMAR JANARDHAN BHANDARI

I.D.: 20143335



INDIAN INSTITUTE OF SCIENCE EDUCATION AND
RESEARCH, PUNE

2020

DEDICATION

I dedicate this thesis to my parents,

Mr. JANARDHAN PENTAPPA BHANDARI

&

Mrs. RENUKA JANARDHAN BHANDARI

DECLARATION

I declare that this written submission represents my ideas in my own words, and wherever other's ideas have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea / data / fact / source in my submission. I understand that violation of the above will cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been appropriately cited or from whom proper permission has not been taken when needed.

A handwritten signature in black ink, appearing to read 'P. Bhandari', written in a cursive style. The signature is positioned above a horizontal line that extends to the right.

Date: 28-04-2020

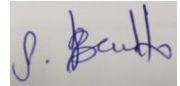
Mr. Pavankumar Bhandari

I.D.: 20143335

CERTIFICATE

I certify that the thesis entitled “*Design, Synthesis, Self-Assembly, And Dis-Assembly Studies Of Monodisperse Protein-Dendron Conjugates*” presented by Mr. Pavankumar Janardhan Bhandari represents his/her original work which was carried out by him/her at IISER, Pune under my guidance and supervision during the period from 25-07-2014 to present. The work presented here or any part of it has not been included in any other thesis submitted previously for the award of any degree or diploma from any other University or institutions. I further certify that the above statements made by him/her in regard to his/her thesis are correct to the best of my knowledge.

Date: 28th April 2020



Dr. Sandanaraj Britto

ACKNOWLEDGMENT

I am greatly indebted to my advisor, Dr. Sandanaraja Britto, for his guidance, support, and constant encouragement over the last five years at IISER, Pune. His support, thoughtful insights, and critical examination of my research work helped me to cross the difficulties that appeared during my research work. The training I got from him has made me confident about my future research endeavors.

It has been an excellent opportunity for me to work at IISER Pune. An excellent interdisciplinary research environment with world-class research facilities at IISER Pune always amazed me. I am thankful to Council of Scientific and Industrial Research (CSIR), Government of India, for research scholarship during the course of Ph.D. I would like to express my sincere gratitude to Dr. Siddhesh Kamat and Dr. Radha Chauhan for being a research advisory committee (RAC) members and providing critical suggestions during the RAC meeting.

I also express my deepest gratitude to all past lab members, for all the help and friendship. I also thank present members, Mohan, and Punita, for their friendship. A special thanks to Mohan, for his help throughout the Ph.D. I am grateful to all my friends for showing confidence and faith in me. Particularly I would always be grateful to Pankaj for directing my interest in science.

I shall always remain indebted to my parents for their unconditional love, blessings, sacrifices, patience, and support. I am blessed to have such parents. The values, ethics, and virtues they have instilled in me have made me achieve whatever I have achieved so far. A special thanks to my sister, Komal, for her constant support, encouragement, which helped me to be focused on work. A special thanks to my brothers-in-law, Manoj, for support. I would like to thank a very special person, my lovely wife, Manisha, for being supportive all the time. Her help during my work, making presentations, and writing thesis made my Ph.D. journey smoother. Also, her encouragement to do new things has inspired me to perform better.

TABLE OF CONTENT

Table of Content	i
Abbreviations and Symbols	v
Synopsis	ix
Chapter 1. Methods to Design Artificial Supramolecular Protein Assemblies	
1.1. Introduction	2
1.2. Genetic Strategies	3
1.2.1. Fusion of Natural Oligomers	3
1.2.2. New Interface Design	7
1.3. Chemical Strategies	11
1.3.1. Receptor-Ligand Interactions Driven Protein Self-Assembly	11
1.3.1.1. Co-factor Reconstitution.....	12
1.3.1.2. Protein Assembly via Engineering Natural Ligand-Protein Interactions	13
1.3.1.3. Synthetic Ligand-Protein Interactions Driven Self-Assembly	15
1.3.2. Metal Ion Induced Protein Assembly	17
1.3.3. Protein-Polymer Conjugates	19
1.3.3.1. Non-Covalent Coupling.....	20
1.3.3.2. Covalent Coupling	23
1.3.3.3. Design of Monodisperse Facially Amphiphilic Proteins	27
1.3.3.4. Protein-Dendron Conjugates.....	29

1.4. Current Challenges and Aim of the Thesis	31
1.5. References	39
Chapter 2. Design, Synthesis, and Self-Assembly Studies of Suite of Monodisperse Facially Amphiphilic Protein-Dendron Conjugates	
2.1. Introduction.....	45
2.2. Result and Discussion	48
2.2.1. Synthesis of Monodisperse Hydrophilic Linker	48
2.2.2. Synthesis of Macromolecular Amphiphilic Activity-based Probes.....	49
2.2.3. Micelle-Assisted Protein Labeling Technology	52
2.2.4. Multi-Step Purification and Characterization of Protein-Dendron Bioconjugates	54
2.2.5. Self-assembly Studies	56
2.3. Conclusion	59
2.4. Experimental Section.....	60
2.5. References	84
2.5. Appendix-I: Characterization Data of Synthesized Compounds.....	87
Chapter 3. Programmed and Sequential Disassembly of Multi-Responsive Supramolecular Protein Assemblies	
3.1 Introduction.....	114
3.2. Results and Discussion.....	116
3.2.1. Macromolecular Design.....	116
3.2.2. Synthesis of Monodisperse Hydrophilic Linker having Protein Reactive Group	117

3.2.3. Synthesis of a Photo-Sensitive Amphiphilic Activity-Based Probe	118
3.2.4. Synthesis of photo-sensitive protein-dendron conjugates	119
3.2.5. Self-Assembly Studies of Photo-Sensitive Protein-Dendron Bioconjugates	122
3.2.6. Dis-assembly Studies of Photo-Sensitive Protein-Dendron Complexes	123
3.3. Conclusion	133
3.4. Experimental Methods	133
3.5. References	156
3.6. Appendix-II: Characterization Data of Synthesized Compounds	159
Chapter 4. Programmed Disassembly of Redox-Responsive Protein-Dendron Complexes	
4.1 Introduction	179
4.2. Results and Discussion	180
4.2.1. Macromolecular Design	180
4.2.3. Synthesis of Redox-Sensitive Amphiphilic Activity-Based Probe	181
4.2.4. Synthesis of Redox-Sensitive Protein-Dendron Conjugates	182
4.2.5. Self-Assembly Studies of Redox-Sensitive Protein-Dendron Conjugate	184
4.2.6. Dis-Assembly Studies of Redox-Sensitive Protein-Dendron Complexes	185
4.3. Conclusion	188
4.4. Experimental Methods	188
4.5 References	199
4.6 Appendix III Characterization Data of Synthesized Compounds	202

Chapter 5. General Conclusions and Future Perspective

5.1. Summary of the Thesis	208
5.2. Limitations	209
5.3. Future Prospective	210
5.3.1. Synthesis of Biocompatible Protein-Dendron Bioconjugates	210
5.3.2. Synthesis of Stimuli-Besponsive biocompatible Protein-Dendron Bioconjugates...	212
List of Publications and Patents	215

ABBREVIATIONS and SYMBOLS

A

AABP	Amphiphilic Activity-based Probes
AFM	Atomic Force Microscopy

B

BSA	Bovine Serum Albumin
-----	----------------------

C

C	Cysteine
CALB	Candida Antarctica
CEG	Hexadeca(ethylene glycol)
Chy	Chymotrypsin
CVs	Column Volumes
ConA	Concanavalin A
CMC	Critical Micelle Concentration

D

Da	Dalton
DAHC	Diammonium Hydrogen Citrate
DHAP	2, 6-Dihydroxyacetophenone
DAST	Diethylaminosulfur Trifluoride
DHFR	Dihydrofolate Reductase
DMAP	4-Dimethylaminopyridine
DMSO	N, N-dimethyl sulfoxide
DMF	Dimethylformamide
DTT	Dithiothreitol
DLS	Dynamic Light Scattering

E

EDTA	Ethylenediaminetetraacetic Acid
Eq	Equivalent
F	
FP	Fluorophosphonate
fM	femtomolar
G	
G1	1 st Generation Dendrimer
G2	2 nd Generation Dendrimer
G3	3 rd Generation Dendrimer
G4	4 rd Generation Dendrimer
GST	Glutathione S-transferase
GFP	Green Fluorescent Protein
H	
H	Histidine
Hz	Hertz
HRP	Horseradish Peroxidase
D _h	Hydrodynamic Diameter
I	
IEX	Ion-Exchange Chromatography
pI	Isoelectric Point
K	
kDa	Kilo Dalton
L	
LCST	Lower Critical Solution Temperature

M

MALDI-ToF	Matrix-assisted Laser Desorption/ionization- Time of Flight
MC	Merocyanine
MeOH	Methanol
mg	Milligram
MHz	Megahertz
MTX	Methotrexate
μm	Micrometer

N

nm	Nanometer
nM	Nanomolar
NB	Nitrobenzyl
NTP	Nucleotide Triphosphates
NDP	Nucleotide Diphosphate
NMR	Nuclear Magnetic Resonance

O

OEG	Oligoethylene glycol
-----	----------------------

P

PPM	Parts Per Million
PDI	Polydispersity Index
PEG	Polyethylene Glycol
PMA	Phosphomolybdic Acid
PAMAM	Poly(amidoamine)
PNIPAAM	Poly(<i>N</i> -isopropyl acrylamide)
ProK	Proteinase K

R

RAFT	Reversible Addition–Fragmentation Chain Transfer
------	--

RBF Round Bottom Flask
RhB Rhodamine-B
RIDC3 Rosetta-Prescribed Surface Mutations to Cytochrome

S

SEC Size Exclusion Chromatography
SEC-MALS Size-Exclusion Chromatography - Multi-Angle Laser Light Scattering
Sub Subtilisin
SP Spiropyran

T

TEG Tetraethylene Glycol
THF Tetrahydrofuran
TLC Thin Layer Chromatography
hHint1 Triad Nucleotide-Binding 1
TFA Trifluoroacetic Acid
Try Trypsin

U

UV Ultraviolet

SYNOPSIS

The fundamental objective of this thesis is to introduce a methodology to design self-assembling, stimuli-responsive, protein-dendron conjugates. The research during my doctoral study was mainly focused on designing a chemical methodology to construct monodisperse, stimuli-responsive, facially amphiphilic protein-dendron conjugates and understand their self-assembly and dis-assembly behavior. In particular, this thesis mainly deals with the construction of protein-dendron assemblies, which respond to both extrinsic and intrinsic stimuli such as pH, light, and redox potential. Besides, accessibility for control over hydrodynamic radius (D_h), oligomeric state, and the molecular weight of the protein-dendron complex has been addressed by re-engineering the components in the molecular design. The protein-dendron system presented in this thesis provides an opportunity to functionalize interior and exterior domains of assemblies with a variety of therapeutic agents. These opportunities could be used in devising antibody or ligand decorated particles with controlled densities, which we expect to find application in the area of vaccine design, targeted drug delivery.

Chapter 1: Methods to Design Artificial Supramolecular Protein Assemblies

Chemical strategies have vastly expanded the scope of protein nanotechnology. Among the various methods, the synthesis of amphiphilic protein-polymer conjugate *via* chemical conjugation of proteins with polymer (hydrophobic or temperature-sensitive polymer) - that serve as a driving force for proteins to self-assemble - has significantly contributed to the construction of artificial ordered protein nanostructures *via* a bottom-up approach. However, a significant disadvantage to the use polymer is that they induce a degree of heterogeneity in the form of polydispersity resulting in heterogeneous samples, which are hard to characterize by standard analytical and biophysical tools.

The replacement of polydisperse linear polymer with monodisperse branched dendrimer is considered to be a viable approach to synthesize protein-polymer conjugates with an ideal polydisperse index (PDI) of 1.00. The existing methodologies present tools

for synthesis of globally hydrophilic protein-dendron conjugates *via* conjugating hydrophilic monodisperse dendron to hydrophilic/amphiphilic proteins. However, the majority of these conjugates cannot self-assemble due to a lack of amphiphilic nature. In this regard, the custom design of protein-dendron amphiphilic macromolecules is very interesting because of their propensity to self-assemble into various nanoscopic objects.

Chapter 2: Design, Synthesis, and Self-Assembly Studies of Suite of Monodisperse Facially Amphiphilic Protein-Dendron Conjugates

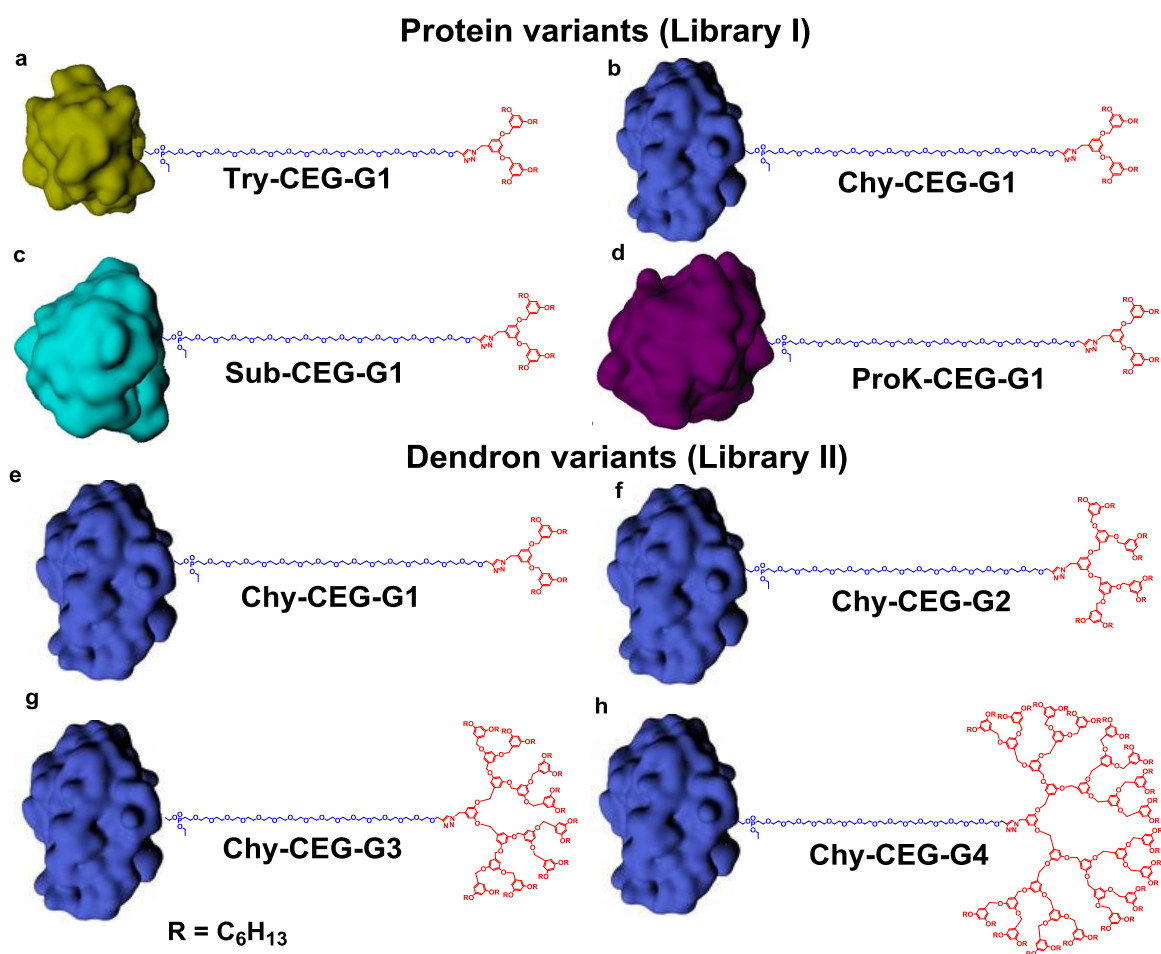


Figure 1 | Chemical structures of protein-dendron bioconjugates composed of globular water-soluble protein (green = trypsin, blue = chymotrypsin, cyan = subtilisin, magenta = proteinase K), hydrophilic CEG (blue), and hydrophobic dendron block (red). Structures of (a) Try-CEG-G1 (b) Chy-CEG-G1 (c) Sub-CEG-G1 (d) ProK-CEG-G1 (e) Chy-CEG-G1 (f) Chy-CEG-G2 (g) Chy-CEG-G3 (h) Chy-CEG-G4.

The novel design strategy to construct monodisperse facially amphiphilic protein-dendron bioconjugates is described in this chapter. These conjugates are synthesized by attaching a macromolecular amphiphilic activity-based probe composed of a fluorophosphonate as a reactive group, hydrophilic linker, and hydrophobic dendrimer, to the serine protease using triton X-100 as a solubilizing agent. The precise choice of linker length and optimized reaction conditions showed the exclusive mono-labeling to the protein (Figure 1). Further, the three-step purification of reaction yielded pure protein-dendron conjugates. The designed protein complexes exhibit both dendrimer generation-dependent as well as protein-dependent self-assembly properties. Most importantly, this method provides exquisite control over size, oligomeric state, and the molecular weight of the protein nanoassemblies by either choosing an appropriate dendron or protein of interest. Thus, this design strategy opens up the possibility to design nearly monodisperse, megadalton-scale protein nanoassemblies through a chemical method.

Chapter 3: Programmed and Sequential Disassembly of Multi-Responsive Supramolecular Protein Assemblies

In chapter 1, we have demonstrated that extremely hydrophobic and highly branched benzyl-ether dendron could be site-specifically attached to a protein of interest. However, one of the major limitations of that study is that custom designed nanoassemblies are static and failed to exhibit dynamic behavior. In this chapter, we report a chemical method for the construction of multi-responsive supramolecular protein nanoassemblies using custom-designed amphiphilic monodisperse protein-based macromolecular synthons. The molecular design of photo-responsive bioconjugate is similar to the previous design except for the photo-responsive group (2-nitrobenzyl derivative) is incorporated between the hydrophilic and hydrophobic portion of bioconjugate. Thus, the newly designed macromolecular synthons have four core structural elements: (i) hydrophilic globular protein, (ii) flexible hydrophilic linker, (iii) photoresponsive group, and (iv) hydrophobic dendron of different generations (Figure 2). Further, self-assembly data interestingly manifest that, installing a photo-sensitive O-nitro benzyl group between the hydrophilic and hydrophobic portion of bioconjugate does not affect the self-assembly.

Exposure of supramolecular protein nanoassemblies to UV-light leads to partial disassembly (Figure 2). The newly formed protein assembly no longer responds to UV-light but could be disassembled at lower pH or treatment with a small molecule. More interestingly, the distribution ratio of assembled versus a disassembled state of protein nanoassemblies after photo-trigger does not depend on dendron generation, nature of linker functionality, the identity of the protein but depends on linker length. Altogether, this work discloses a new chemical method for the rational design of the monodisperse multi-responsive supramolecular system with exquisite control over disassembly processes.

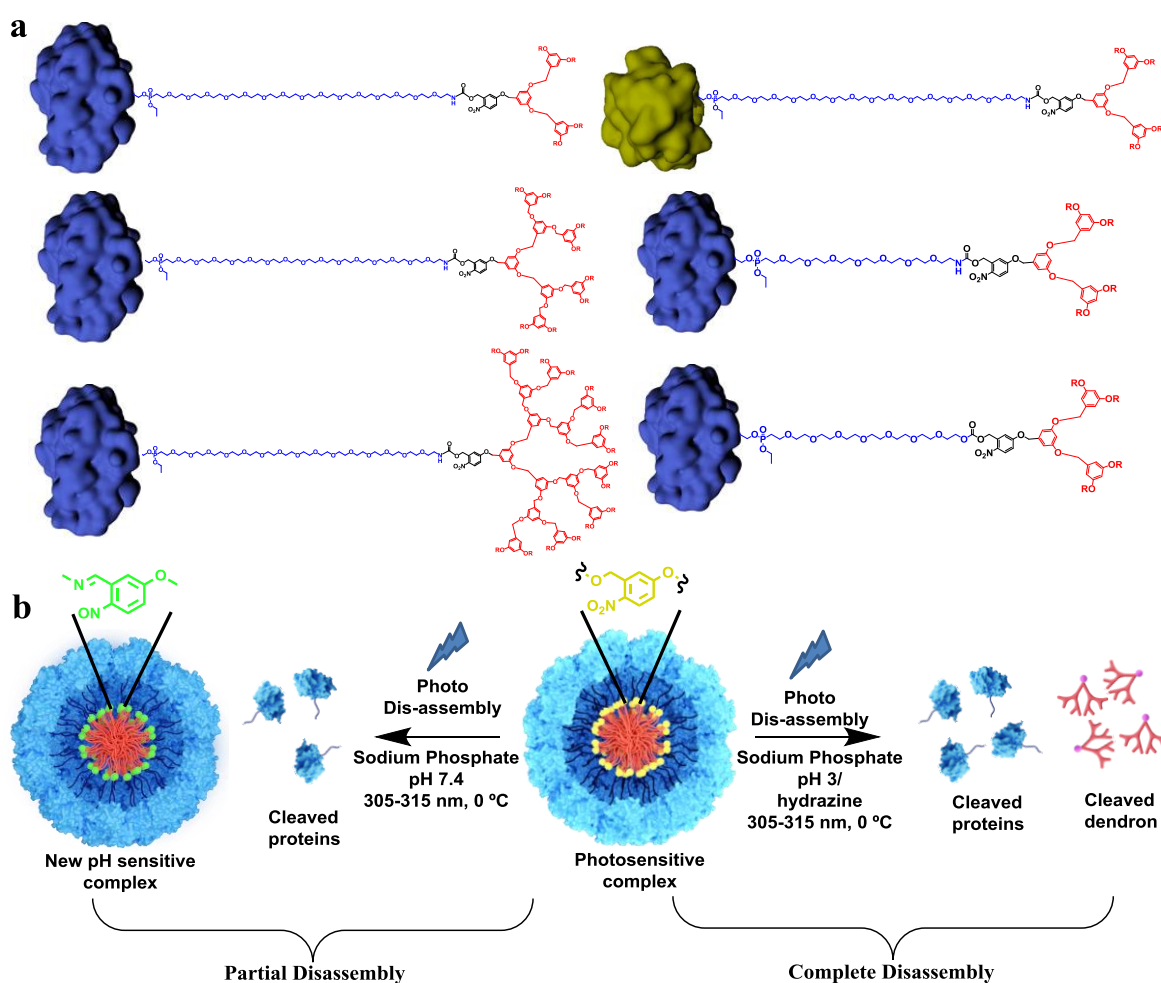


Figure 2 | Schematic representation of chemical structures, self-assembly, and dis-assembly of photo-sensitive protein-dendron bioconjugates. (a) These conjugates are composed of protein (blue = chymotrypsin, green = proteinase K), hydrophilic oligoethylene glycol linker (blue) (CEG = cetylene glycol, OEG = octaethylene glycol), photo-sensitive nitrobenzyl group (black), and

hydrophobic dendron block (red). **(b)** Structures of (i) Chy-CEG-NB-G1 (ii) Chy-CEG-NB-G2 (iii) ChyCEG-NB-G3, (iv) ProK-CEG-NB-G1, (v) Chy-OEG-NB-G1, and (vi) Chy-OEG-NB-O-G1. Schematic representation of disassembly of photo-sensitive protein-dendron at pH 7.4, pH 3, and in the presence of hydrazine at pH 7.4.

Chapter 4: Programmed Disassembly of Redox-Responsive Protein-Dendron Complexes

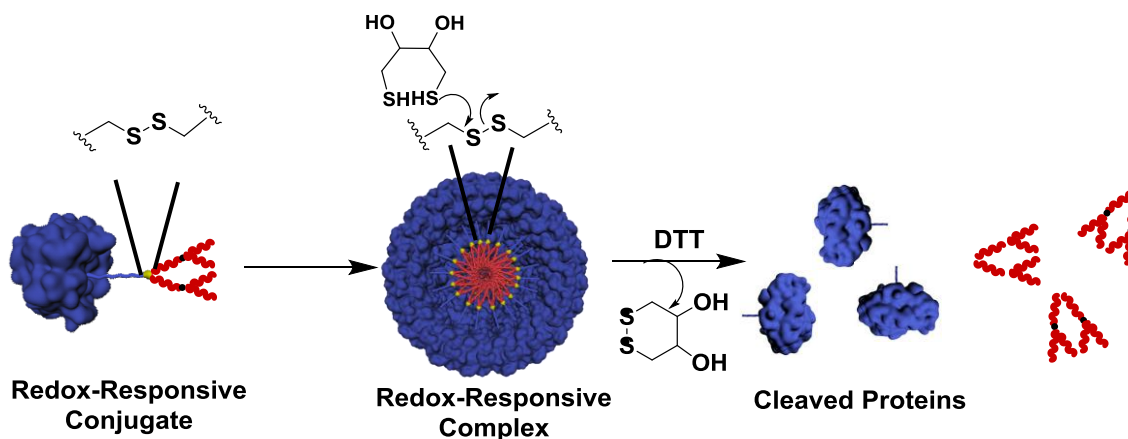


Figure 3 | Schematic representation of self-assembly and disassembly of redox-sensitive complexes.

As mentioned above, we have developed strategies to synthesize static and dual-sensitive protein-dendron bioconjugates. Particularly, in chapter 3, we have used the photosensitive nitrobenzyl group as an external trigger, which provides excellent spatiotemporal control. However, the phototoxicity associated with the UV light and its less penetration depth into the skin are the significant limitations. In this aspect, the development of materials that can respond to physiological changes such as pH, redox, imbalance in enzyme or protein concentration, metabolite are highly relevant for biomedical applications. Towards that goal, in this chapter, we extend our chemical methodology for the design of monodisperse redox-responsive facially amphiphilic protein-dendron bioconjugates.

The disulfide functionality which connects the hydrophobic dendron to the hydrophilic linker and protein would cleave, upon treatment with DTT. This would convert facially amphiphilic protein into hydrophilic protein. The loss of attractive hydrophobic interaction then would lead to disassembly of protein nanoassemblies to consecutive

monomers (Figure 3). These rationally designed redox-responsive protein-dendron complexes have several outstanding advantages compared to the protein-polymer complexes. This is due to the fact that these intrinsic responsive redox-sensitive protein-dendron materials are highly monodisperse. Besides, the system offers the advantage of installing various therapeutic agents to the interior and exterior domains of complexes in order to achieve the specific delivery of cargo.

Chapter 5: General Conclusions and Future Perspective

In this chapter, general conclusion and summary of the thesis is discussed. Besides, advantages, and limitations of presented facially amphiphilic, stimuli-responsive protein-dendron bioconjugates are also described. Next, considering the aim and limitations of our scaffold, detailed future aspects of projects, towards designing biocompatible, drug and antibody functionalized, stimuli-responsive protein-dendron assemblies are presented (Figure 4).

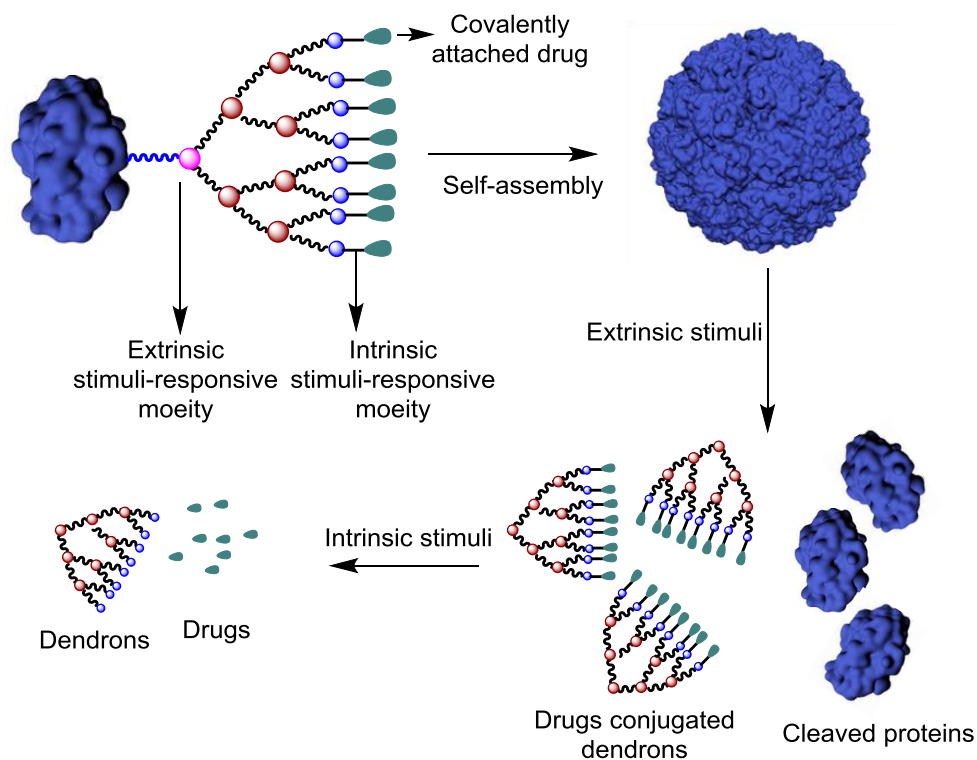


Figure 4 | Schematic representation of the self-assembly and dis-assembly of drug-functionalized stimuli-responsive protein-dendron bioconjugate.

CHAPTER 1

METHODS TO DESIGN ARTIFICIAL SUPRAMOLECULAR PROTEIN ASSEMBLIES

1.1. Introduction

Proteins are Nature's essential building blocks made up of 20 different amino acids. This linear heteropolymer folds into a particular three-dimensional structure. The precise folding of linear polypeptide chain is primarily mediated by precise intramolecular non-covalent interactions between different amino acids which distant to each other in the primary sequence. During the course of evolution, Nature has evolved the ability to utilize these well-folded, three-dimensional proteins to construct a diverse set of higher-order nano and micro-scale protein architectures such as viruses¹, molecular motors², protein cages³, bacterial compartments⁴, and other fascinating architectures.

Protein complexes are crucial for the functioning of living systems because they perform various significant and specific biological functions.⁵ For example, natural protein cages such as virus capsids¹, bacterial microcompartment⁴, and ferritin⁶, act as excellent delivery and storage devices for their respective cargo. As another example, sliding clamps and helicases are biologically relevant protein ring complexes that catalyze the reactions involved in DNA replications and repairing.⁷ In a similar fashion, protein filaments such as actin and tubulin work as track roads for transport of the organelles and lipid vesicles in the cells and maintain the integrity of the cell.⁸ Apart from fascinating functional properties, many of these protein assemblies are dynamic in nature, i.e., they can be dissociated into constitutive monomers in response to a stimulus, and some can be even assembled back to original monodisperse architectures under selective conditions.^{9,10} For example, polymerization and depolymerization dynamics of actin filaments and microtubules is one of the extensively-studied systems. The polymerization of actin and tubulin subunits is initiated by the binding of adenosine triphosphate (ATP) and guanine triphosphate (GTP) to the actin and tubulin subunits, respectively. While in the depolymerization phase, subunits of actin and tubulin hydrolyze their bound nucleotide triphosphates (NTP) to nucleotide diphosphate (NDP) and release the hydrolyzed phosphate (Pi).¹¹ The above mentioned inspirational characteristics of Natural biomolecular assemblies and other exquisite biological structures stimulated an intense interest to create different artificial protein assemblies from non-assembling natural monomeric proteins.¹² The construction of such

artificial protein assemblies not only helps to replicate the dynamics of natural protein assemblies at a molecular scale but also aids in comprehending the self-assembly mechanism of natural protein assemblies. The capability to understand and replicate the dynamics of natural protein assemblies would provide a platform for designing a diverse set of functional nanoscopic biomaterials.

In contrast to the remarkable efforts that have been made to build architectures from DNA¹³, RNA¹⁴, and synthetic peptides¹⁵, the construction of protein assemblies has encountered substantial difficulty, due to the chemical complexity, structural instability, and partial understanding of the rules by which monomeric proteins identifies its cognate partner to form assemblies. Despite these challenges, various efforts along the multiple lines in the last two decades have led to the substantial growth of protein nanotechnology field. Several innovative methodologies have been reported till now to design protein assemblies, including the genetic or chemical modification of viruses or protein cages¹⁶, fusion of natural oligomers based self-assembly¹⁷, new interface guided self-assembly¹⁸, metal ion guided self-assembly¹⁹, host-guest interactions driven self-assembly²⁰, synthetic molecule templated self-assembly²¹, and others²². The above-discussed methods broadly fall under two main broad strategies, i.e., genetic and chemical strategies.

In this chapter, an overview of the different synthetic and genetic methodologies and their development to design protein assemblies are discussed. Besides, challenges and shortcomings in designing monodisperse, dynamic, and functional protein assemblies by using the existing methodologies are described. Finally, the motivation behind the research work presented in this thesis is also discussed.

1.2. Genetic Strategies

1.2.1. Fusion of Natural Oligomers

In 1956, Watson and Crick hypothesized that shells of viruses are constructed from a large number of identical protein molecules, symmetrically assembled around its genome.²³ This is because the symmetrical arrangement of protein molecules requires the

less number of distinct interfacial contacts between individual subunits. Later, it has been shown that not only viruses but many naturally occurring protein assemblies composed of the same or similar subunits, nearly assemble symmetrically.²⁴ Thus, it is evident that during evolution, nature has designed ordered protein assemblies based upon symmetry principles.

This Nature's' approach served as a clue and inspiration to design protein assemblies based upon symmetry principles. The basic design requirement for constructing symmetry-based artificial protein assemblies primarily depends on the synthesis of genetically or chemically fused proteins, which should at least have two subunits with distinctive interaction at the interfaces. Based on this requirement, many strategies have been reported until now, which are discussed below.

The first general strategy to construct protein materials using symmetry is demonstrated in 2001 by Yeates and coworkers. In this strategy, a natural protein monomer, which is capable of forming an oligomer, is tied rigidly to another protein monomer with a different oligomeric state. The resultant fusion protein, therefore, carries two proteins with different oligomeric states; each of the components has a strong tendency to associate with other copies of itself. Also, each of these oligomeric domains contains its own axis of symmetry; hence, the fusion proteins contain two symmetry axis (Figure 1.1a-b).²⁵

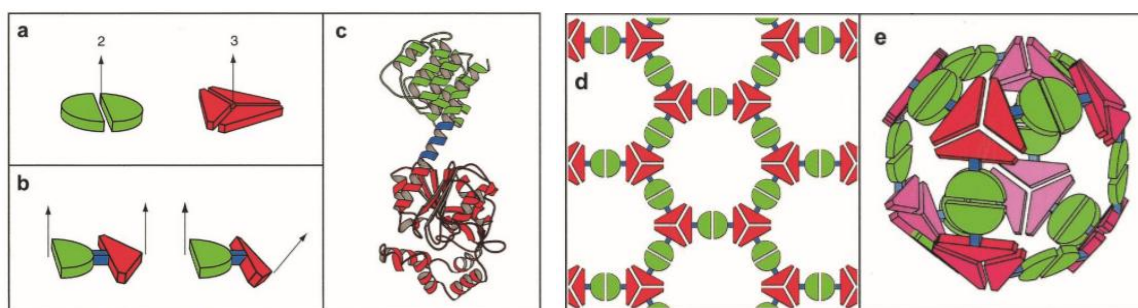


Figure 1.1 | Schematic representation of fusion natural oligomeric strategy to construct protein assemblies. (a) Self-assembling natural oligomers, i.e., dimer (green) and trimer (red). These oligomers possess their own axis of symmetry, as shown above. (b) These natural oligomeric domains are then fused rigidly to each other using a linker (blue). The fusion protein with different geometry of the symmetry axis can be constructed. (c) A ribbon diagram of genetically fused oligomeric domains containing α helical linker, extending from one oligomeric domain to the other. Such continuous α -helix spanning the two domains furnishes the rigidity to the scaffold. (d)

Depending upon the geometry of the symmetry axes, the fusion constructs self-assembles into a particular kind of assembly. 2-D layer structure arises from oligomeric domains shown in the left side of b. (e) Oligomeric domains shown at the right side of b would form a cubic cage assembly. Adapted with permission from ref. [25]. Copyright @ 2001, the National Academy of Sciences.

By using the geometric relationship or relative orientation between each of these symmetry axes, the authors described the construction rule (Table 1.1) to design the various protein architectures.

Table 1.1 | Construction table for designing protein assemblies.

Symmetry	Construction	Geometry of symmetry elements	
Cages and shells			
tetrahedral	Dimer-trimer	54.7	I
octahedral	Dimer-trimer	35.3	I
icosahedral	Dimer-trimer	20.9	I
Double-layer rings			
dihedral	Dimer-trimer	180/n	I
Two-dimensional layer			
$p6^{\parallel}$	Dimer-trimer	0	I
$p321$	Dimer-trimer	90	I
$p3$	Dimer-trimer	0	I
Three-dimensional crystals			
$I2_13$	Dimer-trimer	54.7	I
$P4_132$ or $P4_332$	Dimer-trimer	35.3	I
$P23$	Dimer-trimer	70.5	I
Helical filaments			
Helical [\]	Dimer-trimer	Any angle	I

I and N denote intersecting and nonintersecting axes.

However, the geometric rules described in the paper are only restricted to the combinations of dimer and trimer oligomeric proteins. In order to design the pre-determined protein assemblies using the construction table, it is bare essential to calculate

the relative orientation between two genetically fused domains. This is feasible only if the two oligomeric domains in the fusion protein are held rigidly. Due to free backbone rotation at the point of fusion, it is impossible to predict the relative orientation between the symmetry axes. A solution to this problem was fixed by choosing the individual oligomeric domains whose structure begins and ends with α -helices. That means, both the proteins would carry α helices, one will begin at *C*-terminus of first protein, and another would-be at *N*-terminus of a second protein. The linker, connecting these two oligomeric domains, is also designed in such a way that it strongly favors the α -helical conformation. The overall, newly designed genetically fused protein would possess α -helical conformation extending from within one oligomeric domain to the other oligomeric domain. Such continuous α -helix spanning the two domains provides the rigidity to the scaffold and makes it possible to calculate the relative geometry between the two symmetry axes of oligomeric domains (Figure 1.1).

The design of linear protein filament was aimed by utilizing the proposed geometric relationship or relative orientation between each of these symmetry axes. To do that, a fusion protein from two dimeric oligomerization domains, i.e., influenza virus matrix protein M1 and carboxylesterase, was constructed in such a way that symmetry axes do not intersect. Surprisingly, self-assembly studies of this construct by using electron microscopy revealed the formation of linear protein filaments, which is in agreement with the envisioned protein architectures proposed in the construction table. Further, using a similar strategy, the design of another fusion construct was envisioned. This fusion construct is composed of a trimeric bromoperoxidase enzyme and a dimeric viral matrix protein. The symmetry axes of this construct were found to be intersecting at 54° . This 49 kDa fusion protein construct is expected to assemble as a tetrahedral cage (contains 12-subunits) with a diameter of about 16 nm. Electron microscopy and other biophysical techniques revealed that, although the construct could self-assemble into the expected shape and size, the final product was heterogeneous. However, in the follow-up paper, the same group made few changes in the original design that could potentially remove the steric hindrance. As a result, the modified construct self-assembles into a monodisperse 12-subunit cage. The diameter and central opening of this assembly were found to be 16 nm and 5 nm, respectively.²⁶ A similar strategy was further explored to design a highly porous protein

cube by rigidly fusing two natural protein oligomers, that is, trimeric 2-keto-3-deoxy-6-phosphogalactonate aldolase and the dimeric FkpA protein. This construct self-assembled into a 750 kDa, a cube-shaped cage complex, with a diameter of 22 nm and 13 nm of the large inner cavity. Rewardingly, atomic-level conformity was encountered between the envisioned model and the crystal structure.²⁷

Noble and co-workers made the key advance to the above strategy by genetically fusing higher symmetry oligomers with relatively flexible two or more linkers in such a way that the symmetry axes align with each other. In such a case, the need for a continuous α -helix linker could be bypassed (Figure 1.2). Although the symmetry-matching fusion protein strategy method can be used to create intended extended 2-D and 3-D arrays using a variety of different symmetries, due to the unavoidable constraint of this method, it cannot be used to generate finite structures such as molecular cages.²⁸

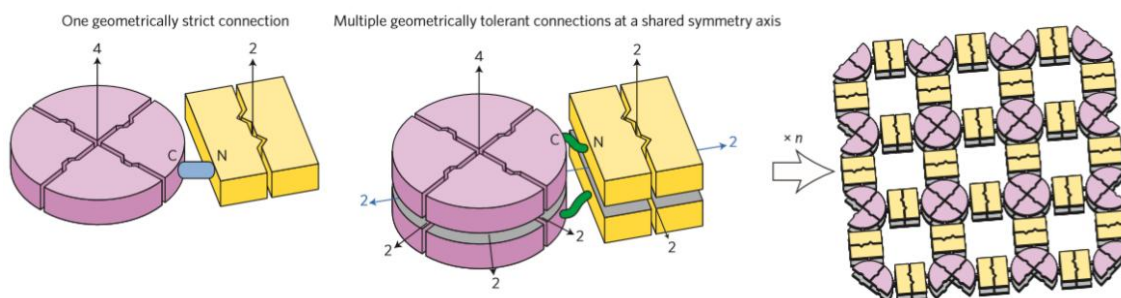


Figure 1.2 | Schematic representation of two different symmetry-based fusion strategies for constructing ordered protein assemblies. Protein oligomers with the different oligomeric states, i.e., dimer (yellow) and tetramer (purple), are connected with each other *via* one rigid linker (left). However, when the same blocks are connected at two positions in such a way that shared symmetry axes align with each other, the need for a continuous α -helix linker could be bypassed (middle). The designed construct then would assemble into extended 2D and 3D arrays (right). Adapted with permission from ref. [17]. Copyright @ 2011, Nature publishing group.

1.2.2. New Interface Design

Although the above-discussed method of fusion of natural oligomeric domains could allow access to the intended design of cages²⁷, 1-D, 2-D²⁸, and 3-D²⁶ protein assemblies. However, the methodology is restricted only to the protein oligomeric domains. Further, the requirement of designing fusion proteins with accurate and specific geometries

obstructs the versatility of this approach for designing novel protein assemblies. These limitations have been surmounted by developing computer programs to establish new contacts onto the protein by mutation.²⁹ The engineered protein then undergoes the symmetric self-association to form the complex. The most common approaches for designing such assemblies are based upon using the natural oligomerization motif as a starting material rather than a natural protein monomer. This is because the natural oligomerization motifs inherently possess one of the modes of interaction at the interfaces without any mutations. Then, the second additional interface can be established by introducing new amino acids *via* mutation using various computer algorithms. Thus, the use of natural oligomerization motifs reduces the novel interfaces that have to be introduced by using computer algorithms.

Grueninger et al., for the first time in 2008, reported the strategy to use the computational design to introduce new contacts at various points on the surface of the natural oligomeric motifs. Using this strategy, they have generated homo-oligomers of double-ring structures by simply employing some nonpolar side chains into the surface of natural monomeric or dimeric proteins.³⁰ This study, for the first time, has demonstrated that proteins oligomers can be re-engineered to form assemblies *via* self-association of new contact interfaces. However, the obtained and intended protein assemblies were not in the best agreement.

Tezcan and coworkers demonstrated the elegant use of modified protein surfaces to design protein assemblies through metal coordination (Read section 1.2.2).³¹ They designed the self-assembling building block by incorporating ten Rosetta-prescribed surface mutations to the cytochrome *cb-562* variant MBPC1 (Figure 1.3a). The resultant protein Rosetta interface designed cytochrome 3 (RIDC3), forms a C₂ symmetric dimer upon binding to Zn²⁺, which is stabilized by mutations incorporated using computer-aided interface design. Further, the solution state self-assembly of RIDC3 was investigated with respect to concentrations of RIDC3, Zn²⁺, and pH. By using this method, authors have shown that the metal-mediated assembly could be used to construct 1-D, 2-D, and 3-D protein assemblies ranging from nano- to micrometer long-range ordering. They have also demonstrated that the protein assemblies could be fine-tuned through the adjustment of the

protein to metal ratio or change in pH of the solution (Figure 1.3).³¹ However, the outcome of protein assemblies prepared by using this metal-mediated strategy could not be predicted computationally.

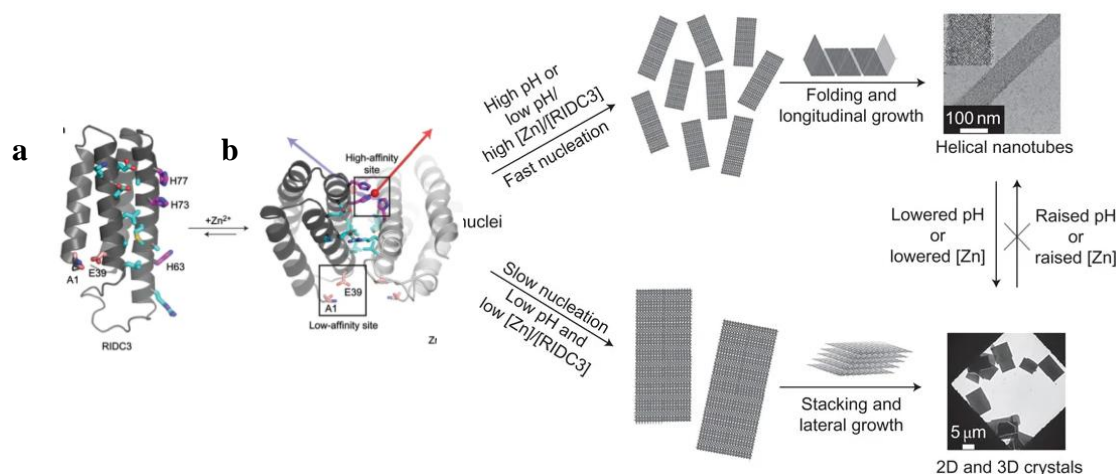


Figure 1.3 | Schematic representation of RIDC3 (a) A ribbon diagram is representing surface residues predicted by Rosetta design to stabilize RIDC3 dimer. Cyan, magenta, and pink sticks represent the high-affinity and the low-affinity Zn binding sites, respectively. (b) The C_2 -symmetric dimer is produced upon Zn (red spheres) and RIDC3 binding. The blue and red arrows symbolize the orthogonal coordination vectors of C_2 -symmetric dimer (middle). Zn-mediated RIDC3 self-assembly at different protein: metal ratio, and pH (right). Adapted with permission from ref. [31]. Copyright @ 2012, Nature Publishing Group.

The significant advancement to this strategy was made by Baker and co-workers. They have developed an impressive general computational method for designing highly-ordered, desired protein architecture with an atomic-level accuracy. The basic working principle of this approach is primarily based upon docking multiple copies of proteins oligomers in the target symmetry architecture. This is performed by aligning the symmetry axis of proteins oligomers and symmetry architecture, followed by finding the complementary low-energy protein-protein interface, which can drive the protein to self-assemble. In one of the examples, authors intended to design of cage-like protein assemblies with octahedral and tetrahedral symmetry. In order to do that, new homodimeric interfaces were introduced on to the two different natural trimeric protein oligomers, which share the same element of symmetry with the intended assemblies.¹⁸ Further, symmetry axes of multiple trimeric building blocks are aligned with the symmetry axes target architecture. Such an organization of building blocks minimizes rigid-body degrees of

freedom. In this case, systematically sampling of the remaining two degrees of freedom was performed to optimize interface separation and face side of trimer using Rosetta design suite by increments of 1 Å and 1°, respectively (Figure 1.4). The initial tetrahedral assembly was slightly less accurate than the design models. However, employing three additional mutations onto the construct resulted in a formation of protein assembly, with architecture similar to the predicted design model at atomic-level resolution.

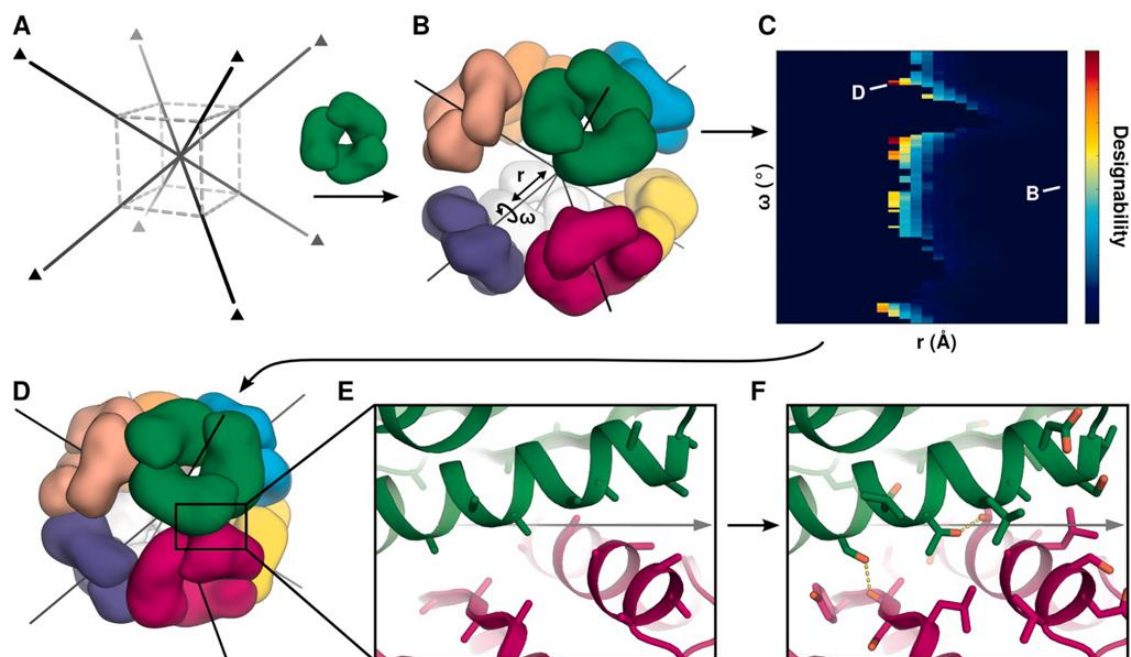


Figure 1.4 | Schematic representation of a computational method for designing protein nanomaterials. (a) Synthesis of protein assembly with octahedral point group symmetry is intended. (b) Then, multiple copies of C_3 -symmetric protein trimers are arranged in the target octahedral point group symmetry architecture. radial displacement (r) and axial rotation (ω) represent two degrees of freedom (c) Appropriateness of configuration for interface design (red: more suitable; blue: less suitable). (d) One of the highly complementary interface protein assembly. (e) Closer view of the interface in (d). (f) Novel low-energy protein-protein interfaces. Adapted with permission from ref. [18]. Copyright @ 2012, Nature Publishing Group.

Apart from designing protein assemblies with a single component, the versatility of this method has been demonstrated by constructing highly accurate two-component co-assembling protein materials.³² Besides the above-mentioned protein cages,³³ this method was also successfully explored for the construction of ordered 2D arrays of protein assemblies.³⁴ However, to realize the desired products, this method often necessitates the

screening of a vast number of design constructs. Also, unintended intermolecular interaction between the subunits leads to the polydisperse, unforeseen protein assemblies.

1.3. Chemical Strategies

Designing protein assemblies utilizing chemical strategies were reported more than three decades ago. However, they did not evolve as mature technologies until now. However, chemical strategies are equipped with unconstrained chemical space (not limited to standard 20 amino acids), which provides a broad platform for building versatile protein architectures with an advanced function. Over the last three decades, the chemical strategies have immensely expanded the scope of protein nanotechnology by offering a variety of design strategies to construct desired protein assembly mainly *via* a bottom-up approach. The efforts in this area are mostly focused on (I) chemical modification of proteins with other building blocks that serve as a driving force for proteins to self-assemble and (II) engineering naturally occurring interactions such as metal–ligand coordination, receptor–ligand interactions, host–guest interactions, inter-protein interactions to create different architectures. In this section, I discuss and categorize the recent reports in this field, based upon the supramolecular interactions employed to drive the self-assembly.

1.3.1. Receptor-Ligand Interactions Driven Protein Self-Assembly

Receptor-ligand interactions play an important role in various biological processes in living systems.³⁵ Interestingly, although the binding affinities of natural receptor-ligand interactions are strong (micromolar (μM) to the femtomolar (fM)), these interactions are often reversible/dynamic in nature. In addition to that, the unique key and lock recognition between these cognate partners, through several non-covalent interactions, makes receptor-ligand interactions very specific. Inspired by these fascinating properties, natural ligand-receptor interactions have been exploited to construct a diverse set of protein self-assemblies. This is primarily achieved by either engineering and/or modulating specific interactions between the receptor and its corresponding ligand. In this section, I will describe about various approaches to design protein assemblies by re-engineering natural

protein-ligand interactions such as hemeprotein-porphyrin, streptavidin-biotin, and lectin-carbohydrates (mannose, glucose, and galactose), respectively.

1.3.1.1. Co-Factor Reconstitution

One such example of natural receptor-ligand interactions are hemeproteins, and their corresponding co-factor, i.e., heme prosthetic groups (porphyrin).³⁶ Hayashi et al. reported impressive strategies to synthesize various 1-D, 2-D, and 3-D protein assemblies through polymerization of the modified heme cofactor and its corresponding apoenzyme.³⁷ In one of the examples, a mutant cytochrome b562, H63C, was expressed by introducing a new cysteine amino acid on the opposite side co-factor binding domain. This unique thiol group was then conjugated to the maleimide-terminated heme cofactor by using thiol-maleimide chemistry. The resultant external heme functionalized cytochrome b562 conjugate was then subjected to sequential denaturation–neutralization cycle to generate 1D nanowires *via* supramolecular polymerization of external porphyrin cofactor and its corresponding apoenzyme (Figure 1.5a).³⁸

Further, to extend this strategy, Hayashi et al. have designed the heme triad molecule with 1,3,5-trisubstituted benzene core as a branching point.³⁹ The addition of this heme triad molecule to the preformed 1D nanowires of cytochrome b562, at a different molar ratio, resulted in the formation of unique 2D hemoprotein networks. This supramolecular 2D assembly is formed mainly by cytochrome-heme and heme-heme interactions (Figure 1.5b). A similar approach was used to design stimuli-responsive heme protein assemblies. Here, the molecular design of the conjugate is similar to the above method except, hydrophilic linker (spanning the reactive head group (maleimide) and heme cofactor) is replaced by a hydrophobic linker which is composed of photoresponsive functional units such as azobenzene, or stilbene.⁴⁰ This heme-azobenzene or stilbene conjugates of apocytochrome b562 upon sequential denaturation–neutralization cycle generated one-dimensional nanowires. However, drastic structural change of the dimensional nanowires to the micelle-type structure was observed by heating the solution at 80 °C, followed by cooling it to 25 °C (Figure 1.5c).

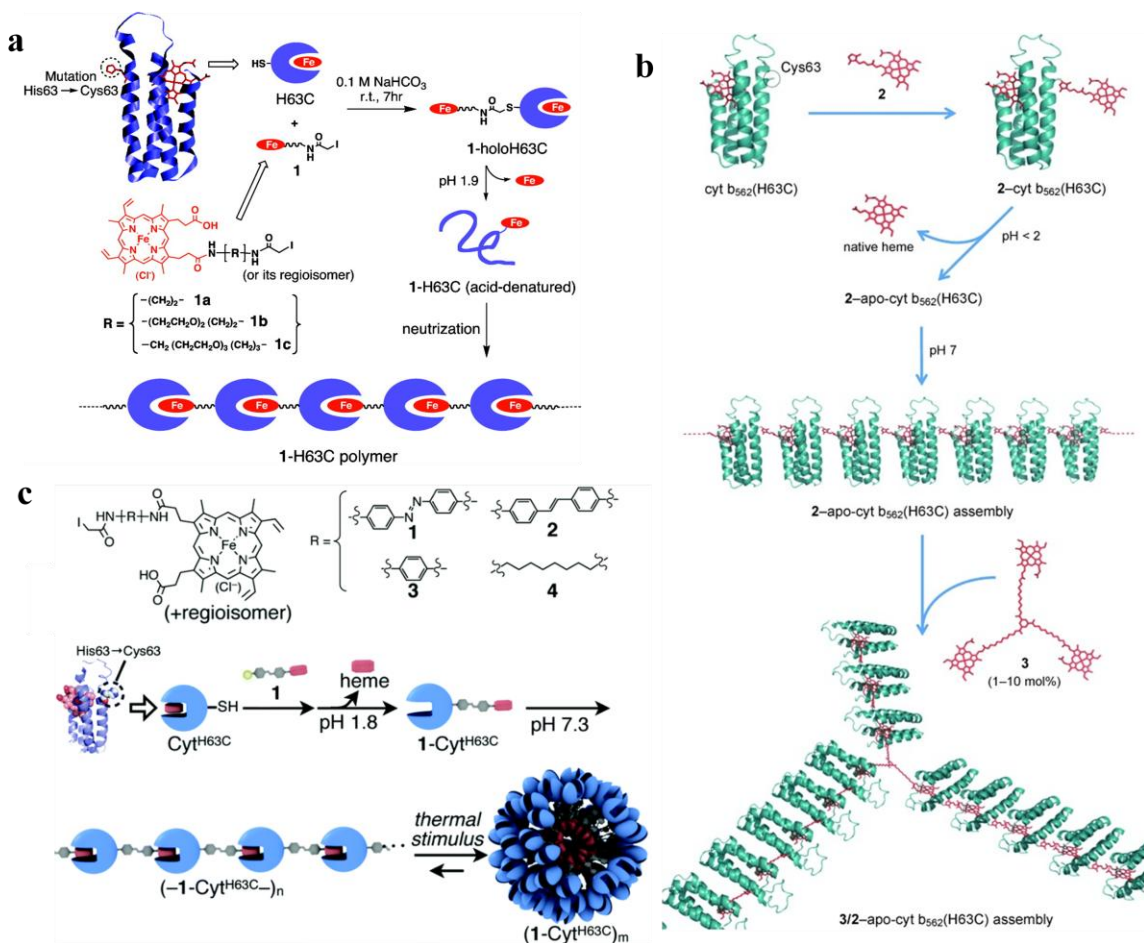


FIGURE 1.5 | (a) Schematic representation representing the mechanism of hemoprotein polymerization to generate 1-D nanowires. Adapted with permission from ref. [38]. Copyright @ 2007, American Chemical Society. (b) Schematic representation of 2-D hemoprotein networks formed *via* protein-ligand and heme-heme interactions. Adapted with permission from ref. [39]. Copyright @ 2009, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. (c) Synthesize thermal-responsive supramolecular protein assembly. Adapted with permission from ref [40]. Copyright @ 2017, Royal Society of chemistry.

1.3.1.2. Protein Assembly *via* Engineering Natural Ligand-Protein Interactions

Streptavidin is the homotetrameric protein with an antiparallel β -barrel-shaped structure. Streptavidin's binding affinity for biotin is one of the strongest protein-ligand interactions known.⁴¹ This highly specific biotin-streptavidin interaction has been exploited for the construction of various protein nanostructures. One such example was reported by Ward et al., where they synthesized a biofunctionalized oligoethylene glycol (OEG) linker composed of bis-biotinylated and terpyridine moieties. The terpyridine functionality of a

molecule, upon chelation with a metal ion *via* ligand-metal ion complexation, yields a linear tetrabiotinylated product. Now, upon the addition of streptavidin, these exposed linear tetrabiotin moiety of the molecule undergo spontaneous polymerization by the specific coupling with streptavidin. This spontaneous polymerization then resulted in the formation of a 1D metal-organic protein framework (Figure 1.6a,b).⁴²

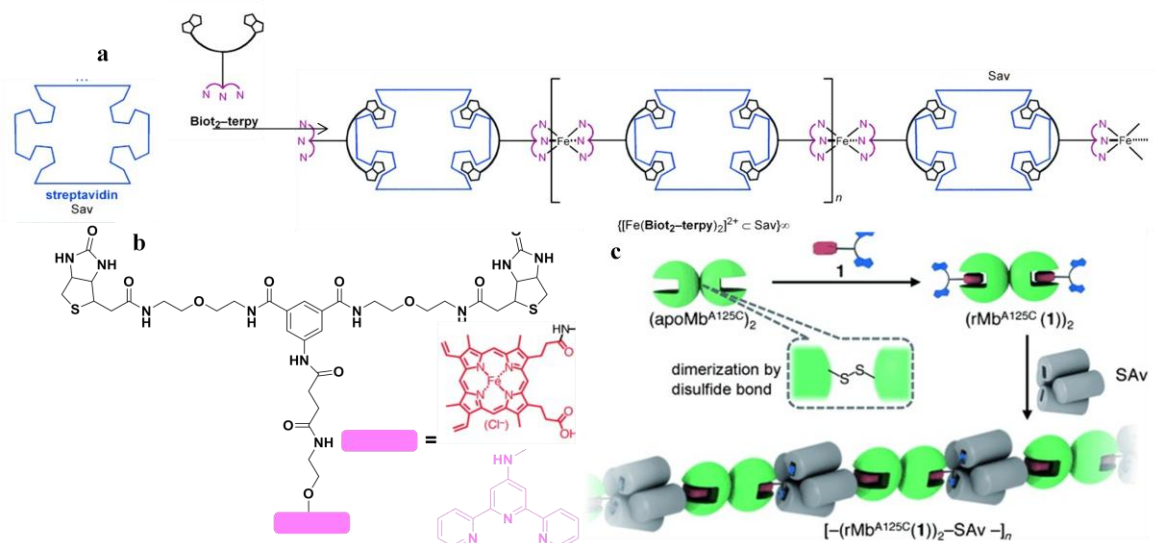


Figure 1.6 | (a) Schematic representation of the supramolecular streptavidin polymerization induced by metal-ligand chelation. Adapted with permission from ref [42]. Copyright @ 2007, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. (b) Chemical structure of divalent linker composed of bis-biotinylated and terpyridine or heme moieties. (c) Schematic representation of the programmed assembly of dimeric myoglobin and streptavidin. Adapted with permission from ref [43]. Copyright @ 2012, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.

In another example, a similar strategy was used to design linear protein assemblies using two orthogonal protein building blocks composed of streptavidin and dimeric myoglobin. In this work, authors have synthesized a target compound incorporating two biotin units tethered through a flexible linker to one heme cofactor-propionate side chain. This, the bis(biotin)-heme cofactor, could then crosslink the proteins streptavidin and dimeric myoglobin to form a linear AAB type supramolecular protein assemblies (Figure 1.6c).⁴³

Lectins and other carbohydrate-binding proteins are the biomacromolecules that bind to specific configurations of sugar molecules and perform various crucial cellular processes.⁴⁴ Jiang and coworkers employed such specific interactions between sugar and its cognate protein partner to establish a novel protocol to design 3-D protein crystalline

frameworks. In one such example, authors have synthesized the ligand composed of monosaccharide and rhodamine groups connected by OEG spacer (Figure 1.7a). The monosaccharide portion of this heterobifunctional ligand first binds to the homotetrameric lectin concanavalin A (ConA) protein. The resultant conjugate then undergoes the dimerization of rhodamine-B (RhB) *via* π - π stacking (Figure 1.7b). These dual supramolecular interactions between the conjugate result in the interpenetrating protein crystalline framework (Figure 1.7c).⁴⁵ Surprisingly, when the protein ConA was replaced with D₂ symmetric homotetrameric soybean agglutinin, a microtubule-like structure was obtained.⁴⁶ Authors have attributed the difference in self-assembly to the distinct protein geometries of ConA (tetrahedral) and soybean agglutinin (nearly planar, D₂ symmetric) (Figure 1.7d).

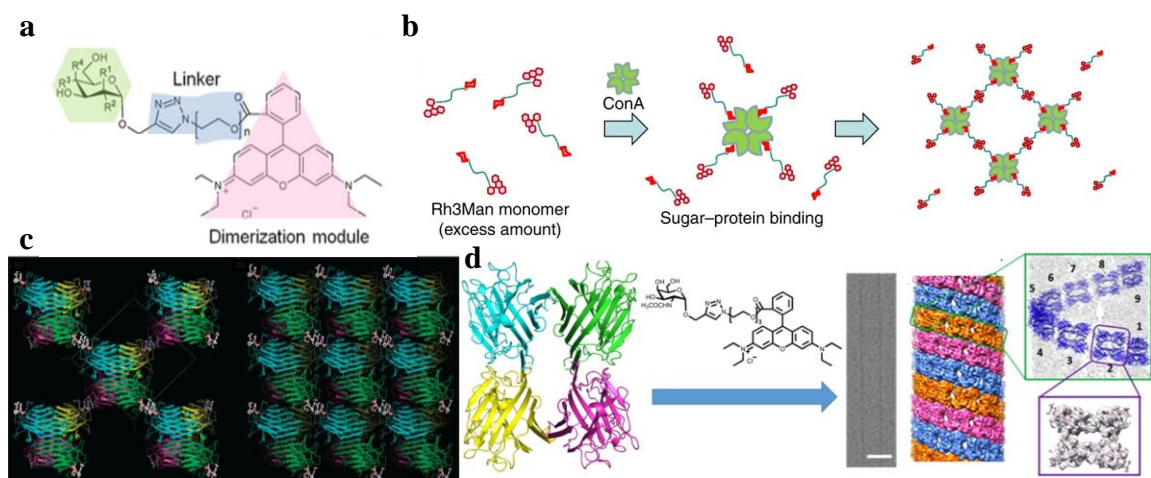


Figure 1.7 | (a) Chemical structure of the self-assembly promoting ligands Adapted with permission from ref. [46]. Copyright @ 2016, American Chemical Society. (b) Possible sequences of Man/ConA interaction and RhB dimerization. (c) The layer structures in the crystals. Adapted with permission from ref. [45]. Copyright @ 2014, American Chemical Society. (d) Schematic representation of helical protein microtubule. Adapted with permission from [46]. Copyright @ 2016, American Chemical Society.

1.3.1.3. Synthetic Ligand-Protein Interactions Driven Self-Assembly

The design of novel protein assemblies could also be achieved by designing high-affinity synthetic ligands that bind to natural proteins. Wagner et al. developed a strategy to synthesize protein assemblies by utilizing the specific interactions between the dihydrofolate reductase (DHFR) fusion proteins and ligand, *i.e.*, methotrexate (MTX). They

have synthesized dihydrofolate reductase (DHFR2) with inter varying domain lengths, and its corresponding bis-MTX ligand. This fusion DHFR2 protein upon binding to the bis-MTX ligand spontaneously forms cyclic structures of size 8 to 20 nm. They have also shown that the diameters of these cyclic assemblies can be modulated by the length of peptide linkers (Figure 1.8).⁴⁷

In the follow-up work, Wagner et al. had extended this strategy to construct enzyme nanorings of size about 10–70 nm by using a fusion construct composed of a DHFR-histidine triad nucleotide-binding 1 (hHint1). Interestingly, the hHint1 enzyme of nanoring exhibited size-dependent enzymatic activity, which ultimately provides the ability to regulate the catalytic activity of an enzyme.⁴⁸ The same authors have also reported the self-assembly of antibody nanorings. Here, they produced dimeric DHFR fusion protein containing antiCD3 single-chain variable fragment (scFv). The oligomerizing of this construct (DHFR2antiCD3) in the presence of bis-MTX ligand generated antibody decorated nanorings. Interestingly, similar to the native antiCD3 antibody, these antibody decorated nanorings were encountered to interact with CD3+ T cells in a tissue-specific manner.⁴⁹

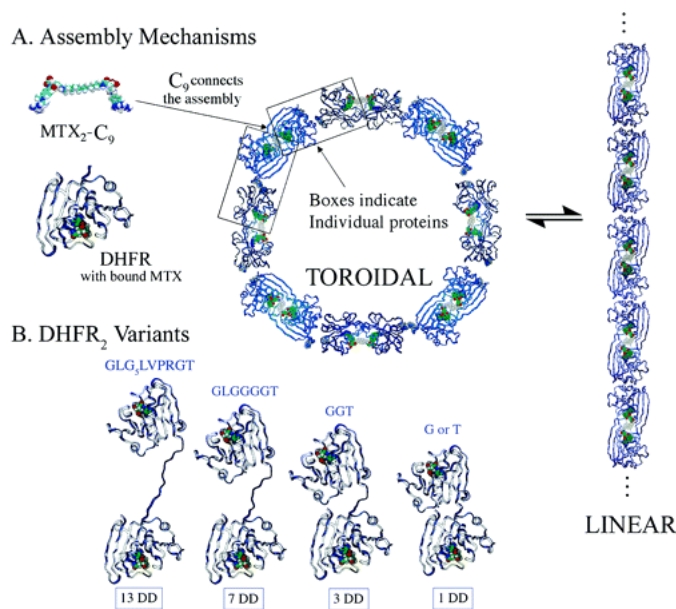


Figure 1.8 | Schematic representation of the design principle to construct protein nanorings by using bis-MTX and DHFR2. Adapted with permission from [47]. Copyright @ 2006, American Chemical Society.

In a further development, Mrksich and co-workers have reported an interesting alternative approach to design giant cyclic molecules of diameters of 10–20 nm. The authors constructed the nanorings by synthesizing fusion protein of *N*-terminal cutinase and *C*-terminal SnapTag domains that react irreversibly with a heterobifunctional ligand composed of *p*-nitrophenyl phosphonate (pNPP) and benzylguanine (BG) groups, respectively. Further, bis-BG and a BG-pNPP linker are used to join these fusion proteins into linear structures that can then react with a bis-pNPP linker to joins the ends to form the cyclic product.⁵⁰

1.3.2. Metal Ion Induced Protein Assembly

Metal ions are frequently found in living systems in association with proteins to perform important biological processes.⁵¹ One such class of proteins are called metalloproteins, where metal ions play a vital role in protein function.⁵² Inspired by the metalloproteins, the scope of metal coordination chemistry has been extended to design protein assemblies.

Tezcan et al. pioneered the use of proteins as a building block in combination with the metal ion to design various protein assemblies.⁵³ They initially constructed structurally more stable cytochrome *cb*-562 protein, which is the variant of four-helix bundle hemeprotein cytochrome *b*-562.⁵⁴ Further, the cytochrome *cb*-562 protein was then modified to install high-affinity two bidentate ligands at *i*, *i*+4 positions of one of its α -helices. At particular Zn^{2+} /protein ratios, this modified protein was found to self-assemble into D_2 -symmetrical tetrameric assembly stabilized by four Zn^{2+} ions.⁵⁵ Based on the foundations set by the above studies, the methodology was further explored to design extended molecular arrays using computational methods (Figure 1.3).³¹ (Read section 1.1.3)

Another impressive study of metal-mediated protein self-assembly was reported by Aida and co-workers. In one of the examples, they have synthesized micrometer-long hollow cylinders protein assembly of chaperonin GroEL. The chaperonin GroEL is a barrel-shaped tetradecameric protein with an inner diameter of 4.5 nm. The authors first created the mutant of GroEL by introducing 14 cysteine residues on each one of its opening

part of the barrel. This mutant protein was then conjugated with photochromic unit spiropyran (SP) *via* thiol-maleimide chemistry. They have demonstrated that the nonionic spiropyran unit appended to GroEL spontaneously transforms the ionic merocyanine (MC) unit in Tris-HCl buffer containing MgCl_2 (Figure 1.9a). The resultant GroEL_{MC} , in presence divalent metal ions (Mg^{2+}), self-assembles into nanofibers *via* the polymerization through several $\text{MC}\cdots\text{Mg}^{2+}\cdots\text{MC}$ bridges.⁵⁶ Interestingly, when strong chelating agent EDTA for Mg^{2+} was added to the solution of GroEL nanofibers, these long cylinders dissociated into monomeric GroEL_{MC} units.

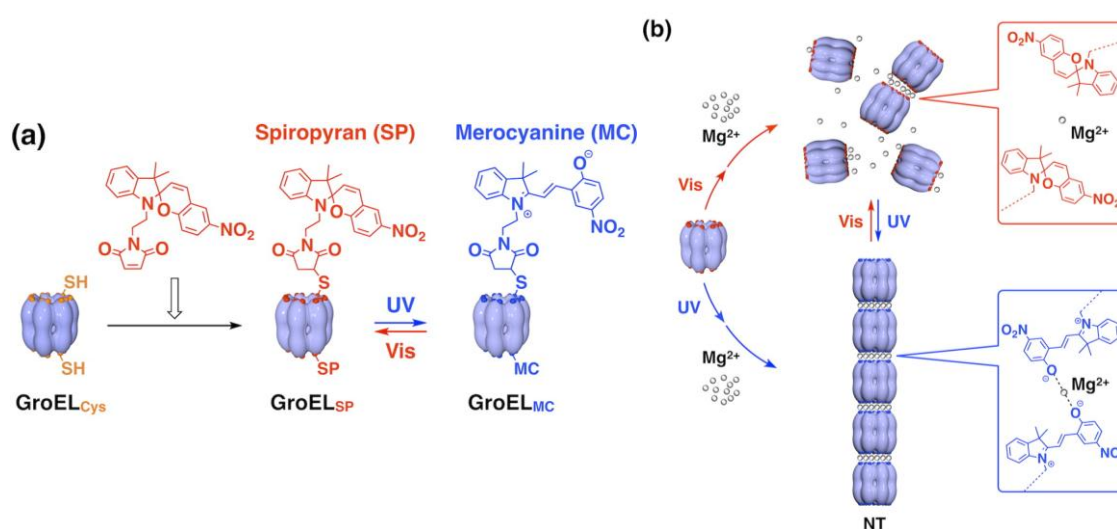


Figure 1.9 | Schematic illustrations of (a) the synthesis of multiple spiropyran labeled GroEL_{SP} , (b) Light-mediated assembly/disassembly GroEL_{SP} and GroEL_{MC} . Adapted with permission from [57]. Copyright @ 2009, American Chemical Society.

In another study, authors envisioned to control the formation and dissociation of these nanotubes by light. This is hypothesized on the fact that $\text{MC} \rightarrow \text{SP}$ and $\text{SP} \rightarrow \text{MC}$ isomerization can be accomplished by exposing the sample to visible light, and UV light, respectively. In order to achieve this, Aida and co-workers first prepared the GroEL nanotubes in Tris-HCl buffer containing MgCl_2 by exposing the sample to UV light, which promotes the $\text{SP} \rightarrow \text{MC}$ isomerization. Surprisingly, in spite of $\text{MC} \rightarrow \text{SP}$ isomerization, formed GroEL nanotubes did not break into consecutive monomers upon exposure to visible light. In sharp contrast to the above example, nanotubes prepared under UV light were barely dis-assembled upon treatment with EDTA. The authors have hypothesized that this unexpected behavior could be because of the formation of radical species upon

photoexcitation of MC, which then undergoes cross-linking to form long nanotubes. In order to prove this, authors have carried out UV-mediated polymerization of GroEL_{MC} in the presence of radical scavenger dithiothreitol (DTT). Surprisingly, with this approach, GroEL nanotubes disassembled into consecutive monomers upon exposure to visible light as a result of MC → SP isomerization. (Figure 1.9b).⁵⁷ Interestingly, these cylinders were also demonstrated to trap denatured proteins. Thus the scaffold can be used as an artificial biocontainer, which can serve as a delivery vehicle for a variety of guest molecules.

Another alternative strategy to construct protein assemblies *via* metal coordination is reported by Liu and co-workers. The authors have designed a C₂ symmetric, homodimeric, globular GlutathioneS-transferase (GST) protein with His-tags. The GST protein is designed in such a way that two of its His-tags are exposed outside of the protein surface. These His-tag groups then undergo chelation with Ni²⁺ to form one dimensional protein nanowires. Then, utilizing the same strategy, the authors have envisioned designing more complex protein nanorings. In order to do that, they engineered C₂ symmetric GST protein with two chelating His-tag groups oriented to form a V-shaped design. This construct then upon chelation with metal ion along with synergistic non-covalent interactions of GST-2His self-assembled into highly ordered protein nanorings.

1.3.3. Protein-Polymer Conjugates

The initial efforts in the area of protein-polymer synthesis have mainly focused upon intending to improve the stability, chemical properties, and functions of proteins by the judicious selection of the polymer and the bioconjugation reaction.⁵⁹ Later, the use of protein-polymer conjugates has also been explored for designing different protein architectures *via* a bottom-up approach. In this approach, the synthetic polymeric domain of protein-polymer conjugate serves as a driving force for protein to self-assemble. Until now, various methodologies are reported for the synthesis of protein-polymer conjugates *via* specific coupling of polymer to the protein of interest. These strategies are broadly divided into two methods, which are discussed below.

- (I) Non-covalent coupling – In this approach, polymers are specifically appended onto proteins mainly by remodeling non-covalent interactions present in natural

enzymes/protein and their corresponding polymer functionalized cofactors/ligands.

- (II) Covalent coupling – In this method, polymers are covalently appended onto the protein. There are mainly two ways to achieve this. (a) A common approach to produce protein-polymer using the covalent coupling method involves a bioconjugation of polymer to a single reactive site on the surface of the protein. (b) Another way to achieve the protein-polymer conjugate is to synthesize protein macroinitiators by installing a heterobifunctional spacer onto the protein. Then the polymer is grown on this protein macroinitiator by situ controlled polymerization.⁶⁰

1.3.3.1. Non-Covalent Coupling

1.3.3.1.1. Affinity Binding

Nolte and co-workers have utilized one of the strongest natural non-covalent interactions, i.e., streptavidin-biotin, to construct giant amphiphiles. The monobiotinylated polystyrene was first synthesized by reacting amine-terminated polystyrene with the carboxyl group of biotin. This biotinylated polystyrene was then subjected to the binding with streptavidin using a monolayer method, where the biotin terminated polystyrene molecule is first distributed on the air/water interface. Then, the streptavidin is added to the above subphase. The atomic force microscopy (AFM), Brewster angle microscopy revealed that out of four biotin-binding pockets of streptavidin molecule, two of them occupied by two biotinylated polystyrene polymer. Then, the remaining vacant binding sites of streptavidin were demonstrated to bind with other pairs of biotin, which are functionalized with iron storage protein ferritin and horseradish peroxidase (HRP). Interestingly, the HRP appended to the streptavidin bioconjugate retained catalytic activity.⁶¹

Stayton and co-workers have reported another example in which streptavidin is used as the building block to design polymer-protein hybrids. Here, the streptavidin and poly(*N*-isopropyl acrylamide) (PNIPAAm) conjugates with low polydispersity were synthesized.⁶² The synthesized polymer-protein conjugates self-assembled to make protein nanoparticle

upon heating, and revert back to the monomeric state upon cooling. This is because, above the lower critical solution temperature (LCST), hydrophilic PNIPAM polymer transforms into a hydrophobic polymer, which then promotes the self-assembly of conjugates *via* hydrophobic interactions (Figure 1.10).⁶² Interestingly, the average hydrodynamic diameters of these particles can be tuned between ~250–900 nm by tuning the concentration, polymer molecular weight, and rate of temperature change.

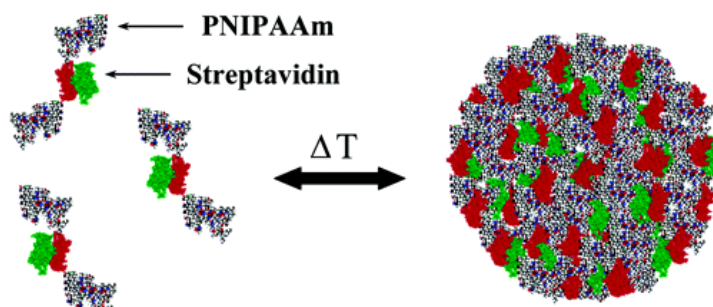


Figure 1.10 | Schematic representation of the hydrophobic interaction-driven self-assembly of PNIPAAm–streptavidin conjugates. Adapted with permission from [62]. Copyright @ 2004, American Chemical Society.

1.3.3.1.2. Co-Factor Reconstitution

Apart from streptavidin-biotin non-covalent interaction, another exciting approach to synthesize giant amphiphiles was reported by Nolte and co-workers *via* the cofactor reconstitution method, in which the cofactor of an enzyme is functionalized by polymer, followed by reconstitution with its corresponding apoenzyme (Figure 1.11). As an example, one of the carboxylic groups of heme-cofactor is functionalized by polystyrene polymer *via* a hydrophilic linker. Further, the THF solution containing heme cofactor-functionalized polystyrene polymer is treated with an HRP apoenzyme to get the desired giant amphiphiles. The self-assembly studies of the giant amphiphiles in water by using electron microscopy unveiled the presence of 80-400 nm diameter aggregates with morphology analogous to vesicles. Surprisingly, assemblies did not show any enzymatic activity reconstitution reaction was performed at 4 °C. However, enzymes recovered substantial enzymatic activity when a similar reconstitution reaction was performed at 22 °C.⁶³ A similar strategy was explored to synthesize giant amphiphiles from oxygen-carrying protein myoglobin, and polystyrene functionalized with haem co-factor. Similar to HRP conjugate, the myoglobin conjugate also found to self-assemble into vesicular structures and retained

its protein activity. This strategy was further extended to construct biohybrid triblock copolymers composed of protein or an enzyme, polystyrene, and polyethylene oxide. Interestingly, depending on the molecular design of the bioconjugate, such as choice of protein and the hydrophobic block length, varieties of assemblies (micellar rods, vesicles, toroids, octopus structures, and spheres with a lamellar surface) are reported.⁶⁴

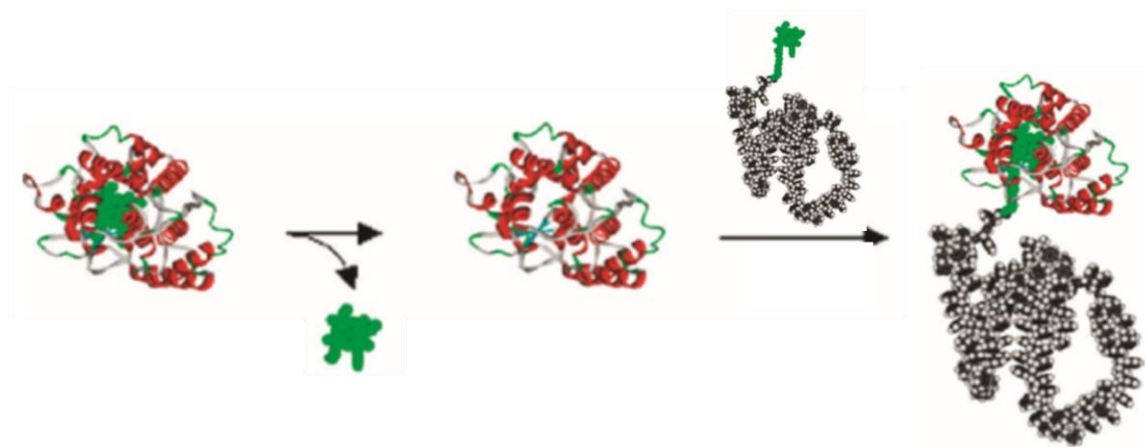


FIGURE 1.11 | Schematic representation of the cofactor reconstitution technique to prepare giant amphiphiles. Adapted with permission from ref. [21]. Copyright @ 2008, American Chemical Society.

Inspired by the seminal work of Nolte and coworkers, Liu and co-workers have reported the synthesis of thermoresponsive protein-polymer bioconjugates *via* cofactor reconstitution.⁶⁵ The thermo-responsive polymer was synthesized by first functionalizing porphyrin moiety with azide-terminated ethylene glycol, followed by a click reaction with alkynyl-PNIPAM. Subsequently, cofactor reconstitution between thermo-responsive cofactor and apomyoglobin afforded well defined protein-polymer bioconjugates. The obtained thermoresponsive bioconjugate was verified to undergo thermal-induced aggregation behavior in aqueous solution (Figure 1.12). This is the first report demonstrating the method to construct thermo-responsive protein assemblies *via* the cofactor reconstitution process.

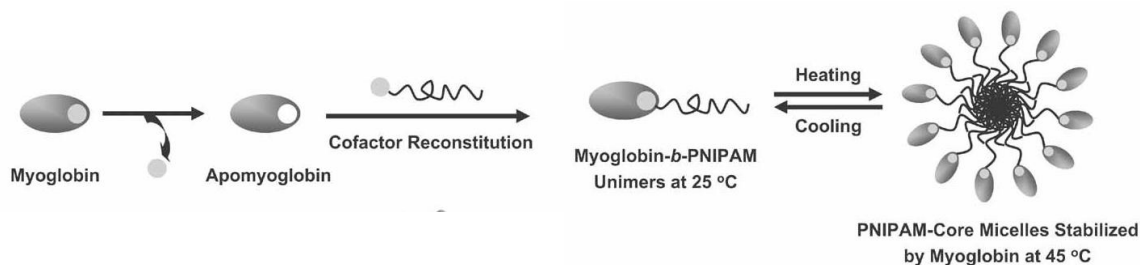


Figure 1.12 | Schematic representation for the synthesis of thermoresponsive nanoassemblies of myoglobin-PNIPAM conjugates *via* the cofactor reconstitution approach. Adapted with permission from ref [65]. Copyright @ 2010, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.

1.3.3.1.3. Metal-to-Ligand Coordination

The use of metal-ligand interactions to create protein assemblies began from simply modifying the proteins/enzymes (Bovine Serum Albumin (BSA)/ Candida Antarctica (CALB)) to create artificial metal coordination points such as terpyridine moiety.⁶⁶ Further, the addition of mono-(terpyridine)-ruthenium(II) functionalized polystyrene to the solution of terpyridine functionalized biohybrids lead to the formation of protein-polymer giant amphiphiles *via* metal-ligand coordination. Interestingly, in this case, distinct self-assembled architectures of giant amphiphiles were encountered compared to its analogous, which is synthesized by direct covalent coupling.

1.3.3.2. Covalent Coupling

Unlike non-covalent coupling strategies, direct covalent bioconjugation provides a diverse platform for the synthesis of desired protein-polymer conjugates. This type of bioconjugation approach involves the direct covalent coupling between the end-functionalized polymer and surface-exposed amino acid of a protein. However, the selection of appropriate bioconjugation reaction for precise control over the placement of the polymer onto protein remains a challenge due to the following reasons.

- (I) The bioconjugation reactions should target only one specific amino acids in the presence of roughly 200 or more amino acids, in order to achieve the synthesis of a defined protein-polymer conjugate. If the target amino acids are present in two or more copies, then it becomes challenging to control the bioconjugation reaction to one particular amino acid.
- (II) This unique amino acid has to be surface exposed and not buried in the interior of the protein structure. Further, modification of such amino acid should not perturb the protein's structure

(III) Most importantly, the bioconjugation reaction conditions should have fast reaction kinetics.

By considering the points mentioned above, significant advances have been made in the field of bioconjugation chemistry. The efforts here are mostly focused upon (a) developing chemical strategies which can specifically target only one copy of surface-exposed amino acid (lysine, glutamate, aspartate, and cysteine) without affecting proteins structure and function, (b) re-engineering of protein to contain a single free unique residue for conjugation to a polymer.⁶⁷

Nolte and co-workers utilized such bioconjugation reactions to construct giant amphiphiles through direct covalent coupling of a hydrophobic polymer to the single amino acid of the protein.⁶⁸ In this study, a specific reduction of surface-exposed disulfide bond of the lipase B from CALB was achieved by using DTT. One of the free thiol was then conjugated to maleimide functionalized polystyrene *via* thiol-maleimide chemistry in a THF/buffer solvent mixture. The self-assembly studies using TEM unveiled the formation of micrometer-long fibers. These fibers are made up of the bundles of rods. The diameter of the smallest rod was found to be around 25 to 30 nm (Figure 1.13). However, these nanoassemblies found to enzymatically less active (6-7% of the activity) compared to native CALB enzyme. This could be because of the perturbation of the enzyme's structure or the unavailability of its active site.

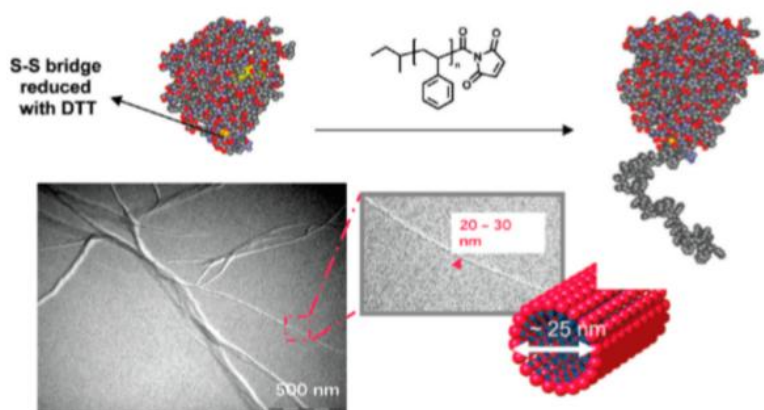


Figure 1.13 | Schematic scheme for the preparation of CALB-polystyrene conjugate, which self-assembles to form micellar rods. Adapted from ref. [68] with permission. Copyright @ 2002, American Chemical Society.

Apart from conjugating hydrophobic polymers to proteins, polymers having LCST property were also conjugated to proteins through either genetic approaches or chemical methodology.⁶⁹ Olsen and co-workers reported the strategy to synthesize high-density protein assemblies from protein-polymer block copolymer conjugates in bulk/thin films.⁷⁰ In one such work, a mutant engineered red fluorescent protein was expressed by introducing a new cysteine amino acid, which was conjugated to PNIPAAm polymer *via* thiol-maleimide chemistry.^{70a} The self-assembly study of this conjugate was performed by evaporating water from the solution of the mCherry-PNIPAAm conjugate. Above the LCST, hydrophilic PNIPAAm turns into the hydrophobic. Therefore, water acts as a non-selective solvent for the polymer, but selective for protein. At room temperature, PNIPAAm is hydrophilic. Thus act as a selective solvent for the system.

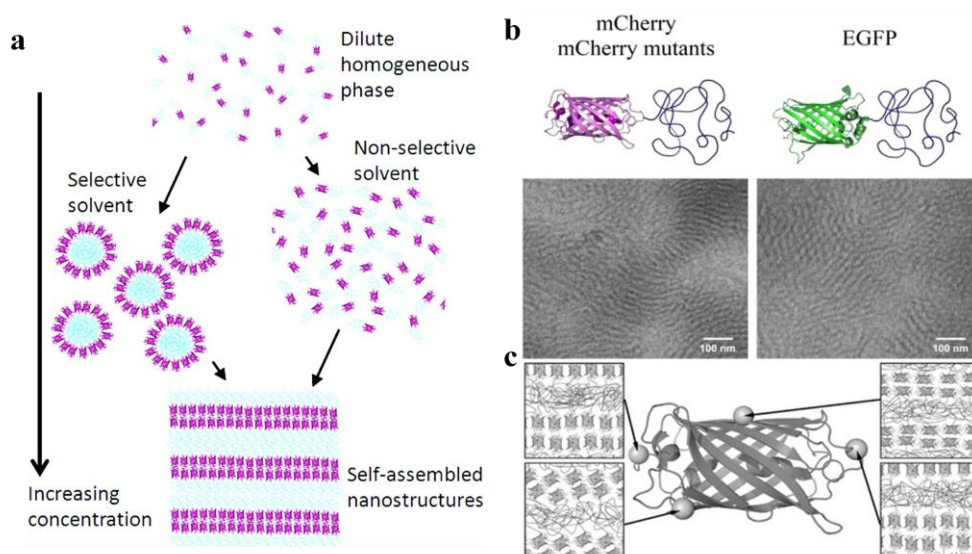


Figure 1.14 | (a) Possible pathways for mCherry–PNIPAM bioconjugates to self-assemble in the water at room-temperature (nonselective solvent) and 40 °C (protein-selective solvent). Adapted with permission from ref. [70a] Copyright @ 2011, American Chemical Society. (b) Similar self-assembly behavior of mCherry–PNIPAM and GFP–PNIPAM conjugates. Adapted with permission from ref. [71] Copyright @ 2014, American Chemical Society. (c) Cartoon schematic showing the effect of each conjugation site on self-assembly. Adapted with permission from ref [73]. Copyright @ 2010, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.

The sample cast at room temperature (non-selective solvent) displayed a relatively intense lamellar and weaker set of hexagonal lattice assemblies. However, poorly ordered assemblies were encountered in protein selective (40 °C) solvent due to the immiscibility of protein and polymer (Figure 1.14a). Further, annealing of the solvent was found to enhance the ordering in assemblies with clear lamellar symmetry indicating hexagonal and lamellar assemblies are in equilibrium. Interestingly, depending upon the conjugate contents (weight percentage), and solvent selectivity, a phase diagram was created. This phase diagram exhibited various morphologies. The strategy to construct dense protein nanostructures has also been extended to investigate the effect of proteins electrostatic interactions,⁷² and surface potential (Figure 1.14b),⁷¹ bioconjugate shape (Figure 1.14c),⁷³ and polymer⁷⁴ on self-assembly.

The examples mentioned above of protein-polymer conjugates rely upon specific bioconjugation of the surface-exposed amino acid of protein with an end-functionalized polymer. This method creates an obstacle to employ such specific bioconjugation chemistry to diverse proteins because the majority of proteins contain several copies of surface-exposed amino acids having a similar reactivity. This might lead to the polydispersity in the product due to the multiple labeling of protein. Hence, the existing methodologies are mostly applicable to few proteins which contain unique surface-exposed amino acids such as GFP, BSA and, lipase. However, incorporation of bio-orthogonal amino acids into proteins provides a unique functional group that would expand the scope of specific bioconjugation chemistry to the distinct proteins. Matyjaszewski et al. reported step-growth polymerization of p-azidophenylalanine integrated green fluorescent protein (GFP) with dialkyne terminated polyethylene glycol (PEG). The resultant bioconjugate undergoes a reversible formation of micron-sized fibers driven by dimerization of GFP *via* hydrophobic interaction.⁷⁵ Although, this method can potentially provide mono-labeled conjugate, the reduced efficiency of unnatural amino acid incorporation into protein limits the use of amber-codon suppression method for protein engineering.⁷⁶ Besides, the chemical diversity of polymeric domain of protein-polymer conjugates has been mostly constricted to only a few linear polymers such as polystyrene and PNIPAAm. Hence, as a result of inherent polydispersity, these linear polymers induce a degree of heterogeneity to the conjugate. Thus, the limited opportunity in the structural diversity of protein and

polymeric domains obstructs the systematic understanding of their structure-property relationship. In addition to that, the resulting heterogeneous protein-polymer bioconjugates mostly lack detailed standard analytical and biophysical characterization.

1.3.3.3. Design of Monodisperse Facially Amphiphilic Proteins

Considering the existing limitations, our group has disclosed the novel design strategy to construct monodisperse facially amphiphilic proteins bioconjugates.⁷⁷ In this work, the site-specific conjugation of proteins was achieved by using an amphiphilic activity-based probe (AABPs), which is composed of a reactive head group, hydrophilic linker, and hydrophobic tails. The reactivity of AABPs towards the protein is accelerated by the geometry of active site. Thus, this strategy gives excellent control over the placement of synthetic moiety at the single active site residue, yielding a well-defined conjugate. The reactive head group of AABP is composed of fluorophosphonate (FP), which reacts with the serine amino acid present in the active site of a serine protease. The rationale behind choosing serine proteases as a target protein is, this class of protease, being the largest proteases family, provides an opportunity to synthesize conjugate with a diverse head group. The oligoethylene glycol was utilized as a linker due to its water solubility and protein repellent properties. Finally, linear 1-tail, 2-tail, and 3-tail alkyl chains were used as a hydrophobic moiety.

The poor solubility of these AABPs in aqueous media posed a problem in bioconjugation. In order to surmount this problem, we invented a micelle-assisted protein labeling (MAPLab) strategy. This technology utilizes small molecule micelle to solubilize amphiphilic probe effectively, where the hydrophobic tail component of a probe would be part of the hydrophobic micellar interior of small molecule micelle, and a hydrophilic part along with protein reactive group would be exposed outside. Further, the reaction mixture was purified by using a three-step purification strategy. First, triton X-100 was removed by using ion-exchange chromatography. Then, to remove unreacted protein from the conjugate, ionic strength of the solution was increased. This led to the self-sorting conjugate through strong hydrophobic interaction, and they eluted at earlier, whereas the unreacted protein eluted later. Using this strategy, we have synthesized a library of facially

amphiphilic protein conjugates, with different protein head group, linker length, and hydrophobic tails. The self-assembling studies of these conjugates by using various complementary techniques such as SEC, DLS, SEC-MALS, and SANS revealed these conjugates self-assemble into protein complexes of defined size and shape driven *via* strong hydrophobicity interactions of tails opposed by electrostatic repulsion between protein head group. Importantly, this method offers control over the hydrodynamic radius, oligomeric state, and the molecular weight of the complex by either tuning protein head group, linker length, and hydrophobicity or branching of tails in the molecular design.

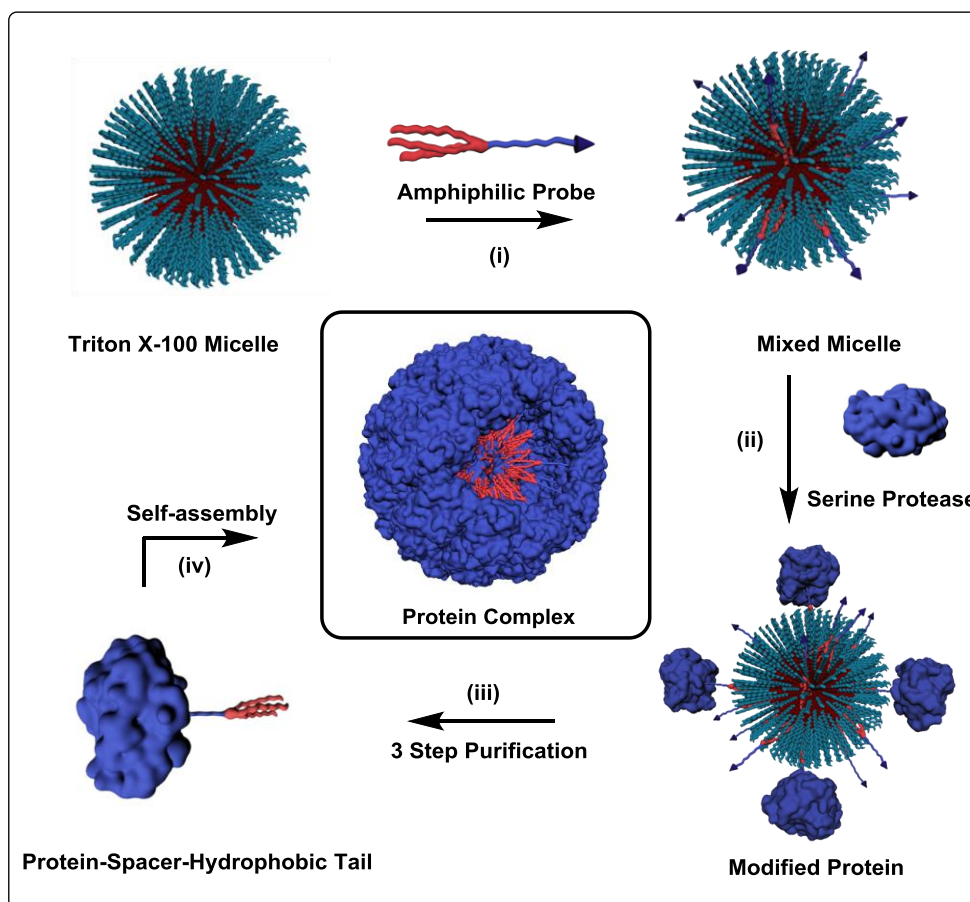


Figure 1.15 | Schematic representation of micelle-assisted protein labelling technology. Adapted with permission from ref [77]. Copyright @ 2018, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.

1.3.3.4. Protein-Dendron Conjugates

An alternative approach to synthesize monodisperse protein conjugates is demonstrated by attaching "hydrophilic dendron" to the proteins. Dendrimers are three dimensional, highly branched, monodisperse, symmetric, multifunctional, nanoscopic polymers synthesized in well-controlled conditions.⁷⁸ The structure of the dendrimer is classified into three main components (Figure 1.16).

- (I) The functional branched core, which determines the shape of the dendrimer.
- (II) The branching units, which contribute to the interactions with guest molecules.
- (III) The surface groups, which provide multivalency to the scaffold.

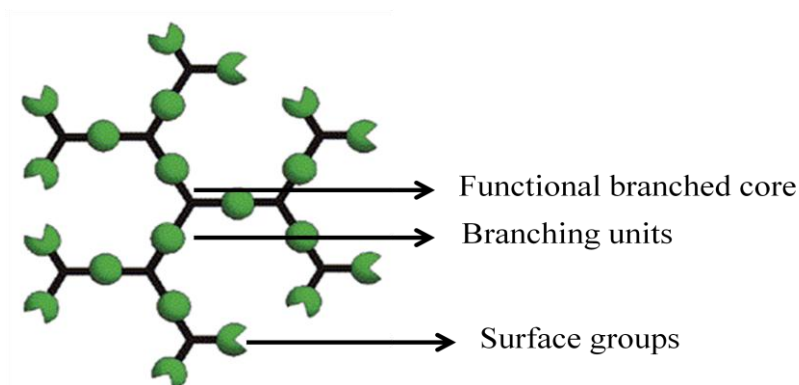


Figure 1.16 | Cartoon representation for the structure of dendrimer and its components. Adapted with permission from ref. [78b] Copyright @ 2001, American Chemical Society.

The synthesis of dendrimer is achieved by using two complementary synthetic approaches, i.e., divergent and convergent approaches. The divergent synthesis of dendrimer was introduced by the seminal work by Tomalia⁷⁹ and Newkome⁸⁰, whereas the convergent synthetic approach was introduced by Fréchet⁸¹. The main difference between the two approaches is in the directionality of dendrimer growth. In a divergent approach, the starting multifunctional unit will become the core of a molecule, and the dendrimer growth proceeds inward to the outward direction (Figure 1.17a). In a convergent approach, a unit that initiates growth will become the exterior, and dendrimer growth proceeds outward to an inward direction (Figure 1.17b).

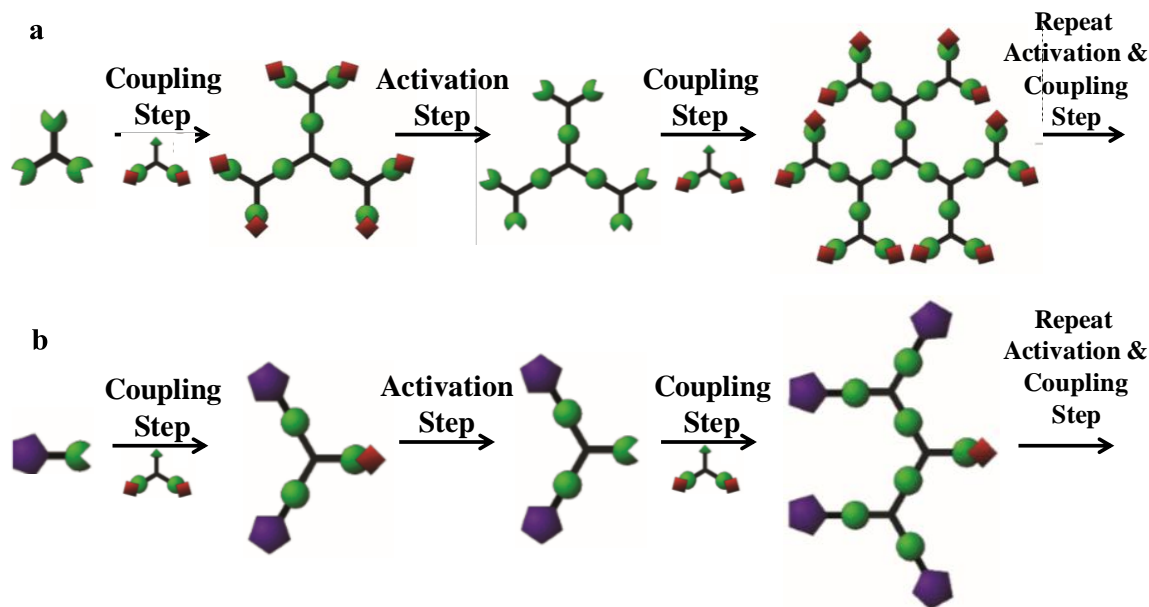


Figure 1.17 | Schematic representation of (a) divergent route and (b) convergent route for the of dendrimer. Adapted with permission from ref. [78b] Copyright @ 2001, American Chemical Society.

Over the last two decades, dendrimers have emerged as a new class of branched, multivalent, monodisperse polymeric scaffold of significant importance. The above mentioned unique features of this macromolecule that make them promising candidates for various applications such as drug delivery, gene delivery, imaging agents. Apart from that, modern advancements in chemical coupling strategies that allow efficient conjugation of self-assembling dendrimers and biological macromolecules such as peptides,⁸¹ nucleic acids,⁸³ carbohydrates⁸⁴ are reported in the literature to obtain various respective assemblies. However, there are only a few reports are available on protein–dendron conjugates. These reports include conjugation of hydrophilic dendrimers to proteins. The conjugation of streptokinase with poly(amidoamine) (PAMAM) dendrimer was found to improve the stability of the protein. The equimolar SK-PAMAM conjugates exhibited the highest enzymatic activity retention compared to traditional polymers such as PEG or dextran.⁸⁵ The same dendrimer was also used to generate dendrimer–antibody conjugates. These conjugates can also be used for the specific delivery of multiple molecules such as a drug, targeted *via* an antibody.⁸⁶ Davis and co-workers have reported the construction of a well-defined hybrid glycoprotein enzyme by conjugating a single cysteine group in the

protein-degrading proteinase, and subtilisin with branched, dendritic saccharide. These hybrid glycoenzymes were found to inhibited bacterial aggregation.⁸⁷ The glycoprotein composed of insulin and sialyloligosaccharides were encountered to show prolonged glucose-lowering effects.⁸⁸

Olli Ikkala and co-workers report other examples of synthesis of precisely defined protein–dendron conjugates by conjugating Newkome-type dendrons containing multiple spermine (natural DNA binding a naturally linear polyamine) units to the HFBI and BSA. Further, DNA binding studies of these conjugates revealed that the HFBI-G1 and BAS-G1 conjugates exhibited weaker binding to DNA at a high salt concentration, whereas BSA-G2 and HFBI-G2 showed less effect by the increase in salt concentration⁸⁹. Although there are few reports available for the synthesis of hydrophilic protein-dendron conjugates, to the best of our knowledge, there are no reports on the synthesis of amphiphilic, stimuli-responsive protein-dendron conjugates.

1.4. Current Challenges and Aim of the Thesis

1.4.1. Background

It is pretty evident from previous studies outlined in previous sections that chemical strategies have vastly expanded the scope of protein nanotechnology. Among the various methods, the synthesis of amphiphilic protein-polymer conjugate *via* chemical conjugation of proteins with polymer (hydrophobic or temperature-sensitive polymer) - that serve as a driving force for proteins to self-assemble - has significantly contributed to the construction of artificial ordered protein nanostructures *via* a bottom-up approach. However, a significant disadvantage to the use polymer is that they induce a degree of heterogeneity in the form of polydispersity resulting in heterogeneous samples, which are hard to characterize by standard analytical and biophysical tools.

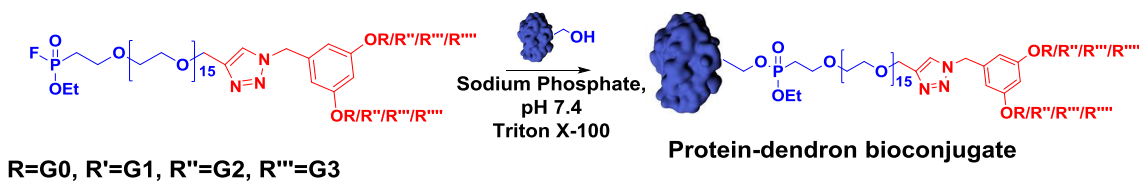
The replacement of polydisperse linear polymer with monodisperse branched dendrimer is considered to be a viable approach to synthesize protein-polymer conjugates with an ideal PDI of 1.00. The existing methodologies present tools for synthesis of globally hydrophilic protein-dendron conjugates *via* conjugating hydrophilic monodisperse

dendron to hydrophilic/amphiphilic proteins. However, the majority of these conjugates cannot self-assemble due to a lack of amphiphilic nature. In this regard, the custom design of protein-dendron amphiphilic macromolecules is very interesting because of their propensity to self-assemble into various nanoscopic objects.

These protein-dendron based biomimetic nanoscopic complexes would be an excellent scaffold for the applications in various fields such as vaccine design, targeted drug delivery, and diagnostic imaging. This is because, unlike other systems, interior domains of these micellar assemblies can be decorated with a variety of therapeutic agents. Similarly, the exterior surface can be site-specifically labeled with targeting units to achieve specific drug delivery to target tissue or cell or organelle. However, to date, there are no reports based on this concept due to challenges associated with the chemical synthesis of amphiphilic protein-dendron bioconjugates. Therefore, the development of chemical strategies that will enable the chemical synthesis of protein-dendron based complexes would be highly beneficial. In this direction, the aim of this research is focused on developing strategies for the construction of monodisperse facially amphiphilic, stimuli-responsive protein-dendron conjugates, and understand their self-assembly and dis-assembly behavior.

1.4.2. Chapter 2: Design, Synthesis and Self-assembly Studies of Suite of Monodisperse Facially Amphiphilic Protein-Dendron Conjugates

In this work, we have established a chemical methodology for the synthesis of monodisperse, facially amphiphilic protein-dendron bioconjugates. These conjugates are synthesized by a site-selectively conjugating macromolecular amphiphilic activity-based probes (AABPs) (composed of a fluorophosphonate as a reactive group, monodisperse unhexa(ethylene glycol linker), and hydrophobic dendrimer of G1-G4 generation) to the catalytic serine residue of four serine protease, i.e., trypsin, chymotrypsin, subtilisin, and proteinase K by using MAPLab technology (Scheme 1).



Scheme 1 | Scheme for the synthesis of protein-dendron bioconjugates.

Then, the reaction mixture was purified using a two-step purification strategy. First, triton X-100 and unreacted probe in the reaction mixture were removed using ion-exchange chromatography. Then, to remove native protein from conjugate, ionic strength of the medium is increased. This leads to the self-sorting of protein-dendron conjugates due to their amphiphilic character and hence eluted at higher elution volume compared to native hydrophilic monomeric protein. MALDI-ToF analyses of purified protein-dendron bioconjugate showed a single peak indicating their monodisperse nature and successful site-specific bioconjugation reaction (Figure 1.18a, b).

Further, the ability of protein-dendron bioconjugates to self-assemble into nanoassemblies was tested by using various complementary techniques such as DLS, SEC, and SEC-MALS. These techniques revealed that the conjugates exhibited generation-dependent as well as protein-dependent self-assembly property. Importantly, the report offers control over hydrodynamic radius (10-18 nm), precise oligomeric state (18-44 mer), and the molecular weight of choice (480-1300 kDa) of the complex by either tuning dendron or protein head group in the molecular design.

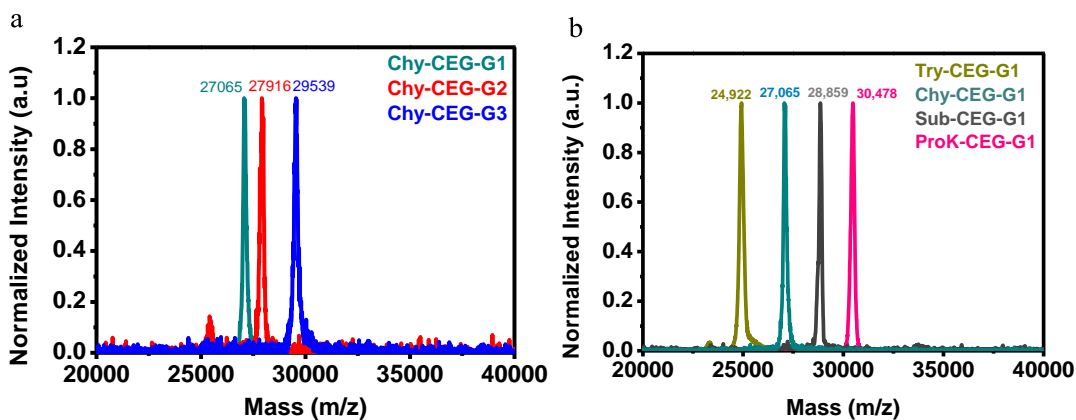
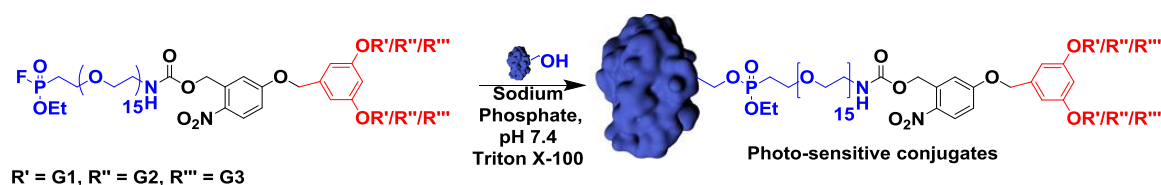


Figure 1.18 | MALDI-ToF characterization of protein-dendron bioconjugates. (a) Purified dendron variants and (b) protein variants.

1.4.3. Chapter 3: Programmed and Sequential Disassembly of Multi-Responsive Supramolecular Protein Assemblies

Although the conjugates mentioned above were capable of forming supramolecular protein assemblies driven through strong hydrophobic interaction, the designed protein complexes were static in nature. In order to surmount this limitation, in this chapter, we provide a rational and straightforward design for the construction of monodisperse photo-responsive protein-dendron assemblies. These photo-responsive protein dendron bioconjugates are mainly composed of has four core structural elements: (i) hydrophilic globular protein, (ii) flexible hydrophilic linker, (iii) photo-responsive group, and (iv) hydrophobic dendrons (Scheme 1.2).



Scheme 1.2 | Scheme for the synthesis of photo-sensitive protein-dendron bioconjugates.

Further, we investigated the self-assembling property of these designed amphiphilic macromolecular synthons. Interestingly, these conjugates also exhibited generation-dependent self-assembly behavior similar to the bioconjugates in chapter 2. The strong hydrophobic interaction of dendrimer is a major driving force for protein-dendron bioconjugates to self-assemble into the complex. Most importantly, the presence of a photo-responsive group did not affect the self-assemblies properties of these conjugates. We hypothesized that the selective clipping of the hydrophobic domain from the rest of the protein conjugate upon photo-irradiation should result in disassembly of the protein complex.

Then, the photo-induced dis-assembly studies were carried out by exposing the Chy-CEG-NB-G1/G2/G3 complexes (in 50 mM sodium phosphate pH7.4) to UV light ($\lambda =$

280 to 360 nm) for 60 mins. Surprisingly, we have observed only partial disassembly of these complexes upon photo-irradiation, as evident from SEC chromatogram (Figure 1.19 a-d).

This could be because, during the photolysis of nitrobenzyl group, these photo-sensitive complexes bearing nitrobenzyl group undergoes rapid rearrangement reactions to form imine functionalized bioconjugate. This imine based bioconjugate preserves the global amphiphilicity, thus prevents the further disassembly of complex (Figure 1.20).

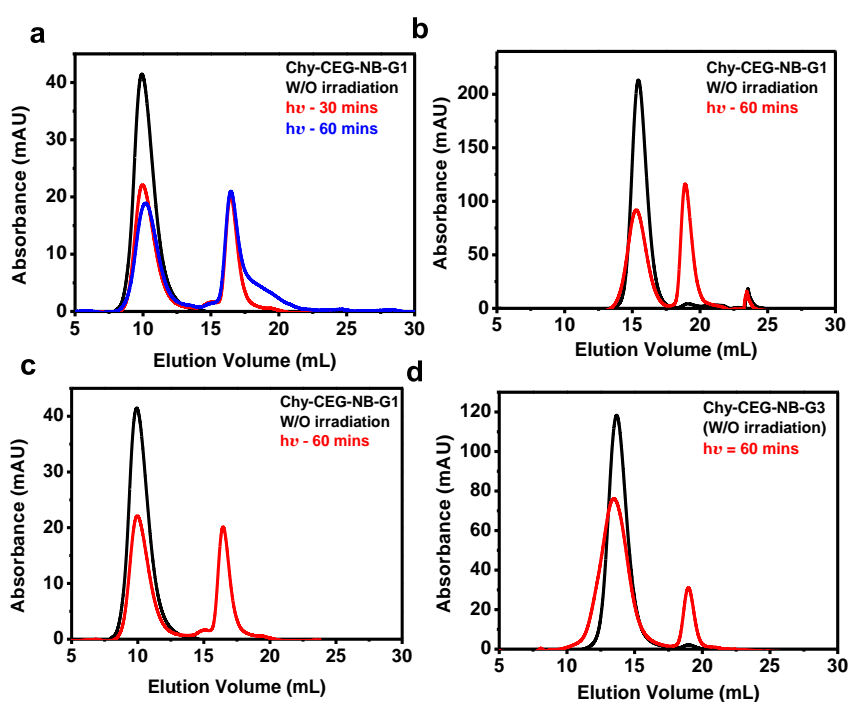


Figure 1.19 | Dis-assembly profile of protein-dendron complexes in 50 mM sodium phosphate pH 7.4 using SEC chromatograms. (a) Time-dependent dis-assembly studies of UV-exposed Chy-CEG-NB-G1(0.5 mg/mL) complex. **(b)** SEC chromatograms of Chy-CEG-NB-G1(2 mg/mL) **(c)** Chy-CEG-NB-G2 (0.8 mg/mL), and **(d)** Chy-CEG-NB-G3 (1.2 mg/mL), W/O UV-exposure (black) and after UV-exposure ($\lambda = 280$ to 360 nm) (red), respectively.

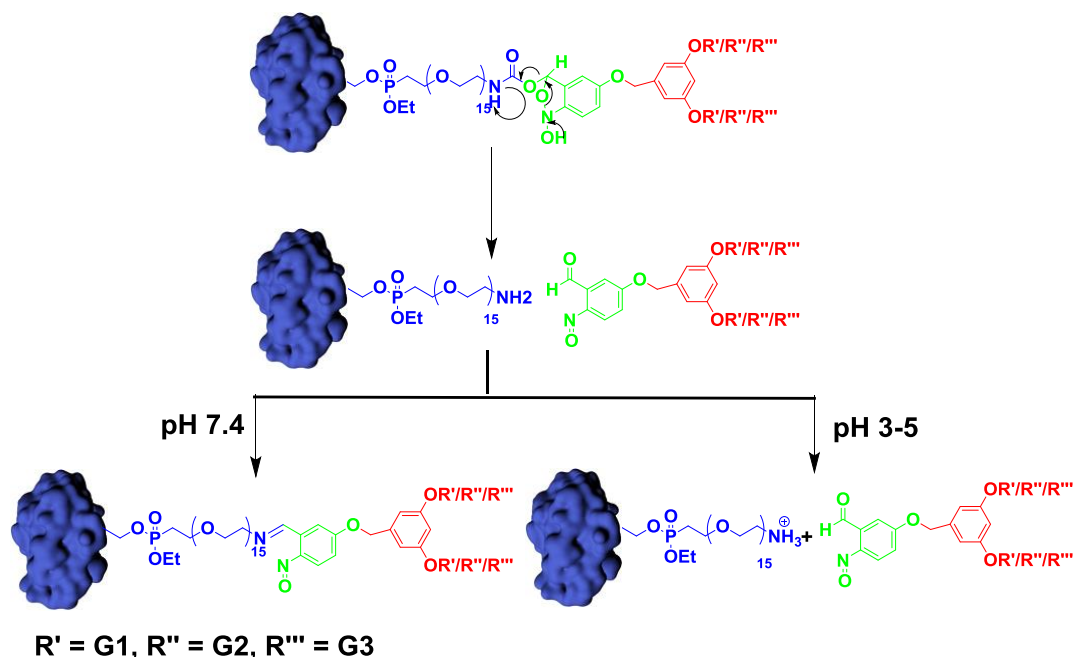


Figure 1.20 | Mechanism for photolysis and imine formation. Photolysis of the o-nitrobenzyl group would result in the formation of 2-nitroso benzaldehyde derivative and protein functionalized by amine-terminated oligoethylene glycol. The newly formed amine-functionalized protein reacts with 2-nitroso benzaldehyde derivative at pH 7.4 to form new protein-dendron conjugate bearing imine bond. The formation of the new complex is inhibited by arresting the nucleophilicity of the newly formed amine *via* protonating it at lower pH (pH 3).

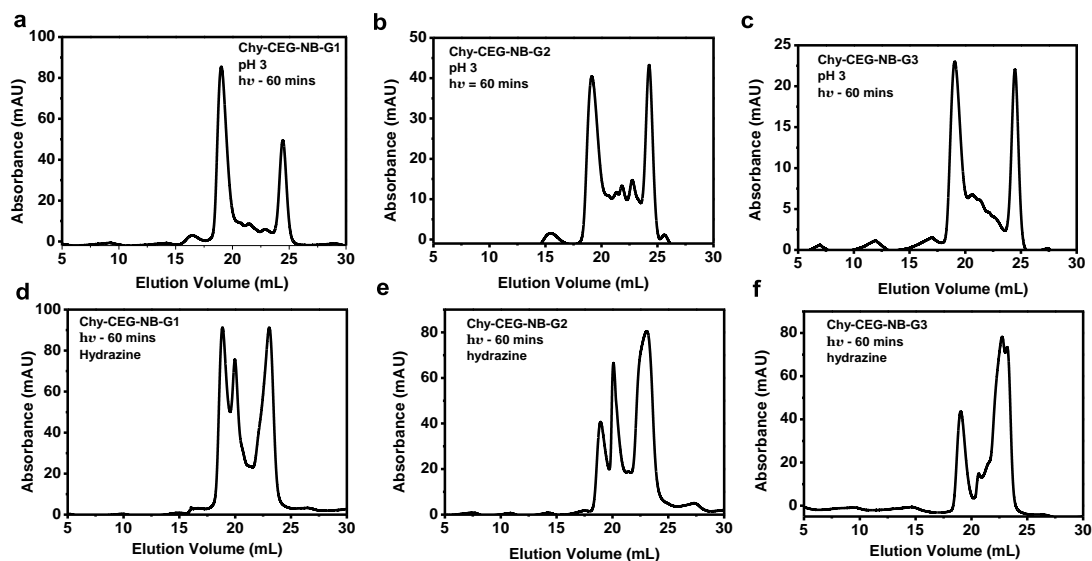
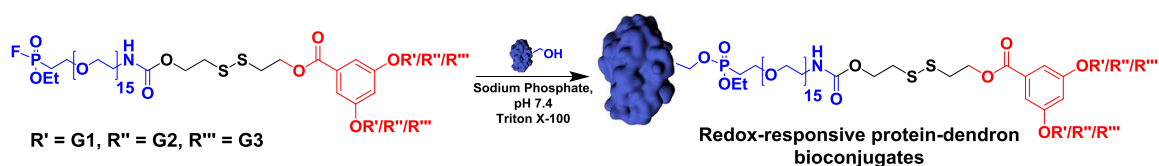


Figure 1.21 | Dis-assembly of protein-dendron complexes at pH 3 and pH 7.4 in the presence of hydrazine. SEC chromatograms of UV-exposed (a) Chy-CEG-NB-G1, (b) Chy-CEG-NB-G2, (c) Chy-CEG-NB-G3, at pH 3, respectively. SEC chromatograms of UV-exposed (d) Chy-CEG-NB-G1, (e) Chy-CEG-NB-G2, (f) Chy-CEG-NB-G3, at pH 7.4 in the presence of hydrazine (200 eq) respectively.

In order to test this hypothesis, we carried out disassembly studies at lower pH (which will block the amine to react with aldehyde) (Figure 1.20). Rewardingly, we have observed complete disassembly of the complex at pH 3 (Figure 1.21 a-c). Next, we envisioned performing disassembly studies in the presence of an excess of hydrazine, which would outcompete with the reaction of primary amine from lysine side chains to form hydrazone (Figure 1.20). Interestingly, we have also observed complete disassembly when irradiation studies were carried out in the presence of hydrazine (200 eq) at pH 7.4 (Figure 1.21 d-f).

1.4.4. Chapter 4: Programmed Disassembly of Redox-Responsive Supramolecular Protein Assemblies

The scope of this strategy was further extended to design redox-responsive protein-dendron complexes. The design principle is the same as previous design expect redox-sensitive moiety is clamped between the hydrophilic and hydrophobic portion of bioconjugate (Scheme 1.3). These complexes upon treatment with DTT lead to cleavage of a disulfide bond, which resulted into a separation of the hydrophobic domain from the rest of protein conjugate (hydrophilic globular domain), that triggered the dis-assembly of a protein complex into constitutive monomers, as the assembly is kept intact by strong hydrophobic interaction (Figure 1.22).



Scheme 1.3 | Scheme for the synthesis of redox-sensitive protein-dendron bioconjugates

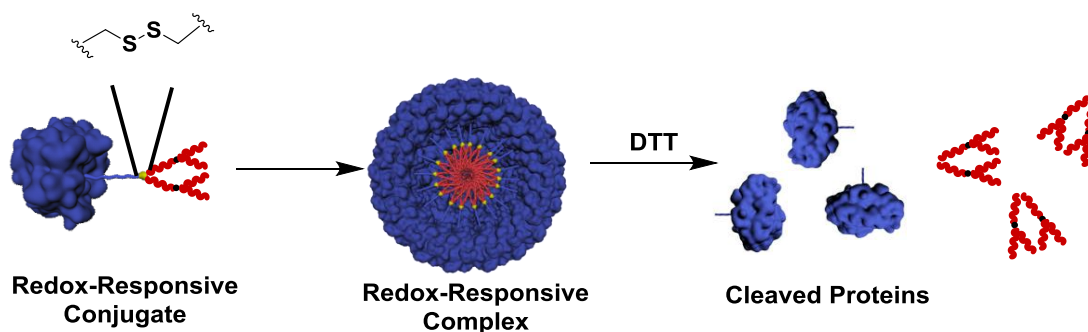


Figure 1.22 | Schematic representation of self-assembly and disassembly of redox-sensitive complexes.

In order to test this hypothesis, we performed dis-assembly studies first with Chy-CEG-SS-G1 in 50 mM sodium phosphate pH 7.4. The Chy-CEG-SS-G1 bioconjugate was treated with the different equivalent of DTT (10, 20, and 30 eq) for 12 hours. The DTT treated samples then subjected to SEC in order to investigate the disassembly profile. As expected, the sample before DTT treatment eluted at 16 mL, indicative of intact protein nanoassemblies. However, the SEC of DTT treated samples revealed that the entire complex is disassembled to its consecutive monomers, even with 10 eq of DTT (Figure 1.23a).

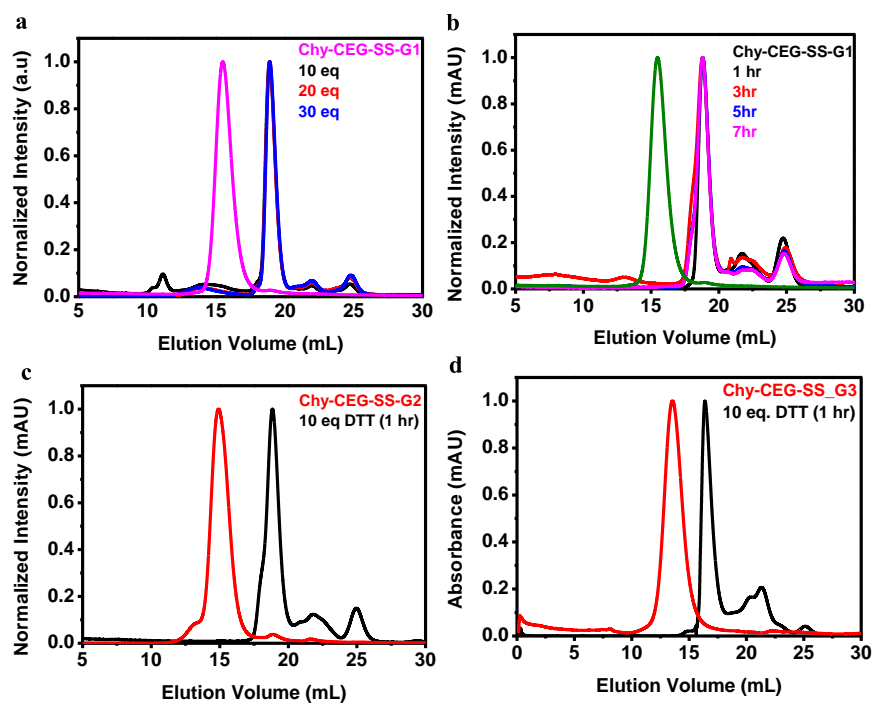


Figure 1.23 | Dis-assembly profile of redox-sensitive protein-dendron complexes in 50 mM sodium phosphate pH 7.4. (a), (b) Eq-dependent and time-dependent dis-assembly studies of DTT-treated Chy-CEG-SS-G1 complex, respectively. **(c)** dis-assembly profile of Chy-CEG-SS-G2, and **(d)** Chy-CEG-SS-G3, W/O DTT-treatment (red), and after DTT-treatment (black), respectively.

Further, the protein complex was incubated with 10 eq of DTT for different time points. The SEC studies reveal that in 60 mins entire complex disassembled as evident from the SEC profiles (Figure 1.23b). After establishing the time and eq of DTT required to achieve the complete dis-assembly, we sought to investigate the dis-assembly behavior of other complexes, i.e., Chy-CEG-SS-G2/G3. We were delighted to notice that Chy-CEG-G2/G3 also could be dis-assembled completely upon treatment with 10 eq. of DTT for 60 mins (Figure 1.23 c, d).

1.5. References

1. Mateu, M. G. *Arch. Biochem. Biophys.* **2013**, *531*, 65–79.
2. Stock, D.; Lelie, A. G.; Walker, J. E. *Science* **1999**, *286*, 1700–1705.
3. Massover, W. H. *Micron.* **1993**, *24*, 389–437.
4. Kerfeld, C. A.; Heinhorst, S.; Cannon, G. C. *Annu. Rev. Microbiol.* **2010**, *64*, 391–408.
5. Pieters, B. J. G. E.; van Eldijk, M. B.; Nolte, R. J. M.; Mecinović, J. *Chem. Soc. Rev.* **2016**, *45*, 24–39.
6. Arosio, P.; Levi, S. *Free Radical Biol Med.* **2002**, *33*, 457–463.
7. Bloom, L. B. (2009) Loading clamps for DNA replication and repair DNA Repair (Amsterdam) *8*, 570–578
8. Van den Ent, F.; Amos, L.; Lowe, J. *Curr. Opin. Microbiol.* **2001**, *4*, 634–638. (b) Rogers, S. L, Gelfand, V I. *Curr Opin Cell Biol.* **2000**, *12*, 57-62.
9. (a) Lin, X.; Xie, J.; Niu, G.; Zhang, F.; Gao, H.; Yang, M.; Quan, Q.; Aronova, M.A.; Zhang, G.; Lee, S.; Leapman, R. *Nano Lett.* **2011**, *11*, 814–819. (b) Chen, H.; Zhang, S.; Xu, C.; Zhao, G. *Chem. Commun.* **2016**, *52*, 7402–7405 (c) De la Escosura, A.; Nolte, R. J. M.; Cornelissen, J. J. L. M. *J. Mater. Chem.* **2009**, *19*, 2274–2278.
10. Marsh, J. A.; Teichmann, S. A. *Annu. Rev. Biochem.* **2015**, *84*, 551–575.
11. Desai, A.; Mitchison, T. J. *Annu. Rev. Cell. Dev. Biol.* **1997**, *13*(1), 83-117.
12. Edwardson, T. G. W.; Hilvert, D. *J. Am. Chem. Soc.* **2019**, *141*, 9432–9443.
13. Seeman, N. C. DNA in a material world. *Nature* **2003**, *421*, 427–431
14. Guo, P. X.; *Nat. Nanotechnol.* **2010**, *5*, 833–842.

15. Ulijn, R. V.; Smith, A. M. *Chem. Soc. Rev.* **2008**, *37*, 664–675.
16. Lee, L. A.; Niu, Z.; Wang, Q. *Nano Res.* **2009**, *2*, 349–364. (b) Ma, Y.; Nolte, R. J. M.; Cornelissen, J. J. L. M. *Adv. Drug Deliv. Rev.* **2012**, *64*, 811–825. (c) Witus, L. S.; Francis, M. B. *Acc. Chem. Res.* **2011**, *44*, 774–783. (d) Uchida, M.; Klem, M.T.; Allen, M.; Suci, P.; Flenniken, M.; Gillitzer, E.; Varpness, Z.; Liepold, L.O.; Young, M.; Douglas, T. *Adv. Mater.*, **2007**, *19*, 1025-1042.
17. Yeates, T. O. *Nat. Nanotechnol.* **2011**, *6*, 541–542 17.
18. King, N. P.; Sheffler, W.; Sawaya, M. R.; Vollmar, B.S.; Sumida, J. P.; André, I.; Gonen, T.; Yeates, T. O.; Baker, D. *Science* **2012**, *336*, 1171–1174.
19. Salgado, E. N.; Radford, R. J.; Tezcan, F. A. *Acc. Chem. Res.* **2010**, *43*, 661–672.
20. Uhlenheuer, D.A., Doctoral dissertation, **2011** PhD Thesis, Eindhoven University of Technology, Eindhoven, The Netherlands.
21. Reynhout, R. C.; Cornelissen, J. J. L. M.; Nolte, R. J. M. *Acc. Chem. Res.* **2009**, *42*, 681–692
22. Luo, Q.; Hou, C. X.; Bai, Y. S.; Wang, R. B.; Liu, J. Q. *Chem. Rev.* **2016**, *116*, 13571–13632
23. Crick, F. H. C.; Watson, J. D. *Nature* **1956**, *177*, 473–475.
24. Goodsell, D. S.; Olson, A. J. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 105–153.
25. Padilla, J. E.; Colovos, C.; Yeates, T. O. *Proc. Natl Acad. Sci. USA.* **2001**, *98*, 2217–2221.
26. Lai, Y. T.; Cascio, D.; Yeates, T. O. *Science* **2012**, *336*, 1129-1129.
27. Lai, Y.-T. et al. *Nat. Chem.* **2014**, *6*, 1065–1071.
28. Sinclair, J. C.; Davies, K. M.; Vénien-Bryan, C.; Noble, M. E. M. *Nat. Nanotechnol.* **2011**, *6*, 558–562
29. Leaver-Fay, A.; Tyka, M.; Lewis, S. M.; Lange, O. F.; Thompson, J.; Jacak, R.; Kaufman, K. W.; Renfrew, P. D.; Smith, C. A.; Sheffler, W.; Davis, I. W. *Methods Enzymol.* **2011**, *487*, 545–574.
30. Grueninger, D.; Treiber, N.; Ziegler, M.O.; Koetter, J.W.; Schulze, M.S.; Schulz, G.E. *Science* **2008**, *319*, 206–209.
31. Brodin, J. D.; Ambroggio, X. I.; Tang, C.; Parent, K. N.; Baker, T. S.; Tezcan, F. A. *Nat. Chem.* **2012**, *4*, 375–382
32. (a) King, N. P.; Bale, J. B.; Sheffler, W.; McNamara, D. E.; Gonen, S.; Gonen, T.; Yeates, T.O.; Baker, D. *Nature* **2014**, *510*, 103-108. (b) Bale, J. B.; Gonen, S.; Liu, Y.; Sheffler, W.; Ellis, D.; Thomas, C., Cascio, D.; Yeates, T.O.; Gonen, T.; King, N. P.; Baker, D. *Science* **2016**, *353*, 389-394.

33. Lai, Y.-T.; et al. *Nat. Chem.* **2014**, *6*, 1065–1071.
34. Gonen, S.; DiMaio, F.; Gonen, T.; Baker, D. *Science*, **2015**, *348*, 1365-1368.
35. Guryanov, I.; Fiorucci, S.; Tennikova, T. *Materials Science and Engineering: C*, **2016**, *68*, 890-903.
36. Chakraborti, A. S. *Mol Cell Biochem*, **2003**, *253*, 49-54.
37. Oohora, K.; Onoda, A.; Hayashi, T. *Chem. Commun.* **2012**, *48*, 11714.
38. Kitagishi, H.; Oohora, K.; Yamaguchi, H.; Sato, H.; Matsuo, T.; Harada, A.; Hayashi, T. *J. Am. Chem. Soc.* **2007**, *129* (34)
39. 39. Kitagishi, H.; Kakikura, Y.; Yamaguchi, H.; Oohora, K.; Harada, A.; Hayashi, T. *Angew. Chem., Int. Ed.* **2009**, *48*, 1271–1274.
40. Oohora, K.; Onuma, Y.; Tanaka, Y.; Onoda, A.; Hayashi, T. *Chem. Commun.* **2017**, *53*, 6879– 6882,
41. Chilkoti, A.; Stayton, P. S. *J. Am. Chem. Soc.*, **1995**, *117*, 10622-10628.
42. Burazerovic, S.; Gradinaru, J.; Pierron, J.; Ward, T. R. *Angew. Chem., Int. Ed.* **2007**, *46*, 5510– 5514
43. Oohora, K.; Burazerovic, S.; Onoda, A.; Wilson, Y. M.; Ward, T. R.; Hayashi, T. *Angew. Chem., Int. Ed.* **2012**, *51*, 3818– 3821,
44. Cecioni, S.; Imberty, A.; Vidal, S. *Chem. Rev.* **2015**, *115*, 525–561.
45. Sakai, F.; Yang, G.; Weiss, M. S.; Liu, Y.; Chen, G.; Jiang, M. *Nat. Commun.* **2014**, *5*, 4634
46. Yang, G.; Zhang, X.; Kochovski, Z.; Zhang, Y.; Dai, B.; Sakai, F.; Jiang, L.; Lu, Y.; Ballauff, M.; Li, X. *J. Am. Chem. Soc.* **2016**, *138*, 1932– 1937.
47. Carlson, J. C. T.; Jena, S. S.; Flenniken, M.; Chou, T.-F.; Siegel, R. A.; Wagner, C. R. *J. Am. Chem. Soc.*, **2006**, *128*, 7630–7638.
48. Chau, T. F.; et al. *ACS Nano* **2008**, *2*, 2519–2525.
49. Li, Q.; So, C. R.; Fegan, A.; Cody, V.; Sarikaya, M.; Vallera, D. A.; Wagner, C. R. *J. Am. Chem. Soc.* **2010**, *132*, 17247– 17257
50. Modica, J. A.; Lin, Y.; Mrksich, M. Synthesis of Cyclic Megamolecules. *J. Am. Chem. Soc.* **2018**, *140* (20), 6391– 6399.
51. Andreini, C.; Bertini, I.; Cavallaro, G.; Holliday, G. L.; Thornton, J. M. *J. Biol. Inorg. Chem.* **2008**, *13*, 1205–1218.
52. Valdez, C. E.; Smith, Q. A.; Nechay, M. R.; Alexandrova, A. N. *Acc. Chem. Res.* **2014**, *47*, 3110–3117.
53. Salgado, E. N.; Radford, R. J.; Tezcan, F. A. *Acc. Chem. Res.* **2010**, *43*, 661–672.

54. Faraone-Mennella, J.; Tezcan, F. A.; Gray, H. B.; Winkler, J. R. *Biochemistry* **2006**, *45*, 10504–10511
55. Salgado, E. N.; Faraone-Mennella, J.; Tezcan, F. A. *J. Am. Chem. Soc.* **2007**, *129*, 13374–13375.
56. Biswas, S.; et al. *J. Am. Chem. Soc.* **2009**, *131*, 7556–7557.
57. Sendai, T.; Biswas, S.; Aida, T. *J. Am. Chem. Soc.* **2013**, *135*, 11509–11512.
58. (a) Zhang, W.; Luo, Q.; Miao, L.; Hou, C.; Bai, Y.; Dong, Z.; Xu, J.; Liu, J. *Nanoscale* **2012**, *4*, 5847–5851. (b) Bai, Y.; Luo, Q.; Zhang, W.; Miao, L.; Xu, J.; Li, H.; Liu, J. *J. Am. Chem. Soc.* **2013**, *135*, 10966–10969.
59. (a) Abuchowski, A.; van Es, T.; Palczuk, N. C.; Davis, F. F. *J. Biol. Chem.* **1977**, *252*, 3578–3581. (b) Abuchowski, A.; McCoy, J. R.; Palczuk, N. C.; van Es, T.; Davis, F. F. *J. Biol. Chem.* **1977**, *252*, 3582–3586. (c) Pelegri-O'Day, E. M., Lin, E. W. & Maynard, H. D. *J. Am. Chem. Soc.* **2014**, *136*, 14323–14332.
60. Heredia, K.L.; Bontempo, D.; Ly, T.; Byers, J.T.; Halstenberg, S.; Maynard, H.D. *J Am Chem Soc*, **2005**, *127*, 16955-16960
61. Hannink, J. M.; Cornelissen, J. J. L. M.; Farrera, J. A.; Foubert, P.; De Schryver, F. C.; Sommerdijk, N. A. J. M.; Nolte, R. J. M. *Angew. Chem., Int. Ed.* **2001**, *40*, 4732–4734.
62. Kulkarni, S.; Schilli, C.; Müller, A.; Hoffman, A.; Stayton, P. *Bioconjugate Chem.* **2004**, *15*, 747–573.
63. (a) Boerakker, M. J.; Hannink, J. M.; Bomans, P. H. H.; Frederik, P. M.; Nolte, R. J. M.; Meijer, E. M.; Sommerdijk, N. A. J. M. *Angew. Chem., Int. Ed.* **2002**, *41*, 4239 (b) Boerakker, M. J.; Botterhuis, N. E.; Bomans, P. H. H.; Frederik, P. M.; Meijer, E. M.; Nolte, R. J. M.; Sommerdijk, N. *Chem.—Eur. J.* **2006**, *12*, 6071–6080
64. Reinhout, I. C.; Cornelissen, J.; Nolte, R. J. M. *J. Am. Chem. Soc.* **2007**, *129*, 2327–2332.
65. Wan, X.; Liu, S. *Macromol. Rapid Commun.* **2010**, *31*, 2070–2076.
66. Velonia, K.; Thordarson, P.; Andres, P. R.; Schubert, U. S.; Rowan, A. E.; Nolte, R. J. M. *Polym. Prepr.* **2003**, *44*, 648–649
67. Stephanopoulos, N.; Francis, M. B. *Nature Chem. Biol.* **2011**, *7*, 876–884.
68. Velonia, K.; Rowan, A.; Nolte, R. *J. Am. Chem. Soc.* **2002**, *124*, 4224–4225.
69. (a) Chen, G.; Hoffman, A. S. *Bioconjugate Chem.* **1993**, *4*, 509-514. (b) Park, W. M.; Champion, J. A. *J. Am. Chem. Soc.* **2014**, *136*, 17906-17909. (c) Hassouneh, W.; Fischer, K.; MacEwan, S.R.; Branscheid, R.; Fu, C.L.; Liu, R.; Schmidt, M.; Chilkoti, A. *Biomacromolecules* **2012**, *13*, 15981605. (d) Dreher, M. R.; et al. *J. Am. Chem. Soc.* **2008**, *130*, 687–694.

70. (a) Thomas, C. S., Glassman, M. J., & Olsen, B. D. *ACS Nano* **2011**, *5*, 5697–5707. (b) Lam, C. N.; Olsen, B. D. *Soft Matter* **2013**, *9*, 2393–2402. (c) Huang, A.; Qin, G.; Olsen, B. D. *ACS Appl. Mater. Interfaces* **2015**, *7*(27), 14660–14669.
71. Lam, C. N.; Kim, M.; Thomas, C. S.; Chan, D.; Sanoja, G. E.; Okwara, C. U.; Olsen, B. D. *Biomacromolecules* **2014**, *15*, 1248.
72. Lam, C. N.; Yao, H.; Olsen, B. D.; *Biomacromolecules*, **2016**, *17*, 2820–2829
73. Huang, A.; Olsen, B. D. *Macromol. Rapid Commun.* **2016**, *37*, 1268–1274,
74. Qin, G.; Glassman, M. J.; Lam, C. N.; Chang, D.; Schaible, E.; Hexemer, A.; Olsen, B. D. *Adv. Funct. Mater.* **2015**, *25*, 729–738
75. Averick, S.; Karácsóny, O.; Mohin, J.; Yong, X.; Moellers, N. M.; Woodman, B. F.; Matyjaszewski, K. *Angew. Chem., Int. Ed.* **2014**, *53*, 8050–8055.
76. Li, X.; Liu, C. C. *ChemBioChem.* **2014**, *15*, 2335–2341.
77. Sandanaraj, B. S.; Reddy, M. M.; Bhandari, P. J.; Kumar, S. and Aswal, V. K. *Chem. Eur. J.* **2018**, *24*, 16085-16096.
78. (a) Fréchet, J. M. J.; Tomalia, D. A. *Dendrimers and Other Dendritic Polymers*, New York: Wiley, 2001. (b) Grayson, S. M. & Fréchet, J. M. J. *Chem. Rev.* **2001**, *101*, 3819–3868. (c) Bosman, A. W., Janssen, H. M. & Meijer, E. W. *Chem. Rev.* **1999**, *99*, 1665–1688.
79. Tomalia, D.A.; Baker, H.; Dewald, J.; Hall, M.; Kallos, G.; Martin, S.; Roeck, J.; Ryder, J.; Smith, P. *Polym. J.*, **1985**, *17*, 117-132
80. Newkome, G. R., Yao, Z., Baker, G. R., & Gupta, V. K. *J. Org. Chem.*, **1985**, *50*, 2003-2004
81. Hawker, C.J.; Fréchet, J.M.J. *J. Am. Chem. Soc.*, **1990**, *112*, 7638-7647
82. (a) Gupta, S.; Pfeil, J.; Kumar, S.; Poulsen, C.; Lauer, U.; Hamann, A.; Hoffmann, U.; Haag, R. *Bioconjugate Chem.* **2015**, *26*, 669–679. (b) Barkley, D. A.; Rokhlenko, Y.; Marine, J. E.; David, R.; Sahoo, D.; Watson, M. D.; Koga, T.; Osuji, C. O.; Rudick, J. G. *J. Am. Chem. Soc.* **2017**, *139*, 15977–15983
83. Dong, Y.; Yang, Y.R.; Zhang, Y.; Wang, D.; Wei, X.; Banerjee, S.; Liu, Y.; Yang, Z.; Yan, H.; Liu, D. *Angew. Chem. Int. Ed.* **2017**, *56*, 1586–1589.
84. (a) Percec, V.; Leowanawat, P.; Sun, H.-J.; Kulikov, O.; Nusbaum, C. D.; Tran, T. M.; Bertin, A.; Wilson, D. A.; Peterca, M.; Zhang, S.; Kamat, N. P.; Vargo, K.; Moock, D.; Johnston, E. D.; Hammer, D. A.; Pochan, D. J.; Chen, Y.; Chabre, Y. M.; Shiao, T. C.; Bergeron-Brlek, M.; André, S.; Roy, R.; Gabius, H.-J.; Heiney, P. A. *J. Am. Chem. Soc.* **2013**, *135*, 9055-9077. (b) Jayaraman, N.; Nepogodiev, S. A.; Stoddart, J. F. *Chem. Eur. J.* **1997**, *3*, 1193–1199.

85. Wang, X.; Inapagolla, R.; Kannan, S.; Lieh-Lai, M.; Kannan, R. M. *Bioconjugate Chem.* **2007**, *18*, 791–799
86. Thomas, T. P.; Patri, A. K.; Myc, A.; Myaing, M. T.; Ye, J. Y.; Norris, T. B.; Baker, J. R., Jr. *Biomacromolecules* **2004**, *5*, 2269–2274.
87. Rendle, P. M.; Seger, A.; Rodrigues, J.; Oldham, N. J.; Bott, R. R.; Jones, J. B.; Cowan, M. M.; Davis, B. G. *J. Am. Chem. Soc.* **2004**, *126*, 4750–4751.
88. Sato, M.; Furuike, T.; Sadamoto, R.; Fujitani, N.; Nakahara, T.; Niikura, K.; Monde, K.; Kondo, H.; Nishimura, S. I. *J. Am. Chem. Soc.* **2004**, *126*, 14013–14022.
89. (a) Kostianen, M. A.; Szilvay, G. R.; Smith, D. K.; Linder, M. B.; Ikkala, O. *Angew. Chem., Int. Ed.* **2006**, *45*, 3538–3542. (b) Kostianen, M. A.; Szilvay, G. R.; Lehtinen, J.; Smith, D. K.; Linder, M. B.; Urtti, A.; Ikkala, O. *ACS Nano* **2007**, *1*, 103–113.

CHAPTER 2

DESIGN, SYNTHESIS AND SELF-ASSEMBLY STUDIES OF SUITE OF MONODISPERSE FACIALLY AMPHIPHILIC PROTEIN-DENDRON CONJUGATES

Part of this work has been published:

B.S. Sandanaraj, **P.J. Bhandari**, M.M. Reddy, A. B. Lohote, B. Sahoo. *ChemBioChem*,
2020, 21, 408-416.

2.1. Introduction

Natural biomolecular assemblies such as viruses,¹ molecular motors,² protein cages,³ and bacterial compartments⁴ are well defined, monodisperse, dynamic architectures that exhibit higher-order functions. These natural biomolecular assemblies serve as an inspiration for the custom design of synthetic supramolecular assemblies using nucleic acids,^{5,6} peptides,⁷ and proteins⁸ as building blocks. Although the majority of naturally occurring biomolecular assemblies are built using proteins as building blocks, biomimicry of this phenomenon is highly challenging. At the same time, nucleic acid^{5,6} and peptides⁷ are routinely used as building blocks for this purpose. However, recent developments in computational protein design and rational approaches using genetic methods have made substantial progress in this field.⁹⁻¹² Although useful, these methods require sophisticated infrastructure, associated with a high failure rate and sometimes yield structures not intended by design¹³. In parallel, chemical strategies are also emerging as complementary methods for the design of protein with improved functions.^{14,15}

Pioneering work by Nolte and others have demonstrated novel chemical strategies for attachment of hydrophobic synthetic polymers on to various proteins exploiting both covalent and non-covalent chemistry.¹⁶⁻²⁰ These amphiphilic bioconjugates have shown to self-assemble into various nanoscopic structures driven *via* hydrophobic interactions. As a complementary method, polymers that exhibit lower-critical solution temperature (LCST) property were used to construct protein-polymer bioconjugates through the biosynthetic and/or chemical methodology. Below LCST, both the protein and polymer domains are hydrophilic, however, above LCST, polymer domain becomes hydrophobic and thereby induce global amphiphilicity, which imparts self-assembling property.²¹⁻²⁸

Among the various bioconjugation methods available to make protein-polymer bioconjugates, chemical modification of cysteine residue of a protein is a popular and straightforward method.^{17, 26-28} However, this method has a little scope because it is applicable to a few proteins having a single cysteine residue on their surface. Genetic incorporation of unnatural amino acids containing bioorthogonal functional groups would expand the scope of site-specific bioconjugation chemistry to a large number of proteins.²⁹

However, the poor efficiency of unnatural amino acid incorporation on to protein and the tedious process associated with this methodology limits the scope of this technology.³⁰ In addition, most of the studies reported so far exclusively use linear polymers to construct protein-polymer bioconjugates. A significant disadvantage to their use is that they induce a degree of heterogeneity in the form of polydispersity. A simple approach to synthesize protein-polymer bioconjugates with a polydispersity index (PDI) of 1 has been demonstrated by site-specific attachment of hydrophilic monodisperse dendron to hydrophilic protein.³¹⁻³⁶ This approach is relatively easy as both dendrimer and protein are soluble in an aqueous medium, and therefore bioconjugation reaction is straight-forward. However, these bioconjugates lack self-assembling properties as expected because of the lack of amphiphilic character.

In this regard, the synthesis of facially amphiphilic biohybrid macromolecules containing both “hydrophilic protein” and “hydrophobic dendron” domains would be an excellent scaffold. The synthetic methodology would provide access to perfectly monodisperse protein-based semi-synthetic macromolecules similar to artificial amphiphilic protein-polymers made through the biosynthetic pathway.^{24,25} In addition, chemical diversity of the second domain is not limited in this approach as opposed to polymers made through the biosynthetic pathway.^{24,25} However, the development of a new chemical methodology to accomplish this goal is a very challenging task for the following reasons; (i) developing site-specific bioconjugation chemistry that can target a wide variety of protein is not easy, because proteins are hetero-polymers made up of twenty amino acids having different functional groups, (ii) considering the hydrophobic nature of dendron, performing bioconjugation reaction in a complete aqueous medium is highly challenging and, (iii) most importantly, purification and analytical characterization of protein-dendron bioconjugate is very difficult as the resultant products are macromolecular/facially amphiphilic in nature and resembles membrane proteins in terms of hydrophathy index.

Our group has initiated a research program to develop chemical methodologies for the design of facially amphiphilic semi-synthetic proteins. In our previous work, we have reported a chemical methodology for the synthesis of novel protein amphiphiles utilizing a micelle-assisted protein labeling technology³⁷. This work laid the foundation for the

synthesis of monodisperse protein amphiphiles. However, the reported methodology was only explored for the bioconjugation of small hydrophobic linear molecules (<1000 Da). In addition, the synthesized protein nanoassemblies were relatively small in size (8-14 nm), having low molecular mass (220 kDa – 660 kDa).³⁷ In order to push the boundaries of our chemical method. Herein, we disclose a novel method for chemical synthesis of a suite of monodisperse protein-dendron conjugates and studied their self-assembling behavior in detail.

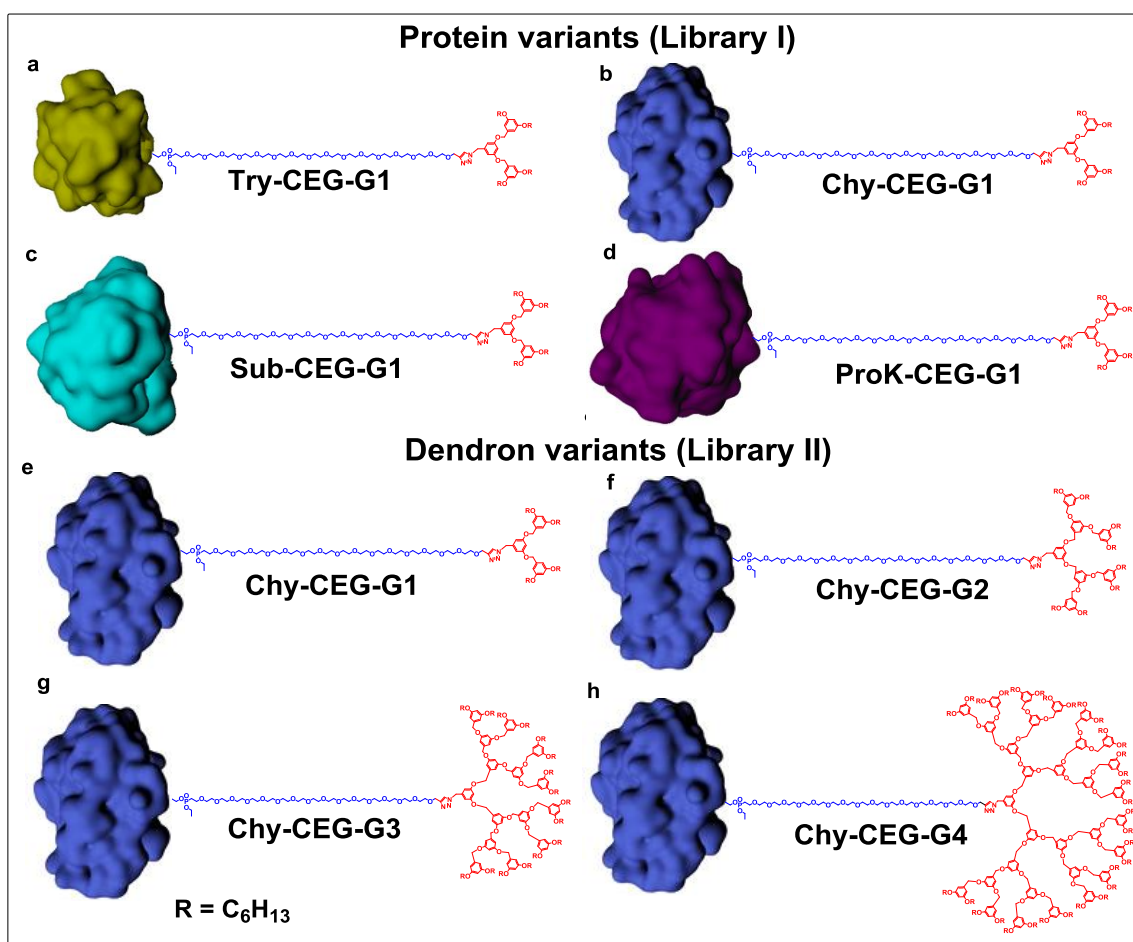


Figure 2.1 | Chemical structures of protein-dendron bioconjugates composed of globular water-soluble protein (green = trypsin, blue = chymotrypsin, cyan = subtilisin, magenta = proteinase K), hydrophilic CEG (blue), and hydrophobic dendron block (red). Structures of (a) Try-CEG-G1 (b) Chy-CEG-G1 (c) Sub-CEG-G1 (d) ProK-CEG-G1 (e) Chy-CEG-G1 (f) Chy-CEG-G2 (g) Chy-CEG-G3 (h) Chy-CEG-G4.

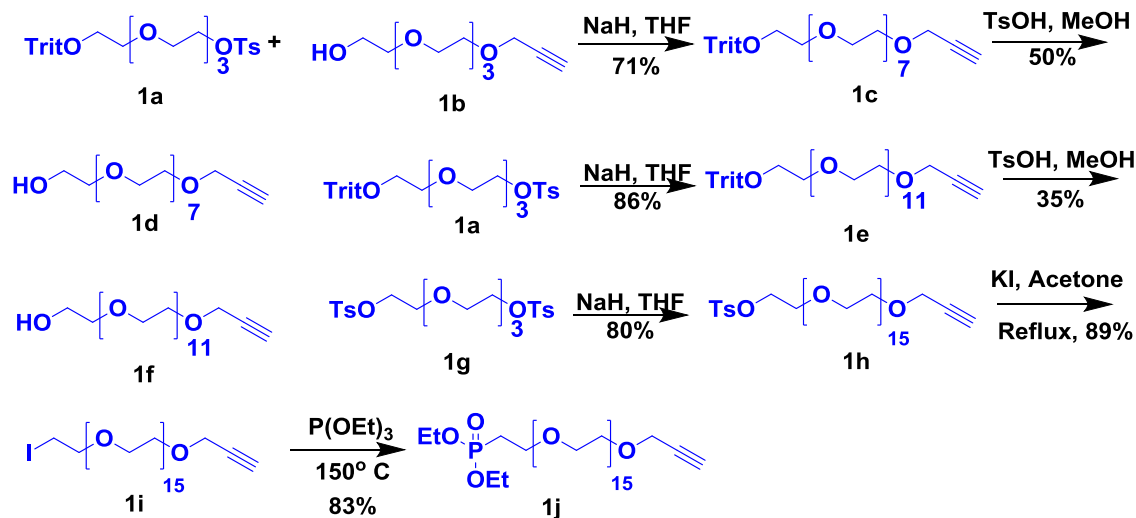
2.2. Result and Discussion

2.2.1. Synthesis of Monodisperse Hydrophilic Linker

From our previous experience, we knew that the direct site-specific attachment of benzyl-ether dendron to a hydrophilic protein is difficult to achieve because of the steric hindrance and orthogonal solubility properties of hydrophilic protein and hydrophobic dendron (Figure 2.1). To overcome this limitation, we hypothesized that bioconjugation of an amphiphilic macromolecular probe containing a hydrophilic monodisperse oligoethylene glycol equipped with the protein reactive functional group covalently attached to hydrophobic dendron block would be a viable approach. To do that, we chose hexadeca(ethylene glycol)/cetylene glycol (CEG) as the linker molecule for the following reasons; (i) it is sufficiently long, flexible and therefore circumvents steric hindrance problem that may arise during bioconjugation reaction of higher-generation dendrons, (ii) it is hydrophilic and has protein-repellent property, thus would not denature the protein.

To synthesize the monodisperse alkyne-terminated diphosphonate derivative of CEG, a strategy by French *et al.* was adopted with a slight modification (Scheme 1.1).³⁸ The multi-step synthesis of the target molecule (1j) were carried out in the following way. Tosyl derivative of tetraethylene glycol (TEG) (1a) on reaction with monopropargyl TEG (1b) using sodium hydride (NaH) in tetrahydrofuran (THF) gave trityl protected octaethylene glycol (1c). The compound 1c on trityl deprotection using p-toluenesulfonic acid (TsOH) in methanol (MeOH) produced alcohol 1d; which further on treatment with compound 1a in the presence of NaH in THF yielded trityl protected dodecaethylene glycol (1e) using standard condition. Subsequently, the trityl deprotection of compound 1e, followed by treatment with ditosyl TEG (1g), gave compound 1h. Further, the obtained tosyl derivative (1h) on reflux in acetone in the presence of potassium iodide afforded corresponding iodide derivative (1i). Finally, diphosphonate ester (1j) was synthesized by refluxing iodide compound (1i) with triethyl phosphite (Scheme 1.1).

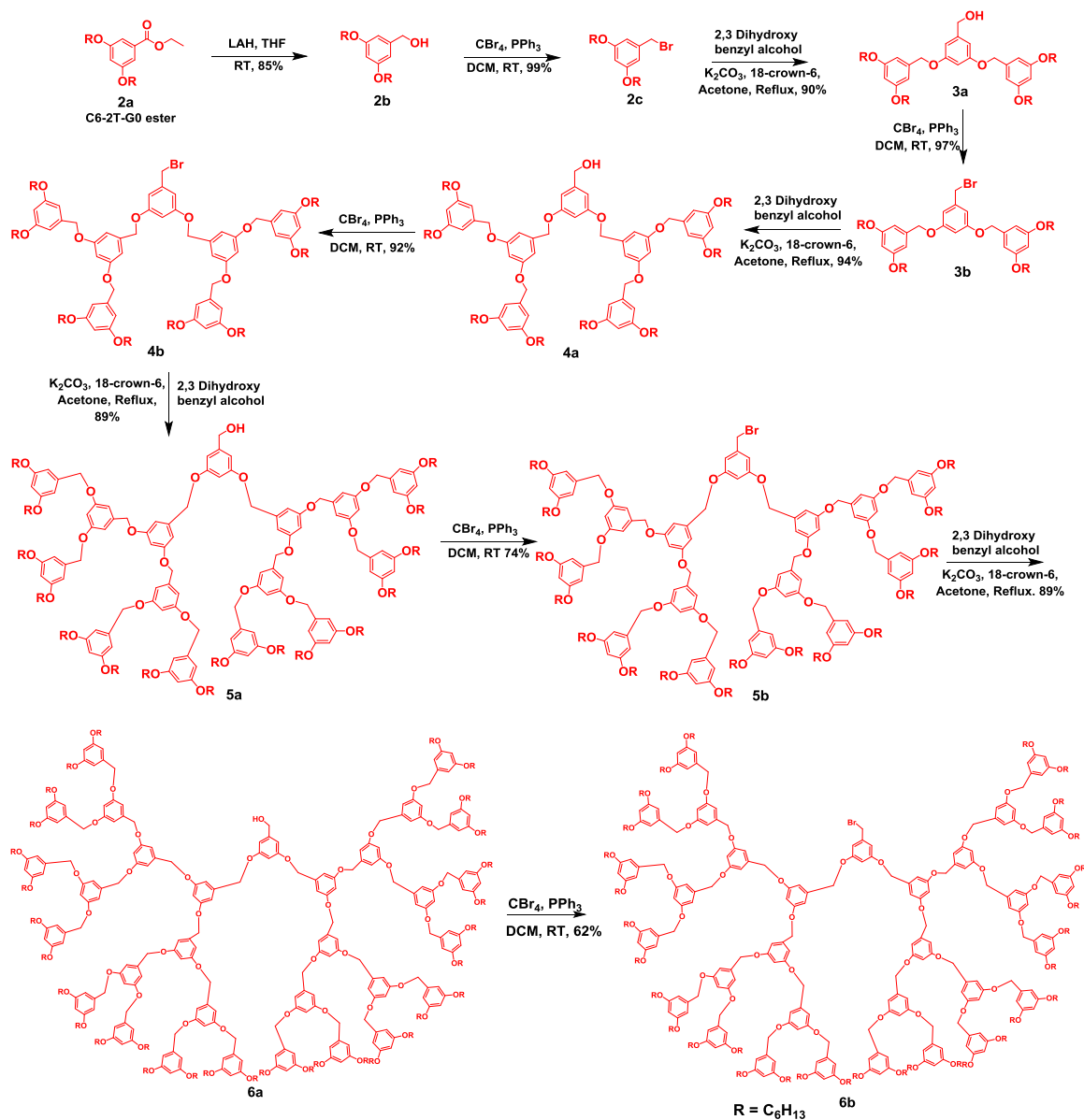
The diphosphonate ester of monodisperse CEG containing sixteen repeating units of ethylene glycol with terminal alkyne group (1j) was synthesized in a total of 12 steps with an overall yield of 2.4% (Scheme 1.1).



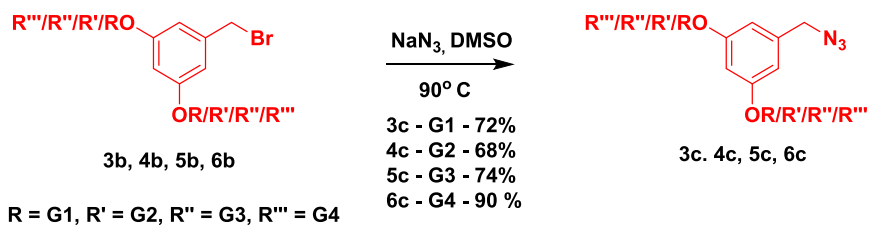
Scheme 1.1 | Synthesis of monodisperse diphosphonate esters of CEG containing terminal alkyne functional group.

2.2.2. Synthesis of macromolecular Amphiphilic Activity-based Probes (AABPs)

Next, we synthesized the benzyl-ether dendrons (G1-G4) with an azide at the focal point by using convergent dendrimer synthesis (Scheme 1.2, 1.3).^{40,41} We choose Fréchet type benzyl-ether dendron as the hydrophobic domain for the following reasons; (i) monodisperse benzyl-ether dendrons having high molecular weight can be easily synthesized, (ii) most importantly, dendrimer of higher-generation adopts the 3-dimensional architecture, and therefore would have pronounced effect on self-assembling behavior.^{40,41}



Scheme 1.2. Scheme for Synthesis of G1-G4 generation bromide.



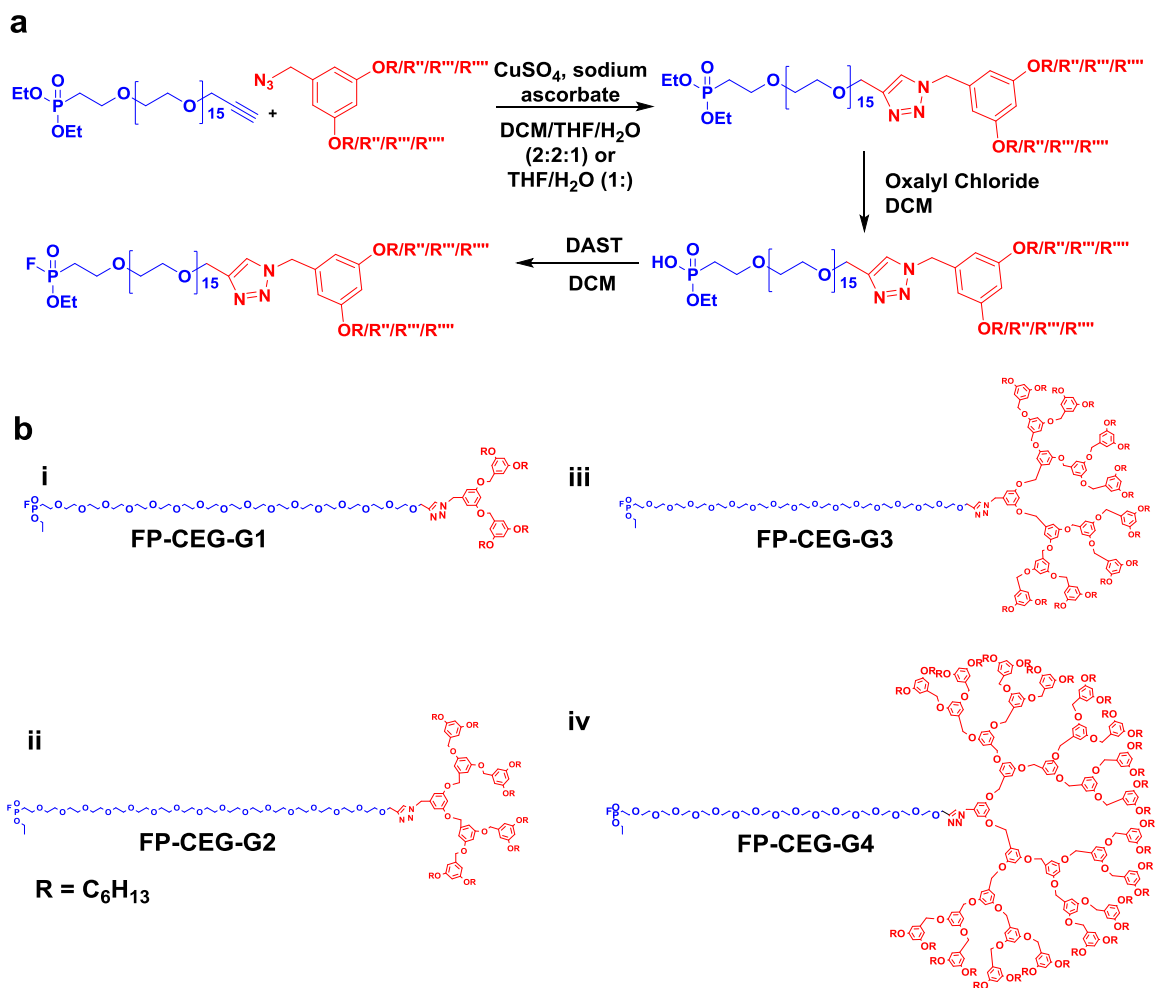
Scheme 1.3. Scheme for the synthesis of G1-G4 generation azide.

Further, hydrophobic G1- and G2-dendritic azide were coupled to alkyne-terminated CEG derivative using click chemistry strategy in the presence of THF/H₂O (1:1)

mixture (Scheme 1.4a).⁴² However, when we attempted the same reaction condition for the coupling of G3- and G4-azides, those reactions did not work. We suspect the extreme hydrophobic character G3- and G4-azide may be the reason for the failure. To mitigate this problem, the polarity of the solvent system was decreased by adding dichloromethane (DCM), and the reaction was attempted again in the new solvent system DCM/THF/H₂O. We were delighted to see that the reaction went smoothly, as evident from MALDI-ToF results. However, the reaction was not quantitative, and some unreacted compound (1j) can be seen in MALDI-ToF results. To separate the unreacted compound 1j, we attempted a normal phase thin layer chromatography (TLC) experiment. However, R_f value difference between the product and the linker molecule same in various solvent combinations. Discouraged by the above results, we again attempted a TLC experiment in reverse phase conditions. This time appreciable R_f difference was observed for all the clicked products due to their amphiphilic nature. Hence, the purification of all the clicked products was achieved using a reverse phase chromatography method. The resultant pure diphosphonate esters were then deprotected using an oxalyl chloride followed by treatment with a diethylaminosulfur trifluoride (DAST) reagent to make macromolecular AABPs with a fluorophosphonate (FP) as a protein reactive group (Scheme 1.4a-b).

Synthesis of monodisperse macromolecular AABPs containing CEG coupled of dendron of a different generation (FP-CEG-G1, FP-CEG-G2, FP-CEG-G3, and FP-CEG-G4 – simply represented as FP-CEG/G1/G2/G3/G4) were achieved in 8, 10, 12, and 14, steps with an overall yield of 1.4%, 1.5%, 1.3%, and 0.8%, respectively.

We chose the activity-based protein labeling method for bioconjugation because this method provides remarkable advantages over other bioconjugation methods such as fast reaction kinetics, mild reaction conditions, chemo- and site-selective modification of the active-site of protein.³⁹ All these features assist in preserving the native structure of protein during bioconjugation reaction and provides an opportunity to design monodisperse protein conjugate with little structural perturbation. Specifically, FP was chosen as a protein reactive functional group because it can target a wide range of serine proteases³⁹; therefore, it provides an opportunity to study the effect of protein on the self-assembly of protein-dendron bioconjugates.

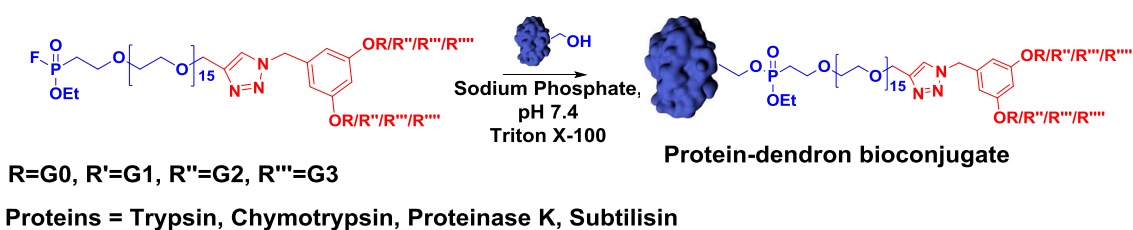


Scheme 1.4 | Synthetic scheme and chemical structures of macromolecular AABP (a) Scheme for the synthesis of macromolecular AABPs. (b) Chemical structures of macromolecular AABPs is composed of monodisperse hydrophilic CEG (blue), and hydrophobic dendrons block (red). Chemical structures of (i) FP-CEG-G1, (ii) FP-CEG-G2, (iii) FP-CEG-G3 and, (iv) FP-CEG-G4.

2.2.3. Micelle-Assisted Protein Labeling (MAPLab) Technology

Considering the amphiphilic nature of macromolecular AABPs, i.e., (FP-CEG-G1/G2/G3/G4), we first attempted the bioconjugation reaction in neat phosphate buffer pH 7.4. Unfortunately, none of the macromolecular AABPs was soluble. Use of water-miscible organic solvent is an alternate approach; however, we consciously decided not to use any organic solvent as it might denature the protein structure and also pose a problem during down-stream purification. Alternatively, we adopted the micelle-assisted bioconjugation method reported by our group.³⁷ Remarkably, all of the macromolecular AABPs are soluble

in the presence of 10X critical micelle concentration (CMC) of triton X-100, including the most hydrophobic probe FP-CEG-G4, presumably due to the formation of a mixed-micelle (Chapter 1.3.3.3.). As a test reaction, labeling of chymotrypsin protein (serine protease) by FP-CEG-G1 probe was attempted in the presence of triton X-100 in 50 mM sodium phosphate buffer at room temperature (Scheme 1.5). The progress of the reaction (formation of protein-dendron bioconjugate) was monitored using MALDI-ToF at different time points. To our delight, we observed two peaks after 12 h; the molecular weight of first peak corresponds to native protein, while the molecular weight of second peak matches with a theoretical molecular weight of protein-dendron bioconjugates (Figure 2.2a). Using a similar procedure, the reactivity of all the macromolecular AABPs was tested. The MALDI-ToF results indicate that all of them, including the most hydrophobic and highly branched FP-CEG-G4, were reactive towards chymotrypsin (Figure 2.2a).



Scheme 1.5 | Scheme for the synthesis of protein-dendron bioconjugates.

Further, to investigate the effect of protein head on protein-dendron self-assembly and to broaden the scope of methodology; conjugation with three other serine protease, i.e., trypsin, proteinase K, subtilisin with FP-CEG-G1 was performed under similar conditions (Scheme 1.5). As expected, FP-CEG-G1 labeled all the serine proteases as evident from MALDI-ToF results (Figure 2.2b). This is understandable because all the proteins used in this study have the same catalytic triad (serine, aspartic acid, and histidine residues) and hence capable of reacting with an FP group in a site-specific manner.³⁹ Thus, unlike previously reported methods, our technology is applicable to a wide variety of proteins and hence amenable for the custom design of protein nanoassemblies with very rich structural diversity.

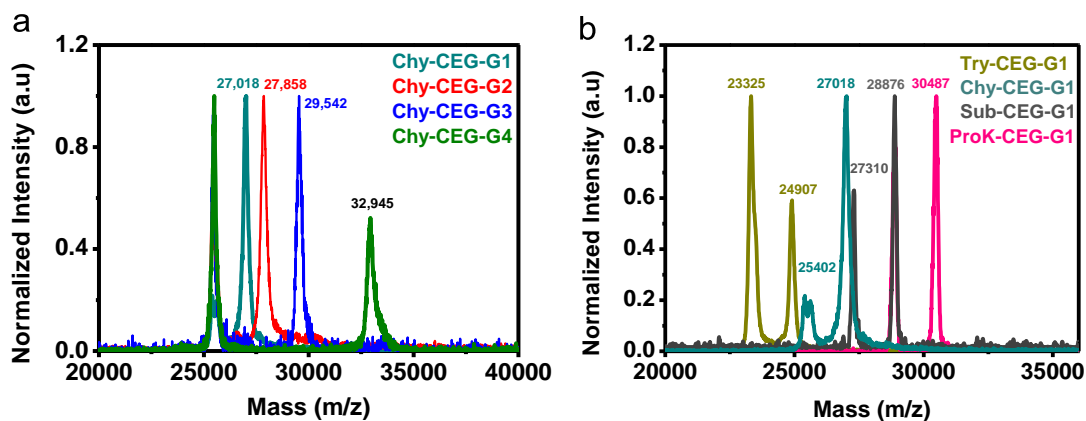


Figure 2.2 | MALDI-ToF characterization of protein-dendron bioconjugates. (a) Reaction mixture dendron variants and (b) protein variants.

2.2.4. Multi-step Purification and Characterization of Protein-Dendron Bioconjugates

We anticipated that the purification and analytical characterization of protein-dendron bioconjugates (Figure 2.1) would be highly challenging, as they are facially amphiphilic bio-hybrid macromolecules and resemble membrane proteins in terms of hydrophathy index. For the purification, the strategy pioneered by our group was followed.³⁷ Triton X-100 and unreacted macromolecular AABPs from the reaction mixture were removed by ion-exchange chromatography (IEX) (Figure 2.3). IEX purification of all the bioconjugates was successful except Chy-CEG-G4, presumably due to its extreme hydrophobicity. The second step involves the removal of native protein from protein-dendron bioconjugate, which is highly challenging because the molecular weight difference between the native protein and protein-dendron bioconjugate is very small. In addition, the surface charge of protein-dendron bioconjugate is same as the native protein. Considering these issues, traditional SEC and ion-exchange chromatography cannot be employed for the purification. In order to solve this problem, we hypothesized that increasing the ionic strength of solution would promote self-sorting of protein-dendron bioconjugate but not the native protein. This is a reasonable assumption because protein-dendron bioconjugate contains two distinct domains, i.e., hydrophilic protein and hydrophobic dendron, whereas native protein does not contain the hydrophobic domain. To do that, SEC studies were

carried out in 200 mM NaCl phosphate buffer pH 7.4. To our delight, we observed two well-resolved peaks with a substantial difference in elution volume (Figure 2.4). MALDI-ToF analysis of the first peak revealed that the first fraction exclusively contains protein-dendron bioconjugate, and the second fraction contains only native protein (Figure 2.5a). Using the same strategy, all the bioconjugates in a library I and II were purified and characterized by using MALDI-ToF (Figure 2.3, 2.4, 2.5).

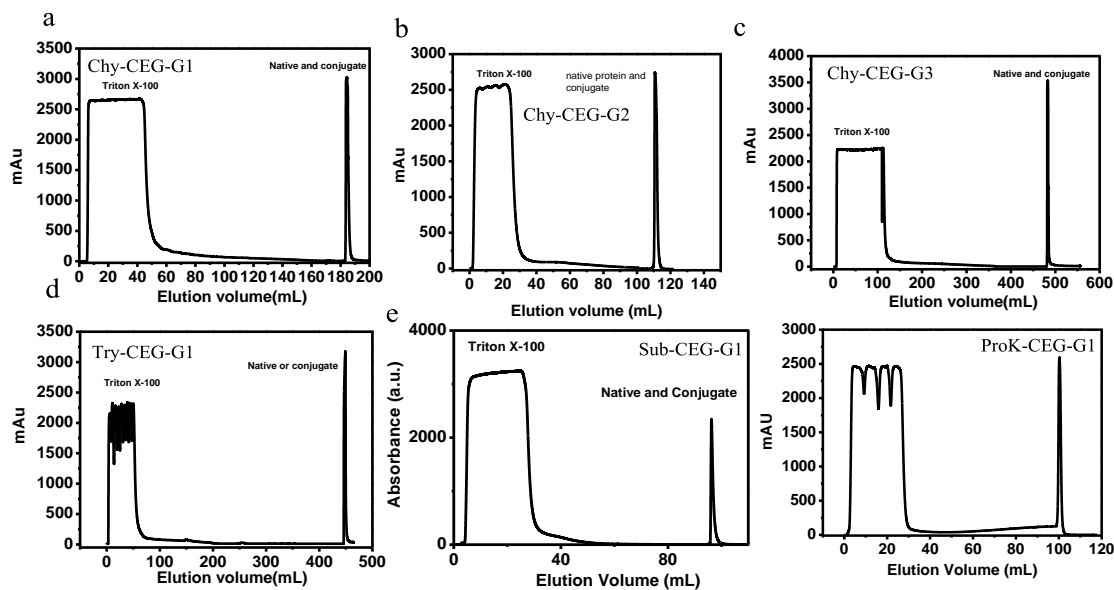


Figure 2.3 | Ion exchange chromatogram. The IEX is performed to remove neutral triton X-100 and excess of macromolecular AABP. IEX of (a) Chy-CEG-G1, (b) Chy-CEG-G2, (c) Chy-CEG-G3, (d) Try-CEG-G1, (e) Sub-CEG-G1, and (f) ProK-CEG-G1, respectively.

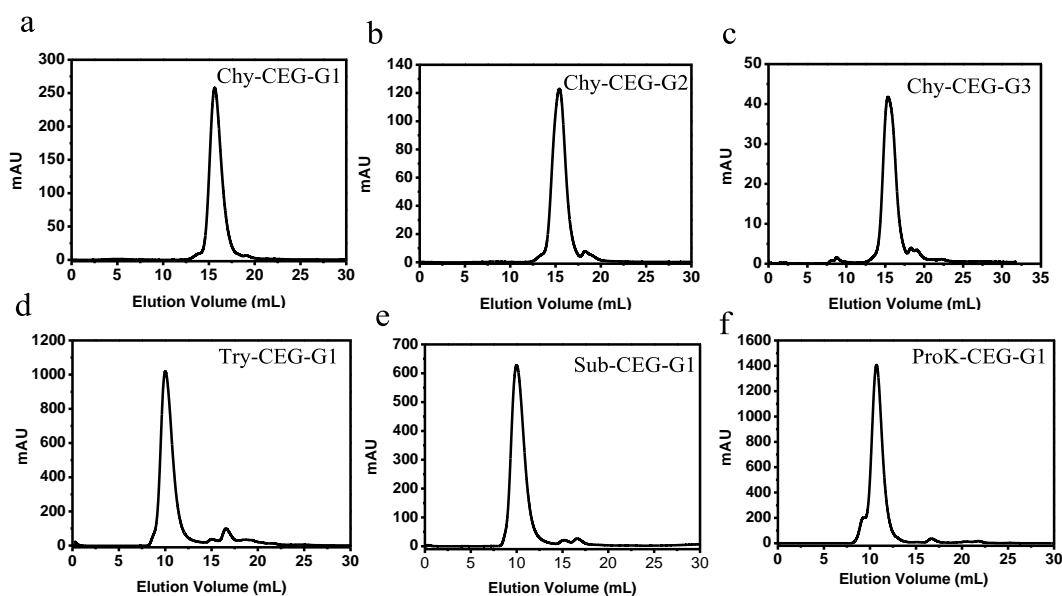


Figure 2.4 | Size exclusion chromatography. The SEC is performed to separate bioconjugates from native protein. SEC chromatogram of (a) Chy-CEG-G1, (b) Chy-CEG-G2, (c) Chy-CEG-G3, (d) Try-CEG-G1, (e) Sub-CEG-G1, and (f) ProK-CEG-G1, respectively.

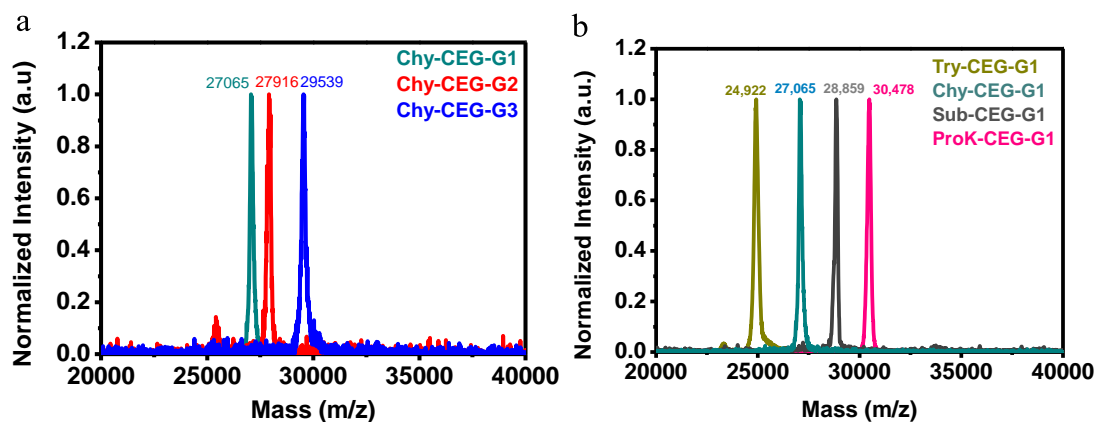


Figure 2.5 | MALDI-ToF characterization of protein-dendron bioconjugates. (a) Purified dendron variants and (b) protein variants.

2.2.5. Self-assembly Studies

Next, the ability of these purified protein-dendron bioconjugates to self-assemble into nanoassemblies was first tested by doing size exclusion chromatography (SEC) studies. All the bioconjugates in the library I and II eluted at lower elution volume indicative of the formation of higher-order protein complexes (Figure 2.6a-b). In addition to SEC studies, we have also carried out dynamic light scattering (DLS) experiments to get information on hydrodynamic diameter (D_h) of protein complexes. The D_h of bioconjugate containing Chy-CEG-G1 self-assemble into a protein complex of ~ 10 nm. An increase in dendron volume in the case of Chy-CEG-G2 resulted in an overall increase in the size of a protein complex by ~ 2 nm. Interestingly, the D_h of Chy-CEG-G3 protein complex was found to be 18 nm, which approximately double the size of protein complex Chy-CEG-G1 (Figure 2.6c). The observed size difference between Chy-CEG-G1 and Chy-CEG-G3 protein complexes is mainly attributed to the increase in hydrophobic dendron volume.⁴³

Further, we set out to examine the effect of protein size and surface charge on D_h of protein complexes. All the bioconjugates in the library I have the identical linker (CEG) and G1 dendron domain, but the protein is systematically varied in the macromolecular design. DLS results revealed that bioconjugates Try-CEG-G1, Sub-CEG-G1, and ProK-

CEG-G1, self-assembled to make protein complexes of sizes 10, 12, and 13 nm, respectively. (Figure 2.6b-d).

In order to further investigate the polydispersity, molecular weight, and oligomeric state of protein complexes, we performed size-exclusion chromatography coupled with multi-angle laser light scattering (SEC-MALS). We were gratified to note that the polydispersity index of protein complexes obtained from SEC-MALS was in the range of 1.001 to 1.003 (Table 1, Figure 2.7). These results indicate all the protein complexes are nearly monodisperse, which is an unexpected finding considering the very flexible nature of the linker and highly branched dendrimer domain. We believe that this behavior may be attributed to the monodisperse nature of protein-dendron bioconjugates.

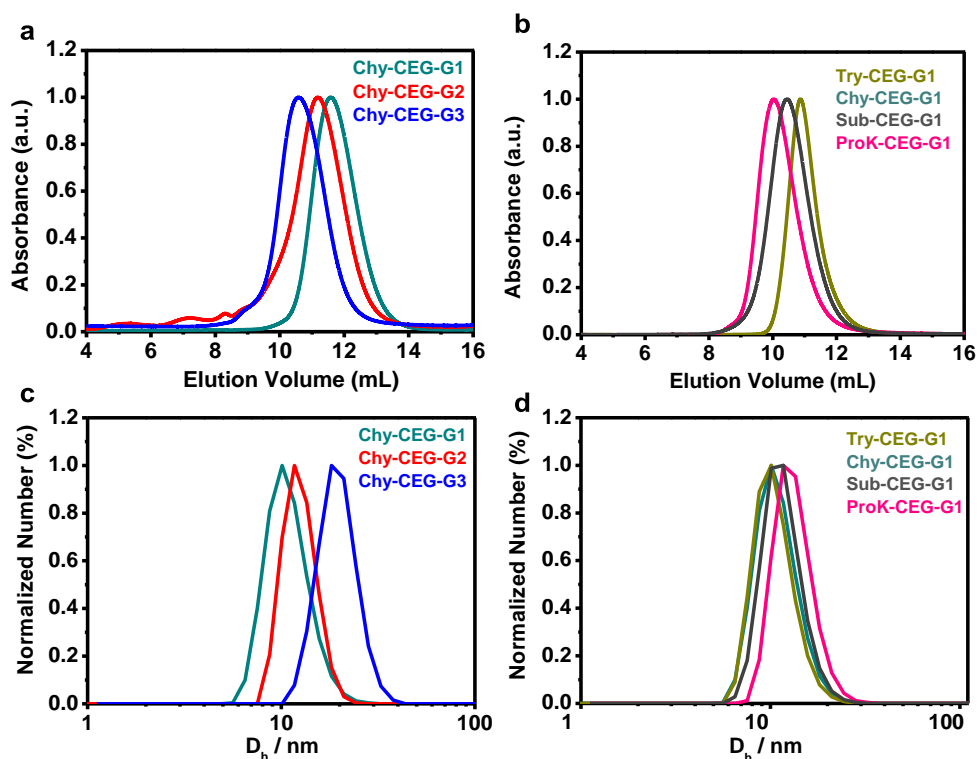


Figure 2.6 | Self-assembly data of protein-dendron conjugates. SEC data for (a) Dendron variants, and (b) Protein variants. Experiments were carried out on a Superdex-200 10/300 GL column, and 500 μL of the sample (5 mg/mL) was injected. DLS data for (c) Dendron variants, and (d) Protein variants.

The above results indicate that the synthesis of monodisperse protein nanoassemblies is feasible, provided all the components of a macromolecular design are perfectly monodisperse, as exemplified in this study. The oligomeric state and the

molecular weight of protein complexes revealed the generation-dependent behavior (library II). Bioconjugate Chy-CEG-G1 forms ~500 kDa protein complex with an average oligomeric state of 18, whereas Chy-CEG-G2 self-assembles into ~520 kDa complex with an oligomeric state of 20 (Figure 2.7a, Table 1). Remarkably, the bioconjugate having the highest dendron volume in a library II, i.e., Chy-CEG-G3 self-assemble into megadalton-scale protein complex having a molecular weight of ~ 1.3 MDa with an average oligomeric state of 44 (Figure 2.7c, Table 1). The obtained SEC-MALS results are in agreement with DLS data and manifest the effect of dendron size on the oligomeric state and size of protein complexes in the library II. Next, SEC-MALS results of the library I revealed that the Try-CEG-G1 also formed 480 kDa protein complex as expected with an average oligomeric state of 18, which is similar to the results obtained with Chy-CEG-G1 (Figure 2.7d, Table 1). The SEC-MALS studies of Sub-CEG-G1 could not be performed due to the instability problem right after ion-exchange purification. We do not know the reason for this behavior, a detailed investigation is underway, and the results will be reported in due course. ProK-CEG-G1 also self-assembled into 840 kDa (13 nm) protein complex with an average oligomeric state of 26 (Figure 2.7e, Table 1). These results indicates the oligomeric state of protein complex strongly depends on the surface charge of native proteins (try = 30.5, chy = 19.5, sub = -6.0, pro = -8.5 mV) .

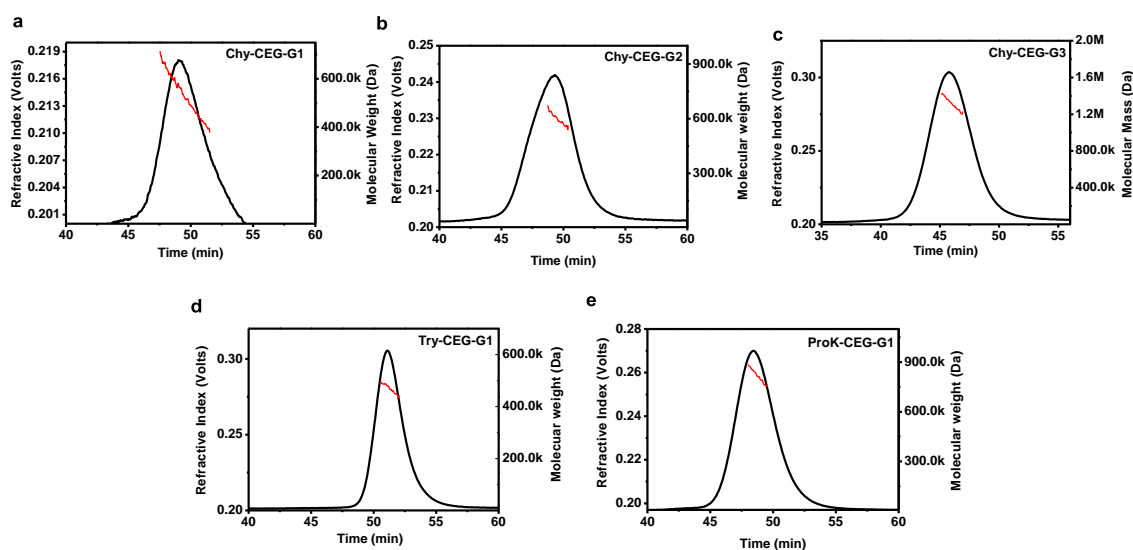


Figure 2.7 | SEC-MALS data of protein-dendron bioconjugates (a) Chy-CEG-G1 (b) Chy-CEG-G2 (c) Chy-CEG-G3, (d) Try-CEG-G1 (e) ProK-CEG-G1, respectively.

Table 1 | Summary of self-assembly studies protein-dendron conjugates.

Protein-dendron conjugate	Elution Vol SEC (mL)	Mol wt from SEC – MALS (kDa)	Oligomeric state from SEC-MALS (mer)	Polydispersity index (PDI)	Hydrodynamic diameter (D_h) DLS (nm)
Chy-CEG-G1	11.6	522	20	1.025	10 ± 3
Chy-CEG-G2	11.2	522	20	1.003	12 ± 2
Chy-CEG-G3	10.6	1320	44	1.003	18 ± 2
Try-CEG-G1	10.9	458	18	1.001	10 ± 2
Sub-CEG-G1	10.5	-----	-----	-----	12 ± 3
Prok-CEG-G1	10.0	842	26	1.002	14 ± 3

2.3. Conclusion

In summary, we have successfully synthesized a suite of monodisperse facially amphiphilic protein-dendron conjugates using a micelle-assisted protein labeling method. For the first time, we have demonstrated that the monodisperse three-dimensional hydrophobic dendrons could be site-specifically conjugated to a globular protein. The designed protein complexes exhibit both dendrimer generation-dependent as well as protein-dependent self-assembly properties. Most importantly, this method provides exquisite control over size, oligomeric state, and the molecular weight of the protein nanoassemblies by either choosing an appropriate dendron or protein of interest and thus open up the possibility to design nearly monodisperse, megadalton-scale protein nanoassemblies through a chemical method. The fact that this method could be applicable to a wide variety of proteins and dendrimers opens up enormous opportunities for this method in the area of vaccine design, targeted drug delivery, and diagnostic imaging.

2.4. Experimental Section

2.4.1. Synthesis, Purification, and Self-Assembly of Protein-Dendron Bioconjugates

2.4.1.2. Protein Modification

Protein (trypsin, chymotrypsin, subtilisin, proteinase K) modification was carried out at the concentration of 100 μM , which was found to be optimum for MALDI-ToF MS monitoring. Triton X-100 was used to solubilize the macromolecular AABPs at 10X of its critical micelle concentration (CMC) or 2% of the total volume of the reaction mixture. Typically, for test reactions, the final volume of the reaction mixture is 1 mL. Protein was weighed (trypsin = 2.3 mg, chymotrypsin = 2.5 mg, subtilisin = 2.7 mg, proteinase K = 2.9 mg) in microcentrifuge tubes and 500 μL 50 mM sodium phosphate pH 7.4 was added and mixed gently with a pipette to make 200 μM solutions. Then, macromolecular AABPs (1 or 2 equivalent) were weighed in a different microcentrifuge tube, followed by the addition of 20 μL triton X-100. This mixture was sonicated until it becomes homogenous (5 minutes). Then, 480 μL of 50 mM sodium phosphate pH 7.4 was added to the triton solution and vortexed for 1 minute. Further, the clear solution of macromolecular AABPs was added to the solution of protein to get 100 μM (1 mL) of final protein concentration and allowed to react for 24 h on rotospin at 20 rpm at 25 $^{\circ}\text{C}$. Scale-up protein modifications were carried out in falcon tubes at a 200 mg scale following the linear scale-up of the procedure mentioned above. Then, the obtained protein-dendron bioconjugates were purified by three-step purification, *i.e.*, IEX, SEC, and desalting, performed using Akta Pure.

2.4.1.2. Matrix Preparation for MALDI-ToF

Molecular weight determination of native proteins and all stages of its modification was followed by MALDI-ToF MS. The samples were analyzed in Linear High Mass mode in AB Sciex 4800 plus MALDI-ToF/ToF analyzer with 4000 Series Explorer as software. Mass was scanned between 10,000 Da and 40,000 Da with focus mass at 25,000-33,000 Da depending on the protein-dendron bioconjugate analyzed.

4 mg of DHAP and 4.5 mg of DAHC were weighed separately in microcentrifuge tubes. 150 μ L of ethanol and 200 μ L of milli-Q water were added to each tube, respectively. Both the solutions were sonicated using bath sonicator for 1 minute and vortexed for another minute. Then 50 μ L of DAHC aqueous solution was transferred to DHAP solution, and the resulting solution was vortexed for one more minute to yield matrix mixture. The sample preparation was carried out in a 500 μ L microcentrifuge tube, 1:1:1 ratio of a protein sample, 2% trifluoroacetic acid (TFA), and matrix mixture were mixed and vortexed for 1 minute. The samples were then kept undisturbed. When the crystallization observed in centrifuge tubes, 0.5 μ L of the samples were spotted on the MALDI plate and air-dried for 15 minutes. The plate was then loaded and fired to get accurate molecular weight both in +1 and +2 states. 100 μ M protein concentration was found to be optimum for MALDI-ToF MS analysis.

2.4.1.3. Monitoring of Modification

To monitor the extent of protein modification, the samples were directly withdrawn from the reaction mixture using a pipette and analyzed. In brief, 2 μ L of the reaction mixture was mixed with 2 μ L of 2% TFA and 2 μ L of matrix mixture as previously stated (2.4.1.2.), vortexed, and spotted on MALDI-ToF MS plate. Please note that the matrix preparation, sample preparation procedures, and analysis procedure remained the same here.

2.4.1.4. Purification of Protein-Dendron Bioconjugates

All the protein-dendron bioconjugates (except Chy-CEG-G4) were purified by three-step purification, *i.e.*, IEX, SEC, and desalting performed using Akta Pure. IEX was performed to remove triton X-100 using either SP sepharose or Q sepharose resins (GE) depending on isoelectric point (pI) and surface charges of proteins (Try: 30.4, Chy: 19.5, Pro K: -8.4, Sub: - 6.2). For example, to purify the reaction mixture of trypsin or chymotrypsin, we used SP sepharose, a cation-exchange resin at pH 7.4, and to purify subtilisin and proteinase K we used Q sepharose, an anion-exchange resin at pH 10. In cation-exchange chromatography, the column was pre-equilibrated with the same buffer (50 mM sodium phosphate pH 7.4) which was used for modification and then a sample was

injected followed by post-injection equilibration for at least 2 Column Volumes (CVs) or until the complete removal of triton X-100 for large scale reactions. The elution of native protein and its corresponding protein-dendron bioconjugate together as the mixture was later achieved using 50 mM sodium phosphate pH 7.4, 1 M NaCl as elution buffer.

In the case of anion-exchange chromatography, the protein reactions were buffer-exchanged first to 50 mM tris base pH 10 (since the buffer used for modification was 50 mM tris base pH 8.5). Then the anion-exchange column was pre-equilibrated using the same buffer (50 mM tris base pH 10), which was used for buffer-exchange, and then the sample was injected, followed by post-injection equilibration for at least 2 CVs or until the complete removal of triton X-100 for large scale reactions. The elution of the native protein and its protein-dendron bioconjugate together as the mixture was later achieved using 50 mM tris base pH 10, 1 M NaCl as elution buffer.

The obtained IEX fractions were subjected to SEC. For the separation of native protein from protein-dendron bioconjugates, 50 mM sodium phosphate pH 7.4, 200 mM NaCl was used as buffer using either Superdex-200 10/300 or Sephacryl-200, GE Healthcare column. The column was pre-equilibrated with 50 mM sodium phosphate with 200 mM NaCl for at least 2CVs, and then the sample was injected, followed by post-injection equilibration with the same for at least 2CVs again or until the complete elution of the proteins. The NaCl was later removed by the Sephadex-G25 desalting column. The column was pre-equilibrated with Milli Q water for at least 2CVs, and then the sample was injected, followed by post-injection equilibration with Milli Q water for at least 2CVs again until the complete elution of the proteins. The desalted fractions were quickly lyophilized and later dissolved in a required buffer when needed.

2.4.1.5. Molecular Weight Determination of Purified Protein-Dendron Bioconjugates

Monitoring of IEX, SEC, and desalted fractions were carried out using the same procedure mentioned for reaction mixtures (2.4.1.3.), except for the addition of triton X-100 in protein fractions. In short, 98 μ L of purified samples from each fraction were mixed with 2 μ L of triton X-100 (2% or 100 times the CMC) in a separate microcentrifuge tube

and vortexed for 4 h. The samples were then analyzed using the same procedure as mentioned above. Please note that the matrix and sample preparation procedure remained the same.

2.4.1.6. Dynamic Light Scattering

The hydrodynamic diameter of native proteins and protein-dendron bioconjugates protein nanoparticles was measured using DLS (Zetasizer Nano 2590, Malvern, UK). Samples (5 mg/mL) were prepared in 50 mM sodium phosphate pH 7.4. 1 mL of sample was taken in disposable polystyrene cells, and then the mean size of the complexes was measured at 90° scattering angle.

2.4.1.7. SEC-MALS

SEC-MALS analyses of the protein-dendron complexes were performed on a Superose-200 10/300 GL column (GE Healthcare) connected to the Postnova AF200AT system equipped with the 21-angle light scattering detector (Postnova PN3621) and a refractive index detector (Postnova PN3150). The system was calibrated with BSA at a concentration of 2 mg/mL; 100 µL of the protein-dendron bioconjugates at concentrations 5 mg/mL were injected, and the molecular weights were calculated using AF2000 software (Postnova).

2.4.2. Synthesis and Purification of Macromolecular AABP and Their Intermediates

All reagents were obtained commercially unless and otherwise stated. The reactions were performed in an oven-dried round bottom flask (RBF) and under nitrogen atmosphere. Air and moisture sensitive solvents were transferred *via* syringe. Reactions were monitored by thin-layer chromatography (TLC), and the developed chromatogram was visualized by ultraviolet (UV) lamp or by phosphomolybdic acid (PMA) staining. Product purification was accomplished by 100-200 mesh size silica gel column chromatography.

All the compounds were characterized by ¹H, ¹³C, and ¹⁹F (in case of fluorinated compounds) nuclear magnetic resonance (NMR) using Bruker or Jeol 400 MHz. ¹H and ¹⁹F

were recorded at an operating frequency of 400 MHz and 100 MHz for ^{13}C using, using TMS as an internal standard. All the ^{13}C Chemical shifts were mentioned in parts per million (PPM) and measured relative to residual CHCl_3 , CH_3OH , or CH_3CN in their deuterated solvent. Coupling constants were reported in Hertz (Hz). Multiplicities were explained as s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, quint = quintet. Mass spectra were obtained with either the MALDI-TOF MS or HRMS. Room temperature varied between 21-35 °C.

2.4.2.1. Synthesis of CEG Spacers

2.4.2.1.1. Procedure for Synthesis of Monotryllolethylene Glycol - Procedure A

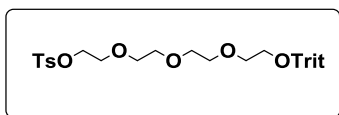
In an oven-dried RBF, monopropargyl oligoethylene glycol (1 eq) and trityl protected tosyl TEG (2 eq) was dissolved in THF under stirring. Then, sodium hydride (NaH) (4 eq) was added in a small portion at 0 °C. The reaction was allowed to react for 12 hours at RT. Upon completion, excess sodium hydride (NaH) was quenched by the dropwise addition of water, and the reaction mixture was extracted with dichloromethane (DCM) for thrice. The combined organic layer was dried over sodium sulphate (NaSO_4) and concentrated under reduced pressure to get the crude residue, which was purified using silica gel column chromatography.

2.4.2.1.2. Procedure for Trityl Deprotection - Procedure B

In an oven-dried RBF, a mixture of monotryl oligoethylene glycol (1 eq) and p-toulenesulfonic acid (TsOH) (1.5 eq) was taken and dissolved in methanol under stirring. The mixture was allowed to react for 12 hours at RT. Upon completion, methanol was evaporated under reduced pressure. To the obtained residue, water was added and extracted thrice in DCM. The combined organic layer was dried over NaSO_4 and concentrated under reduced pressure to get the crude residue, which was purified using silica gel column chromatography.

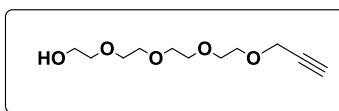
2.4.2.1.3. Synthesis of CEG Spacers and Their Intermediates

Synthesis of compound 1a



Monotrityl TEG (58 g, 132 mmol) was dissolved in THF under stirring. To the above solution, aq. solution of potassium hydroxide (KOH) (26 g, 464 mmol) was added and allowed to stir for 10 minutes. Then, a solution of tosyl chloride (TsCl) (75 g, 398 mmol) in THF was slowly added and allowed to react for 12 hours at RT. Upon completion, the reaction was quenched by the dropwise addition of water and extracted with DCM for thrice. The combined organic layer was dried over Na₂SO₄ and concentrated under reduced pressure to get a crude product which was purified using silica gel column chromatography using Ethyl acetate / Hexane as eluent. The product was obtained as a pale yellow liquid (66 g, 112 mmol, 85%), *R_f* = 0.40 in 50% Ethyl acetate / Hexane. ¹H NMR (400MHz, CDCl₃): δ_H 7.76 (m, *J* = 8.4Hz, 2H), 7.47-7.45 (m, 6H), 7.34-7.18 (m, 11H), 4.11 (t, *J* = 4.8Hz, 2H), 3.67-3.52 (m, 12H), 3.23 (t, *J* = 4.8Hz, 2H), 2.39 (t, *J* = 2.4Hz, 1H). **MALDI-ToF MS** (M+K): 629.23.

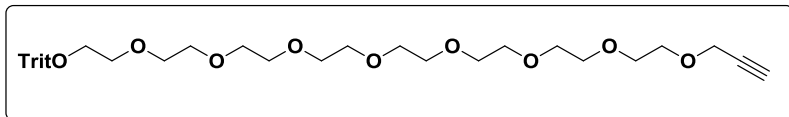
Synthesis of compound 1b



In an oven-dried RBF, TEG (15.0 g, 77 mmol) was dissolved with stirring in THF. Then, NaH (1.23 g, 51 mmol) was added to the flask in small portions at 0° C. After 1 hour, propargyl bromide (6.13 g, 51 mmol) was added dropwise, maintaining the reaction at the same temperature. Then, the reaction was stirred for 12 h at RT. Upon completion, the reaction was quenched by the dropwise addition of water and extracted with DCM for thrice. The combined organic layer was dried over Na₂SO₄ and concentrated under reduced pressure to get a crude product which was purified using silica gel column chromatography using MeOH / DCM as eluent. The product was obtained as pale yellow liquid (4.6 g, 30 mmol, 58%), *R_f* = 0.34 in 5% MeOH / DCM. ¹H NMR (400MHz, CDCl₃): δ_H 3.92 (d, *J* = 2.4Hz, 2H), 3.44-3.36 (m, 15H), 3.31 (t, *J* = 4.4Hz, 3H), 2.34 (t, *J* = 2.4Hz, 1H). ¹³C NMR

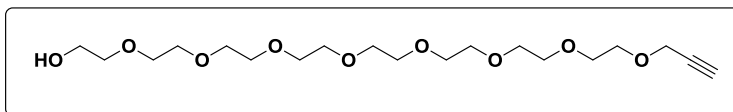
(100MHz, CDCl₃): δ_C 79.15, 74.46, 72.03, 69.95, 69.89, 69.86, 69.70, 69.68, 68.43, 60.82, 57.68, 53.30. **HRMS** (M+Na): 255.12.

Synthesis of compound 1c



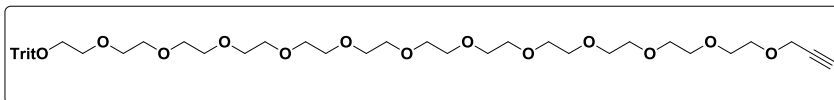
The compound **1c** was prepared by general procedure A, starting from compound **1a** (56 g, 112 mmol), compound **1b** (13 g, 56 mmol), and NaH (5.1 g, 224 mmol) in THF. The product was obtained as a pale yellow liquid (8.2 g, 21 mmol, 71%) after purification by silica gel column chromatography using MeOH / DCM as eluent, $R_f = 0.47$ in 5% MeOH / DCM. **¹H NMR** (400MHz, CDCl₃): δ_H 7.47-7.45 (m, 6H), 7.31-7.20 (m, 9H), 4.20 (d, 2.4Hz, 2H), 3.71-3.61 (m, 27H), 3.23 (t, $J = 4.4$ Hz, 3H), 2.43 (t, $J = 2.4$ Hz, 1H). **¹³C NMR** (100MHz, CDCl₃): 144.28, 128.87, 127.90, 127.06, 77.16, 70.94, 70.83, 70.72, 70.67, 70.60, 70.56, 69.27, 63.48, 58.56, 53.57. **MALDI-ToF MS** (M+K): 689.38.

Synthesis of compound 1d



The compound **1d** was prepared by general procedure B, starting from compound **1c** (13 g, 20 mmol) and TsOH (11.4 g, 60 mmol) in MeOH. The product was obtained as a pale yellow liquid (4 g, 9.8 mmol, 50%) after purification by silica gel column chromatography using MeOH / DCM as eluent, $R_f = 0.43$ in 5% MeOH / DCM. **¹H NMR** (400MHz, CDCl₃) δ_H 4.19 (d, $J = 2.4$ Hz, 2H), 3.72-3.59 (m, 32H), 2.43 (t, $J = 2.4$ Hz, 2H). **¹³C NMR** (100MHz, CDCl₃) δ_C 79.77, 74.67, 72.70, 70.72, 70.67, 70.51, 70.40, 69.22, 61.83, 58.52, 31.07. **MALDI-ToF MS** (M+K): 447.23.

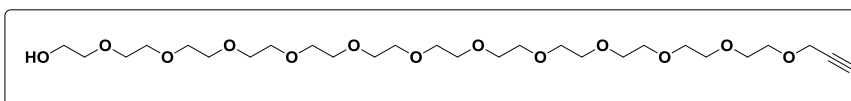
Synthesis of compound 1e



The compound **1e** was prepared by general procedure A, starting from compound **1d** (3.6 g, 8.8 mmol), compound **1a** (10.5 g, 17 mmol) and NaH (1.2 g, 50 mmol) in THF.

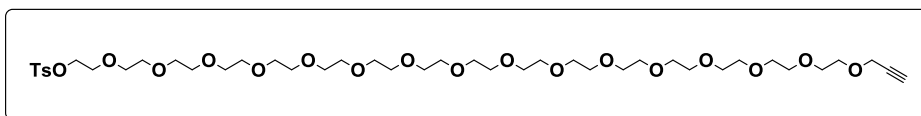
The product was obtained as a pale yellow liquid (6.2 g, 7.5 mmol, 86%) after purification by silica gel column chromatography using MeOH / DCM as eluent, $R_f = 0.47$ in 5% MeOH / DCM. $^1\text{H NMR}$ (400MHz, CDCl_3) δ_{H} 7.50-7.44 (m, 7H), 7.31-7.20 (m, 11H), 4.20 (d, $J=2.4\text{Hz}$, 2H), 3.74-3.52 (m, 54H), 3.23 (t, $J=5.2\text{Hz}$, 2H), 2.43 (t, $J=2.4\text{Hz}$, 2H). $^{13}\text{C NMR}$ (100MHz, CDCl_3) δ_{C} 144.23, 128.83, 127.88, 127.03, 86.63, 74.69, 70.89, 70.82, 70.78, 70.68, 70.52, 69.22, 63.43, 58.52. **MALDI-ToF MS** (M+K): 865.32.

Synthesis of compound 1f



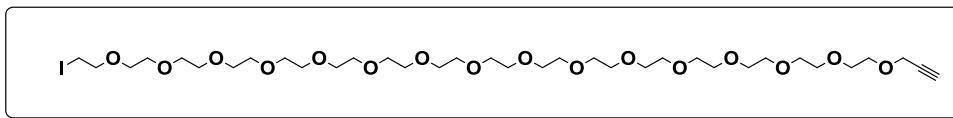
The compound **1f** was prepared by general procedure B, starting from compound **1e** (6.5 g, 7.4 mmol) and TsOH (5.7 g, 21 mmol) in MeOH. The product was obtained as a pale yellow liquid (1.6 g, 2.7 mmol, 35%) after purification by silica gel column chromatography using MeOH / DCM as eluent, $R_f = 0.43$ in 5% MeOH / DCM. **MALDI-ToF MS** (M+K): 623.38.

Synthesis of compound 1h



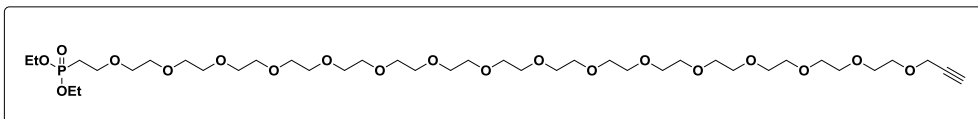
In an oven-dried RBF, the compound **1f** (1 g, 1.3mmol), DMAP (0.08 g, 0.6 mmol), and ditosyl TEG (**1g**) (3 g, 3 mmol) were taken and dissolved in DCM under stirring. The mixture was cooled to 0 °C, and Et_3N (0.9 g, 5 mmol) was added dropwise. The resultant mixture was then stirred for 12 hours at RT. Upon completion, the reaction was quenched by the dropwise addition of water and extracted with DCM for thrice. The combined organic layer was dried over Na_2SO_4 and concentrated under reduced pressure to get a crude product which was purified using silica gel column chromatography. The product was obtained as pale yellow liquid (1 g, 1.0 mmol, 83%), $R_f = 0.45$ in 5% MeOH / DCM. $^1\text{H NMR}$ (400MHz, CD_3OD) δ_{H} 7.83 (d, $J=8\text{Hz}$, 2H), 7.48 (d, $J=8\text{Hz}$, 2H), 4.20-4.15 (m, 4H), 3.69-3.54 (m, 53H), 3.32 (s, 7H), 2.87 (t, $J=2.4\text{Hz}$, 1H), 2.48 (s, 3H). $^{13}\text{C NMR}$ (100MHz, CD_3OD) δ_{C} 146.44, 134.43, 130.83, 129.09, 76.09, 71.56, 71.37, 70.97, 70.10, 69.75, 59.03, 54.83, 49.00, 21.62. **MALDI-ToF MS** (M+K): 953.22.

Synthesis of compound 1i



A mixture of above-obtained tosylate (**1h**) (1 g, 1.0 mmol), KI (0.72 g, 4.3 mmol) was refluxed in acetone for 18 hours. Upon completion, excess KI was filtered and washed thrice with acetone. Collected acetone fraction was evaporated under vacuum to get residue, which was then washed with water and extracted with DCM. The combined organic layer was washed with aqueous Na₂CO₃ and then concentrated under vacuum to get the crude product, which was purified using silica gel column chromatography using MeOH / DCM as eluent. The product was obtained as a pale yellow liquid (0.85 g, 1.0 mmol, 89%), *R_f* = 0.45 in 5% MeOH / DCM. ¹H NMR (400MHz, CD₃OD): δ_H 4.21 (d, *J*=4Hz, 2H), 3.78 (t, *J*=8Hz, 2H), 3.70-3.65 (m, 62H), 3.34-3.31 (m, 6H), 2.88 (t, *J*=4Hz, 1H). ¹³C NMR (100MHz, CD₃OD): δ_C 80.65, 75.96, 73.10, 71.64, 71.56, 71.37, 71.17, 70.11, 59.05, 54.81, 49.00. MALDI-ToF MS (M+K): 909.26.

Synthesis of compound 1j



In an oven-dried RBF, iodide (**1j**) (0.72 g, 0.8 mmol) and triethyl phosphite, P(OEt)₃ (0.5 g, 3.3 mmol) were taken and refluxed for 1 hour at 150 °C. Upon completion of the reaction, the excess P(OEt)₃ was removed under vacuum, and the reaction mixture was directly loaded onto a silica gel column, and the crude mixture was purified using MeOH / DCM as eluent. The product was obtained as a pale yellow liquid (0.80 g, 0.71 mmol, 83%), *R_f* = 0.45 in 5% MeOH / DCM. ¹H NMR (400MHz, CD₃OD): δ_H 4.19 (d, *J*=2Hz, 2H), 4.16-4.03 (m, 4H), 3.76-3.58 (m, 64H), 2.43 (t, *J*=2.4Hz, 1H), 2.15-2.06 (m, 2H), 1.32 (t, *J*=8Hz, 6H). ¹³C NMR (100MHz, CD₃OD): δ_C 109.07, 104.14, 99.85, 99.56, 98.20, 91.48, 87.18, 83.00, 77.16, 54.69, 44.89. MALDI-ToF MS (M+K): 919.18.

2.4.2.2. Synthesis Dendrimer

2.4.2.2.1. Procedure for the Synthesis of G0-G4 Bromide - Procedure C

In an oven-dried RBF, alcohol and tetrabromomethane (CBr₄) were taken and dissolved in DCM. Then a solution of triphenylphosphine (PPh₃) in DCM was added dropwise at 0 °C and allowed to stir for 3 hours at RT. Upon completion of the reaction, DCM was evaporated under reduced pressure. The obtained residue was directly purified using silica gel column chromatography.

2.4.2.2.2. Procedure for the Synthesis of G1-G4 Alcohol – Procedure D

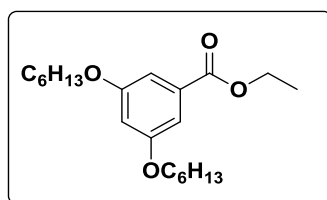
In an oven-dried RBF bromide, 3, 5-dihydroxybenzyl alcohol, potassium carbonate (K₂CO₃), and crown ether (18-crown-6) were taken. The resulting mixture was refluxed for 36 hours. Upon completion, the reaction mixture was cooled to RT. Then acetone was evaporated under reduced pressure. To the obtained residue, water was added and extracted with DCM thrice. The combined organic layer was dried over Na₂SO₄ and concentrated under reduced pressure to get crude product which was purified using silica gel column chromatography.

2.4.2.2.3. Procedure for the Synthesis of G1-G4 Azide - Procedure E

To the mixture of bromide and sodium azide (NaN₃) in oven-dried RBF, DMSO was added and stirred for 28 hours at 80 °C. Upon completion of the reaction, water was added at 0 °C to quench the reaction. The resulting mixture was then extracted in DCM for thrice. The combined organic layer was dried over Na₂SO₄ and concentrated under reduced pressure to get crude product which was purified using silica gel column chromatography.

2.4.2.2.4. Synthesis of Dendrimer Azide and Their Intermediates

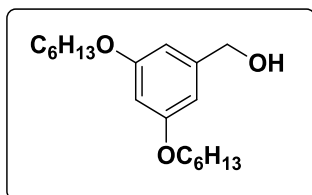
Synthesis of Compound 2a



To the oven-dried RBF ethyl 3, 5-dihydroxy benzoate (5.0 g, 27 mmol), K₂CO₃ (11.5 g, 82 mmol), 1-hexyl bromide (11.5 g, 70 mmol) were taken. DMF was added under

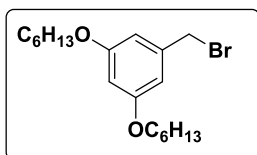
stirring and heated at 75 °C for 12 hours. Upon completion, the reaction mixture was neutralized with a 1N HCl solution. The resulting solution was then extracted thrice with ethyl acetate. The combined organic layers was dried over Na₂SO₄ and concentrated under reduced pressure to get the crude product, which was then purified using silica gel column chromatography using ethyl acetate/hexane to afford **12a** (9.0 g, 25 mmol 96%) as a colorless liquid. $R_f = 0.46$ in 5% ethyl acetate / hexane. ¹H NMR (400MHz, CDCl₃): δ_H 7.16 (d, $J=2.4$ Hz, 2H), 6.63 (t, $J=2.4$ Hz, 1H), 4.34 (q, $J=7.2$ Hz, 2H), 3.97 (t, $J=6.8$ Hz, 4H), 1.77 (quint, $J=6.8$ Hz, 4H), 1.49-1.31 (m, 15H), 0.90 (t, $J=6.8$ Hz, 6H). ¹³C NMR (100MHz, CDCl₃): δ_C 166.67, 160.25, 132.32, 107.74, 106.41, 68.43, 61.21, 31.70, 29.29, 25.83, 22.74, 14.47, 14.18. **MALDI-ToF MS**: (M+K⁺) 389.20.

Synthesis of Compound 2b



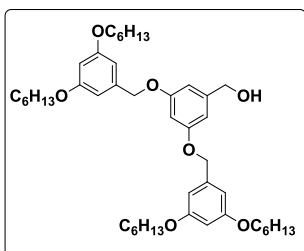
In an oven dried RBF, compound **2a** (40.0 g, 114 mmol) was taken and dissolved in THF with stirring at 0 °C. Then lithium aluminum hydride (LAH) (19.0 g, 500 mmol) was added in small portions, maintaining reaction temperature 0 °C. After 10 minutes, stirring was continued at RT for 2 hours. Upon completion of the reaction, excess of LAH was quenched with dropwise addition of water at 0 °C. Resulting mixture was stirred at RT until off-white precipitate forms. Aqueous hydrochloric acid (HCl) (4N) was then added to get a clear solution. Resulting content was extracted in ethyl acetate thrice. Combined organic layers was dried over Na₂SO₄ and concentrated under reduced pressure to get the crude product, which was purified using silica gel column chromatography to afford **10b** (30.0 g, 97 mmol, 85%) as a colorless liquid. $R_f = 0.32$ in 10% ethyl acetate / hexane. ¹H NMR (400MHz, CDCl₃): δ_H 6.49 (d, $J=2.4$ Hz, 2H), 6.38 (t, $J=2.4$ Hz, 1H), 4.61 (s, 2H), 3.93 (t, $J=6.8$ Hz, 4H), 1.76 (quint, $J=6.4$ Hz, 4H), 1.48-1.39 (m, 4H), 1.37-1.30 (m, 8H), 0.91 (t, $J=6.8$ Hz, 6H). ¹³C NMR (100MHz, CDCl₃): δ_C 160.61, 143.34, 105.13, 100.68, 68.16, 65.52, 31.70, 29.33, 25.84, 22.74, 14.17. **HRMS**: (M+H⁺) 308.24.

Synthesis of Compound 2c



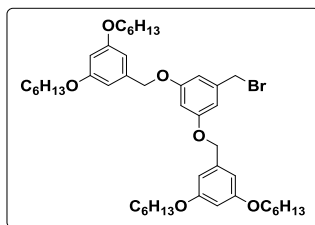
The compound **2c** was synthesized using general procedure C, starting from **2b** (5.0 g, 16 mmol), CBr_4 (6.1 g, 18 mmol) and PPh_3 (4.8 g, 18 mmol) in DCM. The product was obtained as a colorless liquid (6.1 g, 16 mmol, 99%) after purification by silica gel column chromatography using ethyl acetate / hexane as eluent, $R_f = 0.49$ in 5% ethyl acetate / hexane. $^1\text{H NMR}$ (400MHz, CDCl_3): δ_{H} 6.52 (d, $J=2\text{Hz}$, 2H), 6.38 (t, $J=2\text{Hz}$, 1H), 4.41 (s, 2H), 3.92 (t, $J=6.8\text{Hz}$, 4H), 1.76 (quint, $J=6.4\text{Hz}$, 4H), 1.48-1.39 (m, 4H), 1.37-1.30 (m, 8H), 0.91 (t, $J=6.8\text{Hz}$, 6H). $^{13}\text{C NMR}$ (100MHz, CDCl_3): δ_{C} 160.53, 139.66, 107.43, 101.50, 68.21, 33.95, 31.70, 29.31, 25.84, 22.74, 14.18; **MALDI-ToF MS**($\text{M}+\text{Na}^+$): 393.11.

Synthesis of Compound 3a



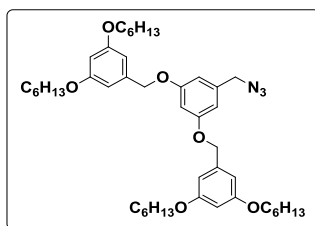
The compound **3a** was synthesized from general procedure D, starting from 3,5-dihydroxybenzyl alcohol (2.3 g, 16 mmol), **2c** (14.0 g, 37 mmol), K_2CO_3 (5.2g, 37 mmol), crown ether (0.70 g, 2 mmol) in acetone. The product was obtained as a yellowish liquid (10.8 g, 15 mmol, 90%) after purification by silica gel column chromatography using ethyl acetate / hexane as eluent. $R_f = 0.2$ in 5% ethyl acetate / hexane. $^1\text{H NMR}$ (400MHz, CDCl_3): δ_{H} 6.60 (d, $J=2\text{Hz}$, 2H), 6.55–6.53 (m, 5H), 6.40 (t, $J=2\text{Hz}$, 2H), 4.95(s, 4H), 4.62 (s, 2H), 3.93 (t, $J=6.8\text{Hz}$, 8H), 1.764 (quint, $J=6.8\text{Hz}$, 8H), 1.48-1.41 (m, 8H) 1.13-1.29 (m, 16H) 0.91 (t, $J=6.8\text{Hz}$, 12H). $^{13}\text{C NMR}$ (100MHz, CDCl_3): δ_{C} 160.60, 160.31, 143.52, 139.14, 105.86, 101.51, 105.89, 70.27, 68.21, 65.48, 31.72, 29.35, 25.82, 22.74, 14.44. **MALDI-ToF MS** ($\text{M}+\text{K}^+$): 759.36.

Synthesis of Compound 3b



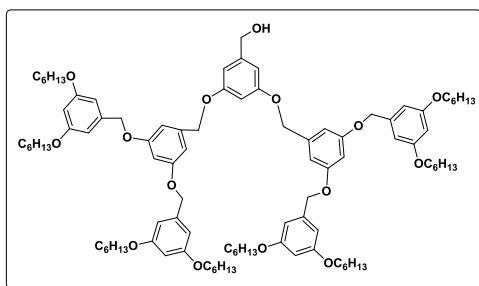
The compound **3b** was synthesized from general procedure C, starting from **3a** (5.0 g, 7 mmol), CBr_4 (3.0 g, 9 mmol) and PPh_3 (2.3 g, 9 mmol) in DCM. The product was obtained as a colorless liquid (5.2 g, 6.6 mmol, 97%) after purification by silica gel column chromatography using ethyl acetate / hexane as eluent, $R_f = 0.5$ in 5% ethyl acetate / hexane. $^1\text{H NMR}$ (400MHz, CDCl_3): δ_{H} 6.60 (d, $J=2\text{Hz}$, 2H), 6.55–6.53 (m, 5H), 6.40 (t, $J=2\text{Hz}$, 2H), 4.95(s, 4H), 4.40 (s, 2H), 3.93 (t, $J=6.8\text{Hz}$, 8H), 1.76 (quint, $J=6.8\text{Hz}$, 8H), 1.48-1.41 (m, 8H) 1.29-1.35 (m, 16H) 0.91 (t, $J=6.8\text{Hz}$, 12H). $^{13}\text{C NMR}$ (100MHz, CDCl_3): δ_{C} 160.60, 160.19, 139.84, 138.91, 108.30, 105.89, 102.38, 101.05, 70.37, 68.21, 65.48, 31.72, 29.35, 25.82, 22.74, 14.44. **MALDI-ToF MS** ($\text{M}+\text{K}^+$): 823.58.

Synthesis of Compound 3c



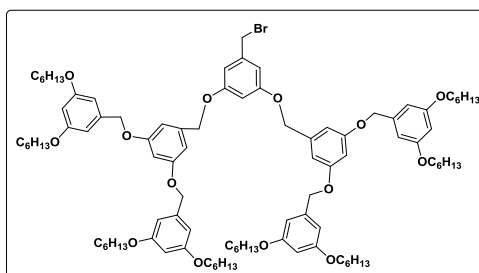
The compound **3c** was synthesized from general procedure E, starting from **3b** (1.0 g, 1.2 mmol), NaN_3 (0.46 g, 6.9 mmol) in DMSO. The product was obtained as a yellowish liquid (0.84 g, 0.94 mmol, 72%) after purification by silica gel column chromatography using ethyl acetate / hexane as eluent. $R_f = 0.32$, solvent = 5% ethyl acetate/hexane. $^1\text{H NMR}$ (400 MHz, CDCl_3): δ_{H} 6.58-6.54 (m, 7H), 6.41 (t, $J=2.4$ Hz, 2H), 4.95 (s, 4H), 4.26 (s, 2H), 3.94 (t, $J=6.8\text{Hz}$, 8H), 1.77 (quint, $J=6.8\text{Hz}$, 8H), 1.49-1.24 (m, 25H), 0.91 (t, $J=6.8\text{Hz}$, 12H). $^{13}\text{C NMR}$ (100 MHz, CDCl_3): δ_{C} 160.64, 160.34, 138.89, 137.70, 107.26, 105.81, 101.93, 100.95, 70.30, 68.19, 60.54, 54.95, 31.71, 29.34, 25.86, 22.74, 14.33, 14.18. **MALDI-ToF** ($\text{M}+\text{K}^+$): 784.44.

Synthesis of Compound 4a



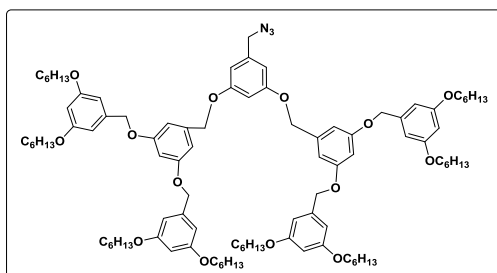
The compound **4a** was synthesized from general procedure D, starting from 3, 5-dihydroxybenzyl alcohol (0.75 g, 5.3 mmol), **3b** (9.8 g, 12 mmol), K₂CO₃ (1.8 g, 13 mmol), crown ether (0.22 g, 0.8 mmol) in acetone. The product was obtained as a yellowish liquid (8.0 g, 5.0 mmol, 94%) after purification by silica gel column chromatography using ethyl acetate / hexane as eluent. $R_f = 0.4$ in 25% ethyl acetate / hexane. ¹H NMR (400MHz, CDCl₃): δ_H 6.71-6.46 (m, 21H), 4.96(s, 12H), 4.61 (s, 2H), 3.97 (t, $J=6.4$ Hz, 16H), 1.81 (quint, $J=6.4$ Hz, 16H), 1.57-1.32 (m, 48H), 0.98 (m, 24H). ¹³C NMR (100MHz, CDCl₃): δ_C 160.61, 160.24, 160.17, 143.58, 139.31, 139.03, 106.42, 105.83, 101.67, 101.32, 100.92, 70.26, 70.08, 68.18, 65.41, 31.71, 29.34, 25.85, 22.74, 14.19. MALDI-ToF MS (M+K⁺): 1585.01.

Synthesis of Compound 4b



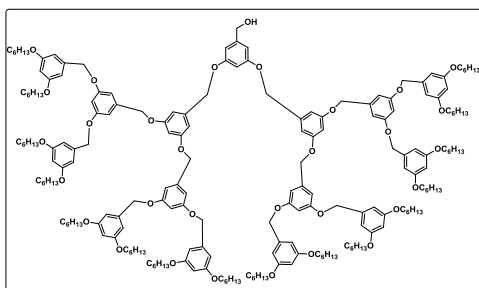
The compound **4b** was synthesized from general procedure C, starting from **4a** (7.5 g, 5.0 mmol), CBr₄ (2.1 g, 6.3 mmol) and PPh₃ (1.7g, 6.3 mmol) in DCM. The product was obtained as a colorless liquid (7.1 g, 4.4 mmol, 92%) after purification by silica gel column chromatography using ethyl acetate / hexane as eluent, $R_f = 0.4$ in 5% ethyl acetate / hexane. ¹H NMR (400MHz, CDCl₃): δ_H 6.69-6.44 (m, 21H), 4.98(s, 12H), 4.43 (s, 2H), 3.96 (t, $J=6.4$ Hz, 16H), 1.78 (quint, $J=6.4$ Hz, 16H), 1.56-1.33 (m, 48H), 0.93 (t, $J=6.8$ Hz, 24H). ¹³C NMR (100MHz, CDCl₃): δ_C 160.63, 160.27, 160.10, 139.89, 139.02, 108.24, 106.52, 105.84, 102.31, 101.78, 100.95, 77.16, 70.28, 70.19, 68.17, 31.71, 29.35, 25.85, 22.77, 14.16. MALDI-ToF MS (M+K⁺): 1648.02.

Synthesis of Compound 4c



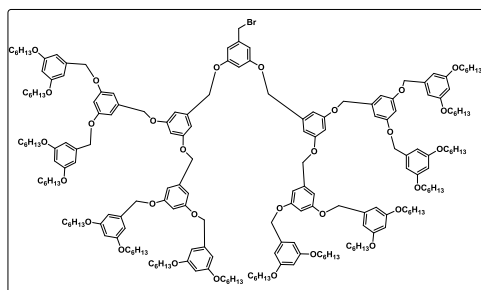
The compound **4c** was synthesized from general procedure E, starting from **4b** (0.80 g, 0.5 mmol), NaN₃ (0.33 g, 5.0 mmol) in DMSO. The product was obtained as a yellowish liquid (0.51 g, 0.3 mmol, 68%) after purification by silica gel column chromatography using ethyl acetate / hexane as eluent. $R_f = 0.3$ in solvent = 5% ethyl acetate/hexane. ¹H NMR (400MHz, CDCl₃): δ_H 6.69-6.42 (m, 21H), 4.99-4.96 (s, 12H), 4.28 (s, 2H), 3.94 (t, $J=6.4$ Hz, 16H), 1.77 (quint, $J=6.4$ Hz, 16H), 1.49-1.31 (m, 48H), 0.91 (t, $J=6.8$ Hz, 24H). ¹³C NMR (100MHz, CDCl₃): δ_C 160.62, 160.28, 139.10, 139.02, 137.77, 107.31, 106.47, 105.83, 101.90, 101.81, 100.94, 77.16, 72.04, 70.26, 70.17, 68.16, 54.90, 31.70, 29.34, 25.84, 22.71, 14.15. MALDI-ToF (M+K⁺): 1610.05.

Synthesis of Compound 5a



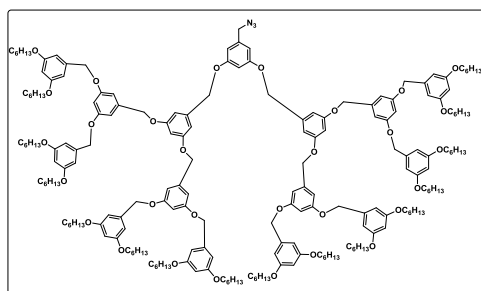
The compound **5a** was synthesized from general procedure D, starting from 3,5-dihydroxybenzyl alcohol (0.20 g, 1.4 mmol), **4b** (5.1 g, 3.1 mmol), K₂CO₃ (0.48 g, 3.4 mmol), crown ether (0.05 g, 0.2 mmol) in acetone. The product was obtained as a yellowish liquid (4.1 g, 1.2 mmol, 89%) after purification by silica gel column chromatography using ethyl acetate / hexane as eluent. $R_f = 0.53$ in 25% ethyl acetate / hexane. ¹H NMR (400MHz, CDCl₃): δ_H 6.72-6.45 (m, 45H), 4.97 (s, 28H), 4.61 (s, 2H), 3.98 (t, $J=6.4$ Hz, 32H), 1.81 (quint, $J=6.4$ Hz, 32H), 1.53-1.31 (m, 96H), 0.97 (t, $J=7.2$ Hz, 48H). MALDI-ToF MS (M+K⁺): 3235.18.

Synthesis of Compound 5b



The compound **5b** was synthesized from general procedure C, starting from **5a** (4.0 g, 1.2 mmol), CBr_4 (0.58 g, 1.7 mmol) and PPh_3 (0.45 g, 1.7 mmol) in DCM. The product was obtained as a colorless liquid (3.0 g, 0.9 mmol, 74%) after purification by silica gel column chromatography using ethyl acetate / hexane as eluent, $R_f = 0.4$ in 5% ethyl acetate / hexane. $^1\text{H NMR}$ (400MHz, CDCl_3): δ_{H} 6.70-6.42 (m, 45H), 4.96 (s, 28H), 4.41 (s, 2H), 3.94 (t, $J=6.4\text{Hz}$, 32H), 1.77 (quint, $J=6.4\text{Hz}$, 32H), 1.53-1.29 (m, 99H), 0.92 (t, $J=7.2\text{Hz}$, 48H). $^{13}\text{C NMR}$ (100MHz, CDCl_3): δ_{C} 160.54, 160.18, 160.13, 160.00, 139.20, 139.14, 138.99, 106.40, 105.73, 101.63, 100.84, 77.16, 70.14, 70.04, 68.05, 31.66, 29.29, 25.80, 22.67, 14.11. **MALDI-ToF MS** ($\text{M}+\text{K}^+$): 3298.08.

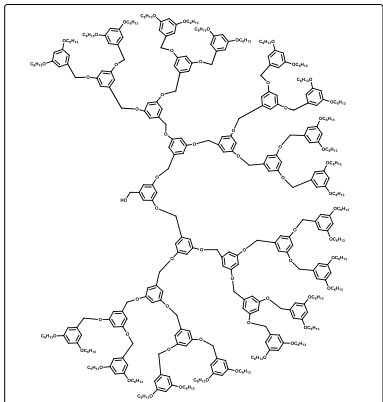
Synthesis of Compound 5c



The compound **5c** was synthesized from general procedure E, starting from **5b** (1.0 g, 0.3 mmol), NaN_3 (0.2 g, 3.0 mmol) in DMSO. The product was obtained as a yellowish liquid (0.58 g, 0.2 mmol, 74%) after purification by silica gel column chromatography using ethyl acetate / hexane as eluent, $R_f = 0.3$ in 5% ethyl acetate/hexane. $^1\text{H NMR}$ (400MHz, CDCl_3): δ_{H} 6.74-6.48 (m, 45H), 4.98 (s, 28H), 4.26 (s, 2H), 3.98 (t, $J=6.4\text{Hz}$, 32H), 1.91-1.79 (m, $J=6.4\text{Hz}$, 32H), 1.53-1.41 (m, 99H), 0.99 (t, $J=7.2\text{Hz}$, 48H). $^{13}\text{C NMR}$ (100MHz, CDCl_3): δ_{C} 160.58, 160.22, 139.20, 139.01, 107.36, 106.45, 105.78, 101.67,

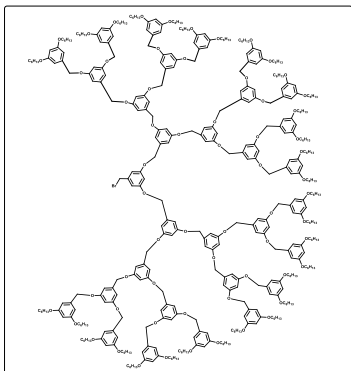
100.89, 77.16, 70.20, 70.11, 68.10, 31.68, 29.32, 25.82, 22.69, 14.13. **MALDI-ToF** (M+Na⁺): 3244.12.

Synthesis of Compound 6a



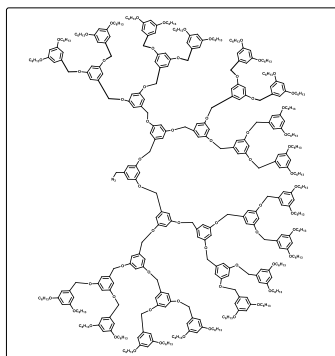
The compound **6a** was synthesized from general procedure D, starting from 3,5-dihydroxybenzyl alcohol (0.1 g, 0.7 mmol), **5b** (4.8 g, 1.4 mmol), K₂CO₃ (0.21 g, 1.5 mmol), crown ether (0.05 g, 0.18 mmol) in acetone. The product was obtained as a yellowish liquid (3 g, 0.4 mmol, 89%) after purification by silica gel column chromatography using ethyl acetate / hexane as eluent. *R_f* = 0.53 in 25% ethyl acetate / hexane. ¹H NMR (400MHz, CDCl₃): δ_H 6.66-6.38 (m, 84H), 5.00-4.96 (3, 53H), 4.54 (s, 2H), 3.89 (t, *J*=6.4Hz, 60H), 1.77 (quint, *J*=6.4Hz, 58H), 1.50-1.26 (m, 194H), 0.88 (t, *J*=7.2Hz, 95H). ¹³C NMR (100MHz, CDCl₃): δ_C 160.62, 160.27, 139.22, 139.05, 106.56, 105.86, 100.95, 77.16, 70.25, 68.17, 31.72, 29.35, 25.86, 22.73, 14.18. **MALDI-ToF MS** (M+K⁺): 6528.56.

Synthesis of Compound 6b



The compound **6b** was synthesized from general procedure C, starting from **6a** (2.9 g, 0.4 mmol), CBr₄ (0.23 g, 0.7 mmol) and PPh₃ (0.18 g, 0.7 mmol) in DCM. The product was obtained as a colorless liquid (1.8 g, 0.27 mmol, 62%) after purification by silica gel column chromatography using ethyl acetate / hexane as eluent, $R_f = 0.4$ in 5% ethyl acetate / hexane. ¹H NMR (400MHz, CDCl₃): δ_H 6.70-6.41 (m, 84H), 4.96-4.94 (m, 52H), 4.38 (s, 2H), 3.93 (t, $J=6.4$ Hz, 60H), 1.75 (quint, $J=6.4$ Hz, 58H), 1.45-1.34 (m, 185H), 0.92 (t, $J=7.2$ Hz, 94H). ¹³C NMR (100MHz, CDCl₃): δ_C 160.59, 160.23, 139.20, 139.02, 106.52, 105.82, 101.70, 100.91, 77.16, 70.22, 70.15, 68.13, 31.70, 29.34, 25.84, 22.72, 14.16. MALDI-ToF MS (M+K⁺): 6592.45.

Synthesis of Compound **6c**



The compound **6c** was synthesized from general procedure E, starting from **6b** (0.8 g, 0.1 mmol) and NaN₃ (0.15 g, 2.2 mmol) in DMSO. The product was obtained as a yellowish liquid (0.7 g, 0.1 mmol, 90%) after purification by silica gel column chromatography using ethyl acetate / hexane as eluent, $R_f = 0.3$ in solvent = 5% ethyl acetate/hexane; ¹H NMR (400MHz, CDCl₃): δ_H 6.68-6.38 (m, 84H), 5.04-4.73 (m, 50H), 4.14 (s, 2H), 3.93 (t, $J=6.4$ Hz, 60H), 1.73 (quint, $J=6.4$ Hz, 64H), 1.42-1.25 (m, 185H), 0.89 (t, $J=7.2$ Hz, 110H). ¹³C NMR (100MHz, CDCl₃): δ_C 160.60, 160.25, 139.03, 106.57, 105.84, 100.91, 77.16, 70.77, 70.23, 68.15, 31.72, 29.84, 29.35, 25.86, 22.74, 14.19. MALDI-ToF MS (M+K⁺): 6537.02.

2.4.2.3. Synthesis of Macromolecular AABP

2.4.2.3.1. General Procedure for Click Reaction - Procedure F

Hydrophobic azide (1 eq) and hydrophilic alkyne (1 eq) were dissolved in degassed THF or THF/DCM (1:1) and stirred until a clear solution was obtained, then degassed water was added and stirred vigorously for 10 more minutes. Freshly prepared 1M sodium ascorbate (0.05 eq) and 1M CuSO₄ (0.1 eq) were added to the reaction mixture at least thrice in intervals of 45 minutes and allowed to react for 16 hours at RT. Upon completion, the reaction mixture was extracted in DCM. The combined organic layer was dried over Na₂SO₄ and concentrated under reduced pressure to get crude product which was purified using reverse-phase chromatography using ACN / H₂O system first (to remove unreacted diphosphonate ester) and CHCl₃ later (to elute click product) followed by normal phase chromatography using MeOH / DCM solvent system.

Note:

MALDI-ToF MS analysis of **2.3.2.3.1.** showed the presence of unreacted diphosphonate ester and click product. Surprisingly both the components were at the same *R_f* in normal-phase TLC with different solvent systems. However, a huge *R_f* difference was observed in reverse phase TLC with 30% ACN / H₂O. Hence, their separation was achieved using reverse phase chromatography. MALDI-ToF MS analysis of purified fractions showed the complete absence of unreacted diphosphonate ester.

2.4.2.3.2. General Procedure for Deprotection - Procedure G

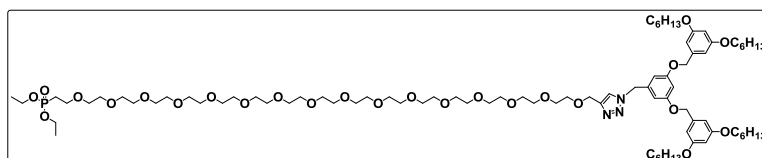
Diphosphonate ester (1 eq) was dissolved in DCM with stirring. Then oxalyl chloride (4 eq) was added dropwise at RT and allowed to react for 18 hours under stirring. Upon completion, excess of oxalyl chloride and DCM were removed under vacuum. Then water was added to the residue and stirred for 5 minutes. The resulting mixture was extracted thrice with DCM. The combined organic layer was dried over Na₂SO₄ and concentrated under vacuum to get crude product which was used for next step without further purification.

2.4.2.3.3. General Procedure for Fluorination - Procedure H

To the stirring solution of monophosphonate ester (1 eq) in DCM, DAST (4 eq) was added dropwise at RT and allowed to react for 4 hours. Excess of DAST and DCM were evaporated under reduced pressure. To the obtained residue, water was added and stirred for 2 more minutes to quench any residual DAST. The reaction mixture was then extracted thrice with DCM. The combined organic layer was dried over Na₂SO₄ and concentrated under vacuum to get the crude product. These final macromolecular AABPs were used for protein modification without further purification.

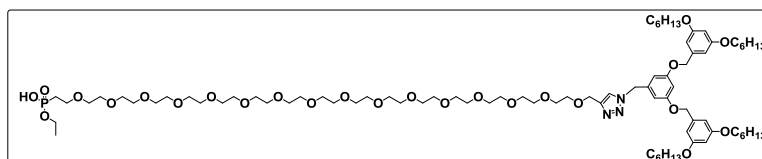
2.4.2.3.4. Synthesis of Macromolecular AABPs and Their Intermediates

Synthesis of compound 7a



The compound **7a** was prepared by general procedure F, starting from **3c** (0.300 g, 0.41 mmol), **1j** (0.370 g, 0.41 mmol), CuSO₄ (3.3 mg, 0.02 mmol), sodium ascorbate (8.2 mg, 0.04 mmol). The product was obtained as a pale yellow liquid (0.38 g, 0.25 mmol, 60%) after purification by reverse phase silica gel column chromatography followed by normal phase silica gel column chromatography using MeOH / DCM as eluent, *R_f* = 0.4 in 5% MeOH / DCM. ¹H NMR (400 MHz, CDCl₃): δ_H 7.51 (s, 1H), 6.55-6.40 (m, 9H), 5.29 (s, 2H), 4.89 (s, 4H), 4.65 (s, 2H), 4.17-4.04 (m, 4H), 3.92 (t, J=6.4Hz, 8H), 3.77-3.51 (m, 58H), 2.19-2.05 (m, 2H), 1.84-1.68 (m, 8H), 1.52-1.21 (m, 50H), 0.89 (m, 14H). ¹³C NMR (100MHz, CDCl₃): δ_C 160.63, 138.64, 107.29, 105.81, 102.11, 100.95, 70.65, 70.29, 69.84, 68.18, 65.22, 61.78, 61.71, 53.55, 31.68, 29.80, 29.32, 25.83, 22.79, 22.71, 16.64, 14.15. MALDI-ToF (M+K⁺): 1665.14.

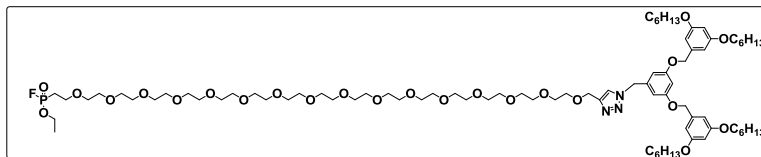
Synthesis of compound 7b



The compound **7b** was prepared by general procedure G, starting from **7a** (0.2 g, 0.13 mmol) and oxalyl chloride (0.1g, 0.8mmol) The product was obtained as a pale

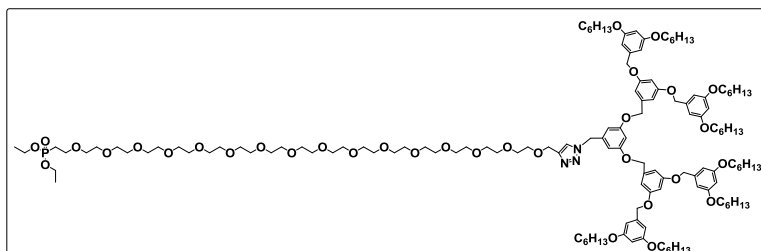
yellow liquid which was used further without purification. **MALDI-ToF** ($M+Na^+$): 1620.31.

Synthesis of compound 7c



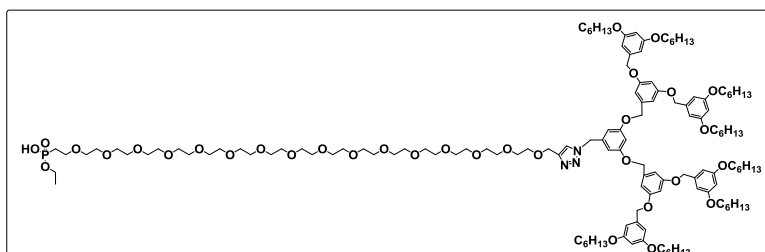
The compound **7c** was prepared by general procedure H, starting from **7b** (0.15 g, 0.1 mmol) and DAST (0.065 g, 0.4mmol). **¹⁹F NMR** (400MHz, CDCl₃): δ_F -59.91, -62.74. **MALDI-ToF** ($M+K^+$): 1638.31.

Synthesis of compound 8a



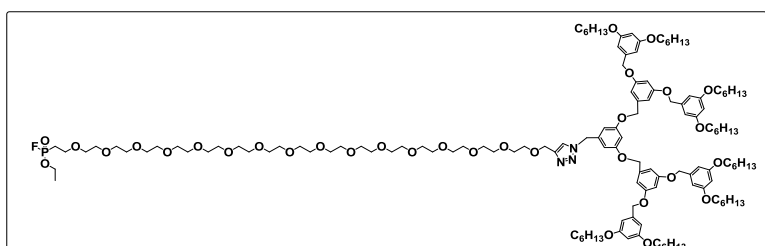
The compound **8a** was prepared by general procedure F, starting from **4c** (0.60g, 0.24 mmol), **1j** (0.22 g, 0.36 mmol), CuSO₄ (1.9mg, 0.012mmol), sodium ascorbate (5 mg, 0.02 mmol). The product was obtained as a pale yellow liquid (0.40 g, 0.16 mmol, 60%) after purification by reverse phase silica gel column chromatography followed by normal phase silica gel column chromatography using MeOH / DCM as eluent, $R_f = 0.4$ in 5% MeOH / DCM. **¹H NMR** (400 MHz, CDCl₃): δ_H 7.53 (s, 1H), 7.51 (s, 1H), 6.65-6.38 (m, 21H), 5.41 (s, 2H), 5.00-4.87 (m, 12H), 4.19-4.07 (4H), 3.92 (t, $J=6.4$ Hz, 16H), 3.76-3.51 (m, 61H), 2.18-2.10 (m, 2H), 1.85-1.63 (m, 22H), 1.45-1.25 (m, 71H), 0.89 (m, $J=7.2$ Hz, 30H). **¹³C NMR** (100MHz, CDCl₃): δ_C 160.64, 160.30, 138.98, 106.43, 105.88, 70.68, 70.59, 70.30, 68.20, 53.57, 31.72, 29.74, 29.35, 25.86, 22.74, 14.18. **MALDI-ToF** ($M+K^+$): 2490.70.

Synthesis of compound 8b



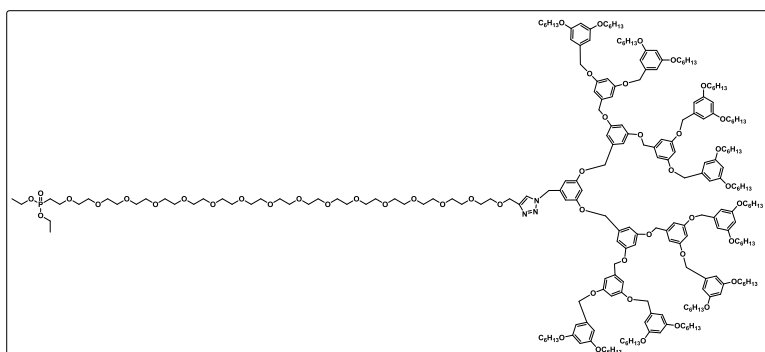
The compound **8b** was prepared by general procedure G, starting from **8a** (0.2 g, 0.08 mmol) and oxalyl chloride (0.05g, 0.3 mmol). The product was obtained as a pale yellow liquid which was used further without purification. **MALDI-ToF** (M+K⁺): 2463.11.

Synthesis of compound 8c



The compound **8c** was prepared by general procedure H, starting from **8b** (0.2 g, 0.08 mmol) and DAST (0.05 g, 0.3 mmol). ¹⁹F NMR (400MHz, CDCl₃): δ_F -59.91, -62.74. **MALDI-ToF** (M+K⁺): 2467.73

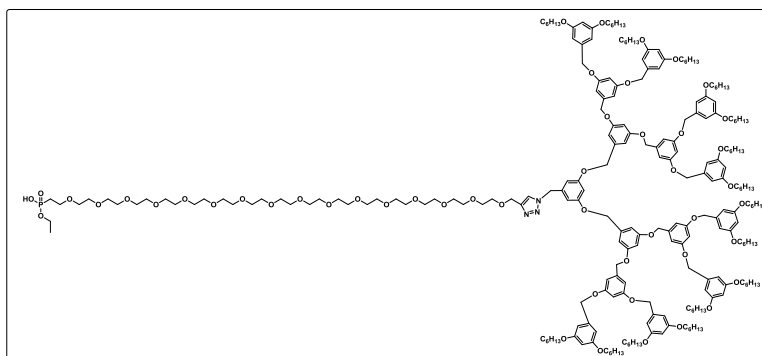
Synthesis of compound 9a



The compound **9a** was prepared by general procedure F, starting from **5c** (1 g, 0.3 mmol), **1j** (0.300 g, 0.3 mmol), CuSO₄ (2.6 mg, 0.016mmol), sodium ascorbate (4.8mg, 0.02mmol). The product was obtained as a pale yellow liquid (0.9 g, 0.2 mmol, 66%) after purification by reverse phase silica gel column chromatography followed by normal phase silica gel column chromatography using MeOH / DCM as eluent, *R_f* = 0.4 in 5% MeOH / DCM.; ¹H NMR (400 MHz, CDCl₃): δ_H 7.50 (s, 1H), 6.74-6.34 (m, 45H), 5.27 (s, 14H),

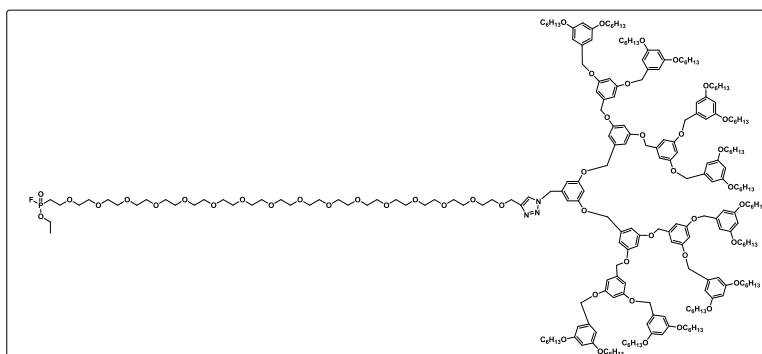
5.00-4.84 (m, 26H), 4.16-4.03 (m, 4H), 3.91 (t, $J=6.4$ H, 32H), 3.71-3.54 (m, 67H), 2.13-2.05 (m, 2H), 1.78-1.67 (m, 32H), 1.49-1.23 (m, 110H), 0.89 (t, $J=6.4$ H, 52H).). ^{13}C NMR (100MHz, CDCl_3): δ_{C} 160.58, 160.22, 139.12, 138.96, 106.52, 105.82, 101.65, 100.87, 70.65, 70.54, 70.23, 70.16, 69.81, 68.12, 31.68, 29.31, 25.82, 22.70, 16.56, 16.50, 14.15. **MALDI-ToF** ($\text{M}+\text{K}^+$): 4137.21.

Synthesis of compound 9b



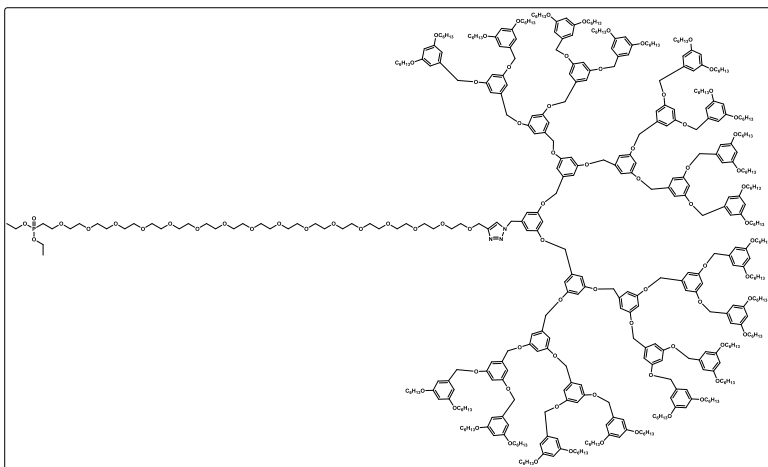
The compound **9b** was prepared by general procedure G, starting from **9a** (0.186 g, 0.04 mmol), oxalyl chloride (0.03 g, 0.18 mmol). **MALDI-ToF** ($\text{M}+\text{K}^+$): 4094.11

Synthesis of compound 9c



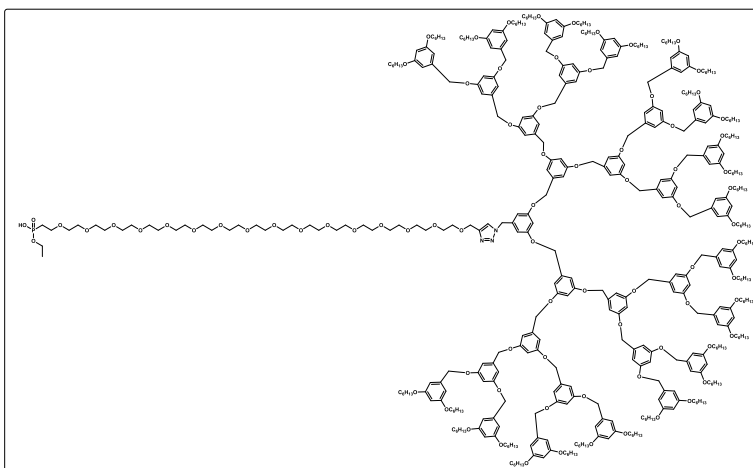
The compound **9c** was prepared by general procedure H, starting from **9b** (0.15 g, 0.03 mmol), DAST (0.03 g, 0.14 mmol); ^{19}F NMR (400MHz, CDCl_3): δ_{F} -59.91, -62.74. **MALDI-ToF** ($\text{M}+\text{K}^+$): 4110.05.

Synthesis of compound 10a



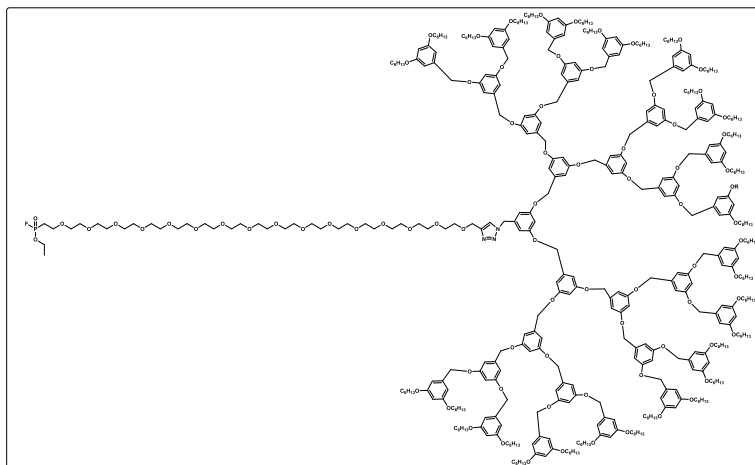
The compound **10a** was prepared by general procedure F, starting from **6c** (0.2 g, 0.031 mmol), **1j** (0.027 g, 0.31 mmol), CuSO₄ (0.2 mg, 0.001mmol), sodium ascorbate (0.4mg, 0.002mmol). The product was obtained as a pale yellow liquid (0.1 g, 0.02 mmol, 59%) after purification by reverse phase silica gel column chromatography followed by normal phase silica gel column chromatography using MeOH / DCM as eluent, $R_f = 0.4$ in 5% MeOH / DCM. ¹H NMR (400 MHz, CDCl₃): δ 7.51 (s, 1H), 6.67-6.37 (m, 84H), 5.00-4.80 (m, 54H), 4.12-4.06 (m, 4H), 3.94-3.83 (m, 67H), 3.71-3.54 (m, 56H), 2.21-2.07 (m, 2H), 1.82-1.62 (m, 60H), 1.45-1.23(m, 260H) 0.98-0.76 (m, 102H). ¹³C NMR (100MHz, CDCl₃): δ_c 160.60, 160.25, 139.03, 106.57, 105.84, 100.91, 77.16, 70.61, 70.19, 68.15, 31.72, 29.84, 29.81, 29.35, 25.86, 22.74, 14.19. MALDI-ToF (M+K⁺): 7435.12

Synthesis of compound 10b



The compound **10b** was prepared by general procedure G, starting from **10a** (0.18 g, 0.02 mmol), oxalyl chloride (0.002 g, 0.09 mmol). MALDI-ToF (M+K⁺): 7400.37

Synthesis of compound **10c**



The compound **10c** was prepared by general procedure HH, starting from **10b** (0.1 g, 0.01 mmol), DAST (0.008 g, 0.054mmol); ^{19}F NMR (400MHz, CDCl_3): δ_{F} -59.91, -62.74.

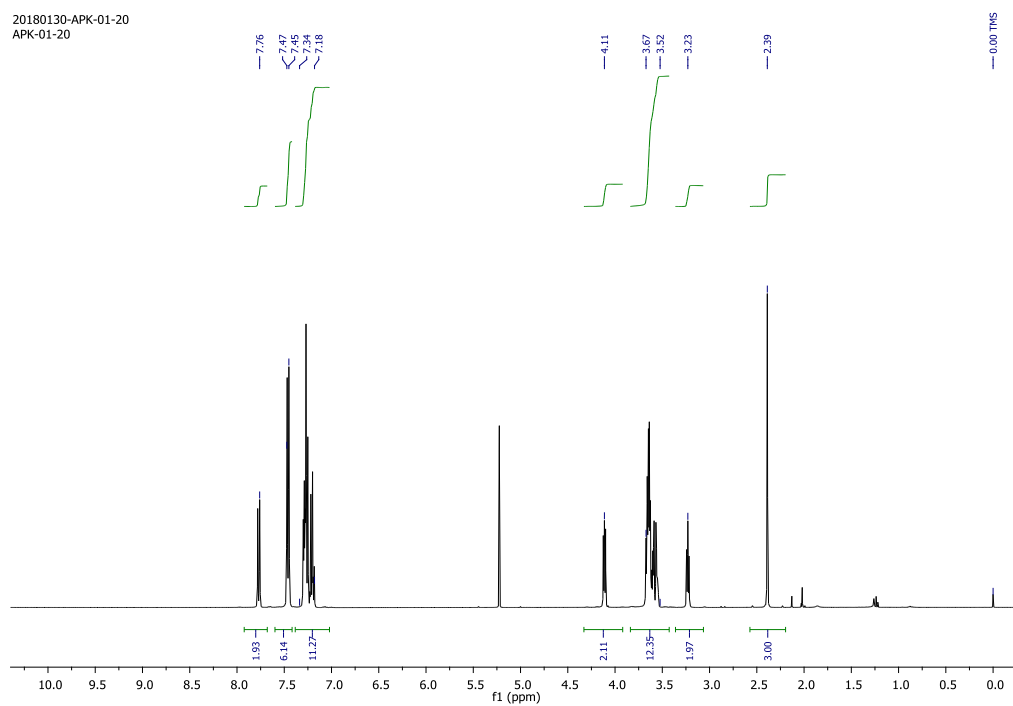
2.5. References

1. Mateu, M. G. *Arch. Biochem. Biophys.* **2013**, *531*, 65–79.
2. Stock, D.; Leslie, A. G. W.; Walker, J. E. *Science* **1999**, *286*, 1700-1705
3. Massover, W. H. *Micron* **1993**, *24*, 389–437.
4. Kerfeld, C. A.; Heinhorst, S.; Cannon, G. C. *Annu. Rev. Microbiol.* **2010**, *64*, 391–408
5. Seeman, N. C. *Nature* **2003**, *421*, 427–431.
6. Guo, P. X. *Nat. Nanotechnol* **2010**, *5*, 833–842.
7. Ulijn, R. V.; Smith, A. M. *Chem. Soc. Rev.* **2008**, *37*, 664–675.
8. Luo, Q.; Hou, C.; Bai, Y.; Wang, R.; Liu, J. *Chem. Rev.* **2016**, *116*, 13571-13632.
9. Lai, Y.-T.; Tsai, K.-L.; Sawaya, M. R.; Asturias, F. J.; Yeates, T. O. *J. Am. Chem. Soc.* **2013**, *135*, 7738–7743
10. King, N. P.; Bale, J. B.; Sheffler, W.; McNamara, D. E.; Gonen, S.; Gonen, T.; Yeates, T. O.; Baker, D. *Nature* **2014**, *510*(7503), 103–108.
11. Worsdorfer, B.; Woycechowsky, K. J.; Hilvert, D. *Science* **2011**, *331*, 589–592.
12. Yeates, T. O. *Annu. Rev. Biophys.* **2017**, *46*, 23–42
13. Huang, P. S.; Boyken, S. E.; Baker, D. *Nature* **2016**, *537*, 320-327.
14. Abuchowski, A.; van Es, T.; Palczuk, N. C.; Davis, F. F. *J. Biol. Chem.* **1977**, *252*, 3578–3581.

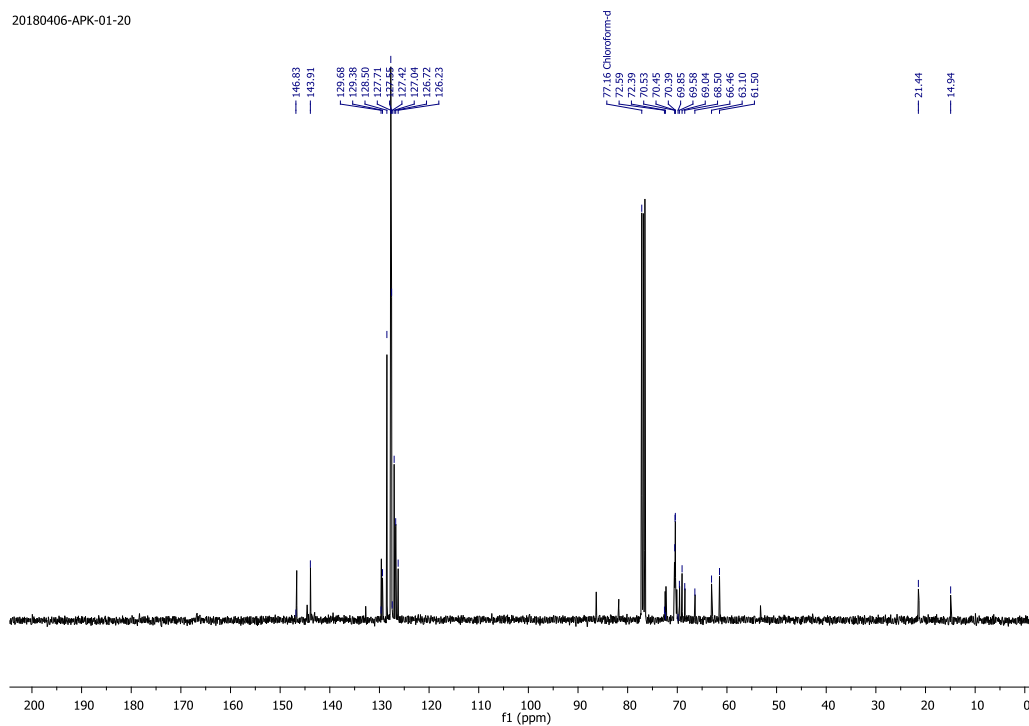
15. Abuchowski, A.; McCoy, J. R.; Palczuk, N. C.; van Es, T.; Davis, F. F. *J. Biol. Chem.* **1977**, *252*, 3582–3586.
16. Hannink, J. M.; Cornelissen, J. J. L. M.; Farrera, J. A.; Foubert, P.; De Schryver, F. C.; Sommerdijk, N. A. J. M.; Nolte, R. J. M. *Angew. Chem., Int. Ed.* **2001**, *40*, 4732-4734.
17. Velonia, K.; Rowan, A. E.; Nolte, R. J. *J. Am. Chem. Soc.* **2002**, *124*, 4224-4225.
18. Boerakker, M. J.; Hannink, J. M.; Bomans, P. H. H.; Frederik, P. M.; Nolte, R. J. M.; Meijer, E. M.; Sommerdijk, N. A. J. M. *Angew. Chem., Int. Ed.* **2002**, *41*, 4239-4241.
19. Reynhout, I. C.; Cornelissen, J. J. L. M.; Nolte, R. J. M. *J. Am. Chem. Soc.* **2007**, *129*, 2327
20. Dirks, A. J. T.; Berkel, S. S. V.; Hatzakis, N. S.; Opsteen, J. A.; Delft, F. L.V.; Cornelissen, J. J. L. M.; Rowan, A. E.; Hest, J. C. M. V.; Rutjes, F. P. J. T.; Nolte, R. J. M. *Chem. Commun.* **2005**, *41*, 4172–4174.
21. Chen, G.; Hoffman, A. S. *Bioconjugate Chem.* **1993**, *4*, 509-514.
22. Kulkarni, S.; Schilli, C.; Müller, A. H.; Hoffman, A. S.; Stayton, P. S. *Bioconjugate Chem.* **2004**, *15*, 747-753.
23. Wan, X. J.; Liu, S. Y. *Macromol. Rapid Commun.* **2010**, *31*, 2070-2076.
24. Park, W. M.; Champion, J. A. *J. Am. Chem. Soc.* **2014**, *136*, 17906-17909.
25. Hassouneh, W.; Fischer, K.; MacEwan, S.R.; Branscheid, R.; Fu, C.L.; Liu, R.; Schmidt, M.; Chilkoti, A. *Biomacromolecules* **2012**, *13*, 15981605.
26. Thomas, C. S.; Glassman, M. J.; Olsen, B. D. *ACS Nano* **2011**, *5*, 5697–5707
27. Lam, C. N.; Olsen, B. D. *Soft Matter* **2013**, *9*, 2393–2402.
28. Huang, A.; Qin, G.; Olsen, B. D. *ACS Appl. Mater. Interfaces* **2015**, *7*(27), 14660–14669.
29. Averick, S.; Karácsony, O.; Mohin, J.; Yong, X.; Moellers, N. M.; Woodman, B. F.; Zhu, W.; Mehl, R. A.; Balazs, A. C.; Kowalewski, T. *Angew. Chem., Int. Ed.* **2014**, *53*, 8050–8055.
30. Li, X., and Liu, C. C. *ChemBioChem.* **2014**, *15*, 2335–2341
31. Kluger, R.; Zhang, J.; *J. Am. Chem. Soc.* **2003**, *125*, 6070-6071 .
32. Van Baal, I.; Malda, H.; Synowsky, S. A.; van Dongen, J. L. J.; Hackeng, T. M.; Merks, M.; Meijer, E. W. *Angew. Chem., Int. Ed.* **2005**, *44*, 5052-5057.
33. Rendle, P. M.; Seger, A.; Rodrigues, J.; Oldham, N. J.; Bott, R. R.; Jones, J. B.; Cowan, M. M.; Davis, B. G. *J. Am. Chem. Soc.* **2004**, *126*, 4750.
34. Sato, M.; Furuike, T.; Sadamoto, R.; Fujitani, N.; Nakahara, T.; Niikura, K.; Monde, K.; Kondo, H.; Nishimura, S. I. *J. Am. Chem. Soc.* **2004**, *126*, 14013–14022.
35. Kostianinen, M. A.; Szilvay, G. R.; Smith, D. K.; Linder, M. B.; Ikkala, O. *Angew. Chem., Int. Ed.* **2006**, *45*, 3538–3542.

36. Kostianen, M. A.; Szilvay, G. R.; Lehtinen, J.; Smith, D. K.; Linder, M. B.; Urtti, A.; Ikkala, O. *ACS Nano* **2007**, *1*, 103–113I
37. Sandanaraj, B. S.; Reddy, M. M.; Bhandari, P. J.; Kumar, S. and Aswal, V. K. *Chem. Eur. J.* **2018**, *24(60)*, 16085-16096.
38. French, A. C.; Thompson, A. L.; Davis, B. G. *Angew. Chem., Int. Ed.* **2009**, *121*, 1274-1278.
39. Liu, Y., Patricelli, M. P. & Cravatt, B. F. *Proc. Natl Acad. Sci. USA*, **1999**, *96*, 14694-14699.
40. Hawker, C. J.; Fréchet, J. M. *J. Am. Chem. Soc.* **1990**, *112*, 76387647.
41. Wooley, K. L.; Hawker, C. J.; Fréchet, J. M. *J. Am. Chem. Soc.* **1991**, *113*, 4252-4261.
42. Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem., Int. Ed.* **2002**, *114*, 2708-2711.
43. Mourey, T. H.; Turner, S. R.; Rubinstein, M.; Fréchet, J. M. J.; Hawker, C. J.; Wooley, K. L. *Macromolecules* **1992**, *25*, 2401-2406 *Macromolecules* **1992**, *25.90*, 2401-2406.

2.5. Appendix-I: Characterization data of synthesized compounds

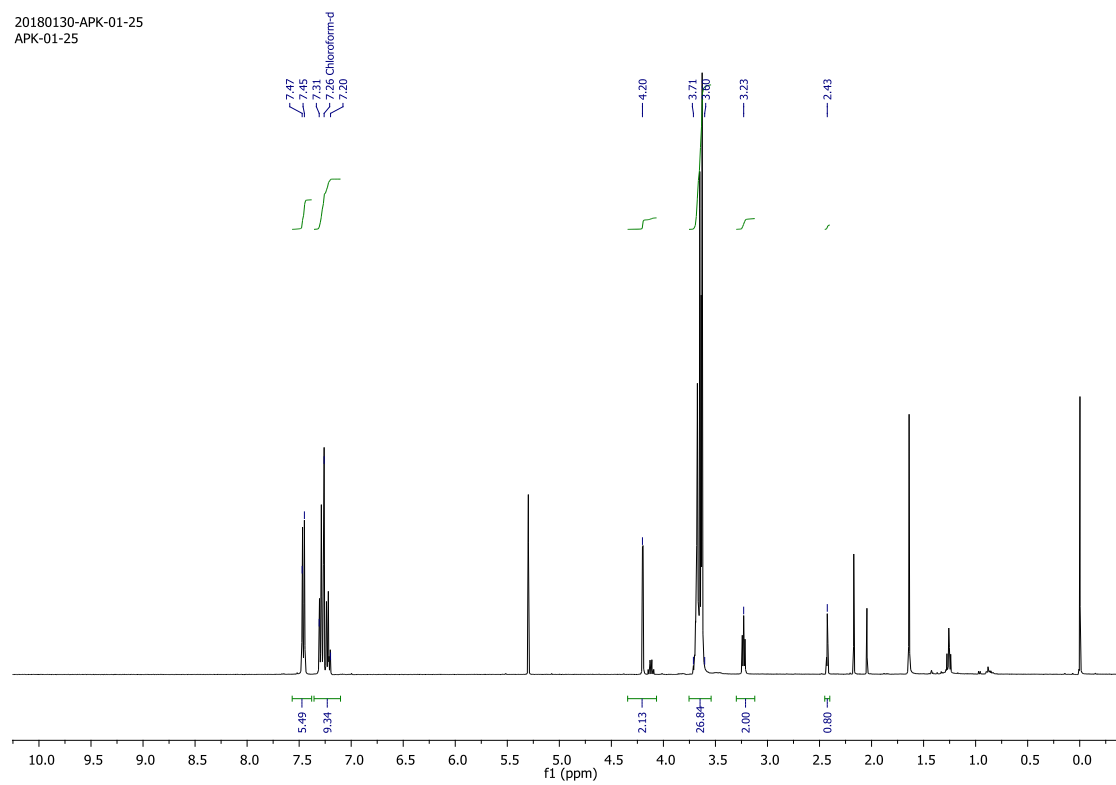


¹H NMR spectrum of compound 1a



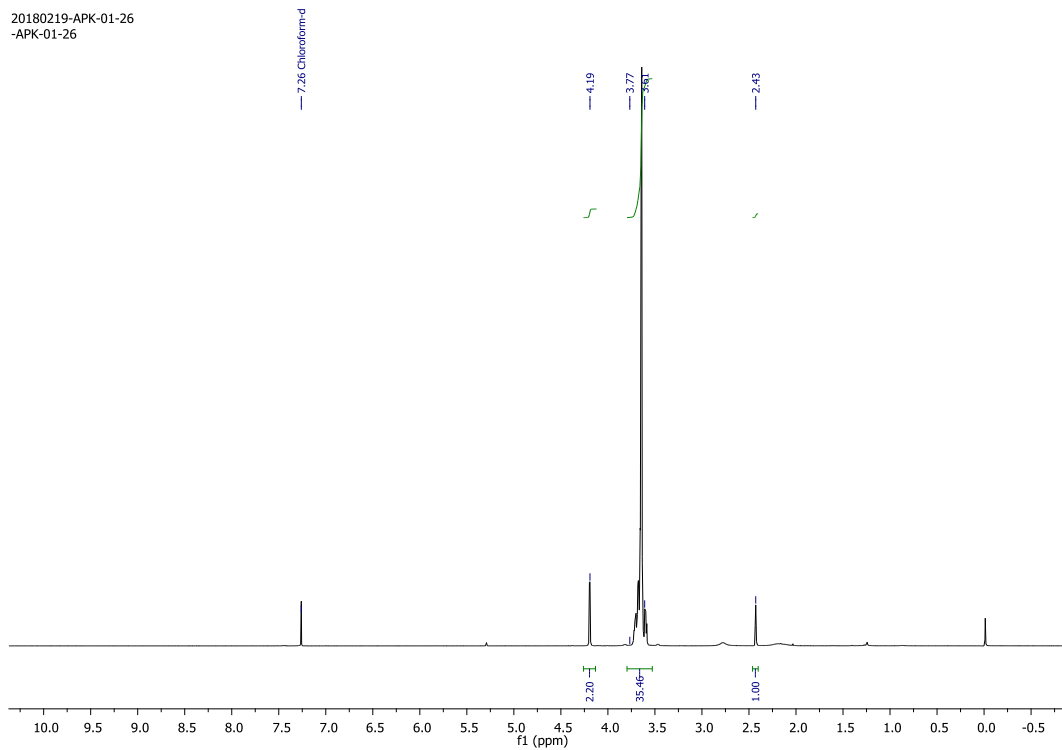
¹³C NMR spectrum of compound 1a

20180130-APK-01-25
APK-01-25



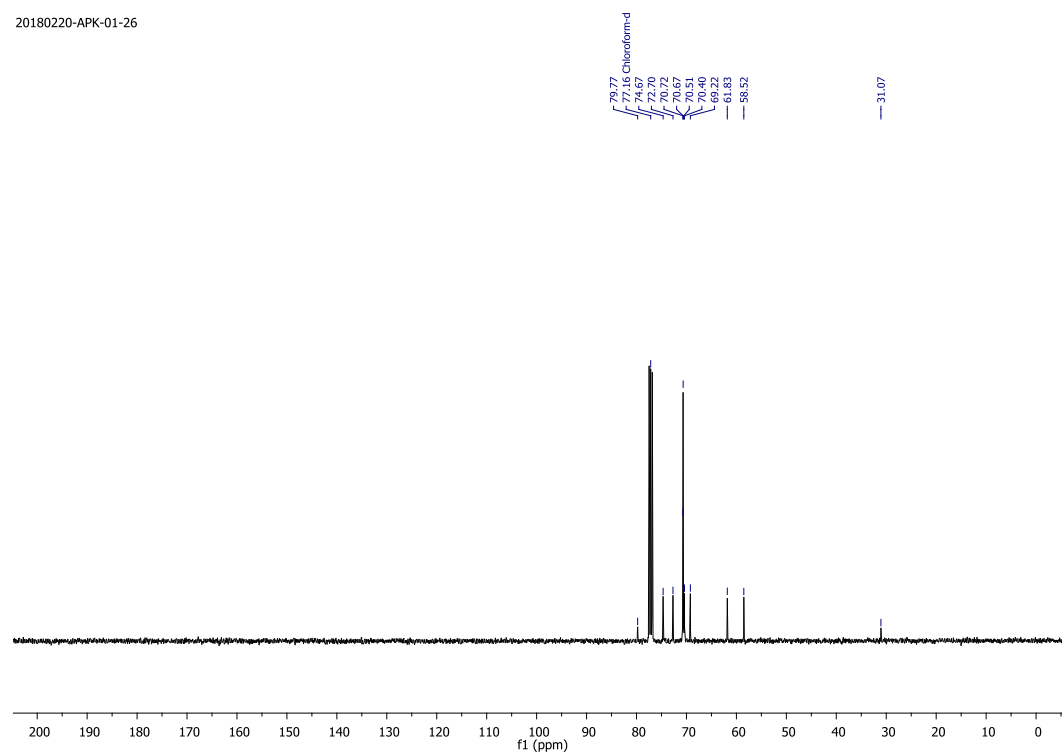
¹H NMR spectrum of compound 1c

20180219-APK-01-26
-APK-01-26



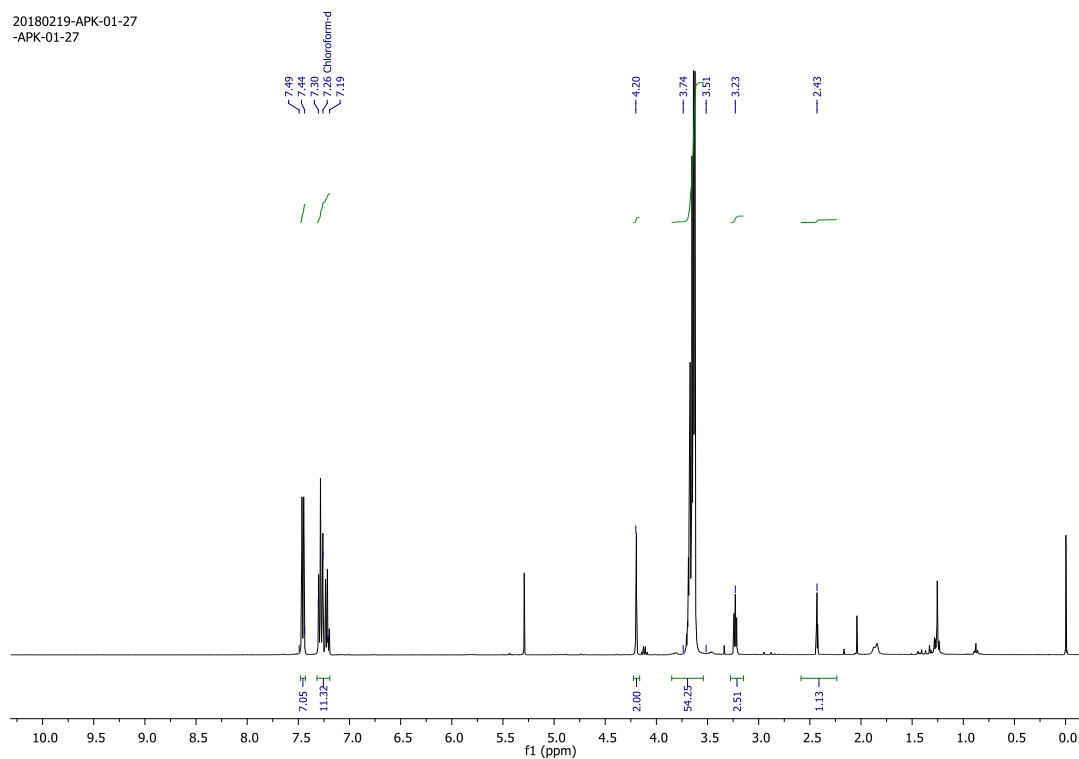
¹H NMR spectrum of compound 1d

20180220-APK-01-26



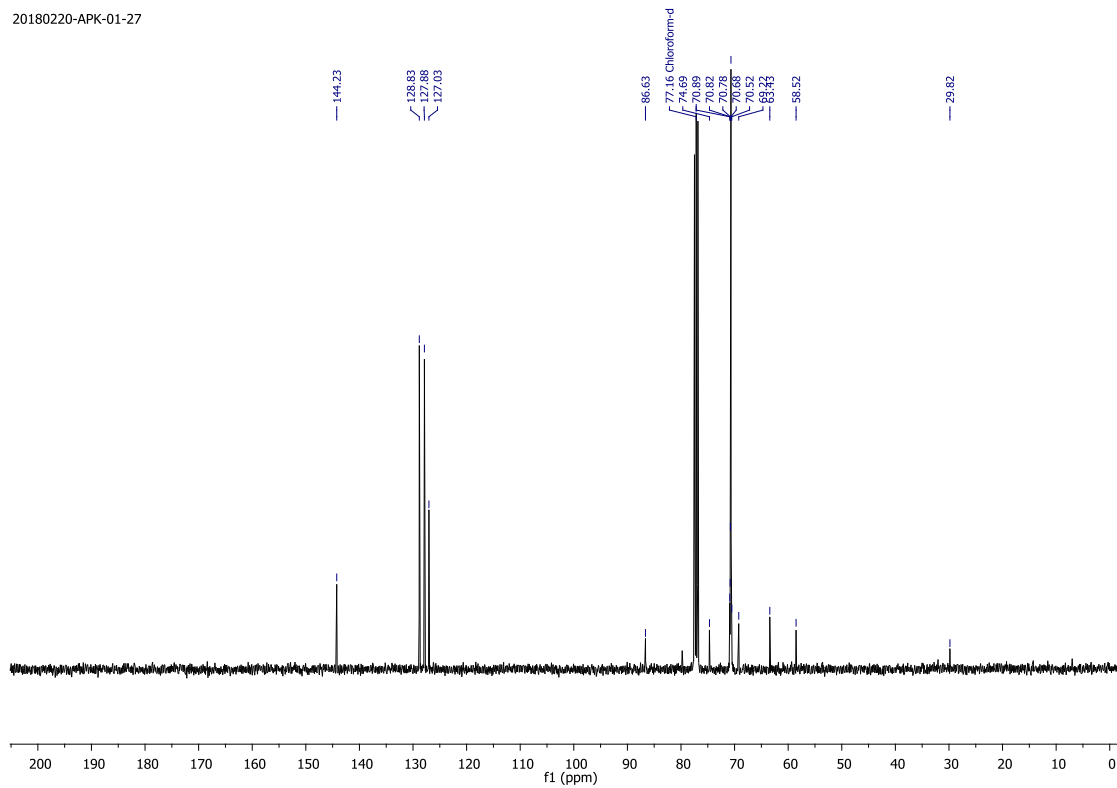
¹³C NMR spectrum of compound 1d

20180219-APK-01-27
-APK-01-27



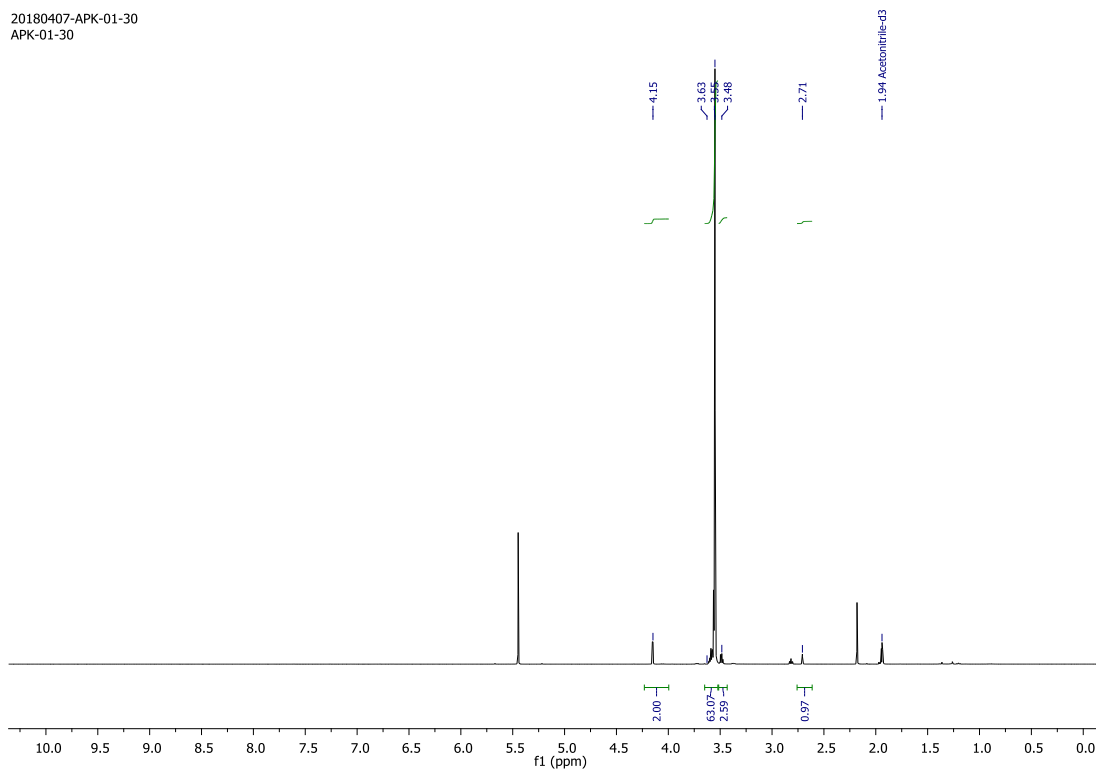
¹H NMR spectrum of compound 1e

20180220-APK-01-27



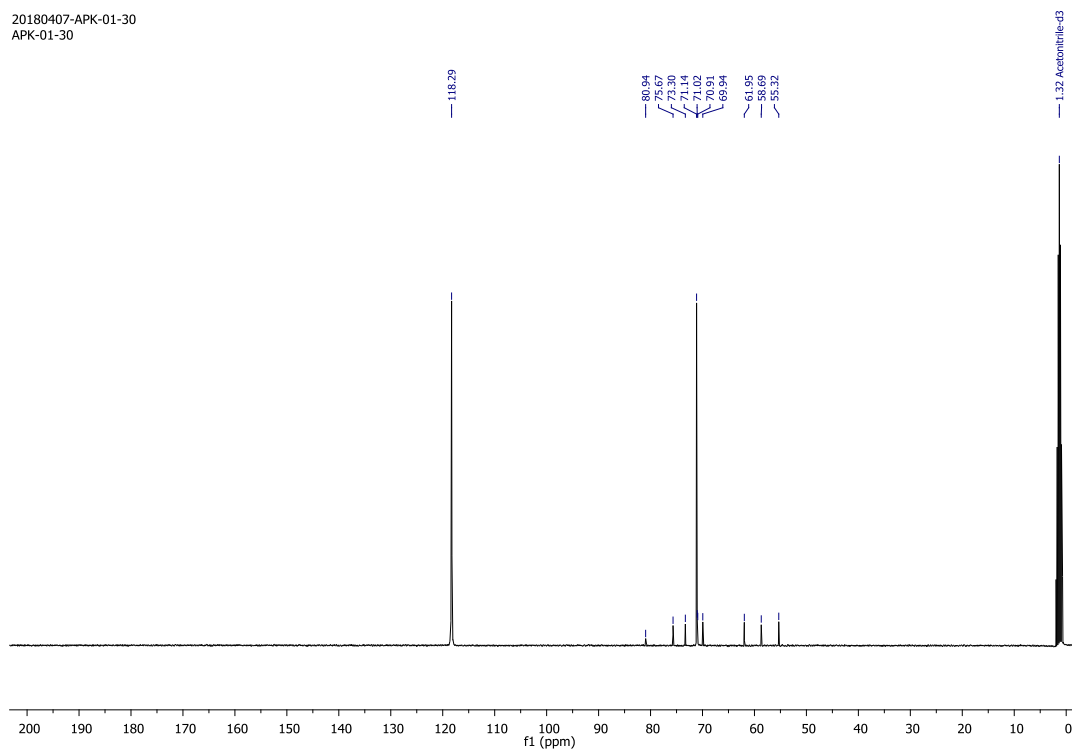
¹³C NMR spectrum of compound 1e

20180407-APK-01-30
APK-01-30

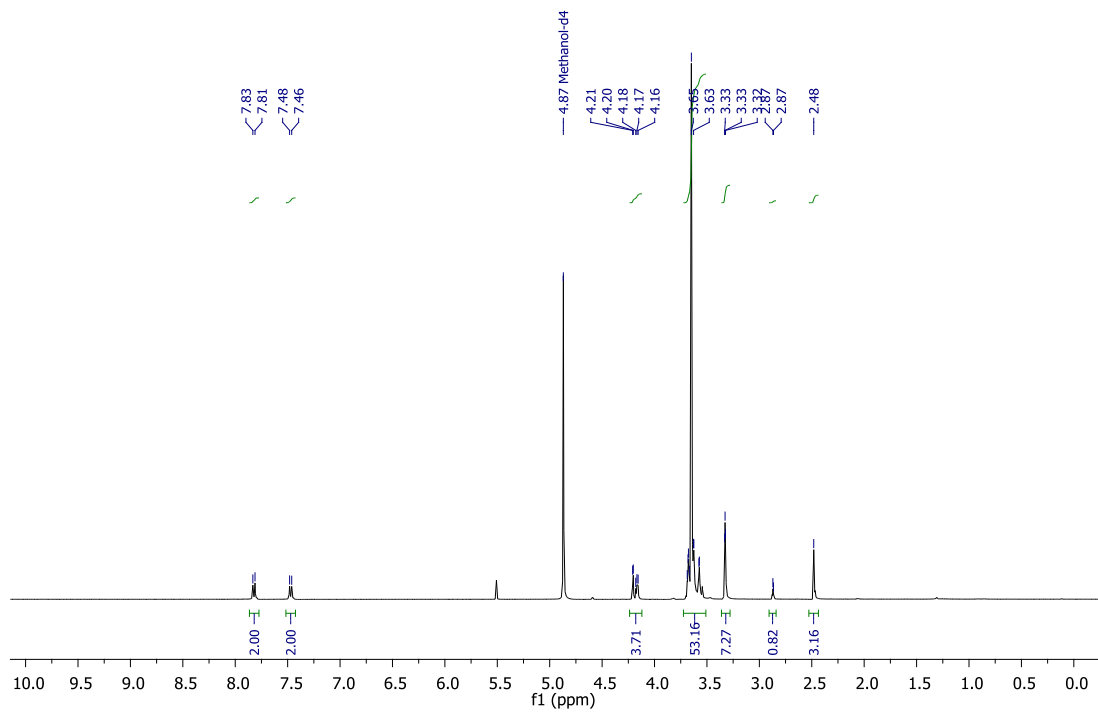


¹H NMR spectrum of compound 1f

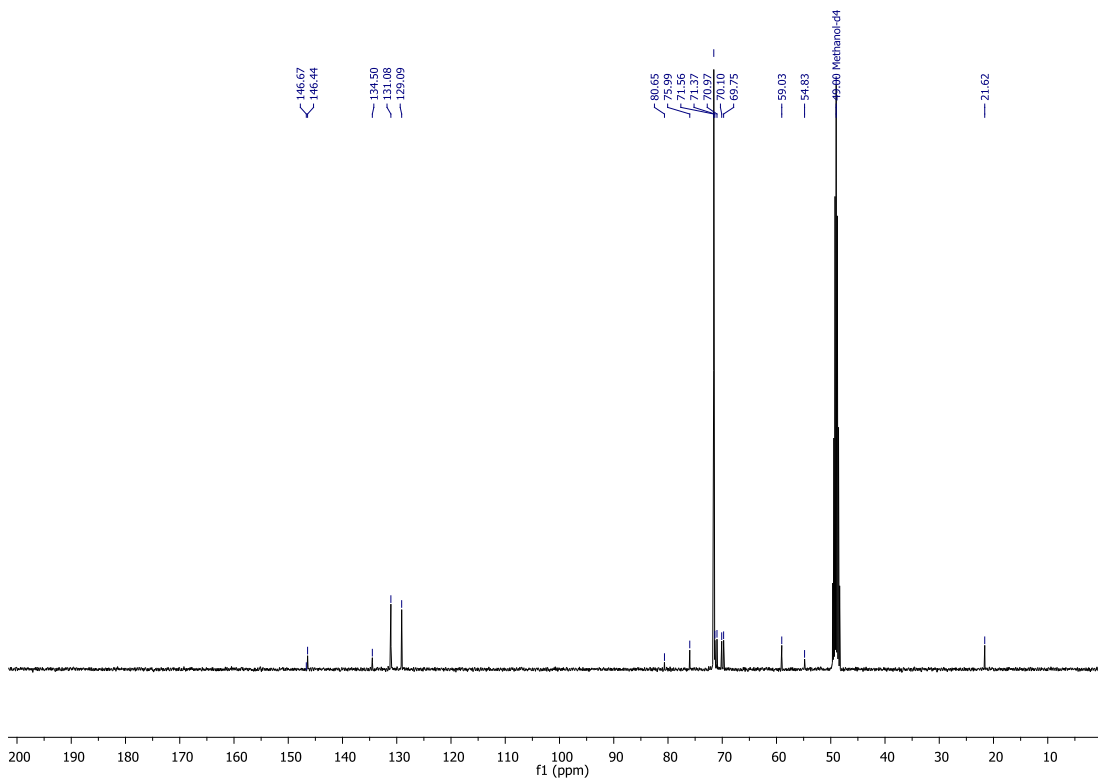
20180407-APK-01-30
APK-01-30



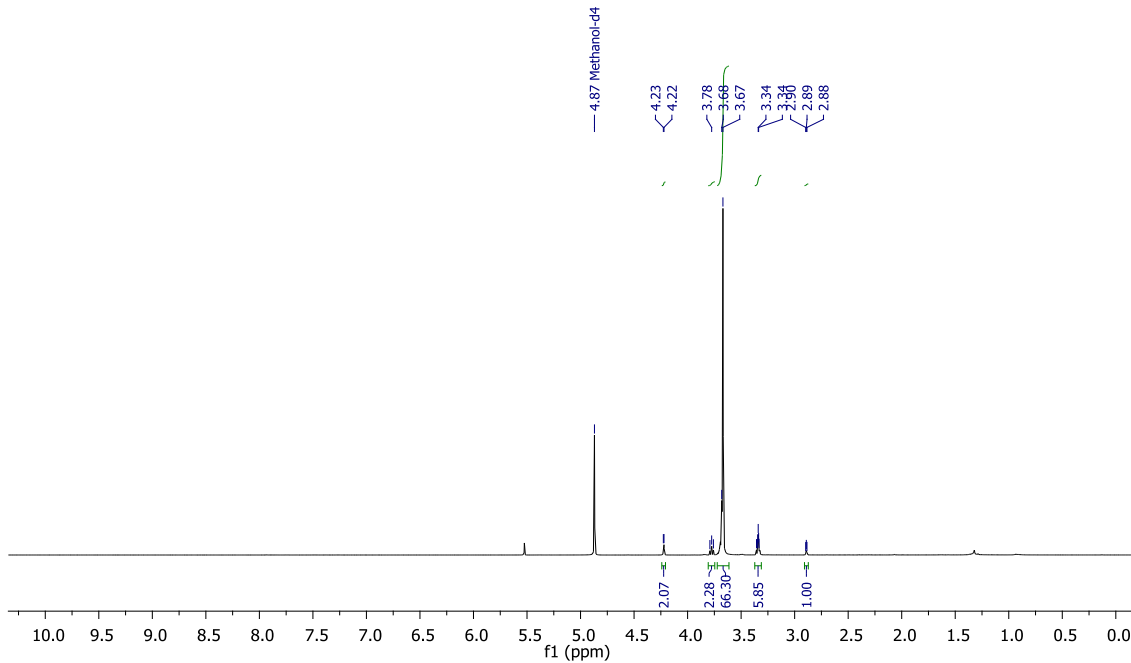
¹³C NMR spectrum of compound 1f



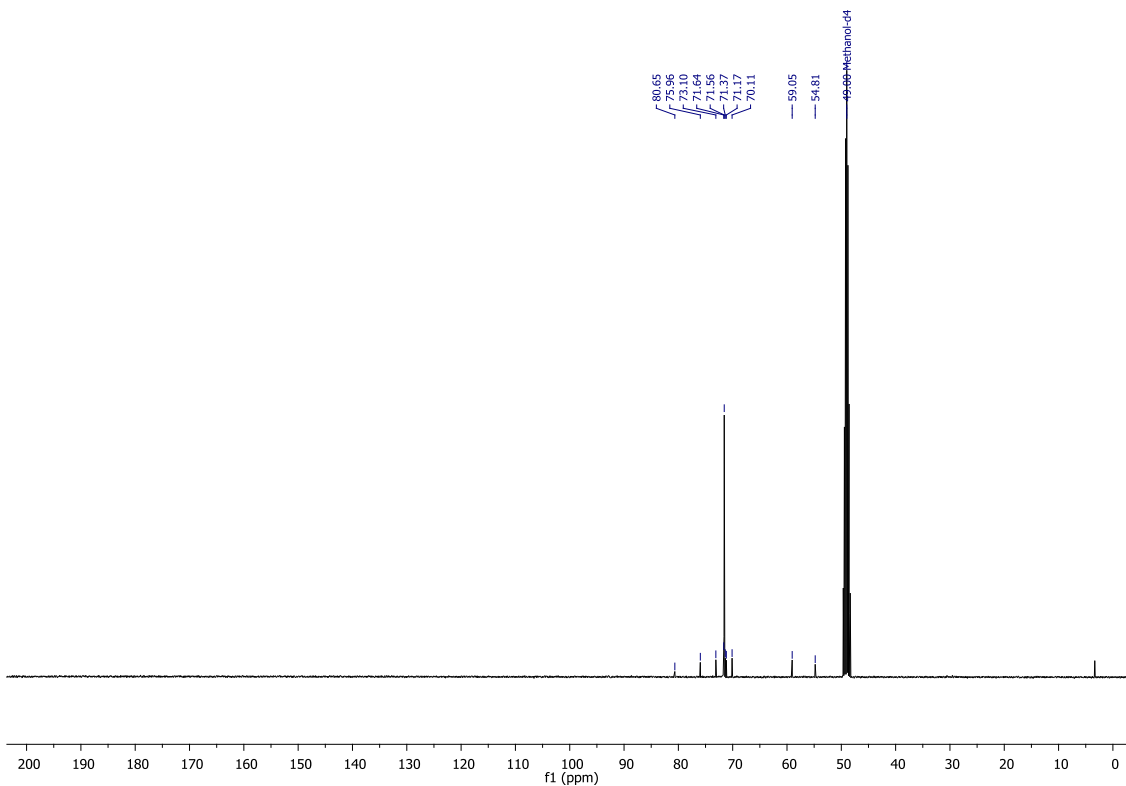
¹H NMR spectrum of compound 1h



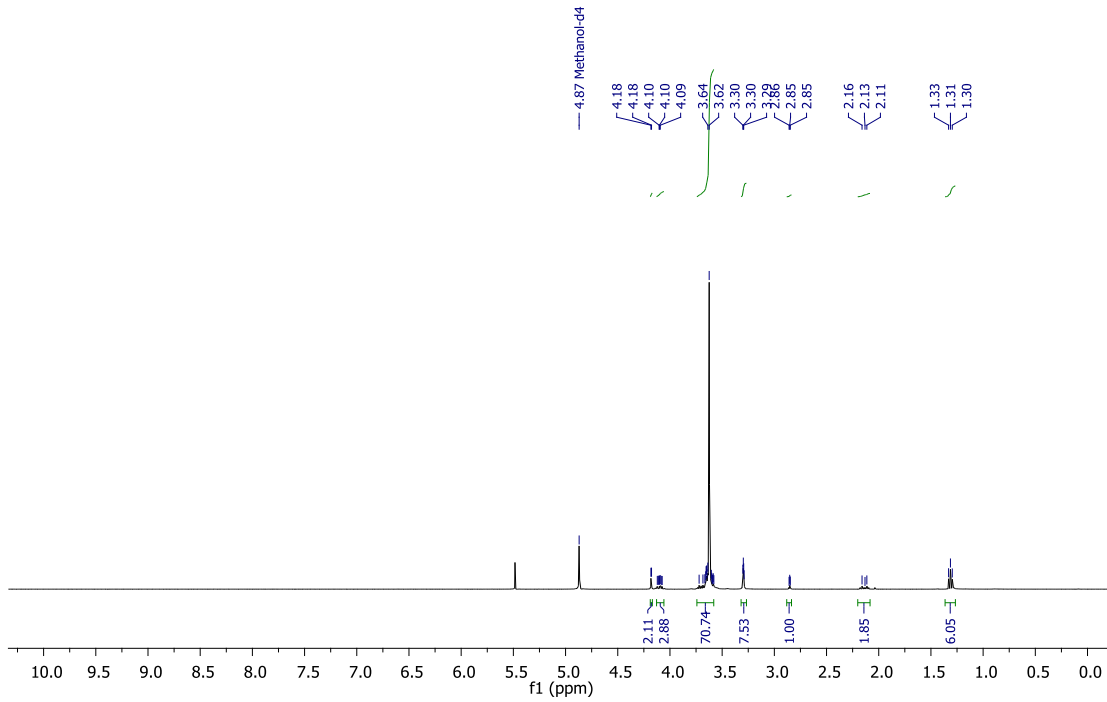
¹³C NMR spectrum of compound 1h



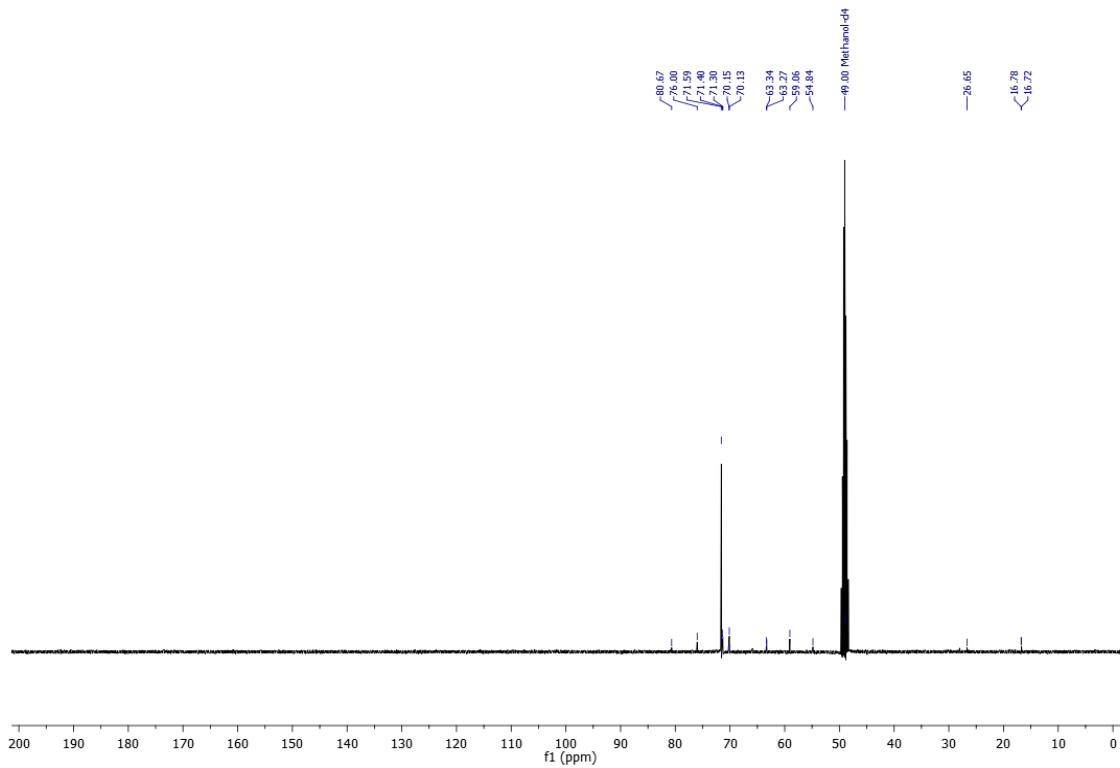
¹H NMR spectrum of compound 1i



¹³C NMR spectrum of compound 1i

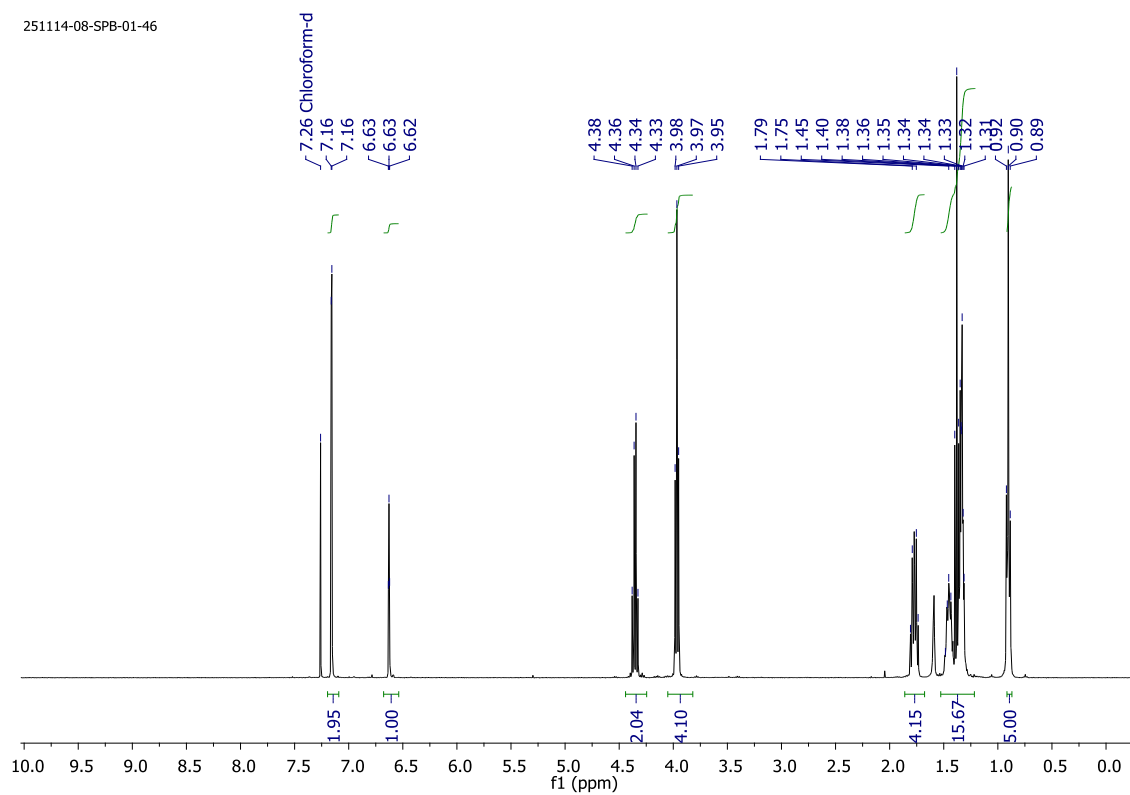


¹H NMR spectrum of compound 1j



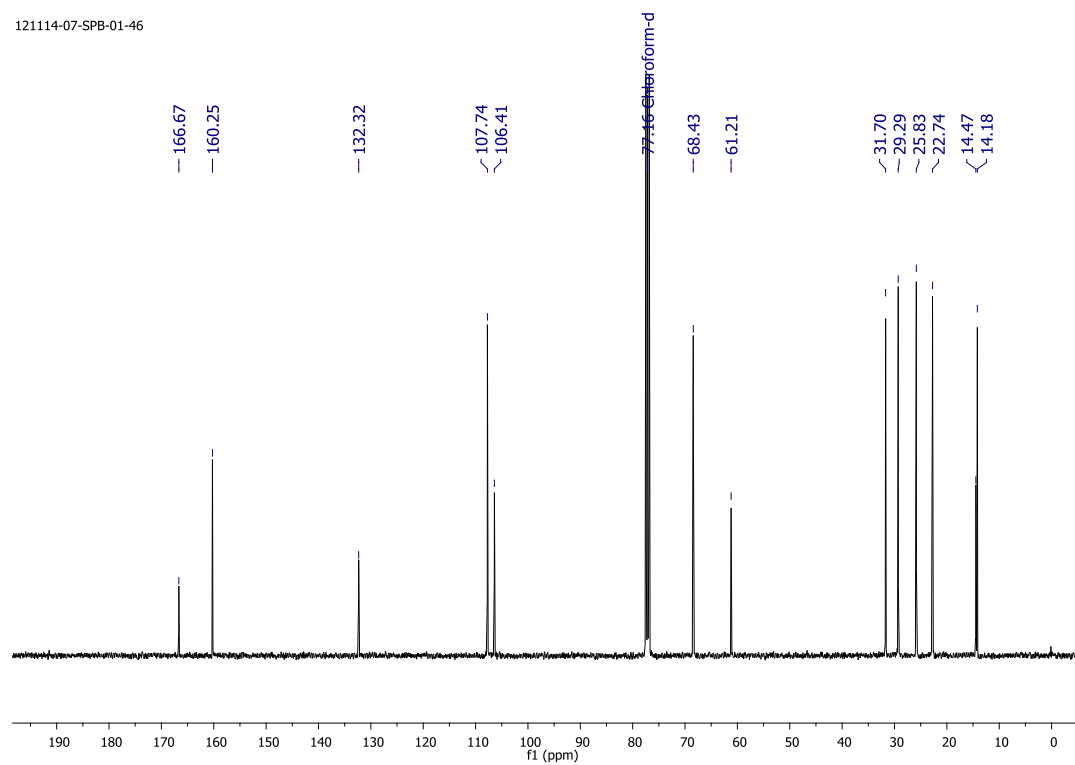
¹³C NMR spectrum of compound 1j

251114-08-SPB-01-46



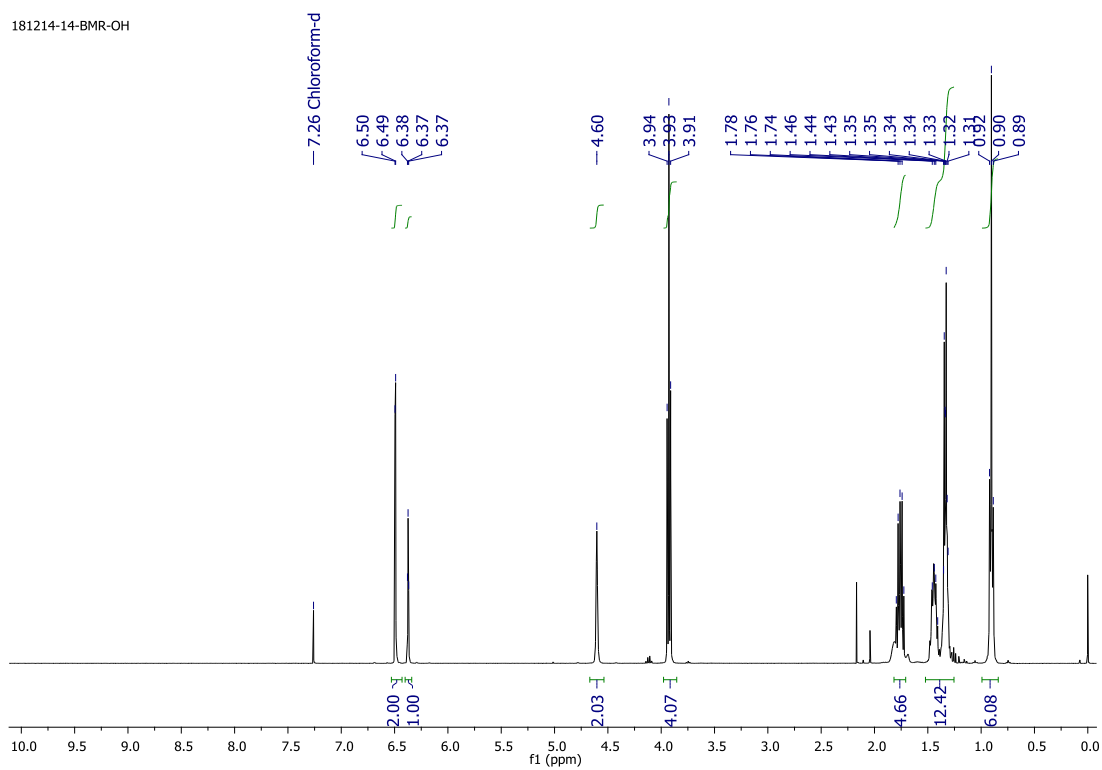
¹H NMR spectrum of compound 2a

121114-07-SPB-01-46



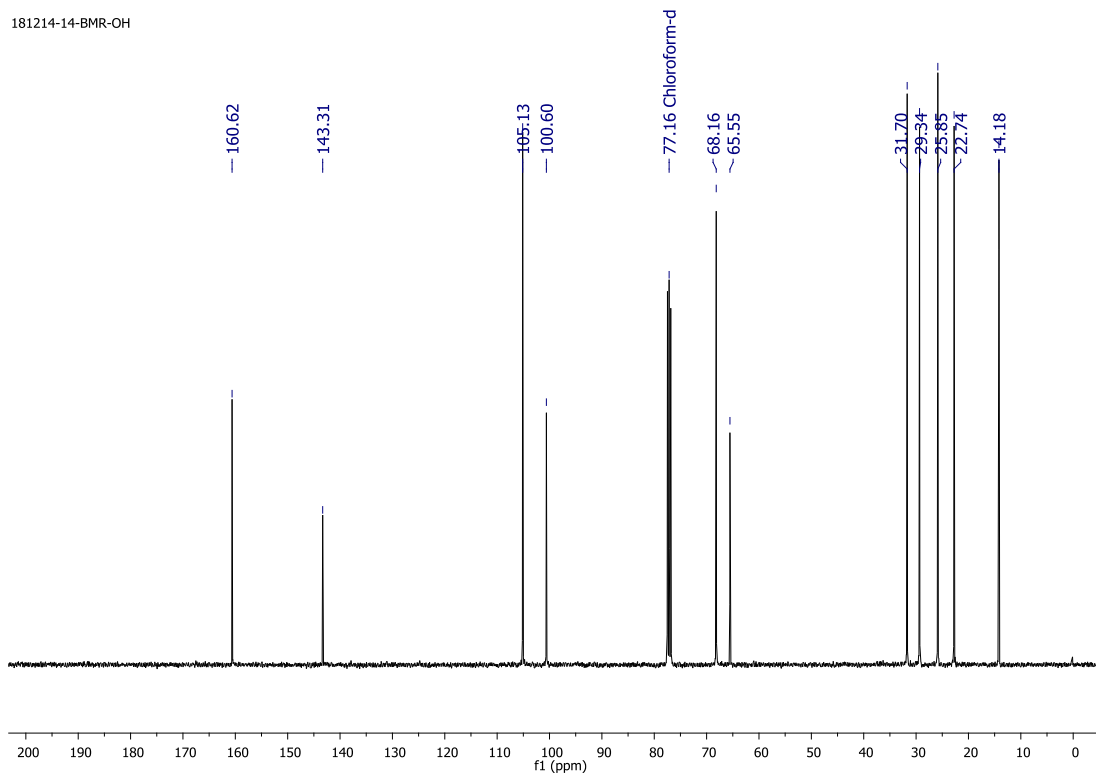
¹³C NMR spectrum of compound 2a

181214-14-BMR-OH



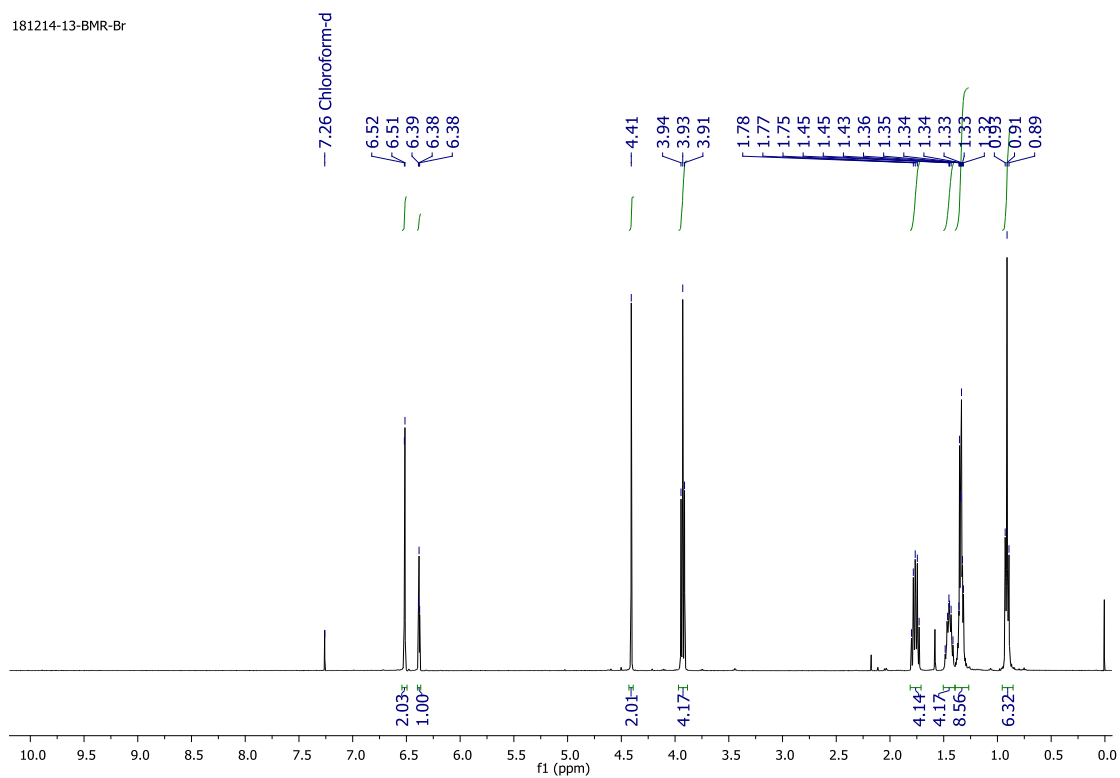
¹H NMR spectrum of compound 2b

181214-14-BMR-OH



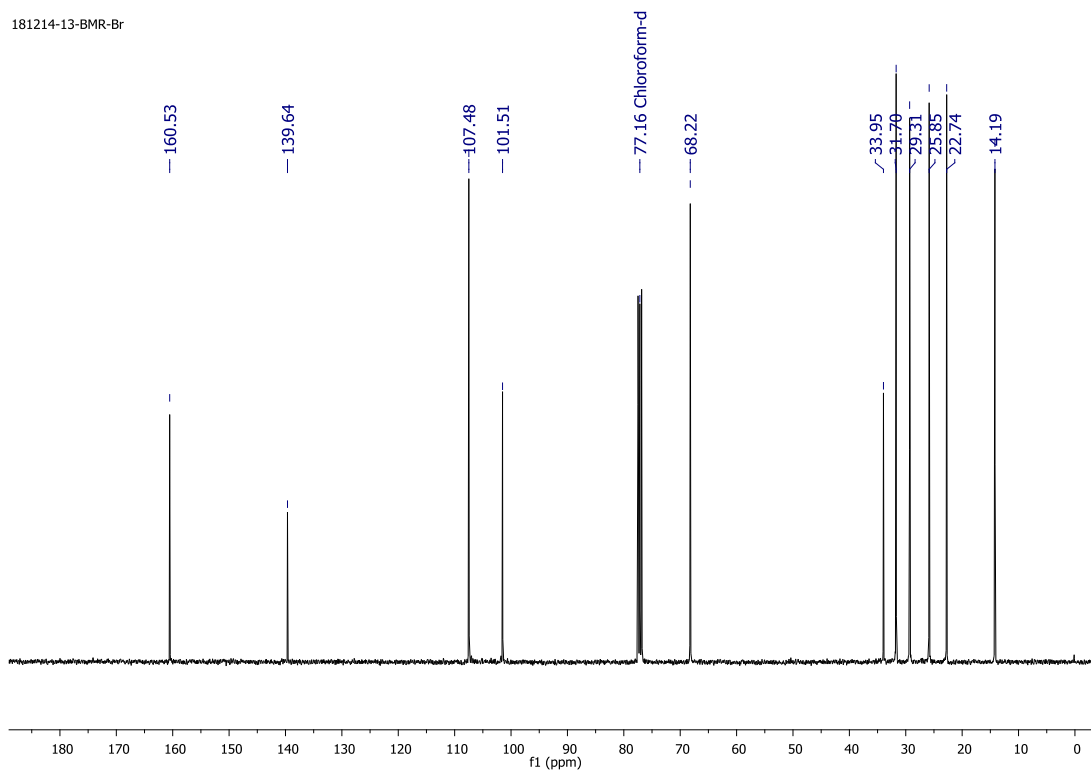
¹³C NMR spectrum of compound 2b

181214-13-BMR-Br



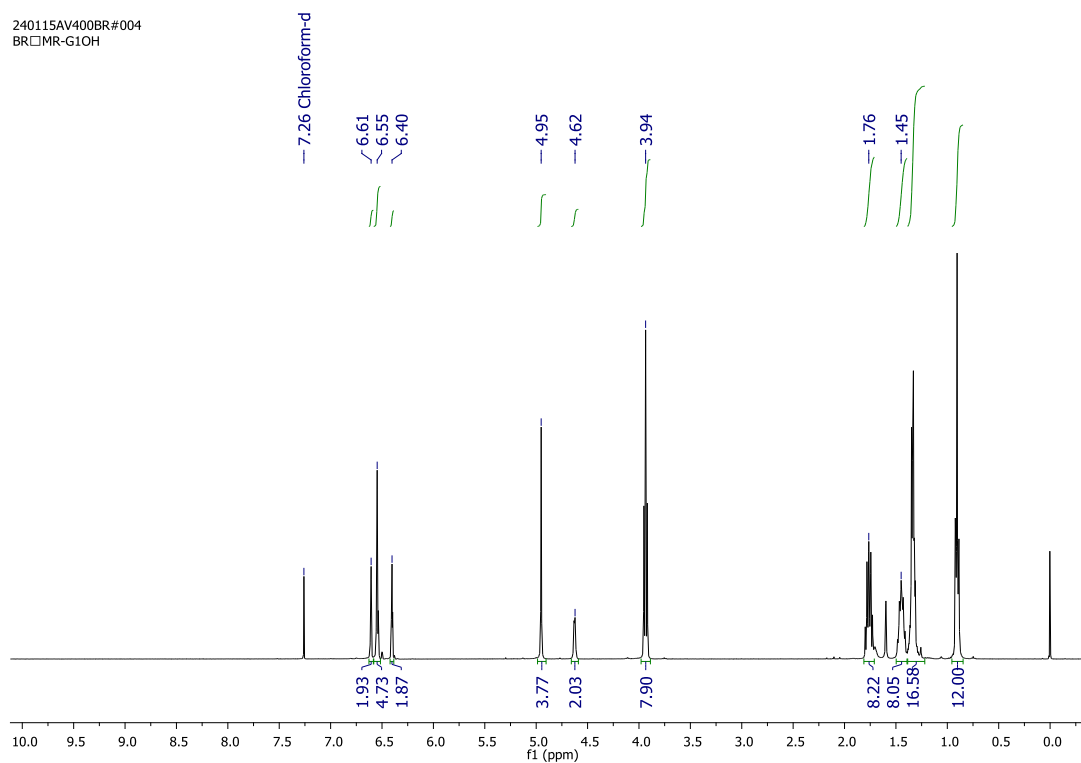
¹H NMR spectrum of compound 2c

181214-13-BMR-Br



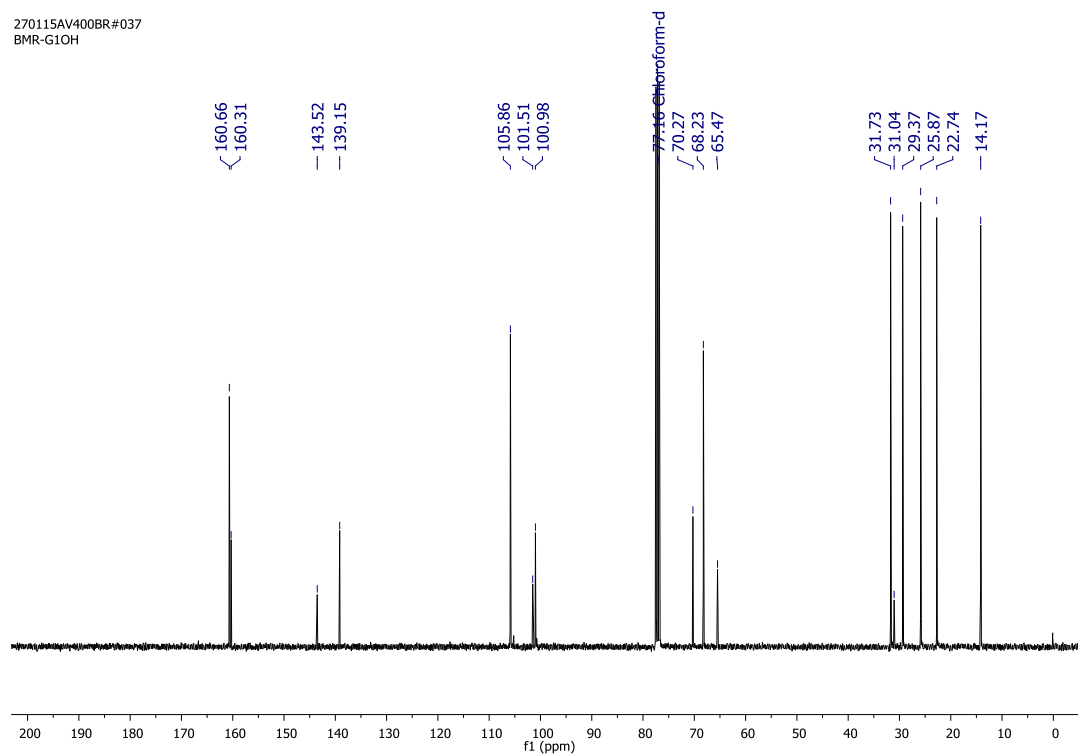
¹³C NMR spectrum of compound 2c

240115AV400BR#004
BRMR-G1OH



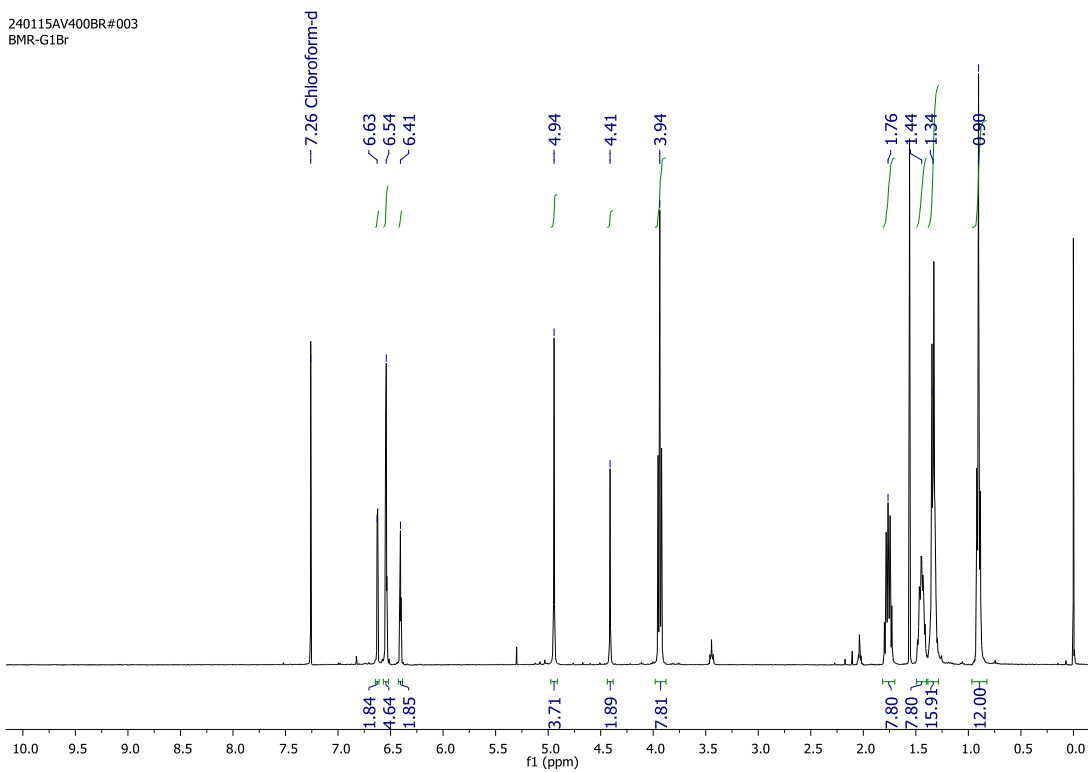
¹H NMR spectrum of compound 3a

270115AV400BR#037
BMR-G1OH



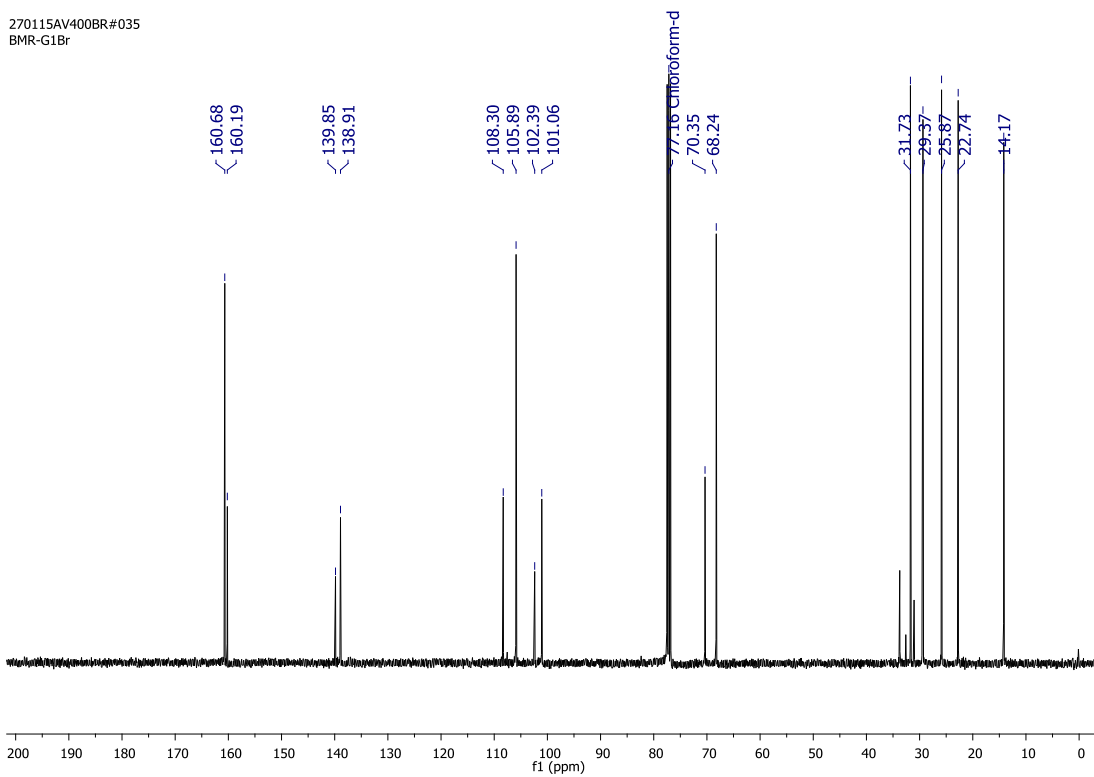
¹³C NMR spectrum of compound 3a

240115AV400BR#003
BMR-G1Br

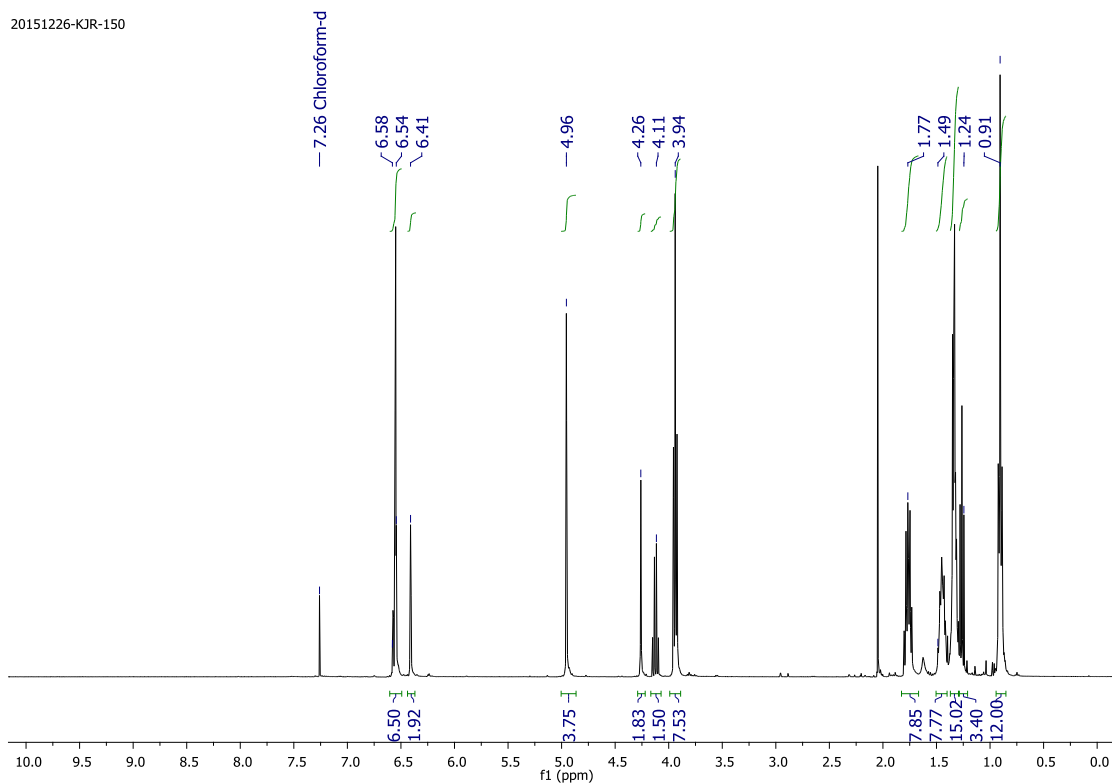


¹H NMR spectrum of compound 3b

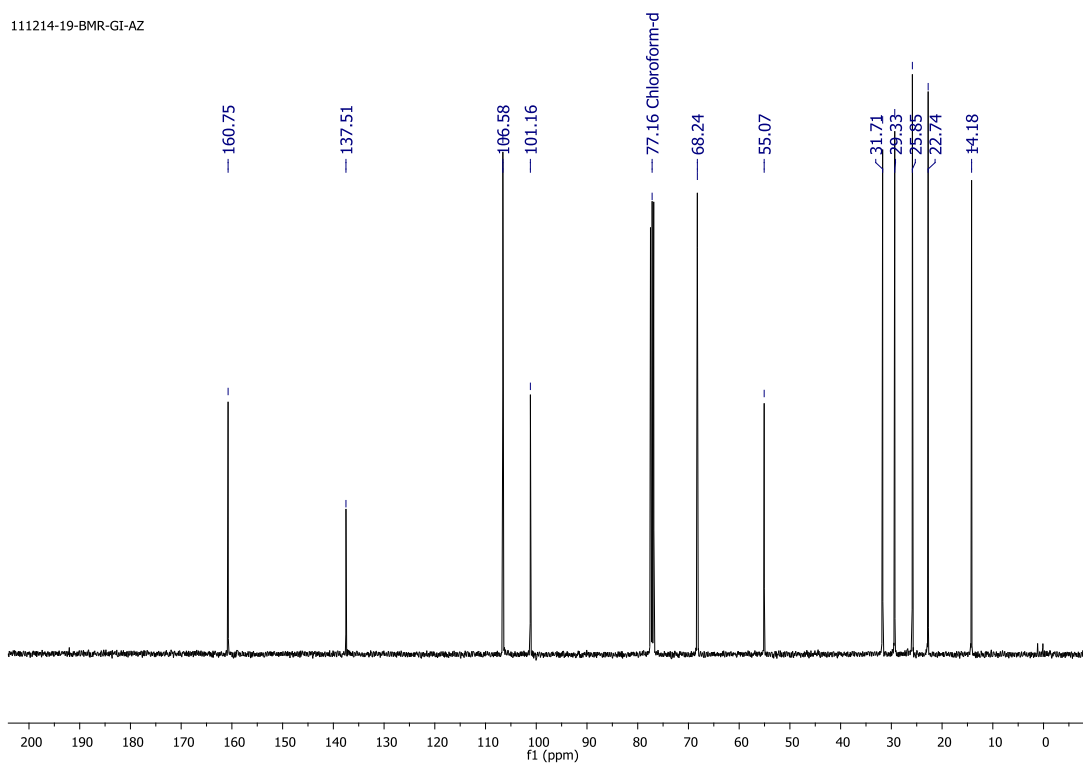
270115AV400BR#035
BMR-G1Br



¹³C NMR spectrum of compound 3b

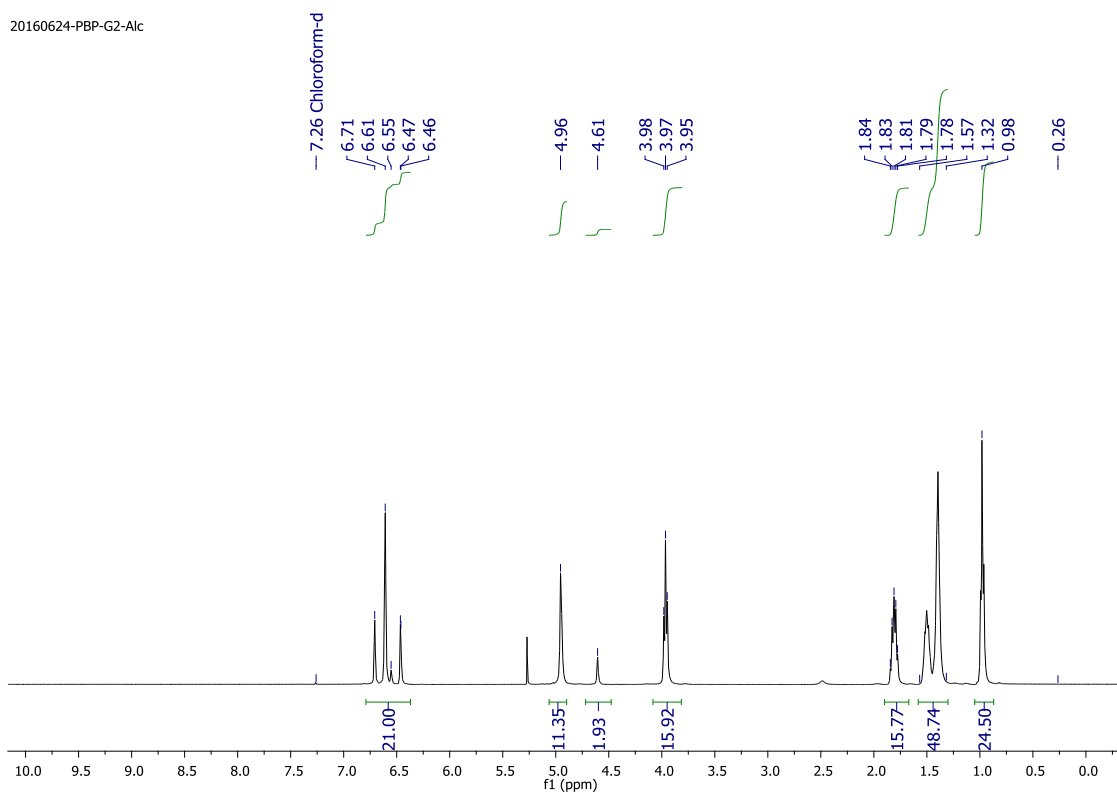


¹H NMR spectrum of compound 3c



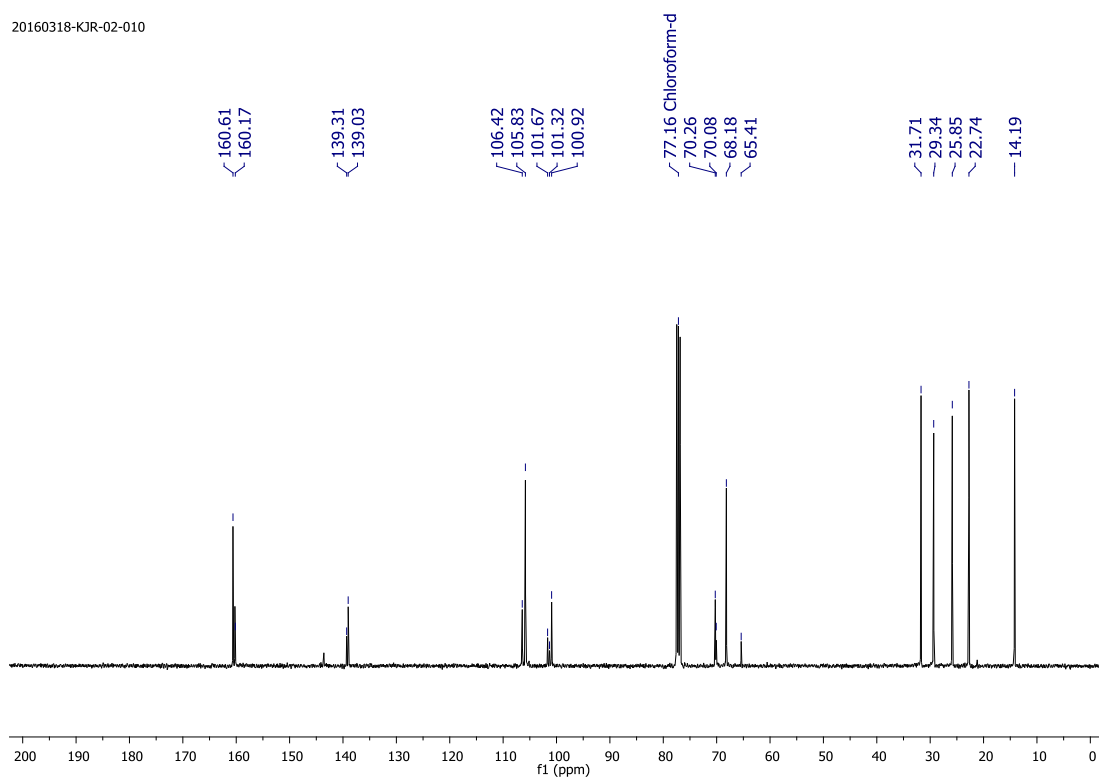
¹³C NMR spectrum of compound 3c

20160624-PBP-G2-Alc



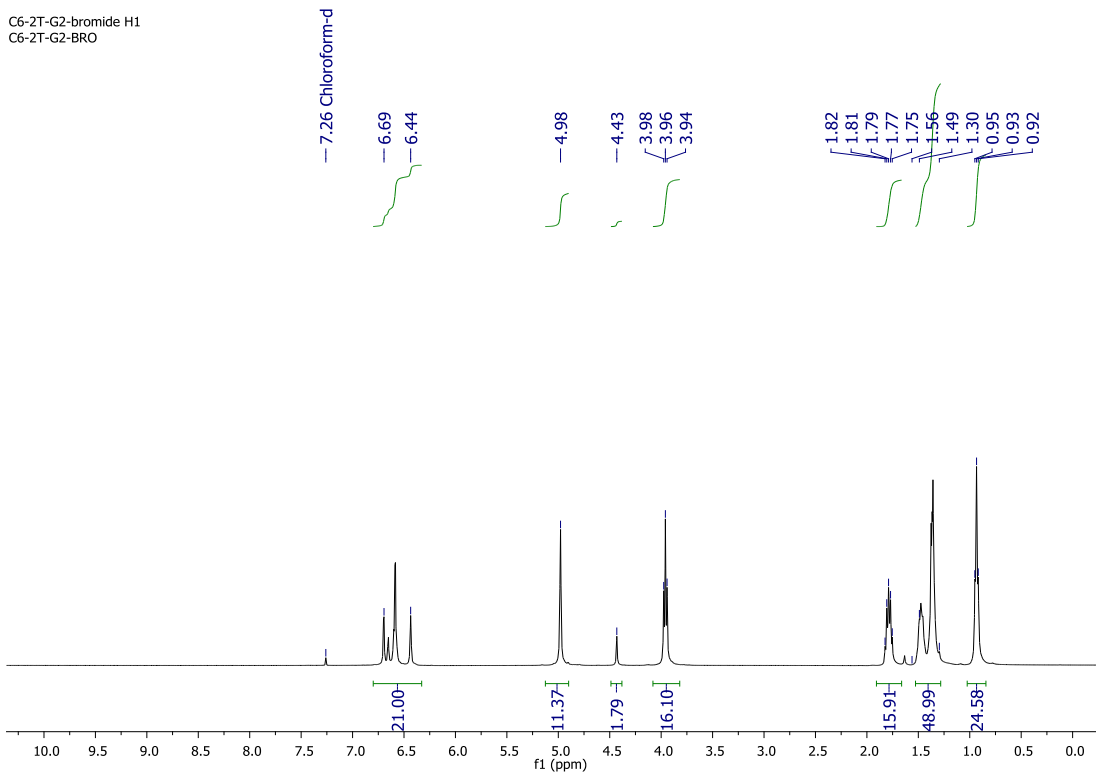
¹H NMR spectrum of compound 4a

20160318-KJR-02-010



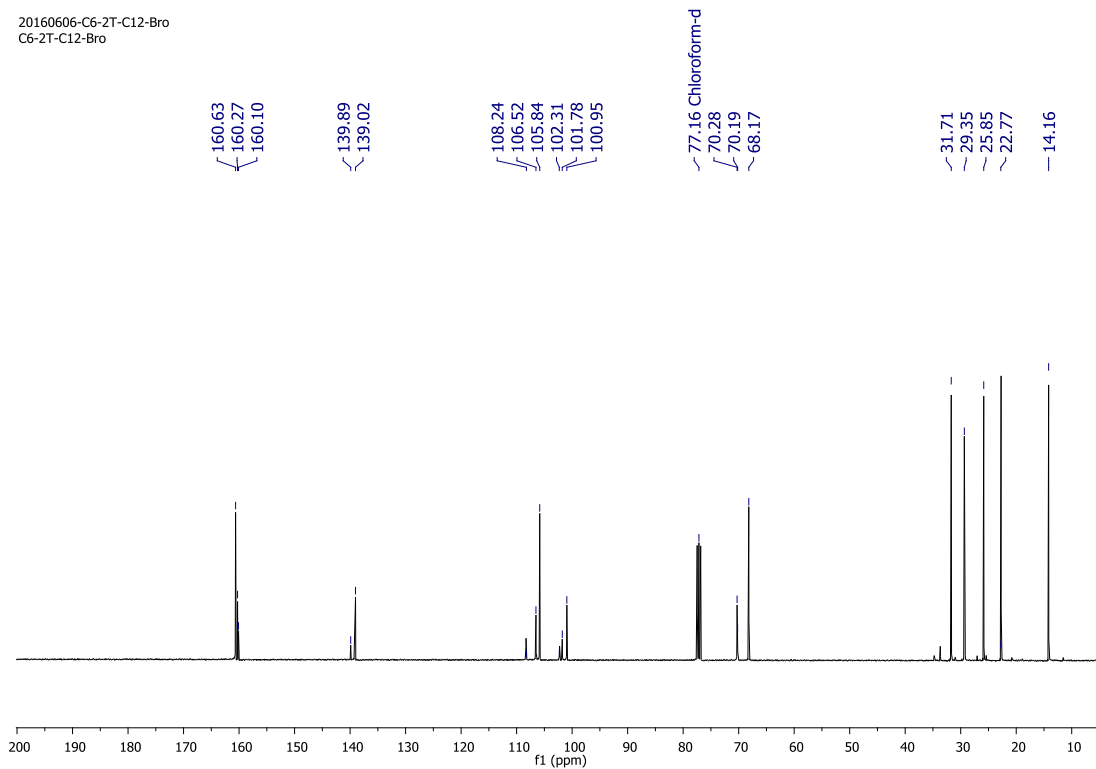
¹³C NMR spectrum of compound 4a

C6-2T-G2-bromide H1
C6-2T-G2-BRO

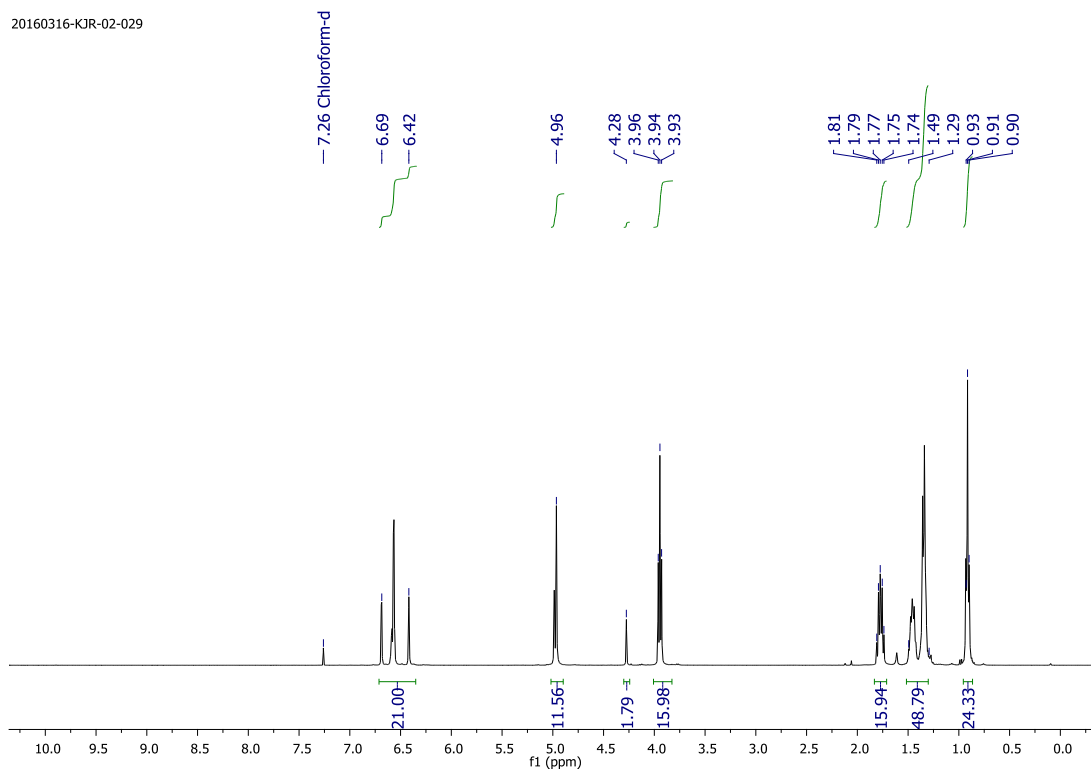
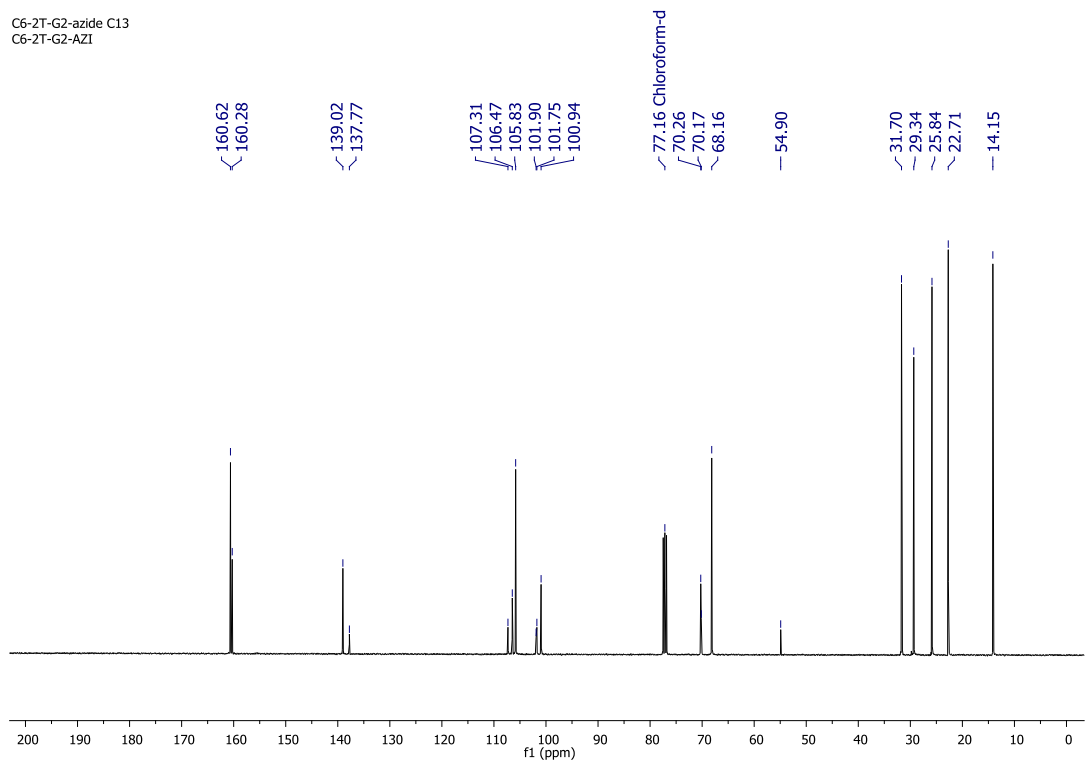


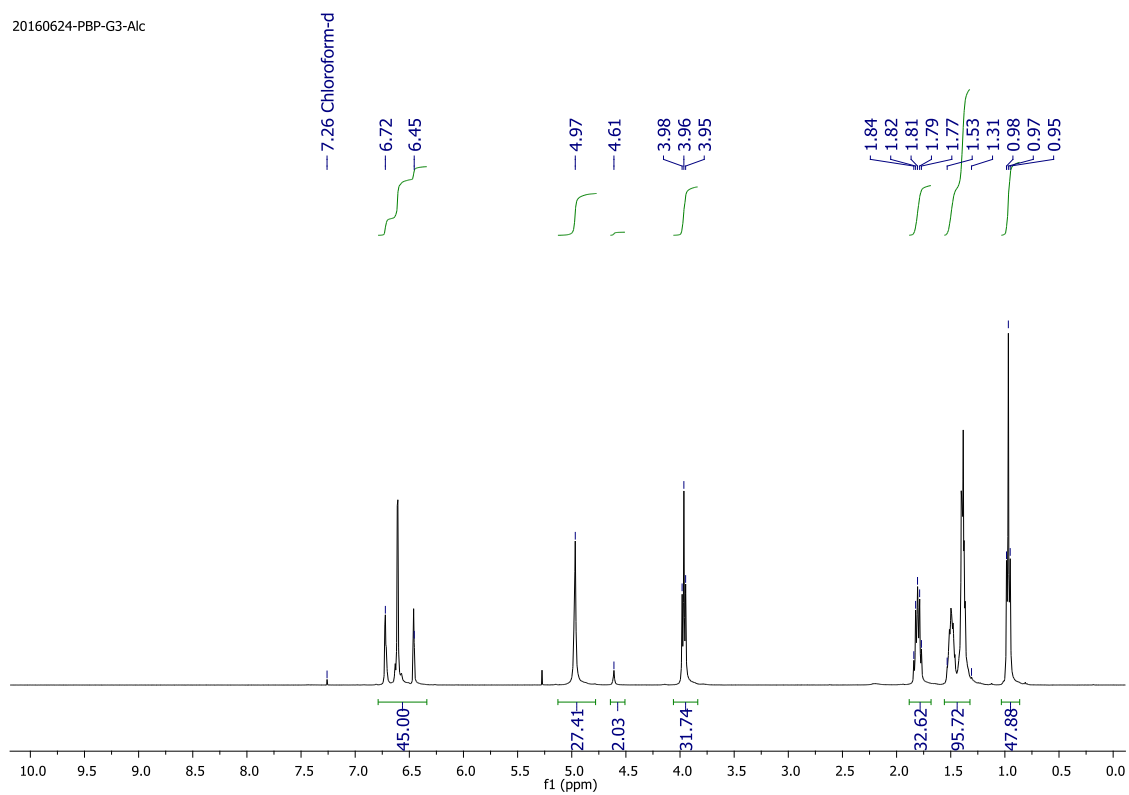
¹H NMR spectrum of compound 4b

20160606-C6-2T-C12-Bro
C6-2T-C12-Bro



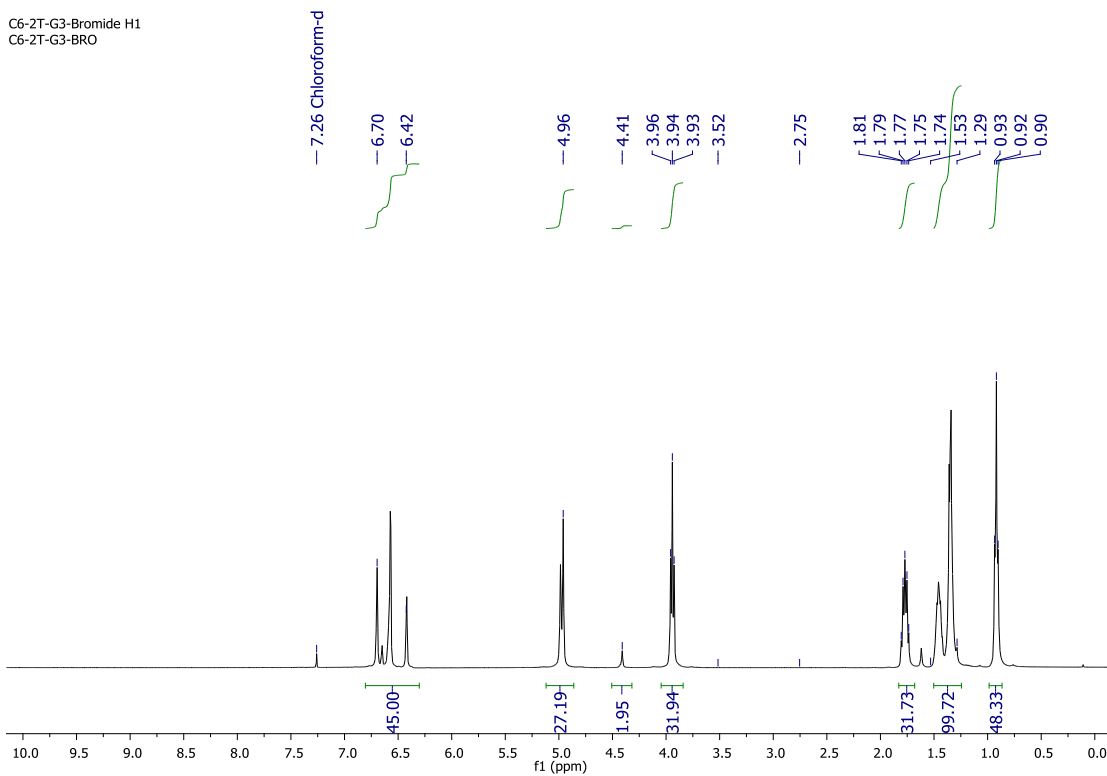
¹³C NMR spectrum of compound 4b

**¹H NMR spectrum of compound 4c**C6-2T-G2-azide C13
C6-2T-G2-AZI**¹³C NMR spectrum of compound 4b**



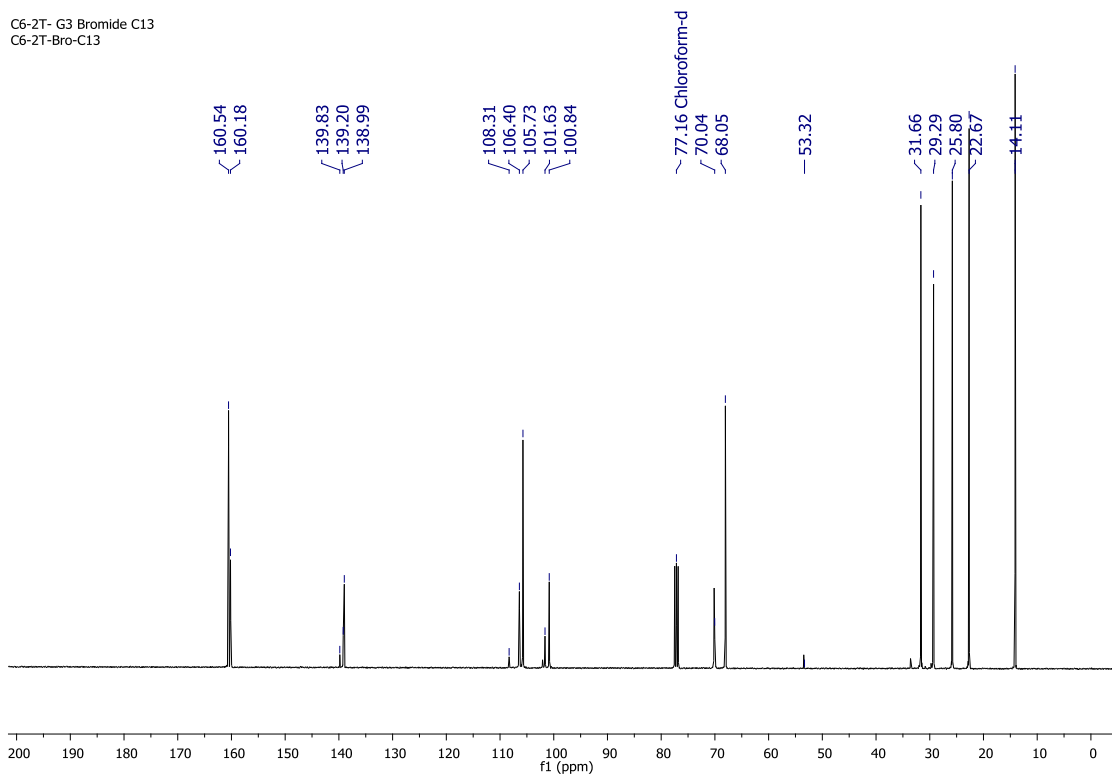
¹H NMR spectrum of compound 5a

C6-2T-G3-Bromide H1
C6-2T-G3-BRO



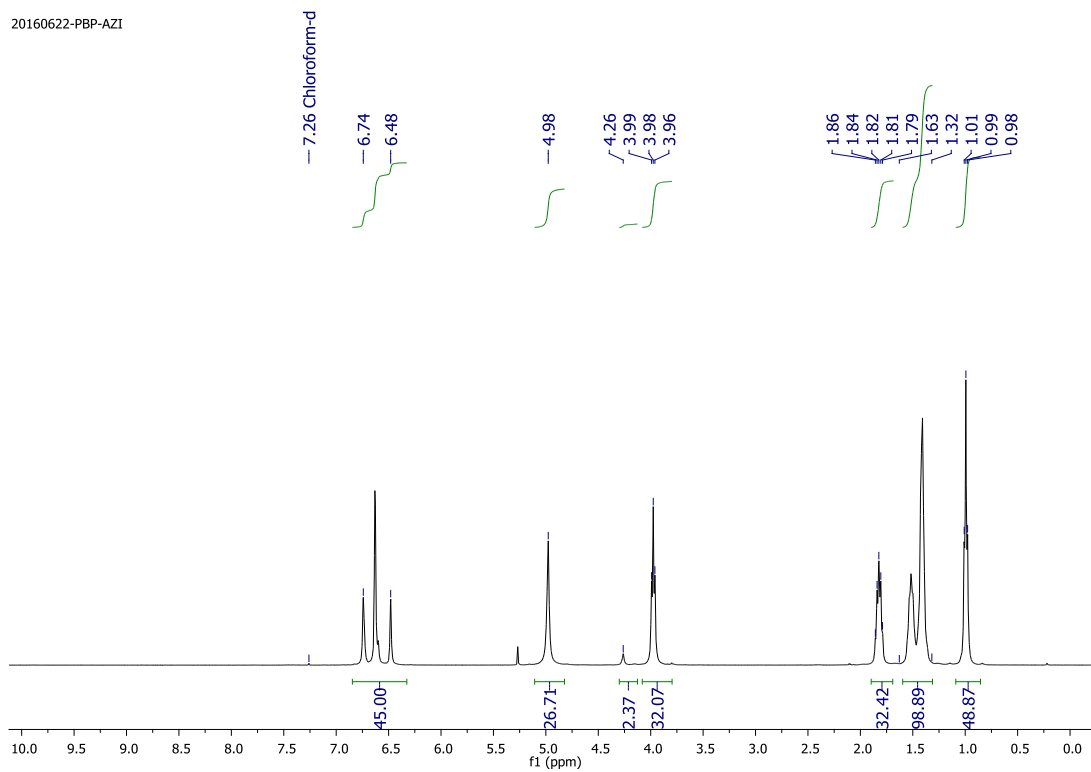
¹H NMR spectrum of compound 5b

C6-2T- G3 Bromide C13
C6-2T-Bro-C13



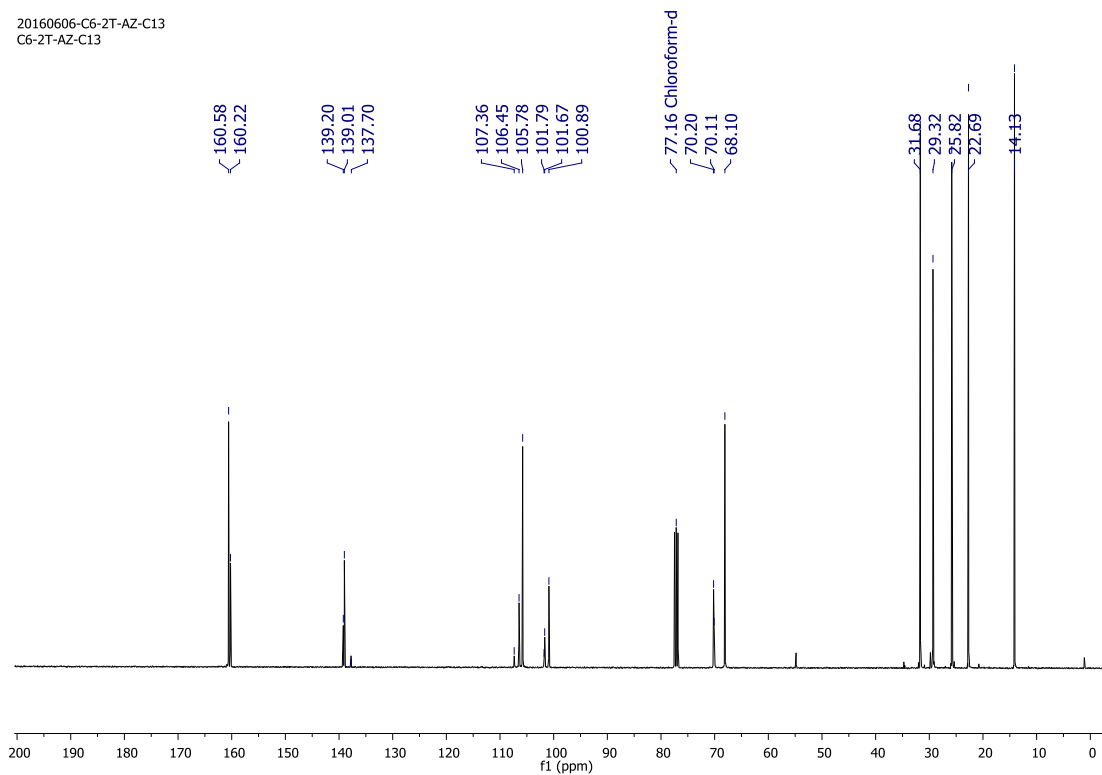
¹³C NMR spectrum of compound 5b

20160622-PBP-AZI



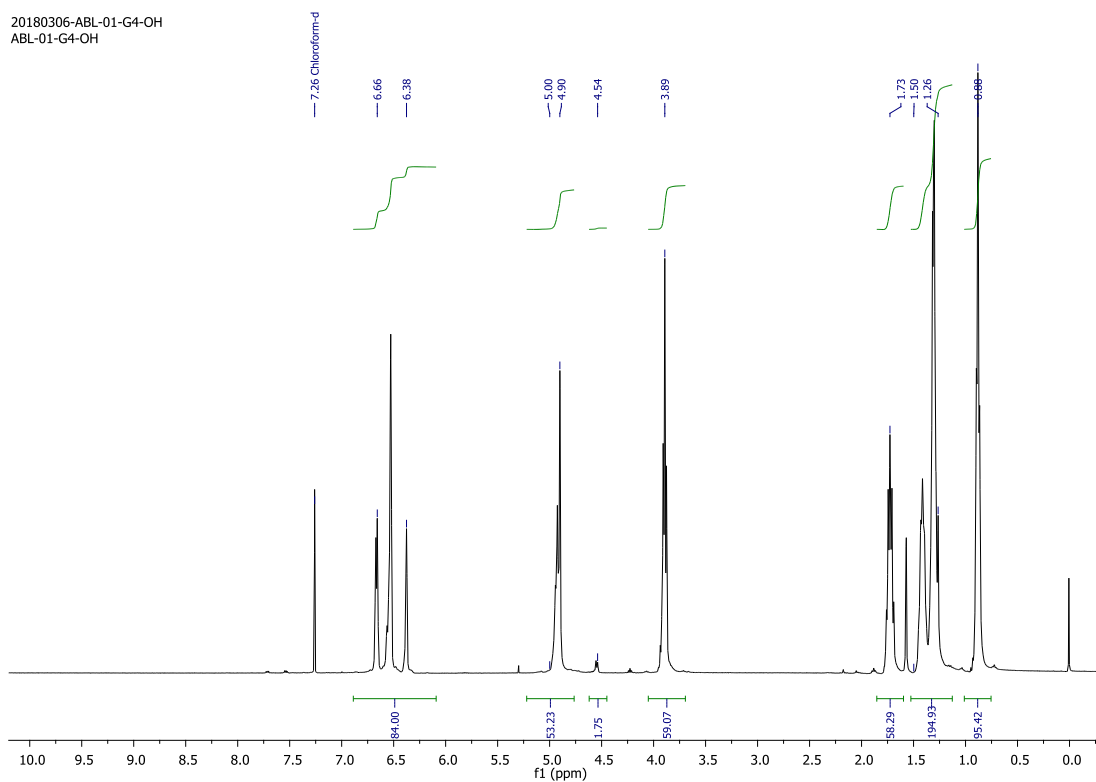
¹H NMR spectrum of compound 5c

20160606-C6-2T-AZ-C13
C6-2T-AZ-C13



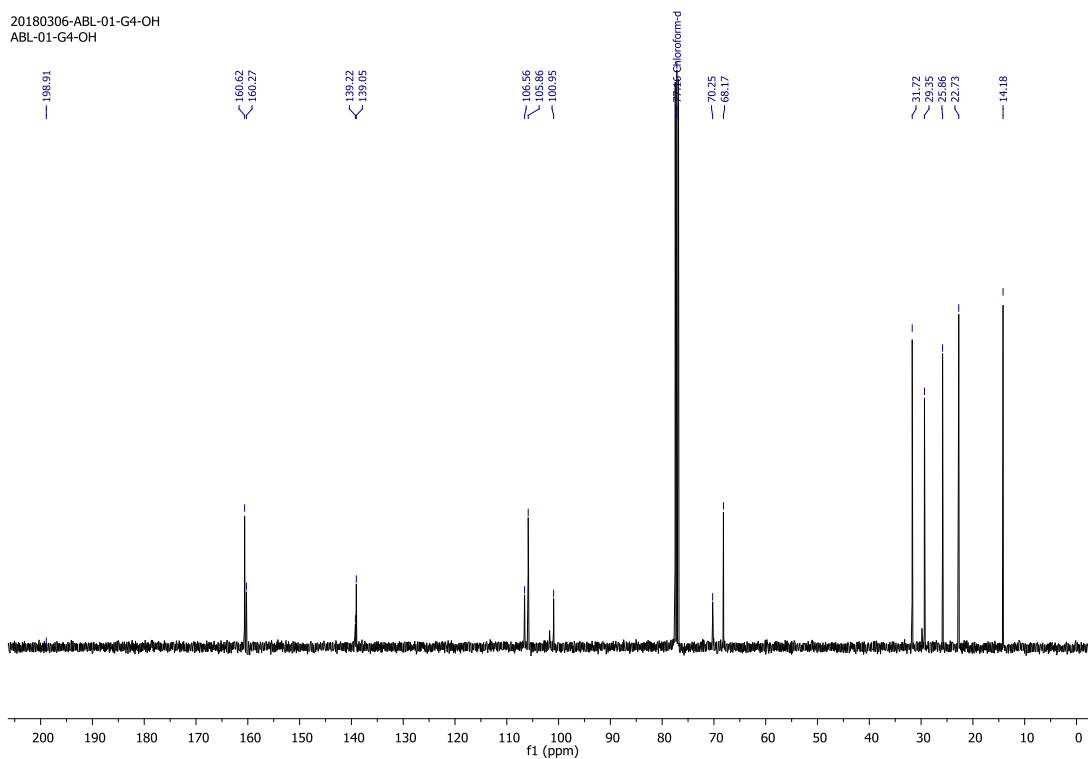
¹³C NMR spectrum of compound 5c

20180306-ABL-01-G4-OH
ABL-01-G4-OH



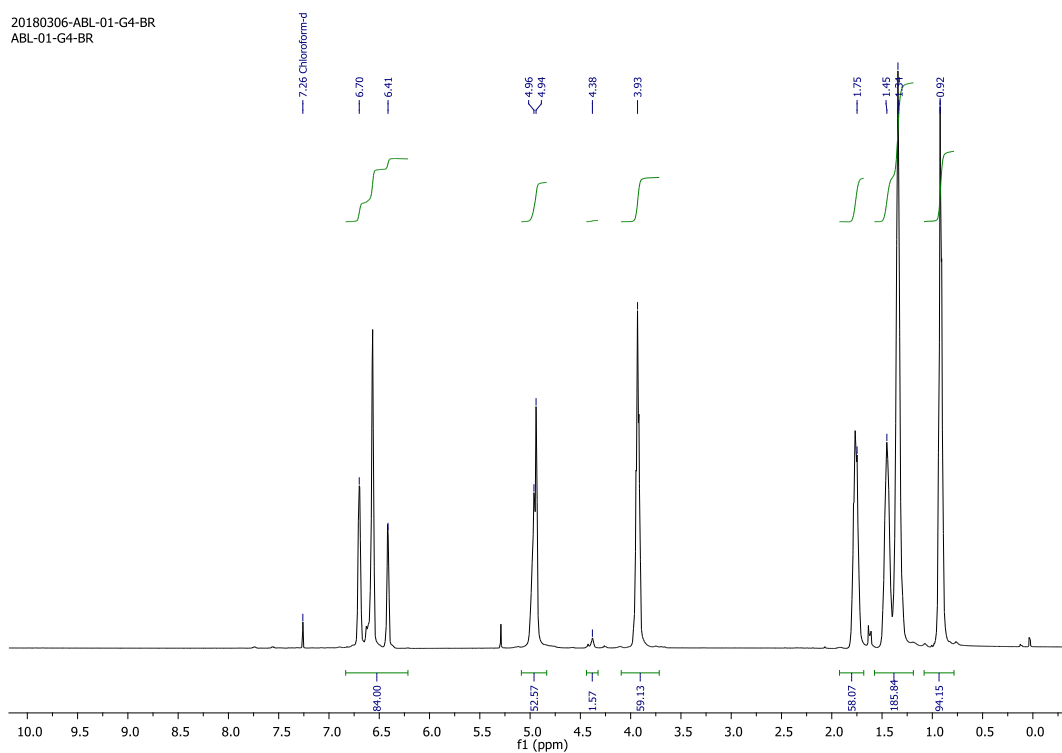
¹H NMR spectrum of compound 6a

20180306-ABL-01-G4-OH
ABL-01-G4-OH



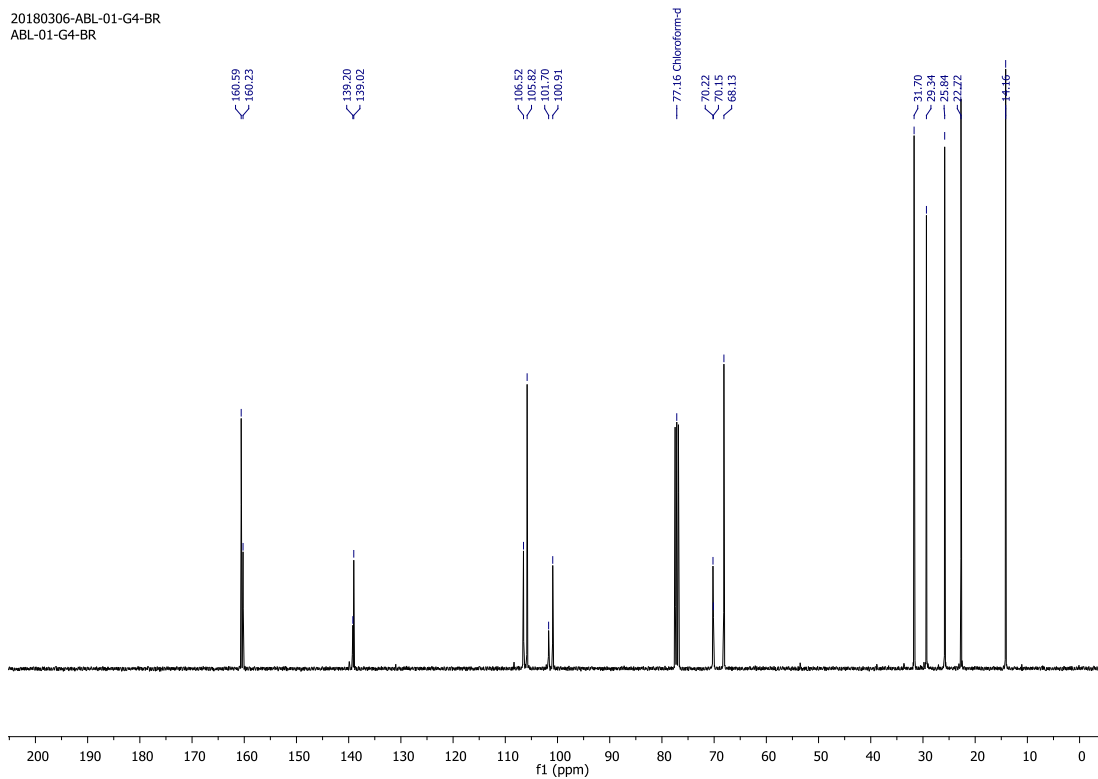
¹³C NMR spectrum of compound 6a

20180306-ABL-01-G4-BR
ABL-01-G4-BR



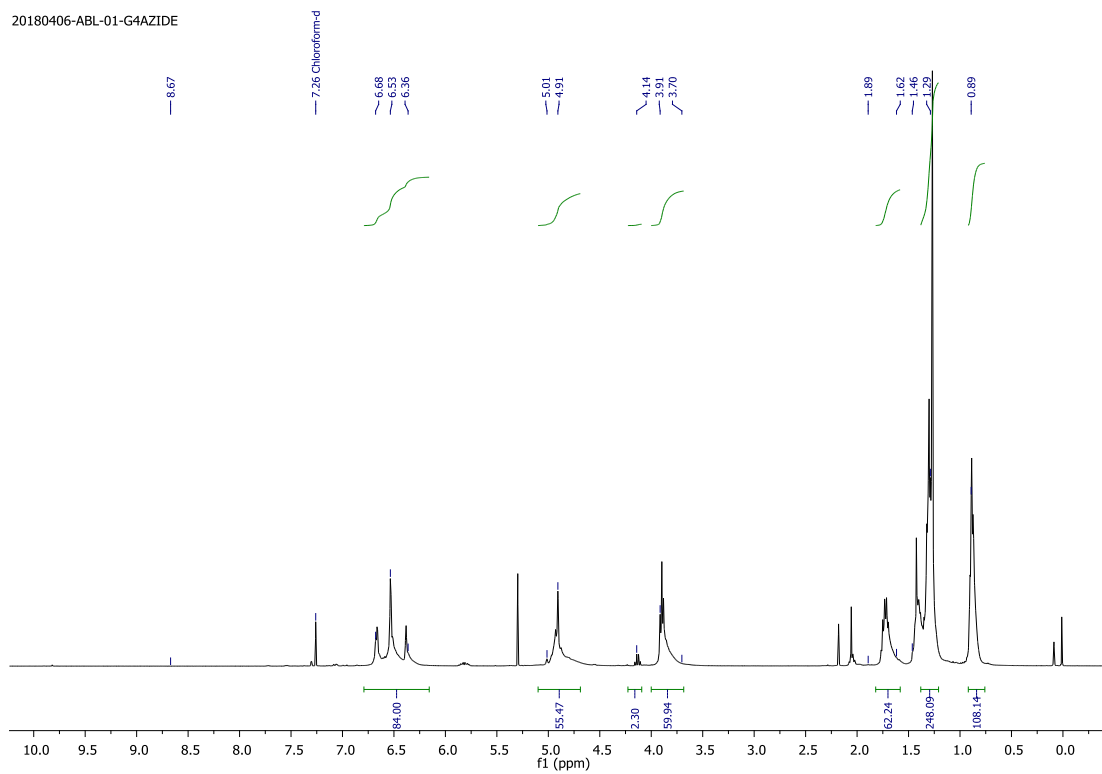
¹H NMR spectrum of compound 6b

20180306-ABL-01-G4-BR
ABL-01-G4-BR



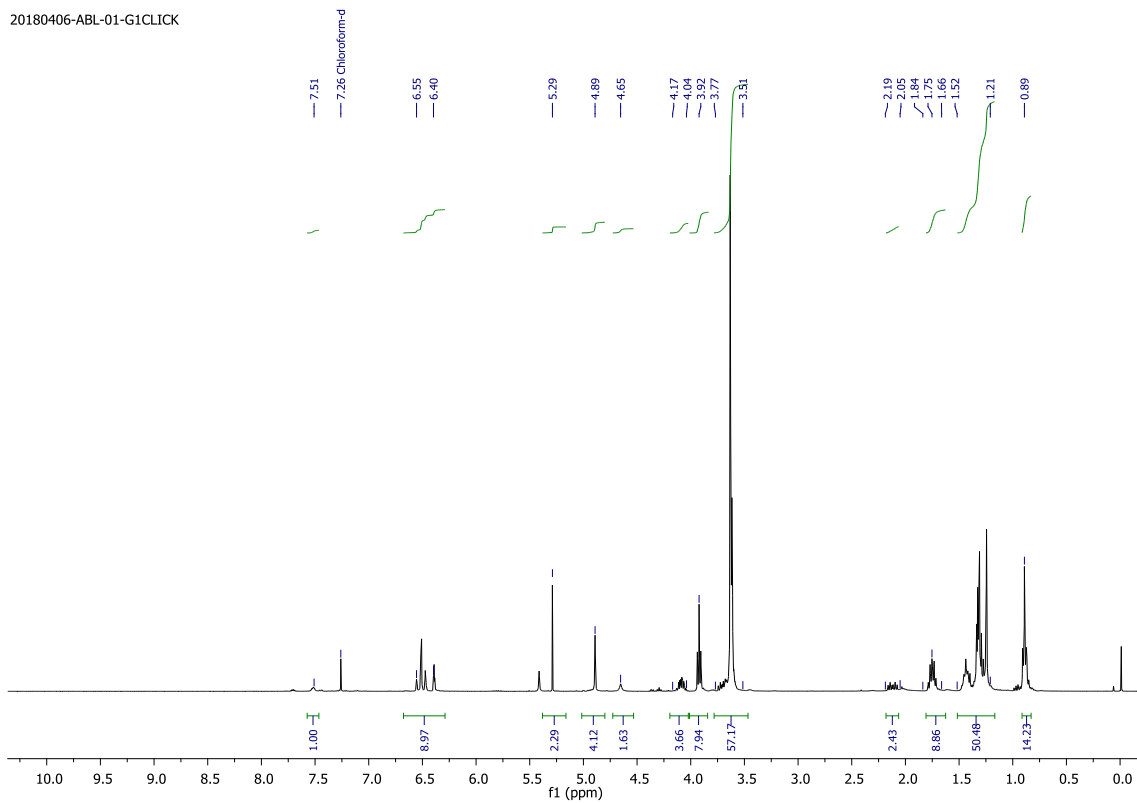
¹³C NMR spectrum of compound 6b

20180406-ABL-01-G4AZIDE



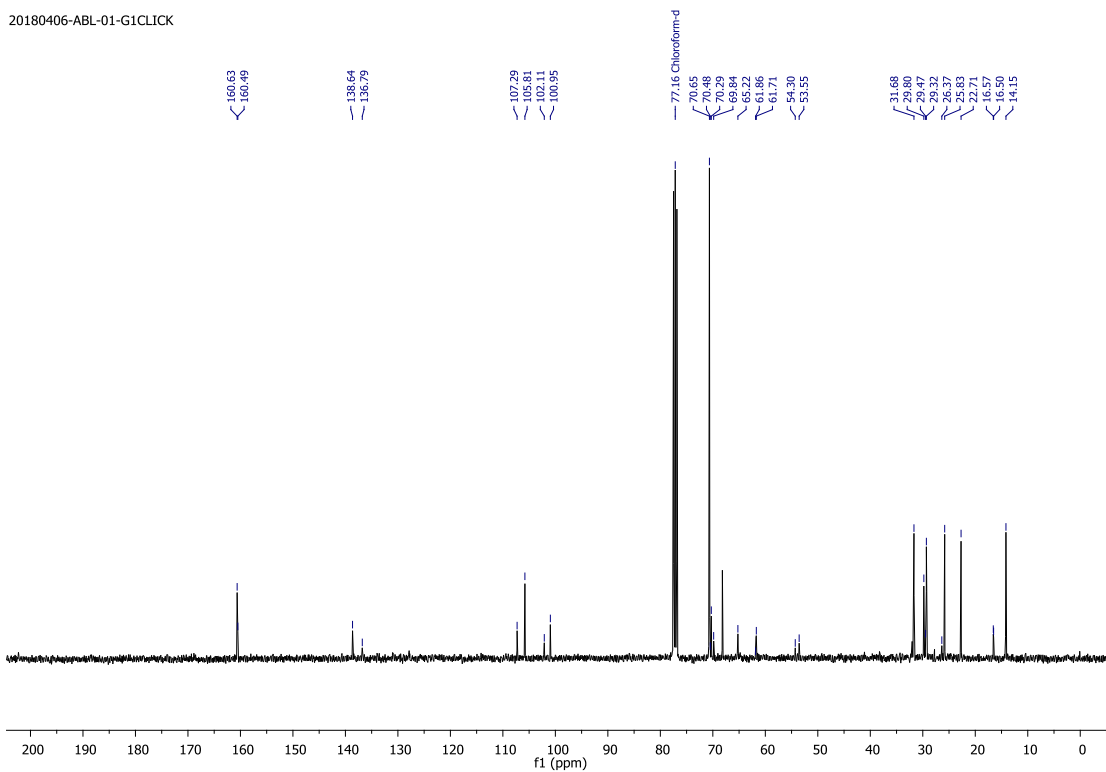
¹H NMR spectrum of compound 6c

20180406-ABL-01-G1CLICK



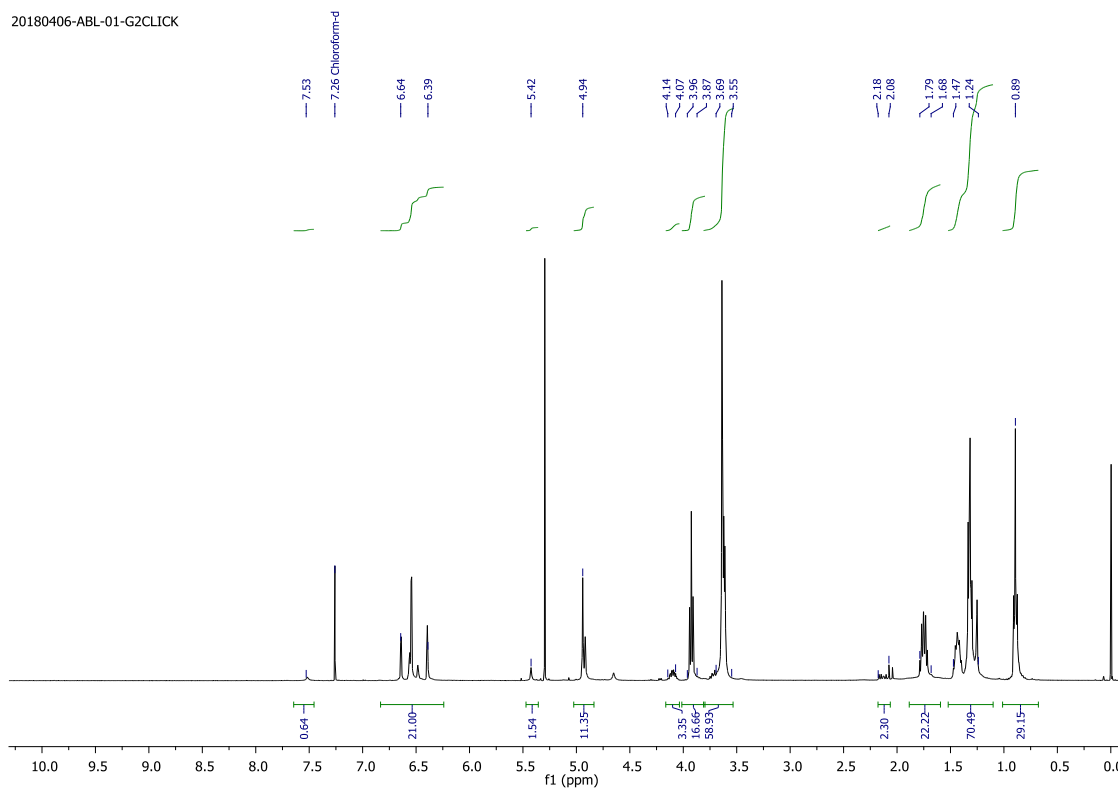
¹H NMR spectrum of compound 7a

20180406-ABL-01-G1CLICK

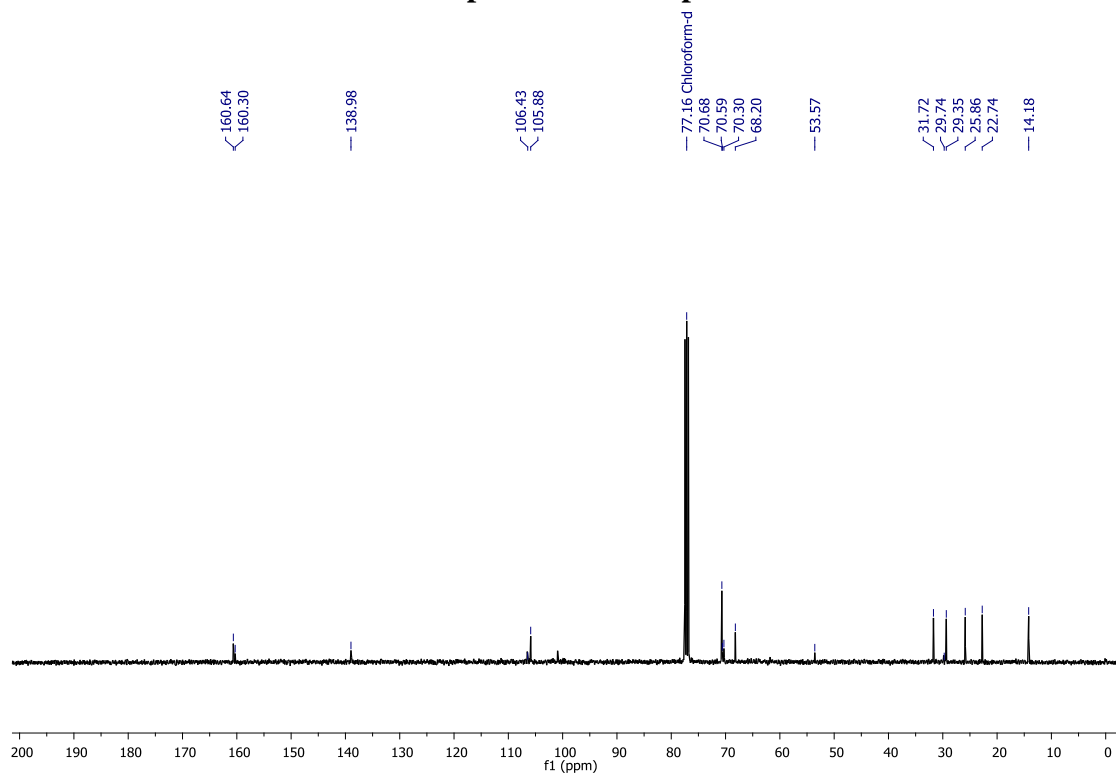


¹³C NMR spectrum of compound 7a

20180406-ABL-01-G2CLICK

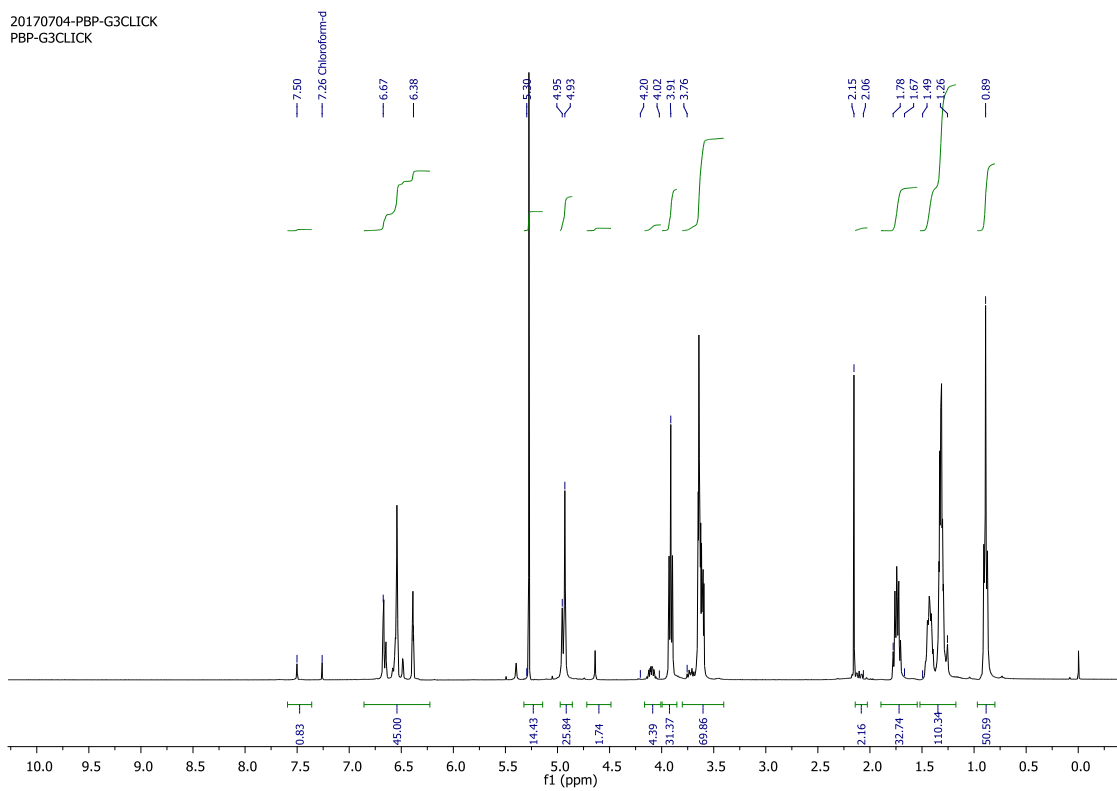


¹H NMR spectrum of compound 8a



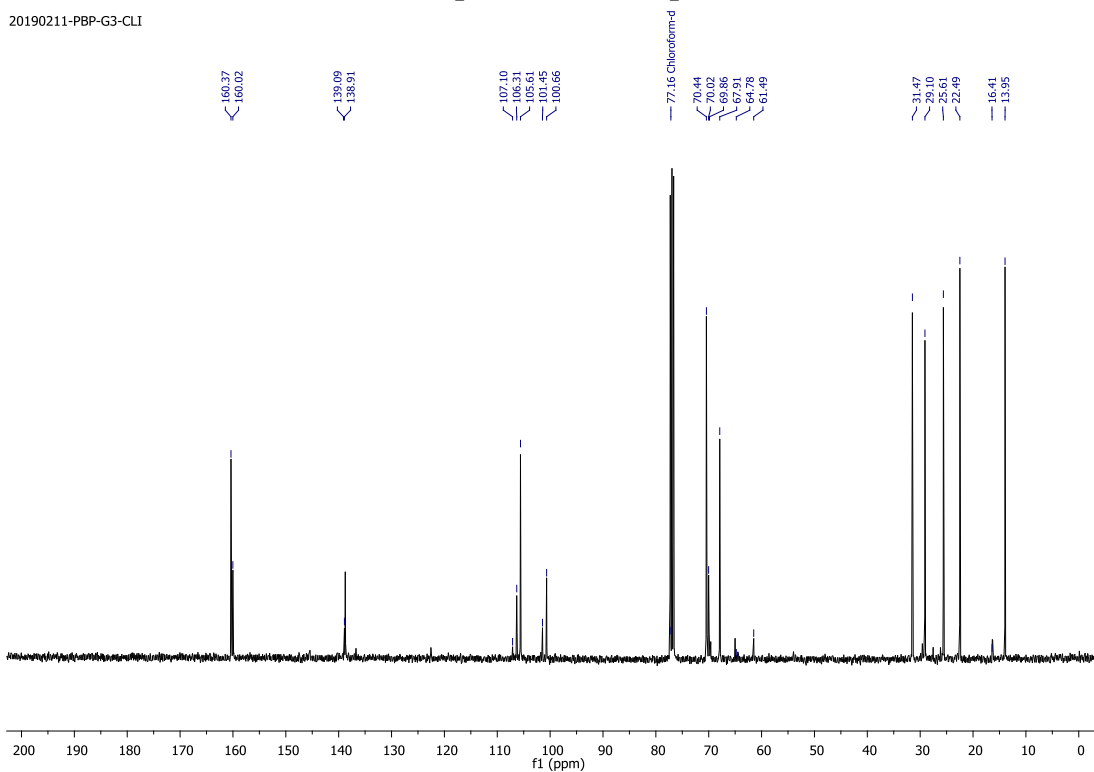
¹³C NMR spectrum of compound 8a

20170704-PBP-G3CLICK
PBP-G3CLICK



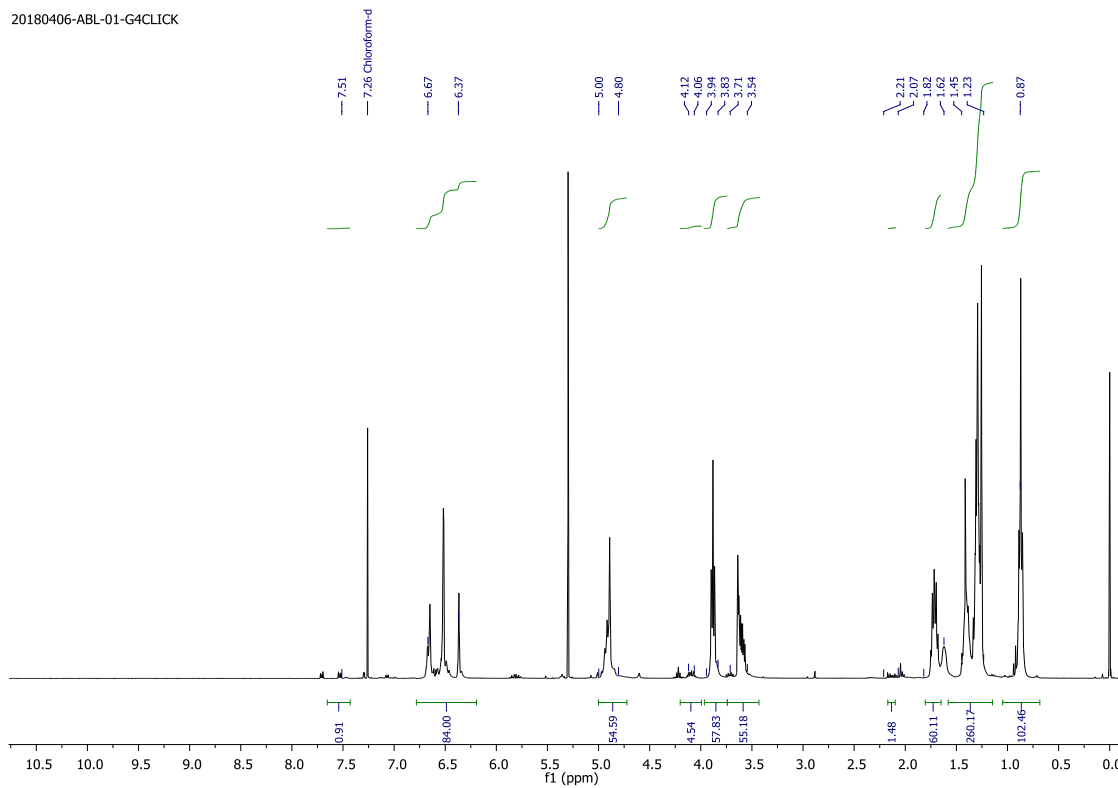
¹H NMR spectrum of compound 9a

20190211-PBP-G3-CLI



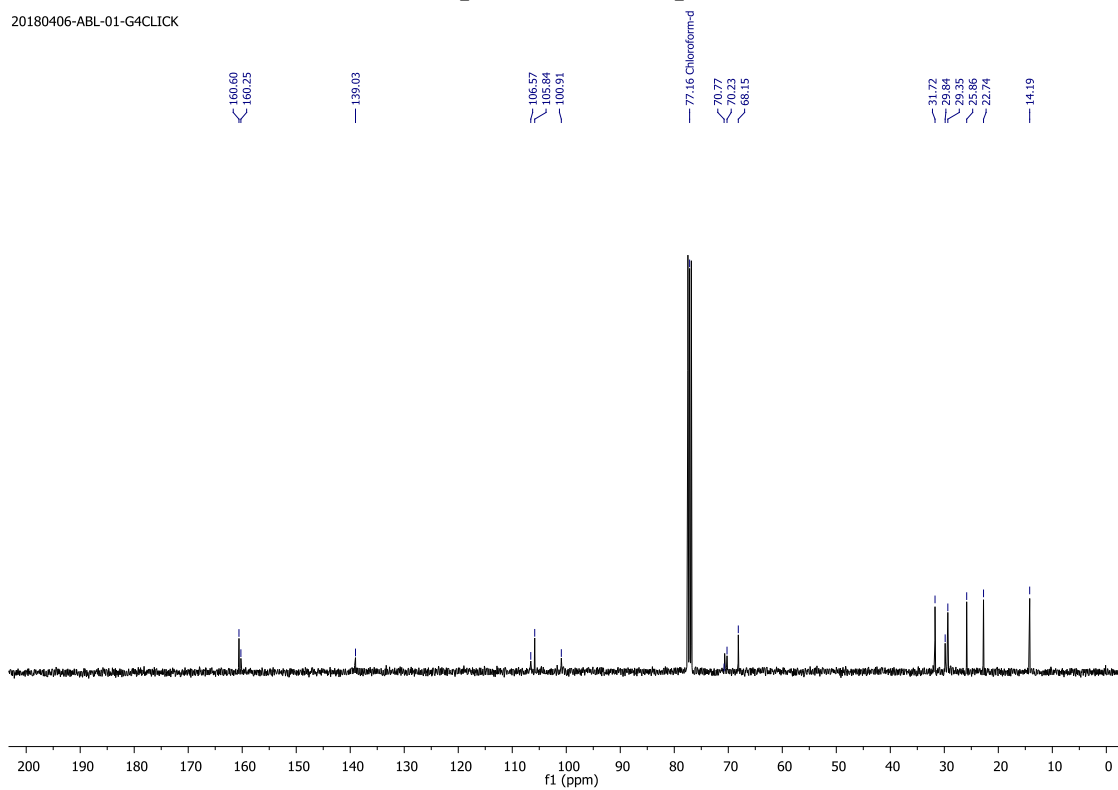
¹³C NMR spectrum of compound 9a

20180406-ABL-01-G4CLICK



¹H NMR spectrum of compound 10a

20180406-ABL-01-G4CLICK



¹³C NMR spectrum of compound 10a

CHAPTER 3

PROGRAMMED AND SEQUENTIAL DISASSEMBLY OF MULTI-RESPONSIVE SUPRAMOLECULAR PROTEIN ASSEMBLIES: A DETAILED MECHANISTIC INVESTIGATION

3.1 Introduction

In living systems, monomeric proteins often organize into highly stable yet dynamic complexes, which then carry out the majority of biological functions such as protein degradation,¹ cell division,² catalysis,³ transportation,⁴ storage containers,⁵ etc⁶. Reproducing these hierarchical biological structures and their dynamics at the molecular scale using synthetic functional materials is regarded to be of paramount importance. There are several efforts made in this direction to achieve this goal. One such impressive method to construct dynamic protein materials is based upon employing the natural protein assemblies to serve as a template.⁷ In this approach, the coat protein of the template is chemically⁸ or genetically modified⁹ to construct biohybrid stimuli-responsive materials. Although useful, the strategy is limited to only a few protein complexes. This is because, most of the robust bioconjugation methods often fail on protein assemblies owing to the complex and dynamic environment of coat proteins, which alters the chemical reactivity of amino acids.^{7a}

An alternative way to construct biohybrid stimuli-responsive materials is to introduce novel interfaces onto the non-assembling monomeric proteins by using computational methods. The construct then undergoes self-association under the selective condition to form a complex.¹⁰ To the best of our knowledge, the only method that describes the synthesis of stimuli-responsive protein assemblies using either rational or computational design is reported by Tezcan and coworkers. Here, they have shown the elegant use of modified protein surfaces to design protein assemblies through metal coordination by incorporating Rosetta-prescribed surface mutations to the cyt *cb*-562 variant MBPC1. They have shown that the properties of protein assemblies can be fine-tuned through the adjustment of the protein: metal ratio or pH.¹¹ However, The limited successes in designing dynamic architectures using computational design could stem from restricted chemical diversity, i.e., standard twenty amino acids.

On the other hand, chemical strategies provide enormous opportunities for the design of stimuli-responsive protein assemblies by site-selective functionalization of native protein using synthetic molecules. With few exceptions,¹² this is primarily achieved by

conjugating polymers having a LCST property to the proteins of interest. Above LCST, the hydrophilic polymer becomes hydrophobic, which induces global amphiphilicity to the bioconjugate, which drives the self-assembly process in a dilute solution.¹³ However, a significant disadvantage to the use of such polymer is that they produce a degree of heterogeneity in the form of polydispersity resulting in polydisperse samples, which are hard to study using standard analytical/biophysical techniques. In this regard, the synthesis of stimuli-responsive protein-dendron complexes would be an excellent scaffold. This is because similar to natural viruses, protein-dendron conjugates provide an opportunity to functionalize the interior of assemblies with a variety of therapeutic agents and exterior with targeting units to achieve specific delivery. Besides, the monodispersity of overall scaffold, and availability of unlimited chemical space, make the scaffold versatile, which can find enormous opportunities for in the area of vaccine design, targeted drug delivery, and diagnostic imaging. Hence, developing methodologies to synthesize protein-dendron stimuli-responsive complexes

In this aspect, we have previously reported the synthesis of protein-dendron bioconjugates, and have systematically studied the effect of dendron size and protein surface charges on the protein self-assembly.¹⁴ This report, for the first time, showed that an extremely hydrophobic and highly branched benzyl-ether dendron could be site-specifically attached to a protein of interest. However, one of the major limitations of that study is that custom-designed nanoassemblies are static and failed to exhibit dynamic behavior. Herein, we extend our chemical methodology for the design of monodisperse photo-responsive facially amphiphilic protein-dendron conjugates. We have used these macromolecular synthons to construct a suite of stimuli-responsive protein assemblies of define sizes. The detailed disassembly studies indicate the designed protein assemblies are not only photo-sensitive but also respond to pH and hydrazine. Besides, we provide detailed mechanistic insight about the disassembly processes.

3.2. Results and Discussion

3.2.1. Macromolecular Design

We have previously reported this technique, where known active-site labeling technology has been utilized to construct monodisperse protein-dendron bioconjugates *via* site-selective labeling of a macromolecular amphiphilic activity-based probe (MAABP) to the active site of various serine proteases.^{14, 15} Although these conjugates were capable of forming supramolecular protein assemblies of defined sizes driven *via* hydrophobic interactions, the complexes failed to exhibit any dynamic properties.

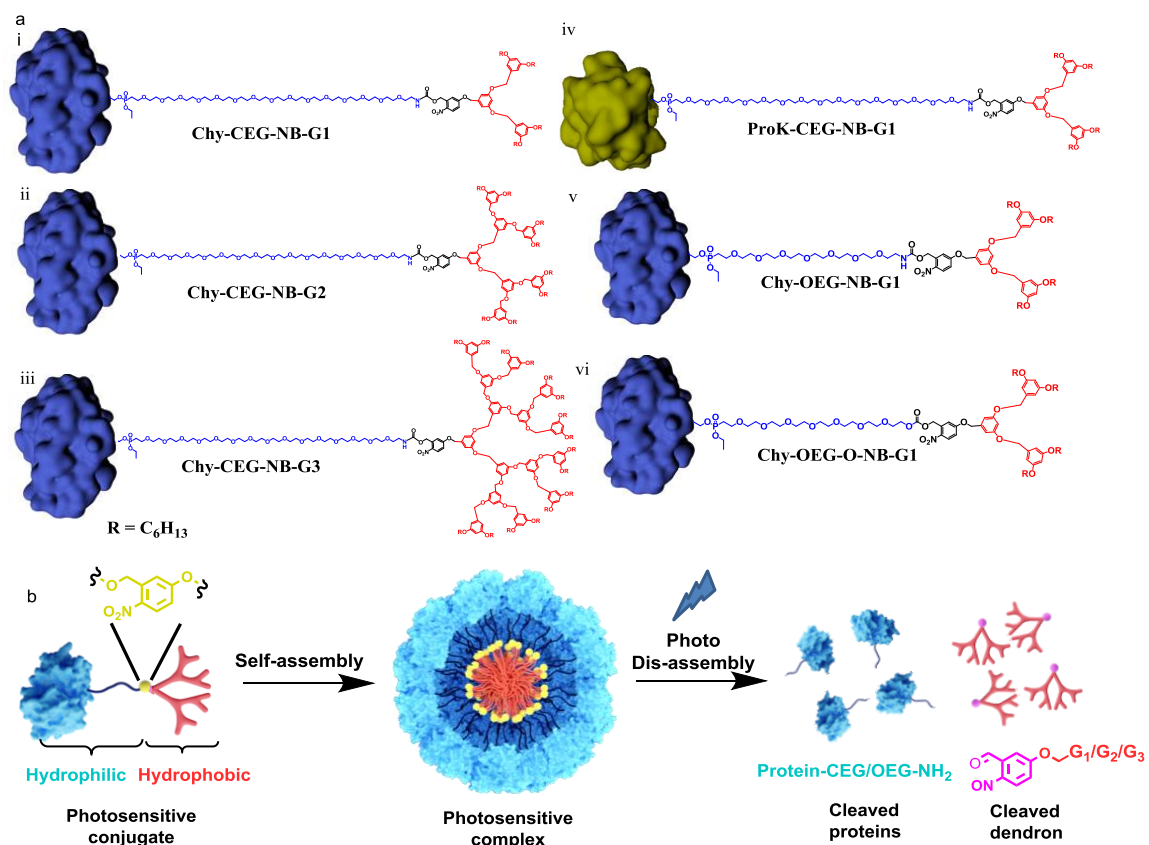


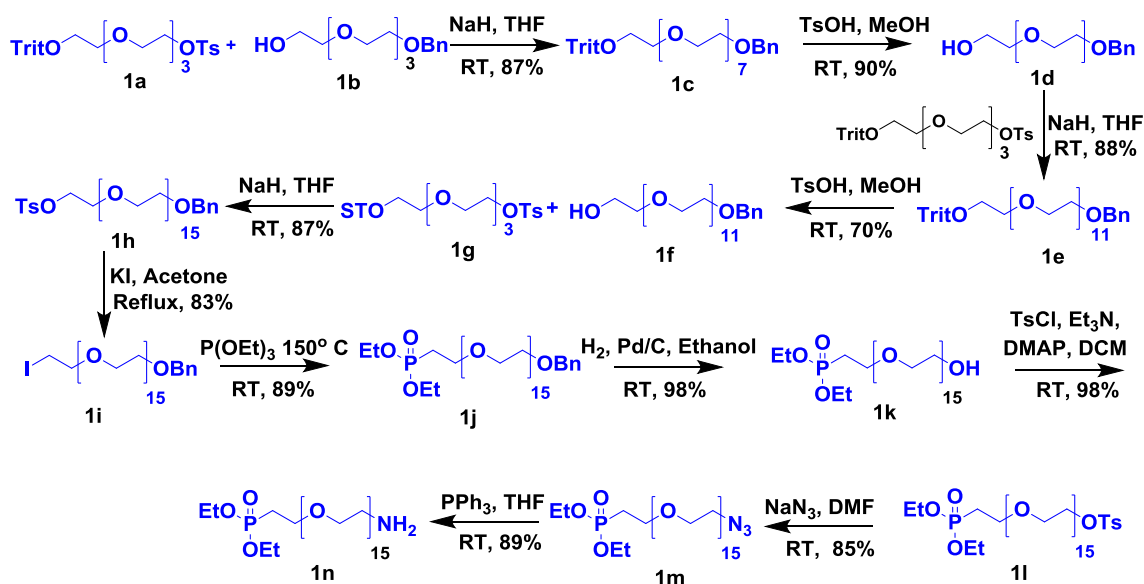
Figure 3.1 | Schematic representation of chemical structures, self-assembly, and dis-assembly of photo-sensitive protein-dendron bioconjugates. These conjugates are composed of protein (blue = chymotrypsin, green = proteinase K), hydrophilic oligoethylene glycol linker (blue) (CEG = cetylene glycol, OEG = octaethylene glycol), photo-sensitive nitrobenzyl group (black), and hydrophobic dendron block (red). (a) Structures of (i) Chy-CEG-NB-G1 (ii) Chy-CEG-NB-G2 (iii) ChyCEG-NB-G3, (iv) ProK-CEG-NB-G1, (v) Chy-OEG-NB-G1, and (vi) Chy-OEG-NB-O-G1. (b)

Schematic representation of self-assembly and irreversible dis-assembly of the photo-sensitive protein-dendron complex upon photo-irradiation.

Herein, we provide a rational and straightforward design for the construction of monodisperse photo-responsive protein-dendron assemblies. The molecular design of photo-responsive bioconjugate is similar to the previous design except for the photo-responsive group (2-nitrobenzyl derivative) is incorporated between the hydrophilic and hydrophobic portion of bioconjugate. The newly designed macromolecular synthons have four core structural elements: (i) hydrophilic globular protein, (ii) flexible hydrophilic linker, (iii) photo-responsive group, and (iv) hydrophobic dendron of different generations (Figure 3.1).

3.2.2. Synthesis of Monodisperse Hydrophilic Linker Having Protein Reactive Group

As mentioned above, the macromolecular photo-sensitive protein-dendron synthons comprise of four core structural elements. We first focused our attention on the construction of hydrophilic linker domain, i.e., amine-terminated diphosphonate ester of unhexa(ethylene glycol)/cetylene glycol (CEG) linker equipped with the protein reactive functional group (Scheme 3.1).



Scheme 3.1 | Scheme for the synthesis of amine-terminated CEG.

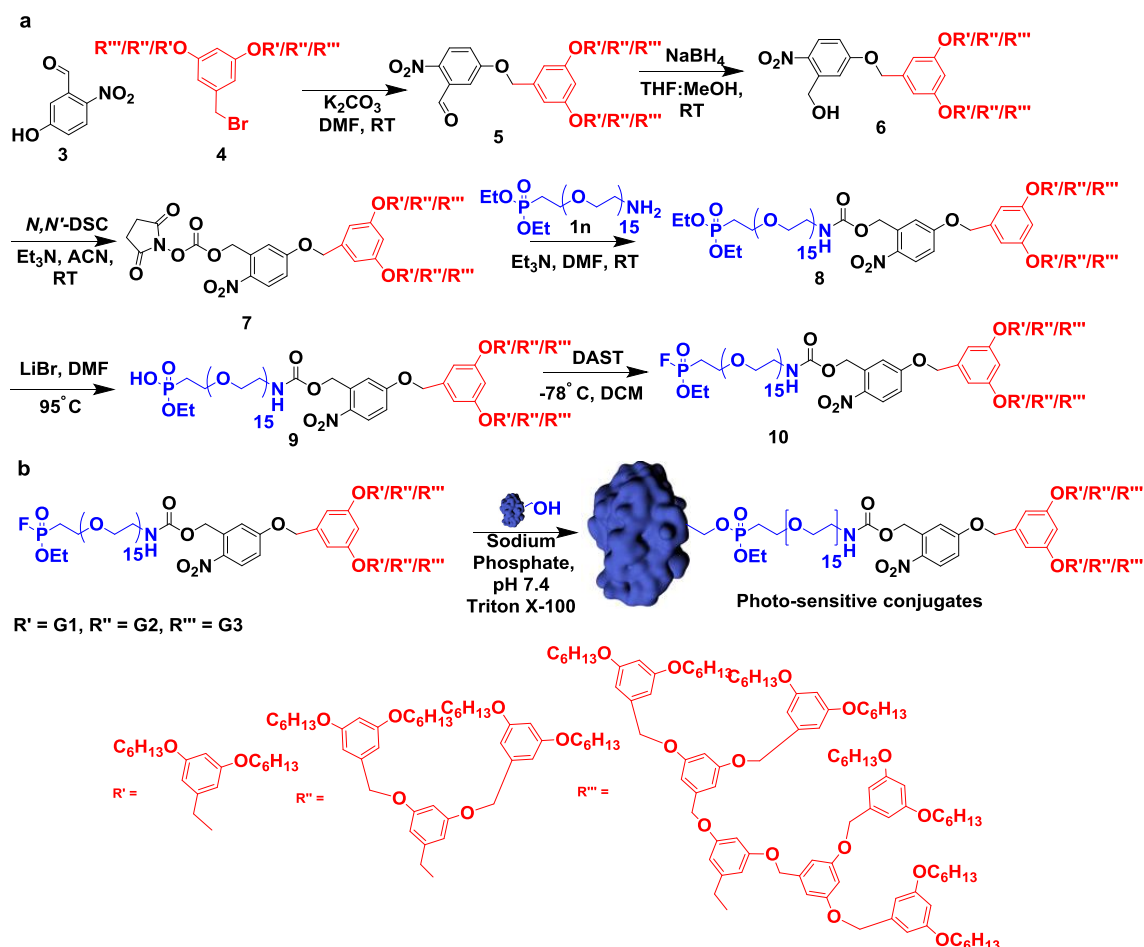
Tosyl derivative of tetraethylene glycol (TEG) (1a) on reaction with monobenzyl TEG (1b) using sodium hydride (NaH) in tetrahydrofuran (THF) gave trityl protected octaethylene glycol (1c). The compound 1c on trityl deprotection using p-toluenesulfonic acid (TsOH) in methanol (MeOH) gave alcohol 1d, which further on treatment with compound 1a in the presence of NaH in THF yielded trityl protected dodecaethylene glycol 1e. Subsequently, trityl deprotection of compound 1e followed by treatment with ditosyl tetraethylene glycol (1g) afforded compound 1h. The obtained tosylate on reflux with KI in acetone yielded corresponding iodo-derivative 1i. Compound 1i on reflux with triethyl phosphite provided compound 1j. Then, the debenzoylation of compound 1j by using hydrogenation reaction in ethanol yielded hydroxyl compound 1k, which upon tosylation in dichloromethane, gave tosylated diphosphonate ester 1l. Compound 1l upon treatment NaN_3 yielded azide 1m. Finally, treating compound 1m with PPh_3 gave compound 1n (Scheme 3.1).

3.2.3. Synthesis of a Photo-Sensitive Amphiphilic Activity-Based Probe

The design strategy for the synthesis of stimuli-responsive macromolecular AABP is shown in Scheme 3.2a. First, we synthesized Fréchet type G1, G2, and G3 dendrons with bromide functionality at the focal point by using a previously reported method.¹⁴ The rationale behind choosing Fréchet type dendrimers are mentioned in chapter 2.2.2. Then, alkylation of G1, G2, or G3 bromide (4) with 5-hydroxy-2-nitrobenzaldehyde (3) in the presence of K_2CO_3 and DMF afforded compound 5. The obtained compound 5 on reduction with sodium borohydride (NaBH_4) yielded alcohol 6, which on treatment with N, N'-disuccinimidyl carbonate (N, N'-DSC) in ACN gave compound 7. The activated ester was then reacted with compound 1n in the presence of Et_3N and DMF to obtain compound 8. Then, the resultant diphosphonate ester 8 was heated in the presence of lithium bromide (LiBr) in DMF to get monophosphonate ester 9, which finally on fluorination using DAST in DCM afforded fluorophosphonate (FP) 10.

Thus, the molecular design of photo-sensitive AABP is composed of FP (reactive head group) functionalized hydrophilic CEG linker, photo-sensitive group, and hydrophobic dendrimer. It is known that the FP group chemo- and site-selectively reacts

with the serine amino acid (part of the catalytic triad of a serine protease). Also, this bioconjugation reaction proceeds under mild reaction conditions with fast reaction kinetics. Inspired by these features, our group has explored this chemistry to design various types of protein assemblies.



Scheme 3.2 | Scheme for the synthesis of photo-sensitive protein-dendron AABPs and bioconjugates. (a) Scheme for the synthesis of photo-sensitive macromolecular AABPs. (b) Scheme for the synthesis of photo-sensitive protein-dendron bioconjugates.

3.2.4. Synthesis of photo-sensitive protein-dendron conjugates

Bioconjugation of photo-sensitive macromolecular AABPs with chymotrypsin was attempted in 50 mM phosphate buffer pH 7.4 using a micelle-assisted protein labeling (MAPLab) technology reported by our group (Chapter 1.3.3.3.).¹⁵ We chose chymotrypsin protein for several reasons: (i) it belongs to classical serine protease family, (ii) it is

commercially available for a reasonable price, and (iii) it is one of well-studied protein. It, therefore, serves as a great model system to probe the hydrophobic effect on protein self-assembly and dis-assembly in an aqueous medium. Next, the extent of bioconjugation reaction was monitored using MALDI-ToF at different time points. After 12 h, a new peak at higher molecular weight is seen in addition to native protein for all the reactions (Figure 3.4a-c).

Subsequently, in order to purify bioconjugate from the reaction mixture, we followed a purification method reported by our group previously.¹⁵ Triton X-100 and unreacted probe in the reaction mixture were removed using ion-exchange chromatography (IEX) (Figure 3.2). Then, to remove native protein from bioconjugate, the ionic strength of the solution was increased. This led to the self-sorting of protein-dendron bioconjugates by the formation of higher-order protein complex driven *via* hydrophobic interaction. Therefore, this complex eluted at lower elution volume, while monomeric native chymotrypsin at eluted later in size-exclusion chromatography (SEC) (Figure 3.3). MALDI-ToF analyses of all the purified protein-dendron bioconjugates showed a single peak indicative of their monodisperse character (Figure 3.4b-d).

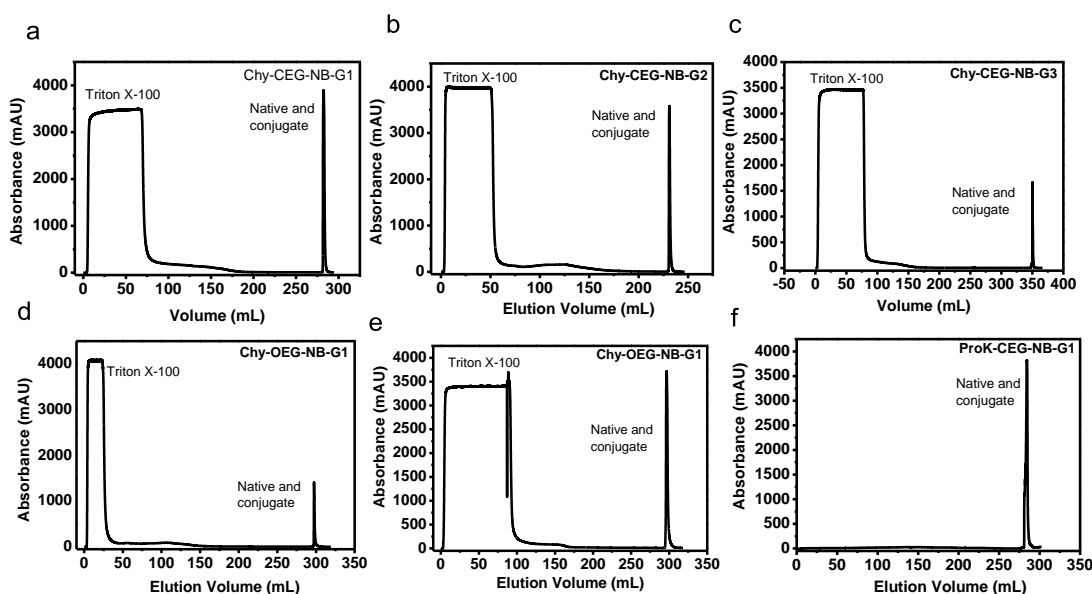


Figure 3.2 | IEX chromatogram. The IEX is performed to remove neutral triton X-100 and excess of macromolecular AABP. IEX of (a) Chy-CEG-NB-G1, (b) Chy-CEG-NB-G2, (c) Chy-CEG-NB-G3, and (d) Chy-OEG-NB-G1, (e) Chy-CEG-NB-O-G1, and (f) ProK-CEG-NB-G1, respectively.

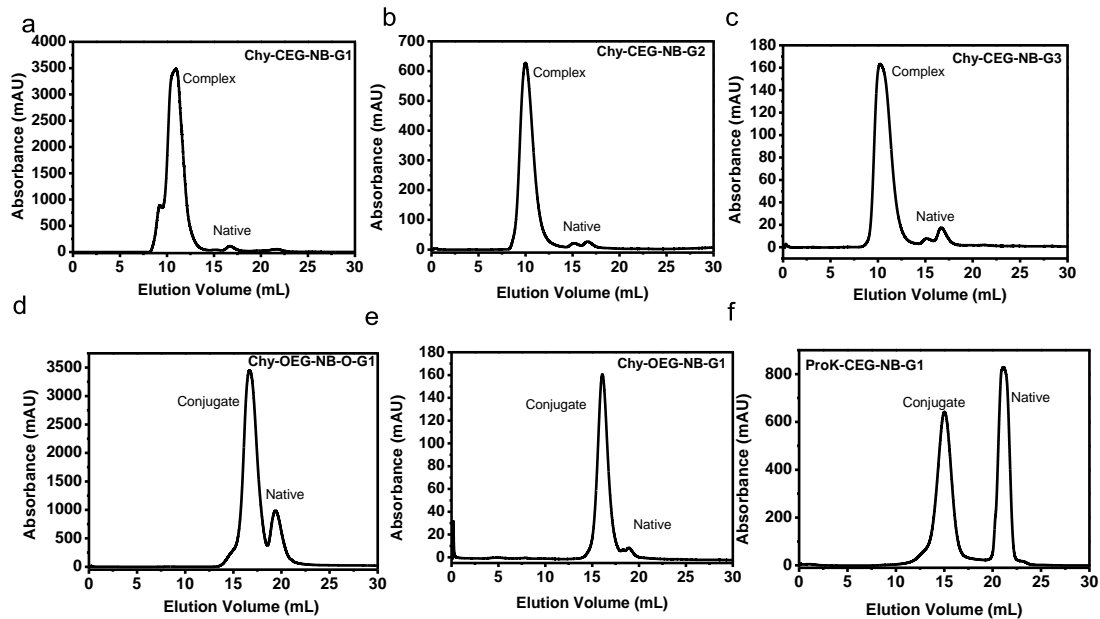


Figure 3.3 | SEC chromatogram. The SEC is performed to separate bioconjugates from native protein. SEC chromatogram of (a) Chy-CEG-NB-G1, (b) Chy-CEG-NB-G2, (c) Chy-CEG-NB-G3, and (d) Chy-OEG-NB-O-G1, (e) Chy-CEG-NB-O-G1, and (f) ProK-CEG-NB-G1, respectively.

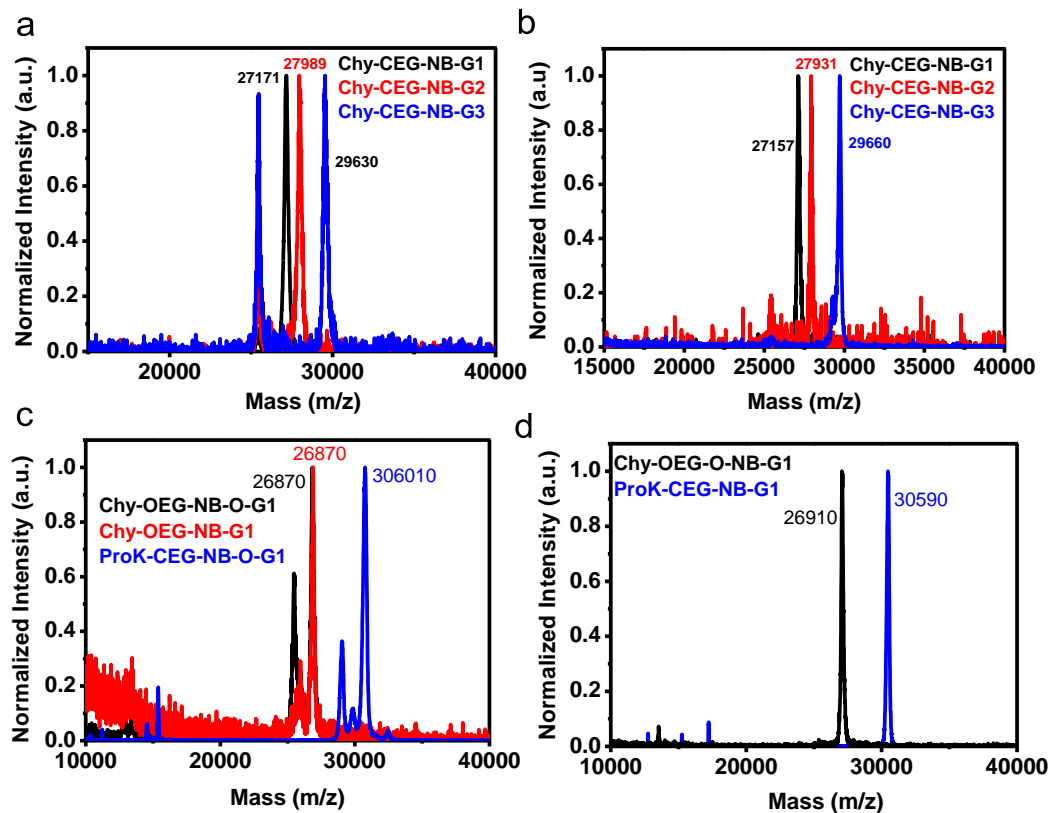


Figure 3.4 | MALDI-ToF characterization photo-sensitive protein-dendron bioconjugates. (a), (c) MALDI-ToF of the reaction mixture, and (b), (d) purified bioconjugates, respectively.

3.2.5. Self-assembly Studies of Photo-Sensitive Protein-Dendron Bioconjugates

After the synthesis of pure monodisperse protein-dendron bioconjugates, we sought to investigate their self-assembly properties by using complementary techniques such as dynamic light scattering (DLS) and SEC. DLS measurements revealed an interesting trend in the hydrodynamic diameter (D_h) of photo-sensitive protein-dendron complexes. The hydrodynamic diameter of the Chy-CEG-NB-G1/G2/G3 complexes was found to be 11, 14, and 18 nm, respectively, indicating that the size of protein-dendron complexes increases with an increase in the size of a dendron (Figure 3.5a). These results are consistent with the previous self-assembly data of protein-dendron bioconjugates (14). The formation of protein nanoassemblies was further verified by using SEC data. We were gratified to observe the same trend in elution volumes (16, 14, 13 mL for Chy-CEG-NB-G1/G2/G3, respectively) by SEC as the D_h of protein assemblies (Figure 3.5b). The above self-assembly data interestingly manifest that, installing a photo-sensitive O-nitro benzyl group between the hydrophilic and hydrophobic portion of bioconjugate does not affect the self-assembly.

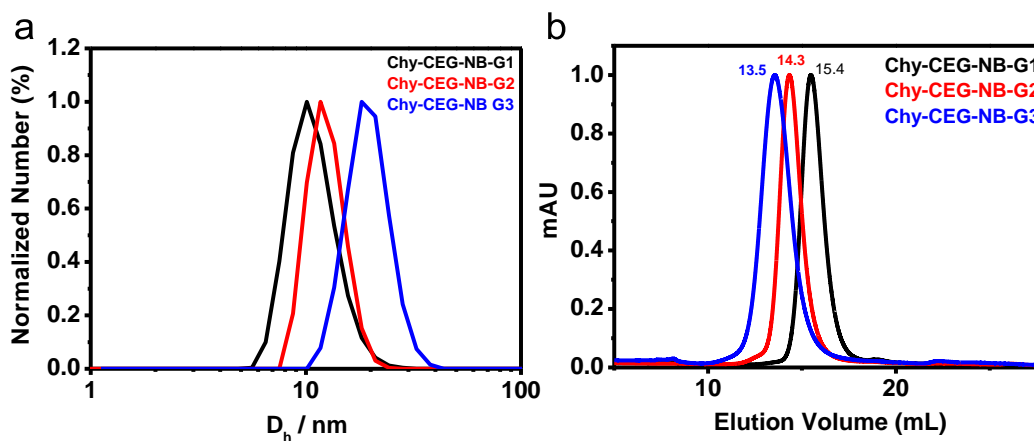


Figure 3.5 | Self-assembly of photo-sensitive protein-dendron bioconjugates. (a) The hydrodynamic diameter of Chy-CEG-NB-G1/G2/G3 complexes is determined using DLS. (b) SEC of Chy-CEG-NB-G1/G2/G3 complexes to determine their elution volume.

3.2.6. Dis-assembly Studies of Photo-Sensitive Protein-Dendron Complexes

After establishing the ability of designed photo-sensitive bioconjugates to self-assemble into nanoassemblies of defined sizes. We set out to examine the capability of custom-designed protein nanoassemblies to disassemble upon photoirradiation. Our hypothesis is, nitrobenzyl group linking hydrophobic (dendron) and hydrophilic domains (protein and linker) of bioconjugate would undergo photolysis reaction upon exposure to light. This would lead to the separation of the hydrophobic domain from the rest of the protein conjugate (Figure 3.1b). As a result, facially amphiphilic protein would convert into hydrophilic protein, which ultimately would lead to disassembly of complex due to the loss of attractive hydrophobic interaction.

To test the hypothesis mentioned above, photo-induced dis-assembly studies were carried out by exposing the Chy-CEG-NB-G1 complex (in 50 mM sodium phosphate pH 7.4) to UV light ($\lambda = 280$ to 360 nm). Subsequently, the photoirradiation sample was then subjected to the SEC in order to investigate the dis-assembly profile. As anticipated, prior to photo-irradiation, the complex eluted at lower elution volume, indicative of intactness of protein nanoassemblies. Serendipitously, the SEC results of light exposed Chy-CEG-NB-G1 complex at different time intervals ($t = 30$ and 60 mins) unveiled the presence of unforeseen peak for intact complex (10 mL in superdex and 16 mL in superpose column) along with expected cleaved monomeric chymotrypsin containing amine-terminated CEG (at 17 mL in superdex and 19 mL in superose column) indicating a partial disassembly of the complex (Figure 3.6a-b). An increase in the UV light exposure time to 4 h did not change the disassembly profile; interestingly, the intensities of cleaved monomer and complex peaks were consistent at all the time points. These results are surprising because this photochemical reaction is extremely fast, and therefore we expect a complete disassembly in that time interval.¹⁶ Next, to examine the perseverance of this dis-assembly behavior across other conjugates, we carried out the disassembly studies of Chy-CEG-NB-G2/G3 complexes in similar conditions and subjected the samples to SEC. Remarkably, the other two bioconjugates behaved similarly except for the ratio of peak 1 to peak 2

progressively increased with respect to dendron generation (G1 to G3) (Figure 3.6c-d, 3.1a).

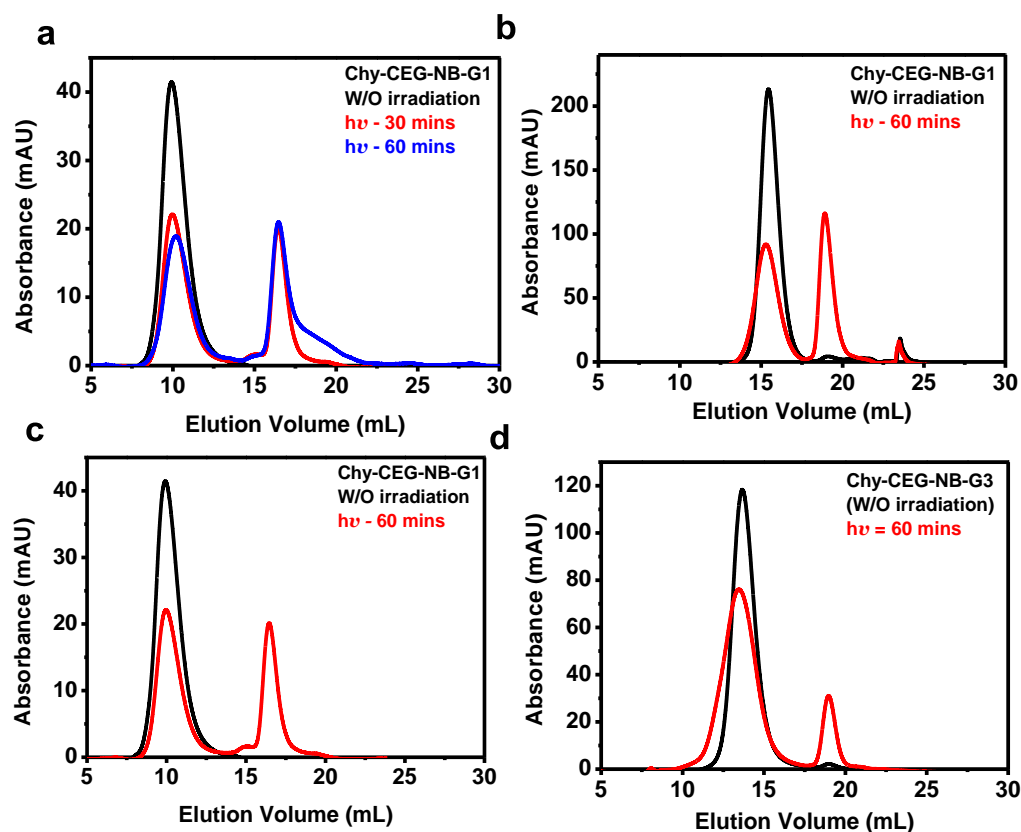


Figure 3.6 | Dis-assembly profile of protein-dendron complexes in 50 mM sodium phosphate pH 7.4 using SEC chromatograms. (a) Time-dependent dis-assembly studies of UV-exposed Chy-CEG-NB-G1(0.5 mg/mL) complex. **(b)** SEC chromatograms of Chy-CEG-NB-G1(2 mg/mL) **(c)** Chy-CEG-NB-G2 (0.8 mg/mL), and **(d)** Chy-CEG-NB-G3 (1.2 mg/mL), W/O UV-exposure (black) and after UV-exposure ($\lambda = 280$ to 360 nm) (red), respectively.

In order to fully understand this serendipitous result, we attempted to characterize each product of the photochemical reaction. To do that, 20 mg Chy-CEG-NB-G1 protein complex was subjected to a photochemical reaction. The light exposed sample was then purified by using SEC to separate intact complex (peak A) and cleaved monomeric protein (peak B) (Figure 3.7a). Further, the isolated intact complex was then again subjected to UV-light irradiation to investigate their photo-responsive property. Surprisingly, the intact complex formed after the first photochemical reaction was not photoreactive, as evident from the very little change in the elution volume in the SEC (Figure 3.7b). Next, the molecular weights of conjugate (part of an intact protein complex, peak A), and monomeric

cleaved protein (peak B) were analyzed using MALDI-TOF. Interestingly, the molecular weight of the monomeric conjugate of peak A was found to be 27160 Da, which is similar to the original Chy-CEG-NB-G1 conjugate, i.e., 27170 Da. In comparison, the molecular weight of the second peak was found to be 26264 Da, which corresponds to the theoretical molecular weight of cleaved monomeric protein, i.e., Chy-CEG-NH₂ (Figure 3.7c-d).

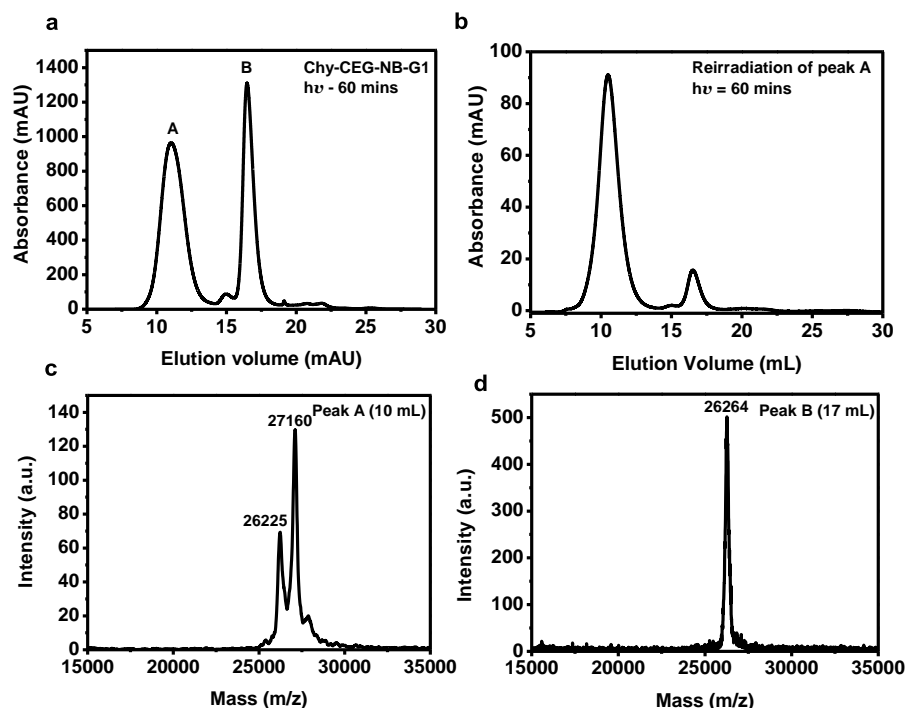


Figure 3.7 | Characterisation of amine-terminated cleaved protein and intact conjugate. (a) SEC chromatogram of UV-exposed Chy-CEG-NB-G1 complex at 20 mg/mL and purification of cleaved proteins (peak B) and intact complexes (peak A). (b) Purified intact protein complex (peak A) was again subjected to UV-exposure, and the exposed sample was subjected to the SEC to investigate the photo-responsive property of the complex. (c) Molecular weights of conjugate forming peak A, and (d) amine-terminated cleaved protein (peak B) using MALDI-ToF spectrometry.

It is apparent from the above studies that although the molecular weight of peak A conjugate is similar to the original conjugate (Chy-CEG-NB-G1), the resultant complex formed after first photochemical reaction is no longer photo-active. The above results manifest two things: (i) the first thing, the partial disassembly is not because of incomplete photolysis reaction, (ii) the second is the photo-sensitive conjugate undergone some rearrangement reaction. Consequently, a new complex is formed, whose elution volume

and molecular weight of its corresponding monomeric conjugate closely match with the starting conjugate, i.e., Chy-CEG-NB-G1.

In order to understand the mechanistic details of these serendipitous results, we carefully examined the reported photolysis mechanism of nitrobenzyl group in the literature. The many groups have proposed that the photolysis mechanism proceeds by the formation of an excited nitro group *via* intramolecular hydrogen abstraction from carbon-hydrogen bond ortho to the nitro group, followed by re-organization of the electron to form aci-nitro derivative. This species then undergoes rapid rearrangement reaction to form aldehyde functionalized nitroso derivative, amine, and CO₂.¹⁶

In our case, the aldehyde functionalized nitroso derivative is appended to the hydrophobic dendron part, and the newly formed amine is hydrophilic (Figure 3.1a). We hypothesized that there could be two competing pathways happening simultaneously after the cleavage of nitrobenzyl functionality; (i) the first possibility is the dissociation of monomeric protein from the protein assembly because it lacks hydrophobic dendron domain and, (ii) second possibility is the reaction of primary amine (part of linker) with nitrosobenzaldehyde derivative (dendron part) to form an imine derivative. The in-situ formation of the imine compound preserves the amphiphilic nature of the bioconjugates and hence prevents further disassembly (Figure 3.8).

However, it is well known that formation of imine reaction is unfavorable in the aqueous medium due to the loss of a water molecule.¹⁷ This is due to the entropic factors, i.e., the formation of one product molecule from two reactant molecules in solution will result in a high loss of entropy. On the contrary, confinement of reacting partners in synthetic receptors, capsules, or micelles has been shown to have a profound influence on the kinetics of reaction by surmounting the entropic problems.¹⁸ Thus, in our case, the photochemical disassembly reactions are occurring in nano-confinements. Hence, the reactions take place under crowded conditions that can increase the local concentration (i.e., aldehyde and amine) of both the reactants and, therefore, the availability of reaction partners.

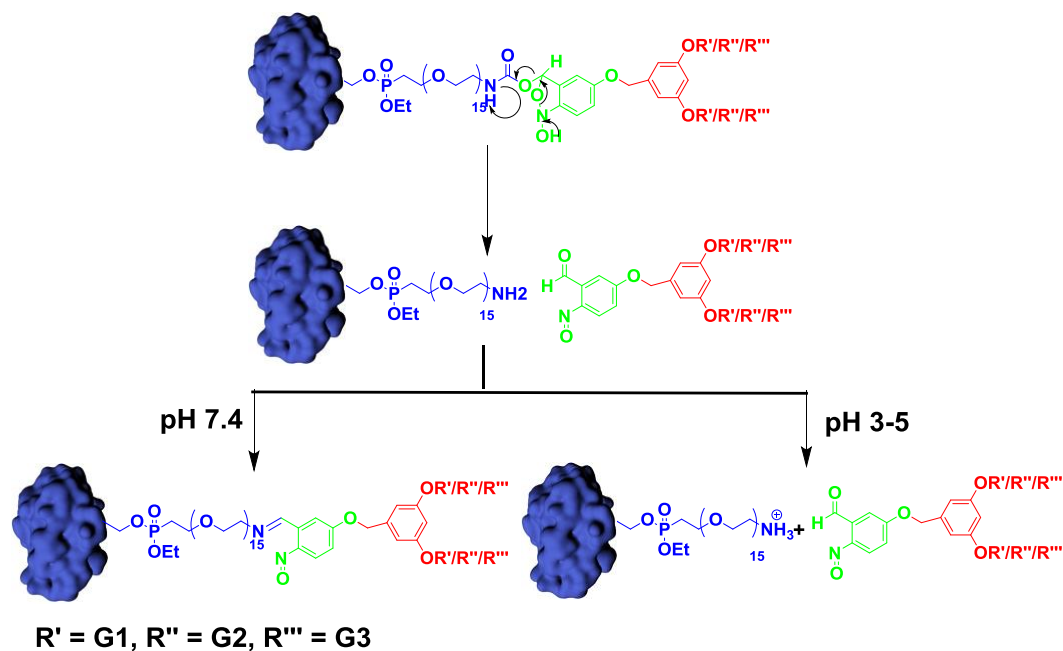


Figure 3.8 | Mechanism for Photolysis and imine formation. Photolysis of the o-nitrobenzyl group would result in the formation of 2-nitroso benzaldehyde derivative and protein functionalized by amine-terminated oligoethylene glycol. The newly formed amine-functionalized protein reacts with 2-nitroso benzaldehyde derivative at pH 7.4 to form new protein-dendron conjugate bearing imine bond. The formation of the new complex is inhibited by arresting the nucleophilicity of the newly formed amine *via* protonating it at lower pH (pH 3).

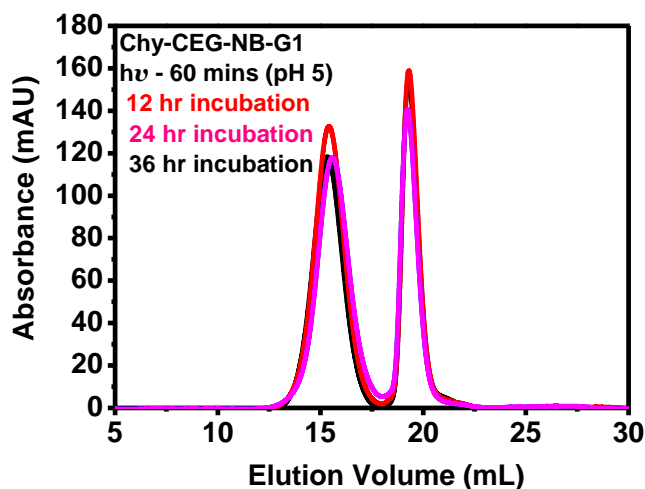


Figure 3.9 | SEC chromatograms of UV-exposed Chy-CEG-NB-G1 complex in 50mM sodium phosphate, pH 5 at different time points.

Next, in order to evaluate the formation of imine in nanoconfinement, we performed the disassembly studies at lower pH. It is well known that the hydrolysis of imines can be

accelerated by protonation of nitrogen in imine (Figure 3.8).¹⁹ Hence, we assayed disassembly studies of the Chy-CEG-NB-G1 complex at pH 5 (Figure 3.9). Interestingly, the disassembly profile at pH 5 was similar to the disassembly at neutral pH (Figure 3.9). This behavior signifies that imines in the micellar environment are stable at pH 5, unlike imines in bulk solution.

Next, we evaluated the disassembly studies of the complex at pH 3. In order to do that, first, we confirmed the stability of the native complex, i.e., Chy-CEG-NB-G1 at pH 3 by using SEC. As anticipated, the assembly is intact; however, there was a slight increase in the size of the assembly as evident from elution volume in the SEC (Figure 3.10).

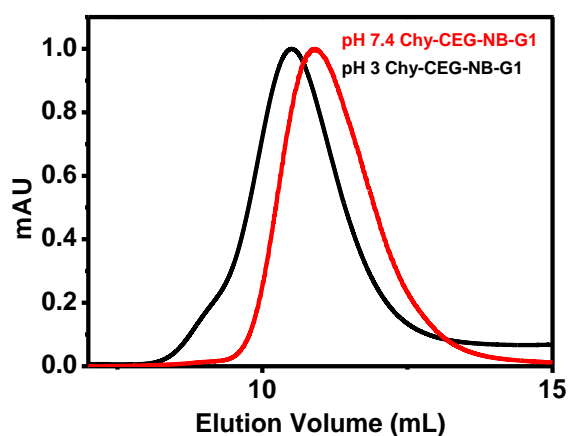


Figure 3.10 | SEC chromatogram of Chy-CEG-NB-G1 at pH 7.4 (red) and pH 3 (black).

After that, we carried out disassembly studies at pH 3. Surprisingly, the complete disappearance of the complex peak (16 mL) was observed, indicating complete disassembly of the Chy-CEG-NB-G1 complex (Figure 3.12a). Interestingly, a new peak at 25 mL is emerged along with a monomeric cleaved protein peak (19 mL). We believe that the new peak is most likely the cleaved dendritic molecule. Next, we subjected other bioconjugates, i.e., Chy-CEG-NB-G2/G3 to photo-irradiation under similar conditions. Interestingly, we observed complete disassembly of these complexes as well, similar to the Chy-CEG-NB-G1 complex (Figure 3.12b-c).

It is apparent from the above disassembly studies (at pH 7.4 and 3) that two competing mechanisms are possible in the nanoconfinement, which determines the

complete or partial disassembly of the photo-sensitive complex (Figure 3.6a-d, 3.12a-c). The first possibility is, the liberated new amine might react with aldehyde and form an imine-based conjugate. While in the second possibility, protonation of liberated amine might occur before it reacts with an aldehyde to form hydrazones (Figure 3.8). Interestingly, at neutral pH, due to the presence of a hydrophobic micro-environment and the amplified nucleophilicity of amine, the hydrazone formation outcompetes the protonation step. However, at lower pH, due to the increased local concentration of H^+ ion, kinetics, the protonation step could be faster than the hydrazone formation. As a result, masked nucleophilic amine will not react with an aldehyde. This increases the probability of dissociation of monomeric protein from the protein assembly. Thus, the above studies manifest that these photo-sensitive complexes sequentially respond to light and pH, respectively (Figure 3.13).

An alternative way to achieve complete disassembly would be to trap the reactivity of newly formed aldehyde. To do that, we envisioned performing disassembly studies in the presence of an excess of hydrazine, which would outcompete the reaction of primary amine (from lysine side chains) with a liberated aldehyde to form hydrazone (Figure 3.11). This is postulated on the fact that hydrazones possess greater intrinsic stability than imines.²⁰

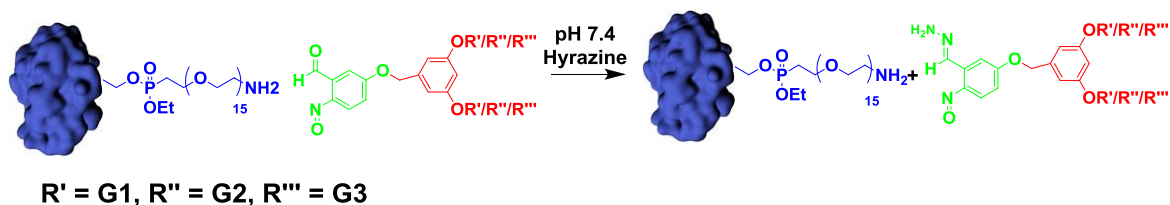


Figure 3.11 | Schematic representation of scheme for inhibiting imine formation by hydrazine. The reactivity of newly formed aldehyde can be arrested by carrying out photochemical reactions in the presence of hydrazine. The hydrazine outcompetes the reaction of linker amine with aldehyde and form hydrazone.

In order to test this hypothesis, photochemical disassembly studies of Chy-CEG-NB-G1/G2/G3 complexes are carried out in the presence of hydrazine (200 eq). Surprisingly, all the complexes were completely disassembled as evident by the disappearance of the peak corresponding to the intact complex in SEC (Figure 3.12d-f). In addition, we separated the intermediate imine protein complex obtained after the photochemical reaction and subjected this sample to reaction with hydrazine. As expected,

we observed complete disassembly (Figure 3.14). This data suggests again supports our hypothesis described in the last section.

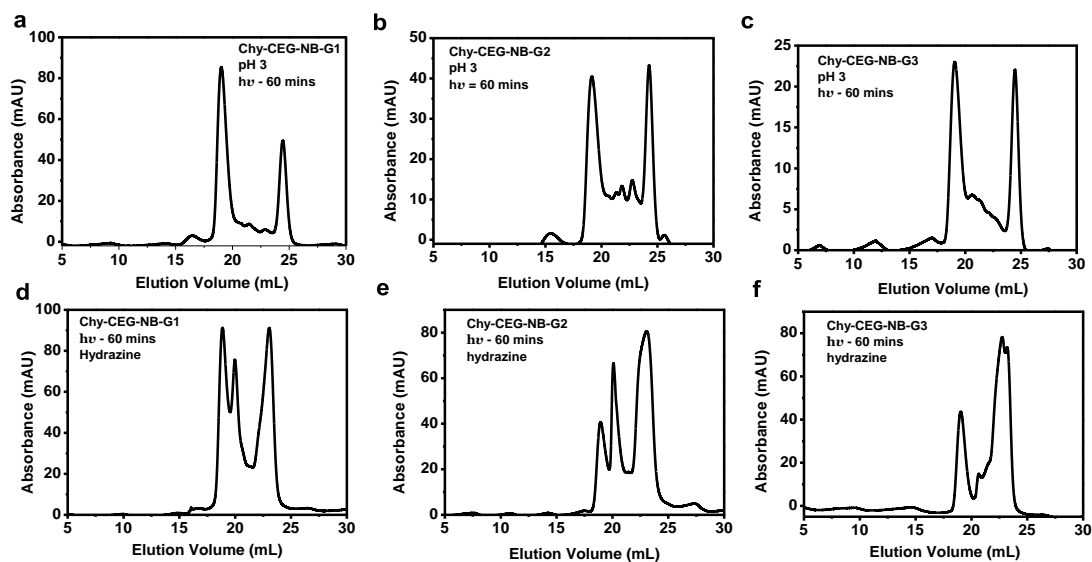


Figure 3.12 | Dis-assembly of protein-dendron complexes at pH 3 and pH 7.4 in the presence of hydrazine. SEC chromatograms of UV-exposed (a) Chy-CEG-NB-G1, (b) Chy-CEG-NB-G2, (c) Chy-CEG-NB-G3, at pH 3, respectively. Dis-assembly of at pH 7.4 in the presence of hydrazine (200 eq) (d) Chy-CEG-NB-G1, (e) Chy-CEG-NB-G2, (f) Chy-CEG-NB-G3, respectively.

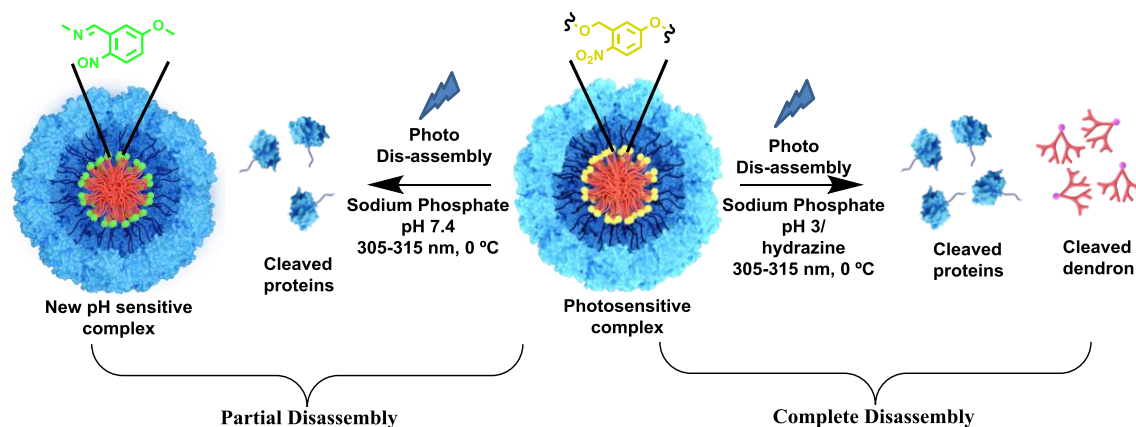


Figure 3.13 | Schematic representation of disassembly of photo-sensitive protein-dendron at pH 7.4, pH 3, and in the presence of hydrazine at pH 7.4.

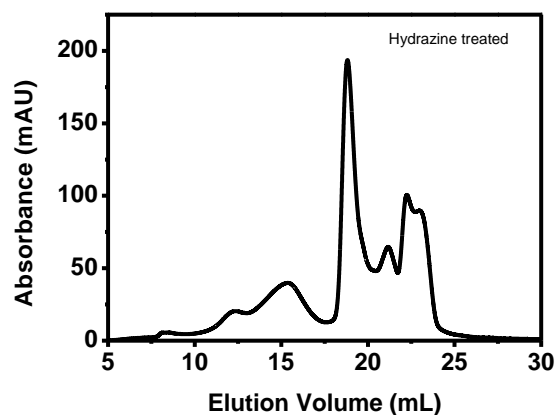
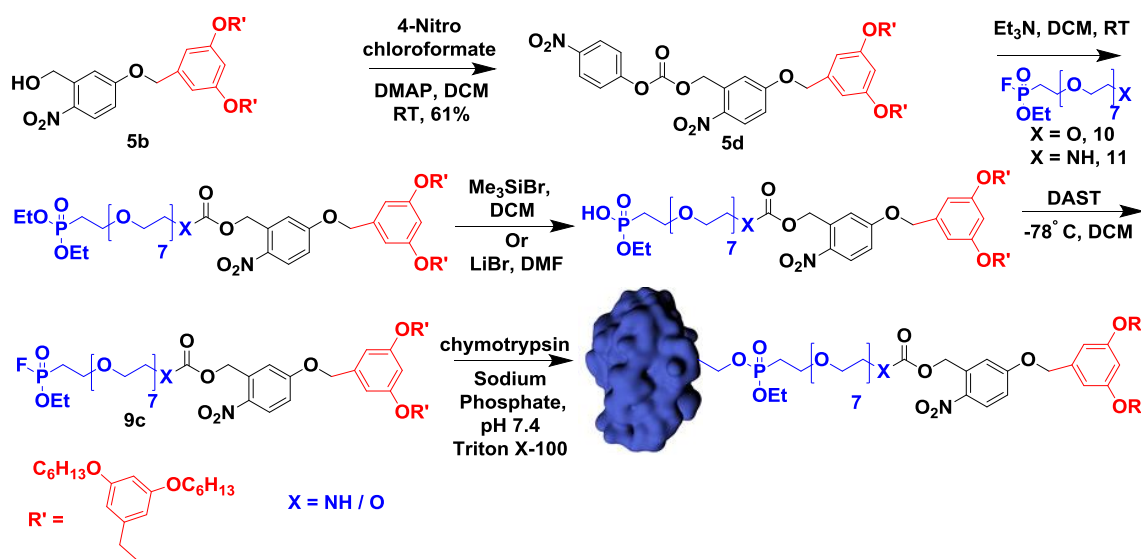


Figure 3.14 | Disassembly profile of hydrazone incubated imine complex. The photoirradiation was first carried out on the Chy-CEG-NB-G1 complex. Further, the intact complex (imine) was purified using SEC, and the purified complex was incubated with hydrazine (200 eq) for 1 hour. The sample was then subjected to SEC.

Are amine functionality from the various lysine side chain of a protein involved also in imine formation? To answer this question, we synthesized two new bioconjugates, Chy-OEG-NH-NB-G1 and Chy-OEG-O-NB-G1 (Figure 3.1a, 3.2d-e, 3.3d-e) (Scheme 3.3). Both these conjugates contain the same linker length (octaethylene glycol), protein (chymotrypsin), and dendrimer (G1) (Figure 3.1a). However, we synthetically mutated carbamate functionality, which connects a hydrophilic and hydrophobic portion of bioconjugates into carbonate functionality. We hypothesized that the conjugate bearing carbonate functionality, i.e., Chy-OEG-O-NB-G1, should completely disassembly in response to light, provided if the liberated aldehyde does not react with amine from lysines of protein. This is postulated on the fact that, upon photoirradiation, the complex would disassemble to yield alcohol rather than amine, which will not react with aldehyde under these reaction conditions. Interestingly, both of these conjugates exhibit similar disassembly profile, indicating that the liberated nitrosobenzaldehyde derivative can also react with lysine residues of the protein surface, which are in close proximity because of nanoscale confinement. (Figure 3.15a).

To support the above argument, we envisioned reducing the linker length in the molecular design, which would lead to a decrease in the proximity between the lysine side chain and the liberated aldehyde functionality. The decrease in proximity should increase in the probability of imine formation, and that should reflect in the disassembly profile. To

test this hypothesis, we performed disassembly studies of Chy-CEG-NB-G1 and Chy-OEG-NB-G1 (Figure 3.1a). To our delight, the complex with a shorter linker, i.e., Chy-OEG-NB-G1 shown the increased imine-based complex formation upon photoirradiation compared to the complex with longer linker (Chy-CEG-NB-G1) (Figure 3.15b). Further, the effect of protein surface charge on disassembly was investigated by subjecting proteinase K (ProK-CEG-NB-G1) complex to photoirradiation and compared with the disassembly of Chy-CEG-NB-G1 (Figure 3.1a, 3.15c). Here, both the conjugates behaved in a similar fashion, indicating imine formation is independent upon the identity of protein.



Scheme 3.3 | Scheme for the synthesis of carbonate/ carbamate functionalized photo-sensitive macromolecular AABPs with octaethylene glycol linker and corresponding protein-dendron bioconjugates.

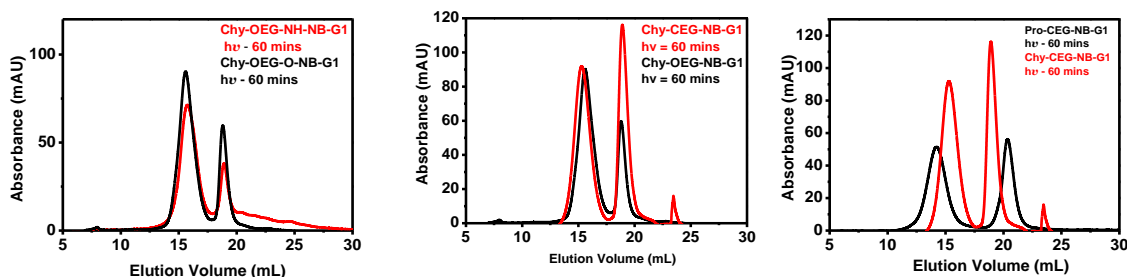


Figure 3.15 | Effect of linkage, linker, and protein surface on dis-assembly of protein-dendron bioconjugates. Dis-assembly profile of (a) CEG-OEG-NB-G1 (red) and Chy-OEG-NB-O-G1 (black), (b) CEG-CEG-NB-G1 (red) and Chy-OEG-NB-G1 (black), and Chy-CEG-G1 (red) and ProK-CEG-NB-G1 (black).

3.3. Conclusion

To best of our knowledge, this is the first detailed mechanistic study of the disassembly processes of a nearly monodisperse multi-responsive protein-based supramolecular system. We have shown that our simple molecular design allows precise installation of the photo-responsive group between the dendron and the linker part. The presence of a photo-responsive group did not adversely affect the self-assembling properties of protein-dendron bioconjugates. Detailed studies reveal that this supramolecular system is not only responding to light but also pH and a small molecule. A detailed mechanistic study reveals that the imine formation is the most likely reason for the behavior of partial disassembly. The ability to design protein assemblies with high precision using an organic chemistry approach offers a novel way to construct multi-responsive supramolecular protein assemblies and also provides opportunities regarding the non-covalent forces that drive both assembly/disassembly processes. One can envision implications of this study in different fields such as controlled drug release, targeted drug delivery, etc. which are current foci of our laboratories.

3.4. Experimental Methods

3.4.1. Synthesis and Characterization of Photo-Sensitive Protein-Dendron Bioconjugates.

3.4.1.1. Protein Conjugation

The bioconjugation reaction of chymotrypsin and proteinase K was performed using the procedure reported by our group. Briefly, a mixture composed of photo-sensitive macromolecular AABPs (1 or 2 equivalent) and 20 μL triton X-100 was sonicated till it becomes homogenous. Then, 480 μL of 50 mM sodium phosphate pH 7.4 was added to the above obtained homogenous solution and vortexed. To this, 500 μL of 200 μM of protein solution (chymotrypsin or proteinase K dissolved in 50 mM sodium phosphate pH 7.4) was added. This reaction mixture was then allowed to react for 24 h on rotor spin at 20 rpm at 25 $^{\circ}\text{C}$.

3.4.1.2. Purification of Photo-Sensitive Protein-Dendron Bioconjugates

All the photo-sensitive protein-dendron conjugates were purified by three-step purification, *i.e.*, IEX, SEC, and desalting. Briefly, neutral triton X-100 and unreacted probe were removed by IEX by using sepharose or Q sepharose resins (GE). In cation-exchange chromatography, the column was pre-equilibrated with the 50 mM sodium phosphate pH 7.4 buffer. Then the reaction mixture was injected, followed by post-injection equilibration until the complete removal of triton X-100. Subsequently, the elution of native protein and corresponding photo-sensitive protein-dendron conjugate together as the mixture was later achieved using 50 mM sodium phosphate pH 7.4, 1 M NaCl as elution buffer.

In the case of anion-exchange chromatography, the protein reactions were buffer-exchanged first to 50 mM tris base pH 10. Then the anion-exchange column was pre-equilibrated using the same buffer (50 mM tris base pH 10), which was used for buffer-exchange. Then the sample was injected, followed by post-injection equilibration until the complete removal of triton X-100. Subsequently, the elution of the native protein and its protein-dendron bioconjugate together as the mixture was later achieved using 50 mM tris base pH 10, 1 M NaCl as elution buffer.

The obtained IEX fractions were subjected to SEC in order to separate native protein from bioconjugates. 50 mM sodium phosphate pH 7.4, 200 mM NaCl was used as buffer using either Superdex-200 10/300 or Sephacryl-200, GE Healthcare column.

3.4.1.3. Matrix Preparation for MALDI-ToF

4 mg of DHAP and 4.5 mg of DAHC were weighed separately in microcentrifuge tubes. 150 μ L of ethanol and 200 μ L of milli-Q water were added to each tube, respectively. Both the solutions were sonicated using bath sonicator for 1 minute and vortexed for another minute. Then 50 μ L of DAHC aqueous solution was transferred to the DHAP solution, and the resulting solution was vortexed for one more minute to yield matrix mixture.

3.4.1.4. Monitoring of Modification

Molecular weights of photo-sensitive protein-dendron conjugates were analyzed by using MALDI-ToF spectrometry (citation). In brief, 2 μL of the reaction mixture or 100 μM purified protein conjugate was mixed with 2 μL of 2% TFA and 2 μL of matrix mixture (1.1.2), vortexed and spotted on MALDI-ToF MS plate. The plate was then loaded and fired to get accurate molecular weight both in +1 and +2 states. 100 μM protein concentration was found to be optimum for MALDI-ToF MS analysis.

3.4.2. Self-Assembly

3.4.2.1. Dynamic Light Scattering

The hydrodynamic diameter of photo-sensitive protein-dendron assemblies was measured using DLS (Zetasizer Nano 2590, Malvern, UK). Samples (5 mg/mL) were prepared in 50 mM sodium phosphate pH 7.4. 1 mL of sample was taken in disposable polystyrene cells, and then the mean size of the complexes was measured at 90° scattering angle.

3.4.2.2. Size Exclusion Chromatography

The elution volume of photo-sensitive protein-dendron assemblies in SEC was determined using Akta Pure and Superose-200 10/300 GE Healthcare column. In order to do that, IEX fractions of bioconjugates were first subjected to the SEC in order to separate native protein from bioconjugates. Then, the 500 μL of the center fraction of the complex peak was again subjected to SEC by using the pre-equilibrated column mentioned above in 50 mM sodium phosphate pH 7.4, 200 mM NaCl.

3.4.3. Programmed Dis-assembly of a Photo-Sensitive Supramolecular Protein Complex

3.4.3.1. Dis-Assembly at pH 7.4

The photo-sensitive protein-dendron conjugates were dissolved in 50 mM sodium phosphate pH 7.4. The samples were then exposed to ultraviolet rays ($\lambda = 280$ to 360 nm) using Sankyo Denki G8T5E UVB linear lamp (8 W output) for 60 minutes at 0 °C. Further, 500 μ L of UV light-exposed sample was withdrawn and subjected to SEC using Superose or Superdex-200 10/300 GE Healthcare column to determine the dis-assembly.

3.4.3.2. Characterization of Cleaved Protein and Intact Conjugate

Irradiation studies with the Chy-CEG-NB-G1 complex at 20 mg/mL under similar conditions. The light exposed sample was then purified by using SEC using Superdex-200 10/300 column to separate intact complex (peak A) and cleaved monomeric protein (peak B). Further, the separated intact complex was then again subjected to ultraviolet rays ($\lambda = 280$ to 360 nm) irradiation for 60 minutes. The irradiated sample was then again subjected to SEC using Superdex-200 10/300 column to investigate their photo-responsive property. Next, the molecular weight of conjugate forming intact complex (peak A), and monomeric cleaved protein (peak B) was probed using MALDI-TOF.

3.4.3.3. Dis-Assembly at pH 5 and 3

The pH of 50 mM sodium phosphate (pH 7.4, 200 mM NaCl) containing protein-dendron conjugates was adjusted to either pH 5 or pH 3 by the dropwise addition of $2N$ HCl. These samples were then exposed to ultraviolet rays ($\lambda = 280$ to 360 nm), as mentioned above, for 60 minutes at 0 °C. Subsequently, 500 μ L of UV light-exposed sample was subjected to the SEC using Superose-200 10/300 GE Healthcare column in order to investigate the dis-assembly.

3.4.3.4. Stability of Photo-Sensitive Protein-Dendron Complexes at pH 3

The pH of 50 mM sodium phosphate (pH 7.4, 200 mM NaCl) containing the Chy-CEG-NB-G1 complex was adjusted to either pH 3 by the dropwise addition of $2N$ HCl. The sample was then subjected to SEC using Superdex-200 10/300 GE Healthcare column.

3.4.3.5. Hydrazine and Light-Induced Dis-Assembly at pH 7.4

To the solution of protein-dendron conjugates in 50 mM sodium phosphate (pH 7.4, 200 mM NaCl), 200 eq of hydrazine solution was added. This solution was then subjected to the ultraviolet rays ($\lambda = 280$ to 360 nm) for 60 minutes at 0 °C. Further, these UV light-exposed samples were subjected to the SEC using Superose-200 10/300 GE Healthcare column in order to investigate the dis-assembly.

3.4.3.6. Hydrazine Induced Dis-Assembly

The photo-sensitive protein-dendron conjugates were dissolved in 50 mM sodium phosphate pH 7.4. The samples were then exposed to ultraviolet rays ($\lambda = 280$ to 360 nm). Further, 500 μ L of UV light-exposed sample was withdrawn and subjected to the purification using SEC using Superose-200 10/300 GE Healthcare column (Fig4a). Subsequently, the purified peak I was incubated with 200 eq of hydrazine for 2 hours. This mixture was then subjected to the SEC in order to investigate the dis-assembly.

3.4.4. Synthesis and Characterization of the Amine-Terminated Linker and Macromolecular AABPs.

3.4.4.1. General

All reagents were obtained commercially unless and otherwise stated. Reactions were performed in an oven-dried round bottom flask (RBF) and under nitrogen atmosphere. Air and moisture sensitive solvents were transferred *via* syringe. Reactions were monitored by thin-layer chromatography (TLC), and the developed chromatogram was visualized by ultraviolet (UV) lamp or by phosphomolybdic acid (PMA) staining. Product purification was accomplished by 100-200 mesh size silica gel column chromatography.

All the compounds were characterized by ^1H , ^{13}C , and ^{19}F (in case of fluorinated compounds) nuclear magnetic resonance (NMR) using Bruker or Jeol 400 MHz. ^1H and ^{19}F were recorded at an operating frequency of 400 MHz and 100 MHz for ^{13}C using, using TMS as an internal standard. All the ^{13}C Chemical shifts were mentioned in parts per million (PPM) and measured relative to residual CHCl_3 , CH_3OH , or CH_3CN in their deuterated solvent. Coupling constants were reported in Hertz (Hz). Multiplicities were

explained as s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, quint = quintet. Mass spectra were obtained with either the MALDI-TOF MS or HRMS. Room temperature varied between 21-35 °C.

3.4.4.2. Synthesis of Amine-Terminated Cetylene and Octaethylene Glycol Spacer

3.4.4.2.1. General Procedures for the Synthesis of Trityl Protected Monobenzyl Oligoethylene Glycol - Procedure A

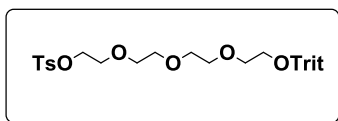
In an oven-dried RBF, monobenzyl oligoethylene glycol (1 eq) and tosyl tetraethylene glycol (2 eq) was dissolved in THF under stirring. Then, sodium hydride (NaH) (4 eq) was added in a small portion at 0 °C. The reaction was allowed to react for 12 hours at RT. Upon completion, excess of NaH was quenched by the dropwise addition of water, and the reaction mixture was extracted with DCM for thrice. The combined organic layer was dried over sodium sulphate (NaSO₄) and concentrated under reduced pressure to get the crude residue, which was purified using silica gel column chromatography.

3.4.4.2.2. General Procedure for the Synthesis of Trityl Deprotection - Procedure B

In an oven-dried RBF, a mixture of monotrityl oligoethylene glycol (1 eq) and p-toulenesulfonic acid (TsOH) (1.5 eq) was taken and dissolved in methanol under stirring. The mixture was allowed to react for 12 hours at RT. Upon completion, methanol was evaporated under reduced pressure. To the obtained residue, water was added and extracted thrice in DCM. The combined organic layer was dried over sodium sulphate (NaSO₄) and concentrated under reduced pressure to get the crude residue, which was purified using silica gel column chromatography.

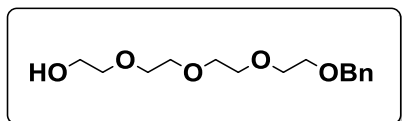
3.4.4.2.3. Synthesis of Cetylene Glycol Spacers and Their Intermediates

Synthesis of compound 1a



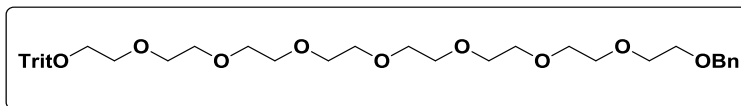
Monotrityl tetraethylene glycol (58 g, 132 mmol) was dissolved in THF under stirring. To the above solution, aq. solution of potassium hydroxide (KOH) (26 g, 464 mmol) was added and allowed to stir for 10 minutes. Then, a solution of TsCl (75 g, 398 mmol) in THF was slowly added and allowed to react for 12 hours at RT. Upon completion, the reaction was quenched by the dropwise addition of water and extracted with DCM for thrice. The combined organic layer was dried over Na₂SO₄ and concentrated under reduced pressure to get a crude product that was purified using silica gel column chromatography using ethyl acetate / Hexane as eluent. The product was obtained as a pale yellow liquid (66 g, 112 mmol, 85%), *R_f* = 0.40 in 50% Ethyl acetate / Hexane. ¹H NMR (400MHz, CDCl₃): δ_H 7.77(d, *J*=8Hz, 2H), 7.47(d, *J*=8Hz, 6H), 7.32-7.20 (m, 11H), 4.14-4.11(t, *J*=8Hz, 2H), 3.70-3.56(m, 12H), 3.24-3.21(t, *J*=8Hz, 2H), 2.42(s, 3H), **MALDI-TOF MS** (M+K): 629.59.

Synthesis of compound 1b



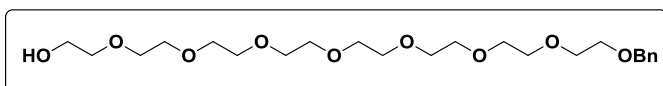
In an oven-dried RBF, tetraethylene glycol (15.0 g, 77 mmol) was dissolved with stirring in THF. Then, NaH (1.23 g, 51 mmol) was added to the flask in small portions at 0° C. After 1 hour, benzyl bromide (6.13 g, 51 mmol) was added dropwise, maintaining the reaction at the same temperature. Then, the reaction was stirred for 12 h at RT. Upon completion, the reaction was quenched by the dropwise addition of water and extracted with DCM for thrice. The combined organic layer was dried over Na₂SO₄ and concentrated under reduced pressure to get a crude product that was purified using silica gel column chromatography using MeOH / DCM as eluent. The product was obtained as pale yellow liquid (4.6 g, 30 mmol, 58%), *R_f* = 0.34 in 5% MeOH / DCM. ¹H NMR (400MHz, CDCl₃): δ_H 7.32-7.26(m, 5H), 4.55(s, 2H), 3.71-3.57 (m, 16H).

Synthesis of compound 1c



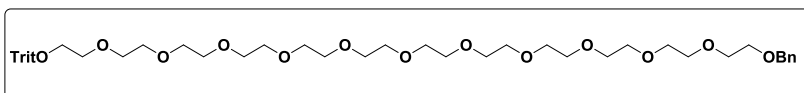
The compound **1c** was prepared by general procedure A, starting from compound **1a** (3.8 g, 6 mmol), compound **1b** (1.1 g, 4 mmol), and NaH (0.3 g, 12 mmol) in THF. The product was obtained as a pale yellow liquid (2.5 g, 3 mmol, 87%) after purification by silica gel column chromatography using MeOH / DCM as eluent, $R_f = 0.47$ in 5% MeOH / DCM. $^1\text{H NMR}$ (400MHz, CDCl_3): δ_{H} 7.48(d, $J=8\text{Hz}$, 6H), 7.36-7.23 (m, 14H), 4.58(s, 2H), 3.70-3.56(m, 31H), 3.27-3.24(t, $J=8\text{Hz}$, 2H), $^{13}\text{C NMR}$ (100MHz, CDCl_3): 144.18, 138.31, 128.77, 128.41, 128.00, 127.81, 126.97, 73.28, 70.72, 70.69, 70.64, 70.61, 69.48, 63.37. **MALDI-TOF MS** ($\text{M}+\text{Na}$): 725.42.

Synthesis of compound **1d**



The compound **1d** was prepared by general procedure B, starting from compound **1c** (2.4 g, 3 mmol) and TsOH (0.5 g, 2 mmol) in MeOH. The product was obtained as a pale yellow liquid (1.4 g, 3 mmol, 90%) after purification by silica gel column chromatography using MeOH / DCM as eluent, $R_f = 0.43$ in 5% MeOH / DCM. $^1\text{H NMR}$ (400MHz, CDCl_3): δ_{H} 7.33-7.26(m, 5H), 4.56(s, 2H), 3.71-3.58 (m, 32H). $^{13}\text{C NMR}$ (100MHz, CDCl_3) δ_{C} 138.21, 128.30, 127.68, 127.53, 73.16, 72.55, 70.57, 70.52, 70.49, 70.23, 69.37, 63.69, 61.58.. **MALDI-ToF MS** ($\text{M}+\text{K}$): 499.23.

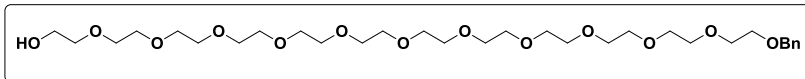
Synthesis of compound **1e**



The compound **1e** was prepared by general procedure A, starting from compound **1d** (0.9 g, 2 mmol), compound **1a** (1.8 g, 3 mmol), and NaH (0.25 g, 10 mmol) in THF. The product was obtained as a pale yellow liquid (1.5 g, 1.7 mmol, 88%) after purification by silica gel column chromatography using MeOH / DCM as eluent, $R_f = 0.47$ in 5% MeOH / DCM. 7.49(d, $J=8\text{Hz}$, 6H), 7.35-7.22(m, 14H), 4.58(s, 2H), 3.70-3.56(m, 47H), 3.27-3.24(t, $J=8\text{Hz}$, 2H). $^{13}\text{C NMR}$ (100MHz, CDCl_3) δ_{C} 144.15, 138.26, 128.75, 128.41,

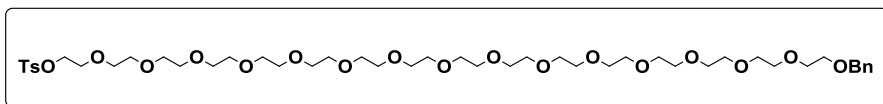
127.81, 127.65, 126.97, 86.58, 73.27, 70.80, 70.71, 70.56, 69.45, 63.36, 53.53. **MALDI-TOF MS** (M+Na): 901.34.

Synthesis of compound **1f**



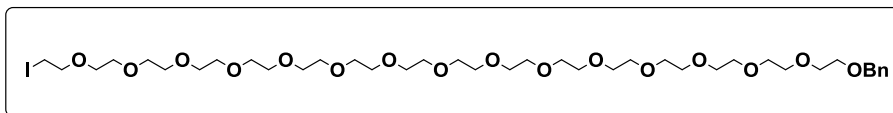
The compound **1f** was prepared by general procedure B, starting from compound **1e** (1.5 g, 1.7 mmol) and TsOH (0.16 g, 0.8 mmol) in MeOH. The product was obtained as a pale yellow liquid (0.98 g, 1.5 mmol, 88%) after purification by silica gel column chromatography using MeOH / DCM as eluent, $R_f = 0.43$ in 5% MeOH / DCM. $^1\text{H NMR}$ (400MHz, CDCl_3): δ_{H} 7.33-7.25(m, 5H), 4.54(s, 2H), 3.70-3.58 (m, 52H). **MALDI-ToF MS** (M+K): 675.28.

Synthesis of compound **1h**



The compound **1h** was prepared by general procedure A, starting from compound **1f** (1 g, 1.5 mmol), **compound 1g** (2.3 g, 5 mmol), and NaH (0.150 g, 6 mmol), The product was obtained as a pale yellow liquid (0.7 g, 0.7 mmol, 70%) after purification by silica gel column chromatography using MeOH / DCM as eluent, $R_f = 0.43$ in 5% MeOH / DCM. **MALDI-ToF MS** (M+K): 1005.23

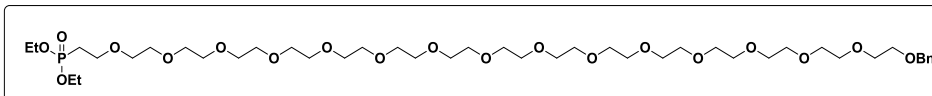
Synthesis of compound **1i**



A mixture of above-obtained tosylate **1h** (0.6 g, 0.6 mmol) and KI (0.3 g, 1.8 mmol) was refluxed in acetone for 18 h. Upon completion, excess KI was filtered and washed thrice with acetone. Collected acetone fraction was evaporated under vacuum to get residue, which was then washed with water and extracted with DCM. The combined organic layer was washed with aqueous Na_2CO_3 and then concentrated under vacuum to get crude product which was purified by silica gel column chromatography using MeOH / DCM as eluent to get pale yellow liquid (0.5 g, 0.5 mmol, 87%), $R_f = 0.45$ in 5% MeOH /

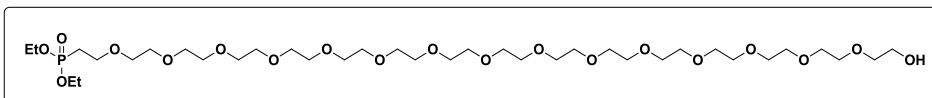
DCM. $^1\text{H NMR}$ (400MHz, CDCl_3): δ_{H} 7.34-7.24(m, 5H), 4.57(s, 2H), 3.75 (t, $J=6.8\text{Hz}$, 2H), 3.69-3.58 (m, 64H), 3.26 (t, $J=6.8\text{Hz}$, 2H). $^{13}\text{C NMR}$ (100MHz, CDCl_3): δ_{C} 138.13, 128.22, 127.59, 127.45, 73.06, 71.81, 70.47, 70.37, 70.06, 69.28, 2.95. **MALDI-ToF MS** (M+K): 961.29.

Synthesis of compound 1j



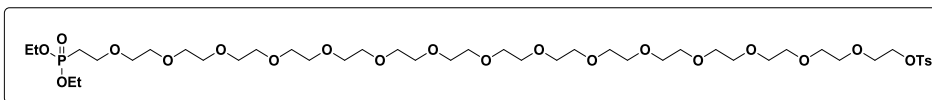
In an oven-dried RBF, iodide **1i** (0.54 g, 0.5 mmol) and $\text{P}(\text{OEt})_3$ (0.27 g, 1.4 mmol) were taken and refluxed for 1 h. Upon completion of the reaction, the excess $\text{P}(\text{OEt})_3$ was removed under vacuum, and the reaction mixture was directly loaded onto a silica gel column, and the crude mixture was purified using MeOH / DCM as eluent. The product was obtained as a pale yellow liquid (0.45 g, 0.4 mmol, 83%), $R_f = 0.45$ in 5% MeOH / DCM. $^1\text{H NMR}$ (400MHz, CDCl_3): δ_{H} 7.39-7.27(m, 5H), 4.56(s, 2H), 4.16-4.05 (m, 4H), 3.81-3.62(m, 64H), 2.17-1.98(m, 2H), 1.32 (t, $J=6.8\text{Hz}$, 6H). $^{13}\text{C NMR}$ (100MHz, CDCl_3): δ_{C} 138.03, 128.10, 127.46, 127.32, 72.94, 70.37, 70.29, 69.93, 69.19, 64.86, 61.39, 61.33, 27.41, 26.02, 16.24, 16.18. **MALDI-ToF MS** (M+K): 971.18.

Synthesis of compound 1k



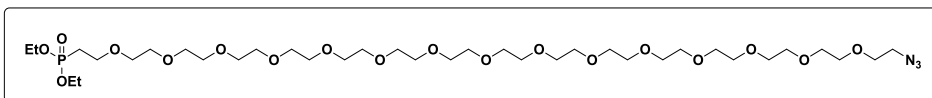
The disphosphonate ester **1j** (0.4 g, 0.4 mmol) was dissolved in alcohol, Pd/C was added and stirred under hydrogen for about 18 hours. Upon completion, the mixture was filtered through celite and washed with alcohol; the filtrate was concentrated under vacuum to get crude mixture **1k** and used for next step without purification. The product was obtained as a pale yellow liquid (0.35 g, 0.4 mmol, 89%), $R_f = 0.45$ in 5% MeOH / DCM. $^1\text{H NMR}$ (400MHz, CDCl_3): δ_{H} 4.07-3.93(m, 4H), 3.66-3.49 (m, 63H), 2.11-1.99(m, 2H), 1.22 (t, $J=6.8\text{Hz}$, 6H). $^{13}\text{C NMR}$ (100MHz, CDCl_3): δ_{C} 70.43, 70.33, 70.14, 70.07, 64.99, 61.56, 61.49, 27.53, 26.14, 16.36, 16.30. **MALDI-ToF MS** (M+K): 981.72.

Synthesis of compound 1l



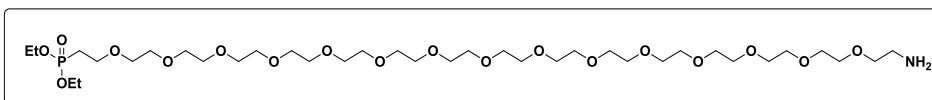
In an oven-dried RBF, compound **1k** (0.33 g, 0.4 mmol), TsCl (0.30 g, 1 mmol), DMAP (0.0079 g, 0.06 mmol) were taken and was dissolved in DCM under at 0 °C. Et₃N (0.16 g, 1.5 mmol) was added to the above mixture and allowed to stir for 12 h at RT. Upon completion, the reaction was quenched with water and extracted in DCM for thrice. The combined organic layer was dried over Na₂SO₄ and concentrated under reduced pressure to get the crude product, which was purified using silica gel column chromatography using MeOH / DCM as eluent to get pale yellow liquid (0.39 g, 4 mmol, 98%), $R_f=0.40$ in 5% MeOH / DCM. ¹H NMR (400MHz, CDCl₃): δ_H 7.79 (d, J=8Hz, 2H), 7.34 (d, J=8Hz, 2H), 4.17-4.07 (m, 6H), 3.77-3.62 (m, 60H), 2.45 (s, 3H), 2.17-2.09 (m, 2H), 1.32 (t, J=6.8Hz, 6H). ¹³C NMR (100MHz, CDCl₃): δ_C 144.89, 133.06, 129.92, 128.07, 77.16, 70.81, 70.62, 70.26, 69.34, 68.76, 65.20, 61.77, 61.71, 27.73, 26.34, 21.74, 16.55, 16.49. **MALDI-ToF MS** (M+K): 1035.13.

Synthesis of compound 1m



In an oven-dried RBF, compound a mixture of **1l** (0.21 g, 0.21 mmol), and NaN₃ (0.03 g, 0.4 mmol) was taken and dissolved in DMF under stirring. The resultant mixture was allowed to react for 12 h. Upon completion of the reaction, DMF was evaporated under reduced pressure. Then, the obtained residue was purified using silica gel column chromatography using MeOH / DCM as eluent to get pale yellow liquid (0.15 g, 0.17 mmol, 85%), $R_f = 0.45$ in 5% MeOH / DCM. The obtained product was used for next step without purification.

Synthesis of compound 1n



In an oven-dried RBF, compound **1m** (2.7 g, 5.2 mmol) was taken and dissolved in THF under stirring. Then, PPh₃ (0.072 g, 0.2 mmol) was added to the above mixture at 0

°C, and the mixture was allowed to react for 18 h at RT. Upon completion of the reaction, a few drops of water were added and again stirred for 1 hour. Then, THF was evaporated under reduced pressure, and the obtained water layer was washed with toluene for at least four times. The water layer was then concentrated under reduced pressure to get the product (0.14 g, 0.15 mmol 89%), which was used without further purification. $R_f=0.59$ in of 2% triethylamine in 5% MeOH / DCM. $^1\text{H NMR}$ (400MHz, D_2O): δ_{H} 4.18- 4.11 (m, 4H), 3.83-3.63 (m, 61H), 2.30-2.19 (m, 2H), 1.32 (t, $J=7.2\text{Hz}$, 6H). $^{13}\text{C NMR}$ (100MHz, CDCl_3): δ_{C} 69.95, 69.61, 69.47, 69.41, 64.25, 63.41, 63.34, 39.59, 25.80, 24.42, 15.64, 15.58. **MALDI-TOF MS (M+K):** 880.45.

3.4.4.3. Synthesis of Photo-Sensitive Macromolecular AABP

3.4.4.3.1. General Procedure for the Synthesis of Aldehyde - Procedure C

In an oven-dried RBF, 5-hydroxy-2-nitrobenzaldehyde (1.1 eq), K_2CO_3 (1.1 eq), and bromide (1 eq) were dissolved in DMF under stirring at RT and allowed to react for 12 h. Upon completion of the reaction, water was added and extracted thrice with DCM. The combined organic layer was dried over Na_2SO_4 and evaporated under vacuum to get crude product which was purified using silica gel column chromatography using ethyl acetate/hexane as eluent.

3.4.4.3.2. General Procedure for the Synthesis of Alcohol - Procedure D

In an oven-dried RBF, the above-obtained aldehyde (1 eq) was dissolved in THF: MeOH (2:1) under stirring. Then, the mixture was cooled to 0° C and NaBH_4 (1.5 eq) was added in small portions and allowed to react for 1 hour. Upon completion, the reaction was quenched with Na_2CO_3 and extracted thrice with ethyl acetate. The combined organic layer was dried over Na_2SO_4 and concentrated under vacuum to get the crude product, which was purified using silica gel column chromatography using ethyl acetate/hexane as eluent.

3.4.4.3.3. General Procedure for the Synthesis of Activated Ester - Procedure E

In an oven-dried RBF, above obtained alcohol (1 eq) and N, N'-DSC (5 eq) were dissolved in ACN under stirring. Then, Et_3N (5 eq) was then added slowly at RT and

allowed to react for 12 h. Upon completion of the reaction, ACN and Et₃N were evaporated under vacuum. The obtained residue was directly purified using silica gel column chromatography using ethyl acetate/hexane as eluent.

3.4.4.3.4. General Procedure for the Synthesis of Diphosphonate Ester - Procedure F

In an oven-dried RBF, the above obtained activates ester (1.1 eq) and compound 1i (1 eq) were dissolved in DMF under stirring. Then, Et₃N (1.1 eq) was added slowly to the reaction mixture and stirred at RT for 12 h. Upon completion of the reaction, DMF and Et₃N were evaporated under vacuum. To the obtained residue, water was added and extracted thrice with DCM. The combined organic layer was dried over Na₂SO₄ and concentrated under vacuum to get the crude product, which was purified using silica gel column chromatography using MeOH / DCM as eluent.

3.4.4.3.5. General Procedure for the Synthesis of Monophosphonate Ester - Procedure G

In an oven-dried RBF, the above-obtained diphosphonate ester (1 eq) was taken, and LiBr (20 eq) was added. To the mixture, DMF was added and heated at 95°C for 20 h. Upon completion, water was added and extracted thrice with ethyl acetate. The water layer was collected, and 2N HCl was added and stirred for another 30 minutes. The mixture was then extracted thrice with ethyl acetate. The combined organic layer was dried over Na₂SO₄ and concentrated under vacuum to get the crude product, which was used without purification.

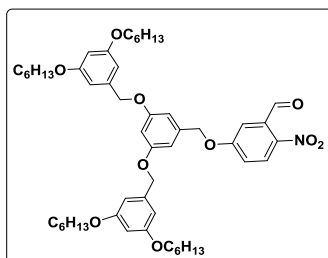
3.4.4.3.6. General Procedure for the Synthesis of Fluorophosphonate Ester - Procedure H

To the stirring solution of the above obtained monophosphonate ester (1 eq) in DCM, DAST (4 eq) was added dropwise at -78 °C and allowed to react for 15 minutes. Upon completion of the reaction, excess of DAST and DCM were evaporated under vacuum. To the obtained residue, water was added and stirred for 2 more minutes to quench any residual DAST. The reaction mixture was then extracted thrice with DCM. The

combined organic layer was dried over Na_2SO_4 and concentrated under vacuum to get the crude product, which was used without purification.

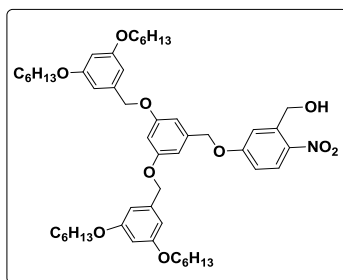
3.4.4.3.7. Synthesis of Photo-Sensitive Macromolecular AABP and Their Intermediates

Synthesis of compound 5a



The compound **5a** was prepared by general procedure C, starting from 5-hydroxy-2-nitrobenzaldehyde **3** (0.7 g, 4 mmol), K_2CO_3 (0.7 g, 5 mmol) and G1 bromide **4a** (3.3 g, 4.2 mmol). The product obtained as pale solid (3.3 g, 3.8 mmol, 90%), $R_f=0.52$ in 20% ethyl acetate/hexane. $^1\text{H NMR}$ (400MHz, CDCl_3): δ_{H} 10.47 (s, 1H), 8.11 (d, $J=8.9$ Hz, 1H), 7.37 (d, $J=8.9, 2.8$ Hz, 1H), 7.14 (d, $J=2.8$ Hz, 1H), 6.61-6.40 (m, 9H), 5.12 (s, 2H), 4.96 (s, 4H), 3.93 (t, $J=6.4\text{Hz}$, 4H), 1.77-1.53 (m, 8H), 1.53-1.26 (m, 24H), 0.91 (t, $J=7.2\text{Hz}$, 12H). $^{13}\text{C NMR}$ (100MHz, CDCl_3): 188.56, 163.05, 160.65, 160.29, 142.48, 138.83, 137.35, 134.40, 127.23, 119.27, 114.52, 106.37, 105.78, 100.85, 70.95, 70.30, 70.01, 68.20, 31.71, 29.34, 25.85, 22.73, 14.18. **MALDI-TOF MS** (M+K): 908.31.

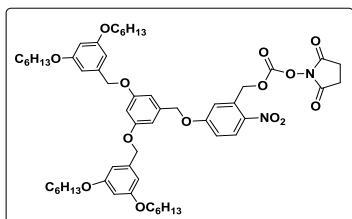
Synthesis of compound 5b



The compound **5b** was prepared by general procedure D, starting from compound **5a** (0.9 g, 1 mmol), NaBH_4 (0.09 g, 1.5 mmol). The product obtained as a pale yellow solid (0.8 g, 0.9 mmol, 90%), $R_f=0.20$ in 20% ethyl acetate / hexane. δ_{H} $^1\text{H NMR}$ (400MHz,

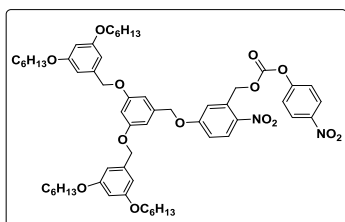
CDCl₃): δ_{H} 8.13 (d, $J=8.9$ Hz, 1H), 7.34 (d, $J=8.9, 2.8$ Hz, 1H), 7.18 (d, $J=2.8$ Hz, 1H), 6.62-6.41 (m, 9H), 5.14 (s, 2H), 4.94 (s, 4H), 3.93 (t, $J=6.4$ Hz, 4H), 1.78-1.52 (m, 8H), 1.53-1.26 (m, 24H), 0.90 (t, $J=7.2$ Hz, 12H). **MALDI-TOF MS** (M+K): 910.43.

Synthesis of compound 5c



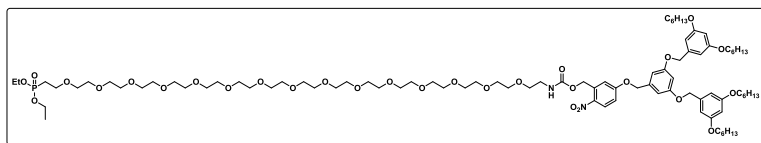
The compound **5c** was prepared by general procedure E, starting from compound **5b** (0.75 g, 0.8 mmol) and *N, N'*-DSC (3.3 g, 12 mmol) and, Et₃N (1.30 g, 12 mmol). The product obtained as a pale yellow solid (0.40 g, 0.4 mmol, 50%), $R_f=0.18$ in 25% ethyl acetate / hexane. **¹H NMR** (400MHz, CDCl₃): δ_{H} 8.21 (d, $J=8.9$ Hz, 1H), 7.19 (dd, $J=8.9, 2.8$ Hz, 1H), 6.96 (d, $J=2.8$ Hz, 1H), 6.69-6.40 (m, 12H), 5.80 (s, 2H), 5.17 (s, 2H), 3.93 (t, $J=6.4$ Hz, 4H), 2.78 (s, 4H), 1.81-1.72 (m, 10H), 1.47-1.28 (m, 35H), 0.9 (t, $J=7.2$ Hz, 18H). **¹³C NMR** (100MHz, CDCl₃): 160.62, 138.91, 113.99, 106.33, 105.82, 100.86, 70.28, 68.23, 51.98, 31.70, 29.33, 25.84, 25.52, 25.22, 22.78, 20.20, 14.22, 13.73. **MALDI-TOF MS** (M+K): 1053.73.

Synthesis of compound 5d



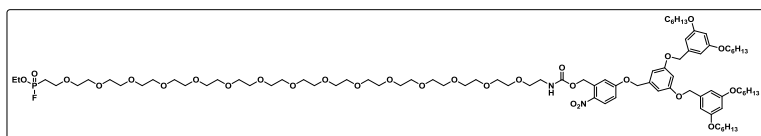
To the mixture of compound **5b** (0.6 g, 0.13 mmol) and pyridine (60 mg, 0.70 mmol) in DCM, a solution of 4-nitrophenyl chloroformate (150 mg, 0.70 mmol) in DCM was added at 0 °C. The mixture was allowed to react for 3 hours at RT. Then, the solvent was evaporated under reduced pressure to get the crude product, which was then purified using silica gel column chromatography (0.4 g, 0.1 mmol, 58%), $R_f=0.20$ in 20% ethyl acetate/hexane. **MALDI-TOF MS** (M+K): 1077.3.

Synthesis of compound 5e



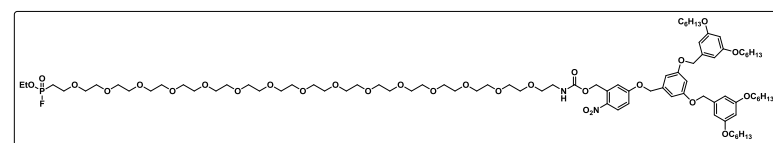
The compound **5d** was prepared by general procedure F, starting from compound **5c** (0.6 g, 0.6 mmol), compound **1n** (0.4 g, 0.04 mmol) and, Et₃N (0.06 g, 0.6 mmol). The obtained product was pale yellow liquid (0.35 g, 0.2 mmol, 35%), *R_f*=0.40 in 5% MeOH / DCM. ¹H NMR (400MHz, CDCl₃) 8.08 (d, *J*=8.9 Hz, 1H), 7.10 (dd, *J*=8.9, 2.8 Hz, 1H), 6.82 (d, *J*=2.8 Hz, 1H), 6.57-6.33 (m, 12H), 5.46 (s, 2H), 5.23 (s, 2H), 4.88 (s, 2H), 3.85 (t, *J*=6.4Hz, 8H), 3.76-3.46 (m, 67H), 2.25-2.11 (m, 2H), 1.72-1.61 (m, 12H), 1.46-1.14 (m, 62H), 0.82 (t, *J*=6.8Hz, 18H). ¹³C NMR (100MHz, CDCl₃): 162.49, 160.42, 138.67, 127.78, 113.14, 106.31, 105.61, 101.68, 100.59, 70.43, 70.21, 70.04, 69.78, 67.95, 36.08, 31.76, 31.48, 31.31, 29.56, 29.11, 25.63, 22.59, 14.04. **MALDI-TOF MS** (M+K): 1777.10.

Synthesis of compound 5f



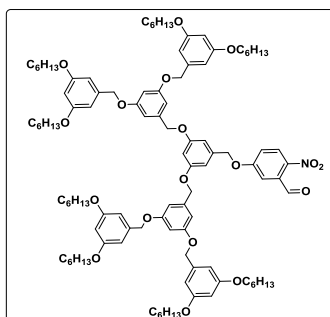
The compound **5e** was prepared by general procedure G, starting from compound **5d** (0.25 g, 0.14 mmol) and, LiBr (0.25 g, 2 mmol) was added. The obtained product was carried to the next step without further purification.

Synthesis of compound 5g



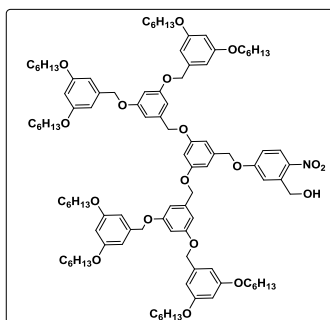
The compound **5f** was prepared by general procedure H, starting from compound **5e** (0.1 g, 0.050 mmol), DAST (0.040 g, 0.2 mmol). The product was utilized for conjugation without further purification. ¹⁹F NMR (400MHz, CDCl₃): δ_F -59.91, -62.74. **MALDI-TOF MS** (M+K): 1753.10.

Synthesis of compound 6a



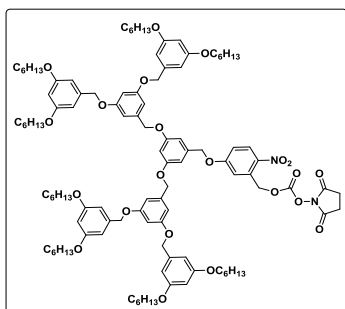
The compound **6a** was prepared by general procedure C, starting from 5-hydroxy-2-nitrobenzaldehyde (**3**) (0.018 g, 0.10 mmol), K_2CO_3 (0.02 g, 0.12 mmol) and G2 bromide (**4b**) (0.17 g, 0.10 mmol). The product obtained as pale solid (0.10 g, 0.06 mmol, 65%), $R_f=0.52$ in 20% ethyl acetate / hexane. 1H NMR (400MHz, $CDCl_3$): δ_H 10.46 (s, 1H), 8.11 (d, $J=8.9$ Hz, 1H), 7.37 (d, $J=8.9, 2.8$ Hz, 1H), 7.12 (d, $J=2.8$ Hz, 1H), 6.65-6.39 (m, 21H), 5.13 (s, 2H), 4.97-4.94 (m, 12H), 3.93 (t, $J=6.4$ Hz, 16H), 1.83-1.70 (m, 16H), 1.46-1.32 (m, 52H), 0.90 (t, $J=7.2$ Hz, 24H). ^{13}C NMR (100MHz, $CDCl_3$): 188.54, 163.02, 160.62, 160.27, 138.95, 134.39, 127.34, 114.43, 70.91, 70.18, 68.17, 31.70, 29.33, 25.85, 22.73, 14.18. MALDI-TOF MS (M+K): 1734.21

Synthesis of compound **6b**



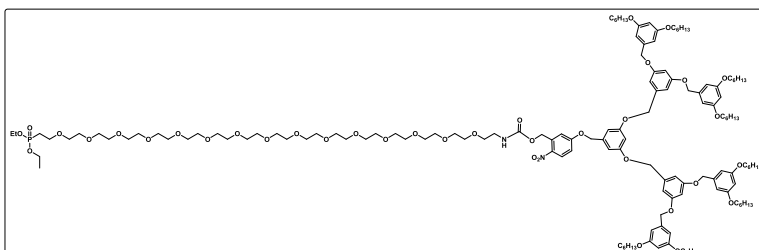
The compound **6b** was prepared by general procedure D, starting from compound **6a** (0.13 g, 0.09 mmol), $NaBH_4$ (0.045 g, 1.1 mmol). The product obtained as a pale yellow solid (0.090 g, 0.056 mmol, 70%), $R_f=0.20$ in 20% ethyl acetate / hexane. 1H NMR (400MHz, $CDCl_3$): δ_H 8.11 (d, $J = 8.9$ Hz, 1H), 7.14 (dd, $J = 8.9, 2.8$ Hz, 1H), 6.86 (d, $J = 2.8$ Hz, 1H), 6.62-6.38 (m, 21H), 5.61 (s, 2H), 5.03 (s, 2H), 4.90 (s, 12H), 3.90 (t, $J=6.4$ Hz, 15H), 1.77-1.70 (m, 16H), 1.45-1.25 (m, 50H), 0.88 (t, $J=7.2$ Hz, 24H). ^{13}C NMR (100MHz, $CDCl_3$): 160.64, 160.28, 138.96, 105.89, 100.92, 99.96, 77.16, 70.31, 68.21, 31.72, 29.35, 25.86, 22.75, 14.19. MALDI-TOF MS (M+K): 1735.30

Synthesis of compound 6c



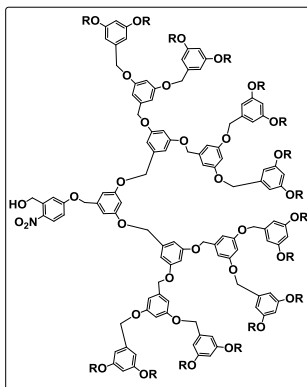
The compound **6c** was prepared by general procedure E, starting from compound **6b** (0.048 g, 0.03 mmol), *N,N'*-DSC (0.2 g, 0.80 mmol) and Et₃N (0.10 g, 0.1 mmol). The product obtained as a pale yellow solid (0.03 g, 0.019 mmol, 66%), *R*_f=0.18 in 25% ethyl acetate / hexane. ¹H NMR (400MHz, CDCl₃): δ_H 8.20 (d, *J*=8.9 Hz, 1H), 7.18 (dd, *J*=8.9, 2.8 Hz, 1H), 6.98 (d, *J*=2.8 Hz, 1H), 6.70-6.40 (m, 21H), 5.80 (s, 2H), 4.98-4.94 (m, 12H), 3.93 (t, *J*=6.4Hz, 16H), 2.71 (s, 4H), 1.84-1.68 (m, 18H), 1.47-1.22 (m, 56H), 0.89 (t, *J*=7.2Hz, 27H). ¹³C NMR (100MHz, CDCl₃): 168.57, 163.39, 160.59, 160.30, 160.21, 139.44, 139.18, 138.94, 106.44, 106.20, 105.83, 101.55, 100.84, 70.60, 70.22, 70.07, 69.19, 68.14, 31.67, 29.30, 25.82, 22.70, 14.14.

Synthesis of compound 6d



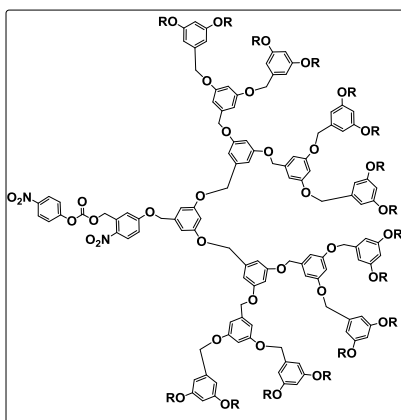
The compound **6d** was prepared by general procedure F, starting from compound **6c** (0.10 g, 0.05 mmol), compound **1n** (0.04 g, 0.05 mmol) and, Et₃N (0.005 g, 0.05 mmol). The product obtained was pale yellow liquid (0.1 g, 0.04 mmol, 60%), *R*_f=0.40 in 5% MeOH / DCM. 8.15 (d, *J* = 8.9 Hz, 1H), 7.17 (dd, *J* = 8.9, 2.8 Hz, 1H), 6.91 (d, *J* = 2.8 Hz, 1H), 6.66-6.39 (m, 27H), 5.53 (s, 2H), 5.08 (s, 2H), 5.00-4.86 (m, 16H), 4.14-4.06 (m, 4H), 3.90 (t, *J*=6.4Hz, 20H), 3.75-3.49 (m, 77H), 2.19-2.06 (m, 2H), 1.86-1.66 (m, 22H), 1.50-1.18 (m, 88H), 0.89 (t, *J*=7.2Hz, 36H). ¹³C NMR (100MHz, CDCl₃): 160.59, 160.19, 138.92, 106.63, 105.84, 100.85, 70.61, 70.25, 70.16, 68.15, 65.19, 61.81, 31.67, 29.79, 29.30, 25.82, 22.70, 16.55, 16.41, 14.15. MALDI-TOF MS (M+K): 2602.39

Synthesis of compound 7b



The compound **7b** was prepared by general procedure D, starting from compound **7a** (0.8 g, 0.2 mmol), NaBH₄ (18 mg, 4 mmol). The product obtained as a pale yellow solid (0.67 g, 0.2 mmol, 86%), *R_f*=0.20 in 20% ethyl acetate / hexane. ¹H NMR (400MHz, CDCl₃): δ_H 8.09 (m, 1H), 7.28 (d, *J*=8Hz, 1H), 6.84 (dd, *J* = 8.9, 2.8 Hz, 1H), 6.68-6.40 (m, 45H), 5.05-4.94 (m, 28H), 3.92 (t, *J*=6.4Hz, 34H), 1.79-1.72 (m, 36H), 1.48-1.27 (m, 106H), 0.90 (t, *J*=6.4Hz, 52H). ¹³C NMR (100MHz, CDCl₃): 160.61, 160.26, 139.17, 138.99, 106.51, 105.86, 101.68, 100.91, 77.16, 70.26, 68.17, 31.83, 29.34, 25.85, 22.73, 14.18.

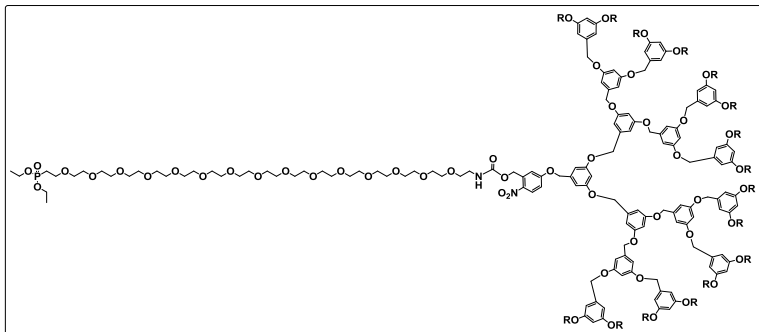
Synthesis of compound 7c



To the mixture of compound **7b** (0.5 g, 0.13 mmol) and pyridine (60 mg, 0.70 mmol) in DCM, a solution of 4-nitrophenylchloroformate (150 mg, 0.70 mmol) in DCM was added at 0 °C. The mixture was allowed to react for 3 hours at RT. Then, the solvent was evaporated under reduced pressure to get crude product, which was then purified using silica gel column chromatography (0.35 g, 0.1 mmol, 68%), *R_f*=0.20 in 20% ethyl acetate /

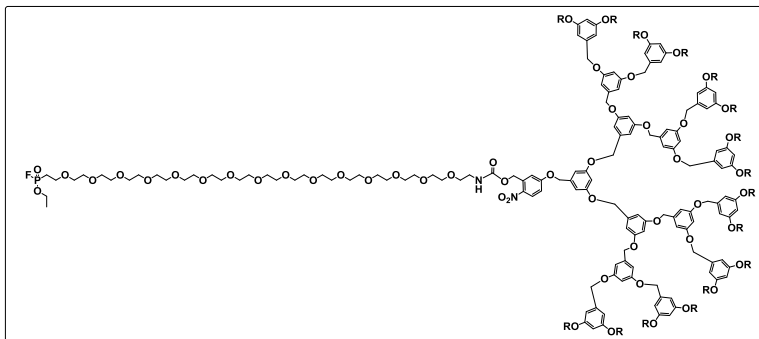
hexane. $^1\text{H NMR}$ (400MHz, CDCl_3): δ_{H} 8.15 (m, 2H), 7.31 (d, $J=8\text{Hz}$, 2H), 7.20 (m, 1H), 6.0 (m, 1H), 6.66-6.39 (m, 46H), 5.08 (s, 2H), 4.96-4.93 (m, 26H), 3.91 (t, $J=6.4\text{Hz}$, 34H), 1.78-1.73 (m, 36H), 1.44-1.26 (m, 144H), 0.89 (t, $J=6.4\text{Hz}$, 64H). $^{13}\text{C NMR}$ (100MHz, CDCl_3): 160.64, 160.30, 139.16, 139.01, 125.42, 121.77, 106.57, 105.89, 101.70, 100.93, 70.29, 70.19, 68.19, 31.72, 29.85, 29.36, 25.86, 22.74, 14.18.

Synthesis of compound 7d



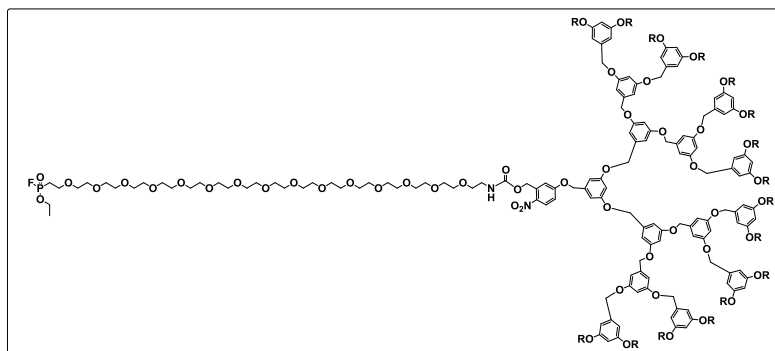
The compound **7d** was prepared by general procedure F, starting from compound **7c** (0.3 g, 0.1 mmol), compound **1n** (0.08 g, 0.1 mmol) and, Et_3N (26 mg, 0.3 mmol). The product obtained was pale yellow liquid (0.2 g, 0.06 mmol, 60%), $^1\text{H NMR}$ (400MHz, CDCl_3): δ_{H} 8.10 (m, 1H), 7.14 (m, 1H), 6.85 (m, 1H), 6.66-6.39 (m, 45H), 5.04 (s, 2H), 4.96-4.93 (m, 26H), 4.15-4.05 (m, 4H), 3.91 (t, $J=6.4\text{Hz}$, 32H), 3.75-3.52 (m, 60H), 2.19-2.09 (m, 2H), 1.78-1.71 (m, 34H), 1.51-1.26 (m, 144H), 0.89 (t, $J=6.4\text{Hz}$, 58H).

Synthesis of compound 7e



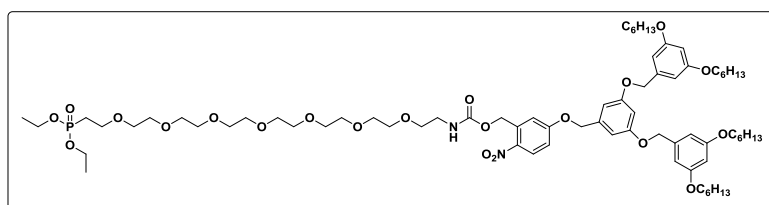
The compound **7e** was prepared by general procedure G, starting from compound **7d** (0.2 g, 0.04 mmol) and LiBr (0.040 g, 0.5 mmol). The obtained product was carried to the next step without further purification.

Synthesis of compound 7f



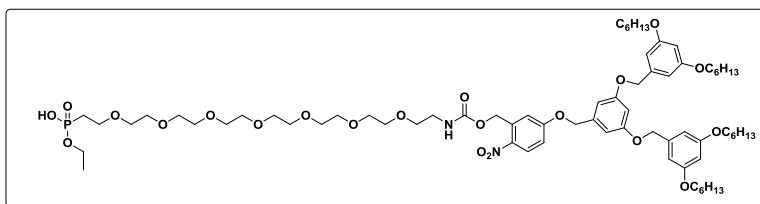
The compound **7f** was prepared by general procedure H, starting from compound **7e** (0.2 g, 0.04 mmol), DAST (0.01 g, 0.05 mmol). The obtained product was utilized for conjugation without further purification. ^{19}F NMR (400MHz, CDCl_3): δ_{F} -59.91, -62.74.

Synthesis of compound 8a



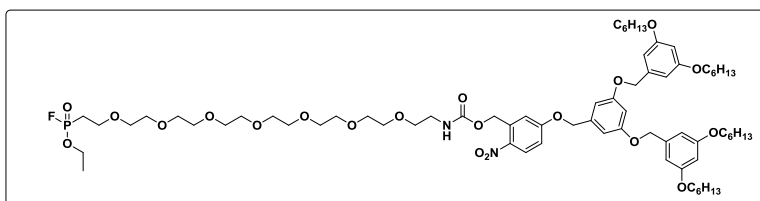
The compound **8d** was prepared by general procedure F, starting from compound **5c** (0.6 g, 0.6 mmol), compound **11** (synthetic scheme and characterization data for compound 11 is reported by our group in ref. no 15) (0.4 g, 0.04 mmol) and, Et_3N (0.06 g, 0.6 mmol). The obtained product was pale yellow liquid (0.35 g, 0.2 mmol, 45%), $R_f=0.40$ in 5% MeOH / DCM. ^1H NMR (400MHz, CDCl_3) 8.14 (d, $J=8.9$ Hz, 1H), 7.23 (d, $J=2.8$ Hz, 1H), 6.85 (dd, $J=8.9$ Hz, 1H), 6.64 (d, $J=2.8$ Hz, 1H), 6.61 (t, $J=2.8$ Hz, 1H), 6.58 (d, $J=2.8$ Hz, 1H), 6.44 (t, $J=2.8$ Hz, 1H), 5.60 (s, 2H), 5.10 (s, 2H), 4.90 (s, 2H), 4.35-4.33 (m, 2H), 4.15-4.04 (m, 4H), 3.95 (t, $J=6.4$ Hz, 8H), 3.75-3.60 (m, 28H), 2.18-2.10 (m, 2H), 1.80-1.72 (m, 10H), 1.48-1.25 (m, 32H), 0.90 (t, $J=6.8$ Hz, 12H). ^{13}C NMR (100MHz, CDCl_3): MALDI-TOF MS (M+K):

Synthesis of compound 8b



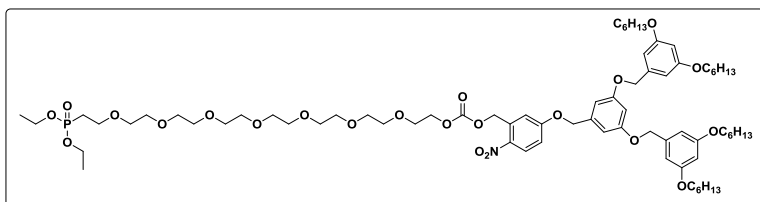
The compound **8e** was prepared by general procedure G, starting from compound **8d** (0.25 g, 0.14 mmol) and, LiBr (0.25 g, 2 mmol) was added. The obtained product was carried to the next step without further purification.

Synthesis of compound **8c**



The compound **8f** was prepared by general procedure H, starting from compound **8e** (0.1 g, 0.050 mmol), DAST (0.040 g, 0.2 mmol). The product was utilized for conjugation without further purification. ^{19}F NMR (400MHz, CDCl_3): δ_{F} -59.91, -62.74.

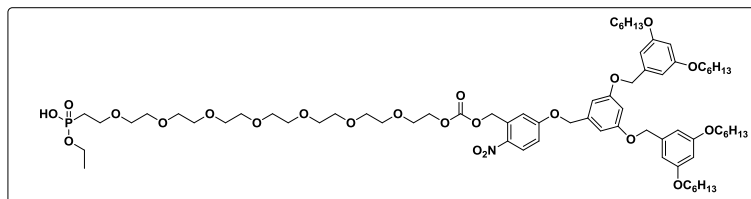
Synthesis of compound **9a**



The compound **9a** was prepared by general procedure F, starting from compound **5d** (0.5 g, 0.5 mmol), compound **10** (synthetic scheme and characterization data for compound 10 is reported by our group in ref. no 15) (0.5 g, 0.04 mmol) and, Et_3N (0.07 g, 0.6 mmol). The obtained product was pale yellow liquid (0.3 g, 0.2 mmol, 40%), $R_f=0.40$ in 5% MeOH / DCM. ^1H NMR (400MHz, CDCl_3) 8.17 (d, $J=8.9$ Hz, 1H), 7.20 (d, $J=2.8$ Hz, 1H), 6.90 (dd, $J=8.9$ Hz, 1H), 6.63 (d, $J=2.8$ Hz, 1H), 6.60 (t, $J=2.2$ Hz, 1H), 6.55 (d, $J=2.2$ Hz, 1H), 6.40 (t, $J=2.2$ Hz, 1H), 5.61 (s, 2H), 5.10 (s, 2H), 4.95 (s, 2H), 4.35-4.33 (m, 2H), 4.13-4.04 (m, 4H), 3.93 (t, $J=6.4$ Hz, 8H), 3.75-3.60 (m, 28H), 2.16-2.08 (m, 2H), 1.80-1.72 (m, 10H), 1.48-1.25 (m, 32H), 0.90 (t, $J=6.8$ Hz, 12H). ^{13}C NMR (100MHz, CDCl_3) 163.20, 160.67, 160.42, 154.83, 140.13, 138.87, 137.82, 135.34, 128.23, 114.43, 113.83, 106.54, 105.88,

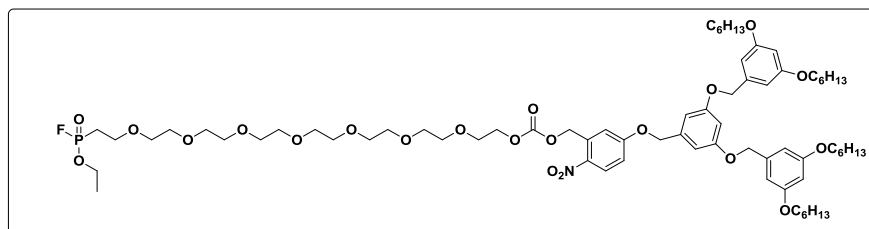
102.07, 100.90, 70.75, 70.68, 70.58, 70.33, 68.99, 68.22, 67.56, 66.51, 65.26, 61.80, 61.74, 31.71, 29.35, 25.86, 22.73, 16.59, 16.53, 14.17.

Synthesis of compound 9b



The compound **8d** was prepared by general procedure J, starting from compound **8c** (0.6 g, 0.6 mmol), compound **5m** (0.4 g, 0.04 mmol) and, Et₃N (0.06 g, 0.6 mmol). The obtained product was pale yellow liquid (0.35 g, 0.2 mmol, 35%), $R_f=0.40$ in 5% MeOH / DCM.

Synthesis of compound 8a



The compound **8d** was prepared by general procedure J, starting from compound **8c** (0.6 g, 0.6 mmol), compound **5m** (0.4 g, 0.04 mmol) and, Et₃N (0.06 g, 0.6 mmol). The obtained product was pale yellow liquid (0.35 g, 0.2 mmol, 35%), $R_f=0.40$ in 5% MeOH / DCM.

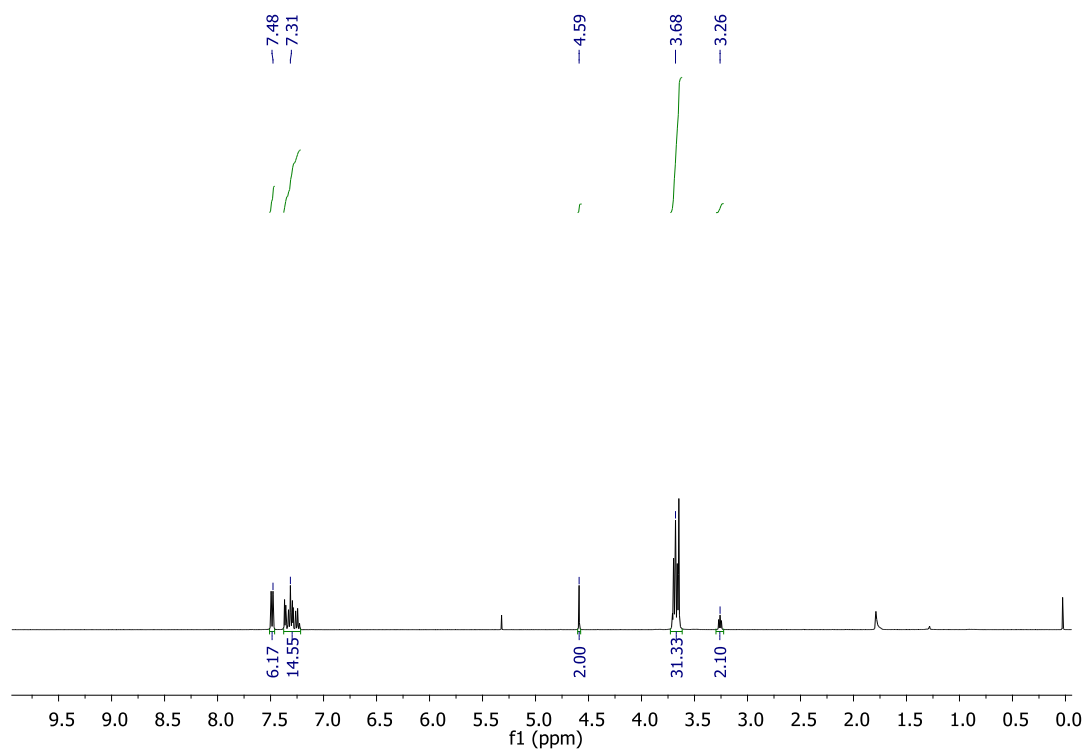
3.5. References

1. Smalle, J.; Vierstra, R. D. *Annu. rev. Plant Biol.* **2004**, *55*, 555–590.
2. Wittmann, T.; Hyman, A.; Desai, A. *Nature Cell Biol.* **2001**, *3*, E28–E34.
3. Kerfeld, C. A.; Heinhorst, S.; Cannon, G. C. *Annu. Rev. Microbiol.* **2010**, *64*, 391–408.
4. McMahon, H. T. *Curr. Biol.* **1999**, *9*, R332–R335
5. Liu, X.; Theil, E. C. *Acc. Chem. Res.* **2005**, *38*, 167–175
6. (a) Saibil, H. *Nat. Rev. Mol. Cell Biol.* **2013**, *14*, 630–642. (b) Kedersha, N. L.; Rome, L. H. *J. Cell Biol.* **1986**, *103*, 699–709.

7. Witus, L. S.; Francis, M. B. *Acc. Chem. Res.* **2011**, *44*, 774–783. (b) Lee, L. A.; Niu, Z. W.; Wang, Q. *Nano Res.* **2009**, *2*, 349 (c) Uchida, M.; Klem, M.T.; Allen, M.; Suci, P.; Flenniken, M.; Gillitzer, E.; Varpness, Z.; Liepold, L.O.; Young, M.; Douglas, T. *Adv. Mater.*, **2007**, *19*, 1025-1042.
8. Matsumoto, N.M.; Prabhakaran, P.; Rome, L.H.; Maynard, H.D. *Acs Nano*, **2013**, *7*, pp.867-874. (b) Kwak, M.; Minten, I. J.; Anaya, D. M.; Musser, A. J.; Brasch, M.; Nolte, R. J. M.; Mullen, K.; Cornelissen J.; Herrmann, A. *J. Am. Chem. Soc.*, **2010**, *132*, 7834–7835 (c) Stephanopoulos, N.; Tong, G. J.; Hsiao, S. C.; Francis, M. B. *ACS Nano* **2010**, *4*, 6014–6020. (d) Schlick, T. L.; Ding, Z.; Kovacs, E. W.; Francis, M. B. *J. Am. Chem. Soc.*, **2005**, *127*, 3718-3723.
9. Douglas, T.; Young, M. *Nature* **1998**, *393*, 6681–6685. (b) Comellas-Aragonès, M.; Engelkamp, H.; Claessen, V. I.; Sommerdijk, N. A.; Rowan, A. E.; Christianen, P. C.; Maan, J. C.; Verduin, B. J.; Cornelissen, J. J.; Nolte, R. J. *Nature Nanotech.* **2007**, *2*, 635–639 (c) Minten, I. J.; Hendriks, L. J. A.; Nolte, R. J. M.; Cornelissen, J. J. *J. Am. Chem. Soc.* **2009**, *131*, 17771–17773.
10. (a) King, N. P.; Bale, J. B.; Sheffler, W.; McNamara, D. E.; Gonen, S.; Gonen, T.; Yeates, T.O.; Baker, D. *Nature* **2014**, *510*, 103-108. (b) Bale, J. B.; Gonen, S.; Liu, Y.; Sheffler, W.; Ellis, D.; Thomas, C., Cascio, D.; Yeates, T.O.; Gonen, T.; King, N. P.; Baker, D. *Science* **2016**, *353*, 389-394.
11. Brodin, J. D.; Ambroggio, X. I.; Tang, C.; Parent, K. N.; Baker, T. S.; Tezcan, F. A. *Nat. Chem.* **2012**, *4*, 375–382
12. Oohora, K.; Onuma, Y.; Tanaka, Y.; Onoda, A.; Hayashi, T. *Chem. Commun.* **2017**, *53*, 6879– 6882. (b) Biswas, S.; et al. *J. Am. Chem. Soc.* **2009**, *131*, 7556–7557. (c) Sendai, T.; Biswas, S.; Aida, T. *J. Am. Chem. Soc.* **2013**, *135*, 11509–11512.
13. Chen, G.; Hoffman, A. S. *Bioconjugate Chem.* **1993**, *4*, 509-514. 22. Kulkarni, S.; Schilli, C.; Müller, A. H.; Hoffman, A. S.; Stayton, P. S. *Bioconjugate Chem.* **2004**, *15*, 747-753. 23. (b) Wan, X. J.; Liu, S. Y. *Macromol. Rapid Commun.* **2010**, *31*, 2070-2076. 24. (c) Park, W. M.; Champion, J. A. *J. Am. Chem. Soc.* **2014**, *136*, 17906-17909. (d) Hassouneh, W.; Fischer, K.; MacEwan, S.R.; Branscheid, R.; Fu, C.L.; Liu, R.; Schmidt, M.; Chilkoti, A. *Biomacromolecules* **2012**, *13*, 15981605
14. Sandanaraj, B. S.; Bhandari, P. J.; Reddy, M. M.; Lohote, A. B.; Sahoo, B. *ChemBioChem*, **2020**, *21*, 408-416.
15. Sandanaraj, B. S.; Reddy, M. M.; Bhandari, P. J.; Kumar, S. and Aswal, V. K. *Chem. Eur. J.* **2018**, *24*, 16085-16096.

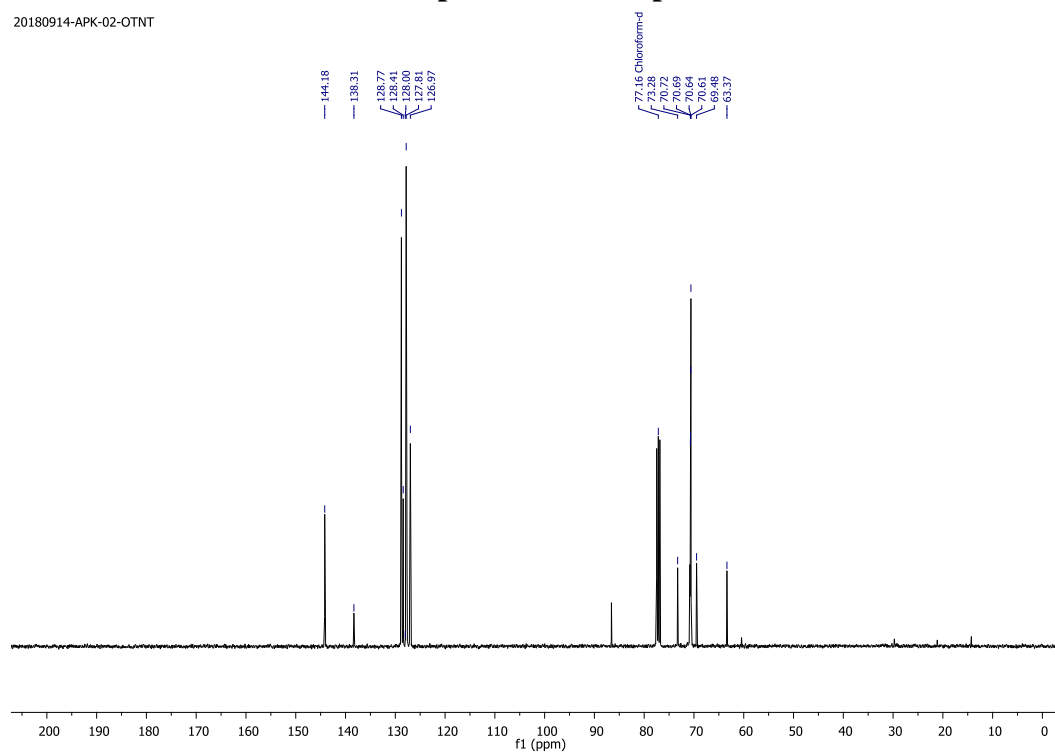
16. (a) Patchornik, A.; Amit, B. W. R. B.; Woodward, R. B. *J. Am. Chem. Soc.* **1970**, *92*, 6333-6335. (b) Cameron, J. F.; Fréchet, J. M. *J. Am. Chem. Soc.* **1991**, *113*, 4303-4313.
17. Layer, R.W. *Chem. reviews*, **1963**, *63*, 489-510.
18. (a) Hof F.; Rebek Jr, J.; *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 4775 (receptor). (b) Kang, J.; Rebek, J. *Nature*, **1997**, *385*, 50-52 (capsules). (c) Dwars, T.; Paetzold, E.; and Oehme, G. *Angew. Chem. Int. Ed.* **2005**, *44*, 7174-7199.
19. Godoy-Alcántar, C.; Yatsimirsky, A. K.; Lehn, J. M.; *J. Phy. Org. Chem.* **2005**, *18*, 979-985.
20. Kalia, J.; Raines, R.T. *Angew. Chem. Int. Ed.* **2008**, *47*, 7523-7526.

3.6. Appendix-II: Characterization data of synthesized compounds

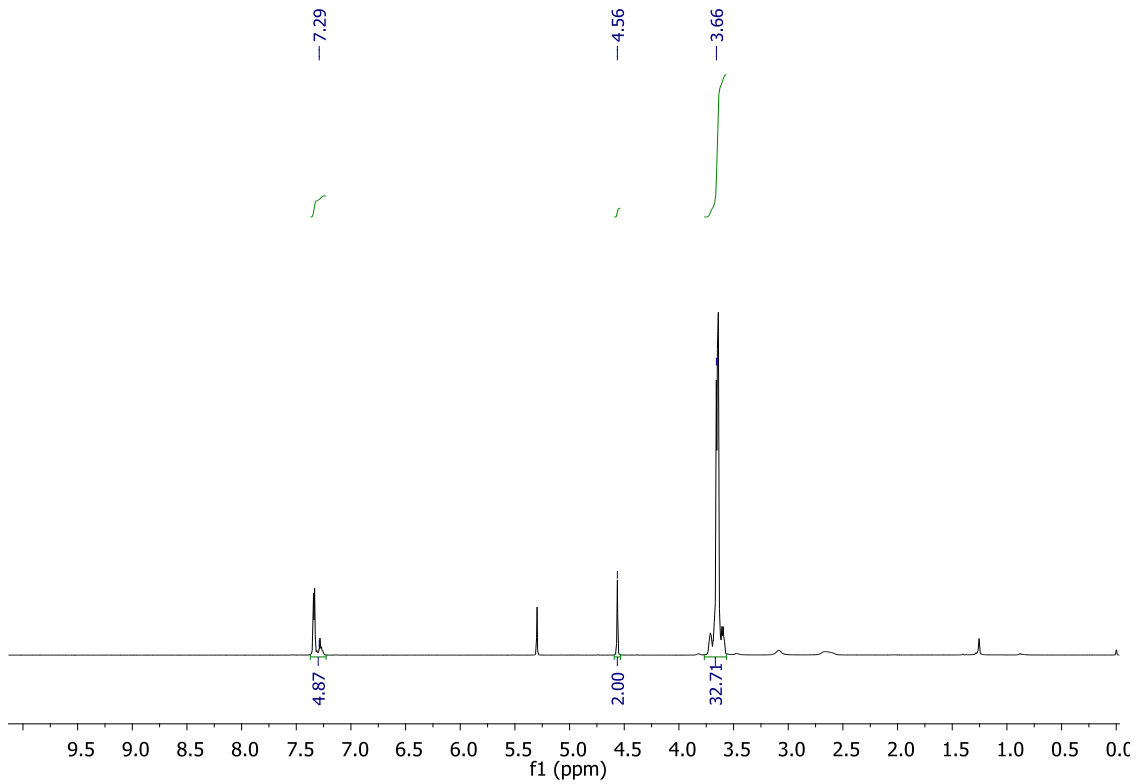


¹H NMR spectrum of compound 1c

20180914-APK-02-OTNT

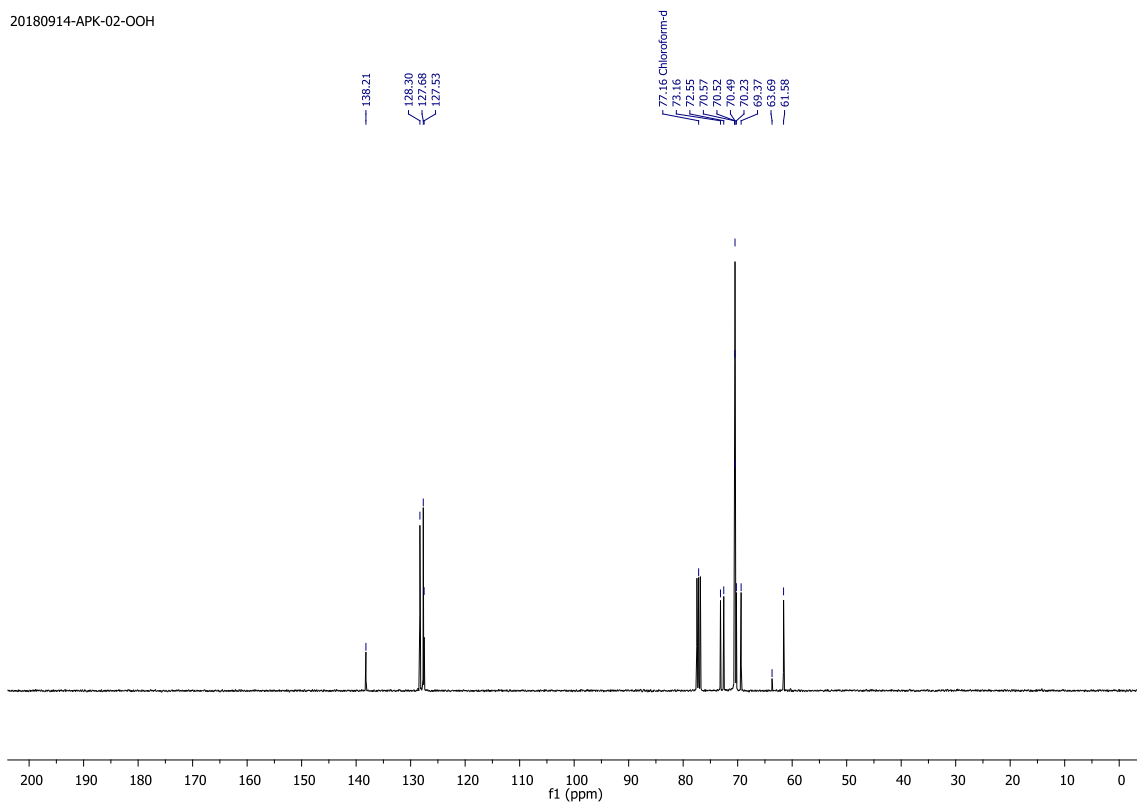


¹³C NMR spectrum of compound 1c

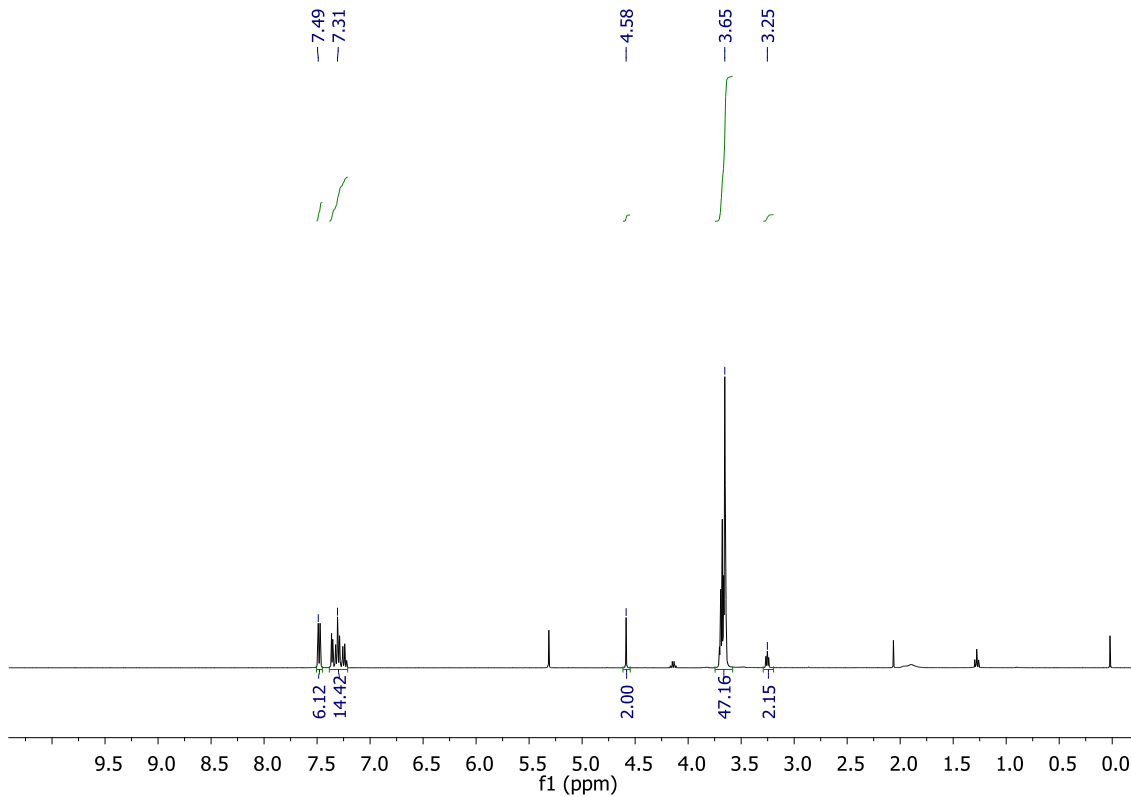


¹H NMR spectrum of compound 1d

20180914-APK-02-OOH

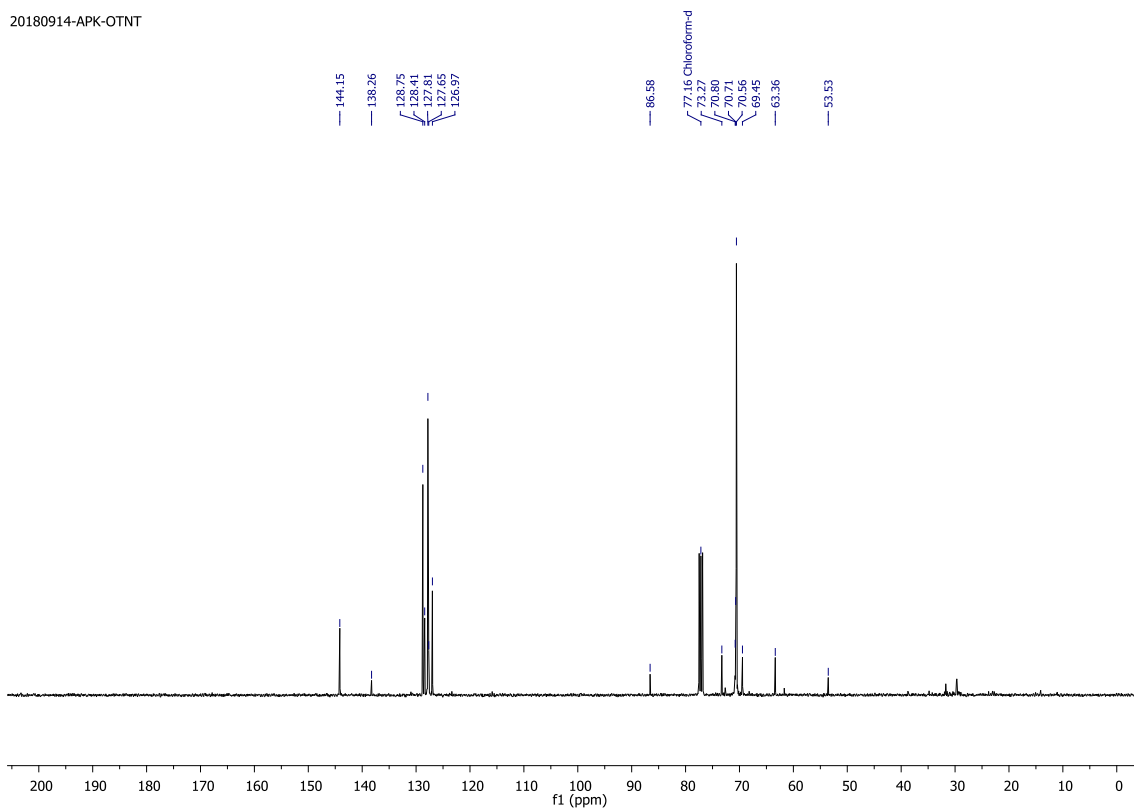


¹³C NMR spectrum of compound 1d

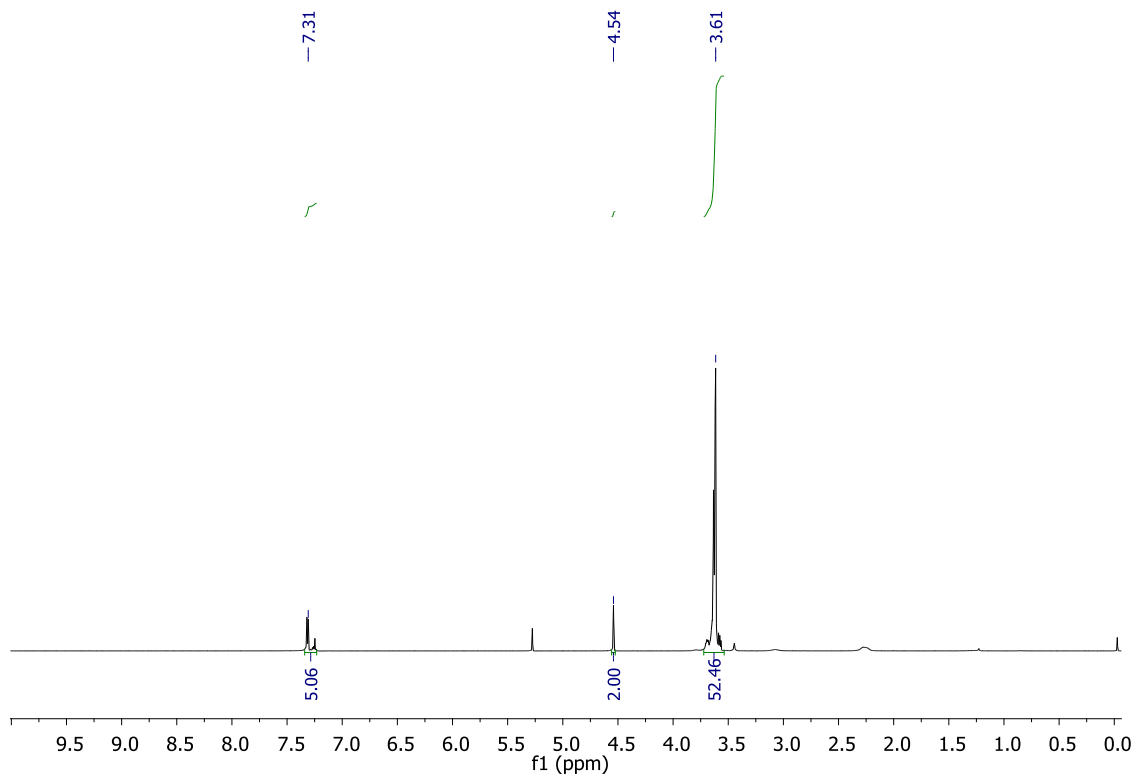


^1H NMR spectrum of compound 1e

20180914-APK-OTNT

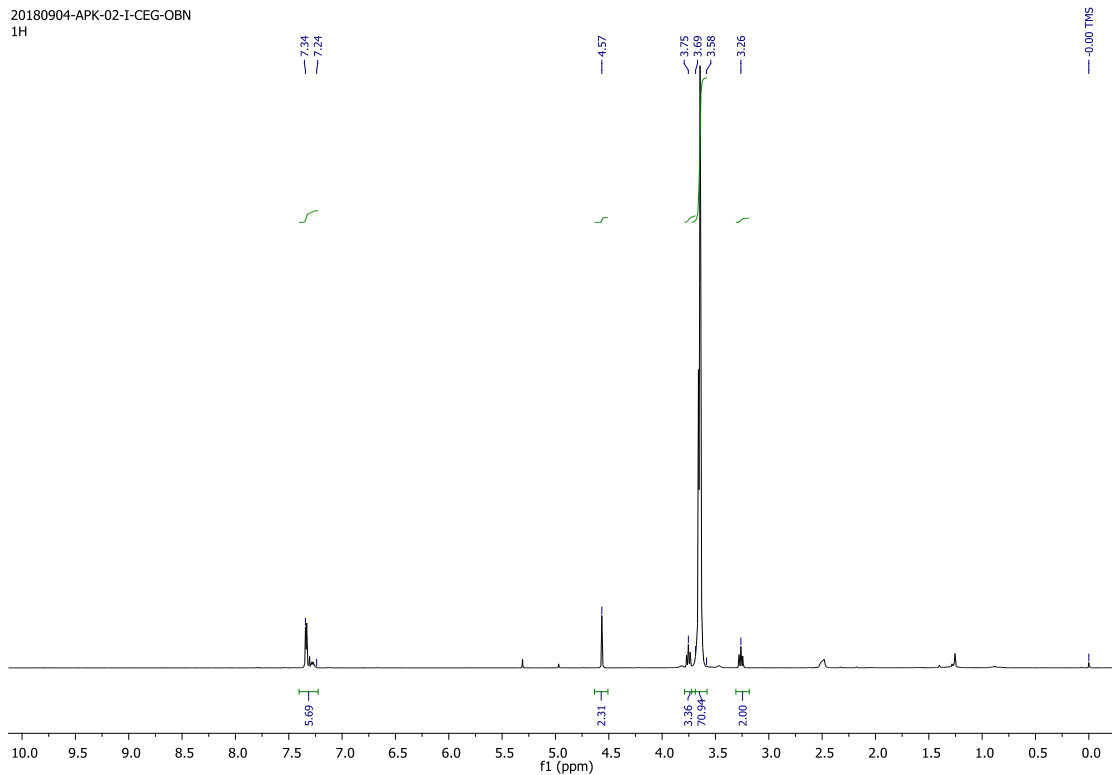


^{13}C NMR spectrum of compound 1e



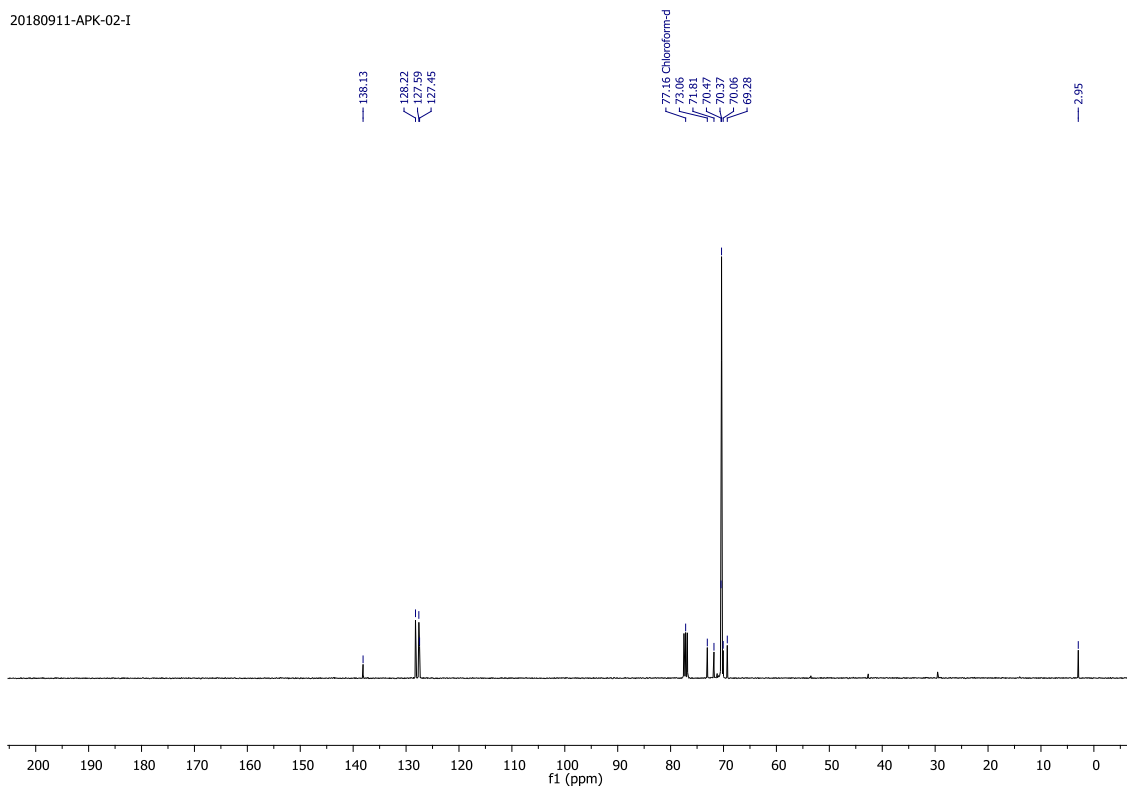
^1H NMR spectrum of compound 1f

20180904-APK-02-I-CEG-OBN
1H



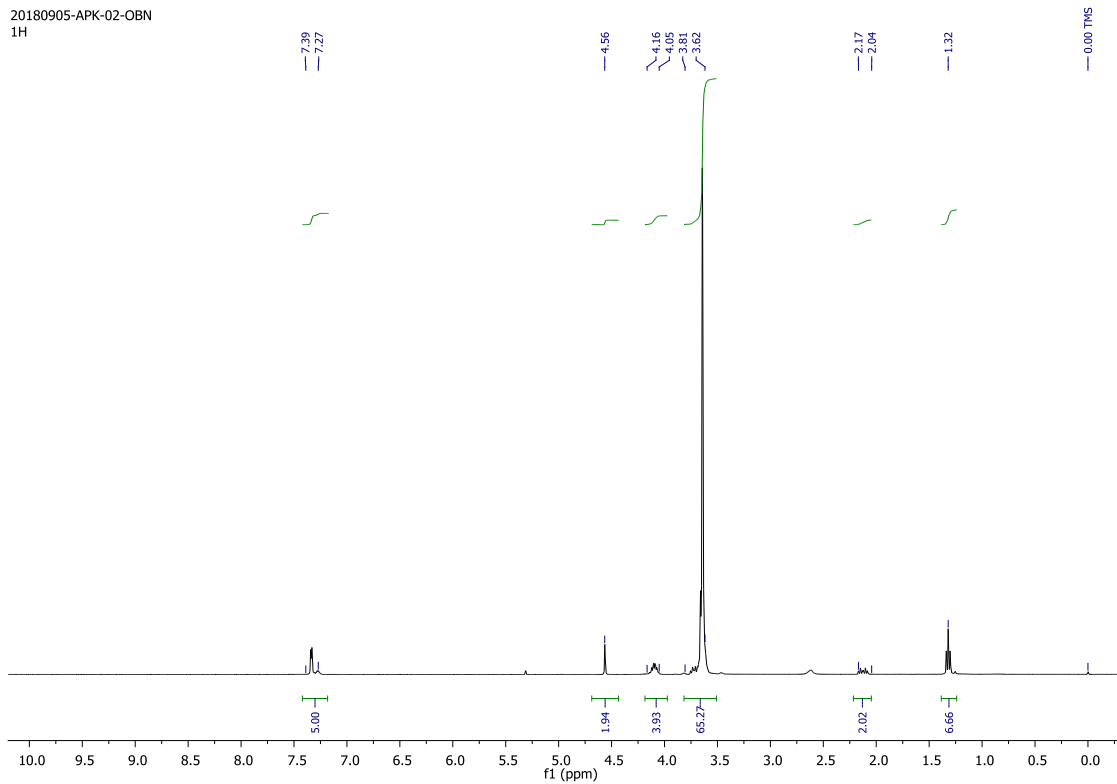
¹H NMR spectrum of compound 1i

20180911-APK-02-I



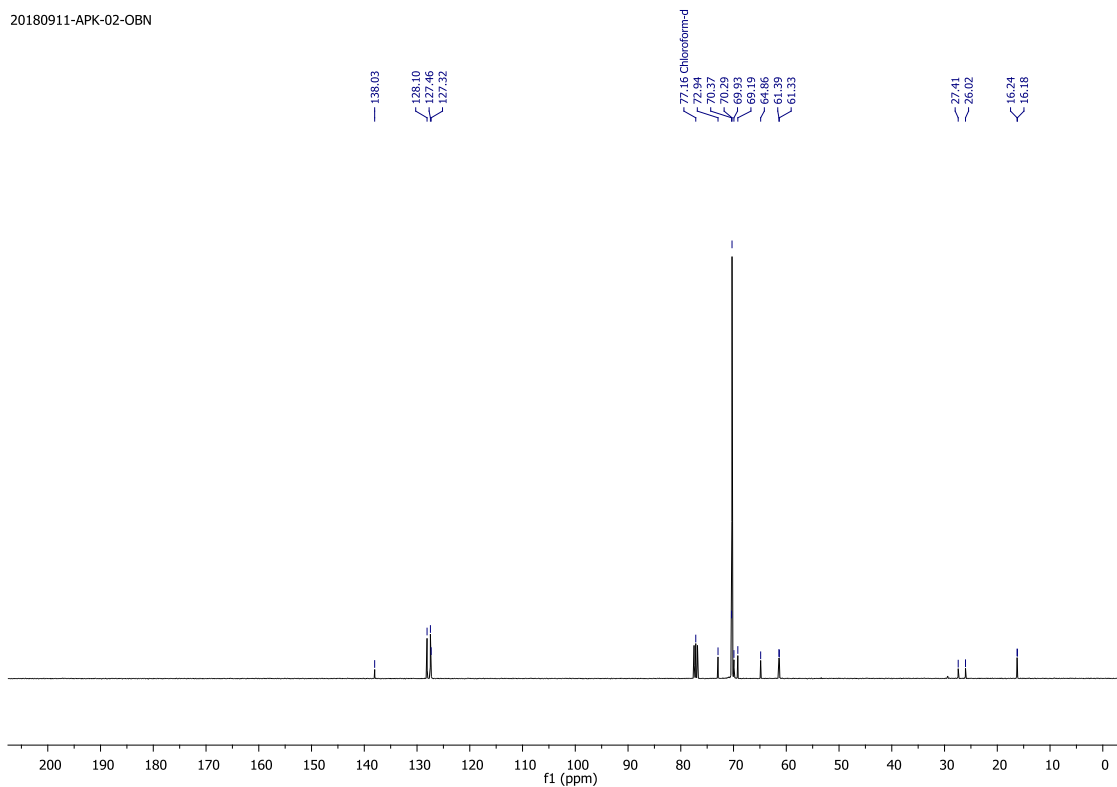
¹³C NMR spectrum of compound 1i

20180905-APK-02-OBN
1H



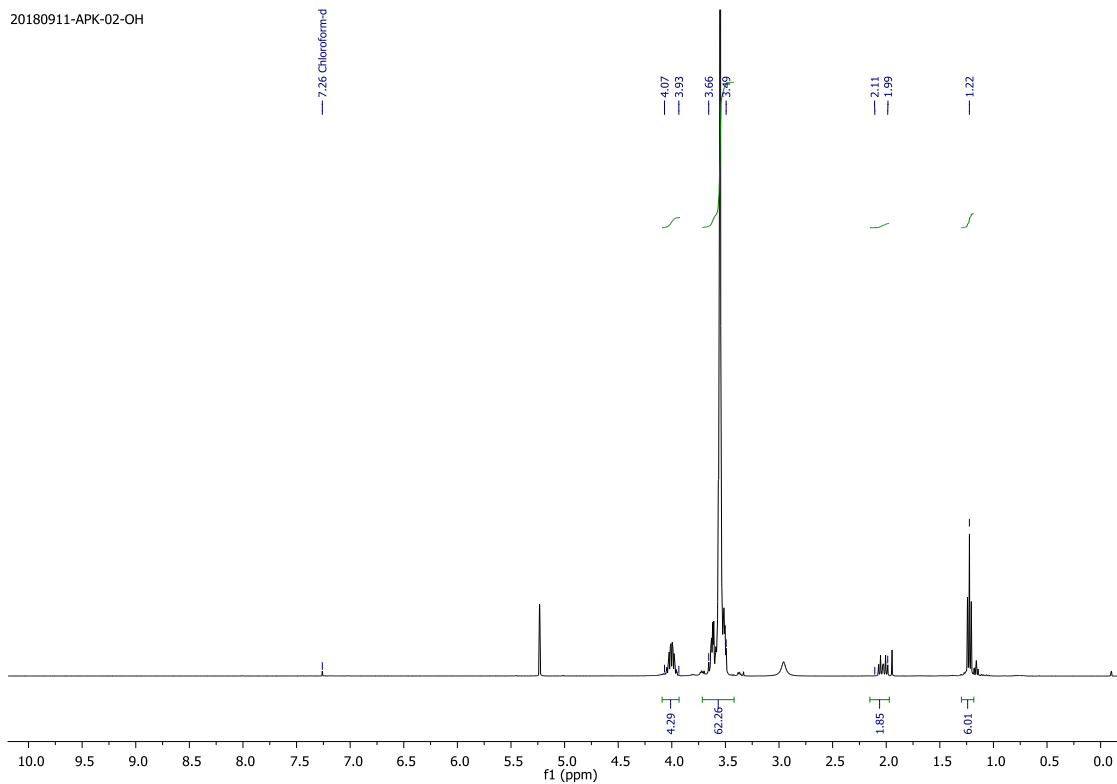
¹H NMR spectrum of compound 1j

20180911-APK-02-OBN



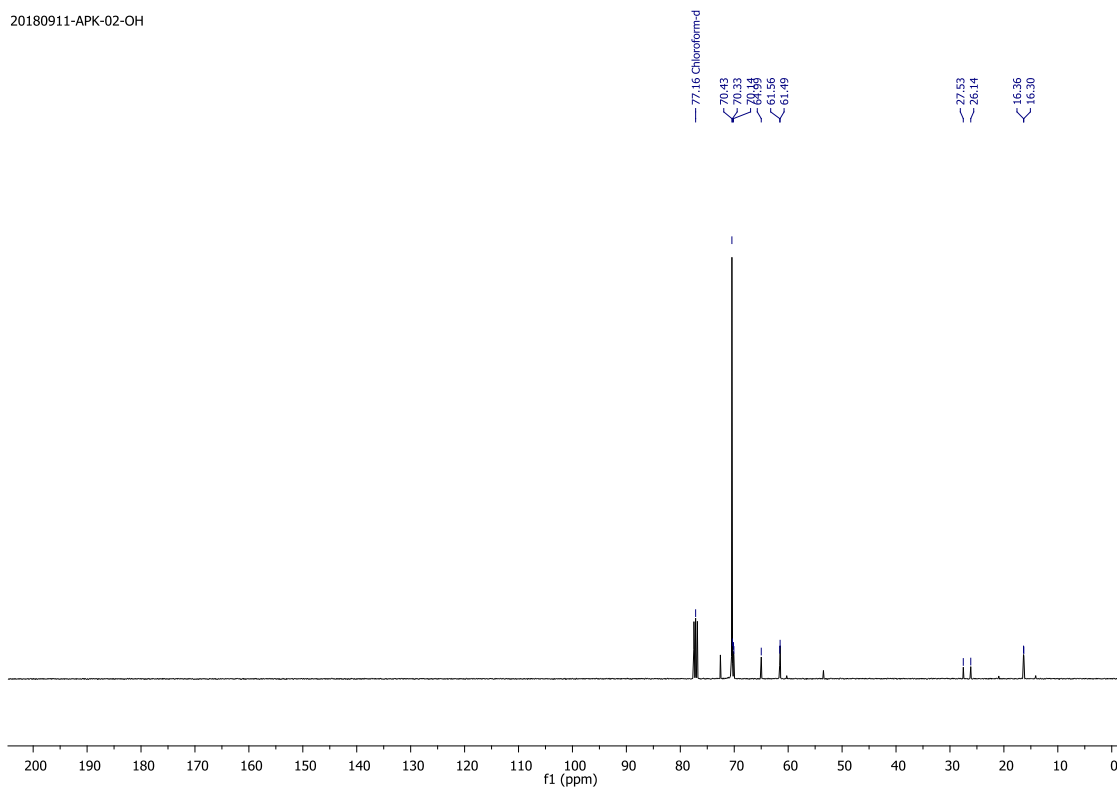
¹³C NMR spectrum of compound 1j

20180911-APK-02-OH



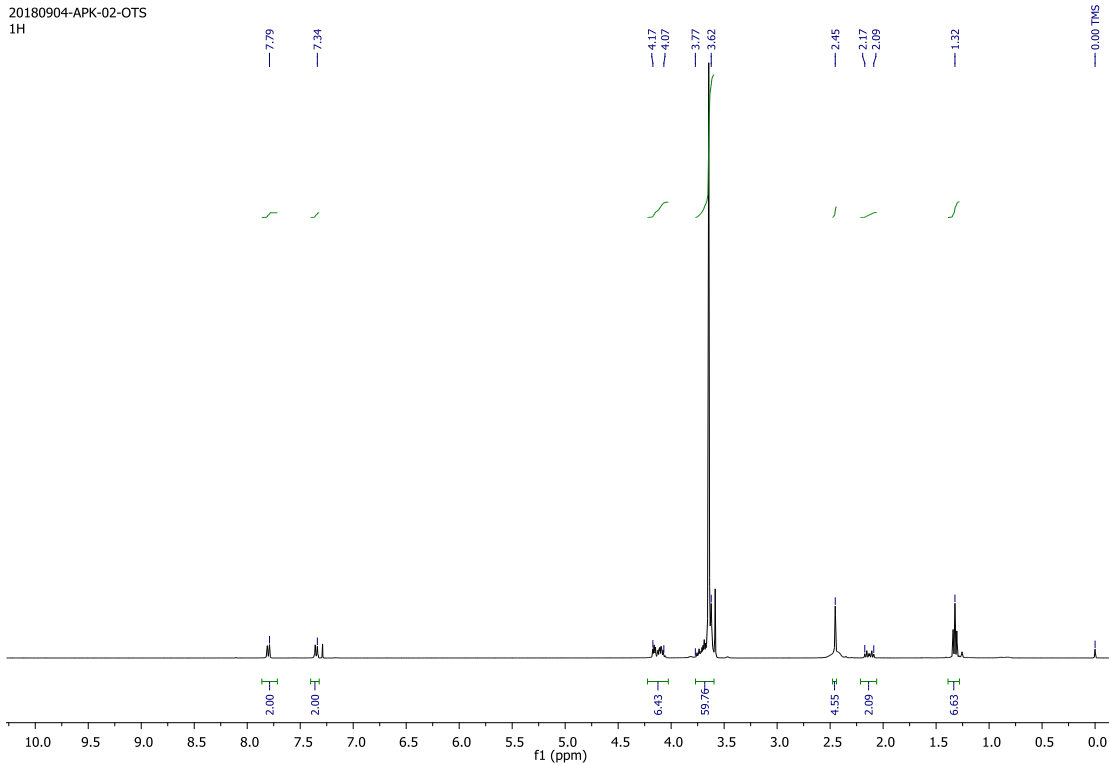
¹H NMR spectrum of compound 1k

20180911-APK-02-OH



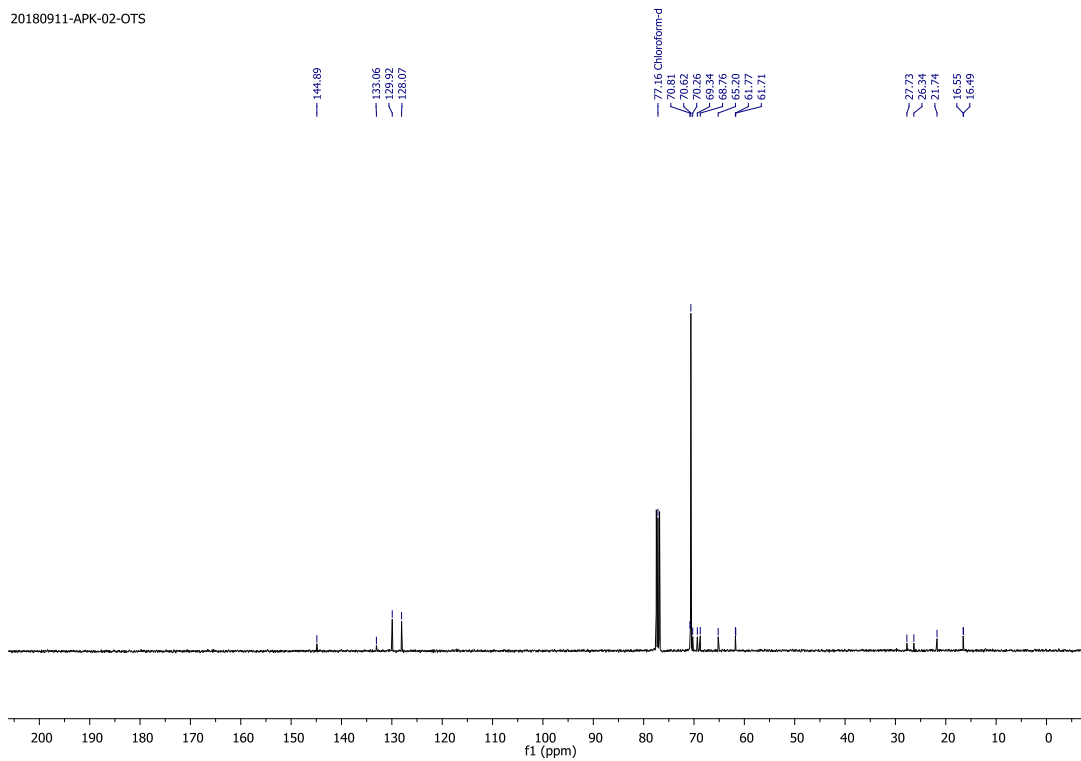
¹³C NMR spectrum of compound 1k

20180904-APK-02-OTS
1H



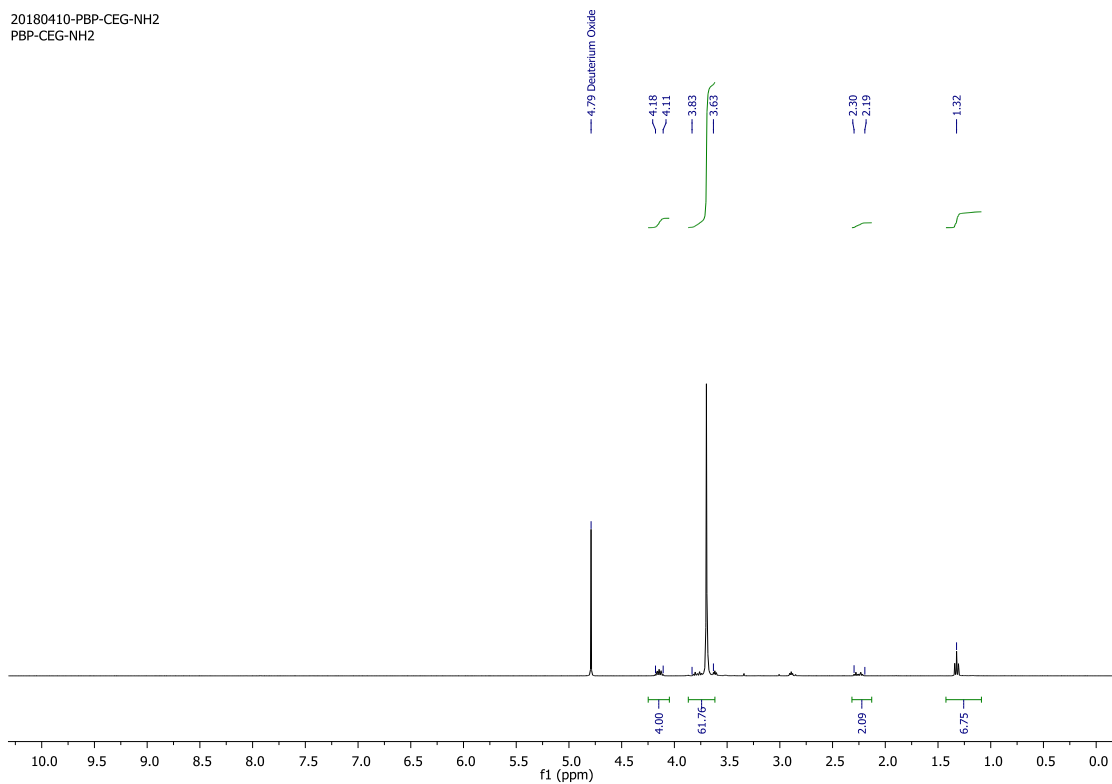
¹H NMR spectrum of compound 11

20180911-APK-02-OTS



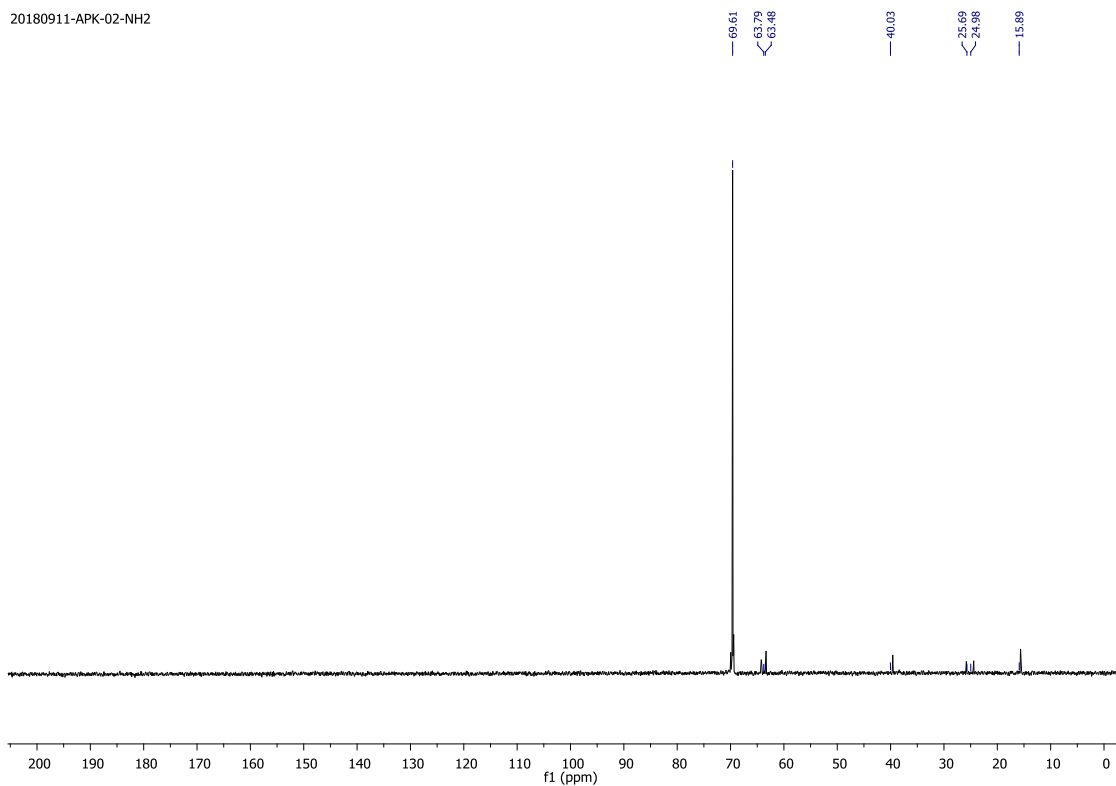
¹³C NMR spectrum of compound 11

20180410-PBP-CEG-NH2
PBP-CEG-NH2



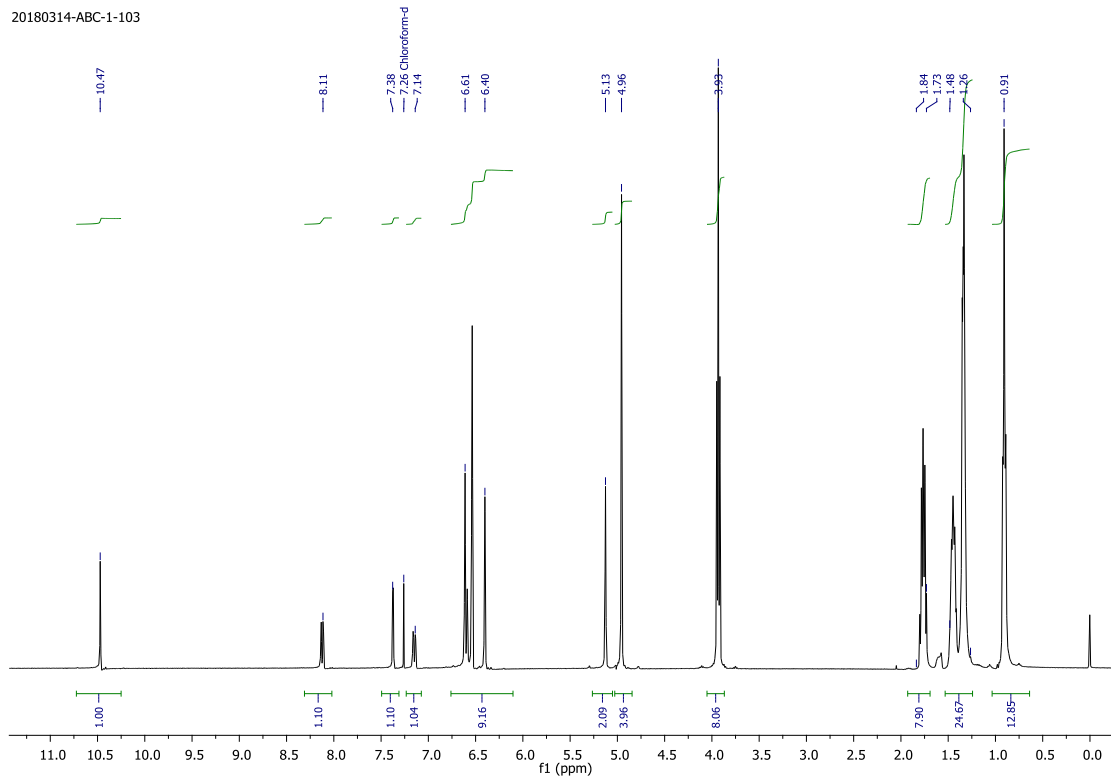
¹H NMR spectrum of compound 1n

20180911-APK-02-NH2



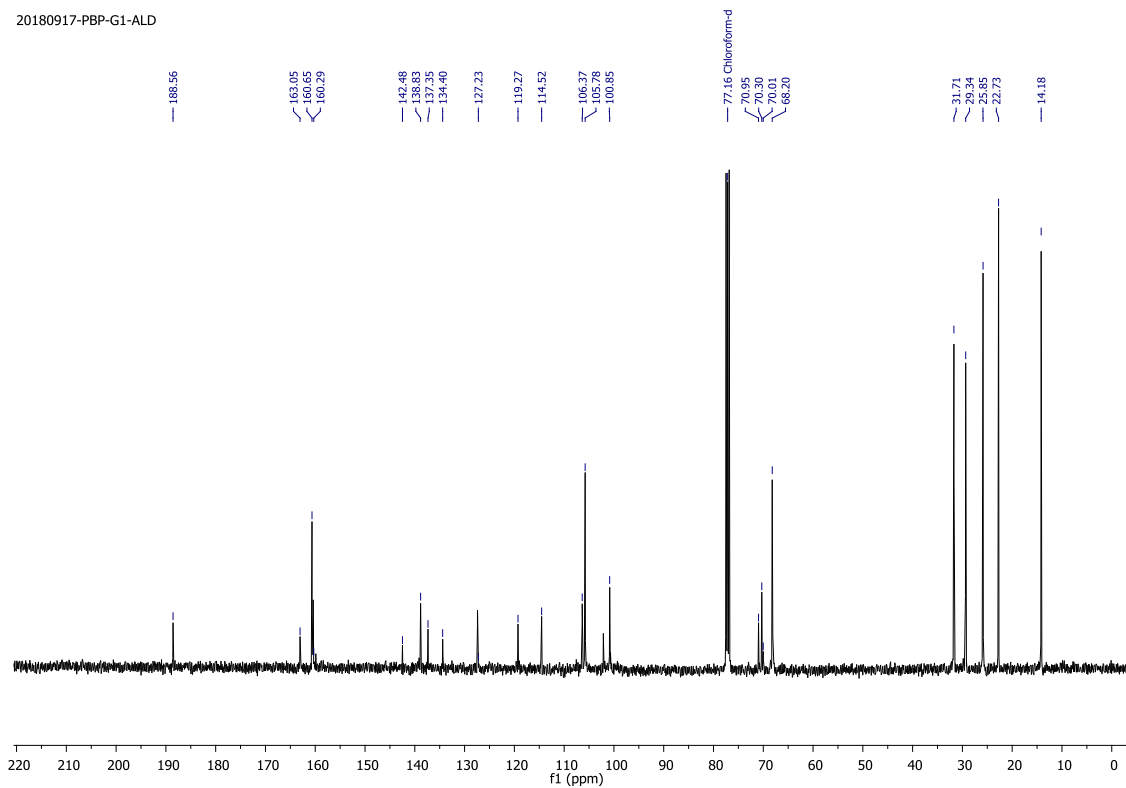
¹³C NMR spectrum of compound 1n

20180314-ABC-1-103

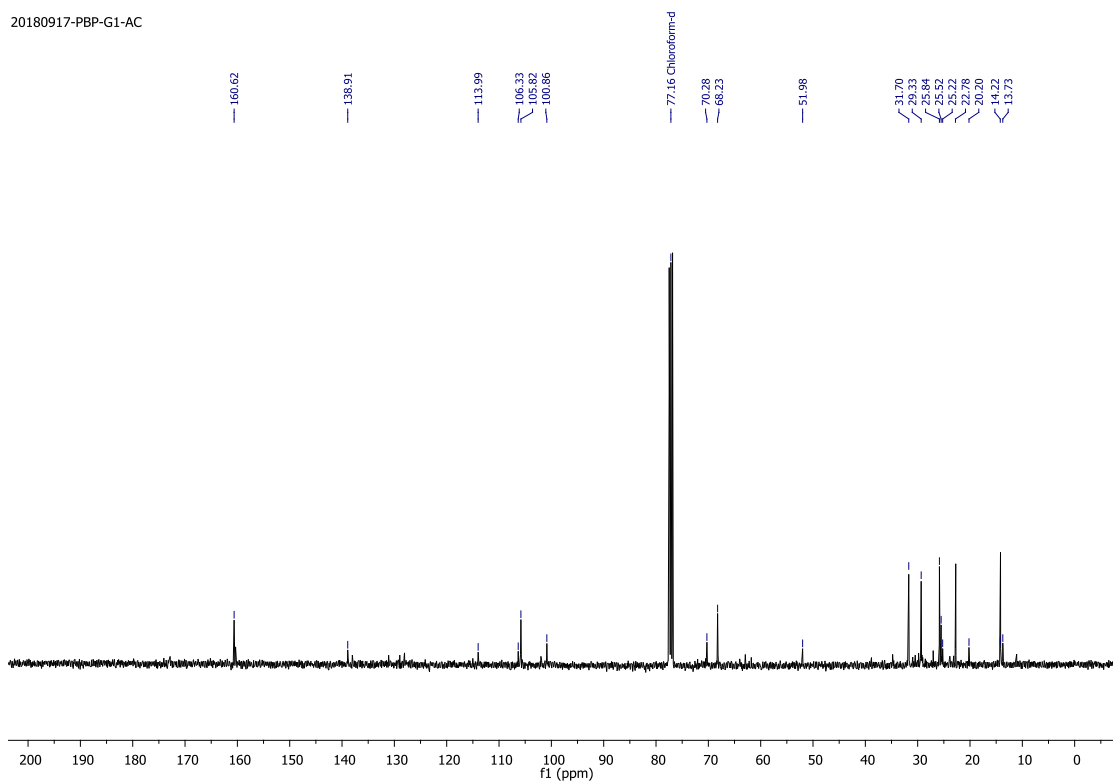


¹H NMR spectrum of compound 5a

20180917-PBP-G1-ALD

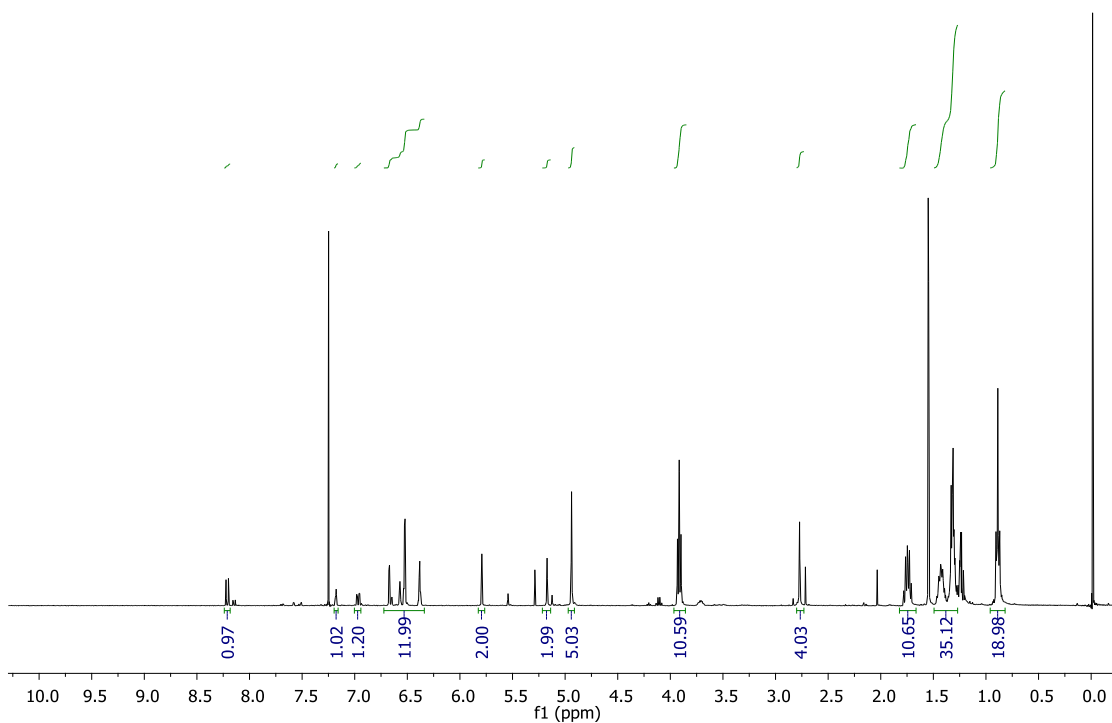


¹³C NMR spectrum of compound 5a



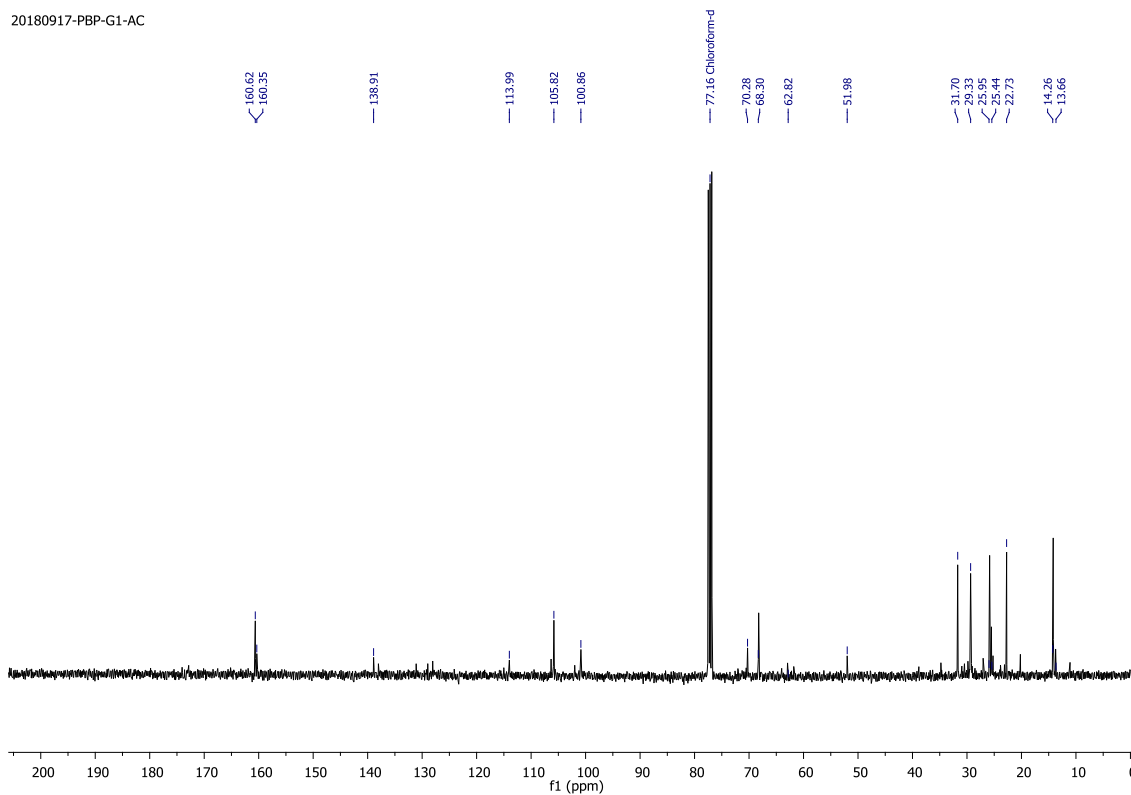
¹³C NMR spectrum of compound 5b

20180316-ABL-01-107



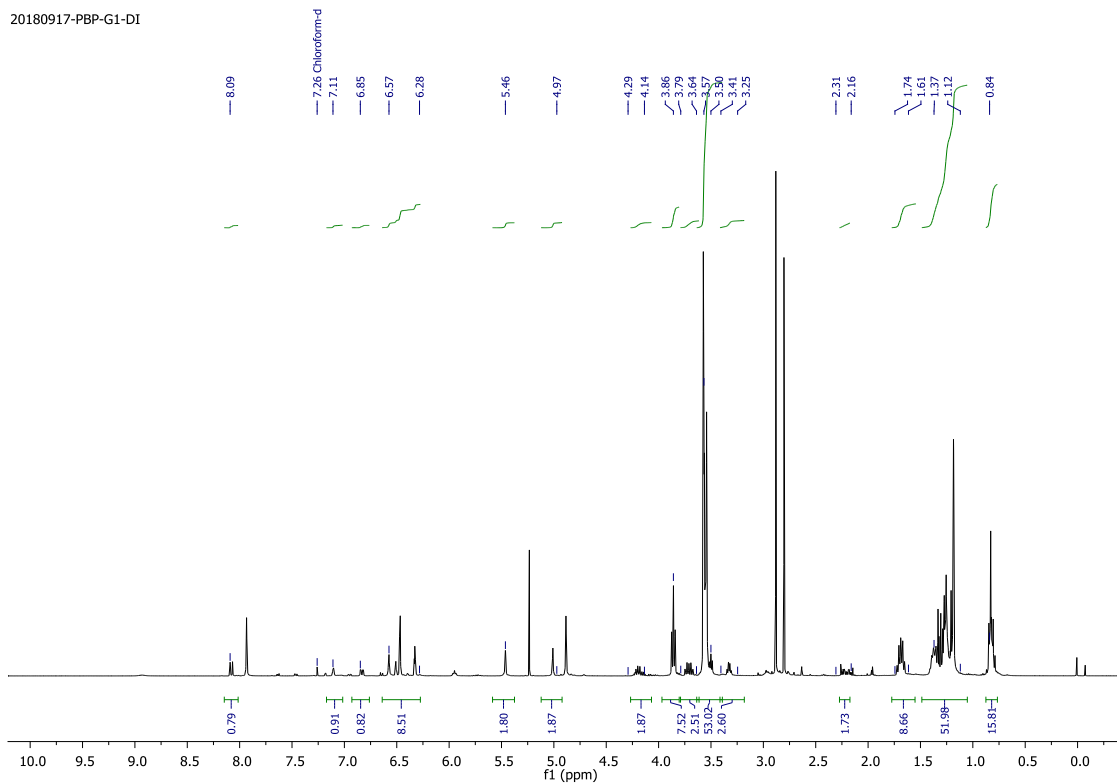
¹H NMR spectrum of compound 5c

20180917-PBP-G1-AC



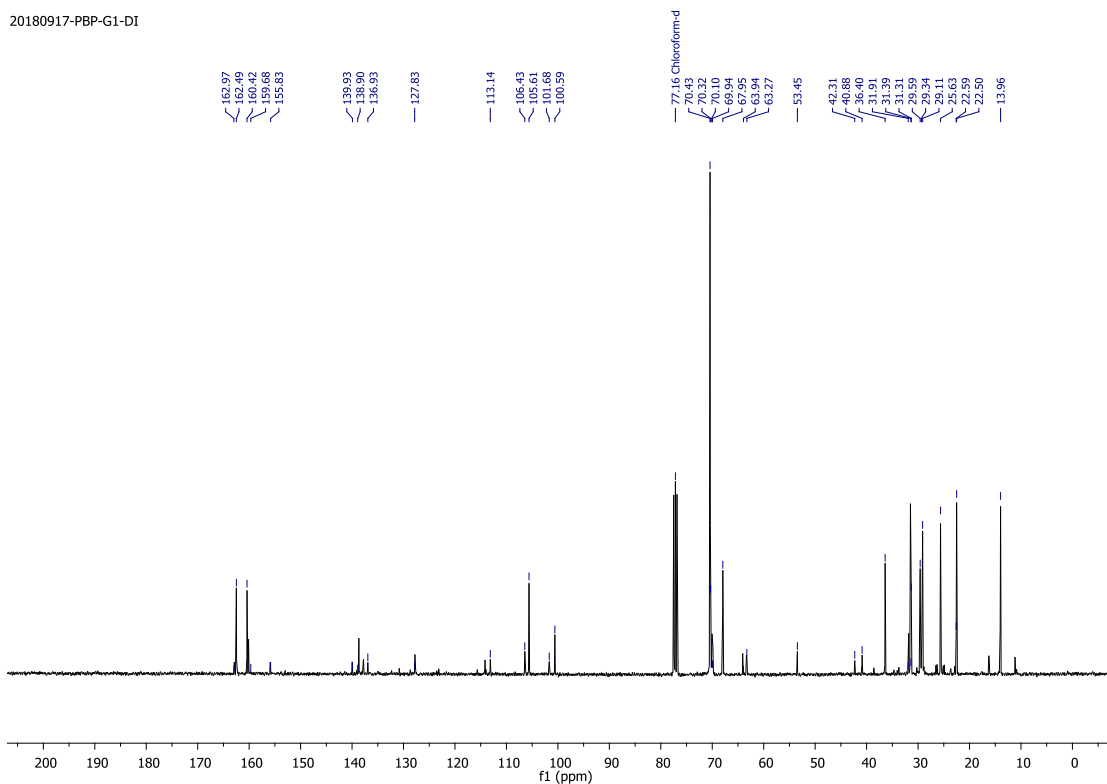
¹³C NMR spectrum of compound 5c

20180917-PBP-G1-DI



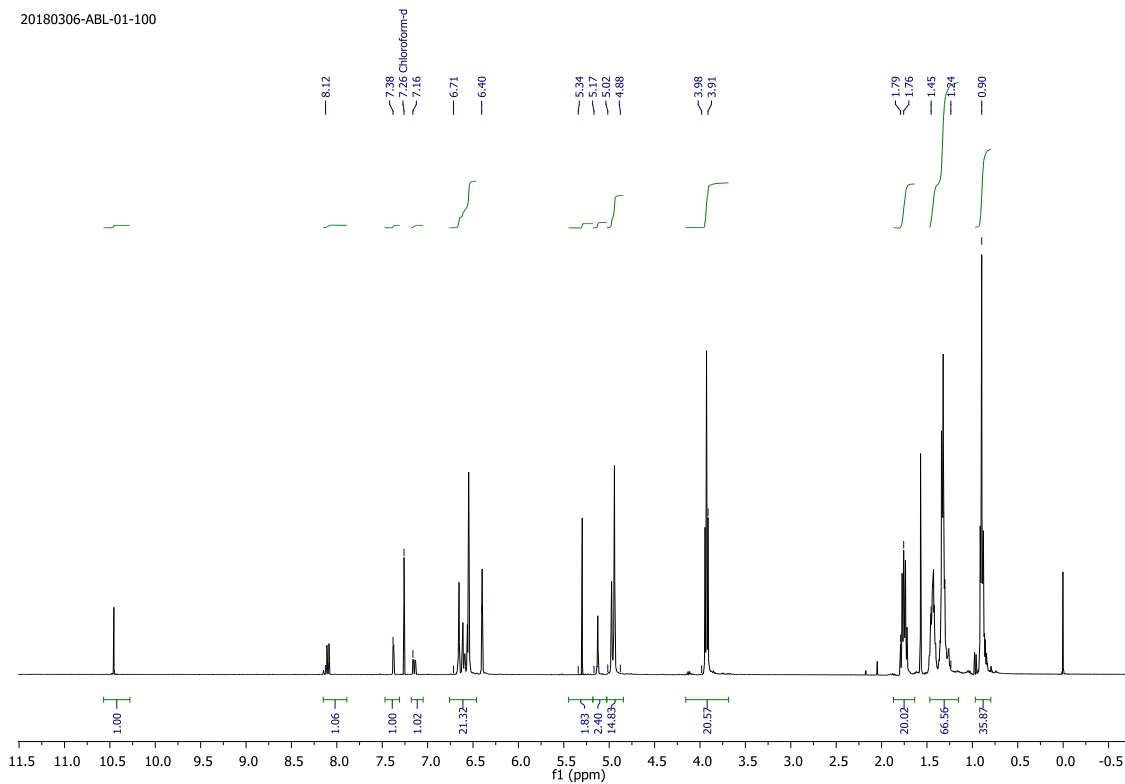
¹H NMR spectrum of compound 5d

20180917-PBP-G1-DI



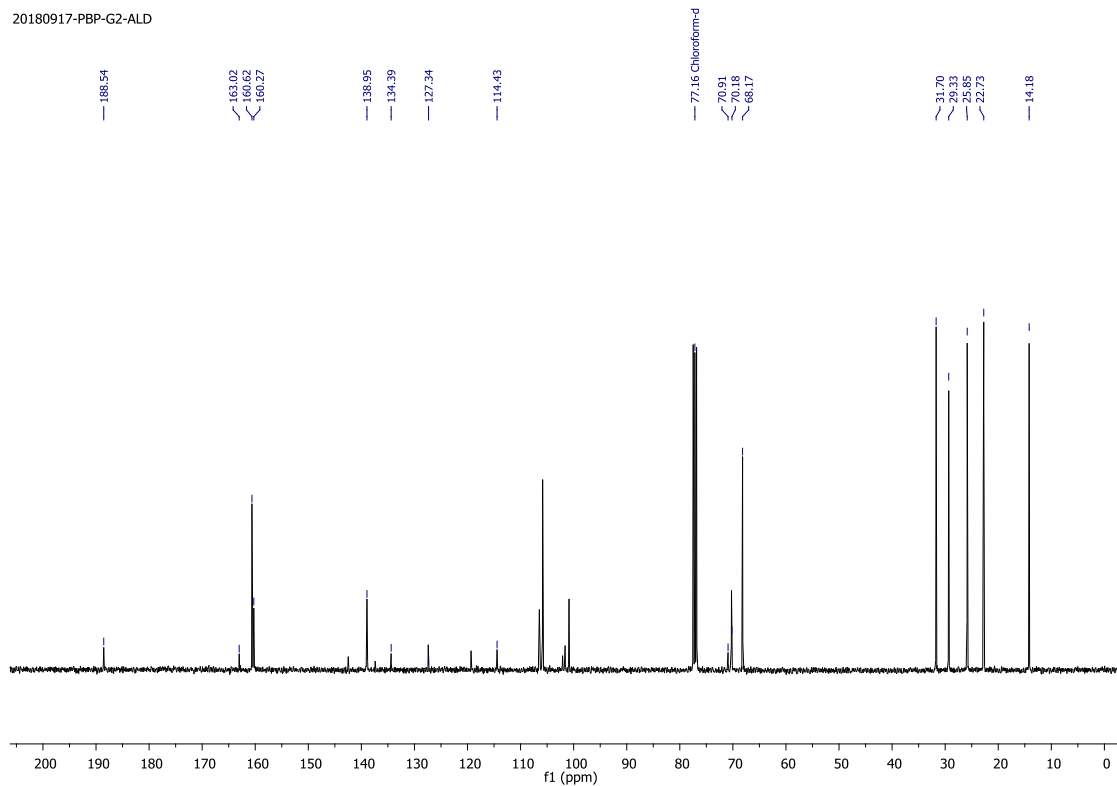
¹³C NMR spectrum of compound 5d

20180306-ABL-01-100

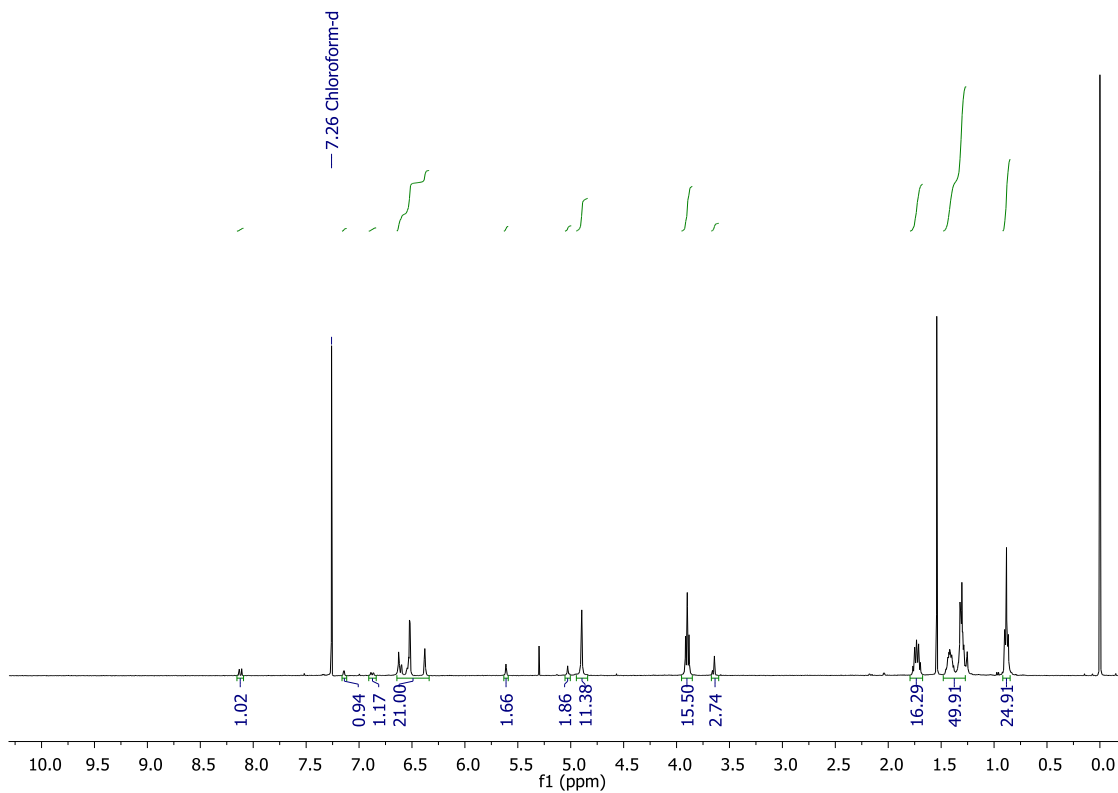


¹H NMR spectrum of compound 6a

20180917-PBP-G2-ALD

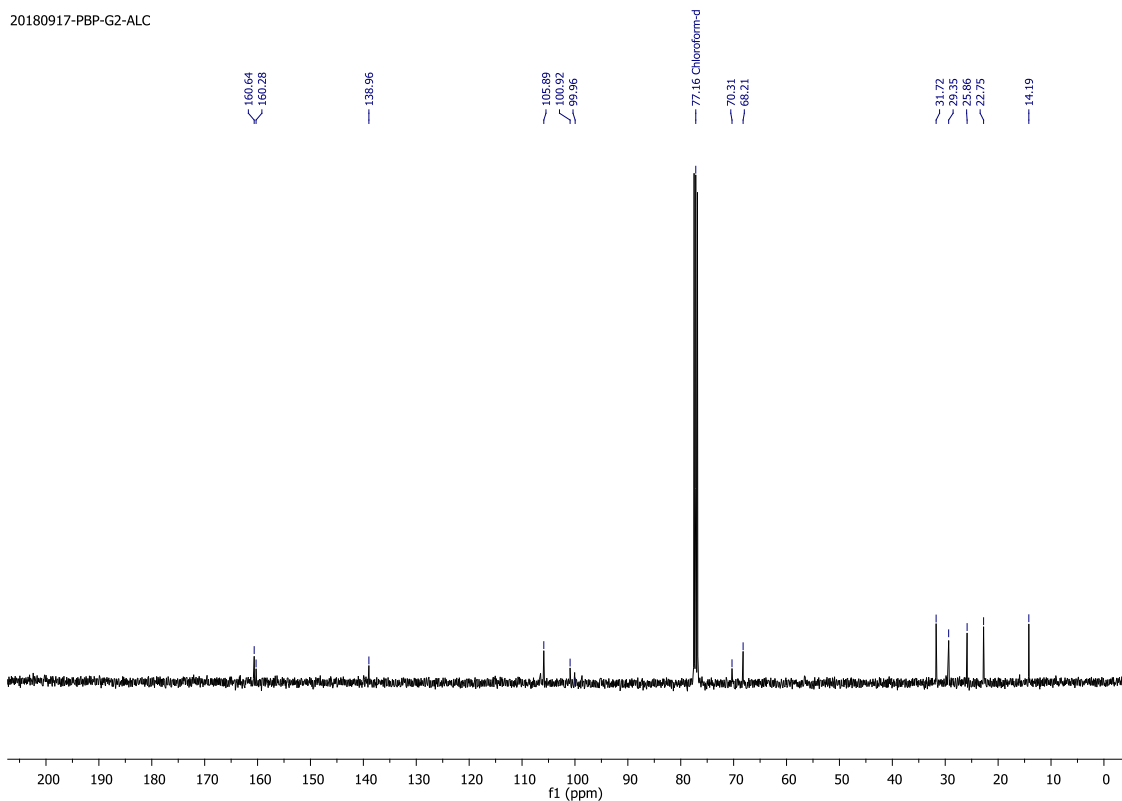


¹³C NMR spectrum of compound 6a



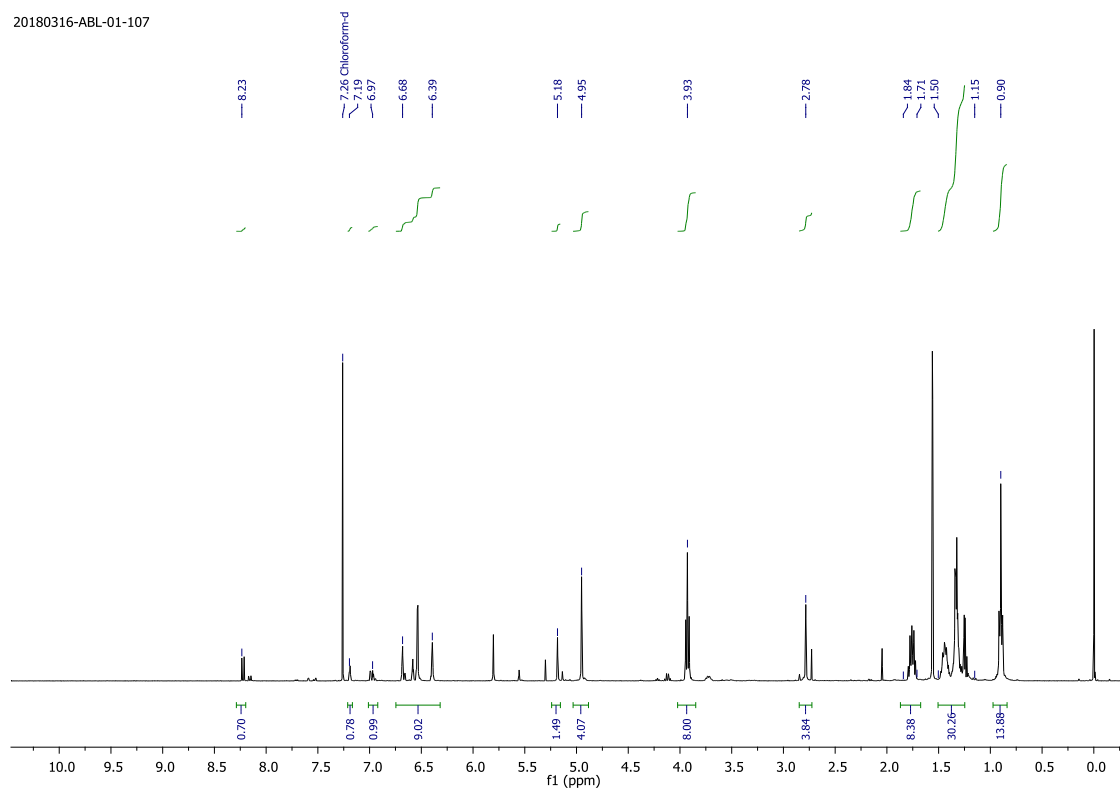
¹H NMR spectrum of compound 6b

20180917-PBP-G2-ALC



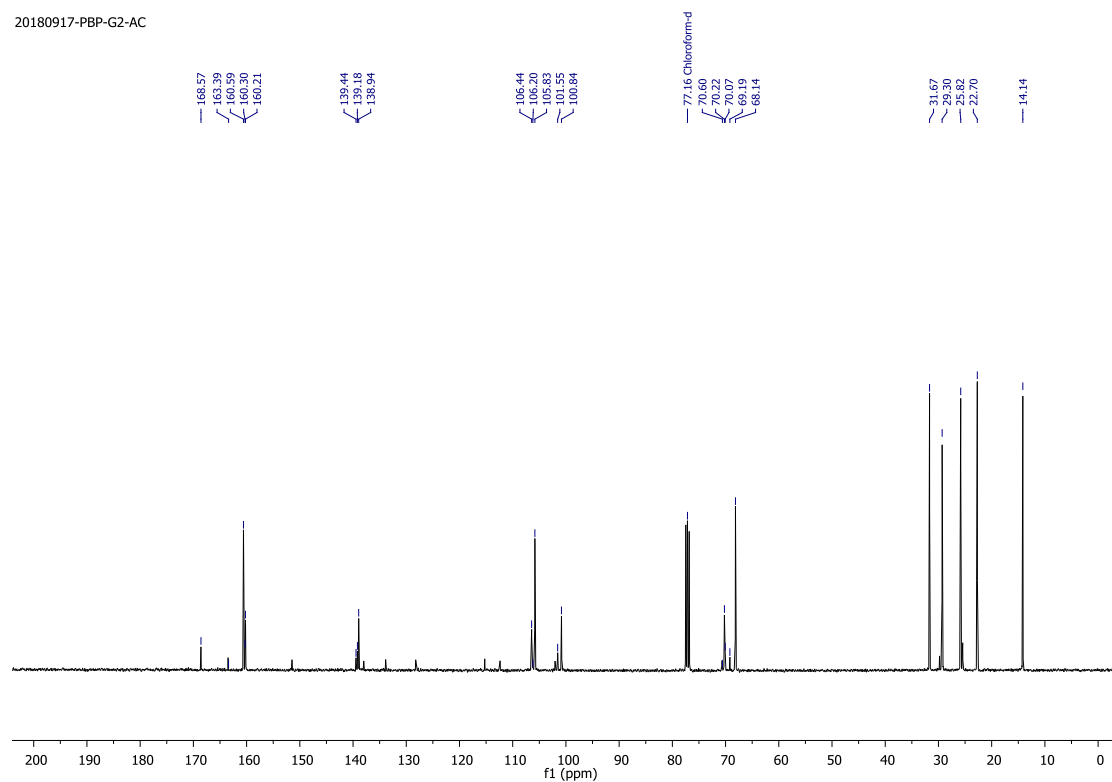
¹³C NMR spectrum of compound 6b

20180316-ABL-01-107



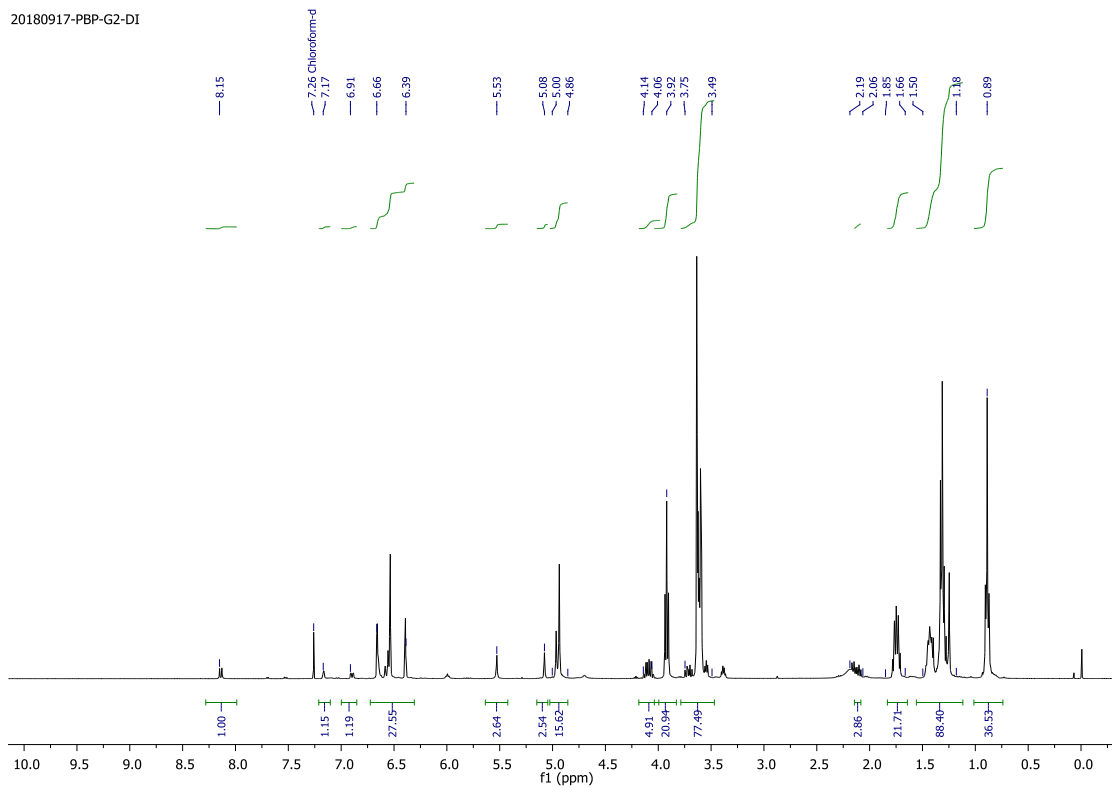
¹H NMR spectrum of compound 6c

20180917-PBP-G2-AC



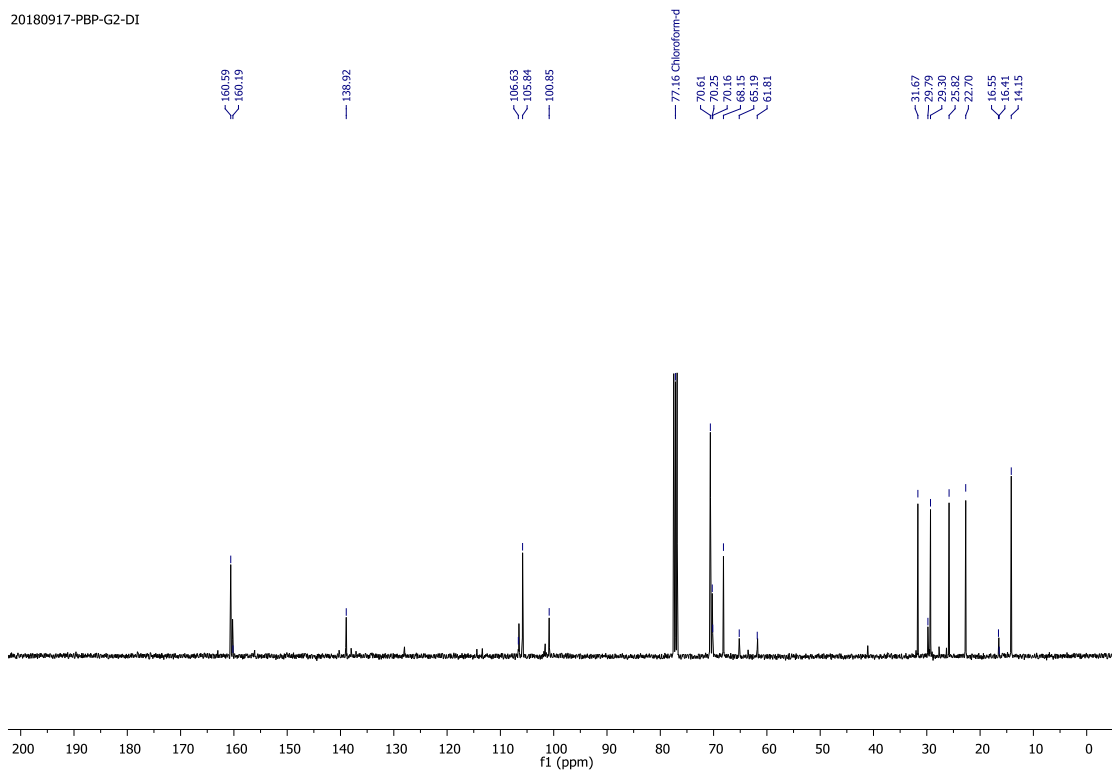
¹³C NMR spectrum of compound 6c

20180917-PBP-G2-D1



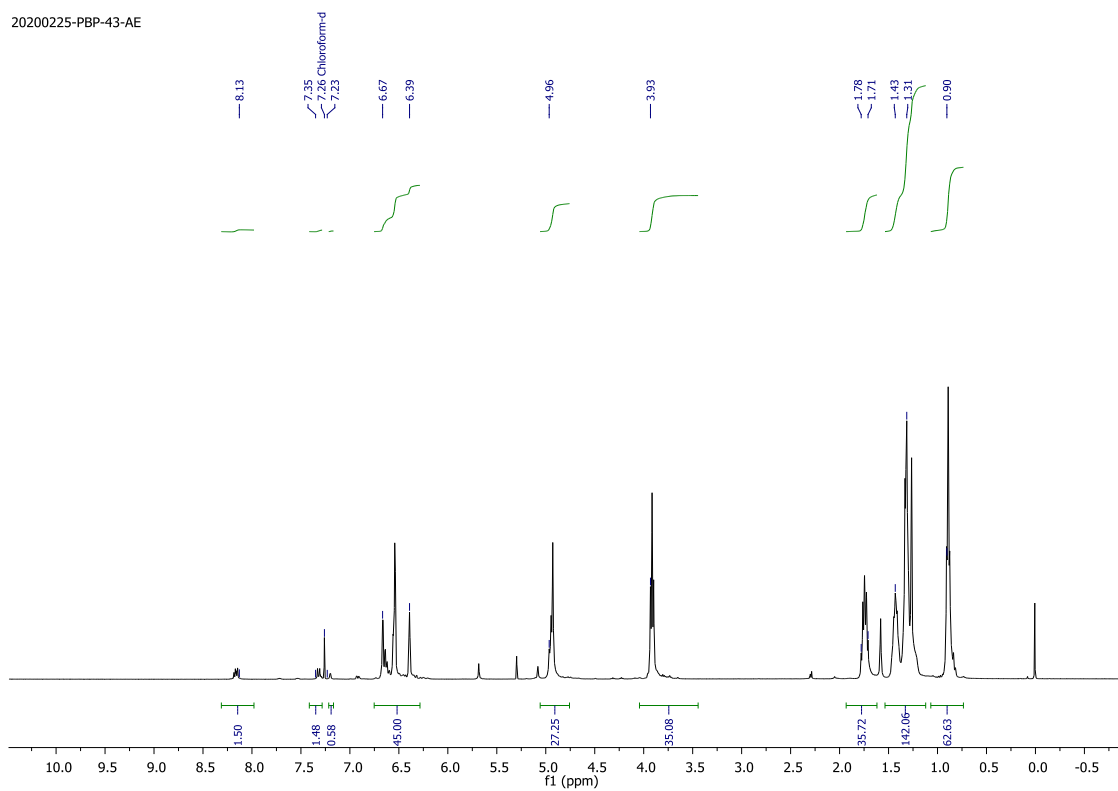
¹H NMR spectrum of compound 6d

20180917-PBP-G2-D1



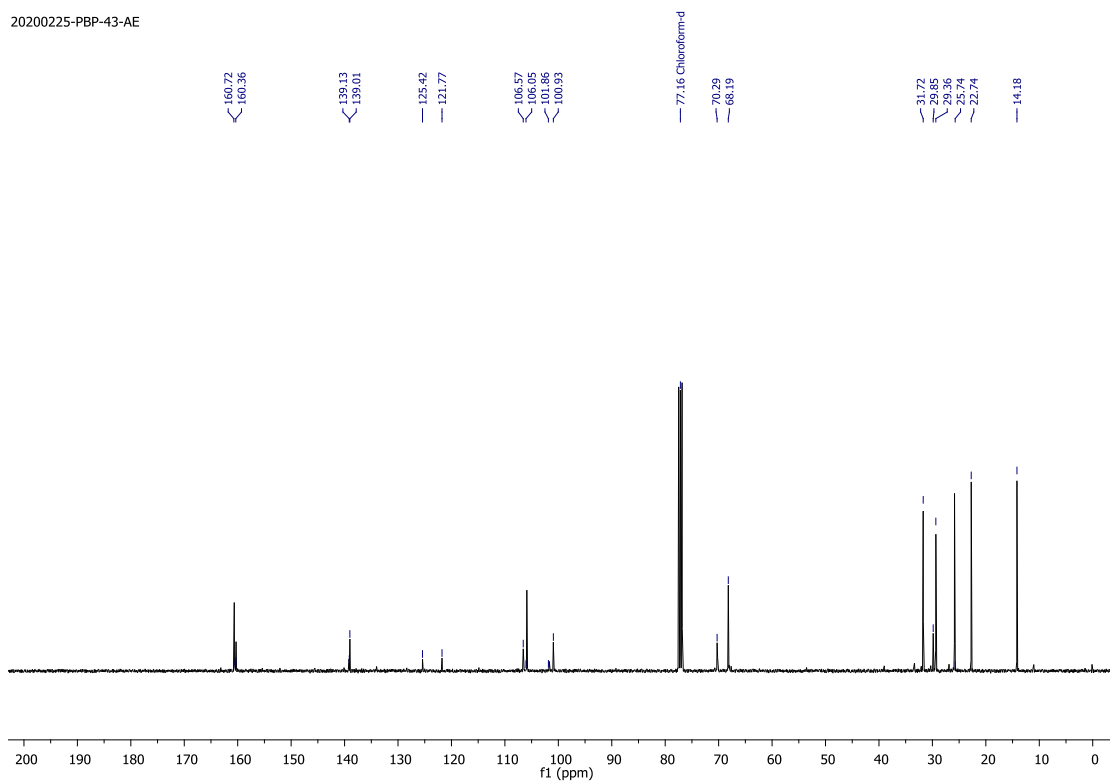
¹³C NMR spectrum of compound 6d

20200225-PBP-43-AE



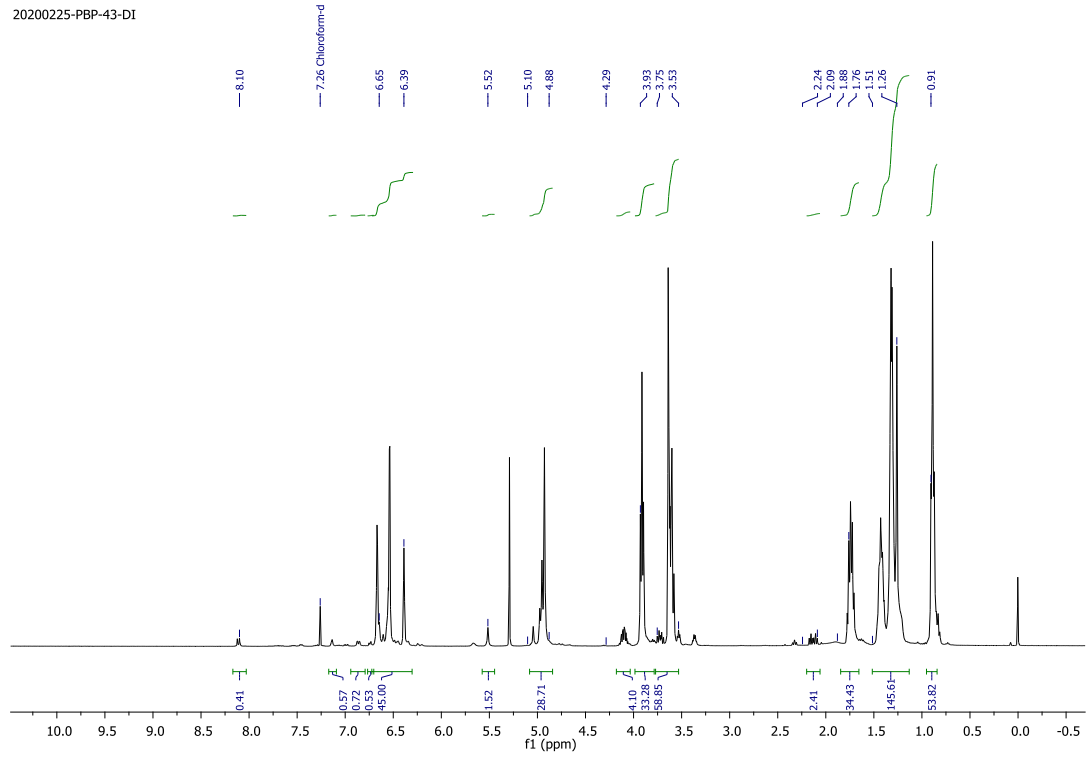
¹H NMR spectrum of compound 7c

20200225-PBP-43-AE



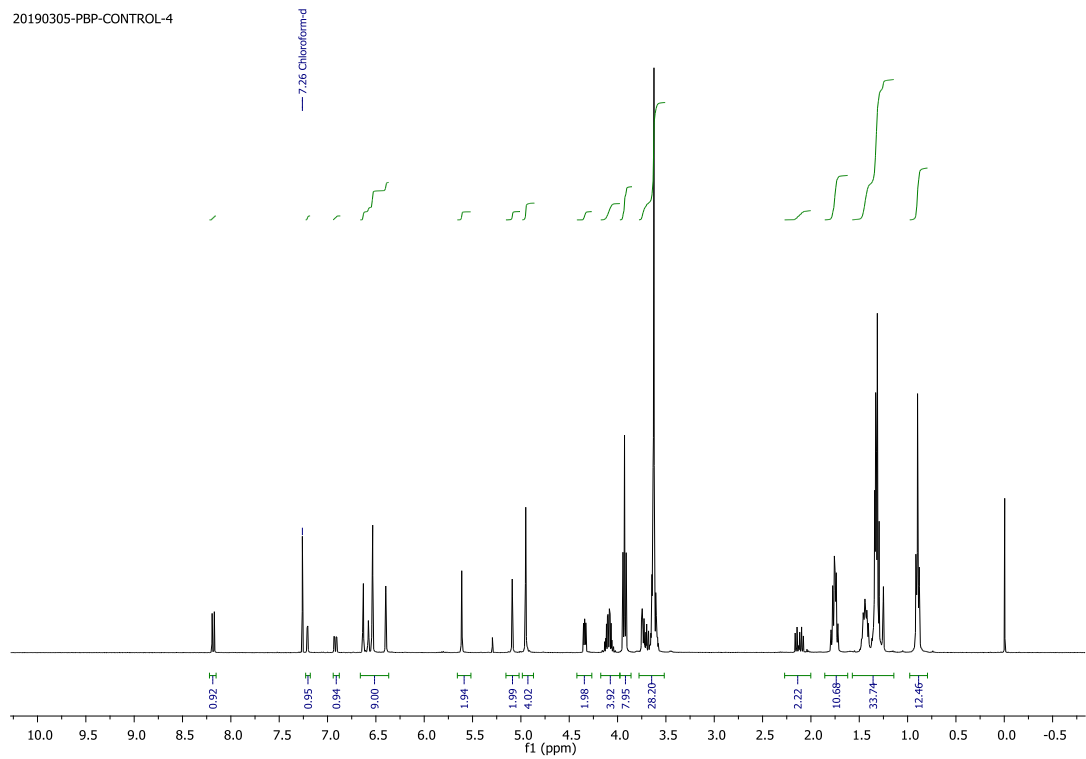
¹³C NMR spectrum of compound 7c

20200225-PBP-43-DI



¹H NMR spectrum of compound 7d

20190305-PBP-CONTROL-4



¹H NMR spectrum of compound 8a

CHAPTER 4

PROGRAMMED DISASSEMBLY OF REDOX-RESPONSIVE SUPRAMOLECULAR PROTEIN ASSEMBLIES

4.1 Introduction

An enormous amount of studies have been reported in the literature to construct nano and microscopic assemblies by using polymeric systems, which respond to a wide range of stimuli such as pH,¹ redox,² temperature,³ and light⁴. In addition to that, various strategies have been reported for the synthesis of smart protein-polymer conjugates *via* conjugation of the stimuli-responsive polymeric domain onto the protein. These conjugates have gained enormous interest in recent years due to its potential application in the area of drug delivery,⁵ *in vivo* imaging,⁶ and biocatalysis⁷. Even though chemical strategies offer unlimited chemical space to construct dynamic protein assemblies; the success in designing protein-polymeric assemblies which can respond to a wide range of stimuli such as redox, pH, metabolites, enzymes, light, and mechanical forces is limited.⁸ This is because the efforts in designing protein assemblies have mostly been focused upon increasing the structural sophistication. As a result, less attention has been paid in designing protein assemblies, which can respond to a wide range of stimuli.

The existing strategies are mainly focused on designing temperature-responsive protein assemblies. This is primarily achieved by a conjugating polymer having LCST property to the protein of interest. Above the LCST, the polymer becomes hydrophobic, the resultant conjugate collapse to form assemblies driven *via* attractive hydrophobic interaction. Whereas, below the LCST, the conjugate will be hydrophilic and hence will exist in monomeric form.⁹ The attractive feature of this method is, it is relatively easy to synthesize because both globular protein and polymer are hydrophilic and hence freely soluble in an aqueous medium at room temperature. On the flip side, this methodology is limited to only polymers having LCST property, thus severely limiting the chemical diversity of the polymers scaffold. In this aspect, the synthesis of a diverse set of stimuli-responsive protein assemblies would be useful not only in reproducing the dynamics of natural protein assemblies but also in creating completely novel, orthogonal, and functional artificial protein assemblies.

In this aspect, we have reported the synthesis of static and dual-sensitive protein-dendron bioconjugates in chapter 2 and 3, respectively.¹⁰ Particularly in chapter 3, we have

shown that during the photolysis of nitrobenzyl group, these photo-sensitive complexes bearing nitrobenzyl group undergoes rapid rearrangement reactions to form imine functionalized bioconjugate, which then self-assemble into complex driven *via* hydrophobic interactions. The use of a photosensitive nitrobenzyl group as an external trigger provides excellent spatiotemporal control. However, the phototoxicity associated with the UV light and its less penetration depth into the skin are the significant limitations.¹¹ In this aspect, the development of materials that can respond to physiological changes such as pH, redox, imbalance in enzyme or protein concentration, metabolite are highly relevant for biomedical applications.¹² Towards that goal, in this chapter, we extend our chemical methodology for the design of monodisperse redox-responsive facially amphiphilic protein-dendron bioconjugates.

4.2. Results and Discussion

4.2.1. Macromolecular Design

We have previously reported (chapter 3) the technique for the construction of multi-responsive protein-dendron assemblies. Further, to extend the scope and diversity of our technology, we here demonstrate the construction of redox -responsive supramolecular protein-dendron complexes. The molecular design of redox-responsive bioconjugate is similar to a photo-responsive protein-dendron bioconjugate, except the (2-nitrobenzyl derivative) is replaced with the redox-sensitive disulfide bond. Thus, the macromolecular design has four core structural elements: (i) hydrophilic globular protein, (ii) flexible hydrophilic linker, (iii) redox-responsive group, and (iv) hydrophobic dendrons (Figure 4.1).

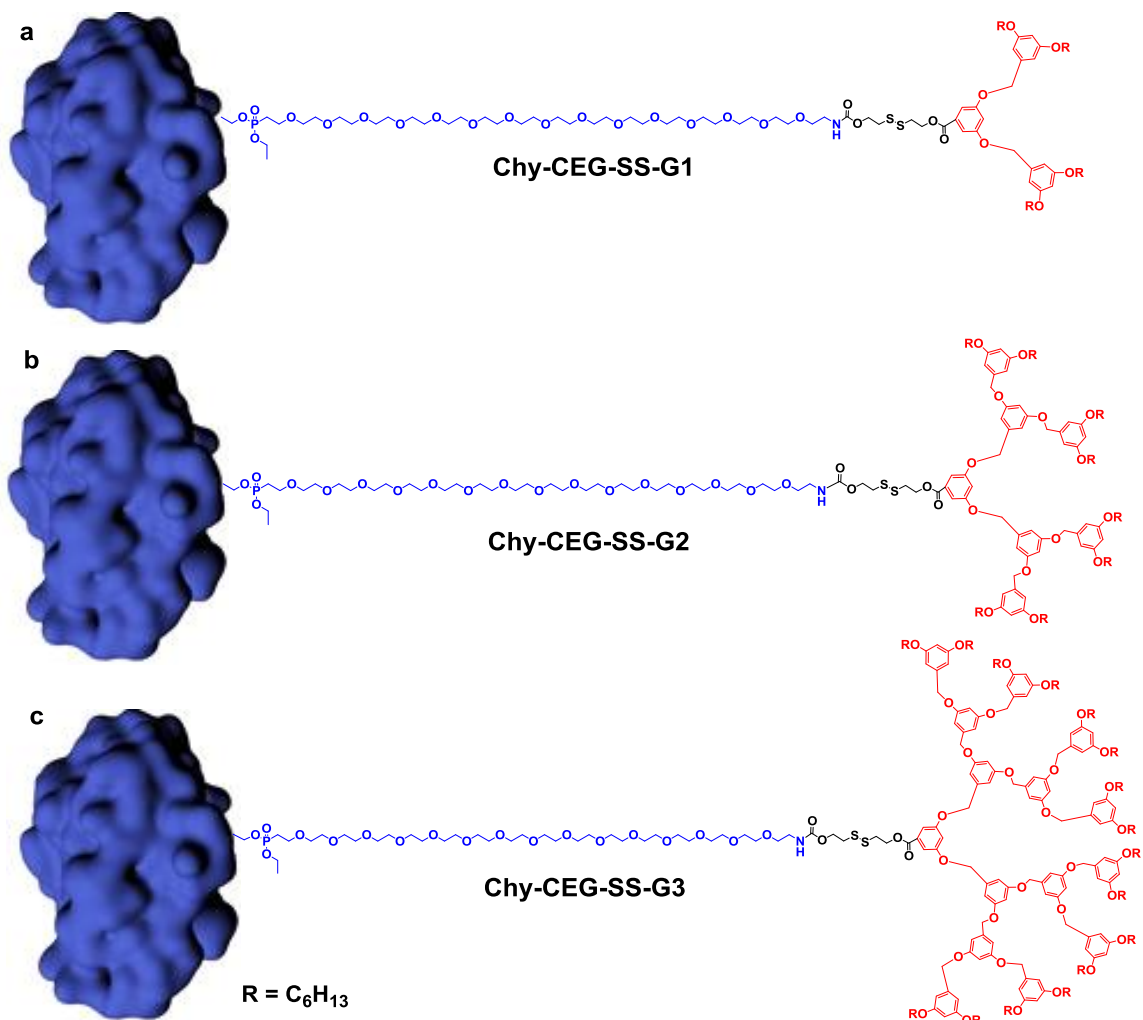
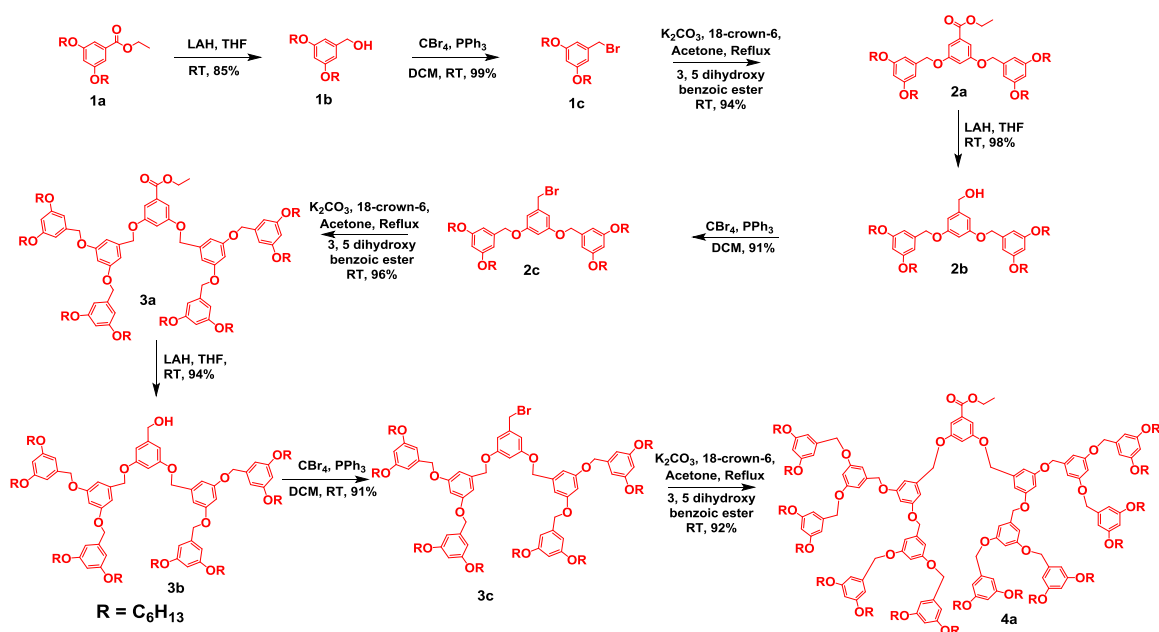


Figure 4.1 | Schematic representation of chemical structures of redox-sensitive protein-dendron bioconjugates. These conjugates are composed of protein (blue = chymotrypsin), hydrophilic oligoethylene glycol linker (blue) (CEG = cetylene glycol), redox-sensitive disulphide bond (black), and hydrophobic dendron block (red). Structures of (a) Chy-CEG-SS-G1 (b) Chy-CEG-SS-G2 (c) ChyCEG-SS-G3.

4.2.3. Synthesis of a Redox-Sensitive Amphiphilic Activity-Based Probe

The design strategy for the synthesis of stimuli-responsive macromolecular AABP is shown in Scheme 4.2a. First, we synthesized G1, G2, and G3 esters by following a previously reported method (Scheme 4.1).¹³ Then, the obtained G1 (2a), G2 (3a), and G3 (4a) esters compounds were subjected to hydrolysis by treatment with NaOH in ethanol to get corresponding acid derivatives (5b). The G1-, G2-, and G3- acid derivatives (5b) were then reacted with 2,2'-disulfanediybis(ethan-1-ol) in the presence of EDC and DMAP in

DCM to get compound 5c followed by activation using N, N'-DSC, in the presence of Et₃N to afford compound 5d. The activated ester 5d was then reacted with linker amine (6) in the presence of Et₃N and DMF to obtain compound 5e. Then, the resultant diphosphonate ester 5e was treated with oxalyl chloride; however, due to the acidic reaction conditions, we observed degraded products by using MALDI-ToF. Then, the diphosphonate ester 5e was heated with lithium bromide (LiBr) in DMF to get monophosphonate ester 5f, which finally on fluorination using DAST in DCM afforded fluorophosphonate 5g (Scheme 4.2).

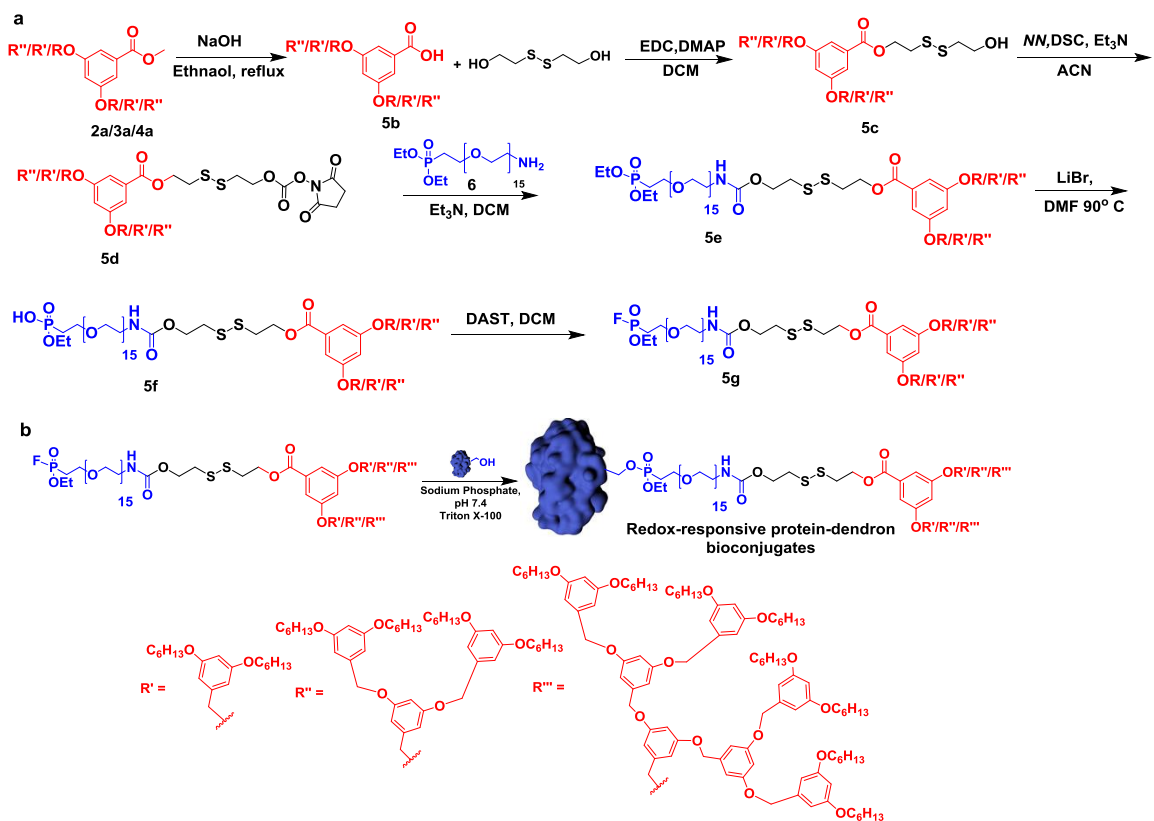


Scheme 4.1 | Scheme for the synthesis of ester-terminated dendrimers.

4.2.4. Synthesis of Redox-Sensitive Protein-Dendron Conjugates

We have followed a similar strategy for the synthesis of redox-responsive protein-dendron bioconjugate.¹⁴ In short, the conjugation of redox-sensitive macromolecular AABPs with chymotrypsin was attempted in 50 mM phosphate buffer pH 7.4 by using a micelle-assisted protein labeling (MAPLab) technology reported by our group. In this approach, the redox-responsive macromolecular AABPs were solubilized in 10X critical micelle concentration (CMC) of triton X-100 at room temperature. Subsequently, obtained homogeneous solution was then treated with chymotrypsin (Scheme 4.2b). Then, the extent of the formation of protein-dendron bioconjugates was monitored using MALDI-ToF at

different time points. After 12 h, a significant peak for all the bioconjugates was observed with minor unreacted native protein (Figure 4.3a).



Scheme 4.2 | Scheme for the synthesis of redox-sensitive protein-dendron AABPs and bioconjugates. (a) Scheme for the synthesis of redox-sensitive macromolecular AABPs. (b) Scheme for the synthesis of redox-sensitive protein-dendron bioconjugates.

Next, the purification of these conjugates is performed by using the previously reported procedure by our group.¹⁴ Briefly, triton X-100 and excess of macromolecular AABP in the reaction mixture were separated by using IEX chromatography. Here cationic or anionic unreacted protein and its corresponding conjugate would stick to the charged column through non-covalent strong electrostatic interactions, and the neutral molecules would pass through the column first (Figure 4.3a-c). Then, the unreacted protein and its corresponding conjugate were separated by using SEC chromatography, where the conjugate being facially amphiphilic forms higher-order complex driven *via* hydrophobic interaction and elutes at an earlier time point, while the monomeric native protein elutes at a later time point (Figure 4.3e-g). MALDI-ToF results of purified conjugates show that they are extremely pure and devoid of any native protein (Figure 4.3b).

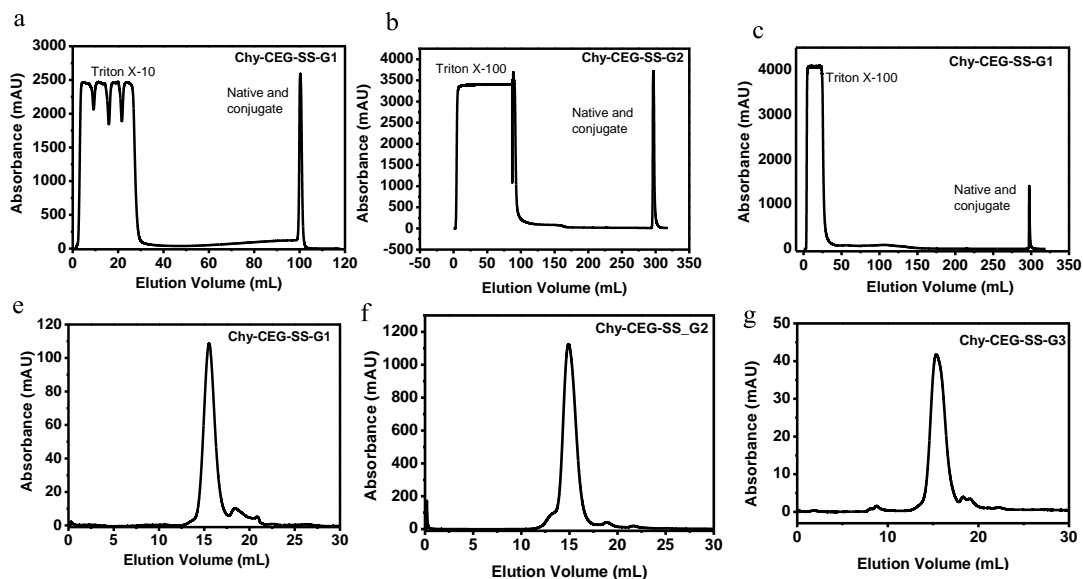


Figure 4.2 | IEX and SEC chromatogram. The IEX is performed to remove neutral triton X-100 and excess of macromolecular AABP. IEX and SEC of (a), (e) Chy-CEG-SS-G1, (b), (f) Chy-CEG-SS-G2, (c), (g) Chy-CEG-SS-G3, respectively.

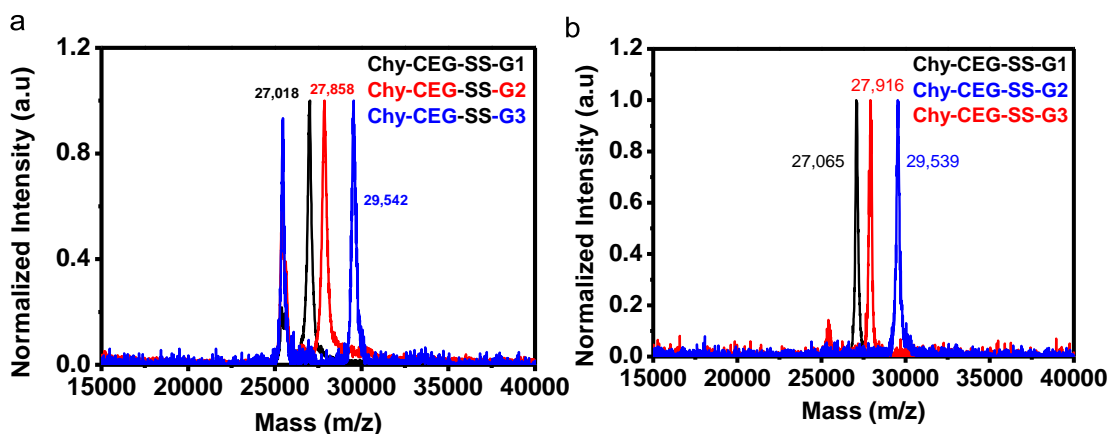


Figure 4.3 | MALDI-ToF characterization redox-sensitive protein-dendron bioconjugates. (a) MALDI-ToF of the reaction mixture (Library I), and (b) purified bioconjugates, respectively.

4.2.5. Self-Assembly Studies of Redox-Sensitive Protein-Dendron Conjugate

After the synthesis, purification, and detailed analytical characterization of protein-dendron bioconjugates, we sought to investigate their self-assembly properties by using complementary techniques such as DLS and SEC (Figure 4.4a-b). The hydrodynamic

diameter of redox-sensitive complexes by using DLS revealed a consistent trend similar to protein-dendron complexes (Chapter 2) and photo-sensitive complexes (Chapter 3). The hydrodynamic diameter of the Chy-CEG-SS-G1/G2/G3 complexes was found to be 12, 15, and 17 nm, respectively, indicating that the size of redox-sensitive protein-dendron complexes increases with an increase in the size of the dendron (Figure 4.4a). The SEC results also manifest the same behavior that the size of the protein complexes heavily depends on the size of the dendron (16, 15, 14 mL for Chy-CEG-SS-G1/G2/G3, respectively) (Figure 4.4a-b).

Both DLS and SEC studies revealed that the incorporation of disulfide moiety does not affect the self-assembly behavior of protein dendron bioconjugates. In addition, protein complexes exhibit narrow and monomodal distribution indicative of their nearly monodisperse character. The DLS and SEC profiles of protein-dendron conjugates mostly resemble self-assembly profiles of natural and synthetic protein cages rather than protein-polymer conjugates.

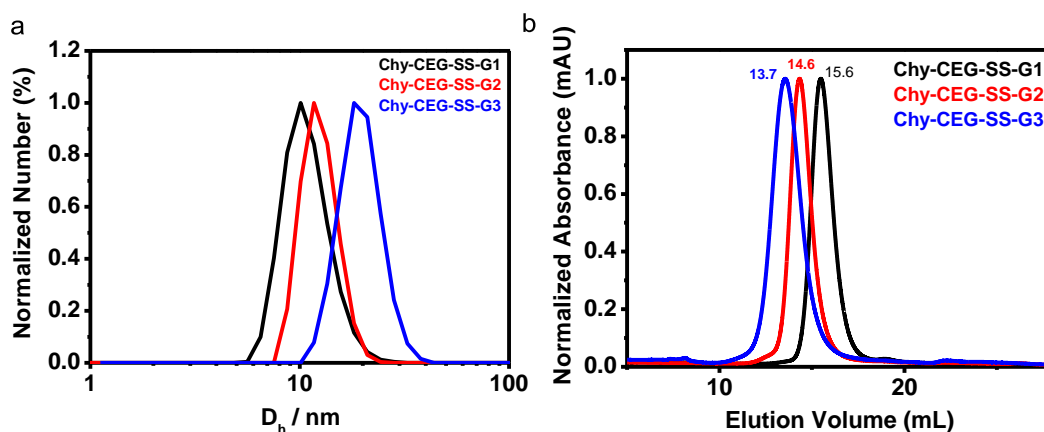


Figure 4.4 | Self-assembly of redox-sensitive protein-dendron bioconjugates. (a) The hydrodynamic diameter of Chy-CEG-SS-G1/G2/G3 complexes is determined using DLS. (b) SEC of Chy-CEG-SS-G1/G2/G3 complexes to determine their elution volume.

4.2.6. Dis-Assembly Studies of Redox-Sensitive Protein-Dendron Complexes

Having established the self-assembling behavior of bioconjugates, we wanted to test whether the custom-designed protein nanoassemblies can be programmed to disassemble.

Our hypothesis is upon treatment with DTT, the disulfide functionality which connects the hydrophobic dendron to the hydrophilic linker and protein would cleave that would convert facially amphiphilic protein into hydrophilic protein, the loss of attractive hydrophobic interaction then would lead to disassembly of protein nanoassemblies (Figure 4.5).¹⁴

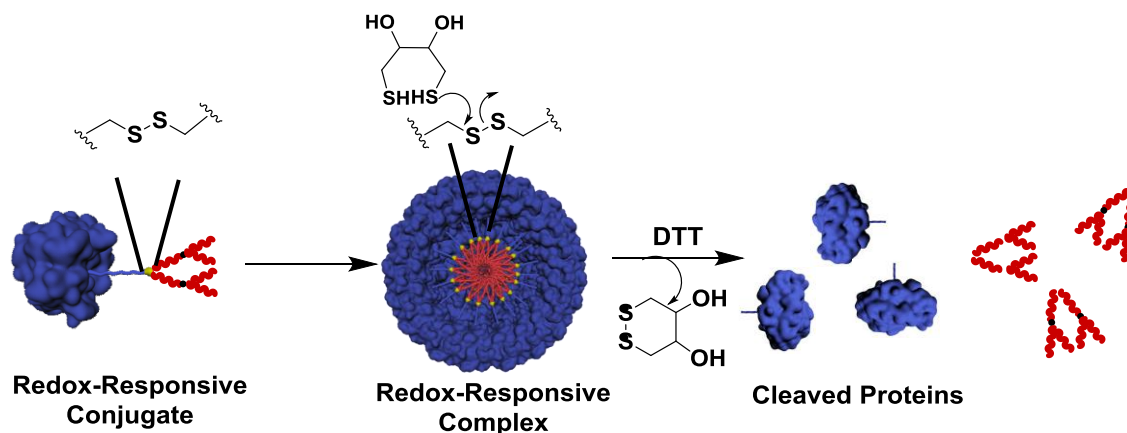


Figure 4.5 | Schematic representation of self-assembly and disassembly of redox-sensitive complexes.

In order to test this hypothesis, we performed dis-assembly studies of Chy-CEG-SS-G1 in 50 mM sodium phosphate pH 7.4. The Chy-CEG-SS-G1 bioconjugate was treated with the different equivalent of DTT (10, 20, and 30 eq) for 12 hours. The DTT treated samples were then subjected to SEC in order to investigate the disassembly profile. As expected, before DTT treatment, the protein complex was intact and therefore eluted at earlier (16 mL) compared to native monomeric protein (19 mL). However, the SEC of DTT incubated samples revealed that the entire complex was disassembled to its consecutive monomers with 10 eq of DTT (Figure 4.6a). This experiment reveals that the 10 eq. of DTT is sufficient to completely disassemble the protein complex.

Next, in order to check whether the incubation period of DTT has any effect on disassembly behavior, the protein Chy-CEG-SS-G1 complex was incubated with 10 eq of DTT for the different incubation period. The SEC results reveal that 60 mins of incubation are sufficient to disassemble the entire complex as evident from their disappearance of peak at 16 mL (Figure 4.6b). After establishing the time and the amount of DTT required to achieve the complete dis-assembly, we sought to investigate the dis-assembly behavior of the other two redox responsive protein complexes, i.e., Chy-CEG-SS-G2/G3. We were

delighted to notice that both Chy-CEG-G2/G3 protein complexes also could be disassembled into respective monomeric proteins completely upon treatment with 10 eq. of DTT for about 60 mins. Based on the above experiments, it is evident that the treatment of 10 eq of DTT for 60 mins is sufficient to disassemble the Chy-CEG-SS-G1/G2G3 complexes.

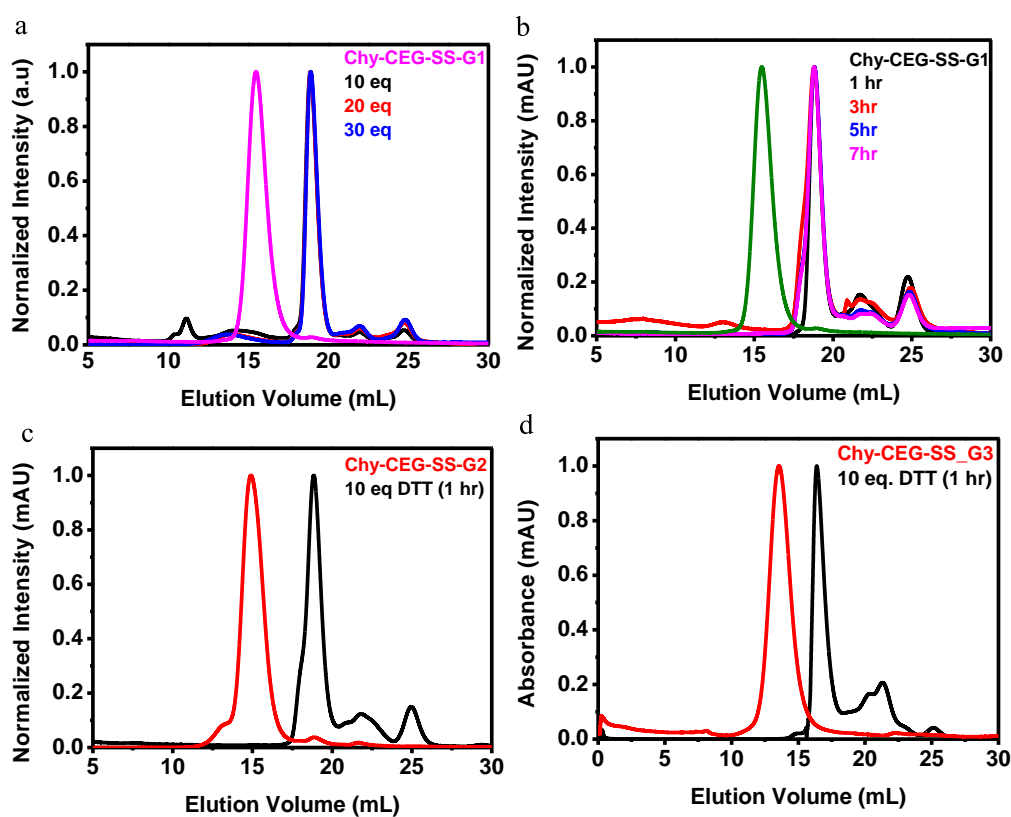


Figure 4.6 | Dis-assembly profile of redox-sensitive protein-dendron complexes in 50 mM sodium phosphate pH 7.4 using SEC chromatograms. (a), (b) Eq-dependent and time-dependent dis-assembly studies of DTT-treated Chy-CEG-SS-G1 complex, respectively. (c) SEC chromatograms of Chy-CEG-SS-G2(2 mg/mL) (d) Chy-CEG-SS-G3 (1.2 mg/mL), W/O DTT-treatment (red) and after DTT-treatment (black), respectively.

These rationally designed redox-responsive protein-dendron complexes have several outstanding advantages compared to the protein-polymer complexes. This is due to the fact that these intrinsic responsive redox-sensitive protein-dendron materials are highly monodisperse. Besides, the system offers the advantage of installing various therapeutic agents to the interior and exterior domains of complexes in order to achieve the specific delivery of cargo. Apart from DTT, the disassembly of these complexes could also be

programmed by using glutathione (GSH). The glutathione is a naturally active reducing molecule composed of glutamate, cysteine, and glycine. This molecule, similar to DTT, selectively breaks the disulfide bond through a thiol-disulfide exchange reaction. The concentration of GSH in the extracellular medium is lower (micromolar). While, the GSH concentration in the cytosolic medium is higher (mM), and enough to cleave the disulfide bond.¹⁵ Therefore, by utilizing this redox-gradient, these intrinsic stimuli-responsive protein-dendron complexes could be used for the specific cytosolic delivery of cargo.

4.3. Conclusion

In summary, we have successfully synthesized monodisperse redox-responsive protein-dendron bioconjugates by using MAPLab technology. These designed conjugates self-exhibits generation-dependent self-assembly properties. The treatment of redox-responsive protein-dendron complexes with DTT led to cleavage of a disulfide bond, which resulted in a separation of the hydrophobic domain from the rest of the protein conjugate (hydrophilic globular domain) and that triggered the dis-assembly of a protein complex into constitutive monomers. These redox-sensitive protein-dendron complexes can be employed for the targeted drug delivery by utilizing the different redox gradient inside and outside of the cell.

4.4. Experimental Methods

4.4.1. Synthesis and Characterization of Redox-Sensitive Protein-Dendron Bioconjugates.

4.4.1.1. Protein Conjugation and Purification of Protein-Dendron Bioconjugates

The bioconjugation reaction of chymotrypsin was performed using the procedure reported by our group. Briefly, 10X CMC of triton X-100 was used to solubilize redox-sensitive macromolecular AABPs (1 or 2 equivalent) and then 480 μ L of 50 mM sodium phosphate pH 7.4 was added to make a final solution to 500 μ L. The homogeneous solution

was then treated with 500 μL of 200 μM of protein solution. This reaction mixture was then allowed to react for 24 h on rotor spin at 20 rpm at 25 $^{\circ}\text{C}$. This reaction mixture was purified by using two-step purification, *i.e.*, IEX, and SEC as mentioned in chapter 2.

4.4.1.4. Monitoring of Modification

Molecular weights of photo-sensitive protein-dendron conjugates were analyzed by using MALDI-ToF spectrometry. The matrix was prepared by the protocol mentioned by our previously (Chapter 2). In brief, 2 μL of the reaction mixture or 100 μM purified protein conjugate was mixed with 2 μL of 2% TFA and 2 μL of matrix mixture, vortexed and spotted on MALDI-ToF MS plate. The plate was then loaded and fired to get accurate molecular weight both in +1 and +2 states. 100 μM protein concentration was found to be optimum for MALDI-ToF MS analysis.

4.4.2. Self-Assembly

4.4.2.1. Dynamic Light Scattering

The hydrodynamic diameter of redox-sensitive protein-dendron assemblies was measured using DLS (Zetasizer Nano 2590, Malvern, UK). Samples (5 mg/mL) were prepared in 50 mM sodium phosphate pH 7.4. 1 mL of sample was taken in disposable polystyrene cells, and then the mean size of the complexes was measured at 90 $^{\circ}$ scattering angle.

4.4.2.2. Size Exclusion Chromatography

The elution volume of redox-sensitive protein-dendron assemblies in SEC was determined using Akta Pure and Superose-200 10/300 GE Healthcare column. In order to do that, IEX fractions of bioconjugates were first subjected to the SEC in order to separate native protein from bioconjugates. Then, the 500 μL of the center fraction of the complex peak was again subjected to SEC by using the pre-equilibrated column mentioned above in 50 mM sodium phosphate pH 7.4, 200 mM NaCl.

4.4.3. Programmed Dis-Assembly of a Redox-Sensitive Supramolecular Protein Complex

The redox-responsive protein complex dissolved in 50 mM sodium phosphate pH 7.4 was treated with the different equivalent of aq. DTT for 12 hours. Then, 500 μ L of this was then subjected to SEC by using Superose column (pre-equilibrated sodium phosphate pH 7.4, 200 mM NaCl) and sodium phosphate pH 7.4, 200 mM NaCl as an elution buffer. By using the same protocol, the complex was incubated with 10 eq of DTT for different time points. Then the sample was subjected to the SEC using the same column and buffer mentioned above.

4.4.4.1. General

All reagents were obtained commercially unless and otherwise stated. Reactions were performed in an oven-dried round bottom flask (RBF) and under nitrogen atmosphere. Air and moisture sensitive solvents were transferred *via* syringe. Reactions were monitored by thin-layer chromatography (TLC), and the developed chromatogram was visualized by ultraviolet (UV) lamp or by phosphomolybdic acid (PMA) staining. Product purification was accomplished by 100-200 mesh size silica gel column chromatography.

All the compounds were characterized by ^1H , ^{13}C , and ^{19}F (in case of fluorinated compounds) nuclear magnetic resonance (NMR) using Bruker or Jeol 400 MHz. ^1H and ^{19}F were recorded at an operating frequency of 400 MHz and 100 MHz for ^{13}C using, using TMS as an internal standard. All the ^{13}C Chemical shifts were mentioned in parts per million (PPM) and measured relative to residual CHCl_3 , CH_3OH , or CH_3CN in their deuterated solvent. Coupling constants were reported in Hertz (Hz). Multiplicities were explained as s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, quint = quintet. Mass spectra were obtained with either the MALDI-TOF MS or HRMS. Room temperature varied between 21-35 $^\circ\text{C}$.

4.4.4.2. Synthesis of Redox-Sensitive Macromolecular AABP

4.4.4.3. General Procedure for the Synthesis of Acid - Procedure A

The ester (1 eq), sodium hydroxide (NaOH) (4 eq) were dissolved in ethanol. The mixture was then refluxed for 4 hours. Upon completion, the reaction was quenched with dropwise addition of water and acidified with conc. HCl. The obtained precipitate was filtered and washed with ethanol for two times. The combined organic layer dried over Na₂SO₄ and concentrated under reduced pressure to get crude, which was utilized for the next reaction without further purification.

4.4.4.4. General Procedure for the Synthesis of Alcohol- Procedure B

In an oven-dried RBF above-obtained acid (1 eq), 2,2'-disulfanediylbis(ethan-1-ol) (2 eq) and DMAP (0.5 eq) were taken and dissolved in DCM under stirring. Then, the solution of EDC in DCM (2 eq) was added slowly to the above mixture and allowed to stir for 12 hours. Upon completion, the reaction was quenched with water, and the resulting content was extracted in DCM thrice. Combined organic layers was dried over Na₂SO₄ and concentrated under reduced pressure to get the crude product, which was purified using silica gel column chromatography.

4.4.4.5. General Procedure for the Synthesis of Activated Ester- Procedure C

In an oven-dried RBF, above-obtained alcohol (1 eq) and N, N'-DSC (5 eq) were dissolved in ACN under stirring. Then, Et₃N (5 eq) was then added slowly at RT and allowed to react for 12 h. Upon completion of the reaction, ACN and Et₃N were evaporated under vacuum. The obtained residue was directly purified using silica gel column chromatography using ethyl acetate/hexane as eluent.

4.4.4.6. General Procedure for the Synthesis of Diphosphonate Ester - Procedure D

In an oven-dried RBF, the above obtained activates ester (1.1 eq) and amine-terminated cetylene glycol (1 eq) were dissolved in DMF under stirring. Then, Et₃N (1.1 eq) was added slowly to the reaction mixture and stirred at RT for 12 h. Upon completion of the reaction, DMF and Et₃N were evaporated under vacuum. To the obtained residue, water was added and extracted thrice with DCM. The combined organic layer was

dried over Na_2SO_4 and concentrated under vacuum to get the crude product, which was purified using silica gel column chromatography using MeOH / DCM as eluent.

4.4.4.7. General Procedure for the Synthesis of Monophosphonate Ester - Procedure E

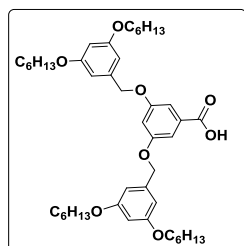
In an oven-dried RBF, the above-obtained diphosphonate ester (1 eq) was taken, and LiBr (20 eq) was added. To the mixture, DMF was added and heated at $95\text{ }^\circ\text{C}$ for 20 h. Upon completion, water was added and extracted thrice with ethyl acetate. The water layer was collected, and 2N HCl was added and stirred for another 30 minutes. The mixture was then extracted thrice with ethyl acetate. The combined organic layer was dried over Na_2SO_4 and concentrated under vacuum to get the crude product, which was used without purification.

4.4.4.8. General Procedure for the Synthesis of Fluorophosphonate - Procedure F

To the stirring solution of the above obtained monophosphonate ester (1 eq) in DCM, DAST (4 eq) was added dropwise at $-78\text{ }^\circ\text{C}$ and allowed to react for 15 minutes. Upon completion of the reaction, excess of DAST and DCM were evaporated under vacuum. To the obtained residue, water was added and stirred for 2 more minutes to quench any residual DAST. The reaction mixture was then extracted thrice with DCM. The combined organic layer was dried over Na_2SO_4 and concentrated under vacuum to get the crude product, which was used without purification.

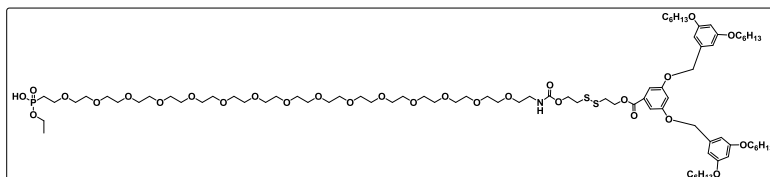
4.4.4.9. Synthesis of Photo-Sensitive Macromolecular AABP and Their Intermediates

Synthesis of compound 6a



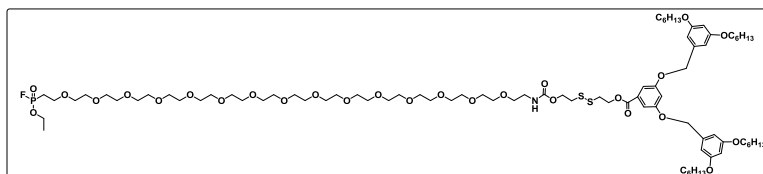
mentioned in chapter 3) (0.6 g, 0.7 mmol) and, Et₃N (0.06 g, 0.6 mmol). The obtained product was pale yellow liquid (0.78 g, 0.4 mmol, 65%), *R_f*=0.40 in 5% MeOH / DCM. ¹H NMR (400MHz, CDCl₃) 7.10 (dd, *J*=8.9, 2.8 Hz, 1H), 6.5. (d, m, 9H), 5.46 (s, 2H), 5.23 (s, 2H), 4.24 (m, 4H), 4.07 (t, *J*=6.4Hz, 8H), 3.77-3.52 (m, 57H), 2.19-2.05 (m, 2H), 1.85-1.60 (m, 12H), 1.44-1.29 (m, 48H), 0.82 (t, *J*=6.8Hz, 9H).

Synthesis of compound 6e



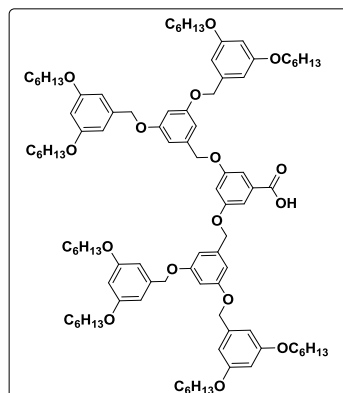
The compound **6e** was prepared by general procedure E, starting from compound **6d** (0.5 g, 0.3 mmol) and, LiBr (0.3 g, 3 mmol) was added. The obtained product was carried to the next step without further purification.

Synthesis of compound 6f



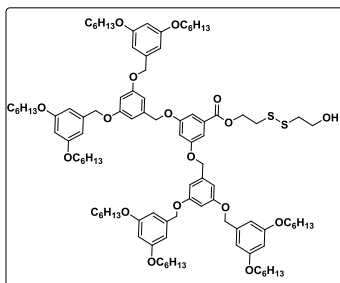
The compound **6f** was prepared by general procedure F, starting from compound **6e** (0.1 g, 0.050 mmol), DAST (0.040 g, 0.2 mmol). The product was utilized for conjugation without further purification. ¹⁹F NMR (400MHz, CDCl₃): δ_F -59.91, -62.74.

Synthesis of compound 7a



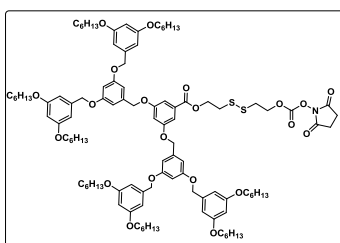
The compound **7a** was prepared by general procedure A, starting from compound **4a** (0.018 g, 0.10 mmol), NaOH (0.17 g, 0.10 mmol). The product obtained as pale solid (0.10 g, 0.06 mmol, 65%), $R_f=0.52$ in 20% ethyl acetate / hexane. $^1\text{H NMR}$ (400MHz, CDCl_3): δ_{H} 7.33 (d, $J=8.9$, 2H), 6.71-6.43 (m, 28H), 5.00-4.94 (m, 20H), 3.97 (t, $J=6.4\text{Hz}$, 26H), 1.80-1.75 (m, 28H), 1.54-1.30 (m, 90H), 0.93 (t, $J=7.2\text{Hz}$, 44H).

Synthesis of compound 7b



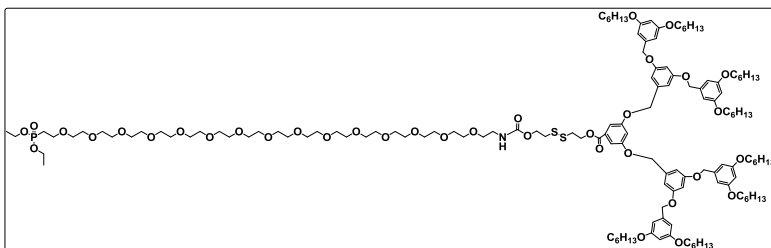
The compound **7b** was prepared by general procedure B, starting from compound **7a** (1 g, 0.6 mmol), 2,2'-disulfanediylobis(ethan-1-ol) (0.2 g, 1.2 mmol). The product obtained as a pale yellow solid (0.4 g, 0.2 mmol, 35%), $R_f=0.20$ in 20% ethyl acetate / hexane. $^1\text{H NMR}$ (400MHz, CDCl_3): δ_{H} 7.29 (d, $J = 2.8$ Hz, 2H), 6.62-6.38 (m, 21H), 5.61 (s, 2H), 5.03-4.89 (m, 12H), 3.94 (t, $J=6.4\text{Hz}$, 14H), 3.06-2.91 (m, 4H), 1.85-1.69 (m, 16H), 1.47-1.22 (m, 53H), 0.90 (t, $J=7.2\text{Hz}$, 24H).

Synthesis of compound 7c



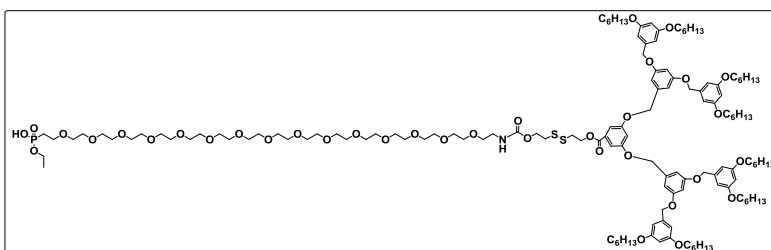
The compound **7c** was prepared by general procedure C, starting from compound **7b** (0.4 g, 0.2 mmol), N,N' -DSC (0.3 g, 1 mmol) and Et_3N (0.1 g, 1 mmol). The product obtained as a pale yellow solid (0.3 g, 0.1 mmol, 66%), $R_f=0.18$ in 25% ethyl acetate / hexane.

Synthesis of compound 7d



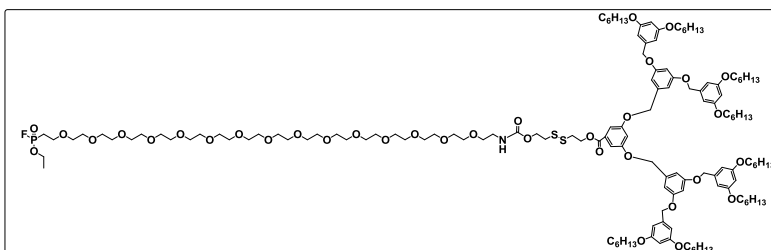
The compound **7d** was prepared by general procedure D, starting from compound **7c** (0.3 g, 0.1 mmol), compound **6** (Synthesis and characterization data for compound **6** is mentioned in chapter 3) (0.09 g, 0.1 mmol) and, Et₃N (0.01 g, 1 mmol). The product obtained was pale yellow liquid (0.3 g, 0.1 mmol, 70%), *R*_f=0.40 in 5% MeOH / DCM. ¹H NMR (400MHz, CDCl₃): δ_H 7.17 (d, 2.8 Hz, 1H), 6.64-6.40 (m, 21H), 5.0-4.82 (m, 12H), 4.20-4.10 (m, 20H), 3.70-3.45 (m, 66H), 2.19-2.06 (m, 2H), 1.86-1.66 (m, 18H), 1.50-1.18 (m, 70H), 0.89 (t, *J*=7.2Hz, 30H). ¹³C NMR (100MHz, CDCl₃):

Synthesis of compound 7e



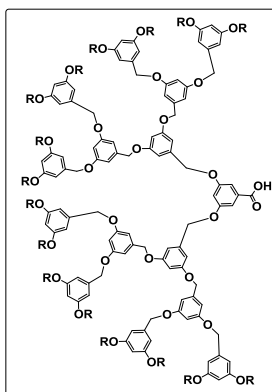
The compound **7e** was prepared by general procedure E, starting from compound **7d** (0.140 g, 0.02 mmol) and, LiBr (0.080 g, 0.5 mmol). The obtained product was carried to the next step without further purification.

Synthesis of compound 7f



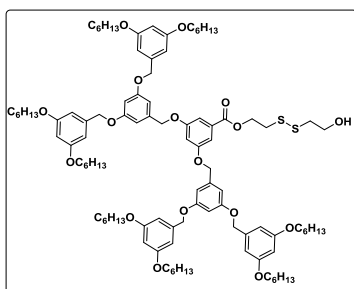
The compound **7f** was prepared by general procedure F, starting from compound **7e** (0.08 g, 0.01 mmol), DAST (0.02 g, 0.05 mmol). The obtained product was utilized for conjugation without further purification. ¹⁹F NMR (400MHz, CDCl₃): δ_F -59.91, -62.74.

Synthesis of compound 8a



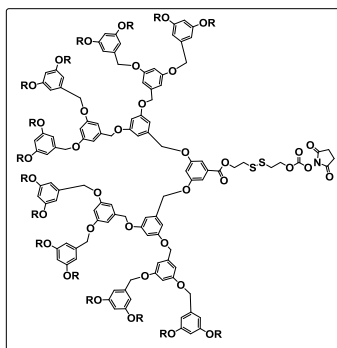
The compound **8a** was prepared by general procedure A, starting from compound **6a** (3 g, 0.9 mmol), and NaOH (0.15 g, 3.6 mmol). The product obtained as pale solid was carried to the next step without purification. $^1\text{H NMR}$ (400MHz, CDCl_3): δ_{H} 7.42 (d, $J = 2.4$ Hz, 2H), 6.73-6.45 (m, 45H), 5.05-4.90 (m, 28H), 3.96 (t, $J = 6.8$ Hz, 34H), 1.83-1.71 (m, 34), 1.49-1.22 (m, 134H), 0.94 (t, $J = 7.2$ Hz, 54H).

Synthesis of compound 8b



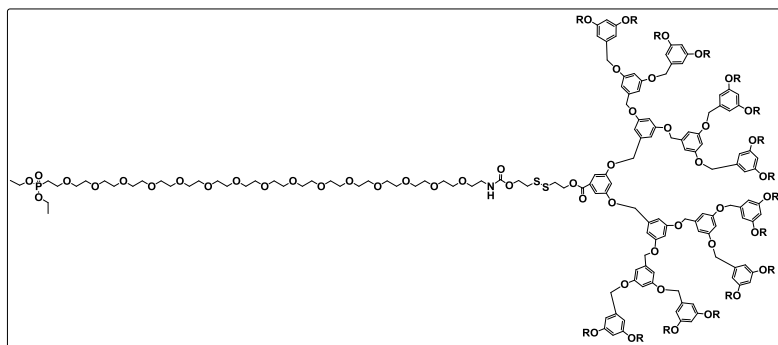
The compound **8b** was prepared by general procedure B, starting from compound **8a** (2 g, 0.6 mmol), 2,2'-disulfanediybis(ethan-1-ol) (0.2 mg, 1.2 mmol). The product obtained as a pale yellow solid (0.7 g, 0.2 mmol, 36%), $R_f=0.20$ in 20% ethyl acetate / hexane. $^1\text{H NMR}$ (400MHz, CDCl_3): δ_{H} 7.30 (d, $J = 2.4$ Hz, 2H), 6.69-6.37 (m, 45H), 5.01-4.94 (m, 28H), 3.92 (t, $J = 6.8$ Hz, 34H), 3.04-2.85 (m, 4H), 1.82-0.86 (m, 270H).

Synthesis of compound 8c



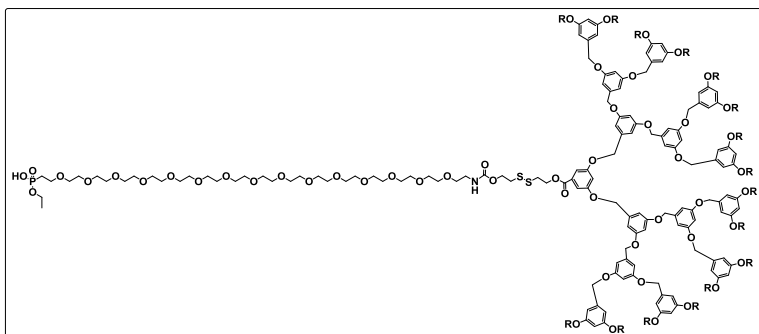
To the mixture of compound **7b** (0.7 g, 0.2 mmol) and pyridine (0.1 g, 1 mmol) in DCM, a solution of 4-nitrophenyl chloroformate (0.2 g, 1 mmol) in DCM was added at 0 °C. The mixture was allowed to react for 3 hours at RT. Then, the solvent was evaporated under reduced pressure to get the crude product, which was then purified using silica gel column chromatography (0.35 g, 0.1 mmol, 55%), $R_f=0.20$ in 20% ethyl acetate/hexane.

Synthesis of compound **8d**



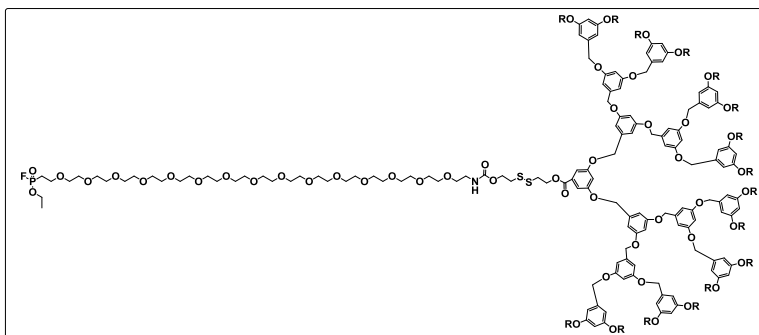
The compound **8d** was prepared by general procedure D, starting from compound **8c** (0.4 g, 0.1 mmol), compound **6** (Synthesis and characterization data for compound **6** is mentioned in chapter 3) (0.09 g, 0.1 mmol) and, Et_3N (0.03 g, 0.3 mmol). The product obtained was pale yellow liquid (0.2 g, 0.07 mmol, 65%), $^1\text{H NMR}$ (400MHz, CDCl_3): δ_{H} 7.3 (d, $J = 2.4\text{Hz}$, 2H), 6.66-6.38 (m, 45H), 5.30 (3, 14H), 4.95-4.93 (m, 26H), 4.20-4.15 (m, 4H), 4.02-3.93 (m, 30H), 3.91-3.76 (m, 70H), 2.19-2.09 (m, 2H), 1.78-1.71 (m, 34H), 1.51-1.26 (m, 144H), 0.89 (t, $J=6.4\text{Hz}$, 58H).

Synthesis of compound **8e**



The compound **8e** was prepared by general procedure E, starting from compound **8d** (0.6 g, 0.04 mmol) and LiBr (0.12 g, 0.5 mmol). The obtained product was carried to the next step without further purification.

Synthesis of compound **8f**



The compound **8f** was prepared by general procedure H, starting from compound **8e** (0.1 g, 0.04 mmol), DAST (0.005 g, 0.05 mmol). The obtained product was utilized for conjugation without further purification. ^{19}F NMR (400MHz, CDCl_3): δ_{F} -59.91, -62.74.

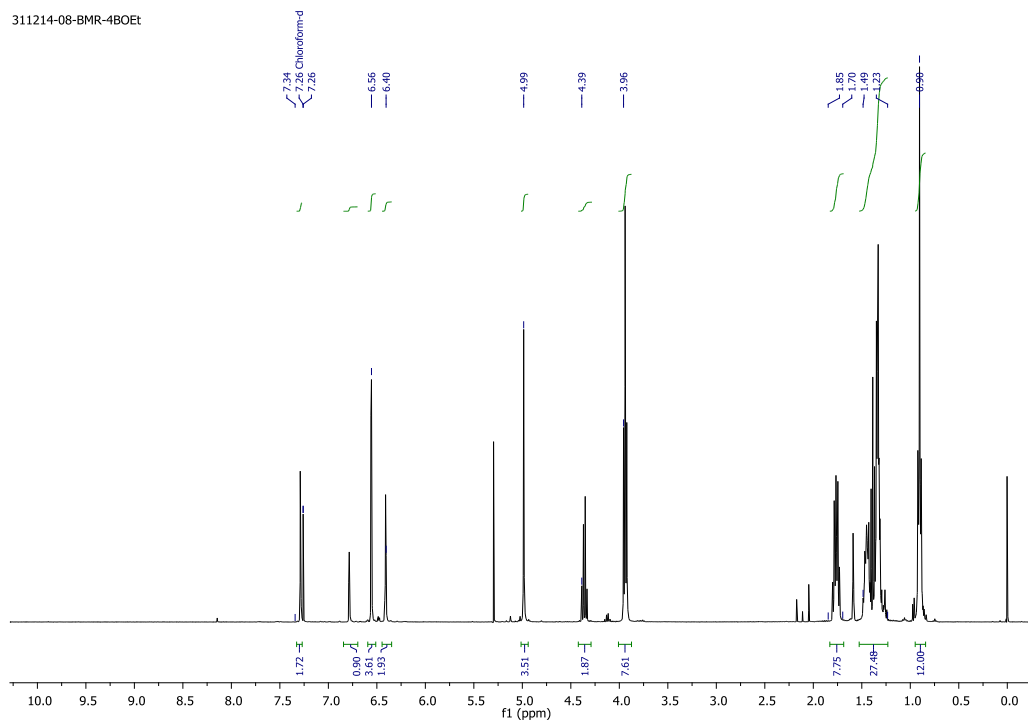
4.5 References

- (a) Lee, E. S.; Na, K.; Bae, Y. H. *J. cont. release* **2005**, *103*, 405-418. (b) Sawant, R. M.; Hurley, J. P.; Salmaso, S.; Kale, A.; Tolcheva, E.; Levchenko, T. S.; Torchilin, V. P. *Bioconjugate chem.* **2006**, *17*, 943-949. (c) Gillies, E. R.; Jonsson, T. B.; Fréchet, J. M. *J. Am. Chem. Soc.* **2004**, *126*, 11936-11943.
- (a) Li, Y.; Lokitz, B. S.; Armes, S. P.; McCormick, C. L. *Macromolecules*, **2006**, *39*, 2726-2728. (b) Ryu, J. H.; Roy, R.; Ventura, J.; Thayumanavan, S. *Langmuir* **2010** *26*, 7086-7092. (c) Cho, H., Bae, J., Garripelli, V. K.; Anderson, J. M.; Jun, H. W.; Jo, S. *Chem. Comm*, **2012**, *48*, 6043-6045.

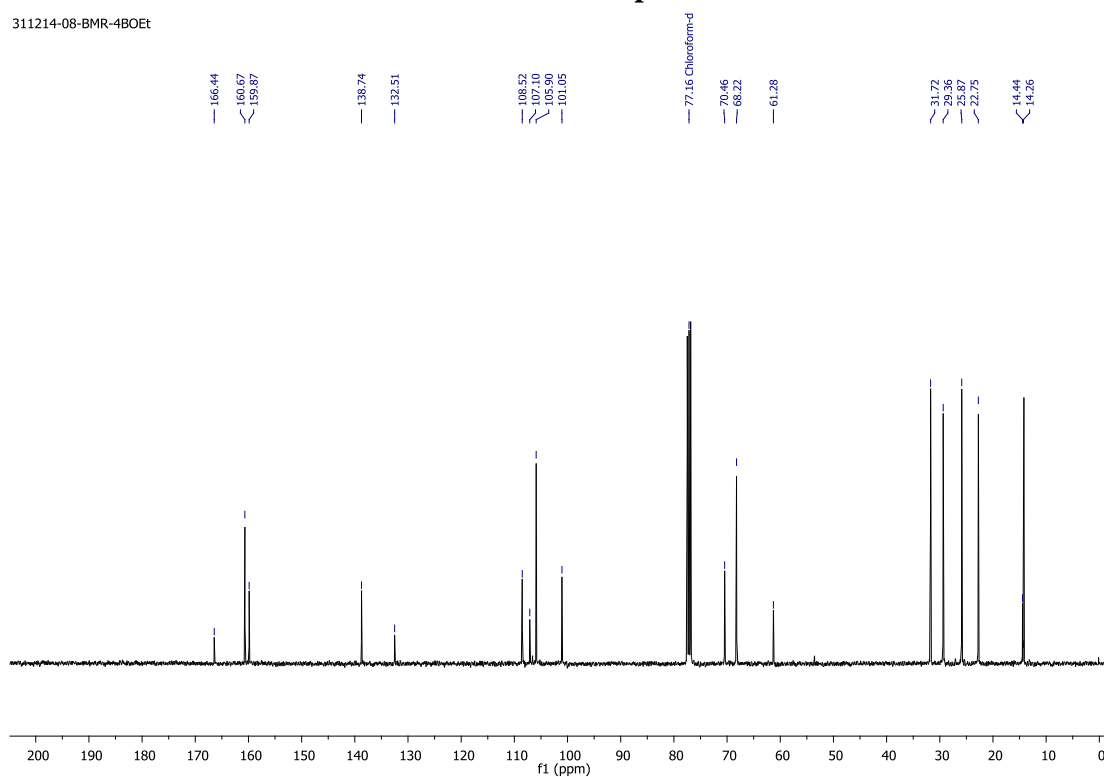
3. (a) Wei, H.; Cheng, S. X.; Zhang, X. Z.; Zhuo, R. X. *Progress in Polymer Science* **2009**, *34*, 893-910. (b) Aathimanikandan, S. V.; Savariar, E. N.; Thayumanavan, S. *J. Am. Chem. Soc.* **2005**, *127*, 14922–14929.
4. (a) Huang, Y.; Dong, R.; Zhu, X.; Yan, D. *Soft Matter*, **2014** *10*, 6121-6138. (b) Guragain, S.; Bastakoti, B. P.; Malgras, V.; Nakashima, K.; Yamauchi, Y. *Chem. Eur. J.* **2015**, *21*, 13164-13174.
5. (a) Ge, J.; Neofytou, E.; Lei, J.; Beygui, R. E.; Zare, R. N. *Small* **2012**, *8*, 3573-3578. (b) Mukherjee, B.; Karmakar, S.; Hossain, C.; Bhattacharya, S. *Protein and peptide letters*, **2014**, *21*, 1121-1128.
6. (a) Kim, W.; Brady, C.; Chaikof, E. L. *Acta biomaterialia*, **2012**, *8*, 2476-2482. (b) Phan, V. G.; Duong, H. T. T.; Thambi, T.; Nguyen, T. L.; Turabee, M. H.; Yin, Y.; Kim, S.H.; Kim, J.; Jeong, J. H.; Lee, D. S. *Biomaterials*, **2019**, *195*, 100-110. (c) Janib, S. M.; Liu, S.; Park, R.; Pastuszka, M. K.; Shi, P.; Moses, A. S.; Orosco, M. M.; Lin, Y. A.; Cui, H.; Conti, P. S.; Li, Z. *Integrative Biology*, **2013**, *5*, 183-194.
7. Basak, S.; Punetha, V. D.; Bisht, G.; Bisht, S. S.; Sahoo, N. G.; Cho, J. W. *Polymer Reviews*, 2015, *55*, 163-198.
8. Oohora, K.; Onuma, Y.; Tanaka, Y.; Onoda, A.; Hayashi, T. *Chem. Commun.* **2017**, *53*, 6879– 6882. (b) Biswas, S.; et al. *J. Am. Chem. Soc.* **2009**, *131*, 7556–7557. (c) Sendai, T.; Biswas, S.; Aida, T. *J. Am. Chem. Soc.* **2013**, *135*, 11509–11512. (d) MacKay, J.A., Chen, M., McDaniel, J.R., Liu, W., Simnick, A.J. and Chilkoti, A., 2009. *Nature materials*, *8*(12), pp.993-999.
9. (a) Chen, G.; Hoffman, A. S. *Bioconjugate Chem.* **1993**, *4*, 509-514. 22. Kulkarni, S.; Schilli, C.; Müller, A. H.; Hoffman, A. S.; Stayton, P. S. *Bioconjugate Chem.* **2004**, *15*, 747-753. 23. (b) Wan, X. J.; Liu, S. Y. *Macromol. Rapid Commun.* **2010**, *31*, 2070-2076. 24. (c) Park, W. M.; Champion, J. A. *J. Am. Chem. Soc.* **2014**, *136*, 17906-17909. (d) Hassouneh, W.; Fischer, K.; MacEwan, S. R.; Branscheid, R.; Fu, C. L.; Liu, R.; Schmidt, M.; Chilkoti, A. *Biomacromolecules* **2012**, *13*, 15981605.
10. Sandanaraj, B. S.; Bhandari, P. J.; Reddy, M. M.; Lohote, A. B.; Sahoo, B. *ChemBioChem*, **2020**, *21*, 408-416.
11. Nakashima, Y.; Ohta, S.; Wolf, A. M. *Free Radical Biol. Med.* **2017**, *108*, 300
12. (a) Stuart, M. A. C.; Huck, W. T.; Genzer, J.; Müller, M.; Ober, C.; Stamm, M.; Sukhorukov, G. B.; Szleifer, I.; Tsukruk, V. V.; Urban, M.; Winnik, F. *Nature materials* **2010**, *9*, 101-113. (b) Galaev, I. Y. and Mattiasson, B. *Trends in biotechnology*, **1999**, *17*, 335-340. (c) Mura, S., Nicolas, J.; Couvreur, P. *Nature materials* **2013**, *12*, 991-1003.

13. (a) Hawker, C. J.; Fréchet, J. M. *J. Am. Chem. Soc.* **1990**, *112*, 76387647. (b) Wooley, K. L.; Hawker, C. J.; Fréchet, J. M. *J. Am. Chem. Soc.* **1991**, *113*, 4252-4261.
14. Sandanaraj, B. S.; Reddy, M. M.; Bhandari, P. J.; Kumar, S. and Aswal, V. K. *Chem. Eur. J.* **2018**, *24*, 16085-16096.
15. Kidd, P.M. *Altern Med Rev*, **1997**, *2*, 155-176.

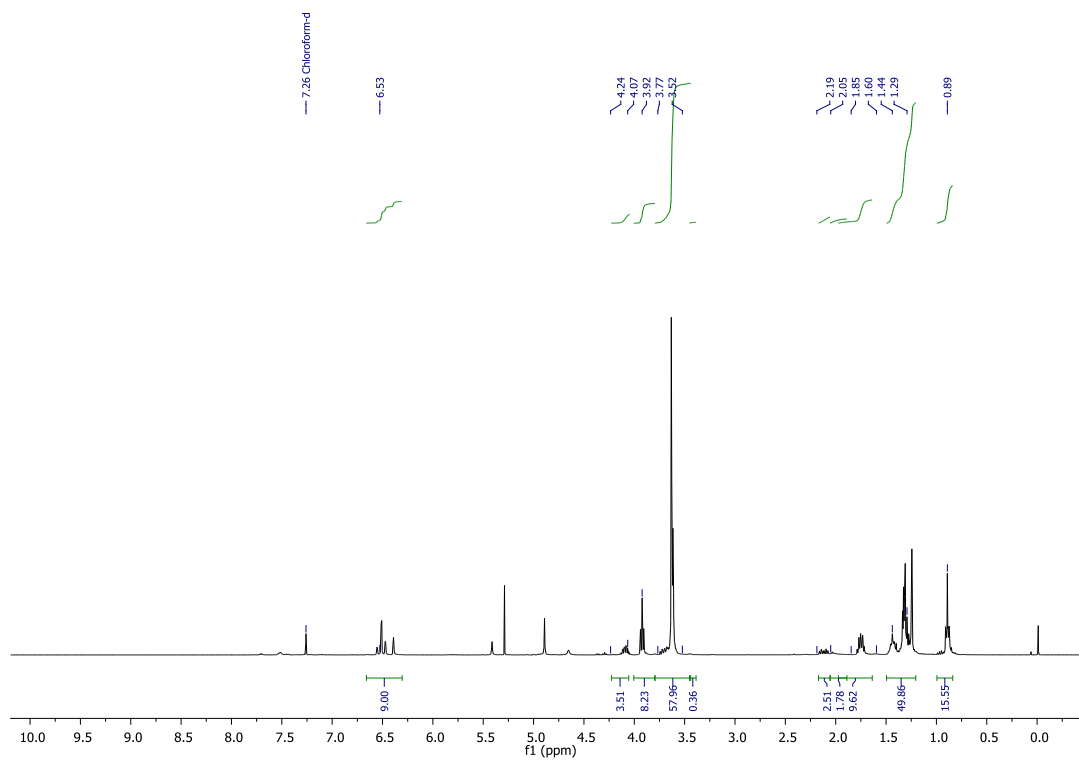
4.6 Appendix III Characterization data of synthesized compounds



¹H NMR of compound 6a

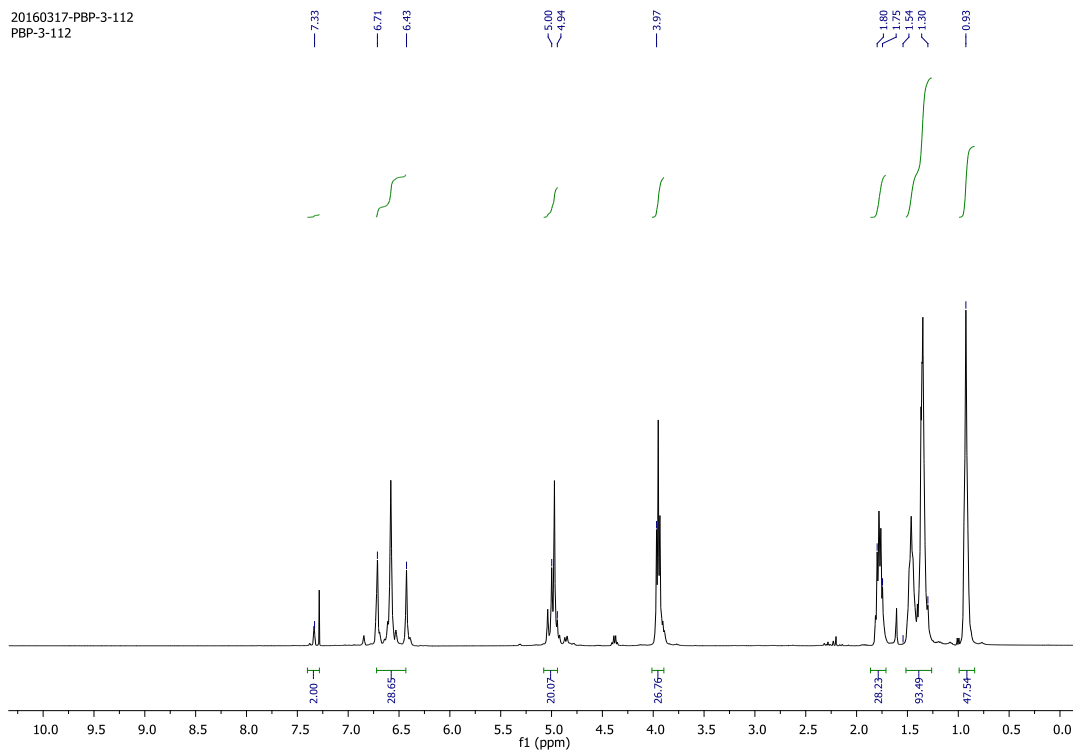


¹³C NMR of compound 6a



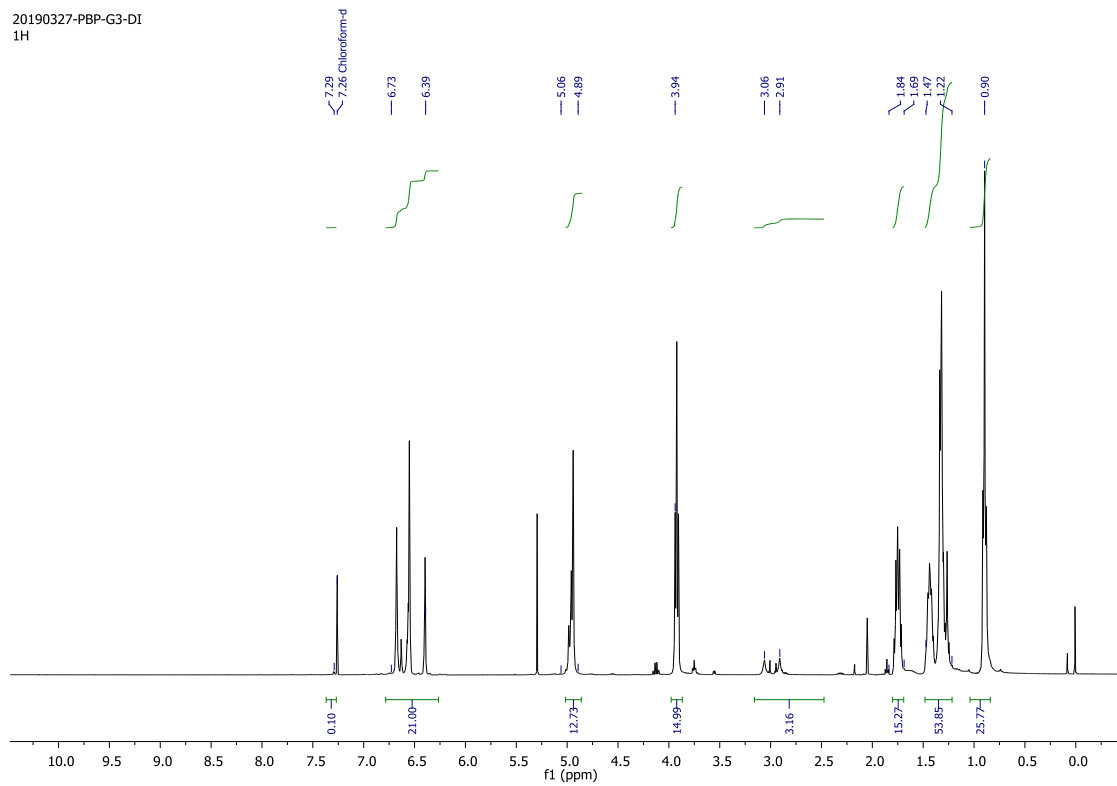
¹H NMR spectrum of compound 6d

20160317-PBP-3-112
PBP-3-112



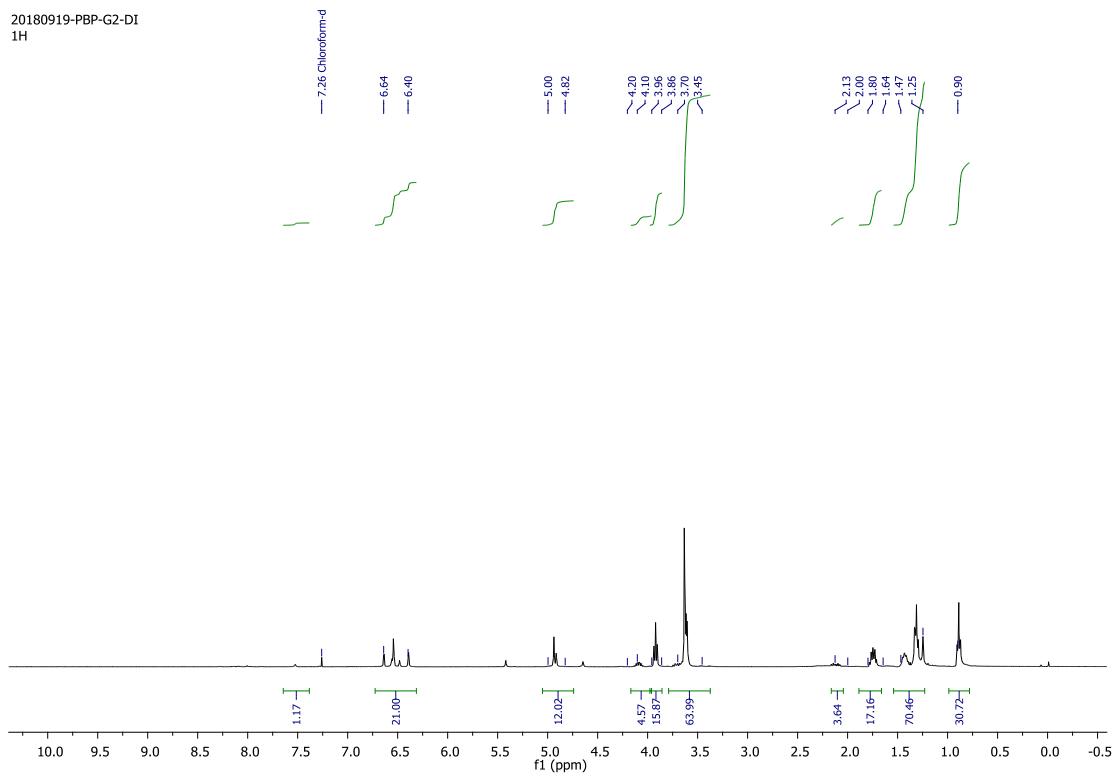
¹H NMR of compound 7a

20190327-PBP-G3-DI
1H



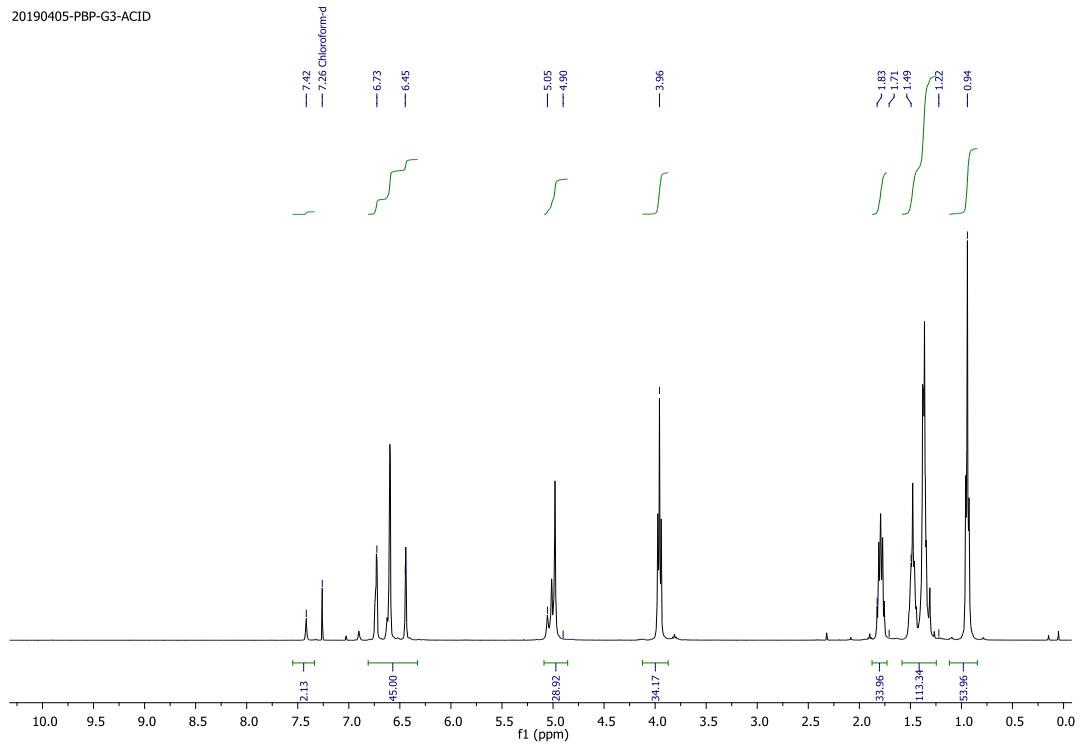
¹H NMR of compound 7b

20180919-PBP-G2-DI
1H



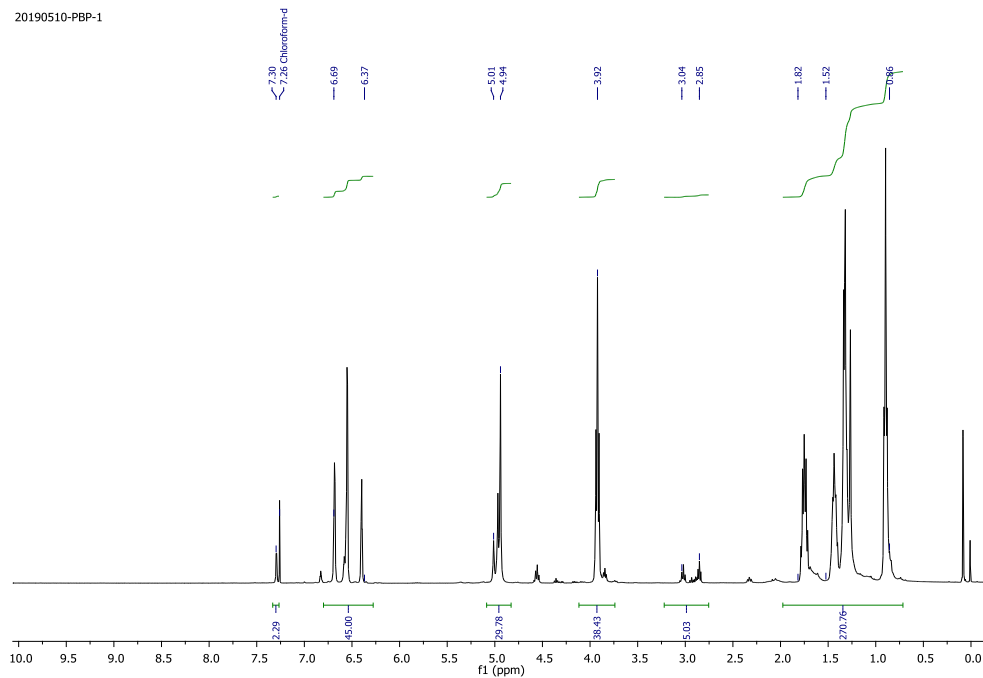
¹H NMR of compound 7d

20190405-PBP-G3-ACID

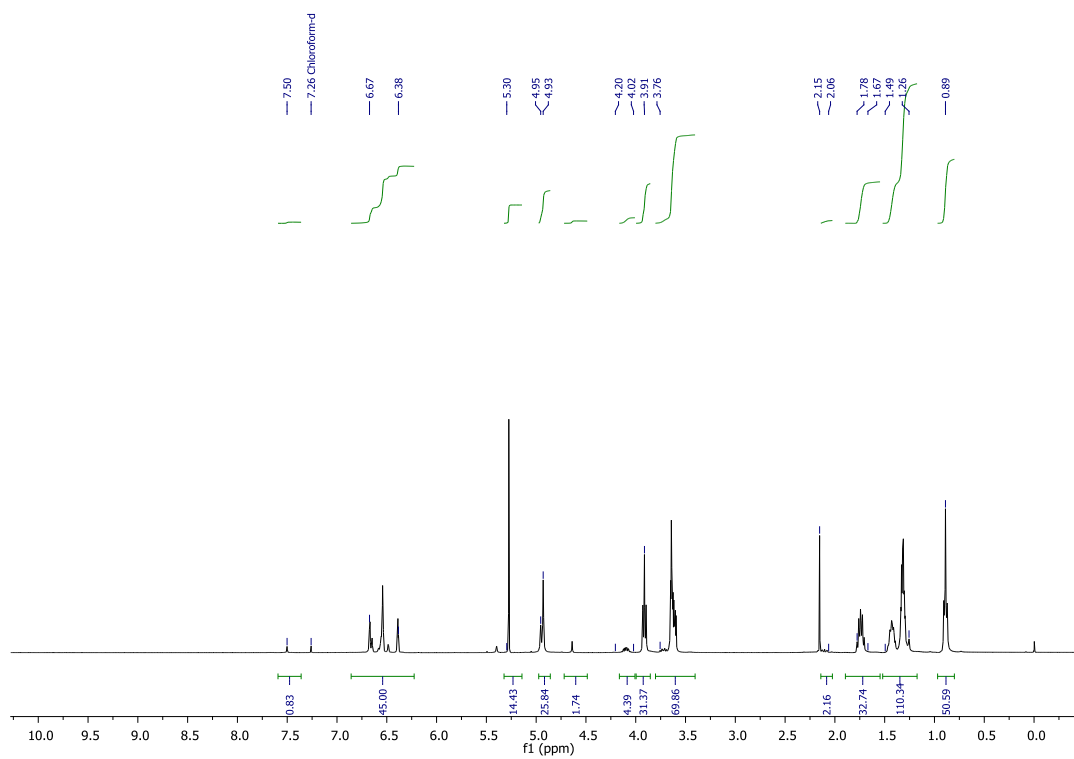


¹H NMR of compound 8a

20190510-PBP-1



¹H NMR of compound 8b



¹H NMR of compound 8d

CHAPTER 5

GENERAL CONCLUSIONS AND FUTURE PERSPECTIVE

5.1. Summary of the Thesis

The field of protein nanotechnology has encountered substantial difficulty, due to the chemical complexity, structural instability, and incomplete understanding of the rules by which proteins recognize each other to form assemblies. Despite these challenges, efforts along the multiple lines in the last two decades have substantially grown this area. In this context, my research work has focused on designing a chemical methodology to construct stimuli-responsive amphiphilic protein- dendron conjugates and understand their self-assembly and dis-assembly behavior. In this direction, we have synthesized monodisperse protein-dendron bioconjugates by conjugating a macromolecular amphiphilic activity-based probe (composed of a fluorophosphonate as a reactive group, hydrophilic linker, and hydrophobic dendrimer) to the serine proteases using a micelle-assisted protein labeling technology. The strategy offers control over hydrodynamic radius (D_h), oligomeric state, and the molecular weight of the complex by either re-engineering dendron or protein head group in the molecular design (Chapter 2).

Further, the scope of this strategy was extended to design multi-responsive protein-based supramolecular systems by incorporating photo-responsive functionality in between the hydrophilic and hydrophobic portion of bioconjugate. These designed synthons self-assemble to make supramolecular protein assemblies of defined sizes driven *via* hydrophobic interaction of hydrophobic dendrimer. These photo-sensitive supramolecular protein assemblies on exposure to UV light surprisingly led to partial disassembly. Surprisingly, the complete disassembly of the photo-sensitive protein complex is achieved by decreasing the pH of the solution or treatment with hydrazine (Chapter 3). Next. The scope of this strategy was further extended to design redox-responsive protein-dendron complexes. The design principle is the same as previous design expect redox-sensitive moiety is clamped between the hydrophilic and hydrophobic portion of bioconjugate. These complexes upon treatment with DTT lead to cleavage of a disulfide bond, which resulted into a separation of the hydrophobic domain from the rest of protein conjugate (hydrophilic globular domain), that triggered the dis-assembly of a protein complex into constitutive monomers, as the assembly is kept intact by strong hydrophobic interaction (Chapter 4).

The methodology provides an opportunity to synthesize monodisperse protein-dendron complexes which can respond to extrinsic and intrinsic stimulus. Also, the dimensions, molecular weight, and oligomeric state of these complexes can be tuned by re-engineering the molecular design. The protein-dendron system presented in this thesis provides an opportunity to functionalize interior and exterior domains of assemblies with a variety of therapeutic agents. These opportunities can be used in devising antibody or ligand decorated particles with controlled densities, which we expect to find application in the area of vaccine design, targeted drug delivery.

5.2. Limitations

Despite the advantages protein-dendron assemblies possess over protein-polymer systems, the presented in this thesis exhibits some limitations, which are discussed point-wise below.

1. The major practical disadvantage of the study carried out in this thesis is the need for multi-step organic synthesis of dendrimer and the monodisperse oligoethylene glycol molecular components. The syntheses of these molecules are both labor and purification intensive.
2. The dendrimer component of protein-dendron bioconjugates is not biocompatible and biodegradable. Thus, as such cannot be explored for biomedical applications.
3. The process for the construction of final protein-dendron bioconjugate is time-consuming.
4. This strategy offers unlimited chemical space to construct defined protein-dendron assemblies with different protein head groups. However, in the thesis, we have explored the effect of only four proteins, i.e., trypsin, chymotrypsin, proteinase K, and subtilisin on the self-assembly.
5. The methodology provides an opportunity to regulate dimensions, oligomeric state, and molecular weight protein complex. However, in order to achieve the desired assemblies, it is necessary to re-engineer the components of molecular design.

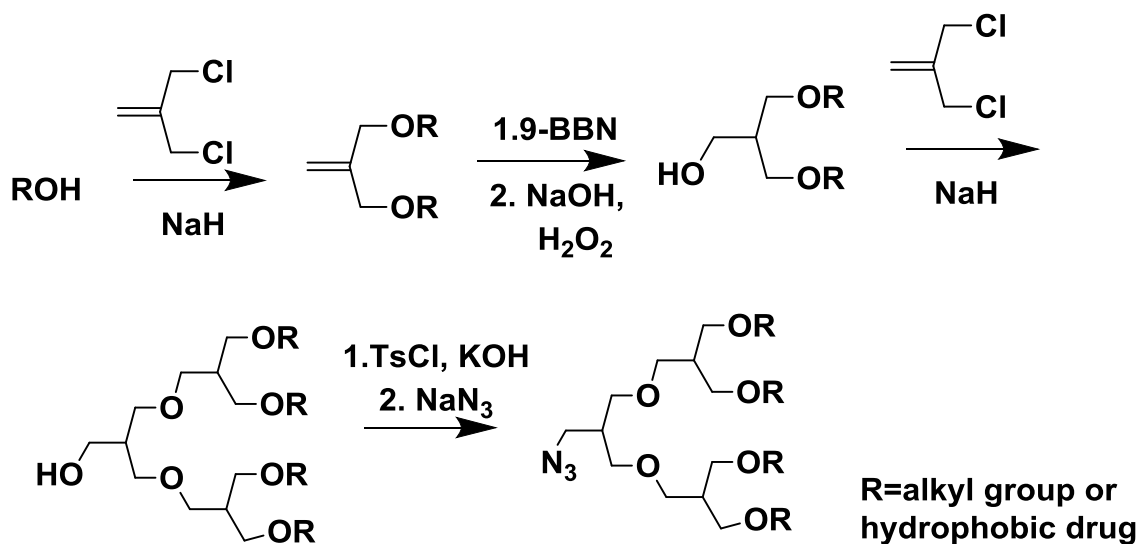
6. We have installed a photosensitive nitrobenzyl group in the molecular design of protein-dendron bioconjugates to impart stimuli-responsive behavior. However, the protein-dendron bioconjugate is associated with the phototoxicity of the UV light, and its less penetration depth into the skin is the significant limitations.
7. Although we thoroughly characterized protein-dendron assemblies in the solution state, however, we lack detailed solid-state self-assembly characterization of the same by using techniques such as TEM, AFM, and SEM.

5.3. Future Prospective

5.3.1. Synthesis of Biocompatible Protein-Dendron Bioconjugates

This thesis is mainly focused upon designing strategies for the synthesis of stimuli-responsive protein-dendron bioconjugates and understanding their self-assembly and disassembly mechanism. However, the ultimate aim of our research program is to utilize this technology for various biomedical applications such as vaccine design, targeted drug or gene delivery, and diagnostic imaging. Thus, considering our aim and the limitations mentioned above of the current scaffold, the future aspects of these projects will be geared towards minimizing the labor and purification intensive synthetic strategies to simplify the design process for the construction of protein-dendron assemblies. Besides, significant attention will be paid for the construction of biocompatible and biodegradable protein-dendron bioconjugates.

The first step towards the solution for disadvantages associated with the current scaffold can be realized by substitution of Fréchet type dendrons with biocompatible dendrimer. In this case, polyglycerol dendrimer serves as a great model. This is because the polyglycerol dendrimer is compact, well-defined, stable, and biocompatible polyether dendrimers, which can be synthesized by less labor and purification intensive strategies. As proof of the model, the synthetic strategy of the polyglycerol dendrimer is decapitated in Scheme 5.1.



Scheme 5.1 | Scheme of the synthetic of polyglycerol dendrimer

Next, the alcohol or the azide functionality can be coupled to the monodisperse oligoethylene glycol by using standard organic chemical strategies. Then full probe with the protein-reactive group will be conjugated to the serine protein of interest by using MapLab technology (Figure 5.1). The resulting protein-dendron scaffold will be purified by our in-house developed purification technique. This designed protein-dendron bioconjugate would then can undergo to form nanoassemblies driven *via* hydrophobic interactions (Figure 5.1).

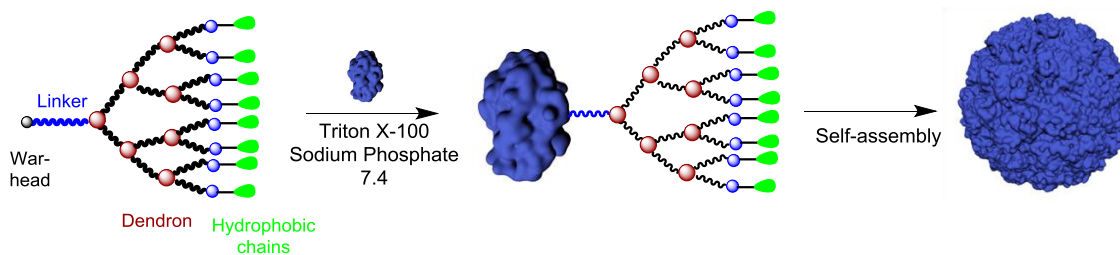


Figure 5.1 | Schematic representation of the synthesis of biocompatible protein-dendron bioconjugate

5.3.2. Synthesis of Stimuli-Responsive Biocompatible Protein-Dendron Bioconjugates

After optimizing the self-assembly properties of the new construct, we envision to utilize them for biomedical applications. In order to do that, we plan to synthesize a drug-functionalized polyglycerol dendrimer.

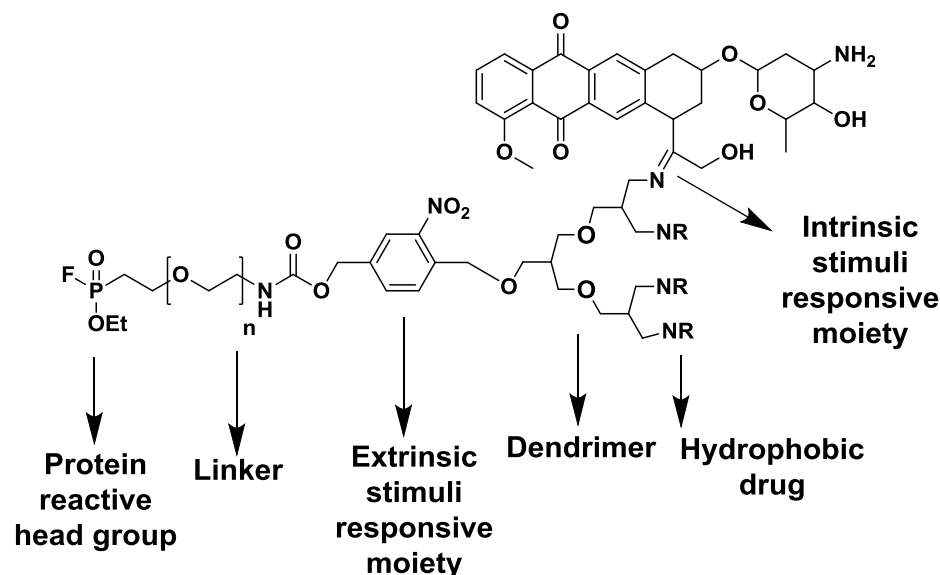


Figure 5.2 | Chemical structure of biocompatible stimuli-responsive protein-dendron bioconjugate

Here, the functionality which connects the drug molecules and hydrophobic dendrimer will be engineered in a way that they would be cleaved in response to physiological imbalance (pH, redox gradient, esterase) and release the drug. Besides, we envision to incorporate the extrinsic stimuli-responsive moiety in between linker and dendron. Thus, the assemblies would respond to the two layers of stimuli. The role of primary extrinsic stimuli would give us an excellent spatiotemporal opportunity to disassemble the drug-loaded particles. Once the first layer of protection is unmasked, then the drug-functionalized hydrophobic dendrimer will be exposed to the bulk medium. Then, the physical imbalance of tissue or cell would trigger the release of a covalently attached drug molecule.

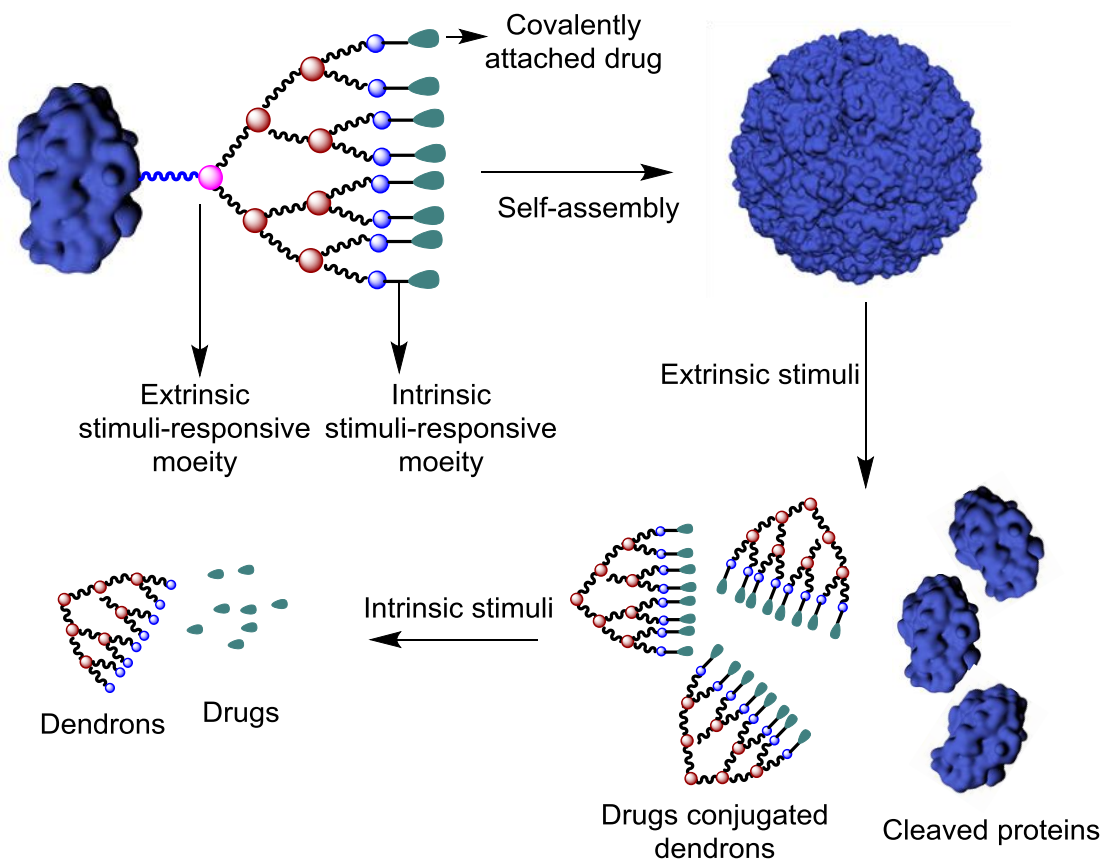


Figure 5.3 | Schematic representation of the self-assembly and dis-assembly of drug-functionalized stimuli-responsive protein-dendron bioconjugate.

Next, the above-mentioned nanoassemblies can also be made target specific by decorating these particles with target specific antibodies.

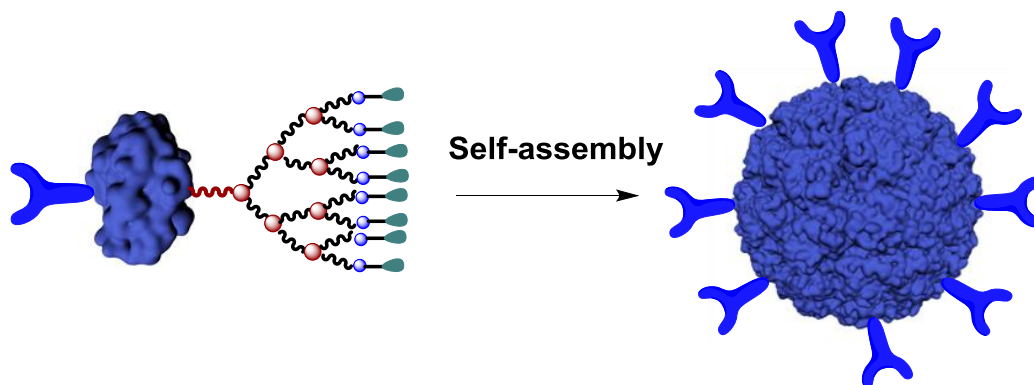


Figure 5.4 | Schematic representation of the self-assembly of drug and antibody-functionalized stimuli-responsive protein-dendron bioconjugate

Towards that goal, we have collaborated with various national and international.

- Utilizing MAPLab™ Technology for design of novel antibody-drug conjugates; Therapeutic Area: Antibody-Drug Conjugates; Partner – Major Pharmaceutical Company in the Greater Boston Area.
- Utilizing MAPLab™ Technology for design of Novel Malaria Vaccines; Therapeutic Area: Vaccine Design; Partner – Gennova Biopharmaceuticals Pvt Ltd – Pune.
- Utilizing MAPLab™ Technology for Novel Drug Delivery Vehicles; Therapeutic Area: Drug Delivery; Partner – Sai Life Sciences Pvt Ltd.
- Utilizing MAPLab™ Technology for design of Novel Influenza Vaccine; Therapeutic Area: Vaccine Design; Partner – MynVax – Bangalore

These collaborations will be mainly focused upon exploring micelle-assisted protein labeling (MAPLab) technology for target-specific drug delivery and vaccine design. The

LIST OF PUBLICATIONS AND PATENTS

1. B.S. Sandanaraj, **P.J. Bhandari**, M.M. Reddy, A. B. Lohote, B. Sahoo. Design, Synthesis, and Self-assembly Studies of Suite of Monodisperse Facially Amphiphilic Protein-Dendron Conjugates. *ChemBioChem*, **2020**, *21*, 408-416.
2. B.S. Sandanaraj, M.M. Reddy*, **P.J. Bhandari** S*. Kumar, V. Aswal. Rational Design of Supramolecular Dynamic Protein Assemblies Using a Micelle-Assisted Activity-based Protein Labeling Technology. *Chem. Eur. J.* **2018**, *24*, 16085-16096. (* - Equal contribution)
3. B.S. Sandanaraj, M.M. Reddy, K. Jagan, **P.J. Bhandari**. Rational Design of Semi-Synthetic Protein Complexes with the Defined Oligomeric State. *Chem. Select.* **2019**, *4*, 6397-6402.
4. **P.J. Bhandari**, B.S. Sandanaraj, M.M. Reddy, M. M, A. B. Lohote. Programmed Disassembly of Multi-Responsive Supramolecular Protein Assemblies. (*manuscript under preparation*).
5. **P.J. Bhandari**, B.S. Sandanaraj. Synthesis and Disassembly Studies of Redox-Responsive Supramolecular Protein Assemblies. (*manuscript under preparation*).
6. B.S. Sandanaraj, **P.J. Bhandari**, M.M. Reddy. Generation-dependent Supramolecular Assemblies of Protein-dendron Conjugates. PCT WO 2019/111282 A1 (Published June 2019).
7. B.S. Sandanaraj, **P.J. Bhandari**, M.M. Reddy. Supramolecular Protein Assemblies with Advanced Functions and Synthesis thereof. US Patent, 2019/0134212 A1 (Published May 2019).
8. B.S. Sandanaraj, M.M. Reddy, **P.J. Bhandari**, K. Jagan. Hydrophobin Mimics: Process for Preparation thereof. US Patent 10,188,136 B2 (Granted – Jan 2019).