Chemoselective Synthesis of Boronopeptides and Their Biological Applications

A Thesis Submitted

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

by

Saurav Chatterjee

(2019CYZ0008)

DEPARTMENT OF CHEMISTRY

INDIAN INSTITUTE OF TECHNOLOGY ROPAR

JUNE 2024

Saurav Chatterjee

Chemoselective Synthesis of Boronopeptides and their Biological Applications

ॐ नमः शिवाय

The universe bows to Lord Shiva.

I bow to Lord Shiva.

Nāsti me jātih nāsti me dharmapanthāh
Nāsti vā bhāsā nāsti me rājyam
Eko'yam paricayah ekaiva vyaktitā
bhāratamātuh sutā bhāratamātuh sutā
Jīvanam mama ātmā vā
Karmam mama śraddhā vā
Bhāratamātre khalu bhāratamātre khalu
Yadi moksaprāptih yadi vā sampraśnam
Ekā pratyuktih ko'si praśnasya
prcchet adhyaksah yadi vā parameśah
Ekā pratyuktih bhāratamātuh sutā
Eko'yam paricayah ekaiva vyaktitā
bhāratamātuh sutā bhāratamātuh sutā

Neither caste, nor religion,
Neither language, nor kingdom (state),
Only one identity I bear,
I am a child of Mother India.
My life, my soul,
My work, my faith,
Is all for Mother India.
Whether I attain Moksha,
or I stand on Judgement Day,
When asked who I am,
By the One Most High,
There will be only one answer:
I am a child of Mother India.
Only one identity I bear,
I am a child of Mother India.

DEDICATED

TO

Late Shri Maniklal Banerjee

My grandfather, who never allowed me to give up

and

Smt. Sipra Chatterjee

My mother, who inspired me to study Chemistry

DECLARATION OF ORIGINALITY

I hereby declare that the work which is being presented in the thesis entitled **Chemoselective Synthesis of Boronopeptides and Their Biological Applications** has been solely authored by me. It presents the result of my own independent investigation/research conducted during the time period from July, 2019 to May, 2024 under the supervision of Dr. Anupam Bandyopadhyay, Assistant Professor, IIT Ropar. To the best of my knowledge, it is an original work, both in terms of research content and narrative, and has not been submitted or accepted elsewhere, in part or in full, for the award of any degree, diploma, fellowship, associateship, or similar title of any university or institution. Further, due credit has been attributed to the relevant state-of-the-art and collaborations (if any) with appropriate citations and acknowledgments, in line with established ethical norms and practices. I also declare that any idea/data/fact/source stated in my thesis has not been fabricated/ falsified/ misrepresented. All the principles of academic honesty and integrity have been followed. I fully understand that if the thesis is found to be unoriginal, fabricated, or plagiarized, the Institute reserves the right to withdraw the thesis from its archive and revoke the associated Degree conferred. Additionally, the Institute also reserves the right to appraise all concerned sections of society of the matter for their information and necessary action (if any). If accepted, I hereby consent for my thesis to be available online in the Institute's Open Access repository, inter-library loan, and the title & abstract to be made available to outside organizations.

Signature

Name: Saurav Chatterjee
Entry Number: 2019CYZ0008
Department: Chemistry
Indian Institute of Technology Ropar
Rupnagar, Punjab 140001

Date: 28.05.2024

ACKNOWLEDGMENTS

I always felt that Ph.D. is not just a degree, but a way of life. It has taught me to have a different perspective about the world around me. I would like to appreciate the person who has influenced me most in this journey, my supervisor Dr. Anupam Bandyopadhyay, an ever-energetic person fascinated by chemical biology. He was always present when the lab needed him and never failed to inspire us with his research acumen. It has been my great privilege to be able to work and learn from him.

I am also thankful to my doctoral committee members – Dr. T. J. Dhilip Kumar, Dr. Yashveer Singh, Dr. Indranil Chatterjee and Dr. Srivatsava Naidu for their valuable inputs and suggestions. I am grateful to the current head of the department Dr. Tharamani C.N., former head of the department Dr. T. J. Dhilip Kumar and other faculty members of Chemistry Department for facilitating the use of various departmental instruments and their valuable suggestions.

It is my privilege to thank the present director of IIT Ropar, Prof. Rajeev Ahuja, and former director of IIT Ropar, Prof. Sarit K. Das, for providing world-class research infrastructures. I acknowledge the NMR, and HRMS facilities of IIT Ropar. I want to thank the convenor of the NMR facility Dr. Manoj Kumar Pandey, for his cooperation and insightful discussion sessions on ^{11}B NMR. Next, I sincerely thank HRMS facility committee members Dr. Avijit Goswami and Dr. Anupam Bandyopadhyay along with the operators Mr. Nazre Heyat, Dr. Nilotpal Singha, Mr. Suryanarayan Parida and Ms. Hanspreet Kaur. I also thank the non-teaching staff especially, Mr. Nagendra, Mr. Sushil, Mrs. Poonam, and Mrs. Samita for their help.

This work would not have been possible without the untiring contribution of our collaborators – Dr. Archana Mukherjee (BARC) whose lab performed the radiolabeling and in-vivo studies and Dr. Srivatsava Naidu (IIT Ropar) whose lab taught us in-vitro cell-culture experiments. I would also like to thank the companies Kashiv Biosciences, Amneal, Raks Pharma, Kayura Effect, and Shivanka who helped us to learn and grow.

Without a lively and healthy working environment, science does not flourish. My lab-mates made sure that there was not a single dull moment. I am very fortunate to have worked with Mr. Arnab Chowdhury, Mr. Nitesh Mani Tripathi, Dr. Basab Kanti Das, Mr. Vinod Gour, Mr. Bhavesh Badgujar, Mr. Ramraj Kameriya, Ms. Neelam Verma, Ms. Sagnika Ghosh, Mr. Neloy Kumar Ghosh, Ms. Deepa Debnath, Mr. Akumlong Pongen, Mr. Dhanjit Sarania, Mr. Amandeep Karangtar, Mr. Bibekananda Pati and Mr. Soumit Datta. The scientific discussions with my batchmates Ms. Sweta Saha, Ms. Akansha Chaturvedi, Mr. Arijit Hazra, Mr. Rakesh Kumar, Mr. Gulshan Singh, Mr. Man Singh, Ms. Ritika Arora, Mr. Vishal Singh, Mr. Priyank Sharma, Mr. Sandeep Patel, Mr. Vatan Chawla, Mr. Shine Kader, Mr. Ganesh More, Mr. Shambhu Nath Chatterjee, and Mr. Anoop Chahal has greatly helped me develop as a researcher.

I would to express my heartfelt gratitude to my parents, Mr. Kabiram and Mrs. Sipra, my sister Mrs. Sarmistha, and my grandparents without whom I would not be here today. They are my support system who provide me courage to face the world. Of special mention, are my close friends Ms. Sweta Saha, with whom I have shared many lovely experiences, Mr. Subham Maity and Mr. Arnab Dutta who I have known since the last 12 years and have helped me to be a better person.

Finally, I would like to thank Lord Shiva for looking out for me. May he bless us all.

CERTIFICATE

This is to certify that the thesis entitled **Chemoselective Synthesis of Boronopeptides and Their Biological Applications**, submitted by **Saurav Chatterjee (2019CYZ0008)** for the award of the degree of **Doctor of Philosophy** of Indian Institute of Technology Ropar, is a record of bonafide research work carried out under my guidance and supervision. To the best of my knowledge and belief, the work presented in this thesis is original and has not been submitted, either in part or full, for the award of any other degree, diploma, fellowship, associateship or similar title of any university or institution.

In my opinion, the thesis has reached the standard fulfilling the requirements of the regulations relating to the Degree.

Signature of the Supervisor

Name: Anupam Bandyopadhyay

Department: Chemistry

Indian Institute of Technology Ropar

Rupnagar, Punjab 140001

LAY SUMMARY

Peptides are short biopolymers of amino acids (typically 2-50) linked through covalent bonds. One can imagine them as building blocks of proteins, similar to what a brick is to a house. From a medicinal chemistry perspective, peptides are being studied for their potential to treat various diseases, including cancer and diabetes. Researchers are exploring ways to design peptides that can specifically target diseased cells while leaving healthy cells unharmed. This targeted approach could lead to more effective and less toxic treatments for patients.

In order for peptides to function effectively outside their native state, chemical modifications are necessary. In this respect, modification with boronic acids, to create 'smart peptides' or boronopeptides are emerging as forerunners. Boronic acid is a compound that contains a boron atom bonded to hydroxyl group(s). In medicinal chemistry, boronic acid is gaining attention for its ability to react with certain molecules in the body, particularly sugars and proteins. This reactivity makes boronic acid useful for designing drugs that can target specific biological targets, such as enzymes or receptors involved in disease processes. Since the last two decades, five boronic acid based drugs have been approved by the Food and Drug Administration (FDA) for treatment of diseases such as multiple myeloma. The multiple arrays of reactivity enabled by the mere presence of boron has led scientists to claim it as a 'magic element'.

Boronopeptides enjoy many perks in the realm of medicinal chemistry such as superior carbohydrate/ glycan recognition (applications in cancer diagnosis), conjugating to drugs/fluorophores in a reversible or irreversible manner, peptide cyclization to increase the targeting ability and its half-life in the body. These benefits can be reaped only if the production matches the demand. Given the current state of affairs, facile synthetic strategies for boronopeptides, which can provide clean and quantitative conversion, are in huge demand. The main objective of my thesis is to try and fill this gap by introducing novel methods for the synthesis of boronopeptides. Keeping the synthetic aspect as a starting point, my thesis also aims to showcase real-time applications in diagnosing cancer cells and in-vivo imaging of bacterial cells. Combating cancer and antimicrobial resistance currently pose a significant challenge to the well-being of humanity, and I hope my research will be a stepping stone to a disease-free utopia.

ABSTRACT

Non-natural late-stage modifications to peptide chains have been in effect since the 1950s to improve stability, binding affinity, and targeting selectivity. Recent research focuses on merging synthetic organic chemistry and chemical biology to develop robust methods for chemical modifications to peptide chains that do not interfere with their native biological functions but rather serve as upgrades. In this regard, boronic acids (BA) pose as an exciting module for the diversification of peptide function due to their extensive bonding nature with endogenous nucleophiles.

In the thesis, we have established two methods to install BA derivatives on cysteine residues in peptides and proteins. The first method employs the well-known thiol-ene click chemistry where an alkene handle appended BA molecule was effectively installed on cysteine residue(s) in the presence of a photoinitiator and 365 nm light. The strategy yielded clean to quantitative conversion with eight different BA derivatives spanning over eight biologically active peptides. The BA derivatives were handpicked based on the diverse reaction genres offered by them, which in turn were demonstrated by four chemical modules – cyclization, functional group conversion, disulfide formation and α-nucleophile conjugation.

The established methodology was diligently used to install BA derivatives, known for their superior sugar binding ability, on programmed cysteine residues placed on wheat germ agglutinin (WGA) peptide which reportedly binds selectively to sialyl-glycans. Currently, the overexpression of sialyl-glycan antigens has been considered as the hallmarks of cancer progression and the in-use clinical techniques involve the use of lectins which are costly and difficult to handle. We hypothesized that a synergistic binding exerted by the designed boronopeptides will prove to be as effective as a lectin, thereby providing a cheaper alternative. NMR and cell imaging studies proved the binding selectivity of the boronopeptides while flow cytometer and EC_{50} studies showed that Wulff boronate conjoined peptides are pretty effective in recognising the glycans. In a separate investigation, we endeavoured to confront the imminent challenge presented by antimicrobial resistance (AMR) by affixing boronic acid probes to ubiquicidine, a widely recognized antimicrobial peptide. The designed peptides show increased salt tolerance and the ability to differentiate sterile inflammation from infection. Radiolabelling with ^{99m}Tc and ^{68}Ga enabled superior imaging techniques to be applied to in vivo studies with mice.

In our second method, we employed practically simple and benign S_N2 reaction to install methyl boronic acid on cysteine residues in peptides and proteins. A quantitative conversion can be achieved within 10 min in water at physiological pH with only 1.5-2 equivalent of the reagent, providing a green technology to modify sensitive biomolecules. This opens up avenues to create multivalent peptides which provide enhanced affinity and specificity.

Keywords: peptide, boronic acid, boronopeptides, thiol-ene click, sialyl-glycan, sialic acid, antimicrobial resistance, halomethyl boronic acid, protein modification

LIST OF PUBLICATIONS

Journal papers

(1) **Chatterjee, S.**; Anslyn, E. V.; Bandyopadhyay, A. Boronic Acid Based Dynamic Click Chemistry: Recent Advances and Emergent Applications. *Chem Sci* **2021**, *12* (5), 1585–1599. https://doi.org/10.1039/d0sc05009a.

(2) **Chatterjee, S.**; Tripathi, N. M.; Bandyopadhyay, A. The Modern Role of Boron as a 'Magic Element' in Biomedical Science: Chemistry Perspective. *Chemical Communications* **2021**, *57* (100), 13629–13640. https://doi.org/10.1039/D1CC05481C.

(3) **Chatterjee, S.**; Bandyopadhyay, A. Cysteine-Selective Installation of Functionally Diverse Boronic Acid Probes on Peptides. *Org Lett* **2023**, *25* (13), 2223–2227. https://doi.org/10.1021/acs.orglett.3c00386.

(4) **Chatterjee, S.**; Chowdhury, A.; Saproo, S.; Mani Tripathi, N.; Naidu, S.; Bandyopadhyay, A. Capturing Sialyl-glycan on Live Cancer Cells by Tailored Boronopeptide**. *Chemistry – A European Journal* **2024**, *30* (8). https://doi.org/10.1002/chem.202303327.

(5) Bhatt Mitra, J.; **Chatterjee, S.**; Kumar, A.; Bandyopadhyay, A.; Mukherjee, A. Integrating a Covalent Probe with Ubiquicidin Fragment Enables Effective Bacterial Infection Imaging. *RSC Med Chem* **2022**, *13* (10), 1239–1245. https://doi.org/10.1039/D2MD00190J.

(6) Bhatt Mitra, J.; **Chatterjee, S.**; Kumar, A.; Khatoon, E.; Chandak, A.; Rakshit, S.; Bandyopadhyay, A.; Mukherjee, A. Expanding a Peptide-Covalent Probe Hybrid for PET Imaging of S. Aureus Driven Focal Infections. *EJNMMI Radiopharm Chem* **2024**, *9* (1), 25. https://doi.org/10.1186/s41181-024-00252-4.

Patents

HALOMETHYL BORONIC ACID FOR CHEMOSELECTIVE AND SITE-SELECTIVE INSTALLATION OF BORONIC ACID AND PRODUCTION OF DEHYDROALANINE IN PEPTIDE AND PROTEIN, **S. Chatterjee**, A. Bandyopdhyay, Indian Institute of Technology (IIT) Ropar, Punjab, India (Patent Application Number 202411017017).

TABLE OF CONTENTS

LIST OF FIGURES

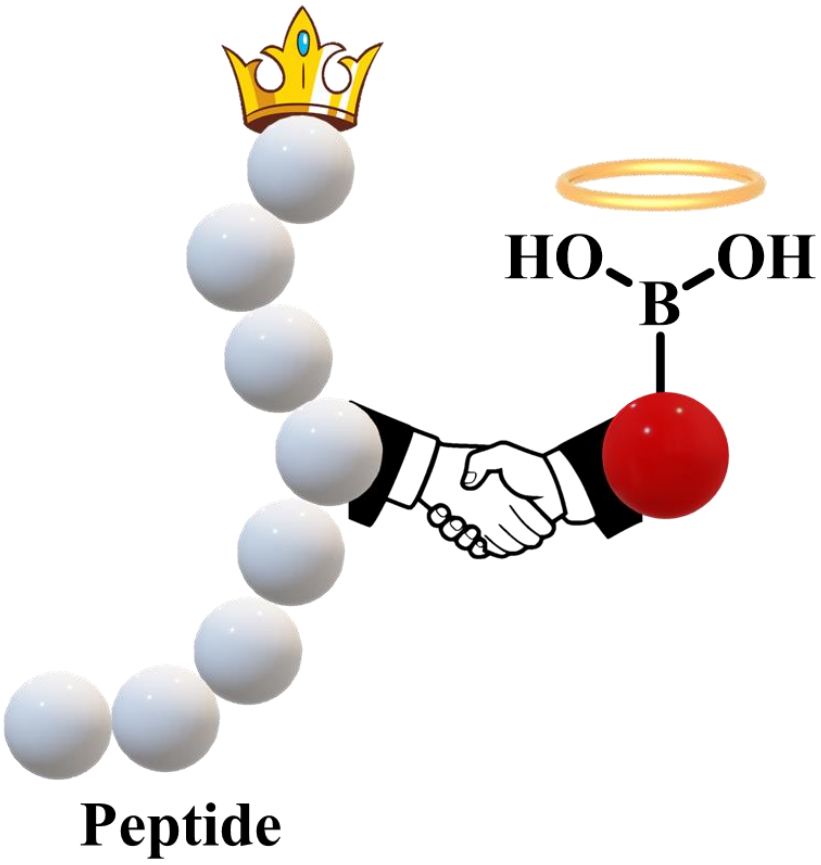
NOTATIONS AND ABBREVIATIONS

Ac	Acetyl
µL	microlitre
2APBA	2-Acetylphenyl boronic acid
2FPBA	2-Formylphenyl boronic acid
AcCys	Acetyl cysteine
ACE	Angiotensin converting enzyme
Acm	Acetamidomethyl
Acp	Amino caproic acid
Aib	α-amino isobutyric acid
AMP	Antimicrobial Peptide
AMPBA	Aminomethyl phenylboronic acid
AMR	Antimicrobial Resistance
BA	Boronic acid
BARC	Bhaba Atomic Research Centre
BASHY	Boronic acid-derived salicylidenehydrazone
BHS	Boron hot-spot
BLB	Blue light black
BNCT	Boron neutron capture therapy
Boc	*tert*-Butyloxycarbonyl
BSA	Bovine serum albumin
CD	Circular dichroism
$CDCl_3$	Chloroform-d_3
COSY	Correlation spectroscopy
CT	Computed Tomography
CTC	Chloro trityl
DCC	Dicyclohexyl carbodiimide
DCM	Dichloromethane
Dha	Dehydroalanine
DiPEA	Diisopropyl ethyl amine
DMF	Dimethyl formamide
DMPA	2,2-Dimethoxy-2-phenyl-acetophenone
DMSO	Dimethyl sulfoxide
DNP	Dinitrophenyl hydrazine
DTT	Dithiothreitol
EC_{50}	Effective concentration
EDT	1,2-Ethanedithiol
eq.	equivalent
ESI	Electrospray ionization
Et_3N	Triethylamine
EtOAc	Ethyl acetate
FBS	Fetal bovine serum
FDA	Food and drug administration
FITC	Fluorescein isothiocyanate
Fmoc	Fluorenylmethoxycarbonyl
GIP	Glucose-dependent insulinotropic peptide
GLP	Glucagon-like peptide
GPCR	G-protein coupled receptors
GSH	Glutathione
HBTU	2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIV	Human immunodeficiency virus
hLF	Human Lactoferrin
HOBt	Hydroxybenzotriazole
hr	hour
HRMS	High resolution mass spectrometry
K_2CO_3	Potassium carbonate
KOAc	Potassium acetate
LC-MS	Liquid chromatography coupled mass spectrometry
LUMO	Lowest unoccupied molecular orbital
m.p.	Melting point
Mcl-1	Myeloid cell leukemia sequence 1
MDM2	Murine double minute 2
MeOH	Methanol
mg	milligram
mL	mililitre
mM	milimolar
mmol	milimole
MRI	Magnetic Resonance Imaging
MS/MS	Tandem mass spectrometry
Na_2CO_3	Sodium carbonate

$NaBH_4$	Sodium borohydride
$NaCNBH_3$	Sodium cyanoborohydride
$NaHCO_3$	Sodium bicarbonate
NCL	Native chemical ligation
Neu	Neuraminic acid
NHS	Positron Emission Tomography
NMP	N-methyl pyrrolidone
NMR	Nuclear magnetic resonance
NODAGA	1, 4, 7 triazacyclononane 1- glutaric acid 4–7 acetic acid
NOTA	1,4,7-triazacyclononane-1,4,7-triacetic acid
PBA	Phenylboronic acid
PBS	Phosphate buffered saline
$Pd(dppf)Cl_2$	(1,1'-Bis(diphenylphosphino) ferrocene)palladium(II) dichloride
PDL	Programmed death ligand
pdt.	product
PET	Positron emission tomography
PG	Phosphatidyl-glycerol
PPI	Protein-protein interaction
ppm	parts per million
RNS	Reactive nitrogen species
ROS	Reactive oxygen species
RP-HPLC	Reverse phase high performance liquid chromatography
SDS-PAGE	Sodium dodecyl sulfate – polyacrylamide gel electrophoresis
Sia	Sialic acid
SHA	Salicylhydroxamic acid
SPPS	Solid phase peptide synthesis
SNA	*Sambucus nigra*
STI	Soft Tissue Infections
SPECT	Single Photon Emission Computed Tomography
sTn	Sialyl-Tn antigen
TCEP	Tris(2-carboxyethyl)phosphine
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TIPS	Triisopropylsilane
TLC	Thin-Layer Chromatography
UBI	Ubiquicidin
VA-044	2,2'-Azobis[2-(2-imidazolin-2-yl) propane]dihydrochloride.
WHO	World Health Organization

CHAPTER 1

INTRODUCTION TO BORONOPEPTIDES



"What is now proved, was once only imagined." *-William Blake, 1757-1827*

1.1 *PEPTIDES*

The birth of protein and peptide chemistry is indeed a fascinating story attributed to sugars, yeasts, and the hunger of a scientist to always ask questions. Around the 1890s, based on his studies on sugars and purines, Emil Fischer was chosen as a successor of August Hofmann as the professor of chemistry at the University of Berlin, Germany. While studying the digestion of sugars by yeast, Fischer observed that only a few stereoisomers of glucose were digested by the enzymes in yeast. He reasoned that the enzyme must contain a specific spatial arrangement of the molecules, enabling it to interact specifically. This spurred him to synthesize a class of molecules composed of amino acids, which the enzymes could break down similarly[1]. In 1901, along with Fourneau, Fischer reported the first synthesis of a dipeptide – glycylglycine, and in 1902, he introduced the term 'PEPTIDES' at a meeting with German chemists. Thus began the remarkable journey of peptide chemistry, influencing lives and research alike since its inception.

The definition of peptides, though ambiguous, is arbitraryly considered as short biopolymers of about 2-50 amino acids[2] linked via amide bonds. The amino acids, in turn, are an asymmetric system comprising an amine and carboxylic acid bonded to an α-chiral carbon. Of the reported 20 naturally occurring amino acids, glycine is the only achiral amino acid. At the same time, among the rest, cysteine stands out with an *R*-configuration as opposed to the commonly observed *S*-configuration for the other monomers. The particular arrangement of amino acids within a peptide sequence makes up the primary structure and endows it with a specific chirality. This arrangement enables a peptide fragment to take up particular secondary structures (α-helix, β-sheets, random coils), which, in an ensemble of polypeptides, further folds into tertiary and quaternary structures to give rise to proteins (Figure 1.1).

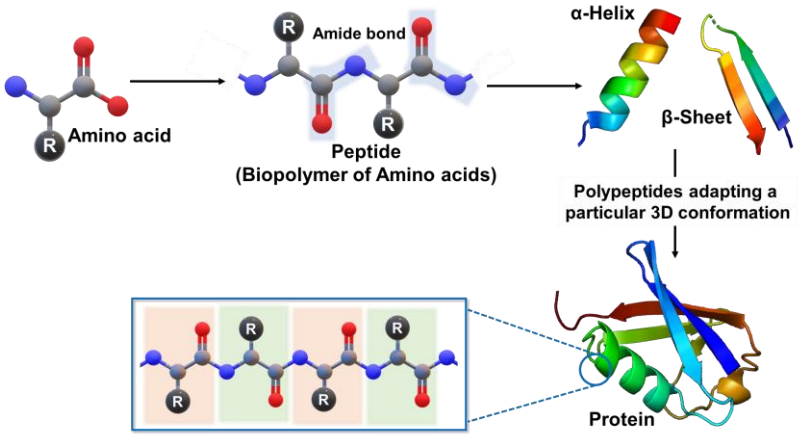


Figure 1.1: Amino acids undergo condensation reactions to form peptide chains which in turn polymerize to form proteins (blue ball represents nitrogen, red ball represents oxygen and grey ball represents carbon; hydrogen atoms have not been shown for clarity).

1.2 *ROLE OF PEPTIDES IN BIOLOGICAL SYSTEM*

Given the overwhelming role played by peptides in biological systems, they have been divided into various classes, such as antibiotics, toxins, insecticides, vaccines, thyroid hormones, gastrointestinal peptides, and neuropeptides, to name a few[3]. Broadly, they can be classified into two main types – endogenous and exogenous peptides[4]. As evident from the name, endogenous peptides are present in the body and produced in different cell types depending on their function. Neural cells produce peptides that have analgesic or opioid effects, while the peptides produced by immune cells play a role in inflammation and antimicrobial activities. The exogenous peptides, on the other hand, have to be assimilated from dietary supplements and medications.

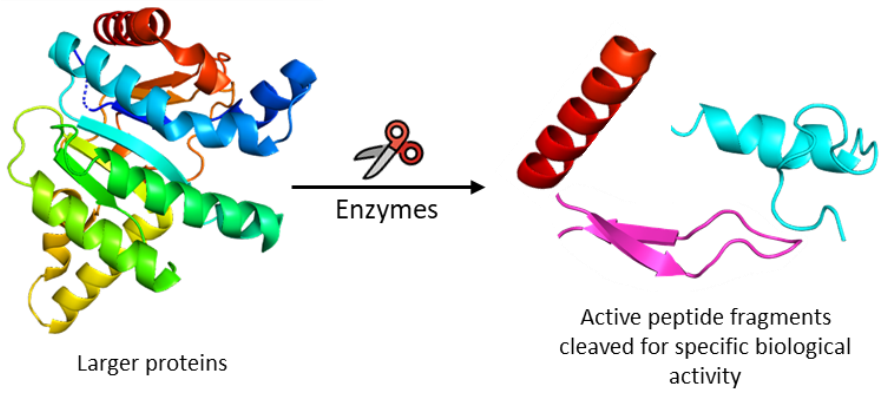


Figure 1.2: Cartoon representation of enzymatic cleavage of biologically active peptide sequences cleaved from proteins.

Active peptide sequences buried within mature proteins are released in the body most commonly by enzyme hydrolysis (mostly digestive enzymes such as pepsin, trypsin, and chymotrypsin), as represented in Figure 1.2. For example, trypsin is responsible for the formation of ACE (angiotensin-converting enzyme) inhibitory peptides and calcium-binding phosphatases, while papain and pronase have been employed for the enzymatic hydrolysis of soy flour and wheat[4]. Such enzymatic methods are also used to obtain active peptides (antimicrobial, antihypertensive, etc.) from foods such as meat and milk. Likewise, in response to electrical impulses to the hypothalamus, the production of oxytocin as a pre-propeptide is triggered, which is then further matured into the active sequence responsible for parturition, lactation, maternal behavior, and sexual functions in men.

The peptides exert the desired functions by binding to their respective receptors on the cell surface, preferably by a specific non-covalent interaction. In the uterus, the oxytocin couples to the oxytocin receptor, belonging to the G-protein coupled receptor (GPCR) superfamily, which, through a cascade of events, induces the release of Ca^{2+} from cells. This increased mobility of Ca^{2+} induces the smooth uterine contraction of the muscles[5].

1.3 *PEPTIDES IN DRUG DESIGN*

The history of peptide synthesis, especially towards drug design, is incomplete without the names of some of the giants of the 20th century, all of them Nobel Prize winners – de Vigneaud, Banting and Macleod, Schally and Guillemin, Sanger, and Merrifield. The first landmark was the isolation of insulin peptide by Frederick Banting in 1921[6] from the porcine pancreas. During this period, diabetes mellitus, particularly Type 1 diabetes,

was considered fatal, and the isolation of this peptide hormone was a major landmark in putting people's faith in scientific research. Hailed as a 'miracle drug', it was commercialized within a year of its discovery by Eli Lilly and is still in prevalent use. The second wave of peptide drug development began in the late 1980s, spearheaded by the success of Eli Lilly, who introduced human insulin produced by recombinant technology. In the recombinant technology, the human insulin gene was introduced into a bacterial cell, thus allowing the production of large quantities in a laboratory setting. Synthetic gonadotropin-releasing hormones leuprolide and goserelin followed suit, and with the advances in structural biology, recombinant biologics, new synthetic and analytical techniques have led to the approval of about 80 peptide drugs on the global market[7] with several others in the pipeline[8]. Isolation and marketing of short peptide hormones such as oxytocin and vasopressin by Vigneaud in the 1950s also helped popularize the importance of peptides as drug candidates. This further spurred research to develop techniques for robust synthesis and purification of peptides.

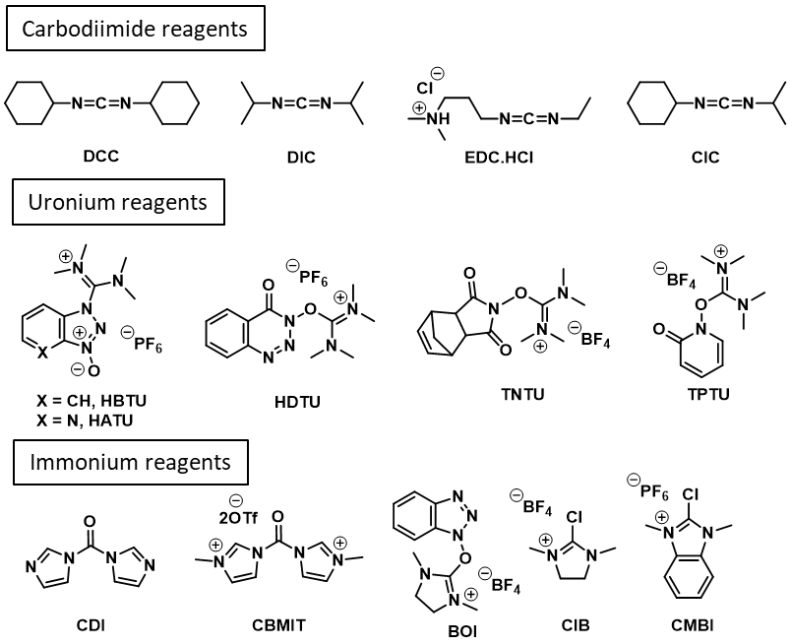


Figure 1.3: List of some of the developed coupling reagents.

Due to the presence of both amine and carboxylic acid functionality in an amino acid, it was necessary to design protecting groups for the sequential synthesis of the peptide chain. A significant milestone was the development of benzyloxycarbonyl protecting group (Bergmann and Zervas, 1932)[9] and was followed by the discovery of *tert*-butyloxycarbonyl (Boc) group in 1957[10] and 9-flyorenylmethoxycarbonyl (Fmoc) in 1970[11] by Carpino and group. To facilitate the efficient coupling of acid and amine, coupling reagents were designed (Figure 1.3) and the Sheehan group introduced N,N'-dicyclohexylcarbodiimide (DCC)[12] as the first diimide-based coupling reagent in 1955. This was followed by the phosphonium reagents in the early 1970s by Castro[13], which, like DCC showed the propensity to racemize. This was solved when HOBt was developed as a racemization suppression reagent (König and Geiger, 1970[14]). Thus, the modified coupling agents were coupled with the acid activating compound and racemization suppressor into one reagent, leading to uronium reagents such as HBTU in 1978 by Gross[15]. Xu further modified the uronium reagents to introduce immonium reagents in 2000, followed by imidazolium and organophosphorous reagents[16]. However, steady progress of peptide synthesis was still hindered due to the laborious solution-phase protocol, which required months to years of work for longer peptides. A major breakthrough was achieved when Robert Bruce Merrifield invented

the solid-phase peptide synthesis in 1963[17] which simplified the protection-deprotection and purification steps involved in a typical synthesis cycle by a huge margin. Merrifield was awarded the Nobel Prize in Chemistry in 1984 for his contributions to peptide chemistry. Presently, the synthetic procedure has beautifully matured into microwave assisted synthetic techniques[18] using DIC/Oxyma as the coupling agent requiring only few minutes coupling time, typically useful in an industrial setting, or the advanced flow synthesis which can synthesize proteins in a short amount of time[19]. The advent of high-performance liquid chromatography for large-scale purification and the Edman degradation[20] method for the determination of the amino acid sequence further supported the advancement of peptides in the pharmaceutical industry.

The wide availability of peptides as drug candidates in the market came with its own problems. The significant developments for peptides as therapeutics occurred during the 1970-1980s when the approval of ~20 orally available small molecule drugs was the norm. In such a setting, peptides with very low oral bioavailability did not appeal to the general patient. To top it off, peptides, or more correctly, peptides derived from native sequences, suffered from a low probability of crossing the gut barrier, low plasma stability, and thus short circulation times. These factors apparently overshadowed its advantages, such as superior selectivity, higher affinity, lower off-target toxicity, and low immunogenicity. Given their size, peptides enjoy a niche between small molecule drugs and biologics with maximum application found in protein-protein inhibitions (PPIs, covering a contact area of 1500-3000 Å^2) and as agonists (requiring only 5-20% receptor occupancy). Currently, peptide drugs occupy about 5% of the global pharmaceutical market, with global sales exceeding USD 50 billion in 2019[7]. With a 10% increase in their compound annual growth rate for the foreseeable future, the importance of peptide drugs will only keep rising with time[21] (Figure 1.4).

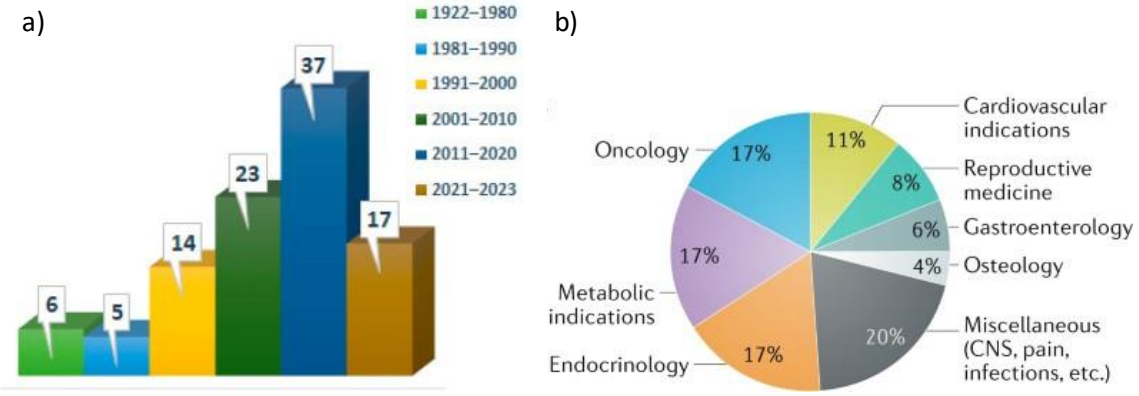


Figure 1.4: a) FDA approved peptides from 1922-2023 (Reproduced with permission from Musaimi, *Cancers*, 2024); b) Most commonly targeted indications of peptide therapy (Reproduced with permission from Muttenthaler et al., *Nat. Rev. Drug Discov.*, 2021).

1.4 *CHEMICAL MODIFICATIONS OF PEPTIDES IN DRUG DESIGN*

Peptides are usually constituted of 20 naturally occurring amino acids. The primary structure of peptide dictates the particular order of arrangement of the amino acids and, as such, determines the secondary structure and the degree of hydrophobicity of the sequence. More often than not, the biological function exerted by a peptide sequence depends on the secondary structural arrangement within a protein backbone, such as α-helix, β-sheet, or a rigid turn, which enables it to interact tightly with a receptor. For instance, the helical

conformation of the p53 domain interacts with MDM2/X to inhibit the apoptotic pathway. Such epitopes lose their native conformation and bioactivity when disconnected from the protein. Therefore, chemical or non-natural modifications to the peptide chain are crucial for further activity modulation. Some of the major drawbacks experienced by peptide drugs include low oral bioavailability (about 90% of all peptide drugs are delivered by injection), low probability of crossing the gut barrier, poor in vivo stability, and higher production cost (compared to small molecules). Current research on peptide drugs evolves around backbone modification, stapling strategies, side-chain modifications, inversing chirality of residues, pegylation or fatty acid, and protein conjugation to improve their target specificity, affinity, in vivo stability, and druggability[22] (Figure 1.5a). One example is the FDA approved peptide drug tirzepatide, marketed by Eli Lilly, for treating type 2 diabetes[23]. It is the only dual glucagon-like peptide-1 (GLP-1) and glucose-dependent insulinotropic peptide (GIP) receptor agonist that can significantly reduce glycemic levels and improve insulin sensitivity, as well as reduce body weight by more than 20% and improve lipid metabolism. A C20 fatty diacid is installed on the Lys20 residue via a hydrophilic linker, coupled with two α-amino isobutyric acids (Aib) residues serve to drastically increase the half-life of the peptide to five days by binding it to albumin and limiting its subcutaneous injection to only once a week (Figure 1.5b).

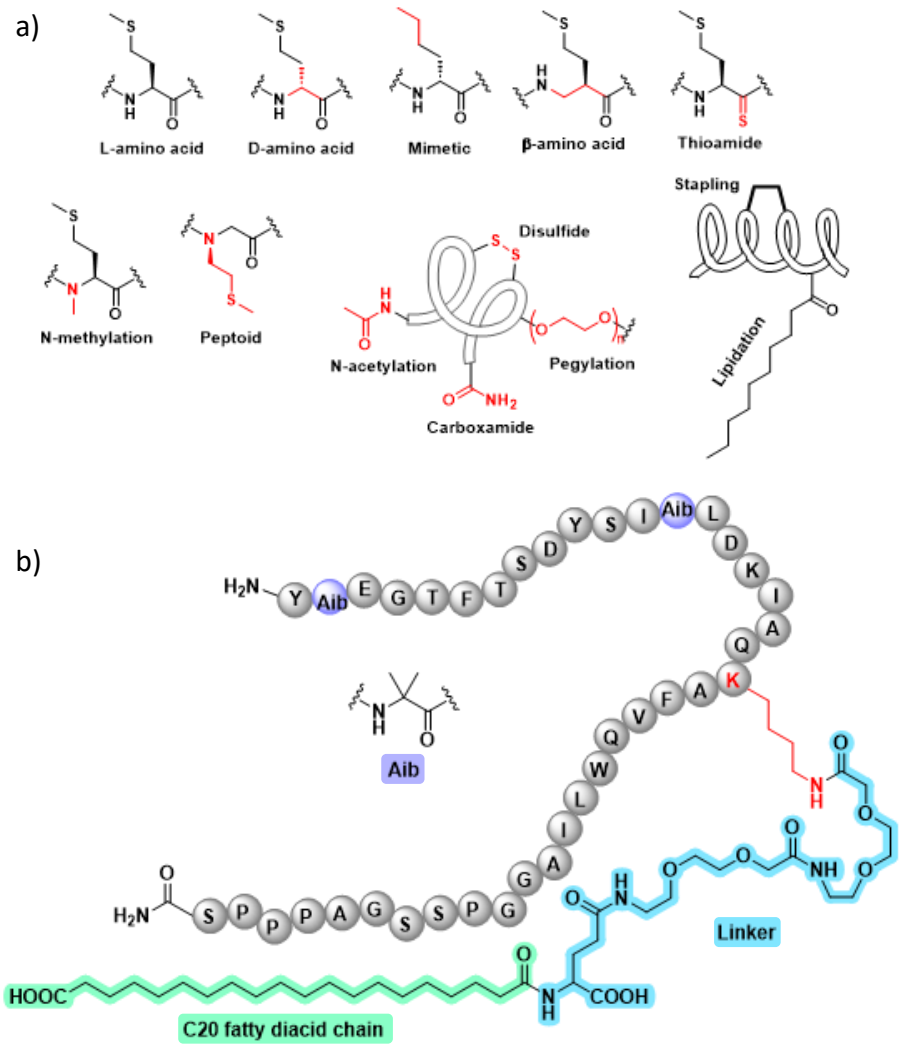


Figure 1.5: a) Well- validated chemical modifications used in peptide drug development to increase metabolic stability and bioavailability; b) Structural features of tirzepatide.

Pertinent to chemical modifications of peptides, bioconjugation and bioorthogonal chemistry have shown remarkable importance. Bioconjugation generally refers to forming a stable covalent bond between the native form of biomolecule and a suitable chemical entity. This is achieved by a direct covalent linkage with suitably reacting functional groups on the biomolecules in a chemoselective manner. Due to the presence of multiple functional groups in a biomolecule, it is imperative to design chemical strategies that allow modification to only a particular functionality in the presence of the other reactive groups, known as the chemoselective method (Figure 1.6a). Two different classes of protein/peptide modification in the laboratory have facilitated enormous opportunity in chemical biology and therapeutic endeavor. First: introduction of unnatural amino acids (UAAs) to a native protein, which requires a genetic code expansion technology and a specialized laboratory for chemical mutagenesis[24,25]. With limited applications to this, similar modification can also be achieved by utilizing specific enzymes. Second: late-stage chemical modification of a recombinant protein by targeting natural residues in a mild condition and site-selective manner[26,27]. To date, late-stage chemical modification of protein has shown multiple opportunities to incorporate chemical and functional diversity. These modifications enable attributing stability, installing secondary handle, presenting reporters, preparing drug conjugates, and beyond[28]. In this respect, dynamic covalent bonding strategies have garnered much attention due to their reversible nature which greatly reduces off-target toxicity[29]. Even in a protein, the microenvironment of a particular residue will be different as compared to its counterparts due to the folding in the quaternary structure. Not only side chain modification, but backbone modification i.e. amide backbone of the biomolecules can also be accessed, for eg., cyclosporine, a naturally available cyclic peptide with N-methylated backbone from fungi was the first orally available peptide drug for immunosuppression (Figure 1.6b). Thus, continuous research is essential to develop new and better reaction strategies.

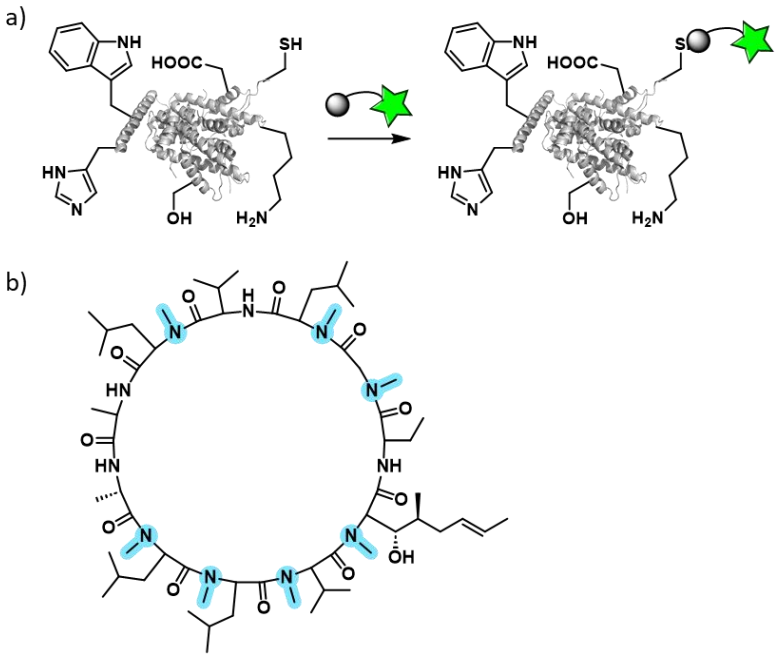


Figure 1.6: a) Chemoselective reaction strategy for modification of a single amino acid residue; b) Structure of cyclosporine with N-methylated backbone highlighted in blue.

1.5 *IMPORTANCE OF BORONIC ACIDS IN BIOMEDICAL APPLICATIONS*

A pool of synthetic modifications is strategically employed to provide the peptide a functionality tailored to its need. Case in point, 'the lesser the better' – simpler covalent modifications which can provide multiple advantages are considered superior since they do not tend to interfere with the peptide's native targeting ability. The boronic acid-based modification strategy has garnered much attention in this regard, especially over the past decade. With five FDA approved boronic acid drugs[30,31] starting with Bortezomib[32] in 2003, a first-in-class proteasome inhibitor used in the clinic as an anticancer agent in patients with multiple myeloma, followed by Tavaborole[33] in 2014, a leucyl-tRNA synthetase inhibitor commercialized for the treatment of onychomycosis, and Ixazomib (2015), Crisaborole (2016), Vaborbactam (2017) (Figure 1.7a). Ixazomib became the first orally available proteasome inhibitor due to the derivatization of the boronic acid moiety with citric acid. The unique chemistry offered by boron stems mainly from the availability of a vacant *p*-orbital, which enables it to dynamically change its coordination from sp^2 to sp^3 in the presence of endogenous nucleophiles such as alcohols and amines at physiological pH (Figure 1.7b) to form cyclic or linear derivatives. The most well-known and widely explored reaction of boronic acid is the dynamic covalent bonding with *cis*-diols to form boronate esters[34] (Figure 1.7c). Raines group has exploited the boronate ester formation to show that it helps in cell penetration of biomacromolecules by binding to cell-surface glycans[35,36]. This has become a powerful synthetic strategy to construct self-organizing systems, sensors, and many sets of functional materials. The current research on boronic acid-based pharmaceutics has skyrocketed, due to the vast array of interactions and reactivity it offers in a biological setting[37]. Therefore, scientists have coined the term boron as a 'magic element' in biomedical applications[38].

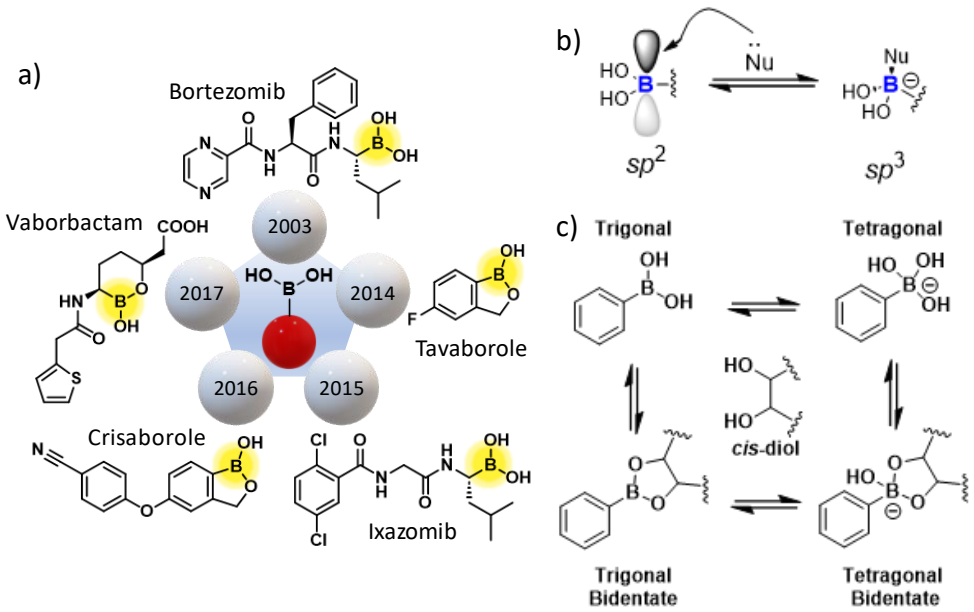


Figure 1.7: a) FDA approved boronic acid drugs; b) Interconversion of boronic acid from trigonal (sp^2) to tetragonal (sp^3) form; c) Dynamic covalent interactions of boronic acid at pH 7.4 with *cis*-diols.

The growth of boron in medical applications since the past century was mainly hindered due to a misconception that it was potentially toxic. However, boron is a micronutrient and is required in mammals for bone health, and brain function, as well as for reducing cancer risk[38]. Boron supplements are also known to compensate for Vitamin D deficiency by inhibiting an enzyme responsible for the inactivation of Vitamin D.

Recent reports have also shown that the lethal dose of boron in mammals is comparable to that of table salt. This, along with the fact that boron is an essential constituent for maintaining the structural integrity of cell walls in plants, clearly reflects the importance of boron in living systems. Boronic acid is often considered the most popular form of boron, with boric acid being used as an antiseptic since the 19th century. In synthetic organic chemistry, boronic acids are indispensable for forging C-X or C-C bonds, while materials chemistry utilizes the dynamic covalent bonding properties of boron to create stimuli-responsive networks or self-healing materials[30]. The dynamic bonding of boronic acids with *cis*-diols at physiological pH requires a careful understanding of the thermodynamic factors. Trigonal boronic acid exists in an ionic equilibrium with its tetrahedral boronate anion in aqueous media (ionization constant K_a). Therefore, the binding between boronic acid and *cis*-diol exists in equilibrium processes of the trigonal planar boronic ester form (K_{trig}) and the tetrahedral boronate anion form (K_{tet}), both of which are in an ionic equilibrium process (K_a'). Primarily, a drop in pH disrupts the association between boronic acid and *cis*-diol. In general, boronate anion has a better tendency to associate with diols ($K_{tet} > K_{trig}$). Therefore, a medium having pH higher than the p*K*a of boronic acid favors the association of *cis*-diol with the majority of boronic acid-containing compounds owing to their existence in the anionic form. However, the optimal binding pH may not always be higher than the p*K*a of boronic acid. The p*K*a of boronic acid is referred to as the pH at which a boronic acid exists in 50% boronate anion form (sp^3) of the boronic acid. Apart from these basics, the thermodynamic stability of boronate ester and its association ability to boronate ester rely on multiple parameters, such as the nature of diols, substitution in boronic acid derivatives, solvent systems, and many other conditions under considerable debate. Yet, it is sometimes troublesome to envisage the relationship among association constant, pH, and p*K*a while considering boronate ester formation. As a notable phenomenon, cyclic boronate esters are usually more acidic than free boronic acid due to ring strain. Thus, neutral BA (trigonal planar, sp^2) species dominate in anionic form (tetrahedral, sp^3) upon complex formation with *cis*-diol, in general, at physiological pH 7.4.

The affinity of boronic acid binding can be further finetuned by tinkering with the chemical environment of the moiety, particularly for aryl boronic acid systems. The presence of an aldehyde or acetyl group *ortho* to the boronic acid moiety (2-formylphenyl boronic acid 2FPBA and 2-acetylphenyl boronic acid 2APBA, respectively) enables the system to form thermodynamically stable imine effectively conjugates with a K_d of 10-15 mM. Functioning as a Lewis acid, the boronic acid catalyzes the imine formation's condensation step, accelerating the reaction (Figure 1.8a). This strategy has been employed to conjugate with lysyl-PG and therefore label bacterial membranes[39], peptide cyclization[40] intracellular protein delivery[41] and increasing residence time of antimicrobial peptide (AMP) conjugated 2APBA on bacterial membranes for better PET and SPECT imaging[42,43], among other commendable applications[44–46]. The *ortho*-carbonyl boronic acid systems have also been used to conjugate with α-nucleophiles such as hydrazine[47,48], hydrazide[48], oxime[48], sulfonyl hydrazide[49] and semicarbazide[50] (Figure 1.8a), and also β-amino nucleophiles such as cysteamine and cysteine[51–54] (Figure 1.8b) with sufficiently fast kinetics of the order of 10^3 M^{-1} s^{-1} all proceeding via iminoboronate formation. Remarkably, the presence of an electron-withdrawing group such as nitrogen (Wulff-boronates, Figure 1.8c) serves to lower the pK_a (~5.3) of the boronic acid group, thus facilitating tight *cis*-diol conjugation at physiological pH. o-Aminomethyl boronic acids, more commonly known as Wulff-boronates, have been extensively studied and marketed by the James group for recognizing sugars (Eversense and Glysure)[55–58]. Benzoxaboroles, which can be easily synthesized from 2FPBA precursor, possess similar sugar-binding properties due to their higher ring strain, which can be released after binding to sugar and

converting to the tetrahedral sp^3 form (Figure 1.8d). Aryl boronic acids have also been explored to conjugate with salicylhydroxamic acids to form salicylhydroxamic-boronate (SHAB) conjugates[30,59] with fast kinetics and a ready dissociation at pH <5 (Figure 1.8e). This has been found to be applicable to affinity chromatography[60] and synthesizing HIV barrier gels[61,62]. ^{10}B containing compounds are also reported to have a high neutron fluence. As such, irradiation of a concentrated compound in cancer cells with low-energy thermal neutrons results in a localized thermal fission, resulting in cell death. This radio-therapeutic technique finds application in boron neutron capture therapy (BNCT), which treats malignant tumors[63,64] (Figure 1.8f).

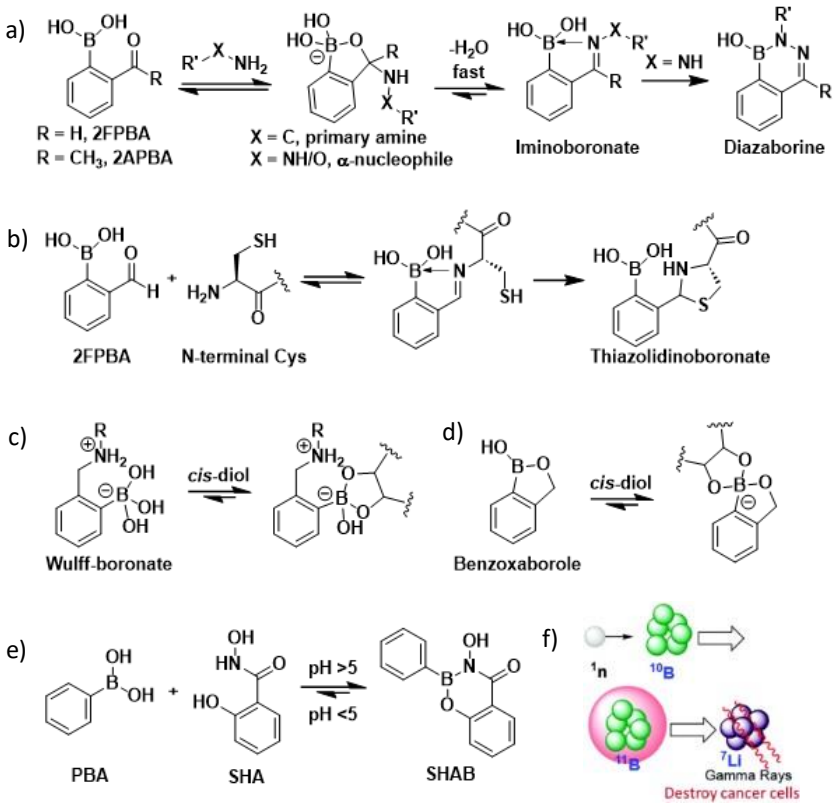


Figure 1.8: a) Dynamic conjugation of 2FPBA and 2APBA to iminoboronates and diazaborines at pH 7.4; b) Conjugation of 2FPBA with Cys at pH 7.4; Superior recognition of *cis*-diols by c) Wulff-boronates and d) Benzoxaboroles at physiological pH; e) SHAB conjugation; f) Principle of BNCT (Reproduced from Chatterjee et. al., *Chem. Comm*., 2021.

The versatile iminoboronate chemistry has motivated scientists to further classify it based on thermodynamic considerations:

Spontaneously reversible: First demonstrated by Bandyopadhyay et al. in 2015[39], the iminoboronate formation with primary amines is reversible with a K_d of ~10 mM under physiological conditions, and a simple dilution below this concentration range can lead to its easy reversibility without requiring any small molecule inhibitors. The reversibility of the iminoboronate conjugate enables the final product to exchange with other amines, a desirable feature for biomolecular recognition, avoiding off-target modification. Irreversible covalent drug design sometimes suffers from off-target modification, a shortcoming resolved by the exchangeable feature of iminoboronates, leading to the suitably applied reversible covalent inhibition of a protein target. A suitable application was shown by Bandyopadhyay et al. in selectively labelling S. *aureus* by targeting the primary amines attached to Lys-PG expressed on the cell surface by a cationic peptide tagged with 2APBA[39] and eight years later, a similar strategy was employed to improve the PET and SPECT imaging

of bacterial infection in mice models[42,43]. Akçay et al. showed that small molecule inhibitors of Mcl-1, an oncogenic target when attached to 2APBA or 2FPBA, exhibited 20-50 times better binding affinity when compared to control molecules since the design utilized capturing of non-catalytic Lys residues in the binding pocket to enhance significantly the binding affinity[65]. It has also been used in stimuli-responsive macrocyclization of peptides in biological systems[40].

Kinetically labile: The synergistic influence of multiple structural effects between boronic acids and nucleophiles leads to the formation of thermodynamically stable but kinetically labile products. The Gois group showed the formation of a stable three-component assembly with hydrazones of various phenyl glyoxylic acids, N, N- diethyl salicylaldehyde and phenylboronic acid (PBA) derivatives to create boronic acid salicylidenehydrazone (BASHY) dyes for applications in live cell imaging techniques[66]. To address the high reversibility of hydrazides with 2FPBA/APBA, the Bane group employed α-amino hydrazides where the additional coordination provided by the α-amine led to the formation of monomeric diazaborines with hydrazides[67]. Another remarkable application was to direct the iminoboronate formation with specific amino acid side chains by installing a 'boron hot spot' (BHS) at desired sites on a peptide[68]. 4-Carboxamido 3-hydroxyquinolin-2(1H)-one with a glycyl linker at the amide NH_2 (4-Gly-3HQ), having a hydroxyl group that can coordinate to boron atom, was chosen as the BHS with an affinity of K_a = 699 ± 2 M^{-1} for PBA. The BHS, having a maleimide group, was installed on a peptide containing an N-terminal Cys residue, a step that also allowed its reversibility in the presence of glutathione. They reported the reaction of BHS–Cys with 2-FPBA (20 mM, pH 7) proceeded with a K_d of 1.7 × 10^{-2} mM at 37 °C. An important aspect of this work is that iminoboronate formation only occurs with the N-terminal amine, even in the presence of other amine residues.

Repurposed: In this system, the dynamic iminoboronate is trapped through intramolecular nucleophilic addition. In other words, the final product does not possess the iminoboronate structure. This approach generates kinetically inert, thermodynamically stable heterocycles under a pH range of 6–10. Starting was the thiazolidinoboronate formation with N-terminal Cys and 2FPBA, which created a stable tricyclic core by the proximity-driven attack of sulfur on the formed iminoboronate[69]. It is stable to the presence of other thiol moieties and pretty stable to dissociation at slightly acidic pHs. The iminoboronate products of hydrazine and semicarbazide are kinetically trapped, but now, they are formed by stable, aromatic diazaborine. The aromatic ring is formed by the NH group adjacent to the imine nitrogen, forming a bond with the boron atom in a second dehydration step, a step which largely depends on the electron density of the α-N. Diazaborines are therapeutically very important scaffolds with antibacterial activity, as demonstrated by Kocak et al. in developing antibacterial and anti-quorum-sensing material[70]. Research has also been shown to determine the absolute configuration of mono-alcohols[71], sensing semicarbazides as a toxic food contaminant[72] and as a conjugate in live-cell fluorescent sensing applications[30].

The sensitivity of boronic to external oxidative conditions (Figure 1.9a) has also been utilized to design self-immolative linkers in applications as a prodrug. In a cancer cell, the concentration of reactive oxygen species (ROS) such as HOO˙ and reactive nitrogen species (RNS) such as NO˙ can be 10-40 times higher than a normal cell. As such, boronic acid gets oxidized to the corresponding phenol, a feature that has been utilized in designing prodrugs containing a phenol moiety, such as SN-38[34] (Figure 1.9b). This strategy was also applied to design fluorescent turn-on probes upon self-immolation (Figure 1.9c).

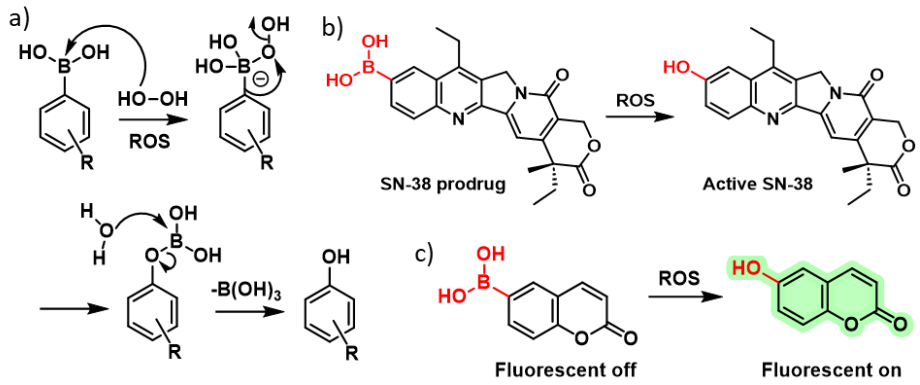


Figure 1.9: a) Mechanism for the oxidation of phenyl boronic acids to phenols in the presence of ROS, and its utility in designing b) prodrug of SN-38, and c) fluorescent turn-on probe.

Thus, given the enormous applications of the boronic acid moiety[73] and the shortage of methodologies available for the installation of diverse boronic acid derivatives on biomolecules, it is essential to develop facile and economic strategies for rapid and smooth chemoselective installation of boronic acid derivatives. This will help foster both research in academia and industry likewise.

1.6 *RATIONALE AND LAYOUT OF THE THESIS*

Given the diverse nature of bonding offered by boronic acid moieties spanning from medicinal chemistry to material science, it seems reasonable to endow peptides with the power of boronic acids to obtain boronopeptides as 'smart acting peptides'. In a biological setting, such peptides can be easily finetuned to obtain desirable biorthogonal traits, such as meeting the needs for covalent drug discovery, mimicking bioisosters, drug delivery, and post-translational modifications[74]. To better cope with the imminent demand, advanced synthetic strategies must evolve while considering the atom economy and the sustainability of such reactions. In Chapter 2, we explore thiol-ene click chemistry as a tool to help in the late-stage installation of different boronic acid derivatives on biologically active peptides. Chapter 3 is divided into two parts which delve into the application of synthesized boronopeptides – Part A deals with the capturing of sialyl-glycans on cancer cells, while Part B talks about the advantage of attaching a covalent boronic acid probe to an antimicrobial peptide to aid in radiotherapy applications. Finally, Chapter 4 covers a very effortless installation of methyl boronic acid on peptides and proteins by a straightforward and rapid S_N2 reaction in water at physiological conditions.

1.7 *REFERENCES*

(1) *Emil Fischer | Nobel Prize Winner, Organic Chemist & Synthesizer | Britannica*. https://www.britannica.com/biography/Emil-Fischer (accessed 2024-05-20).

(2) Union, I. U. of P. and A. C. International Union of Pure Joint Commission on Biochemical Nomenclature * Nomenclature and Symbolism For. *Pure & Applied Chemistry* **1984**, *56* (5), 595–624.

(3) Suresh Babu, V. V. One Hundred Years of Peptide Chemistry. *Resonance* **2011**, *16* (7), 640–647. https://doi.org/10.1007/s12045-011-0071-7.

(4) Akbarian, M.; Khani, A.; Eghbalpour, S.; Uversky, V. N. Bioactive Peptides: Synthesis, Sources, Applications, and Proposed Mechanisms of Action. *Int J Mol Sci* **2022**, *23* (3). https://doi.org/10.3390/ijms23031445.

(5) Arrowsmith, S.; Wray, S. Oxytocin: Its Mechanism of Action and Receptor Signalling in the Myometrium. *J Neuroendocrinol* **2014**, *26* (6), 356–369. https://doi.org/https://doi.org/10.1111/jne.12154.

(6) Scott, D. A.; Best, C. H. The Preparation of Insulin. *Ind Eng Chem* **1925**, *17* (3), 238–240. https://doi.org/10.1021/ie50183a004.

(7) Muttenthaler, M.; King, G. F.; Adams, D. J.; Alewood, P. F. Trends in Peptide Drug Discovery. *Nat Rev Drug Discov* **2021**, *20* (4), 309–325. https://doi.org/10.1038/s41573-020-00135-8.

(8) Wang, L.; Wang, N.; Zhang, W.; Cheng, X.; Yan, Z.; Shao, G.; Wang, X.; Wang, R.; Fu, C. Therapeutic Peptides: Current Applications and Future Directions. *Signal Transduct Target Ther* **2022**, *7* (1), 48. https://doi.org/10.1038/s41392-022-00904-4.

(9) Bergmann, M.; Zervas, L. Über Ein Allgemeines Verfahren Der Peptid-Synthese. *Berichte der deutschen chemischen Gesellschaft (A and B Series)* **1932**, *65* (7), 1192–1201. https://doi.org/https://doi.org/10.1002/cber.19320650722.

(10) Anderson, G. W.; Mcgregor, A. C. T-Butyloxycarbonylamino Acids and Their Use in Peptide Synthesis. *J. Am. Chem. Soc* **1957**, *79*, 6180.

(11) Carpino, L. A.; Han, G. Y. The 9-Fluorenylmethoxycarbonyl Function, a New Base-Sensitive Amino-Protecting Group Sir: *J. Am. Chem. Soc* **1970**, *92*, 5748–5749.

(12) Sheehan, J. C.; Hess, G. P. A NEW METHOD OF FORMING PEPTIDE BONDS. *J. Am. Chem. Soc* **1955**, *77*, 1067–1068.

(13) Han, S. Y.; Kim, Y. A. Recent Development of Peptide Coupling Reagents in Organic Synthesis. *Tetrahedron* **2004**, *60* (11), 2447–2467. https://doi.org/10.1016/j.tet.2004.01.020.

(14) König, W.; Geiger, R. Eine Neue Methode Zur Synthese von Peptiden: Aktivierung Der Carboxylgruppe Mit Dicyclohexylcarbodiimid Unter Zusatz von 1-Hydroxy-Benzotriazolen. *Chem Ber* **1970**, *103* (3), 788–798. https://doi.org/https://doi.org/10.1002/cber.19701030319.

(15) Dourtoglou, V.; Ziegler, J.-C.; Gross, B. L'hexafluorophosphate de O-Benzotriazolyl-N,N-Tetramethyluronium: Un Reactif de Couplage Peptidique Nouveau et Efficace. *Tetrahedron Lett* **1978**, *19* (15), 1269–1272. https://doi.org/https://doi.org/10.1016/0040-4039(78)80103-8.

(16) Albeiicio, F.; Chinchilla, R.; Dodsworth, D. J.; Nájera, C. NEW TRENDS IN PEPTIDE COUPLING REAGENTS. *Org Prep Proced Int* **2001**, *33* (3), 203–303. https://doi.org/10.1080/00304940109356592.

(17) Merrifield, R. B. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *J Am Chem Soc* **1963**, *85* (14), 2149–2154. https://doi.org/10.1021/ja00897a025.

(18) Erdélyi, M.; Gogoll, A. Rapid Microwave-Assisted Solid Phase Peptide Synthesis. *Synthesis (Stuttg)* **2002**, No. 11, 1592–1596. https://doi.org/10.1055/s-2002-33348.

(19) Hartrampf, N.; Saebi, A.; Poskus, M.; Gates, Z. P.; Callahan, A. J.; Cowfer, A. E.; Hanna, S.; Antilla, S.; Schissel, C. K.; Quartararo, A. J.; Ye, X.; Mijalis, A. J.; Simon, M. D.; Loas, A.; Liu, S.; Jessen, C.; Nielsen, T. E.; Pentelute, B. L. Synthesis of Proteins by Automated Flow Chemistry. *Science* **2020**, *368* (6494), 980–987. https://doi.org/10.1126/science.abb2491.

(20) Bradley, C. V; Williams, D. H.; Hanley, M. R. Peptide Sequencing Using the Combination of Edman Degradation, Carboxypeptidase Digestion and Fast Atom Bombardment Mass Spectrometry. *Biochem Biophys Res Commun* **1982**, *104* (4), 1223–1230. https://doi.org/https://doi.org/10.1016/0006-291X(82)91381-X.

(21) Chandarana, C.; Juwarwala, I.; Shetty, S.; Bose, A. Peptide Drugs: Current Status and Treatment for the Treatment of Various Diseases. *Current Drug Research Reviews*. 2024, pp 1–14. https://doi.org/http://dx.doi.org/10.2174/0125899775295960240406073630.

(22) Boutureira, O.; Bernardes, G. J. L. Advances in Chemical Protein Modification. *Chem Rev* **2015**, *115* (5), 2174–2195. https://doi.org/10.1021/cr500399p.

(23) Chavda, V. P.; Ajabiya, J.; Teli, D.; Bojarska, J.; Apostolopoulos, V. Tirzepatide, a New Era of Dual-Targeted Treatment for Diabetes and Obesity: A Mini-Review. *Molecules* **2022**, *27* (13), 4315. https://doi.org/10.3390/molecules27134315.

(24) Kim, C. H.; Axup, J. Y.; Schultz, P. G. Protein Conjugation with Genetically Encoded Unnatural Amino Acids. *Curr Opin Chem Biol* **2013**, *17* (3), 412–419. https://doi.org/10.1016/j.cbpa.2013.04.017.

(25) Chalker, J. M.; Davis, B. G. Chemical Mutagenesis: Selective Post-Expression Interconversion of Protein Amino Acid Residues. *Curr Opin Chem Biol* **2010**, *14* (6), 781–789. https://doi.org/10.1016/j.cbpa.2010.10.007.

(26) deGruyter, J. N.; Malins, L. R.; Baran, P. S. Residue-Specific Peptide Modification: A Chemist's Guide. *Biochemistry* **2017**, *56* (30), 3863–3873. https://doi.org/10.1021/ACS.BIOCHEM.7B00536.

(27) Koniev, O.; Wagner, A. Developments and Recent Advancements in the Field of Endogenous Amino Acid Selective Bond Forming Reactions for Bioconjugation. *Chem Soc Rev* **2015**, *44* (15), 5495–5551. https://doi.org/10.1039/c5cs00048c.

(28) Reddy, N. C.; Kumar, M.; Molla, R.; Rai, V. Chemical Methods for Modification of Proteins. *Org Biomol Chem* **2020**. https://doi.org/10.1039/d0ob00857e.

(29) Marek Zegota, M.; Wang, T.; Seidler, C.; Yuen Wah Ng, D.; Ling Kuan, S.; Weil, T. "Tag and Modify" Protein Conjugation with Dynamic Covalent Chemistry. **2018**. https://doi.org/10.1021/acs.bioconjchem.8b00358.

(30) Chatterjee, S.; Anslyn, E. V.; Bandyopadhyay, A. Boronic Acid Based Dynamic Click Chemistry: Recent Advances and Emergent Applications. *Chem Sci* **2021**, *12* (5), 1585–1599. https://doi.org/10.1039/d0sc05009a.

(31) Messner, K.; Vuong, B.; Tranmer, G. K. The Boron Advantage: The Evolution and Diversification of Boron's Applications in Medicinal Chemistry. *Pharmaceuticals* **2022**, *15* (3), 264. https://doi.org/10.3390/ph15030264.

(32) Kane, R. C.; Bross, P. F.; Farrell, A. T.; Pazdur, R. Velcade®: U.S. FDA Approval for the Treatment of Multiple Myeloma Progressing on Prior Therapy. *Oncologist* **2003**, *8* (6), 508–513. https://doi.org/10.1634/theoncologist.8-6-508.

(33) Gupta, A. K.; Versteeg, S. G. Tavaborole – a Treatment for Onychomycosis of the Toenails. *Expert Rev Clin Pharmacol* **2016**, *9* (9), 1145–1152. https://doi.org/10.1080/17512433.2016.1206467.

(34) António, J. P. M.; Russo, R.; Carvalho, C. P.; Cal, P. M. S. D.; Gois, P. M. P. Boronic Acids as Building Blocks for the Construction of Therapeutically Useful Bioconjugates. *Chem Soc Rev* **2019**, *48* (13), 3513–3536. https://doi.org/10.1039/c9cs00184k.

(35) Andersen, K. A.; Smith, T. P.; Lomax, J. E.; Raines, R. T. Boronic Acid for the Traceless Delivery of Proteins into Cells. *ACS Chem Biol* **2016**, *11* (2), 319–323. https://doi.org/10.1021/acschembio.5b00966.

(36) Ghosh, P. Boronic Acid-Linked Cell-Penetrating Peptide for Protein Delivery. *ACS Omega* **2024**. https://doi.org/10.1021/acsomega.3c09689.

(37) Longwitz, L.; Leveson-Gower, R. B.; Rozeboom, H. J.; Thunnissen, A.-M. W. H.; Roelfes, G. Boron Catalysis in a Designer Enzyme. *Nature* **2024**, *41* (November 2023), 39–41. https://doi.org/10.1038/s41586-024-07391-3.

(38) Chatterjee, S.; Tripathi, N. M.; Bandyopadhyay, A. The Modern Role of Boron as a 'Magic Element' in Biomedical Science: Chemistry Perspective. *Chemical Communications* **2021**, *57* (100), 13629–13640. https://doi.org/10.1039/D1CC05481C.

(39) Bandyopadhyay, A.; McCarthy, K. A.; Kelly, M. A.; Gao, J. Targeting Bacteria via Iminoboronate Chemistry of Amine-Presenting Lipids. *Nat Commun* **2015**, *6*, 1–9. https://doi.org/10.1038/ncomms7561.

(40) Bandyopadhyay, A.; Gao, J. Iminoboronate-Based Peptide Cyclization That Responds to PH, Oxidation, and Small Molecule Modulators. *J Am Chem Soc* **2016**, *138* (7), 2098–2101. https://doi.org/10.1021/jacs.5b12301.

(41) Rong, G.; Zhou, X.; Hong, J.; Cheng, Y. Reversible Assembly of Proteins and Phenolic Polymers for Intracellular Protein Delivery with Serum Stability. *Nano Lett* **2024**. https://doi.org/10.1021/acs.nanolett.4c00937.

(42) Bhatt Mitra, J.; Chatterjee, S.; Kumar, A.; Bandyopadhyay, A.; Mukherjee, A. Integrating a Covalent Probe with Ubiquicidin Fragment Enables Effective Bacterial Infection Imaging. *RSC Med Chem* **2022**, *13* (10), 1239–1245. https://doi.org/10.1039/D2MD00190J.

(43) Bhatt Mitra, J.; Chatterjee, S.; Kumar, A.; Khatoon, E.; Chandak, A.; Rakshit, S.; Bandyopadhyay, A.; Mukherjee, A. Expanding a Peptide-Covalent Probe Hybrid for PET Imaging of S. Aureus Driven Focal Infections. *EJNMMI Radiopharm Chem* **2024**, *9* (1), 25. https://doi.org/10.1186/s41181-024-00252-4.

(44) Cal, P. M. S. D.; Vicente, J. B.; Pires, E.; Coelho, A. V.; Veiros, L. F.; Cordeiro, C.; Gois, P. M. P. Iminoboronates: A New Strategy for Reversible Protein Modification. *J Am Chem Soc* **2012**, *134* (24), 10299–10305. https://doi.org/10.1021/ja303436y.

(45) Faustino, H.; Silva, M. J. S. A.; Veiros, L. F.; Bernardes, G. J. L.; Gois, P. M. P. Iminoboronates Are Efficient Intermediates for Selective, Rapid and Reversible N-Terminal Cysteine Functionalisation. *Chem Sci* **2016**, *7* (8), 5052–5058. https://doi.org/10.1039/c6sc01520d.

(46) Reja, R. M.; Wang, W.; Lyu, Y.; Haeffner, F.; Gao, J. Lysine-Targeting Reversible Covalent Inhibitors with Long Residence Time. *J Am Chem Soc* **2022**, *144* (3), 1152–1157. https://doi.org/10.1021/jacs.1c12702.

(47) Dilek, O.; Lei, Z.; Mukherjee, K.; Bane, S. Rapid Formation of a Stable Boron–Nitrogen Heterocycle in Dilute, Neutral Aqueous Solution for Bioorthogonal Coupling Reactions. *Chemical Communications* **2015**, *51* (95), 16992–16995. https://doi.org/10.1039/C5CC07453C.

(48) Bandyopadhyay, A.; Gao, J. Iminoboronate Formation Leads to Fast and Reversible Conjugation Chemistry of A-Nucleophiles at Neutral PH. *Chemistry – A European Journal* **2015**, *21* (42), 14748–14752. https://doi.org/10.1002/chem.201502077.

(49) Chowdhury, A.; Chatterjee, S.; Kushwaha, A.; Nanda, S.; Dhilip Kumar, T. J.; Bandyopadhyay, A. Sulfonyl Diazaborine 'Click' Chemistry Enables Rapid and Efficient Bioorthogonal Labeling**. *Chemistry – A European Journal* **2023**, *29* (41). https://doi.org/10.1002/chem.202300393.

(50) Cambray, S.; Bandyopadhyay, A.; Gao, J. Fluorogenic Diazaborine Formation of Semicarbazide with Designed Coumarin Derivatives. *Chemical Communications* **2017**, *53* (93), 12532–12535. https://doi.org/10.1039/c7cc07389e.

(51) Bandyopadhyay, A.; Cambray, S.; Gao, J. Fast and Selective Labeling of N-Terminal Cysteines at Neutral PH: Via Thiazolidino Boronate Formation. *Chem Sci* **2016**, *7* (7), 4589–4593. https://doi.org/10.1039/c6sc00172f.

(52) Zheng, M.; Haeffner, F.; Gao, J. N-Terminal Cysteine Mediated Backbone-Side Chain Cyclization for Chemically Enhanced Phage Display †. *Chem Sci* **2022**. https://doi.org/10.1039/d2sc03241d.

(53) Li, K.; Weidman, C.; Gao, J. Dynamic Formation of Imidazolidino Boronate Enables Design of Cysteine-Responsive Peptides. *Org Lett* **2018**, *20* (1), 20–23. https://doi.org/10.1021/acs.orglett.7b03116.

(54) Gao, J.; Li, K.; Wang, W. Fast and Stable N-Terminal Cysteine Modification via Thiazolidino Boronate Mediated Acyl Transfer. *Angewandte Chemie International Edition* **2020**, 1–6. https://doi.org/10.1002/anie.202000837.

(55) Lacina, K.; Skládal, P.; James, T. D. Boronic Acids for Sensing and Other Applications - a Mini-Review of Papers Published in 2013. *Chem Cent J* **2014**, *8* (1), 1–17. https://doi.org/10.1186/s13065-014-0060-5.

(56) James, T. D.; Phillips, M. D.; Shinkai, S. Boronic Acids in Saccharide Recognition. **2006**. https://doi.org/10.1039/9781847557612.

(57) Sun, X.; Zhai, W.; Fossey, J. S.; James, T. D. Boronic Acids for Fluorescence Imaging of Carbohydrates. *Chemical Communications* **2016**, *52* (17), 3456–3469. https://doi.org/10.1039/c5cc08633g.

(58) Sun, X.; Chapin, B. M.; Metola, P.; Collins, B.; Wang, B.; James, T. D.; Anslyn, E. V. The Mechanisms of Boronate Ester Formation and Fluorescent Turn-on in Ortho-Aminomethylphenylboronic Acids. *Nat Chem* **2019**, *11* (9), 768–778. https://doi.org/10.1038/s41557-019-0314-x.

(59) Stolowitz, M. L.; Ahlem, C.; Hughes, K. A.; Kaiser, R. J.; Kesicki, E. A.; Li, G.; Lund, K. P.; Torkelson, S. M.; Wiley, J. P. Phenylboronic Acid-Salicylhydroxamic Acid Bioconjugates. 1. A Novel Boronic Acid Complex for Protein Immobilization. *Bioconjug Chem* **2001**, *12* (2), 229–239. https://doi.org/10.1021/bc0000942.

(60) Wiley, J. P.; Hughes, K. A.; Kaiser, R. J.; Kesicki, E. A.; Lund, K. P.; Stolowitz, M. L. Phenylboronic Acid - Salicylhydroxamic Acid Bioconjugates. 2. Polyvalent Immobilization of Protein Ligands for Affinity Chromatography. *Bioconjug Chem* **2001**, *12* (2), 240–250. https://doi.org/10.1021/bc000086l.

(61) Jay, J. I.; Shukair, S.; Langheinrich, K.; Hanson, M. C.; Cianci, G. C.; Johnson, T. J.; Clark, M. R.; Hope, T. J.; Kiser, P. F. Modulation of Viscoelasticity and HIV Transport as a Function of PH in a Reversibly Crosslinked Hydrogel. *Adv Funct Mater* **2009**, *19* (18), 2969–2977. https://doi.org/10.1002/adfm.200900757.

(62) Mahalingam, A.; Jay, J. I.; Langheinrich, K.; Shukair, S.; McRaven, M. D.; Rohan, L. C.; Herold, B. C.; Hope, T. J.; Kiser, P. F. Inhibition of the Transport of HIV in Vitro Using a PH-Responsive Synthetic Mucin-like Polymer System. *Biomaterials* **2011**, *32* (33), 8343–8355. https://doi.org/10.1016/j.biomaterials.2011.05.001.

(63) Lamba, M.; Goswami, A.; Bandyopadhyay, A. A Periodic Development of BPA and BSH Based Derivatives in Boron Neutron Capture Therapy (BNCT). *Chemical Communications* **2021**, *57* (7), 827–839. https://doi.org/10.1039/D0CC06557A.

(64) Xu, H.; Liu, J.; Li, R.; Lin, J.; Gui, L.; Wang, Y.; Jin, Z.; Xia, W.; Liu, Y.; Cheng, S.; Yuan, Z. Novel Promising Boron Agents for Boron Neutron Capture Therapy: Current Status and Outlook on the Future. *Coord Chem Rev* **2024**, *511*, 215795. https://doi.org/https://doi.org/10.1016/j.ccr.2024.215795.

(65) Akçay, G.; Belmonte, M. A.; Aquila, B.; Chuaqui, C.; Hird, A. W.; Lamb, M. L.; Rawlins, P. B.; Su, N.; Tentarelli, S.; Grimster, N. P.; Su, Q. Inhibition of Mcl-1 through Covalent Modification of a Noncatalytic Lysine Side Chain. *Nat Chem Biol* **2016**, *12* (11), 931–936. https://doi.org/10.1038/nchembio.2174.

(66) Santos, F. M. F.; Rosa, J. N.; Candeias, N. R.; Carvalho, C. P.; Matos, A. I.; Ventura, A. E.; Florindo, H. F.; Silva, L. C.; Pischel, U.; Gois, P. M. P. A Three-Component Assembly Promoted by Boronic Acids Delivers a Modular Fluorophore Platform (BASHY Dyes). *Chemistry - A European Journal* **2016**, *22* (5), 1631–1637. https://doi.org/10.1002/chem.201503943.

(67) Gu, H.; Chio, T. I.; Lei, Z.; Staples, R. J.; Hirschi, J. S.; Bane, S. Formation of Hydrazones and Stabilized Boron–Nitrogen Heterocycles in Aqueous Solution from Carbohydrazides and Ortho-Formylphenylboronic Acids. *Org Biomol Chem* **2017**, *15* (36), 7543–7548. https://doi.org/10.1039/C7OB01708A.

(68) Russo, R.; Padanha, R.; Fernandes, F.; Veiros, L. F.; Corzana, F.; Gois, P. M. P. Engineering Boron Hot Spots for the Site-Selective Installation of Iminoboronates on Peptide Chains. *Chemistry – A European Journal* **2020**, *26* (66), 15226–15231. https://doi.org/10.1002/chem.202002675.

(69) Bandyopadhyay, A.; Cambray, S.; Gao, J. Fast and Selective Labeling of N-Terminal Cysteines at Neutral PH via Thiazolidino Boronate Formation. *Chem Sci* **2016**, *7* (7), 4589–4593. https://doi.org/10.1039/C6SC00172F.

(70) Kocak, G.; Cicek, H.; Ceylan, Ö.; Bütün, V. Antimicrobial and Anti-Quorum-Sensing Properties and Paint Film Usage of Novel Diazaborine-Based Copolymers. *J Appl Polym Sci* **2019**, *136* (1). https://doi.org/10.1002/app.46907.

(71) Shimo, S.; Takahashi, K.; Iwasawa, N. 1,2-Dihydro-1-Hydroxy-2,3,1-Benzodiazaborine Bearing an Acridine Moiety as a Circular Dichroism Probe for Determination of Absolute Configuration of Mono-

Alcohols. *Chemistry - A European Journal* **2019**, *25* (15), 3790–3794. https://doi.org/10.1002/chem.201900350.

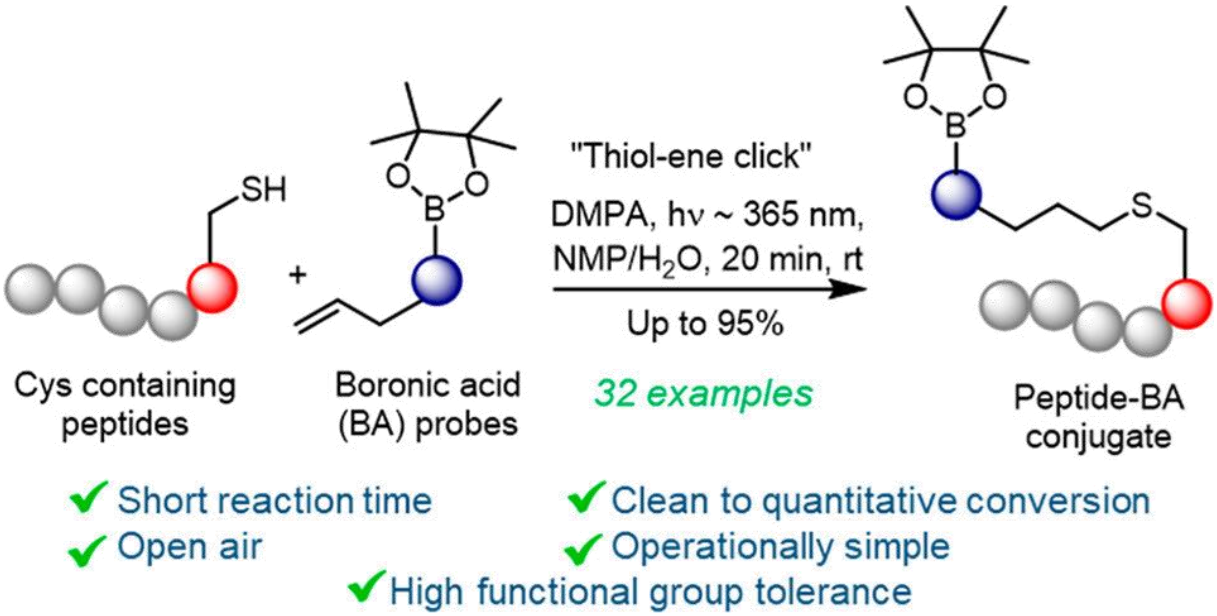
(72) Kong, X.; Sun, Z.; Liu, S.; Luo, X.; Li, G. Determination of Semicarbazide in Foodstuffs by HPLC with Fluorescence Detection Using 2-Formylphenylboronic Acid as Derivatization Reagent. *Chromatographia* **2019**, *82* (7), 1051–1058. https://doi.org/10.1007/s10337-019-03734-9.

(73) Ball, Z. T. Protein Substrates for Reaction Discovery: Site-Selective Modification with Boronic Acid Reagents. *Acc Chem Res* **2019**, *52* (3), 566–575. https://doi.org/10.1021/acs.accounts.8b00626.

(74) Ahangarpour, M.; Kavianinia, I.; Harris, P. W. R.; Brimble, M. A. Photo-Induced Radical Thiol–Ene Chemistry: A Versatile Toolbox for Peptide-Based Drug Design. *Chem Soc Rev* **2021**, *50* (2), 898–944. https://doi.org/10.1039/D0CS00354A.

CHAPTER 2

CYSTEINE-SELECTIVE INSTALLATION OF FUNCTIONALLY DIVERSE BORONIC ACID PROBES ON PEPTIDES



This work has been published in Organic Letters (a journal of American Chemical Society) and can be cited as "**Chatterjee, S.**; Bandyopadhyay, A. Cysteine-Selective Installation of Functionally Diverse Boronic Acid Probes on Peptides. *Org Lett* **2023**, *25* (13), 2223–2227."

"Electron poor but chemically rich element with a storied past continues to surprise with a surge of new developments." *-Stephen Ritter, Boron chemistry branches out, 2016*

2.1 *INTRODUCTION*

The past two decades have witnessed an enormous research interest linked to boronic acid (BA) in chemical biology, medicinal chemistry, and biomaterial designing[1]. BA is an attractive moiety due to its wide-range reactivity and interactivity with biologically important nucleophiles[2]. BA has been harnessed as bioisosteres of carboxylic acids to improve pharmacological behavior[3]. Aromatic BA having an *ortho*-carbonyl group has lately been the subject of extraordinary claims in bioconjugation and biorthogonal chemistry[4]. Nevertheless, chemistries enabling the late-stage incorporation of functionally variable BA building blocks onto peptides remain limited. Disconnected epitope fragments from protein sequences often require non-native modifications to help restore or improve their native activity. Incorporating non-natural BA-containing amino acids or building blocks during the conventional peptide synthesis is fairly expensive, often requiring gram-scale synthesis of the BA modified amino acid. In our hands we have observed that BA compounds are difficult to purify by conventional normal phase column purification by silica gel, resulting in only about 60-70% recovery from the crude. In this respect, the direct late-stage site-selective modification of native peptides at the reactive amino acid side chains has revealed numerous advantages since it requires a small amount of the reagent to modify the peptides, thus providing a more economical process. Of special mention, the main essence of a strategic modification to the peptide chain is to solely aid the peptide's biological activity and not interfere with its native activity.

The main hindrance in designing such a modification strategy is to circumvent the various reactive natural amino acid side chains (sulfhydryl, primary amine, alkyl and aromatic alcohol, alkyl carboxylic acid, guanidium and imidazole) and target a chemically distinct residue (chemoselective) and if possible, only at desired sites (site-selective), provided multiple copies of the residue of interest is available. By nature's grace, a difference in the pK_a values of the side-chains provides the perfect opportunity for chemoselective modification under physiological conditions. Cysteine (Cys) has about <2% natural abundance on protein surfaces[5] thus presenting a lucrative site for chemical modifications with the sulfhydryl group possessing a high nucleophilicity (pK_a ~ 8.5) at pH 7.4 where the other functional groups such as amines, alcohols, acids are generally less reactive[6]. Several different conjugation strategies have been developed with Cys residue[7] such as maleimide, thiol-ene/yne, alkylation, 2-cyanobenzothiazoles, and Michael addition (Figure 2.1a). The thiol-ene click chemistry, a radical mediated reaction between an alkene and a sulfhydryl group is a well explored reaction in chemical biology due to the fast kinetics and compatibility to physiological conditions[8,9]. The generated thioether linkages are very stable compared to maleimide conjugation which can undergo exchange reaction[10] or the Michael addition which can reportedly revert to the starting materials by a retro-Michael reaction[11]. Some popular drugs such as Montelukast and Ranitidine contain this thioether linkage[12]. Mechanistically, a photoinitiator such as DMPA (2,2-Dimethoxy-2-phenylacetophenone) or the water-soluble VA-044 (2,2'-azobis[2-(2-imidazolin-2-yl)propane]- dihydrochloride) helps in the generation of an initial thiyl radical (by formation of a carbon-centered radical, Figure 2.1b) in the presence of UV-irradiation (365 nm)[13,14]. Thiyl radicals are pretty stable in water because the bond dissociation energy of alkyl R-SH (~87 kcal/mol) is weaker than H-OH (~120 kcal/mol)[15]. The generated radicals then react with an alkenyl moiety in an anti-Markovnikov fashion to form a carbon-centered radical thioether, leading to the desired product via a chain-propagation step through another sulfhydryl molecule (Figure 2.1c). Competing side-

reactions such as homopolymerization or chain termination via disulfide formation can be minimized by a wise choice of the alkene handle, which should form an extremely stable radical species[6,16,17]. For the first time, we explored the compatibility and efficacy of the radical-mediated thiol−ene "click" reaction for the installation of the chemically diverse BA repertoire onto synthetically programmed cysteine (Cys) residue(s) in biologically active peptides. Further, a few synthetic boronopeptides were customized for peptide cyclization, conjugation, and alteration of BA functionality.

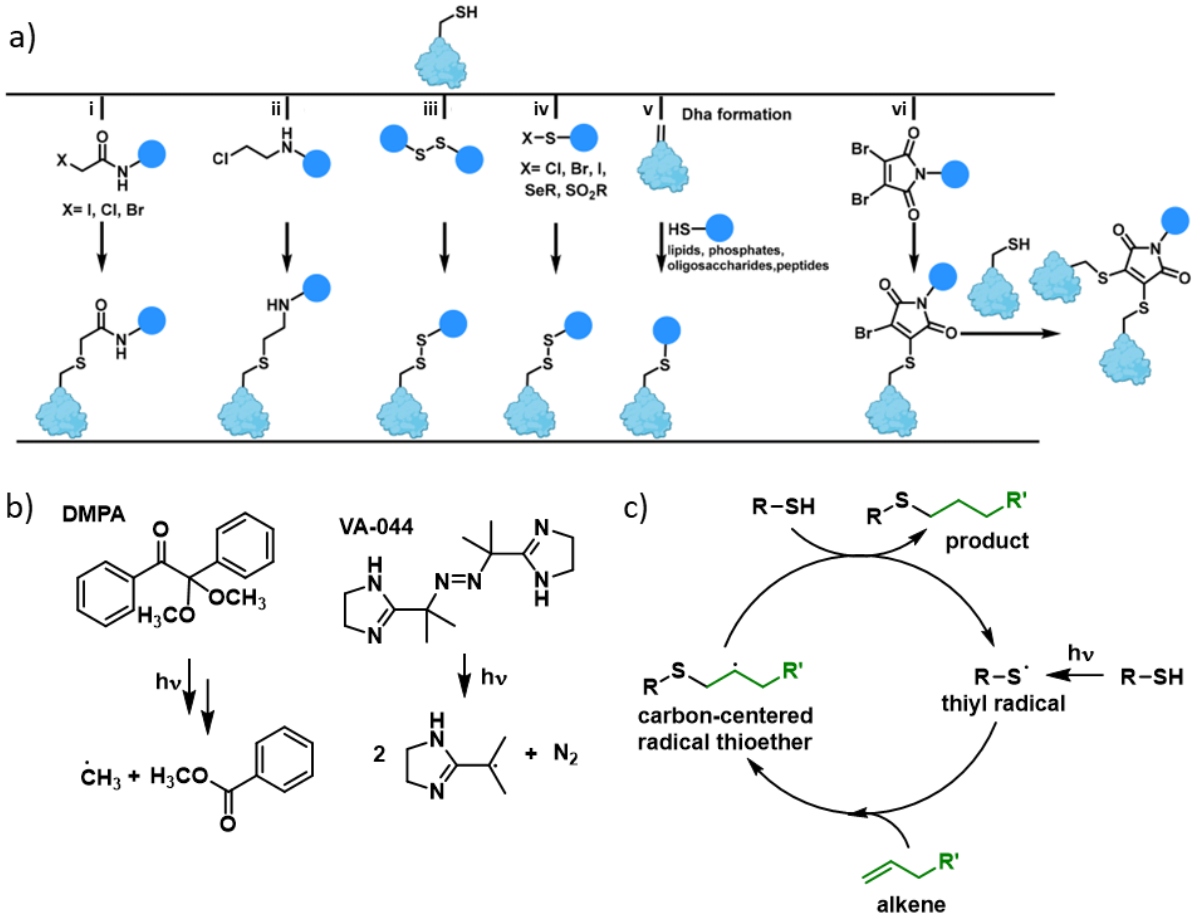


Figure 2.1: a) Site-selective methods for the modification of cysteine by (i) α-halocarbonyls, (ii) alkylation through a S_N2 reaction with haloalkyl reagents, (iii) sulfhydryl-disulfide (SH/SS) exchange, (iv) disulfide formation at cysteine, (v) dehydroalanine (Dha) formation and its modification with functionalized thiols, (vi) dibromomaleimides as useful scaffolds for peptide modification (Reproduced with permission from Ahangarpour et al., *Org. Biomol. Chem.*, 2023); b) Mechanism of radical thiol-ene chemistry; c) Photoinitiator DMPA and VA-044 generates a carbon centered radical upon irradiation.

2.2 *LITERATURE REVIEW*

In 2017, the Baran group[18] demonstrated a decarboxylative borylation method to synthesize α-amino boronic acid on biologically active peptides (Figure 2.2a). The Davis group[19] (Figure 2.2b) and Roelfes et al.[20] recently demonstrated Cu(II) catalyzed facile incorporation of minimal boryl moiety via a Dha precursor, presenting peptide and protein. This "tag-and- modify"[21] approach has been applied in multiple chemo- and regioselective modifications of peptides and proteins. The thiol-ene click chemistry has been widely explored for post-

translational modifications to peptides and proteins, such as lipidation, glycosylation, and peptide stapling[6]. We have examined the existing methods[22,23], which have not yet explored the installation scope for biomedically important and recently developed BA building blocks (vide infra) that are of prodigious importance in multidisciplinary research[1,2,4,24]. In addition, conjugation of BA via the known azide–alkyne "click" reaction is shown to degrade BA, and adding fluoride ions to this reaction can recover a moderate yield for specific BA derivatives[25].

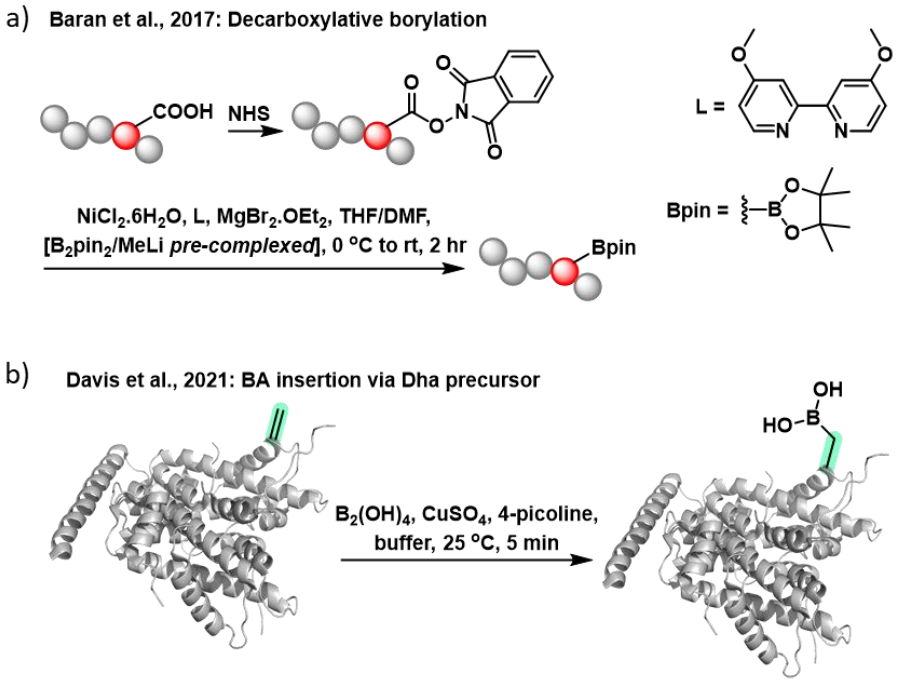


Figure 2.2: Previous work on late-stage installation of BA by a) Decarboxylative borylation; b) Dha precursor.

2.3 *RESULTS AND DISCUSSION*

2.3.1 *RATIONALE FOR CHOOSING THE DIFFERENT BORONIC ACID DERIVATIVES*

To demonstrate the utility of the thiol-ene click chemistry, we probed eight chemically diverse and widely used BA building blocks to demonstrate their installation on Cys-containing peptides. BA1 is a simple derivative of phenyl boronic acid (PBA), the most popular BA derivative, installed on peptides or materials to explore sugar recognition[26]. BA2 is an alkyl BA reported to be more stable in oxidative conditions than aryl BA[27]. A Wulff-boronate BA4[28] and benzoxaborole BA7[29] were also synthesized, showing superior binding to *cis*-diols at physiological pH. 2FPBA and 2APBA derivatives, named BA5 and BA6, respectively[1], were also synthesized, demonstrating a multitude of reactions with primary amines (forming iminoboronates) and α-nucleophiles (forming diazaborines). BA3 is a simple aryl boronic acid with a vinyl alkene handle that is readily available commercially. BA8 is a recently reported pyridine BA for selective recognition of sialic acid[30]. The structures of BA1-8 have been shown in Figure 2.3.

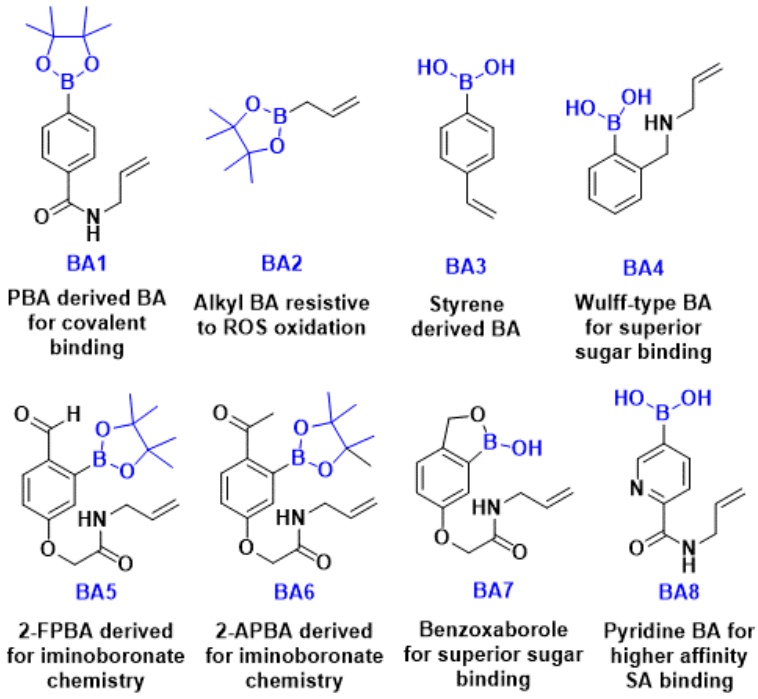


Figure 2.3: Synthesized boronic acid derivatives and the rationale for choosing them.

2.3.2 *SYNTHESIS OF BORONIC ACID DERIVATIVES AND PEPTIDES*

BA2 and BA3 were procured from commercially available sources, and the other six BA derivatives were synthesized in the lab with allylamine as the alkene handle of choice. BA1 and BA8 were synthesized from commercially available carboxylic acid precursors and coupled with allylamine through simple amide coupling. The reductive amination of Schiff-base formed between 2-formylphenyl boronic acid (2FPBA) and allylamine yielded BA4. BA5 and BA6 were obtained by a five-step synthetic route, starting from 2,4-dihydroxy benzaldehyde and 2,4-dihydroxy acetophenone, respectively. The intermediate steps involved alkylation of acid precursor, triflation, Miyaura borylation, ester hydrolysis, and finally amide coupling with allylamine. Reduction of aldehyde in BA5 by $NaBH_4$ yielded the benzoxaborole derivative, BA7. Notably, the synthetic BA building blocks have been stable at room temperature for months, demonstrating these reagents' easy accessibility.

All the peptides except GSH were synthesized by SPPS using the standard Fmoc-deprotection strategy. Eight biologically active peptides with variable lengths and amino acid compositions were chosen for the study. Some native sequences have been altered by adding a Cys residue to enable the thiol−ene click reaction. Oxytocin has been shown as the linear protected form. The native names for all the peptide sequences have been used for ease of understanding.

2.3.3 *OPTIMIZATION OF THE THIOL-ENE CLICK REACTION*

We first optimized the thiol−ene reaction between N-acetyl cysteine (AcCys) and BA1 at an equal stoichiometry (~200 mM each) by employing a few standard solvents reported earlier for the thiol−ene click reaction[13,14]. In our study, N-methyl-2-pyrrolidone (NMP) in an acidified condition with 1% TFA offered the cleanest conversion with a quantitative conversion in 20 min (Figure 2.4a). 2,2-Dimethoxy-2-phenyl-acetophenone (DMPA) was used as

the radical initiator with the help of a BLB light source (~365 nm) to perform the reactions. We noticed that free amine impurities in NMP inhibit thiol−ene reactions to some extent. The presence of free amine impedes the conversion yield in thiol−ene reaction, which has been well studied[31]. Thereby, the addition of TFA certainly improved the efficiency of reactions by quenching the free amines. The reaction profile appears extremely clean to deliver a quantitative product when analyzed by HPLC, ^{1}H NMR analysis, along with mass spectrometry (Figure 2.4b). The optimized conditions also delivered a quantitative conjugation of other BA derivatives (BA2−8) with AcCys, confirmed by the disappearance of the alkene protons in crude ^{1}H NMR studies. The UV-active products were further purified by RP-HPLC to verify their identity via ^{1}H NMR and HRMS.

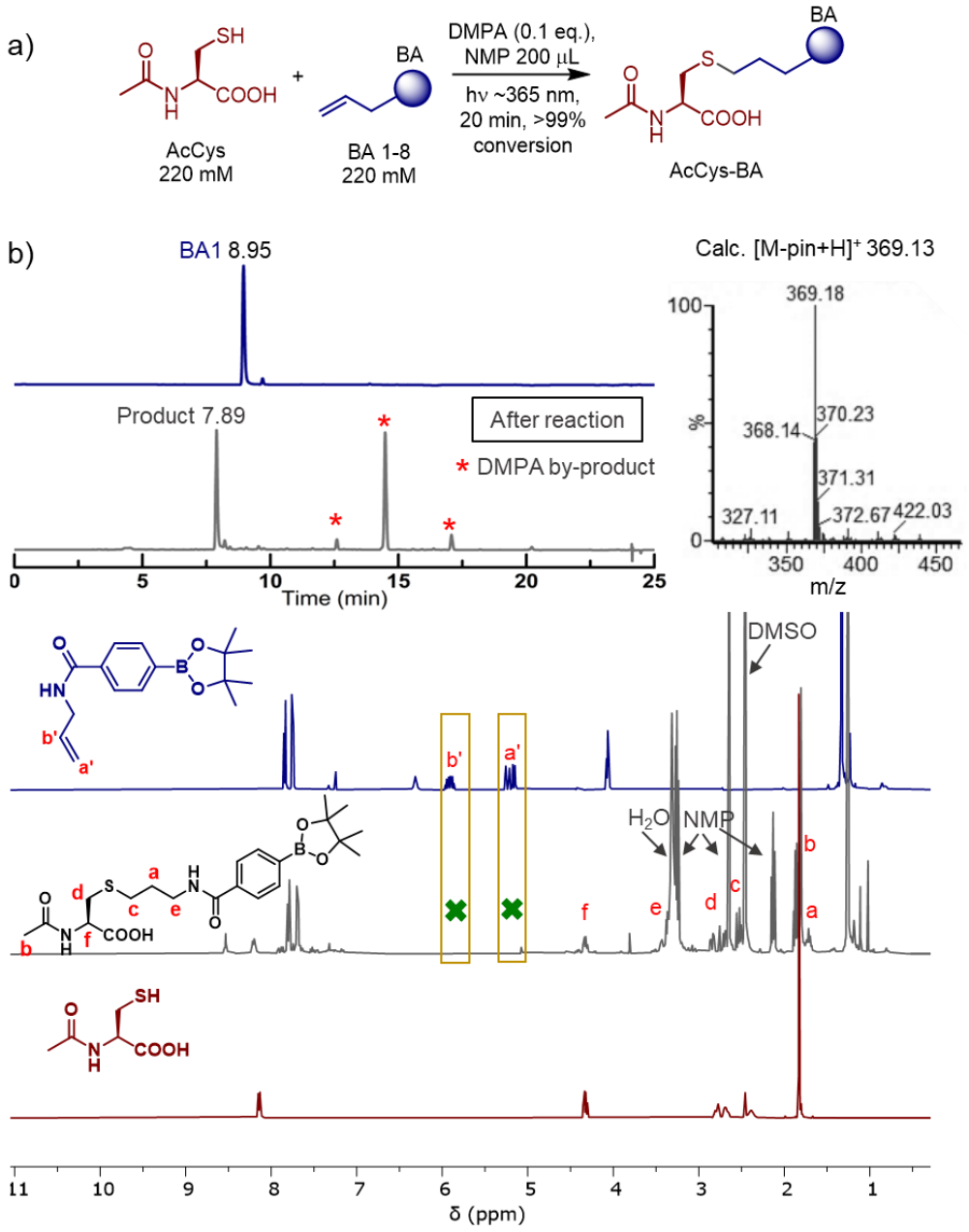


Figure 2.4: a) General scheme of the thiol−ene reaction with acetylcysteine (AcCys); b) HPLC, mass, and NMR (in DMSO-d_6) data of crude product obtained via thiol−ene reaction between AcCys and BA1. NMR datum highlights quantitative product formation assigned by the disappearance of the alkene group's chemical shifts (marked a′ and b′).

The reaction conditions were further optimized with oxytocin as the model peptide since it possesses a well-balanced hydrophilic and hydrophobic profile. A 1:1 mixture of H_2O:NMP (1% TFA) was chosen as the solvent of choice to improve the conversion yield of peptides. Very gratifyingly, a smooth reaction was observed at 10 mM concentration of the peptide. In previous literature reports, the use of additives such as DTT or TCEP was recommended to improve the yield[13]. Still in our hand, we observed that the use of TCEP induces the oxidation of sulfur in the formed thioether linkage. Since disulfide bonds, if formed (by coupling of two thiyl radicals in the chain termination step), are reported to regenerate the thiyl radicals by homolytic fission[32], we omitted the use of disulfide-reducing additives. While further optimizing the peptide concentration, it was found that the reaction was gradually compromised at lower concentrations (<10 mM) of peptide, even with an extended reaction time. For example, we observed an excellent reaction conversion (~93%) in a control reaction at 10 mM of oxytocin, with 1.2 eq. BA1 and 0.8 eq. DMPA. But the conversion was reduced to ~76% and ~53% when 3 mM and 1 mM of oxytocin were used, respectively. After numerous optimization trials, we retrieved ~73% conversion at 1 mM peptide concentration using ~3 eq. BA1 and ~2.66 eq. of DMPA, denoting a 20% higher conversion. Unfortunately, these optimization conditions did not improve the reaction outcome for hydrophobic peptides at lower concentrations (LeuEnk and p53 at 1−3 mM). We postulated that the reaction stops in the chain propagation step at a lower thiol concentration (<8−10 mM). Thus, complete consumption of the starting material was not observed. Therefore, we postulated that increasing overall thiol concentration in the reaction (total ~10 mM) by adding a thiol-containing small molecule along with a hydrophobic peptide (~1−3 mM) may restore the higher conversion, hoping that the additional thiol will help in the chain propagation of the generated thiyl radicals. To our delight, the synthesis of LeuEnk-BA1 conjugate was improved from ~40% to ~80% by using GSH as an additive with an optimized stoichiometry of BA1 (1.2 eq. of the total thiol concentration) and DMPA (0.8 eq. of the total thiol concentration). This protocol helped us improve the reaction conversion for hydrophobic peptide sequences at relatively lower concentrations (~1−3 mM). GSH-BA conjugation occurred independently, and the desired material was easily separable from hydrophobic peptide conjugates by RP-HPLC.

With the optimized reaction conditions in hand, we successfully installed all BAs on GSH with excellent conversion, indicated via HPLC and mass spectra. Additionally, the UV-active products were purified and analyzed by ^{1}H NMR to confirm their identity. The reaction conditions offer clean reactions with relatively polar peptides at about 10 mM concentration. A near complete peptide conjugation was obtained in all BAs except for the aminomethyl boronic acid derivative BA4. Among the 32 examples (Figure 2.5), BA4 attributed moderate conversion, possibly due to the neighboring amine group that interfered with the reaction. Remarkably, the optimized reaction conditions efficiently labeled bivalent Cys residues programmed on CysLeuEnk with BAs. We observed the oxidation of thioether linkage (S=O) in very low amounts (up to 10%) as the only side product in some BA conjugations. Neither oxidation of BA to phenol[27] nor any side reactions with BA moiety were observed in LC−MS, which highlights the better prospects of this reaction for BA installation compared to the azide−alkyne reaction with BA substrate[25]. Altogether, the systematic optimization of the radical "thiol−ene" reaction in our studies offered improved compatibility and higher conversion to prepare a boron−peptide conjugate.

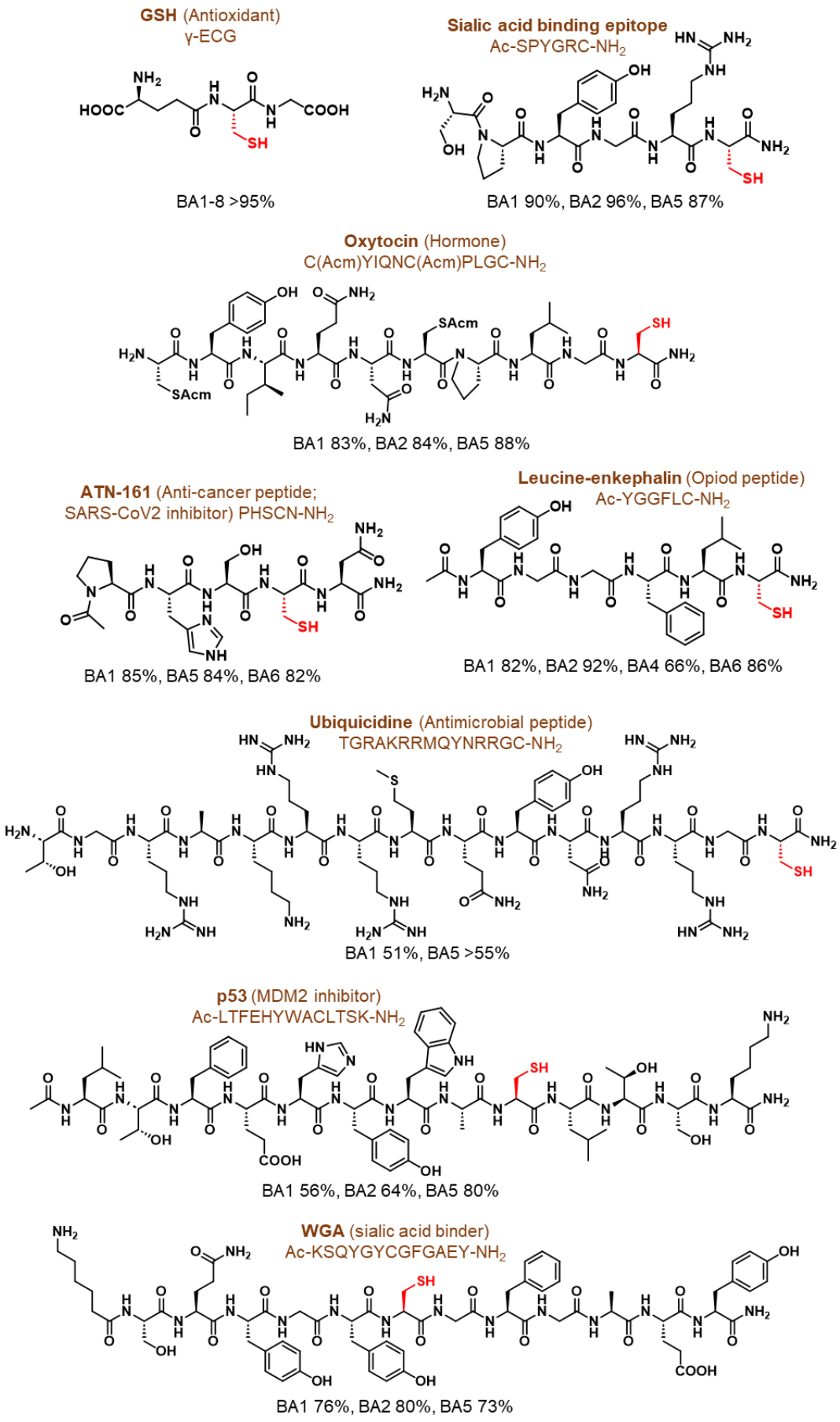
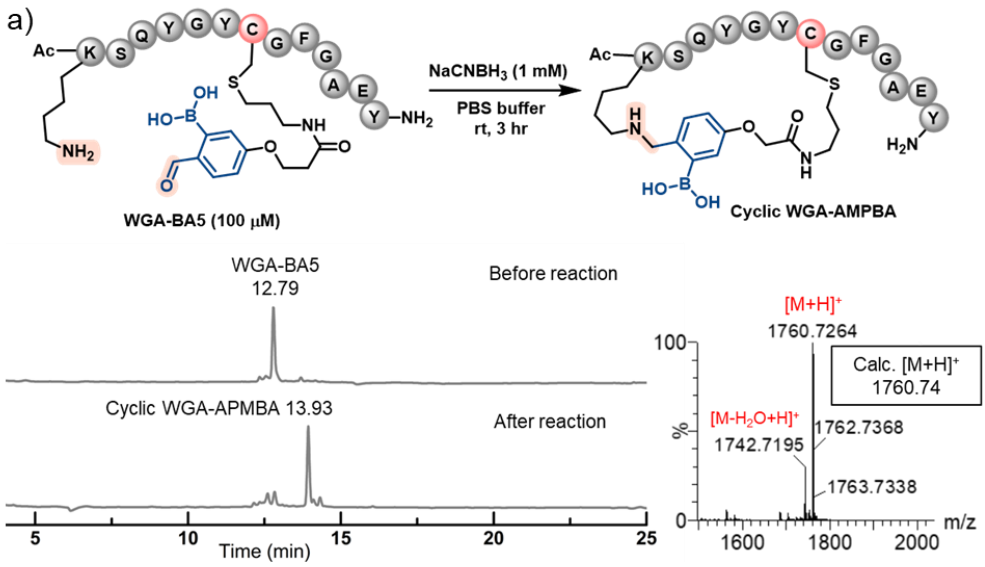


Figure 2.5: General scheme of BA installation on peptides via the thiol−ene click reaction. The chemical structure of bioactive peptides and their conversion in thiol−ene reaction with the mentioned BA derivatives.

2.3.4 *CHEMICAL MODULES OF THE SYNTHESIZED BORONOPEPTIDES*

The merged field of boron chemistry and peptides has become prevalent in medicinal chemistry[1,33] and therapeutic applications. Hence, we projected utilizing this boronopeptide preparation method and modulating it with the hope to gain ample opportunities in the biomedical arena through conjugation, cyclization, and boron-derivative modification. We demonstrated such advantages by performing four prototypical reactions.

1. The BA5 moiety in WGA peptide was readily cyclized with N-terminus lysine (Lys) residue via an iminoboronate intermediate at physiological pH, previously reported by Bandyopadhyay et al.[34]. The iminoboronate formation between BA5 and Lys residues is a spontaneous process, and reducing agents such as $NaCNBH_3$ can trap the resulting imine bond in the same pot at physiological pH (Figure 2.6a).
2. We were inquisitive about whether the I_2 mediated solution phase peptide cyclization, through disulfide formation, could be carried out with post BA-modified peptide. A clean and quantitative conversion with Acm protected oxytocin proves the formation of disulfide, in which the BA moiety was untouched by I_2 (Figure 2.6b).
3. A quantitative transformation of ATN-BA5 to ATN-BA7 privileged the facile alteration of BA functional moiety. The 2-FPBA moiety on the ATN peptide was effortlessly converted to a benzoxaborole by the treatment of $NaBH_4$, supported by HPLC and mass spectra (Figure 2.6c). This transformation opens up an alternative route for benzoxoborole-based peptide drug design.
4. Lastly, an efficient conjugation of Oxy-BA5 with a sulfonyl hydrazide derivative was shown in PBS buffer at physiological pH (Figure 2.6d). This conjugated product, a diazaborine scaffold, is reported to display potent antibacterial properties[35]. In this regard, we anticipate diazaborine formation can be used to conjugate drugs and aid antibacterial pharmacophores in peptides.



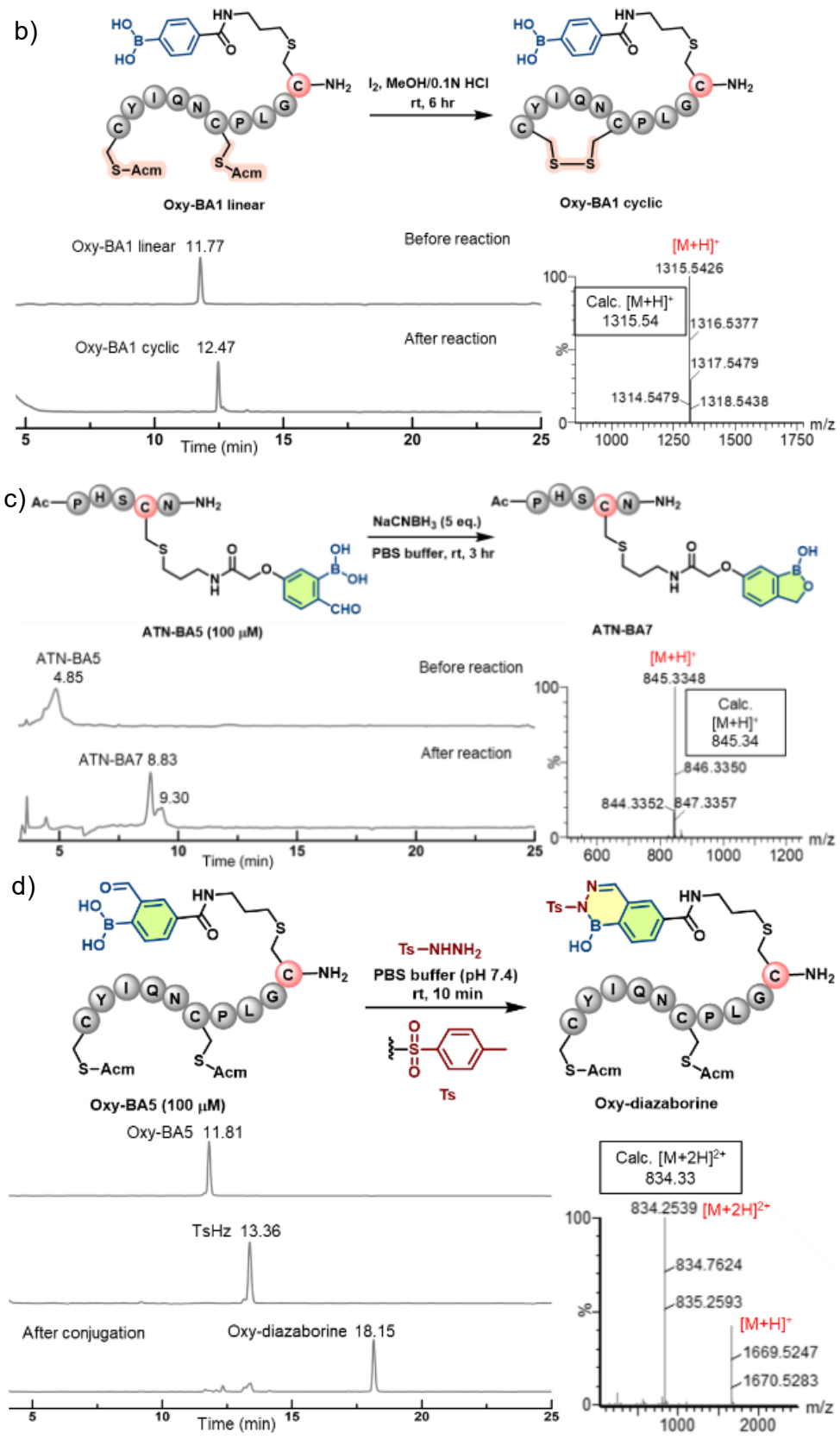


Figure 2.6: Chemical modules of boronopeptides a) Cyclization of WGA-BA5 via iminoboronate formation (AMPBA stands for aminomethyl phenylboronic acid); (b) Cyclization of oxytocin via disulfide bond formation to generate native bioactive structure; (c) Functional group conversion of 2FPBA group to benzoxaborole moiety on peptide; (d) Conjugation of 2FPBA functionalized peptide with α-nucleophiles for diazaborine formation.

2.3.5 *STABILITY STUDIES OF SYNTHESIZED BORONOPEPTIDES*

Given the huge importance of boronopeptides in medicinal chemistry applications, it was imperative to carry out stability studies of the synthesized boronopeptides. Serum stability of some synthesized boronopeptides was carried out following previously reported protocol[36]. LeuEnk and p53 were considered as controls, and nearly 80% of the peptide was found to be degraded after 24 h. Corresponding BA modified peptides showed higher stability to some extent, compared to native peptides, with about the improvement of ~40% half-life. . Similar results were obtained with SPYG and GSH boronopeptides (Figure 2.7). In our hands, boronopeptides were found to be bench stable in a lyophilized form at room temperature for 3−4 days.

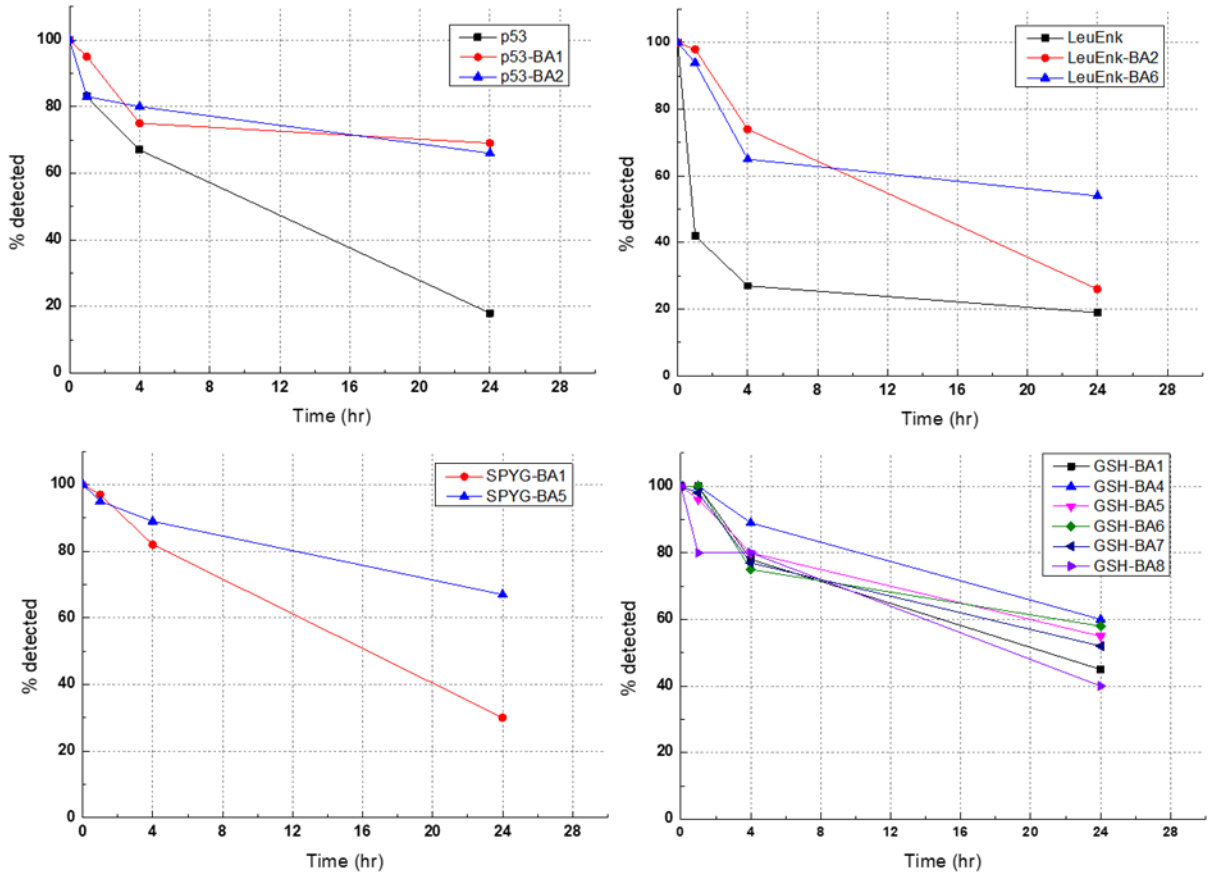


Figure 2.7: Serum stability studies of some BA-modified and unmodified peptides shown, which were carried out at intervals of 0, 1, 4, and 24 hr.

2.4 *SUMMARY AND CONCLUSION*

To summarize, we have shown the installation of functionally versatile BA building blocks onto bioactive native peptides of variable lengths via a thiol−ene radical reaction. Of special mention, we observed the reaction efficiency relies on the light source, wattage, and the irradiation distance from the light source to the reaction pot. In our hands, reported optimization conditions leveraged moderate to excellent conversion without compromising BA identity. Notably, the method enables labeling BA through a traceless thioether linkage as a final product that may bring advantages in particular biomedical applications.. The optimization of the reaction also supports the reaction practicability at higher concentrations of the peptide, provided there is no solubility problem. A clean reaction profile for peptides at a lower concentration (~1 mM) is feasible by changing reactant stoichiometry and

adding GSH. Most BA derivatives are known to stick in regular silica columns, so their purification by normal phase poses a problem. On the bright side, this labeling strategy requires a few milligrams of BA counterpart, which can benefit the research and development level. Further, the work showed the customizable aspects of boronopeptides through their smooth conversion to peptide cyclization, functional alteration, and conjugation. We admit that the method may have limitations for a single-site protein modification; however, peptide reagents are worth comparing to proteins due to easier chemical modularity, accessibility, analysis, storage, and reduced immunogenicity. We believe this method of preparing boronopeptides can be advantageous in wide applications.

2.5 *EXPERIMENTAL SECTION*

2.5.1 *CHEMICALS AND GENERAL METHODS*

All the chemicals were purchased from commercial sources and used without further purification. The reagents for peptide synthesis such as Fmoc-protected amino acids, 2-(1*H*-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), and polystyrene rink amide AM resin (loading capacity 0.67 mmol/g) were purchased from Chem-Impex Int'l Inc (Wood Dale, IL).

All reactions were monitored by thin-layer chromatography (TLC) on aluminium-backed silica gel 60 F_{254}, 0.2 mm plates (Merck), and compounds were visualized under UV light (254 nm) or charred with appropriate solutions – ninhydrin charring to visualize amines, DNP charring for aldehydes, $KMnO_4$ charring for phenolic OH, and curcumin or anisaldehyde charring for boronic acids. Synthetic compounds were purified on 60-120 or 100-200 silica gel purchased from Finar.

2.5.2 *INSTRUMENTS*

NMR data were recorded on a 400 MHz Jeol JNM ECS400 NMR spectrometer. Mass-spec data were generated by a Waters XEVO G2-XS QTOF mass spectrometer. HPLC analysis and peptide purification were carried out on a Shimadzu Prominence UFLC system. UV-Vis studies were carried out on a Shimadzu UV-2600 system. Melting point studies were carried out in Stuart SMP30.

2.5.3 *GENERAL METHODS FOR LC-MS AND HPLC ANALYSIS*

LC-MS: Waters Acquity UPLC C18 (1.7 µm, 2.1×50 mm) analytical column and the mobile phase acetonitrile-water (0.1% HCOOH) with a flow rate of 0.4 mL/min were used. The gradient: gradient 5% to 35% acetonitrile in 2.5 min, then a gradient from 35% to 95% acetonitrile in 2 min followed by isocratic 95% over a min, and then column equilibration at the initial isocratic gradient 5% acetonitrile over 1.5 min. All MS analyses were performed in ESI +ve mode unless otherwise specified.

HPLC methods for reaction monitoring and peptide purity analysis (unless otherwise specified): Waters Reliant C18 (5 µm, 4.6×250 mm) analytical column and mobile phase acetonitrile-water (0.05% TFA) with a flow rate of 1 mL/min were used. The gradient used: isocratic 10% acetonitrile over 1 min, then a gradient from 10% to 50% acetonitrile in 19 min, then a gradient from 50% to 90% acetonitrile over 2 min, followed by isocratic 90%acetonitrile over 3 min, then column equilibration at the initial isocratic gradient 10% acetonitrile over 5 min.

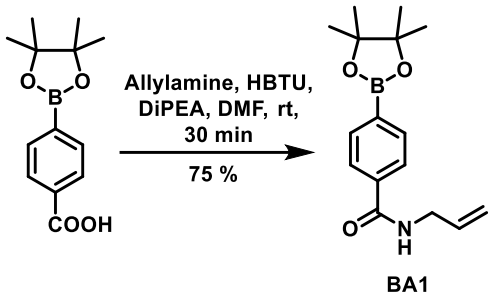
Some shifts in the retention times can be observed when comparingthe crude reaction mixture profile and the purified product profile. This observation is probably due to the presence of NMP and other solutes in the reaction

mixture. Similarly, analysis performed on different days can show similar outcomes. The reaction conversion has been calculated by comparing the peak area of the peptide, except for the GSH peptide, before and after the reaction. In the case of GSH and AcCys reactions, we have monitored the peak area of UV-active BA compounds before and after the reaction. For the other peptides, since the major contribution to UV activity at 220 nm is from the peptide backbone, the conversion was calculated from the peak area at 220 nm.

HPLC method for peptide purification through semi-preparative (unless otherwise specified): Shimadzu shim-pack GIST C18 (5 µm, 10×250 mm) semi-preparative column and mobile phase acetonitrile-water (0.05% TFA) with a flow rate 4 mL/min were used. The gradient used: isocratic 10% acetonitrile over 2 min, then a gradient from 10% to 50% acetonitrile over 26 min, followed by a gradient from 50% to 90% acetonitrile over 3 min, then isocratic 90% acetonitrile over 5 min, followed by column equilibration at the initial isocratic gradient 10% acetonitrile over 5 min.

2.5.4 *SYNTHESIS OF BORONIC ACID DERIVATIVES*

2.5.4.1 *SYNTHESIS OF BA1*



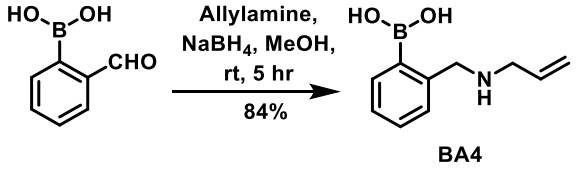
4-Carboxylphenylboronic acid pinacol ester (372 mg, 1.5 mmol), HBTU (570 mg, 1.5 mmol, 1 eq.), and DiPEA (385 µL, 2.25 mmol, 1.5 eq.) were dissolved in 1.5 mL DMF and stirred at room temperature. To this solution, allylamine (123 µL, 1.65 mmol, 1.1 eq.) was added slowly after 5 min, and the reaction mixture was further stirred for 30 min. TLC analysis indicated the completion of the reaction. The reaction mixture was acidified with chilled 50 mL 1N HCl and extracted with 2×50 mL EtOAc. The organic layers were washed with chilled 50 mL 10% $NaHCO_3$, then combined and washed with brine. Upon evaporation of the organic layer under reduced pressure, it was purified through silica gel using 30% EtOAc/Hexane to yield 320 mg of **BA1** (75% yield) as a white powder, mp: 161-163 °C.

^{1}H NMR (400 MHz, $CDCl_3$): δ 7.84 (d, *J* = 8.1 Hz, 2H), 7.75 (d, *J* = 8.1 Hz, 2H), 6.32 (s, 1H), 5.96-5.87 (m, 1H), 5.26-5.15 (m, 2H), 4.08 – 4.05 (m, 2H), 1.33 (s, 12H).

^{13}C NMR (101 MHz, $CDCl_3$): δ 167.3, 136.7, 135.0, 134.2, 126.1, 116.9, 84.2, 42.5, 29.8, 25.0.

HRMS-ESI$^+$ (*m/z*): $[M\text{-pin}+H]^+$ Calc. 206.0988, Obs. 206.0986.

2.5.4.2 *SYNTHESIS OF BA4*



BA4

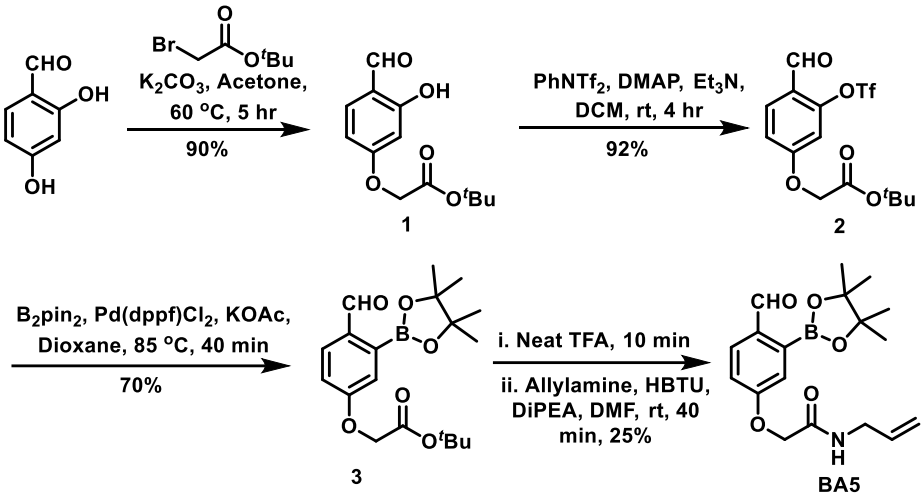
2-FPBA (75 mg, 0.5 mmol) and allylamine (38 µL, 0.5 mmol, 1 eq.) were dissolved in 3.7 mL dry MeOH and stirred at room temperature. After 5 min, some 4 Å molecular sieves (10% w/v) were added. After cooling the solution for 5 min at 0 °C, $NaBH_4$ (21 mg, 0.55 mmol, 1.1 eq.) was added and further stirred for 15 min. TLC analysis indicated consumption of 2-FPBA. The reaction mixture was filtered, and MeOH was evaporated under reduced pressure. The residue was suspended in 50 mL 1N HCl and washed with 50 mL EtOAc to remove the organic by-products. The volume of the aqueous layer was reduced to 25 mL, neutralized to pH 7 by ammonia solution and washed with 3×100 mL EtOAc. The organic layer was washed with brine, evaporated to 75 mL and kept in a beaker where the desired product crystallized overnight. The crystals were collected, dissolved in 0.1 N HCl and lyophilized to obtain 80 mg of **BA4** (84% yield) as a white powder, mp: 228-230 °C.

^{1}H NMR (400 MHz, DMSO-d_6) δ 8.66 (s, 3H), 7.81-7.79 (m, 1H), 7.45-7.35 (m, 3H), 5.92-5.81 (m, 1H), 5.47-5.37 (m, 2H), 4.24 (t, *J* = 6.0 Hz, 2H), 3.57 (q, *J* = 5.7 Hz, 2H).

^{13}C NMR (101 MHz, DMSO-d_6): δ 137.2, 136.1, 131.6, 130.6, 129.6, 128.8, 123.2, 50.6, 49.4.

HRMS-ESI^{+} (*m/z*): $[M+H]^{+}$ Calc. 192.1196, Obs. 192.1196.

2.5.4.3 *SYNTHESIS OF BA5*



*Synthesis of **1***

2,4-Dihydroxybenzaldehyde (2.2 g, 16 mmol), *tert*-butyl bromoacetate (2.5 mL, 16.8 mmol, 1.05 eq.) and K_2CO_3 (2.2g, 16 mmol, 1 eq.) were mixed in 32 mL acetone and refluxed at 60 °C for 5 hr. The reaction mixture was then cooled and filtered to remove K_2CO_3. The acetone was evaporated under vacuum and purified through silica gel using 5% EtOAc/Hexane to obtain 3.63 g of **1** (90% yield) as a white powder.

^{1}H NMR (400 MHz, $CDCl_3$): δ 11.42 (s, 1H), 9.71 (s, 1H), 7.44 (d, *J* = 8.7 Hz, 1H), 6.55 (dd, *J* = 8.6, 2.5 Hz, 1H), 6.35 (d, *J* = 2.4 Hz, 1H), 4.54 (s, 2H), 1.47 (s, 9H).

^{13}C NMR (101 MHz, $CDCl_3$): δ 194.6, 166.9, 165.0, 164.4, 135.5, 115.8, 108.6, 101.5, 83.1, 65.6, 28.1.

HRMS-ESI$^+$ (*m/z*): $[M+H]^+$ Calc. 253.1076, Obs. 253.1074.

*Synthesis of **2***

1 (2.8 g, 11.15 mmol), $PhNTf_2$ (4.38 g, 12.26 mmol, 1.1 eq.), and DMAP (136.2 mg, 1.1 mmol, 0.1 eq.) were dissolved in 20 mL dry DCM and stirred at room temperature for 5 min. Et_3N (3.1 mL, 22.3 mmol, 2 eq.) was added dropwise to the reaction mixture when the colour of the solution changed from light yellow to deep yellow. It was further stirred for 4 hr, and then DCM was evaporated. The residue was dissolved in 300 mL EtOAc, washed with 120 mL 1N HCl, and 150 mL 10% Na_2CO_3 to remove the by-product, and the combined organic layers were washed with brine. Upon evaporation of the solvent, the product was purified through silica gel using 8% EtOAc/Hexane to obtain a 3.94 g of **2** (92% yield) as a white solid.

^{1}H NMR (400 MHz, $CDCl_3$): δ 10.12 (s, 1H), 7.94 (d, *J* = 8.7 Hz, 1H), 6.99 (dd, *J* = 8.7, 2.3 Hz, 1H), 6.86 (d, *J* = 2.2 Hz, 1H), 4.61 (s, 2H), 1.48 (s, 9H).

^{13}C NMR (101 MHz, $CDCl_3$): δ 185.4, 166.4, 163.6, 151.2, 132.3, 122.5, 120.3, 117.1, 114.8, 109.0, 83.6, 66.0, 28.0.

HRMS-ESI$^+$ (*m/z*): $[M+H]^+$ Calc. 385.0569, Obs. 385.0566.

*Synthesis of **3***

2 (1.06 g, 2.76 mmol), $Pd(dppf)Cl_2$ (80.8 mg, 0.11 mmol, 0.04 eq.), B_2pin_2 (1.05 g, 4.14 mmol, 1.5 eq.), and KOAc (811.4 mg, 8.28 mmol, 3 eq.) was taken together in a 50 mL oven-dried round-bottomed flask. 16.5 mL of dry dioxane was added, and the mixture was purged with argon for 20 min. The reaction mixture was then heated at 85 °C for 40 min in a stirring condition. After completion of the reaction, as indicated by TLC, the reaction mixture was filtered through a celite bed. Upon evaporation of the solvent, the residue was purified through flash chromatography using 10% EtOAc/Hexane to obtain 1 g of **3** as a yellowish solid (70% yield).

^{1}H NMR (400 MHz, $CDCl_3$): δ 10.39 (s, 1H), 7.92 (d, *J* = 8.6 Hz, 1H), 7.26 (d, *J* = 2.7 Hz, 1H), 7.02 (dd, *J* = 8.5, 2.6 Hz, 1H), 4.59 (s, 2H), 1.47 (s, 9H), 1.36 (s, 12H).

^{13}C NMR (101 MHz, $CDCl_3$): δ 193.2, 167.3, 161.5, 135.4, 130.4, 120.4, 117.2, 84.5, 83.6, 82.9, 65.6, 28.1, 25.0.

HRMS-ESI$^+$ (*m/z*): $[M\text{-pin-}H_2O+H]^+$ Calc. 263.1085, Obs. 263.1098.

*Synthesis of **BA5***

1 g of **3** was dissolved in 5 mL cold neat TFA, which formed a yellow-coloured solution. It was then stirred at room temperature for 10 min when the colour changed from yellow to brown to deep blue. TLC analysis indicated the completion of the reaction. Residual TFA was evaporated on a rotary evaporator by diluting with 100 mL DCM, and 594 mg of grey solid was obtained by precipitation using 10% EtOAc/Hexane. The crude product was used for the next step without further purification.

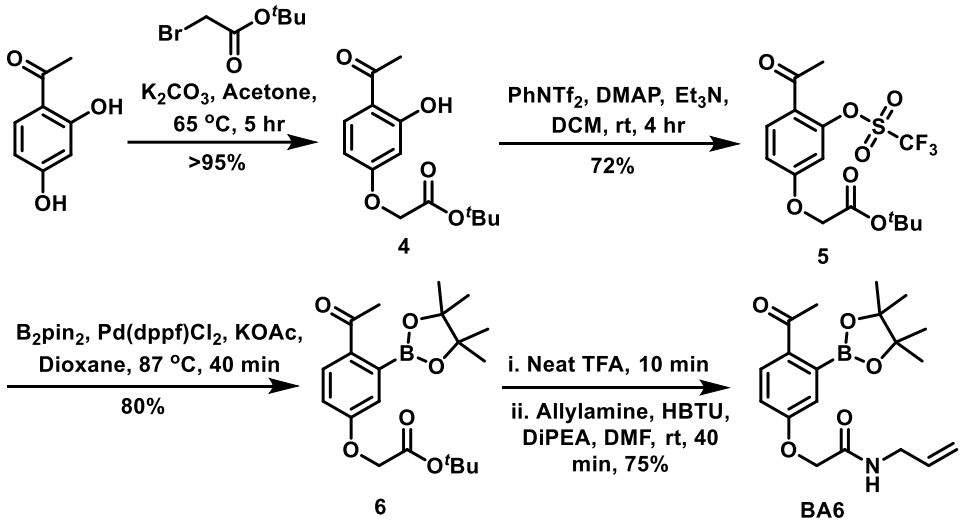
336 mg of the acid derivative (1.5 mmol) and HBTU (568.5 mg, 1.5 mmol, 1 eq.) were dissolved in 1.5 mL DMF and stirred at room temperature. To this solution, DiPEA (385 µL, 2.25 mmol, 1.5 eq.) and allylamine (123 µL, 1.65 mmol, 1.1 eq.) were added slowly with a 5 min interval in a stirring condition. The TLC analysis indicated that the reaction was completed after 40 minutes.. The reaction mixture was diluted in 200 mL EtOAc, and the organic layer was washed with 75 mL chilled 1N HCl and 100 mL chilled 5% $NaHCO_3$, followed by brine wash. After evaporation of the solvent, the residue was purified through flash chromatography using 40% EtOAc/Hexane to obtain 116 mg of **BA5** (35% yield) as a yellowish gummy solid, which solidified on refrigeration, mp: 118-120 °C.

^{1}H NMR (400 MHz, $CDCl_3$): δ 10.40 (s, 1H), 7.96 (d, *J* = 8.6 Hz, 1H), 7.34 (d, *J* = 2.7 Hz, 1H), 7.05 (dd, *J* = 8.7, 2.7 Hz, 1H), 6.64 (s, 1H), 5.88-5.80 (m, 1H), 5.24 – 5.14 (m, 2H), 4.59 (s, 2H), 3.99-3.96 (m, 2H), 1.38 (s, 12H).

^{13}C NMR (101 MHz, $CDCl_3$): δ 193.0, 167.2, 160.5, 135.9, 133.5, 130.8, 121.0, 117.1, 116.6, 84.7, 67.2, 41.5, 38.8, 25.0.

HRMS-ESI$^+$ (*m/z*): $[M\text{-pin-}H_2O+H]^+$ Calc. 246.0940, Obs. 246.0938.

2.5.4.4 *SYNTHESIS OF BA6*



*Synthesis of **4***

2', 4'-Dihydroxyacetophenone (1.97 g, 13 mmol), *tert*-butyl bromoacetate (2.1 mL, 14.3 mmol, 1.1 eq.) and K_2CO_3 (1.8 g, 13 mmol, 1 eq.) were suspended in 26 mL acetone and refluxed at 65 °C for 5 hr. The reaction mixture was cooled and filtered to remove K_2CO_3. The organic layer was concentrated and filtered through silica gel using 8% EtOAc/Hexane to obtain 3.38 g of **4** (98% yield) as an off-white powder.

^{1}H NMR (400 MHz, $CDCl_3$): δ 7.81 (d, *J* = 8.8 Hz, 1H), 6.91 (dd, *J* = 8.8, 2.5 Hz, 1H), 6.79 (d, *J* = 2.4 Hz, 1H), 4.56 (s, 2H), 2.56 (s, 3H), 1.46 (s, 9H).

^{13}C NMR (101 MHz, $CDCl_3$): δ 202.8, 167.1, 165.1, 164.2, 132.6, 114.6, 107.9, 101.7, 83.0, 65.5, 28.1, 26.4.

HRMS-ESI$^+$ (*m/z*): $[M+H]^+$ Calc. 267.1232, Obs. 267.1229.

*Synthesis of **5***

4 (3.3 g, 12.4 mmol), $PhNTf_2$ (4.87 g, 13.6 mmol, 1.1 eq.), and DMAP (151.5 mg, 1.24 mmol, 0.1 eq.) were dissolved in 23 mL dry DCM and stirred at room temperature for 5 min. Et_3N (3.5 mL, 24.8 mmol, 2 eq.) was added dropwise to the reaction mixture when the colour of the solution changed from light yellow to deep yellow. It was further stirred for 4 hr, and then DCM was evaporated. The residue was dissolved in 300 mL EtOAc and washed with 120 mL 1N HCl, 150 mL 10% Na2CO3. The combined organic layers were washed with brine. Upon evaporation of the solvent, the product was purified through silica gel using 10% EtOAc/Hexane to obtain a 3.54 g of **5** (72% yield) as a white solid.

^{1}H NMR (400 MHz, $CDCl_3$): δ 7.64 (d, *J* = 8.9 Hz, 1H), 6.48 (dd, *J* = 9.0, 2.5 Hz, 1H), 6.34 (d, *J* = 2.5 Hz, 1H), 4.53 (s, 2H), 2.55 (s, 3H), 1.48 (s, 9H).

^{13}C NMR (101 MHz, $CDCl_3$): δ 195.1, 166.6, 161.8, 148.4, 132.8, 124.9, 120.2, 117.1, 113.9, 109.7, 83.4, 65.9, 29.2, 28.0.

HRMS-ESI$^+$ (*m/z*): $[M+H]^+$ Calc. 399.0725, Obs. 399.0724.

*Synthesis of **6***

5 (0.995 g, 2.5 mmol), $Pd(dppf)Cl_2$ (73.2 mg, 0.1 mmol, 0.04 eq.), B_2pin_2 (0.953 g, 3.75 mmol, 1.5 eq.), and KOAc (612.5 mg, 6.25 mmol, 2.5 eq.) was taken together in a 25 mL oven-dried round-bottomed flask. 15 mL of dry dioxane was added, and the mixture was purged with argon for 20 min. The reaction mixture was then heated at 87 °C for 40 min with stirring. After completion of the reaction, as indicated by TLC, the reaction mixture was filtered through a celite bed. Upon evaporation of the solvent, the residue was purified through silica gel using 12% EtOAc/Hexane to obtain **6** as 670 mg yellowish solid (80% yield).

^{1}H NMR (400 MHz, $CDCl_3$): δ 7.76 (d, *J* = 8.6 Hz, 1H), 6.92 (d, *J* = 2.6 Hz, 1H), 6.87 (dd, *J* = 8.7, 2.7 Hz, 1H), 4.55 (s, 2H), 2.54 (s, 3H), 1.46 (s, 9H), 1.41 (s, 12H).

^{13}C NMR (101 MHz, $CDCl_3$): δ 198.6, 167.4, 161.5, 134.3, 130.7, 117.8, 114.7, 83.7, 83.2, 82.9, 77.4, 77.1, 76.8, 65.7, 28.1, 25.0, 24.6.

HRMS-ESI$^+$ (*m/z*): $[M\text{-pin-}H_2O+H]^+$ Calc. 277.1247, Obs. 277.1239.

*Synthesis of **BA6***

670 mg of **6** was dissolved in 3 mL cold neat TFA, which formed a yellow-coloured solution. It was then stirred at room temperature for 10 min. TLC analysis indicated the completion of the reaction. Residual TFA was evaporated on a rotary evaporator with 100 mL DCM, and 216 mg of yellow solid was obtained by precipitation using 10% EtOAc/Hexane, which was used for the next step without further purification.

216 mg of the acid derivative (0.91 mmol) and HBTU (345 mg, 0.91 mmol, 1 eq.) were dissolved in 1 mL DMF and stirred at room temperature. To this solution, DiPEA (232 µL, 1.36 mmol, 1.5 eq.) and allylamine (75 µL, 1 mmol, 1.1 eq.) were added with a 2 min interval in stirring condition. The reaction was completed after 40 min, as

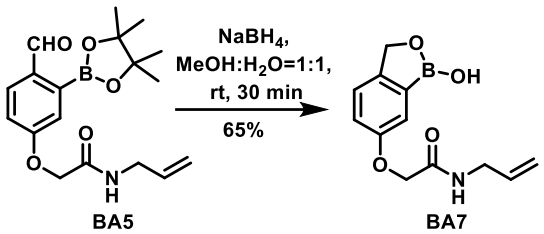
indicated by TLC analysis. The reaction mixture was diluted in EtOAc (200 mL), and the organic layer was washed with 75 mL chilled 1N HCl and 100 mL chilled 5% $NaHCO_3$, followed by brine wash. After evaporation of the solvent, the residue was purified through silica gel using 40% EtOAc/Hexane to obtain 245 mg of **BA6** (75% yield) as a yellowish gummy solid, which solidified on refrigeration, mp: 108-110 °C.

^{1}H NMR (400 MHz, 5% $CD_3OD/CDCl_3$): δ 7.81 (d, *J* = 8.6 Hz, 1H), 7.01 (d, *J* = 2.6 Hz, 1H), 6.89 (dd, *J* = 8.5, 2.6 Hz, 1H), 6.59 (brs, 1H), 5.89-5.79 (m, 1H), 5.23 – 5.14 (m, 2H), 4.57 (s, 2H), 3.99 – 3.96 (m, 2H), 2.56 (s, 3H), 1.43 (s, 12H).

^{13}C NMR (101 MHz, 5% $CD_3OD/CDCl_3$): δ 168.0, 163.2, 160.7, 134.5, 133.3, 131.1, 118.5, 116.8, 113.9, 84.0, 66.9, 38.5, 24.8.

HRMS-ESI$^+$ (*m/z*): $[M\text{-pin-}H_2O+H]^+$ Calc. 260.1094, Obs. 260.1104.

2.5.4.5 *SYNTHESIS OF BA7*



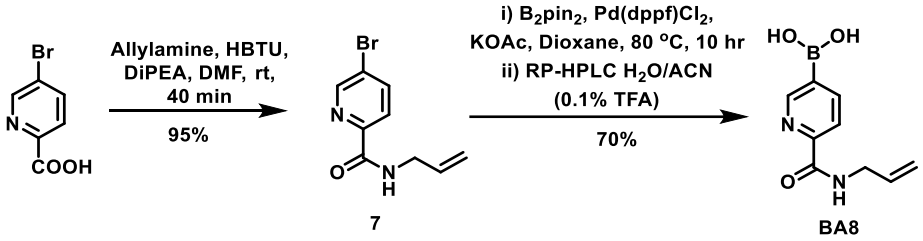
BA5 (14 mg, 0.05 mmol) and $NaBH_4$ (9.5 mg, 0.25 mmol, 5 eq.) were taken in 1 mL 1:1 MeOH:H_2O and stirred at room temperature for 15 min. MeOH was evaporated, and the aqueous solution was acidified with 20 mL 1N HCl and washed with 2×50 mL EtOAc. The organic layer was further washed with brine and evaporated to obtain the benzoxaborole **BA7** as a white solid with a 65% yield (8 mg), mp: 148-150 °C. We believe there was some mechanical loss of the product during the workup.

^{1}H NMR (400 MHz, $CDCl_3$): δ 7.23 (d, *J* = 8 Hz, 1H), 7.16 (d, *J* = 2.0 Hz, 1H), 7.02 (dd, *J* = 8.3, 2.6 Hz, 1H), 6.83 (s, 1H), 5.86-5.76 (m, 1H), 5.19 – 5.09 (m, 2H), 5.00 (s, 2H), 4.49 (s, 2H), 3.94-3.92 (m, 2H).

^{13}C NMR (101 MHz, $CDCl_3$): δ 156.7, 147.3, 133.5, 122.5, 118.9, 116.8, 114.7, 83.0, 71.0, 67.5, 41.3.

HRMS-ESI$^+$ (*m/z*): $[M+H]^+$ Calc. 248.1094, Obs. 248.1096.

2.5.4.6 *SYNTHESIS OF BA8*



*Synthesis of **7***

5-Bromopicolinic acid (505 mg, 2.5 mmol) and HBTU (950 mg, 2.5 mmol, 1 eq.) were dissolved in 2 mL DMF. Then, DiPEA (0.64 mL, 3.75 mmol, 1.5 eq.) and allylamine (0.2 mL, 2.75 mmol, 1.1 eq.) were added at 5 min intervals to the reaction mixture. TLC analysis confirmed the completion of the reaction after 35 min of stirring. 100 mL EtOAc was added to the reaction mixture, and the organic layer was washed with 50 mL chilled 1 N HCl and 50 mL chilled 5% $NaHCO_3$, followed by brine wash. After evaporation of the solvent, the residue was purified through silica gel using 10% EtOAc/Hexane to obtain 570 mg of **7** (95% yield) as a yellow oil.

^{1}H NMR (400 MHz, $CDCl_3$): δ 8.57 (dd, *J* = 2.3, 0.7 Hz, 1H), 8.07 (dd, *J* = 8.3, 0.7 Hz, 1H), 7.95 (dd, *J* = 8.3, 2.3 Hz, 1H), 5.91 (ddt, *J* = 17.1, 10.3, 5.6 Hz, 1H), 5.24 (ddd, *J* = 17.1, 3.0, 1.7 Hz, 1H), 5.16 (ddd, *J* = 10.2, 2.8, 1.4 Hz, 1H), 4.10 – 4.04 (m, 3H).

^{13}C NMR (101 MHz, $CDCl_3$): δ 163.5, 149.4, 148.4, 140.1, 133.9, 124.1, 123.8, 116.7, 41.9.

HRMS-ESI$^+$ (*m/z*): $[M+H]^+$ Calc. 240.9977, Obs. 240.9977.

*Synthesis of **BA8***

7 (360 mg, 1.5 mmol), $Pd(dppf)Cl_2$.DCM (61.2 mg, 0.075 mmol, 0.05 eq.), B_2pin_2 (952 mg, 3.75 mmol, 2.5 eq.), and KOAc (441 mg, 4.5 mmol, 3 eq.) were added together in a 25 mL oven-dried round-bottomed flask. 9 mL of dry dioxane was added, and the suspension was purged with argon for 20 min. The reaction mixture was then heated at 80 °C for 10 hr in stirring conditions. The reaction was monitored by HPLC, and the product formation was confirmed by LCMS. TLC analysis indicated consumption of 7. The reaction mixture was then filtered through a pad of celite. Upon evaporation of the solvent, 100 mg of the residue was purified through HPLC to obtain 65 mg of **BA8** (TFA salt) as a yellow powder (70% yield) after lyophilization, mp: 112-114 °C. Method: Mobile phase acetonitrile-water (0.05% TFA) with a gradient of 15-95% acetonitrile over 30 min. We observed the removal of the pinacol group from **BA8** due to the use of the acid mobile phase in HPLC purification.

^{1}H NMR (400 MHz, 10% $CD_3OD/CDCl_3$): δ 8.81 (s, 1H), 8.19 (d, *J* = 7.7 Hz, 1H), 8.09 (d, *J* = 7.8 Hz, 1H), 5.94-5.85 (m, 1H), 5.26-5.21 (m, 1H), 5.16 – 5.14 (m, 1H), 4.05 (d, *J* = 5.5 Hz, 3H).

^{13}C NMR (101 MHz, 10% $CD_3OD/CDCl_3$): δ 164.6, 153.0, 143.3, 133.8, 121.5, 116.7, 84.7, 41.8.

HRMS-ESI$^+$ (*m/z*): $[M+H]^+$ Calc. 207.0941, Obs. 207.0939.

2.5.5 *SYNTHESIS OF BIOLOGICALLY ACTIVE PEPTIDES*

Except for GSH, all the other peptides were synthesized by standard Fmoc-SPPS on Rink amide polystyrene AM resin. Three equivalents of the commercially available amino acids, HBTU (1 eq.) as a coupling agent and DiPEA (1.5 eq.) as a base, were used for the coupling reaction for 15 min. Fmoc deprotection was achieved using 3 mL 20% piperidine/DMF twice for 3 min each, followed by washing with 3 mL DMF six times. The peptides were cleaved off of the resin and globally deprotected with reagent K (82.5% TFA, 5% H_2O, 2.5% EDT, 5% Thioanisole and 5% phenol). Precipitation with chilled diethyl ether gave the crude peptides, which were purified by RP-HPLC. The purity of all peptides was determined by using LC-MS (Waters) and HPLC (Shimadzu) to be >94%. The characterization data are presented in Table 1.

Table 2.1: List of biologically active peptides used (except for GSH, all were synthesized). Acp refers to 6-Aminocaproic acid.

Sl. No.	**Name**	**Sequence**	**Length**	**Calc. *m/z***	**Obs. *m/z***
1.	Glutathione (GSH)	γ-ECG	3-mer	-	-
2.	SPYG	Ac-SPYGRC-NH_2	6-mer	$[M+H]^+$ 724.33	$[M+H]^+$ 724.31
3.	Oxytocin	C(Acm)YIQNC(Acm)PLG C-NH_2	10-mer	$[M+2H]^{2+}$ 627.77	$[M+2H]^{2+}$ 627.77
4.	Leucine enkephalin (LeuEnk)	Ac-YGGFLC-NH_2	6-mer	$[M+H]^+$ 700.31	$[M+H]^+$ 700.29
5.	Cys Leucine enkephalin (CysLeuEnk)	Ac-CYGGFLC-NH_2	7-mer	$[M+H]^+$ 803.32	$[M+H]^+$ 803.30
6.	Wheat Germ Agglutinin, 62-73 (WGA)	Ac-KSQYGYCGFGAEY-NH_2	13-mer	$[M+H]^+$ 1513.64	$[M+H]^+$ 1513.59
7.	p53	Ac-LTFEHYWACLTSK-NH_2	13-mer	$[M]^+$ 1638.79	$[M]^+$ 1638.73
8.	Ubiquicidine (UBI)	TGRAKRRMQYNRRGC-NH_2	15-mer	$[M+3H]^{3+}$ 618.40	$[M+3H]^{3+}$ 618.01
9.	ATN-161	Ac-PHSCN-NH_2	5-mer	$[M+H]^+$ 599.24	$[M+H]^+$ 599.24

2.5.6 *EXPERIMENTAL SETUP*

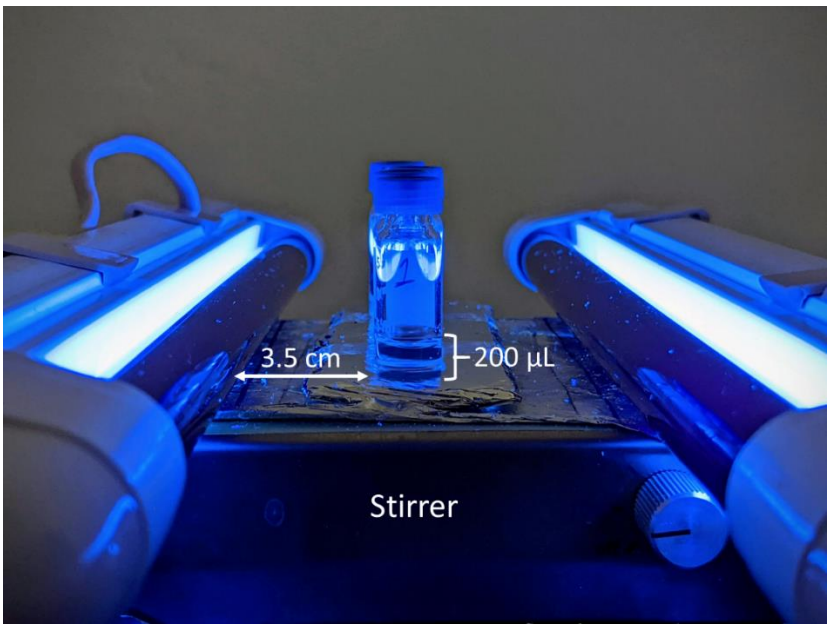


Figure 2.8: Experimental setup of thiol-ene click reactions.

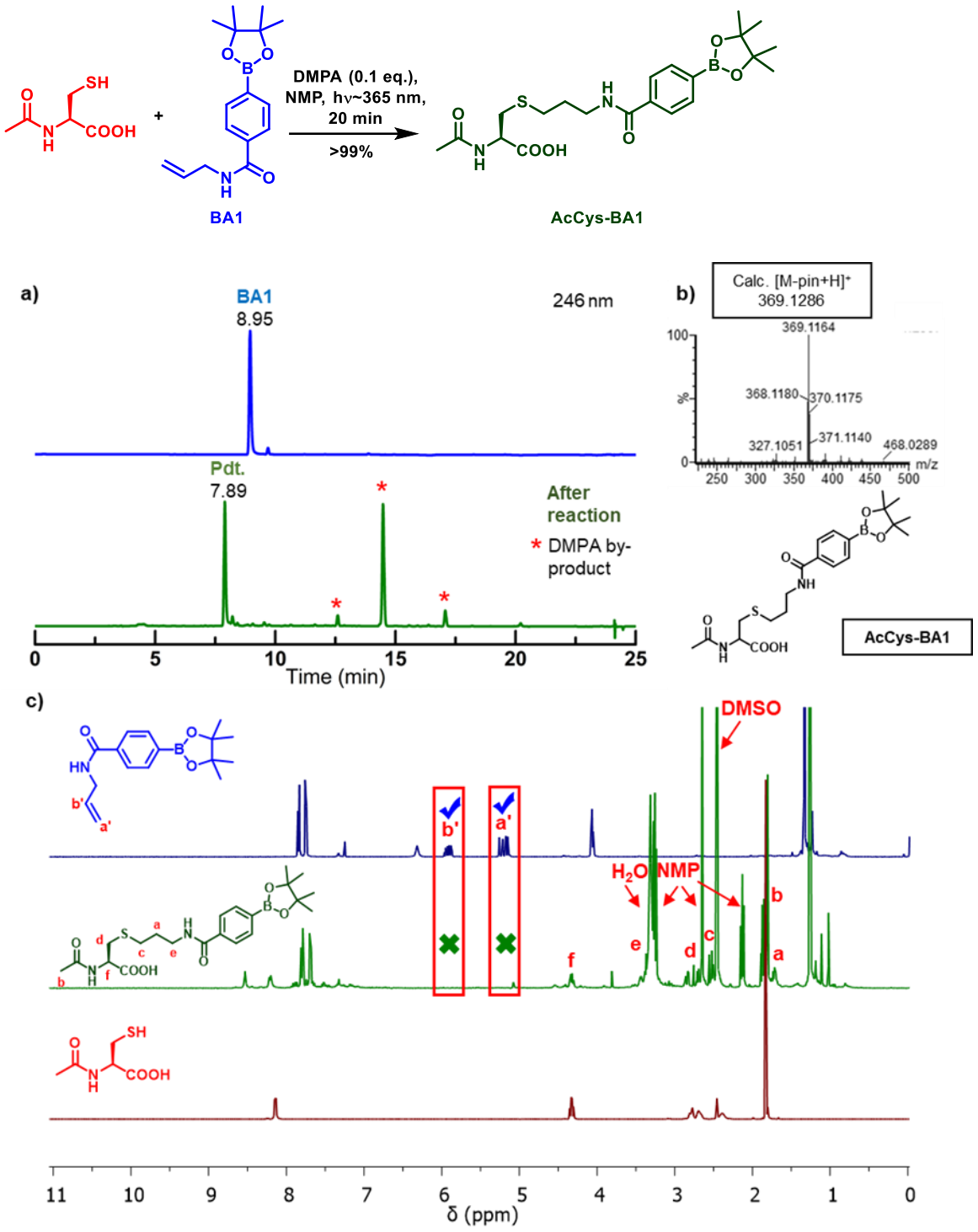
All the thiol-ene click reactions were performed in 1.5 mL glass vials with a minimum of 200 µL reaction volume. They were stirred for 20 min at room temperature with constant stirring. The vials were kept in the center of two Philips TL 8W BLB (black light blue) light sources, 7 cm apart with λ~365 nm. NMP was used as supplied, and H_2O with 1% TFA was used unless otherwise specified.

2.5.7 *GENERAL PROCEDURE FOR THIOL-ENE CLICK REACTIONS WITH ACETYL CYSTEINE (AcCYS)*

AcCys-OH (220 mM, 1.1 eq), BA1-BA8 (200 mM, 1eq), and DMPA (20 mM) were dissolved in NMP (1% TFA) and stirred for 20 min in the presence of hν ~365 nm light source in dark. The reaction mixture was diluted to 1 mL using 1 N HCl and washed with 2 mL EtOAc. The crude residue was subjected to analytical HPLC, NMR (solvent DMSO-d_6), and ESI-MS for product confirmation and understanding conversion directly from the reaction mixture. The HPLC analysis was carried out at the absorption maxima of the respective BA substrates introduced in the reaction. The term 'pdt.' is an abbreviation for the product in all the HPLC spectra. The residue was further purified through RP-HPLC, and their respective purified ^{1}H NMR data are presented. Complete loss of the pinacol group (~ δ 1.2-1.4) was observed after the HPLC purification. This statement can be verified by comparing the NMR data of the crude sample with that obtained after purification by HPLC; for instance, compare Figure 2.9c & d. Since AcCys-BA2 is UV inactive, conjugation of BA2 with Fmoc-Cys was performed for the convenience of purification through RP-HPLC and isolation of the product. All data corroborate with each other and are provided in Figures 2.9-2.16. All the experiments were repeated twice, and the best data are presented here.

Method for analytical HPLC (Shimadzu Prominence UFLC): Waters Sunfire C18 (5 µm, 4.6×250 mm) analytical column and mobile phase acetonitrile-water (0.05% TFA) with a flow rate of 1 mL/min were used. Gradient used: isocratic 10% acetonitrile for 1 min, then a gradient from 10% to 90% acetonitrile in 19 min, then isocratic 90% acetonitrile for a min, then column equilibration at an initial isocratic gradient 10% acetonitrile over 4 min.

Reaction with BA1


DMPA (0.1 eq.),
NMP, hν~365 nm,
20 min
>99%
BA1
AcCys-BA1
a)
BA1
8.95
246 nm
Pdt.
7.89
After reaction
* DMPA by-product
Time (min)
b)
Calc. [M-pin+H]+
369.1286
369.1164
368.1180
370.1175
327.1051
371.1140
468.0289
m/z
AcCys-BA1
c)
DMSO
H2O
NMP
δ (ppm)

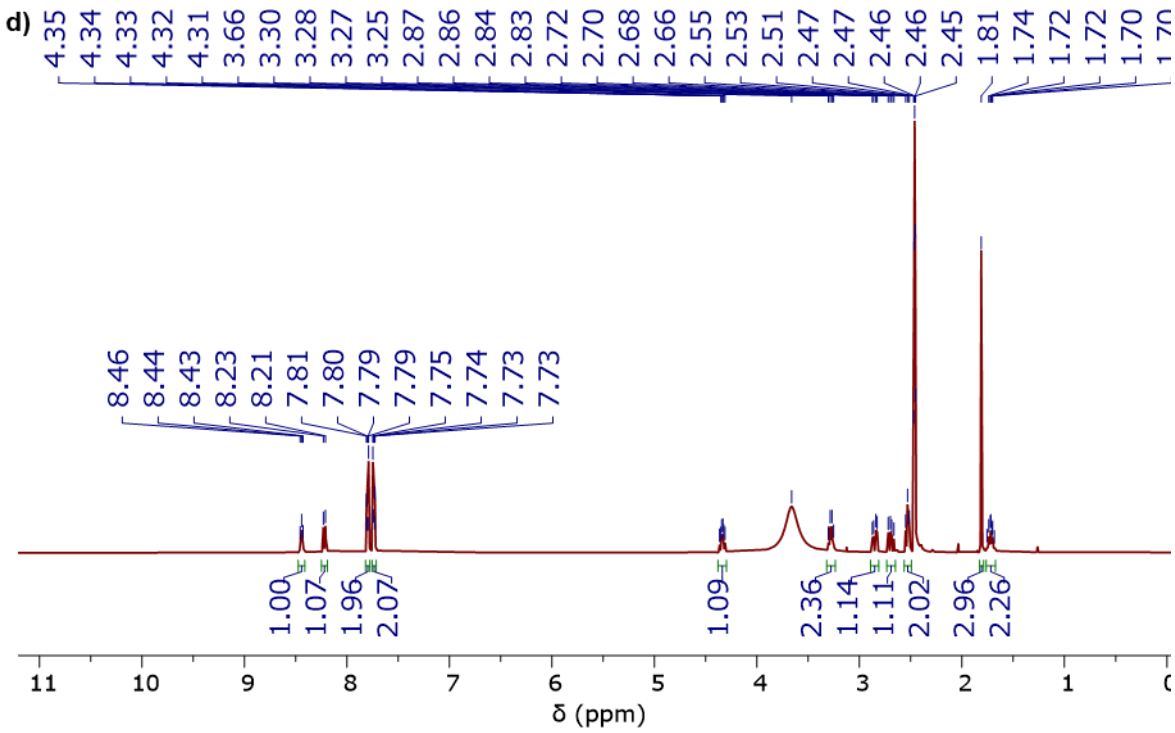
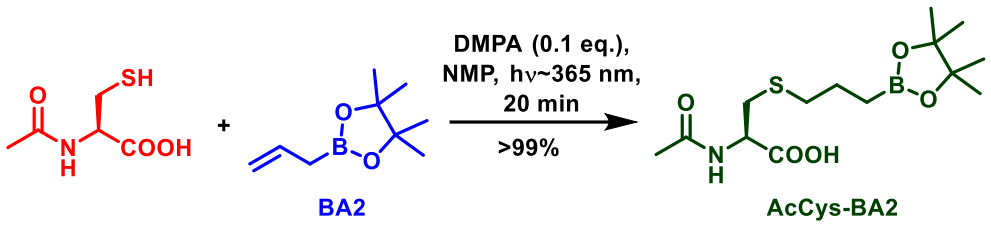
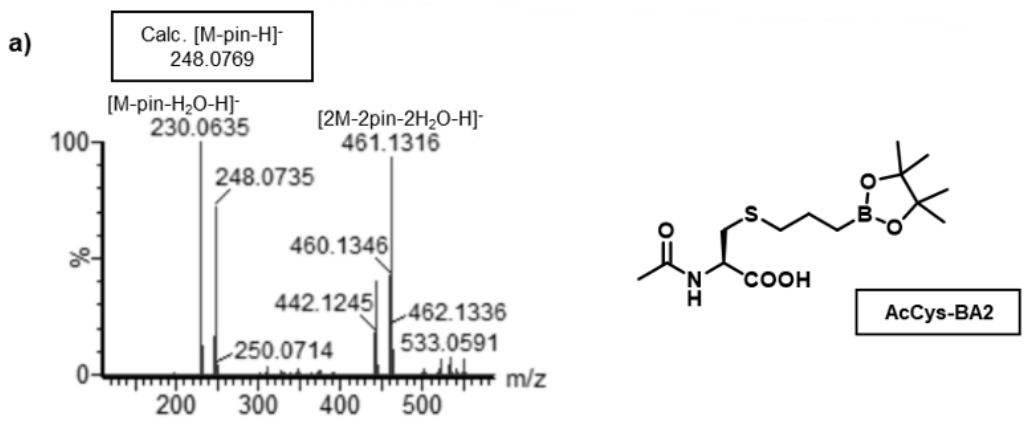


Figure 2.9: a) HPLC trace at 246 nm; b) Mass data; c) NMR stacked spectra showing the consumption of allylic protons (marked a' and b') in the crude spectra of the product. Peaks for water and NMP solvents have been assigned in the crude NMR. Small peaks appeared at aromatic region baselines due to photoinitiator (DMPA); d) NMR spectra of pure product in DMSO-d_6 (400 MHz). 73% yield

Reaction with BA2





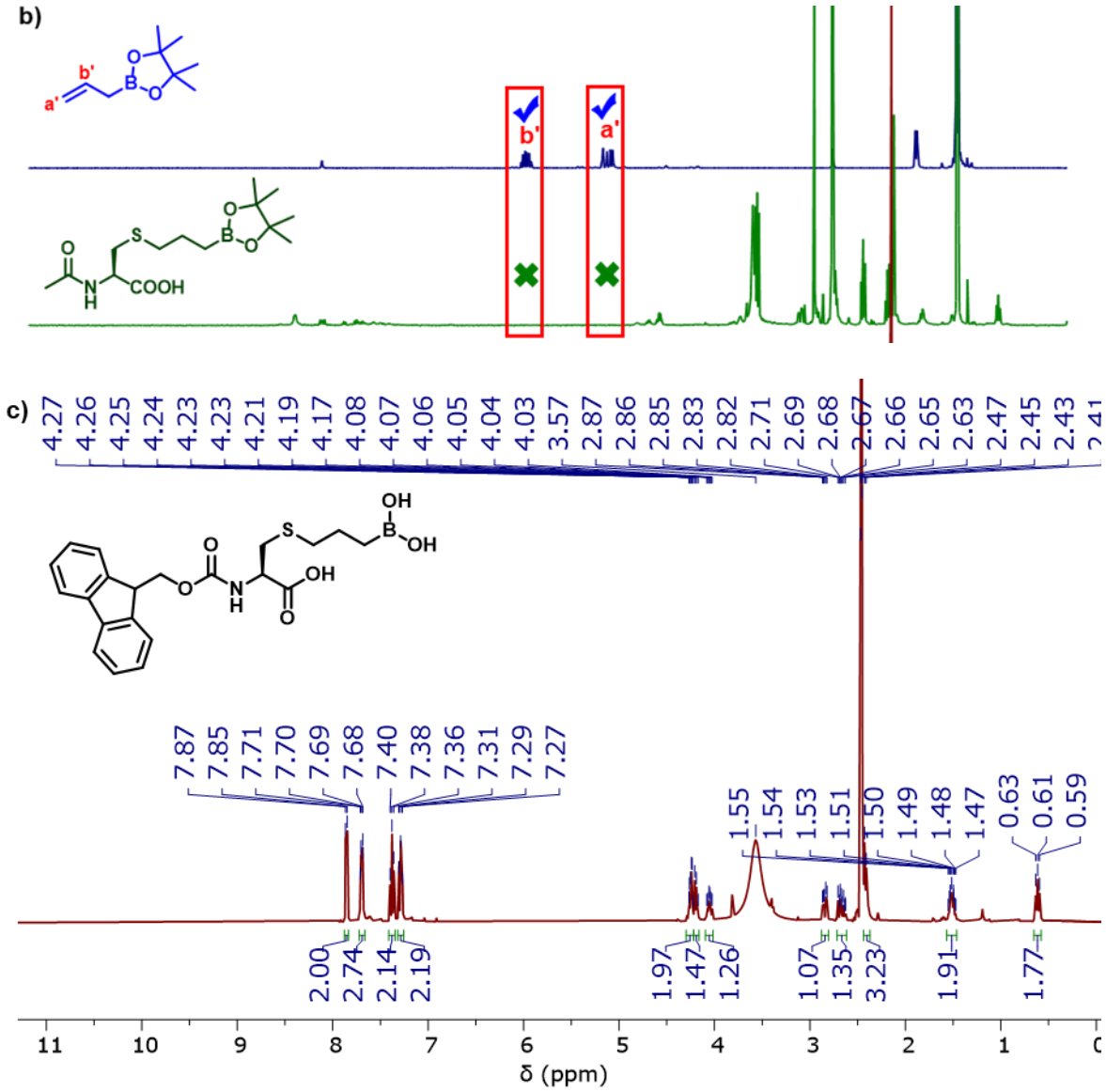
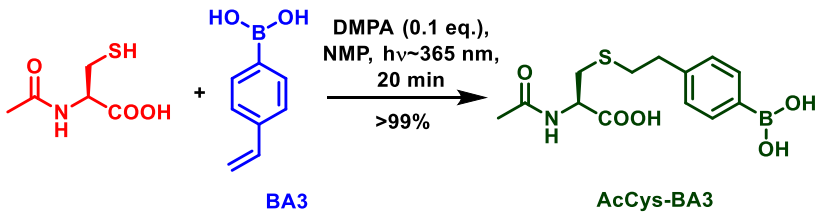


Figure 2.10: a) Mass data in ESI^- (-ve mode); b) NMR stacked spectra showing the consumption of allylic protons (marked a' and b') in the crude spectra of the product. Small peaks appeared at aromatic region baselines due to photoinitiator (DMPA); c) Since AcCys-BA2 is UV inactive, conjugation of BA2 with Fmoc Cys was performed, and the obtained product was purified through RP-HPLC. NMR data of FmocCys-BA2 is presented here in DMSO-d_6 (400 MHz). 72% yield

Reaction with BA3



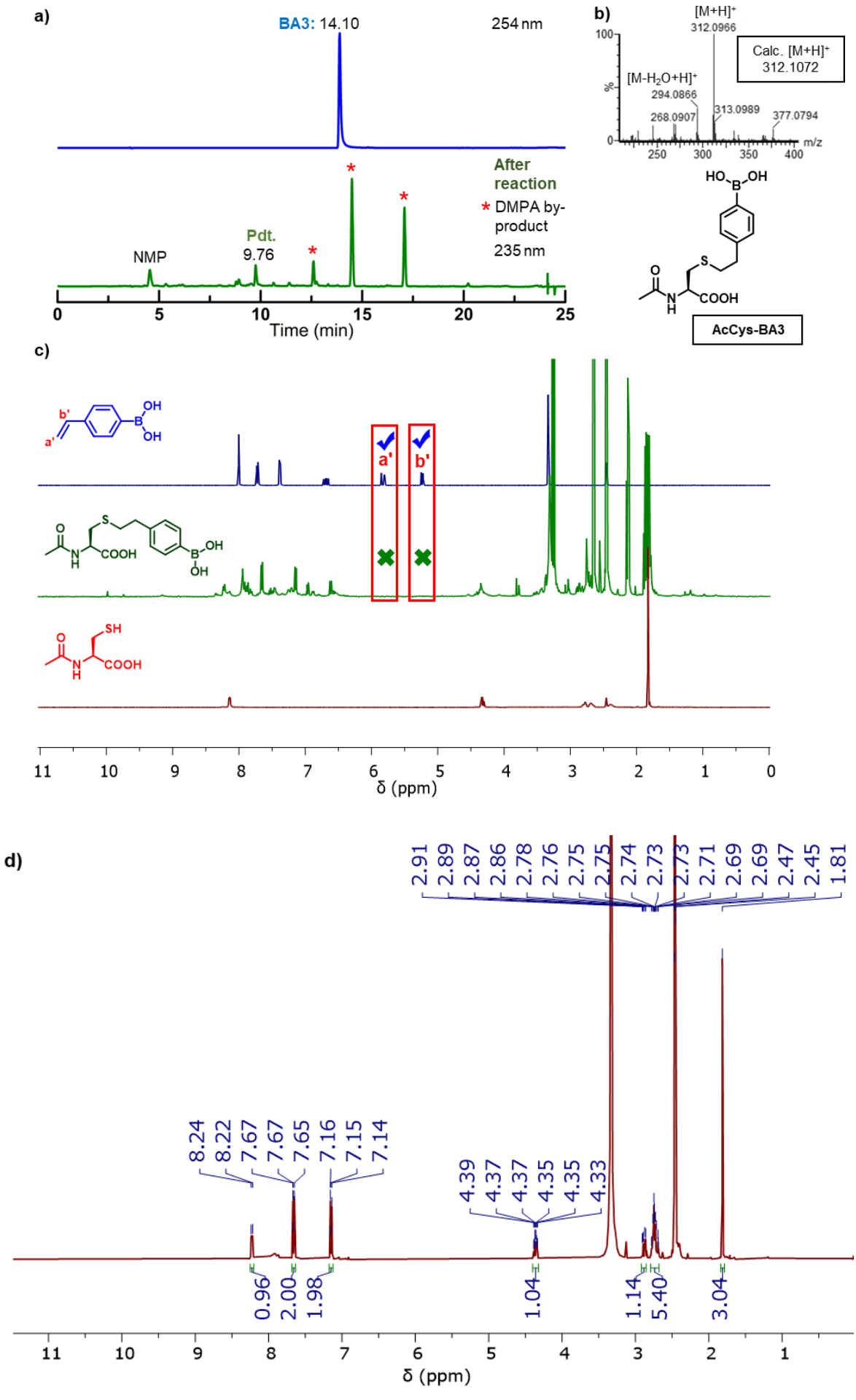
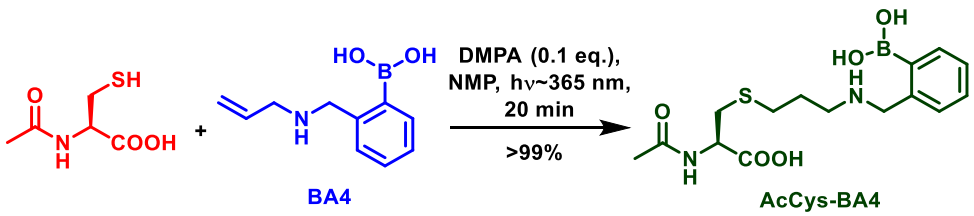
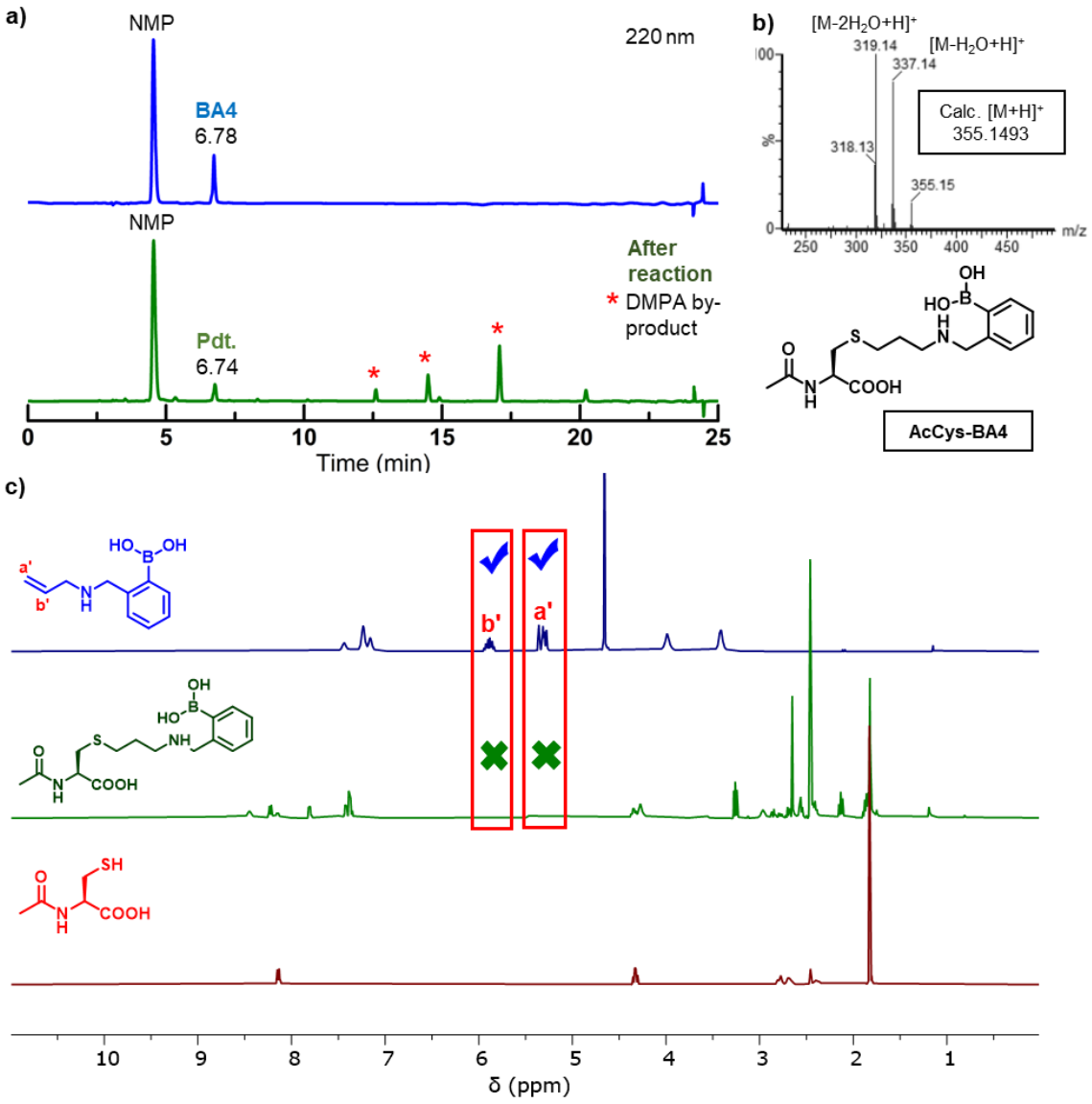

a)
BA3: 14.10
254 nm
After reaction
* DMPA by-product
235 nm
NMP
Pdt.
9.76
Time (min)
b)
[M+H]+
312.0966
Calc. [M+H]+
312.1072
[M-H2O+H]+
294.0866
268.0907
313.0989
377.0794
m/z
AcCys-BA3
c)
a'
b'
δ (ppm)
d)
δ (ppm)

Figure 2.11: a) HPLC trace at 254 nm of BA3 and 235 nm of reaction mixture; b) Mass data; c) NMR stacked spectra showing the consumption of allylic protons (marked a' and b') in the crude spectra of the product. Small peaks appeared at aromatic region baselines due to photoinitiator (DMPA); d) NMR spectra of pure product are in DMSO-d_6 (400 MHz). 76% yield.

Reaction with BA4





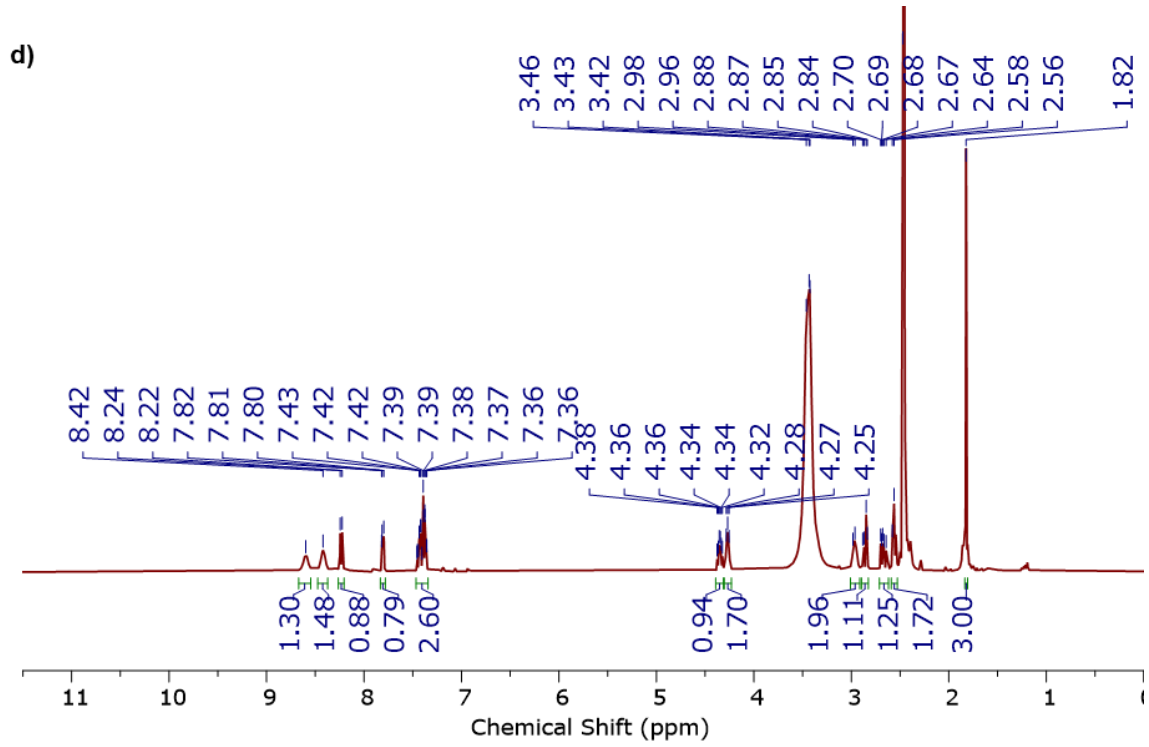
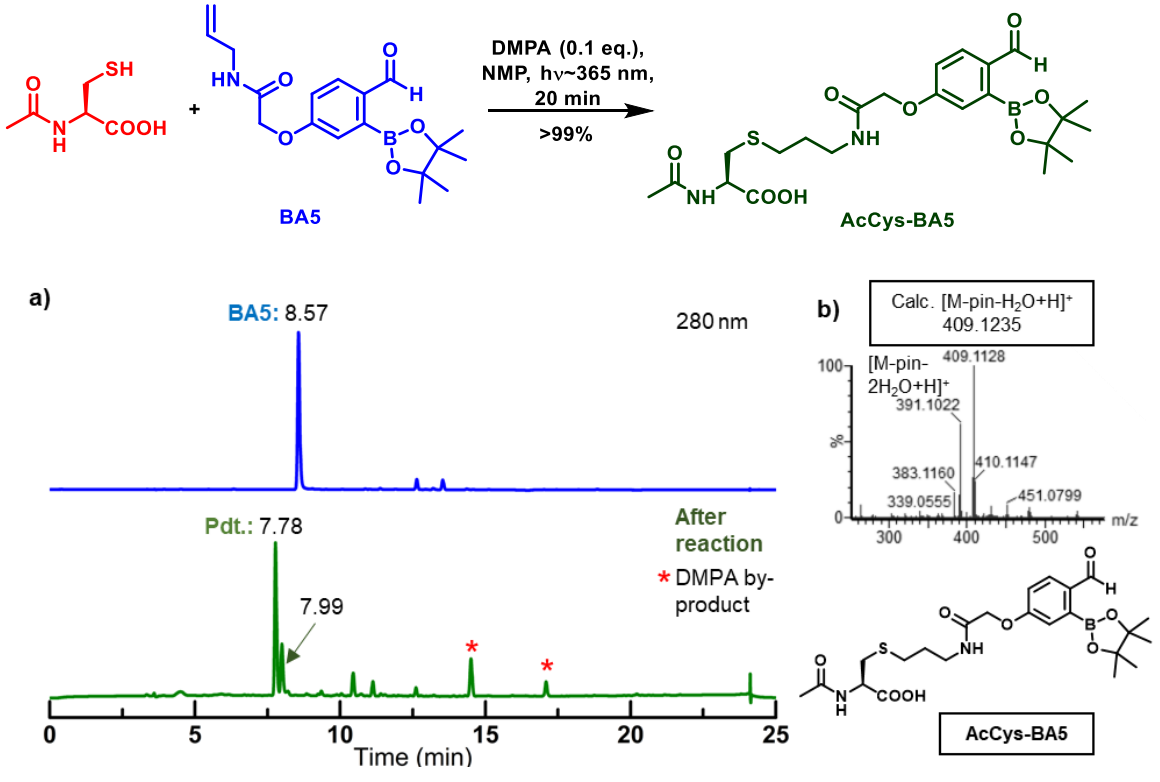


Figure 2.12: a) HPLC trace at 220 nm; b) Mass data; c) NMR stacked spectra showing the consumption of allylic protons (marked a' and b') in the crude spectra of the product. Small peaks appeared at aromatic region baselines due to photoinitiator (DMPA); d) NMR spectra of pure product in DMSO-d_6 (400 MHz). 75% yield

Reaction with BA5



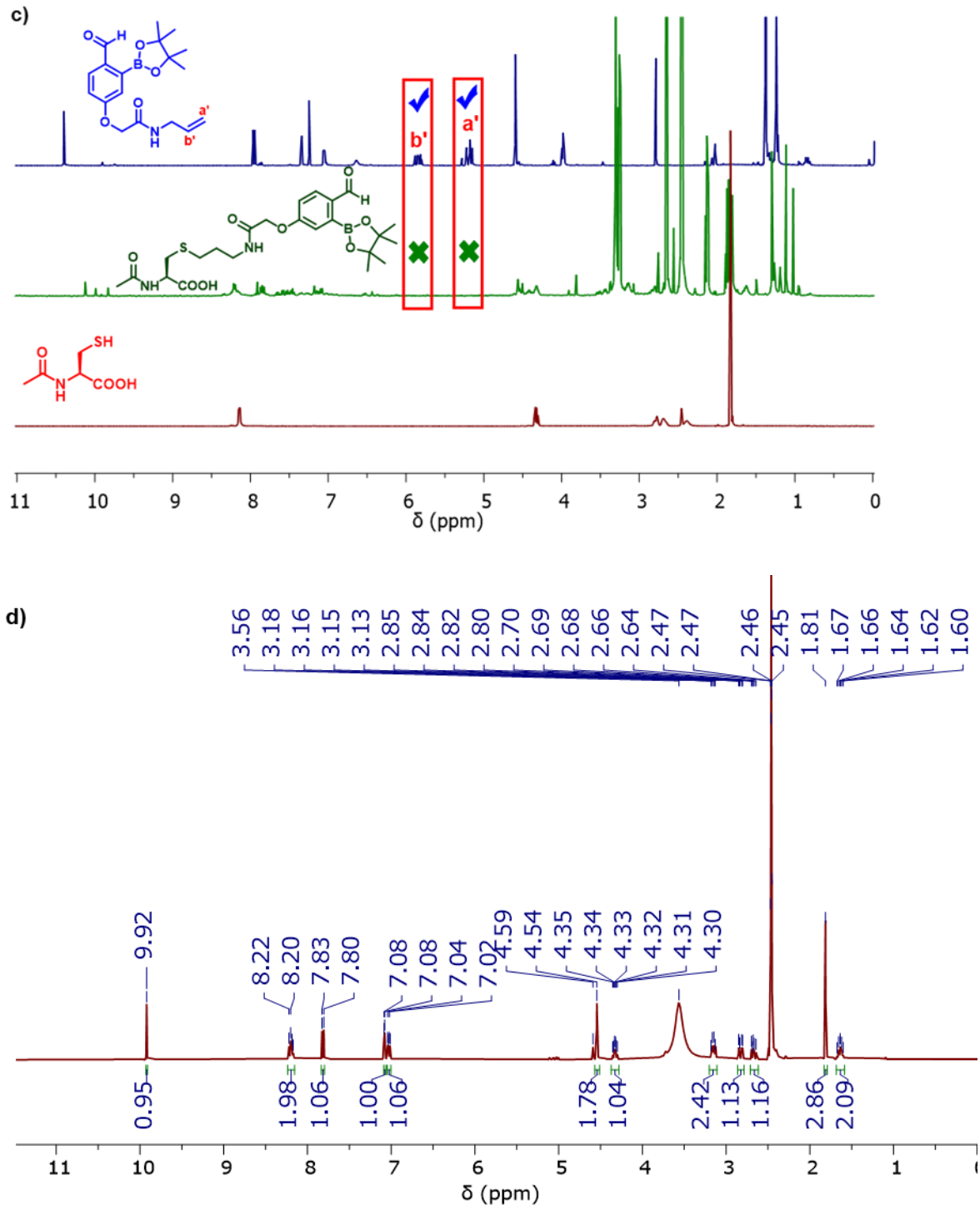
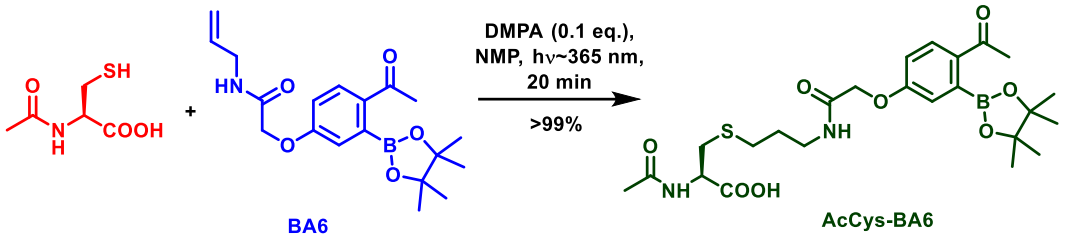
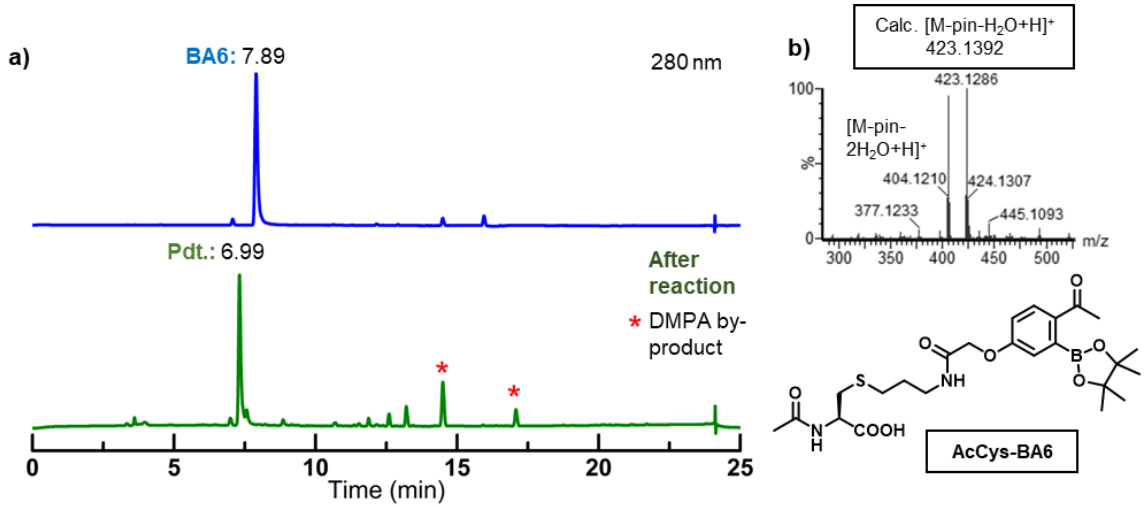
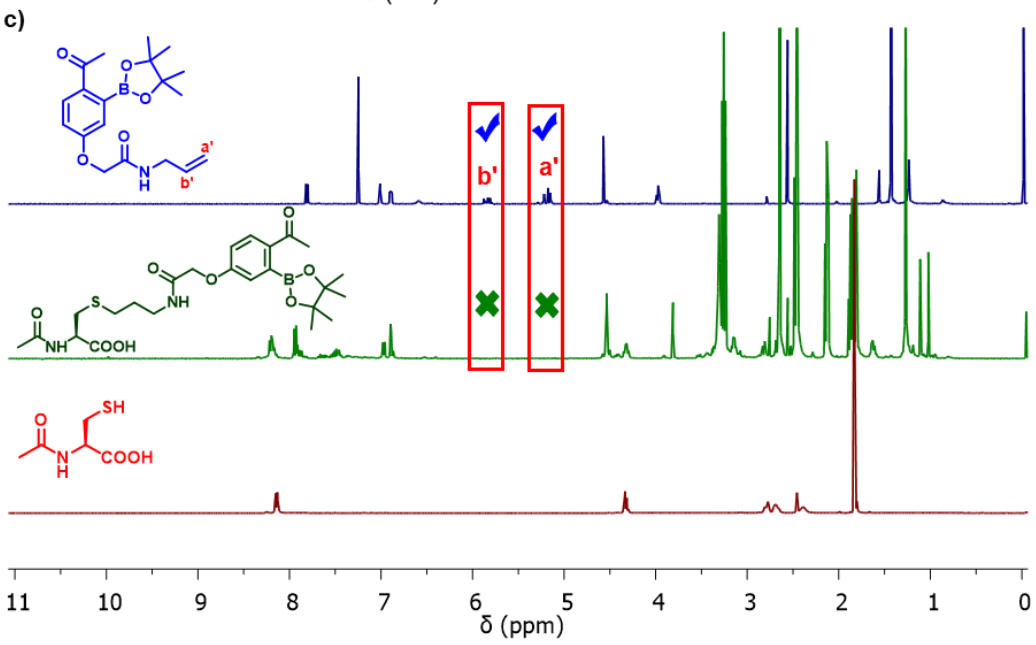


Figure 2.13: a) HPLC trace at 280 nm; b) Mass data; c) NMR stacked spectra showing the consumption of allylic protons (marked a' and b') in the crude spectra of the product. Small peaks appeared at aromatic region baselines due to photoinitiator (DMPA); d) NMR spectra of pure product in DMSO-d_6 (400 MHz). 69% yield

Reaction with BA6







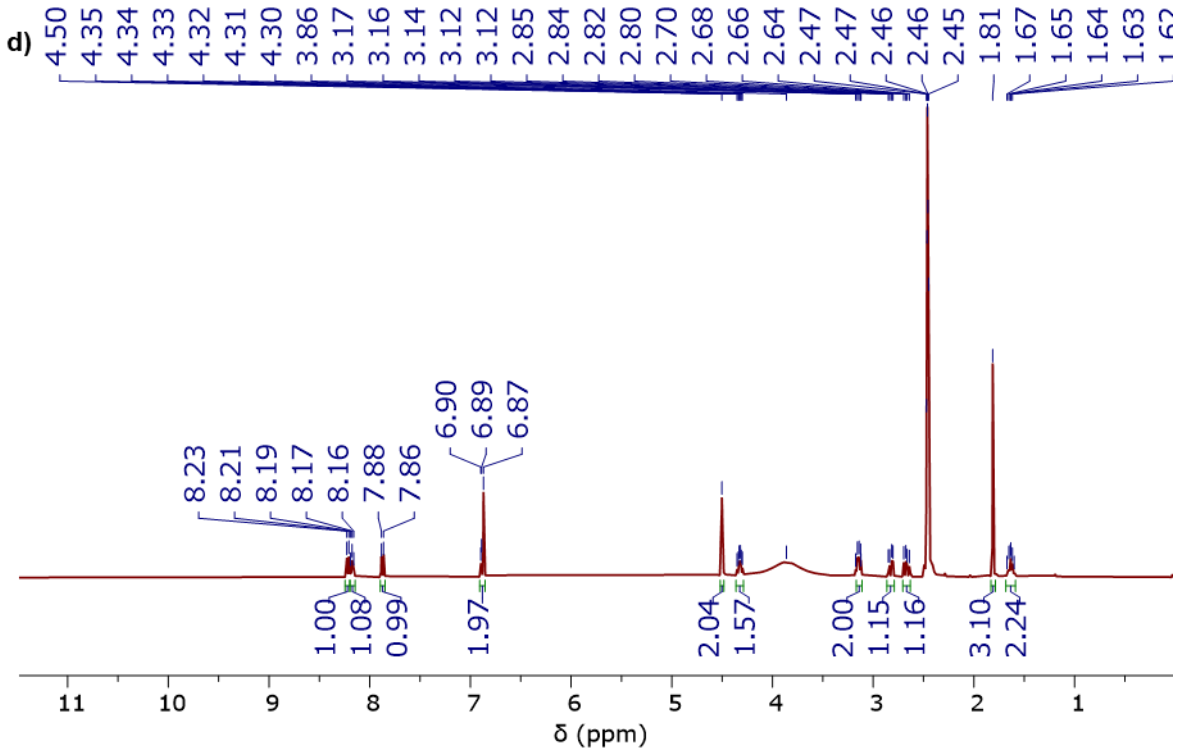
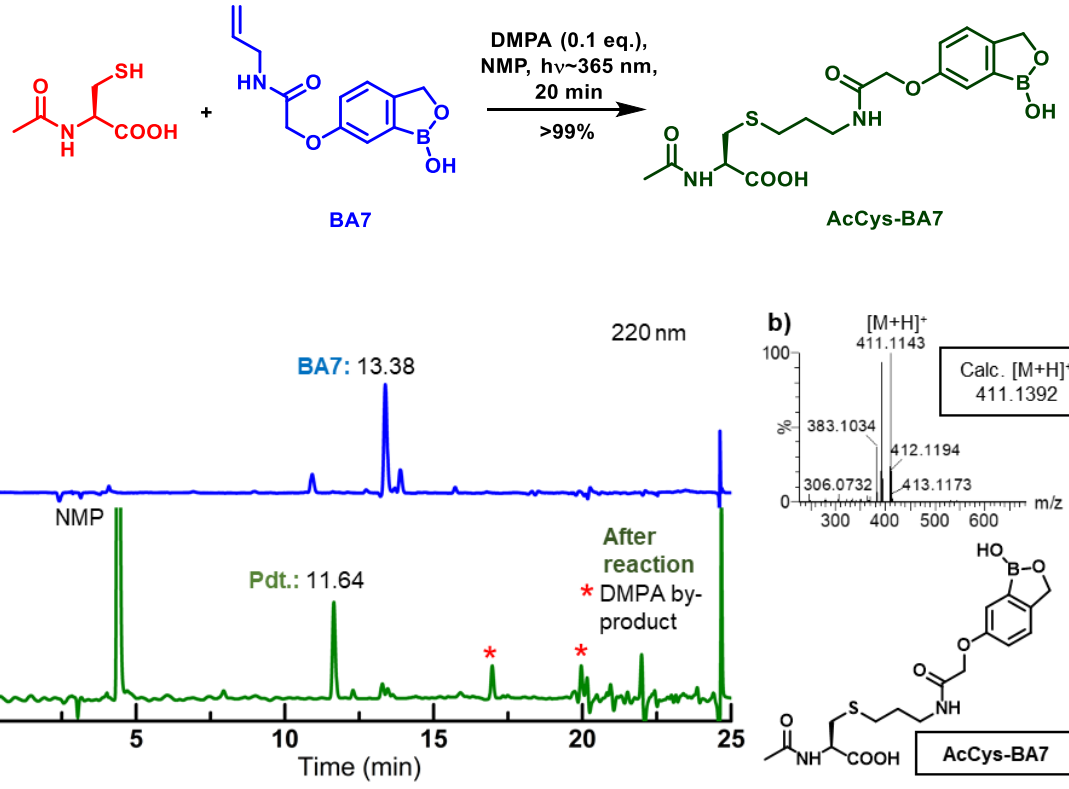


Figure 2.14: a) HPLC trace at 280 nm; b) Mass data; c) NMR stacked spectra showing the consumption of allylic protons (marked a' and b') in the crude spectra of the product. Small peaks appeared at aromatic region baselines due to photoinitiator (DMPA); d) NMR spectra of pure product in DMSO-d_6 (400 MHz). 66% yield

Reaction with BA7



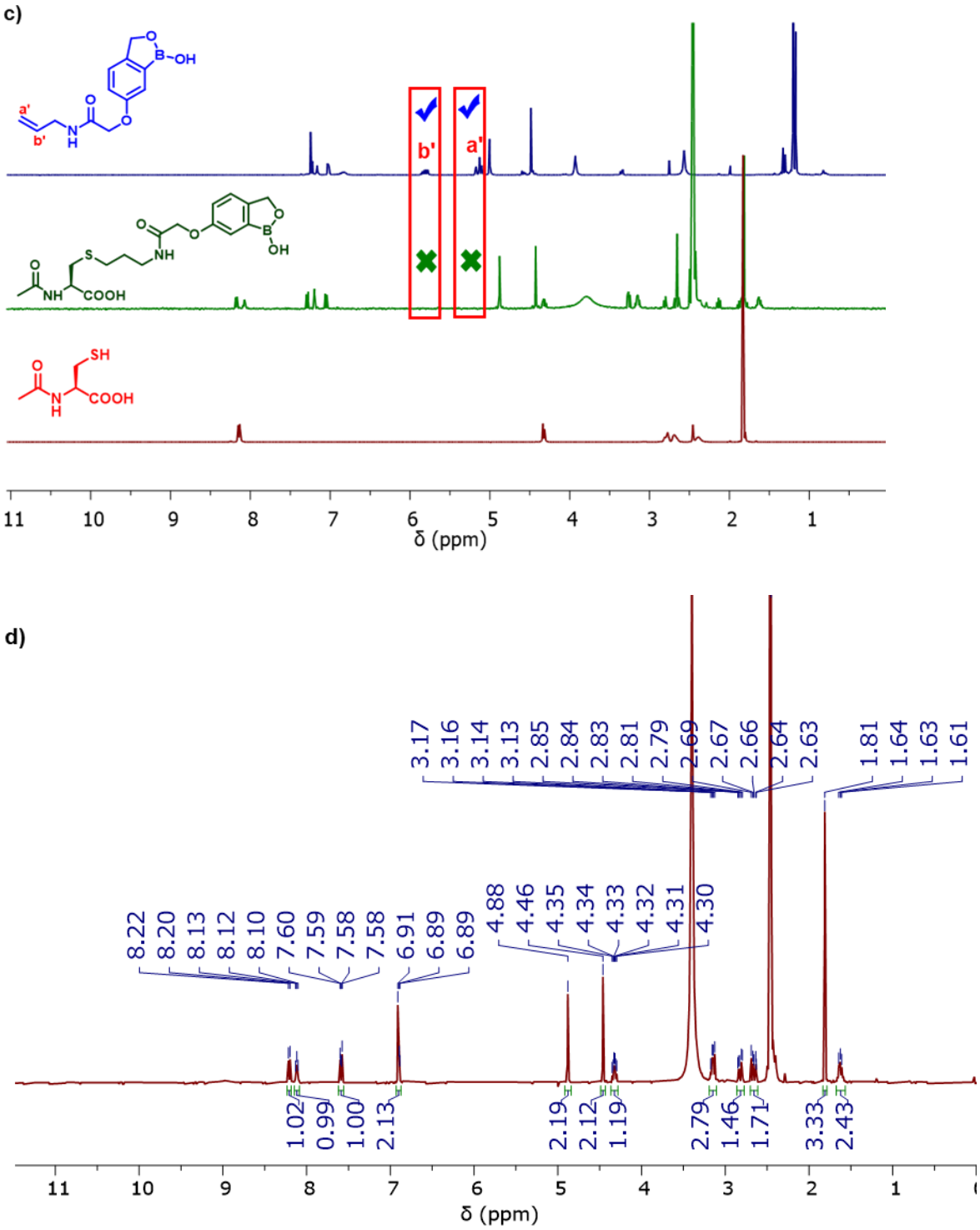
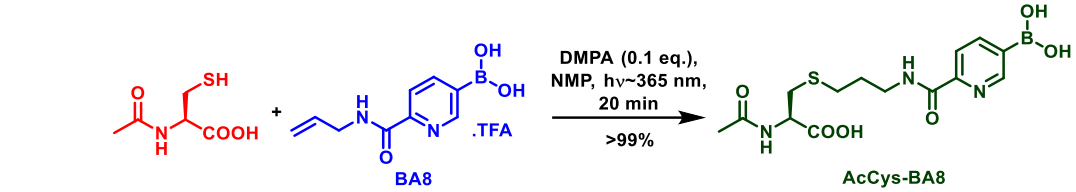
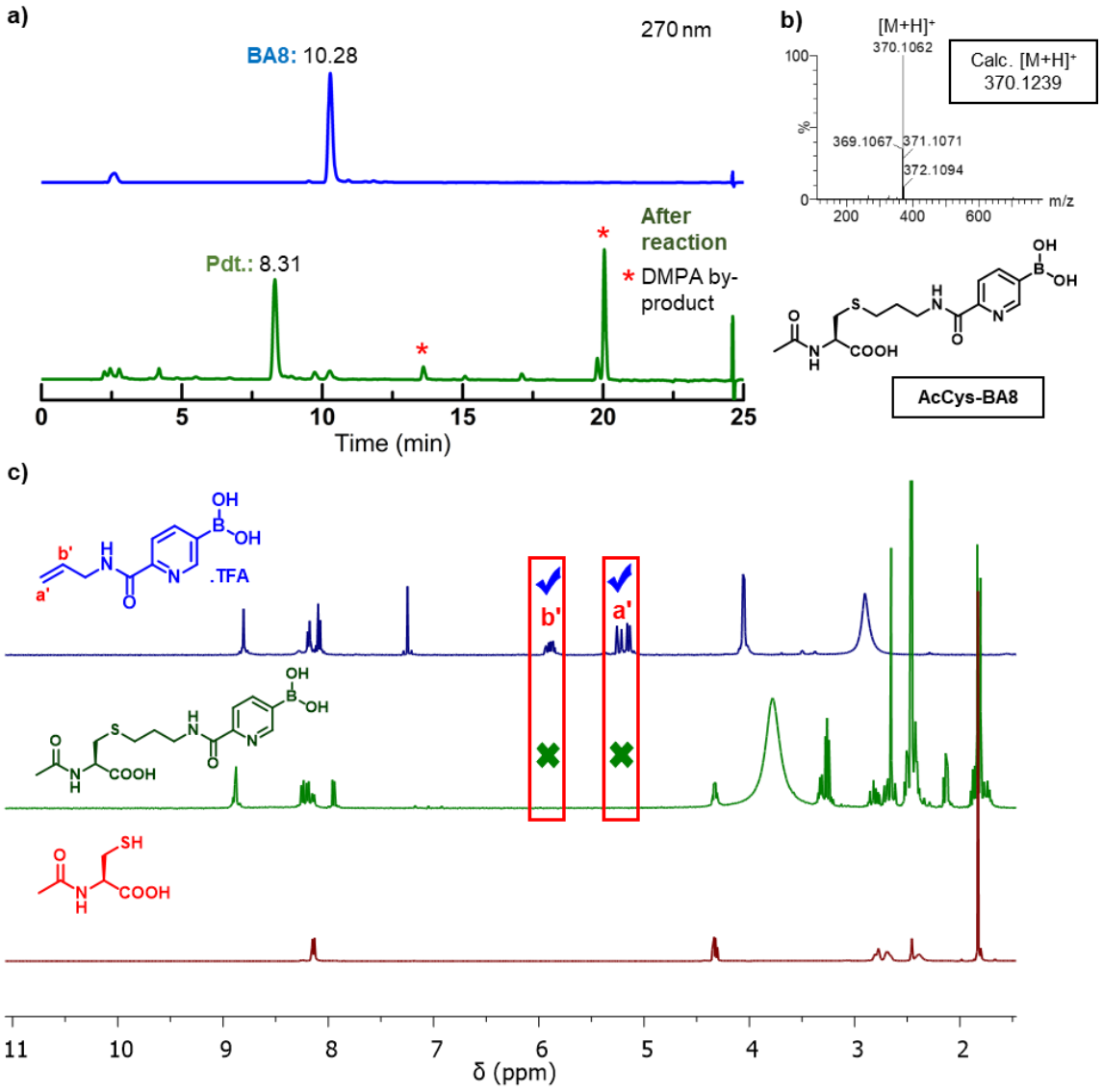


Figure 2.15: a) HPLC trace at 220 nm; b) Mass data; c) NMR stacked spectra showing the consumption of allylic protons (marked a' and b') in the crude spectra of the product. Small peaks appeared at aromatic region baselines due to photoinitiator (DMPA); d) NMR spectra of pure product in DMSO-d_6 (400 MHz). 66% yield

Reaction with BA8





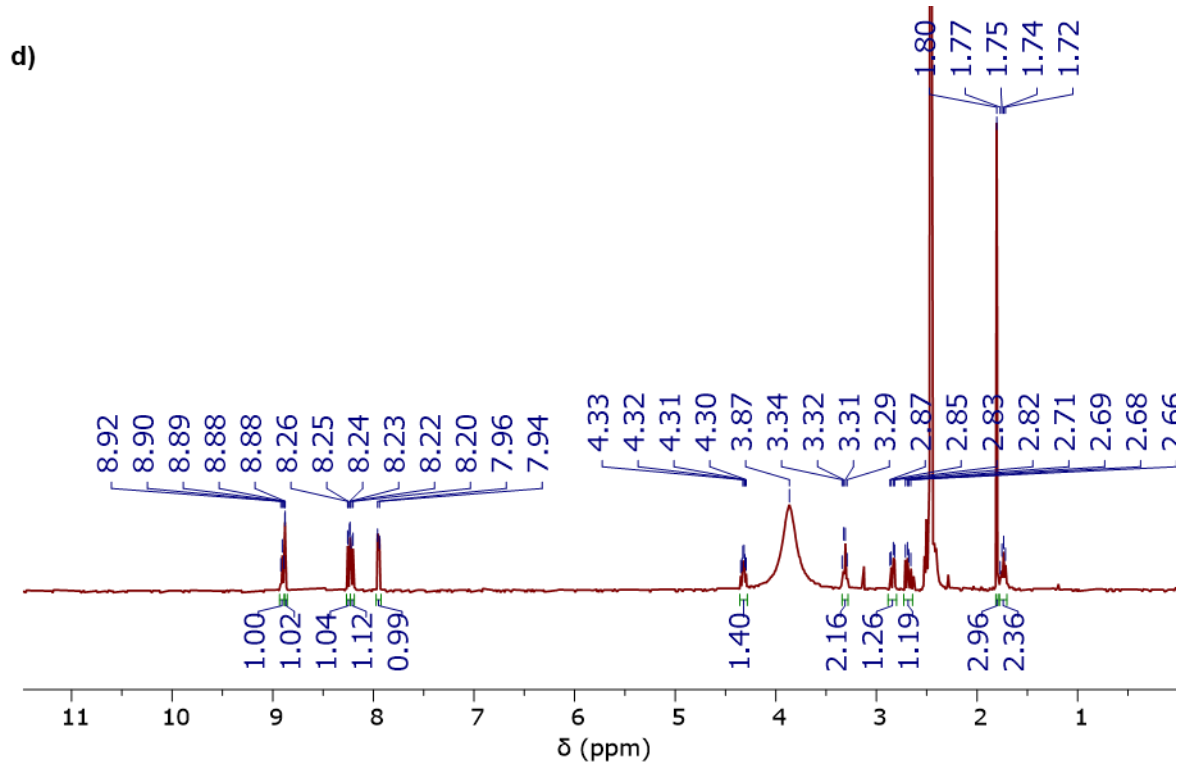
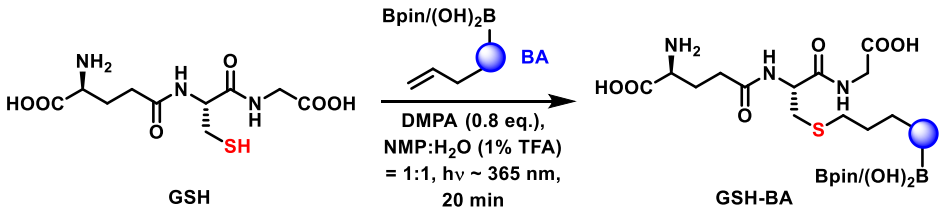


Figure 2.16: a) HPLC trace at 270 nm; b) Mass data; c) NMR stacked spectra showing the consumption of allylic protons (marked a' and b') in the crude spectra of the product. Small peaks appeared at aromatic region baselines due to photoinitiator (DMPA); d) NMR spectra of pure product in DMSO-d_6 (400 MHz). 64% yield

2.5.8 *REACTION PROCEDURE FOR THIOL-ENE CHEMISTRY WITH PEPTIDES*

2.5.8.1 *REACTIONS WITH GSH*



General procedure with GSH: 12 mM of GSH was reacted with 10 mM BA1, BA3-BA8, and 8 mM DMPA for 20 min in NMP:H_2O (1% TFA) = 1:1 as the solvent and hν ~365 nm light source. Reaction with BA2 was not performed since both GSH and BA2 are UV-inactive. The HPLC trace and ESI-MS data for the reactions are provided. The HPLC analysis was carried out at the absorption maxima of the respective substrates for a clear presentation of data. The reaction mixture was further purified through RP-HPLC, and the respective ^{1}H NMR data were recorded. Complete loss of the pinacol group (~ δ 1.2-1.4) was observed after the HPLC purification. We obtained a few milligrams of purified product, so the obtained yields are not reported as the weight measurement will not be accurate. All the experiments were repeated twice, and the best data are presented here. Since the removal of the pinacol ester group is prevalent in the mass spectrometer, boronic acid has been shown in the structures rather than the pinacol ester.

GSH-BA1

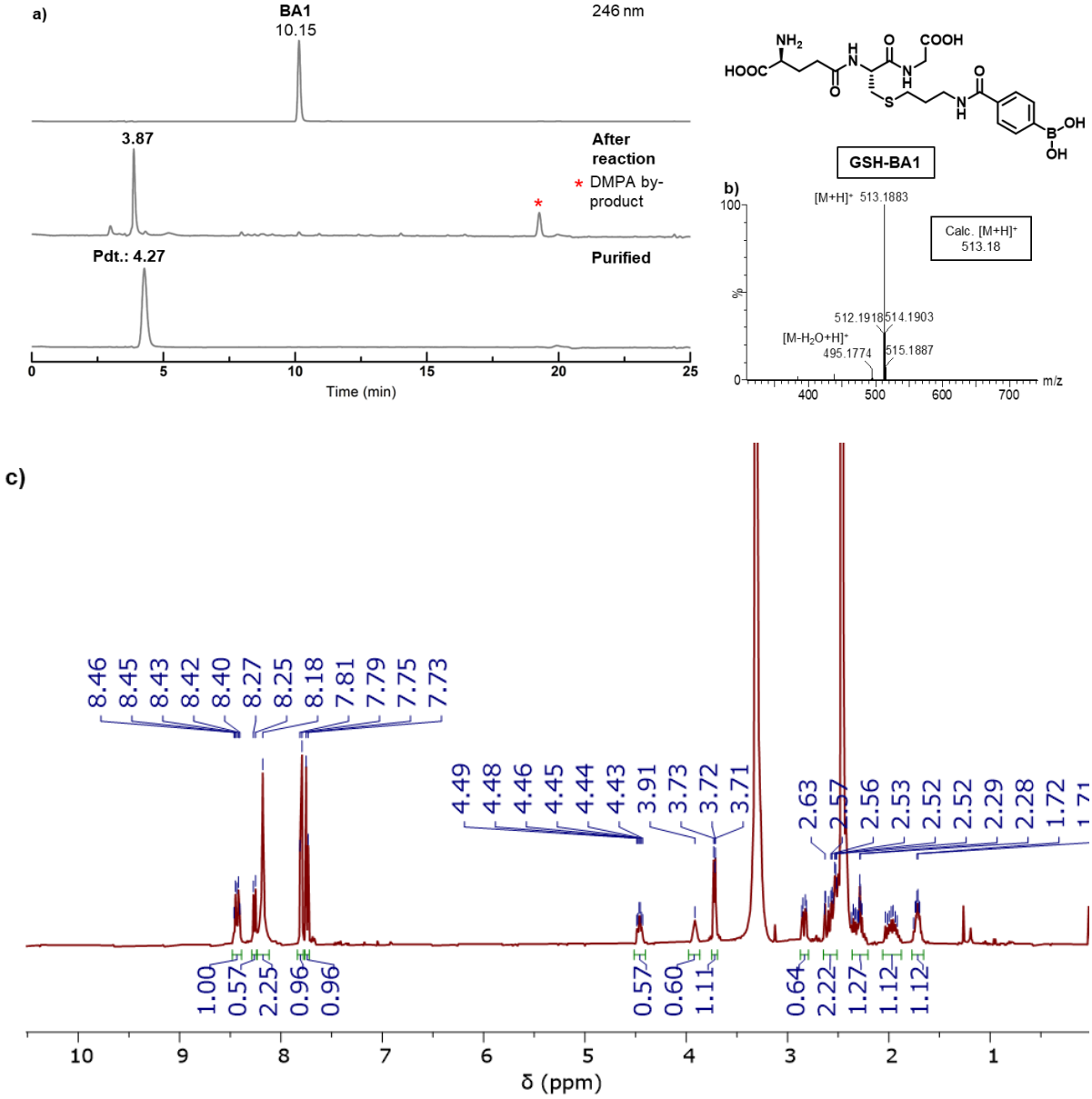


Figure 2.17: a) HPLC trace at 246 nm determines >95% conversion.; b) Mass data; c) NMR spectra of pure product in DMSO-d_6 (400 MHz). 62% yield

GSH-BA3

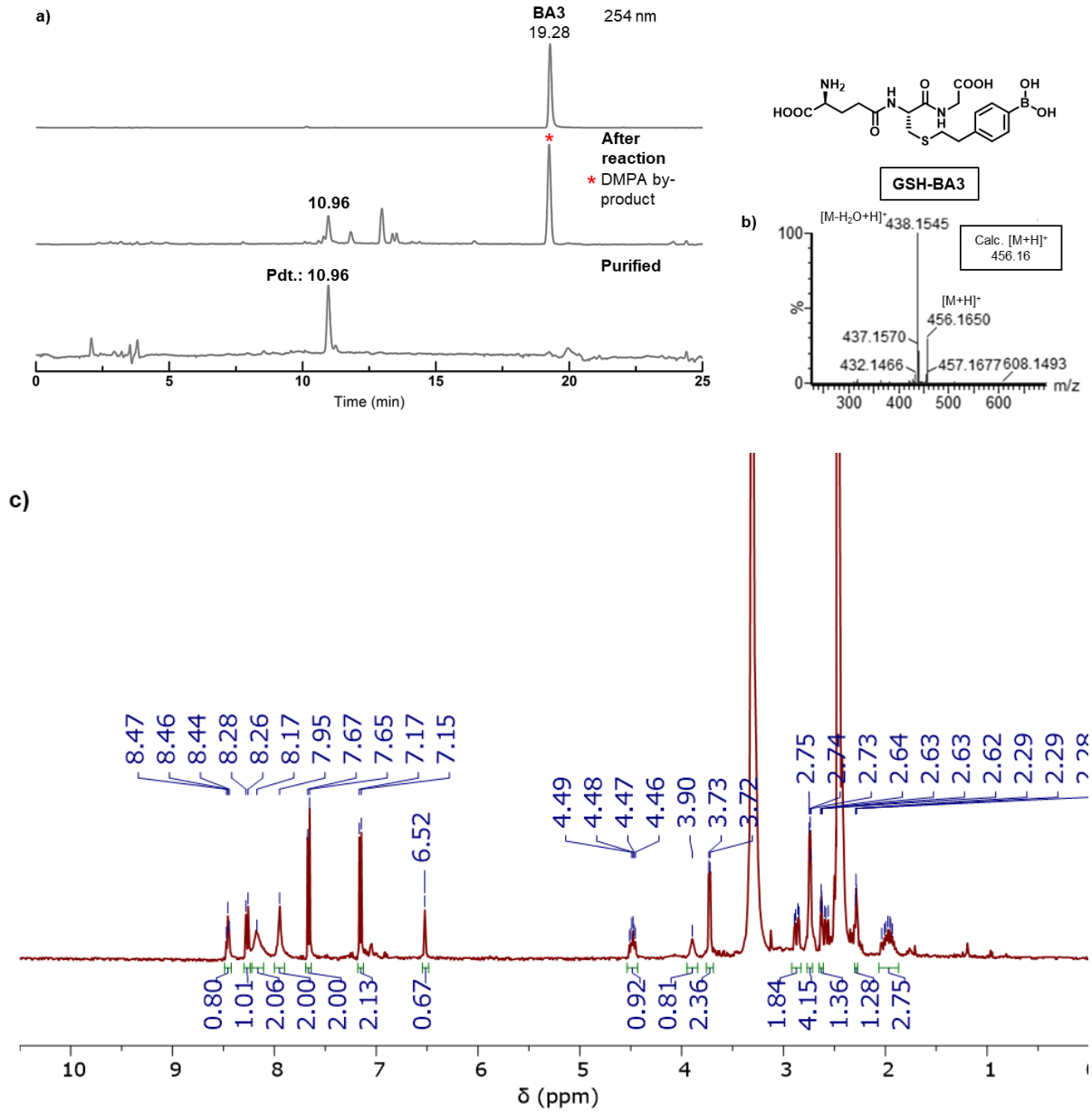


Figure 2.18**:** a) HPLC trace at 254 nm determines >95% conversion; b) Mass data; c) NMR spectra of pure product in DMSO-d_6 (400 MHz). 70% yield

GSH-BA4

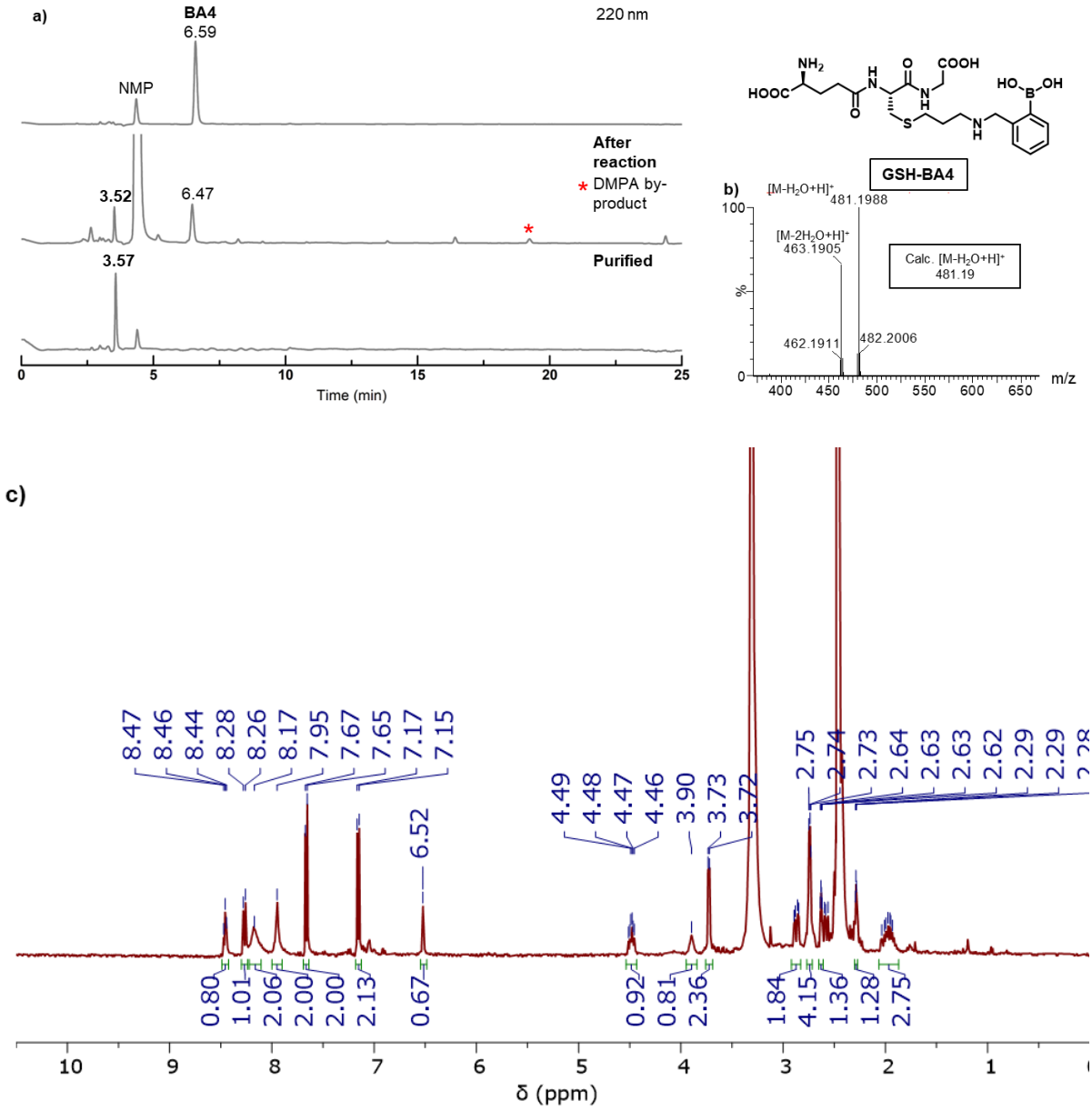


Figure 2.19: a) HPLC trace at 220 nm determines 75% conversion; b) Mass data; c) NMR spectra of pure product in DMSO-d_6 (400 MHz). 55% yield

GSH-BA5

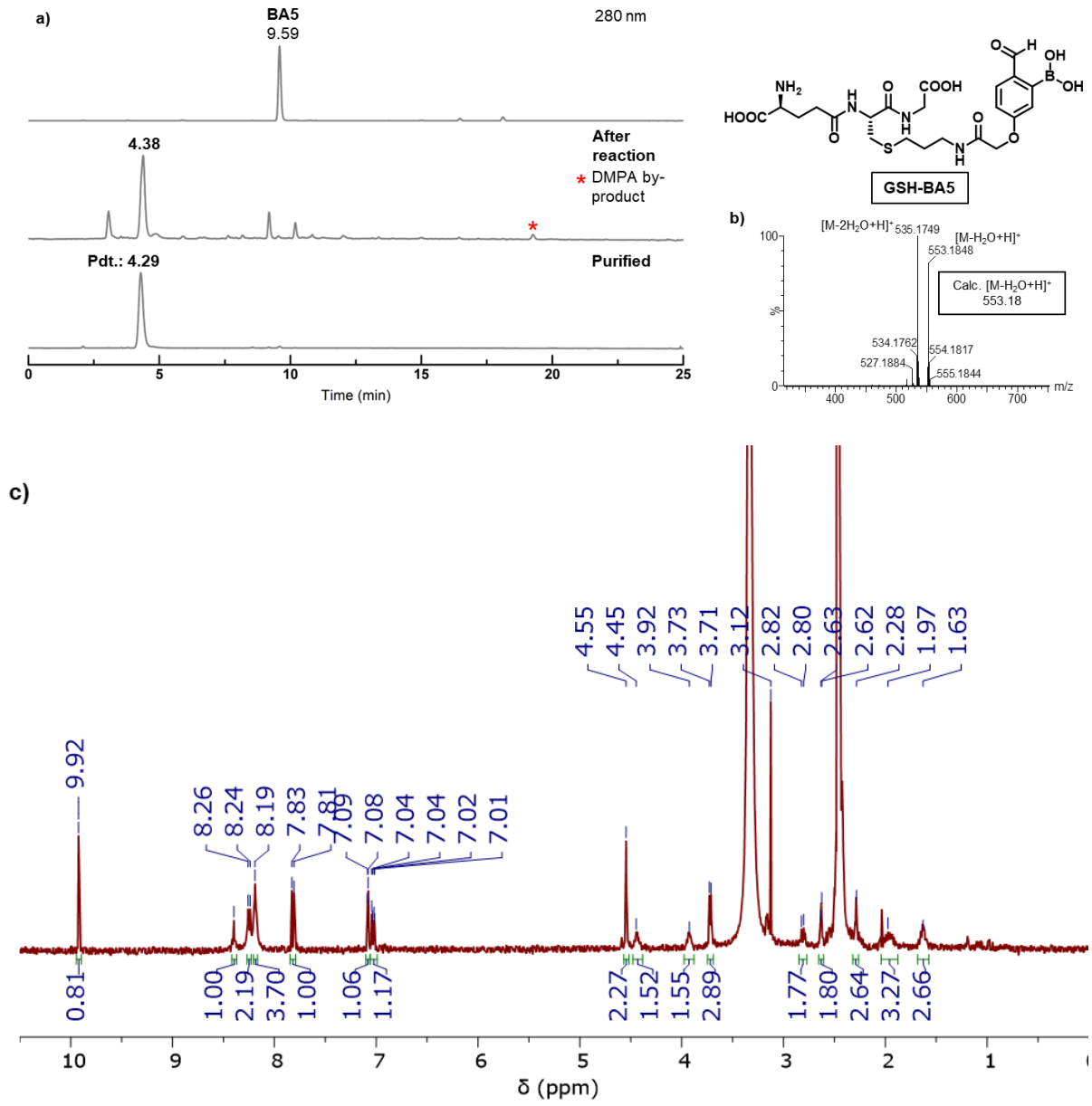


Figure 2.20: a) HPLC trace at 280 nm determines >95% conversion; b) Mass data; c) NMR spectra of pure product in DMSO-d_6 (400 MHz). 68% yield

GSH-BA6

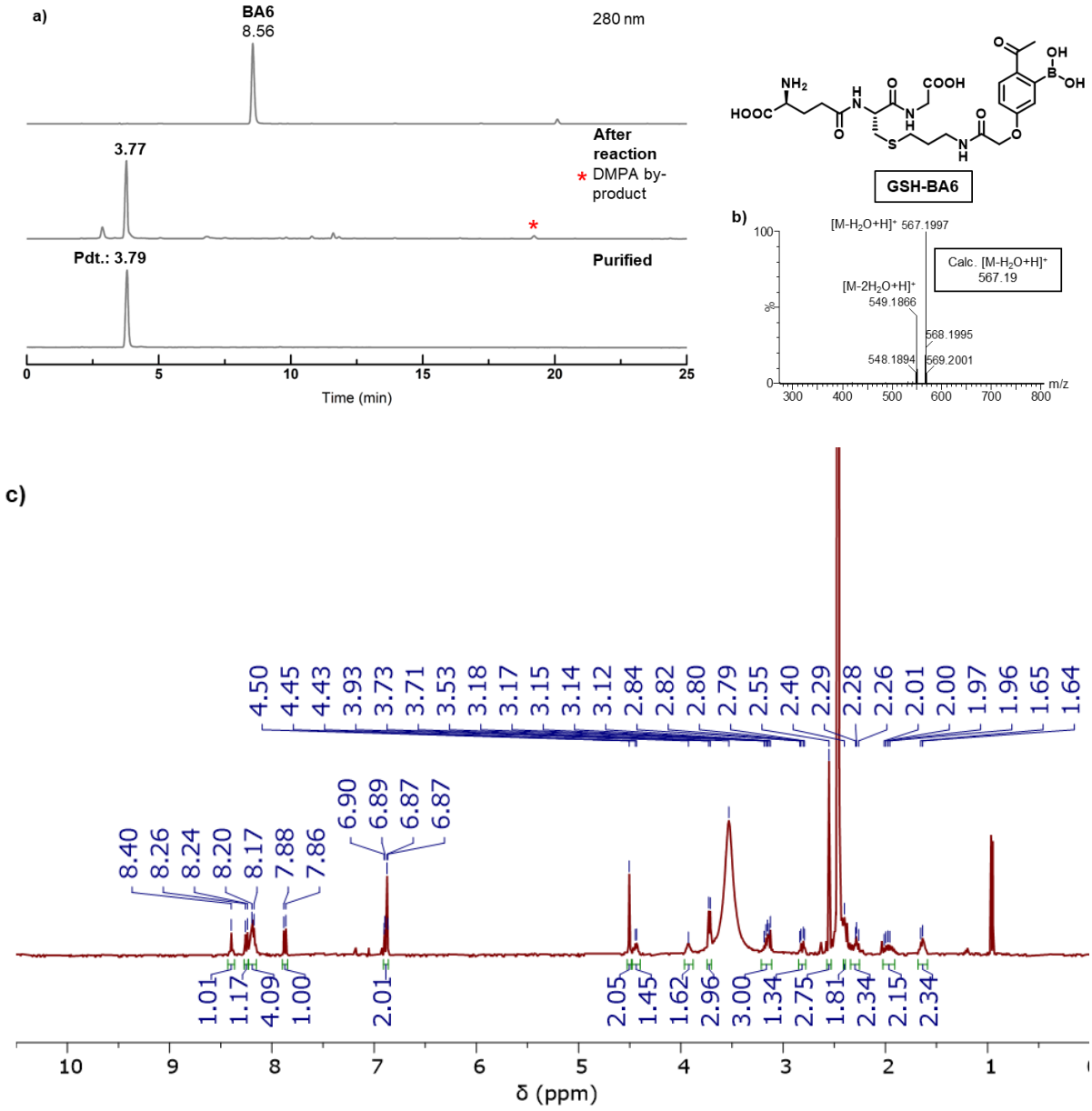


Figure 2.21: a) HPLC trace at 280 nm determines >95% conversion; b) Mass data; c) NMR spectra of pure product in DMSO-d_6 (400 MHz). 72% yield

GSH-BA7

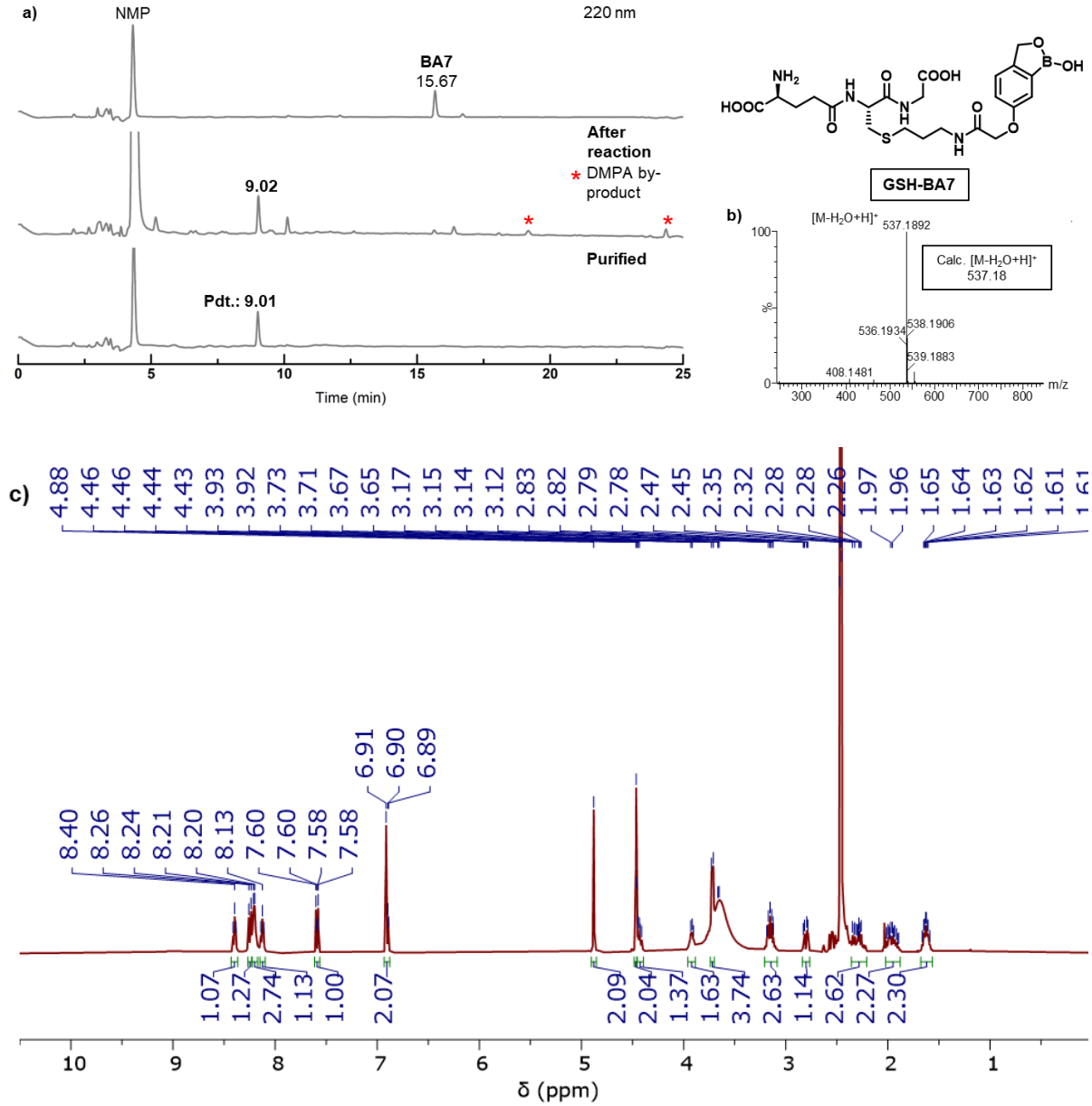


Figure 2.22: a) HPLC trace at 220 nm determines >95% conversion; b) Mass data; c) NMR spectra of pure product in DMSO-d_6 (400 MHz). 64% yield

GSH-BA8

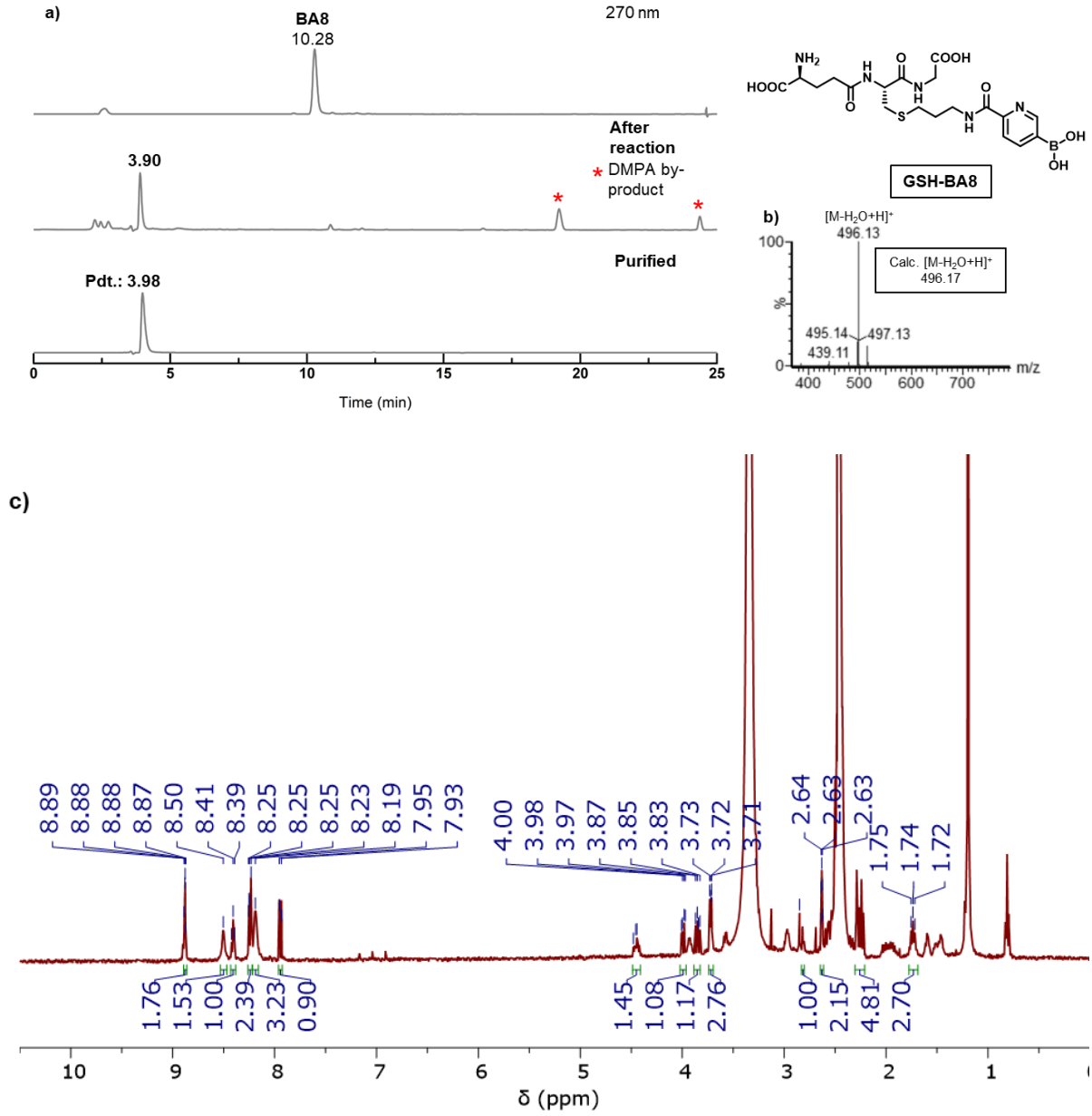
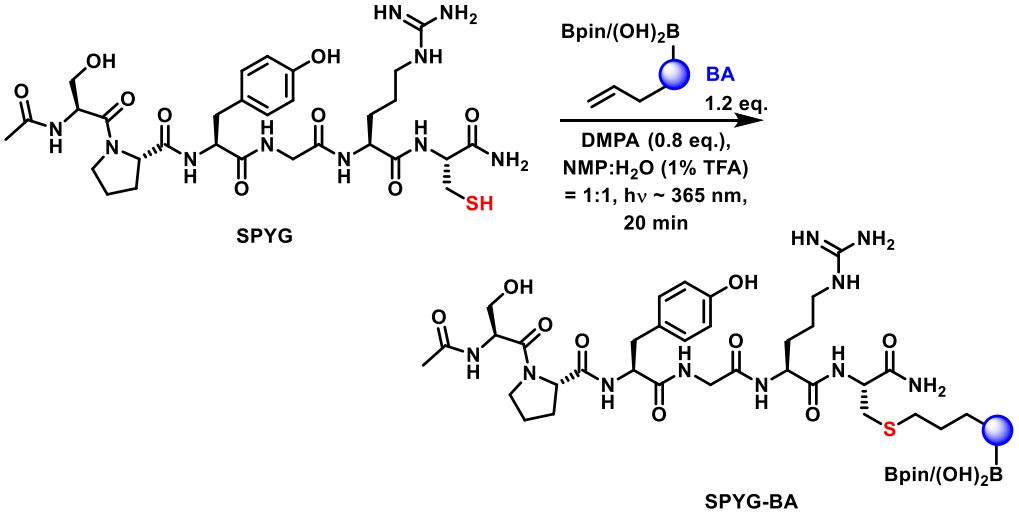


Figure 2.23: a) HPLC trace at 270 nm determines >95% conversion; b) Mass data; c) NMR spectra of pure product in DMSO-d_6 (400 MHz). 65% yield

General procedure with SPYG, Oxytocin, UBI, ATN-161, and WGA: A particular peptide (10 mM) was reacted with 12 mM BA derivatives and 8 mM DMPA for 20 min in NMP: H_2O (1% TFA) = 1:1 as the solvent and hv ~365 nm light source. The HPLC trace and ESI-MS data for the reactions are provided. In some reactions, an oxidized product (S=O) was formed in trace amounts (denoted in brackets in the figure captions). It was confirmed by mass spectra and eluted just after the major product in the HPLC, as indicated by a green arrow in the HPLC chromatogram.

2.5.8.2 *REACTIONS WITH SPYG*



SPYG-BA1

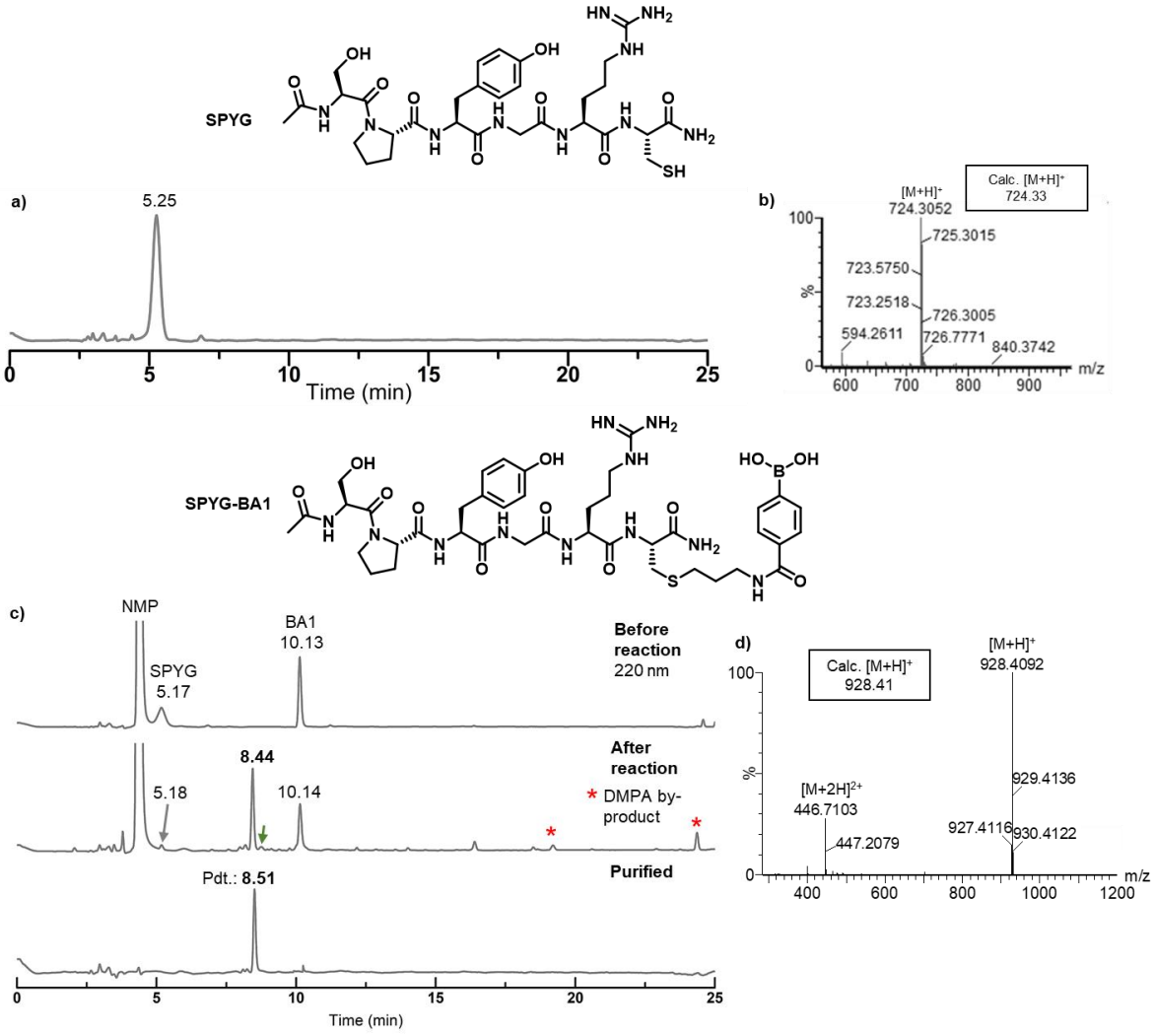


Figure 2.24: SPYG data: a) HPLC trace at 220 nm (94% purity); b) Mass data; SPYG-BA1 data: c) HPLC trace at 220 nm, 90% (4% oxidised product) conversion; d) Mass data. 90% (4% oxidized product) conversion

SPYG-BA2

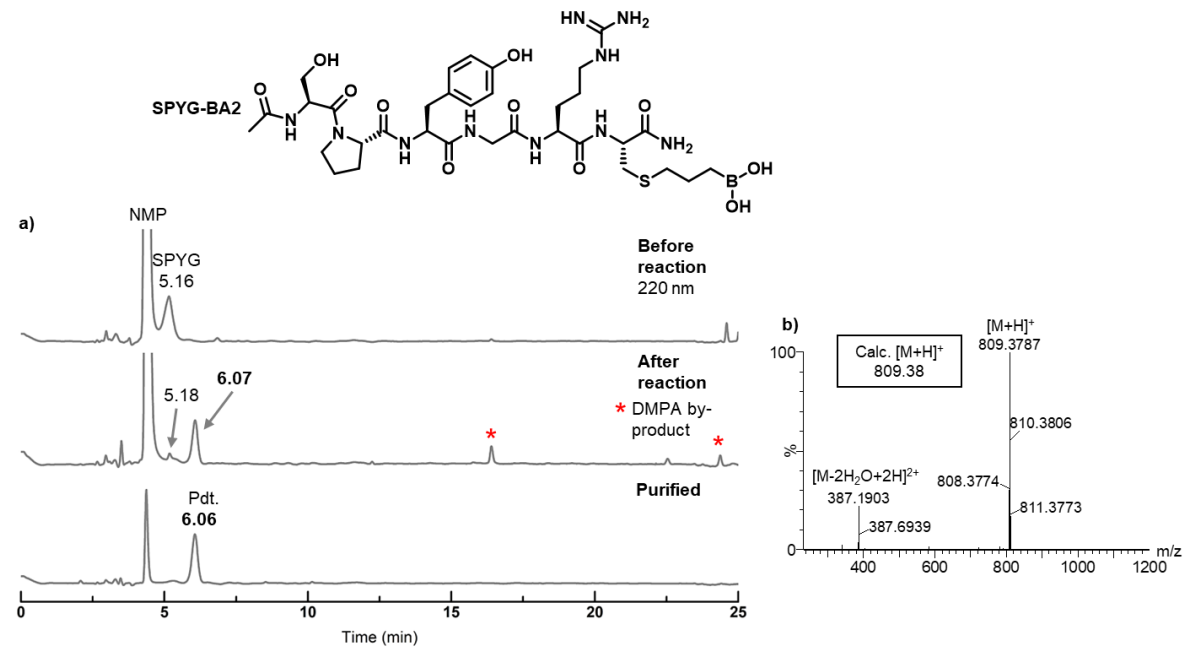


Figure 2.25: a) HPLC trace at 220 nm, >95% conversion (peak ~4.5 min is due to residual NMP); b) Mass data. 70% yield

SPYG-BA5

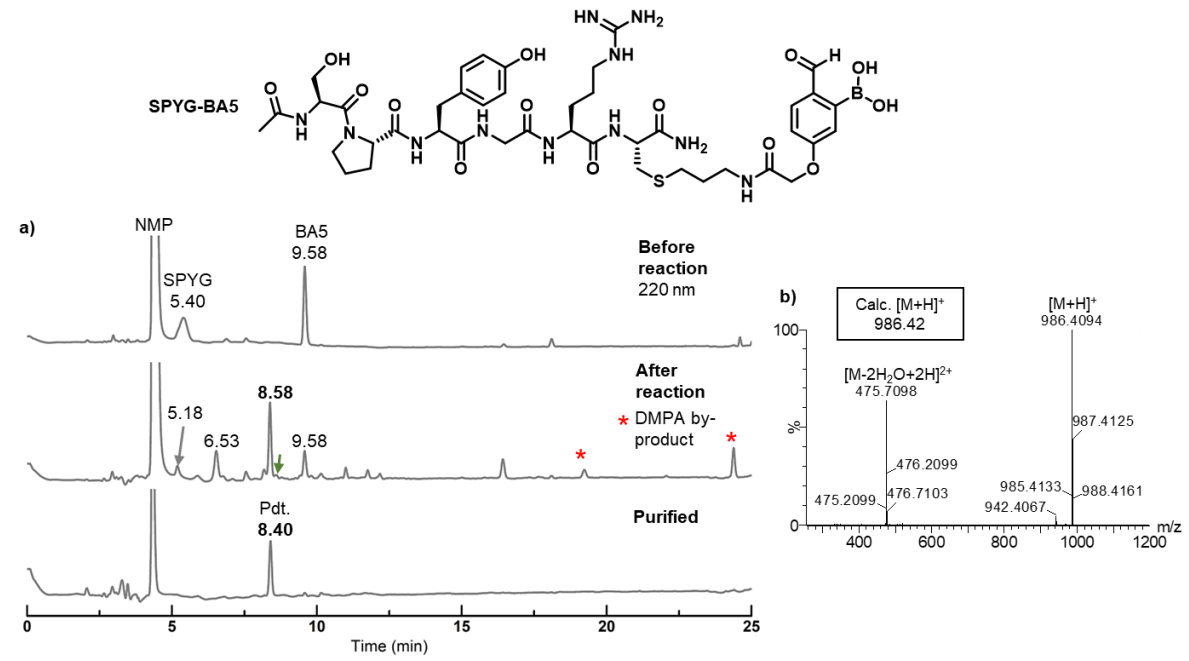
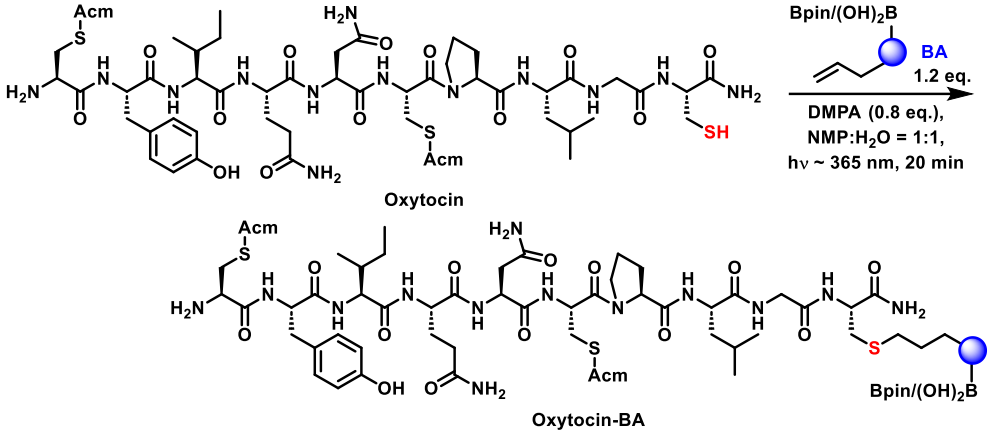


Figure 2.26: a) HPLC trace at 220 nm, 87% (5% oxidized product) conversion (peak ~4.5 min is due to residual NMP); b) Mass data. 62% yield

2.5.8.3 *REACTIONS WITH OXYTOCIN*



Oxytocin-BA1

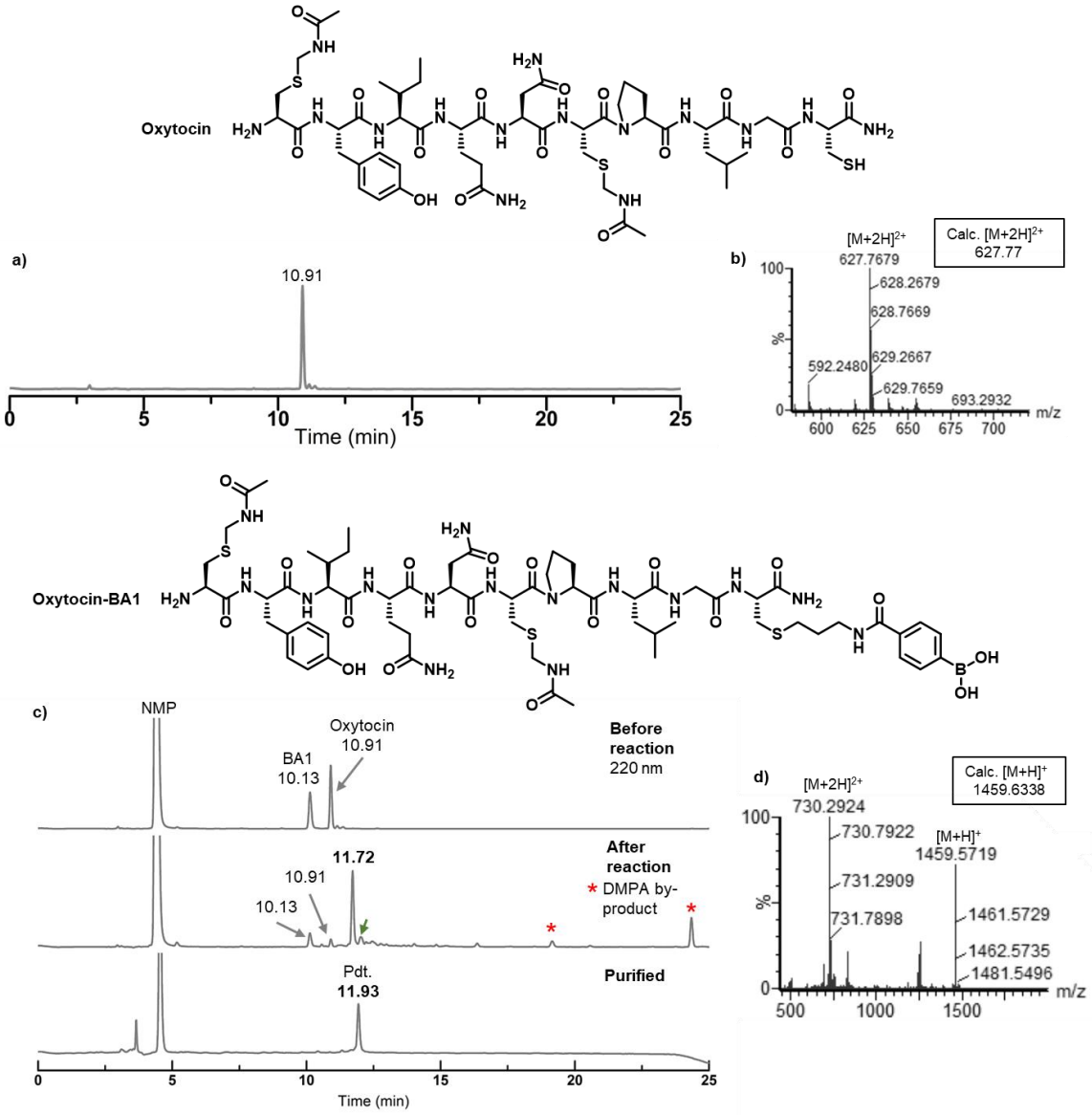


Figure 2.27: Oxytocin data: a) HPLC trace at 220 nm (95% purity); b) Mass data; Oxytocin-BA1 data: c) HPLC trace at 220 nm; 83% (10% oxidised product) conversion; d) Mass data. 61% yield

Oxytocin-BA2

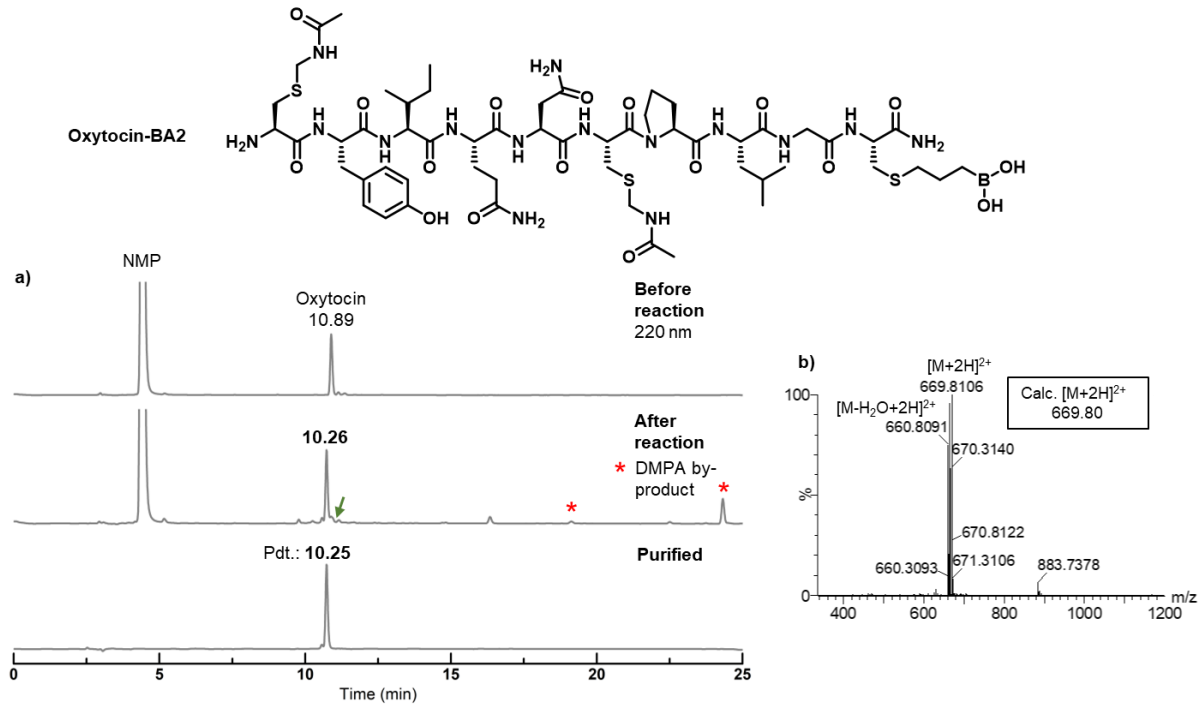


Figure 2.28: a) HPLC trace at 220 nm; 84% (10% oxidised product) conversion; b) Mass data. 62% yield

Oxytocin-BA5

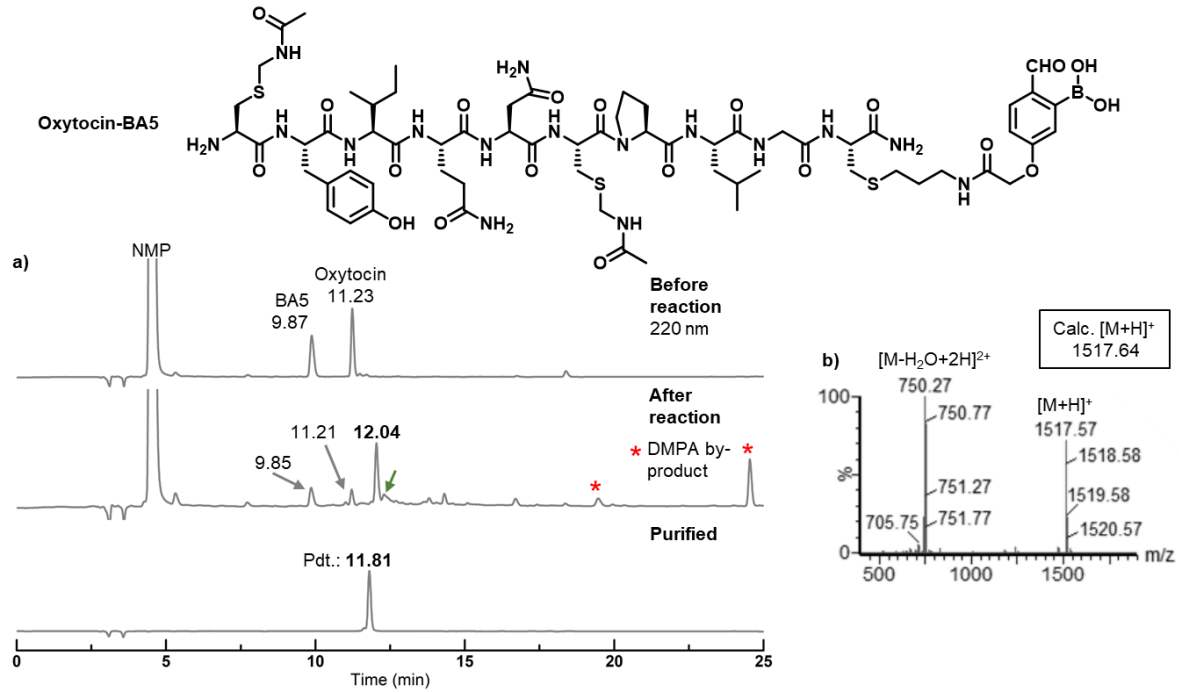


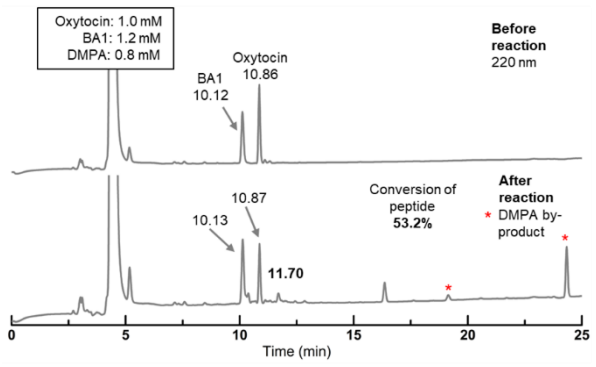
Figure 2.29: a) HPLC trace at 220 nm; 88% (5% oxidised product) conversion; b) Mass data. 67% yield

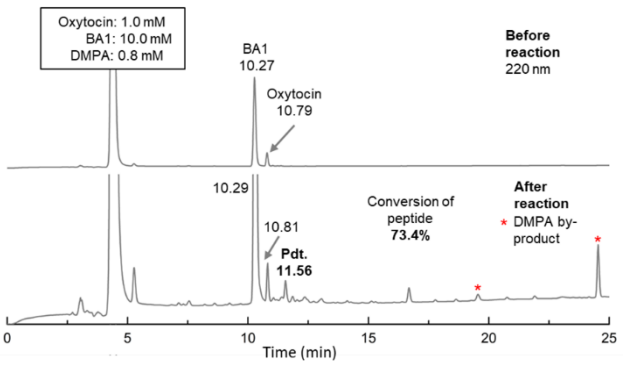
OPTIMIZATION OF THIOL-ENE REACTION AT LOW MILLIMOLAR CONCENTRATION

We investigated the most satisfactory conditions required for the thiol-ene click reaction to perform at low millimolar concentration. Therefore, the reaction conditions were optimized by taking oxytocin as the model peptide and BA1. At 10 mM peptide concentration, we have achieved a conversion of ~93% with BA1 in 20 min. Considering the excellent conversion as a standard, we tried to check whether a similar conversion could be achieved at lower peptide concentrations. However, we found the reaction is less productive at low concentrations. Therefore, we started optimizing reactions at lower peptide concentrations given higher concentrations of BA/DMPA or longer reaction times. Table 2.2**,** Figures 2.30 and 2.31 present all the data analyzed by analytical RP-HPLC. In these studies, we first kept 1.2 eq. BA and 0.8 eq. DMPA constant, the reaction was carried out with 5 mM, 3 mM, and 1 mM peptide concentration, and a linear decrease in the peptide conversion were observed (Entry 1, 2, 6). Next, we focused on improving the conversion rate of 3 mM and 1 mM peptide concentrations. We found that increasing the BA concentration to ~3 eq. At the same time, we kept DMPA constant, increasing the conversion by ~10% (Entry 3, 7) only. We speculated that quenching of the photocatalyst might be the possible reason for the lower conversion. Using ~2.66 eq. of DMPA and ~3 eq. of BA led to an excellent conversion of >90% (Entry 4). We were curious to check whether longer reaction times improve the conversion. Keeping peptide concentration at 3 mM, we studied the peptide conversion at intervals of 15 min and up to 60 min. All the reactions proceeded to completion within 15 min, and longer reaction times did not improve the product conversion (Figure 2.31). All the experiments were repeated twice, which are consistent, and one of the results is presented.

Table 2.2. Optimization table for Oxytocin. Time = 20 min.

Sl. No.	Oxytocin	BA1	DMPA	GSH	Peptide conversion
1.	5 mM	6.0 mM	4.0 mM	-	86%
2.	3 mM	3.6 mM	2.4 mM	-	76%
3.	3 mM	10.0 mM	2.4 mM	-	82%
4.	3 mM	10.0 mM	8.0 mM	-	92%
5.	1 mM	1.2 mM	0.8 mM	-	53%
6.	1 mM	10.0 mM	0.8 mM	-	73%





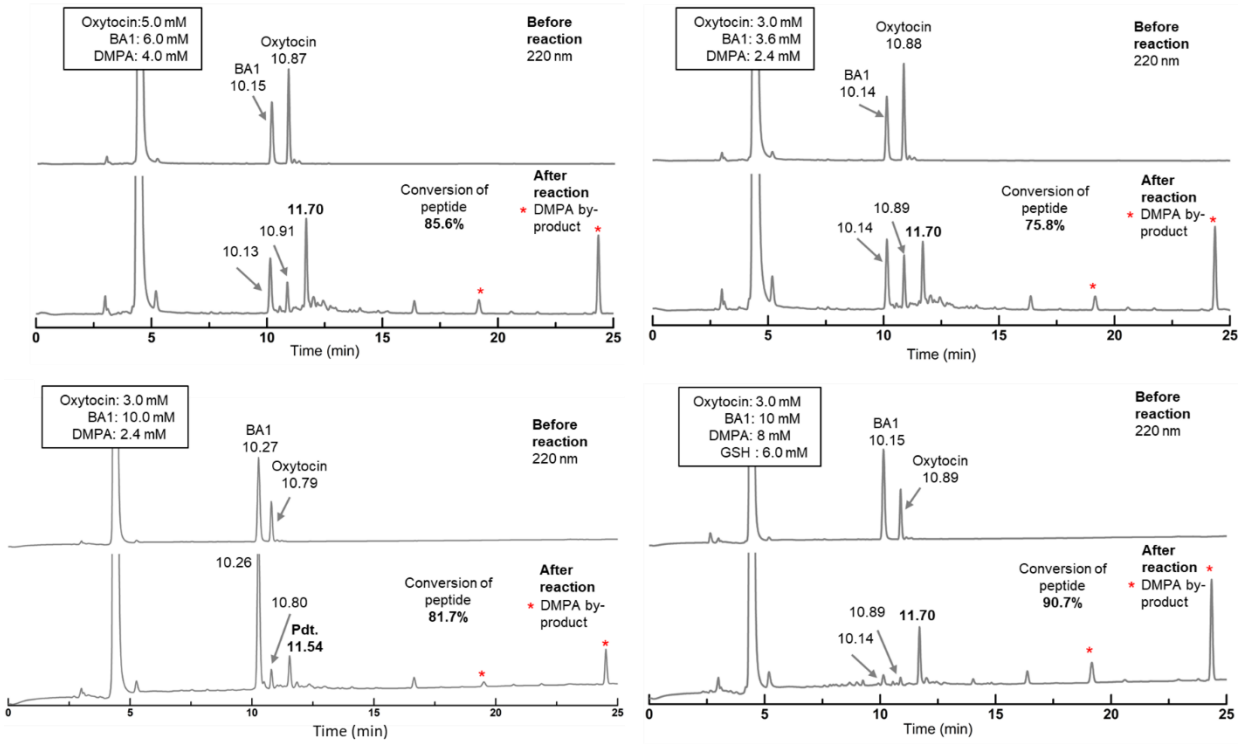
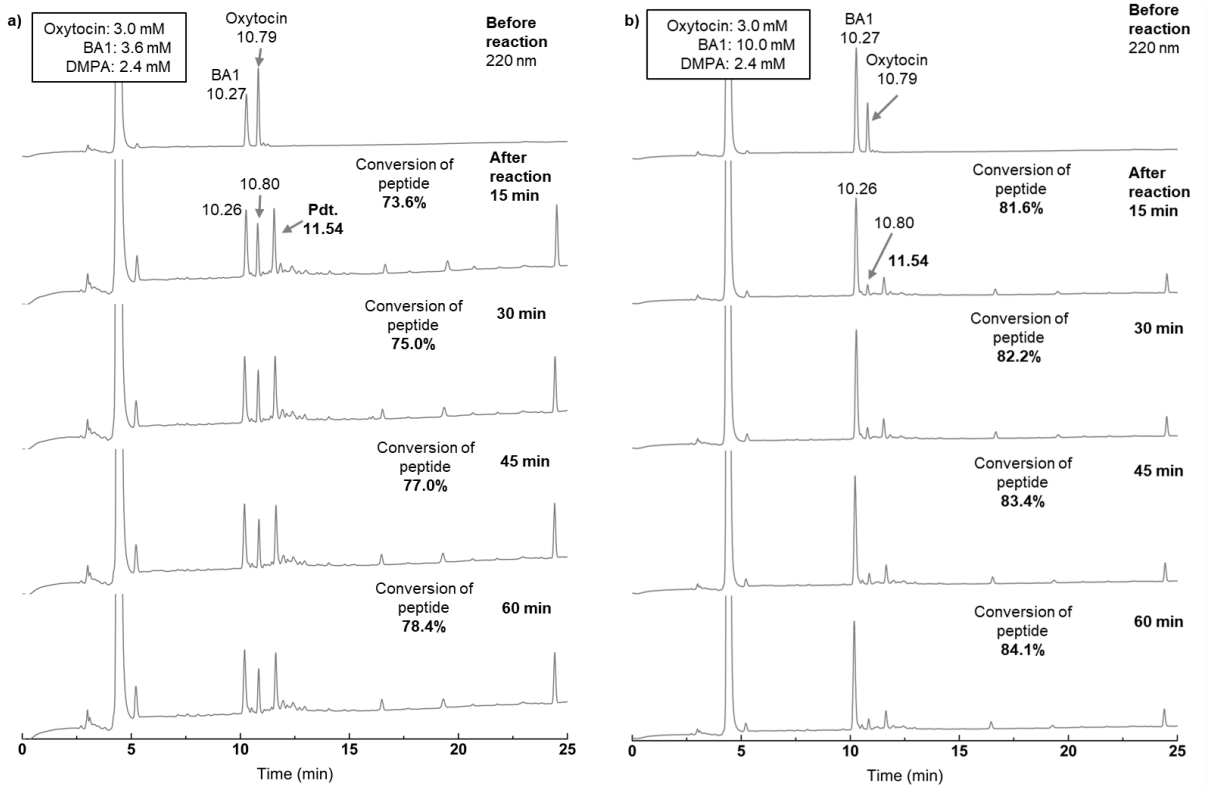


Figure 2.30: HPLC traces at 220 nm for the optimization reactions of Oxytocin. The peak observed at ~4.8 min is due to NMP.



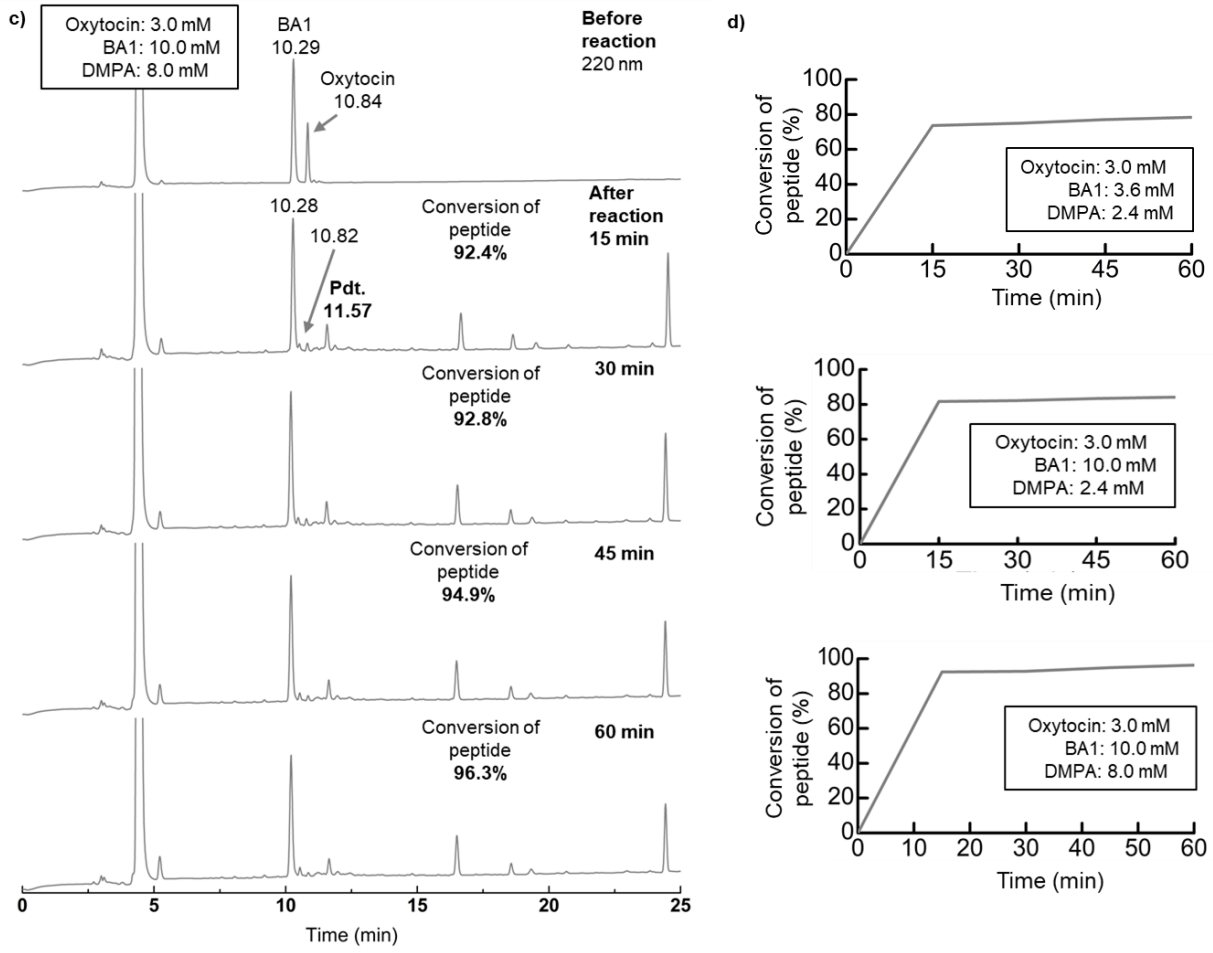
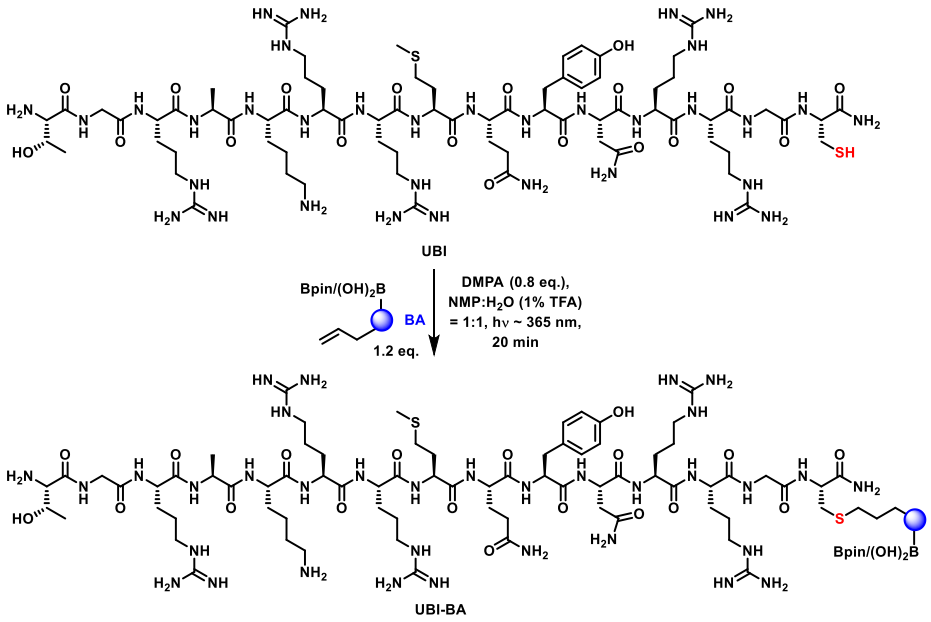


Figure 2.31: Optimization reactions of Oxytocin at 3 mM showing the conversion rate of the peptide with time. Reactions (a), (b) and (c) have been summarised graphically in (d). The peak observed at ~4.8 min is due to NMP. For clarity, DMPA peaks have not been marked.

2.5.8.4 *REACTIONS WITH UBI*



UBI-BA1

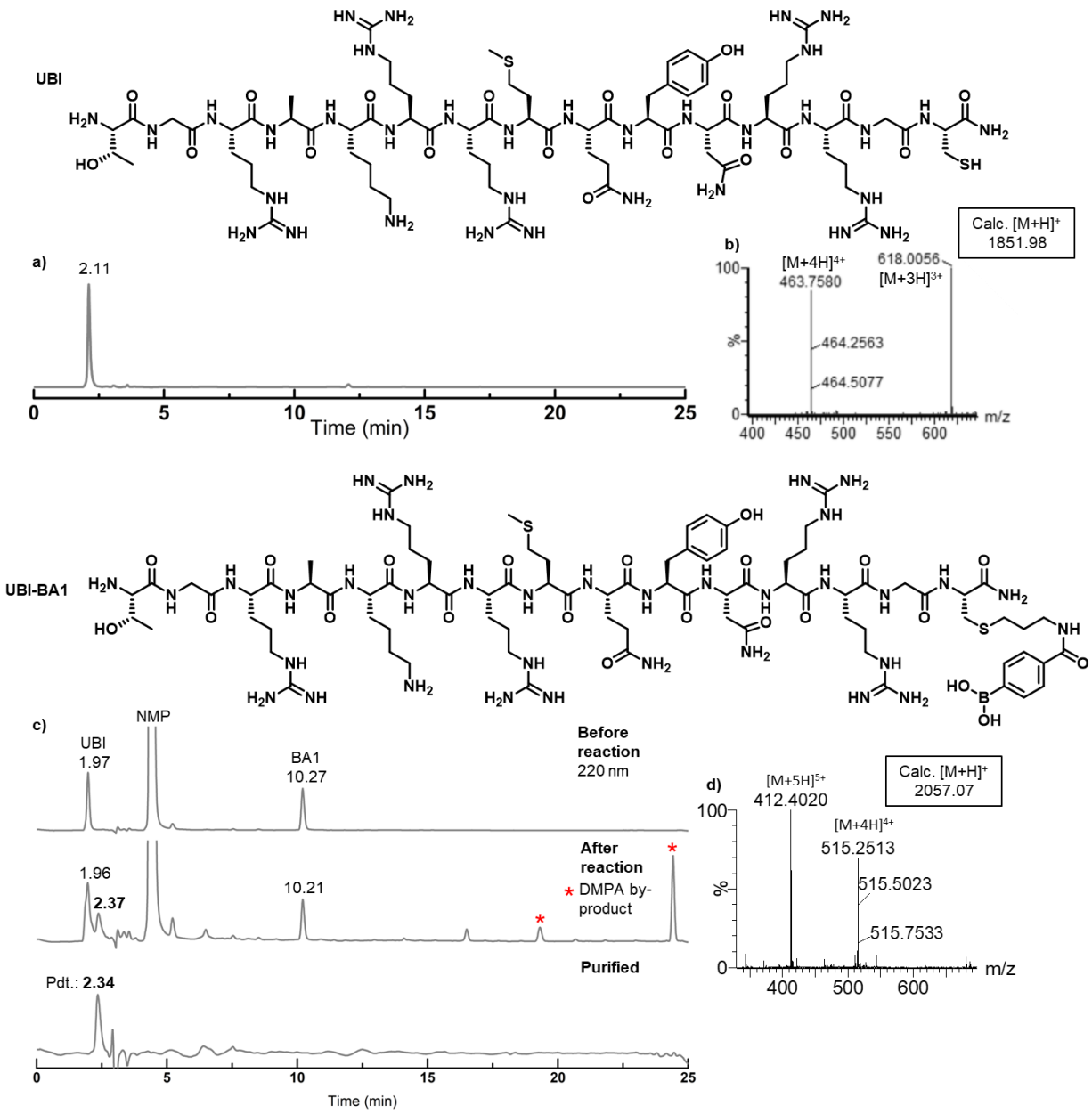


Figure 2.32: UBI data: a) HPLC trace at 220 nm (>95% purity); b) Mass data; UBI-BA1 data: c) HPLC trace at 220 nm, 47.5% conversion; d) Mass data. 38% yield

UBI-BA5

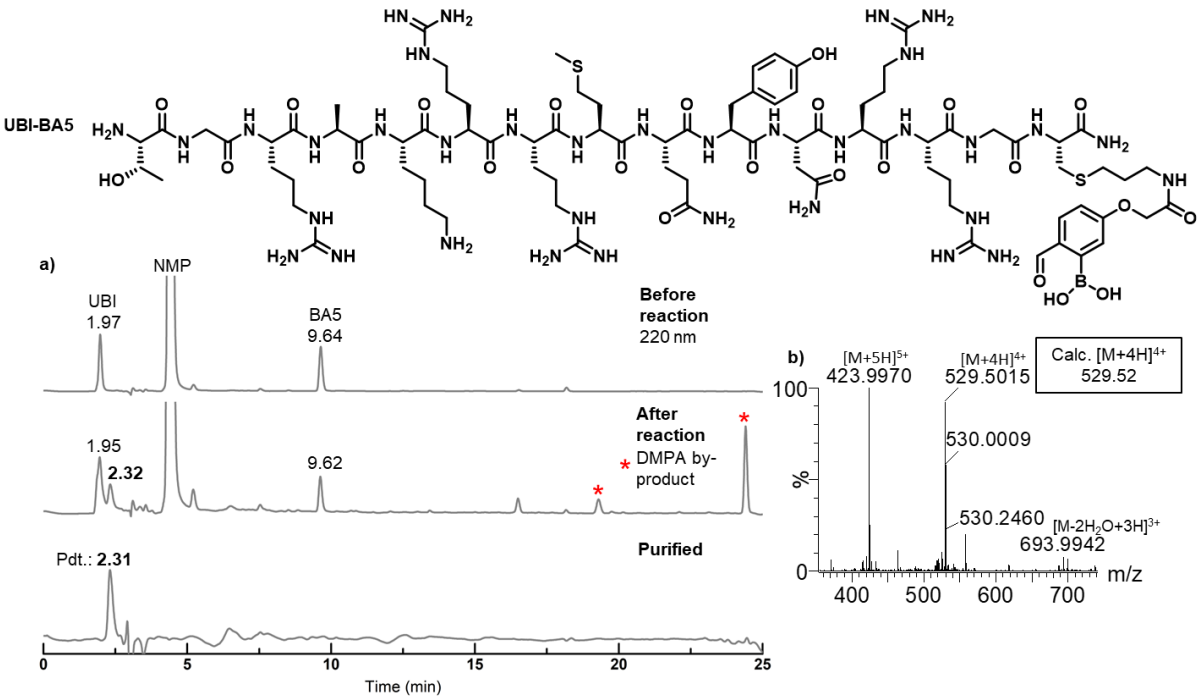
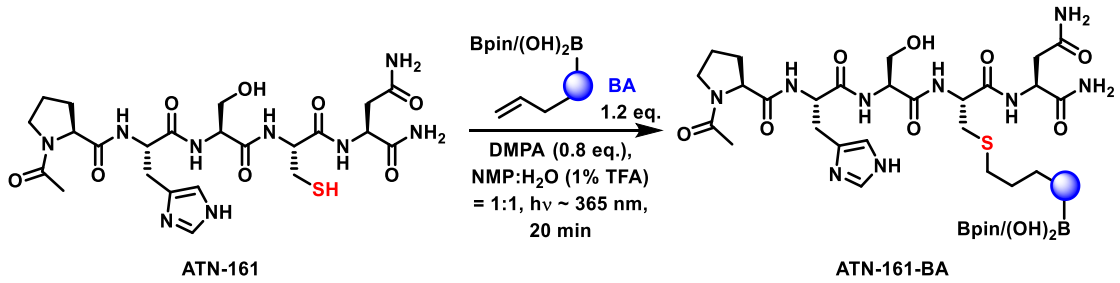
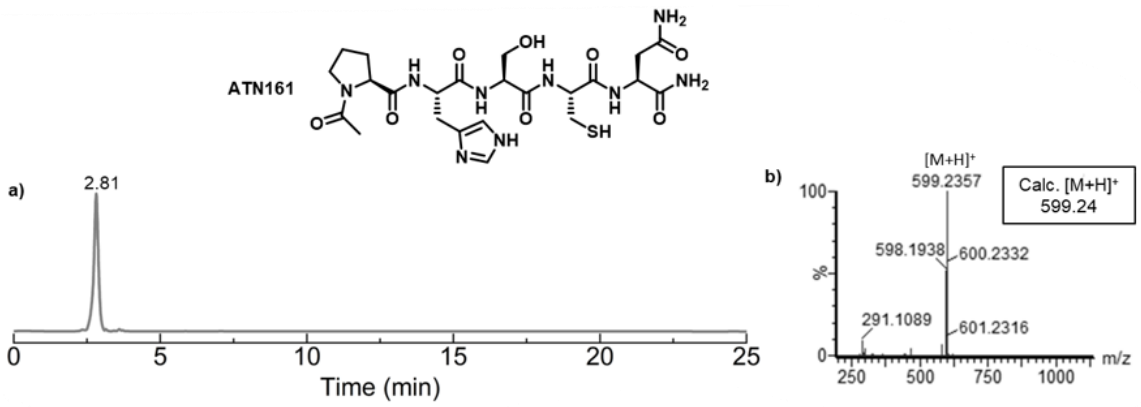


Figure 2.33**:** a) HPLC trace at 220 nm, 54.6% conversion; b) Mass data. 45% yield

2.5.8.5 *REACTIONS WITH ATN-161*



ATN161-BA1



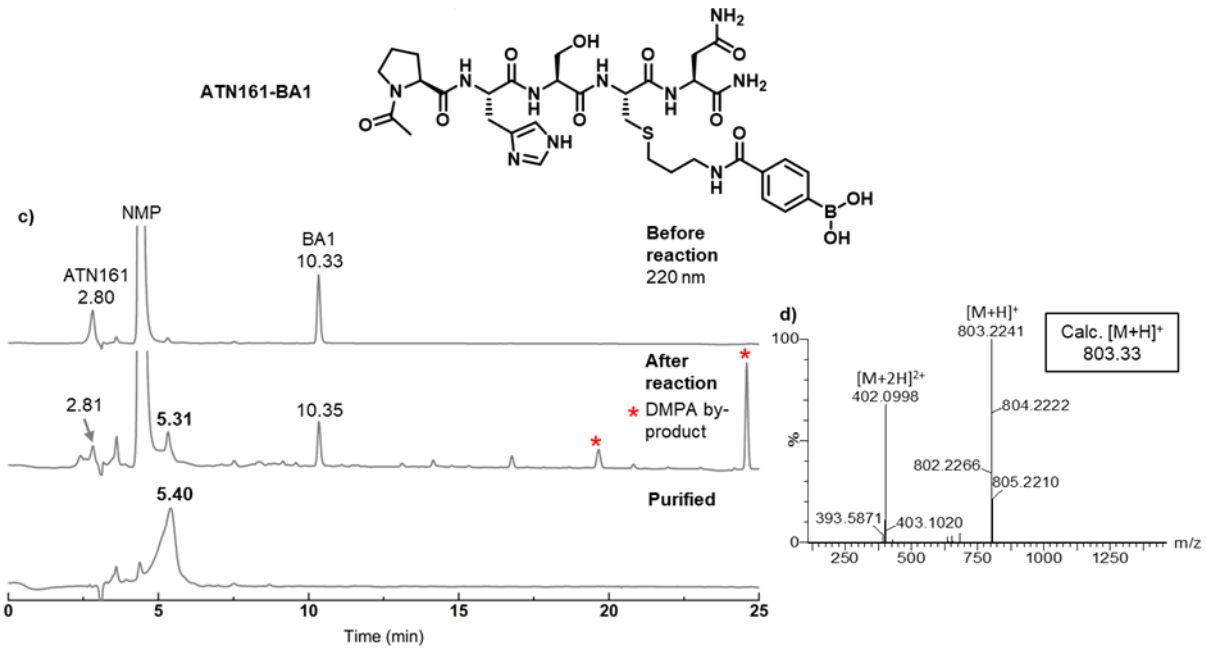


Figure 2.34: ATN161 data: a) HPLC trace at 220 nm (95% purity); b) Mass data; ATN161-BA1 data: c) HPLC trace at 220 nm, 85% conversion (ATN161-BA peptides dragged in column purification, thus purification step was difficult); d) Mass data. 60% yield

ATN161-BA5

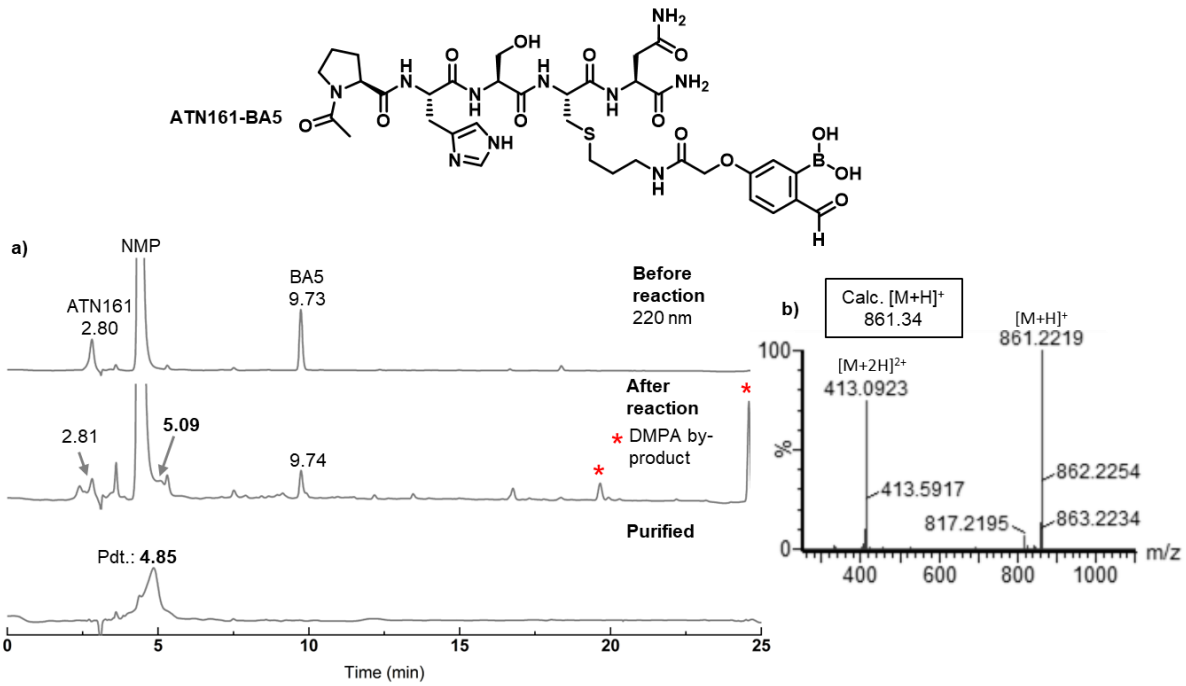


Figure 2.35: a) HPLC trace at 220 nm, 84% conversion; b) Mass data. 62% yield

ATN161-BA6

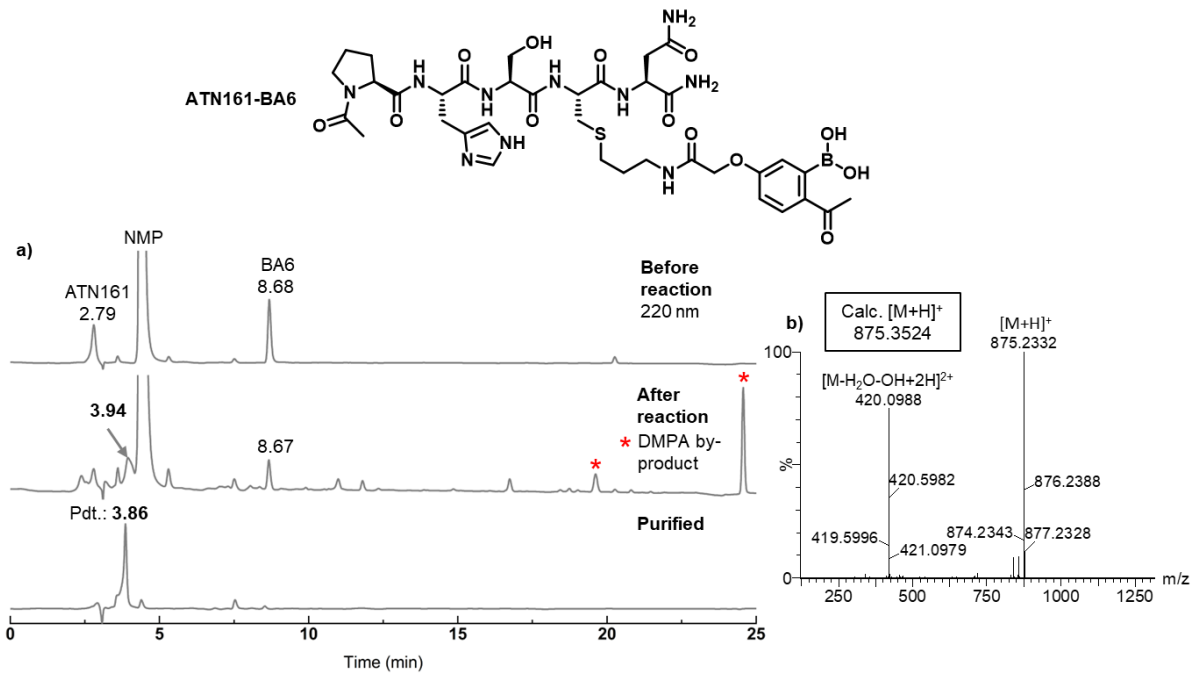
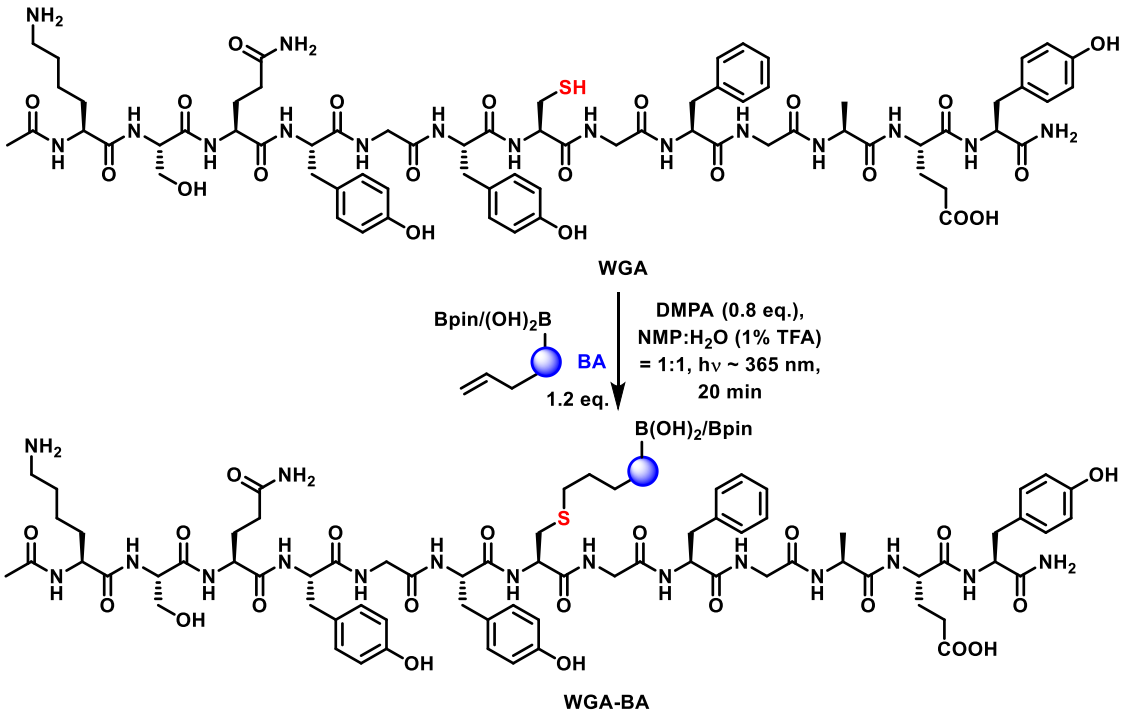


Figure 2.36: a) HPLC trace at 220 nm, 82% conversion; b) Mass data. 63% yield

2.5.8.6 *REACTIONS WITH WGA*



WGA-BA1

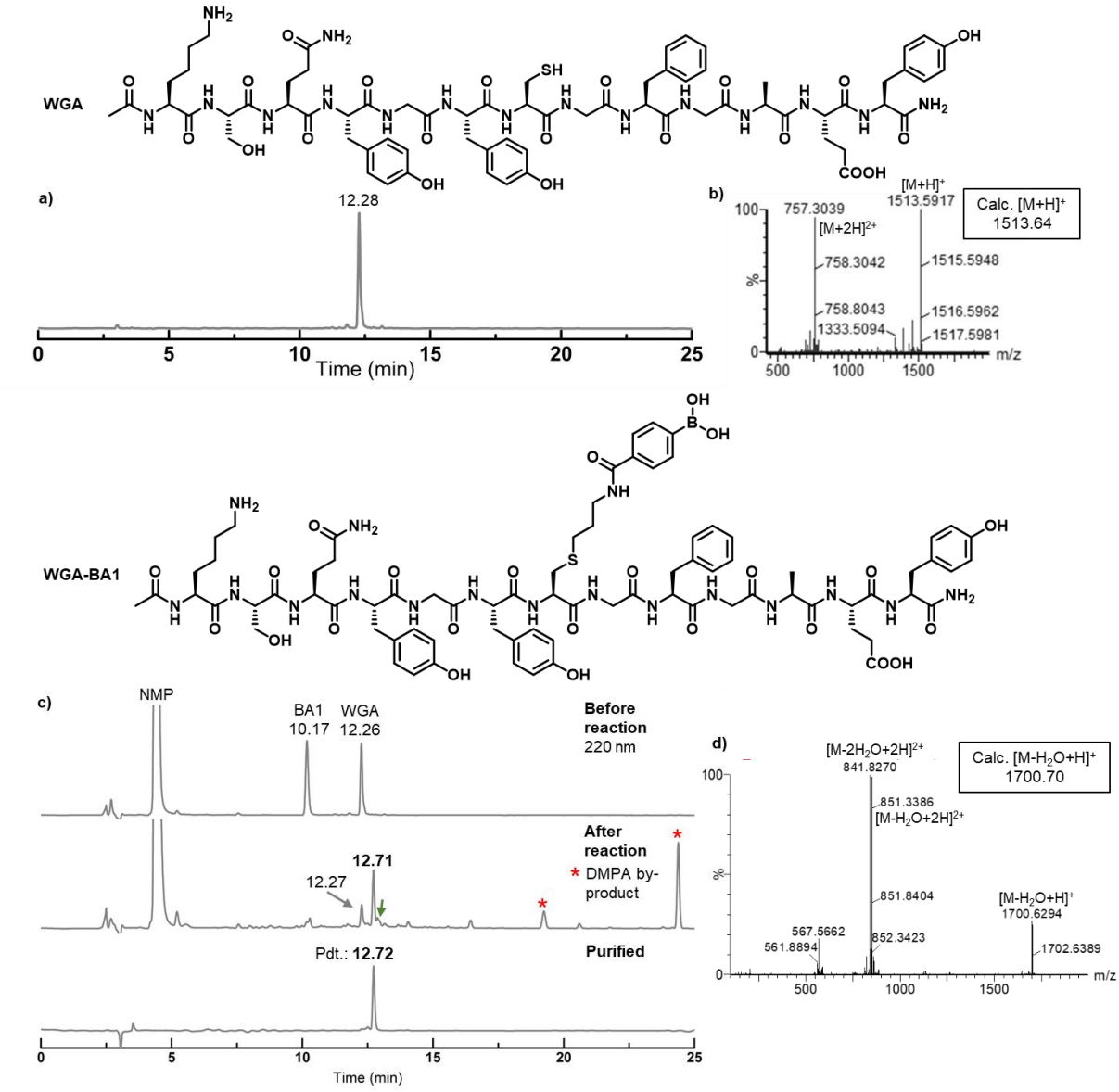


Figure 2.37: WGA data: a) HPLC trace at 220 nm (92% purity); b) Mass data; WGA-BA1 data: c) HPLC trace at 220 nm, 76% conversion; d) Mass data. 52% yield

WGA-BA2

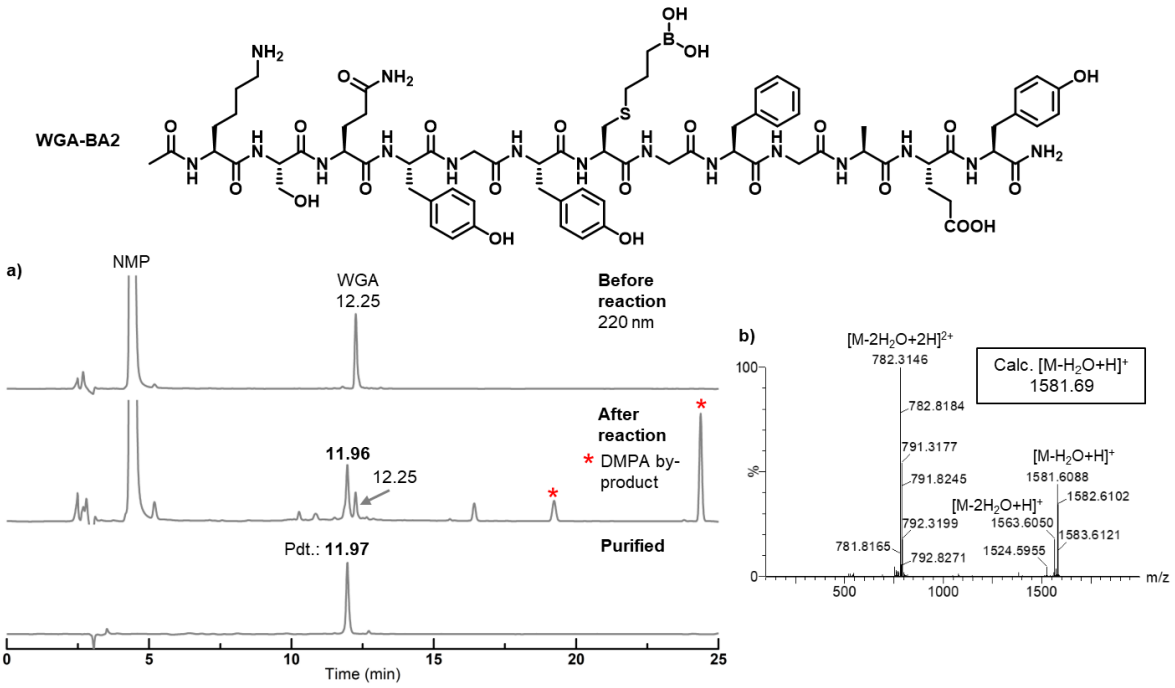


Figure 2.38: a) HPLC trace at 220 nm, 80% conversion; b) Mass data. 57% yield

WGA-BA5

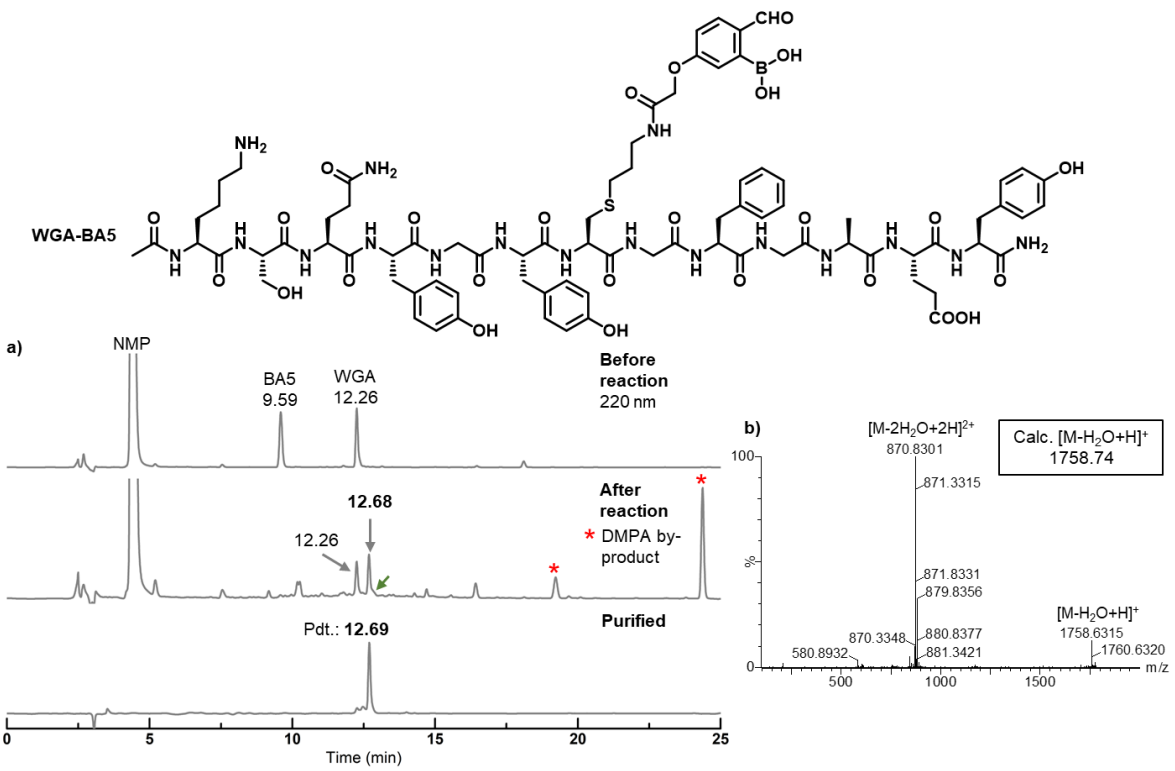
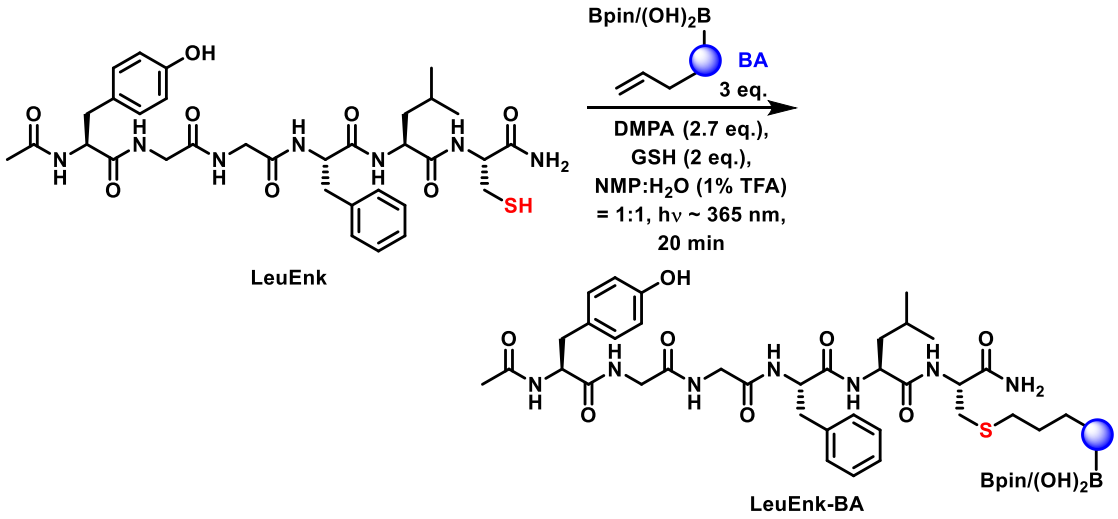


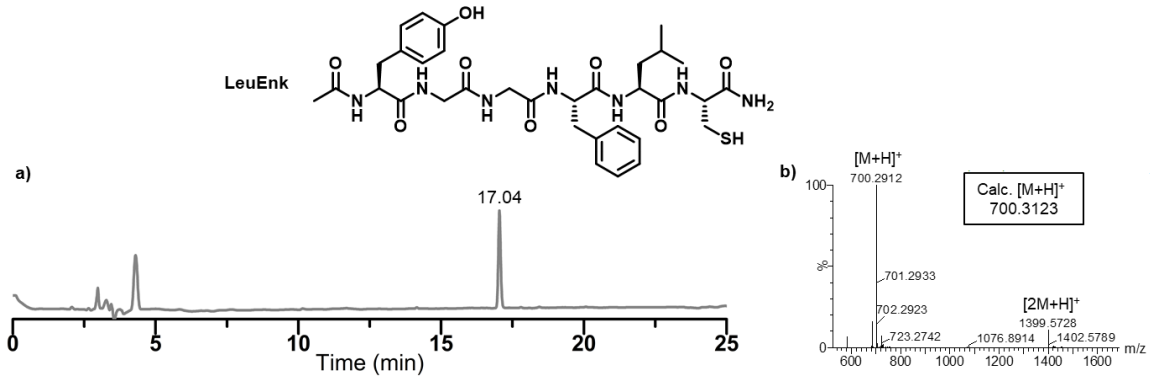
Figure 2.39: a) HPLC trace at 220 nm, 73% conversion; b) Mass data. 50% yield

General procedure with LeuEnk, CysLeuEnk and p53: All reactions were performed at 3 mM concentration due to counter the issue of peptide solubility. Unfortunately, the optimized reaction conditions did not help improve conversion (Figure 2.40e). We reasoned that adding GSH as an additive can help continue the chain propagation steps (as discussed in Section 2.3.3) and, eventually, result in a better conversion yield. Peptides (3 mM) were reacted with 10 mM of BA derivatives in the presence of 6 mM GSH and 8 mM DMPA for 20 min in NMP:H_2O (1% TFA) = 1:1 as the solvent and hν ~365 nm light source. The HPLC trace and ESI-MS data for the reactions are provided. In some reactions, an oxidized product (S=O) was formed in trace amounts (denoted in brackets in the figure captions). It was confirmed by mass spectra and eluted just after the major product peak in the HPLC chromatogram, as indicated by a green arrow in the HPLC trace. The product formation with GSH is also denoted in the spectra. In some cases, a clear separation between the parent peptide and product was not observed during purification by RP-HPLC; hence the mass of the parent peptide can also be observed in the data.

2.5.8.7 *REACTIONS WITH LEUENK*



LeuEnk-BA1



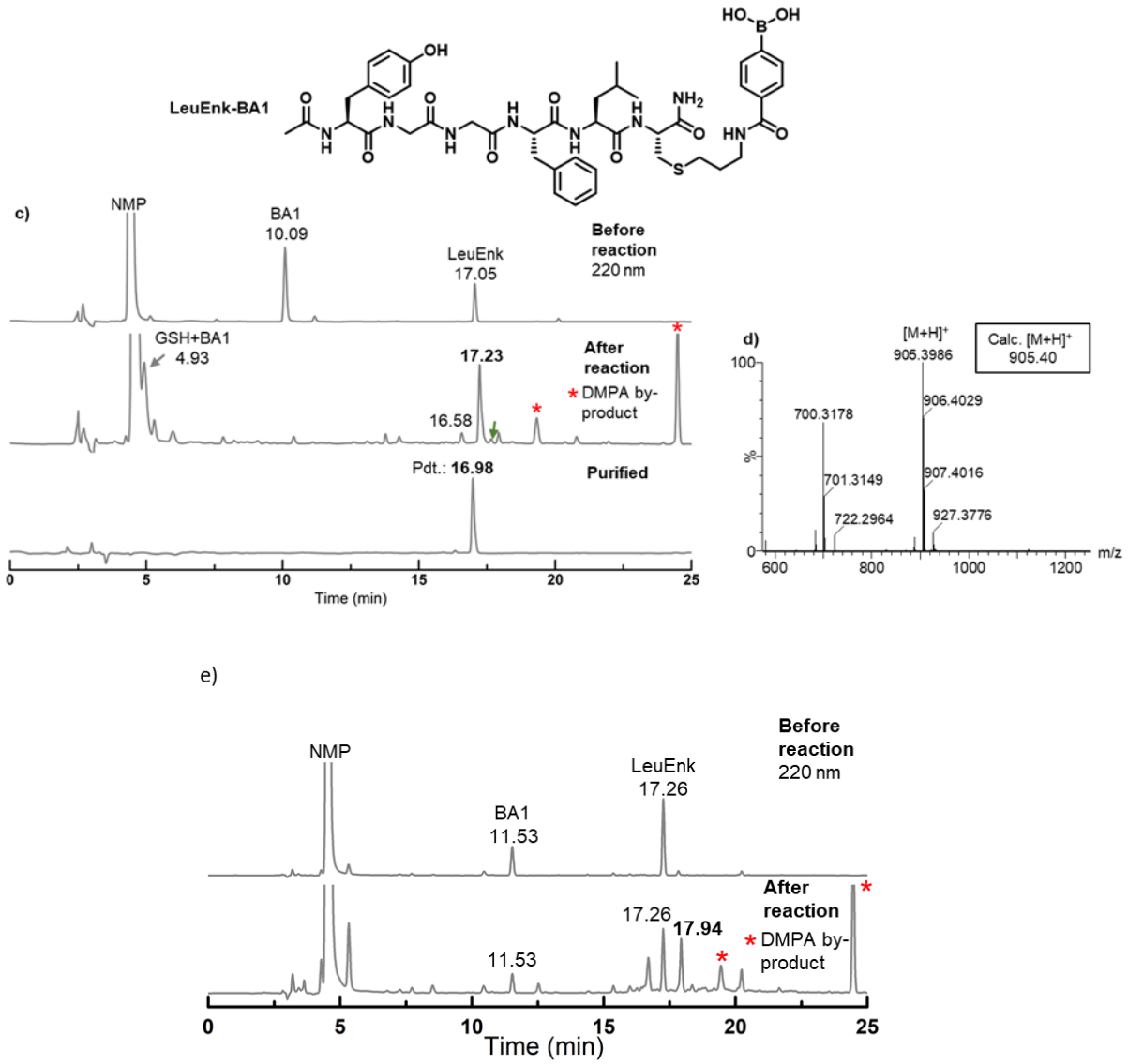


Figure 2.40: LeuEnk data: a) HPLC trace at 220 nm (>95% purity, the peak at ~4.5 min is from column contamination); b) Mass data; LeuEnk-BA1 data: c) HPLC trace at 220 nm, 72% (8% oxidised product) conversion. 60% yield. The retention times have changed as the analysis was performed with a new batch of mobile phase on different days. The product formation was confirmed by LC-MS; d) Mass data (the product purification was not clean, so the mass of LeuEnk can be seen); e) Thiol-ene reaction without GSH. The reaction conversion is 60%, and product conversion is 40% by HPLC area. The reaction is unclean.

LeuEnk-BA2

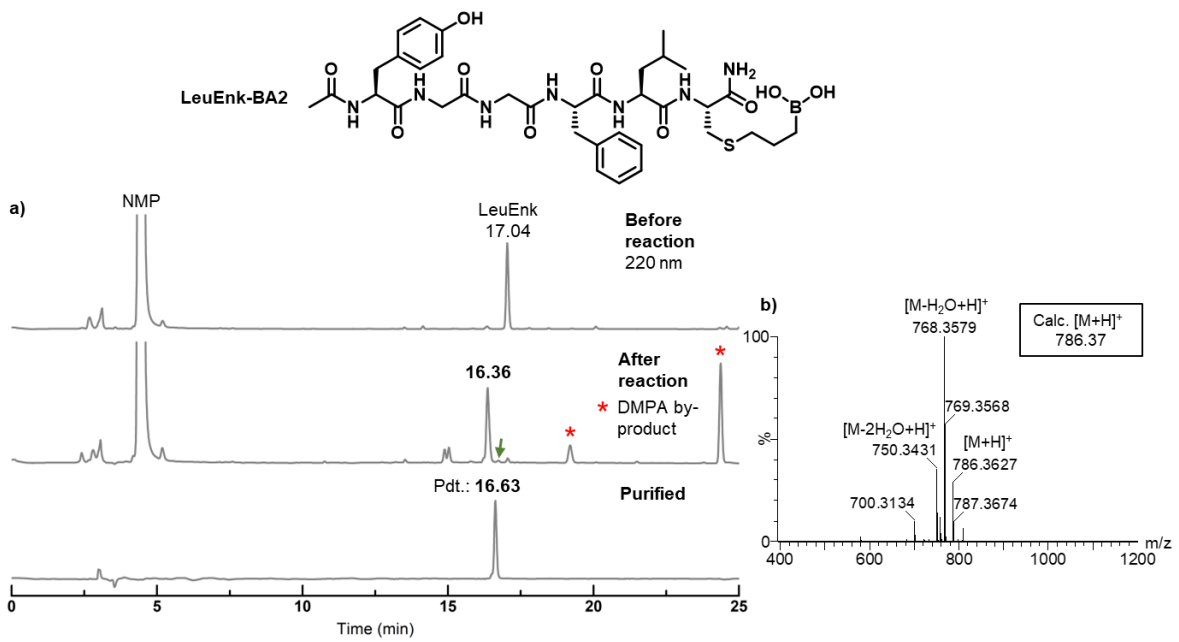


Figure 2.41: a) HPLC trace at 220 nm, 92% (5% oxidised product) conversion; b) Mass data. 78% yield

LeuEnk-BA4

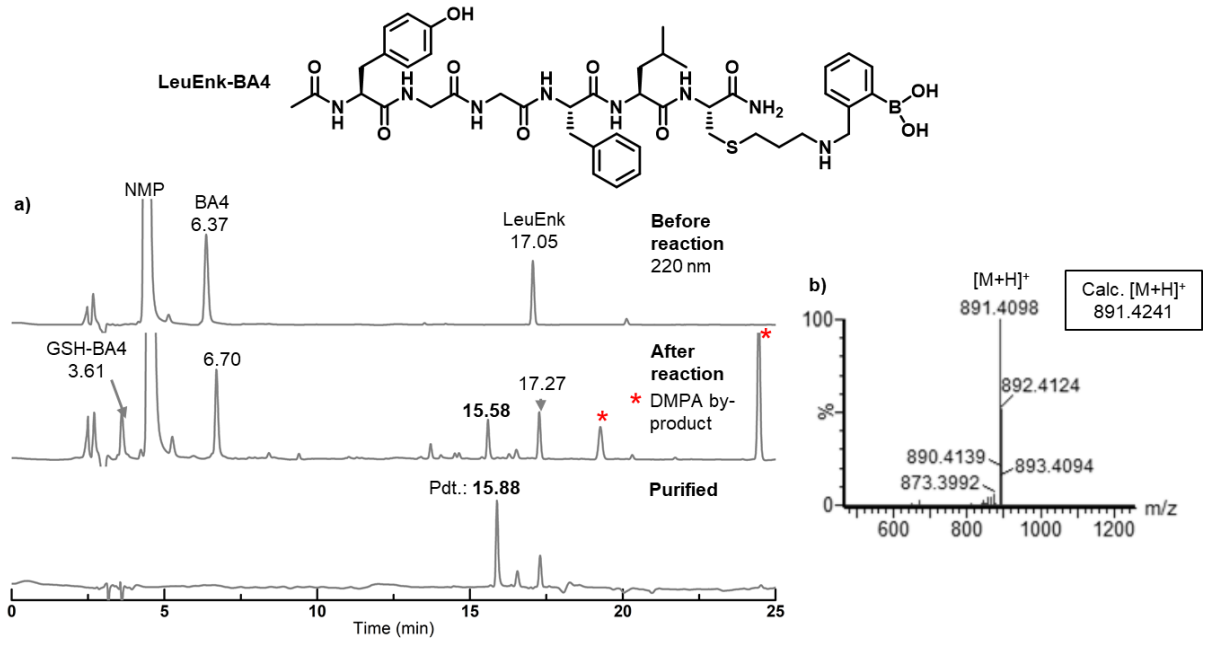


Figure 2.42: a) HPLC trace at 220 nm, 66% conversion; b) Mass data. 52% yield

LeuEnk-BA6

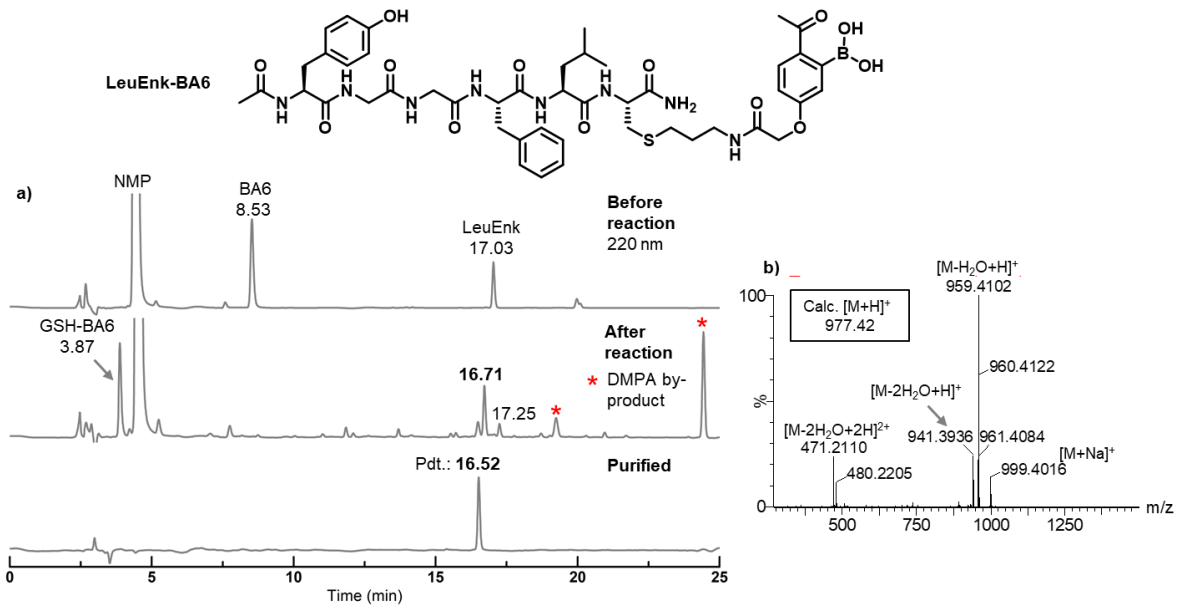
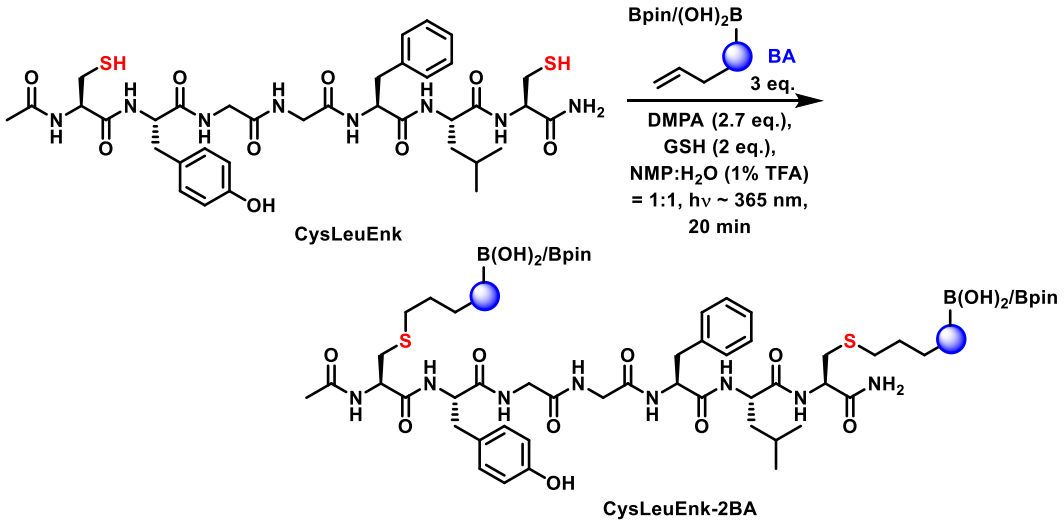


Figure 2.43: a) HPLC trace at 220 nm, 86% conversion; b) Mass data. 61% yield

2.5.8.8 *REACTIONS WITH CYSLEUENK*



CysLeuEnk-BA1

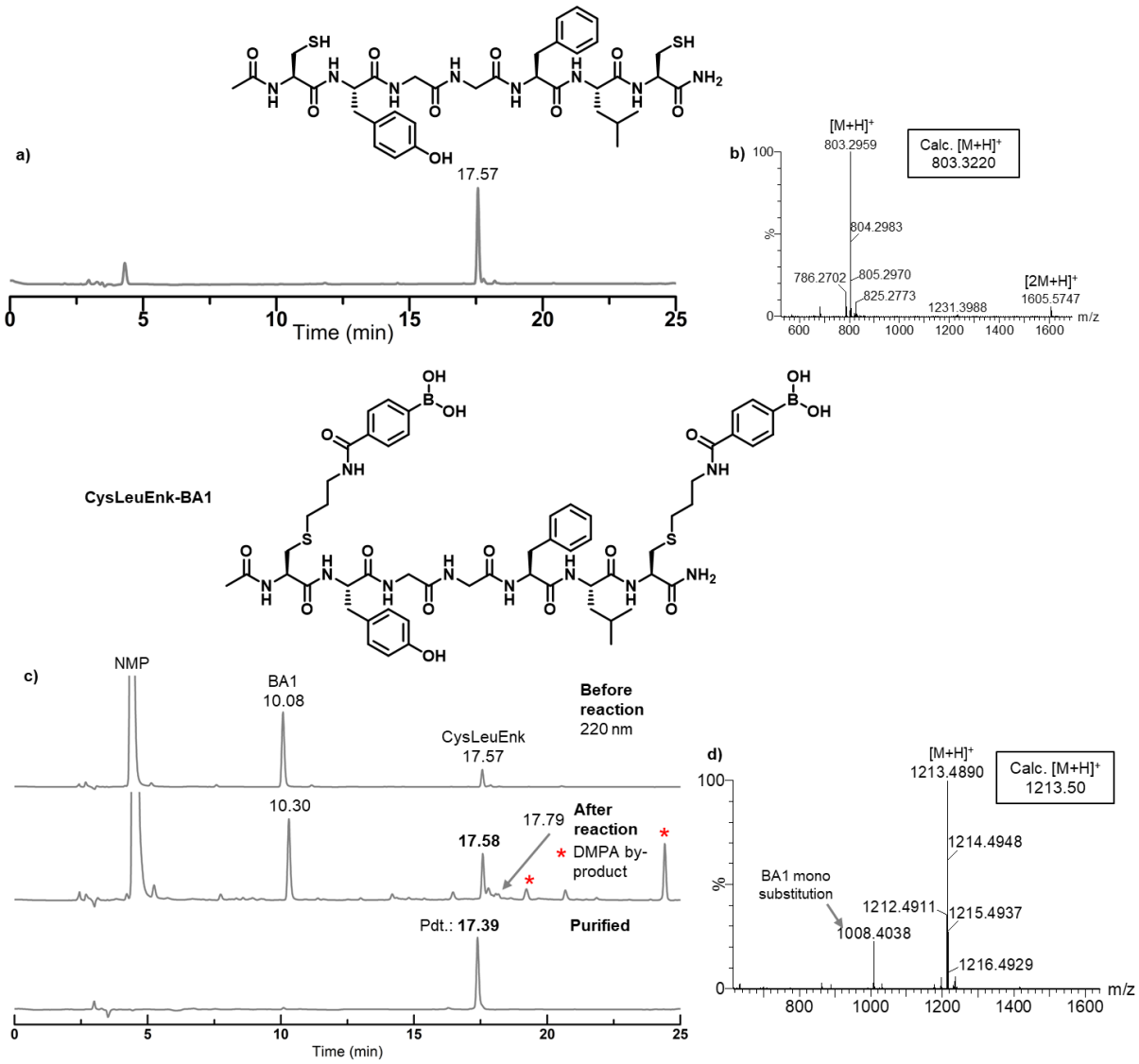


Figure 2.44: CysLeuEnk data: a) HPLC trace at 220 nm (>95% purity, the peak at ~4.5 min is from column contamination); b) Mass data; CysLeuEnk-BA1 data: c) HPLC trace at 220 nm, 68% conversion; d) Mass data. 47% yield

CysLeuEnk-BA2

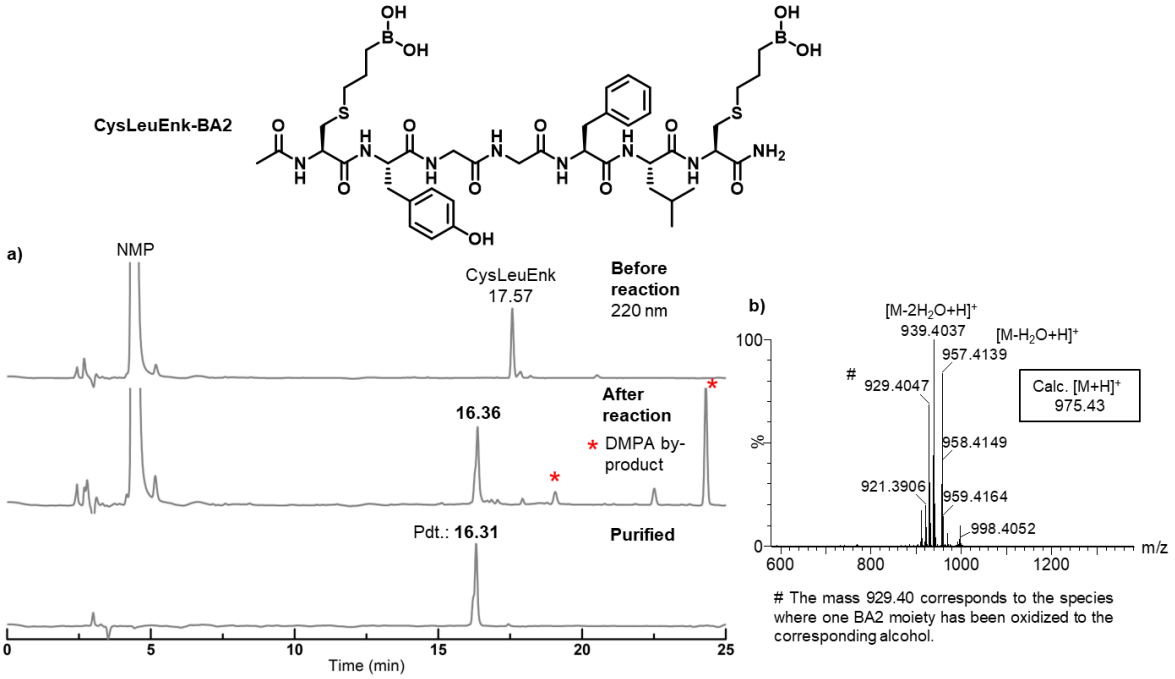


Figure 2.45: a) HPLC trace at 220 nm, >99% conversion; b) Mass data. 62% yield (# The mass 929.40 corresponds to the species where one BA2 moiety has been oxidized to the corresponding alcohol.)

CysLeuEnk-BA5

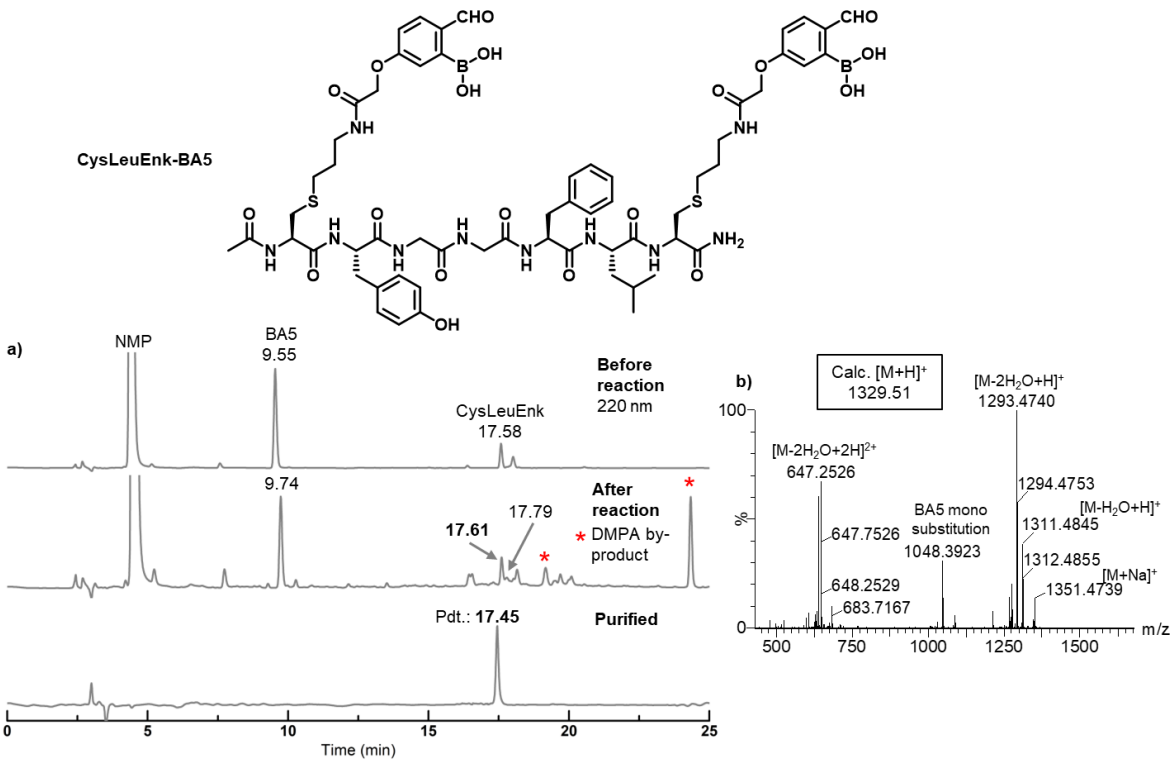
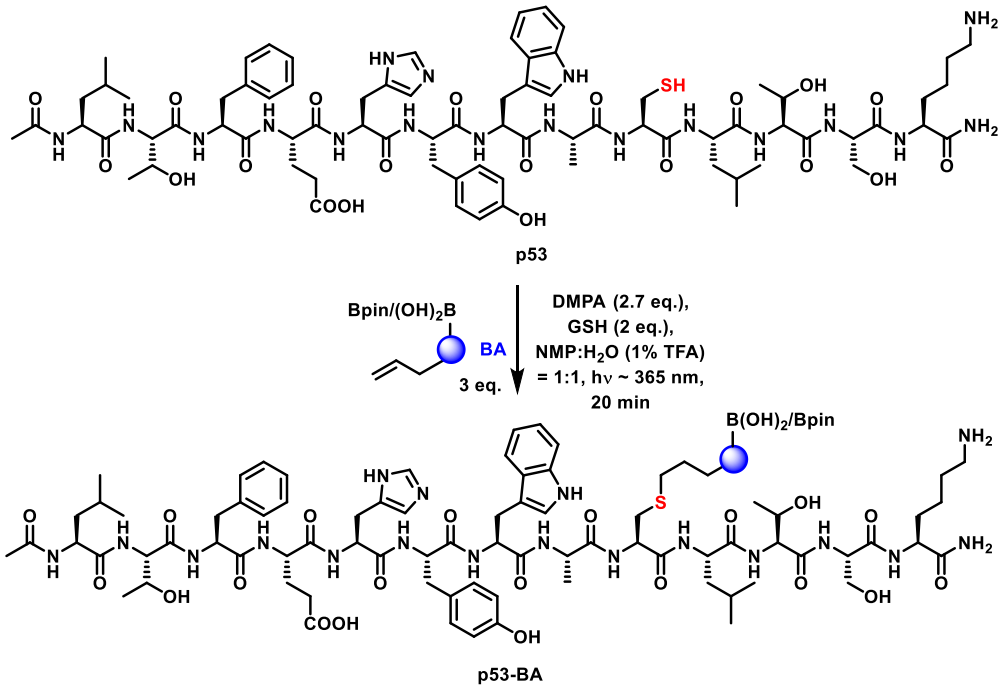
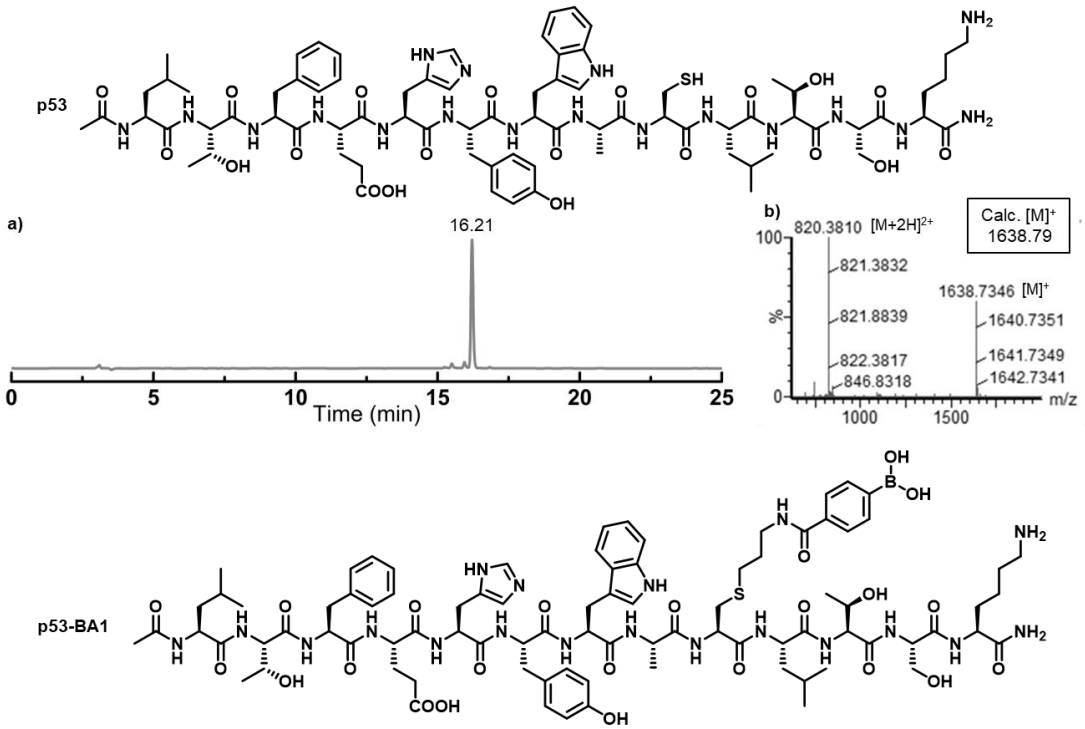


Figure 2.46: a) HPLC trace at 220 nm, 88% conversion; b) Mass data. 61% yield

2.5.8.9 *REACTIONS WITH p53*



p53-BA1



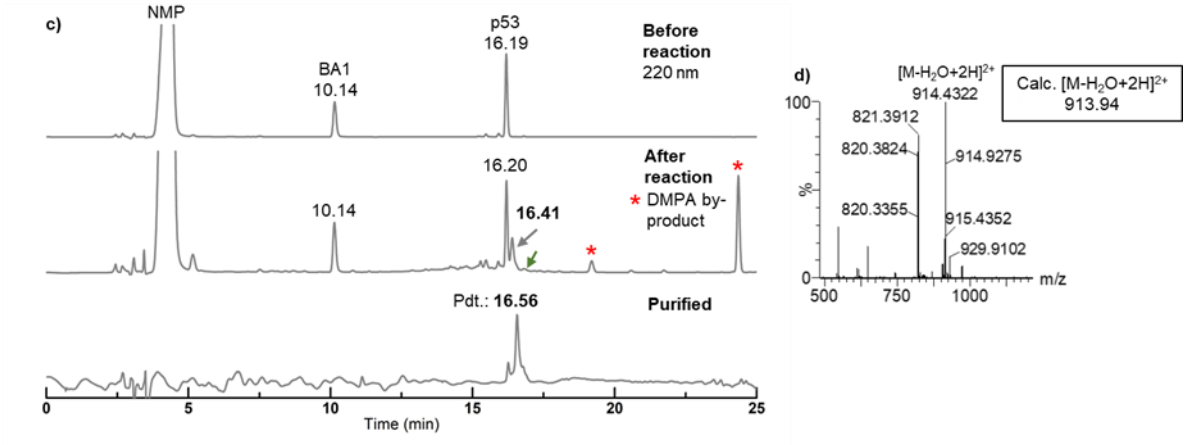


Figure 2.47: p53 data: a) HPLC trace at 220 nm (94% purity); b) Mass data; p53 -BA1 data: c) HPLC trace at 220 nm, 56% (10% oxidised product) conversion; d) Mass data. 40% yield

p53-BA2

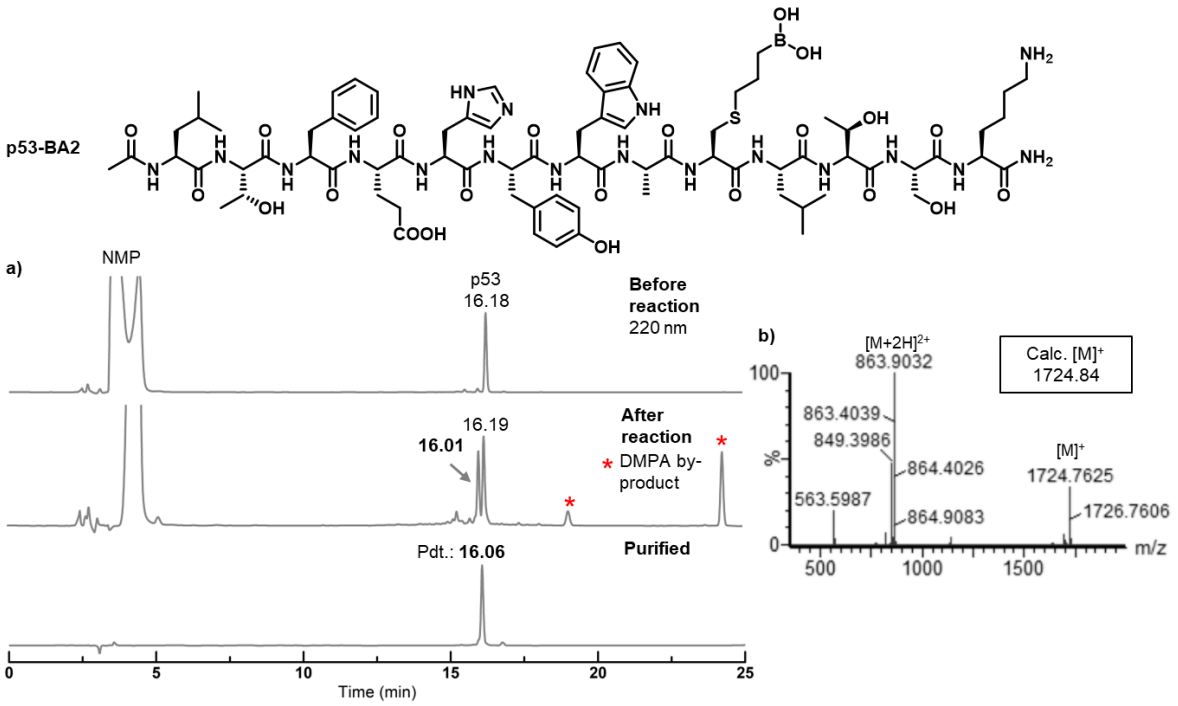


Figure 2.48: a) HPLC trace at 220 nm, 64% conversion; b) Mass data. 45% yield

p53-BA5

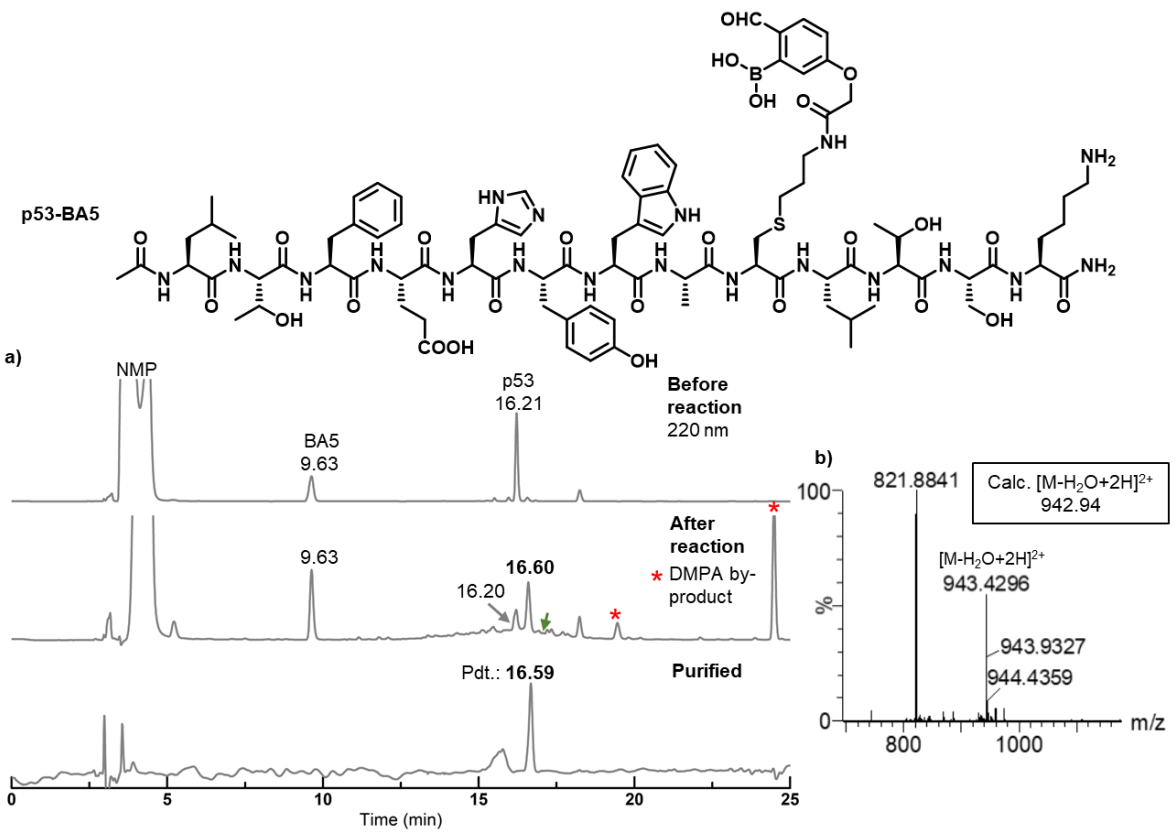


Figure 2.49: a) HPLC trace at 220 nm, 80% (10% oxidised product) conversion; b) Mass data. 45% yield

2.5.9 *CHEMICAL MODULATION OF BORONOPEPTIDE*

2.5.9.1 *SIDE-CHAIN TO SIDE-CHAIN CYCLIZATION OF THE PEPTIDE*

100 µM of WGA-BA5 in PBS buffer (pH 7.0) was treated with 10 eq. of $NaCNBH_3$ and stirred at room temperature for 3 hr. Thereafter, the crude sample was acidified and directly subjected to HPLC and mass analysis without purification, revealing a near quantitative conversion (Figure 2.6a).

2.5.9.2 *CYCLIZATION OF OXYTOCIN PEPTIDE*

0.5 mg of Oxy-BA1 was dissolved in 1 mL 80% MeOH/0.1N HCl, to which a solution of 5 mM I_2 in MeOH was added dropwise till the yellow colour of I_2 persisted. The mixture was stirred at room temperature for 6 hr, and the crude sample was subjected to Analytical HPLC and mass analysis without purification. A clean conversion was observed without side reactions or modifications to the BA moiety (Figure 2.6b).

2.5.9.3 *LABELLING OF PEPTIDE VIA DIAZABORINE FORMATION*

p-Toluenesulfonyl hydrazide (TsHz) was prepared according to previously reported literature.[7]

To a 100 µM solution of Oxy-BA5 in PBS buffer at physiological pH, 5 eq. of TsHz was added and stirred at room temperature for 2 hr and then subjected to HPLC and mass analysis without purification. The reaction proceeded to complete conversion (Figure 2.6c).

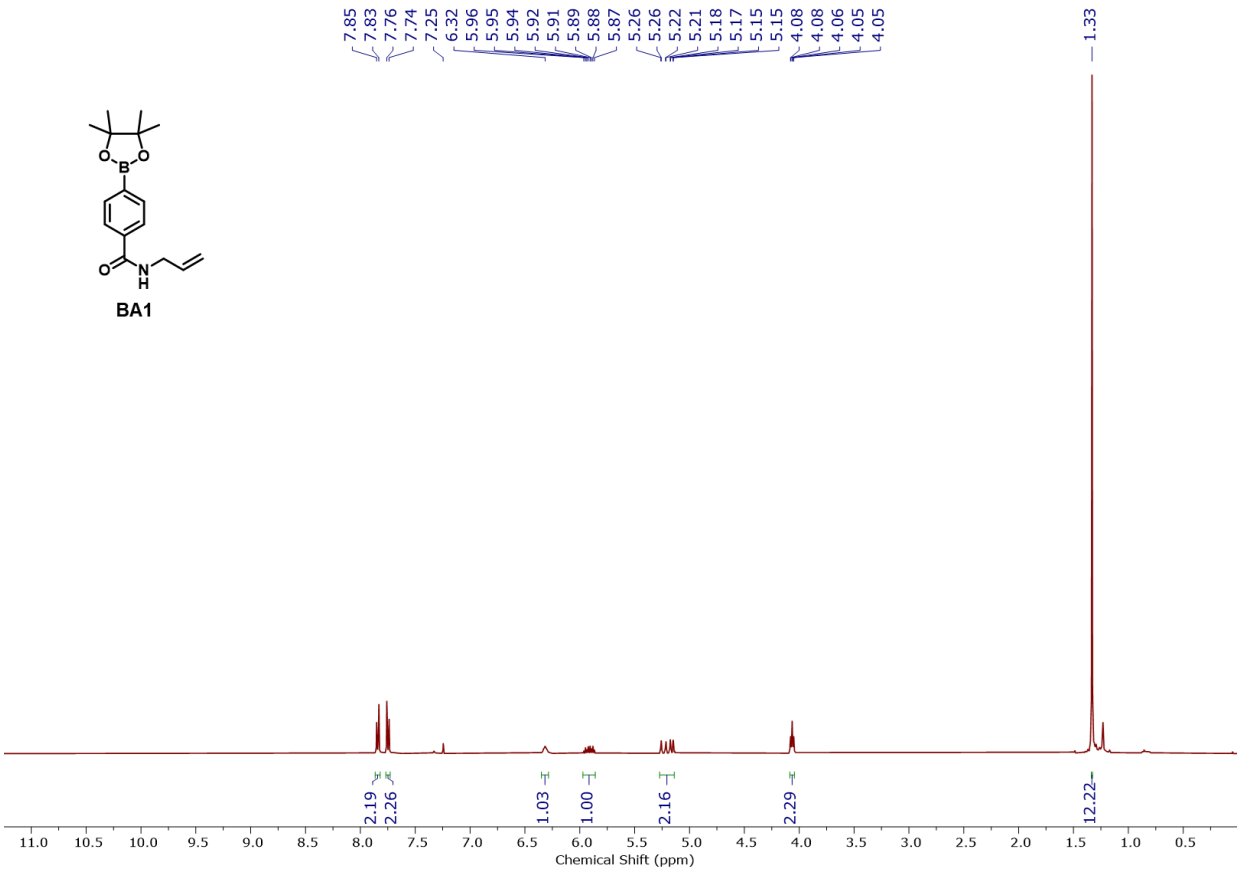
2.5.9.4 *BENZOXABOROLE FUNCTIONALIZATION*

To a solution of ATN-BA5 (100 µM) in PBS buffer (pH 7) was added 5 eq. of $NaBH_4$ and stirred at room temperature for 3 hr. It was then acidified and subjected to HPLC and mass analysis without purification. The complete conversion was observed. However, a neighbouring peak to the expected product at 9.30 min was observed in the HPLC trace, which had the same *m/z* value as the product. This observation may require further investigation to understand the observation clearly (Figure 2.6d).

2.5.10 *SERUM STABILITY STUDIES*

The serum stability of boronopeptides was carried out according to the reported literature[8]. Human serum freshly extracted from an individual was diluted to 25% with PBS (10 mM, pH 7.4) and kept at 37 °C. Peptides (6 µL, 12 mM stock in DMSO) were added to 300 µL of the serum and incubated at 37 °C. 60 µL aliquots were withdrawn at 0, 1, 4, and 24 hr intervals and added to 30 µL 1:1 MeOH/10% trichloroacetic acid (TCA) and 110 µL H_2O. The resulting suspension was vortexed and spun at 4000 rpm for 5 min. The supernatant was freeze-dried, and the resulting residue was reconstituted in 1:1 ACN/H_2O. The mixture was again spun down, and the supernatant was subjected to analytical HPLC analysis. The sample collected at 0 hr was considered the initial standard, against which the peak areas of samples at different time intervals were compared as shown in Figure 2.7. All data were taken in duplicate.

2.5.11 *NMR DATA*



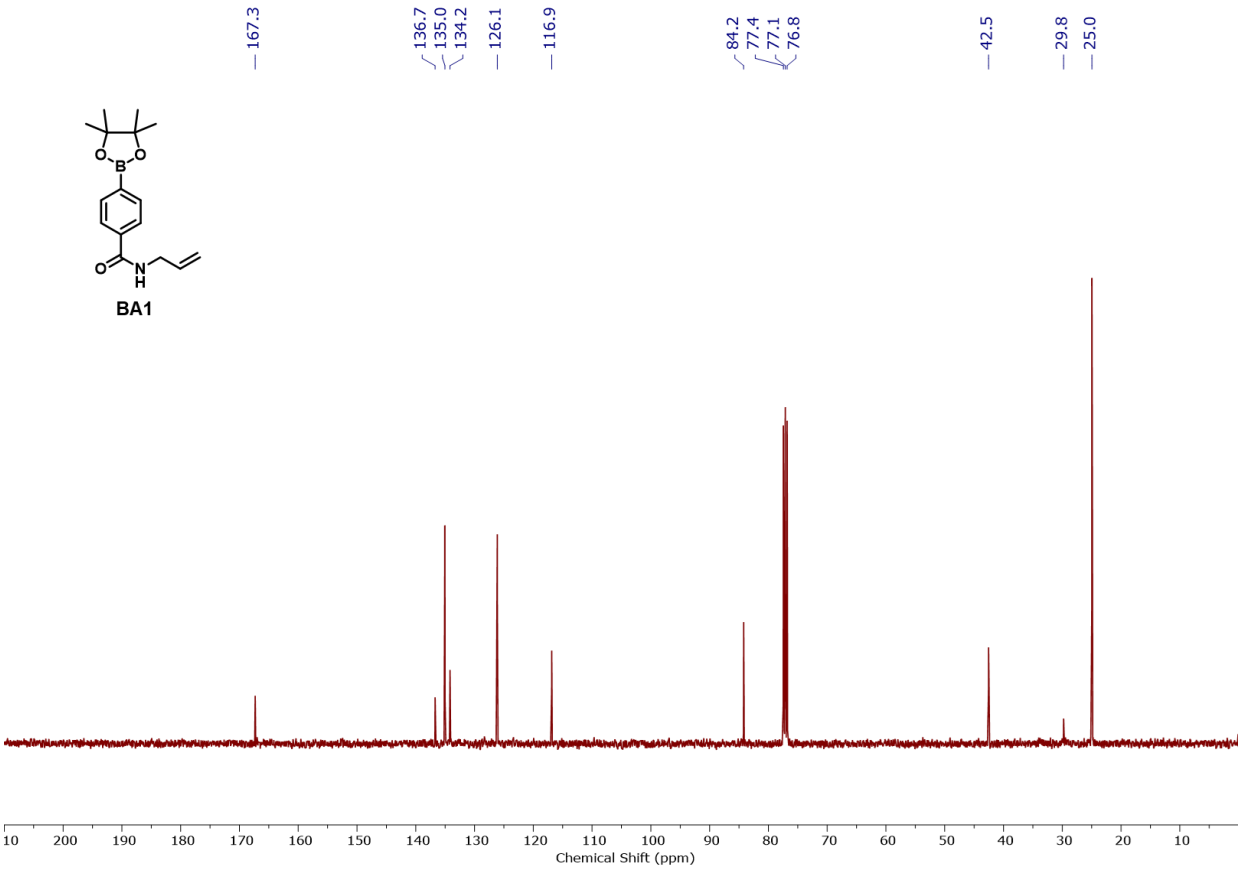
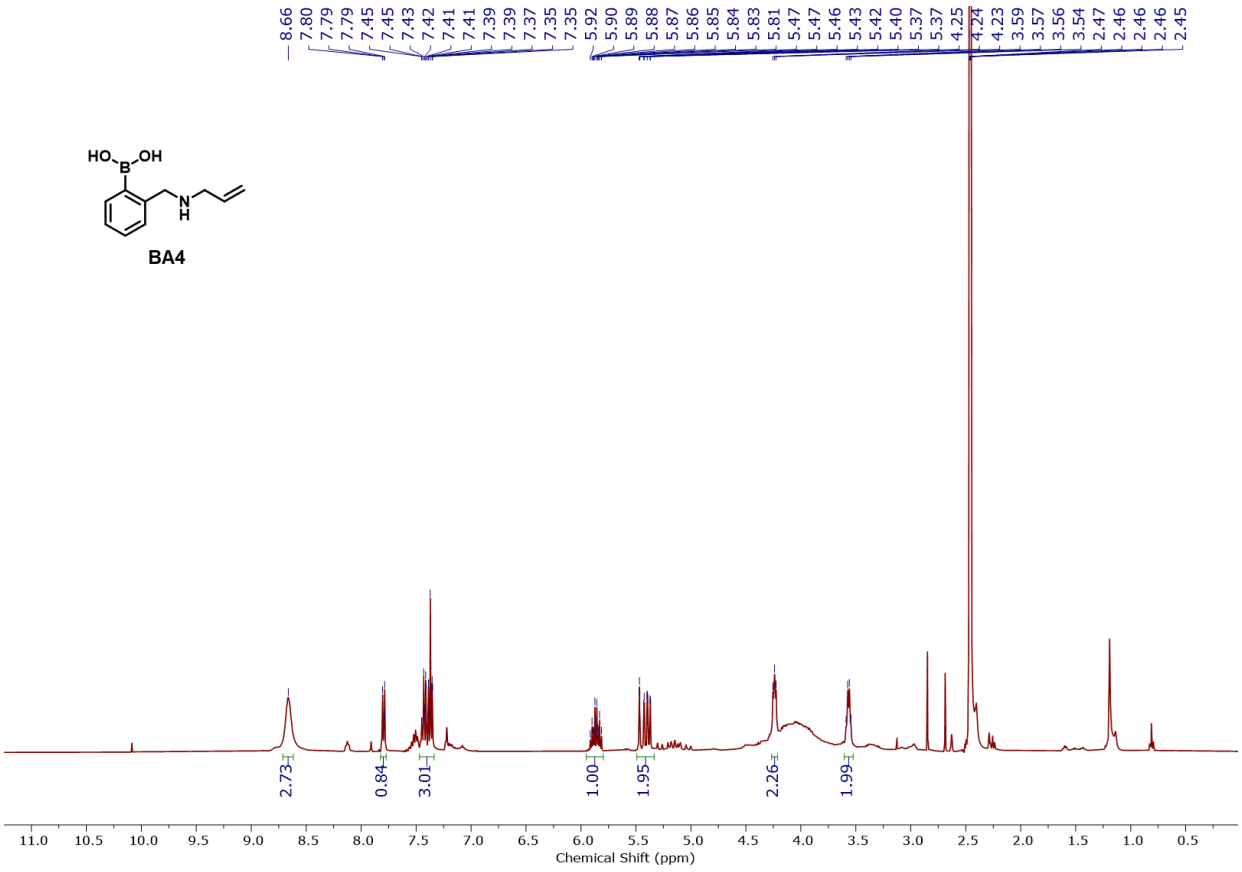


Figure 2.50: ^{1}H (400 MHz) and ^{13}C (101 MHz) NMR spectra of BA1 in $CDCl_3$.



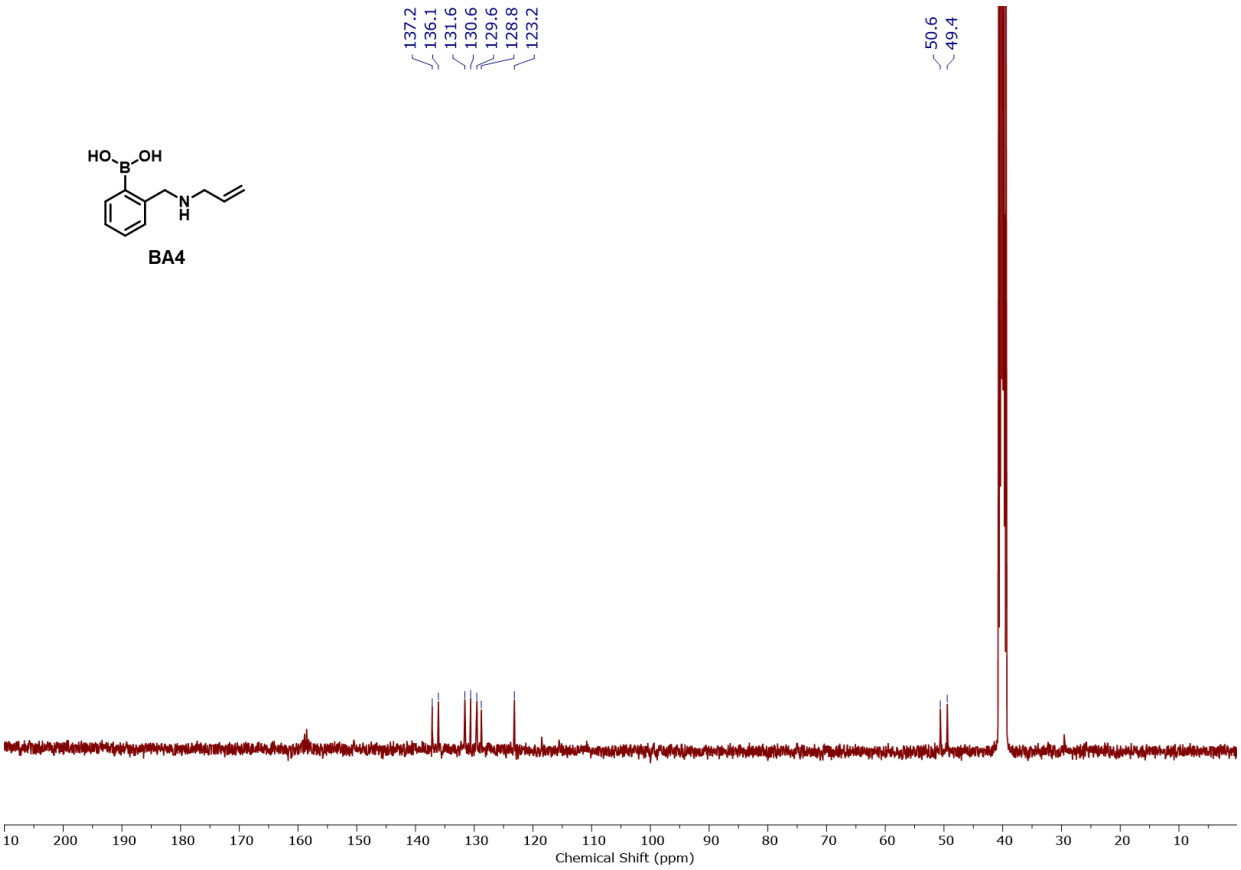
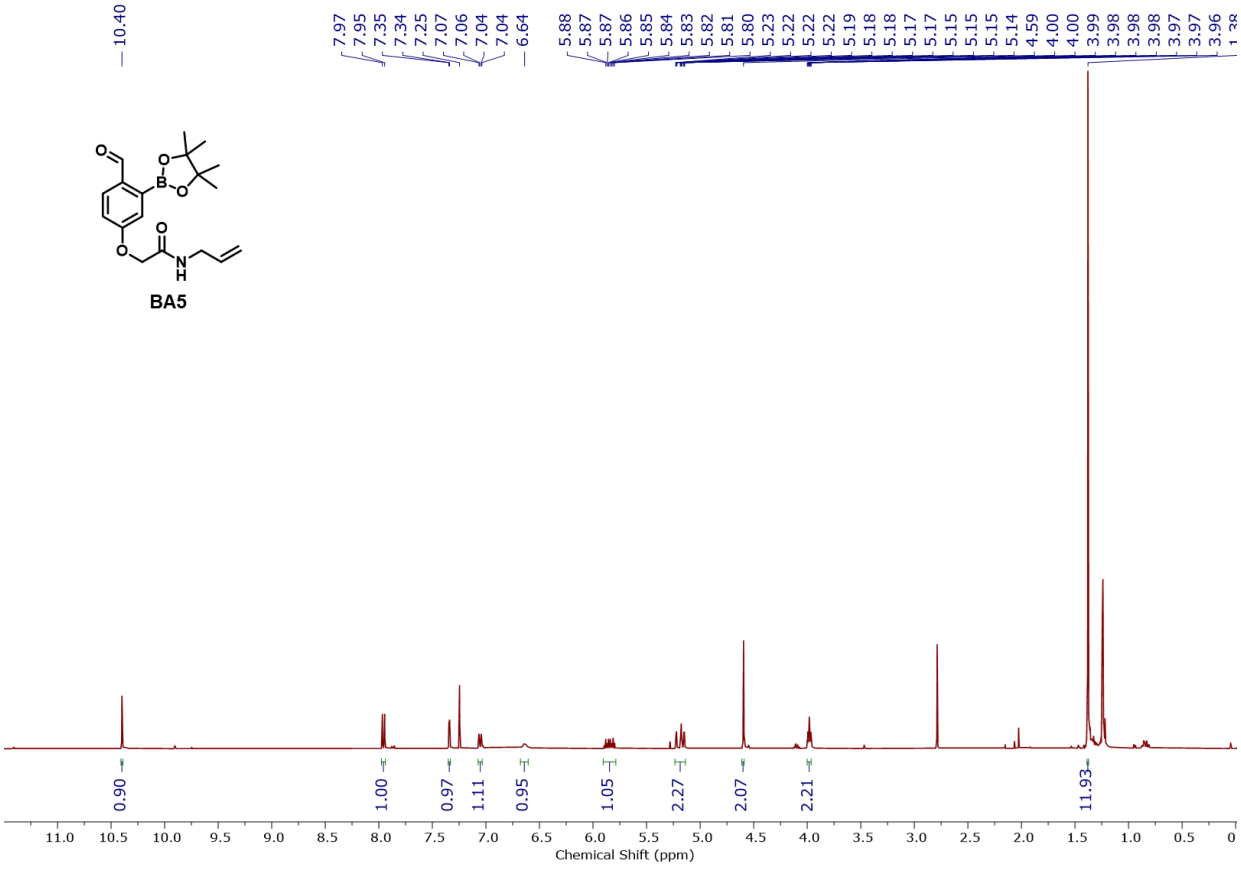


Figure 2.51: ^{1}H (400 MHz) and ^{13}C (101 MHz) NMR spectra of crude BA4 in DMSO-d_6.



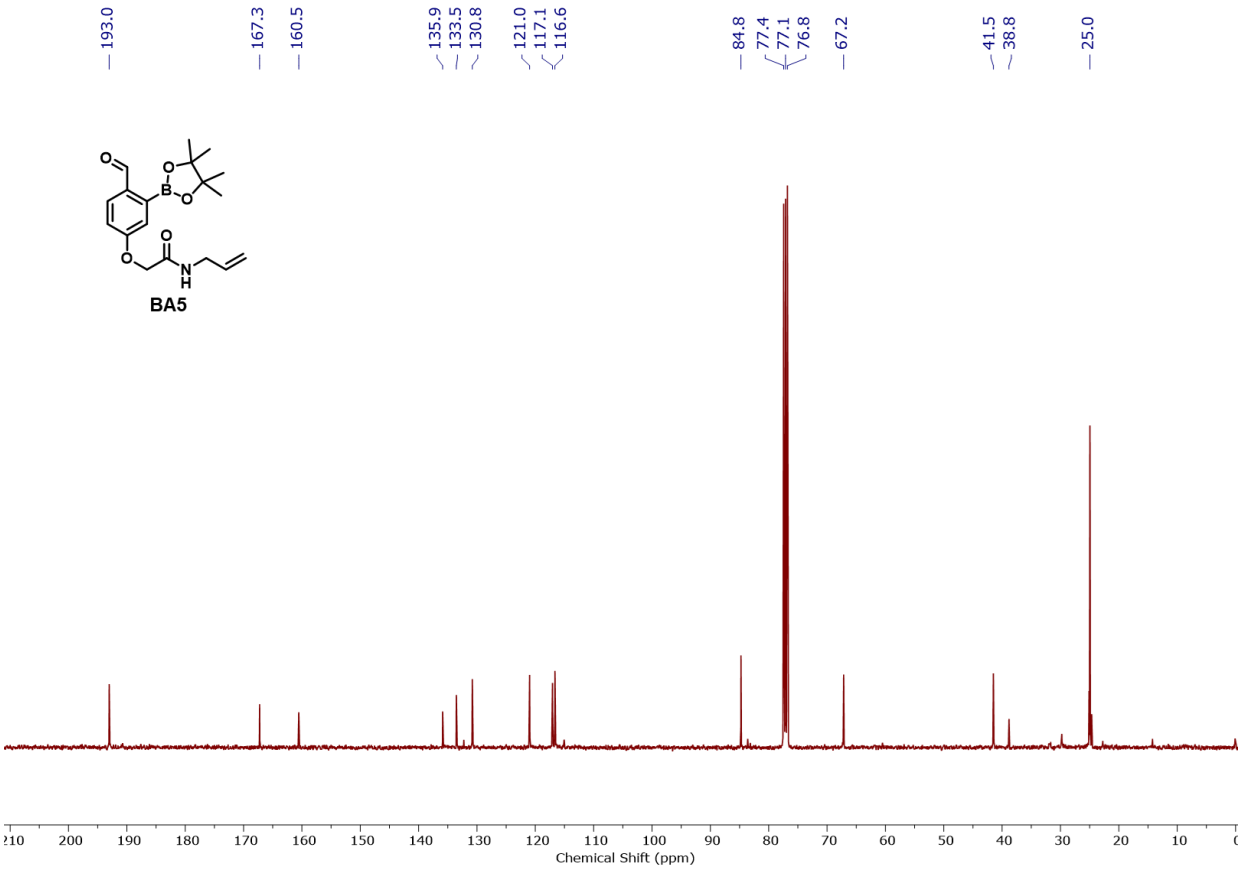
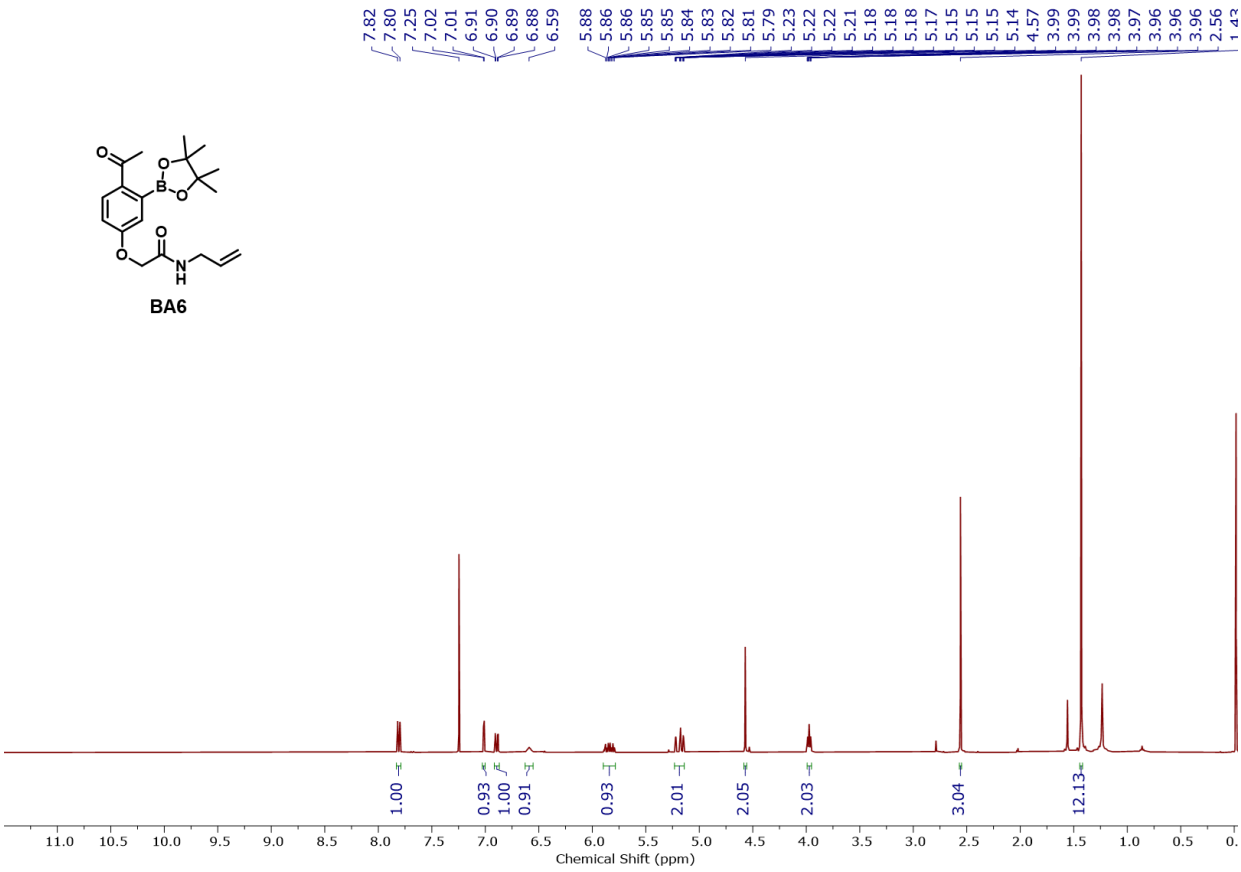


Figure 2.52: ^{1}H (400 MHz) and ^{13}C (101 MHz) NMR spectra of BA5 (having residual B_2pin_2) in 5% $CD_3OD/CDCl_3$.



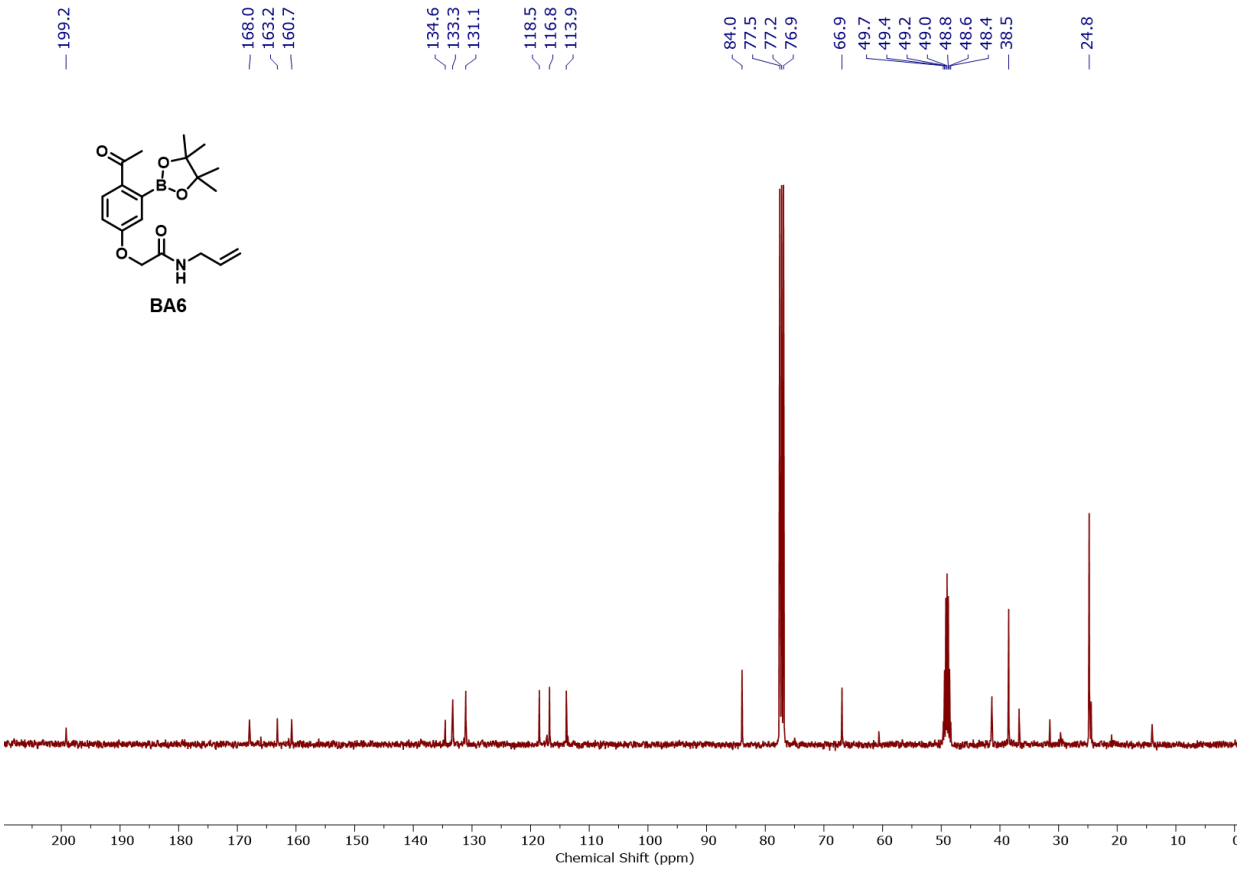
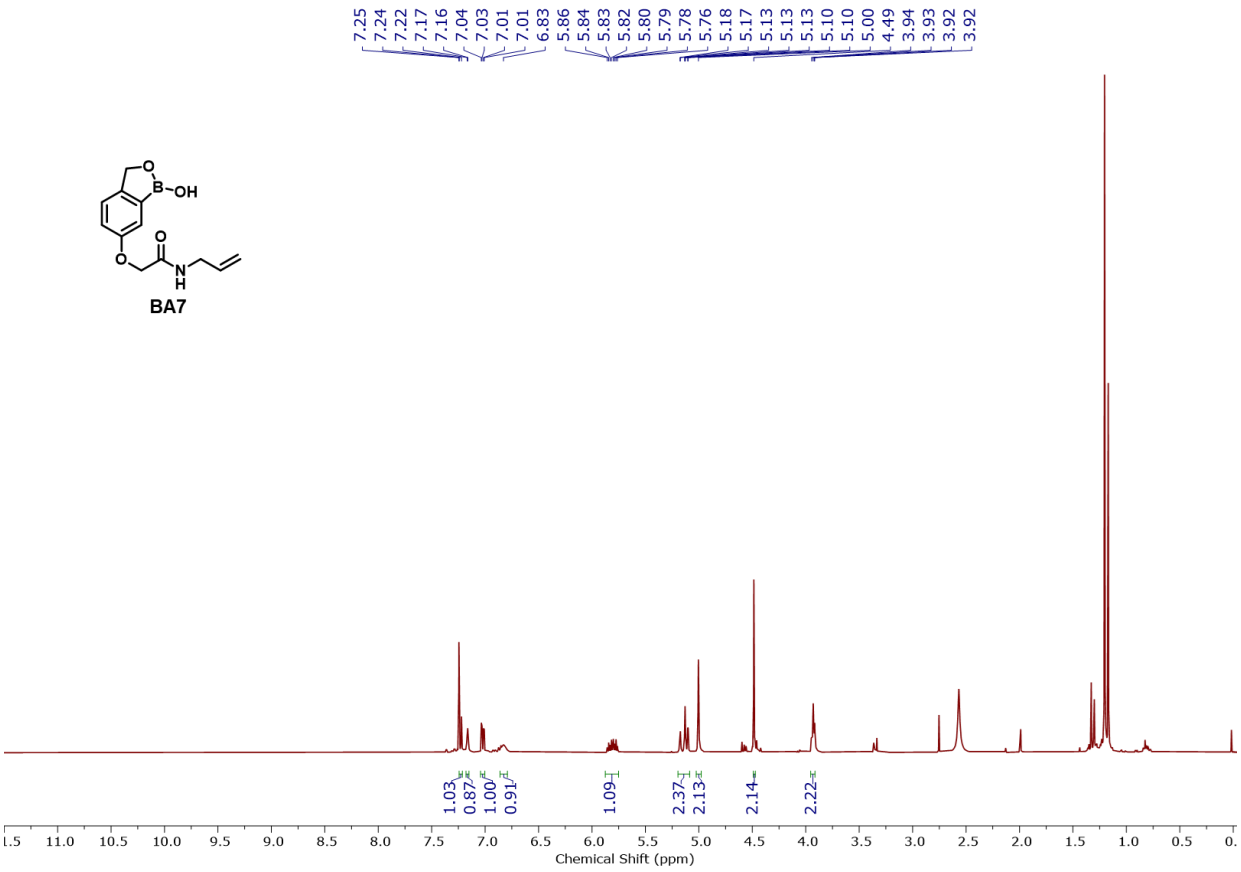


Figure 2.53: ^{1}H (400 MHz) and ^{13}C (101 MHz) NMR spectra of BA6 (having residual B_2pin_2) in 5% $CD_3OD/CDCl_3$.



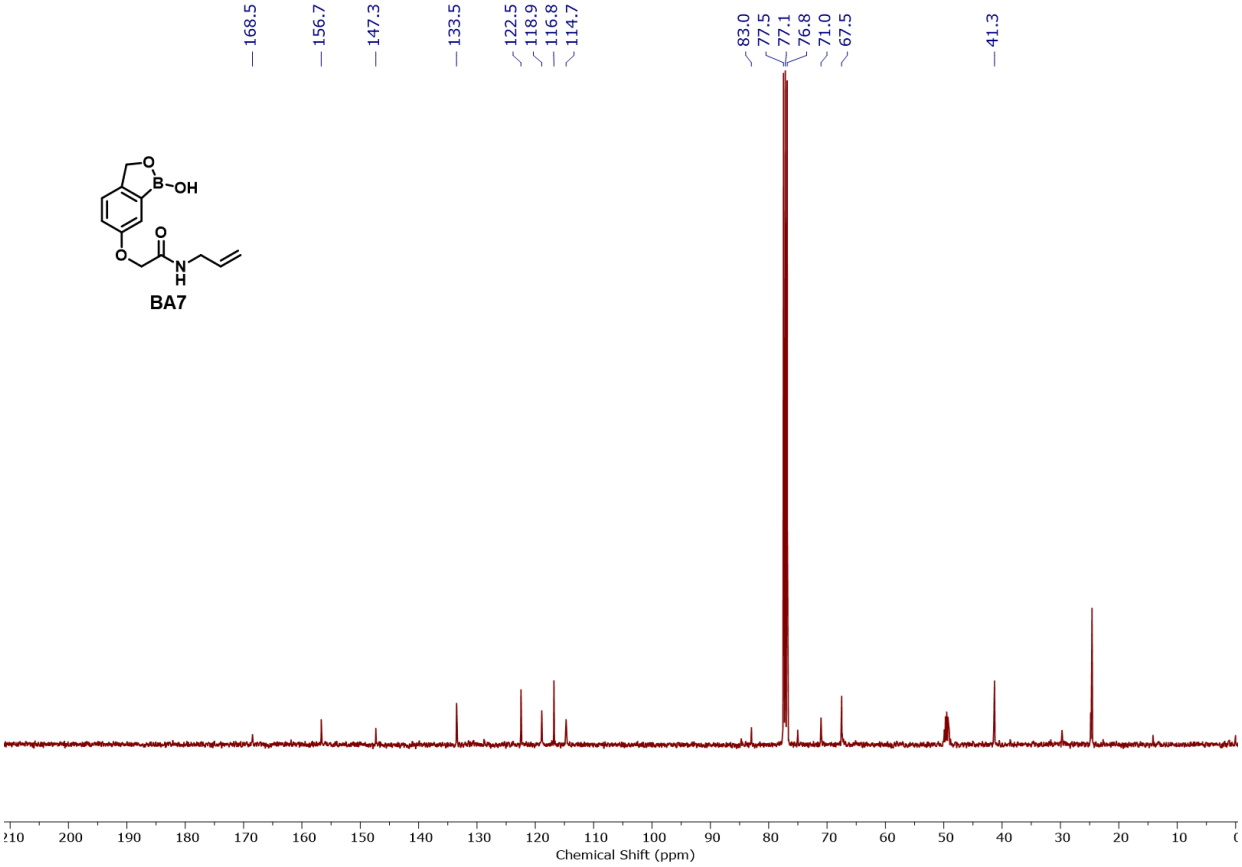
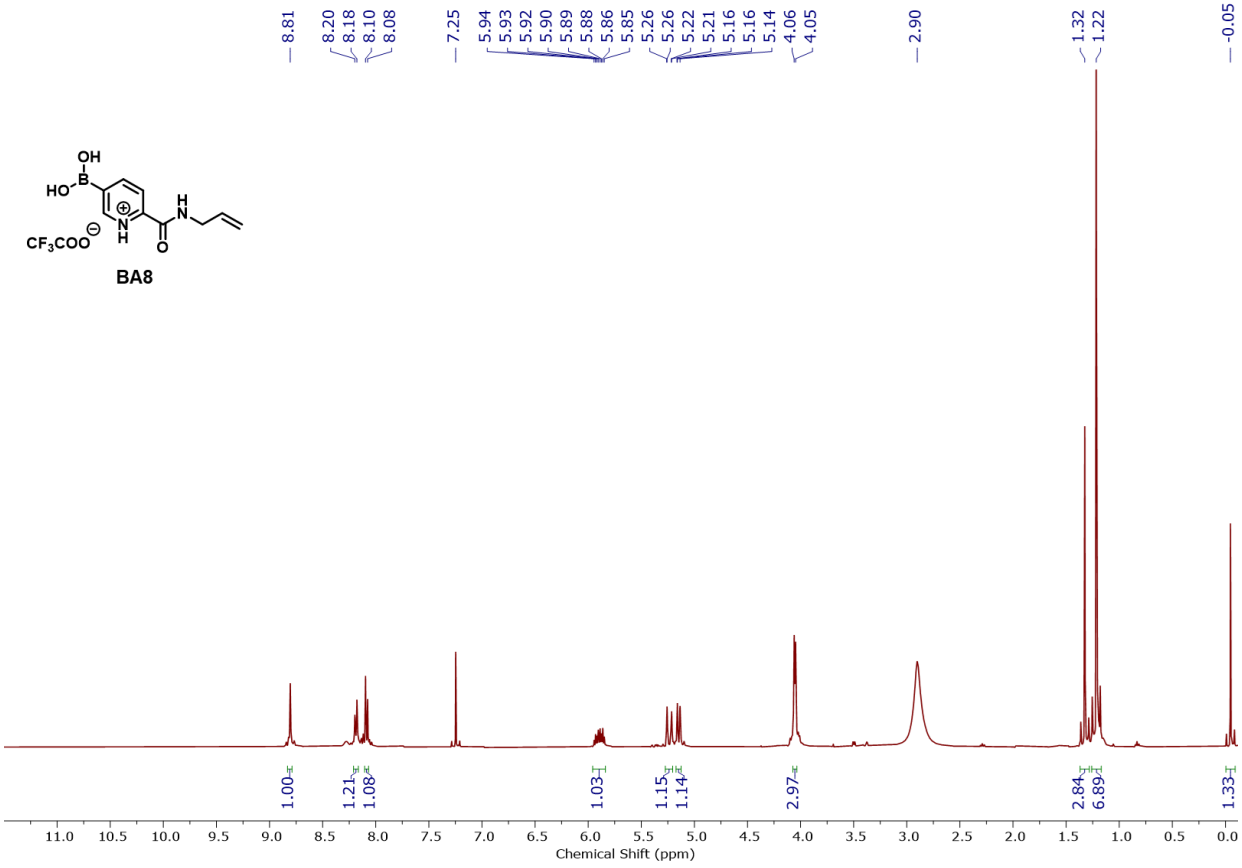


Figure 2.54: ^{1}H (400 MHz) and ^{13}C (101 MHz) NMR spectra of crude BA7 in $CDCl_3$.



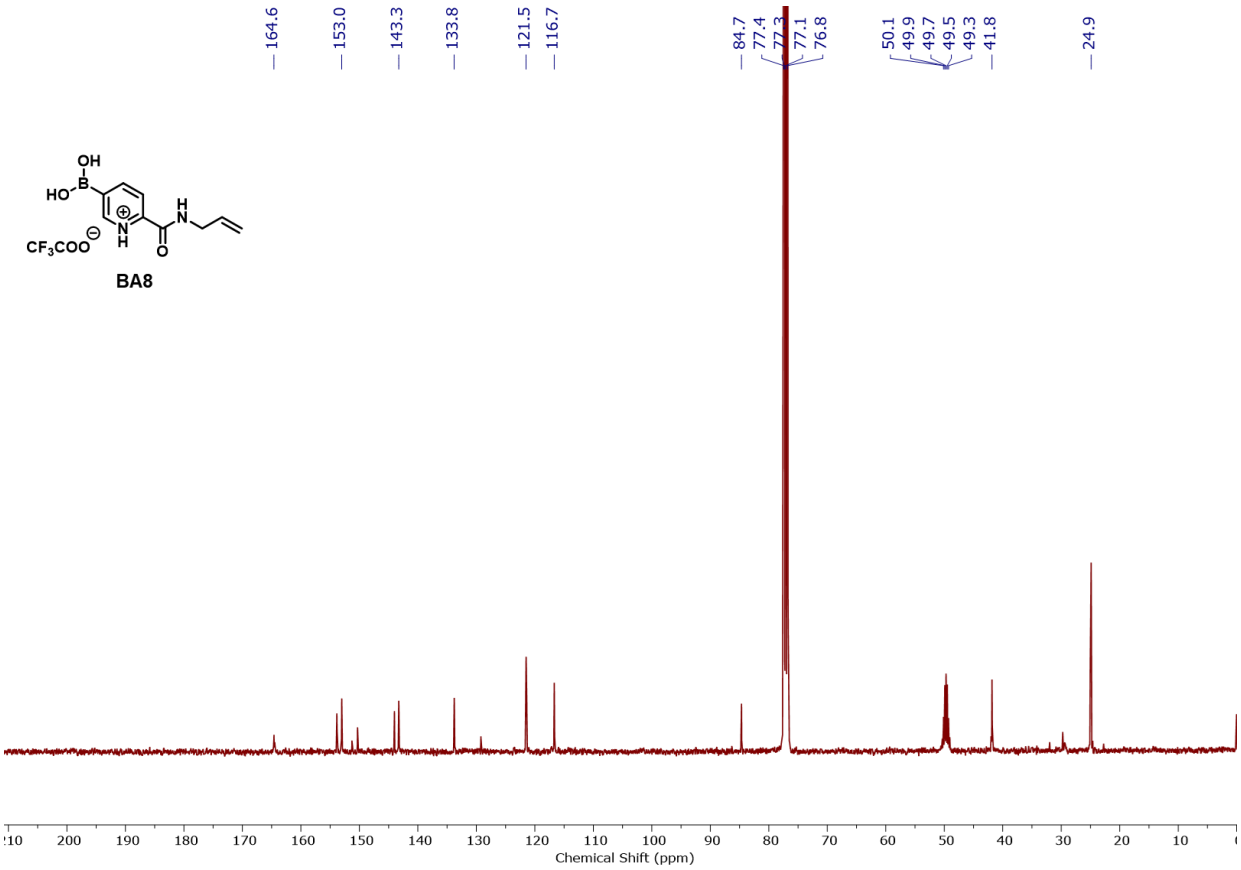


Figure 2.55: ^{1}H (400 MHz) and ^{13}C (101 MHz) NMR spectra of BA8 in 10% $CD_3OD/CDCl_3$.

2.6 *REFERENCES*

(1) Chatterjee, S.; Anslyn, E. V.; Bandyopadhyay, A. Boronic Acid Based Dynamic Click Chemistry: Recent Advances and Emergent Applications. *Chem Sci* **2021**, *12* (5), 1585–1599. https://doi.org/10.1039/d0sc05009a.

(2) António, J. P. M.; Russo, R.; Carvalho, C. P.; Cal, P. M. S. D.; Gois, P. M. P. Boronic Acids as Building Blocks for the Construction of Therapeutically Useful Bioconjugates. *Chem Soc Rev* **2019**, *48* (13), 3513–3536. https://doi.org/10.1039/c9cs00184k.

(3) Ballatore, C.; Huryn, D. M.; Smith, A. B. Carboxylic Acid (Bio)Isosteres in Drug Design. *ChemMedChem* **2013**, *8* (3), 385–395. https://doi.org/10.1002/cmdc.201200585.

(4) Chatterjee, S.; Tripathi, N. M.; Bandyopadhyay, A. The Modern Role of Boron as a 'Magic Element' in Biomedical Science: Chemistry Perspective. *Chemical Communications* **2021**, *57* (100), 13629–13640. https://doi.org/10.1039/D1CC05481C.

(5) Miseta, A.; Csutora, P. Relationship Between the Occurrence of Cysteine in Proteins and the Complexity of Organisms. *Mol Biol Evol* **2000**, *17* (8), 1232–1239. https://doi.org/10.1093/oxfordjournals.molbev.a026406.

(6) Ahangarpour, M.; Kavianinia, I.; Harris, P. W. R.; Brimble, M. A. Photo-Induced Radical Thiol–Ene Chemistry: A Versatile Toolbox for Peptide-Based Drug Design. *Chem Soc Rev* **2021**, *50* (2), 898–944. https://doi.org/10.1039/D0CS00354A.

(7) Boutureira, O.; Bernardes, G. J. L. Advances in Chemical Protein Modification. *Chem Rev* **2015**, *115* (5), 2174–2195. https://doi.org/10.1021/cr500399p.

(8) Hoyle, C. E.; Bowman, C. N. Thiol–Ene Click Chemistry. *Angewandte Chemie International Edition* **2010**, *49* (9), 1540–1573. https://doi.org/https://doi.org/10.1002/anie.200903924.

(9) Hoyle, C. E.; Lowe, A. B.; Bowman, C. N. Thiol-Click Chemistry: A Multifaceted Toolbox for Small Molecule and Polymer Synthesis. *Chem Soc Rev* **2010**, *39* (4), 1355–1387. https://doi.org/10.1039/B901979K.

(10) Lahnsteiner, M.; Kastner, A.; Mayr, J.; Roller, A.; Keppler, B. K.; Kowol, C. R. Improving the Stability of Maleimide–Thiol Conjugation for Drug Targeting. *Chemistry – A European Journal* **2020**, *26* (68), 15867–15870. https://doi.org/https://doi.org/10.1002/chem.202003951.

(11) Conlon, J. M. Purification of Naturally Occurring Peptides by Reversed-Phase HPLC. *Nat Protoc* **2007**, *2* (1), 191–197. https://doi.org/10.1038/nprot.2006.437.

(12) Xiao, Q.; Tong, Q.-X.; Zhong, J.-J. Recent Advances in Visible-Light Photoredox Catalysis for the Thiol-Ene/Yne Reactions. *Molecules*. 2022. https://doi.org/10.3390/molecules27030619.

(13) Wang, Y.; Chou, D. H. C. A Thiol-Ene Coupling Approach to Native Peptide Stapling and Macrocyclization. *Angewandte Chemie - International Edition* **2015**, *54* (37), 10931–10934. https://doi.org/10.1002/anie.201503975.

(14) Wang, Y.; Bruno, B. J.; Cornillie, S.; Nogieira, J. M.; Chen, D.; Cheatham, T. E.; Lim, C. S.; Chou, D. H. C. Application of Thiol–Yne/Thiol–Ene Reactions for Peptide and Protein Macrocyclizations. *Chemistry - A European Journal* **2017**, *23* (29), 7087–7092. https://doi.org/10.1002/chem.201700572.

(15) Moosmann, B.; Hajieva, P. Probing the Role of Cysteine Thiyl Radicals in Biology: Eminently Dangerous, Difficult to Scavenge. *Antioxidants*. 2022. https://doi.org/10.3390/antiox11050885.

(16) Sinha, A. K.; Equbal, D. Thiol−Ene Reaction: Synthetic Aspects and Mechanistic Studies of an Anti-Markovnikov-Selective Hydrothiolation of Olefins. *Asian J Org Chem* **2019**, *8* (1), 32–47. https://doi.org/https://doi.org/10.1002/ajoc.201800639.

(17) Povie, G.; Tran, A.-T.; Bonnaffé, D.; Habegger, J.; Hu, Z.; Le Narvor, C.; Renaud, P. Repairing the Thiol-Ene Coupling Reaction. *Angewandte Chemie International Edition* **2014**, *53* (15), 3894–3898. https://doi.org/https://doi.org/10.1002/anie.201309984.

(18) Li, C.; Wang, J.; Barton, L. M.; Yu, S.; Tian, M.; Peters, D. S.; Kumar, M.; Yu, A. W.; Johnson, K. A.; Chatterjee, A. K.; Yan, M.; Baran, P. S. Decarboxylative Borylation. *Science (1979)* **2017**, *356* (6342). https://doi.org/10.1126/science.aam7355.

(19) Mollner, T. A.; Isenegger, P. G.; Josephson, B.; Buchanan, C.; Lercher, L.; Oehlrich, D.; Hansen, D. F.; Mohammed, S.; Baldwin, A. J.; Gouverneur, V.; Davis, B. G. Post-Translational Insertion of Boron in Proteins to Probe and Modulate Function. *Nat Chem Biol* **2021**, *17* (12), 1245–1261. https://doi.org/10.1038/s41589-021-00883-7.

(20) de Vries, R. H.; Viel, J. H.; Kuipers, O. P.; Roelfes, G. Rapid and Selective Chemical Editing of Ribosomally Synthesized and Post-Translationally Modified Peptides (RiPPs) via Cu II -Catalyzed B-Borylation of Dehydroamino Acids. *Angewandte Chemie International Edition* **2021**, *60* (8), 3946–3950. https://doi.org/10.1002/anie.202011460.

(21) Chalker, J. M.; Bernardes, G. J. L.; Davis, B. G. A "Tag-and-Modify" Approach to Site-Selective Protein Modification. *Acc Chem Res* **2011**, *44* (9), 730–741. https://doi.org/10.1021/ar200056q.

(22) Ball, Z. T. Protein Substrates for Reaction Discovery: Site-Selective Modification with Boronic Acid Reagents. *Acc Chem Res* **2019**, *52* (3), 566–575. https://doi.org/10.1021/acs.accounts.8b00626.

(23) Waddington, M. A.; Zheng, X.; Stauber, J. M.; Hakim Moully, E.; Montgomery, H. R.; Saleh, L. M. A.; Král, P.; Spokoyny, A. M. An Organometallic Strategy for Cysteine Borylation. *J Am Chem Soc* **2021**, *143* (23), 8661–8668. https://doi.org/10.1021/jacs.1c02206.

(24) Cambray, S.; Gao, J. Versatile Bioconjugation Chemistries of Ortho-Boronyl Aryl Ketones and Aldehydes. *Acc Chem Res* **2018**, *51* (9), 2198–2206. https://doi.org/10.1021/acs.accounts.8b00154.

(25) Dai, C.; Cheng, Y.; Cui, J.; Wang, B. Click Reactions and Boronic Acids: Applications, Issues, and Potential Solutions. *Molecules* **2010**, *15* (8), 5768–5781. https://doi.org/10.3390/molecules15085768.

(26) James, T. D.; Phillips, M. D.; Shinkai, S. Boronic Acids in Saccharide Recognition. **2006**. https://doi.org/10.1039/9781847557612.

(27) Maslah, H.; Skarbek, C.; Pethe, S.; Labruère, R. Anticancer Boron-Containing Prodrugs Responsive to Oxidative Stress from the Tumor Microenvironment. *Eur J Med Chem* **2020**, *207*, 112670. https://doi.org/10.1016/j.ejmech.2020.112670.

(28) Sun, X.; Chapin, B. M.; Metola, P.; Collins, B.; Wang, B.; James, T. D.; Anslyn, E. V. The Mechanisms of Boronate Ester Formation and Fluorescent Turn-on in Ortho-Aminomethylphenylboronic Acids. *Nat Chem* **2019**, *11* (9), 768–778. https://doi.org/10.1038/s41557-019-0314-x.

(29) Liu, C. T.; Tomsho, J. W.; Benkovic, S. J. The Unique Chemistry of Benzoxaboroles: Current and Emerging Applications in Biotechnology and Therapeutic Treatments. *Bioorg Med Chem* **2014**, *22* (16), 4462–4473. https://doi.org/10.1016/J.BMC.2014.04.065.

(30) Matsumoto, A.; Stephenson-Brown, A. J.; Khan, T.; Miyazawa, T.; Cabral, H.; Kataoka, K.; Miyahara, Y. Heterocyclic Boronic Acids Display Sialic Acid Selective Binding in a Hypoxic Tumor Relevant Acidic Environment. *Chem Sci* **2017**, *8* (9), 6165–6170. https://doi.org/10.1039/c7sc01905j.

(31) Love, D. M.; Kim, K.; Goodrich, J. T.; Fairbanks, B. D.; Worrell, B. T.; Stoykovich, M. P.; Musgrave, C. B.; Bowman, C. N. Amine Induced Retardation of the Radical-Mediated Thiol-Ene Reaction via the Formation of Metastable Disulfide Radical Anions. *Journal of Organic Chemistry* **2018**, *83* (5), 2912–2919. https://doi.org/10.1021/acs.joc.8b00143.

(32) Gammelgaard, S. K.; Petersen, S. B.; Haselmann, K. F.; Nielsen, P. K. Direct Ultraviolet Laser-Induced Reduction of Disulfide Bonds in Insulin and Vasopressin. *ACS Omega* **2020**, *5* (14), 7962–7968. https://doi.org/10.1021/acsomega.9b04375.

(33) Tan, Y.; Wu, J.; Song, L.; Zhang, M.; Hipolito, C. J.; Wu, C.; Wang, S.; Zhang, Y.; Yin, Y. Merging the Versatile Functionalities of Boronic Acid with Peptides. *Int J Mol Sci* **2021**, *22* (23). https://doi.org/10.3390/ijms222312958.

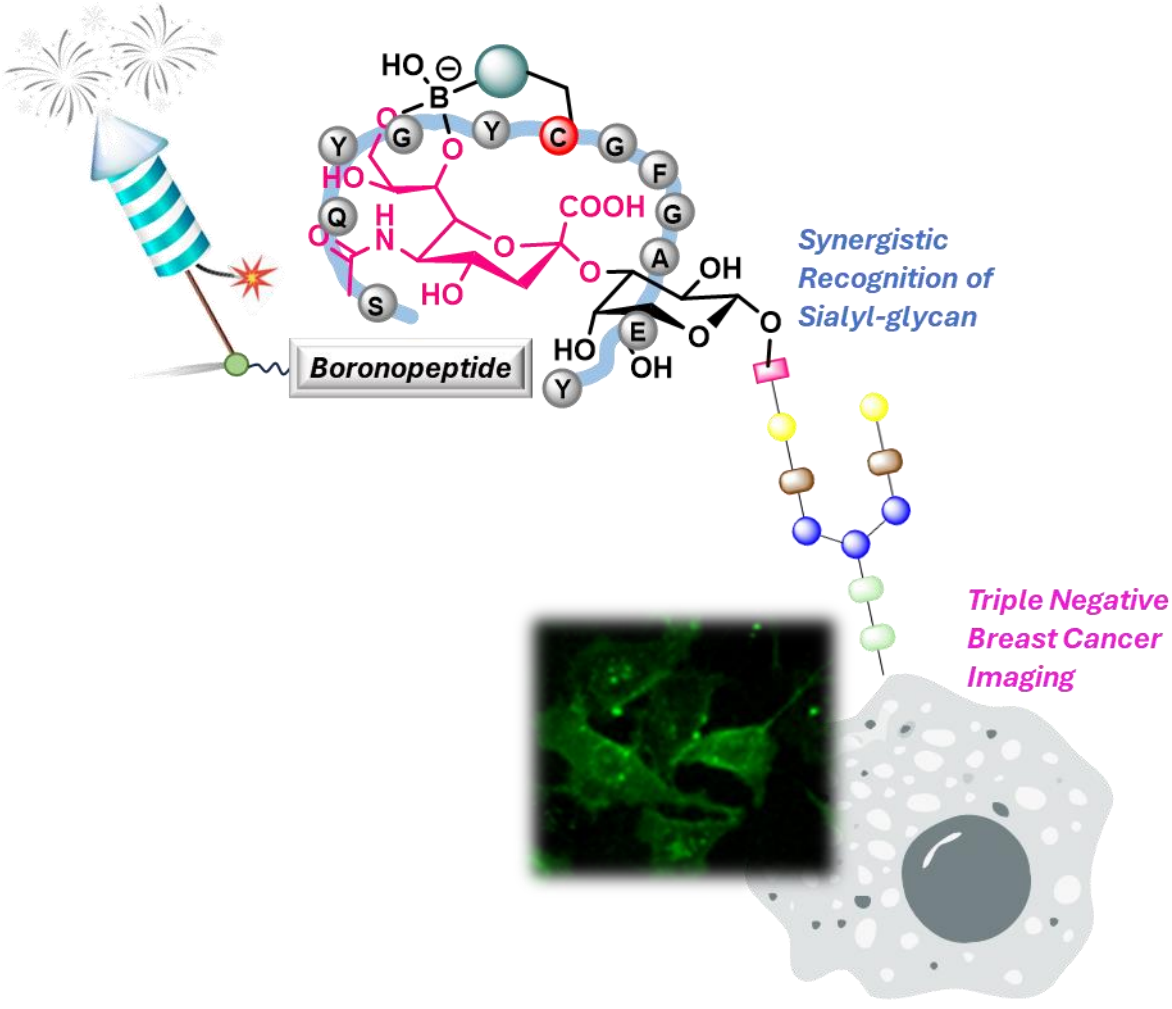
(34) Bandyopadhyay, A.; Gao, J. Iminoboronate-Based Peptide Cyclization That Responds to PH, Oxidation, and Small Molecule Modulators. *J Am Chem Soc* **2016**, *138* (7), 2098–2101. https://doi.org/10.1021/jacs.5b12301.

(35) Kocak, G.; Cicek, H.; Ceylan, Ö.; Bütün, V. Antimicrobial and Anti-Quorum-Sensing Properties and Paint Film Usage of Novel Diazaborine-Based Copolymers. *J Appl Polym Sci* **2019**, *136* (1). https://doi.org/10.1002/app.46907.

(36) Harris, P. W. R.; Siow, A.; Yang, S.-H.; Wadsworth, A. D.; Tan, L.; Hermant, Y.; Mao, Y.; An, C.; Hanna, C. C.; Cameron, A. J.; Allison, J. R.; Chakraborty, A.; Ferguson, S. A.; Mros, S.; Hards, K.; Cook, G. M.; Williamson, D. A.; Carter, G. P.; Chan, S. T. S.; Painter, G. A.; Sander, V.; Davidson, A. J.; Brimble, M. A. Synthesis, Antibacterial Activity, and Nephrotoxicity of Polymyxin B Analogues Modified at Leu-7, <scp>d</Scp> -Phe-6, and the N-Terminus Enabled by S-Lipidation. *ACS Infect Dis* **2022**, *8* (12), 2413–2429. https://doi.org/10.1021/acsinfecdis.1c00347.

CHAPTER 3A

CAPTURING SIALYL-GLYCAN ON LIVE CANCER CELLS BY TAILORED BORONOPEPTIDE



This work has been published in Chemistry – A European Journal (a journal of Chemistry Europe) and can be cited as "**Chatterjee, S.;** Chowdhury, A.; Saproo, S.; Mani Tripathi, N.; Naidu, S.; Bandyopadhyay, A. Capturing Sialyl-glycan on Live Cancer Cells by Tailored Boronopeptide. *Chemistry – A European Journal* **2024**, *30* (8)."

"Glycans are as universal in nature as nucleic acids, proteins, lipids and metabolites, and as essential to the existence of all known living organisms. But glycans are also the most structurally diverse and rapidly evolving major class of molecules. Taken together with much greater technical difficulties in their analysis, one can understand why the knowledge base regarding these major building blocks of life has lagged so far behind." *-Ajit Varki, Biological Role of Glycans, 2017*

3A.1 *INTRODUCTION*

3A.1.1 *ROLE OF SIALIC ACIDS IN PHYSIOLOGY AND IMPORTANCE OF SIALIC ACID RECOGNITION*

The term 'glycan' usually refers to compounds consisting of a large number of monosaccharides linked usually through O-glycosidic linkages[1], although more broadly, it also refers to the typical glycoconjugates found in living systems such as glycoproteins or glycolipids. Research in the first half of the 20th century mainly focused on elucidating the structure and biochemistry of simple and complex glycans by Nobel Laureates such as Fischer, Landsteiner, Haworth, Cori, Leloir, and Palade. Beyond the previously known roles of glycans in energy generation and metabolism, extensive research has shown that these are involved in cell-cell interactions, protecting proteins from degradation, and notably in cell signalling[2]. Given the remarkable diversity in the structural and functional aspects of glycans, the discussions on glycans are somewhat biased toward the author. Still, they can be broadly classified into four major classes – structural roles, molecular mimicry of host glycans, extrinsic, and intrinsic recognition[3] (Figure 3A.1).

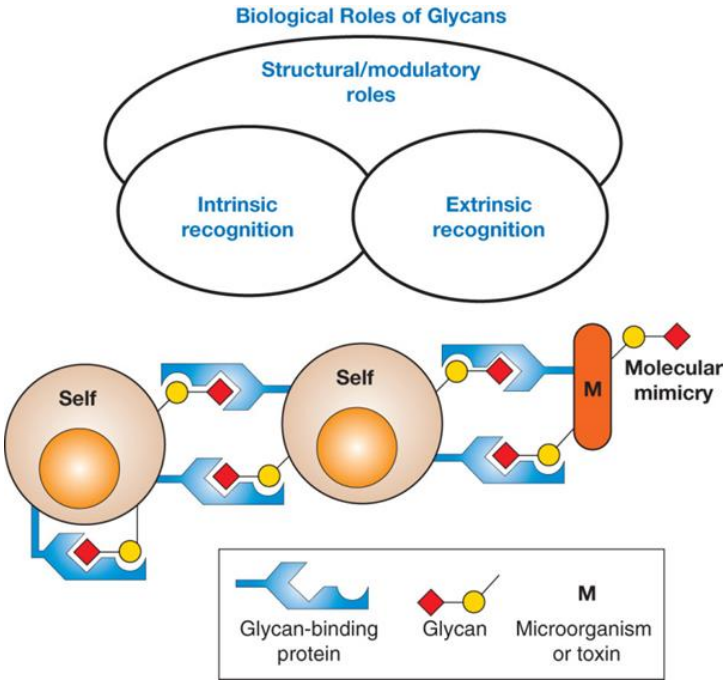


Figure 3A.1: General classification of the biological roles of glycans – structural roles involve providing rigidity to cell walls or exoskeletons, microorganisms often mimic host-cell glycans to evade immune response, extrinsic recognition of host-cell glycans by pathogens determines whether they will have a beneficial or detrimental effect on the host, and intrinsic recognition of terminal glycans leads to triggering of functions such as endocytosis (Reproduced with permission from Varki, *Glycobiology*, 2017) .

Aberrant glycosylation is a classical hallmark of malignant tumor progression. Since glycosylation is not template-driven, the mechanisms producing such glycosylations are broad and unclear[4]. But regardless of the change in the underlying glycan structure, the upregulation and alteration of terminal sialic acid (Sia) structures remains a hallmark of cancer. It is responsible for the reduced adhesion of tumor cells to the extracellular matrix or increased plasticity of the cells within the tissue matrix and to avoid detection by immune cells possessing Sia recognizing lectins[5]. The Sia is a subclass of a superfamily called nonulosonic acids, the simplest one being the neuraminic acid (Neu). The Sia family shares a characteristic nine carbon chain backbone with a carboxylic acid residue on the anomeric carbon (Figure 3A.2). It is perhaps the most versatile of the glycan platforms, and the structural flexibility provided by the Sia enables it to store potential chemical information rendering it to be described as sugar codes. Currently, three sialyl-glycan antigens expressed in particular cancers have been widely researched as glycan biomarkers[6] (Sialyl-T_N, Sialyl-Lewis X, and Sialyl-Lewis A shown in Figure 3A.2), and their high expression elucidates cancer progression at multiple levels[7]. In addition, aberrant sialylation and glycosylation have been identified as general molecular indicators for cancer progression[8,9]. Existing clinical tests are often expensive, labor-intensive, and technically challenging[10], though they are substantially valuable. The significance of sialyl-glycan in pathophysiology demands extensive research on designing inexpensive probes for the potent sensing sialyl-antigens.

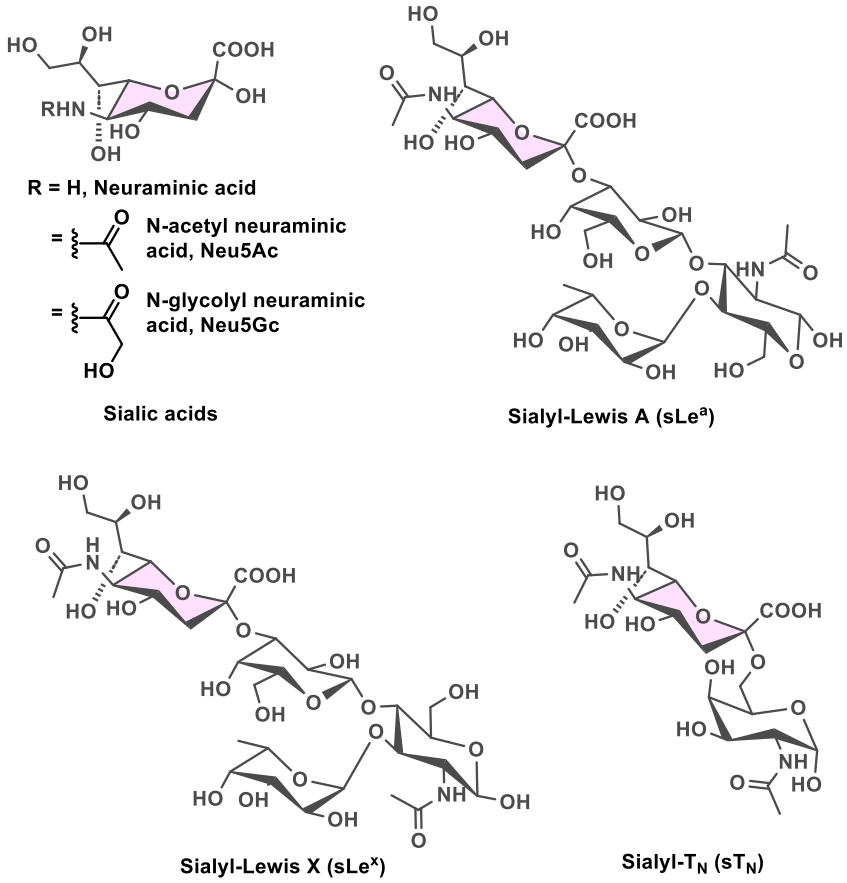


Figure 3A.2: Structures of sialic acids and the commonly researched sialyl antigens (the sialic acid $^2C_5$ pyranose ring has been coloured lavender for ease of understanding).

3A.1.2 *BORONIC ACIDS AS SUITABLE CANDIDATES FOR AUGMENTING SIALIC ACID RECOGNITION*

Typical protein-protein interactions have binding affinities in the nanomolar range, while studies have found glycan-protein interactions to be in the micromolar or even millimolar range[3]. This frequently necessitates the need for multivalent avidity to generate effective biological functions. In biology, multivalency is the general state at the cellular level, hence, a high-affinity interaction between protein and glycan is observed in nature. Also, in a biological settings, one must keep in mind that the extensive solvation imposed by water molecules hijack non-covalent binding partners, which troubles a sensor to function efficiently. In this regard, boronic acid (BA)-based small molecules named 'boronolectins' have significantly improved the sensing capability of carbohydrates[11,12]. More specifically, active peptide fragments derived from glycan-recognizing proteins, when operating outside of their native environment, can be 'evolved' into a better binder by augmenting it with a boronic acid handle. Given the advancements in designing boronic acid derivatives finetuned to specific applications [13], we propose that research on boronopeptide probes may unravel the long-term challenges of developing inexpensive glycan-specific probes. Although at physiological pH, simple boronic acid moieties do not show binding selectivity to sialyl-glycans[14]. For instance, Matsumoto et al. have shown that 5-boronopicolinic acid shows high selectivity and affinity at acidic pH but not physiological pH[15]. Of course, 5-boronopicolinic acid merits its application for specific tumor imaging. The concept of designing boronopeptides for selective recognition of sialic acid (Sia) is previously documented to support this hypothesis[2,16–20]. Nonetheless, we noticed a lack of glycan selectivity studies only with the native peptide sequences.

We aimed to design a boronopeptide covalent receptor bearing a monovalent boronic acid probe, which is highly selective for the recognition of sialyl-glycans because multiple boronic acid probes installed in a peptide sequence may cause off-target modifications[2]. The design of such boronopeptide receptors necessitates the choice of a highly selective Sia binding peptide sequence or scaffold, which is crucial for the application in a complex biological milieu. Based on this postulate, we searched through lectins that bind to sialyl-glycan and their topology of the binding site. These studies motivated us to pick the wheat germ agglutinin (WGA) protein and its peptide fragment (62–73, named WGA peptide, Figure 3A.3a) that reportedly binds to Sia with K_d ~0.8 mM[21]. However, the binding selectivity of WGA peptide with other monosaccharides present in vertebrates was never disclosed. We also explored the selectivity of WGA peptide to Sia recognition and its binding modulation abilities to sialyl-glycan when chemically distinct boronic acid derivatives were installed chemoselectively on a cysteine (Cys) residue. We postulated that the tailored boronopeptides would enhance binding affinity in cancer cell surface labeling by the synergistic recognition governed by selective non-covalent interactions and cis-diol boronic acid conjugation (Figure 3A.3b).

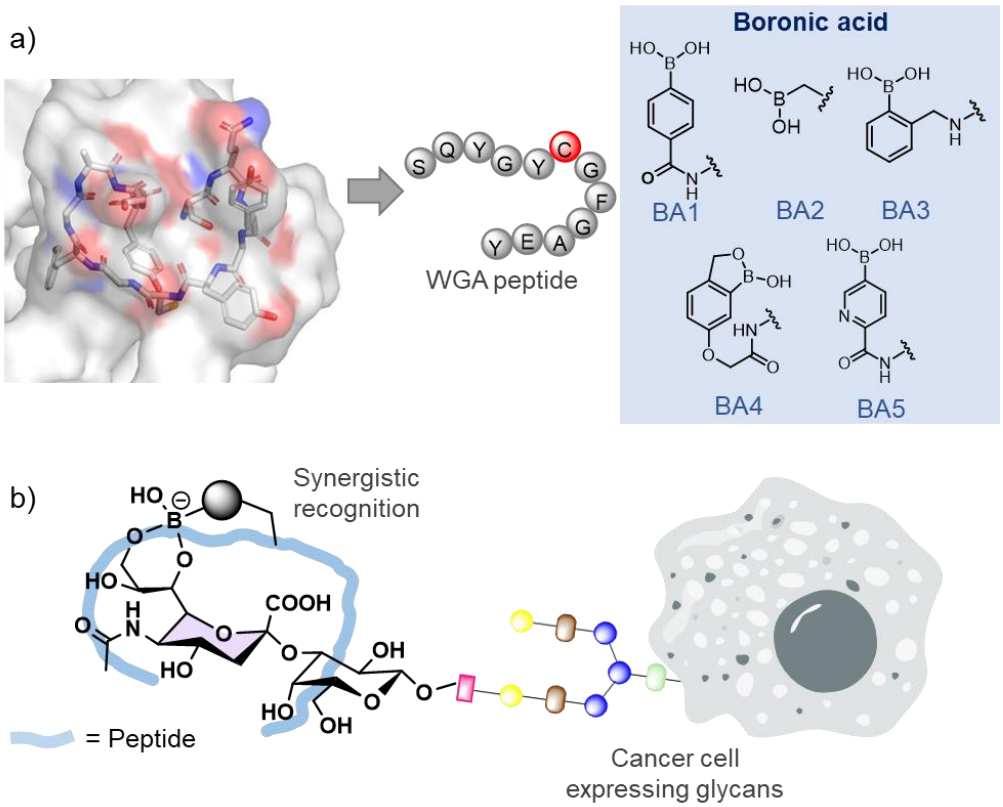
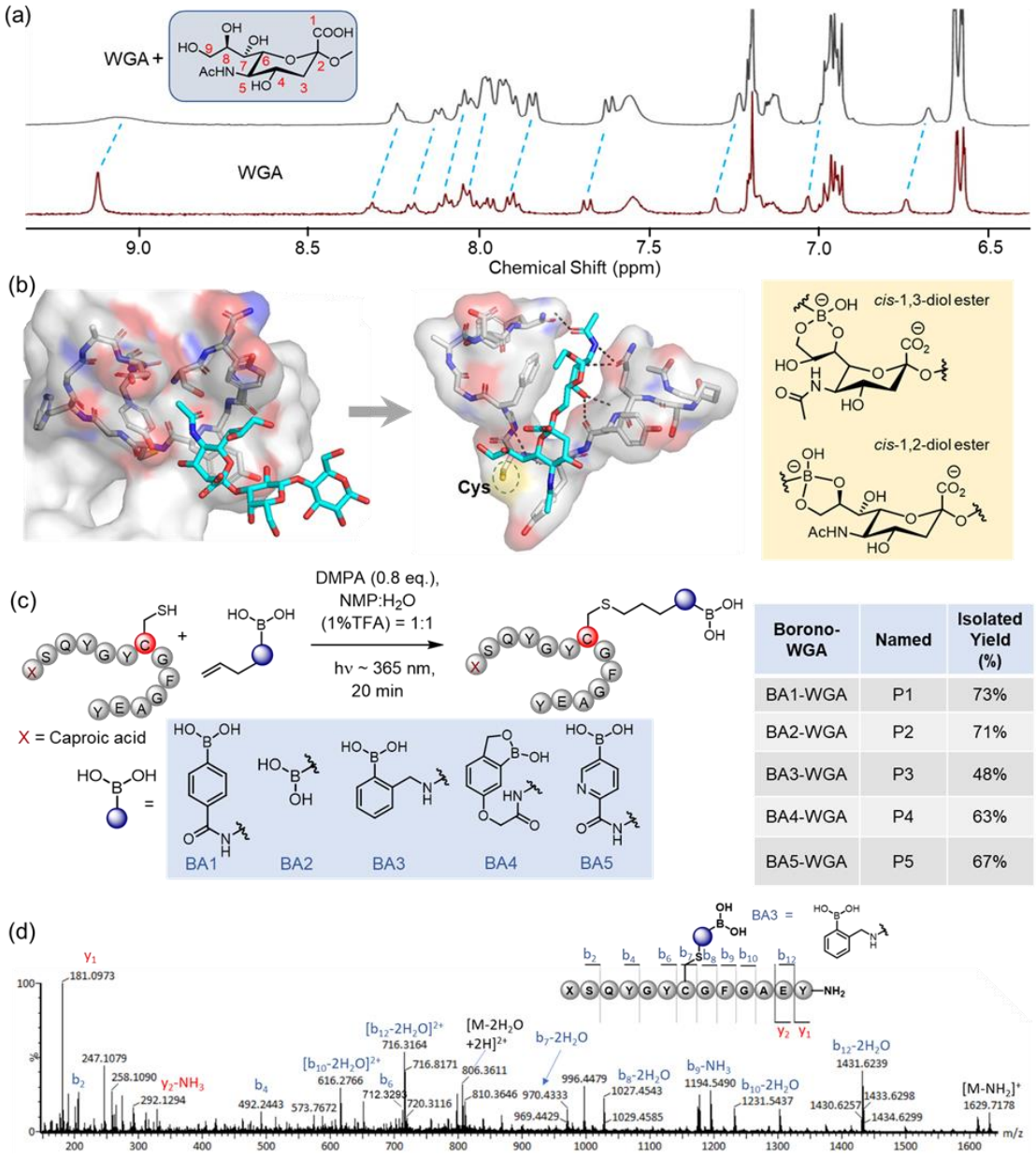


Figure 3A.3: a) The protein structure (PDB: 1WGC) highlights the binding epitope of sialyl-glycan. The sequence of WGA peptide and structure of boronic acid derivatives used for binding modulation of the WGA epitope sequence is presented; b) The anticipated synergistic binding of WGA boronopeptide to sialyl-glycan present on a cancer cell surface. In the binding trajectory, the blue ribbon represents the WGA peptide and the boronic acid is shown complexed with Sia.

3A.2 *RESULTS AND DISCUSSION*

3A.2.1 *DESIGN OF BORONO-WGA PEPTIDE*

We first investigated the binding selectivity of WGA peptide to ten abundant monosaccharides present in vertebrates via ^{1}H-NMR studies. When WGA was incubated with Sia, significant changes in chemical shifts of WGA amide protons were observed (Figure 3A.4a), which were absent from the other nine abundant monosaccharides in vertebrates. In the NMR study, we used methylated Sia, which means the anomeric -OH of Sia was protected with a methyl group because protecting the anomeric -OH eliminates the ambiguous binding mode of boronic acid to Sia C1/C2. Therefore, the later NMR investigations with designed borono-WGA peptides would be unambiguous. It has been reported that BA exclusively captures the α-hydroxy acid group at pH < 8 when the carboxyl group or anomeric -OH of Sia is unprotected[22]. The protection of anomeric -OH with methyl group uniquely mimics the binding mode via cis-diol, either C7/C9 or C8/C9 of the Sia glycerol chain, preferentially C7/C9 because of the formation of relatively stable six-membered cyclic boronate ester (Figure 3A.4b)[23].



Borono-WGA	Named	Isolated Yield (%)
BA1-WGA	P1	73%
BA2-WGA	P2	71%
BA3-WGA	P3	48%
BA4-WGA	P4	63%
BA5-WGA	P5	67%

Figure 3A.4: a) Stacked ^{1}H-NMR data of WGA peptide with (top) and without (bottom) Sia at 2 mM in DMSO-d_6. Dotted blue lines designate a sharp change in chemical shifts with Sia only (full NMR shown later); (b) The design strategy of borono-WGA peptide in-silico: Crystal structure of WGA protein (PDB: 1WGC) with N-acetyl neuraminyl-lactose (left) inspired docked structure between WGA peptide and sTn antigen showing a proximal distance of Cys residue to cis-diols (middle). An anticipated *cis*-1,2 or 1,3-diol conjugation with BA is represented by chemical structures (right); (c) General scheme for the installation of BA1-5 on WGA peptide and their isolated yield after HPLC purification. (d) MS/MS data for WGA-BA3 (P3) peptide clearly demonstrate the cysteine selective installation of BA3 probe to WGA.

Further, we were interested in looking at the conformational integrity of the trimmed WGA peptide sequence, and it was compared with the WGA protein. A structure-minimizing simulation has revealed a few deviations in the structural backbone and side chain. Interestingly, molecular docking between in-silico adopted WGA structure and clinically important tumor biomarkers, Sialyl-T_N, showed a significant binding (Figure 3A.4b). Nonetheless, the Sia-moiety recognition pattern by the WGA peptide was found to be different from that of the WGA protein. We arbitrarily chose sTn to understand the free WGA peptide

recognition pattern since terminal Sia follows acetyl-galactosamine or galactose frequently, except for polysialic acid in glycome. The best-docked pose between WGA and sTn suggested that installing BA derivatives in the middle of the peptide sequence would enable boronate ester formation with the cis-diol present in sialic acid (Figure 3A.4b). Hence, we postulated that installing an appropriate BA derivative on WGA would help recognize sialyl-glycan tightly via cooperative molecular recognition.

3A.2.2 *SYNTHESIS OF BORONO-WGA PEPTIDES AND CONTROL PEPTIDES*

Encouraged by the NMR data and computational models, chemically diverse and widely used five BA building blocks (BA1-BA5, Figure 3A.3a) were installed on the Cys site of the WGA peptide chemoselectively using recently reported late-stage BA installation method[24] (P1-P5) (Figure 3A.4c). A caproic acid residue was incorporated at the N-terminus of the WGA peptide for labeling FITC. BA-modified peptides were isolated after HPLC purification in a moderate yield, and their identity was characterized using mass spectrometry. A well-known signature of water loss for BA-containing molecules[24] was observed in mass spec data. MS/MS studies of P3 support the cysteine selective installation of the BA3 probe to WGA (Figure 3A.3d) by thiol-ene click chemistry. A detailed comparison of MS/MS data P3 and the WGA peptide is shown in the Experimental section.

A few controls were included to carry out the cell surface sialyl-glycan binding studies with the borono-WGA peptides. Therefore, Cys alkylated-WGA peptide (P0) was considered a control to compare binding advances with borono-WGA peptides to probe the binding affinity of WGA alone. We have included the scrambled peptide sequence (P6, Figure 3A.5a) as a negative control, which is not expected to bind to Sia. All the peptides were labeled with FITC through 6-amino caproic acid incorporated at the N-terminus and further purified via RP-HPLC to enable their visualization on the cell surface binding.

3A.2.3 *THE APTITUDE OF SIALYL-GLYCAN RECOGNITION BY BORONO-WGA PEPTIDES*

Next, FITC-labeled borono-WGA were subjected to the cell surface of two cancer cell lines to assess the binding enhancement to sialyl-glycan, compared to WGA peptide. Adenocarcinoma human alveolar basal epithelial cells (A549) and epithelial human breast cancer cell line (MDA-MB231) were treated with all peptides (~10 μg/mL in PBS) tabulated, including controls at ice bath temperature. We chose *Sambucus nigra* lectin (SNA) as a positive control (~10 μg/mL), which is reported to bind to α-2,6 linked sialic acids preferentially on the cell surface glycans[25]. Peptide and SNA-treated cell lines were immediately washed in PBS (pH 7.4) and subjected to quantitative analysis in flow cytometry. We chose this particular concentration (~10 μg/mL) because SNA shows >90% binding, as revealed by titration in flow cytometer experiments. Initially, we wanted to compare and explore the efficacy of the designed boronopeptides' staining with SNA at this concentration.

Borono-WGA peptides (P1–P5) exhibited a particular hierarchy of staining efficiency, similar for both cell lines in the flow cytometry assay. Peptide P0 (Figure 3A.5b–c) showed slight staining as it is a weak binder of Sia; however, scrambled WGA (P6, Figure 3A.5d) did not endorse substantial contrast in the labeling of the cell lines compared to untreated cells. Boronopeptide, P1, P2, and P4 showed a better binding ability

than P0. In our design, P3 and P5 revealed a substantial binding enhancement with both cell lines, in which P3 offered the best sialyl-glycan recognition. A clear comparison is shown in bar graphs for both cell lines. To clarify the contribution of BA3 individual binding to sialyl- glycan, we have included the BA3-installed scrambled WGA (P7) in the experiment of A549 labeling. Interestingly, P7 did not exhibit significant binding, which was apparent when staining was compared with the scrambled WGA (P6) and P3 (Figure 3A.5d). It showed non-significant, non-specific staining similar to the scrambled WGA (P6). However, the positive control (SNA) showed high binding ability on the A549 cell surface compared to MDA-MB231 (Figure 3A.5b–c). The flow cytometry experiments undoubtedly demonstrate higher affinity binding through synergistic recognition corroborated by boronic acid installed WGA peptide, which is difficult to achieve by WGA peptide alone or individual boronic acid probe BA3.

We further sought to experiment on determining EC_{50} for P3 and P5 on the cell surface because they exhibited similar staining efficiency on both cell lines. A549 cells were incubated with serial concentrations (0.2, 0.5, 1, 2, 5, 10, 20 μg/mL) of P0, P1, P3, and P5 in an ice bath and subjected to flow cytometer analysis after a couple of stringent washes. To compare the data in a better resolution, we have also included P0 and P1 in this EC_{50} study quantitatively. The data exhibited a clear picture that P3 enables the best binding with EC_{50} 3.9 μg/mL, while P5 showed a slightly weaker binding affinity with EC_{50} 6.1 μg/mL (Figure 3A.5e). Compared to P0 (EC_{50} >20 μg/mL), the addition of the BA3 probe on WGA at the appropriate position revealed synergistic binding and increment of binding efficiency >5 folds. It has been reported that a monomeric state of BA shows principally weak binding for glycans' recognition compared to lectin, which is also evident in our experiment. However, probes containing suitably organized multivalent BA showed higher binding affinity in biomedical applications[11]. Still, they limit broad applications because of poor glycan selectivity in general[19]. In our studies, P3 showed the highest affinity in the panel of designed borono-WGA peptides due to a better accessibility and association ability of the sialic acid diol by BA3-moiety (Wulff-type boronate)[12,26] attached to WGA. Perhaps an additional ionic interaction between the benzylamino group of BA3 and carboxylate in Sia favors tight recognition. It has been shown that Wulff-type boronates show a better association with cis-diol over other popular boronic acid derivatives used in carbohydrate recognition[27]. Although boronopeptide association with sialyl-glycan involves the synergy between non-covalent interactions and diol complexation, we found that the BA3 probe installed on WGA associates better in cell surface sialyl-glycan recognition synergistically. The positive control (SNA) showed a high contrast of labeling when compared with the best labeling agent (P3) for A549 staining (Figure 3A.5d), which is consistent under an epifluorescence microscope. The P6 served as an appropriate negative control (Figure 3A.5d–e) in the experiments demonstrating the true binding events on the cell surface. We found that the P3 is stable in the open air at physiological pH even after 48 h, which depicted good chemical stability of the probe. However, WGA-BA3 (P3) showed only ~45% degradation in serum compared to near complete degradation for WGA after 24 hr, demonstrating BA3 installation improved the stability of the WGA peptide in serum. These systematic live cell experiments illustrate the future perspective of capturing sialyl-glycan by boronopeptides, thus mimicking lectin-like recognition properties.

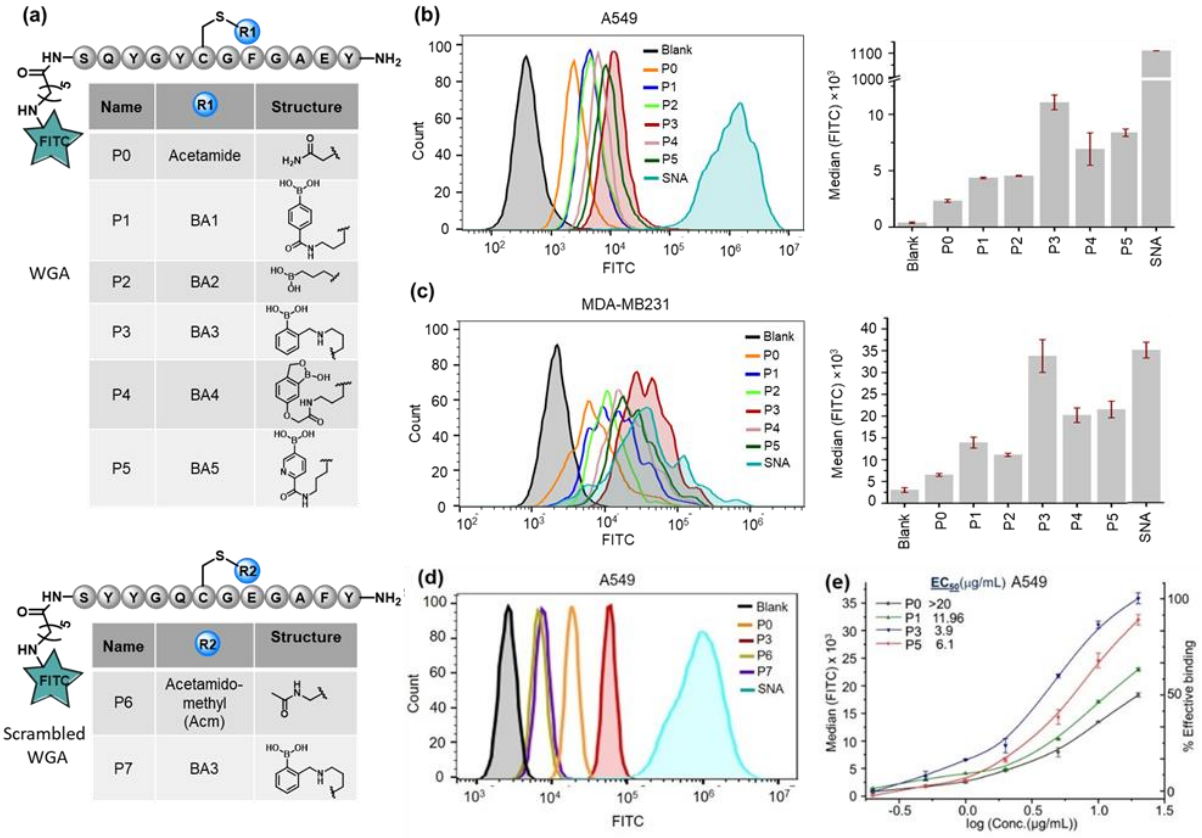


Figure 3A.5: a) Structure and tabulated list of FITC labeled WGA, borono-WGA (**P1–5**) and control peptides (**P6–7**) were synthesized for cell binding assays. A quantitative staining comparison in flow cytometry among **P0** to **P5** peptides experimented on b) A549 and c) MDA-MB-231. Corresponding error bar graphs (SD) represent clear comparisons of staining; d) A clear contrast of binding comparison among untreated cells, control peptides (**P6–7**), positive control (SNA), and **P3** on the surface of A549 cells: quantitative analysis was performed in flow cytometry; e) Determination of cell surface binding median effective concentration (EC_{50}) of boronopeptide **P1**, **P3**, and **P5**, which was compared with WGA peptide (**P0**). The flow cytometry data revealed more than five-fold increments of EC_{50} for **P3** over **P0**. All flow cytometer experiments are technical replicates (n=1, triplicated data represented with error bar graphs by calculating standard deviation). However, some crucial flow experiments were repeated during EC50 measurements and found to be reproducible. One set of data is presented in b–d as histograms to compare the flow cytometer readout.

3A.2.4 *SELECTIVITY TO SIALYL-GLYCAN AND SYNERGISTIC BINDING*

It is essential to comprehend the binding selectivity of the peptide P3 in the presence of abundant serum proteins for further applications. Thus, the staining attainability of P3 with A549 cells was measured in the presence of FBS (5% and 10%). Flow cytometry data demonstrates ~50% inhibition of staining efficiency with 5% FBS, which is nearly unaffected by the increase of FBS concentration to 10% (Figure 3A.6a). Consequences are perhaps not surprising as FBS contains sialyl-glycan proteins[28], which cause inhibition in binding. However, the dynamic nature of boronate ester conjugate between P3 and sialyl-glycan establishes a binding equilibrium between serum proteins and cell surface glycans. We presumed that the P3 binding association equilibrium favored the cell surface due to the high local concentration of sialyl-glycan, which in turn did not compromise the staining efficacy by the rise in serum percentage. Dynamic covalent probes

show advantages in lowering off-target modification in general, and we observed through the selective modification of densely populated silyl-glycan on cancer cell surfaces in the presence of serum.

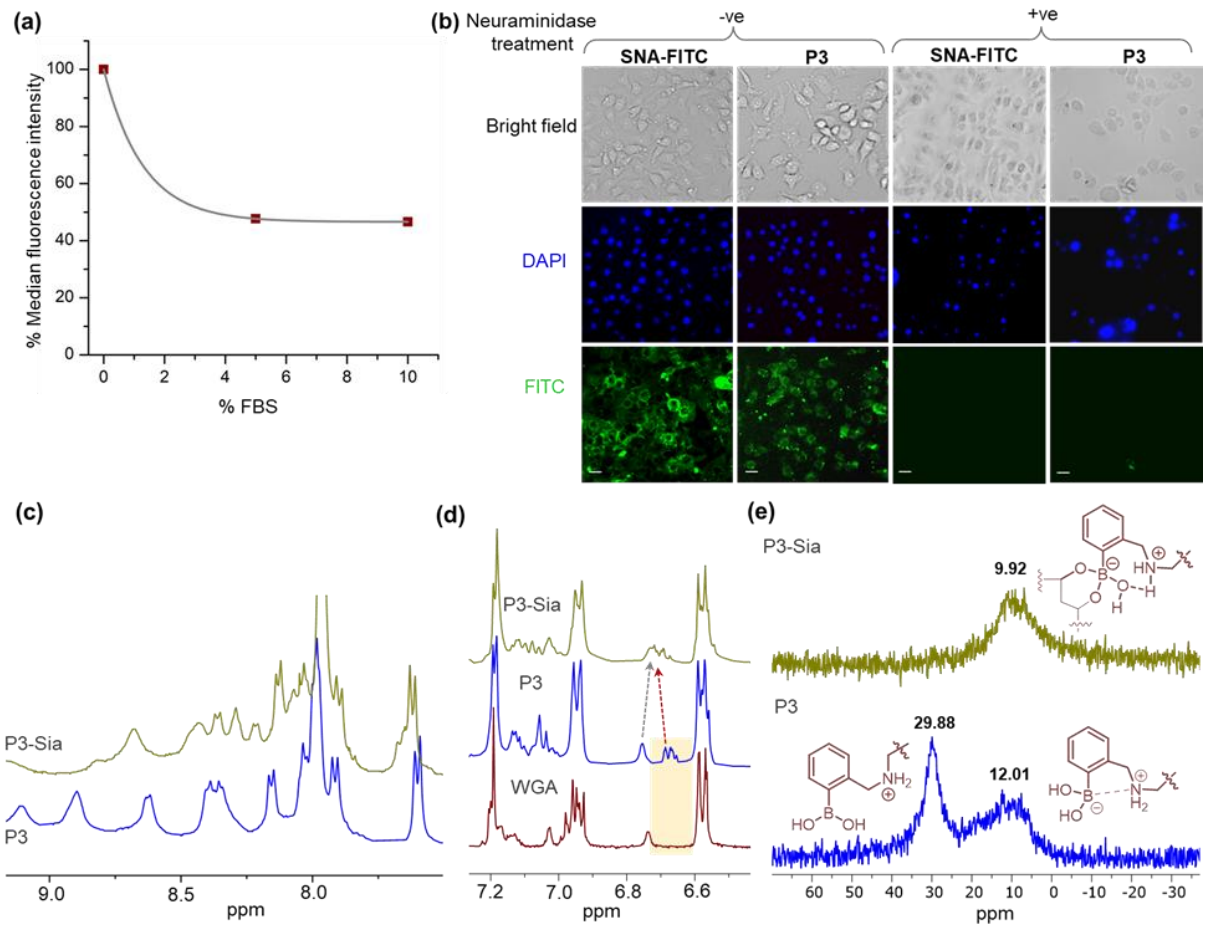


Figure 3A.6: a) FBS inhibition assay of **P3** staining with A549 cells, in which fluorescent intensity was normalized to 100% with no FBS. The experiments were performed twice. Nearly similar staining inhibitions have resulted in median values; b) Epifluorescence microscopic image of MDA-MB231 after treatment with SNA and **P3**. Neuraminidase-treated MDA-MB231 failed to stain with P3 and positive control (SNA-FITC), replicated in the FITC channel. (The scale bar is 20µm); c) ^{1}H NMR data demonstrate that the BA3-modified WGA peptide sequence retains the binding ability with Sia; d) Stacking spectra of WGA, **P3** and **P3-Sia** complex and highlighting the clear downfield shift (marron arrow) of the prominent peak (absent in WGA shown in gold box) associated with BA3 moiety; e) ^{11}B-NMR of P3 and P3-Sia complex (1:1, 1 mM) depicts boronate ester formation, studied in a quartz NMR tube.

Interestingly, MDA-MB231 failed to stain with P3 under a fluorescence microscope when cells were pre-treated with Neuraminidase[29], highlighting P3 binding selectivity on terminal Sia. The results were compared with Neuraminidase untreated, freshly grown MDA-MB231 cells (Figure 3A.6b). Neuraminidase selectively cleaves sialic acid from glycan terminals. Therefore, the absence of Sia on the cell surface fails to demonstrate P3 binding, which is analogous to the positive control lectin SNA-FITC (Figure 3A.6b). The data unambiguously prove that P3 selectively and tightly binds to sialyl-glycan on the cell surface. Finally, NMR analysis shows that the BA3 modified WGA retains its binding selectivity with sialic acid. The P3 showed changes in amide resonances when Sia was bound, evident via the clustering of NH-protons (Figure 3A.6c). However, these changes are absent with other abundant sugars, analogous to the initially studied

WGA peptide. These apparent chemical shifts indicate conformational changes and the adoption of a unique structure in the P3 sequence due to Sia recognition. In addition, a characteristic aromatic peak at ~6.8 ppm associated with BA3 moiety, which is absent in the WGA parent peptide, shows a slight downfield shift (Figure 3A.6d) upon complex formation, endorsing participation in Sia binding. However, this downfield shift in the aromatic region was absent in pure covalent diol complexation between BA3 and Sia. We reason that such a chemical shift at ~6.8 ppm resulted from synergistic binding to Sia by P3 due to the preorganization of the peptide backbone and side chains. The chemical shift in ^{11}B NMR (from ~30 and ~12 ppm to ~10 ppm) unambiguously supports boronate ester formation between P3 and Sia (Figure 4e). Such chemical shifts are associated with N–B interactions and boronate ester formation, as reported by the Ansyln group[30,31]. ^{1}H and ^{11}B NMR data collectively admit our hypothesis of the synergistic binding module by WGA-BA3 and narrate why P3 shows a higher binding affinity than the WGA peptide. Due to the poor solubility of P3 in PBS buffer, these NMR investigations were performed in DMSO at a millimolar concentration. These collective investigations strongly suggest that P3 selectively recognizes sialyl-glycan through a synergistic mechanism.

3A.3 *CONCLUSIONS*

To summarize, we showed that the WGA peptide selectively binds to Sia via NMR studies, and in-silico investigation of the WGA binding module to sialyl-glycan has assisted in correctly positioning BA derivatives in designing borono-WGA peptides. Further, quick installation of BA derivatives on WGA peptide by thiol-ene reaction has enabled screening of sialyl-glycan recognition on two cancer cell lines, A549 and MDA-MB231. The in vitro studies of BA3-modified WGA (P3) demonstrated more than five-fold higher recognition ability than the WGA peptide based on EC50 of cell surface binding. In addition, P3 showed sialyl-glycan selectivity on par with SNA lectin on MDA- MB231. We are still uncertain about the binding specificity of P3 among the isomers of sialyl-glycans at this early investigation stage. It is also challenging to attain the selectivity of sialyl- glycan binding in a complex glycome by synthetic molecules. The research on unraveling the structure of the P3-Sia complex is ongoing in our laboratory by combining 2D-NMR and MD simulation to construct superior probes for biomedical applications, which could be healthier than widely used lectin[32]. Encouragingly, this synergistic binding strategy could be extended by contemplating explicitly designed peptide ligands in selective and efficient labeling of specific glycan antigens.

3A.4 *EXPERIMENTAL SECTION*

3A.4.1 *CHEMICALS, CELLS AND GENERAL METHODS*

All the chemicals were purchased from commercial sources and used without further purification. The reagents for peptide synthesis such as Fmoc-protected amino acids, 2-(1*H*-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), and polystyrene rink amide AM resin (loading capacity 0.67 mmol/g) were purchased from Chem-Impex Int'l Inc (Wood Dale, IL). Neuraminidase (from Clostridium perfringens, *C. welchii*, Type V lyophilized powder), all monosaccharides, were purchased

from Sigma Aldrich and 4,6-diamidino-2-phenylindole dihydrochloride (DAPI, Molecular Biology Grade), were purchased from SRL.

Roswell Park Memorial Institute (RPMI-1640) media for the A549 cell line, Dulbecco's Modified Eagle Medium (DMEM) media for the MDA-MB-231 cell line, Fetal bovine serum (FBS) and Trypsin–EDTA Solution 1X was purchased from Gibco. PenStrep, Dulbecco phosphate-buffered saline (DPBS, pH 7.4), and Dimethyl sulfoxide (DMSO, Molecular Biology grade) were purchased from Himedia. FITC-Sambucus Nigra Lectin (SNA, EBL) was purchased from Vector Laboratories. Cancer cell lines are the source of ATCC: MDA-MB-231 (HTB-26) and A549 (CCL-185).

All reactions were monitored by thin-layer chromatography (TLC) on aluminium-backed silica gel 60 F_{254}, 0.2 mm plates (Merck), and compounds were visualized under UV light (254 nm) or charred with appropriate solutions – ninhydrin charring to visualize amines, DNP charring for aldehydes, $KMnO_4$ charring for phenolic OH, and curcumin or anisaldehyde charring for boronic acids. Synthetic compounds were purified on 60-120 or 100-200 silica gel purchased from Finar.

3A.4.2 *INSTRUMENTS*

NMR data were recorded on a 400 MHz Jeol JNM ECS400 NMR spectrometer. Mass-spec data were generated by a Waters XEVO G2-XS QTOF mass spectrometer. HPLC analysis and peptide purification were carried out on a Shimadzu Prominence UFLC system. UV-Vis studies were carried out on a Shimadzu UV-2600 system. Melting point studies were carried out in Stuart SMP30. Flow cytometry analyses were carried out on a BD Accuri C6 Plus. The EC_{50} studies were carried out on a Sysmex Cube 6 system (FSC voltage 250, SSC voltage 350, FL1 550). Fluorescence images were taken on a Leica DMi8A microscope.

3A.4.3 *GENERAL METHODS FOR LC-MS AND HPLC ANALYSIS*

LC-MS: Waters Acquity UPLC C18 (1.7 µm, 2.1×50 mm) analytical column and the mobile phase acetonitrile-water (0.1% HCOOH). All LC analyses were carried out in Method A, and all MS analyses were performed in ESI +ve mode unless otherwise specified.

Method A: Flow rate 0.4 mL/min. The gradient used: gradient 5% to 35% ACN in 2.5 min, then a gradient from 35% to 95% ACN in 2 min followed by isocratic 95% over a min, and then column equilibration at the initial isocratic gradient 5% ACN over 1.5 min.

Method B: Flow rate 0.25 mL/min. The gradient used: isocratic 5% ACN for 0.6 min, followed by the gradient 5% to 60% ACN till 11.5 min, then a gradient from 60% to 90% ACN in 1 min, followed by the gradient to 5% in 0.5 min and then a column equilibration at the initial isocratic gradient 5% ACN over 3 min.

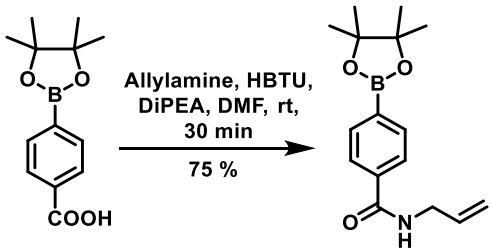
HPLC methods for reaction monitoring and peptide purity analysis (unless otherwise specified): Waters Reliant C18 (5 µm, 4.6×250 mm) analytical column and mobile phase acetonitrile-water (0.05% TFA) with a flow rate of 1 mL/min were used. The gradient used: isocratic 10% acetonitrile over 1 min, then a gradient from 10% to 50% acetonitrile in 19 min, then a gradient from 50% to 90% acetonitrile over 2 min, followed

by isocratic 90%acetonitrile over 3 min, then column equilibration at the initial isocratic gradient 10% acetonitrile over 5 min.

HPLC method for peptide purification through semi-preparative (unless otherwise specified): Shimadzu shim-pack GIST C18 (5 µm, 10×250 mm) semi-preparative column and mobile phase acetonitrile-water (0.05% TFA) with a flow rate 4 mL/min were used. The gradient used: isocratic 10% acetonitrile over 2 min, then a gradient from 10% to 50% acetonitrile over 26 min, followed by a gradient from 50% to 90% acetonitrile over 3 min, then isocratic 90% acetonitrile over 5 min, followed by column equilibration at the initial isocratic gradient 10% acetonitrile over 5 min. In the case of WGA-BA3, we used 0.1% AcOH instead of 0.05% TFA for better separation.

3A.4.4 *SYNTHESIS OF BORONIC ACID DERIVATIVES AND 2-OME Neu5Ac (Sia)*

3A.4.4.1. *SYNTHESIS OF **BA1***



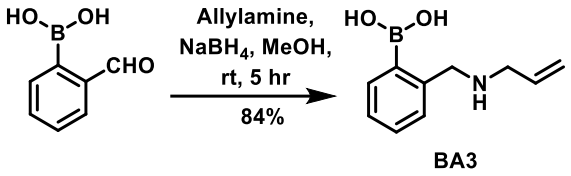
4-Carboxylphenylboronic acid pinacol ester (372 mg, 1.5 mmol), HBTU (570 mg, 1.5 mmol, 1 eq.), and DiPEA (385 µL, 2.25 mmol, 1.5 eq.) were dissolved in 1.5 mL DMF and stirred at room temperature. To this solution, allylamine (123 µL, 1.65 mmol, 1.1 eq.) was added slowly after 5 min, and the reaction mixture was further stirred for 30 min. TLC analysis indicated the completion of the reaction. The reaction mixture was acidified with chilled 50 mL 1N HCl and extracted with 2×50 mL EtOAc. The organic layers were washed with chilled 50 mL 10% $NaHCO_3$, then combined and washed with brine. Upon evaporation of the organic layer under reduced pressure, it was purified through silica gel using 30% EtOAc/Hexane to yield 320 mg of **BA1** (75% yield) as a white powder, mp: 161-163 °C.

^{1}H NMR (400 MHz, $CDCl_3$): δ 7.84 (d, *J* = 8.1 Hz, 2H), 7.75 (d, *J* = 8.1 Hz, 2H), 6.32 (s, 1H), 5.96-5.87 (m, 1H), 5.26-5.15 (m, 2H), 4.08 – 4.05 (m, 2H), 1.33 (s, 12H).

^{13}C NMR (101 MHz, $CDCl_3$): δ 167.3, 136.7, 135.0, 134.2, 126.1, 116.9, 84.2, 42.5, 29.8, 25.0.

HRMS-ESI^{+} (m/z): $[M\text{-pin}+H]^{+}$ Calc. 206.0988, Obs. 206.0986.

3A.4.4.2. *SYNTHESIS OF **BA3***



BA3

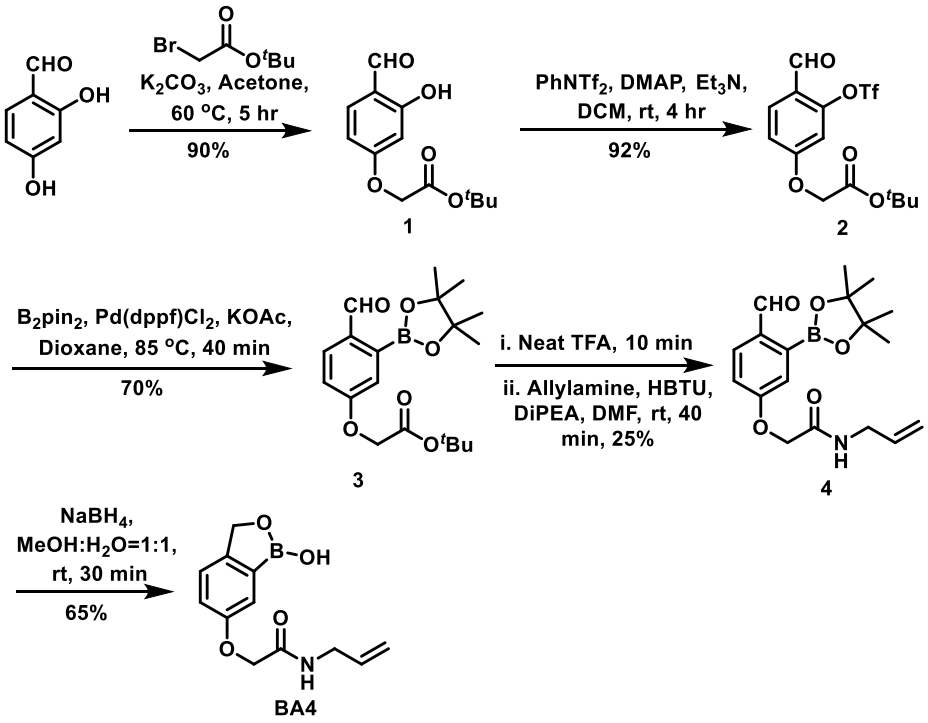
2-FPBA (75 mg, 0.5 mmol) and allylamine (38 µL, 0.5 mmol, 1 eq.) were dissolved in 3.7 mL dry MeOH and stirred at room temperature. After 5 min, some 4 Å molecular sieves (10% w/v) were added. After cooling the solution for 5 min at 0 °C, $NaBH_4$ (21 mg, 0.55 mmol, 1.1 eq.) was added and further stirred for 15 min. TLC analysis indicated consumption of 2-FPBA. The reaction mixture was filtered, and MeOH was evaporated under reduced pressure. The residue was suspended in 50 mL 1N HCl and washed with 50 mL EtOAc to remove the organic by-products. The volume of the aqueous layer was reduced to 25 mL, neutralized to pH 7 by ammonia solution and washed with 3×100 mL EtOAc. The organic layer was washed with brine, evaporated to 75 mL and kept in a beaker where the desired product crystallized overnight. The crystals were collected, dissolved in 0.1 N HCl and lyophilized to obtain 80 mg of **BA4** (84% yield) as a white powder, mp: 228-230 °C.

^{1}H NMR (400 MHz, DMSO-d_6) δ 8.66 (s, 3H), 7.81-7.79 (m, 1H), 7.45-7.35 (m, 3H), 5.92-5.81 (m, 1H), 5.47-5.37 (m, 2H), 4.24 (t, J = 6.0 Hz, 2H), 3.57 (q, J = 5.7 Hz, 2H).

^{13}C NMR (101 MHz, DMSO-d_6): δ 137.2, 136.1, 131.6, 130.6, 129.6, 128.8, 123.2, 50.6, 49.4.

HRMS-ESI$^+$ (*m/z*): $[M+H]^+$ Calc. 192.1196, Obs. 192.1196.

3A.4.4.3. *SYNTHESIS OF **BA4***



BA4

*Synthesis of **1***

2,4-Dihydroxybenzaldehyde (2.2 g, 16 mmol), *tert*-butyl bromoacetate (2.5 mL, 16.8 mmol, 1.05 eq.), and K_2CO_3 (2.2g, 16 mmol, 1 eq.) were mixed in 32 mL acetone and refluxed at 60 °C for 5 hr. The reaction mixture was then cooled and filtered to remove K_2CO_3. The acetone was evaporated under vacuum and purified through silica gel using 5% EtOAc/Hexane to obtain 3.63 g of **1** (90% yield) as a white powder.

^{1}H NMR (400 MHz, $CDCl_3$): δ 11.42 (s, 1H), 9.71 (s, 1H), 7.44 (d, *J* = 8.7 Hz, 1H), 6.55 (dd, *J* = 8.6, 2.5 Hz, 1H), 6.35 (d, *J* = 2.4 Hz, 1H), 4.54 (s, 2H), 1.47 (s, 9H).

^{13}C NMR (101 MHz, $CDCl_3$): δ 194.6, 166.9, 165.0, 164.4, 135.5, 115.8, 108.6, 101.5, 83.1, 65.6, 28.1.

HRMS-ESI$^+$ (*m/z*): $[M+H]^+$ Calc. 253.1076, Obs. 253.1074.

*Synthesis of **2***

1 (2.8 g, 11.15 mmol), $PhNTf_2$ (4.38 g, 12.26 mmol, 1.1 eq.), and DMAP (136.2 mg, 1.1 mmol, 0.1 eq.) were dissolved in 20 mL dry DCM and stirred at room temperature for 5 min. Et_3N (3.1 mL, 22.3 mmol, 2 eq.) was added dropwise to the reaction mixture when the colour of the solution changed from light yellow to deep yellow. It was further stirred for 4 hr, and then DCM was evaporated. The residue was dissolved in 300 mL EtOAc, washed with 120 mL 1N HCl, and 150 mL 10% Na_2CO_3 to remove the by-product, and the combined organic layers were washed with brine. Upon evaporation of the solvent, the product was purified through silica gel using 8% EtOAc/Hexane to obtain a 3.94 g of **2** (92% yield) as a white solid.

^{1}H NMR (400 MHz, $CDCl_3$): δ 10.12 (s, 1H), 7.94 (d, *J* = 8.7 Hz, 1H), 6.99 (dd, *J* = 8.7, 2.3 Hz, 1H), 6.86 (d, *J* = 2.2 Hz, 1H), 4.61 (s, 2H), 1.48 (s, 9H).

^{13}C NMR (101 MHz, $CDCl_3$): δ 185.4, 166.4, 163.6, 151.2, 132.3, 122.5, 120.3, 117.1, 114.8, 109.0, 83.6, 66.0, 28.0.

HRMS-ESI$^+$ (*m/z*): $[M+H]^+$ Calc. 385.0569, Obs. 385.0566.

*Synthesis of **3***

2 (1.06 g, 2.76 mmol), $Pd(dppf)Cl_2$ (80.8 mg, 0.11 mmol, 0.04 eq.), B_2pin_2 (1.05 g, 4.14 mmol, 1.5 eq.), and KOAc (811.4 mg, 8.28 mmol, 3 eq.) was taken together in a 50 mL oven-dried round-bottomed flask. 16.5 mL of dry dioxane was added, and the mixture was purged with argon for 20 min. The reaction mixture was then heated at 85 °C for 40 min in a stirring condition. After completion of the reaction, as indicated by TLC, the reaction mixture was filtered through a celite bed. Upon evaporation of the solvent, the residue was purified through flash chromatography using 10% EtOAc/Hexane to obtain 1 g of **3** as a yellowish solid (70% yield).

^{1}H NMR (400 MHz, $CDCl_3$): δ 10.39 (s, 1H), 7.92 (d, *J* = 8.6 Hz, 1H), 7.26 (d, *J* = 2.7 Hz, 1H), 7.02 (dd, *J* = 8.5, 2.6 Hz, 1H), 4.59 (s, 2H), 1.47 (s, 9H), 1.36 (s, 12H).

^{13}C NMR (101 MHz, $CDCl_3$): δ 193.2, 167.3, 161.5, 135.4, 130.4, 120.4, 117.2, 84.5, 83.6, 82.9, 65.6, 28.1, 25.0.

HRMS-ESI^{+} (*m/z*): [M-pin-H_2O+H]$^{+}$ Calc. 263.1085, Obs. 263.1098.

*Synthesis of **4***

1 g of **3** was dissolved in 5 mL cold neat TFA, which formed a yellow-coloured solution. It was then stirred at room temperature for 10 min when the colour changed from yellow to brown to deep blue. TLC analysis indicated the completion of the reaction. Residual TFA was evaporated on a rotary evaporator by diluting with 100 mL DCM, and 594 mg of grey solid was obtained by precipitation using 10% EtOAc/Hexane. The crude product was used for the next step without further purification.

336 mg of the acid derivative (1.5 mmol) and HBTU (568.5 mg, 1.5 mmol, 1 eq.) were dissolved in 1.5 mL DMF and stirred at room temperature. To this solution, DiPEA (385 µL, 2.25 mmol, 1.5 eq.) and allylamine (123 µL, 1.65 mmol, 1.1 eq.) were added slowly with a 5 min interval in a stirring condition. The reaction was completed after 40 minutes, as indicated by the TLC analysis. The reaction mixture was diluted in 200 mL EtOAc, and the organic layer was washed with 75 mL chilled 1N HCl and 100 mL chilled 5% $NaHCO_3$, followed by brine wash. After evaporation of the solvent, the residue was purified through flash chromatography using 40% EtOAc/Hexane to obtain 116 mg (35% yield) as a yellowish gummy solid, which solidified on refrigeration, mp: 118-120 °C.

^{1}H NMR (400 MHz, $CDCl_3$): δ 10.40 (s, 1H), 7.96 (d, *J* = 8.6 Hz, 1H), 7.34 (d, *J* = 2.7 Hz, 1H), 7.05 (dd, *J* = 8.7, 2.7 Hz, 1H), 6.64 (s, 1H), 5.88-5.80 (m, 1H), 5.24 – 5.14 (m, 2H), 4.59 (s, 2H), 3.99-3.96 (m, 2H), 1.38 (s, 12H).

^{13}C NMR (101 MHz, $CDCl_3$): δ 193.0, 167.2, 160.5, 135.9, 133.5, 130.8, 121.0, 117.1, 116.6, 84.7, 67.2, 41.5, 38.8, 25.0.

HRMS-ESI^{+} (*m/z*): [M-pin-H_2O+H]$^{+}$ Calc. 246.0940, Obs. 246.0938.

*Synthesis of **BA4***

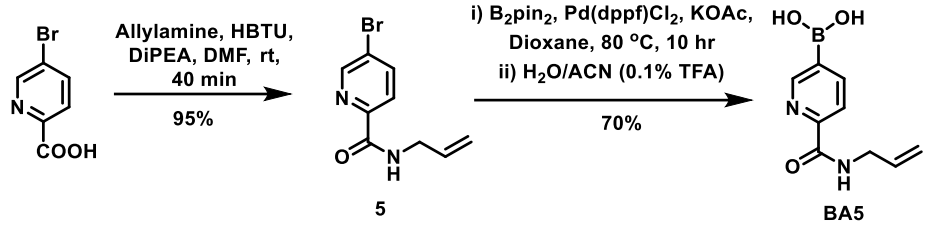
4 (14 mg, 0.05 mmol) and $NaBH_4$ (9.5 mg, 0.25 mmol, 5 eq.) were taken in 1 mL 1:1 MeOH:H_2O and stirred at room temperature for 15 min. MeOH was evaporated, and the aqueous solution was acidified with 20 mL 1N HCl and washed with 2×50 mL EtOAc. The organic layer was further washed with brine and evaporated to obtain the benzoxaborole **BA4** as a white solid with a 65% yield (8 mg), mp: 148-150 °C. We believe there was some mechanical loss of the product during the workup.

^{1}H NMR (400 MHz, $CDCl_3$): δ 7.23 (d, *J* = 8 Hz, 1H), 7.16 (d, *J* = 2.0 Hz, 1H), 7.02 (dd, *J* = 8.3, 2.6 Hz, 1H), 6.83 (s, 1H), 5.86-5.76 (m, 1H), 5.19 – 5.09 (m, 2H), 5.00 (s, 2H), 4.49 (s, 2H), 3.94-3.92 (m, 2H).

^{13}C NMR (101 MHz, $CDCl_3$): δ 156.7, 147.3, 133.5, 122.5, 118.9, 116.8, 114.7, 83.0, 71.0, 67.5, 41.3.

HRMS-ESI^{+} (*m/z*): [M+H]$^{+}$ Calc. 248.1094, Obs. 248.1096.

3A.4.4.4. *SYNTHESIS OF **BA5***



*Synthesis of **5***

5-Bromopicolinic acid (505 mg, 2.5 mmol) and HBTU (950 mg, 2.5 mmol, 1 eq.) were dissolved in 2 mL DMF. Then, DiPEA (0.64 mL, 3.75 mmol, 1.5 eq.) and allylamine (0.2 mL, 2.75 mmol, 1.1 eq.) were added at 5 min intervals to the reaction mixture. TLC analysis confirmed the completion of the reaction after 35 min of stirring. 100 mL EtOAc was added to the reaction mixture, and the organic layer was washed with 50 mL chilled 1 N HCl and 50 mL chilled 5% $NaHCO_3$, followed by brine wash. After evaporation of the solvent, the residue was purified through silica gel using 10% EtOAc/Hexane to obtain 570 mg of **7** (95% yield) as a yellow oil.

^{1}H NMR (400 MHz, $CDCl_3$): δ 8.57 (dd, *J* = 2.3, 0.7 Hz, 1H), 8.07 (dd, *J* = 8.3, 0.7 Hz, 1H), 7.95 (dd, *J* = 8.3, 2.3 Hz, 1H), 5.91 (ddt, *J* = 17.1, 10.3, 5.6 Hz, 1H), 5.24 (ddd, *J* = 17.1, 3.0, 1.7 Hz, 1H), 5.16 (ddd, *J* = 10.2, 2.8, 1.4 Hz, 1H), 4.10 – 4.04 (m, 3H).

^{13}C NMR (101 MHz, $CDCl_3$): δ 163.5, 149.4, 148.4, 140.1, 133.9, 124.1, 123.8, 116.7, 41.9.

HRMS-ESI$^+$ (*m/z*): $[M+H]^+$ Calc. 240.9977, Obs. 240.9977.

*Synthesis of **BA5***

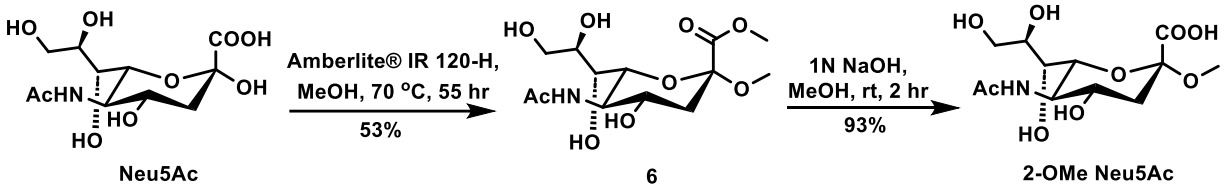
5 (360 mg, 1.5 mmol), $Pd(dppf)Cl_2$.DCM (61.2 mg, 0.075 mmol, 0.05 eq.), B_2pin_2 (952 mg, 3.75 mmol, 2.5 eq.), and KOAc (441 mg, 4.5 mmol, 3 eq.) were added together in a 25 mL oven-dried round-bottomed flask. 9 mL of dry dioxane was added, and the suspension was purged with argon for 20 min. The reaction mixture was then heated at 80 °C for 10 hr in stirring conditions. The reaction was monitored by HPLC, and the product formation was confirmed by LCMS. TLC analysis indicated consumption of **5**. The reaction mixture was then filtered through a pad of celite. Upon evaporation of the solvent, 100 mg of the residue was purified through HPLC to obtain 65 mg of **BA5** (TFA salt) as a yellow powder (70% yield) after lyophilization, mp: 112-114 °C. Method: Mobile phase acetonitrile-water (0.05% TFA) with a gradient of 15-95% acetonitrile over 30 min. We observed the removal of the pinacol group from **BA5** due to the use of the acid mobile phase in HPLC purification.

^{1}H NMR (400 MHz, 10% $CD_3OD/CDCl_3$): δ 8.81 (s, 1H), 8.19 (d, *J* = 7.7 Hz, 1H), 8.09 (d, *J* = 7.8 Hz, 1H), 5.94-5.85 (m, 1H), 5.26-5.21 (m, 1H), 5.16 – 5.14 (m, 1H), 4.05 (d, *J* = 5.5 Hz, 3H).

^{13}C NMR (101 MHz, 10% $CD_3OD/CDCl_3$): δ 164.6, 153.0, 143.3, 133.8, 121.5, 116.7, 84.7, 41.8.

HRMS-ESI$^+$ (*m/z*): $[M+H]^+$ Calc. 207.0941, Obs. 207.0939.

3A.4.4.5. *SYNTHESIS OF 2-OMe Neu5Ac (Sia)*



It was synthesized according to previously reported literature[33].

*Synthesis of **6***

2 g N-acetylneuraminic acid (Neu5Ac, 6.5 mmol) was added to 200 mL methanol. To the resulting suspension, 8 g Amberlite® IR 120-H was added, and the mixture refluxed at 70 ºC for 55 hr. Reaction completion was confirmed by TLC (Mobile phase EtOAc:MeOH:H_2O = 3:1:1, R_f = 0.5, $KMnO_4$ charring), and the solution was filtered to remove the solid resin. The solvent was evaporated to obtain a residue, which was further subjected to column purification using 100-200 silica gel packed in 1% MeOH/EtOAc. The gradient was increased to 15% to elute the hydrophobic impurities. The product was eluted with 15% MeOH/EtOAc with 2% H_2O and evaporated to obtain 1.16 g of white powder (53% yield). The product characterizations match the previous report.

^{1}H NMR (400 MHz, DMSO-d_6): δ 8.06 (s, 1H), 4.85 – 4.27 (m, 4H), 3.64 (s, 3H), 3.62 – 3.44 (m, 5H), 3.19 (d, *J* = 13.4 Hz, 2H), 3.11 (s, 3H), 2.11 (dd, *J* = 12.8, 4.8 Hz, 1H), 1.83 (s, 3H), 1.50 – 1.39 (m, 1H).

HRMS-ESI$^-$ (*m/z*): $[M-H]^-$ Calc. 336.1295, Obs. 336.1295.

*Synthesis of **Sia***

1 g of **6** (3 mmol) was dissolved in 125 mL methanol and stirred at room temperature. 19 mL of 1N NaOH was added to the solution and further stirred for 2 hr when TLC analysis (mobile phase EtOAc:MeOH:H_2O = 3:1:1, R_f = 0.3, $KMnO_4$ charring) indicated completion of the reaction. The reaction mixture was acidified with Amberlite® IR 120-H to avoid the formation of undesired mineral salts upon solvent evaporation. After filtration of the resin, the solvent was evaporated to obtain 0.9 g of yellowish solid (93%). The product characterizations match the previous report.

^{1}H NMR (400 MHz, DMSO-d_6): δ 8.10 – 7.89 (m, 1H), 3.84 – 3.69 (m, 2H), 3.65 – 3.45 (m, 5H), 3.38 – 3.26 (m, 1H), 3.22 (d, *J* = 9.6 Hz, 1H), 3.12 (s, 3H), 2.10 (dd, *J* = 12.9, 5.0 Hz, 1H), 1.83 (s, 3H), 1.43 (dd, *J* = 12.8, 11.3 Hz, 1H).

HRMS-ESI$^-$ (*m/z*): $[M-H]^-$ Calc. 322.1138, Obs. 322.1138.

3A.4.5 *SYNTHESIS OF PEPTIDES*

All the peptides were synthesized by standard Fmoc-SPPS on Rink amide polystyrene AM resin. Three equivalents of the commercially available amino acids, HBTU (1 eq.) as a coupling agent and DiPEA (1.5 eq.) as a base, were used for the coupling reaction for 15 min. Fmoc deprotection was achieved using 3 mL 20% piperidine/DMF twice for 3 min each, followed by washing with 3 mL DMF six times. The peptides

were cleaved off the resin and globally deprotected with reagent K (82.5% TFA, 5% H_2O, 2.5% EDT, 5% Thioanisole and 5% phenol). Precipitation with chilled diethyl ether gave the crude peptides, which were purified by RP-HPLC in ACN/H_2O (0.1% AcOH) by dissolving and injecting 20 mg at a time. The purity of all peptides was determined by using LC-MS (Waters) and HPLC (Shimadzu) to be >94%. The characterization data are presented in Table 3A.1.

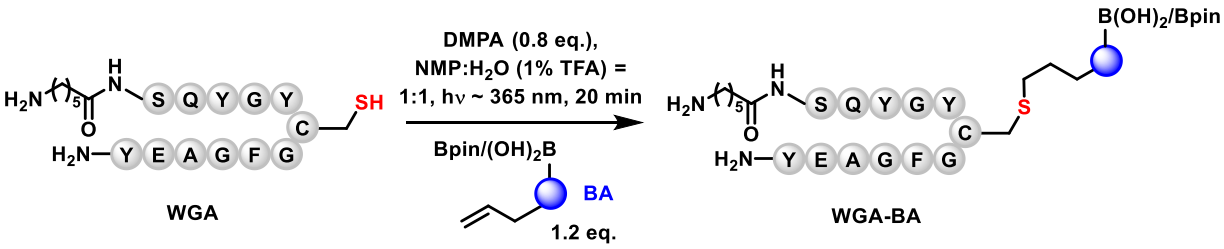
Table 3A.1. List of synthesized biologically active peptides. Acp refers to 6-Aminocaproic acid.

Sl. No.	Name	Sequence	Length	Calc. *m/z*	Obs. *m/z*
1.	Wheat Germ Agglutinin, 62-73 (WGA)	AcpSQYGYCGFGAEY-NH_2	13-mer	$[M+H]^+$ 1456.62	$[M+H]^+$ 1456.62
2.	Scrambled WGA (sWGA)	AcpSYYGQCGEGAFY-NH_2	13-mer	$[M+H]^+$ 1456.62	$[M+H]^+$ 1456.59

3A.4.6 *INSTALLATION OF BA DERIVATIVES ON PEPTIDES BY THIOL-ENE*

A particular peptide (10 mM) was reacted with 12 mM BA derivatives and 8 mM DMPA for 20 min in NMP: H_2O (1% TFA) = 1:1 as the solvent and hν~365 nm light source. The HPLC trace and ESI-MS data for the reactions are provided.

3A.5.6.1 *REACTIONS WITH WGA*



WGA-BA1

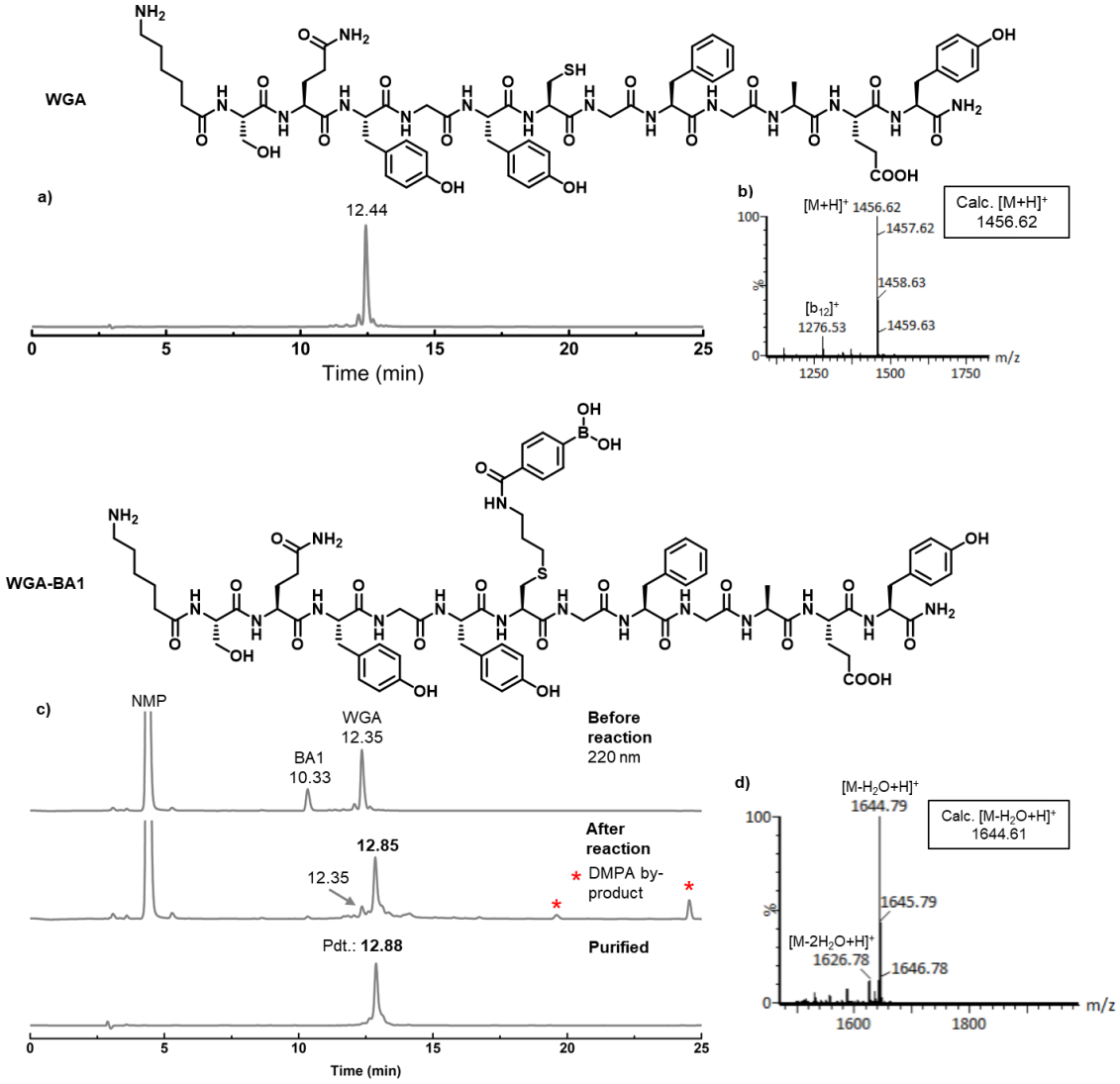


Figure 3A.7: WGA data: a) HPLC trace at 220 nm (92% purity); b) Mass data; WGA-BA1 data: c) HPLC trace at 220 nm, 86.4% conversion; d) Mass data. 55% yield.

WGA-BA2

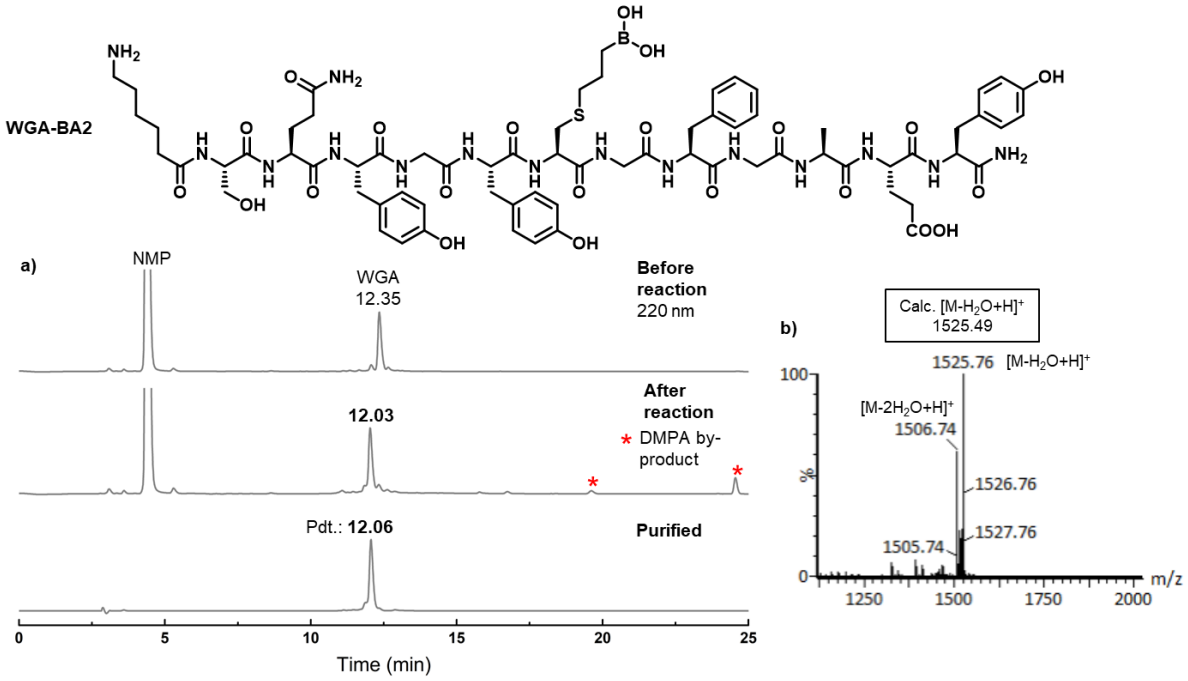


Figure 3A.8: WGA-BA2 data: a) HPLC trace at 220 nm, >95% conversion; b) Mass data. 60% yield.

WGA-BA3

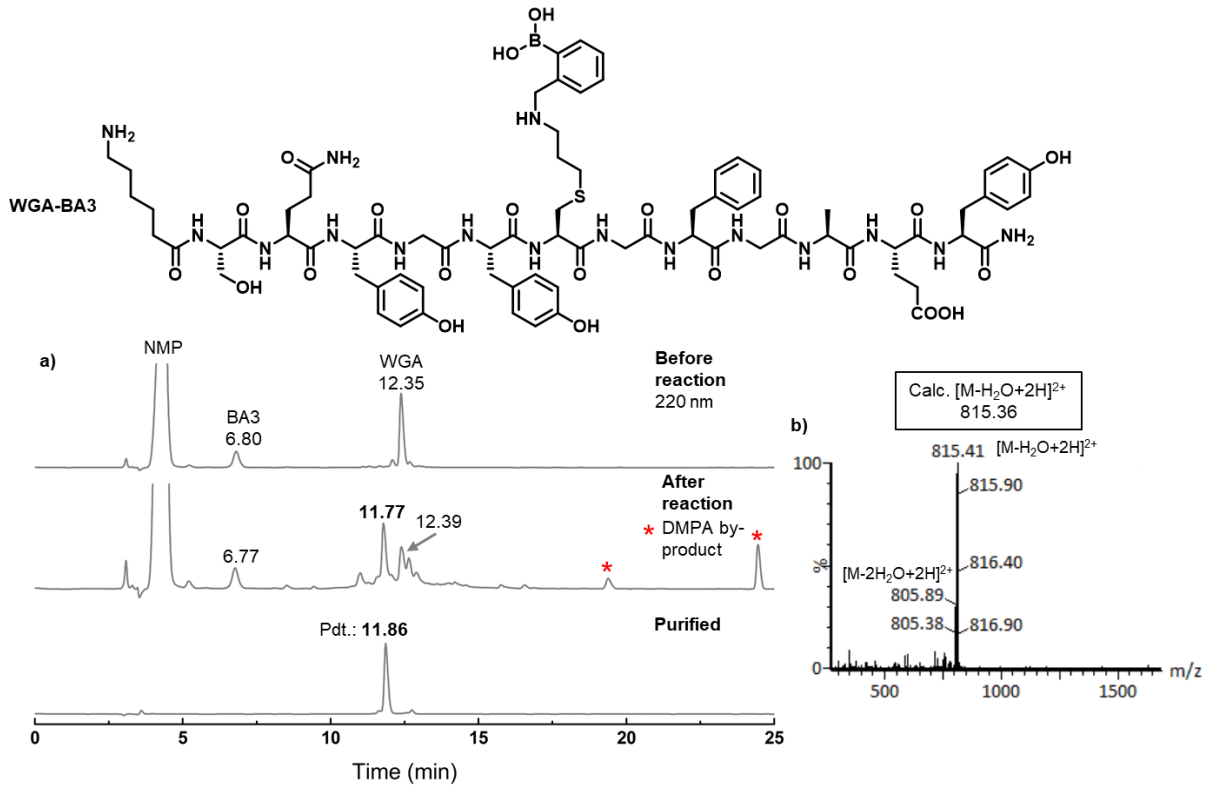


Figure 3A.9: WGA-BA3 data: a) HPLC trace at 220 nm, 77.6% conversion; b) Mass data. 50% yield. A loss of at least two neutral molecules of water is a commonly observed phenomenon in the mass spectra of peptides containing Wulff-type boronates[34] (although it can also be observed in other boronopeptides[24]).

WGA-BA4

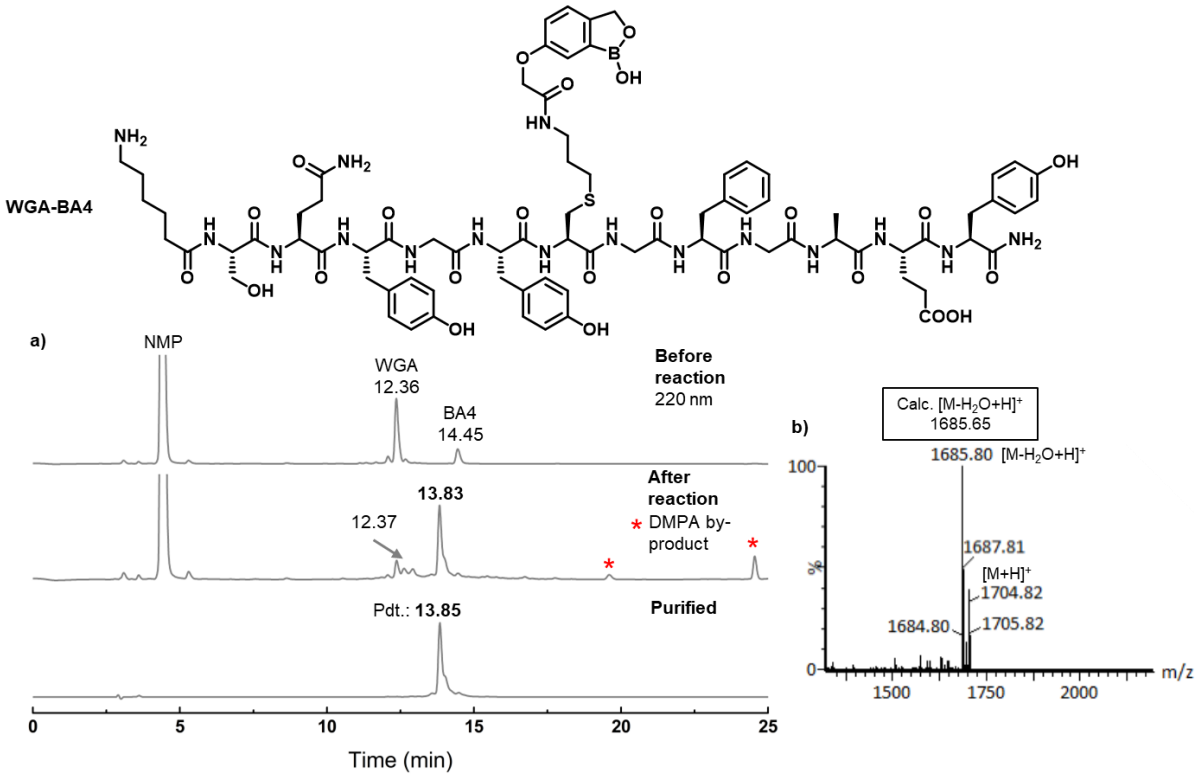


Figure 3A.10: WGA-BA4 data: a) HPLC trace at 220 nm, 86% conversion; b) Mass data. 55% yield.

WGA-BA5

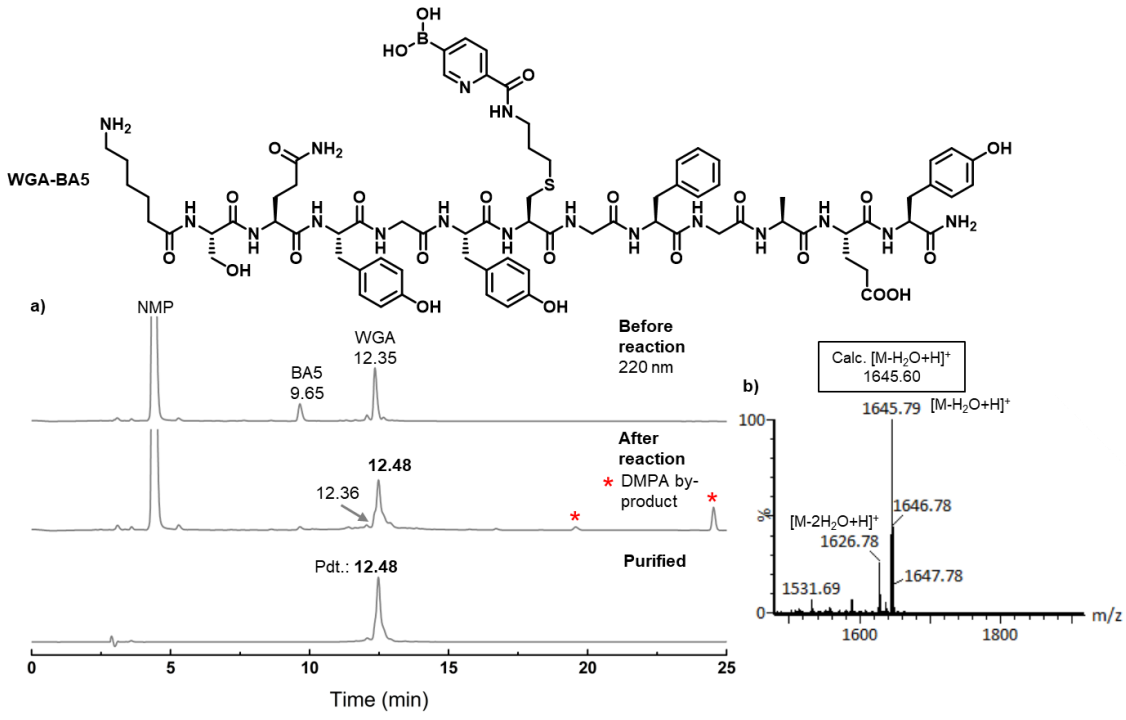


Figure 3A.11: WGA-BA5 data: a) HPLC trace at 220 nm, 88% conversion; b) Mass data. 57% yield.

3A.4.6.1 *REACTIONS WITH SCRAMBLED WGA (sWGA)*

Post deprotection of the N-terminal Fmoc group, the peptides were labelled with 2 eq. FITC and 2 eq. DipEA as the base. They were further purified after cleavage by RP-HPLC to obtain the pure peptide. Stock solutions were prepared in DMSO (MB grade) and their concentrations determine by UV-Vis (ε_{FITC} = 73,000 M^{-1} s^{-1}).

Synthesis of scrambled WGA-FITC (P6)

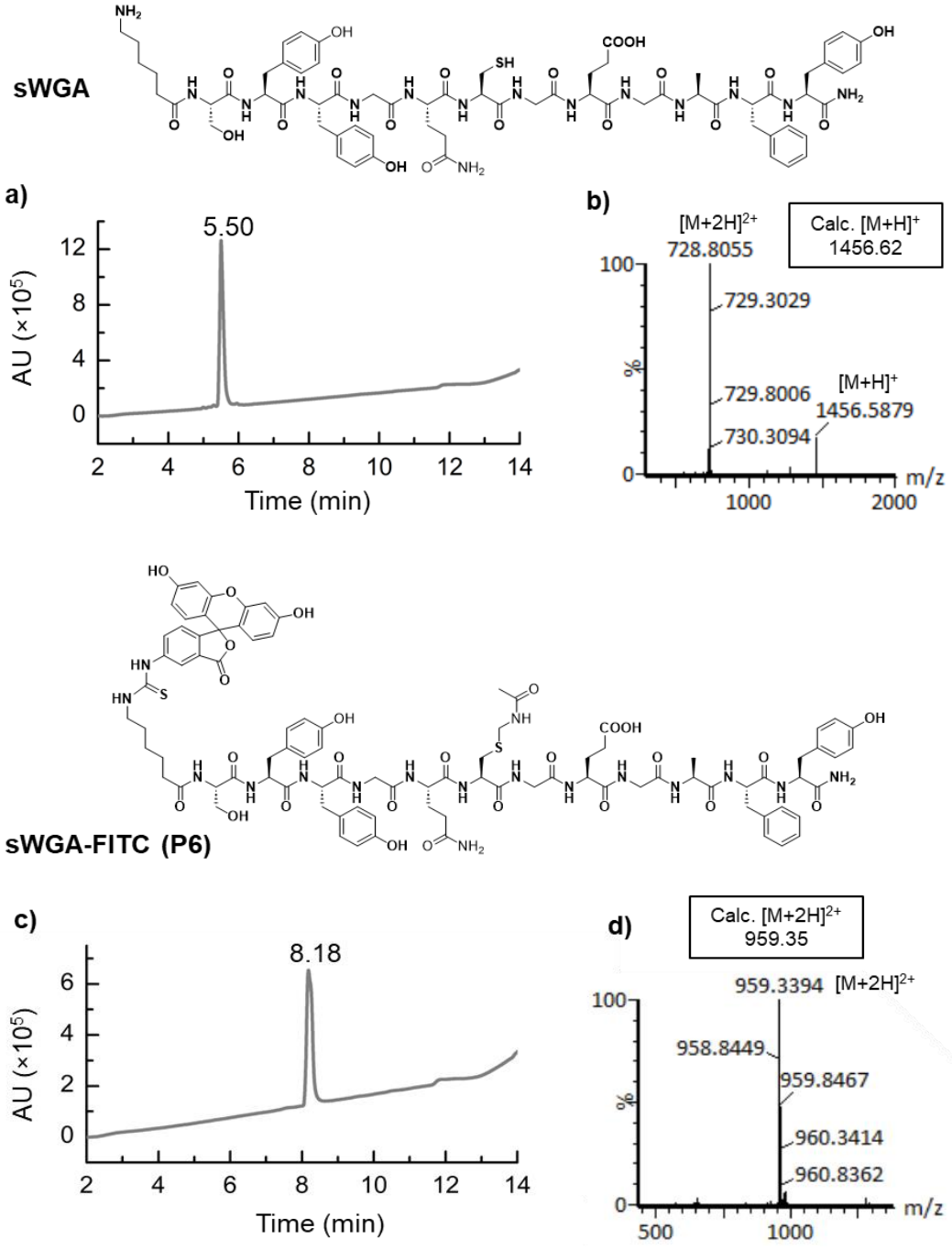


Figure 3A.12: Scrambled WGA data: a) LC trace at 220 nm, 92% conversion; b) Mass data. 35% yield; sWGA-FITC data: c) LC trace at 220 nm, 94% conversion; d) Mass data. 30% yield.

Synthesis of scrambled WGA-BA4-FITC (P7)

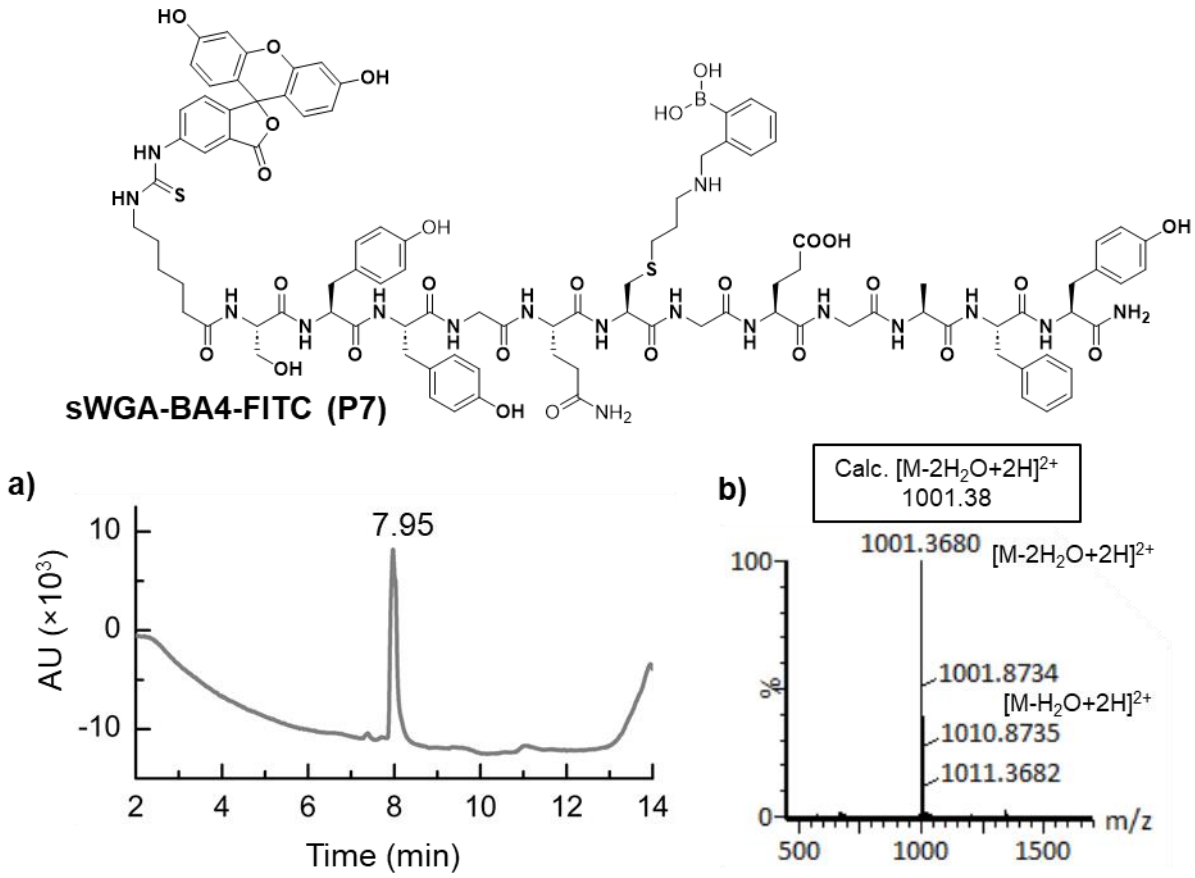


Figure 3A.13: sWGA-BA4-FITC (P7) data: c) LC trace at 220 nm, 60% conversion; d) Mass data. 65% yield.

3A.4.7. *MS/MS ANALYSIS*

On comparing the fragmentation pattern of WGA and WGA-BA3, it can be observed that b7-b12 fragments of WGA appear as singly charged species while the b7-b12 fragments of WGA-BA3 appear with the loss of two water molecules. Such behaviour can be typically observed for Wulff-type boronic acids. This observation unambiguously proves the installation of BA3 on the Cys residue.

a) WGA Precursor ion $[M+H]^+$: 1456.62

Sequence	Ions	
	Calc. *m/z* (*z*=1)	Obs. *m/z* (*z*=1)
AcpS	b_2 201.1239	b_2 201.0873
AcpSQY	b_4 492.2458	b_4 492.2460
AcpSQYGY	b_6 712.3306	b_6 712.3297
AcpSQYGYC	b_7 815.3398	b_7 815.3415
AcpSQYGYCG	b_8 872.3613	b_8 872.3616
AcpSQYGYCGF	b_9 1019.4297	b_9 1019.4310
AcpSQYGYCGFG	b_{10} 1076.4511	b_{10} 1076.4514
AcpSQYGYCGFGA	b_{11} 1147.4883	b_{11} 1147.4897
AcpSQYGYCGFGAE	b_{12} 1276.5308	b_{12} 1276.5328
Y	y_1 181.0972	y_1 181.0973

b) WGA-BA3 Precursor ion $[M+H]^+$: 1647.73

Sequence	Ions	
	Calc. *m/z* (*z*=1)	Obs. *m/z* (*z*=1)
AcpS	b_2 201.1239	b_2 201.0873
AcpSQY	b_4 492.2458	b_4 492.2443
AcpSQYGYC(BA3)	b_7-$2H_2O$ 970.4516	b_7-$2H_2O$ 970.4333
AcpSQYGYC(BA3)G	b_8-$2H_2O$ 1027.4730	b_8-$2H_2O$ 1027.4543
AcpSQYGYC(BA3)GF	b_9-NH_3 1194.5414	b_9-NH_3 1194.5490
AcpSQYGYC(BA3)GFG	b_{10}-$2H_2O$ 1231.5629	b_{10}-$2H_2O$ 1231.5437
AcpSQYGYC(BA3)GFGAE	b_{12}-$2H_2O$ 1431.6426	b_{12}-$2H_2O$ 1431.6239
Y	y_1 181.0972	y_1 181.0973
EY	y_2-NH_3 292.1397	y_2-NH_3 292.1294

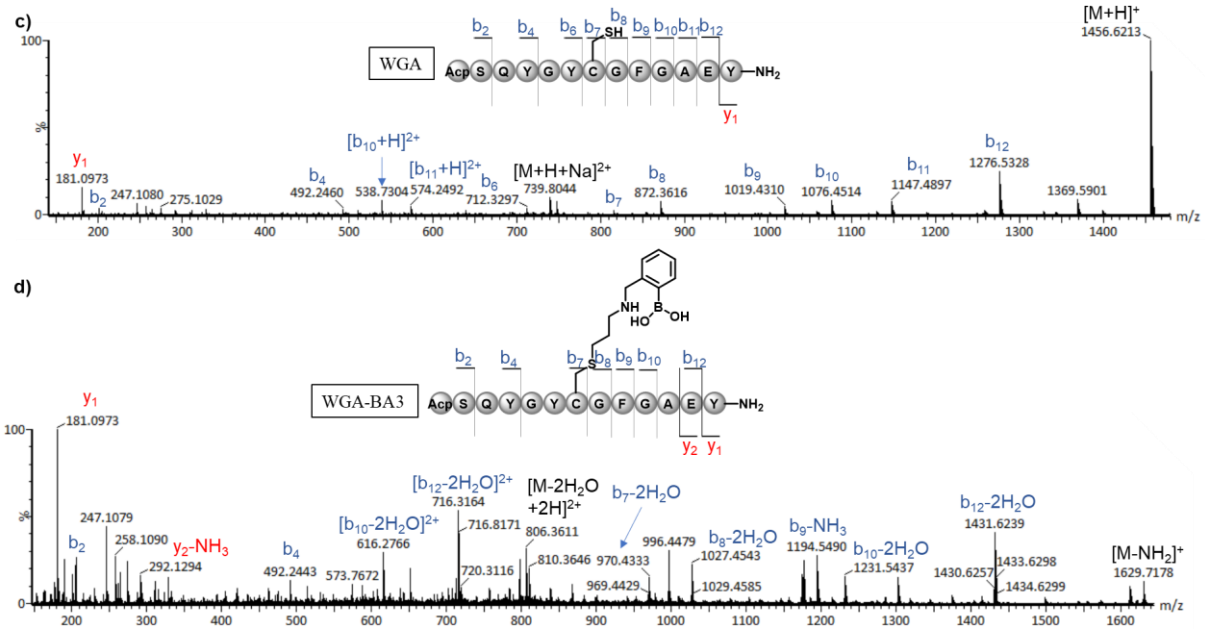


Figure 3A.14: MS/MS pattern tabulated for a) WGA, b) WGA-BA3; MS/MS spectra for c) WGA, d) WGA-BA3.

3A.4.8. *COMPUTATIONAL DESIGN OF BORONO-WGA PEPTIDE WITH sTn*

3A.4.8.1 *Protein Structure Preparation*

The dodecamer WGA peptide (62-73) was derived from wheat germ agglutinin protein (Wright C.S., 1990, PDB ID: 1WGC)[21,35] and was prepared for GLIDE/XP (Extra Precision) protocol using Maestro (Schrödinger, LLC, New York, NY, 2021) protein preparation wizard. This includes removing crystal water molecules, assigning protonation state, and minimization using the Optimized Potential Liquid Simulation 3e (OPLS3e) force field. All the residues forming Hydrogen bonds within a distance of 3.5 Å with dodecamer WGA peptide were considered, including Ser62, Gln63, Tyr64, Tyr66, Phe69, Ala71, and Tyr73. These active site residues were used for receptor grid generation.

Desmond, an MD simulation software, was employed to model the dodecamer WGA peptide in an aqueous environment. The model constitutes an orthorhombic box of water with a minimum size to contain the model, ensuring a 10 Å distance from the edge of the box. TIP3P was used as the solvent model, and 0.15 M NaCl was added to neutralize it. The system was minimized to 1000 steps, holding all the peptide atoms. The minimized system was subjected to MD simulations using the NPT ensemble and periodic boundary conditions for 100 ns. The Martyna-Tobias-Klein method was used to control the pressure, which allows the pressure of the system to be at 1.01 bar by using the isotropic coupling method. The Nose-Hoover thermostat was applied to control the temperature at 300 K. The trajectories and other parameters were saved every 100 and 1.2 ps, respectively, to return 1000 frames.

3A.4.8.1 *Ligand Structure Preparation*

An examined sTn (sialyl-Tn antigen) was built and optimized with ChemDraw Ultra 2010 (PerkinElmer, Inc) using MMFF94x force field until a minimum root-mean-square gradient was reached. A conformational search for sTn ligand preparation was carried out using the Maestro LigPrep method. The molecules were ionized at a pH of 7.0 by Epik. The number of sTn ligands, including tautomers, rises to around 141. The library was used after minimization by OPLS3e without further modifications.

3A.4.8.1 *Molecular Docking*

We performed molecular docking to study binding interaction using the Glide/XP (eXtra Precision) docking protocol in the Schrodinger suite. The protein preparation wizard module prepared the simulated 12-mer WGA peptide and then docked it with the sTn ligand prepared using the LigPrep module in the Schrodinger suite. The lowest observed GLIDE/XP docking score for the WGA-sTn complex was -3.894 kJ/mol (Gscore,XP = -3.894). The key residues involved Tyr and Ser with sTn interactions were observed. For visualization and analysis of the molecules, the PyMOL tool was used.

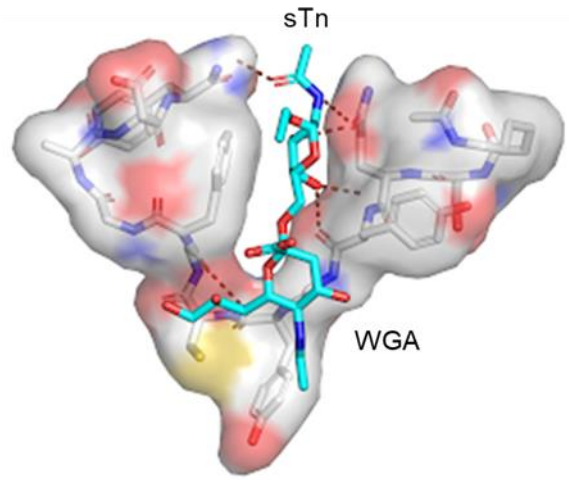


Figure 3A.15: A structural overview of WGA-sTn complex showing hydrogen bonds of SA-OH with Tyr-OH and hydrophobic interaction of C-backbone of SA with key residues.

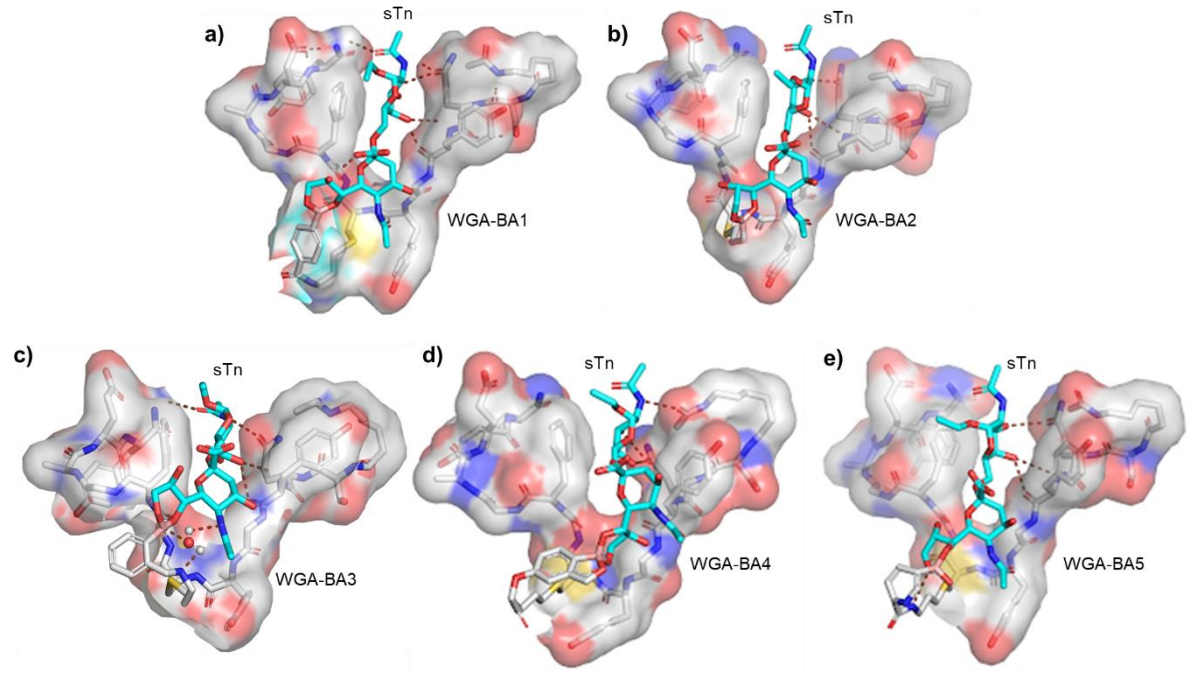
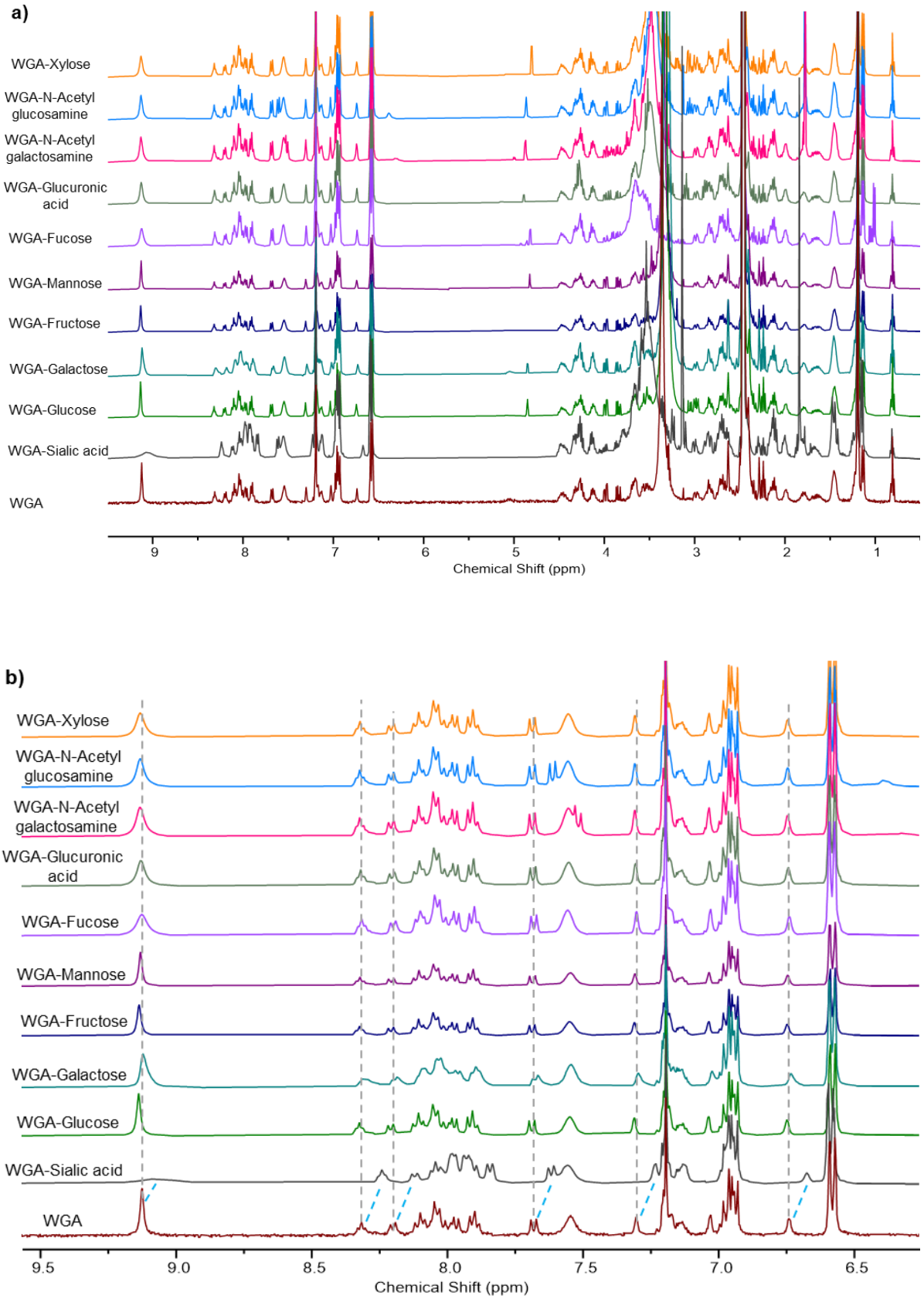


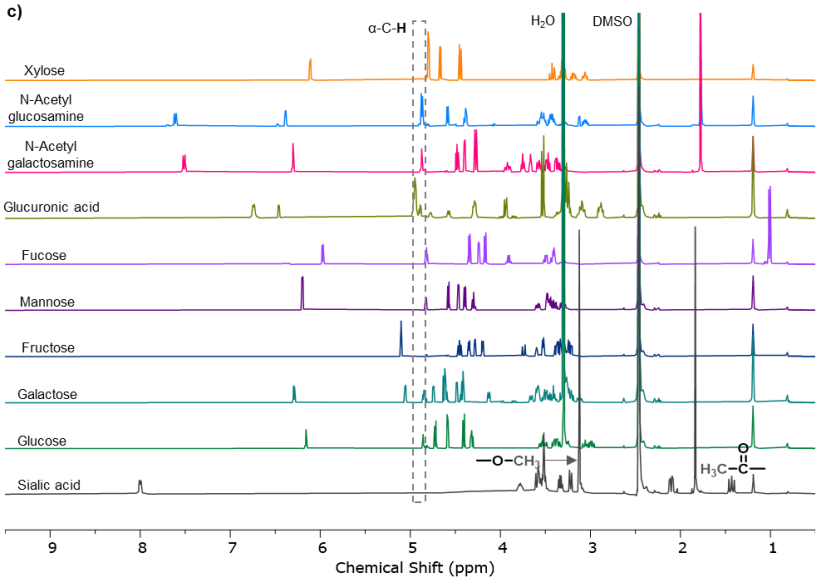
Figure 3A.16: A structural overview of borono-WGA peptide interactions with sTn, a) WGA-BA1 interaction with sTn; b) WGA-BA2 interaction with sTn; c) WGA-BA3 interaction with sTn; d) WGA-BA4 interaction with sTn; e) WGA-BA5 interaction with sTn.

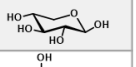
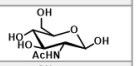
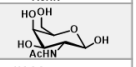
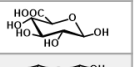
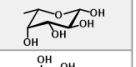
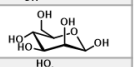
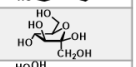
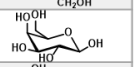
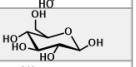
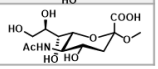
3A.4.9. *BINDING SELECTIVITY OF WGA PEPTIDE WITH ABUNDANT SUGARS*

Clear evidence for the selective binding of the WGA peptide sequence (SQYGYCGFGAEY-NH_2) to sialic acid was not found in the literature, where a pool of abundant sugar has been considered for comparison. The given status has inspired us to examine the selectivity of WGA peptide to sialic acid through ^{1}H NMR. We evaluated ten biologically abundant sugars (Figure 3A.17c) to understand the preferential binding study to WGA peptide at a concentration (~2 mM) well above the reported K_d to sialic acid (~ 0.79 mM). Nearly a 1:1 ratio of peptide and a particular sugar were dissolved in DMSO-d_6 and subjected to 400 MHz NMR with 1024 scans. We observed a significant change in WGA amide proton when incubated with sialic acid. Stacked NMR spectra are shown in Figure 3A.17a. The peptide has shown solubility issues in MeOH-d_4 and even in water:DMSO-d_6 (1:1). Therefore, we have performed binding selectivity studies in DMSO-d_6. WGA peptide shows a great hydrogel property: a few pictures are shown below (Figure 3A.17d). To see the change in proton NMR of Sia incubated with WGA-BA3, we performed an NMR study also at 2 mM in DMSO-d_6. Similarly, stacked NMR spectra of WGA-BA3 with different sugars have also been presented in Figure 3A.17e. To further corroborate the binding of BA moiety with *cis*-diol of Sia, we performed ^{11}B NMR studies by incubating 1 mM each of WGA-BA3 and Sia in DMSO-d_6. 12011 scans were required to obtain a clear spectrum, as shown. It can be clearly seen that WGA-BA3 shows two peaks at ~30 ppm, a signature of trigonal BA, and at ~12 ppm, a signature of solvent-inserted or N-coordinated BA. Upon tight boronate ester formation, a single peak is observed at ~10 ppm.

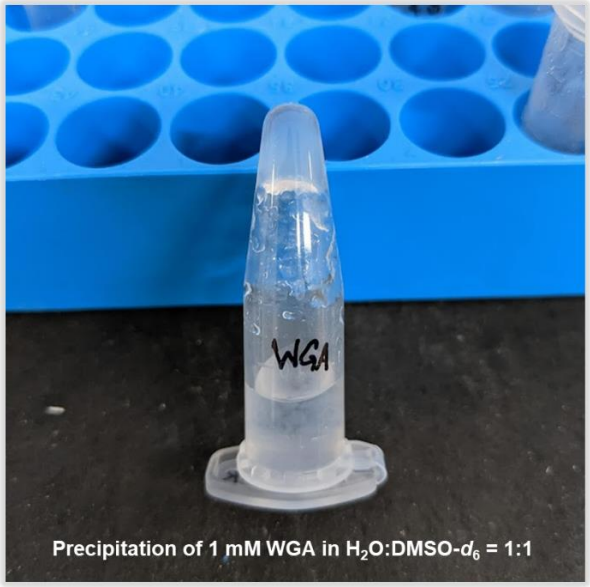

a)
WGA-Xylose
WGA-N-Acetyl glucosamine
WGA-N-Acetyl galactosamine
WGA-Glucuronic acid
WGA-Fucose
WGA-Mannose
WGA-Fructose
WGA-Galactose
WGA-Glucose
WGA-Sialic acid
WGA
9
8
7
6
5
4
3
2
1
Chemical Shift (ppm)
b)
WGA-Xylose
WGA-N-Acetyl glucosamine
WGA-N-Acetyl galactosamine
WGA-Glucuronic acid
WGA-Fucose
WGA-Mannose
WGA-Fructose
WGA-Galactose
WGA-Glucose
WGA-Sialic acid
WGA
9.5
9.0
8.5
8.0
7.5
7.0
6.5
Chemical Shift (ppm)

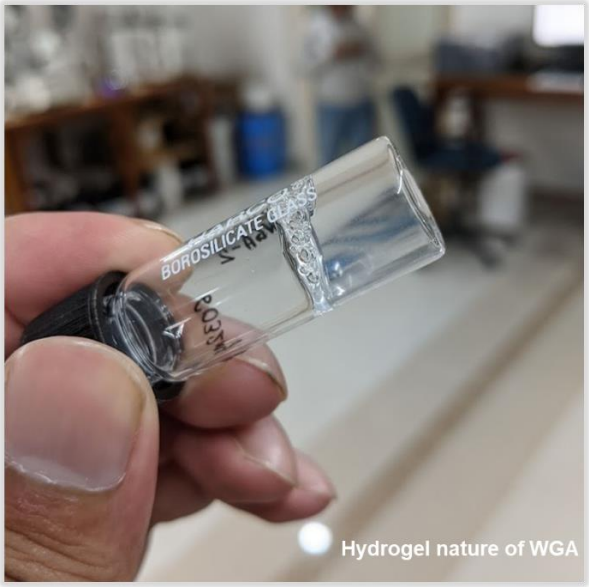
c)



Sugar	Abbrv.	Structure
Xylose	Xyl	
N-Acetyl glucosamine	GluNAc	
N-Acetyl galactosamine	GalNAc	
Glucuronic acid	GlcA	
Fucose	Fuc	
Mannose	Man	
Fructose	Fru	
Galactose	Gal	
Glucose	Glu	
Sialic acid	Sia	

d)





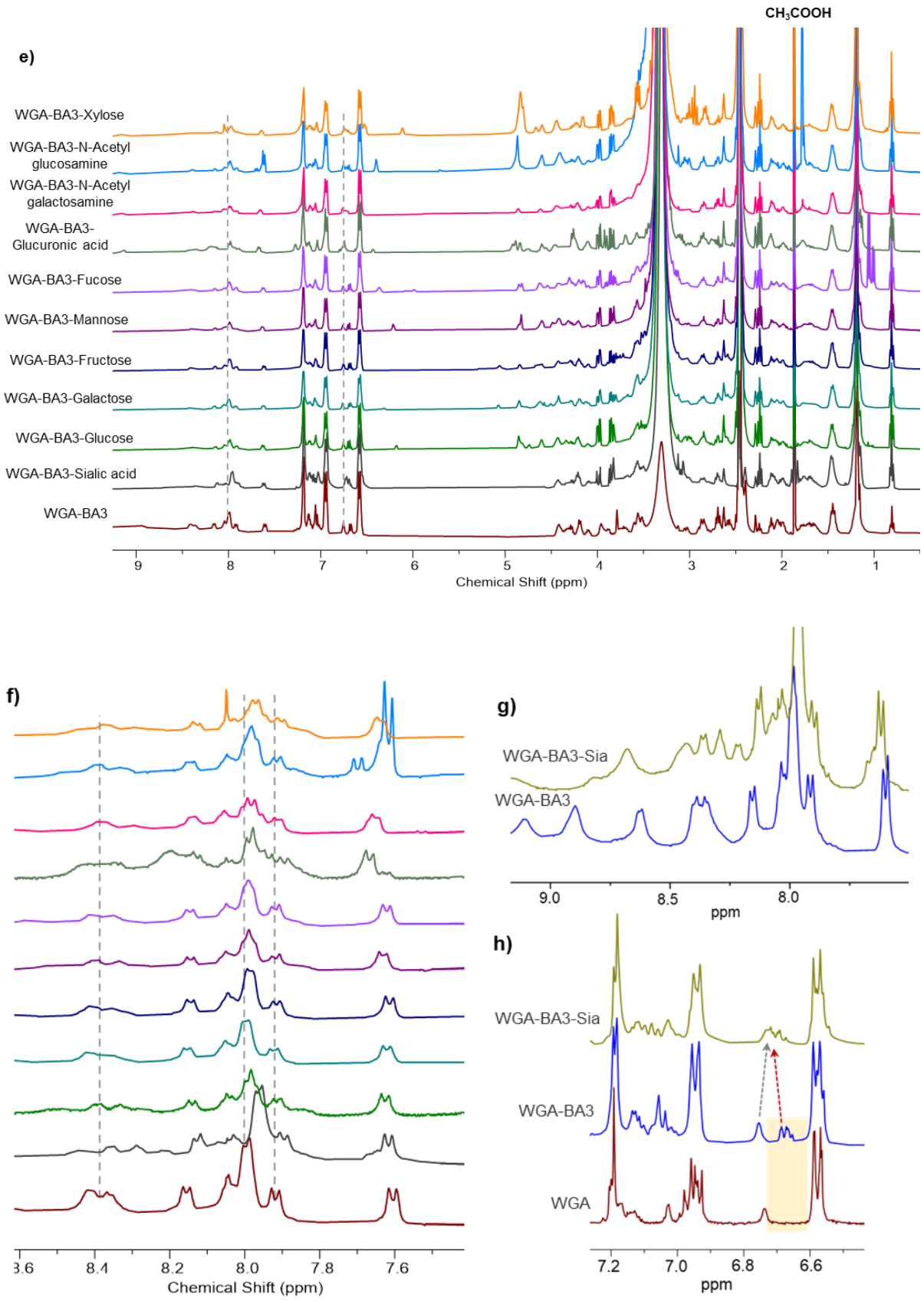


Figure 3A.17: a) Stacked ^{1}H NMR spectra of WGA and different sugars incubated with WGA in DMSO-d_6 and b) expanded spectra; c) Stacked ^{1}H NMR spectra of sugars used for study in DMSO-d_6; d) Pictorial images denoting insolubility of 1 mM WGA in H_2O:DMSO-d_6 = 1:1 and hydrogel nature; e) Stacked ^{1}H-NMR spectra of 1 mM WGA-BA3 with different sugars incubated in DMSO-d_6; f) Expanded aromatic

region showing changes in the spectra; g) ^{1}H-NMR data demonstrate that the BA3-modified WGA peptide sequence retains the binding ability with Sia. h) Stacking spectra of WGA, WGA-BA3, and WGA-BA3-Sia complex and highlighting the clear downfield shift (marron arrow) of the prominent peak (absent in WGA shown in gold box) associated with BA3 moiety.

10 mM of both BA3 and Sia was incubated in 30% D_2O/PBS pH 7.4 for recording ^{1}H and ^{11}B NMR.

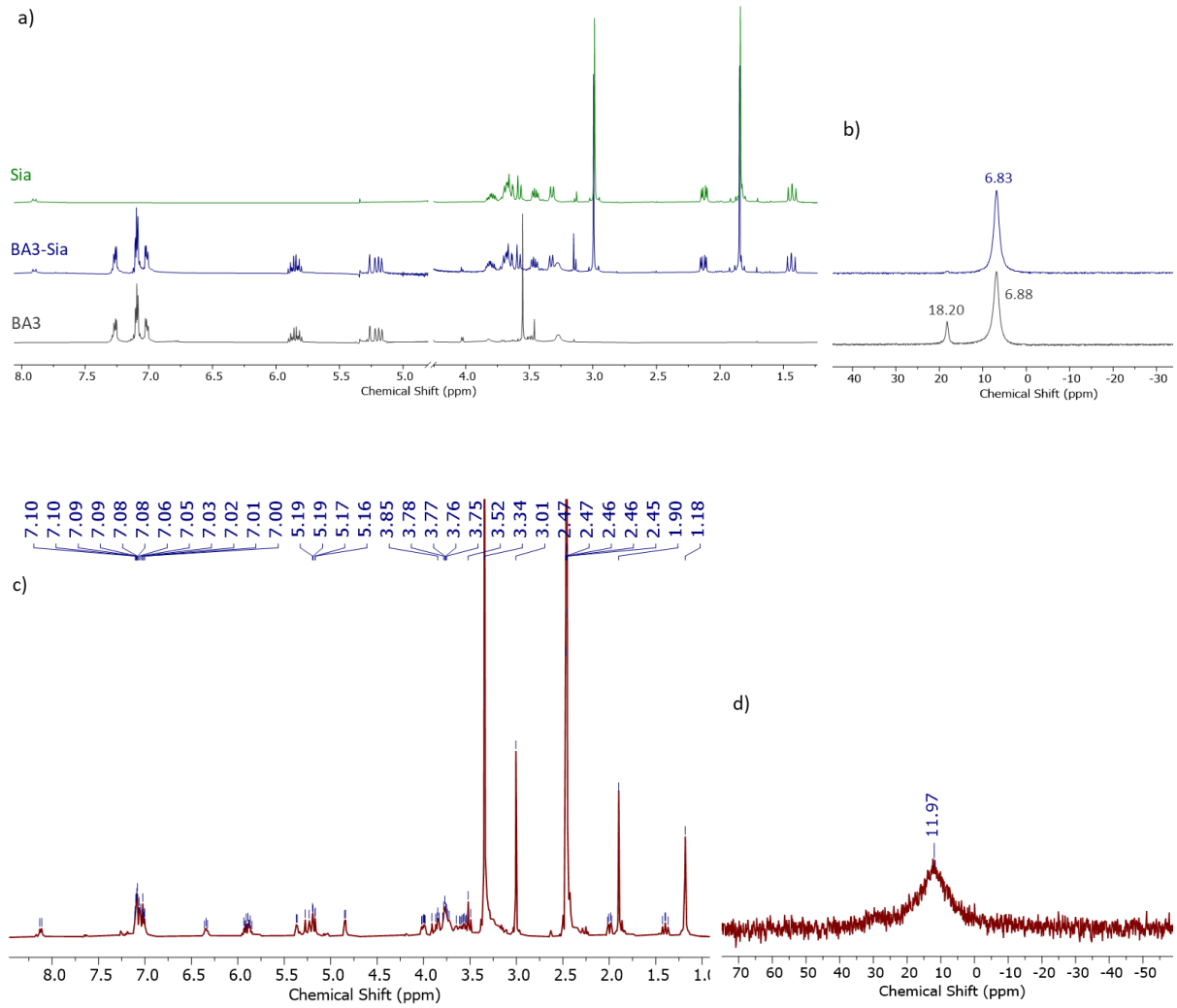


Figure 3A.18: a) Stacked ^{1}H NMR spectra of Sia, BA3-Sia, and BA3 in D_2O/H_2O; b) Stacked ^{11}B NMR spectra of BA3-Sia and BA3 in D_2O/H_2O; c) ^{1}H and d) ^{11}B NMR spectra of BA3-Sia in DMSO-d_6.

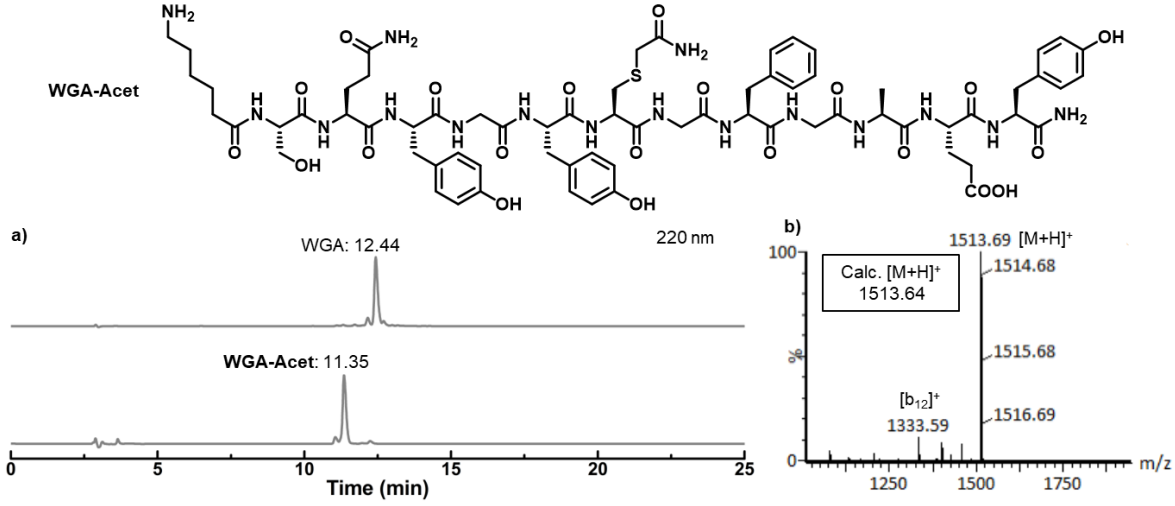
3A.4.10. *FITC LABELLING OF PEPTIDES*

The characterization data of synthesized peptides are presented in Table 3A.2.

Table 3A.2: FITC labelled peptides and their mass spec data.

Sl. No.	Labelled peptide	Dubbed	Calc. mass	Obs. mass
1.	WGA-FITC	P0	$[M+2H]^{2+}$ 952.34	$[M+2H]^{2+}$ 952.37
2.	WGA-BA1-FITC	P1	$[M-2H_2O+H]^{2+}$ 1008.49	$[M-2H_2O+H]^{2+}$ 1008.38
3.	WGA-BA2-FITC	P2	$[M-2H_2O+H]^{2+}$ 948.93	$[M-2H_2O+H]^{2+}$ 948.37
4.	WGA-BA3-FITC	P3	$[M-H_2O+2H]^{2+}$ 1010.50	$[M-H_2O+2H]^{2+}$ 1010.34
5.	WGA-BA4-FITC	P4	$[M-H_2O+2H]^{2+}$ 1038.50	$[M-H_2O+2H]^{2+}$ 1038.40
6.	WGA-BA5-FITC	P5	$[M-2H_2O+H]^{2+}$ 1008.98	$[M-2H_2O+H]^{2+}$ 1008.89
7.	sWGA-FITC	P6	$[M+2H]^{2+}$ 959.35	$[M+2H]^{2+}$ 959.354
8.	sWGA-BA3-FITC	P7	$[M-2H_2O+2H]^{2+}$ 1001.38	$[M-2H_2O+2H]^{2+}$ 1001.37

WGA-FITC (P0)



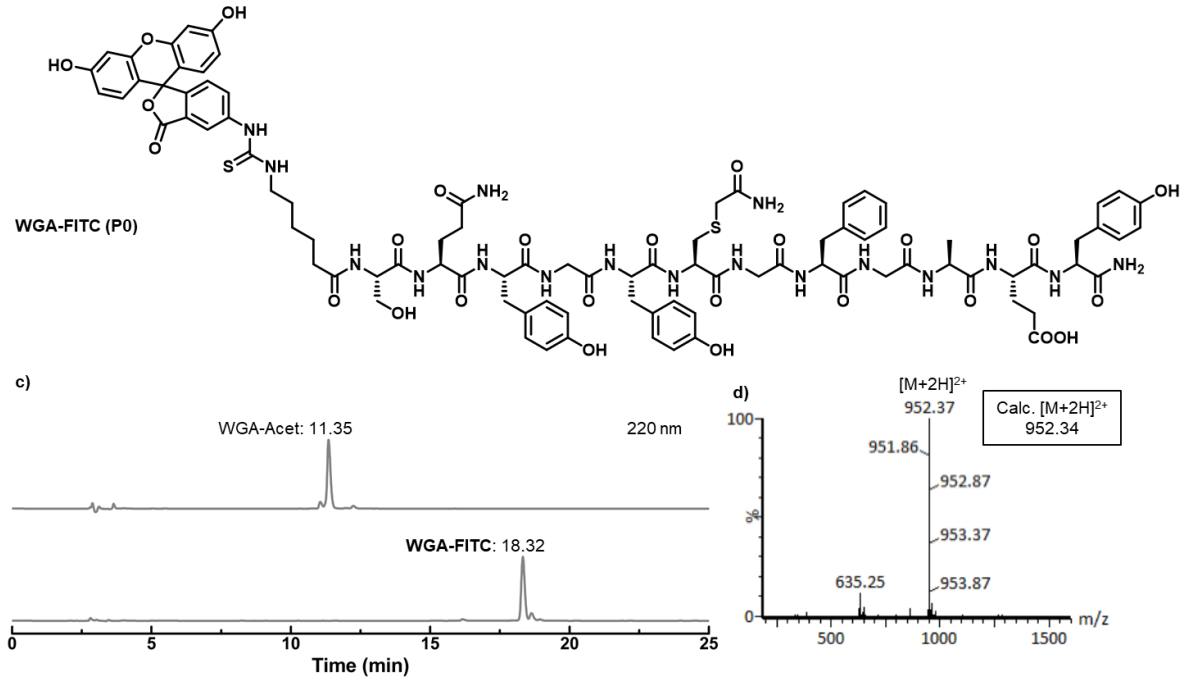


Figure 3A.19: Modification of WGA to WGA-Acet a) HPLC analysis, >90% purity; b) Mass data. Acet refers to the acetamide group. Labeling of WGA-Acet by FITC a) HPLC analysis, >92% purity; b) Mass data.

WGA-BA1-FITC (P1)

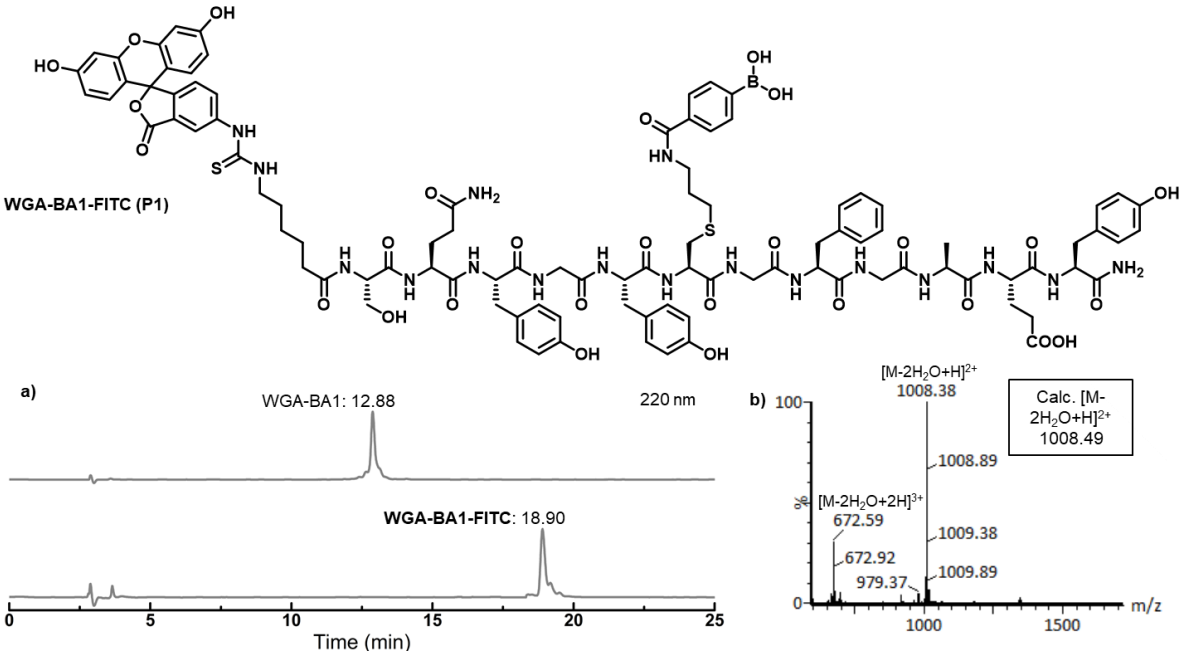


Figure 3A.20: WGA-BA1 labelling by FITC a) HPLC analysis, >90% purity; b) Mass data.

WGA-BA2-FITC (P2)

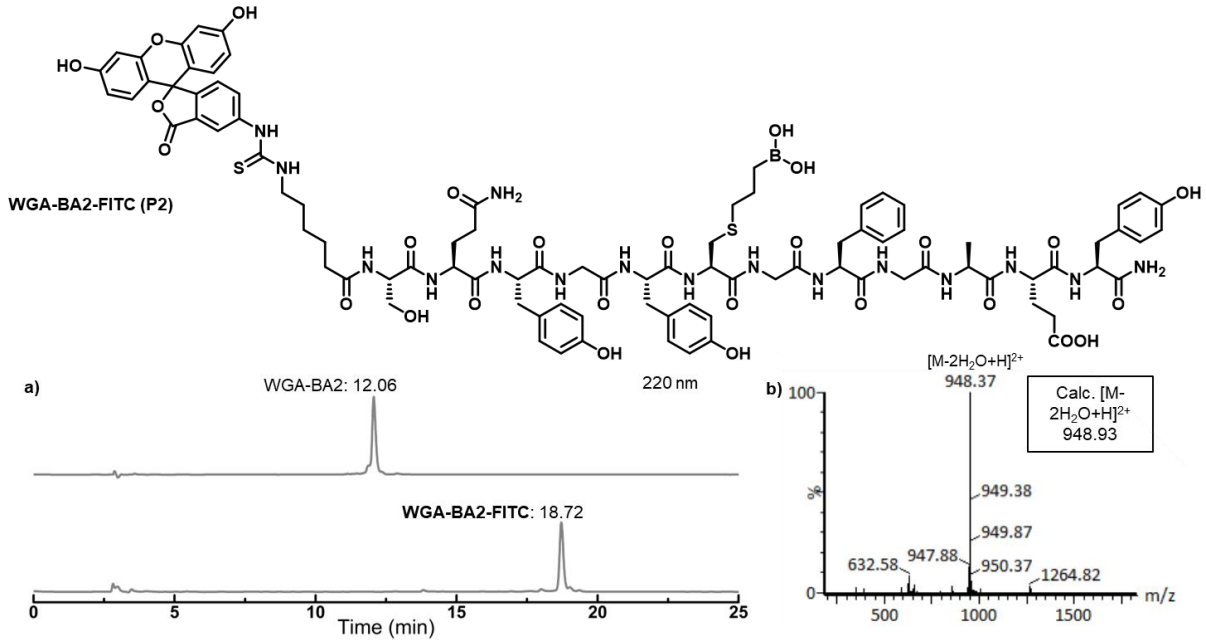


Figure 3A.21: WGA-BA2 labelling by FITC a) HPLC analysis, >94% purity; b) Mass data.

WGA-BA3-FITC (P3)

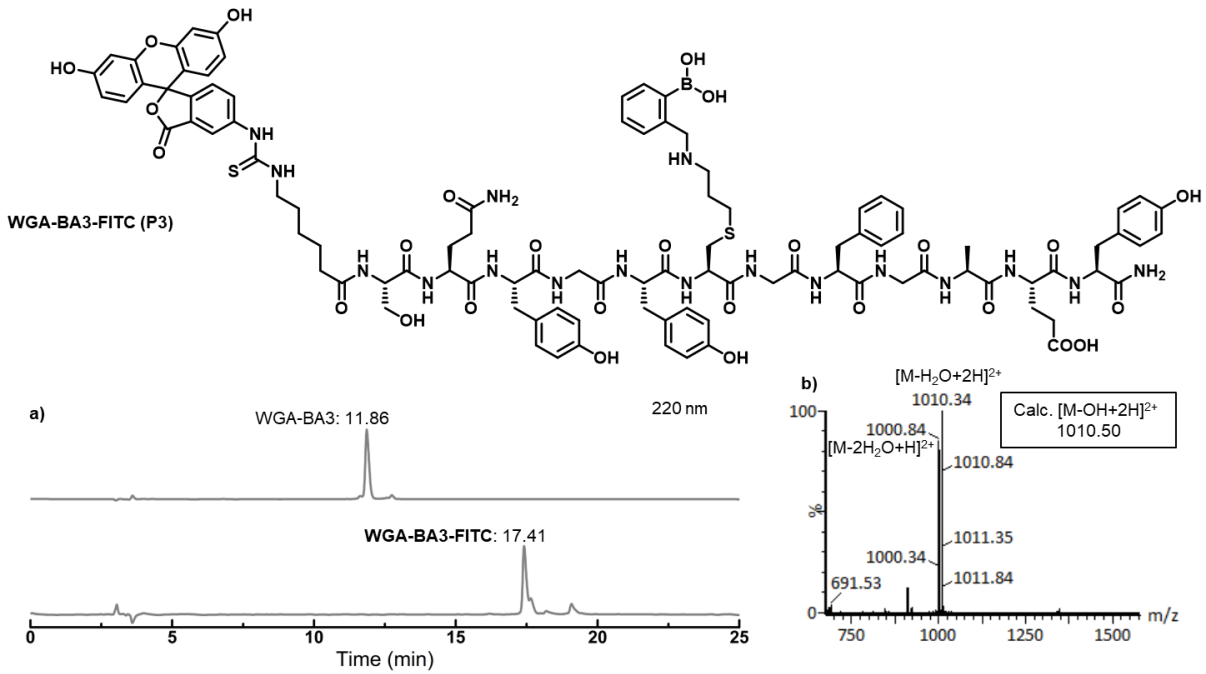


Figure 3A.22: WGA-BA3 labelling by FITC a) HPLC analysis, 90% purity; b) Mass data.

WGA-BA4-FITC (P4)

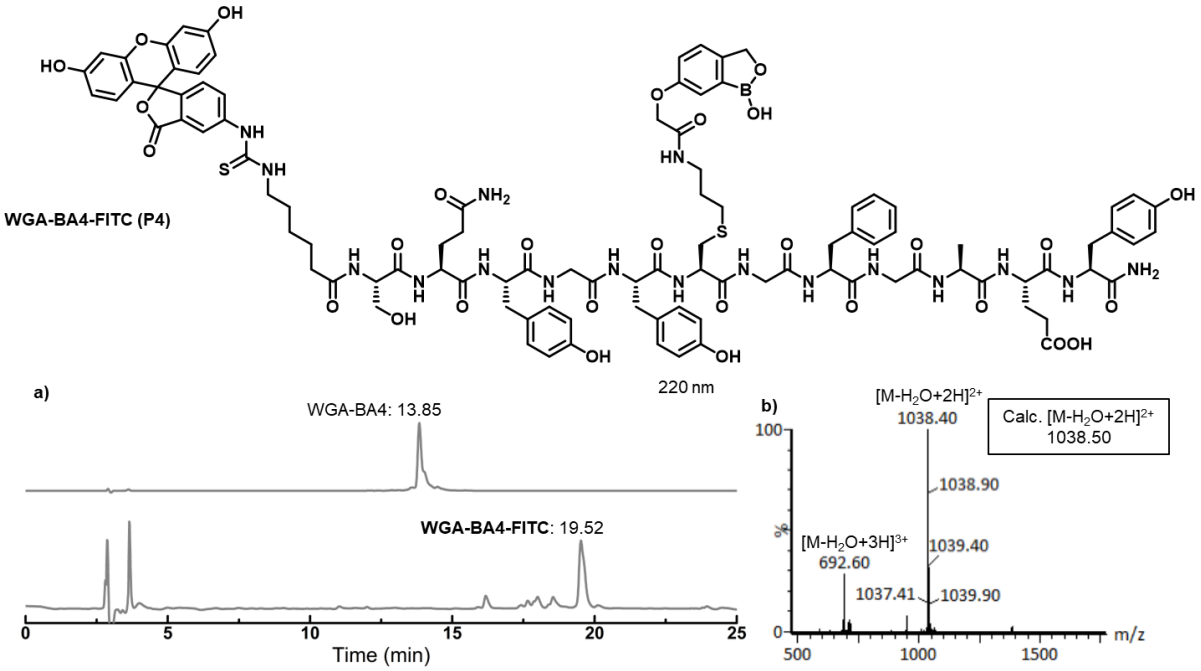


Figure 3A.23: WGA-BA4 labelling by FITC a) HPLC analysis, >85% purity; b) Mass data.

WGA-BA5-FITC (P5)

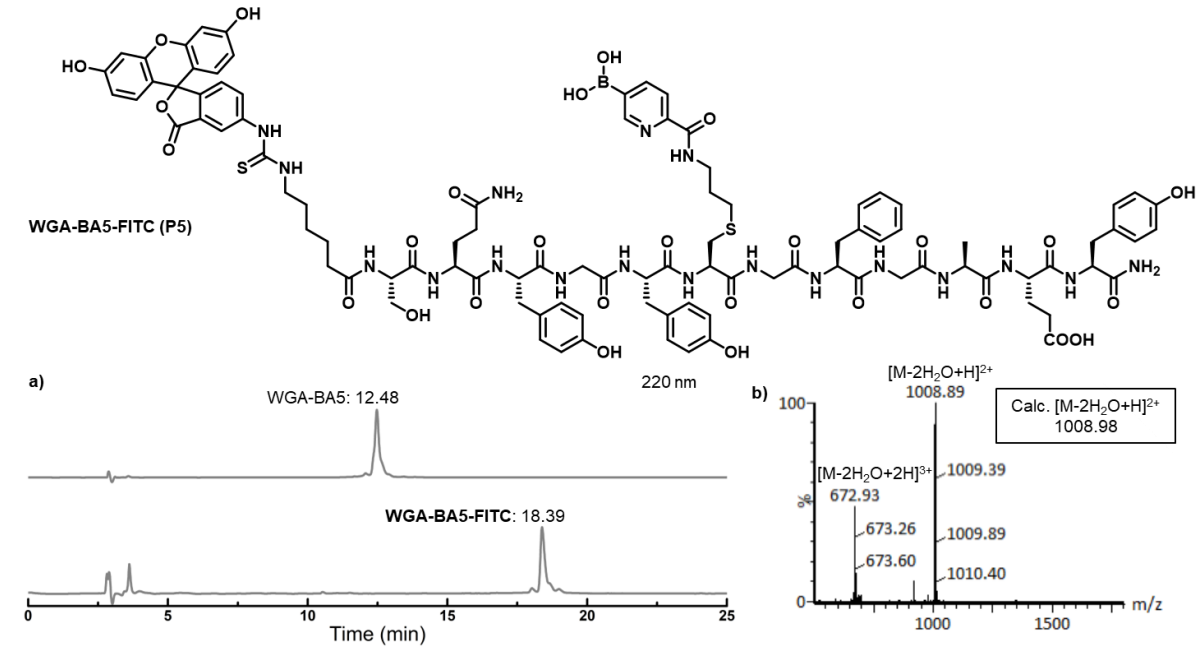


Figure 3A.24: WGA-BA5 labelling by FITC a) HPLC analysis, >90% purity; b) Mass data.

3A.4.11. *EXPERIMENTS ON CANCER CELLS*

The flow cytometry results, along with their median values, are presented in Figure 3A.5 of the main text and Figure 3A.25a and 3A.27. The median fluorescence intensity of FBS-treated cell samples was extracted (Figure 3A.25b) and plotted against FBS concentration in Figure 3A.25c. The percentage inhibition was calculated by determining the fractional median fluorescence intensities at 5% and 10% with respect to 100% FBS. Each experiment was triplicated, and their staining intensity is presented in a bar graph (Figure 3A.26 and 3A.28). All flow cytometer experiments are technical replicates (n =1, triplicated data represented with error bar graphs by calculating standard deviation). However, some crucial flow experiments were repeated during EC_{50} measurements and found to be reproducible. One set of data is presented in the main text Figure 3A.5b-d as histograms to compare the flow cytometer readout.

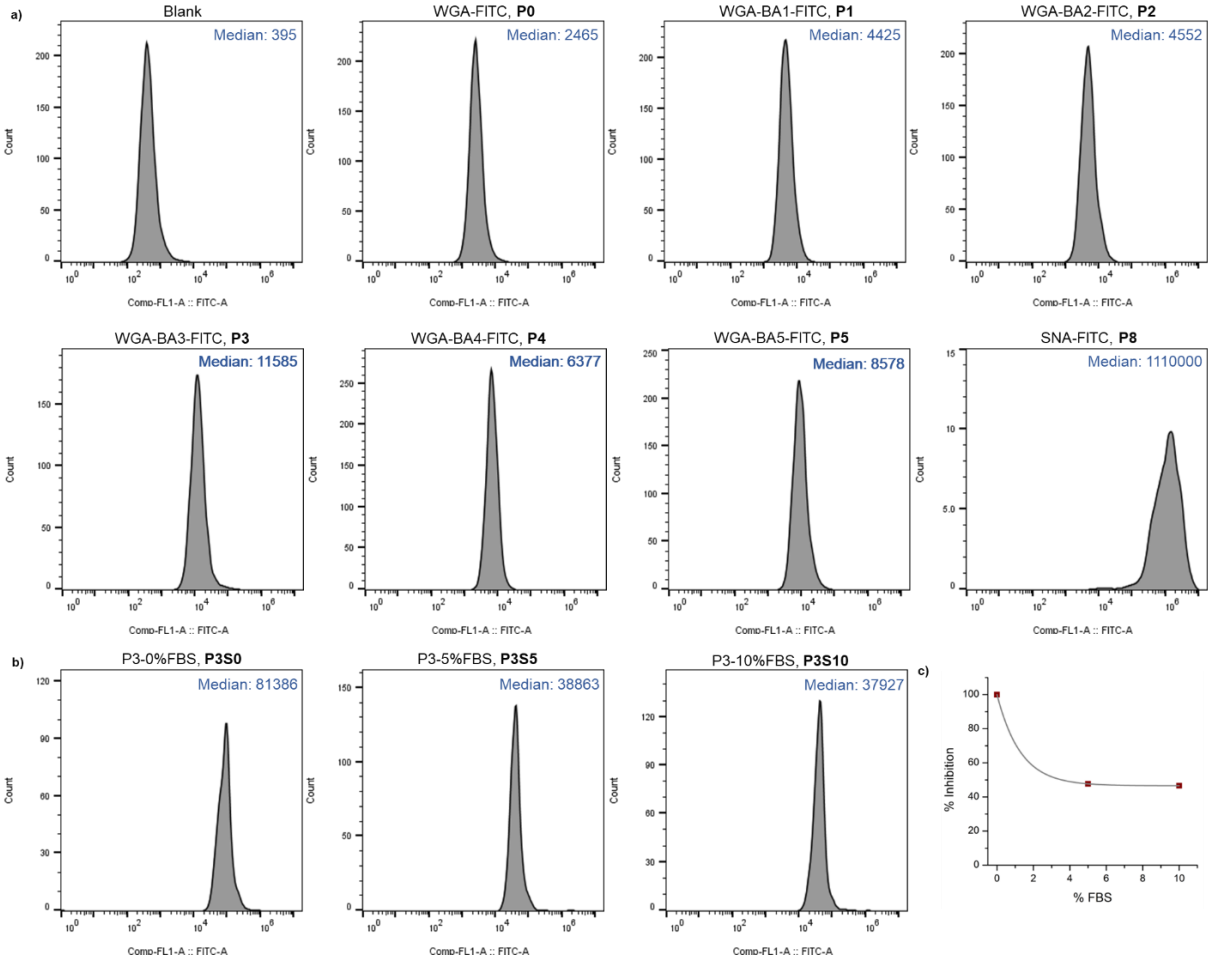


Figure 3A.25: Flow cytometry data of A549 stained cells with a) ~10 µg/mL FITC labelled WGA and SNA probes; b) ~10 µg/mL P3 in the presence of FBS. The median values are shown at the top right corner for each experiment; c) FBS inhibition assay of P3 staining with A549 cells.

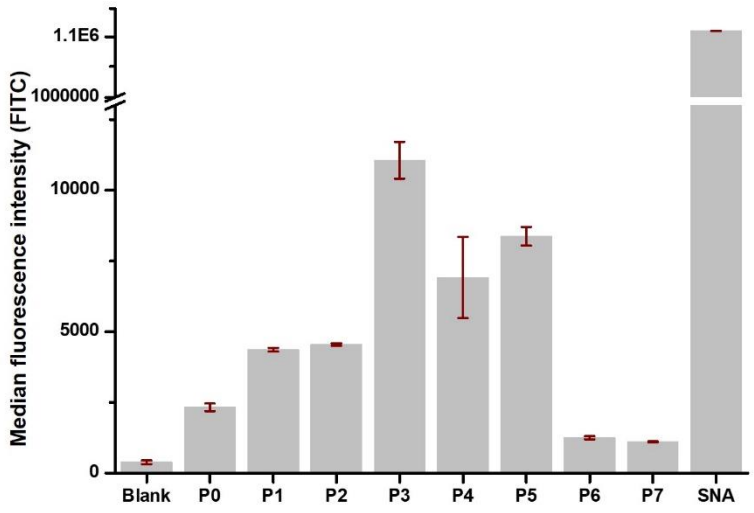


Figure 3A.26: Comparative flow cytometry analysis of the median fluorescence intensities (FITC) of P0-SNA on the A549 cell line.

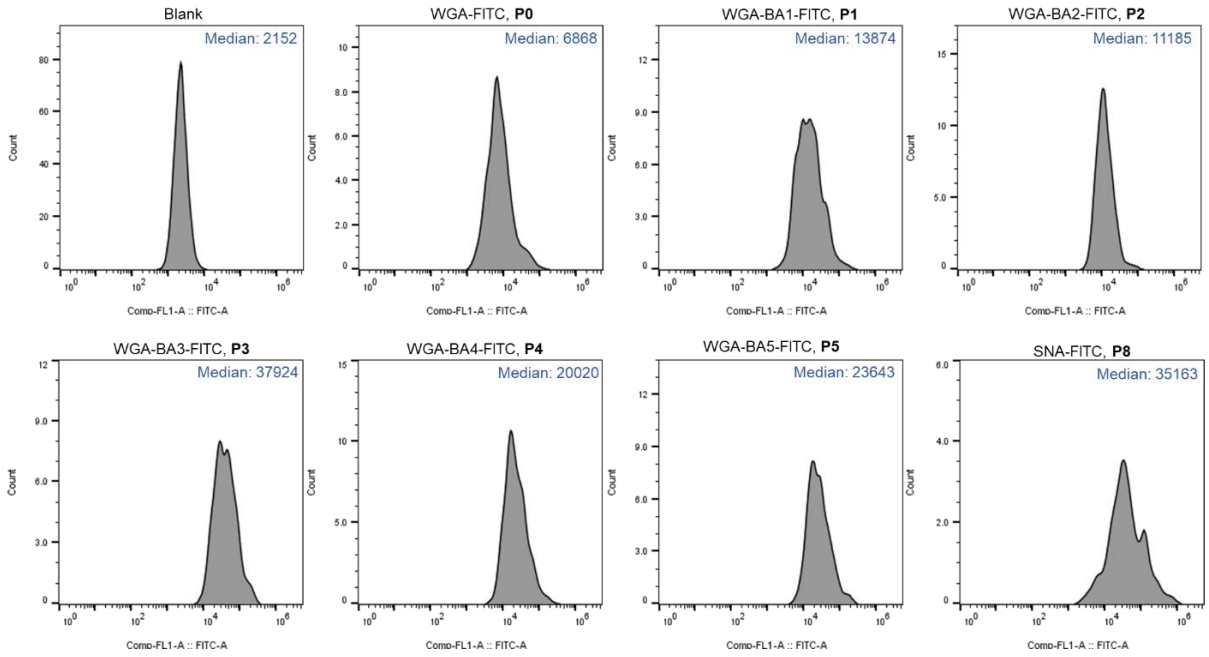


Figure 3A.27: Flow cytometry data of MDA-MB-231 cells with ~10 µg/mL FITC labelled WGA and SNA probes. The median values are shown at the top right corner for each experiment.

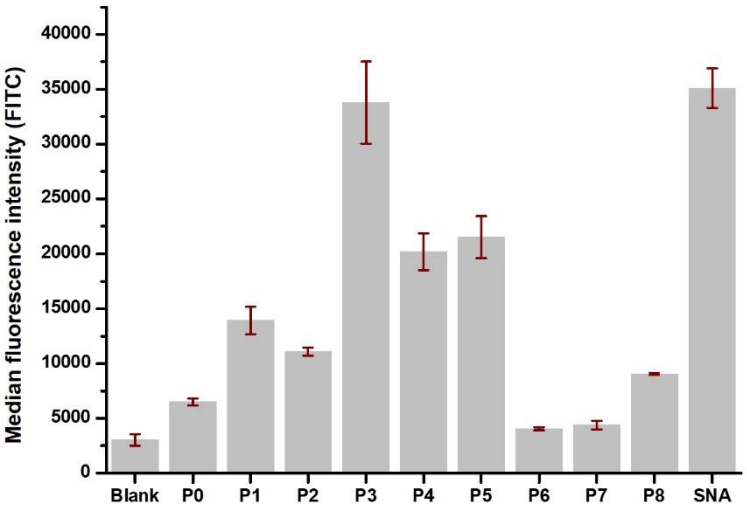


Figure 3A.28: Comparative flow cytometry analysis of the median fluorescence intensities (FITC) of P0-SNA on MDA-MB-231 cell line.

Microscopic analysis of stained cells

The microscopic results for MDA-MB231 are shown in Figure 3A.6b of the main text, and for A549 are shown in Figure 3A.29. For better representations, a portion of the captured frames is shown with the maintained scale (20 µm bar in the image).

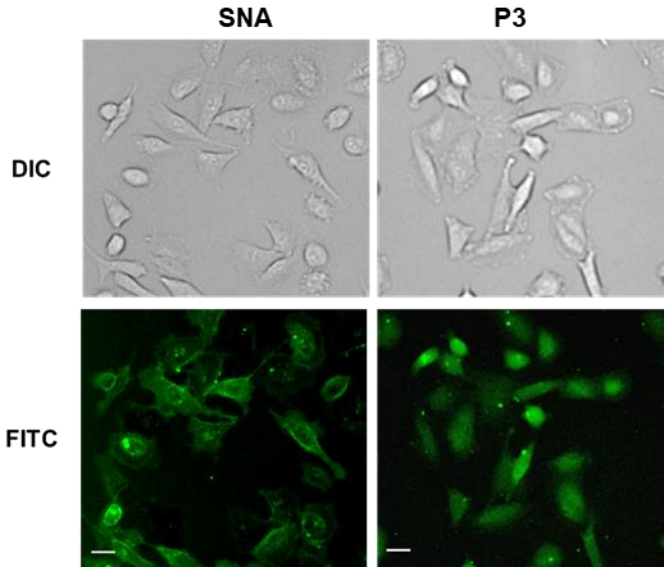


Figure 3A.29: Epifluorescence images with A549. Cells were cultured in Petri dishes and treated with SNA FITC (~10 µg/mL, 1mL) and P3 (~10 µg/mL, 1mL). The white bar is a scale of 20 µm.

3A.4.12. *CHEMICAL AND SERUM STABILITY OF WGA-BA3*

Chemical stability: To check the stability of the peptide at physiological pH and open air, 50 µM of WGA-BA3 was taken in PBS buffer pH 7.4 and kept at room temperature. LCMS analysis was performed after 24

hr intervals for 3 days to check the stability of the peptide in the physiological buffer. The data in Figure 3A.30 shows that the peptide (R_t 2.53 min) is stable under these conditions. No significant degradation was observed under these conditions. The peaks at 4.49 min and 4.59 min are column contaminants.

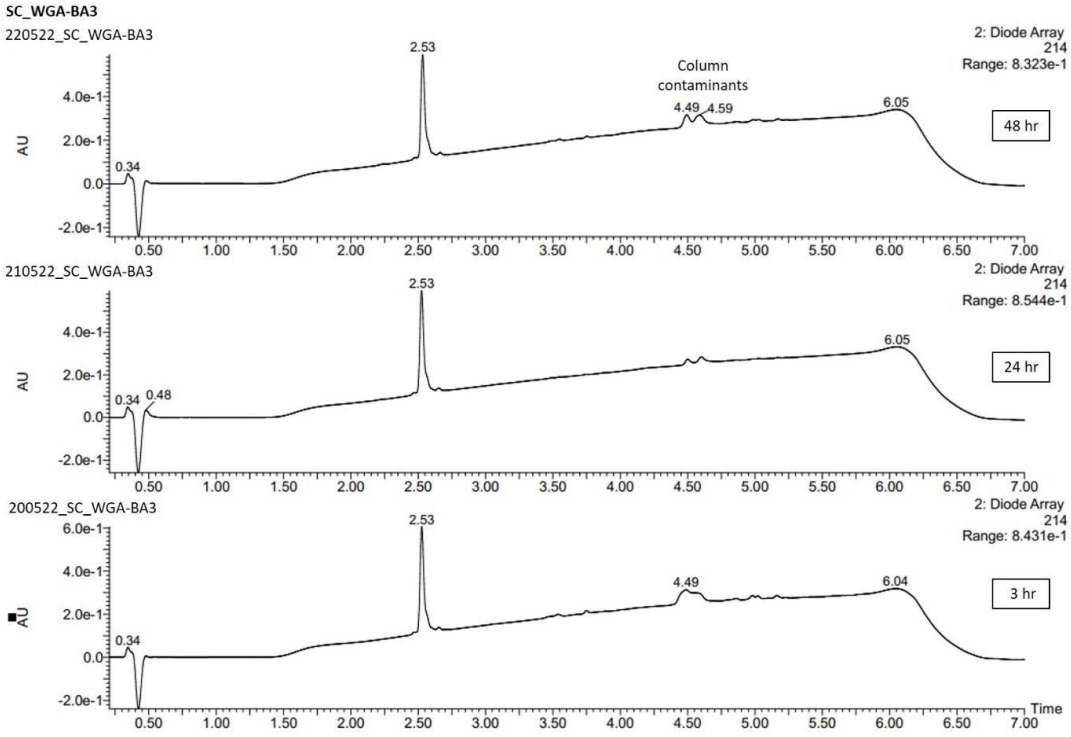


Figure 3A.30: Chemical stability study of WGA-BA3 in PBS 7.4 and open air.

Serum stability: The serum stability of WGA and WGA-BA3 was carried out according to the reported literature[36]. Human serum freshly extracted from an individual was diluted to 25% with PBS (10 mM, pH 7.4) and kept at 37 °C. 200 µM peptides (10 µL, 16 mM stock in DMSO) were added to 800 µL of 25% serum and incubated at 37 °C. 60 µL aliquots were withdrawn at 0, 1, 4, and 24 hr intervals and added to 30 µL 1:1 MeOH/10% trichloroacetic acid (TCA) and 110 µL H_2O. The resulting suspension was vortexed and spun at 4000 rpm for 5 min. The supernatant was freeze-dried, and the resulting residue was reconstituted in 1:1 ACN/H_2O. The mixture was again spun down, and the supernatant was subjected to analytical HPLC analysis. The sample collected at 0 hr was considered the initial standard, against which the peak areas of samples at different time intervals were compared. All data were measured in duplicate.

From the degradation pattern, it can be observed that the installation of BA3 remarkably improves the stability of the WGA peptide. About 90% of the WGA peptide was found to degrade after 1 hr, but only about 10% of the WGA-BA3 peptide degraded after 1 hr. After 24 hr, near complete degradation of WGA was observed, and for WGA-BA3, ~46% of degradation was observed.

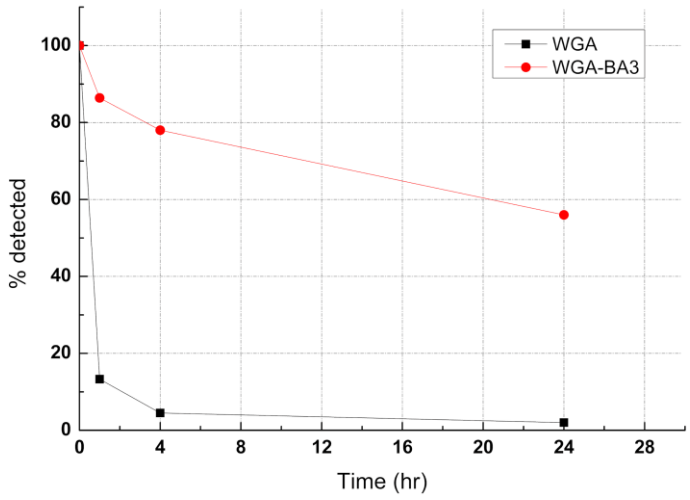
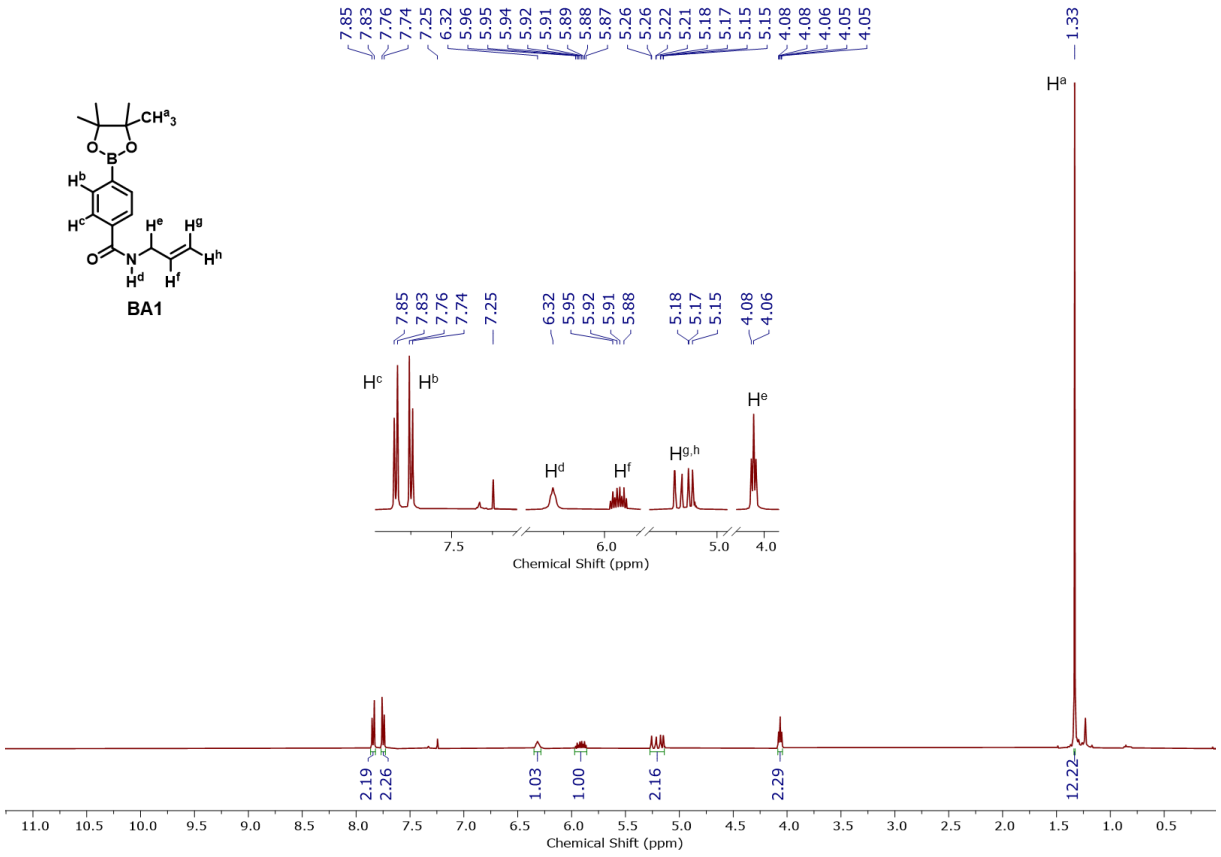


Figure 3A.31: Serum stability study of WGA and WGA-BA3 carried out at an interval of 0, 1, 4, and 24 hr.

3A.5 *NMR DATA*



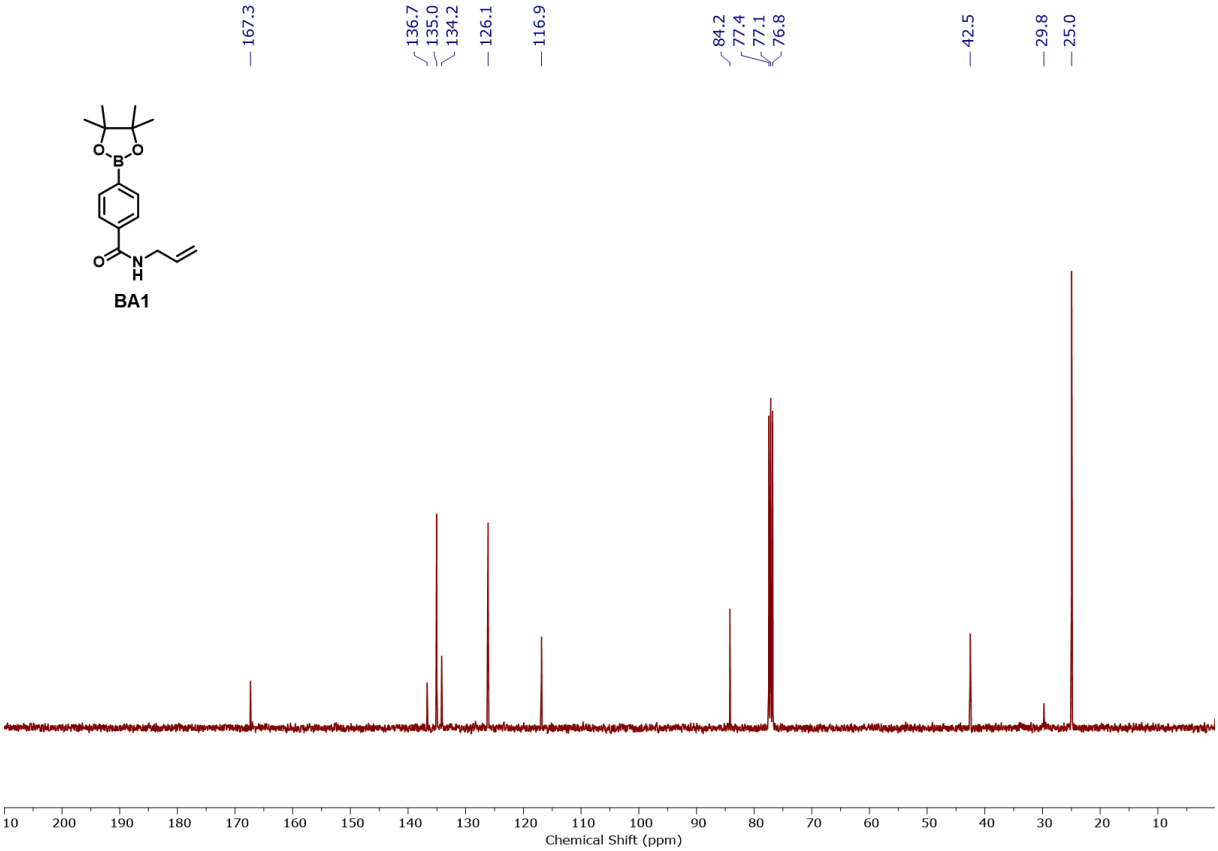
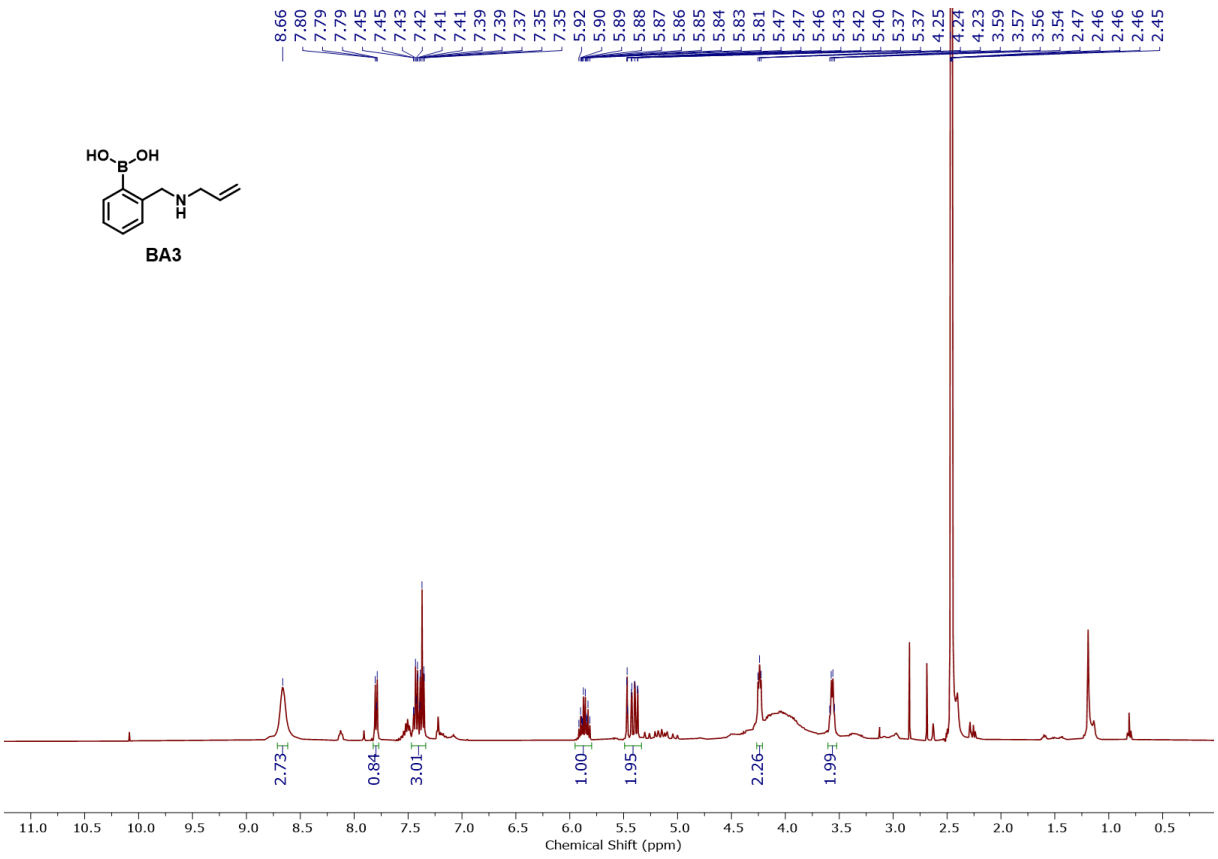


Figure 3A.32: ^{1}H (400 MHz) and ^{13}C (101 MHz) NMR spectra of BA1 in $CDCl_3$.



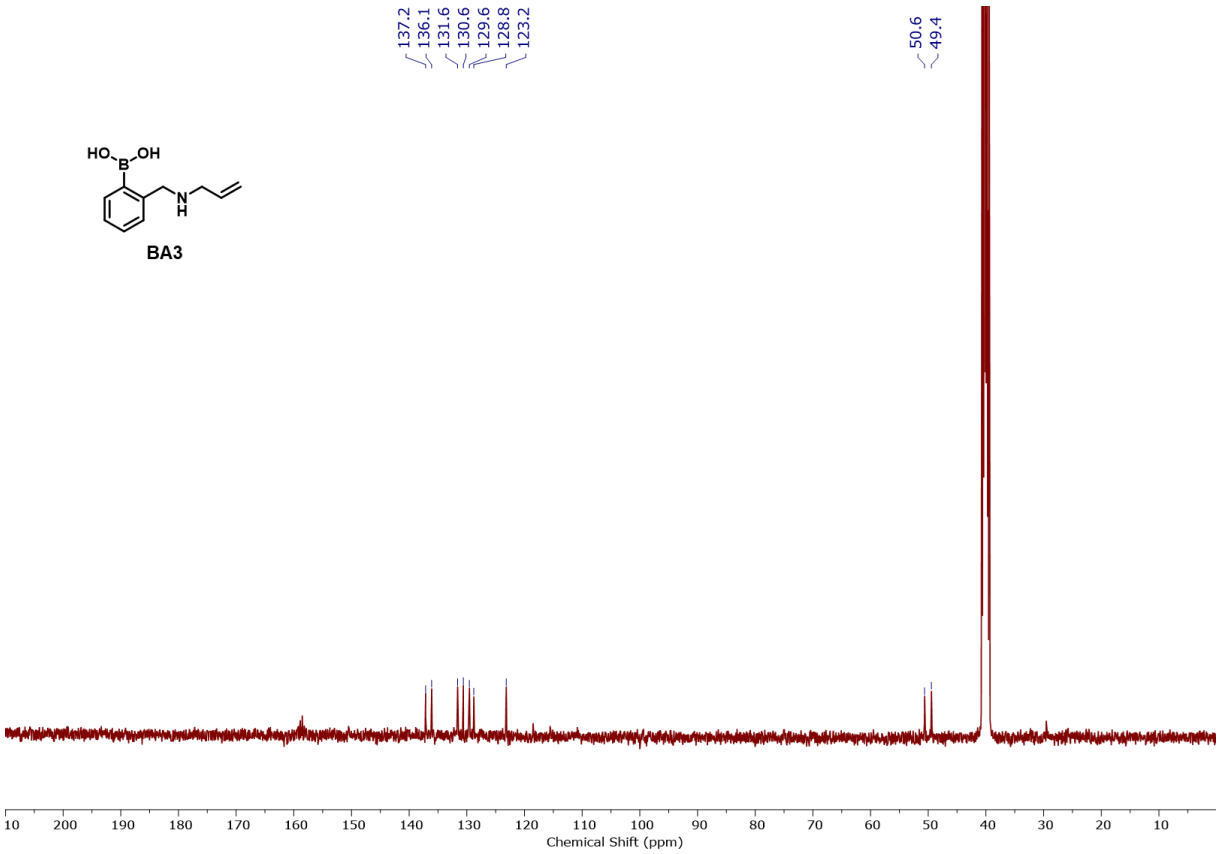
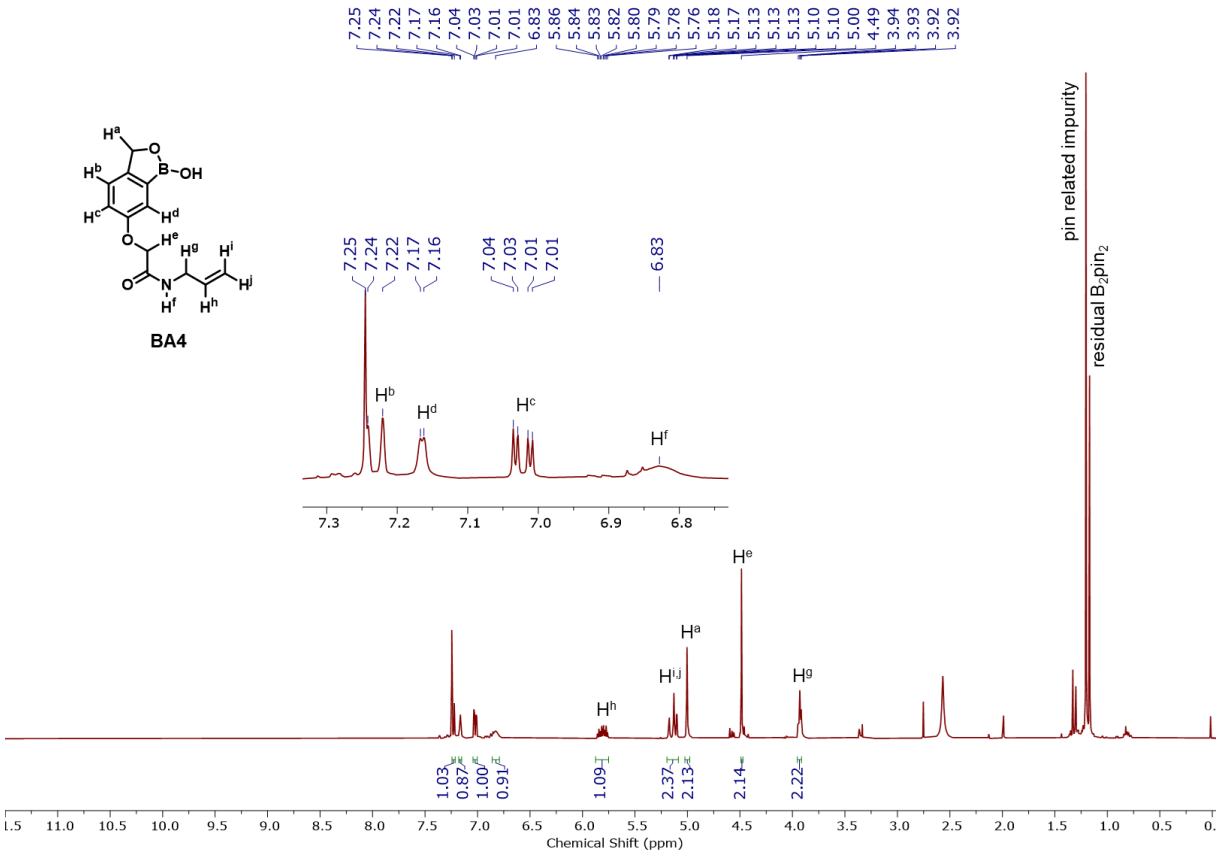


Figure 3A.33: ^{1}H (400 MHz) and ^{13}C (101 MHz) NMR spectra of BA3 in DMS)-d_6.



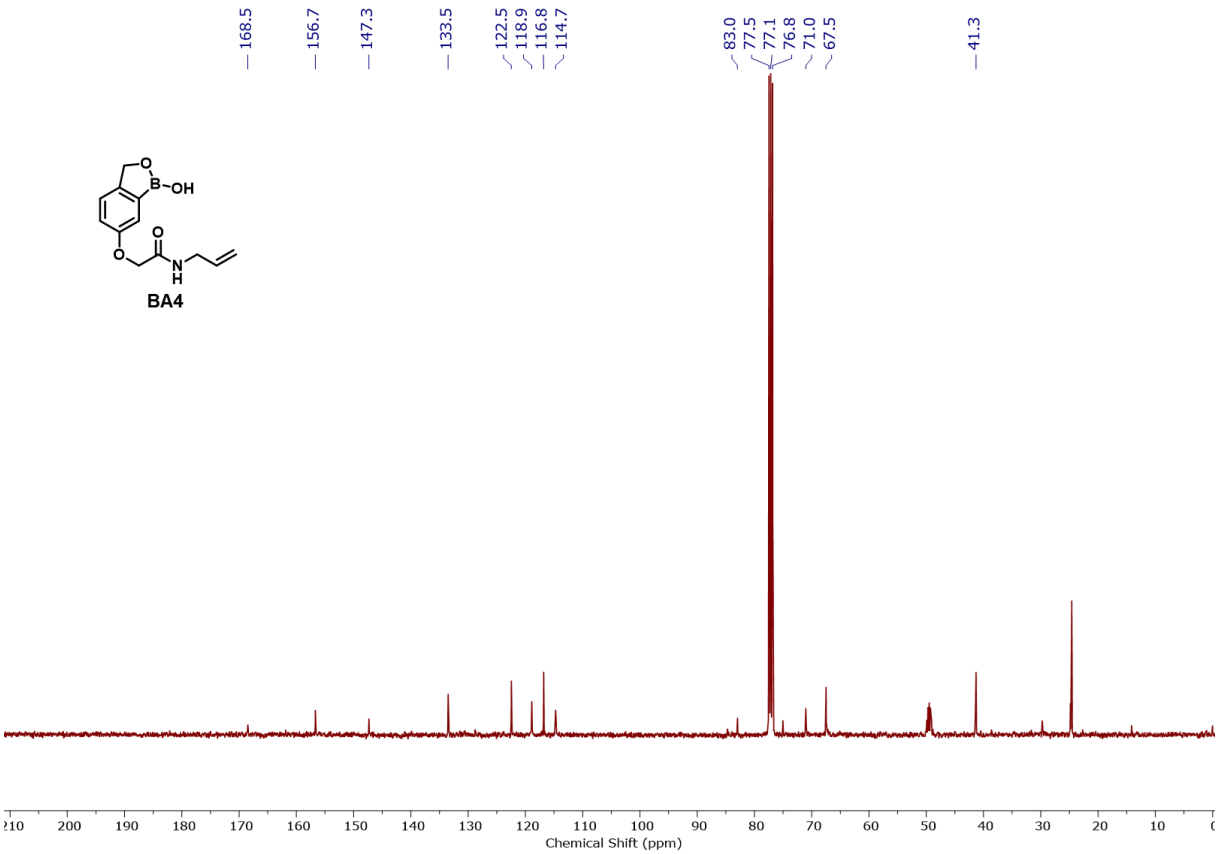
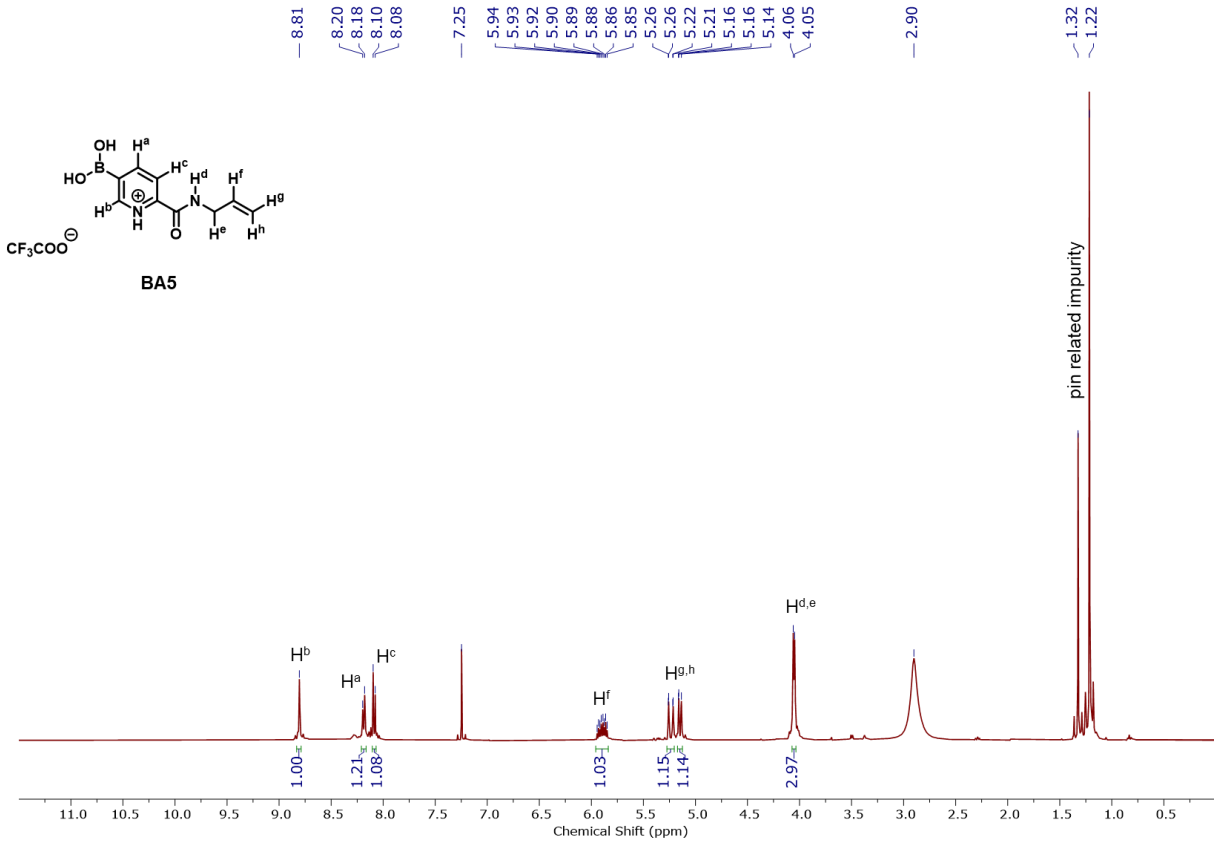


Figure 3A.34: ^{1}H (400 MHz) and ^{13}C (101 MHz) NMR spectra of crude BA4 in $CDCl_3$.



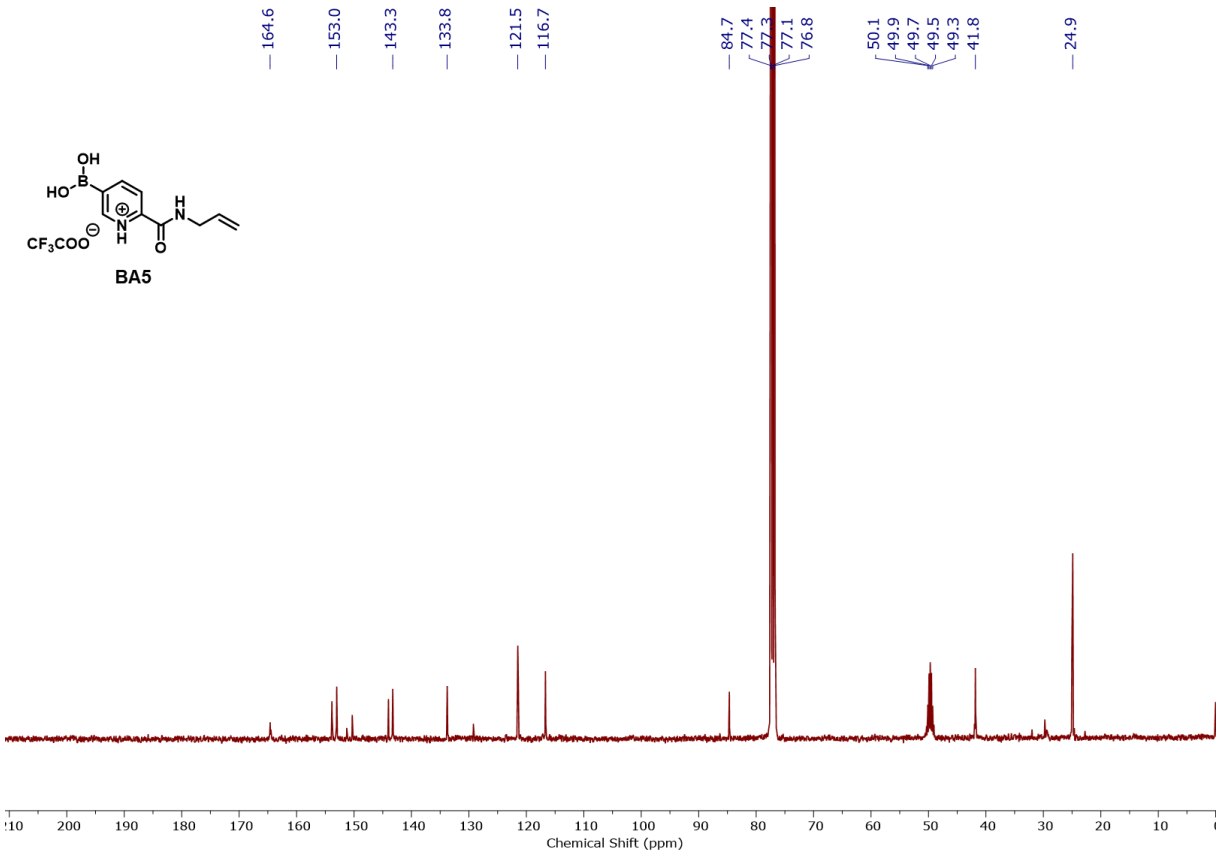


Figure 3A.35: ^{1}H (400 MHz) and ^{13}C (101 MHz) NMR spectra of BA5 in 10% $CD_3OD/CDCl_3$.

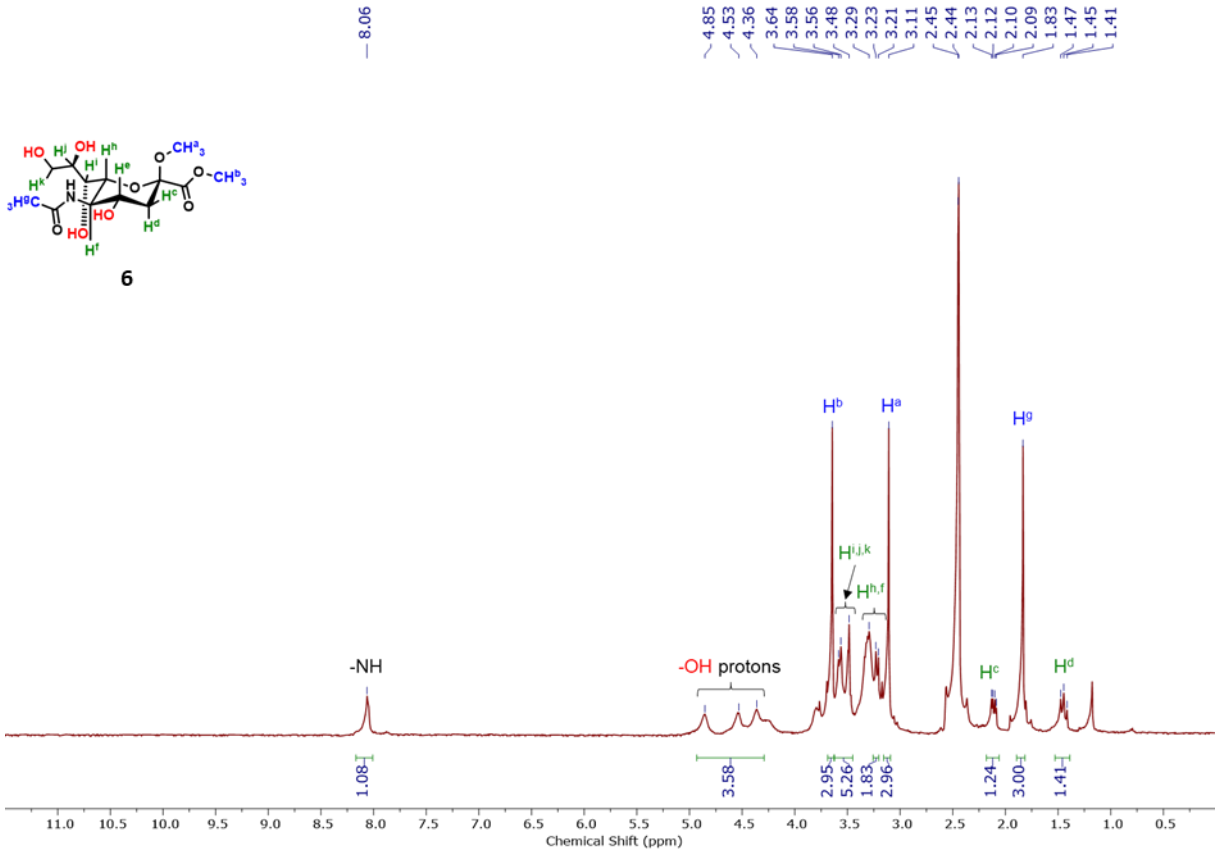


Figure 3A.36: ^{1}H (400 MHz) spectra of compound 6 in DMSO-d_6.

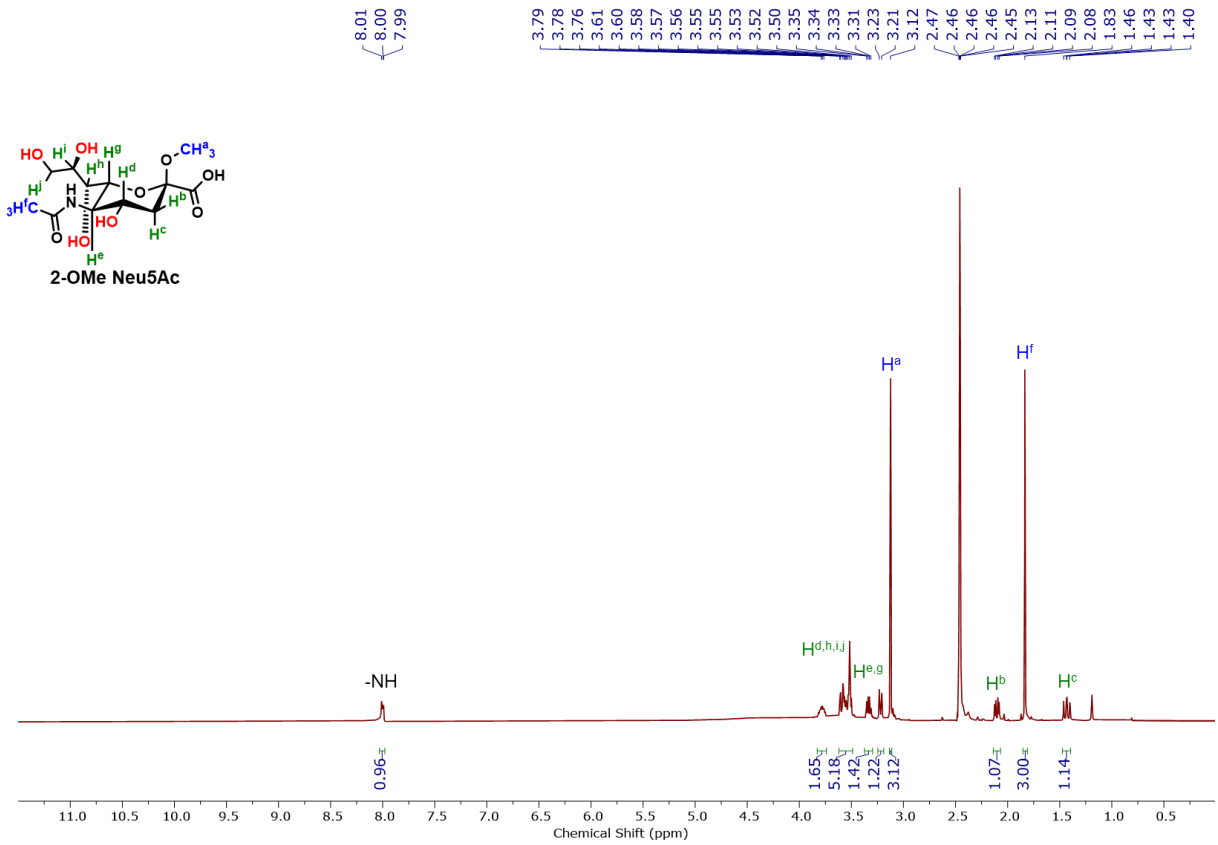


Figure 3A.37: ^{1}H (400 MHz) spectra of compound 2-OMe Neu5Ac in DMSO-d_6.

3A.6 *REFERENCES*

(1) 2019, I. U. of P. and A. C. (IUPAC) Y. R.-. Glycans. https://doi.org/10.1351/goldbook.G02645 WV - 3.0.1.

(2) Bicker, K. L.; Sun, J.; Lavigne, J. J.; Thompson, P. R. Boronic Acid Functionalized Peptidyl Synthetic Lectins: Combinatorial Library Design, Peptide Sequencing, and Selective Glycoprotein Recognition. *ACS Comb Sci* **2011**, *13* (3), 232–243. https://doi.org/10.1021/CO100054E.

(3) Varki, A. Biological Roles of Glycans. *Glycobiology* **2017**, *27* (1), 3–49. https://doi.org/10.1093/glycob/cww086.

(4) Boligan, K. F.; Mesa, C.; Fernandez, L. E.; von Gunten, S. Cancer Intelligence Acquired (CIA): Tumor Glycosylation and Sialylation Codes Dismantling Antitumor Defense. *Cellular and Molecular Life Sciences* **2015**, *72* (7), 1231–1248. https://doi.org/10.1007/s00018-014-1799-5.

(5) Pearce, O. M. T.; Läubli, H. Sialic Acids in Cancer Biology and Immunity. *Glycobiology* **2015**, *26* (2), 111–128. https://doi.org/10.1093/glycob/cwv097.

(6) Silsirivanit, A. *Glycosylation Markers in Cancer*, 1st ed.; Elsevier Inc., 2019; Vol. 89. https://doi.org/10.1016/bs.acc.2018.12.005.

(7) Reily, C.; Stewart, T. J.; Renfrow, M. B.; Novak, J. Glycosylation in Health and Disease. *Nat Rev Nephrol* **2019**, *15* (6), 346–366. https://doi.org/10.1038/s41581-019-0129-4.

(8) Hollingsworth, M. A.; Swanson, B. J. Mucins in Cancer: Protection and Control of the Cell Surface. *Nat Rev Cancer* **2004**, *4* (1), 45–60. https://doi.org/10.1038/nrc1251.

(9) Munkley, J. The Glycosylation Landscape of Pancreatic Cancer (Review). *Oncol Lett* **2019**, *17* (3), 2569–2575. https://doi.org/10.3892/ol.2019.9885.

(10) Wu, L.; Qu, X. Cancer Biomarker Detection: Recent Achievements and Challenges. *Chem Soc Rev* **2015**, *44* (10), 2963–2997. https://doi.org/10.1039/C4CS00370E.

(11) Matsumoto, A.; Miyahara, Y. 'Borono-Lectin' Based Engineering as a Versatile Platform for Biomedical Applications. *Sci Technol Adv Mater* **2018**, *19* (1), 18–30. https://doi.org/10.1080/14686996.2017.1411143.

(12) Williams, G. T.; Kedge, J. L.; Fossey, J. S. Molecular Boronic Acid-Based Saccharide Sensors. *ACS Sens* **2021**, *6* (4), 1508–1528. https://doi.org/10.1021/acssensors.1c00462.

(13) Chatterjee, S.; Tripathi, N. M.; Bandyopadhyay, A. The Modern Role of Boron as a 'Magic Element' in Biomedical Science: Chemistry Perspective. *Chemical Communications* **2021**, *57* (100), 13629–13640. https://doi.org/10.1039/D1CC05481C.

(14) Bandyopadhyay, A.; Gao, J. Targeting Biomolecules with Reversible Covalent Chemistry. *Curr Opin Chem Biol* **2016**, *34*, 110–116. https://doi.org/10.1016/j.cbpa.2016.08.011.

(15) Matsumoto, A.; Stephenson-Brown, A. J.; Khan, T.; Miyazawa, T.; Cabral, H.; Kataoka, K.; Miyahara, Y. Heterocyclic Boronic Acids Display Sialic Acid Selective Binding in a Hypoxic Tumor Relevant Acidic Environment. *Chem Sci* **2017**, *8* (9), 6165–6170. https://doi.org/10.1039/c7sc01905j.

(16) Chaudhary, P. M.; Murthy, R. V.; Yadav, R.; Kikkeri, R. A Rationally Designed Peptidomimetic Biosensor for Sialic Acid on Cell Surfaces. *Chemical Communications* **2015**, *51* (38), 8112–8115. https://doi.org/10.1039/c5cc01662b.

(17) Otremba, T.; Ravoo, B. J. Linear and Cyclic Carbohydrate Receptors Based on Peptides Modified with Boronic Acids. *ChemistrySelect* **2016**, *1* (9), 2079–2084. https://doi.org/10.1002/slct.201600567.

(18) Kowalczyk, W.; Sanchez, J.; Kraaz, P.; Hutt, O. E.; Haylock, D. N.; Duggan, P. J. The Binding of Boronated Peptides to Low Affinity Mammalian Saccharides. *Peptide Science* **2018**, *110* (3). https://doi.org/10.1002/pep2.23101.

(19) Zou, Y.; Broughton, D. L.; Bicker, K. L.; Thompson, P. R.; Lavigne, J. J. Peptide Borono Lectins (PBLs): A New Tool for Glycomics and Cancer Diagnostics. *ChemBioChem* **2007**, *8* (17), 2048–2051. https://doi.org/10.1002/cbic.200700221.

(20) Bicker, K. L.; Sun, J.; Harrell, M.; Zhang, Y.; Pena, M. M.; Thompson, P. R.; Lavigne, J. J. Synthetic Lectin Arrays for the Detection and Discrimination of Cancer Associated Glycans and Cell Lines. *Chem Sci* **2012**, *3* (4), 1147–1156. https://doi.org/10.1039/c2sc00790h.

(21) Wright, C. S. 2·2 Å Resolution Structure Analysis of Two Refined N-Acetylneuraminyl-Lactose-Wheat Germ Agglutinin Isolectin Complexes. *J Mol Biol* **1990**, *215* (4), 635–651. https://doi.org/10.1016/S0022-2836(05)80174-3.

(22) Pasquale, A. Di; Tommasone, S.; Xu, L.; Ma, J.; Mendes, P. M. Cooperative Multipoint Recognition of Sialic Acid by Benzoboroxole-Based Receptors Bearing Cationic Hydrogen-Bond Donors. *Cite This: J. Org. Chem* **2020**, *85*, 8338. https://doi.org/10.1021/acs.joc.0c00173.

(23) Dale, J. 191. The Stereochemistry of Polyborate Anions and of Borate Complexes of Diols and Certain Polyols. *Journal of the Chemical Society (Resumed)* **1961**, No. 0, 922–930. https://doi.org/10.1039/JR9610000922.

(24) Chatterjee, S.; Bandyopadhyay, A. Cysteine-Selective Installation of Functionally Diverse Boronic Acid Probes on Peptides. *Org Lett* **2023**, *25* (13), 2223–2227. https://doi.org/10.1021/acs.orglett.3c00386.

(25) Nicholls, J. M.; Bourne, A. J.; Chen, H.; Guan, Y.; Peiris, J. M. Sialic Acid Receptor Detection in the Human Respiratory Tract: Evidence for Widespread Distribution of Potential Binding Sites for Human and Avian Influenza Viruses. *Respir Res* **2007**, *8* (1), 1–10. https://doi.org/10.1186/1465-9921-8-73/FIGURES/4_611.

(26) Liu, Z.; He, H. Synthesis and Applications of Boronate Affinity Materials: From Class Selectivity to Biomimetic Specificity. *Acc Chem Res* **2017**, *50* (9), 2185–2193. https://doi.org/10.1021/ACS.ACCOUNTS.7B00179/ASSET/IMAGES/LARGE/AR-2017-00179S_0009.JPEG.

(27) Brooks, W. L. A.; Deng, C. C.; Sumerlin, B. S. Structure-Reactivity Relationships in Boronic Acid-Diol Complexation. *ACS Omega* **2018**, *3* (12), 17863–17870. https://doi.org/10.1021/acsomega.8b02999.

(28) Sparbier, K.; Asperger, A.; Resemann, A.; Kessler, I.; Koch, S.; Wenzel, T.; Stein, G.; Vorwerg, L.; Suckau, D.; Kostrzewa, M. Analysis of Glycoproteins in Human Serum by Means of Glycospecific Magnetic Bead Separation and LC-MALDI-TOF/TOF Analysis with Automated Glycopeptide Detection. *J Biomol Tech* **2007**, *18* (4), 252–258.

(29) Gatt, S.; Gazit, B.; Barenholz, Y. Effect of Bile Salts on the Hydrolysis of Gangliosides, Glycoproteins and Neuraminyl-Lactose by the Neuraminidase of Clostridium Perfringens. *Biochemical Journal* **1981**, *193* (1), 267–273. https://doi.org/10.1042/BJ1930267.

(30) Zhu, L.; Shabbir, S. H.; Gray, M.; Lynch, V. M.; Sorey, S.; Anslyn, E. V. A Structural Investigation of the N−B Interaction in an o -(N , N -Dialkylaminomethyl)Arylboronate System. *J Am Chem Soc* **2006**, *128* (4), 1222–1232. https://doi.org/10.1021/ja055817c.

(31) Sun, X.; Chapin, B. M.; Metola, P.; Collins, B.; Wang, B.; James, T. D.; Anslyn, E. V. The Mechanisms of Boronate Ester Formation and Fluorescent Turn-on in Ortho-Aminomethylphenylboronic Acids. *Nat Chem* **2019**, *11* (9), 768–778. https://doi.org/10.1038/s41557-019-0314-x.

(32) Bojar, D.; Meche, L.; Meng, G.; Eng, W.; Smith, D. F.; Cummings, R. D.; Mahal, L. K. A Useful Guide to Lectin Binding: Machine-Learning Directed Annotation of 57 Unique Lectin Specificities. *ACS Chem Biol* **2022**, *17* (11), 2993–3012. https://doi.org/10.1021/acschembio.1c00689.

(33) Buchini, S.; Gallat, F. X.; Greig, I. R.; Kim, J. H.; Wakatsuki, S.; Chavas, L. M. G.; Withers, S. G. Tuning Mechanism-Based Inactivators of Neuraminidases: Mechanistic and Structural Insights. *Angewandte Chemie - International Edition* **2014**, *53* (13), 3382–3386. https://doi.org/10.1002/ANIE.201309675.

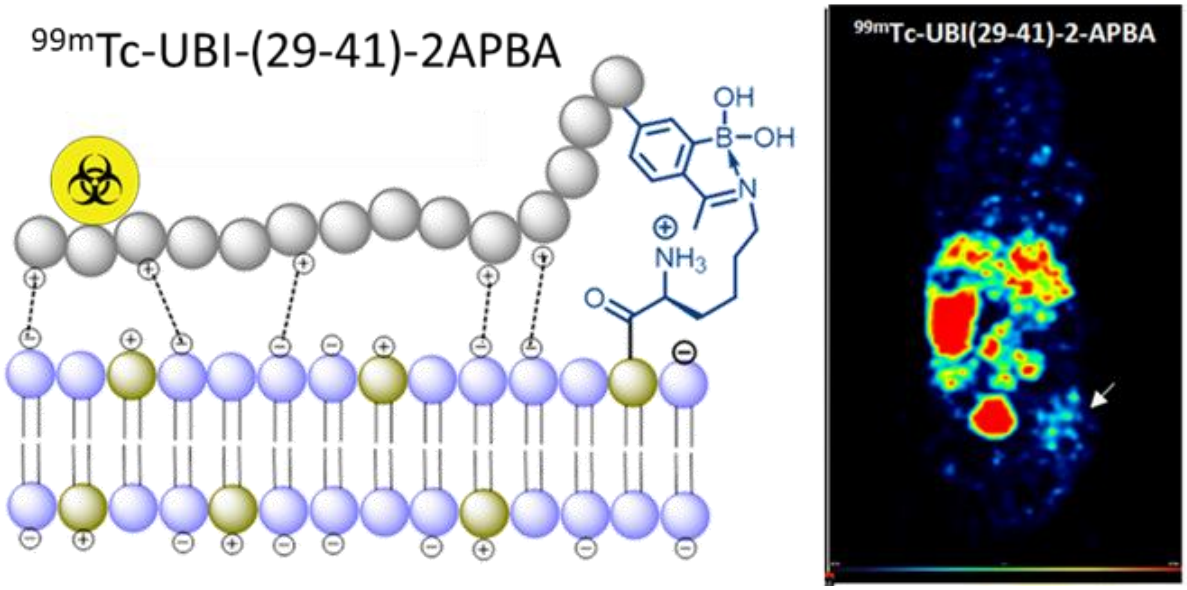
(34) Hernandez, E. T.; Kolesnichenko, I. V.; Reuther, J. F.; Anslyn, E. V. An Efficient Methodology to Introduce O-(Aminomethyl)Phenyl-Boronic Acids into Peptides: Alkylation of Secondary Amines. *New Journal of Chemistry* **2016**, *41* (1), 126–133. https://doi.org/10.1039/C6NJ02862D.

(35) Heerze1, L. D.; Smith, R. H.; Wang, N.; Armstrong, G. D. Utilization of Sialic Acid-Binding Synthetic Peptide Sequences Derived from Pertussis Toxin as Novel Anti-Inflammatory Agents. *Glycobiology* **1995**, *5* (4), 427–433. https://doi.org/10.1093/glycob/5.4.427.

(36) Harris, P. W. R.; Siow, A.; Yang, S.-H.; Wadsworth, A. D.; Tan, L.; Hermant, Y.; Mao, Y.; An, C.; Hanna, C. C.; Cameron, A. J.; Allison, J. R.; Chakraborty, A.; Ferguson, S. A.; Mros, S.; Hards, K.; Cook, G. M.; Williamson, D. A.; Carter, G. P.; Chan, S. T. S.; Painter, G. A.; Sander, V.; Davidson, A. J.; Brimble, M. A. Synthesis, Antibacterial Activity, and Nephrotoxicity of Polymyxin B Analogues Modified at Leu-7, <scp>d</Scp> -Phe-6, and the N-Terminus Enabled by S-Lipidation. *ACS Infect Dis* **2022**, *8* (12), 2413–2429. https://doi.org/10.1021/acsinfecdis.1c00347.

CHAPTER 3B

INTEGRATING A COVALENT PROBE WITH UBIQUICIDINE FRAGMENT ENABLES EFFECTIVE BACTERIAL INFECTION IMAGING



This two-part work has been published in RSC Medicinal Chemistry (a journal of Royal Society of Chemistry) and can be cited as "Bhatt Mitra, J.; **Chatterjee, S**.; Kumar, A.; Bandyopadhyay, A.; Mukherjee, A. Integrating a Covalent Probe with Ubiquicidin Fragment Enables Effective Bacterial Infection Imaging. *RSC Med Chem* **2022**, *13* (10)" and in EJNMMI Radiopharmacy and Chemistry (a journal of Springer Nature) and can be cited as "Bhatt Mitra, J.; **Chatterjee, S**.; Kumar, A.; Khatoon, E.; Chandak, A.; Rakshit, S.; Bandyopadhyay, A.; Mukherjee, A. Expanding a Peptide-Covalent Probe Hybrid for PET Imaging of S. Aureus Driven Focal Infections. *EJNMMI Radiopharm Chem* **2024**, *9* (1), 25."

"The greatest possibility of evil in self-medication [with penicillin] is the use of too-small doses, so that, instead of clearing up the infection, the microbes are educated to resist penicillin and a host of penicillin-fast organisms is bred out which can be passed on to other individuals and perhaps from there to others until they reach someone who gets a septicemia or a pneumonia which penicillin cannot save.

In such a case the thoughtless person playing with penicillin treatment is morally responsible for the death of the man who finally succumbs to infection with the penicillin-resistant organism. I hope this evil can be averted." *-Sir Alexander Fleming, 1945*

3B.1 *INTRODUCTION*

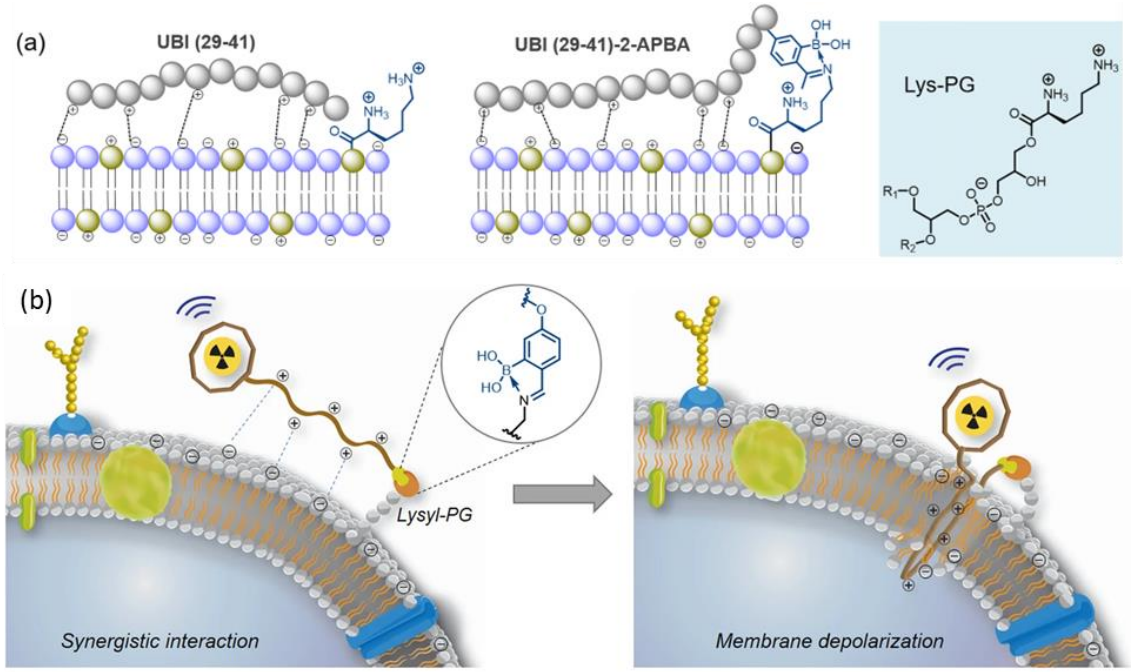
Bacterial infections are one of the significant causes of mortality and morbidity worldwide. It was reported that in 2017 alone, 20% (11.0 million) of global deaths were due to bacterial sepsis[1]. Approximately 5 million deaths were attributed to bacterial antimicrobial resistance (AMR) worldwide in 2019[2]. The evolution of drug-resistant pathogens demands focused research on designing new antimicrobials to overcome the crisis. Moreover, the development of fast and non-invasive methods for early detection of infection might play a pivotal role in managing different types of focal infections such as fever of unknown origin (FUO), osteomyelitis, and soft tissue infections (STI). Recently, World Health Organization (WHO) has declared antimicrobial resistance (AMR) as a challenge to human health[3]. To curb the spread of AMR there is an urgent need for minimizing the prescription of unnecessary antibiotics.

In recent years, significant research work on the development of non-invasive techniques for imaging bacterial infections has been reported in preclinical as well as clinical settings[4–6]. While structural imaging techniques such as magnetic resonance imaging (MRI) and computed tomography (CT) provide excellent structural resolution for detecting advanced infections, the identification of infectious foci at an early stage remains unattainable through these methods. This limitation arises because they solely rely on late-occurring anatomic changes in disease pathophysiology, such as the host immune response and tissue damage[6,7]. Direct imaging of bacteria using functional imaging modalities i.e., single photon emission computed tomography (SPECT) and positron emission tomography (PET), are being envisaged for non-invasive and efficient detection of focal infections. Strategies based on a variety of molecules, i.e., bacteria-specific metabolites (sorbitol, maltotriose), antibiotics (ciprofloxacin, rifampicin), antibodies against specific proteins present on microbial surfaces (1D9), and antimicrobial peptides (AMPs) such as ubiquicidin (UBI) and, human lactoferrin (hLF) have been employed to achieve this objective[6,8]. AMPs are attractive candidates for developing infection imaging radiopharmaceuticals due to their versatility and desirable pharmacokinetics. An ideal radiotracer for facilitating the expeditious diagnosis of focal infection should selectively accumulate in infection foci, be non-toxic, cost-effective, capable of distinguishing sterile inflammation from infection, and suitable for immunocompromised patients[9]. The urgent need for such an ideal agent in clinical diagnosis persists, which is currently unavailable, except for some success with SPECT and PET-based tracers[5,9–12].

An antimicrobial peptide (AMP) serving as a targeting vector for formulating an imaging agent is an attractive strategy for the specific detection of bacterial infections. Multiple groups have shown that fragments derived from Ubiquicidin originally isolated from mouse macrophages serve as a promising candidate for PET and SPECT imaging probe[5,10,13–15]. Multiple studies have demonstrated that ^{99m}Tc labeled UBI (29–41) can

successfully differentiate between bacterial infection and sterile inflammation[10,16]. Further, a recent systematic review examining the progress of ubiquicidin-based PET imaging agents provided strong case for the use of [^{68}Ga]Ga-UBI as a PET based infection imaging agent[17]. AMP derived peptide UBI (29-41) has shown promise as PET as well as SPECT based infection. The authors have previously demonstrated that ubiquicidin derived peptides selectively interact with the bacterial membrane through electrostatic interaction[18]. We hypothesize that detection sensitivity could be improved by modifying the peptide to avoid solely electrostatic dependent interaction with bacterial membrane. In this regard, the stapling of antimicrobial peptides, covalent probe installation, and non-natural amino acid incorporation are being pursued to develop a new generation of antibacterials and imaging agents for clinical translation[19].

The covalent probe 2-acetylphenyl boronic acid (2APBA) forms an iminoboronate linkage with the amine-presenting bacterial phospholipid, lysyl-phosphatidyl-glycerol (Lysyl-PG)[20] (Figure 3B.1a). Importantly, lysylation of phosphatidyl-glycerol (PG) is a widely recognized lipid mutation strategy employed by *S. aureus* to evade the binding of cationic AMPs[21]. Therefore, conjugating antimicrobial peptides with 2APBA can potentially facilitate the detection of AMP-resistant bacteria (Figure 3B.1b). The peptide UBI was finally labelled with ^{99m}Tc, and the complex was assessed in infection bearing mice to confirm its potential for in situ imaging of infection via SPECT. For another study in PET imaging, UBI was conjugated with the bifunctional chelator 1,4,7-triazacyclononane-1,4,7-triacetic acid (NOTA) and its derivative 1, 4, 7 triazacyclononane 1-glutaric acid 4–7 acetic acid (NODAGA) to form thermodynamically stable complexes with ^{68}Ga[7] (Figure 3B.1c). Moreover, these macrocyclic chelators contribute to forming hydrophilic ^{68}Ga-complexes facilitating their faster renal clearance, ultimately resulting in improved target to non-target ratios. Hence, UBI fragments conjugated to macrocyclic chelator NODAGA were used in this study for their evaluation as potential infection imaging probes[7]. Since 2APBA is attached to peptide as an acetyl-cysteine conjugate (AcCys-2APBA), the latter was used as a control in the present study.



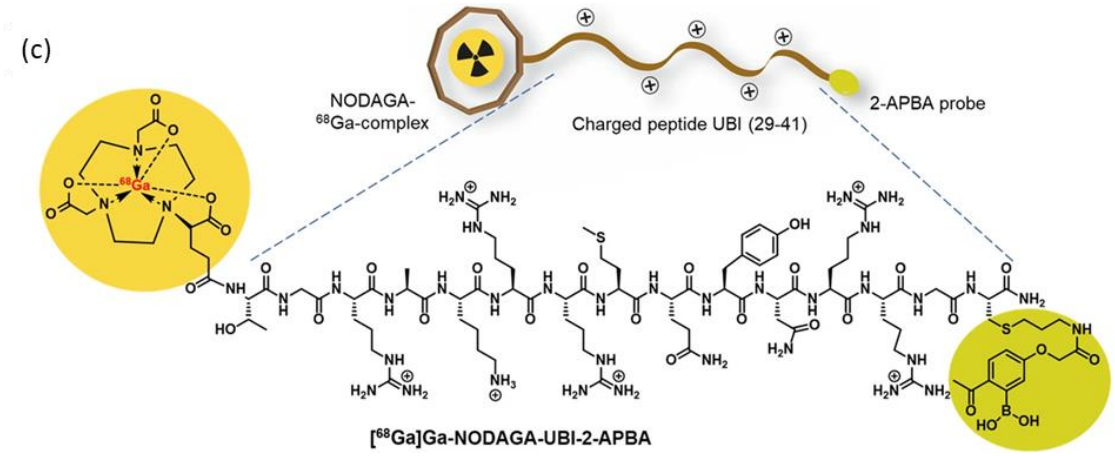


Figure 3B.1: a) A cartoon representation of synergistic interaction. Iminoboronate formation with Lysyl-PG is shown; b) Synergistic interaction of UBI-APBA for bacterial detection and membrane depolarization; c) Chemical structure of [^{68}Ga] Ga-UBI-2APBA formulation.

***Declaration*:** The radionuclide labelling to peptide and all biological studies with radionuclide labelleld peptide were performed by our collaborator at BARC, Mumbai. The key data has been highlighted to complete the outcome of chapter 3B.

3B.2 *RESULTS AND DISCUSSION*

3B.2.1. *SYNTHESIS OF PEPTIDES AND INSTALLATION OF NODAGA*

The synthesis of the peptide was carried out using the conventional Fmoc-based solid-phase peptide synthesis. Further, the 2APBA moiety was installed on a programmed cysteine moiety at the C-terminus via thiol-ene click reaction (Figure 3B.2a). An alkene handle was positioned with 2APBA moiety as per the reported procedure[22]. A glycine residue was placed between the UBI (29-41) sequence and at the C-terminus 2-APBA probe conjugated on the cysteine residue. We hypothesize that the glycine spacer would allow both functional parts to interact on the bacterial plasma membrane without interference. The UBI (29-41)-2APBA was purified through HPLC, confirmed its identity via ESI-HRMS, and further subjected to studies. As the UBI (29-41)-2APBA peptide contains a lysine residue, it is expected to form a cyclic structure through intramolecular spontaneous iminoboronate formation, which was confirmed by the UV-vis experiment (Figure 3B.2b). A clear UV-vis red shift from λ_{max} ~275 nm to ~282 nm was observed, previously reported by Bandyopadhyay et al.[20] This cyclic structure may restrict unwanted modification with lysine side chains of protein during in vivo imaging. Further, NODAGA (^{t}Bu)$_3$ was coupled on-resin at the N-terminus for ^{68}Ga applications (Figure 3B. 2c).

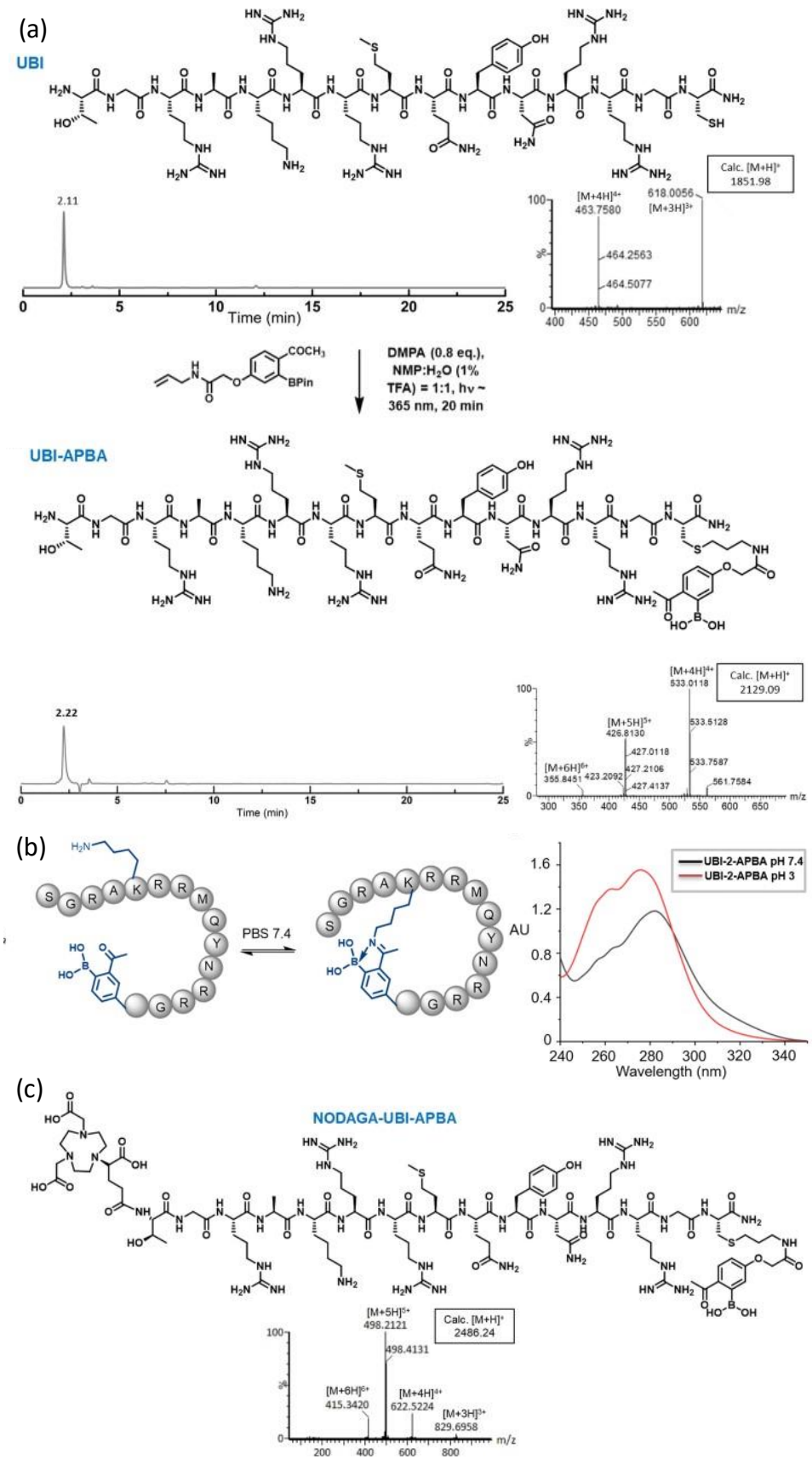


Figure 3B.2: a) Installation of 2-APBA probe on UBI (29-41) by thiol–ene click reaction; b) UBI cyclization at physiological pH through iminoboronate is demonstrated by UV-vis spectra; c) NODAGA-UBI-APBA data: its chemical structure, and mass datum confirms the product's identity.

3B.2.2. *STUDIES WITH ^{99m}UBI AND ^{99m}UBI-2APBA*

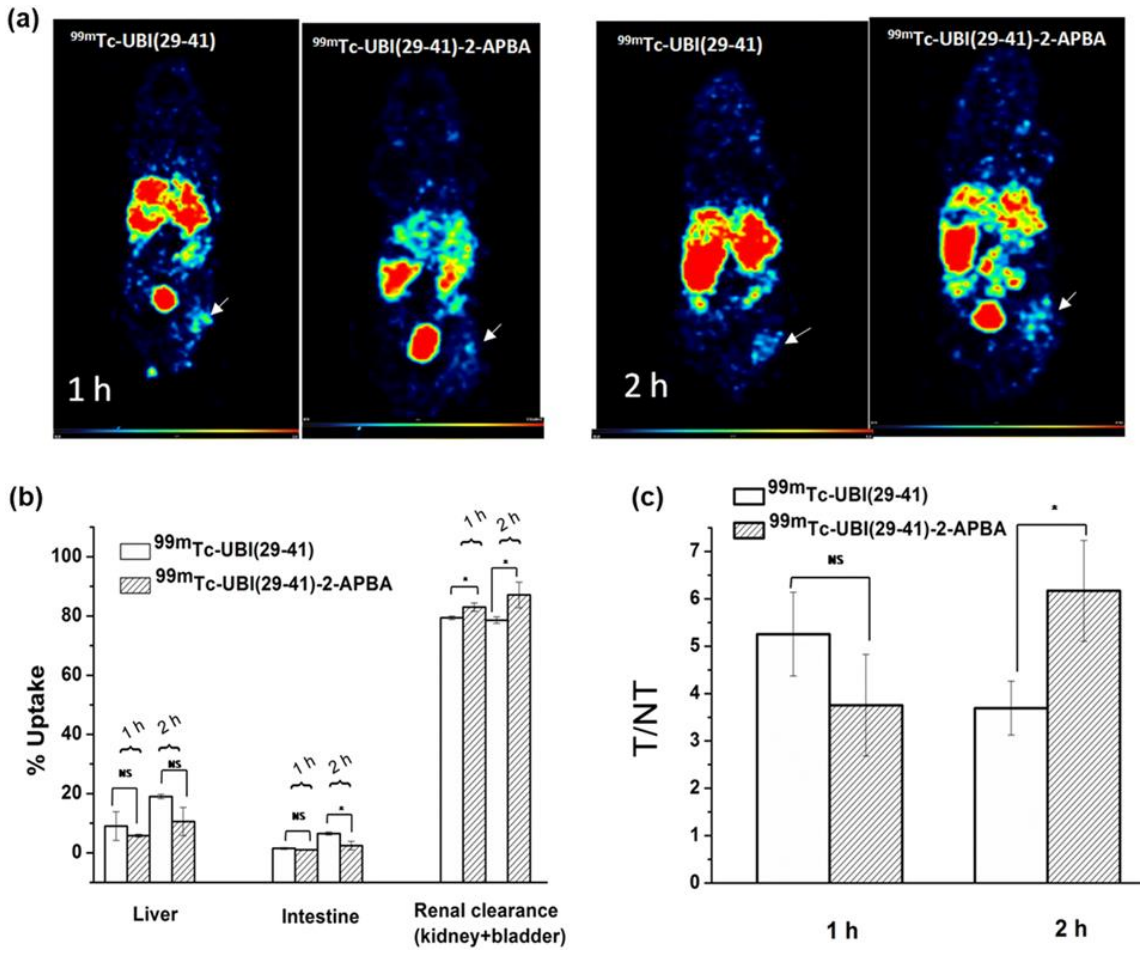


Figure 3B.3: In vivo studies a) SPECT images of mice model (Balb/c) of infection at various time points showing uptake of ^{99m}Tc-UBI (29-41) and ^{99m}Tc-UBI (29-41)-2APBA in infected (right thigh, arrow) *vs.* inflamed regions (left thigh); b) % uptake in vital organs and renal clearance; c) T/NT (target to non-target) ratio at 1 h and 2 h post injection (p.i.). * indicates $p < 0.05$.

The peptides were labelled with ^{99m}Tc at room temperature, pH ~10 using direct labeling method[23,] and all the complexes were obtained with high radiochemical purity. The serum stability of the complexes was tested before demonstrating their in-vitro uptake assay. To test the potential of ^{99m}Tc-UBI complexes to detect *S. aureus* cells in vivo, SPECT imaging was carried out in mice models of infection (right thigh muscle) and sterile inflammation (left thigh muscle). The uptake of UBI complexes was observed at the site of infection (right thigh) (Figure 3B.3). The target to non-target ratio or T/NT (ratio of tracer uptake in infected right thigh vs. inflamed left thigh) is depicted in Figure 3B.3a. It was observed that the value of T/NT for ^{99m}Tc-UBI (29-41) was maximum (T/NT: 5.2 ± 0.9) at 1 h but declined at 2 h (T/NT: 3.7 ± 0.6), indicating faster clearance from infection foci. On the other hand, the accumulation of ^{99m}Tc-UBI (29-41)-2APBA at the target site continued to improve for 2 h (T/NT: 6.2 ± 1). These observations can be attributed to the more efficient membrane–peptide interaction achieved by introducing the 2-APBA group to UBI (29-41), corroborating the data from in vitro uptake assays. Apart from the target site, the complexes were taken up by the liver, intestine, kidney, and bladder. The % uptake in these organs was compared for both complexes, and it was found that the distribution of radiotracers in the liver was not significantly different at either time point. This observation was surprising as the introduction of 2APBA to UBI (29-41) dramatically increased the lipophilicity of the

complex, as evident from the logP values. This observation could be explained by the intramolecular iminoboronate linkage formed by 2APBA with the lysine present in UBI (29-41) at physiological pH. Due to the "stapling" of the UBI (29-41)-2APBA conjugate, the newly introduced 2APBA and lysine residue in the peptide are unavailable for interaction with blood proteins. This phenomenon is further bolstered by the peptide conjugate's fast renal clearance, which was significantly better than the 99mTc-UBI (29-41) (Figure 3B.3b-c). All animal procedures were performed in accordance with the national guidelines for the care and use of laboratory animals and approved by the institutional animal ethics committee of Bhabha Atomic Research Centre, India.

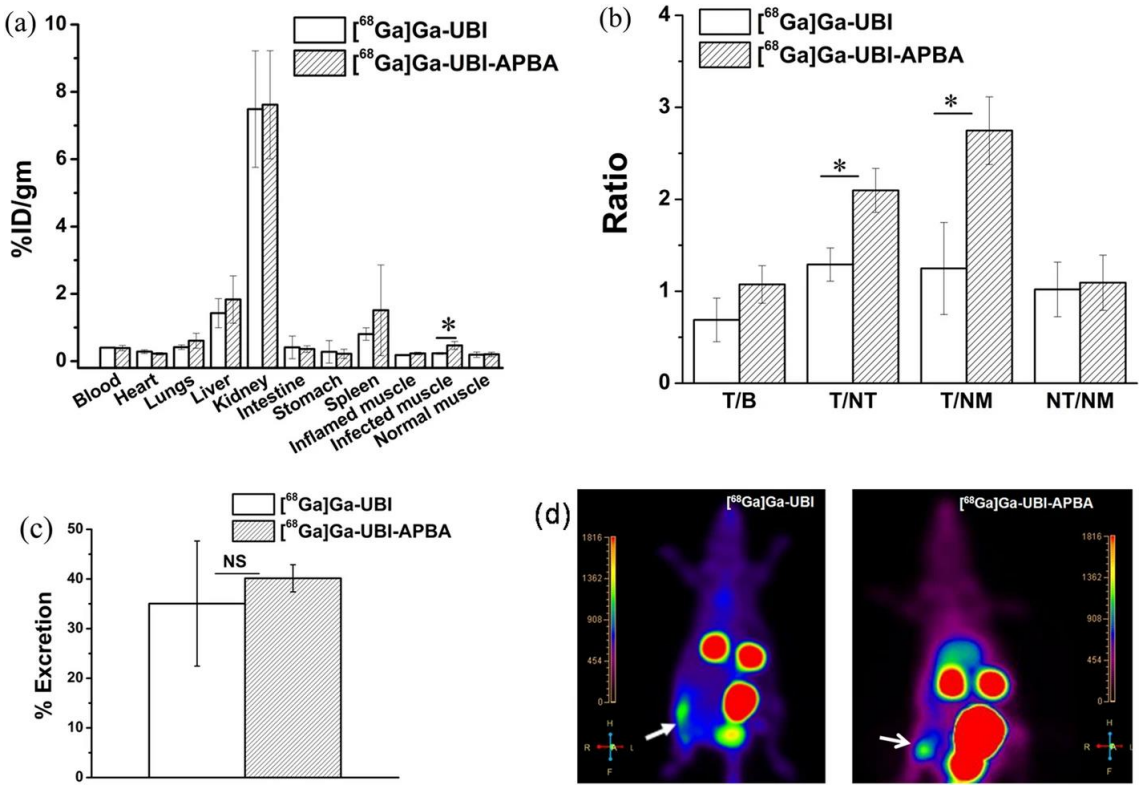


Figure 3B.4: In vivo studies in rat model of *S. aureus* infection, a) Ex-vivo biodistribution of [^{68}Ga] Ga-UBI complexes in rat model of infection and inflammation; b) Bar graph comparing the efficacy of in vivo detection of *S. aureus* by ^{68}Ga tracers. T = % ID/gm at the target or site of infection, B = Blood, NT = non-target or site of sterile inflammation, NM = Normal muscle; c) Bar graph showing % excreted activity at 1 h post injection (p.i.) for respective [^{68}Ga]Ga-UBI complexes; d) PET images showing accumulation of tracer at the site of infection (white arrow, right thigh muscle) in rat model of infection and no accumulation in sterile inflammation (left thigh muscle), images show renal clearance through kidneys and bladder. *Indicates a p value of < 0.05 ($n = 4$), NS denotes non-significant difference.

3B.2.3. *STUDIES WITH [^{68}Ga]Ga-UBI AND [^{68}Ga]Ga-UBI-2APBA*

The NODAGA attached peptides were labelled with ^{68}Ga, and their in vitro uptake assay was analyzed. We observed that the complex showed minimal uptake in splenocytes, indicating the selectivity of the complexes towards *S. aureus*, also confirmed by the toxicity studies. The ^{68}Ga complexes were assessed in rat models of infection and inflammation through ex-vivo biodistribution and PET imaging. Both tracers exhibited the

highest accumulation in kidneys, liver, and spleen, in that order (Figure 3B.4a). Accumulation in kidneys and liver suggests clearance primarily through the renal route, followed by the hepatobiliary pathway. As previously mentioned, a primary renal clearance was anticipated for these ^{68}Ga complexes due to their high hydrophilicity. In all other organs, the accumulation of tracers was minimal. The bar graph in Figure 3B.4b illustrates the accumulation of tracers at the site of infection (T = Target) in relation to blood (B), sterile inflammation (NT = Non-target), and normal muscle (NM). It is evident that [^{68}Ga]Ga-UBI-2APBA could distinguish the target site significantly better compared to [68Ga]Ga-UBI, as reflected in the improved T/NT ratio (Figure 3B.4b). Figure 3B.4c shows that up to 40% of the activity was cleared from the rats within 1 h of injection, primarily through the renal route. To summarize, ex-vivo biodistribution data established [^{68}Ga]Ga-UBI-APBA as a superior tracer for imaging *S. aureus* driven focal infection as compared to [^{68}Ga]Ga-UBI. The PET images of rats injected with [^{68}Ga]Ga-UBI tracers show that both the tracers could identify the target site as well as distinguish it from the site of sterile inflammation (Figure 3B.4d). This observation corroborated the biodistribution data. Moreover, clearance of activity largely through the renal route was observed, as evident from uptake in the kidneys and urinary bladder.

3B.3 *CONCLUSIONS*

Radiolabeled UBI derivatives have shown promising results for in situ detection of *S. aureus* driven infection by SPECT and PET imaging techniques. We observed that ^{99m}Tc-labeled UBI-2APBA and [^{68}Ga] Ga-UBI-2APBA showed significantly higher uptake compared to their unconjugated counterparts in bacterial cells. Uptake studies in mouse splenocytes, combined with our toxicity data in human erythrocytes and HEK 293 cells, suggest that 2APBA conjugation did not compromise the selectivity of UBI towards *S. aureus* cells, rather it improved. Therefore, our preliminary data confirm the potential of the UBI-2APBA peptide-covalent probe hybrid for diagnosing bacterial infections. Our study also showed that the conjugation of the covalent probe, 2APBA, to UBI, drastically impacted the growth and survival of *S. aureus* by depolarization of the bacterial membrane. Hence, we propose that the conjugation of 2APBA could effectively enhance the antibacterial activity of existing AMPs without introducing toxicity against host cells.

3B.4 *EXPERIMENTAL SECTION*

3B.4.1 *CHEMICALS, CELLS AND GENERAL METHODS*

All the chemicals were purchased from commercial sources and used without further purification. The reagents for peptide synthesis such as Fmoc-protected amino acids, 2-(1*H*-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), and polystyrene rink amide AM resin (loading capacity 0.67 mmol/g) were purchased from Chem-Impex Int'l Inc (Wood Dale, IL). *S. aureus* (ATCC 25923) and brain heart infusion broth for in-vitro studies were procured from Himedia (Mumbai, India).

All reactions were monitored by thin-layer chromatography (TLC) on aluminium-backed silica gel 60 F_{254}, 0.2 mm plates (Merck), and compounds were visualized under UV light (254 nm) or charred with appropriate solutions – ninhydrin charring to visualize amines, DNP charring for aldehydes, $KMnO_4$ charring for phenolic

OH, and curcumin or anisaldehyde charring for boronic acids. Synthetic compounds were purified on 60-120 or 100-200 silica gel purchased from Finar.

3B.4.2 *INSTRUMENTS*

NMR data were recorded on a 400 MHz Jeol JNM ECS400 NMR spectrometer. Mass-spec data were generated by a Waters XEVO G2-XS QTOF mass spectrometer. HPLC analysis and peptide purification were carried out on a Shimadzu Prominence UFLC system. UV-Vis studies were carried out on a Shimadzu UV-2600 system.

3B.4.3 *GENERAL METHODS FOR LC-MS AND HPLC ANALYSIS*

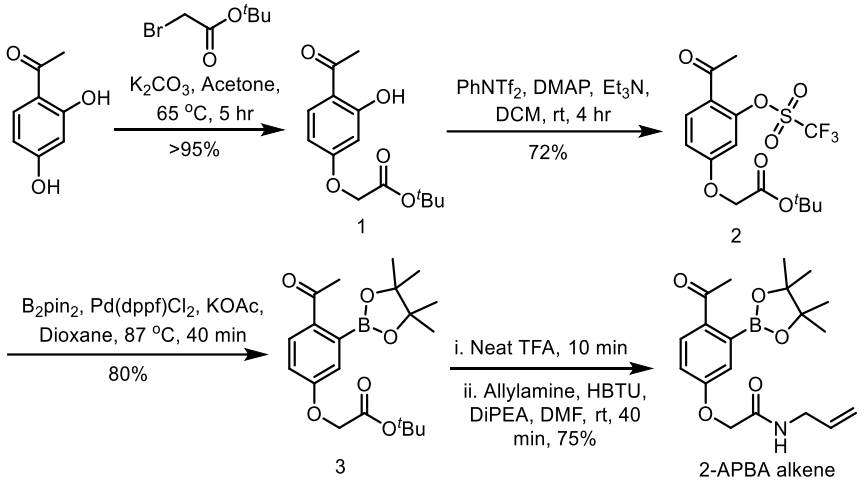
LC-MS: Waters Acquity UPLC C18 (1.7 µm, 2.1×50 mm) analytical column and the mobile phase acetonitrile-water (0.1% HCOOH) with a flow rate of 0.4 mL/min were used. The gradient: gradient 5% to 35% acetonitrile in 2.5 min, then a gradient from 35% to 95% acetonitrile in 2 min followed by isocratic 95% over a min and then column equilibration at the initial isocratic gradient 5% acetonitrile over 1.5 min. All MS analyses were performed in ESI +ve mode unless otherwise specified.

Method for analytical RP-HPLC: Waters Reliant C18 (5 µm, 4.6×250 mm) analytical column using mobile phase ACN-H_2O (0.05% TFA) with a 1 mL/min flow rate. The gradient used: isocratic 10% ACN over 1 min, then a gradient from 10% to 50% ACN in 19 min, then a gradient from 50% to 90% ACN over 4 min, followed by isocratic 90% ACN over 2 min, then column equilibration at the initial isocratic gradient 10% ACN over 5 min.

Method for semi-prep RP-HPLC: Shimadzu shim-pack GIST C18 (5 µm, 10×250 mm) semi-preparative column using mobile phase ACN-H_2O (0.05% TFA) with a flow rate 4 mL/min. The gradient used: isocratic 5% ACN over 2 min, then a gradient from 5% to 30% ACN over 26 min, then a gradient from 30% to 90% ACN over 3 min, then isocratic 90% ACN over 5 min, followed by column equilibration at the initial isocratic gradient 5% ACN over 5 min.

3B.4.4 *SYNTHESIS AND CHARACTERIZATION OF UBI (29-41)-2APBA*

3B.4.4.1. *SYNTHESIS OF 2APBA alkene*



*Synthesis of **1***

2,4-Dihydroxybenzaldehyde (1.97 g, 13 mmol), *tert*-butyl bromoacetate (2.1 mL, 14.3 mmol, 1.1 eq.) and K_2CO_3 (1.8 g, 13 mmol, 1 eq.) were suspended in 26 mL acetone and refluxed at 65 °C for 5 hr. The reaction mixture was cooled and filtered to remove K_2CO_3. The organic layer was concentrated and filtered through silica gel to obtain 3.38 g (98%) white powder.

^{1}H NMR (400 MHz, $CDCl_3$): δ 7.81 (d, *J* = 8.8 Hz, 1H), 6.91 (dd, *J* = 8.8, 2.5 Hz, 1H), 6.79 (d, *J* = 2.4 Hz, 1H), 4.56 (s, 2H), 2.56 (s, 3H), 1.46 (s, 9H).

^{13}C NMR (101 MHz, $CDCl_3$): δ 202.8, 167.1, 165.1, 164.2, 132.6, 114.6, 107.9, 101.7, 83.0, 65.5, 28.1, 26.4.

HRMS-ESI$^+$ (*m/z*): $[M+H]^+$ Calc. 267.1232, Obs. 267.1229.

Synthesis of 2

1 (3.3 g, 12.4 mmol), $PhNTf_2$ (4.87 g, 13.6 mmol, 1.1 eq.), and DMAP (151.5 mg, 1.24 mmol, 0.1 eq.) were dissolved in 23 mL dry DCM and stirred at room temperature for 5 min. Et_3N (3.5 mL, 24.8 mmol, 2 eq.) was added dropwise to the reaction mixture when the colour of the solution changed from light yellow to deep yellow. It was further stirred for 4 hr, and then DCM was evaporated. The residue was dissolved in 300 mL EtOAc, washed with 120 mL 1N HCl, 150 mL 10% Na_2CO_3 and the combined organic layers was washed with brine. Upon evaporation of the solvent, the product was purified through silica gel using 10% EtOAc/Hexane to obtain a 3.54 g of **2** (72% yield) as a white solid.

^{1}H NMR (400 MHz, $CDCl_3$): δ 7.64 (d, *J* = 8.9 Hz, 1H), 6.48 (dd, *J* = 9.0, 2.5 Hz, 1H), 6.34 (d, *J* = 2.5 Hz, 1H), 4.53 (s, 2H), 2.55 (s, 3H), 1.48 (s, 9H).

^{13}C NMR (101 MHz, $CDCl_3$): δ 195.1, 166.6, 161.8, 148.4, 132.8, 124.9, 120.2, 117.1, 113.9, 109.7, 83.4, 65.9, 29.2, 28.0.

HRMS-ESI$^+$ (*m/z*): $[M+H]^+$ Calc. 399.0725, Obs. 399.0724.

Synthesis of 3

2 (0.995 g, 2.5 mmol), $Pd(dppf)Cl_2$ (73.2 mg, 0.1 mmol, 0.04 eq.), B_2pin_2 (0.953 g, 3.75 mmol, 1.5 eq.), and KOAc (612.5 mg, 6.25 mmol, 2.5 eq.) was taken together in a 25 mL oven-dried round-bottomed flask. 15 mL of dry dioxane was added, and the mixture was purged with argon for 20 min. The reaction mixture was then heated at 87 °C for 40 min with stirring. After completion of the reaction, as indicated by TLC, the reaction mixture was filtered through a celite bed. Upon evaporation of the solvent, the residue was purified through silica gel using 12% EtOAc/Hexane to obtain **6** as 670 mg yellowish solid (80% yield).

^{1}H NMR (400 MHz, $CDCl_3$): δ 7.76 (d, *J* = 8.6 Hz, 1H), 6.92 (d, *J* = 2.6 Hz, 1H), 6.87 (dd, *J* = 8.7, 2.7 Hz, 1H), 4.55 (s, 2H), 2.54 (s, 3H), 1.46 (s, 9H), 1.41 (s, 12H).

^{13}C NMR (101 MHz, $CDCl_3$): δ 198.6, 167.4, 161.5, 134.3, 130.7, 117.8, 114.7, 83.7, 83.2, 82.9, 77.4, 77.1, 76.8, 65.7, 28.1, 25.0, 24.6.

HRMS-ESI$^+$ (*m/z*): $[M\text{-pin-}H_2O+H]^+$ Calc. 277.1247, Obs. 277.1239.

Synthesis of 2APBA alkene

670 mg of **3** was dissolved in 3 mL cold neat TFA, which formed a yellow-coloured solution. It was then stirred at room temperature for 10 min. TLC analysis indicated the completion of the reaction. Residual TFA was evaporated on a rotary evaporator with 100 mL DCM, and 216 mg of yellow solid was obtained by precipitation using 10% EtOAc/Hexane, which was used for the next step without further purification.

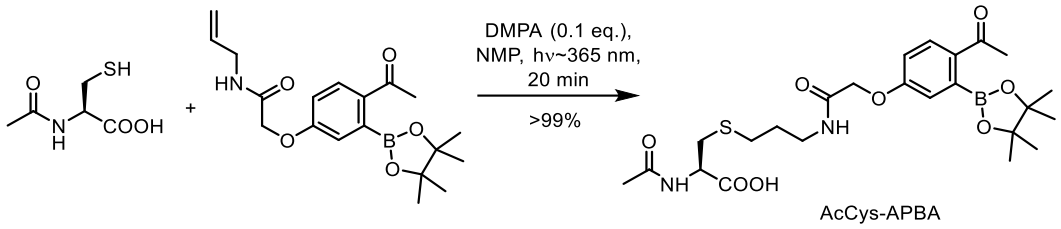
216 mg of the acid derivative (0.91 mmol) and HBTU (345 mg, 0.91 mmol, 1 eq.) were dissolved in 1 mL DMF and stirred at room temperature. To this solution, DiPEA (232 µL, 1.36 mmol, 1.5 eq.) and allylamine (75 µL, 1 mmol, 1.1 eq.) were added with a 2 min interval in stirring condition. The reaction was completed after 40 min, as indicated by TLC analysis. The reaction mixture was diluted in EtOAc (200 mL), and the organic layer was washed with 75 mL chilled 1N HCl and 100 mL chilled 5% $NaHCO_3$, followed by brine wash. After evaporation of the solvent, the residue was purified through silica gel using 40% EtOAc/Hexane to obtain 245 mg (75% yield) as a yellowish gummy solid, which solidified on refrigeration.

^{1}H NMR (400 MHz, 5% $CD_3OD/CDCl_3$): δ 7.81 (d, *J* = 8.6 Hz, 1H), 7.01 (d, *J* = 2.6 Hz, 1H), 6.89 (dd, *J* = 8.5, 2.6 Hz, 1H), 6.59 (brs, 1H), 5.89-5.79 (m, 1H), 5.23 – 5.14 (m, 2H), 4.57 (s, 2H), 3.99 – 3.96 (m, 2H), 2.56 (s, 3H), 1.43 (s, 12H).

^{13}C NMR (101 MHz, 5% $CD_3OD/CDCl_3$): δ 168.0, 163.2, 160.7, 134.5, 133.3, 131.1, 118.5, 116.8, 113.9, 84.0, 66.9, 38.5, 24.8.

HRMS-ESI$^+$ (*m/z*): $[M\text{-pin-}H_2O+H]^+$ Calc. 260.1094, Obs. 260.1104.

3B.4.4.2. *Synthesis of AcCys-2APBA*



AcCys-OH (220 mM, 1.1 eq), APBA (200 mM, 1eq), and DMPA (20 mM) were dissolved in NMP (1% TFA) and stirred for 20 min in the presence of hν ~365 nm light source in a dark place. The reaction mixture was diluted to 1 mL with 1 N HCl and washed with 2 mL EtOAc. The crude residue was purified through RP-HPLC, and the NMR data for the same are presented.

Method for analytical HPLC (Shimadzu Prominence UFLC): Waters Sunfire C18 (5 µm, 4.6×250 mm) analytical column using mobile phase ACN-H_2O (0.05% TFA) with a flow rate of 1 mL/min. The gradient used: isocratic 10% ACN for 1 min, then a gradient from 10% to 90% ACN in 19 min, then isocratic 90% ACN for a min, then column equilibration at an initial isocratic gradient 10% ACN over 4 min.

ESI-MS analysis: Calc. $[M\text{-pin-}H_2O+H]^+$ 423.14, Obs. 423.13

Analytical RP-HPLC: R_t = 6.99 min

3B.4.5.3 *SYNTHESIS OF PEPTIDES*

Peptide synthesis was carried out using a standard SPPS Fmoc-deprotection strategy on Rink amide resin. Three equivalents of the commercially available amino acids, HBTU as a coupling agent and DiPEA as a base were used for the coupling reaction for 15 min. Fmoc deprotection was achieved using 3 mL 20% piperidine/DMF twice for 3 min each, followed by washing with 3 mL DMF six times. NODAGA-(^{t}Bu)$_3$ was coupled to the N-terminal. The peptides were cleaved off the resin and globally deprotected with reagent K (82.5% TFA, 5% H_2O, 2.5% EDT, 5% Thioanisole and 5% phenol). Cleavage protocol for the NODAGA derivative required repeated deprotection steps at elevated temperatures (50 °C). Precipitation with chilled diethyl ether gave the crude peptides, which were purified by RP-HPLC. The mass and purity of the peptide were determined by using LC-MS (Waters) and HPLC (Shimadzu) to be >95%.

ESI-MS analysis: Calc. $[M+3H]^{3+}$ 618.40, Obs. 618.01

Analytical RP-HPLC for UBI-2APBA: R_t = 2.1 min

Analytical RP-HPLC for NODA-UBI-2APBA: R_t = 2.1 min

Installation of 2-APBA alkene on UBI (29-41)

2-APBA alkene was installed on the C-terminal Cys residue of UBI by thiol-ene click reaction. 10 mM peptide, 12 mM 2-APBA alkene and 8 mM DMPA were dissolved in 200 µL1:1 NMP: H_2O (1% TFA) in a 1.5 mL clear glass vial. It was kept in the center between two Philips TL 8W BLB light sources, 7 cm apart with λ~365 nm for 20 min. The reaction mixture was then diluted with 800 µL 20% ACN/H_2O (0.05% TFA), filtered through a 0.22 µm nylon filter, and subjected to purification by RP-HPLC.

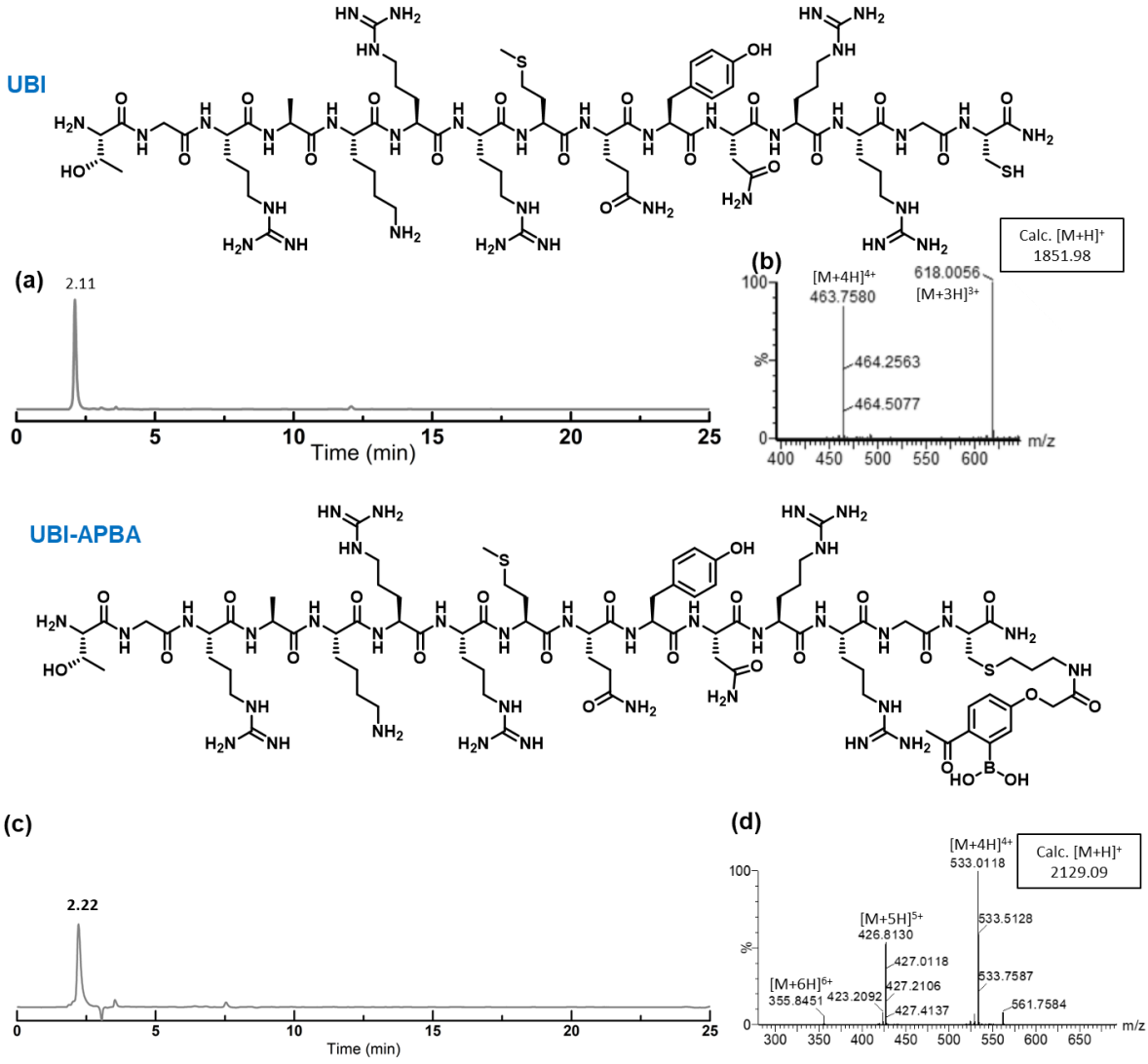
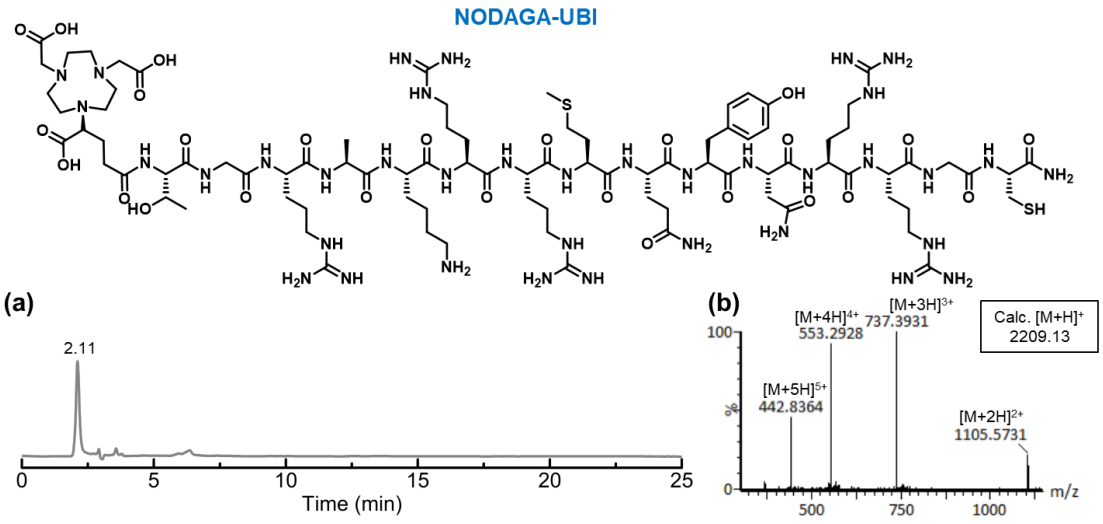


Figure 3B.5: UBI peptide data: a) HPLC purity at 220 nm; b) Mass data; UBI-APBA peptide data: c) HPLC chromatogram at 220 nm showing >95% purity of the product; d) Mass datum confirms the product's identity.



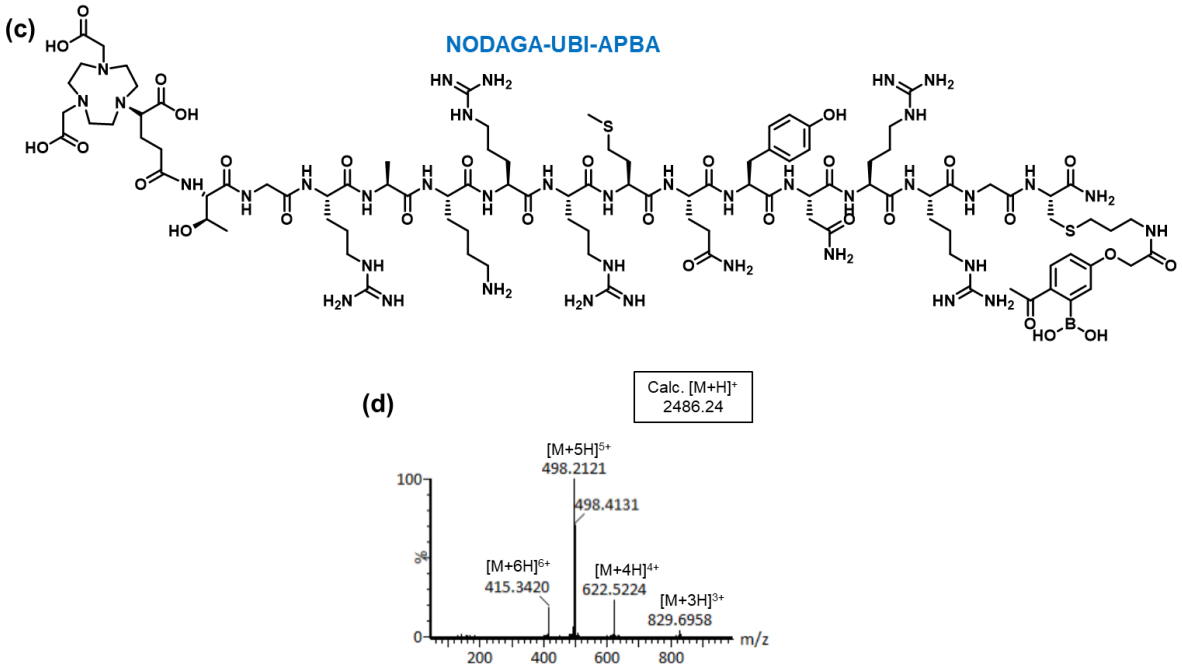
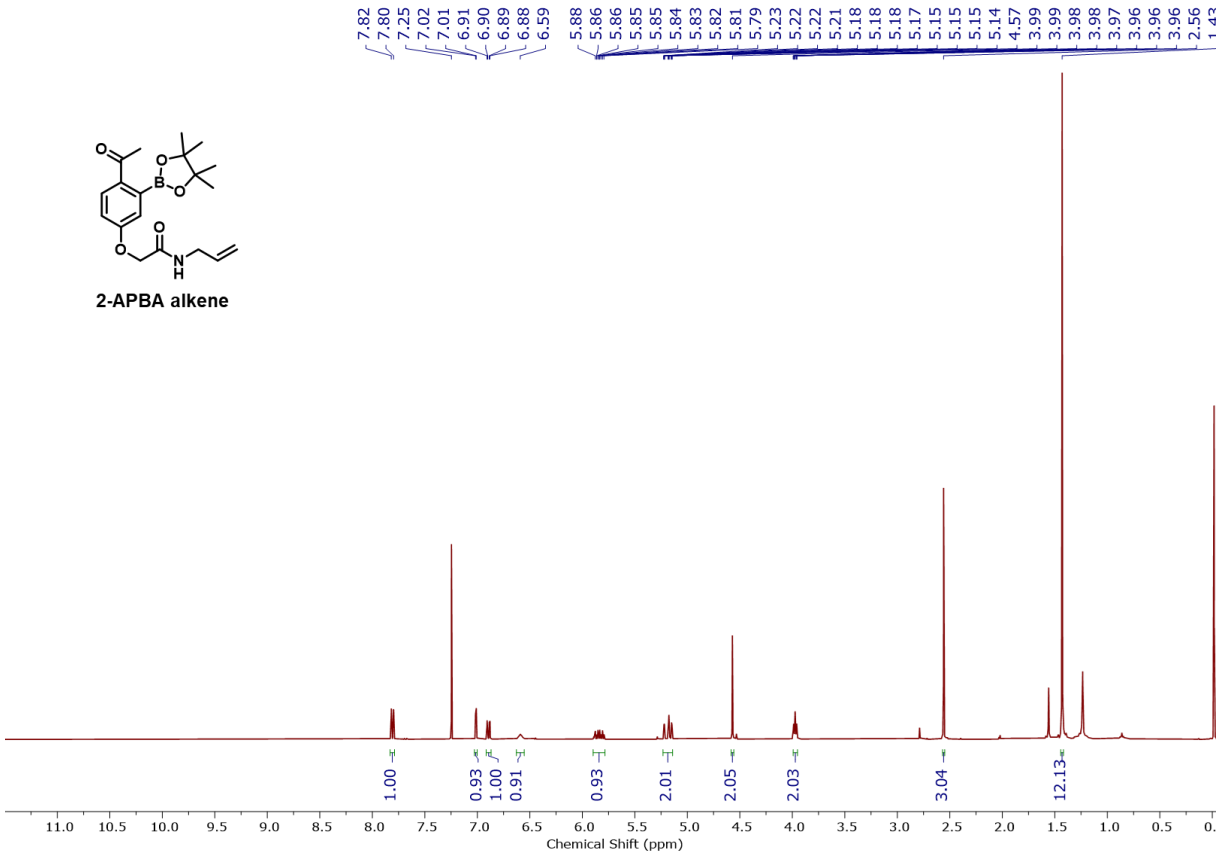


Figure 3B.6: Chemical structure of NODAGA-UBI: a) HPLC chromatogram determines purity at 220 nm; b) Mass datum confirms peptide's identity. (c) Chemical structure of NODAGA-UBI-APBA and its identity in d) mass spectrometry.

3B.5 *NMR DATA*



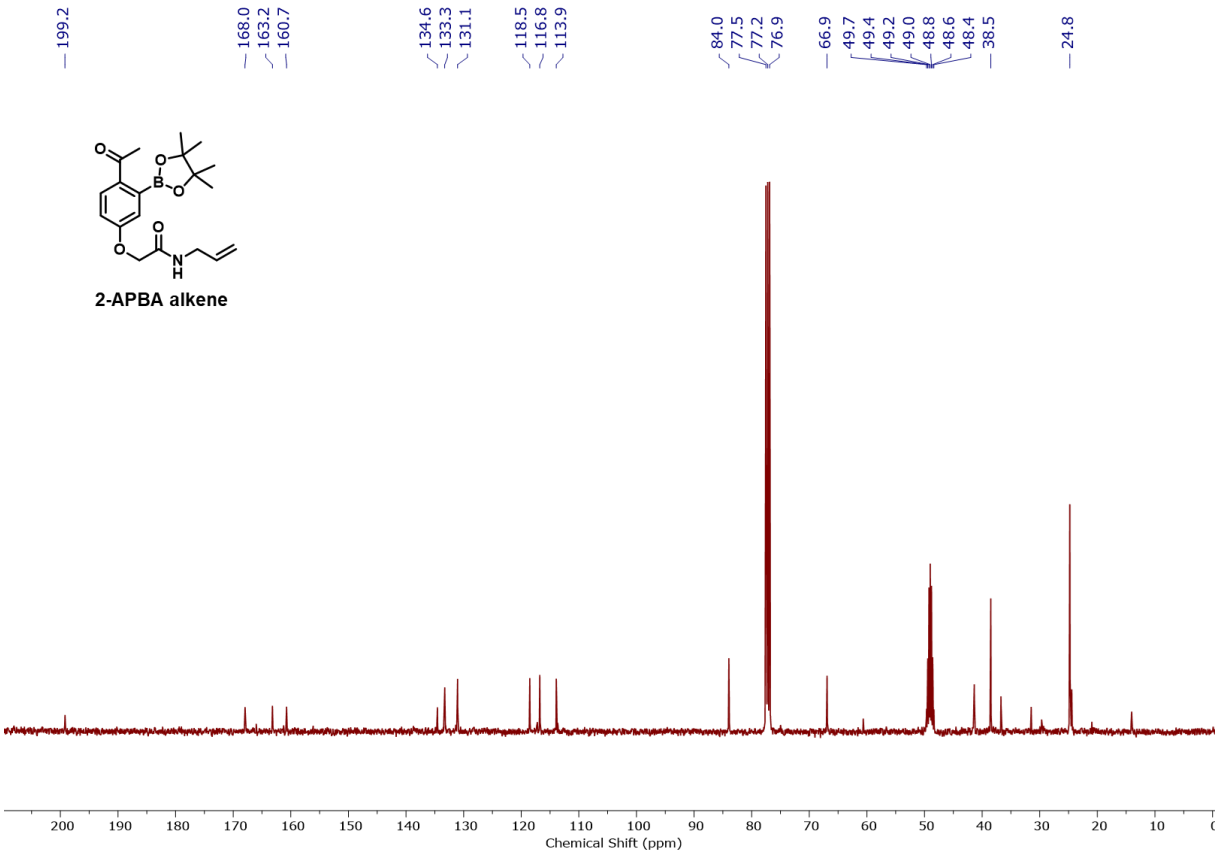


Figure 3B.7: ^{1}H and ^{13}C NMR spectra of compound 2-APBA alkene in 5% $CD_3OD/CDCl_3$.

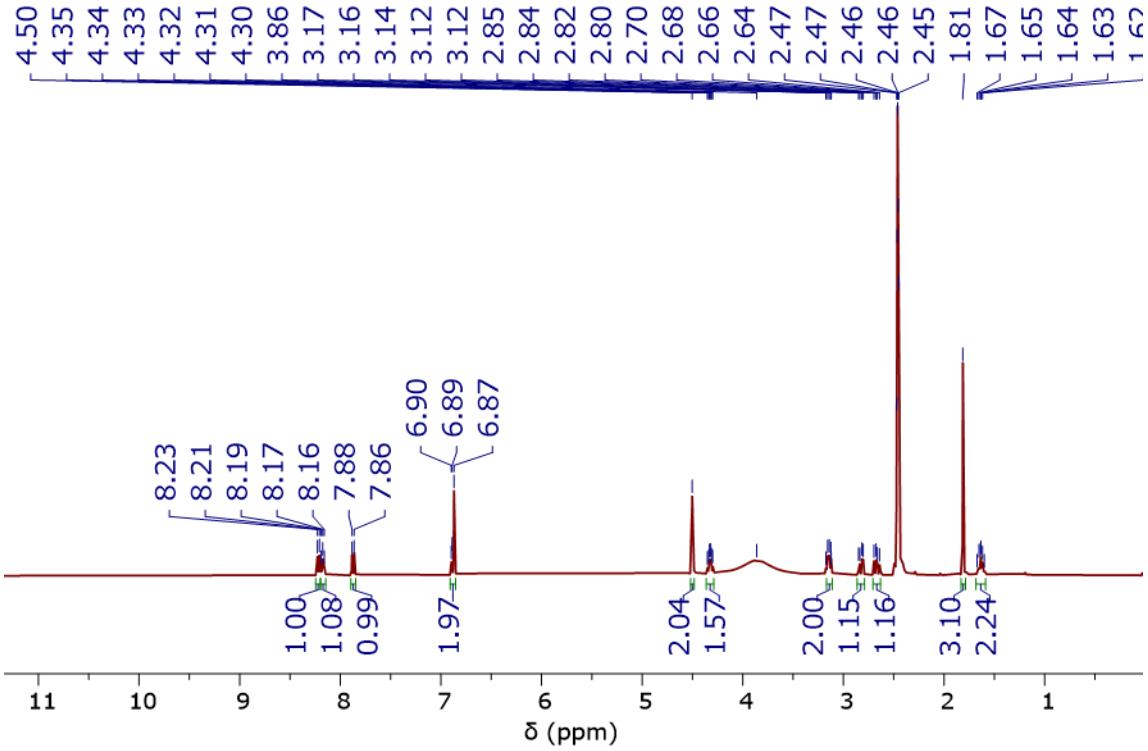


Figure 3B.8: a) NMR spectrum of pure AcCys-2-APBA in DMSO-d_6.

3B.6 *REFERENCES*

(1) Rudd, K. E.; Johnson, S. C.; Agesa, K. M.; Shackelford, K. A.; Tsoi, D.; Kievlan, D. R.; Colombara, D. V.; Ikuta, K. S.; Kissoon, N.; Finfer, S.; Fleischmann-Struzek, C.; Machado, F. R.; Reinhart, K. K.; Rowan, K.; Seymour, C. W.; Watson, R. S.; West, T. E.; Marinho, F.; Hay, S. I.; Lozano, R.; Lopez, A. D.; Angus, D. C.; Murray, C. J. L.; Naghavi, M. Global, Regional, and National Sepsis Incidence and Mortality, 1990–2017: Analysis for the Global Burden of Disease Study. *The Lancet* **2020**, *395* (10219), 200–211. https://doi.org/10.1016/S0140-6736(19)32989-7.

(2) Ranjbar, R.; Alam, M. Antimicrobial Resistance Collaborators (2022). Global Burden of Bacterial Antimicrobial Resistance in 2019: A Systematic Analysis. *Evidence Based Nursing* **2024**, *27* (1), 16 LP – 16. https://doi.org/10.1136/ebnurs-2022-103540.

(3) Walsh, T. R.; Gales, A. C.; Laxminarayan, R.; Dodd, P. C. Antimicrobial Resistance: Addressing a Global Threat to Humanity. *PLoS Med* **2023**, *20* (7), e1004264.

(4) van Oosten, M.; Hahn, M.; Crane, L. M. A.; Pleijhuis, R. G.; Francis, K. P.; van Dijl, J. M.; van Dam, G. M. Targeted Imaging of Bacterial Infections: Advances, Hurdles and Hopes. *FEMS Microbiol Rev* **2015**, *39* (6), 892–916. https://doi.org/10.1093/femsre/fuv029.

(5) Ebenhan, T.; Sathekge, M. M.; Lengana, T.; Koole, M.; Gheysens, O.; Govender, T.; Zeevaart, J. R. 68Ga-NOTA-Functionalized Ubiquicidin: Cytotoxicity, Biodistribution, Radiation Dosimetry, and First-in-Human PET/CT Imaging of Infections. *Journal of Nuclear Medicine* **2018**, *59* (2), 334 LP – 339. https://doi.org/10.2967/jnumed.117.200048.

(6) Polvoy, I.; Flavell, R. R.; Rosenberg, O. S.; Ohliger, M. A.; Wilson, D. M. Nuclear Imaging of Bacterial Infection: The State of the Art and Future Directions. *Journal of Nuclear Medicine* **2020**, *61* (12), 1708 LP – 1716. https://doi.org/10.2967/jnumed.120.244939.

(7) Rakesh, K.; Sandip, B.; Drew, T.; Vivek, A.; Hongming, Z.; Abass, A. Role of Modern Imaging Techniques for Diagnosis of Infection in the Era of 18F-Fluorodeoxyglucose Positron Emission Tomography. *Clin Microbiol Rev* **2008**, *21* (1), 209–224. https://doi.org/10.1128/cmr.00025-07.

(8) Signore, A.; Artiko, V.; Conserva, M.; Ferro-Flores, G.; Welling, M. M.; Jain, S. K.; Hess, S.; Sathekge, M. Imaging Bacteria with Radiolabelled Probes: Is It Feasible? *Journal of Clinical Medicine*. 2020. https://doi.org/10.3390/jcm9082372.

(9) Salmanoglu, E.; Kim, S.; Thakur, M. L. Currently Available Radiopharmaceuticals for Imaging Infection and the Holy Grail. *Semin Nucl Med* **2018**, *48* (2), 86–99. https://doi.org/https://doi.org/10.1053/j.semnuclmed.2017.10.003.

(10) de Murphy, C. A.; Gemmel, F.; Balter, J. Clinical Trial of Specific Imaging of Infections. *Nucl Med Commun* **2010**, *31* (8).

(11) Bhatt, J.; Mukherjee, A.; Shinto, A.; Koramadai Karuppusamy, K.; Korde, A.; Kumar, M.; Sarma, H. D.; Repaka, K.; Dash, A. Gallium-68 Labeled Ubiquicidin Derived Octapeptide as a Potential Infection Imaging Agent. *Nucl Med Biol* **2018**, *62–63*, 47–53. https://doi.org/https://doi.org/10.1016/j.nucmedbio.2018.04.003.

(12) Mukherjee, A.; Bhatt, J.; Shinto, A.; Korde, A.; Kumar, M.; Kamaleshwaran, K.; Joseph, J.; Sarma, H. D.; Dash, A. 68Ga-NOTA-Ubiquicidin Fragment for PET Imaging of Infection: From Bench to Bedside. *J Pharm Biomed Anal* **2018**, *159*, 245–251. https://doi.org/https://doi.org/10.1016/j.jpba.2018.06.064.

(13) Vallejo, E.; Martinez, I.; Tejero, A.; Hernandez, S.; Jimenez, L.; Bialostozky, D.; Sanchez, G.; Ilarraza, H.; Ferro-Flores, G. Clinical Utility of 99mTc-Labeled Ubiquicidin 29–41 Antimicrobial Peptide for the Scintigraphic Detection of Mediastinitis after Cardiac Surgery. *Arch Med Res* **2008**, *39* (8), 768–774. https://doi.org/https://doi.org/10.1016/j.arcmed.2008.09.002.

(14) Shinto, A. S.; Mukherjee, A.; Karuppusamy, K. K.; Joseph, J.; Bhatt, J.; Korde, A.; Upadhya, I.; Arjun, C.; Samuel, G.; Dash, A. Clinical Utility of 99mTc-Ubiquicidin (29–41) as an Adjunct to Bone Scan in Differentiating Infected versus Noninfected Loosening of Prosthesis before Revision Surgery. *Nucl Med Commun* **2017**, *38* (4).

(15) Sathekge, M.; Garcia-Perez, O.; Paez, D.; El-Haj, N.; Kain-Godoy, T.; Lawal, I.; Estrada-Lobato, E. Molecular Imaging in Musculoskeletal Infections with 99mTc-UBI 29-41 SPECT/CT. *Ann Nucl Med* **2018**, *32* (1), 54–59. https://doi.org/10.1007/s12149-017-1219-7.

(16) Welling, M. M.; Paulusma-Annema, A.; Balter, H. S.; Pauwels, E. K. J.; Nibbering, P. H. Technetium-99m Labelled Antimicrobial Peptides Discriminate between Bacterial Infections and Sterile Inflammations. *Eur J Nucl Med* **2000**, *27* (3), 292–301. https://doi.org/10.1007/s002590050036.

(17) Marjanovic-Painter, B.; Kleynhans, J.; Zeevaart, J. R.; Rohwer, E.; Ebenhan, T. A Decade of Ubiquicidin Development for PET Imaging of Infection: A Systematic Review. *Nucl Med Biol* **2023**, *116–117*, 108307. https://doi.org/https://doi.org/10.1016/j.nucmedbio.2022.11.001.

(18) Bhatt Mitra, J.; Sharma, V. K.; Mukherjee, A.; Garcia Sakai, V.; Dash, A.; Kumar, M. Ubiquicidin-Derived Peptides Selectively Interact with the Anionic Phospholipid Membrane. *Langmuir* **2020**, *36* (1), 397–408. https://doi.org/10.1021/acs.langmuir.9b03243.

(19) Migoń, D.; Neubauer, D.; Kamysz, W. Hydrocarbon Stapled Antimicrobial Peptides. *Protein J* **2018**, *37* (1), 2–12. https://doi.org/10.1007/s10930-018-9755-0.

(20) Bandyopadhyay, A.; McCarthy, K. A.; Kelly, M. A.; Gao, J. Targeting Bacteria via Iminoboronate Chemistry of Amine-Presenting Lipids. *Nat Commun* **2015**, *6*, 1–9. https://doi.org/10.1038/ncomms7561.

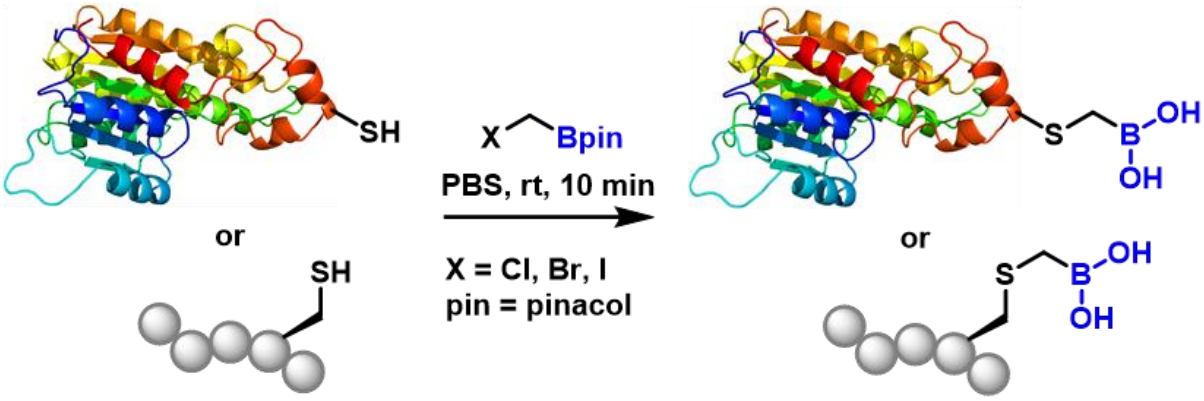
(21) Roy, H. Tuning the Properties of the Bacterial Membrane with Aminoacylated Phosphatidylglycerol. *IUBMB Life* **2009**, *61* (10), 940–953. https://doi.org/https://doi.org/10.1002/iub.240.

(22) Chatterjee, S.; Bandyopadhyay, A. Cysteine-Selective Installation of Functionally Diverse Boronic Acid Probes on Peptides. *Org Lett* **2023**, *25* (13), 2223–2227. https://doi.org/10.1021/acs.orglett.3c00386.

(23) Arjun, C.; Mukherjee, A.; Bhatt, J.; Chaudhari, P.; Repaka, K. M.; Venkatesh, M.; Samuel, G. Studies on Batch Formulation of a Kit for the Preparation of the 99mTc-Ubiquicidin (29-41): An Infection Imaging Agent. *Applied Radiation and Isotopes* **2016**, *107*, 8–12. https://doi.org/https://doi.org/10.1016/j.apradiso.2015.09.009.

CHAPTER 4

GETTING BORONATED TO PEPTIDES AND PROTEINS: A CATALYST AND METAL-FREE APPROACH



- Fast kinetics (10^2 M^{-1} s^{-1})
- No metals or additives
- ₹ Commercially available reagents
- Clean to quantitative conversion

"Electron poor but chemically rich element with a storied past continues to surprise with a surge of new developments". *- On boron, by Stephen Ritter, 2016*

4.1 *INTRODUCTION*

With the ongoing discussion in Chapter 1 and 2 on the importance of boronopeptides, it becomes imperative to develop benign or greener methods to install boronic acid derivatives on biomolecules. As opposed to peptides, proteins are chemically more sensitive to small environmental changes owing to the presence of tertiary or quaternary structures. Hence, a complete aqueous solution, physiological pH, and optimum temperature is essential when designing protein modifications in a normal laboratory setting. An efficient conjugation strategy should not alter the native conformation of a biomolecule, which, in part, can be achieved by the insertion of 'minimal functional groups'.

With a vacant *p*-orbital, boron is uniquely capable of demonstrating different dynamic bonding natures that are not available with the common elements of life such as carbon, oxygen, nitrogen, phosphorous, or sulfur. In response to the presence of endogenous nucleophiles, boron can change its bonding nature from sp^2 to sp^3 via Lewis acid-base pairing. Interestingly, ligand binding abilities can be achieved even in the presence of large amounts of water as the solvent, acting as a Lewis base. The wide variety of binding modes can also lead to tunable de novo function, proteo- or thermolytic stability, or redox-responsive mutation[1]. Boronic acids are also considered as the bioisostere for carboxylic acids[2] with undeniable roles in medicinal chemistry[3]. Due to the interaction of boron with amines and carboxylic acids, synthesis and isolation of alkylborono amino acid comes with its own challenges, thus prompting scientists to look for 'late-stage installation' techniques. The previous reports on the programmable insertion of boronic acids to biomolecules involve the synthesis of non-natural boronic acid containing amino acid[4], de novo synthesis of proteins containing C-B bonds[5], use of metals such as Ni, Cu[1,6,7], or pre-requisition of dehydroalanine (Dha) for installation[1]. Very often, it is observed that such methods also employ the use of short wavelength light with higher concentrations of peptides[8] and lower conversion yields with undesirable side-product formation. Such methods come with their own challenges – expensive and cumbersome synthesis of unnatural boronic acid amino acids, inhibitory effect of boron on protein translation which hinders the expansion of genetic code[9], and the loss in stereochemistry from Dha mediated installation (Figure 4.1a). The lack of a simple method to form C-B bonds often acts as a roadblock to the exploration of boronic acid chemistry in chemical biology.

Taking cue from the widely used haloamides for the alkylation of Cys in aqueous solution, we hypothesized that such a simple S_N2 reaction can be extrapolated to install the isosteric boronic acids. In this regard, we turned to cheap and commercially available halomethyl boronic acids (hBA), popularized in organic asymmetric synthesis by Matteson[10], to fill in for the role of a 'very simple installation strategy'. We hypothesized that the presence of a neighboring vacant *p*-orbital might influence the behavior of the electrophilic methylene carbon. Observation of the LUMO orbitals of the hBA (exemplified by iodo derivative) showed an orbital overlap of methylene and boron, which was absent in the case of iodoacetamide as shown in Figure 4.1b. This prompted us to go ahead with the reaction strategy and we developed a green method to install methyl boronic acids on programmed cysteine residues in peptides and proteins, thus presenting a cost-effective and operationally simple procedure of modification with boronic acid, eliminating the need for catalysts, additives, or metals.

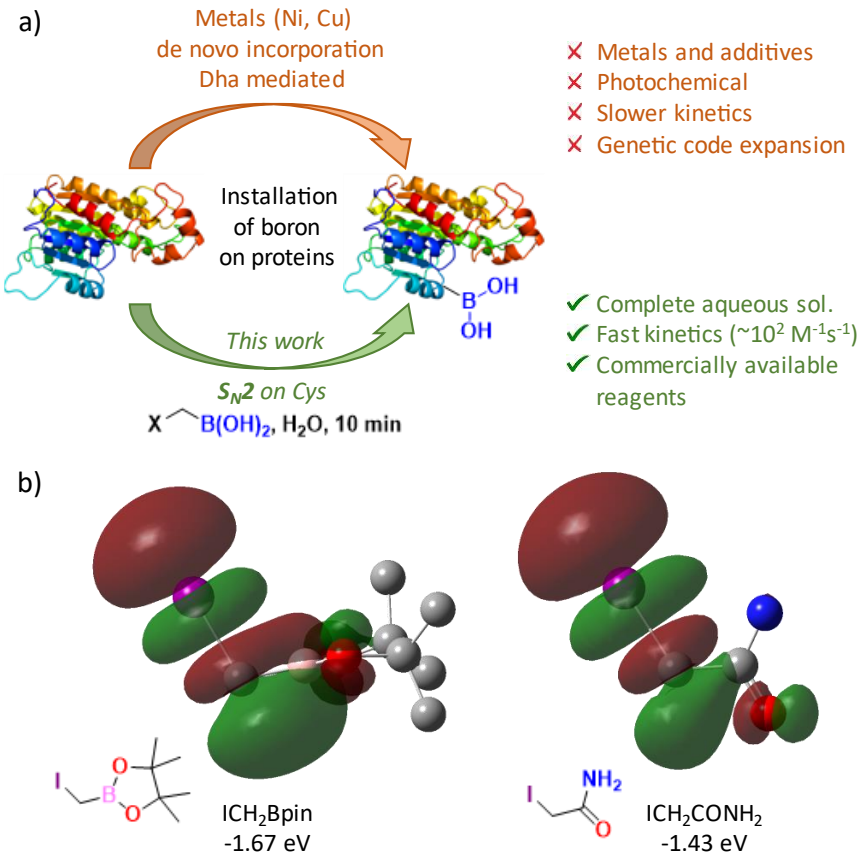


Figure 4.1. a) Comparison of installation strategies of boryl group on proteins; b) Theoretical calculation comparing the LUMO orbitals of ICH_2Bpin and iodoacetamide (ICH_2CONH_2). An orbital overlap between boron and methylene carbon can be observed for the boron reagent.

4.2 *RESULTS AND DISCUSSION*

To optimize the reaction conditions, the leucine enkephalin (LeuEnk, YGGFLC) peptide was chosen as the model sequence. Chloro, bromo, and iodo derivatives of methyl boronic acid (1.5 eq.) were reacted with LeuEnk (100 µM) in PBS buffer pH 7.4 at room temperature, and their kinetics were measured. As expected, the chloro derivative showed the slowest kinetics (6.5 M^{-1} s^{-1}), reaching about 80% product conversion in 4 hr. The bromo and iodo derivatives revealed a >95% conversion within 5 min and 2 min respectively (10^2 M^{-1} s^{-1}) as shown in Figure 4.2a. It was also observed that the iodo derivative led to bis-alkylation at the same time due its higher reactivity, hence offered dehydroalanine side products. Therefore, due to the optimum reactivity of the bromo derivative, it was considered as the reagent of choice for the rest of the studies. The borono-LeuEnk (LeuEnk-BA) was purified by RP-HPLC (Figure 4.2b) and confirmed by LC-MS, MS/MS and NMR studies. Loss of water molecules, typical in the mass spectrometric pattern of boronic acids was also observed[8,11]. ^{11}B NMR showed a trigonal (sp^2) boronic acid species[12] (Figure 4.2c). In ^1H NMR and ^1H-^1H COSY studies, shifts in the Cys β-CH_2 and α-NH protons were observed, which clearly indicated the chemoselective installation on Cys residue (Figure 4.2d).

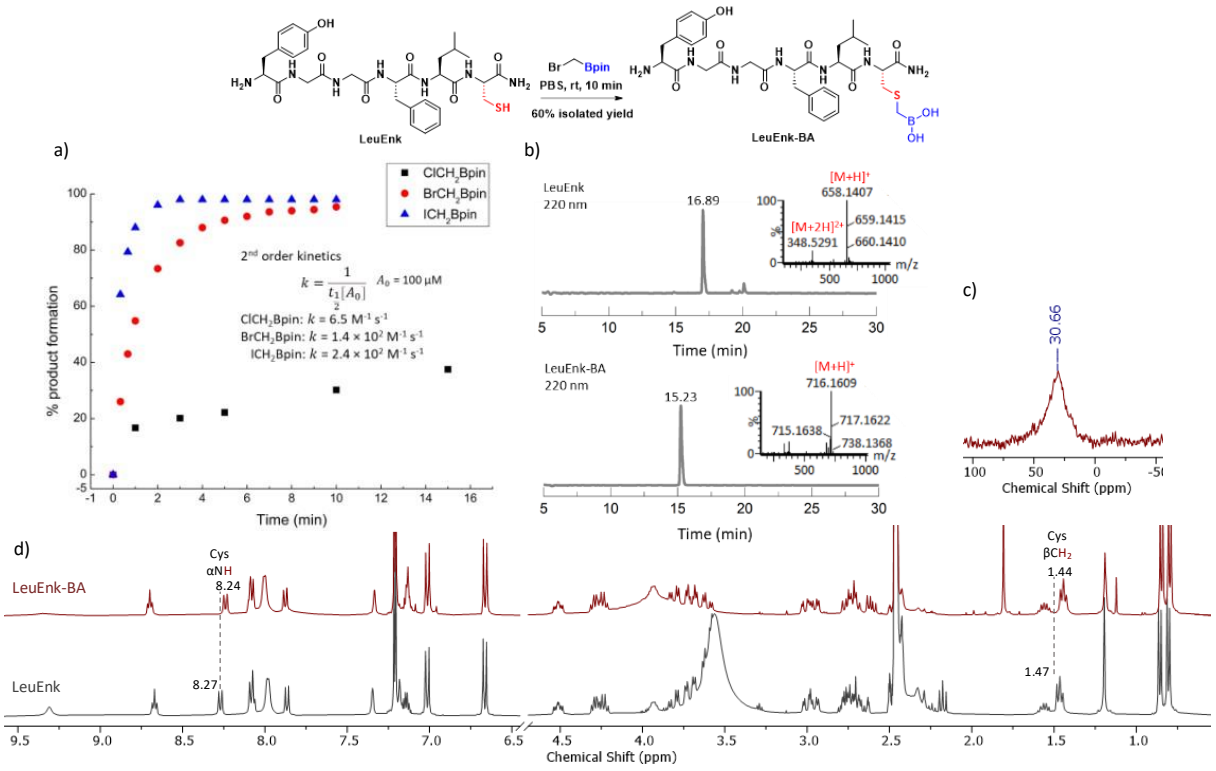


Figure 4.2: General scheme for the reaction of $BrCH_2Bpin$ with LeuEnk and a) Confirmation by HPLC and mass spec; b) Kinetic studies with chloro, bromo and iodo derivatives of halomethyl BA; c) ^{11}B NMR of 6 mM LeuEnk-BA in DMSO-d_6; and d) ^{1}H NMR of 6 mM LeuEnk stacked with LeuEnk-BA in DMSO-d_6 showing a clear shift in the protons of Cys αNH and αCH_2.

pH dependent studies revealed that the reaction at pH 6 was slightly compromised at 90% conversion within 30 min and remained so till 1 hr. Reactions at pH 7.4 and 8.5 showed similar conversion rates. To find the optimum concentration of the reagent required for a clean conversion, different concentrations ranging from 100 µM (equimolar) to 500 µM (5 eq.) were analyzed and it was found that a 150 µM (1.5 eq.) concentration of bromomethyl boronopinacol ester gave the cleanest conversion. It is not surprising that a higher concentration of activated haloalkyl agents could always lead to over-alkylation. A consistent result was also obtained with HEPES and Tris buffers, indicating that the bioconjugation method is irresponsive to standard buffer counterions in the reaction. The reaction strategy was also compared with other commonly applied Cys alkylating agents such as iodoacetic acid (an isostere of halomethyl boronic acid), iodoacetamide, and benzyl iodide. Despite an extended reaction time (1 hr), keeping the other conditions same, other than iodoacetamide which showed a ~40% conversion, the other two reagents failed to show any detectable conjugated product. This clearly demonstrates the superiority of the halo methyl boronic acid reagents with respect to alkylation of Cys at neutral conditions. We also wanted to check the effectiveness of the reaction in presence of disulfide reducing agents such as tris(2-carboxyethyl)phosphine, TCEP, since it is reported to react with halo derivatives to some extent[13,14]. Given the higher reactivity of halomethyl boronic acids, it was not surprising to see that indeed TCEP quenches the halo derivatives and a higher concentration of the reagent is required for attaining a similar conversion. In the presence of 1 mM TCEP (a concentration commonly used for reducing proteins)[15], 750 µM of the reagent was required to achieve the similar conversion. A ^{1}H NMR study of 1:1 mixture of TCEP and $BrCH_2Bpin$ in 30% D_2O/PBS pH 7.4 also showed a clean shift of the methylene protons of $BrCH_2Bpin$. Hence, in reactions use of TCEP, a higher concentration of the boronic acid reagent

can be used since procedures involving removal of TCEP from protein mixtures such as dialysis is very time-consuming and labour intensive[16].

We then delved into the substrate scope of the reaction and subjected to 11 different peptides with the optimised methodology (Figure 4.3). The peptides ranged from short (4-mer) to long (36-mer) chains with the 20-natural amino acids designed in hydrophilic as well as hydrophobic sequences. An excellent conversion (>90%) was observed with all the peptide sequences. In substrate scopes, we also checked the reaction chemoselectivity in presence of α-nucleophiles, since a hydrazide moiety at the C-terminal is often used in native chemical ligation (NCL) strategy. Gratifyingly, a clean conversion was observed for the SFLLR peptide. We then proceeded to synthesize CS and BCS peptides[7], and cyclized them via NCL. We also modified the well-known lanreotide peptide (LAN), an analogue of somatostatin that contains a disulfide linkage. First, the disulfide was reduced by treating with 1 mM TCEP for 1 hr, followed by 750 µM of the boronic acid reagent, which yielded a clean conversion to the disubstituted product. A 20 µM concentration of the p53 peptide sequence containing some of the well-known reactive side chains was successfully modified at the Cys residue, as confirmed by MS/MS analysis. Overall, the designed methodology was selective towards the Cys residue and did not depend on the hydrophobicity of the peptide sequence or the amino acid side-chains. A quantitative conversion was observed with all the different peptides ranging from 20-100 µM. Though the reagent was quenched with TCEP to some extent, but a higher equivalent provided similar conversion yields.

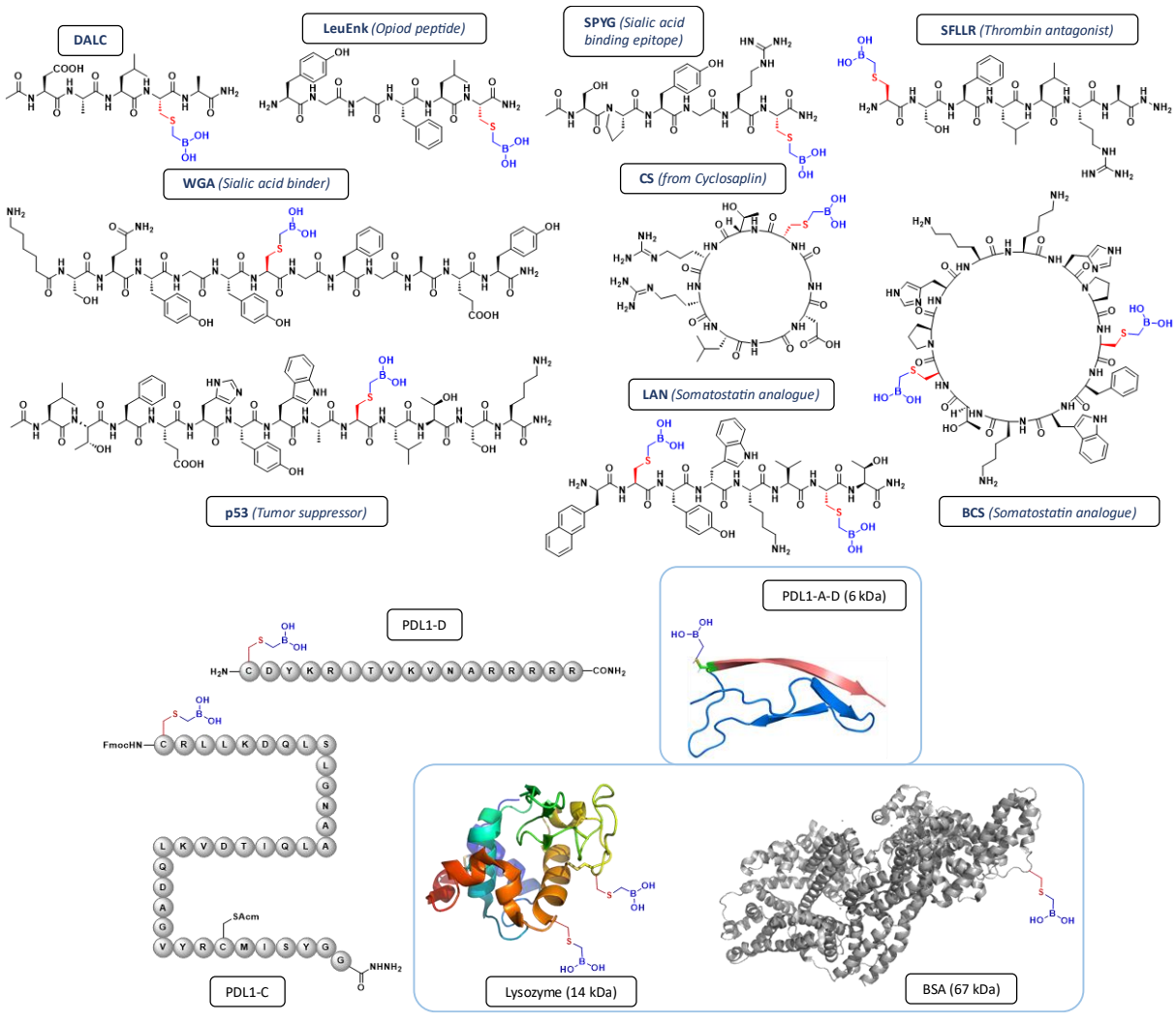


Figure 4.3. Structures of the peptides and proteins (in blue box) used for demonstration.

Having established the reaction strategy with peptides, we moved to test it out on proteins. We synthesized a short protein sequence (51-mer) by combining two fragments (sequence number Ala1-Ala34 and Cys35-Arg51, namely PDL-1-A and PDL-1-D, respectively[17]) of programmed death ligand (PDL1) peptide by NCL. The 6 kDa protein was selectively modified at the Cys residue using only a 10 µM concentration of the substrate. Though the LC spectra did not show a clear peak shift of the conjugated product from the starting material, the MS spectra clearly revealed the full conversion of the peptide. Lysozyme (14 kDa) was subjected for TCEP treatment followed by successful modification of two Cys residues with the boronic acid reagent. A 10 µM solution of the BSA protein was also successfully modified at one Cys residue with the developed strategy (Figure 4.4a). The protein modifications were all confirmed by mass spec by deconvoluting the obtained MS spectra (for BSA see Figure 4.4b). To further confirm the installation of boronic acid moiety, we decided to conjugate a reported ligand salicylhydroxamic acid (SHA)[18,19] to the boronic acid handle. The SHA ligand was tagged with a fluorophore (SHA-FITC) to enable its visualization. Post reaction, an SDS-PAGE analysis was performed where the modified and unmodified proteins were run with and without SHA-FITC. In line with our expectations, only SHA-FITC incubated boronic acid-modified protein (BSA-BA-Fl) showed a bright green fluorescence, while the other lanes used as controls did not show any detectable fluorescence (Figure 4.4c). Altogether the results prove the successful and chemoselective installation of the boronic acid moiety into peptides and proteins.

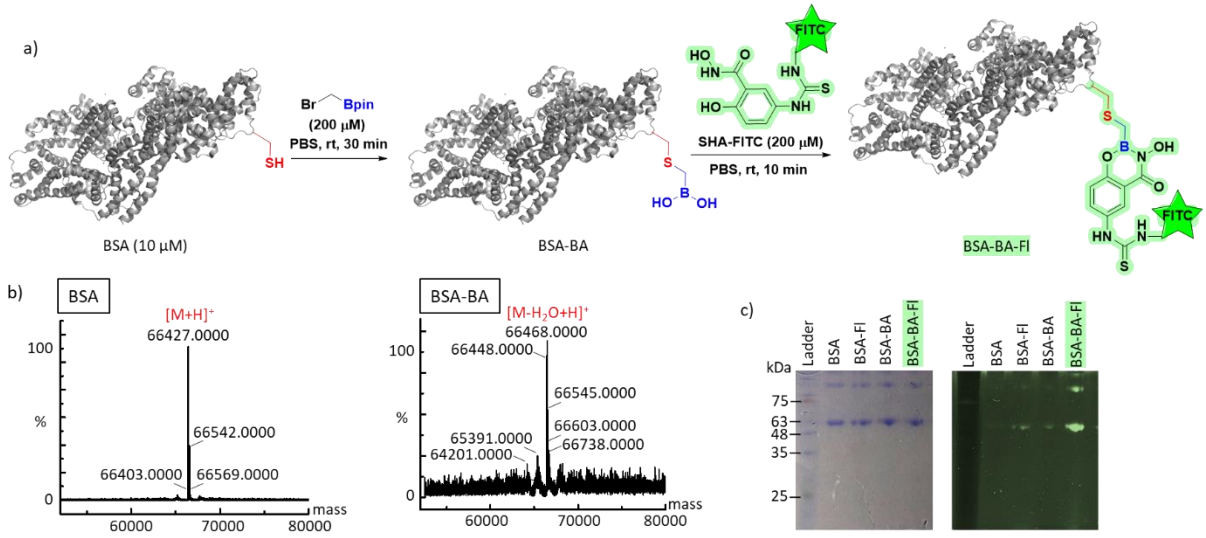


Figure 4.4. a) Installation of BA on BSA and subsequent labelling by a fluorophore SHA-FITC; b) Deconvoluted mass spec data for BSA and BSA-BA; c) SDS-PAGE analysis showing the visible protein bands (left) and fluorescent band (right) where BSA = unmodified BSA protein, BSA-Fl = unmodified BSA incubated with SHA-FITC, BSA-BA = protein modified with BA handle, BSA-BA-Fl = modified BSA-BA protein incubated with SHA-FITC.

To achieve high affinity, nature frequently employs multivalent avidity, since single-site affinity is typically poor; at the cellular level, practically nothing is present as a single copy[20]. So, we were interested to see if multivalent installation of boronic acid is possible with our developed strategy. We designed a 9-mer peptide sequence containing 4 Cys residues. Three variations of this sequence were synthesized where two, three and all four Cys residues were free while other Cys residues were protected with Acm. The peptides were first

treated with TCEP to eliminate any unwanted disulfide formation in solution, following which 750 µM, 1 mM and 2 mM of the boronic acid reagent was incubated to achieve bis, tris and tetra installation (Figure 4.5). This paves the way to obtain multivalent boronic acid modifies peptides with commercially available reagents.

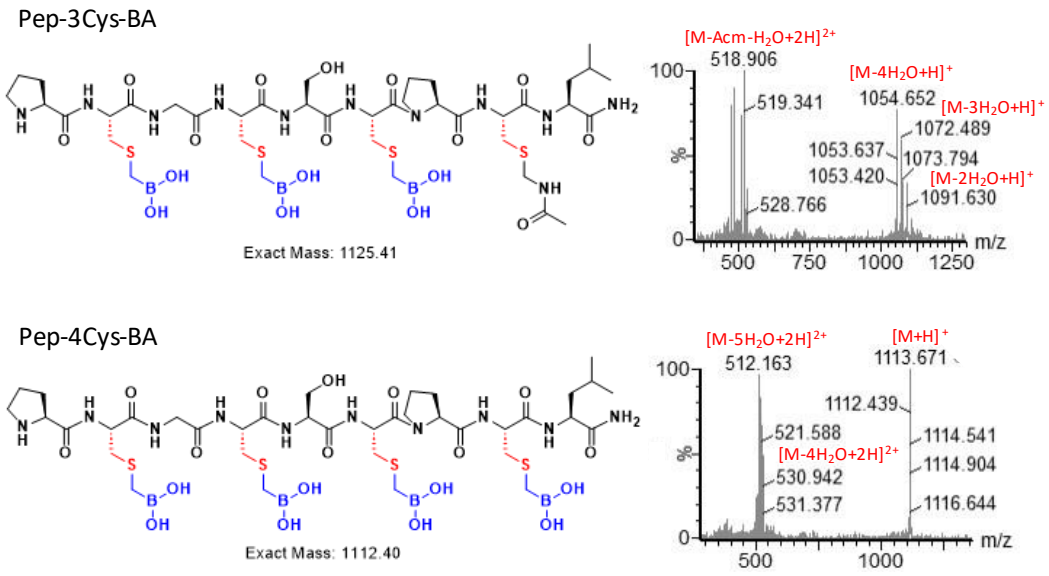


Figure 4.5. Multivalent installation of tris and tetra BA installation with mass spec data.

4.3 *CONCLUSIONS*

Given the multifarious applications of boronopeptides in chemical biology and medicinal chemistry, it would not be wrong to say that evolution in the synthetic strategies for conjugation is necessary, keeping in mind the need for a green synthetic approach. To this end, we have developed a method for the facile installation of a minimal alkyl boronic acid handle on peptides and proteins in an aqueous solution at a physiological pH. Using commercially available cheap reagents renders this strategy easily operable, eliminating any difficult technical steps. A 1.5 eq. of the reagent can modify biomolecules as low as 10 µM with high kinetics, providing a quantitative conversion to a chemoselective conjugate. This strategy is applicable to both peptides and proteins with high efficiency and we anticipate that the multivalent installation strategy will help explore new frontiers in chemical biology applications.

4.4 *EXPERIMENTAL SECTION*

4.4.1 *CHEMICALS AND GENERAL METHODS*

All the chemicals were purchased from commercial sources and used without further purification. Halomethyl boronic acid reagents were purchased from GLR Innovations, India. The reagents for peptide synthesis such as Fmoc-protected amino acids, 2-(1*H*-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), and polystyrene rink amide AM resin (loading capacity 0.67 mmol/g) were purchased from Chem-Impex Int'l Inc (Wood Dale, IL).

All reactions were monitored by thin-layer chromatography (TLC) on aluminium-backed silica gel 60 F_{254}, 0.2 mm plates (Merck), and compounds were visualized under UV light (254 nm) or charred with appropriate solutions – ninhydrin charring to visualize amines, DNP charring for aldehydes, $KMnO_4$ charring for phenolic OH, and curcumin or anisaldehyde charring for boronic acids. Synthetic compounds were purified on 60-120 or 100-200 silica gel purchased from Finar.

4.4.2 *INSTRUMENTS*

NMR data was recorded on a 400 MHz Jeol JNM ECS400 NMR spectrometer. ^{1}H- and ^{13}C-NMR spectra were referenced to TMS. Chemical shifts are given in parts per million (ppm) and coupling constants are declared in Hertz (Hz). The multiplicities of the signals are reported as singlet (s), doublet (d), triplet (t), quartet (q), pentet (pent), multiplet (m), broad (br) or any combination of these. Spectra were analysed using MestReNova.

Mass-spec data was generated by a Waters XEVO G2-XS QTOF mass spectrometer or by Waters SQ equipped with a Waters Acquity UPLC. Separation was achieved using a Waters Acquity UPLC C18 (1.7 µm, 2.1×50 mm) analytical column using solvent A: water (0.1% formic acid) and solvent B: acetonitrile (0.1% formic acid) as mobile phase at a flow rate of 0.25 ml/min and running a 16 min linear gradient as follows:

LCMS_16min_60: 5% solvent B for 0.5 min, 5 to 60% solvent B over 11 min, 60 to 95% solvent B over 1 min, and 5% solvent B for 3 min. MS conditions: Scan time 0.5 – 13 min, Cone voltage of 25 V, Ionization mode as ES+, Data as continuum and a Scan time of 1 sec.

LCMS_16min_80: 5% solvent B for 0.5 min, 5 to 80% solvent B over 11 min, 80 to 95% solvent B over 1 min, and 5% solvent B for 3 min. MS conditions: Scan time 0.5 – 15 min, Cone voltage of 25 V, Ionization mode as ES+, Data as centroid and a Scan time of 1 sec.

Spectra generated from Waters XEVO G2-XS QTOF were deconvoluted using MassLynx 4.1 (Waters) and the "MaxEnt1" deconvolution algorithm with the following settings: resolution: 1.0 Da per channel; damage model: uniform Gaussian; width at half height: 0.4 Da; minimum intensity ratios: 33% (left) and 33% (right); and iterate to convergence.

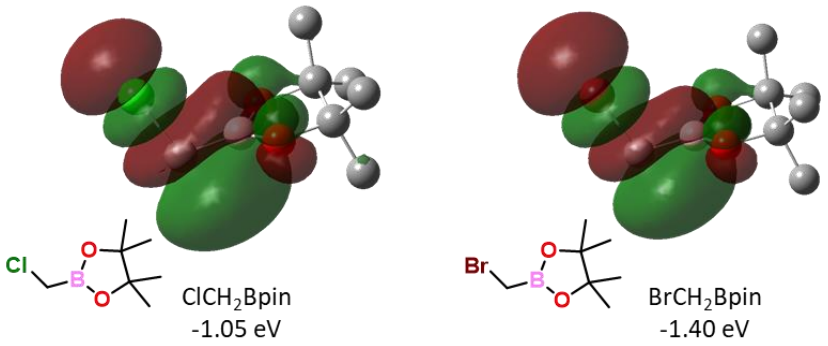
HPLC analysis and peptide purification were carried out on a Shimadzu Prominence UFLC system. Separation was achieved using a Waters Reliant C18 (5 µm, 4.6×250 mm) analytical column using solvent A: water (0.1% TFA) and solvent B: acetonitrile (0.1% TFA) as mobile phase at a flow rate of 1 ml/min and running a 35 min linear gradient as follows: 10% solvent B for 1 min, 10 to 45% solvent B over 24 min, 45 to 90% solvent B over 3 min, isocratic 90% solvent B over 3 min and 10% solvent B for 3.5 min. Peptide purification was carried out on a Shimadzu shim-pack GIST C18 (5 µm, 10×250 mm) semi-preparative column and mobile phase ACN-H_2O (0.1% TFA) with a flow rate 4 mL/min were used. The gradient was optimized for each peptide.

UV data was generated from a Shimadzu UV-2600 system. Flow cytometry analyses were carried out on a BD Accuri C6 Plus. SDS-PAGE was run in a Bio-Rad Mini-PROTEAN® Tetra Cell and visualized by in-gel

fluorescence using a BioRad ChemiDocTM imager. CD spectra of KLA peptides were recorded on a JASCO J-1500 Circular Dichroism Spectrometer, Easton, MD, USA.

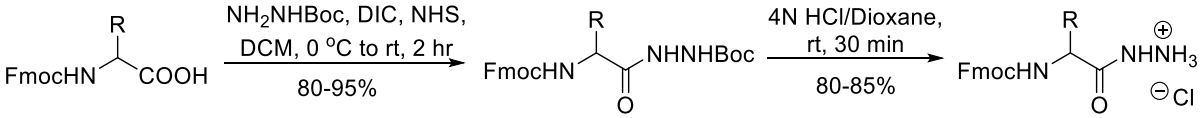
4.4.3 *THEORETICAL CALCULATIONS*

The energy calculation and small molecule optimization was carried out on Gaussian 09 program using a high-performance computing cluster facility of IIT Ropar at the B3LYP level of theory with basis set 6-31G/LANL2DZ. The conductor-like polarizable continuum model (CPCM) solvation model was employed to simulate systems in the aqueous solution. GaussView 6.0 software program were used for the visualization and analysis of the molecular structures (isovalue = 0.02).



4.4.4 *CHEMICAL SYNTHESIS OF SMALL MOLECULES*

Synthesis of Hydrazide derivatives



General procedure: Fmoc-protected amino acid (1 eq.) and NHS (1 eq.) was dissolved in 10% DMF (HPLC gr.)/dry DCM to a final concentration of 200 mM. After cooling the reaction mixture in an ice-bath for 10 min, DIC (1 eq.) was added dropwise. Immediate precipitation of diisopropylurea was observed indicating the onset of active ester formation. The ice-bath was removed and the reaction was allowed to proceed for 40 min. TLC analysis revealed complete consumption of the acid derivative. Solid NH_2NHBoc (1.1 eq.) was subsequently added to the active ester and the reaction was allowed to proceed for 1 hr when TLC analysis revealed completion of the reaction. The reaction mixture was filtered through a silica plug and worked up with EtOAc/1N HCl after evaporation of DCM. The organic layer was washed with brine and evaporated to obtain the desired compound as a mixture with diisopropylurea. The product was confirmed by LC-MS and proceeded for the next step without any further purification.

Waters Single quad using LCMS_16min_60

Fmoc-Gly-NHNHBoc: LC R_t 11.01 min; MS Calc. $[M\text{-}^tBu+H]^+$ 356.12, Obs. $[M\text{-}^tBu+H]^+$ 356.20

Fmoc-Ala-NHNHBoc: LC R_t 11.42 min; MS Calc. $[M\text{-}^tBu+H]^+$ 370.14, Obs. $[M\text{-}^tBu+H]^+$ 370.19

Fmoc-Pro-NHNHBoc: LC R_t 11.96 min; MS Calc. $[M\text{-}^tBu+H]^+$ 396.16, Obs. $[M\text{-}^tBu+H]^+$ 396.30

Boc deprotection of the synthesized amino acids was achieved by treating with 4N HCl/Dioxane at a concentration of about 100 mg/mL for 30 min when TLC analysis revealed completion of reaction. The solvent was evaporated using DCM (20 mL × 3) and washed with hexane to obtain the hydrazides as HCl salt. The amino acid hydrazides were used as such for peptide synthesis without any further purification. To determine the percentage purity of the compounds, UV measurements were carried out by comparing with the respective Fmoc amino acids used as starting materials.

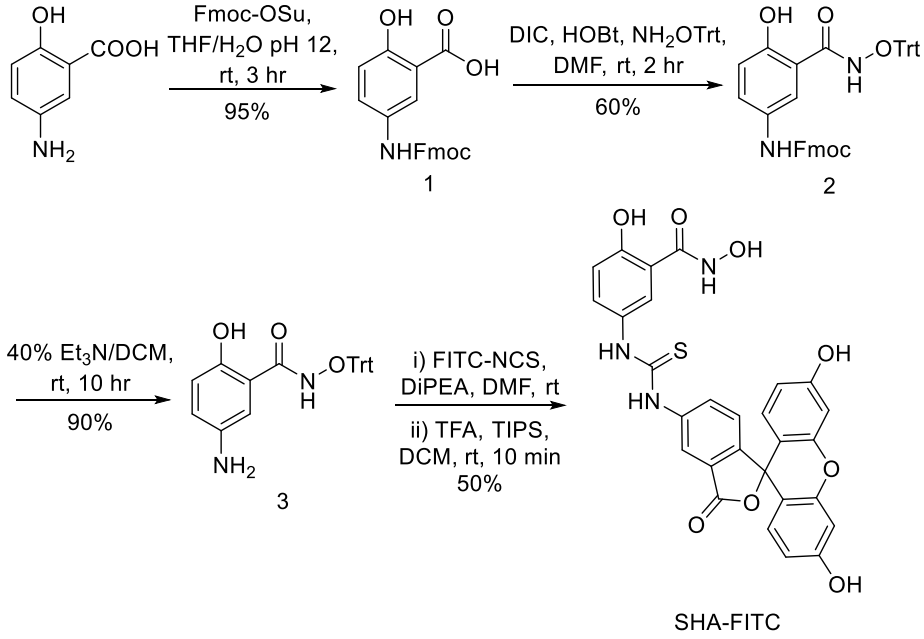
Waters Single quad using LCMS_16min_60

Fmoc-Gly-$NHNH_2$: LC R_t 8.16 min; MS Calc. $[M+H]^+$ 312.13, Obs. $[M+H]^+$ 312.19

Fmoc-Ala-$NHNH_2$: LC R_t 8.54 min; MS Calc. $[M+H]^+$ 326.15, Obs. $[M+H]^+$ 326.11

Fmoc-Pro-$NHNH_2$: LC R_t 8.92 min; MS Calc. $[M+H]^+$ 352.17, Obs. $[M+H]^+$ 352.28

Synthesis of SHA-FITC



*Synthesis of **1***: 5-Amino salicylic acid (306 mg, 2 mmol) was dissolved in 5 mL 10% Na_2CO_3 solution and stirred in an ice-bath. Fmoc-OSu (674 mg, 2 mmol, 1 eq.) was dissolved in 5 mL THF and added to the amine slowly. After complete addition the ice-bath was removed and the reaction mixture was stirred at room temperature for 3 hr when TLC analysis revealed consumption of Fmoc-OSu. The reaction mixture was then acidified to pH 2 using 4N HCl and then extracted with EtOAc (100 mL × 2). The organic layer was washed with brine and evaporated to obtain a crude 720 mg (95% yield) light brown product which was used in the next step without further purification.

^{1}H NMR (400 MHz, DMSO-d_6): δ 9.61 (s, 1H), 7.94 (s, 1H), 7.87 (d, *J* = 7.5 Hz, 2H), 7.70 (d, *J* = 7.4 Hz, 2H), 7.51 (s, 1H), 7.39 (t, *J* = 7.4 Hz, 2H), 7.31 (t, *J* = 7.3 Hz, 3H), 6.85 (d, *J* = 9.1 Hz, 1H), 4.42 (d, *J* = 6.7 Hz, 2H), 4.26 (t, *J* = 6.6 Hz, 1H).

Waters SQ using LCMS_16min_80: LC R_t 9.92 min; MS Calc. $[M+H]^+$ 376.12, Obs. $[M+H]^+$ 376.11

*Synthesis of **3***: 720 mg of **1** (1.9 mmol) and HOBt.H_2O (323 mg, 2.1 mmol, 1.1 eq.) was dissolved in 9.5 mL DMF (HPLC gr.) and stirred in an ice-bath. DIC (300 µL, 1.9 mmol, 1 eq.) was added dropwise to the reaction mixture and the formation of DIC urea was observed immediately. The ice-bath was removed and after 5 min, NH_2OTrt (784 mg, 2.85 mmol, 1.5 eq.) was added to initiate amide formation. After 2 hr, TLC analysis revealed consumption of **1**. The reaction was diluted into a separating flask containing 100 mL EtOAc and 100 mL chilled 1N HCl. The extracted organic layer was further washed with 1N HCl (75 mL × 2) followed by brine and evaporated to obtain **2** as a yellowish sticky solid which was proceeded for next step.

Waters SQ using LCMS_16min_80: LC R_t 13.01 min; MS Calc. $[M+H]^+$ 376.12, Obs. $[M+H]^+$ 376.11

700 mg of **2** (1.1 mmol) was subjected for Fmoc deprotection in 20 mL 50% Et_3N/DCM for 10 hr at room temperature. After completion of the reaction as confirmed by TLC (50% EtOAc/Hexane), DCM was evaporated under reduced pressure and the resulting crude was purified through 100-200 silica gel using 3% MeOH/DCM (0.1% Et_3N) as the eluent. 400 mg (90% yield) of brownish semi-solid was obtained. ^{1}H NMR revealed the presence of Et_3N in the compound which could not be removed by washing with organic solvents or subjecting for lyophilization, hence we assumed the compound exists as a co-crystal with Et_3N.

^{1}H NMR (400 MHz, $CDCl_3$): δ 7.44 (d, *J* = 1.8 Hz, 2H), 7.42 (s, 3H), 7.27 – 7.15 (m, 10H), 6.77 (d, *J* = 2.8 Hz, 1H), 6.60 (d, *J* = 8.6 Hz, 1H), 6.52 (dd, *J* = 8.5, 2.8 Hz, 1H).

Waters SQ using LCMS_16min_80: LC R_t 7.71 min; MS Calc. $[M+H]^+$ 411.17, Obs. $[M+H]^+$ 411.24

*Synthesis of **SHA-FITC***: For labelling of FITC to **3**, to 10 mg of the amine (24 µmol) in 100 µL of DMF (HPLC gr.), was added 8.5 µL DiPEA (2 eq., 48 µmol) and 9.7 mg FITC (1 eq., 24 µmol), The resulting red solution was stirred for 3 hr and monitored by TLC in 10% MeOH/DCM. The reaction mixture was dissolved in 10 mL EtOAc and washed with 8 mL × 2 chilled 1N HCl sol. followed by washing with brine and evaporating the organic layer to obtain 50 mg crude.

The product was subjected for Trt deprotection in 0.3 mL TIPS, 1 mL TFA and 5 mL DCM for 10 min at room temperature. The volatiles were then evaporated and the obtained crude was dissolved in 30% ACN/H_2O (0.1% TFA) and purified by RP-HPLC (gradient 10-60% ACN over 26 min) and lyophilized to obtain 7 mg yellow powder (50% yield). A stock solution was prepared with 2 mg in DMSO (MB grade) and its concentration was determined by UV-Vis (ε_{FITC} = 73,000 M^{-1} s^{-1}).

^{1}H NMR (400 MHz, DMSO-d_6): δ 12.14 (s, 1H), 11.34 (s, 1H), 10.11 (s, 2H), 10.02 (s, 1H), 9.88 (s, 1H), 8.09 (d, *J* = 2.0 Hz, 1H), 7.74 (dd, *J* = 8.0, 2.2 Hz, 1H), 7.64 (d, *J* = 2.8 Hz, 1H), 7.33 (dd, *J* = 8.8, 2.6 Hz, 1H), 7.16 (d, *J* = 8.5 Hz, 1H), 6.88 (d, *J* = 9.0 Hz, 1H), 6.63 (d, *J* = 2.0 Hz, 2H), 6.54 (d, *J* = 2.4 Hz, 3H).

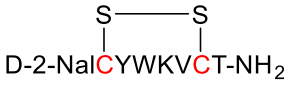
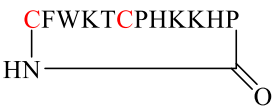
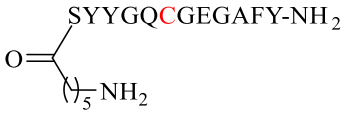
^{13}C NMR (101 MHz, DMSO-d_6) δ 180.8, 169.0, 166.0, 160.0, 157.5, 152.4, 141.8, 131.8, 129.6, 126.9, 124.5, 117.9, 114.6, 113.1, 110.1, 102.8, 29.6.

Waters SQ using LCMS_16min_60: LC R_t 8.36 min; MS Calc. $[M+H]^+$ 558.10, Obs. $[M+H]^+$ 558.06

4.4.5 *SYNTHESIS OF PEPTIDES*

The peptides were synthesized on rink amide AM resin (loading capacity 0.67 mmol/g) or CTC resin (1.5 mmol/g) using standard Fmoc-deprotection strategy. The resin was used with the specified loading or downloaded to a lower loading as required. Three equivalents of the commercially available amino acids, HBTU as a coupling agent and DiPEA as a base, were used for the coupling reaction for 15 min. Fmoc deprotection was achieved using 3 mL 20% piperidine/DMF twice for 3 min each, followed by washing with 3 mL DMF six times. The peptides were cleaved off of the resin and globally deprotected with reagent B (88% TFA, 5% phenol, 5% H_2O and 2% TIPS) for shorter 6 to 8-mer peptides or with reagent K (82.5% TFA, 5% H_2O, 2.5% EDT, 5% thioanisole and 5% phenol) for longer sequences. A typical time of 2 hr was provided for cleavage. Precipitation with chilled diethyl ether gave the crude peptides, which were purified by RP-HPLC. The purity of the peptides was determined by using LC-MS (Waters) and HPLC (Shimadzu) to be >90%. The characterization data are presented in Table 4.1. All the peptides were purified by semi-preparative RP-HPLC and lyophilized to obtain them as white powder with an isolated yield of 25-35%.

Table 4.1. List of synthesized peptides and proteins (entry 13) with their biological activity (if any).

Sl. No.	**Name**	**Sequence**	**Length**	**Calc. *m/z***	**Obs. *m/z***
1.	TestPep	Ac-DALCA-NH_2	5-mer	$[M+H]^+$ 533.24	$[M+H]^+$ 533.41
2.	SPYG (Sialic acid binding epitope)	Ac-SPYGRC-NH_2	6-mer	$[M+H]^+$ 723.32	$[M+H]^+$ 723.45
3.	LeuEnk (opiod peptide)	YGGFLC-NH_2	6-mer	$[M+H]^+$ 658.30	$[M+H]^+$ 658.14
4.	Thrombin antagonist	CSFLLRA-$NHNH_2$	7-mer	$[M+H]^+$ 823.46	$[M+H]^+$ 823.79
5.	LAN (Somatostatin analogue)		8-mer	$[M+H]^+$ 1096.33	$[M+H]^+$ 1096.85
6.	LAN-B	D-2-NalCYWKVCT-NH_2	8-mer	$[M+H]^+$ 1098.49	$[M+H]^+$ 1098.47
7.	CS (from cyclosaplin)		8-mer	$[M+H]^+$ 859.42	$[M+H]^+$ 859.76
8.	BCS (Somatostatin analogue)		12-mer	$[M+3H]^{3+}$ 498.58	$[M+3H]^{3+}$ 498.17
9.	WGA		13-mer	$[M+H]^+$ 1456.62	$[M+H]^+$ 1456.59
10.	p53 (MDM2 inhibitor)	Ac-LTFEHYWACLTSK-NH_2	13-mer	$[M+H]^+$ 1640.90	$[M+H]^+$ 1640.79
11.	PDL1-D	CDYKRITVKVNARRRRR-NH_2	17-mer	$[M+4H]^{4+}$ 548.33	$[M+4H]^{4+}$ 548.32

12.	PDL1-C	FmocCRLLKDQLSLGNAALQITD VKLQDAGVYRC(Acm)MISYGG-$NHNH_2$	36-mer	$[M+5H]^{5+}$ 845.24	$[M+5H]^{5+}$ 845.22
13.	PDL1-A-D	AFTVTVPKDLYVVEYG SNMTIEC(Acm)KFPVEKQLDLA CDYKRITVKVNARRRRR-NH_2	51-mer	$[M+5H]^{5+}$ 1219.26	$[M+5H]^{5+}$ 1219.23

On-resin disulfide mediated peptide cyclization of LAN: This was performed according to the reported literature[21]. 0.05 mmol of peptidyl resin was swelled in DCM and washed with DMF. 23.8 mg sodium persulfate ($Na_2S_2O_8$) (0.1 mmol, 2 eq.) was dissolved in 40 µL of distilled water and mixed with 50 mg iodine (0.2 mmol, 4 eq.) dissolved in 455 µL DMF to a final conc. of 400 mM of iodine. Then the peptidyl resin was treated with this iodine-sodium persulfate mixture at a final concentration 400 mM for 15 min. After washing with DMF (3 mL × 2), it was further washed with 2% ascorbic acid (wt./vol.) in DMF (3 mL × 2) to remove the excess iodine. The resin was further washed with DMF (3 mL × 2) and DCM (3mL × 4) and then subjected for global deprotection.

Coupling of the first hydrazide derivatized amino acid to CTC resin: Peptides **4** (CSFLLRA-$NHNH_2$), **7** (CTRRLGDG-$NHNH_2$), **8** (CFWKTCPHKKHP-$NHNH_2$) and **PDL1-A** (AFTVTVPKDLYVVEYGSNMTIEC(Acm)KFPVEKQLDLA-$NHNH_2$) were synthesized on the CTC resin. In a typical procedure, 0.1 mmol of the resin was swelled in 2 mL of dry DCM under argon atmosphere for 15 min after which it was drained off. 2 eq. of the respective amino acid was dissolved in 2 mL 1:1 DMF (HPLC gr.)/dry DCM and 3 eq. DiPEA was added to it. Argon was bubbled through the solution and then it was added to the resin and stirred for 2 hr at room temperature under argon atmosphere. After 2 hr, the amino acid solution was drained and the resin was washed with DCM. A capping solution composed of MeOH:DiPEA:DCM (2:1:17) was administered (3 mL × 3) to cap the unreacted centres. The rest of the amino acids were coupled according to the procedure described before.

Procedure for peptide cyclization and native chemical ligation (NCL): This was performed according to the reported literature[22]. The crude peptide hydrazide (**7**, **8**, **PDL1-A**) was dissolved in 0.2 M phosphate buffer + 6 M Gdn.HCl (final pH 3) at a concentration of 3 mM. The solution was cooled to -15 °C by employing an ice-salt bath and the whole system was kept in a freezer at the same temperature. Intermittent shaking of the reaction vessel was carried out throughout the reaction. $NaNO_2$ (7 eq.) was dissolved in the same buffer and added dropwise to oxidize the peptide hydrazide to peptide azide. After 40 min, MPAA (50 eq.) dissolved in 0.2 M phosphate buffer + 6 M Gdn.HCl (final pH 7) with equal volume was added to the reaction mixture to quench the excess $NaNO_2$ and convert the peptide azide to the corresponding peptide thioester. The reaction mixture was kept at room temperature for 10 min to ensure complete conversion.

For peptide cyclization (**7**, **8**), the reaction mixtures were then diluted with 0.2 M phosphate buffer (final pH 7) to a concentration of 50 µM and agitated for 2 hr. LC-MS analysis revealed formation of the desired product. For NCL, PDL1-D (1.5 eq.) was dissolved in 0.2 M phosphate buffer + 6 M Gdn.HCl (final pH 7) with equal volume as the reaction mixture of PDL1-A. The two solutions were mixed and the pH of the solution was adjusted to 7. After agitating the mixture for 4 hr, LC-MS analysis revealed completion of reaction. Aliquots for LC-MS analysis was treated with TCEP before injection.

The whole procedure is depicted in Figure 4.6.

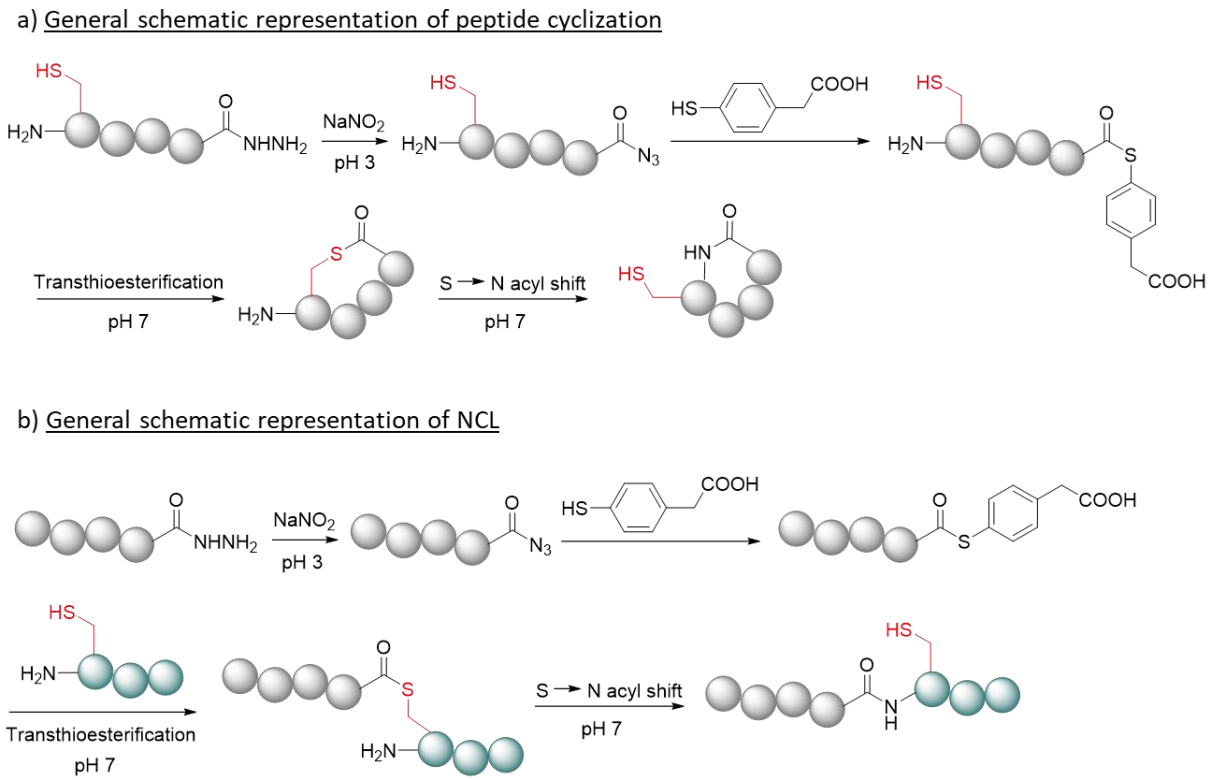


Figure 4.6. a) General schematic representation of peptide cyclization as followed for peptide **7** and **8**; b) General schematic representation of NCL between two peptides (labelled grey and green).

4.4.6. *MODEL REACTIONS OF XCH_2Bpin WITH LEUENK*

Stock solution of the peptide was freshly prepared in DMSO and XCH_2Bpin (X =Cl, Br, I) were prepared in ACN. 100 µM solution of the peptides in 1X PBS pH 7.4 was reacted with 150 µM of XCH_2Bpin for the kinetic studies. From a 2 mL reaction mixture, 150 µL aliquots were withdrawn, acidified with 2 µL of 10% TFA/H_2O at regular time intervals and analysed in analytical RP-HPLC. The percentage product formation was calculated and plotted against time to obtain a curve which was fitted with 2nd order kinetics equation to obtain the rate constants. The chloro derivative showed the slowest kinetics, reaching to 80% product conversion in 4 hr with a rate constant of 6.5 M^{-1} s^{-1} (Figure 4.6a) while the bromo and iodo derivatives showed a similar kinetics ~10^2 6.5 M^{-1} s^{-1}, leading to >95% product conversion within 5 min (Figure 4.6b-c). Due to the higher reactivity of the iodo derivative, a small fraction of over alkylation was observed after 10 min (Figure 4.6d) and thus the bromo derivative was chosen as the reagent of choice for further studies.

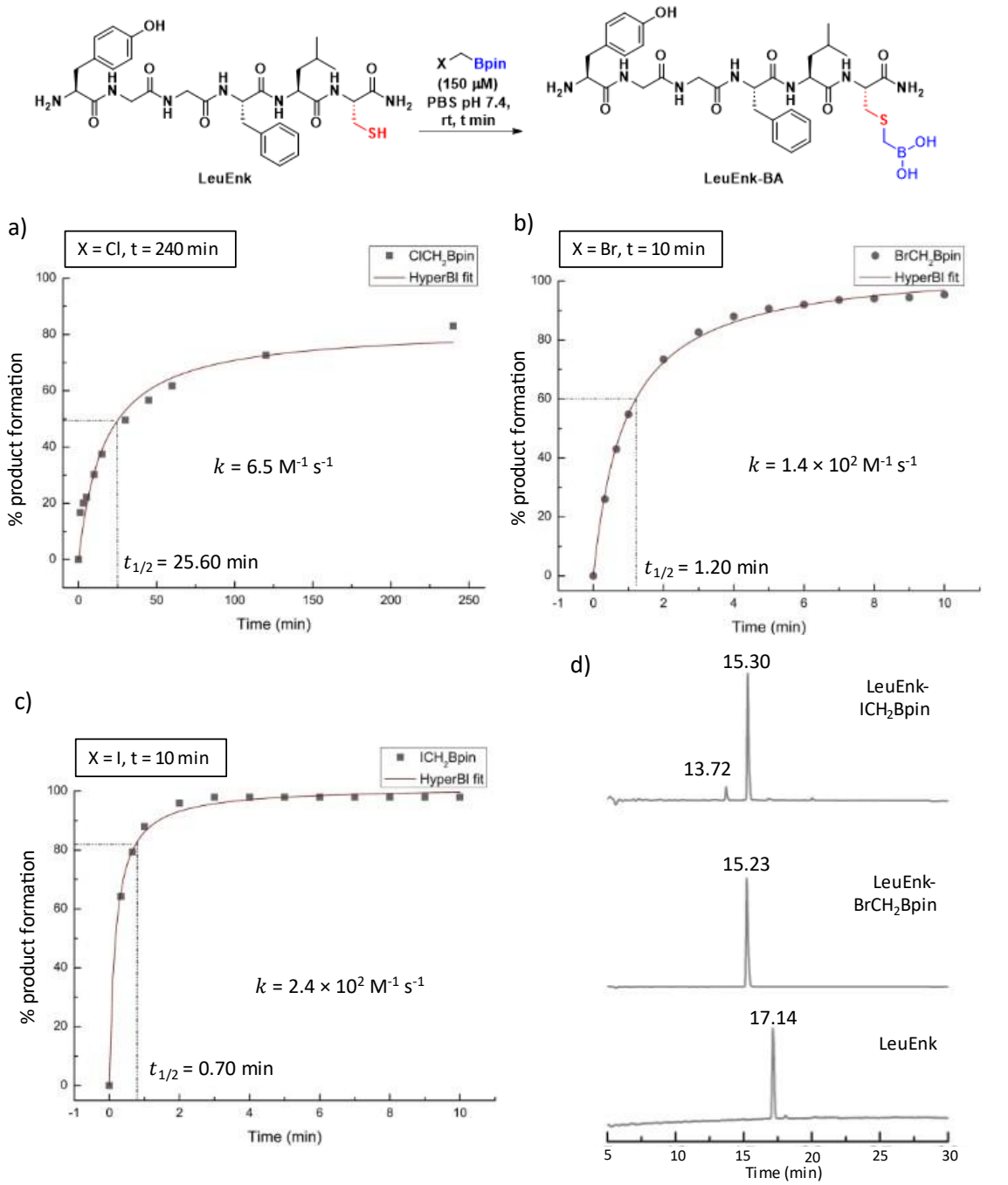


Figure 4.7: General reaction scheme of LeuEnk and reaction kinetics with a) $ClCH_2Bpin$; b) $BrCH_2Bpin$; c) ICH_2Bpin and d) HPLC data showing overalkylation with ICH_2Bpin.

To confirm the site-selective installation of the BA on the Cys residue, a large-scale reaction was performed in a 5 mM scale for 5 min and quenched with 30 µL 6N HCl to pH 3 following which the product was isolated after purification by RP-HPLC in ACN/H_2O (0.1% TFA) with a gradient of 10-45% ACN over 26 min. The product was further confirmed by LC-MS/MS where the fragmentation patterns of b and y ions confirmed the installation on the Cys residue (Figure 4.7a). This was also validated by modification on the p53 peptide where MS/MS was performed (Figure 4.19). ^1H NMR studies coupled with ^1H-^1H COSY in DMSO-d_6 (Figure 4.7b) confirmed the shift in Cys β-CH_2 and α-NH protons when compared to the starting peptide.

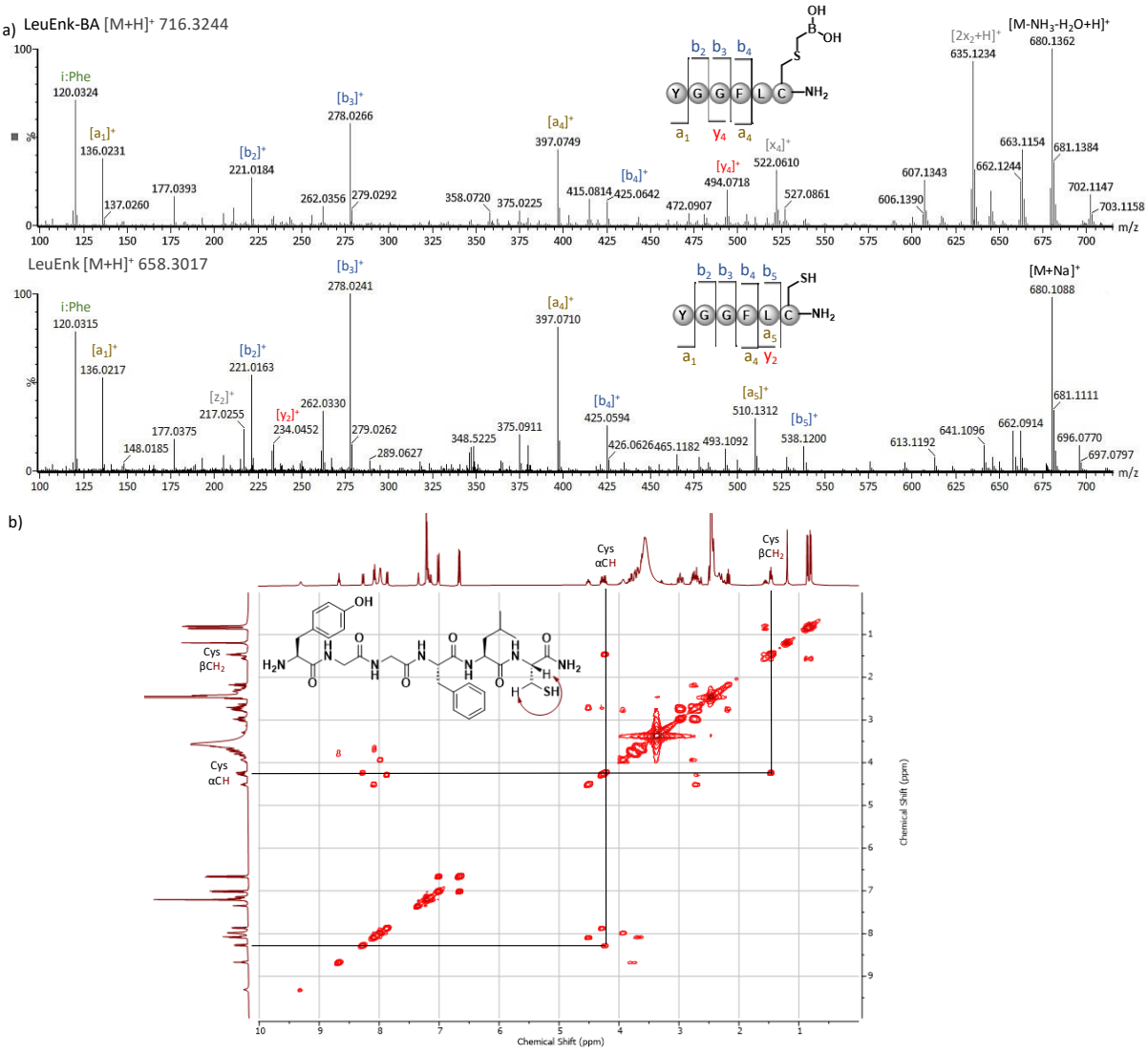


Figure 4.8: a) MS/MS data for LeuEnk and LeuEnk-BA; b) ^{1}H-^{1}H COSY of LeuEnk in DMSO-d_6 to determine the Cys protons.

pH-dependent kinetics with $BrCH_2Bpin$ was performed to determine the range of pH in which this reaction can be performed. It was observed that at a slightly acidic pH 6.2, the reaction was slightly retarded with the product formation reaching about 90% in 45 min while the reactions at pH 7.4 and 8.5 showed a similar kinetic profile (Figure 4.8a). A concentration dependent study of the reagent was also performed with 100 µM LeuEnk. As shown in Figure 4.8b, 1 eq. of the reagent shows about 80% formation of the product, while 1.5 eq. and 2 eq. leads to clean formation to the conjugate. But administration of 5 eq. of the reagent leads to formation of over alkylated product within 10 min.

We were curious to see how this reagent fares when compared to commonly used alkylating agents such as iodoacetamide ICH_2CONH_2, iodoacetic acid ICH_2COOH (a isostere of $BrCH_2B(OH)_2$) and benzyl iodide ICH_2Ph in PBS pH 7.4 for 1 hr using 100 µM LeuEnk and 200 µM of the reagents (Figure 4.8c). $BrCH_2Bpin$ showed a complete conversion to the product within 10 min while the other reagents failed to do so. Among the three control reagents, iodoacetamide showed the best reactivity with about 40% conversion to the product while iodoacetic acid and benzyl iodide did not show any detectable conversion of the peptide. This clearly demonstrates the superiority of the boronic acid reagent. To some extent this study also shows that multiple

electronic factors at the electrophilic methylene centre is at play which dictates the rate of nucleophilic attack of the cysteine thiolate.

To check the compatibility of the reaction in different buffers we performed the reaction in commonly used HEPES and Tris buffer at pH 7.4. Excellent conversion was observed in all the buffers within 10 min (Figure 4.9a). Efficiency of the reaction was also monitored in presence of a disulfide reducing agent tris(2-carboxyethyl)phosphine, TCEP. It was observed that the boronic acid reagent was quenched in the presence of TCEP and thus a higher concentration of the reagent was required to obtain a similar result. To a 100 µM sol. of LeuEnk, 1 mM TCEP was incubated for 1 hr after which $BrCH_2Bpin$ was incubated at different concentrations ranging from 200 µM to 1000 µM. It was found that a minimum of 750 µM of the reagent was required to mediate a complete conversion (Figure 4.9b). A ^1H NMR study between TCEP and $BrCH_2Bpin$ was performed at a concentration of 5 mM in 30% D_2O/H_2O at pH 7.4 where a clear shift of the methylene protons in the boronic acid reagent was observed (Figure 4.9c).

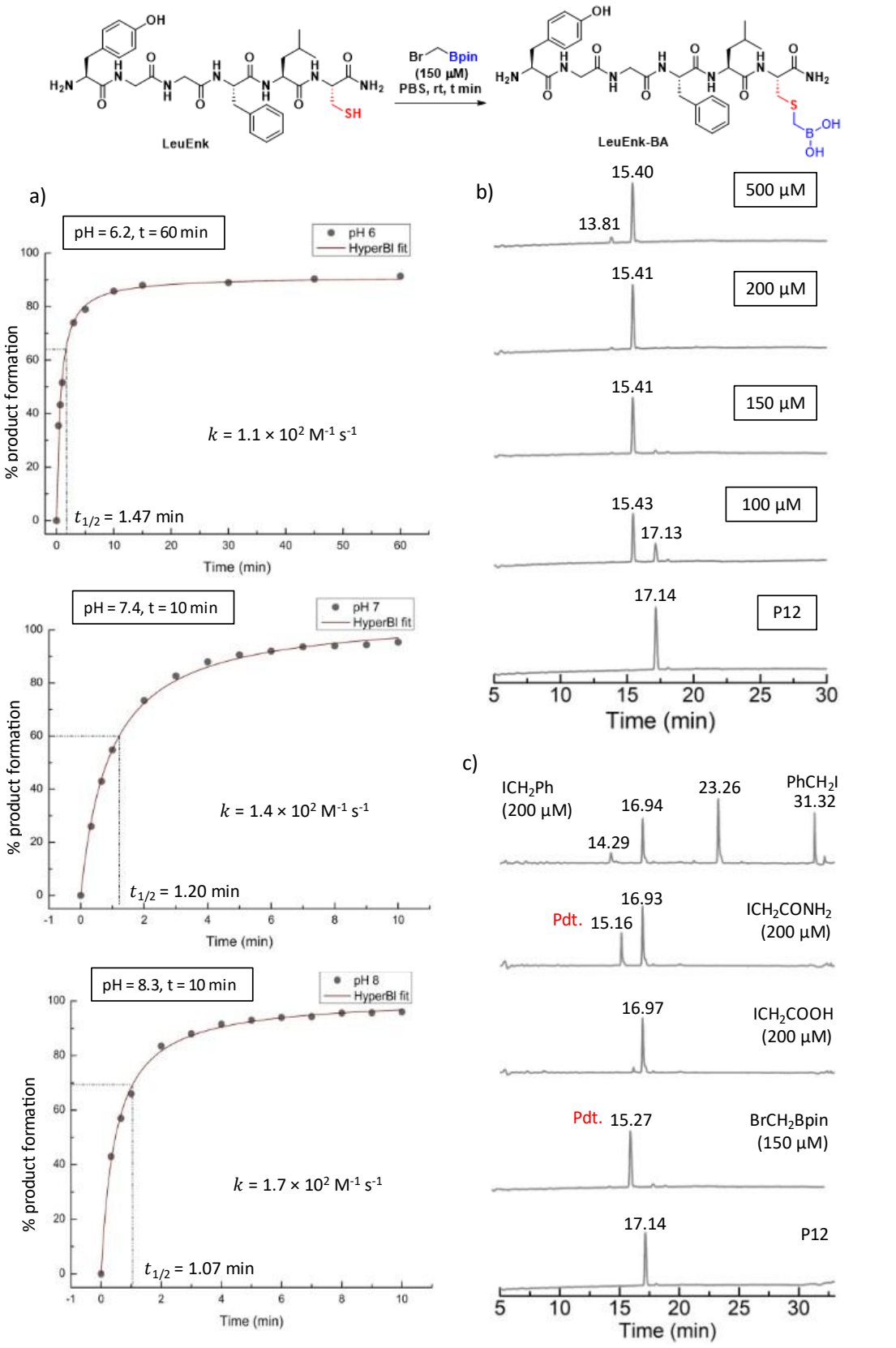


Figure 4.9. a) pH dependent kinetic studies of LeuEnk with $BrCH_2Bpin$; b) Concentration dependent study showing 150 µM optimum concentration of the reagent with 100 µM of LeuEnk; c) Control reactions with commonly used alkylating agents. ICH_2COOH, ICH_2CONH_2 and ICH_2Ph were incubated for 1 hr, keeping other conditions the same.

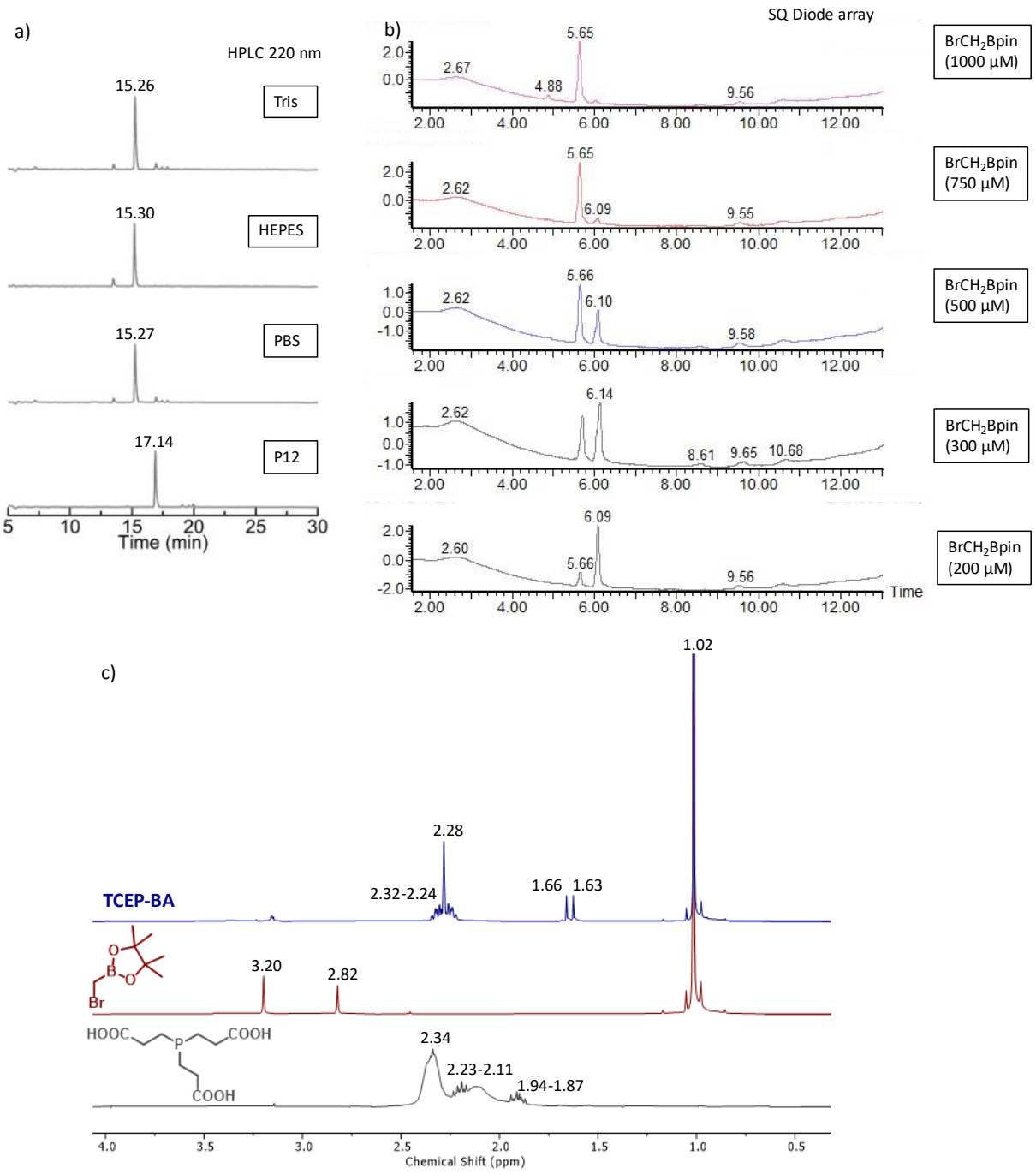


Figure 4.10. a) Reaction of 100 µM LeuEnk with 150 µM $BrCH_2Bpin$ in different buffers for 10 min at pH 7.4; b) Concentration dependent study of $BrCH_2Bpin$ with 100 µM of LeuEnk in the presence of 1 mM TCEP; c) Stacked ^1H NMR data of the reaction of 5 mM TCEP with $BrCH_2Bpin$ in 30% D_2O/H_2O at pH 7.4. The two different peaks for the methylene protons of $BrCH_2Bpin$ may be attributed to the dynamic sp^3 boronate ester formation in an aqueous solution. No other peaks were observed in the rest of the spectrum hence an expanded region from 0.5-4 ppm is presented.

4.4.7. *REACTIONS WITH PEPTIDES AND PROTEINS*

100 µM of the peptides were taken in 1X PBS pH 7.4 and incubated with 150 µM of $BrCH_2Bpin$ for 10 min following which it was quenched with 10% TFA/H_2O to pH 3 and subjected for LC-MS (Waters SQ, method

file LCMS_16min_60) or analytical HPLC. For hydrophobic peptides, a 50 µM concentration was taken while a concentration of only 10 µM was kept for the proteins.

4.4.7.1 *TestPep*

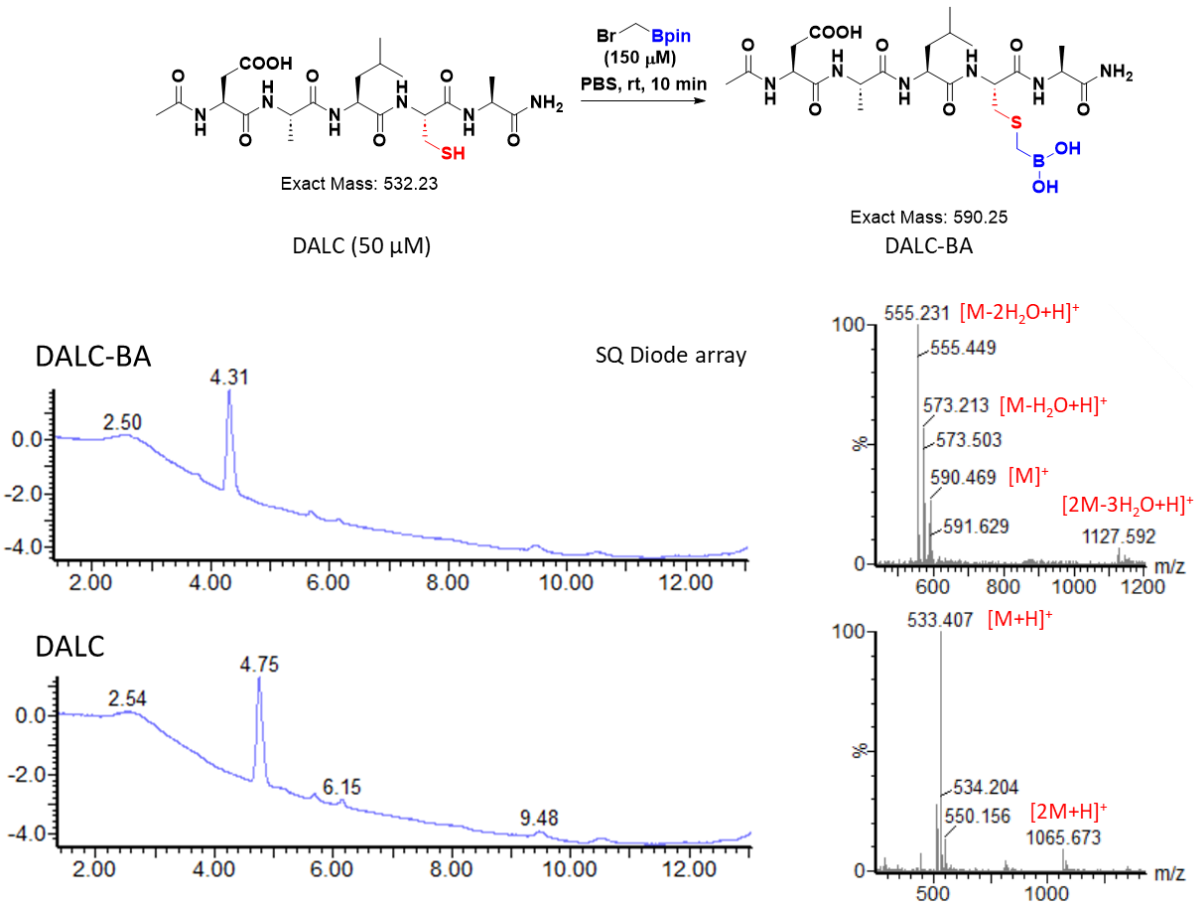


Figure 4.11. Reaction of 50 µM TestPep with 150 µM $BrCH_2Bpin$ showing the LC-MS profile and the respective MS spectra.

4.4.7.2 *SPYG*

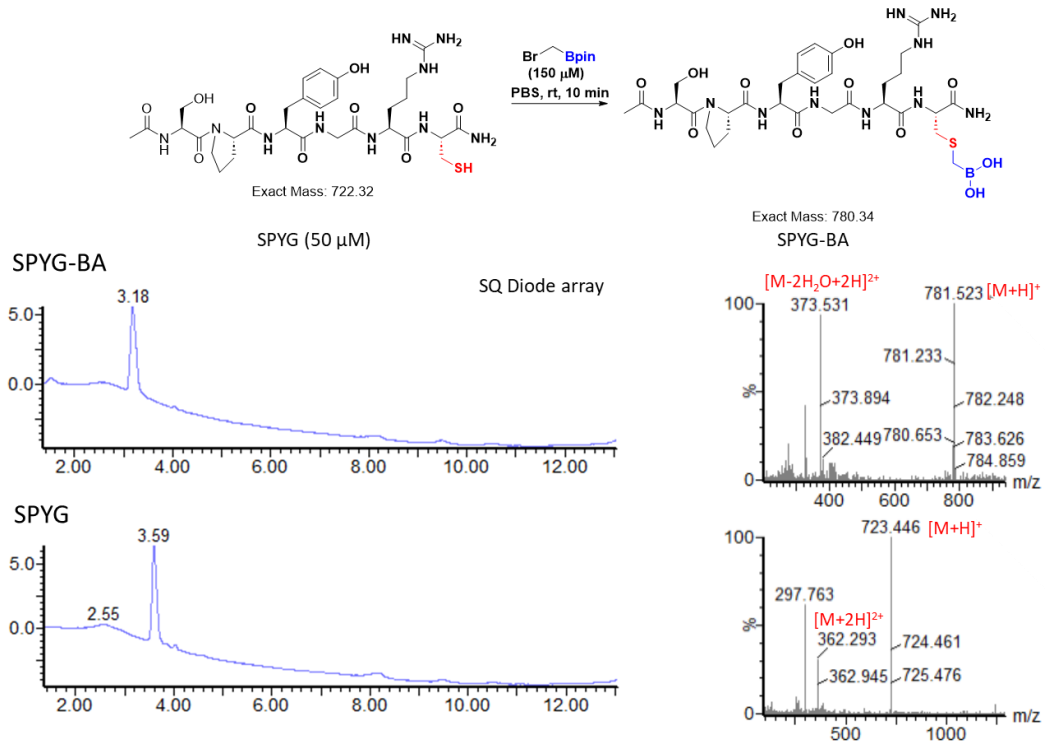


Figure 4.12. Reaction of 50 µM SPYG with 150 µM $BrCH_2Bpin$ showing the LC-MS profile and the respective MS spectra.

4.4.7.3 *SFLLR*

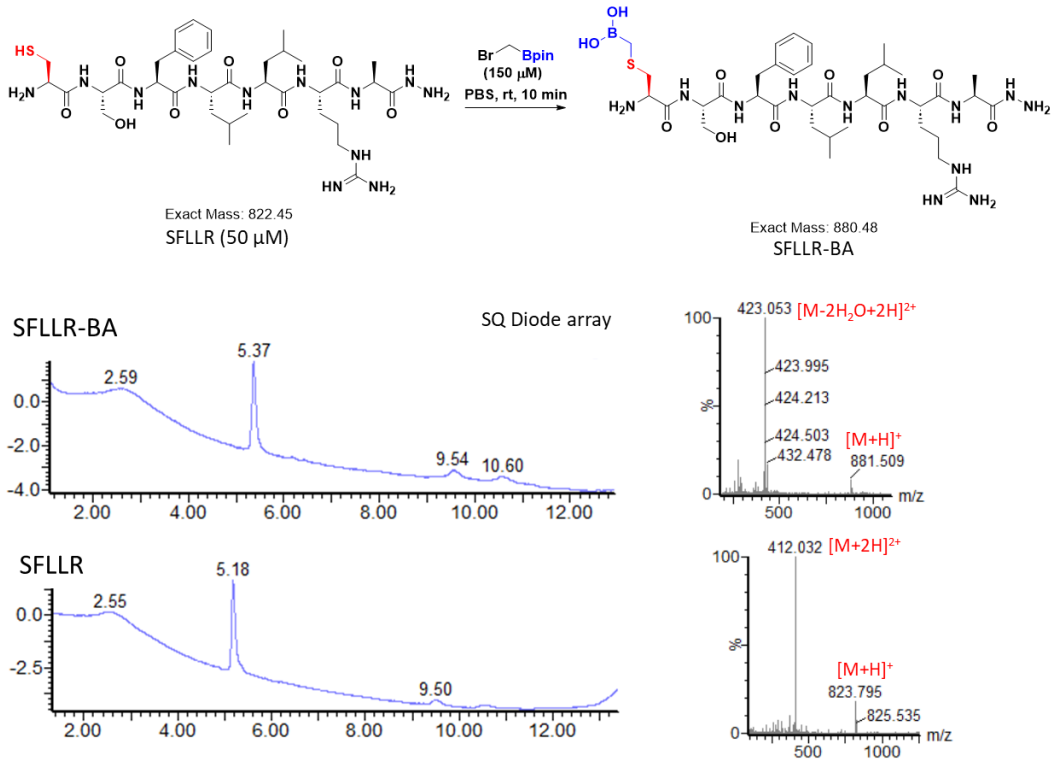


Figure 4.13. Reaction of 50 µM SFLLR with 150 µM $BrCH_2Bpin$ showing the LC-MS profile and the respective MS spectra.

4.4.7.4 *Lanreotide (LAN)*

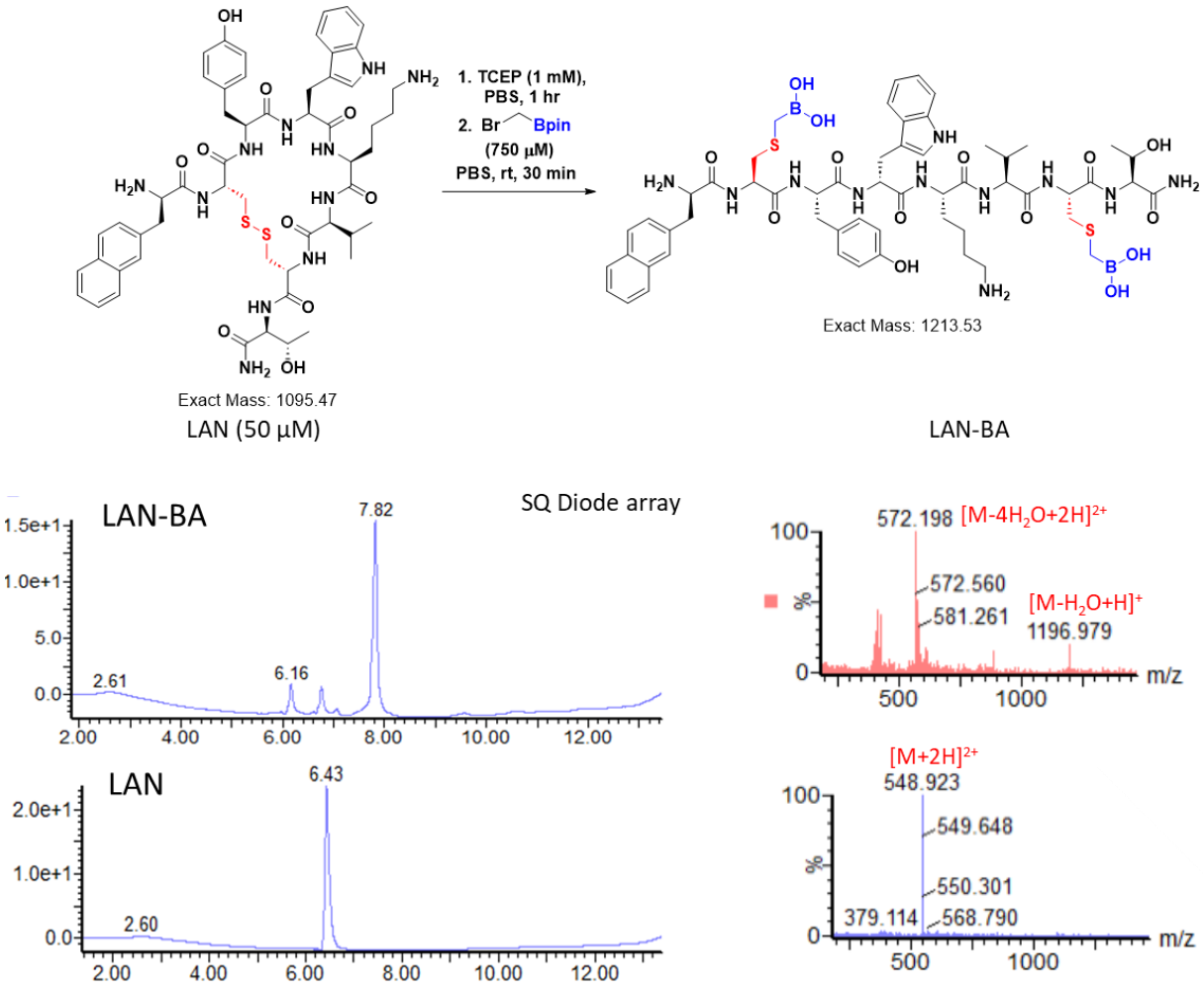


Figure 4.14. Reaction of 50 µM LAN with 1 mM TCEP followed by 750 µM $BrCH_2Bpin$ showing the LC-MS profile and the respective MS spectra. A bis installation was successfully achieved.

4.4.7.5 *Lanreotide impurity B (LAN-B)*

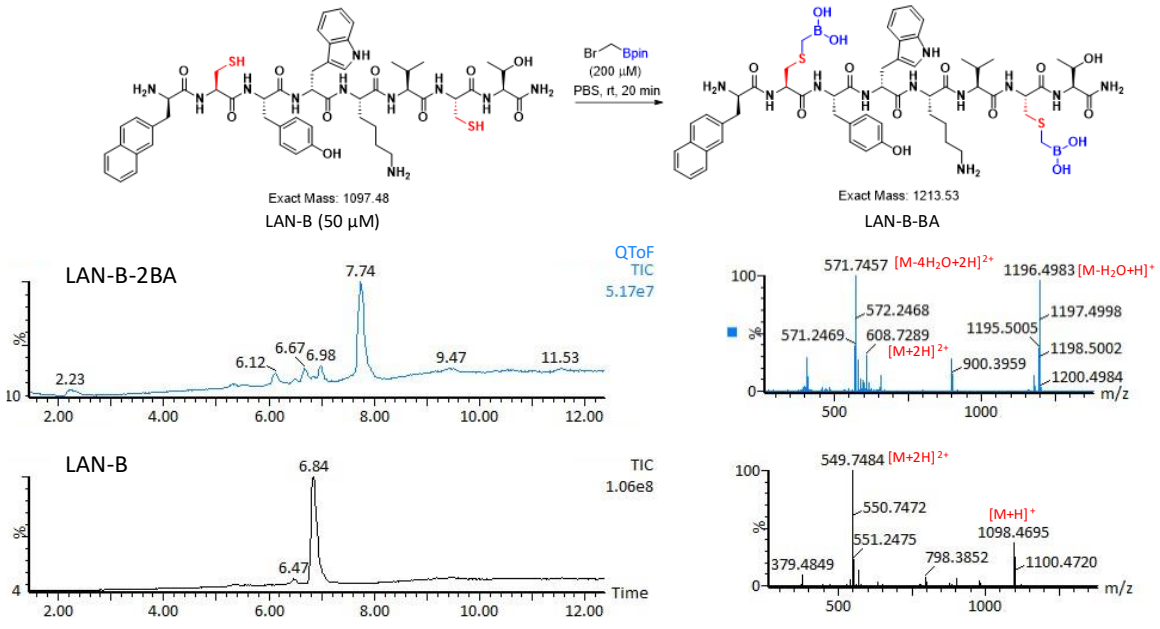


Figure 4.15. Reaction of 50 µM LAN-B with 200 µM $BrCH_2Bpin$ showing the LC-MS profile and the respective MS spectra.

4.4.7.5 *CS (from cyclosaplin)*

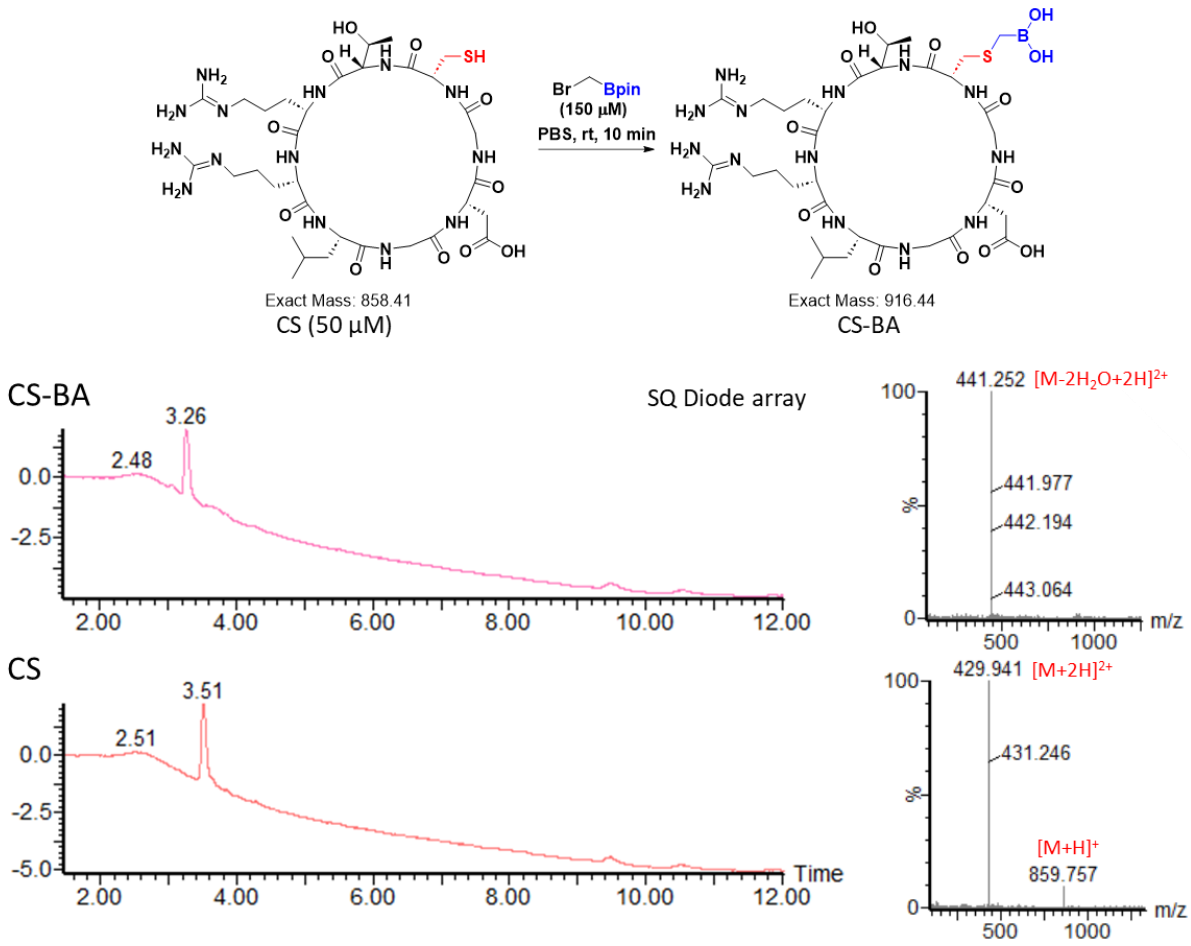


Figure 4.16. Reaction of 50 µM CS with 150 µM $BrCH_2Bpin$ showing the LC-MS profile and the respective MS spectra.

4.4.7.6 *BCS (somatostatin analogue)*

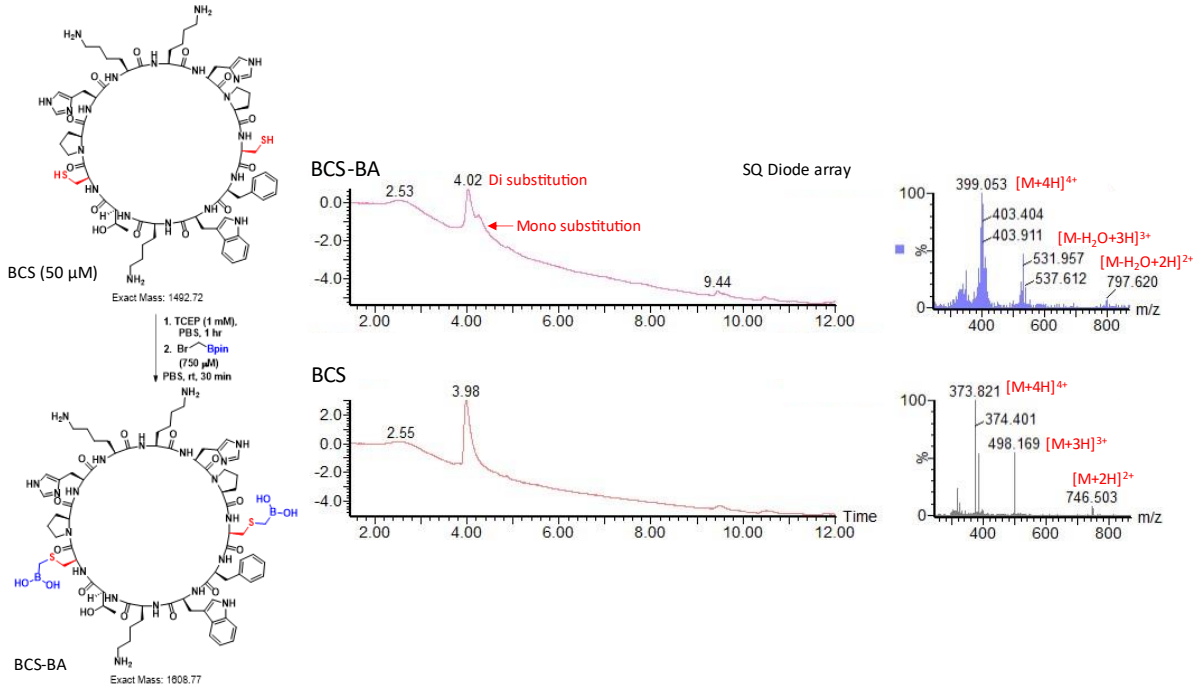


Figure 4.17. Reaction of 50 µM BCS with 1 mM TCEP followed by 750 µM $BrCH_2Bpin$ showing the LC-MS profile and the respective MS spectra.

4.4.7.7 *WGA (sialic acid binder)*

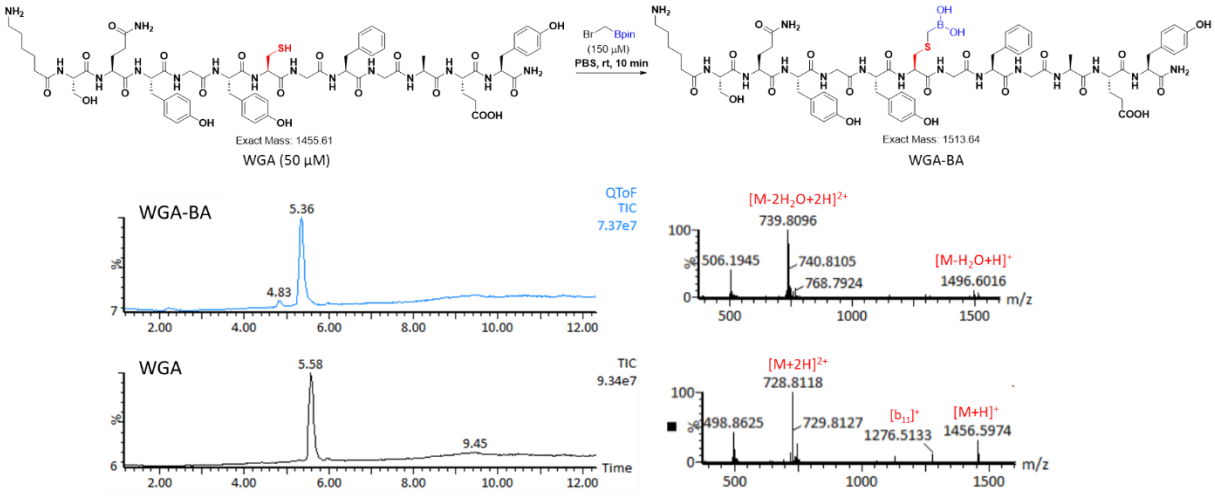


Figure 4.18. Reaction of 50 µM WGA with 150 µM $BrCH_2Bpin$ showing the LC-MS profile and the respective MS spectra.

4.4.7.8 *p53 (tumor suppressor)*

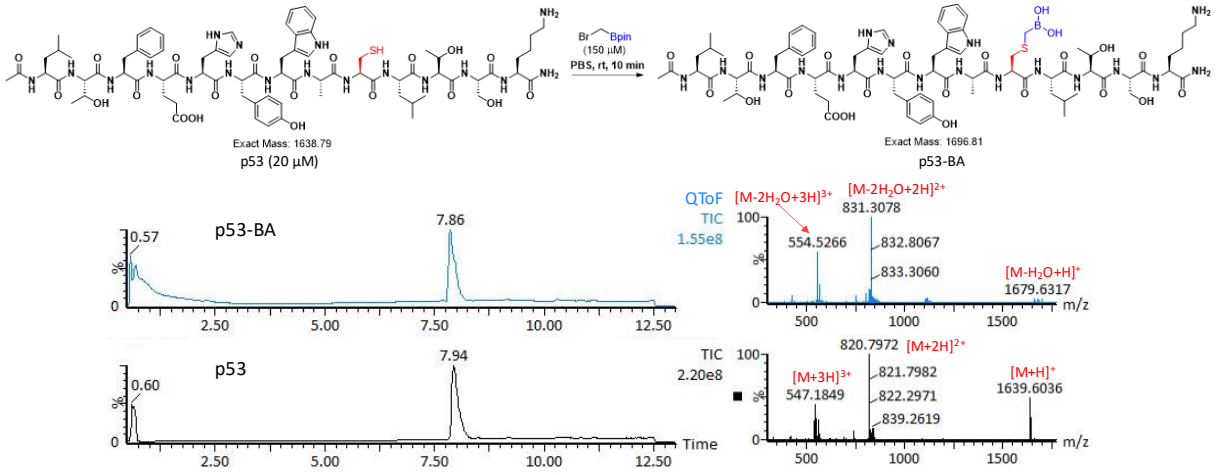


Figure 4.19. Reaction of 20 µM p53 with 150 µM $BrCH_2Bpin$ showing the LC-MS profile and the respective MS spectra.

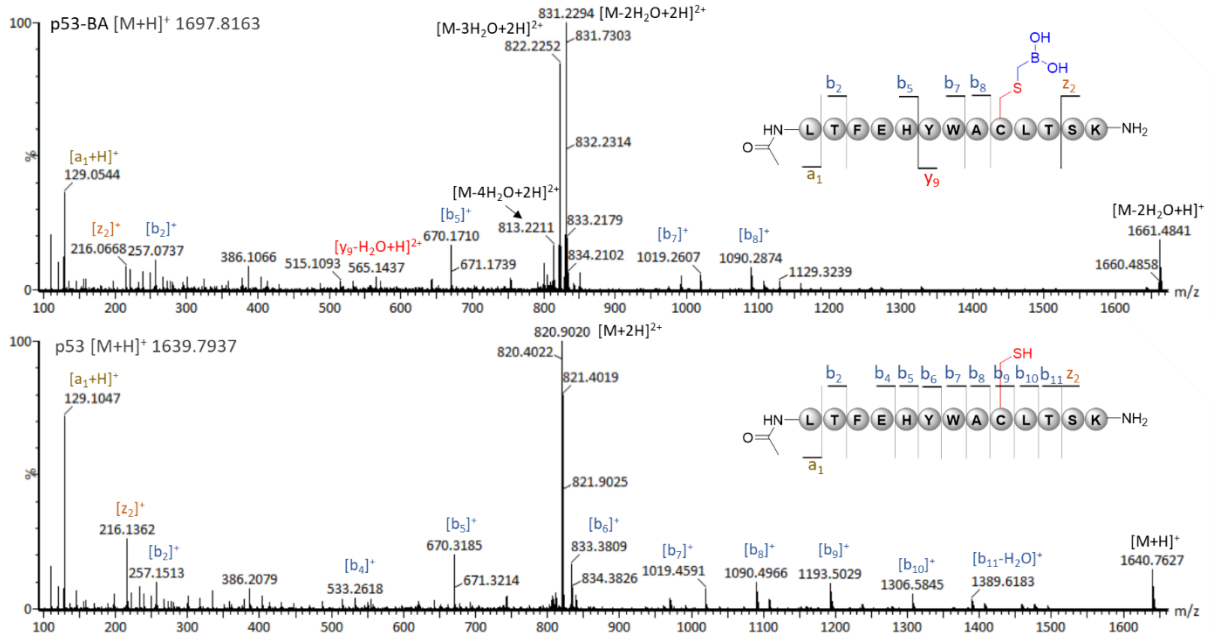


Figure 4.20. MS/MS spectra comparing p53 with p53-BA.

4.4.7.9 *PDL1-D (programmed death ligand)*

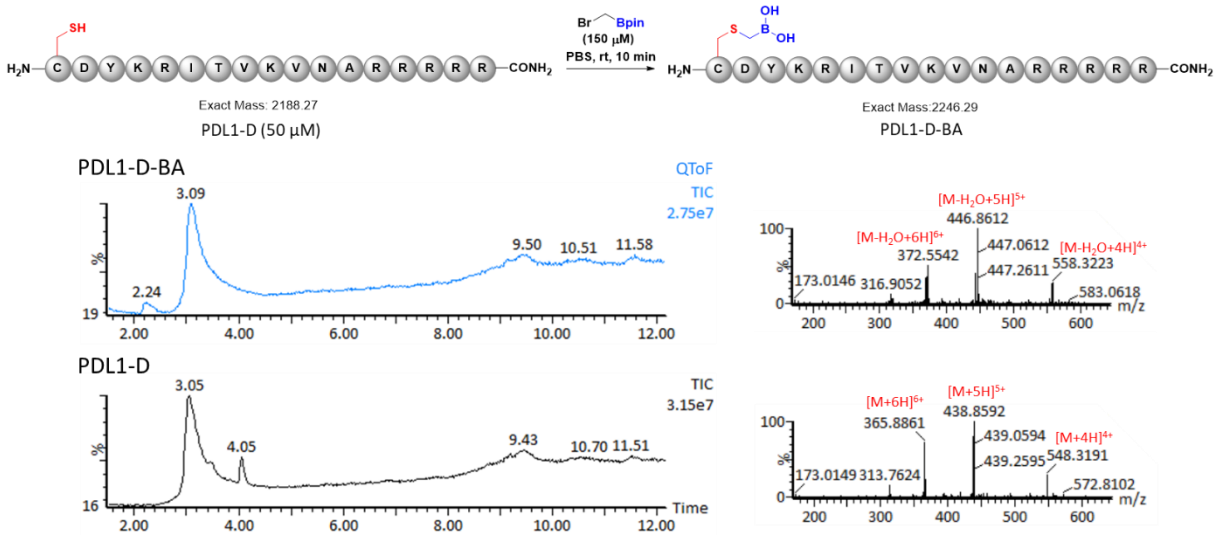


Figure 4.21. Reaction of 100 µM PDL1-D with 150 µM $BrCH_2Bpin$ showing the LC-MS profile and the respective MS spectra.

4.4.7.10 *PDL1-C (programmed death ligand)*

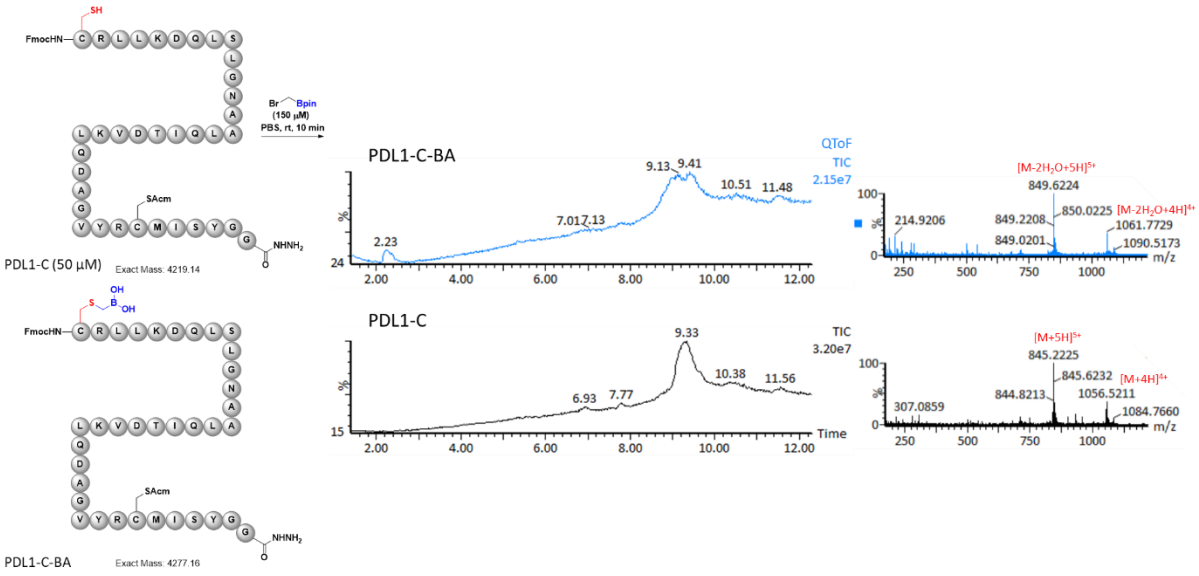


Figure 4.22. Reaction of 50 µM PDL1-C with 150 µM $BrCH_2Bpin$ showing the LC-MS profile and the respective MS spectra.

4.4.7.11 *PDL1-A-D*

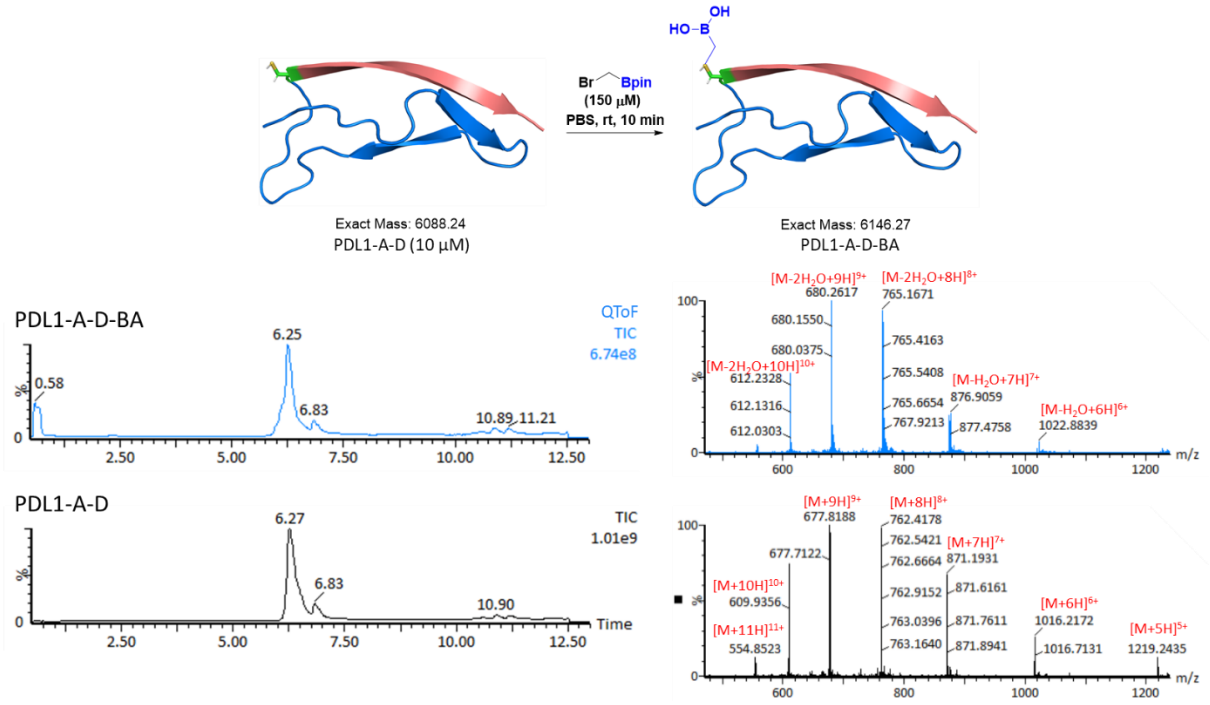


Figure 4.23. Reaction of 10 µM PDL1-A-D with 150 µM $BrCH_2Bpin$ showing the deconvoluted MS spectra.

4.4.7.12 *Lysozyme*

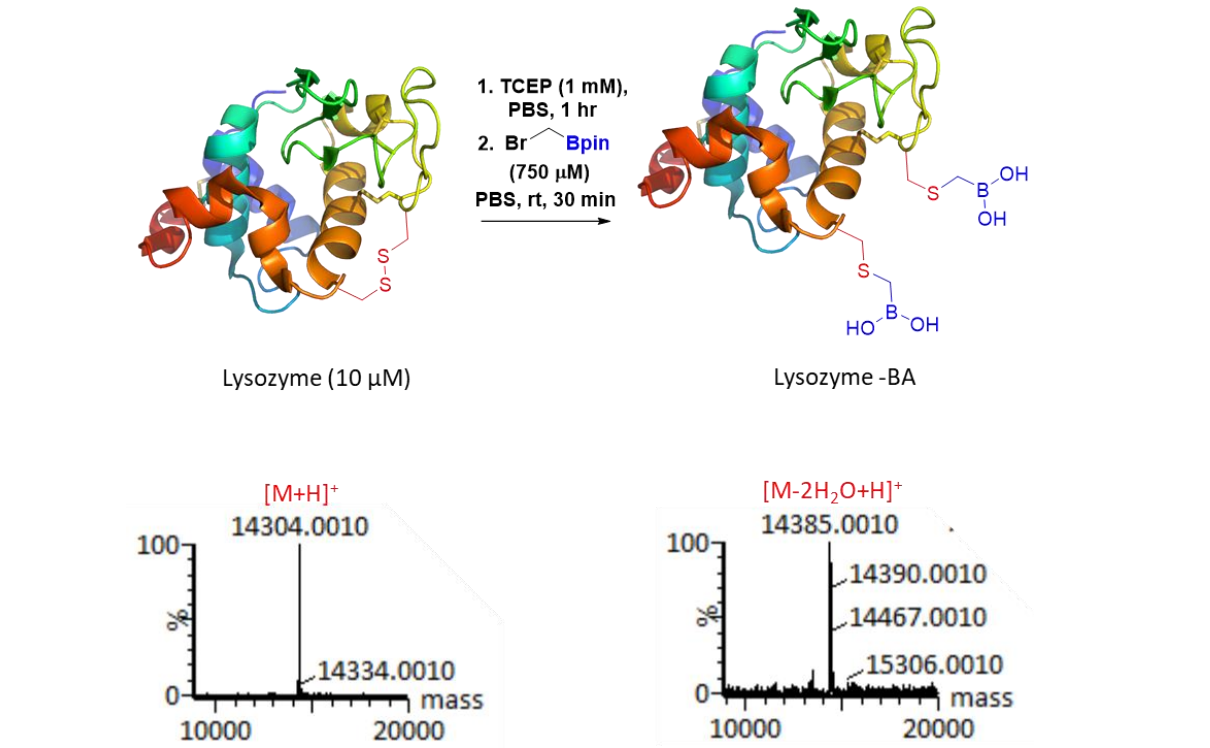


Figure 4.24. Reaction of 10 µM Lysozyme with 1 mM TCEP followed by 750 µM $BrCH_2Bpin$ showing the deconvoluted MS spectra.

4.4.7.13 *BSA*

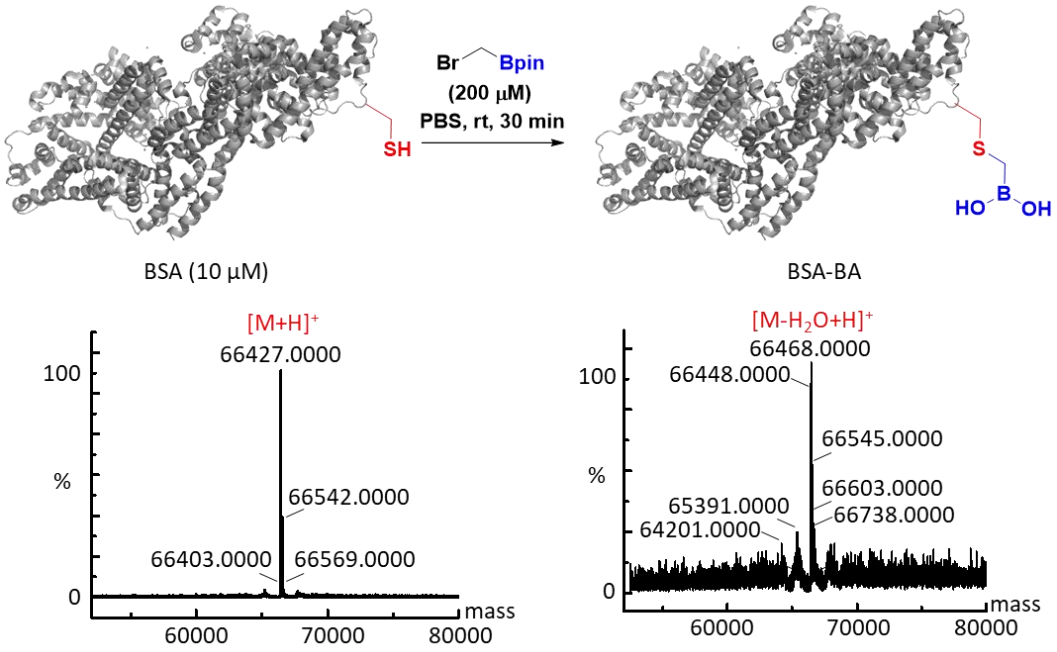
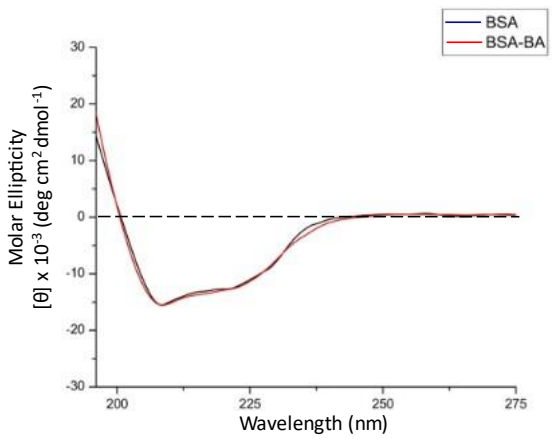


Figure 4.25. Reaction of 10 µM BSA with 200 µM $BrCH_2Bpin$ showing the deconvoluted MS spectra.

4.4.8 *SDS-PAGE AND CIRCULAR DICHROISM ANALYSIS OF BSA CONJUGATION*

SDS-PAGE is a separation technique that separates proteins on the basis of their mass. SDS-PAGE is used to monitor the purification process and to determine the homogeneity and apparent molecular mass of proteins. SDS–PAGE was performed on an electrophoresis apparatus (Bio-Rad Laboratories), with a 12% (*w*/*v*) acrylamide separating gel and a 5% (*w*/*v*) stacking Quick SDS-PAGE Gel Preparation Kit, according Laemmli's method[23]. SDS-PAGE was performed by electrophoresis, at a constant voltage of 100 V, when the bands entered the separation gel, we adjusted the voltage to 120 V for electrophoresis. 20 µL of a 10 µM protein solution was dropcasted onto the wells. Following the run, the protein bands in the gel were visualized by in-gel fluorescence followed by staining with Coomassie brilliant blue dye R-250 for 60 min.



CD-spectra were recorded using a quartz cell with a 1 mm path length between 200 and 250 nm at room temperature. A 10 µM protein concentration was taken in TFE/PBS (10 mM, pH 7.4) mixed solvent (1:1, v/v). Three scans with a scan speed of 50 nm/min were averaged for each measurement. CD spectra were expressed as the mean residue ellipticity. No loss in structure of BSA was observed due to the substitution of boronic acid.

Figure 4.26. CD spectra of BSA and BSA-BA.

4.4.9 *MULTIVALENT INSTALLATION OF BORONIC ACIDS*

A short peptide sequence (PCGCSCPCI-NH_2) was synthesized keeping two, three and four Cys free while the other Cys residues were protected with acetamidomethyl (Acm) protecting groups to confirm the installation of bis, tris and tetra boronic acid handles respectively. Gratifyingly installation of multiple boronic acid residues could be achieved with this reaction strategy using TCEP as a disulfide reducing agent. The reaction scheme and confirmation by LC-MS has been shown in the Figure 4.25-4.27.

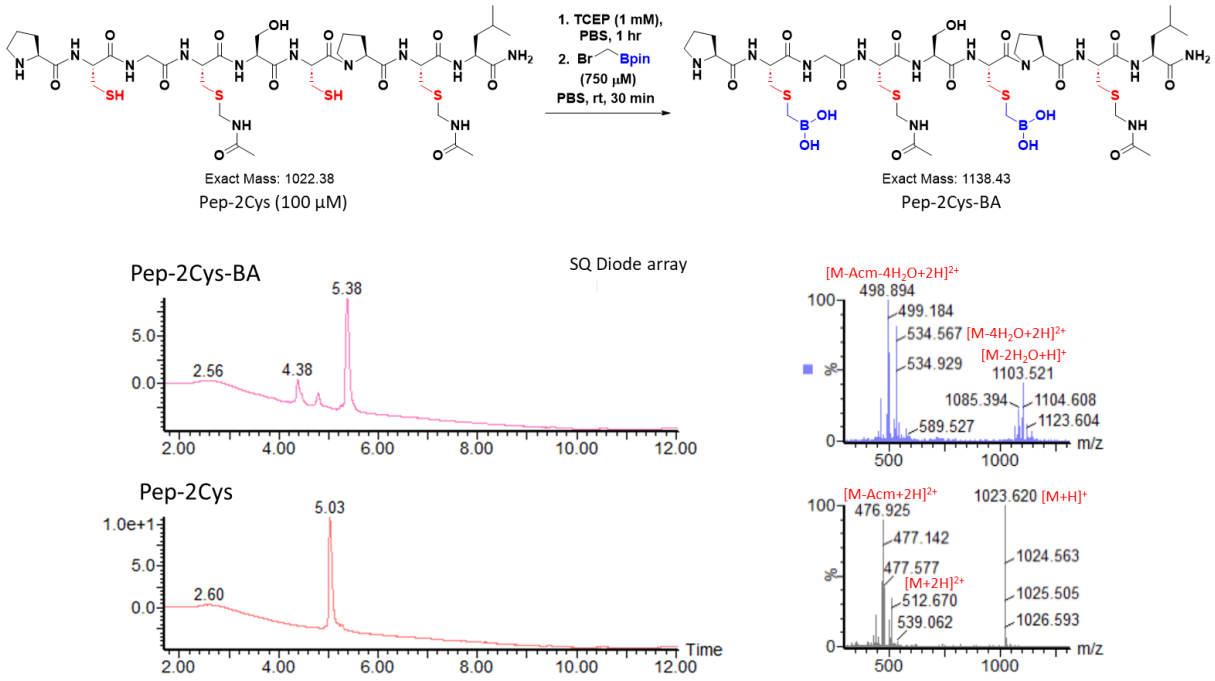


Figure 4.27. Reaction of 100 µM Pep-2Cys (having 2 free Cys residue) with 1 mM TCEP followed by 750 µM $BrCH_2Bpin$ showing the LC-MS profile and the respective MS spectra.

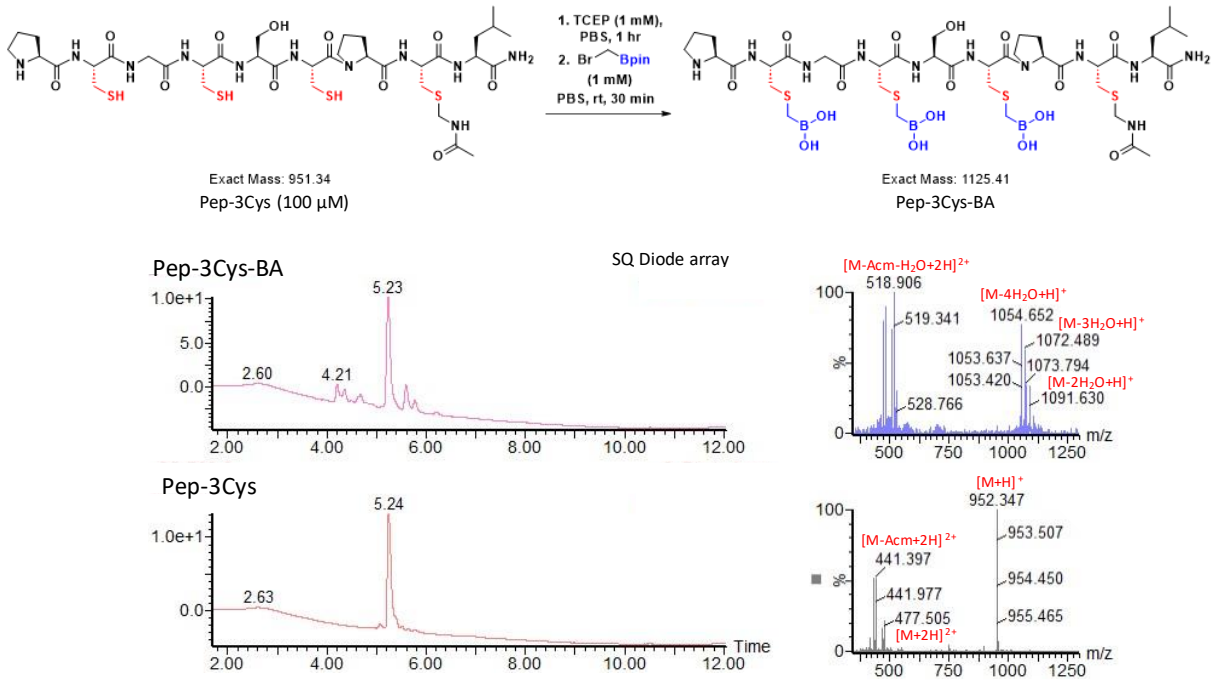


Figure 4.28. Reaction of 100 µM Pep-3Cys (having 3 free Cys residue) with 1 mM TCEP followed by 1 mM $BrCH_2Bpin$ showing the LC-MS profile and the respective MS spectra.

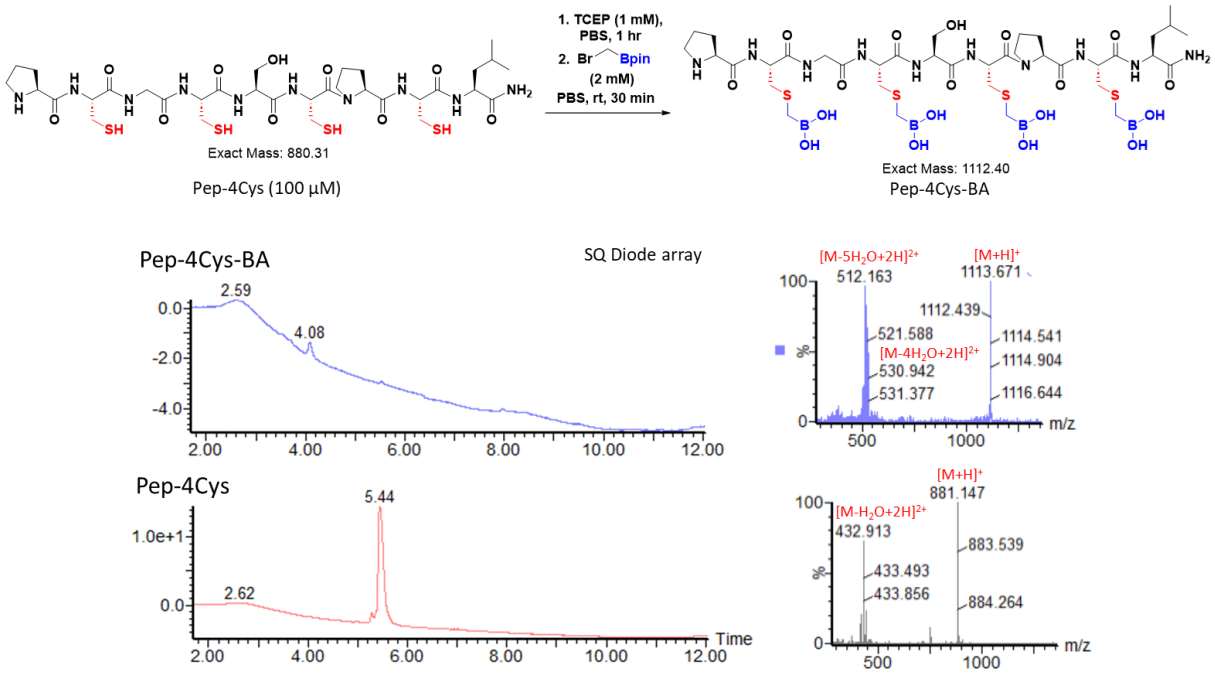


Figure 4.29. Reaction of 100 µM Pep-3Cys (having 3 free Cys residue) with 1 mM TCEP followed by 2 mM $BrCH_2Bpin$ showing the LC-MS profile and the respective MS spectra.

4.5. *NMR DATA*

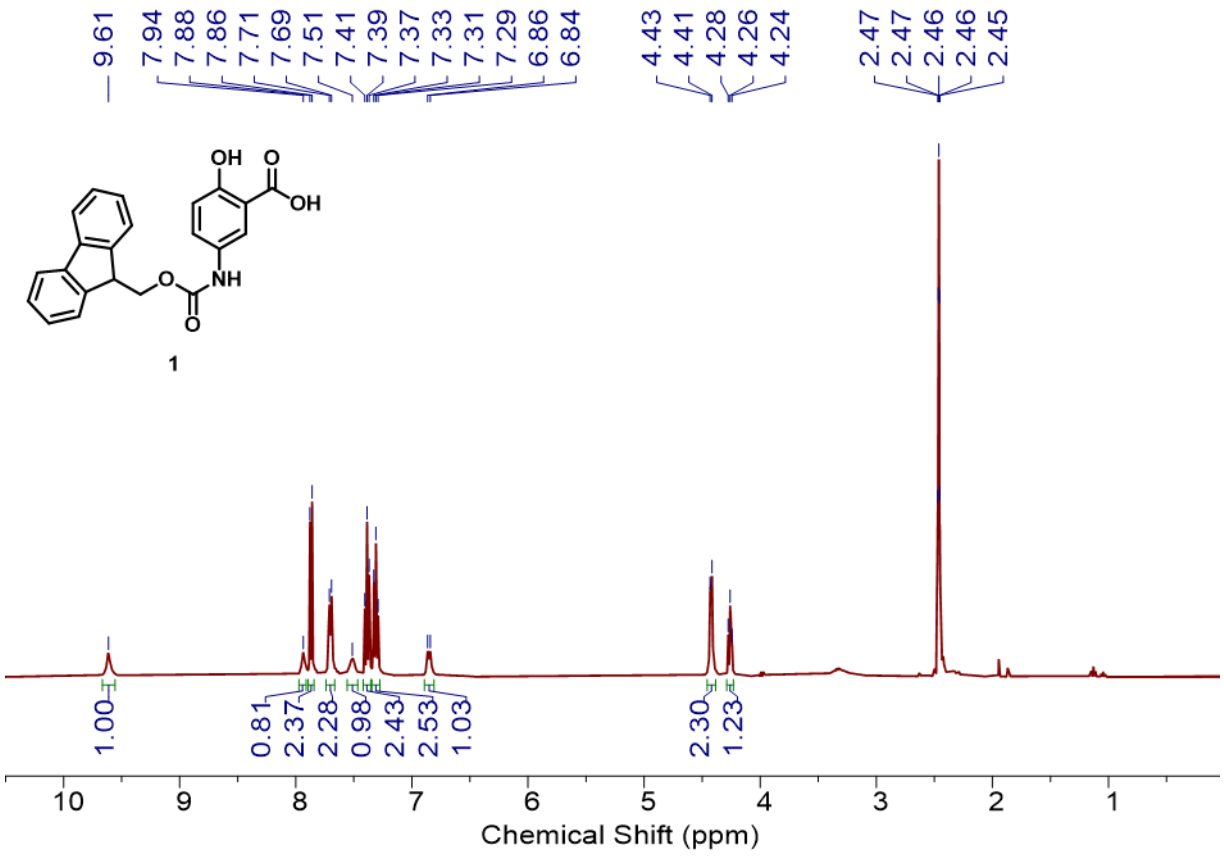


Figure 4.30: ^1H NMR of **1** in DMSO-d_6, 400 MHz.

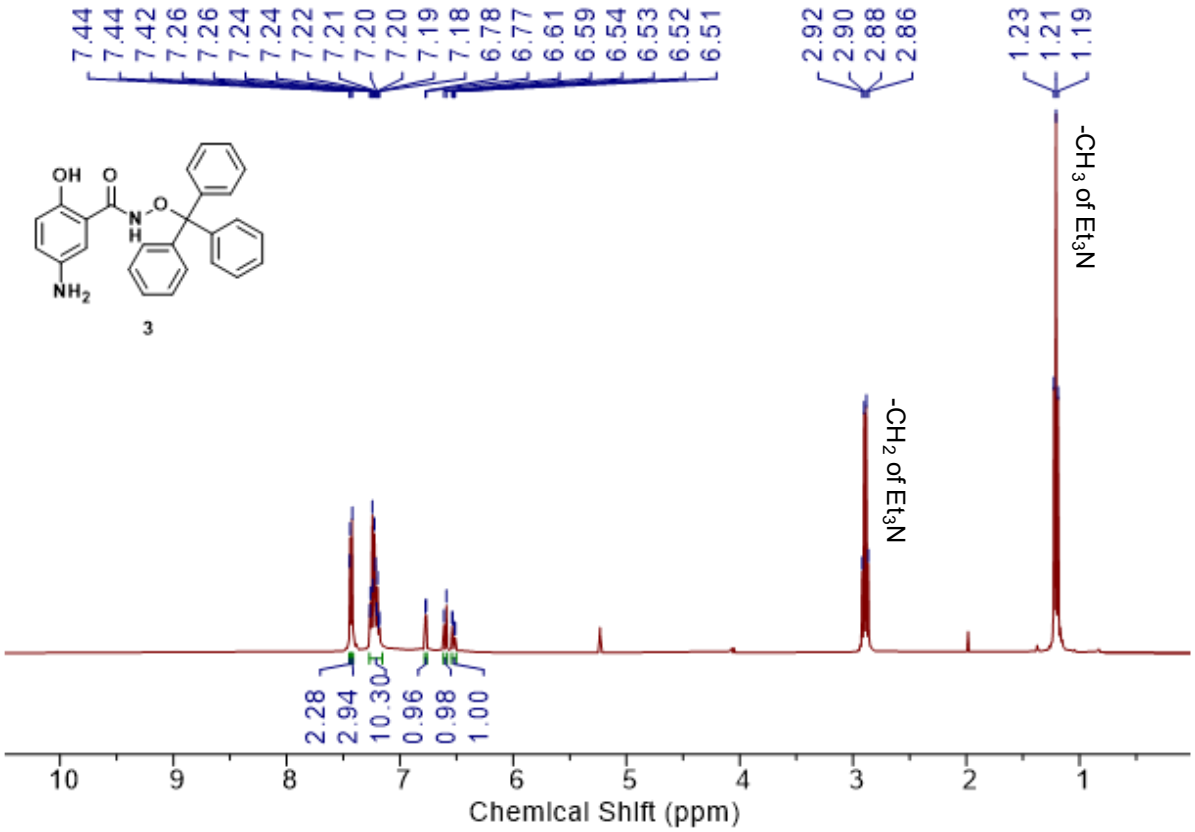
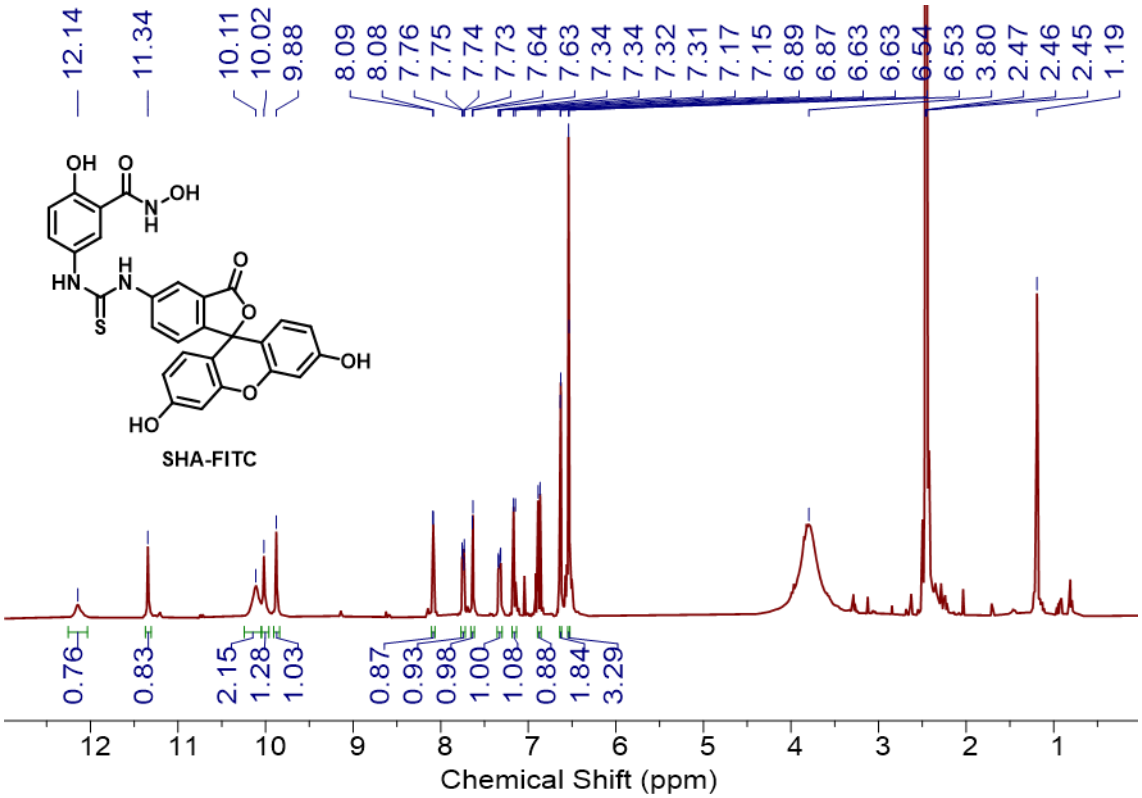


Figure 4.31: ^{1}H NMR of **3** in $CDCl_3$, 400 MHz.



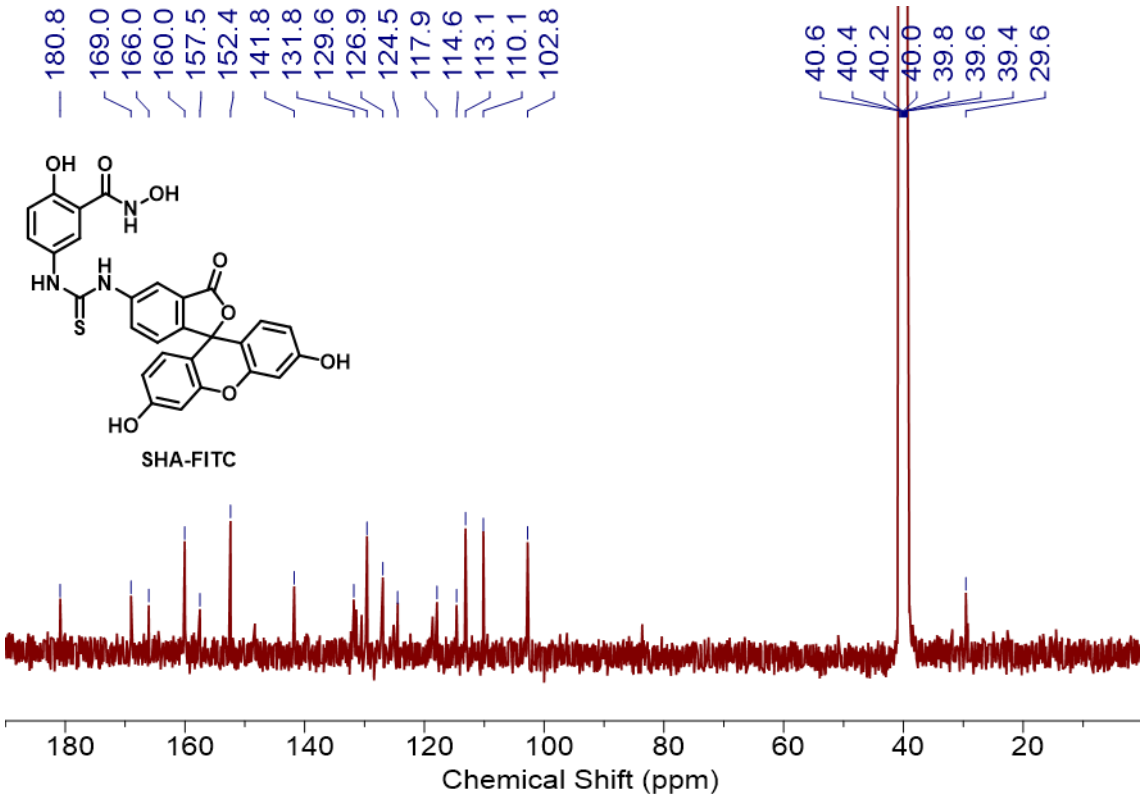


Figure 4.32: ^{1}H NMR of **SHA-FITC** in DMSO-d_6, 400 MHz.

4.6 *REFERENCES*

(1) Mollner, T. A.; Isenegger, P. G.; Josephson, B.; Buchanan, C.; Lercher, L.; Oehlrich, D.; Hansen, D. F.; Mohammed, S.; Baldwin, A. J.; Gouverneur, V.; Davis, B. G. Post-Translational Insertion of Boron in Proteins to Probe and Modulate Function. *Nat Chem Biol* **2021**, *17* (12), 1245–1261. https://doi.org/10.1038/s41589-021-00883-7.

(2) Ballatore, C.; Huryn, D. M.; Smith III, A. B. Carboxylic Acid (Bio)Isosteres in Drug Design. *ChemMedChem* **2013**, *8* (3), 385–395. https://doi.org/https://doi.org/10.1002/cmdc.201200585.

(3) Silva, M. P.; Saraiva, L.; Pinto, M.; Sousa, M. E. Boronic Acids and Their Derivatives in Medicinal Chemistry: Synthesis and Biological Applications. *Molecules* **2020**, *25* (18), 4323. https://doi.org/10.3390/molecules25184323.

(4) Pal, A.; Bérubé, M.; Hall, D. G. Design, Synthesis, and Screening of a Library of Peptidyl Bis(Boroxoles) as Oligosaccharide Receptors in Water: Identification of a Receptor for the Tumor Marker TF-Antigen Disaccharide. *Angewandte Chemie International Edition* **2010**, *49* (8), 1492–1495. https://doi.org/10.1002/anie.200906620.

(5) Zhu, F.; Miller, E.; Powell, W. C.; Johnson, K.; Beggs, A.; Evenson, G. E.; Walczak, M. A. Umpolung AlaB Reagents for the Synthesis of Non-Proteogenic Amino Acids, Peptides and Proteins**. *Angewandte Chemie International Edition* **2022**, *61* (31), e202207153. https://doi.org/https://doi.org/10.1002/anie.202207153.

(6) Fu, X. P.; Yuan, Y.; Jha, A.; Levin, N.; Giltrap, A. M.; Ren, J.; Mamalis, D.; Mohammed, S.; Davis, B. G. Stereoretentive Post-Translational Protein Editing. *ACS Cent Sci* **2023**, *9* (3), 405–416. https://doi.org/10.1021/acscentsci.2c00991.

(7) Jing, R.; Powell, W. C.; Fisch, K. J.; Walczak, M. A. Desulfurative Borylation of Small Molecules, Peptides, and Proteins. *J Am Chem Soc* **2023**, *145* (41), 22354–22360. https://doi.org/10.1021/jacs.3c09081.

(8) Chatterjee, S.; Bandyopadhyay, A. Cysteine-Selective Installation of Functionally Diverse Boronic Acid Probes on Peptides. *Org Lett* **2023**, *25* (13), 2223–2227. https://doi.org/10.1021/acs.orglett.3c00386.

(9) Brustad, E.; Bushey, M. L.; Lee, J. W.; Groff, D.; Liu, W.; Schultz, P. G. A Genetically Encoded Boronate-Containing Amino Acid. *Angewandte Chemie - International Edition* **2008**, *47* (43), 8220–8223. https://doi.org/10.1002/anie.200803240.

(10) Matteson, D. S. α-Halo Boronic Esters: Intermediates for Stereodirected Synthesis. *Chem Rev* **1989**, *89* (7), 1535–1551. https://doi.org/10.1021/cr00097a009.

(11) Chatterjee, S.; Chowdhury, A.; Saproo, S.; Mani Tripathi, N.; Naidu, S.; Bandyopadhyay, A. Capturing Sialyl-glycan on Live Cancer Cells by Tailored Boronopeptide**. *Chemistry – A European Journal* **2024**, *30* (8). https://doi.org/10.1002/chem.202303327.

(12) Zhu, L.; Shabbir, S. H.; Gray, M.; Lynch, V. M.; Sorey, S.; Anslyn, E. V. A Structural Investigation of the N−B Interaction in an o -(N , N -Dialkylaminomethyl)Arylboronate System. *J Am Chem Soc* **2006**, *128* (4), 1222–1232. https://doi.org/10.1021/ja055817c.

(13) Graham, D. E.; Harich, K. C.; White, R. H. Reductive Dehalogenation of Monobromobimane by Tris(2-Carboxyethyl)Phosphine. *Anal Biochem* **2003**, *318* (2), 325–328. https://doi.org/https://doi.org/10.1016/S0003-2697(03)00239-2.

(14) Shafer, D. E.; Inman, J. K.; Lees, A. Reaction of Tris(2-Carboxyethyl)Phosphine (TCEP) with Maleimide and α-Haloacyl Groups: Anomalous Elution of TCEP by Gel Filtration. *Anal Biochem* **2000**, *282* (1), 161–164. https://doi.org/https://doi.org/10.1006/abio.2000.4609.

(15) Getz, E. B.; Xiao, M.; Chakrabarty, T.; Cooke, R.; Selvin, P. R. A Comparison between the Sulfhydryl Reductants Tris(2- Carboxyethyl)Phosphine and Dithiothreitol for Use in Protein Biochemistry. *Anal Biochem* **1999**, *273* (1), 73–80. https://doi.org/10.1006/abio.1999.4203.

(16) Harper, S.; Speicher, D. W. Purification of Proteins Fused to Glutathione S-Transferase BT - Protein Chromatography: Methods and Protocols; Walls, D., Loughran, S. T., Eds.; Humana Press: Totowa, NJ, 2011; pp 259–280. https://doi.org/10.1007/978-1-60761-913-0_14.

(17) Chang, H. N.; Liu, B. Y.; Qi, Y. K.; Zhou, Y.; Chen, Y. P.; Pan, K. M.; Li, W. W.; Zhou, X. M.; Ma, W. W.; Fu, C. Y.; Qi, Y. M.; Liu, L.; Gao, Y. F. Blocking of the PD-1/PD-L1 Interaction by a D -Peptide Antagonist for Cancer Immunotherapy. *Angewandte Chemie - International Edition* **2015**, *54* (40), 11760–11764. https://doi.org/10.1002/anie.201506225.

(18) Stolowitz, M. L.; Ahlem, C.; Hughes, K. A.; Kaiser, R. J.; Kesicki, E. A.; Li, G.; Lund, K. P.; Torkelson, S. M.; Wiley, J. P. Phenylboronic Acid-Salicylhydroxamic Acid Bioconjugates. 1. A Novel Boronic Acid Complex for Protein Immobilization. *Bioconjug Chem* **2001**, *12* (2), 229–239. https://doi.org/10.1021/bc0000942.

(19) Wiley, J. P.; Hughes, K. A.; Kaiser, R. J.; Kesicki, E. A.; Lund, K. P.; Stolowitz, M. L. Phenylboronic Acid - Salicylhydroxamic Acid Bioconjugates. 2. Polyvalent Immobilization of Protein Ligands for Affinity Chromatography. *Bioconjug Chem* **2001**, *12* (2), 240–250. https://doi.org/10.1021/bc000086l.

(20) Varki, A. Biological Roles of Glycans. *Glycobiology* **2017**, *27* (1), 3–49. https://doi.org/10.1093/glycob/cww086.

(21) Chowdhury, A.; Gour, V.; Das, B. K.; Chatterjee, S.; Bandyopadhyay, A. Rapid and Highly Productive Assembly of a Disulfide Bond in Solid-Phase Peptide Macrocyclization. *Org Lett* **2023**, 3–7. https://doi.org/10.1021/acs.orglett.3c00078.

(22) Chang, H.-N.; Liu, B.-Y.; Qi, Y.-K.; Zhou, Y.; Chen, Y.-P.; Pan, K.-M.; Li, W.-W.; Zhou, X.-M.; Ma, W.-W.; Fu, C.-Y.; Qi, Y.-M.; Liu, L.; Gao, Y.-F. Blocking of the PD-1/PD-L1 Interaction by a D -Peptide Antagonist for Cancer Immunotherapy . *Angewandte Chemie* **2015**, *127* (40), 11926–11930. https://doi.org/10.1002/ange.201506225.

(23) LAEMMLI, U. K. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* **1970**, *227* (5259), 680–685. https://doi.org/10.1038/227680a0.

Curriculum Vitae

1. **Name:** Saurav Chatterjee
2. **Father's Name:** Kabiram Chatterjee
3. **Mother's Name:** Sipra Chatterjee
4. **DOB:** 25.11.1993
5. **Address:** 70, Hill View Park (North), P.O. – Chelidanga, Asansol – 713304, West Bengal, India.
6. **Email id:** saurav.19cyz0008@iitrpr.ac.in, sauravda.chatterjee25@gmail.com
7. **Contact No.:** +91 933 036 8248
8. **Nationality:** Indian
9. **Religion:** Hindu
10. **Category:** Unreserved
11. **Marital Status:** Unmarried
12. **Languages known:** Bengali, English, Hindi
13. **Educational Qualifications**

Sl. No.	Degree	Name of the Institution	Board/ University	Year of Passing	Marks	Class/ Division
1.	10th	St. Vincent's High School	CISCE	2010	89.71%	A
2.	12th	Burnpur Riverside School	CBSE	2012	89.16%	A
3.	B.Sc.	St. Xavier's College (Autonomous)	University of Calcutta	2015	76.00% CGPA 7.48	I
4.	M.Sc.	Rajabajar Science College	University of Calcutta	2017	72.60% CGPA 4.63	I
5.	Ph.D.	Indian Institute of Technology, Ropar	--	2024	CGPA 8.40	--

14. **Research Experience:**

(1) January 2017 – May 2017: A project report submitted for the degree of M.Sc in Chemistry entitled 'An efficient green synthesis of 1*H*-indol-4-ol derivatives using magnetically separable carbocatalyst Fe_3O_4@GO under solvent free condition' under the supervision of Prof. A. Pramanik, University of Calcutta.

(2) October 2017 – February 2018: Research work entitled 'Synthesis and biological evaluation of polyhydroxylated oxindole derivatives as potential antileishmanial agent' under the supervision of Prof. S. S. Adhikari, University of Calcutta.

(3) April 2018 – March 2019: 1-year JRF in the project entitled 'Synthetic modification of fluorescent probe NAP-RD for selective detection of stannous ion in biological system' funded by **Colgate-Palmolive Company, USA** under the supervision of Prof. S. S. Adhikari, University of Calcutta.

(4) July 2019 – July 2024: Pursuing Ph.D. in Chemical Biology at IIT Ropar under the supervision of Dr. Anupam Bandyopadhyay, working on engineering peptides for therapeutic applications.

15. **Patents published:**

(1) SYNTHESIS AND ANTI-LEISHMANIAL PROFILE OF NOVEL WATER SOLUBLE FERROCENYLQUINOLINE DERIVATIVES, C. Pal, S. Adhikari, D. Mukherjee, Md Yousuf, S. Dey; **S. Chatterjee**, Rajabazar Science College, Kolkata (Indian Patent Application No. 201831027970 of 25.07.2018).

(2) HALOMETHYL BORONIC ACID FOR CHEMOSELECTIVE AND SITE-SELECTIVE INSTALLATION OF BORONIC ACID AND PRODUCTION OF DEHYDROALANINE IN PEPTIDE AND PROTEIN, **S. Chatterjee**, A. Bandyopdhyay, Indian Institute of Technology (IIT) Ropar, Punjab, India (Patent Application Number 202411017017).

16. **Papers published:**

(1) Yousuf, M.; Mukherjee, D.; Dey, S.; **Chatterjee, S.**; Pal, A.; Sarkar, B.; Pal, C.; Adhikari, S. Synthesis and Biological Evaluation of Polyhydroxylated Oxindole Derivatives as Potential Antileishmanial Agent. *Bioorg Med Chem Lett* **2018**, *28* (6), 1056–1062. https://doi.org/10.1016/j.bmcl.2018.02.023.

(2) **Chatterjee, S.**; Anslyn, E. V.; Bandyopadhyay, A. Boronic Acid Based Dynamic Click Chemistry: Recent Advances and Emergent Applications. *Chem Sci* **2021**, *12* (5), 1585–1599. https://doi.org/10.1039/d0sc05009a.

(3) Chowdhury, A.*; **Chatterjee, S.***; Pongen, A.; Sarania, D.; Tripathi, N. M.; Bandyopadhyay, A. NSite-Selective, Chemical Modification of Protein at Aromatic Side Chain and Their Emergent Applications. *Protein Pept Lett* **2021**, *28*, 1–21. https://doi.org/10.2174/0929866528666210129152535. (* refers to equal contribution)

(4) **Chatterjee, S.**; Tripathi, N. M.; Bandyopadhyay, A. The Modern Role of Boron as a 'Magic Element' in Biomedical Science: Chemistry Perspective. *Chemical Communications* **2021**, *57* (100), 13629–13640. https://doi.org/10.1039/D1CC05481C.

(5) Bhatt Mitra, J.; **Chatterjee, S.**; Kumar, A.; Bandyopadhyay, A.; Mukherjee, A. Integrating a Covalent Probe with Ubiquicidin Fragment Enables Effective Bacterial Infection Imaging. *RSC Med Chem* **2022**, *13* (10), 1239–1245. https://doi.org/10.1039/D2MD00190J.

(6) Chowdhury, A.; Gour, V.; Das, B. K.; **Chatterjee, S.**; Bandyopadhyay, A. Rapid and Highly Productive Assembly of a Disulfide Bond in Solid-Phase Peptide Macrocyclization. *Org Lett* **2023**, 3–7. https://doi.org/10.1021/acs.orglett.3c00078.

(7) **Chatterjee, S.**; Bandyopadhyay, A. Synthesis and Characterization of Two Potential Impurities (Des-Ethyl-Ganirelix) Generated in the Ganirelix Manufacturing Process. *Journal of Peptide Science* **2023**, *29* (9), 1–9. https://doi.org/10.1002/psc.3489.

(8) **Chatterjee, S.**; Bandyopadhyay, A. Cysteine-Selective Installation of Functionally Diverse Boronic Acid Probes on Peptides. *Org Lett* **2023**, *25* (13), 2223–2227. https://doi.org/10.1021/acs.orglett.3c00386.

(9) Chowdhury, A.; **Chatterjee, S.**; Kushwaha, A.; Nanda, S.; Dhilip Kumar, T. J.; Bandyopadhyay, A. Sulfonyl Diazaborine 'Click' Chemistry Enables Rapid and Efficient Bioorthogonal Labeling. *Chemistry – A European Journal* **2023**, *29* (41). https://doi.org/10.1002/chem.202300393.

(10) Jadav, R.; Kameriya, R.; **Chatterjee, S.**; Gour, V.; Purohit, P.; Bandyopadhyay, A. Identification, Synthesis, and Characterization of an Unprecedented N -(2-carboxyethyl) Adduct Impurity in an Injectable Ganirelix Formulation. *Journal of Peptide Science* **2024**, *30* (5). https://doi.org/10.1002/psc.3564.

(11) **Chatterjee, S.**; Chowdhury, A.; Saproo, S.; Mani Tripathi, N.; Naidu, S.; Bandyopadhyay, A. Capturing Sialyl-glycan on Live Cancer Cells by Tailored Boronopeptide**. *Chemistry – A European Journal* **2024**, *30* (8). https://doi.org/10.1002/chem.202303327.

(12) Bhatt Mitra, J.; **Chatterjee, S.**; Kumar, A.; Khatoon, E.; Chandak, A.; Rakshit, S.; Bandyopadhyay, A.; Mukherjee, A. Expanding a Peptide-Covalent Probe Hybrid for PET Imaging of S. Aureus Driven Focal Infections. *EJNMMI Radiopharm Chem* **2024**, *9* (1), 25. https://doi.org/10.1186/s41181-024-00252-4.

17. **Conferences Attended:**

(1) International Symposium on Recent Trends and Future Opportunities in Pharmaceuticals
November 10-12, 2022
Venue: National Institute of Pharmaceutical Education and Research (NIPER), Mohali, Punjab
Poster presented

(2) CRIKC Symposium Recent Advances in Bioorganic and Medicinal Chemistry 2022
19th November, 2022
Venue: National Institute of Pharmaceutical Education and Research (NIPER), Mohali, Punjab
Poster presented and awarded **Best Poster**

(3) 9th Indian Peptide Symposium and Indian Peptide Show
February 23-25, 2023
Venue: BITS Pilani, K. K. Birla Goa Campus, Goa
Poster presented and awarded **Best Poster**

(4) 4th Student IPS Symposium sIPS – 2024
4th March, 2024
Venue: Gujarat Biotechnology University, GIFT City, Gujarat
Poster presented and awarded **Best Poster**

(5) ChemFest 2024, 4th March, 2024
Venue: Indian Institute of Technology, Ropar, Punjab
Poster presented and awarded **Best Poster**

(6) European Peptide Synthesis Conference 2024
29th April – 1st May, 2024
Venue: Prague Congress Centre, Prague, Czech Republic
Poster presented

18. **Analytical skills:** Mass spec (LCMS, MS/MS), NMR, HPLC, UV-Vis.

19. **Achievements:** GATE-2019, SET(WB)-2018 qualified.

20. **Computer Skills:** JAVA, FORTRAN 77, ChemDraw, SciFinder, MNova, Pymol, FlowJo, Gaussian.

21. **Hobbies:** Reading books, and playing cricket & carrom.

22. **Website links:** Lab webpage, Google scholar, LinkedIn, Twitter, ORCiD, Scopus