10-15-2022

Synthesis of Novel DNA-binding Polyamides to Prevent Cancer-related Gene Overexpression

Huy Q. Nguyen
University of Missouri-St. Louis, hqnwv3@umsystem.edu

Synthesis of Novel DNA-binding Polyamides to Prevent Cancer-related Gene Overexpression

Huy Quang Nguyen

Bachelor of Science in Biochemistry and Biotechnology, UMSL, 2016

Master of Science in Chemistry, UMSL, 2021

A Dissertation Submitted to the Graduate school of

The University of Missouri-Saint Louis

In partial fulfillment of the requirement for the degree of

Doctor of Philosophy in Chemistry

with an emphasis in Organic Chemistry

December 2022

Advisory Committee

Professor James K. Bashkin, Ph.D. (Chairperson)

Professor Eike B. Bauer, Ph.D.

Professor James S. Chickos, Ph.D.

Professor Alexei V. Demchenko, Ph.D.

Abstract

Pyrrole-imidazole polyamides (PIPs) are nanomolecular compounds designed to fit in the tight space of DNA minor grooves. As analogs of Netropsin and Distamycin A, PIPs are specifically designed to recognize the base pairs of DNA sequences. PIPs have many biological applications, such as regulating gene expression, biochemistry pathway regulations, and suppressing the development of cancer cells. SETMAR gene is the chimeric fusion of the SET domain with the mariner transposase. Its protein (Metnase) has functions involving DNA repairs in the NHEJ pathway, regulating gene expression, DNA decatenation, etc. Despite not having an active transposable element, SETMAR still has an unknown function in cancer development. Some reports show that SETMAR mRNA increases by 70 times in glioblastoma cancer cells, as well as in other types of cancer cells. Understanding the mechanism of SETMAR in the cancer cell is essential, and suppressing the overexpression of SETMAR is a goal for anti-cancer development.

Our goal in this project was to design new sequences of polyamides based on the previously discovered conserved SETMAR gene binding locations at the TIR (Terminal Inverted Repeat) sequence. We have tested the binding affinity of our polyamides with our collaborators' assistance through surface plasmon resonance, fluorescence anisotropy, and competitive DNA binding. We modified the polyamide sequence by substituting pyrrole with β-alanine to reduce the rigidity of the continuous chain of heterocycles and introduced an additional charge to the N-terminus of our polyamide sequences.

For my family

Acknowledgments

First and foremost, I would like to express my appreciation to my advisor chair Dr. James K. Bashkin. He has been the most inspiring and patient teacher I have ever seen. He has continuously educated and supported us whenever we needed him or whatever trouble I have been having over the past years in his research group. He always shared his knowledge and experience in academics, research, and teaching experiments. He allowed me to try new challenges even when I thought I was not smart enough to try them. I was blessed to be able to do my organic chemistry and analytical chemistry projects and participate in cell culture projects.

I want to show my appreciation to my honorable committee members: Dr. Eike Bauer, Dr. Alexei Demchenko, and Dr. Chickos, for their time, helpful comments, critics, and suggestions toward my research. Each committee member is a master in their field, and each has inspired me differently during my years in the Graduate Program. Additionally, I appreciate the Department Chair, Dr. Stine, for his support of our research lab. I am also grateful to have known and learned from many of the faculty in the Chemistry Department, Dr. Nichols for his knowledge of Alzheimer's and motivation, Dr. Wong for his advice in bioinformatics, Dr. Hamper for always being cheerful and assisting us with HPLC problems, Dr. D'Souza for his teaching experience, and so on. I am in debt to so many faculty and staff in this department.

I also appreciate the motivation and inspiration from Dr. Michael Mannino, Dr. Catherine Alex, Dr. Yashpal Singh, and Dr. Satsawat Visansirikul from the Demchenko lab for encouraging me to challenge myself and aim for a higher goal in academics. I want to thank you, Dr. Patricia Parker and Dr. Samoa Asigau, for inspiring me in my research in science.

Special thanks to all the old and current group members, whom I am blessed to have had: Dr. Ali Sabi, Ruth Kvistad, Nicole Miller, Afnan Hussaini, Aman Chishti, Thomas Perrot, Harper Wieldt, Jessica Thompson, Annie Migdal, Jacques Kamaru, and Zahria Patrick for being able to assist. My friends and colleagues: Deva, Saroj, Shikha, and Ganesh, thank you for always helping and being there. I want to show my appreciation to Dr. Rensheng Luo for his help with the 600 MHz NMR: he showed me how to troubleshoot, use the program for NMR analysis, and refill liquid N_2 to the NMR machine.

I want to show my appreciation to our collaborators, Dr. Millie Georgiadis and her group, Dr. W David Wilson and his colleague Dr. Anaya Paul for their assistance in studying the effects of my compounds, as well as for suggestions and ideas for my research.

I’m grateful to acknowledge the financial support from the UMSL Chemistry and Biochemistry Department for allowing me to complete my research projects. I’m thankful and appreciative to all faculty and staff at UMSL.

Last but never least, from the top of my head to my myocardium, I thank the woman that has made me a complete person, accepted me for who I am, and stood by me (safely quarantined) throughout the three times with COVID. I’m so grateful to call you my wife, Hailey. I thank my family, Ba, Mẹ, grandma, and my little sister Dr. Hao Nguyen for loving me unconditionally. Thank you to my aunt for being the role of a mom in the United States.

Table of Content

List of figures and tables

Abbreviation

A	Adenine
AA#	Amino acid
Ac_2O	acetic anhydride
AcOH	Acetic Acid
ACN	Acetonitrile
ARE	Androgen-response-element
Boc	tert-Butyloxycarbonyl
β	Beta-Alanine
Bp	Base pairs
br	broad
C	Cytosine
$CHCl_3$	Chloroform
CH_2CL_2 (or DCM)	Dichloromethane
$ClCOCCl_3$	Trichloroacetyl chloride
CFN file	CSBio save file
Chl	Chlorambucil
d	doublet
δ	chemical shift in NMR
DIPEA	N,N-Diisopropylethylamine
DMF	Dimethyl Formamide
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic acid
Et	Ethyl
EtOAc	Ethyl Acetate
Ets-1	ETS proto-Oncogene 1, Transcription Factor
G	Guanine
γ	Gamma-aminobutyric acid (GABA)

GRE_____**G**lucocorticoid **R**esponse **E**lement

HATU_____**H**exafluorophosphate **A**zabenzotriazole **T**ertramethyl **U**ronium.

HCl_____Hydrochloric acid

HCOOH_____Formic Acid

HIF_____Hypoxia-inducible factor

HTH_____Helix turn helix

HPLC_____**H**igh **P**erformance **L**iquid **C**hromatography

HPV_____human papillomavirus

HOBt_____Hydroxybenzotriazole

HRE_____Hypoxia response element

Hsmar_____homosapien mariner

Im N-methylimidazole

LC-MS_____Liquid Chromatography-Mass Spectrometry

LEF-1 _____Lymphoid-enhancer-binding-factor-1

M_____Molarity

Me_____Methyl

MeOH_____Methanol

MVA_____Measuring vessel A

MVB_____Measuring vessel B

MW_____Molecular weight

*m/z*_____Mass to charge ratio

µL_____Microliter

NH_4HNCO_2_____Ammonium formate

NHEJ_____nonhomologous end-joining

NMR_____Nuclear Magnetic Resonance

Pa(s) or PA(s)_____Polyamide(s)

PAM resin_____**P**henyl**a**cetamido**m**ethyl resin

PIP(s) ---------- **P**yrrole-**I**midazole **P**olyamide(s)

PyBOP ---------- **B**enzotriazol-1-yl**o**xytri**py**rrolidino**p**hosphonium hexafluorophosphate.

Pd/C ---------- Palladium on Carbon

Py ---------- N-Methylpyrrole

SETMAR ---------- SET domain fuses mariner (MAR) transposase gene

t ---------- triplet

T ---------- Thymine

Ta ---------- 3,3-Diamino-N-methyldipropylamine

TEA ---------- Triethylamine

TFA ---------- Trifluoroacetic acid

TFIIIA ---------- Transcription factor IIIA

TIR ---------- Terminal Inverted Repeats

TVA ---------- Transfer vessel A

TVB ---------- Transfer vessel B

UMSL ---------- University of Missouri-Saint Louis

s ---------- singlet

CHAPTER 1. INTRODUCTION

1.1. Pyrrole-Imidazole-Polyamides

1.1.1 DNA

DNA is the essential molecule that contains and encodes genetic information for all living beings. Its structure consists of two strands of nucleotides twisted around together to form a structure known as the double helix. The double helix is formed by hydrogen bonding between the Watson-Crick's base pairs linked to the sugar backbone of DNA. They pair with each other accordingly to their structure; Adenine (A) pairs with Thymine (T), while Guanine (G) pairs with Cytosine (C). B-form DNA has a right-handed helix structure, which contains a wide major groove and a narrow minor groove. (**Figure 1.1**) The minor groove of DNA provides the opportunity for a hydrogen-bond donor and hydrogen-bond acceptors. This makes the minor grooves of DNA a good target for DNA binding. The DNA binding compounds vary from enzyme and protein binding to DNA for gene expression to regulate transcription and other biological functions of DNA.[1,23]

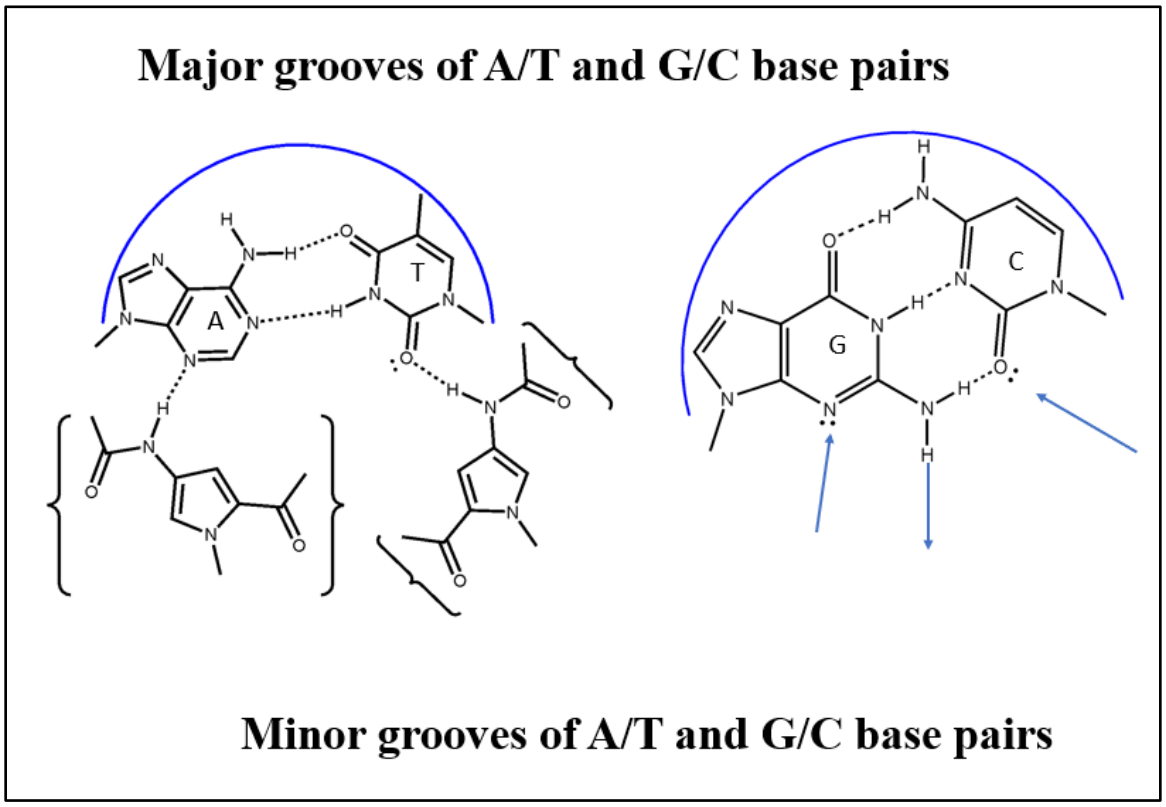


Figure 1.1. *The structure of DNA base pairs and the location of major and minor grooves provide an opportunity for minor groove DNA binding molecules. The arrows pointing toward the lone pair's electron indicate the location of Hydrogen bond acceptors, while the arrows pointing away from H indicate the location of Hydrogen bond donors.*

1.1.2. Pyrrole-imidazole polyamides

Pyrrole-imidazole polyamides (PIPs) are nanomolecular compounds designed specifically to recognize and bind to the minor groove of DNA. They are a class of molecules that can be uptake by cellular and regulate gene expression to act as treatment and therapy for specific diseases such as cancer and other infectious diseases.[4] PIPs are polymers consisting of small nitrogen-containing five-member heterocycle aromatic rings linked by amide bonds. This allows them to selectively recognize and bind to the specific molecule of DNA base pairs in the minor groove of DNA. [4,5] They are analogs of Netropsin and Distamycin A, pyrrole-containing compounds that bind to the minor groove of DNA[6,2,7]. PIPs can be taken up by and permeate through the cell membrane to recognize and selectively bind to AT-rich regions of DNA without causing denaturation of the nucleic acid.[8]

Based on the X-ray structure of netropsin and distamycin, these molecules fit in the tight space of the DNA minor groove. They were studied and found to widen the minor groove and form ligand: DNA complexes in the A/T-rich region of the DNA minor groove, as shown in **Figure 1.2.** Netropsin can only form a 1:1 complex with DNA, but distamycin A can be found to either form a 1:1 complex or a 2:1 complex.[2,9–11]

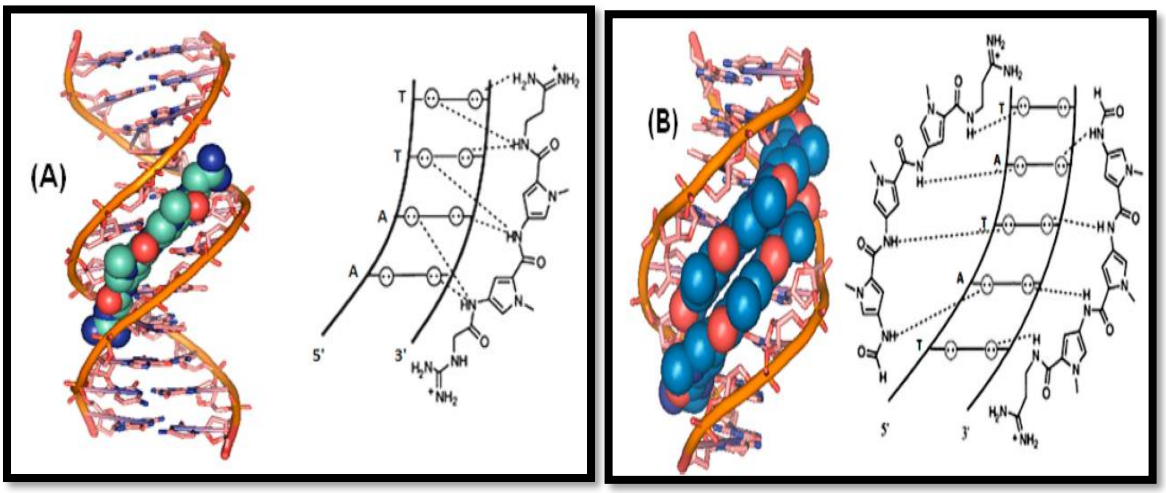


Figure 1.2 : (A) *Structure of Netropsin binding to the minor groove of DNA in the complex of 1:1 ratio, and right (B) is the structure of Distamycin A binding with the ratio of 2:1. It has indicated the NH group forming a hydrogen bond to the AT base pair of the minor groove.*[11,10] Reprinted with permission from {Alniss HY. Thermodynamics of DNA Minor Groove Binders. *J Med Chem.*

2019;62(2):385-402. doi:10.1021/acs.jmedchem.8b00233}. Copyright {2019} American Chemical Society.

The study of Netropsin and Distamycin A shows that these molecules consist of two and three molecules of N-methylpyrrole, favorably forming hydrogen bonds to the minor groove of DNA, specifically in the AT-rich region of the DNA. In 1985, Dickerson and coworkers successfully solved the X-ray crystal structure of the Netropsin molecule binding to the DNA sequence 5'-C-G-C-G-A-A-T-T-^{Br}C-G-C-G-3'; they discovered that the binding of drug molecules Netropsin and Distamycin A to DNA is highly base-specific; Based on their observations, pyrroles on netropsin can bind to the A/T base pairs, but it was not the case in the presence of G/C base pairs.[9] However, in the absence of the NH_2 group on the C-2 position of the purine ring, G/C can bind to netropsin just as they would with T/A or A/T base pair. They hypothesized that substituting imidazole for a pyrrole could allow the NH_2 group of Guanine C2 to form a hydrogen bond and further stabilize the binding toward the DNA minor groove.[9] In 1986, Lown and his coworkers proved that substituting imidazole for a pyrrole would allow the recognition and binding of guanine via hydrogen bonding.[12]

Based on Dickerson and Lown's work of Pyrrole and Imidazole in the study of DNA minor groove binding, in 1986, Dervan and coworkers successfully designed and synthesized a sequence based on the netropsin structure that selectively recognized the DNA minor groove. With the success of their studies, Dervan envisioned a new generation of sequence-specific DNA binding molecules that could surpass the low molecular weight natural products.[13]. In 1992, Dervan and his coworkers presented a Distamycin A analog, Im-Py-Py-Dp, specifically able to recognize the 5'-TGTCA-3' and forming a 2:1 ligand: DNA complex. The study proposed and demonstrated the DNA recognition rule base on the interaction of molecule pyrrole and imidazole binding to the base pair of DNA previously hypothesized and proved by Dickerson and Lown.[14–16] It can also be understood as the polyamide recognition rule: the Py/Py can recognize A/T and T/A without distinguishing them, Py/Im can recognize C/G base pairs, and Im/Py can recognize the G/C base pair. However, this recognition rule does not apply to large polyamides based on the result of our past collaboration with Dr. Dupureur's group, as the larger polyamides tend to provide high tolerance for mismatched pairs.[17,18]

1.1.3. Hairpin polyamides.

It is possible to link between the C-terminus of one linear polyamide and the N-terminus of another linear polyamide by, for example, gamma (γ)-butyric acid or GABA (known as γ-turn). Their role was to enhance the binding affinity and specificity of polyamides. After the successes of

simple multimer polyamides, Dervan and coworkers wished to increase the binding affinity of PIPs and the length of the target DNA sequence. With the new hairpin motif, a polyamide can be synthesized with additional flexibility and still match the target sequence with a higher binding affinity than previous designs.[19,20] In 1999, Dervan reported using the γ-turn and β-tail (**Figure1.3**). They observed that both β and γ expressed strong binding affinity toward A/T and T/A base pairs, and the connection between two strands of polyamides by the linker γ-turn enhanced the binding affinity >100 fold. In addition, Sugiyama reported that substituting a β-alanine for a pyrrole can also increase the binding affinity of the PIPs and the specificity of the sequence and binding to mismatched DNA sequence.[21] At the same time, the presence of β-alanine increased the binding affinity and recognition and facilitated solid phase synthesis of polyamides.

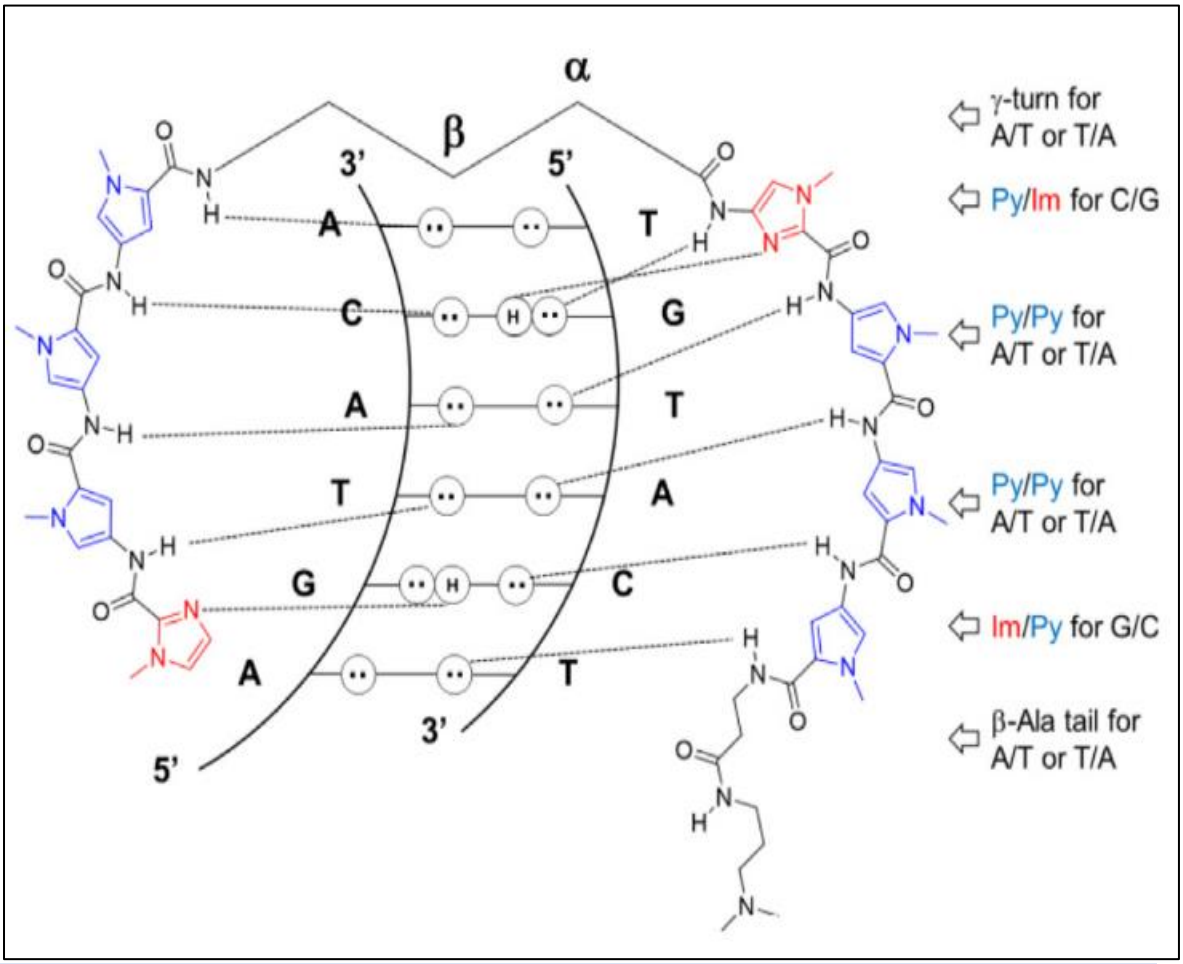


Figure 1.3. *The DNA binding hairpin polyamides' model shows the recognition of base pairs with Py/Py recognize A/T or T/A, Py/Im recognize C/G, Im/Py recognized G/C, γ-turn, and β can recognize A/T or T/A.*[22] [Permission number 5416741158991]

1.1.4. Polyamides applications

Despite having a large molecular weight, PIPs can permeate and be taken up by cells and delivered to the target sequence in the nucleus; designing drugs, and therapeutic agents to study and treat DNA-related diseases has been the goal of many PIPs researchers. Dervan and Gottesfeld's first approach to PIPs cell uptake was to study the regulation of gene expression, specifically in the effect of hairpin PIPs in inhibiting gene expression by blocking the transcription factor TFIIIA binding site.[23] Using the information from this study, Dickinson showed that hairpin PIPs also target the DNA sequence adjacent to the binding site of transcription factors Ets-1 and LEF-1 (lymphoid enhancer-binding factor-1), along with TATA box binding protein, which inhibits the binding, prevents the transcription of HIV-1, and inhibits the replication of the virus.[24]

Hairpin PIPs were also designed to inhibit the binding of hypoxia-inducible factor (HIF) toward hypoxia response element (HRE), which reduces the gene expression and the vascular endothelial growth activity factor, an essential factor in tumor angiogenesis.[25] Additional study with a tumor xenograft model showed that PIPs recognizes and binds to the HRE, reduce the expression of proangiogenic and prometastatic factor, inhibiting tumor growth by suppressing the development of tumor blood vessels. Additionally, PIPs also target the chemotherapy-resistant variant of myeloma cells, which provides a promising approach to developing anti-chemotherapy resistance drug candidates.[26] PIPs have also been reported to target the androgen-response element (ARE) and inhibit the expression of androgen-induced genes. They are also able to bind to the glucocorticoid response element (GRE) to inhibit the interaction of DNA and the glucocorticoid receptor, which represses the gene expression.[27]. However, the effect of polyamides on ARE can also cause replication stress, which is one of the mechanisms of the PIPs anti-cancer side effects.[28]

PIPs have also been found to target the genome of some viruses, such as the human papillomavirus (HPV). Parekh and coworkers have reported that PIPs prevent the E2 protein responsible for transcription from binding to the HPV DNA sequence[29]. On the other hand, Bashkin and his coworkers had demonstrated significant antiviral activity against high-risk HPV16, 18, and 31. They studied large hairpin polyamides designed to target the A/T rich region near the binding site of the human papillomavirus (HPV) E1 or E2 protein, and the hairpin polyamides are shown to have good antiviral activity without any detectable cell, tissue, or animal toxicity.[17,30] The polyamides were later found to bind all over the HPV genome rather than to the originally designed sites.

Aside from PIPs being able to regulate gene expression by interfering with protein-DNA interactions, it has been found that PIPs can regulate gene expression by sequence-specific DNA

alkylation. Sugiyama and his coworker reported conjugation of the pyrrole-imidazole polyamide sequences to an analog of natural products such as chlorambucil (Chl) or seco-1-chloromethyl-5-hydroxy-1,2-dihydro-3H-benzeindole (CBI), allowing the conjugated polyamide to alkylate the specific sequence and silence the gene expression of that target.[31,32] Dervan and coworkers showed that Chl-conjugated hairpin polyamides could also down-regulate the gene expression of the histone H4c gene, thus inhibiting the growth of several cancer cells.[33–35]

1.2 SETMAR

1.2.1 Human DNA transposon

DNA transposons are transposable elements that can be re-localized and integrated into target DNA sequences[36]. Since Watson and Crick's Model of DNA structures, knowledge of studying the characteristics of DNA has been studied and discovered continuously over the past decades. DNA contains multiple functions besides genetic information and is essential for all life.

Around the 1950s, Barbara McClintock proposed transposable elements based on her research on maize chromosome discovery, which later led to transposon's discovery in the 1970s[37,38]. Transposable elements or transposons are also known as jumping genes due to the distinct behavior of moving the DNA sequence from one location into another sequence[36]. In 1989, Finnegan classified transposable elements into 2 classes: Class I-RNA transposons or retrotransposons, and Class II-DNA transposons.[36,39,40]

Retrotransposons transpose through an intermediate RNA sequence, which is later reverse-transcribed back into a DNA sequence via retro-transcription. This type of transposable element is common in many eukaryotes. The new DNA sequence will be integrated into the target DNA. Its transposing mechanism is also called the "copy and paste" mechanism.[36] DNA transposons, found in almost all eukaryotes, have a different transposing mechanism compared to the class I counterpart that is mediated by a single transposase enzyme able to recognize the TIRs (Terminal Inverted Sequences) and excise the sequence from both ends.[36] The mechanism of DNA transposons is known as the "cut and paste" mechanism.[36] Two functional domains in DNA transposons: an N-terminal DNA binding domain (DBD), able to recognize and bind to the TIR end, and a C-terminal catalytic domain that catalyzes the DNA cleavage and transferring the DNA strands during transposition(**Figure 1.4**).[41,36]

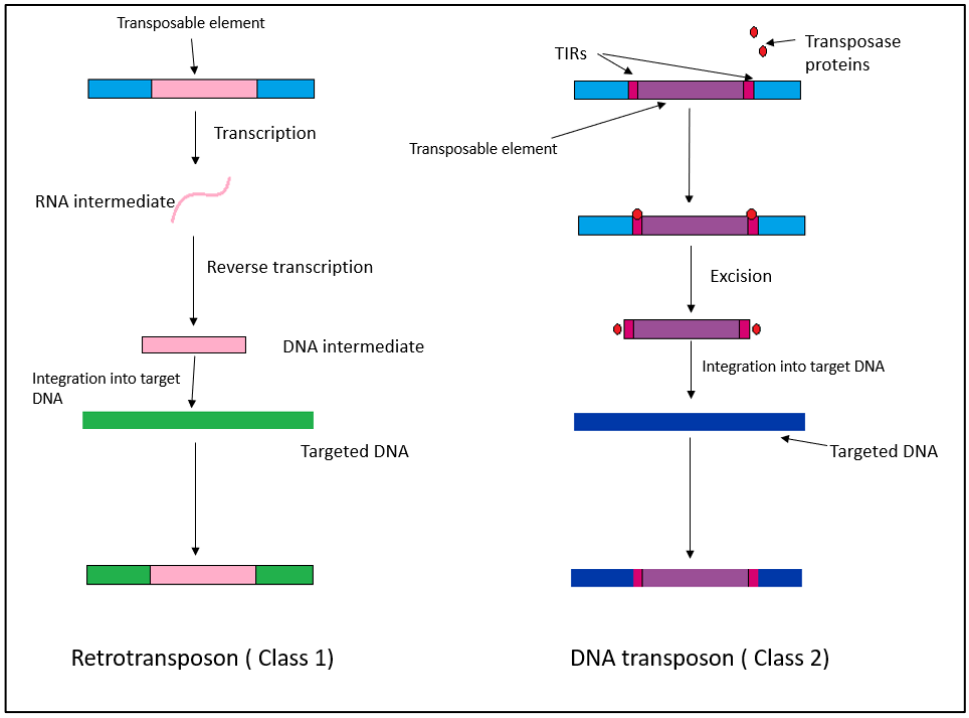


Figure 1.4: *Self-replication mechanisms of 2 types of transposable elements, "copy and paste" with class I (retrotransposon) and "cut and paste" with class II (DNA transposon) (Figure is adapted from* Ågren and Clark et al., 2018).[42]

In the human genome, more than 380,000 DNA transposon copies have been identified and classified into 125 families.[43,44] Thanks to the analysis of DNA transposons, transposons arose in primate evolution around 65-80 million years ago, after the divergence of the primate ancestor; during that time, about 74,000 transposable elements were inserted. About 40-65 million years ago, the number of DNA transposons integrated was around 23,000, with the emergence of the Prosimian Primate. Since the emergence of the New World monkey, no new DNA transposons have been integrated, and no functional DNA transposons have been found.[43,45]

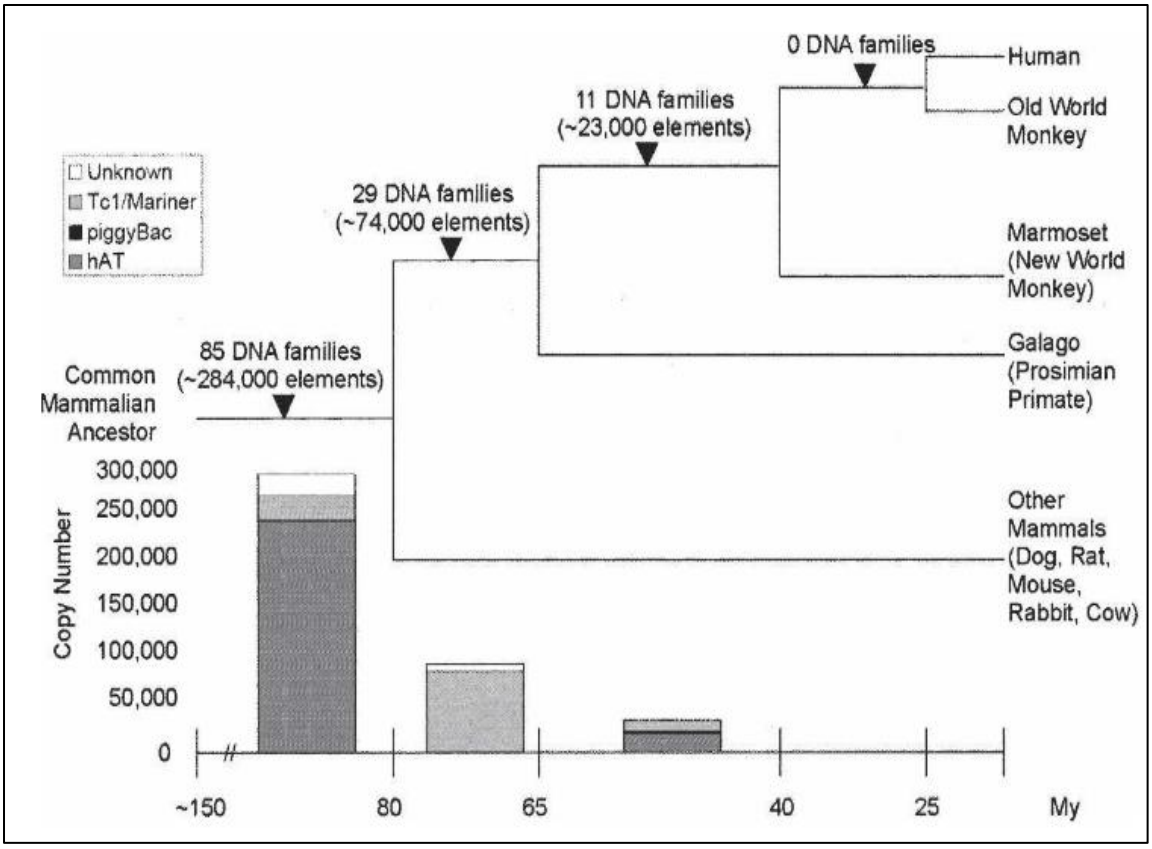


Figure 1.5: *The evolution activity of DNA transposon through the evolution of primates; the bar graph indicates the copy number of active DNA Transposable elements. No new TEs were found to be active since the emergence of the new world monkeys*.[44] Copyright to Cold Spring Harbor Laboratory Press.

1.2.2. Hsmar1

Since transposons can actively relocate their sequence in the genome, their cut-and-paste mechanism is the leading cause of animal mutation and evolution.[46] The Tc1 element/mariner superfamily of DNA transposon is found in many organisms, from fungi to humans[40,47,48,49]. In humans, Tc1/mariner is responsible for one-third of all human TEs. Within this superfamily of transposons, there are two different families of Tc1/mariner, specifically in the human genome: Hsmar1(Homosapien mariner 1) and Hsmar2 (Homosapien mariner 2).[40,48–53] Hsmar1 encoded for a single enzyme known as transposase, responsible for flanking its sequence at either end by the 30 bp Terminal Inverted Region (TIR). It has two functional domains: a DNA binding domain (DBD) and a conserved catalytic domain. The DBD consists of two helix turn helix motif (HTH) sequences responsible for site-specific minor groove DNA binding, while the catalytic domain contains a DD34D motif that is important for Hsmar1 transposition (**Figure 1.6**).[45,49,52]

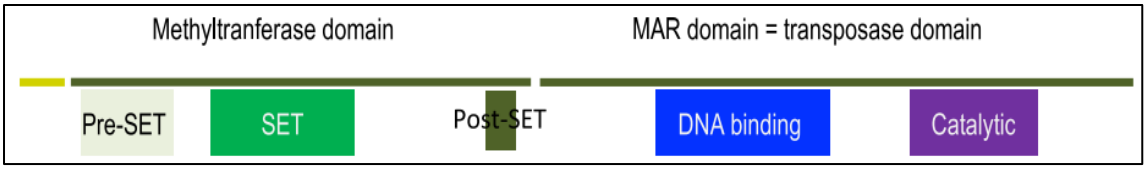


***Figure 1.6**: the representation of SETMAR consists of the SET domain and Hsmar1(MAR) domain.[54] Figure adapted from Gouillou et al. 2022.*

Robertson and his peers also performed the analysis of the Hsmar1 genome; they discovered that the evolutionary history of Hsmar1 was active during the early primate evolution around 40-65 million years ago, and it was changing over the last 37 million years ago, alongside the evolutionary history of human DNA transposons, with the minimum of at least 200 Hsmar1 copies that were integrated into the original primate genome.[52] Most of the current Hsmar1 copies have lost their transposition activity, except for one Hsmar1 copy. However, throughout the evolution of the human gene, one copy of Hsmar1 has been conserved with its entire transposase coding sequence and fused downstream to the SET region (Suppressor of variegation [Su(var)3-9, Enhancer of zeste[E(z)], and Trithorax protein of Drosophila melanogaster), which is responsible for histone lysine H3 K79 methylation.[55-56] The SET domain (SET) is responsible for histone H3 methylation and fused with the mariner Hsmar1 transposase domain (MAR), forming a chimeric gene known as SETMAR (**Figure 1.6**).[45,57,58] This gene is only present in the lineage of anthropoid primates such as old and new world monkeys, apes, chimpanzees, and humans, but it is not found in non-primate mammals and prosimians.[45,52]

1.2.3. SETMAR functions

The fusion of the SET domain and MAR domain gave rise to the chimeric gene SETMAR and its counterpart SETMAR protein (Metnase); studies have shown that SETMAR functions do not stop at methylation of Histone H3 at lysines K4 and K36. SETMAR is also involved in nonhomologous end-joining DNA repair (NHEJ).[54,58,59] In humans, the SETMAR gene can be found on the short arm of human chromosome 3p26.1, which is a region that is linked to multiple types of diseases: leukemia, non-Hodgkin's lymphoma, prostate cancer, and glioblastoma[60–62] Data showed that the MAR domain is essential for DNA repair and responsible for the repair of collapsed DNA replication forks. Although there are multiple variants of the SETMAR gene, only a single protein named Metnase, consisting of 671 amino acids, has been studied so far.[63] As the fusion of SET and Hsmar1, SETMAR also has some functions of Hsmar1 transposase, such as specific binding to the Hsmar1 transposon end, DNA cleavage, and integration. SETMAR has multiple links to the development of many types of cancer, and it is suggested to have an essential role in cancer progression.

The overexpression of SETMAR is found to suppress chromosome translocations, and an increased amount of SETMAR mRNA is located in multiple cancer cells.[62,64] This proposes the relationship between the overexpression of SETMAR toward cancer development, and the SETMAR may be a good target for anti-cancer drugs.[62]

1.2.4 SETMAR gene as DNA binding target

SETMAR has correlations to multiple diseases, bringing up the need to study the mechanism of SETMAR. However, the research on SETMAR in a normal cell is challenging since this protein is only present in simian primates and not in any other animal model due to the lack of Hsmar1 TIR in the genome. Georgiadis and Chen pointed out that the many ways of studying the function of SETMAR rely on gene knockdown or gene overexpression. They also pointed out the irony of SETMAR TIR in nonhomologous end joining as through the evolutionary history of conserved human transposon Hsmar1. The DBD of Hsmar1 has lost the ability to perform TIR-specific DNA cleavage, even though SETMAR needs the TIR-specific DNA binding activity to function in NHEJ, chromosome decatenation, and restarting of stalled replication forks.[65]. Georgiadis and her coworker sought out to identify the function of SETMAR-TIR interactions. They investigated the crystal structure of SETMAR-DNA complexes and discovered the interaction of SETMAR-TIR (**Figure 1.7**).

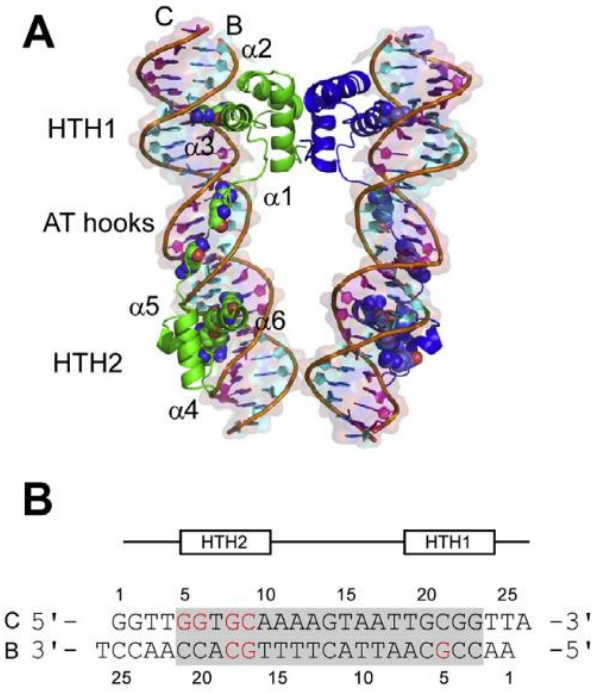


Figure 1.7. *A) the interaction of SETMAR protein (green) binding with the SETMAR gene. B) the diagram of the site-specific DNA binding position of HTH1 and HTH2. The red letters are the specific interaction site on the TIR sequence.*[65] *(Figure adapted from Georgiadis et al. 2022).*

According to the structure (**Figure 1.7A)**, the interaction of the SETMAR protein and the TIR DNA was mediated by the two domains on the Hsmar1 gene, HTH1 and HTH2. Five specific sites were also identified from this study, most of which are located in the HTH2 domain. Interestingly, this TIR sequence is conserved throughout evolutionary history, as indicated by Georgiadis, and the sequence-specific recognition occurs within the 19bp TIR element (5'-GGTG-CAAAAGTAATTGCGG-3').[65,66] This sequence will be the target of this study, as we synthesize molecules designed to modulate protein-DNA binding interactions and control the expression of SETMAR.

1.3. Bibliography

1. Wemmer DE, Dervan PB. Targeting the minor groove of DNA. *Curr Opin Struct Biol*. 1997;7(3):355-361. doi:https://doi.org/10.1016/S0959-440X(97)80051-6

2. Kopka ML, Yoon C, Goodsell D, Pjura P, Dickerson RE. The molecular origin of DNA-drug specificity in netropsin and distamycin. *Proc Natl Acad Sci U S A*. 1985;82(5):1376-1380. doi:10.1073/pnas.82.5.1376

3. Cabeen M, Aaron G, Christine H, et al. *DNA Structure & Chemistry*. Harvard University; 2020. https://projects.iq.harvard.edu/lifesciences1abookv1

4. Blackledge MS, Melander C. Programmable DNA-binding small molecules. *Bioorganic Med Chem*. 2013;21(20):6101-6114. doi:10.1016/j.bmc.2013.04.023

5. White S, Szewczyk JW, Turner JM, Baird EE, Dervan PB. Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands. *Nature*. 1998;391(6666):468-471. doi:10.1038/35106

6. Bailly C, Chaires JB. Sequence-Specific DNA Minor Groove Binders. Design and Synthesis of Netropsin and Distamycin Analogues. *Bioconjug Chem*. 1998;9(5):513-538. doi:10.1021/bc980008m

7. Coll M, Frederick CA, Wang AH, Rich A. A bifurcated hydrogen-bonded conformation in the d(A.T) base pairs of the DNA dodecamer d(CGCAAATTTGCG) and its complex with distamycin. *Proc Natl Acad Sci U S A*. 1987;84(23):8385-8389. doi:10.1073/pnas.84.23.8385

8. Fechter EJ, Olenyuk B, Dervan PB. Sequence-specific fluorescence detection of DNA by

polyamide thiazole orange conjugates. *J Am Chem Soc*. 2005;127(47):16685-16691. doi:10.1021/ja054650k

9. Kopka ML, Yoon C, Goodsell D, Pjura P, Dickerson RE. Binding of an Antitumour Drug to DNA Netropsin and CGCGAATTCGCG. Published online 1985.

10. Alniss HY. Thermodynamics of DNA Minor Groove Binders. *J Med Chem*. 2019;62(2):385-402. doi:10.1021/acs.jmedchem.8b00233

11. Peter B. Dervan. Molecular Recognition of DNA by small molecule. 2001;9(July):2-57.

12. J.William L, Krowicki K, Bhat G. Molecular recognition between oligopeptides and nucleic acids: novel imidazole-containing oligopeptides related to netropsin that exhibit altered DNA sequence specificity. *Biochemistry*. 1986;25:7408–7416. doi:10.1002/(SICI)1099-1409(200006/07)4:4<393::AID-JPP227>3.3.CO;2-2

13. Design of Sequence-Specific DNA-Binding Molecules Author (s): Peter B . Dervan Published by : American Association for the Advancement of Science Stable URL : http://www.jstor.org/stable/1696414. *Am Assoc Adv Sci*. 1986;232(4749):464-471.

14. Mrksich M, Wade WS, Dwyer TJ, Geierstanger BH, Wemmer DE, Dervan PB. Antiparallel side-by-side dimeric motif for sequence-specific recognition in the minor groove of DNA by the designed peptide 1-methylimidazole-2-carboxamide netropsin. *Proc Natl Acad Sci U S A*. 1992;89(16):7586-7590. doi:10.1073/pnas.89.16.7586

15. Wade WS, Mrksich M, Dervan PB. Design of Peptides That Bind in the Minor Groove of DNA at 5′-(A,T)G(A,T)C(A,T)-3′ Sequences by a Dimeric Side-by-Side Motif. *J Am Chem Soc*. 1992;114(23):8783-8794. doi:10.1021/ja00049a006

16. Koeller KJ. DNA Binding Polyamides and the Importance of DNA Recognition in their use as Gene-Specific and Antiviral Agents. *Med Chem (Los Angeles)*. 2014;04(02):338-344. doi:10.4172/2161-0444.1000162

17. Castaneda CH, Scuderi MJ, Edwards TG, et al. Improved antiviral activity of a polyamide against high-risk human papillomavirus: Via N-terminal guanidinium substitution. *Medchemcomm*. 2016;7(11):2076-2082. doi:10.1039/c6md00371k

18. Song Y, Niederschulte J, Bales KN, Park AH, Bashkin JK, Dupureur CM. DNA binding thermodynamics and site stoichiometry as a function of polyamide size. *Biochimie*. 2019;165:170-178. doi:10.1016/j.biochi.2019.07.021

19. Trauger JW, Baird EE, Dervan PB. Extended hairpin polyamide motif for sequence-specific recognition in the minor groove of DNA. *Chem Biol*. 1996;3(5):369-377. doi:10.1016/S1074-5521(96)90120-9

20. White S, Baird EE, Dervan PB. On the pairing rules for recognition in the minor groove of DNA by pyrrole-imidazole polyamides. *Chem Biol*. 1997;4(8):569-578. doi:10.1016/S1074-5521(97)90243-X

21. Han YW, Kashiwazaki G, Morinaga H, et al. Effect of single pyrrole replacement with β-alanine on DNA binding affinity and sequence specificity of hairpin pyrrole/imidazole polyamides targeting 5'-GCGC-3'. *Bioorg Med Chem*. 2013;21(17):5436-5441. doi:10.1016/j.bmc.2013.06.005

22. Kawamoto Y, Bando T, Sugiyama H. Sequence-specific DNA binding Pyrrole–imidazole polyamides and their applications. *Bioorganic Med Chem*. 2018;26(8):1393-1411. doi:10.1016/j.bmc.2018.01.026

23. Neely L, Trauger JW, Baird EE, Dervan PB, Gottesfeld JM. Importance of minor groove binding zinc fingers within the transcription factor IIIA-DNA complex. *J Mol Biol*. 1997;274(4):439-445. doi:10.1006/jmbi.1997.1411

24. Dickinson LA, Gulizia RJ, Trauger JW, et al. Inhibition of RNA polymerase II transcription in human cells by synthetic DNA-binding ligands. *Proc Natl Acad Sci U S A*. 1998;95(22):12890-12895. doi:10.1073/pnas.95.22.12890

25. Olenyuk BZ, Zhang GJ, Klco JM, Nickols NG, Kaelin WG, Dervan PB. Inhibition of vascular endothelial growth factor with a sequence-specific hypoxia response element antagonist. *Proc Natl Acad Sci U S A*. 2004;101(48):16768-16773. doi:10.1073/pnas.0407617101

26. Mysore VS, Szablowski J, Dervan PB, Frost PJ. A DNA-binding Molecule Targeting the Adaptive Hypoxic Response in Multiple Myeloma Has Potent Antitumor Activity. *Mol Cancer Res*. 2016;14(3):253-266. doi:10.1158/1541-7786.MCR-15-0361

27. Yao EH, Fukuda N, Ueno T, et al. A pyrrole–imidazole polyamide targeting transforming growth factor-β1 inhibits restenosis and preserves endothelialization in the injured artery. *Cardiovasc Res*. 2008;81(4):797-804. doi:10.1093/cvr/cvn355

28. Martínez TF, Phillips JW, Karanja KK, et al. Replication stress by Py–Im polyamides

induces a non-canonical ATR-dependent checkpoint response. *Nucleic Acids Res*. 2014;42(18):11546-11559. doi:10.1093/nar/gku866

29. Schaal TD, Mallet WG, McMinn DL, et al. Inhibition of human papilloma virus E2 DNA binding protein by covalently linked polyamides. *Nucleic Acids Res*. 2003;31(4):1282-1291. doi:10.1093/nar/gkg206

30. Edwards TG, Vidmar TJ, Koeller K, Bashkin JK, Fisher C. DNA Damage Repair Genes Controlling Human Papillomavirus (HPV) Episome Levels under Conditions of Stability and Extreme Instability. *PLoS One*. 2013;8(10). doi:10.1371/journal.pone.0075406

31. Oyoshi T, Kawakami W, Narita A, Bando T, Sugiyama H. Inhibition of transcription at a coding sequence by alkylating polyamide. *J Am Chem Soc*. 2003;125(16):4752 – 4754. doi:10.1021/ja029196o

32. Shinohara K ichi, Sasaki S, Minoshima M, Bando T, Sugiyama H. Alkylation of template strand of coding region causes effective gene silencing. *Nucleic Acids Res*. 2006;34(4):1189-1195. doi:10.1093/nar/gkl005

33. Jespersen C, Soragni E, James Chou C, Arora PS, Dervan PB, Gottesfeld JM. Chromatin structure determines accessibility of a hairpin polyamide–chlorambucil conjugate at histone H4 genes in pancreatic cancer cells. *Bioorg Med Chem Lett*. 2012;22(12):4068-4071. doi:https://doi.org/10.1016/j.bmcl.2012.04.090

34. Lin J, Hiraoka K, Watanabe T, et al. Identification of binding targets of a pyrrole-imidazole polyamide KR12 in the LS180 colorectal cancer genome. *PLoS One*. 2016;11(10):1-19. doi:10.1371/journal.pone.0165581

35. Wang YD, Dziegielewski J, Wurtz NR, Dziegielewska B, Dervan PB, Beerman TA. DNA crosslinking and biological activity of a hairpin polyamide-chlorambucil conjugate. *Nucleic Acids Res*. 2003;31(4):1208 – 1215. doi:10.1093/nar/gkg215

36. Wicker T, Sabot F, Hua-Van A, et al. Reply: A unified classification system for eukaryotic transposable elements should reflect their phylogeny. *Nat Rev Genet*. 2009;10(4):276. doi:10.1038/nrg2165-c4

37. Ravindran S. Barbara McClintock and the discovery of jumping genes. *Proc Natl Acad Sci U S A*. 2012;109(50):20198-20199. doi:10.1073/pnas.1219372109

38. McCLINTOCK B. The origin and behavior of mutable loci in maize. *Proc Natl Acad Sci*

U S A. 1950;36(6):344-355. doi:10.1073/pnas.36.6.344

39. Finnegan DJ. Eukaryotic transposable elements and genome evolution. *Trends Genet*. 1989;5(C):103-107. doi:10.1016/0168-9525(89)90039-5

40. Robertson HM. The Tcl-mariner superfamily of transposons in animals. *J Insect Physiol*. 1995;41(2):99-105. doi:https://doi.org/10.1016/0022-1910(94)00082-R

41. Sinzelle L, Izsvák Z, Ivics Z. Molecular domestication of transposable elements: From detrimental parasites to useful host genes. *Cell Mol Life Sci*. 2009;66(6):1073-1093. doi:10.1007/s00018-009-8376-3

42. Ågren JA, Clark AG. Selfish genetic elements. *PLOS Genet*. 2018;14(11):e1007700. doi:10.1371/journal.pgen.1007700

43. Feschotte C, Pritham EJ. DNA transposons and the evolution of eukaryotic genomes. *Annu Rev Genet*. 2007;41(35):331-368. doi:10.1146/annurev.genet.40.110405.090448

44. Pace JK, Feschotte C. The evolutionary history of human DNA transposons: Evidence for intense activity in the primate lineage. *Genome Res*. 2007;17(4):422-432. doi:10.1101/gr.5826307

45. Cordaux R, Udit S, Batzer MA, Feschotte C. Birth of a chimeric primate gene by capture of the transposase gene from a mobile element. *Proc Natl Acad Sci U S A*. 2006;103(21):8101-8106. doi:10.1073/pnas.0601161103

46. Ayarpadikannan S, Kim HS. The Impact of Transposable Elements in Genome Evolution and Genetic Instability and Their Implications in Various Diseases. *Genomics Inform*. 2014;12(3):98. doi:10.5808/gi.2014.12.3.98

47. Robertson HM. The mariner transposable element is widespread in insects. *Nature*. 1993;362(6417):241-245. doi:10.1038/362241a0

48. Plasterk RH. The Tc1/mariner transposon family. *Curr Top Microbiol Immunol*. 1996;204:125-143. doi:10.1007/978-3-642-79795-8_6

49. Lampe DJ, Akerley BJ, Rubin EJ, Mekalanos JJ, Robertson HM. Hyperactive transposase mutants of the Himar1 mariner transposon. *Proc Natl Acad Sci U S A*. 1999;96(20):11428-11433. doi:10.1073/pnas.96.20.11428

50. Smit AF, Riggs AD. Tiggers and DNA transposon fossils in the human genome. *Proc Natl*

Acad Sci U S A. 1996;93(4):1443-1448. doi:10.1073/pnas.93.4.1443

51. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature*. 2001;409(6822):860-921. doi:10.1038/35057062

52. Robertson HM, Zumpano KL. Molecular evolution of an ancient mariner transposon, Hsmar1, in the human genome. *Gene*. 1997;205(1-2):203-217. doi:10.1016/s0378-1119(97)00472-1

53. Robertson HM, Martos R. Molecular evolution of the second ancient human mariner transposon, Hsmar2, illustrates patterns of neutral evolution in the human genome lineage. *Gene*. 1997;205(1-2):219-228. doi:10.1016/s0378-1119(97)00471-x

54. Lié O, Renault S, Augé-Gouillou C. SETMAR, a case of primate co-opted genes: towards new perspectives. *Mob DNA*. 2022;13(1):1-14. doi:10.1186/s13100-022-00267-1

55. Qian C, Zhou MM. SET domain protein lysine methyltransferases: Structure, specificity and catalysis. *Cell Mol Life Sci*. 2006;63(23):2755-2763. doi:10.1007/s00018-006-6274-5

56. Kuzmichev A, Nishioka K, Erdjument-Bromage H, Tempst P, Reinberg D. Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. *Genes Dev*. 2002;16(22):2893-2905. doi:10.1101/gad.1035902

57. Liu D, Bischerour J, Siddique A, Buisine N, Bigot Y, Chalmers R. The Human SETMAR Protein Preserves Most of the Activities of the Ancestral Hsmar1 Transposase . *Mol Cell Biol*. 2007;27(3):1125-1132. doi:10.1128/mcb.01899-06

58. Kim HS, Kim SK, Hromas R, Lee SH. The set domain is essential for metnase functions in replication restart and the 5' end of ss-overhang cleavage. *PLoS One*. 2015;10(10):1-16. doi:10.1371/journal.pone.0139418

59. Chen Q. A ROLE FOR SETMAR IN GENE REGULATION : INSIGHTS FROM STRUCTURAL ANALYSIS OF THE DNA-BINDING DOMAIN IN COMPLEX WITH DNA. 2016;(August).

60. Higgins JJ, Pucilowska J, Lombardi RQ, Rooney JP. Candidate genes for recessive non-syndromic mental retardation on chromosome 3p (MRT2A)*. *Clin Genet*. 2004;65(6):496-500. doi:https://doi.org/10.1111/j.0009-9163.2004.00267.x

61. Kaur E, Nair J, Ghorai A, et al. Inhibition of SETMAR-H3K36me2-NHEJ repair axis in

residual disease cells prevents glioblastoma recurrence. *Neuro Oncol*. 2020;22(12):1785-1796. doi:10.1093/neuonc/noaa128

62. Dussaussois-Montagne A, Jaillet J, Babin L, et al. SETMAR isoforms in glioblastoma: A matter of protein stability. *Oncotarget*. 2017;8(6):9835-9848. doi:10.18632/oncotarget.14218

63. Lee SH, Oshige M, Durant ST, et al. The SET domain protein Metnase mediates foreign DNA integration and links integration to nonhomologous end-joining repair. *Proc Natl Acad Sci U S A*. 2005;102(50):18075-18080. doi:10.1073/pnas.0503676102

64. Wray J, Williamson EA, Chester S, et al. The transposase domain protein Metnase/SETMAR suppresses chromosomal translocations. *Cancer Genet Cytogenet*. 2010;200(2):184-190. doi:10.1016/j.cancergencyto.2010.04.011

65. Chen Q, Bates AM, Hanquier JN, et al. Structural and genome-wide analyses suggest that transposon-derived protein SETMAR alters transcription and splicing. *J Biol Chem*. 2022;298(5):101894. doi:10.1016/j.jbc.2022.101894

66. Chen Q, Georgiadis M. Crystallization of and selenomethionine phasing strategy for a SETMAR-DNA complex. *Acta Crystallogr Sect Struct Biol Commun*. 2016;72:713-719. doi:10.1107/S2053230X16012723

67. Swalley SE, Baird EE, Dervan PB. Effects of γ-turn and β-tail amino acids on sequence-specific recognition of DNA by hairpin polyamides. *J Am Chem Soc*. 1999;121(6):1113-1120. doi:10.1021/ja9830905

68. Liu B, Wang S, Aston K, et al. β-Alanine and N-terminal cationic substituents affect polyamide-DNA binding. *Org Biomol Chem*. 2017;15(46):9880-9888. doi:10.1039/c7ob02513k

69. Tellier M, Chalmers R. The roles of the human SETMAR (Metnase) protein in illegitimate DNA recombination and non-homologous end joining repair. *DNA Repair (Amst)*. 2019;80(March):26-35. doi:10.1016/j.dnarep.2019.06.006

70. Wang S, Aston K, Koeller KJ, et al. Modulation of DNA-polyamide interaction by β-alanine substitutions: A study of positional effects on binding affinity, kinetics and thermodynamics. *Org Biomol Chem*. 2014;12(38):7523-7536. doi:10.1039/c4ob01456a

71. Jaramillo D, Liu Q, Aldrich-Wright J, Tor Y. Synthesis of N-Methylpyrrole and N-

Methylimidazole Amino Acids Suitable for Solid-Phase Synthesis. *J Org Chem*. 2004;69(23):8151-8153. doi:10.1021/jo048686r

72. Baird EE, Dervan PB. Solid Phase Synthesis of Polyamides Containing Imidazole and Pyrrole Amino Acids. 1996;7863(9):6141-6146.

73. Rucker VC, Melander C, Dervan PB. Influence of beta -Alanine on Hairpin Polyamide Orientation in the DNA Minor Groove. 2003;86.

74. Chenoweth DM, Harki DA, Dervan PB. Solution-Phase Synthesis of Pyrrole - Imidazole Polyamides. Published online 2009:7175-7181.

75. Reeves R, Nissen MS. The A·T-DNA-binding domain of mammalian high mobility group I chromosomal proteins. A novel peptide motif for recognizing DNA structure. *J Biol Chem*. 1990;265(15):8573-8582. doi:10.1016/s0021-9258(19)38926-4

76. Fields GB. Introduction to peptide synthesis. *Curr Protoc protein Sci*. 2002;Chapter 18:Unit 18.1. doi:10.1002/0471140864.ps1801s26

77. Greenberg WA, Baird EE, Dervan PB. A comparison of H-pin and hairpin polyamide motifs for the recognition of the minor groove of DNA. *Chem - A Eur J*. 1998;4(5):796-805. doi:10.1002/(SICI)1521-3765(19980515)4:5<796::AID-CHEM796>3.0.CO;2-G

78. Xiao J, Yuan G, Huang W, Chan ASC, Lee KLD. A convenient method for the synthesis of DNA-recognizing polyamides in solution. *J Org Chem*. 2000;65(18):5506-5513. doi:10.1021/jo000135n

79. Liu B, Wang S, Aston K, et al. β-Alanine and N-terminal cationic substituents affect polyamide–DNA binding. *Org Biomol Chem*. 2017;15(46):9880-9888. doi:10.1039/c7ob02513k

80. Gumpper RH, Li W, Castañeda CH, Scuderi MJ, Bashkin JK, Luo M. A Polyamide Inhibits Replication of Vesicular Stomatitis Virus by Targeting RNA in the Nucleocapsid. *J Virol*. 2018;92(8). doi:10.1128/jvi.00146-18

81. Khan GS, Shah A, Zia-Ur-Rehman, Barker D. Chemistry of DNA minor groove binding agents. *J Photochem Photobiol B Biol*. 2012;115:105-118. doi:10.1016/j.jphotobiol.2012.07.003

82. Cho J, Parks ME, Dervan PB. Cyclic polyamides for recognition in the minor groove of

DNA. *Proc Natl Acad Sci U S A*. 1995;92(22):10389-10392. doi:10.1073/pnas.92.22.10389

83. Tellier M. Structure, Activity, and Function of SETMAR Protein Lysine Methyltransferase. *Life*. 2021;11(12). doi:10.3390/life11121342

84. Tellier M, Chalmers R. Human SETMAR is a DNA sequence-specific histone-methylase with a broad effect on the transcriptome. *Nucleic Acids Res*. 2019;47(1):122-133. doi:10.1093/nar/gky937

Chapter 2: Design and Synthesis of novel DNA binding polyamides based on the SETMAR TIR crystal structure using Solid phase peptide synthesis.

2.1. Introduction

The feature of site-specific selective of pyrrole-imidazole polyamides (PIPs) has been studied over decades, and their applications varied on the designed functions and characteristics of the researcher are far from reaching the limit. The hairpin shape of PIPs allows the compound to fit in the tight space of the DNA minor groove while the building blocks of polyamides recognize the site with good selectivity. Many PIPs were designed to target the DNA sequence responsible for certain diseases such as cancer, birth defects, and hormonal diseases. These research areas allow scientists to choose the combination of PIPs building blocks to design and study the binding affinity of the designed compound and the target DNA sequence. [2,9–11]

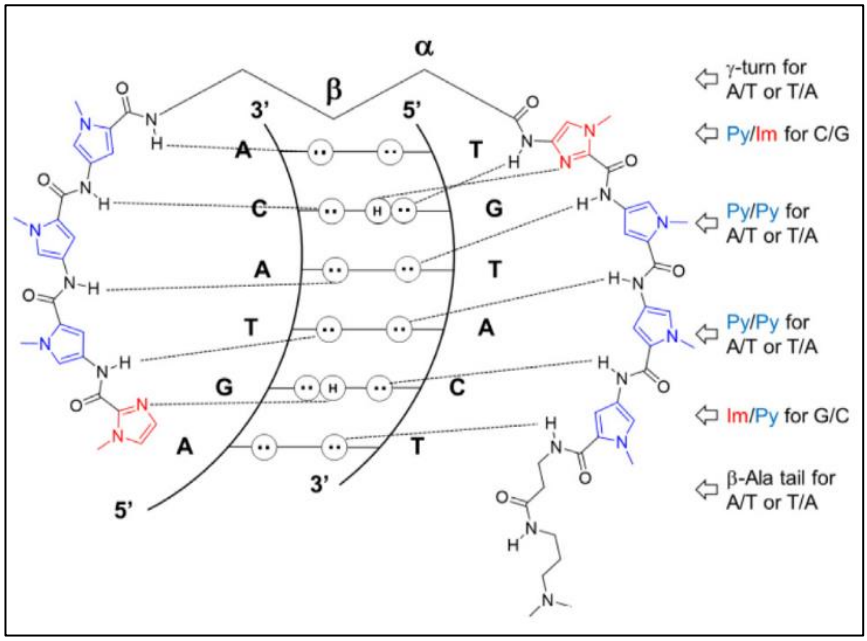


Figure 2.1 *The DNA binding hairpin polyamides' model shows the recognition of base-pairs with Py/Py recognize A/T or T/A, Py/Im recognize C/G, Im/Py recognized G/C, γ-turn, and β can recognize A/T or T/A.* [22]

Pyrrole-imidazole polyamides are polymers of multiple nitrogen-containing heterocycles, such as N-methylpyrrole (Py) and N-methylimidazole (Im), linked by amide bonds. PIPs are analogs of natural products netropsin and distamycin A, known for their binding to the minor groove of the A/T-rich region. [2,9–11] Based on the polyamide pairing rules of Dervan and his coworkers, Py/Py recognizes T/A and A/T in a non-discriminating manner, Py/Im can recognize C/G, and Im/Py recognizes G/C (**Figure 2.1**). [6,1] β and γ-turn can also recognize T/A or A/T (**Figure 2.1**)

.[67,68] Over thousands of polyamides have been designed and synthesized, and they have become important tools in minor groove DNA binding studies.

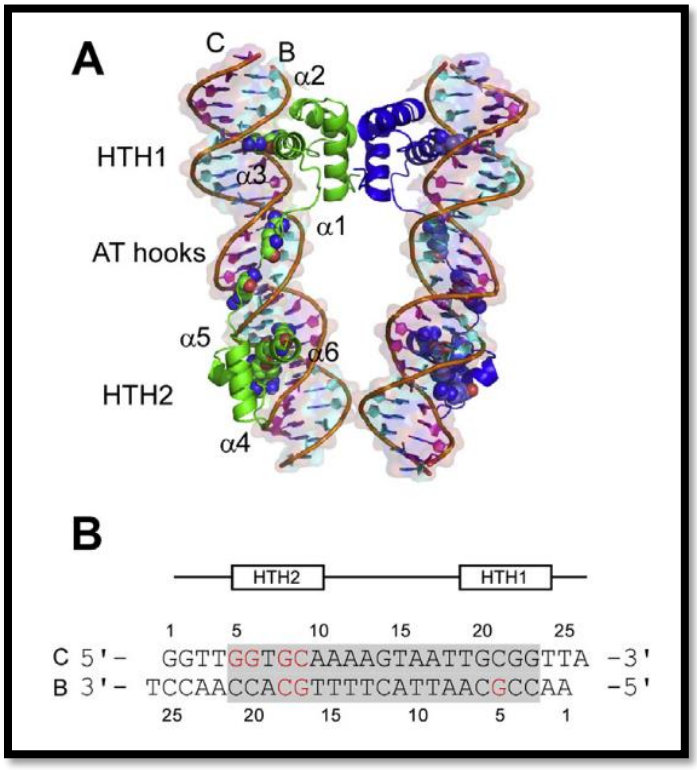


Figure 2.2 *A) the interaction of SETMAR protein (green) binding with the SETMAR gene. B) the diagram of the site-specific DNA binding position of HTH1 and HTH2. The red letters are the specific interaction site on the TIR sequence*[65] (Georgiadis et al. 2022).

SETMAR is known as a chimeric gene due to the fusing of a human transposon domain and a SET histone methyltransferase gene.[45,69] The chimeric gene encodes for a single protein known as Metnase. Despite SETMAR's role in the non-homologous end-joining pathway and gene regulation, many of its functions are still unclear. Understanding the role of SETMAR: TIR binding is a priority in developing an anticancer drug for diseases relating to this gene.[62] Georgiadis and coworkers have solved the crystal structure of SETMAR and the SETMAR-TIR site-specific interaction location. They learned that the SETMAR protein bind to the conserved TIR regions mediated by the HTH2 domain of the Hsmar1 at the 5'-GGTGC-3' sequence (**Figure 2.2)**.[65,66]

The discovery of the SETMAR-TIR site-specific sequence intrigued us in synthesizing new PIP sequences specifically designed to bind to the 4 locations on the TIR sequence of the Hsmar1 domain. In this work, we designed PIPs to target and recognize the first ten conserved base pairs (G/C-G/C-T/A-G/C-C/G) on the TIR of Hsmar1(**Figure 2.3)**. This study will expand what

we know about polyamides and their minor groove DNA binding effects on the sequence of SETMAR.

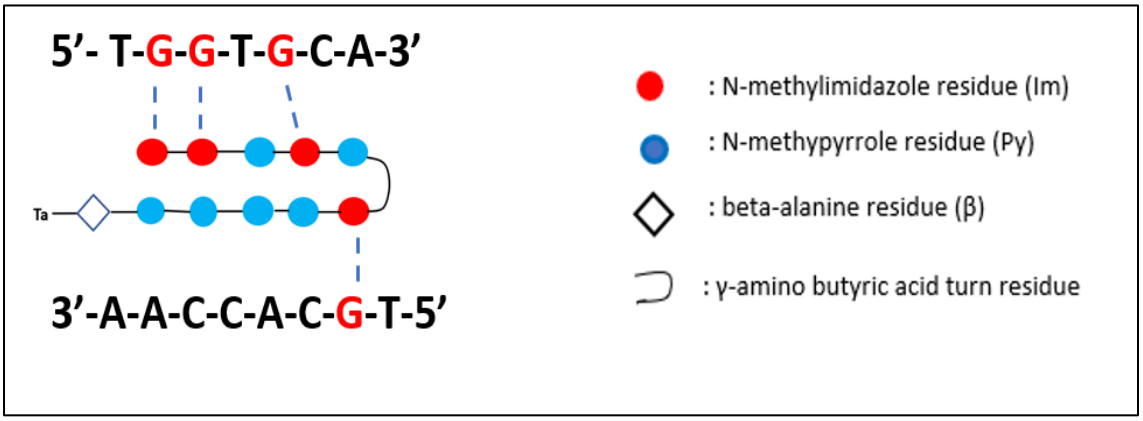


Figure 2.3: *Designed sequence of PIPs based on the polyamide binding rule and the discovered SETMAR: TIR binding study*.

Five polyamides were successfully designed and synthesized for this chapter study (**Table 2.1**). Since β-alanine can also recognize and bind to A or T of DNA, similar to pyrrole, substituting β-alanine from a pyrrole will allow us to try various combinations of polyamide sequences.[16,67,70] Some polyamides were designed to have a single mismatch to act as control samples and compare their results with the tested polyamide sequences. All building blocks of PIPs can be obtained either from commercially available sources or by solution-phase synthesis. The building blocks were linked with the solid-phase peptide synthesis method.

Compound	N-terminus	1	2	3	4	5	6	7	8	9	10	11	C-terminus
1094	**desIm**	**Im**	**β**	**Im**	Py	**γ**	Py	Py	Py	Py	Py	**β**	Ta
1110	**desIm**	**Im**	Py	**Im**	Py	**γ**	Py	Py	Py	Py	Py	**β**	Ta
1123	**desIm**	**Im**	Py	**Im**	Py	**γ**	**Im**	Py	Py	Py	Py	**β**	Ta
1131	**desIm**	**Im**	**β**	**Im**	Py	**γ**	**Im**	Py	Py	β	Py	**β**	Ta
1132	**desIm**	**Im**	Py	**Im**	**β**	**γ**	**Im**	Py	Py	Py	Py	**β**	Ta

Table 2.1. *Sequences of polyamides designed to recognize and bind to SETMAR TIR DNA minor groove specifically*.

Our collaborator, Dr. Millie Georgiadis, and her researchers carried out the biochemical part of this project, and the finished compounds were sent to their lab to test for competitive DNA binding via EMSA and binding effects with Fluorescence Anisotropy. Our second collaborator, Dr. W. David Wilson, and his group carried out the biophysical part of this project. They perform Surface Plasmon Resonance to study the binding affinity of the SETMAR TIR DNA with polyamides.

This study aimed to design and synthesize long-chain PIPs to recognize the discovered SETMAR TIR sequence using the solid phase peptide synthesis method. We synthesized hairpin polyamides by linking the building blocks starting with β-alanine PAM (phenyl-acetaminomethyl) resin. All building blocks have the -Boc (tert-butyloxycarbonyl) group to protect the N-terminus except for des-aminoimidazole (desIm). Each coupling step utilized either a monomer or, when possible, a dimer for better coupling efficiency.

Solid-phase peptide synthesis

All polyamide sequences in this project have been prepared with the assistance of the peptide synthesizer CSBio model CS136XT. In this project, all resin actions were carried out in a 20ml reaction vessel scale. The device is connected to a computer where the command program is set up and controls the action of the synthesizer. CSBio's reaction vessel (RV) is also connected to a heating device that can keep the RV at a constant 35°C throughout the processes of the coupling program. Boc-β-alanine PAM resin (typically 1g, 0.23 meq/g) was transferred into the RV as the initial step, PyBOP in DMF is prepared as an important coupling reagent and usually kept in reagent bottle R4, and PyBOP/DMF is always renewed for every coupling sequencing program. All building blocks were added to an AA# vessel before the command programmed the CSBio to measure the solvent with MVA (Measuring vessel A), followed by transferring the solution to TVA (Transfer vessel A). The building blocks were then activated with DMF solvents and PyBOP/DMF as the amide coupling reagents before transferring into RV for coupling. RV inverts continuously for a minimum of 180 min to ensure the coupling is completed. The process can be viewed as a coupling cycle (**Figure 2.4**). A continuous cycle starting with the Deprotection of the Boc group, followed by activation of building blocks and amide bond couplings, "cap," prevents the formation of unwanted material, such as side products and unreacted amine that can react in the next coupling step. Upon completion of the couplings, the PIPs resin was cleaved from the PAM resin with an amine nucleophile to release the polyamides, and that would be the final step in the synthesis procedure of PIPs synthesis with solid phase peptide synthesis.

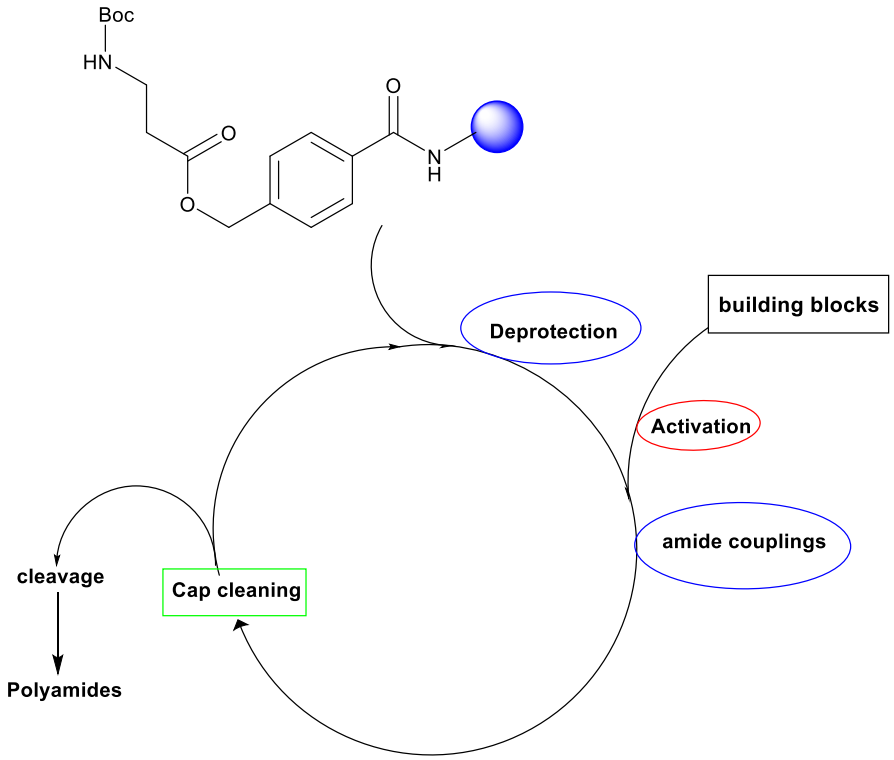


Figure 2.4 *The solid-phase synthesis process cycle was used for this project*.

2.2 Method

Material

Boc- β-alanine-PAM resin was ordered from Peptides International (Louisville, KY), 3-(dimethylamino)-1-propylamine, N-methyl-imidazole, N-methyl-pyrrole, piperidine, indole, ammonium formate, 20% palladium/carbon (20% Pd/C), dimethyl sulfoxide, HCl, and NaOH were purchased from Sigma Aldrich (St. Louis, MO). 1-Methyl-4-nitro-imidazole-2-carboxylate (NO_2-Im-COOMe) was ordered from AC Pharmachem, Inc (Worcester, MA).

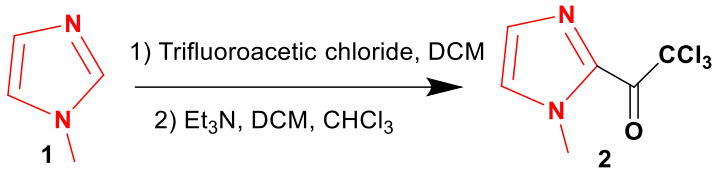
4-tert-butoxycarbonylaminobutyric acid (Boc-γ-COOH), 1-ethyl-4-amino-β-alanine-carboxylate (H_2N-β-COOEt), 4-tert-butoxy-β-alanine-carboxylic acid (Boc-β-COOH), trifluoroacetic acid (TFA), formic acid (FA), hexafluorophosphate azabenzotriazole tetramethyl uronium (HATU), benzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate (PyBOP), hydroxybenzotriazole (HOBT), dimethyl formamide (DMF), and methanol (MeOH) were acquired from Oakwood Chemical (Estill, SC). Triethylamine, N, N-diisopropylethylamine (DIPEA), formic acid, and acetic anhydride were purchased from Fisher Scientific.

Instrumentation

^{1}H-NMR and ^{13}C-NMR spectra were recorded in DMSO_d_6 at 300_MHz, and 600_MHz, respectively. ^{1}H-NMR peaks are referenced to DMSO at 2.5 ppm, and ^{13}C-NMR spectra in DMSO at 39.5 ppm; coupling constants J are reported in Hz. The FID data were processed and analyzed with the software SpinWorks 4. High-resolution mass spectrometry was carried out on Bruker LC/MS instrument. HPLC was carried out in the ThermoFisher U-HPLC, and the data was processed with Chromeleon7 and Agilent HPLC. ThermoFisher preparative HPLC carried out purifications of polyamides.

2.2.1 Synthesis of precursor des-Im-COCCl$_3$ (2, Scheme 2.1)

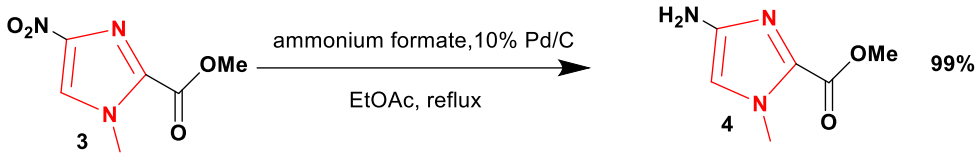
A solution of N-methyl-imidazole (33.09 g, 182 mmol) in DCM (45mL) was added drop-wise into a mixture solution of trifluoroacetic chloride (15 g, 182 mmol) in DCM (107 mL), stirred at 0 ^{o}C, and the solution was stirred under inert Nitrogen gas over 18 h to room temperature at 23^{o}C. The reaction temperature was reduced to 0 ^{o}C then triethylamine (Et_3N, 18.42 g, 182 mmol) in DCM was added drop-wise and stirred under nitrogen over 18 h. The compound was then checked for completion with TLC. The mixture was washed with chloroform, recrystallized in hot hexane, then obtained (33.12 g, 182 mmol,80% yield) off-white to yellow crystalline solid.[71]



Scheme 2.1. Synthesis of desIm-COCCl$_3$ **2**.

2.2.2 Synthesis of precursor H$_2$N-Im-COOMe (4, Scheme 2.2)

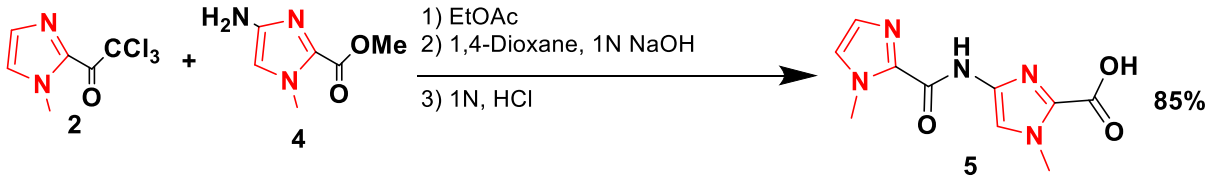
A solution of NO_2-Im-COOMe (20 g, 107.88 mmol) **3** was added to an ethyl acetate solution (206 mL) and stirred at room temperature for 30 min. Catalyst 20% Pd/C (1g) was added, followed by ammonium formate (10.20 g, 161.82 mmol) to the mixture, and the reaction refluxed at 70 ^{o}C for 1- 2 h. The progress of the reaction was monitored with TLC and by staining with $KMnO_4$. All catalyst and side products were removed via filtration through Celite. The solution solvent was concentrated in vacuo to afford the product as a yellow-orange solid **4** (16.55 g, 107.88 mmol, 99% yield).[71]



Scheme 2.2. Synthesis of Methyl-2-amino-1-H-imidazole 5-carboxylate **4**.

2.2.3 Synthesis of dimer building block desIm-Im-COOH (5, Scheme 2.3)

A mixture of desIm-$COCCl_3$ (3.627 g, 15.94 mmol) **2** and H_2N-Im-COOMe (2.474 g, 15.94 mmol) **4** were combined and added to EtOAc (67 mL), and the reaction was stirred under N_2 at room temperature for 18 h. The reaction was monitored with TLC and HPLC until completion. The reaction was filtered by vacuum filtration to collect a yellow solid, and the filtrate solvent was later removed by vacuum to afford a yellow solid. The two solids were combined and dissolved in 2,4-dioxane followed by the addition of 1N NaOH (74 mL); the reaction was stirred well under N_2 for 30 min at room temperature, later the temperature was increased to 45 ^oC; the reaction was checked for completion with HPLC, and the solvent was removed under vacuum, followed by adding 1M HCl until the pH reaches 4.25. The reaction was vacuum filtered to afford 3.37 g of yellow solid **5** (15.94 mmol, 85% yield)**.**

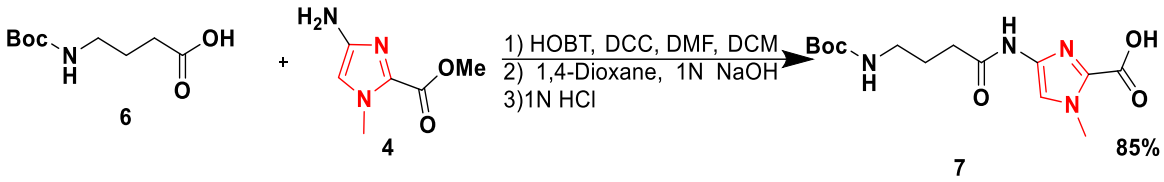


Scheme 2.3. Synthesis of desIm-Im-COOH **5.**

2.2.4 Synthesis of dimer building block Boc-γ-Im-COOH (7, Scheme 2.4)

In a 250 mL round bottom flask, Boc-γ-COOH **6 (**4.98 g, 24.5 mmol**)** was dissolved in DMF (20 mL), followed by HOBT (5.0 g, 37 mmol), and DCC (8.32 g, 37 mmol) in DCM (66 mL), and stirred well under nitrogen gas for 20 min. After 30 min at room temperature, the mixture formed DCU (dicyclohexylurea) and turned into a white solution; the reaction was run overnight at room temperature. After 12 h, H_2N-Im-COOMe **4** (4.75 g, 30.6 mmol) was added in DCM (20 mL), and the reaction was stirred well under nitrogen gas. The reaction was monitored by TLC and HPLC, and the reaction was completed after 18 h at room temperature. The mixture was filtered to remove the solid DCU, and a rotary evaporator was used to remove the DCM solvent. EtOAc (80

mL) was added to the flask after evaporation, and then the mixture was washed with H_2O, aq $NaHCO_3$, 4% HCl, and brine. The organic solvent was dried with anhydrous Na_2SO_4 and gravity filtered to collect an orange solution. The solvent was removed by rotary evaporation to afford an oily solution. The oily mixture was immediately dissolved in 2,4-dioxane (40 mL) and stirred under N_2 for 30 min before adding 1 N NaOH (40 mL). The reaction was heated in an oil bath with a temperature controlled at 45 °C for 1 to 2 h. The reaction was checked for completion with analytical HPLC. The reaction was removed from heat and cooled at room temperature for 10 min. Then the pH was reduced to around 3.5- 4.5 using 1N HCl to afford 6.7 g of off-white solid **7** (24.5mmol, 85% yield).[71]

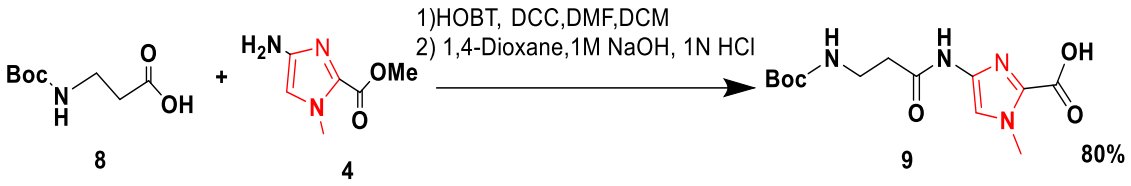


Scheme 2.4 Synthesis of Boc-γ-Im-COOH **7.**

2.2.5 Synthesis of dimmer building block Boc-β-Im-COOH (9, Scheme 2.5)

In a 250 mL round bottom flask, Boc-β-COOH **8** (5.42 g, 28.6 mmol) was dissolved in DMF (104 mL), followed by HOBT (5.7 g, 42.9 mmol), and DCC (8.8 g, 42.9 mmol) in DCM (66 mL), stir well under nitrogen gas for 20 min. After 30 min at room temperature, the mixture formed DCU and turned to a white solution, let the reaction ran overnight at room temperature. After 12 h, H_2N-Im-COOMe **4 (**6.6 g, 42.9 mmol) was dissolved in DCM (20 mL), and the reaction was stirred well under nitrogen gas. The reaction was monitored by TLC and HPLC, and the reaction was completed after 18 h at room temperature. The mixture was filtered to remove the solid DCU, and the DCM solvent was removed by rotary evaporation. EtOAc (70 mL) was added to the flask after vacuo, and then the mixture was washed with H_2O, aq $NaHCO_3$, 4% HCL, and brine.

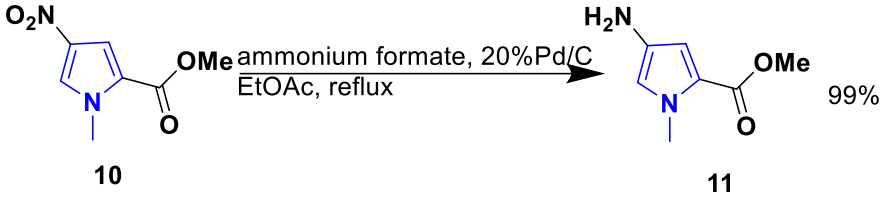
The organic solvent was dried with anhydrous Na_2SO_4 and gravity filtered to collect a red solution. The solvent was removed by rotary evaporation to afford an oil solution. The oil mixture was immediately dissolved in 2,4-dioxane and stirred under N_2 for 30 min before adding 1M NaOH. The reaction was heated in an oil bath with a temperature controlled at 45°C for 1 to 2 h and checked for completion with analytical HPLC. The reaction was removed from heat and settled at room temperature for 10 min, reducing the pH to around 3.5-4.2 to afford red solid **9** (7.15 g, 28.6 mmol, 80% yield).



Scheme 2.5. Synthesis of Boc-β-Im-COOH **9**.

2.2.6 Synthesis of precursor H_2N-Py-COOMe (11, Scheme 2.6)

A solution of NO_2-Py-COOMe **10** (20 g, 108 mmol) was added in a solution of 400 mL ethyl acetate, followed by 1g of 20%Pd/C was stirred at room temperature for 30 min. Catalyst Pd/C was added, followed by ammonium formate (10.53 g, 162 mmol) was added to the mixture, and the reaction was refluxed at 70°C for 45 min. Reaction completion was checked with TLC and stained with $KmnO_4$. All catalyst and side products were removed via filtration through Celite. The solution was concentrated *in vacuo* to afford (16.55 g, 108 mmol, 99% yield) of off-white solid **17.** (**Scheme 2.6)**.

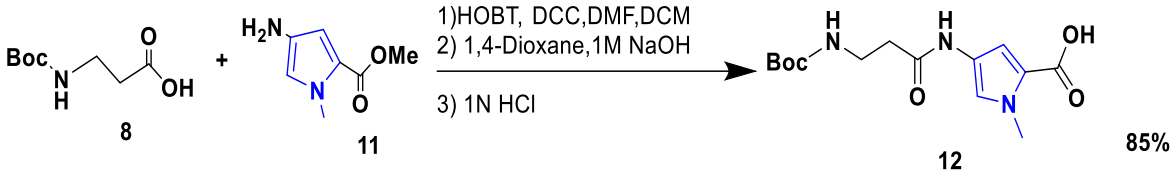


Scheme 2.6. Synthesis of H_2N-Py-COOMe **11.**

2.2.7 Synthesis of dimer building block Boc-β-Py-COOH (12, Scheme 2.7)

In a 250 mL round bottom flask, Boc-β-COOH **8** (5.0 g, 20.6 mmol) as dissolved in DMF (20mL), followed by HOBT (5.0 g, 37.0 mmol), and DCC (7.6 g, 37.0 mmol) in 66 mL DCM, stir well under nitrogen gas for 20 min. After 30 min at room temperature, the mixture formed DCU and turned white solution, let the reaction ran overnight at room temperature. After 12 H, add H_2N-Py-COOMe **11** (5.6 g, 37.0 mmol) in 20 mL DCM, and the reaction was stirred well under Nitrogen gas. The reaction was monitored by TLC and HPLC; the reaction was completed after 18 h at room temperature. The mixture was filtered to remove the solid DCU, and a rotary evaporator was used to remove the DCM solvent. EtOAc was added to the flask, and then the mixture was washed with H_2O, aq $NaHCO_3$, 4% HCL, and brine. The organic solvent was dried with anhydrous Na_2SO_4 and gravity filtered to collect a clear solution. A rotary evaporator was used to remove the solution's solvent to afford an oil. The oil mixture was immediately dissolved in 2,4-dioxane and stirred under

N_2 for 30 min before adding 1N NaOH. The reaction was heated in an oil bath with a temperature controlled at 45°C for 1 to 2 h and checked for completion with analytical HPLC. The reaction was removed from heat and settled at room temperature for 10 min, reducing the pH to around 3.75-4.5 to afford an off-white solid **12** (8.1g, 20.6 mmol, 85% yield), **(Scheme 2.7)**.



Scheme 2.7 Synthesis of Boc-β-Py-COOH **12**.

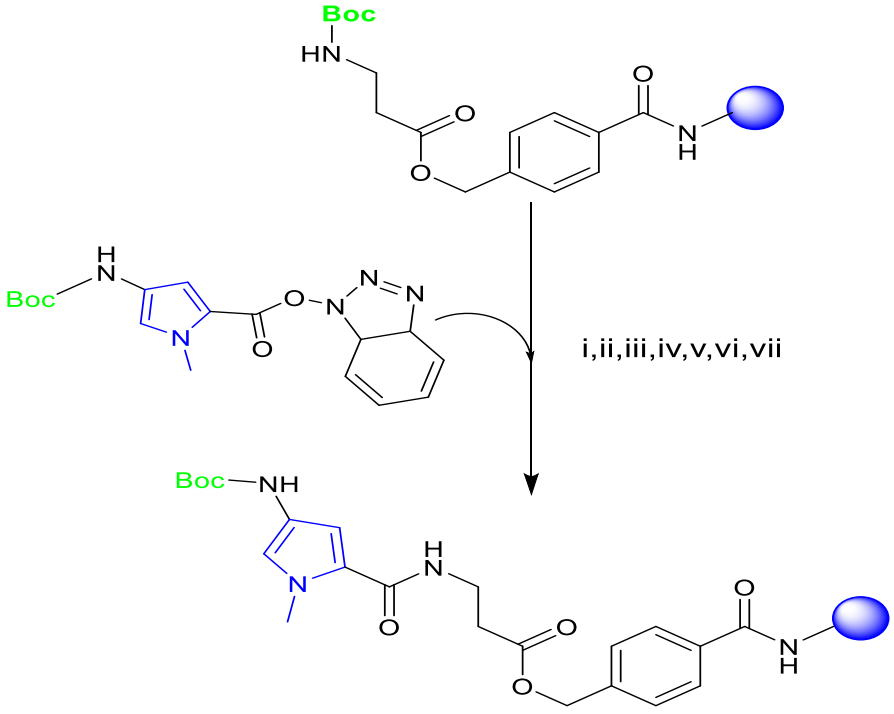
2.2.8 Synthesis of desIm-Im-β-Im-Py-γ-Py-Py-β-Py-β-Ta. (PA1094)

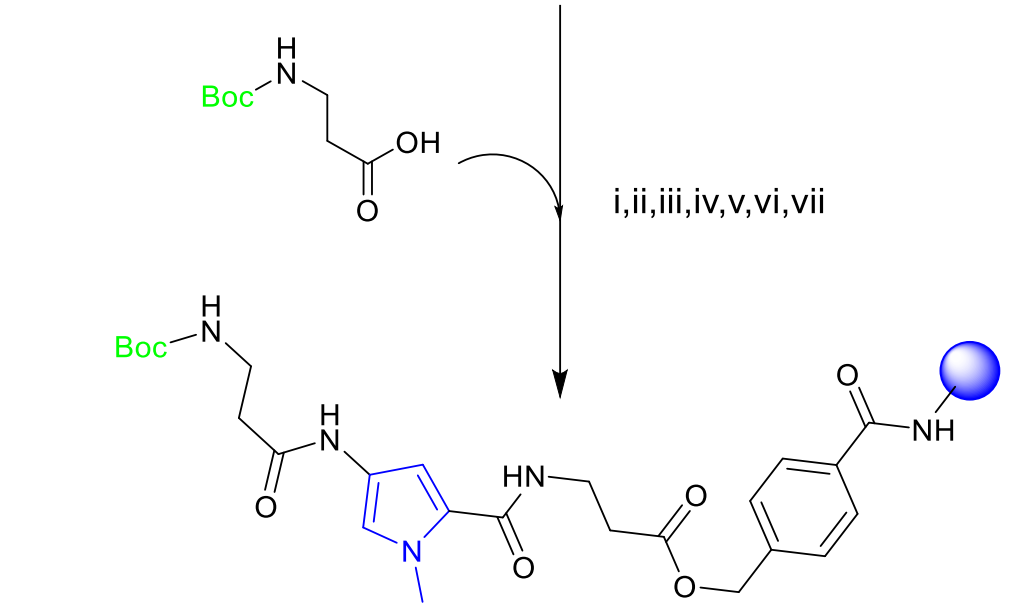
Synthesis of polyamide **1094** resin started by initiating the reaction sequences program with the CSBio program. The CSBio reaction vessel was set at a temperature of 35 °C throughout the whole process. Boc-β-alanine-PAM resin (1000 mg) was loaded into the RV. PyBOP (benzotriazole-1-yl-oxytripyrrolelidinophosphonium hexafluorophosphate (11 gram) was dissolved in DMF (40 mL) and filled in the R4 solvent bottle. After the program started the first resin wash and deprotection of the Boc group for 1 h, a solution of the monomer Boc-Py-Obt (508 mg,1.84 mmol) in DMF (6.5 mL) was stirred to dissolve and transferred into AA#3 fully. The coupling program directed the synthesizer to transfer the solution from AA#3 to TVA and then MVA to begin the activation before transferring to RV for coupling for 180 min, followed by capping with acetic anhydride for 17 min to remove impurities such as an unreacted amine. After completing the first sequence, the program proceeded to "cap" to remove any nonreacted amine before moving on to the next sequence. Each sequence took around 6 h to complete.

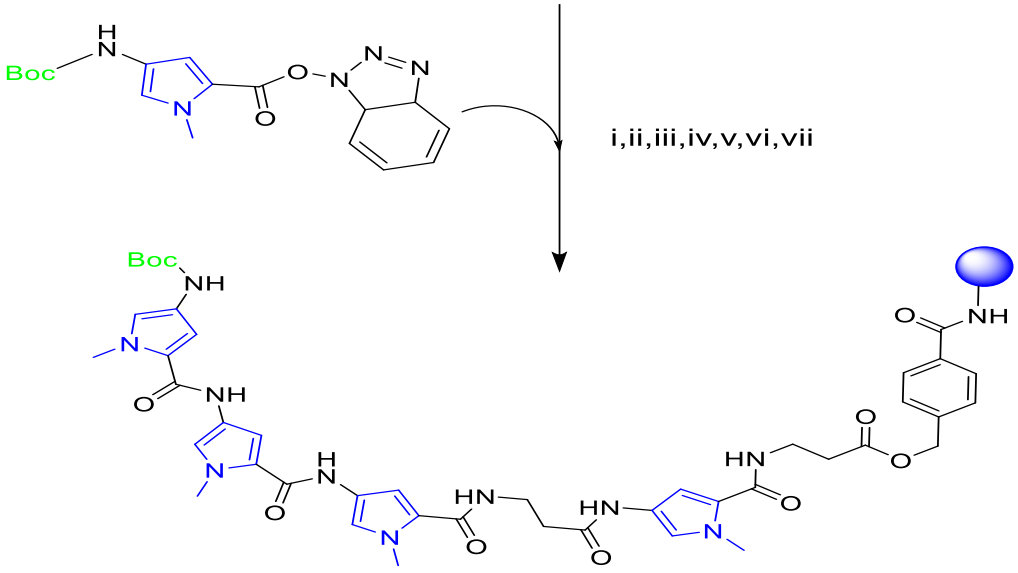
After completing the first sequence, the programmed system transferred a solution of monomer Boc-β-COOH (358 mg, 1.84 mmol) in DMF (6.5 mL), was well dissolved by vortex and loaded in AA#4 to TVA and MVA for sequence coupling following up by capping. After sequence 3 plus cap was completed, a solution of Boc-Py-Py-COOH (514 mg, 1.38 mmol) in DMF (6.5 mL) was well dissolved by vortex, loaded in AA#5, transferred to TVA, then MVA before transferred to RV for coupling reaction for 180 min following by capping with acetic anhydride for 17 min. After sequence 4 plus cap was completed, a solution of Boc-Py-OBt (508 mg, 1.84 mmol) in DMF (6.5 mL) was well dissolved by vortex; loaded in AA#6, transferred to TVA then MVA before transferred to RV for coupling reaction for 180 min following by capping with acetic anhydride for 17 min. After sequence 5 was finished, the programed system was initiated and followed by

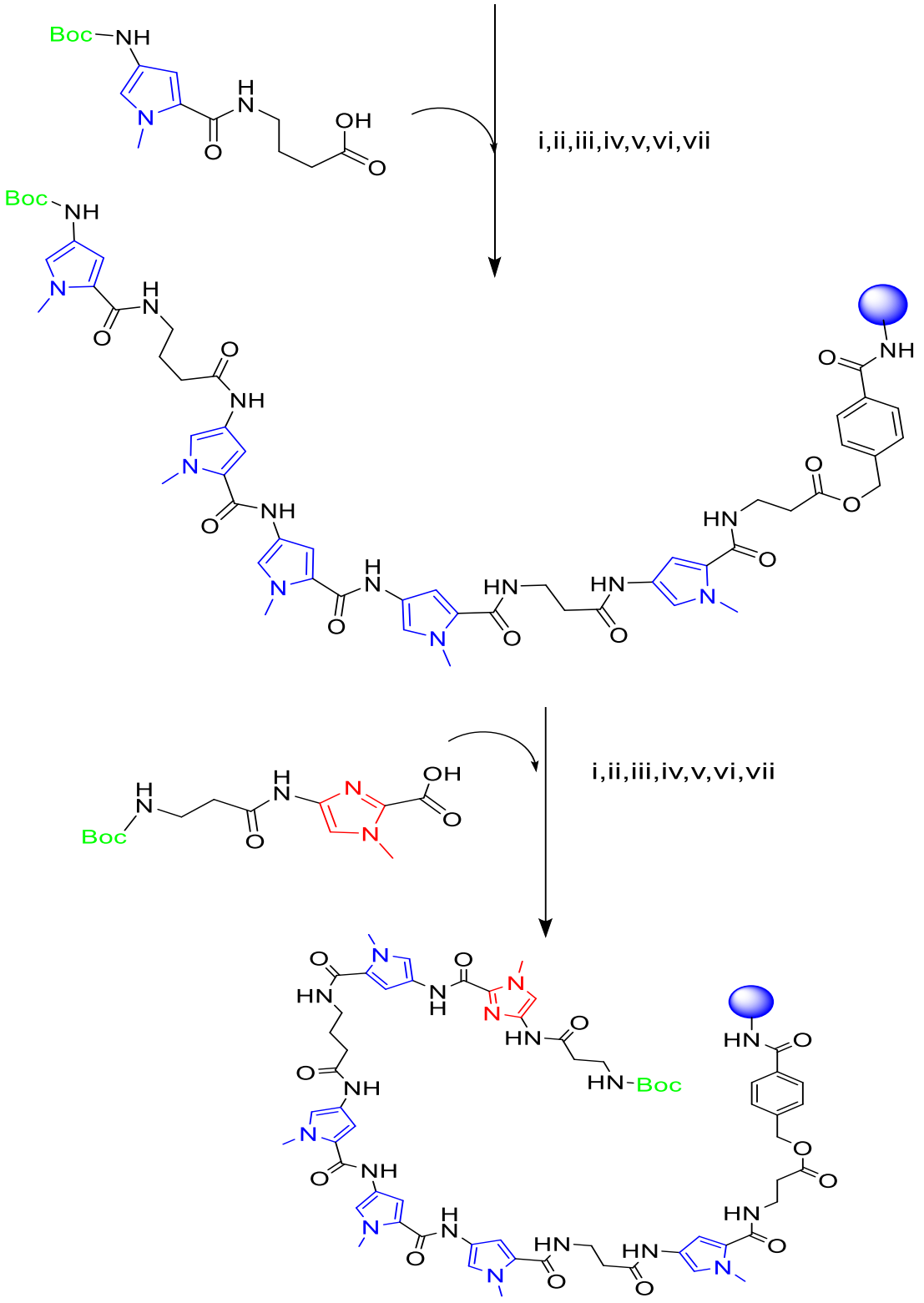
transferring a well-dissolved by vortex solution of Boc-Py-γ-COOH (462 mg, 1.38 mmol) in DMF (6.5 mL) from AA#6 to TVA, then MVA before transferred to RV for coupling reaction for 180 min following by capping with acetic anhydride for 17 min. After sequence 6 was finished, the programed system was initiated and followed by transferring a well-dissolved, vortexed solution of dimer Boc-β-Im-COOH (444 mg, 1.38 mmol) in DMF (6.5 mL) and DIPEA (1mL) from AA#7 to TVA then MVA before transferred to RV for coupling reaction for 180 min following by capping with acetic anhydride for 17 min. After sequence 7 was finished, the programed system was initiated and followed by transferring a well dissolved by vortex solution of dimer desIm-Im-COOH **5** (472mg, 1.38 mmol) and DIPEA (1ml) in DMF (6.5 mL) from AA#8 to TVA then MVA before transferred to RV for coupling reaction for 180 min following by capping with acetic anhydride for 17 min.

The sequence was finished after 37 h, 0.1 mg of resin was cleaved with 100 µL Ta (3,3-diamino-N-methyldipropylamine) solution for 2h at 45°C (**Scheme 2.8**), then the heated solution was collected and filtered before being analyzed with a Bruker LC-MS instrument. The exact mass was determined to be $[M+H]^+$, 1458.73 m/z, which confirmed the presence of the desired product. All the resin was carefully transferred from the RV into a beaker using a small spatula, and the resin was washed three-time with DCM and MeOH, then dried under vacuum for 30 min until reaching dryness. The resin was then collected in a 20mL scintillation vial and stored in the freezer before cleavage the next day.

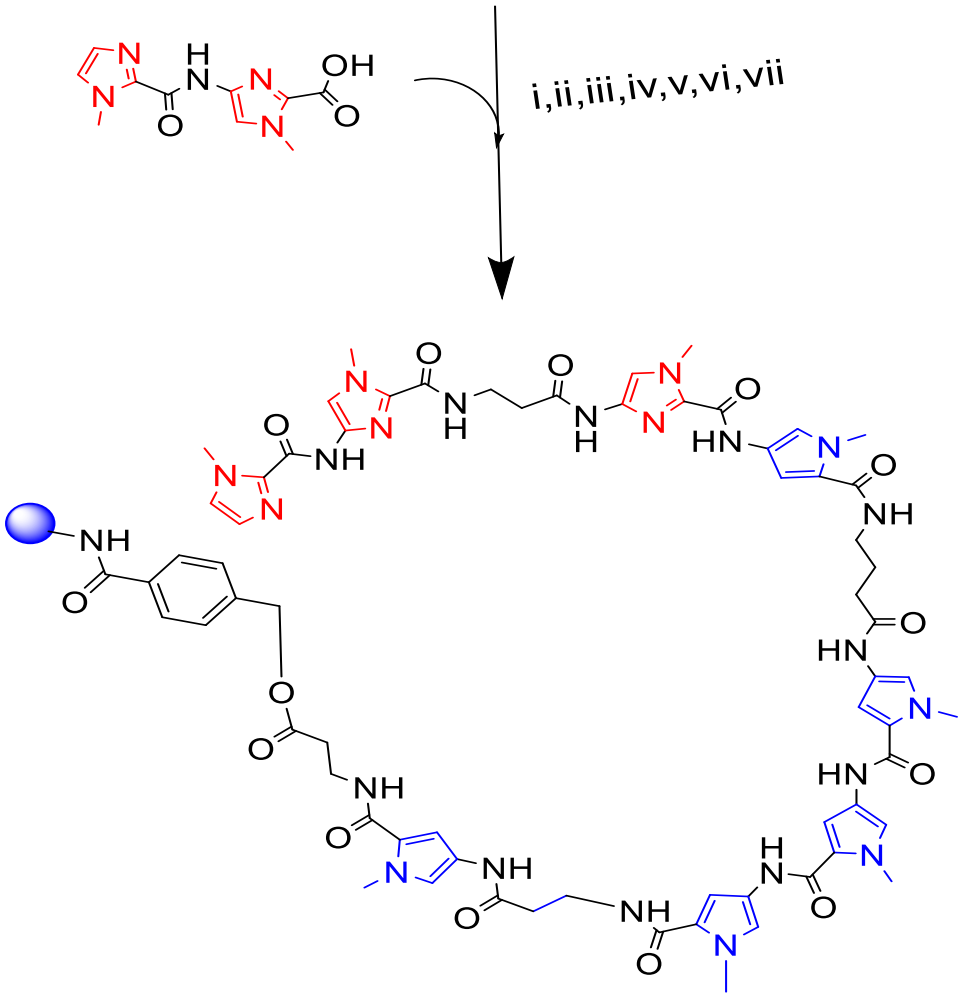



Boc
H
N
OH
O
i,ii,iii,iv,v,vi,vii
Boc
H
N
H
N
O
HN
O
N
O
O
O
NH


Boc
H
N
N
N
O
O
N
i,ii,iii,iv,v,vi,vii
Boc
NH
N
O
NH
N
H
N
O
HN
HN
O
O
HN
O
O
O
NH



Scheme 2.8 (caption on next page)



Scheme 2.8. *Scheme of synthesis polyamide* **1094** *resin by applying solid-phase synthesis method. Started from a commercially acquired Boc-β-PAM resin, and each sequence contains multiple steps from (i) to (viii). (i) wash the resin with DCM and DMF, (ii) deprotection with 60% TFA/ DCM/ 0.5 M Indole, (iii) 25% piperidine/ DMF; (iv) building blocks with the structures shown on the scheme, (v) 0.5M PyBOP/ DMF as coupling reagents, (vi) DIPEA, (vii) DMSO, (viii) capping step with acetic anhydride for 17 min.*

Synthesis of 1094 program sequence

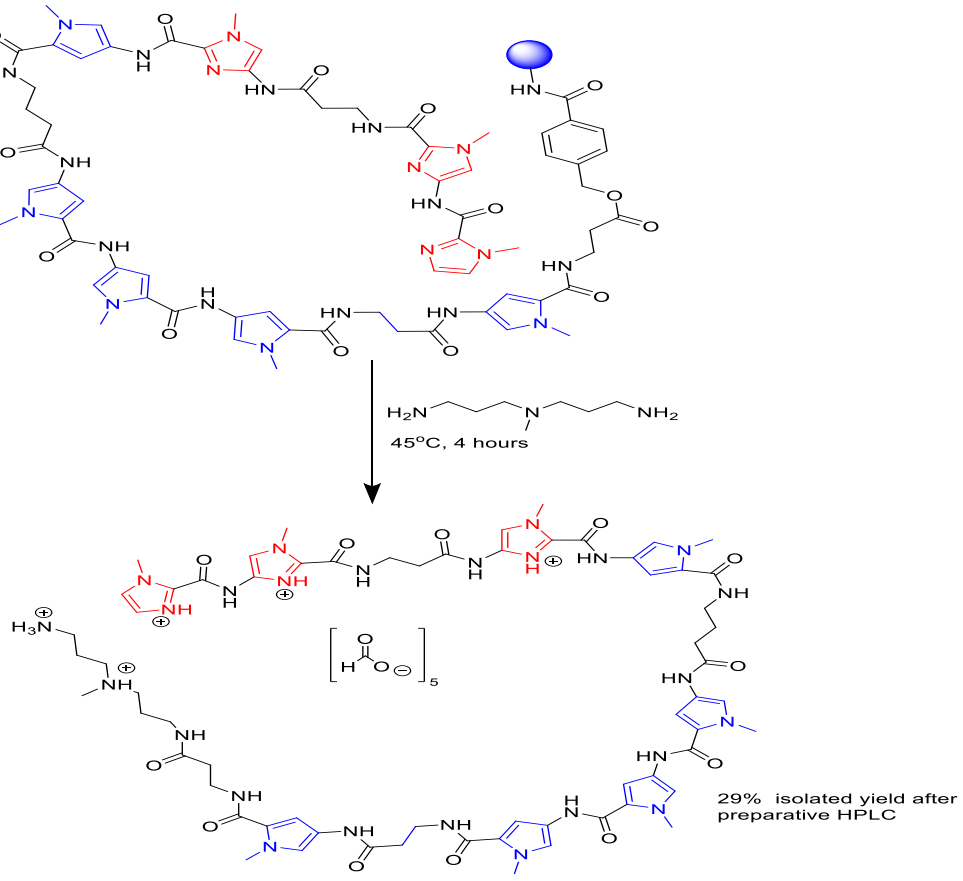
Polyamide **1094** was synthesized with 14 consecutive programs. The total coupling time was 37 h **(Table 2.2)**.

Table 2.2. Program sequence of **1094** resin synthesis.

Sequence step	Building block(s) or type of sequence	Additional Solvent mixture	Applied program file
1	Boc-Py-Obt	(6.5 mL) DMF	Add-DIPEA-180-min-DI-PEA-Flows.CFN
2	Capping	None	Cap-17-min-continues.CFN
3	Boc-β-COOH	(6.5 mL) DMF	Add-DIPEA-180-min-DI-PEA-Flows.CFN
4	Capping	None	Cap-17-min-continues.CFN
5	Boc-Py-Py-COOH	(6.5 mL) DMF	Add-DIPEA-180-min-DI-PEA-Flows.CFN
6	Capping	None	Cap-17-min-continues.CFN
7	Boc-Py-OBt	(6.5 mL) DMF	Add-DIPEA-180-min-DI-PEA-Flows.CFN
8	Capping	None	Cap-17-min-continues.CFN
9	Boc Py-γ-COOH	(6.5 mL) DMF	Add-DIPEA-180-min-DI-PEA-Flows.CFN
10	Capping	None	Cap-17-min-continues.CFN
11	Boc-β-Im-COOH	(6.5 mL) DMF+DI-PEA (1mL)	Add-DIPEA-180-min-DI-PEA-Flows.CFN
12	Capping	None	Cap-17-min-continues.CFN
13	desIm-Im-COOH	(6.5 mL) DMF+DI-PEA (1mL)	Add-DIPEA-180-min-DI-PEA-Flows.CFN
14	14-Capping	None	Cap-17-min-continues.CFN

Cleavage of PAM resin using Ta as a nucleophile.

The next day, the resin was cleaved with aminolysis in 2000 µL of Ta in a 5 mL centrifuge tube at 45ºC in a heating device Isotemp Fisher Scientific for 4 h, and the separation of polyamides solution from the PAM resin support was performed in a 20 µM polyethylene frit placed inside a disposable polypropylene syringe by washing with MeOH and H_2O. The polyamide solution was concentrated in a rotary evaporator to afford **1094** (31 mg, 29% yield) (**Scheme 2.9**).



Scheme 2.9. Cleavage of PA **1094** resin using Ta as the nucleophile.

Purification of the compound.

The polyamide solution was diluted with a mixture of 200 µL DMSO and H_2O/0.2% HCOOH (200 µL), then purified by preparative HPLC using a Phenomenex Luna 250x30 mm, 5 µM, 100 Å, C18 column maintained at 25 ºC. The organic phase was 0.2% HCOOH in MilliQ water and 100% HPLC grade MeOH. The applied gradient was 10% MeOH for 8 min, followed by a ramp rate to 90% Methanol over 35.6 min at a 20 mL/minute flow rate. All collected fractions were analyzed and selected for 95% or higher purity with analytical HPLC using ThermoFisher-

UHPLC. High-purity fractions were collected for rotary evaporation to remove MeOH from the solution before lyophilization.

Lyophilization of compound 1094

Compound **1094** was diluted with a 1:1 mixture of ACN (5 mL): and 0.1% HCOOH in H_2O (5 mL). The samples were frozen before the lyophilization process with ground dry ice until an even coated of frozen material results at an acute angle in the vial. The frozen sample was covered with porous Kimwipe paper and secured by rubber bands. The sample was settled in a lyophilizer glass vessel and connected to a vacuum port. The frozen polyamides sublimed as an off-white solid powder after 48 h. The lyophilizer vacuum was set at 300 torr within 30 min prior to the lyophilization process, and the inside temperature was set at -83 to -85ºC.

Characterization of compound

Analytical HPLC characterization was performed with a C12 Phenomenex Juniper Proteo column maintained at 40ºC. The mobile phase consisted of 0.1% formic acid in water(A) and HPLC grade ACN (B). The applied gradient consisted of 5%B for 0.75 min, followed by a ramp to 60% B over 6.5 min at 2.0 mL/min. The retention was 3.357 min (**Figure 2.13**).

Characteristic results of 1094

PA1094- desIm-Im-Py-Im-β-γ-Im-Py-Py-Py-Py-β-Ta $(HCOO^-)^6$

^{1}H-NMR (600 MHz, DMSO-d_6) δ=10.37 (s, 1H), 9.94 (s, 6H), 8.47 (br, 4H), 8.22 (s, 1H), 8.08- 8.03 (t, J= 31.4 Hz, 3H), 7.95-7.93 (d, J=4.48 Hz, 1H), 7.50 (s, 1H), 7.45 (s, 1H), 7.42 (s, 1H), 7.21 (s, 1H), 7.19 (s, 1H), 7.18 (s, 1H), 7.15 (s, 1H), 7.11 (s, 1H), 7.03 (s, 1H), 7.02 (s, 1H), 6.95 (s, 1H), 6.85 (s, 1H), 6.82 (s, 1H), 6.62 (s, 1H), 3.96 (s, 1H), 3.93 (s, 1H), 3.92 (s, 2H), 3.81 (s, 2H), 3.80 (s, 2H), 3.78 (s, 3H), 3.49-3.46 (q, J= 20.67 Hz, 3H), 3.41-3.38 (q, J= 19.56 Hz, 3H), 3.33- 3.29 (q, J= 18.99 Hz, 2H), 3.20-3.16 (q, J= 19.56 Hz, 2H), 3.03-.3.01 (q, J=18.99 Hz, 2H), 2.76 (br, 6H), 2.59-2.51 (t, J= 12.85 Hz, 2H), 2.32- 2.29 (m, J= 43.02Hz, 4H), 2.21-2.19 (t, J= 12.85 Hz, 6H), 2.07 (s, 1H), 2.05 (s, 1H), 1.77 (s, 1H), 1.65 (s, 2H), 1.49 (s, 1H).

^{13}C-NMR (151 MHz, DMSO-d_6) δ= 165.56, 165.53, 158.30, 155.85, 137.85, 13.64, 127.63, 127.0, 126.94, 126.11, 117.62, 113.89, 54.49, 54.32, 53.94, 41.29, 41.24, 39.93, 39.79, 39.65, 39.51, 39.37, 39.23, 39.10, 39.06, 37.56, 37.14, 36.60, 35.09, 34.91, 34.85, 26,67, 24.57.
HRMS (ESI) was calculated for **1094**-$C_{26}C_{86}N_{26}O_{12}$, $[M+H]^+$,1459.69 m/z found 1458.73 m/z.

HPLC purity: 94%

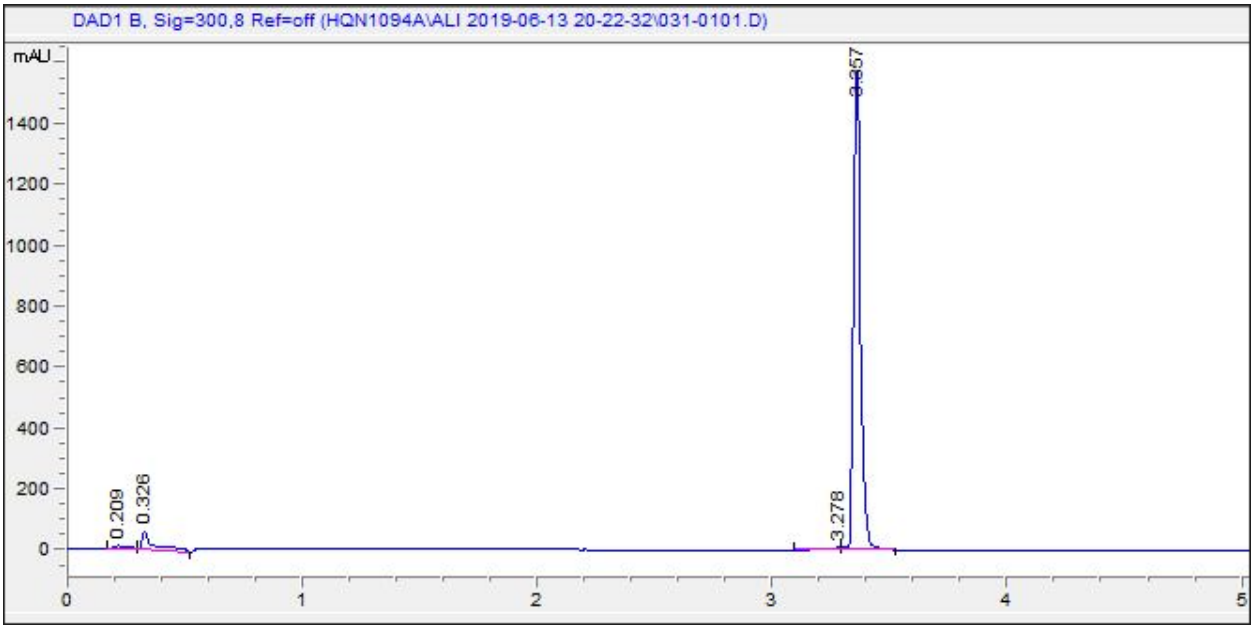


Figure 2.5: Analytical HPLC purity of compound **1094** with λ monitored at 300nm. Compound purity 94%, retention time 3.357 min.

2.2.9 Synthesis of desIm-Im-Py-Im-Py-γ-Py-Py-Py-Py-β-Ta. (PA1110)

Synthesis of 1110 resin.

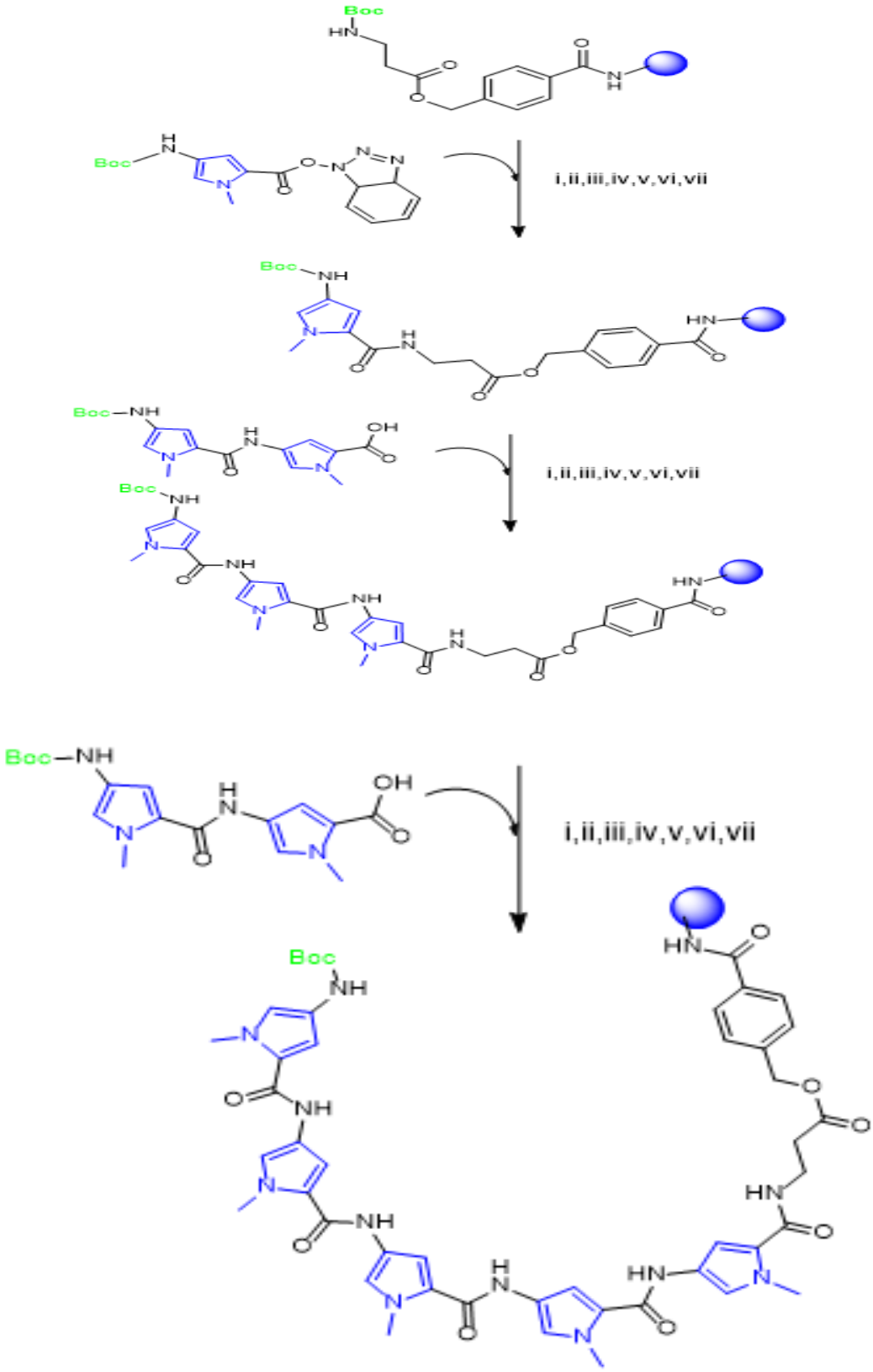
Synthesis of polyamides **1110** resin started by initiating the reaction sequences program with the CSBio program. The CSBio reaction vessel was set at a temperature of 35°C throughout the whole process. Boc-β-alanine-PAM resin (1000 mg) was loaded into the RV. PyBOP (11 g) was dissolved in DMF (40 mL) and filled in the R4 solvent bottle. After the program started the first wash and deprotection reaction for 1 h, a solution of a monomer Boc-Py-Obt (630 mg, 1.84 mmol) in DMF (6.5 mL) was stirred to dissolve and transferred into AA#3 fully. The coupling program directed the synthesizer to transfer the solution from AA#3 to TVA and then MVA to begin the activation before transferring to RV for 180 min, followed by capping with acetic anhydride for 17 min to remove impurities such as unreacted amines.

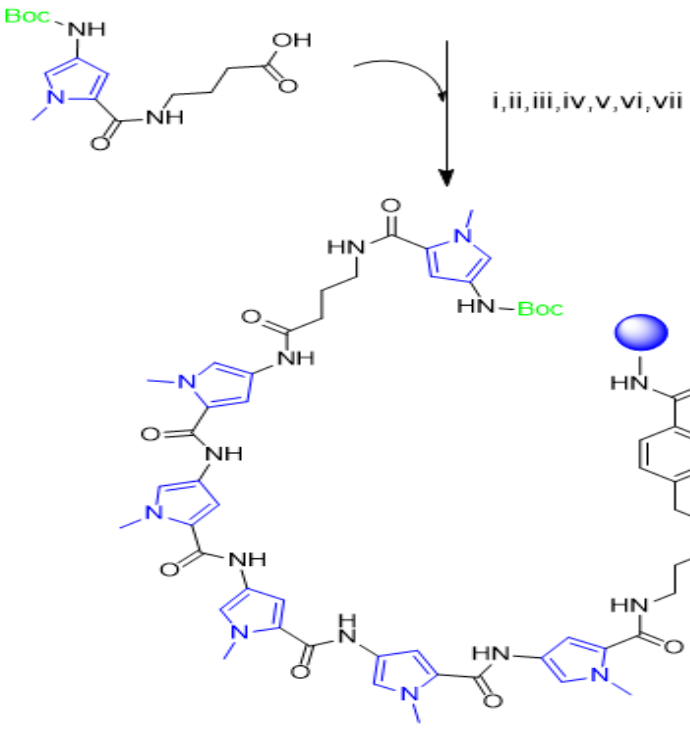
After completing the first sequence, the "cap" program removed any nonreacted amine before moving on to the following sequence. Each sequence took around 6 h total since it took around 3 h to wash the resin and deprotect the protection group before the coupling new building blocks

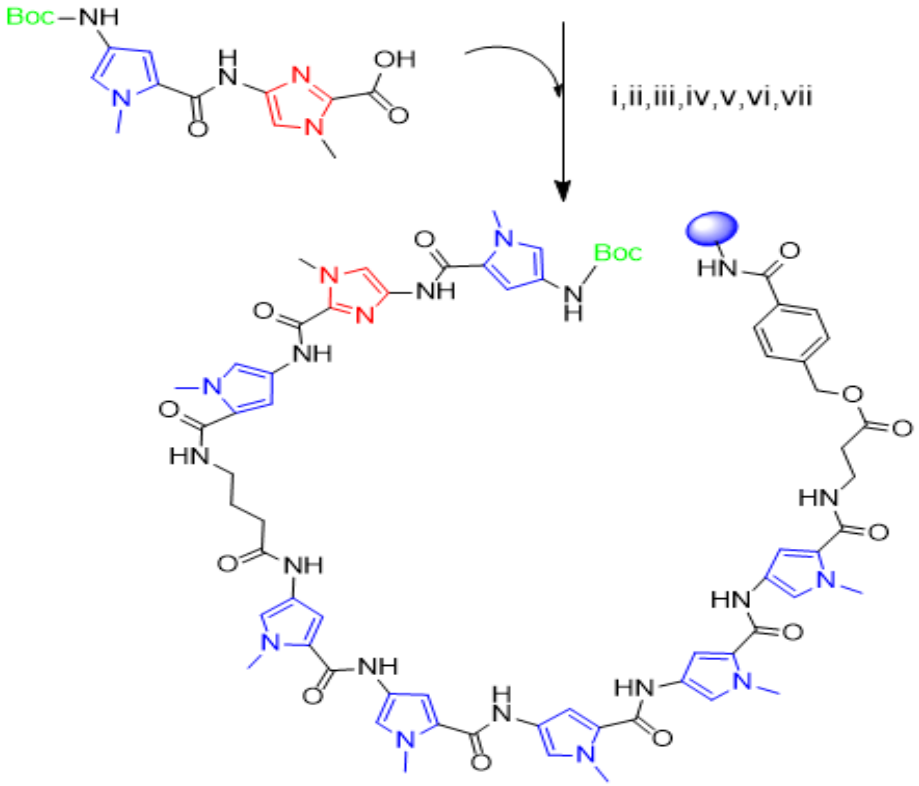
After completing the first sequence, the programmed system transferred a solution of dimer Boc-Py-Py-COOH (480 mg, 1.38 mmol) in DMF (6.5 mL), was well dissolved by vortex and loaded in AA#4 to TVA and MVA for sequence coupling following up by capping. After sequence 3 plus cap was completed, a solution of Boc-Py-Py-COOH (480 mg, 1.38 mmol) in DMF was well dissolved by vortex; loaded in AA#5 transfer to TVA then MVA before transferred to RV for coupling reaction for 180 min followed by capping with acetic anhydride for 17 min. After sequence 3 plus cap was completed, a solution of Boc-Py-γ-COOH (430 mg, 1.38 mmol) in DMF (6.5mL) was well dissolved by vortex; loaded in AA#6 transfer to TVA then MVA before transferred to RV for coupling reaction for 180 min following by capping with acetic anhydride for 17 min.

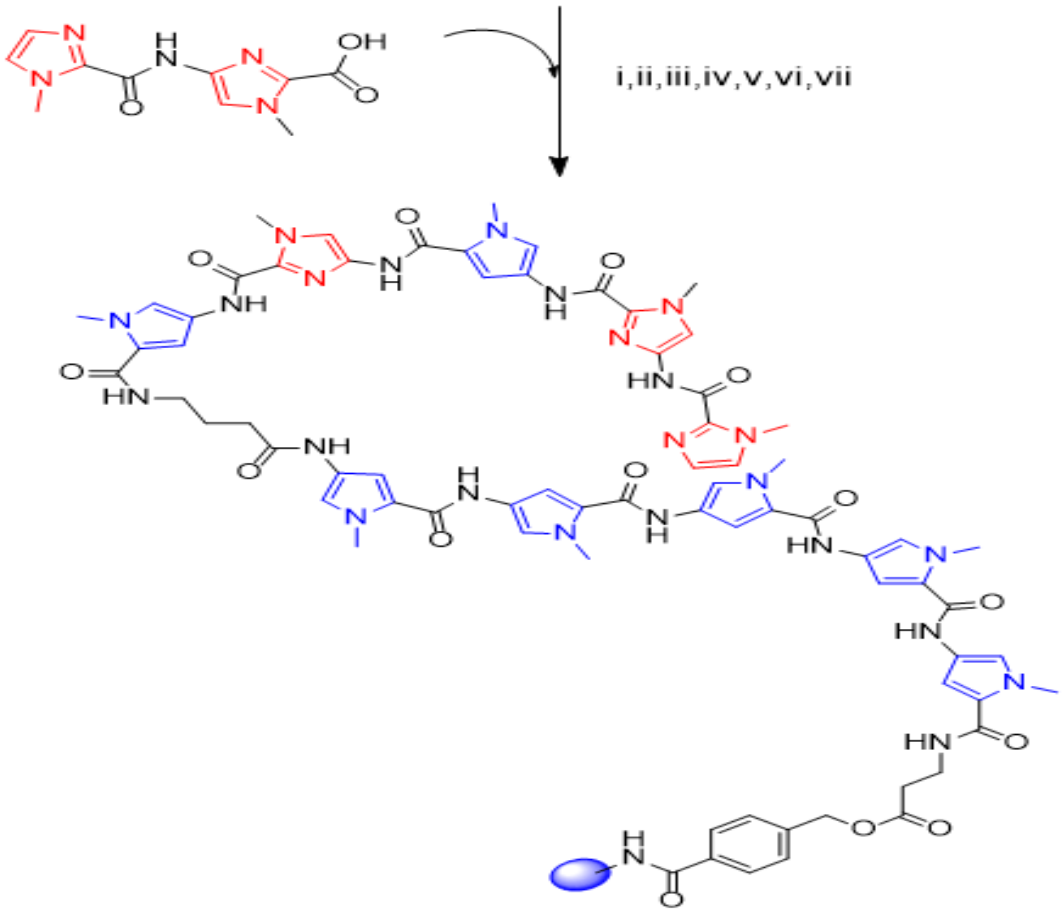
After sequence 4 was finished, the programed system was initiated and followed by transferring a well dissolved by vortex solution of Boc-Py-Im-COOH (481 mg, 1.38 mmol) in DMF (6.5 mL) and DIPEA (1 mL) from AA#6 to TVA then MVA before transferred to RV for coupling reaction for 180 min following by capping with acetic anhydride for 17 min. After sequence 5 was finished, the programed system was initiated and followed by transferring a well-dissolved by vortex solution of dimer desIm-Im-COOH **5** (330 mg, 1.38 mmol) and DIPEA (1ml) in DMF (6.5 mL) from AA#7 to TVA then MVA before transferred to RV for coupling reaction for 180 min following by capping with acetic anhydride for 17 min.

The sequence was finished after 35 h. 0,1 mg of resin was cleaved with 100 μL Ta solution for 2 h at 45°C, and the heated solution was collected and filtered before being analyzed with a Bruker LC-MS instrument. The exact mass was determined to be $[M+H]^+$ 1510.21 m/z, which confirmed the presence of the desired product. All the resin was carefully transferred from the RV into a beaker using a small spatula, and the resin was washed three-time with DCM and MeOH, then dried under vacuum for 30 min until reaching dryness. The resin was then collected in a 20 mL scintillation vial and stored in the freezer before cleavage. (**Scheme 2.10**)


Boc
HN
O
O
O
N
H
Boc
N
H
N
N
N
O
i,ii,iii,iv,v,vi,vii
Boc
NH
N
H
N
HN
O
O
Boc
NH
N
H
N
OH
O
i,ii,iii,iv,v,vi,vii
Boc
NH
N
NH
NH
N
HN
O
O
Boc
NH
N
H
N
OH
O
i,ii,iii,iv,v,vi,vii
Boc
NH
N
NH
N
NH
N
H
N
HN
O
O
HN
O
O
HN
O
O
HN
N


Boc
NH
OH
NH
i,ii,iii,iv,v,vi,vii
HN
HN–Boc
NH
HN
O
NH
HN
HN


Boc–NH
OH
i,ii,iii,iv,v,vi,vii
Boc
NH
HN
HN
HN
HN



Scheme 2.10 *Scheme of synthesis polyamide **1110** resin with solid-phase synthesis method. It started from a commercially acquired Boc-β-PAM resin, and each sequence contains multiple steps from (i) to (viii). (i) wash the resin with DCM and DMF, (ii) deprotection with 60% TFA/ DCM/ 0.5 M Indole, (iii) 25% piperidine/ DMF; (iv) building blocks with the structures shown on the scheme,(v) 0.5M PyBOP/DMF as coupling reagents, (vi) DIPEA, (vii) DMSO, (viii) capping step with acetic anhydride for 17 min.*

Synthesis of 1110 program sequence

Polyamide **110** was synthesized with 12 consecutive programs. The total coupling time was 37 h **(Table 2.3)**.

Table 2.3. Program sequence of **1110** resin synthesis.

Building block(s) and step(s)	Additional Solvent mixture	Applied program
1-Boc-Py-Obt	(6.5 mL) DMF	Add-DIPEA-180-min-DIPEA-Flows.CFN
2-Capping	None	Cap-17-min-continues.CFN
3-Boc-Py-Py-COOH	(6.5 mL) DMF	Add-DIPEA-180-min-DIPEA-Flows.CFN
4-Capping	None	Cap-17-min-continues.CFN
5-Boc-Py-Py-COOH	(6.5 mL) DMF	Add-DIPEA-180-min-DIPEA-Flows.CFN
6-Capping	None	Cap-17-min-continues.CFN
7-Boc-Py-γ-COOH	(6.5 mL) DMF	Add-DIPEA-180-min-DIPEA-Flows.CFN
8-Capping	None	Cap-17-min-continues.CFN
9-Boc Py-Im-COOH	(6.5 mL) DMF+ DIPEA (1mL)	Add-DIPEA-180-min-DIPEA-Flows.CFN
10-Capping	None	Cap-17-min-continues.CFN
11-desIm-Im-COOH	(6.5 mL) DMF+ DIPEA (1mL)	Add-DIPEA-180-min-DIPEA-Flows.CFN
12-Capping	None	Cap-17-min-continues.CFN

Cleavage of PAM resin using Ta as a nucleophile.

The next day, 380 mg resin was cleaved by aminolysis with 2000 µL of Ta at 45°C in a heating device Isotemp Fisher Scientific for 4 h, and the separation of polyamides solution from the PAM resin support was performed in a 20 µM polyethylene frit placed inside a disposable polypropylene syringe by washing with MeOH and H_2O. The polyamide solution was concentrated in a rotary evaporator to afford 28 mg PA **1110** (87.4 mmol, 21% yield) (**Scheme 2.11**).

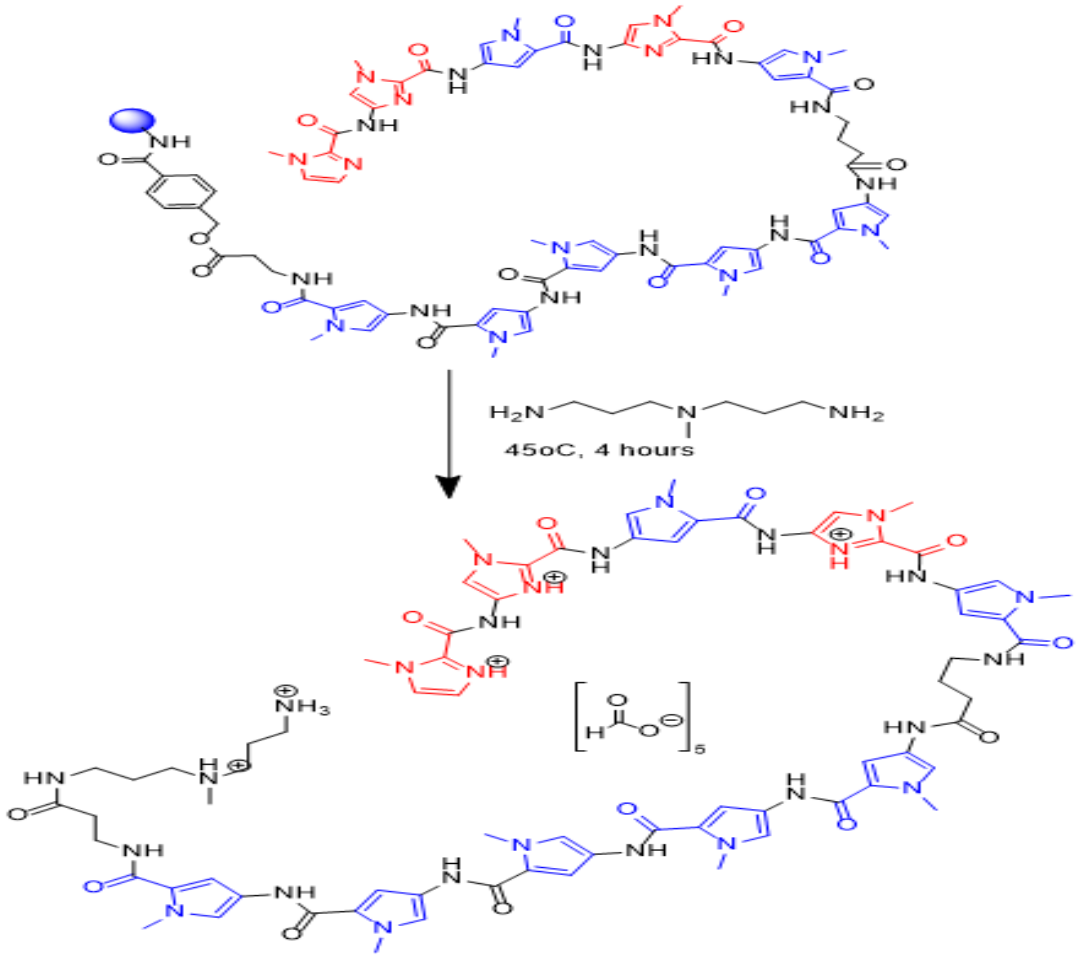


Figure 2.6 Cleavage of **1110** from PAM resin via Ta as the nucleophile.

Purification of the compound.

The polyamides solution was diluted with a mixture of 200 µL DMSO and H_2O/0.2% HCOOH (200 µL), then purified by preparative HPLC using a Phenomenex Luna 250x30 mm, 5 µM, 100Å, C18 column maintained at 25 °C. The organic phase was 0.2% HCOOH in MilliQ water and 100% HPLC grade MeOH.

The applied gradient was 10% MeOH for 10 min, followed by a ramp rate to 90% Methanol over 35.6 min at a 20 mL/min flow rate. All collected fractions were analyzed and selected for 95% or higher purity with analytical HPLC using ThermoFisher-UHPLC; analysis was processed and analyzed with the program Chromeleon7. High-purity fractions were collected for rotary evaporation to remove MeOH before lyophilization.

Lyophilization of compound

Compound **1110** was diluted with a 1:1 mixture of ACN (5 mL): and 0.2% HCOOH in H_2O (5mL). The samples were frozen before the lyophilization process with ground dry ice until an even coated of frozen material results at an acute angle in the vial. The frozen sample was covered with porous Kimwipe paper and secured by a rubber band. The sample was settled in a lyophilizer glass vessel and connected to a vacuum port. The frozen polyamides sublimed as an off-white solid powder after 48 h. The lyophilizer vacuum was set at 300 torr within 30 min prior to the lyophilization process, and the inside temperature was set at -73 to -75ºC.

Characterization of compound

Analytical HPLC characterization was performed with a C12 Phenomenex Juniper Proteo column maintained at 40ºC. The mobile phase consisted of 0.1% formic acid in water (A) and HPLC grade ACN (B). The applied gradient consisted of 5% B for 0.75 min, followed by a ramp to 60% B over 6.5 min at 2.0 mL/min. The retention was 2.68 min.

Characterization result of PA1110

PA **1110,** desImImPyImPyγPyPyPyPyPyβTa $(HCOO^-)_6$

^{1}H-NMR (600 MHz, DMSO-d_6) δ= 10.31 (s, 1H), 9.96 (s, 4H), 8.52 (s, 6H), 8.01- 7.98 (d, J= 17.18 Hz, 2H), 7.55- 7.52 (d, J= 18.02Hz, 2H), 7.44- 7.40 (d, J=26.0 Hz, 2H), 7.23 (s, 3H), 7.17-7.13 (t, J=20.4 Hz, 3H), 7.06- 7.03 (d, J=9.13 Hz, 3H), 6.97 (s, 1H), 6.88 (s, 1H), 6.83 (s, 1H), 3.99 (s, 4H), 3.96 (s, 2H), 3.85 (s, 2H), 3.83-3.81 (t, J=11.57 Hz, 6H), 3.79- 3.78 (d, J=10.96 Hz, 4H), 3.36- 3.32 (q, J= 19.49 Hz, 2H), 3.22-3.18 (q, J= 20.09 Hz, 2H), 3.05-3.01 (q, J= 20.71 Hz, 2H), 2.77 (s, 6H), 2.47 (s, 4H), 2.33 (br, 9H), 2.29- 2.25 (t, J= 14.61 Hz, 3H), 2.24- 2.21 (t, J= 14.01 Hz, 2H), 2.07 - 2.05 (d, J= 6.09 Hz, 6H), 1.79- 1.75 (q, J= 28.02 Hz, 2H), 1.67- 1.63 (q, J= 21.32 Hz, 6H), 1.51- 1.47 (t, J= 28.62 Hz, 2H).

^{13}C-NMR (151 MHz, DMSO-d_6) δ= 170.9, 169.7, 166.6, 165.8, 161.7, 161.7, 159.2, 159.0, 158.9, 158.9, 156.2, 156.2, 156.1, 138.3, 136.5, 135.2, 134.8, 134.6, 128.2, 127.5, 123.8, 123.2, 123.2, 123.2, 123.2, 122.7, 122.6, 122.6, 122.5, 121.8, 121.6, 120.5, 119.9, 118.9, 118.9, 118.9, 118.6, 118.4, 115.2, 114.5, 106.2, 105.3, 105.2, 105.2, 104.7, 104.7, 104.5, 55.0, 54.8, 54.4, 41.8, 41.7, 40.5, 40.4, 40.2, 40.1, 40.0, 39.8, 39.7, 39.5, 36.7, 36.5, 36.4, 36.0, 35.9, 35.6, 35.5, 35.4, 33.8 25.2, 26.18, 27.2.

HRMS (ESI) was calculated for **1110**, $C_{71}H_{87}N_{27}O_{12}$, $[M+H]^+$ 1510.71 m/z, and found 1510.21 m/z.

HPLC purity: 94% (Figure **2.6**)

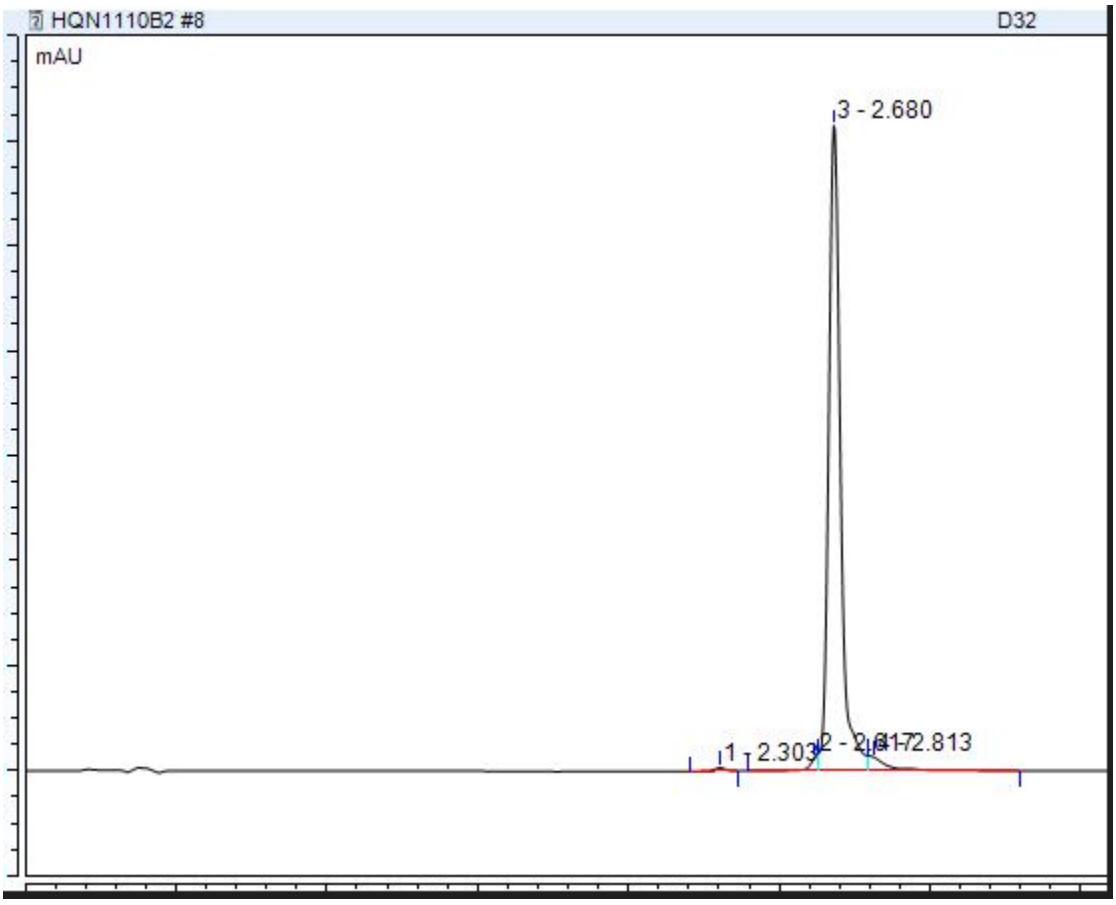


Figure 2.6: Analytical HPLC purity of compound **1110** with λ monitored. Compound purity 94%, retention time 2.680 min.

2.2.10 Synthesis of desIm-Im-Py-Im-Py-γ-Im-Py-Py-Py-Py-β-Ta (1123)

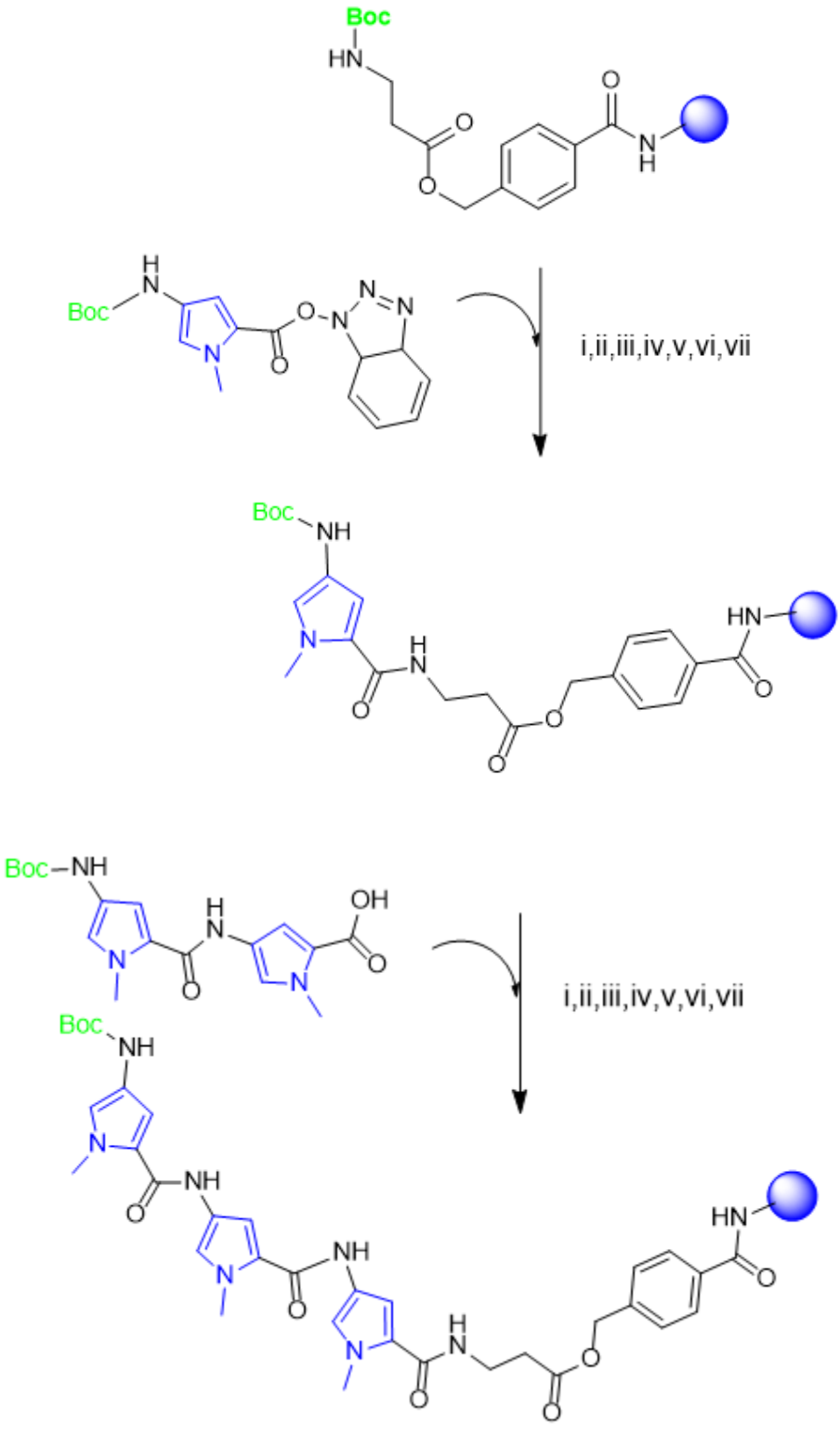
Synthesis of 1123 resin

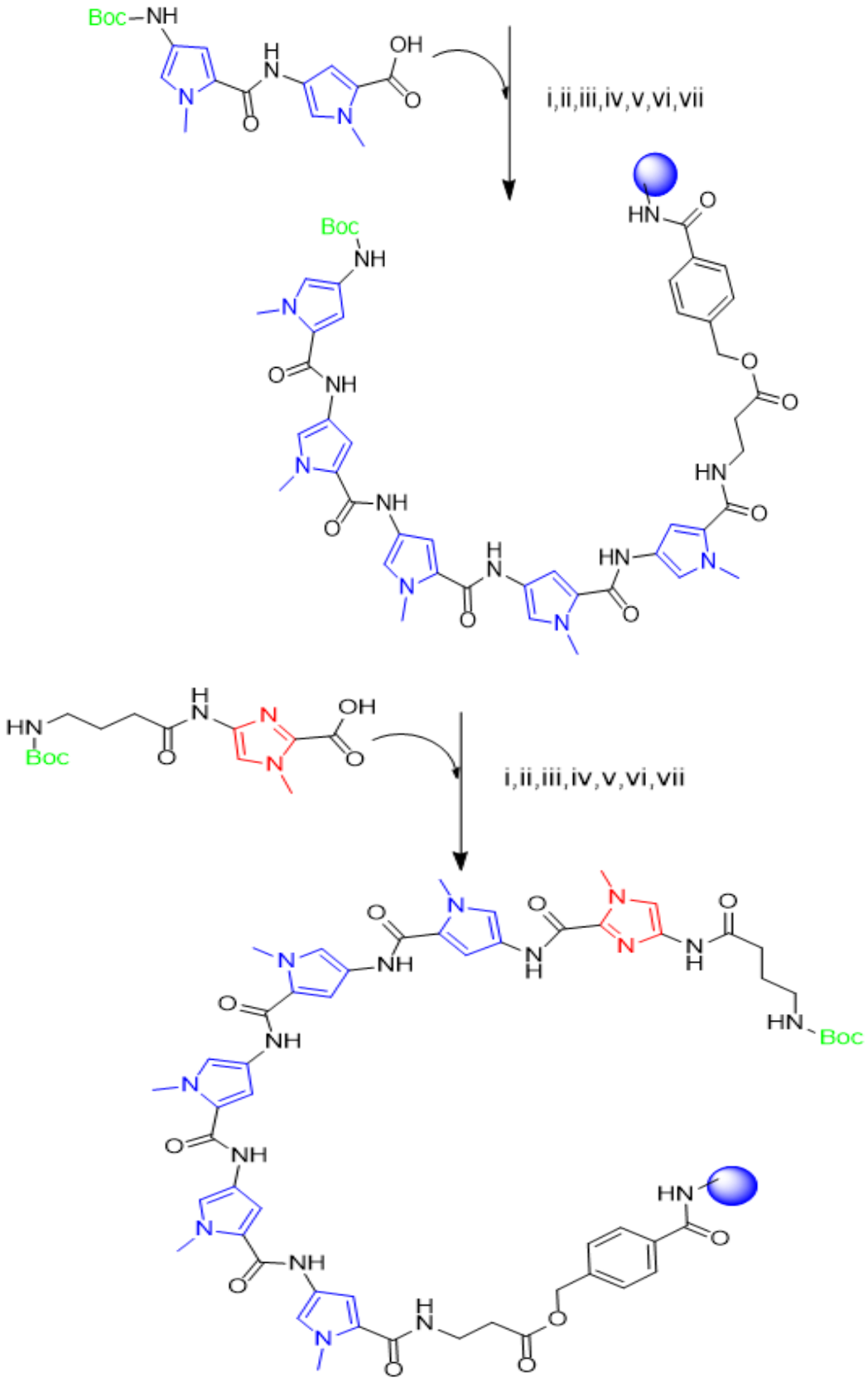
Synthesis of polyamide **1123** resin started by initiating the reaction sequences program with the CSBio program. The CSBio reaction vessel was set at a temperature of 35°C throughout the whole process. Boc-β-alanine-PAM resin (1000 mg) was loaded into the RV. PyBOP (11 g) was dissolved in DMF (40 mL) and filled in the R4 solvent bottle. After the program started the first wash and deprotection reaction for 1 h, a solution of Boc-Py-Py-COOH (500 mg, 1.38 mmol) in DMF (6.5 mL) was stirred to dissolve and transferred into AA#3 entirely. The coupling program directed the synthesizer to transfer the solution from AA#3 to TVA and then MVA to begin the activation before

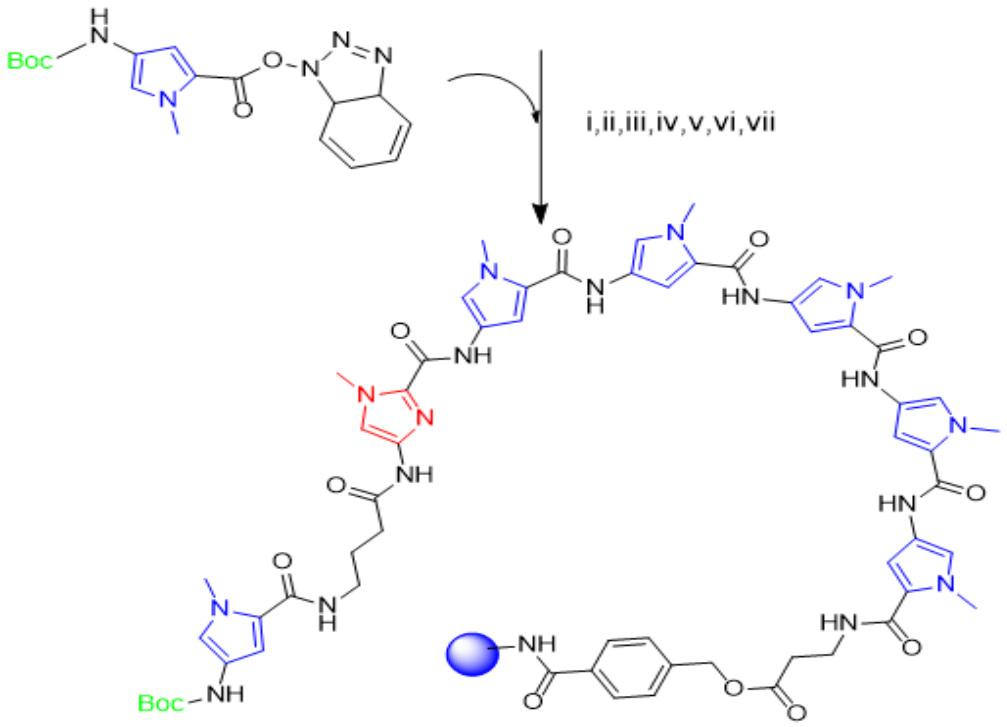
transferring to RV for coupling for 180 min, followed by capping with acetic anhydride for 17 min to remove impurities such as unreacted amines. After completing the first sequence, the program proceeded to "cap" to remove any nonreacted amine before moving on to the next sequence. Each sequence took around 6 h to complete, following by another 17 min of capping.

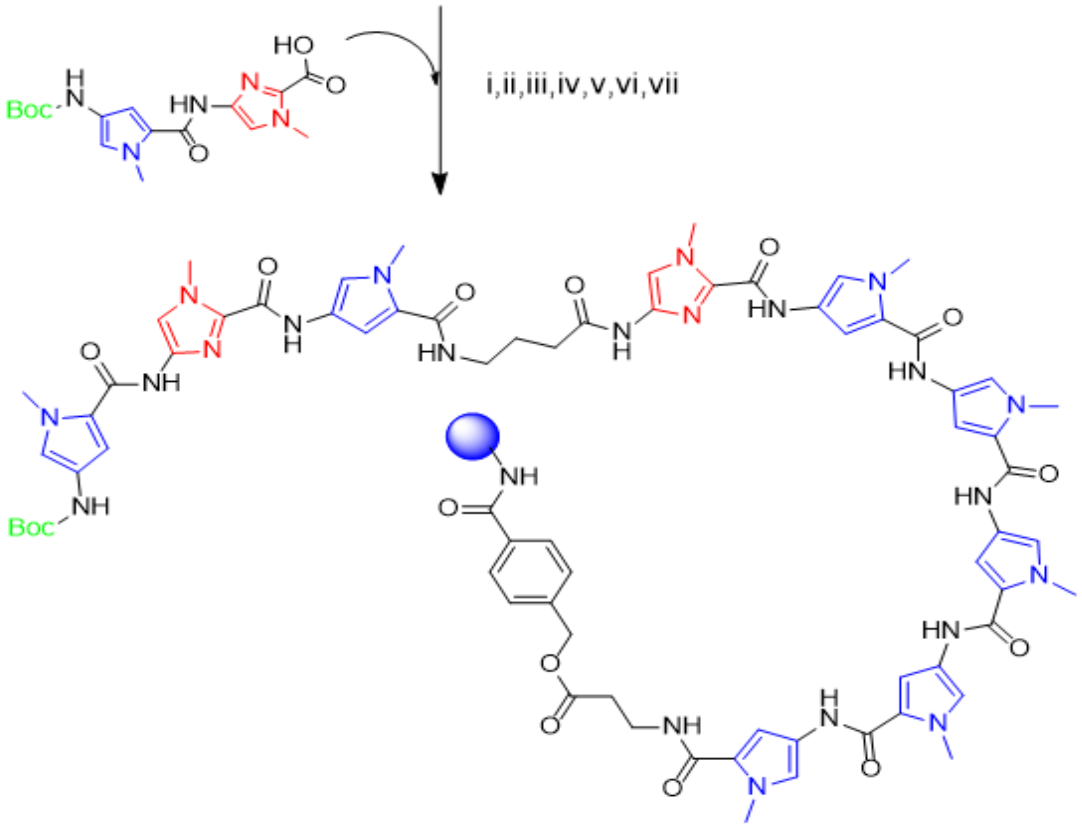
After completing the first sequence, the programmed system transferred a solution of Boc-Py-Py-COOH (501 mg, 1.38 mmol) in DMF (6.5 mL), was well dissolved by vortex and loaded in AA#4 to TVA and MVA for sequence coupling followed up by capping. After sequence 3 plus cap was completed, a solution of Boc-γ-Im-COOH (450 mg, 1.38 mmol) in DMF (6.5 mL) and DIPEA (1 mL) was well dissolved by vortex; loaded in AA#5 transferred to TVA then MVA before transferred to RV for coupling reaction for 180 min followed by capping with acetic anhydride for 17 min. After sequence 4 plus cap was completed, a solution of Boc-Py-OBt (660 mg, 1.84 mmol) in DMF (6.5 mL) was well dissolved by vortex; loaded in AA#6 transferred to TVA then MVA before transferred to RV for coupling reaction for 180 min followed by capping with acetic anhydride for 17 min. After sequence 5 was finished, the programed system was initiated and followed by transferring a well dissolved by vortex solution of Boc-Py-Im-COOH (500 mg, 1.38 mmol) in DMF (6.5 mL) and DIPEA (1 mL) from AA#6 to TVA, then MVA before transferred to RV for coupling reaction for 180 min following by capping with acetic anhydride for 17 min. After sequence 6 was finished, the programed system was initiated and followed by transferring a well dissolved by vortex solution of dimer desIm-Im-COOH **5** (345 mg, 1.38 mmol) and DIPEA (1ml) in DMF (6.5 mL) from AA#7 to TVA, then MVA before transferred to RV for coupling reaction for 180 min following by capping with acetic anhydride for 17 min. (Scheme **2.12**)

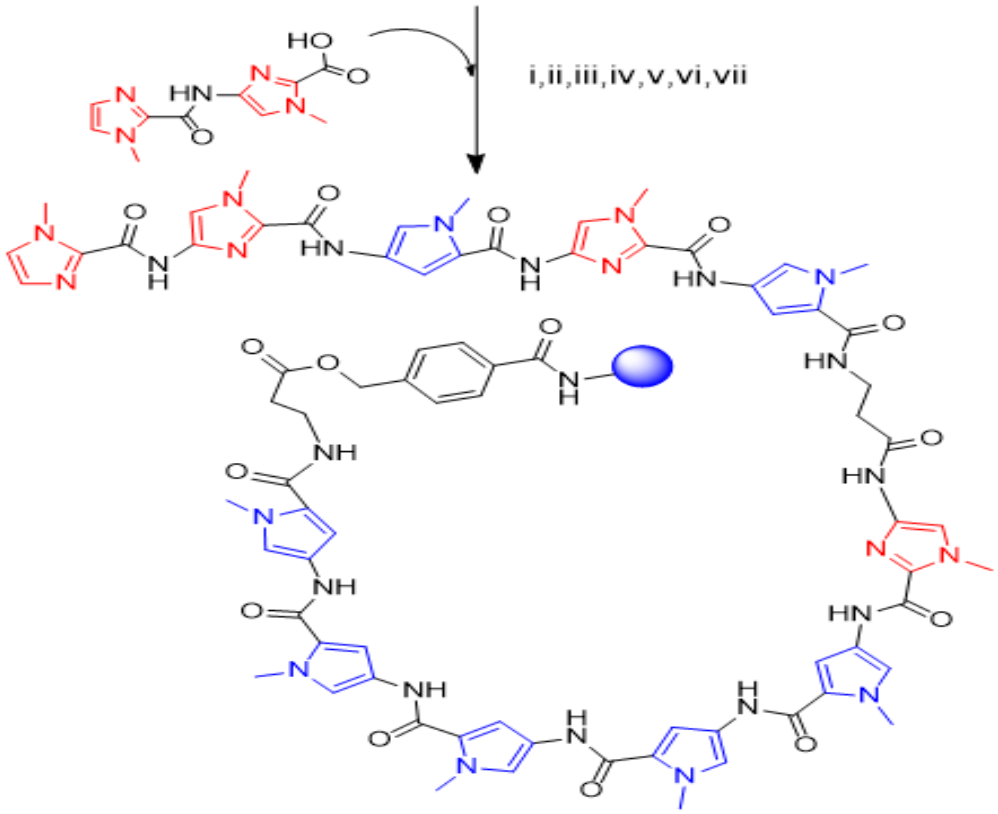
The sequence was finished after 30 h. 0,1 mg of resin was cleaved with 100 μL Ta solution for 2h at 45^{o}C, then the heated solution was collected and filtered before being analyzed with a Bruker LC-MS instrument. The exact mass was determined to be $[M+H]^+$ as 1511.76 m/z, which confirmed the presence of the desired product. All the resin was carefully transferred from the RV into a beaker using a small spatula, and the resin was washed three-time with DCM and MeOH, then dried under vacuum for 30 min until reaching dryness. The resin was then collected in a 20 mL scintillation vial and stored in the freezer before cleavage the next day.


Boc
HN
O
O
O
N
H
Boc
H
N
N
N
N
O
O
i,ii,iii,iv,v,vi,vii
Boc
NH
HN
N
H
N
O
O
O
O
Boc
NH
N
H
N
OH
O
N
O
i,ii,iii,iv,v,vi,vii
Boc
NH
N
NH
NH
N
N
HN
O
H
N
O
O
O


Boc–NH
OH
i,ii,iii,iv,v,vi,vii
Boc
HN
HN–Boc


Boc
i,ii,iii,iv,v,vi,vii
Boc


Boc
HO
i,ii,iii,iv,v,vi,vii
Boc



Scheme 2.12 *Scheme of synthesis polyamide **1123** resin by applying solid-phase synthesis method. Started from a commercially acquired Boc-β-PAM resin, each sequence containing multiple steps from (i) to (viii). (i) wash the resin with DCM and DMF, (ii) deprotection with 60% TFA/ DCM/ 0.5 M Indole, (iii) 25% piperidine/ DMF; (iv) building blocks with the structures shown on the scheme, (v) 0.5M PyBOP/ DMF as coupling reagents, (vi) DIPEA, (vii) DMSO, (viii) capping step with acetic anhydride for 17 min.*

Polyamide **1123** resin was synthesized with 12 consecutive programs. The total coupling time was 30 h **(Table 2.4)**.

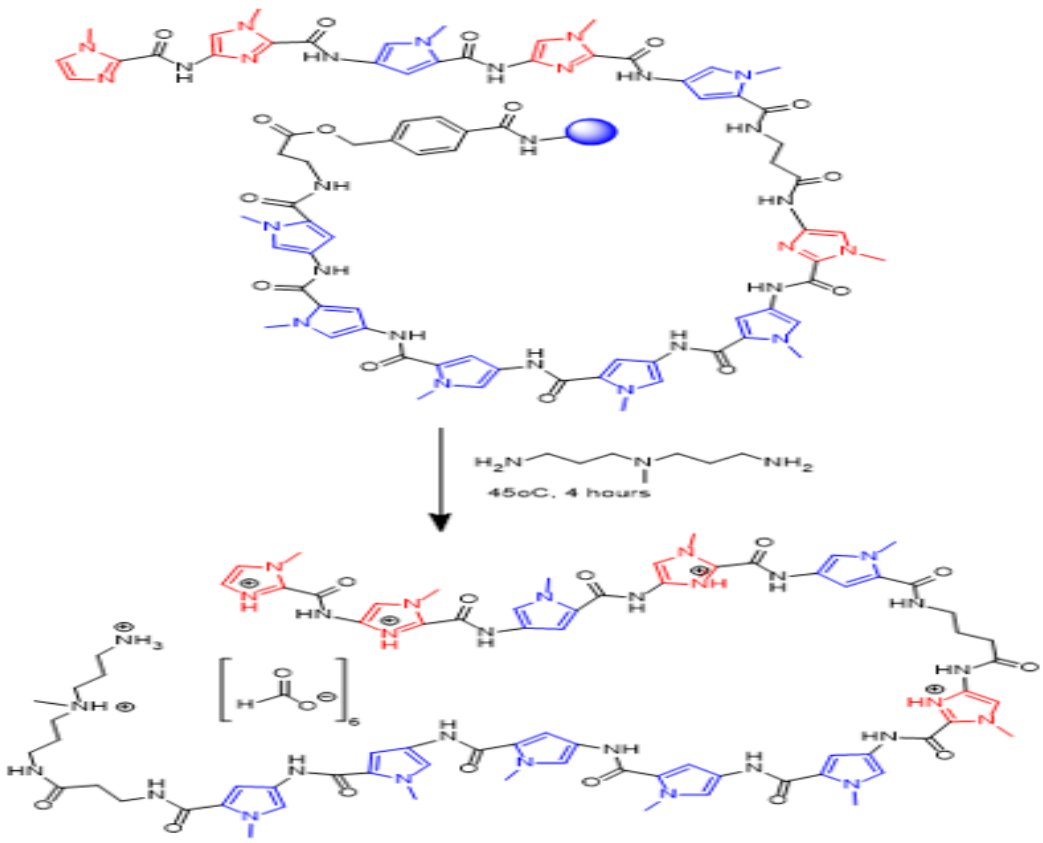
Table 2.4. Program sequences of **1123** resin synthesis.

Building block(s) and step(s)	Additional Solvent mix-ture	Applied program
1-Boc-Py-Py-COOH	(6.5 mL) DMF	Add-DIPEA-180-min-DIPEA-Flows.CFN
2-Capping	None	Cap-17-min-continues.CFN

3-Boc-Py-Py-COOH	(6.5 mL) DMF	Add-DIPEA-180-min-DIPEA-Flows.CFN
4-Capping	None	Cap-17-min-continues.CFN
5-Boc-γ-Im-COOH	DMF (6.5 mL) + DIPEA (1mL)	Add-DIPEA-180-min-DIPEA-Flows.CFN
6-Capping	None	Cap-17-min-continues.CFN
7-Boc-Py-OBt	(6.5 mL) DMF	Add-DIPEA-180-min-DIPEA-Flows.CFN
8-Capping	None	Cap-17-min-continues.CFN
9-Boc Py-Im-COOH	6.5 mL DMF + DIPEA (1mL)	Add-DIPEA-180-min-DIPEA-Flows.CFN
10-Capping	None	Cap-17-min-continues.CFN
11-desIm-Im-COOH	6.5 mL DMF + DIPEA (1mL)	Add-DIPEA-180-min-DIPEA-Flows.CFN
12-Capping	None	Cap-17-min-continues.CFN

Cleavage of PAM resin using Ta as a nucleophile.

The next day, the 383 mg resin was cleaved via aminolysis in 2000 μL of Ta at 45°C in a heating device Isotemp Fisher Scientific for 4 h, and the separation of polyamides solution from the PAM resin support was performed in a 20 μM polyethylene frit placed inside a disposable polypropylene syringe by washing with MeOH and H_2O. The polyamide solution was concentrated in the rotary evaporator to afford a white solid of **1123** (33 mg, 0.087 mmol, 25% yield) (**Scheme 2.13**).



Scheme 2.13 Cleavage of **1123** from PAM resin using Ta as the nucleophile.

Purification of compound

The polyamides solution was diluted with a mixture of DMSO (200 µL) and H_2O/0.2% HCOOH (200 µL), then purified by preparative HPLC using a Phenomenex Luna 250x30 mm, 5 µM, 100Å, C18 column maintained at 25ºC. The organic phase was 0.2% HCOOH in MilliQ water and 100%HPLC grade MeOH. The applied gradient was 10% MeOH for 8 min, followed by a ramp rate to 90% Methanol over 35.6 min at a 20 mL/ min flow rate. All collected fractions were analyzed and selected for 95% or higher purity with analytical HPLC using ThermoFisher -UHPLC; analysis was processed and analyzed with the program Chromeleon 3.0. High-purity fractions were collected for rotary evaporation to remove MeOH before the lyophilization.

Lyophilization of compound

Compound **1123** was diluted with a 1:1 mixture of ACN (5mL): and 0.2% HCOOH in H_2O(5mL). The samples were frozen before the lyophilization process with ground dry ice until an even coated of frozen material results at an acute angle in the vial. The frozen sample was covered with porous Kimwipe paper and secured by a rubber band. The sample was settled in a lyophilizer glass vessel

and connected to a vacuum port. The frozen polyamide solutions sublimed as an off-white solid powder after 48 h. The lyophilizer vacuum was set at 300 torr within 30 min prior to the lyophilization process, and the inside temperature was set at -83 to -85°C

Characterization of the compound.

Analytical HPLC characterization was performed with a C12 Phenomenex Juniper Proteo column maintained at 40°C. The mobile phase consisted of 0.1% formic acid in water (A) and HPLC grade ACN (B). The applied gradient consisted of 5% B for 0.75 min, followed by a ramp to 60% B over 6.5 min at 2.0 mL/min. The retention was 3.17 min.

Characterization result of PA 1123

PA-**1123**, desImImPyImPyγImPyPyPyPyβTa $(HCOO^-)_6$

^{1}H-NMR (600 MHz, DMSO-d_6) δ= 10.28 (s, 1H), 10.24 (s, 1H), 9.96- 9.93 (J= 18.3 Hz, t, 4H), 8.36 (s, 8H), 8.07- 8.04 (t, J= 13.15 Hz, 1H), 8.01- 7.99 (t, J= 13.15 Hz, 2H), 7.96-7.94 (t, J=11.34 Hz, 2H), 7.56 (s, 1H), 7.53 (s, 1H), 7.44 (s, 1H), 7.39 (s, 1H), 7.26 (s, 1H), 7.21-7.17 (d, J= 9.55 Hz, 3H), 7.11 (s, 1H), 7.14 (s, 1H), 7.06 (s, 1H), 7.04 (s, 1H), 6.97 (s, 1H), 6.83 (s, 1H), 3.98 (s, 2H), 3.93 (s, 1H), 3.84 (s, 1H), 3.83- 3.82 (d, J= 4.69 Hz, 3H), 3.78-3.77 (d, J= 7.38 Hz, 3H), 3.35-3.32 (q, J=19.73 Hz, 3H), 3.20- 3.16 (q, J= 21.05 Hz, 3H), 3.05- 3.01 (q, J= 19.73 Hz, 3H), 2.82-2.78 (q, J=19.73 Hz, 6H), 2.77- 2.75 (t, J= 14.02 Hz, 9H), 2.40 (s, 2H), 2.33 (s, 4H), 2.32- 2.31 (t, J= 12.84 Hz, 8H), 2.07 (s, 8H), 1.68- 1.64 (q, J=26.48 Hz, 10 H).

^{13}C-NMR (151 MHz, DMSO-d_6) δ= 170.9, 169.7, 166.6 161.7, 161.7, 159.2, 159.0, 158.9, 158.9, 156.2, 156.2, 156.1, 138.3, 136.5, 135.2, 134.8, 134.6, 128.2, 128.1, 127.5, 123.8, 123.2, 123.2, 122.7, 122.7, 122.7, 122.7, 122.6, ,122.6, 122.6, 121.8, 121.6, 118.9, 117.4, 118.4, 115.1, 114.5, 106.2, 105.2, 105.3, 105.2, 105.2, 104.7, 104.7, 104.6, 104.5, 54.9, 54.8, 54.4, 41.8, 41.7, 40.5, 40.4, 40.2, 40.1, 39.9, 39.8, 39.7, 39.5, 39.4, 39.4, 39.3, 39.3, 39.3, 39.3, 38.7, 38.0, 37.6, 37.1, 36.6, 36.5, 37.0, 35.9, 35.9, 35.6, 35.6, 35.4, 27.2, 26.2, 25.2.

HRMS (ESI) was calculated for **1123**, $[M+H]^+$, 1511.71 m/z, and found to be 1511.76 m/z.

HPLC purity: 94%

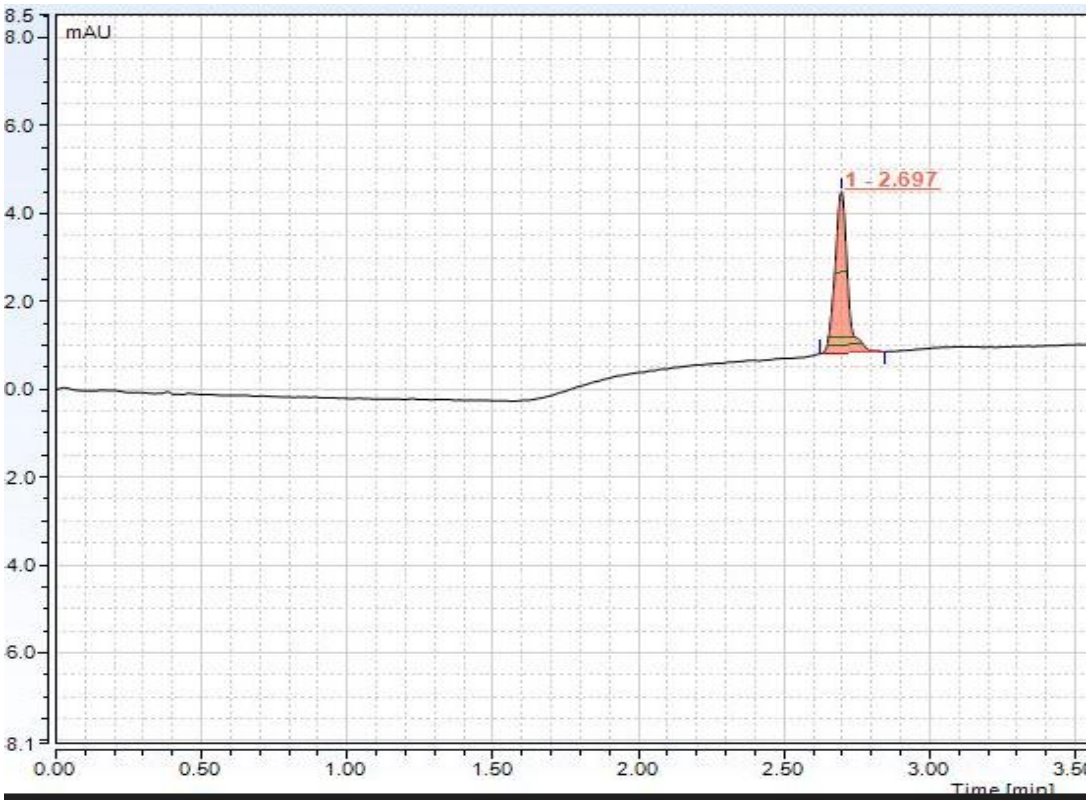


Figure 2.7 Analytical HPLC purity of compound **1123** with λ monitored. Compound purity 94%, retention time 2.697 min.

2.2.11 Synthesis of desIm-Im-β-Im-Py-γ-Im-Py-Py-β-Py-β-Ta (1131)

Synthesis of 1131 resin

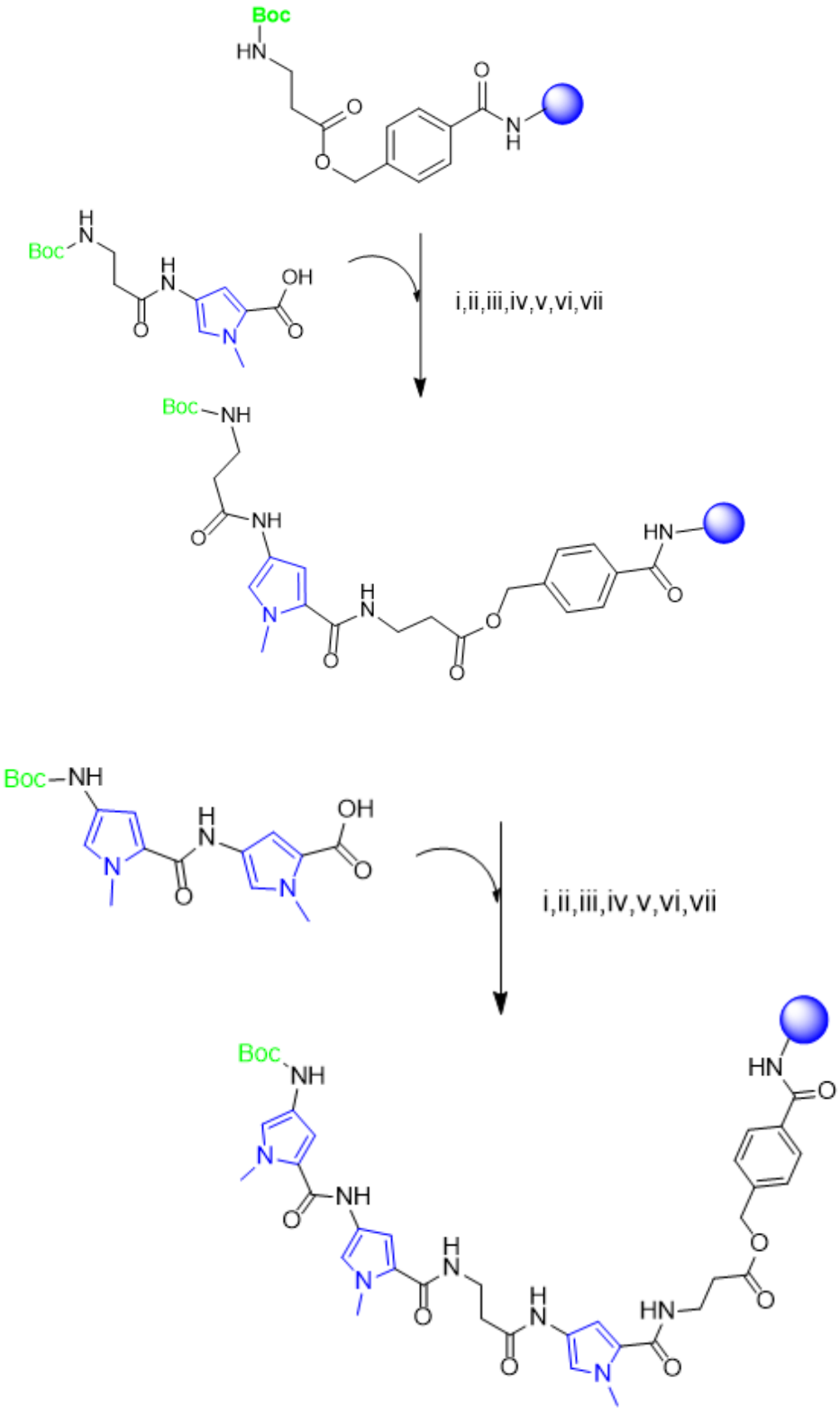
Synthesis of polyamides **1131** resin started by initiating the reaction sequences program with the CSBio program. The CSBio reaction vessel was set at a temperature of 35 ºC throughout the whole process. Boc-β-alanine-PAM resin (1000 mg) was loaded into the RV. PyBOP (11 g) was dissolved in DMF (40 mL) and filled in the R4 solvent bottle.

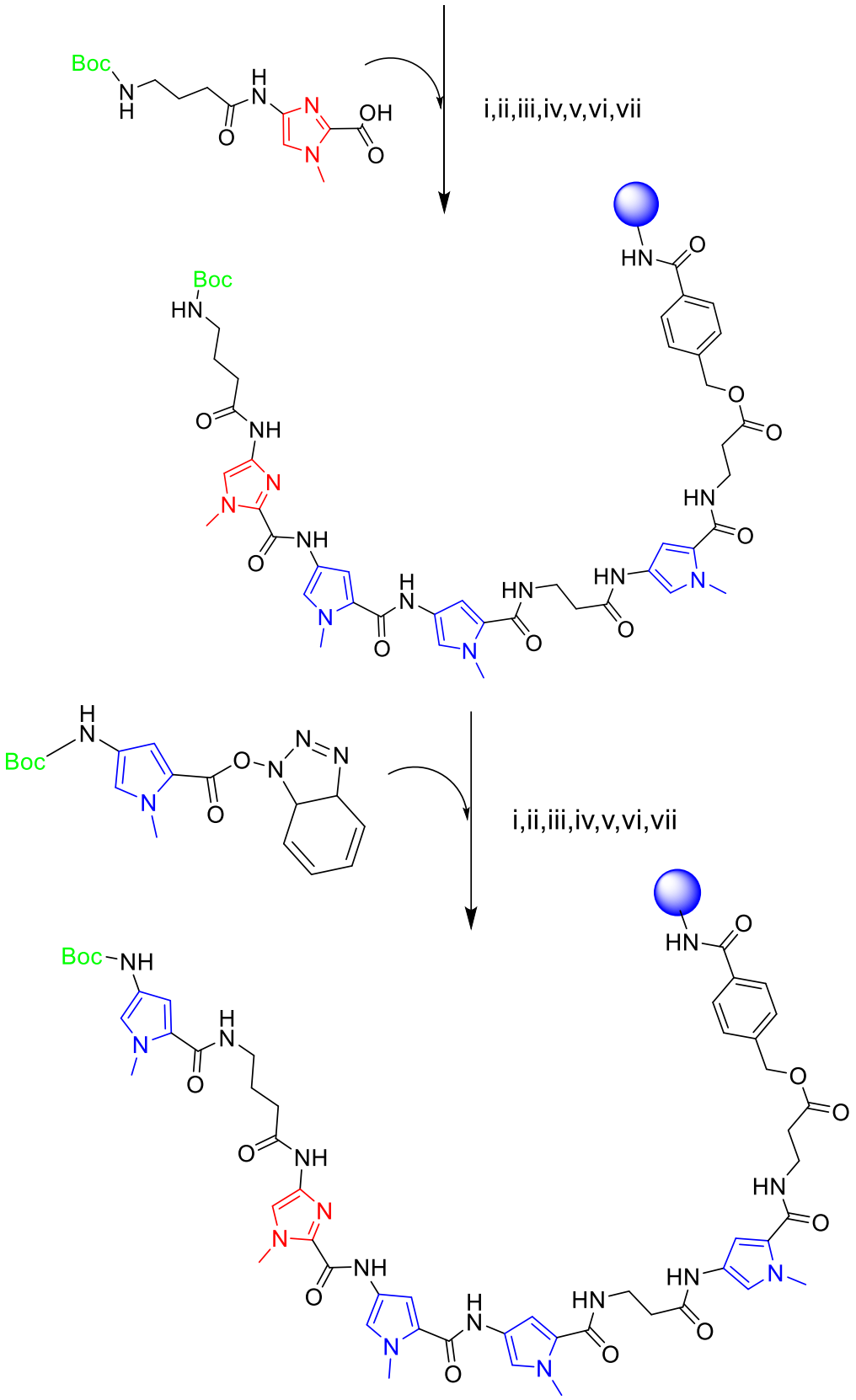
After the program started the first wash and deprotection reaction for 1h, a solution of a monomer Boc-β-Py-COOH (430 mg, 1.38 mmol) in DMF (6.5 mL) was stirred to fully dissolved and transferred into AA#3. The coupling program directed the synthesizer to transfer the solution from AA#3 to TVA and then MVA to begin the activation before transferring to RV for coupling for 180 min, followed by capping with acetic anhydride for 17 min to remove impurities such as unreacted amines. After completing the first sequence, the program proceeded to "cap" to remove any nonreacted amine before moving on to the next sequence. Each sequence took around 6h to complete.

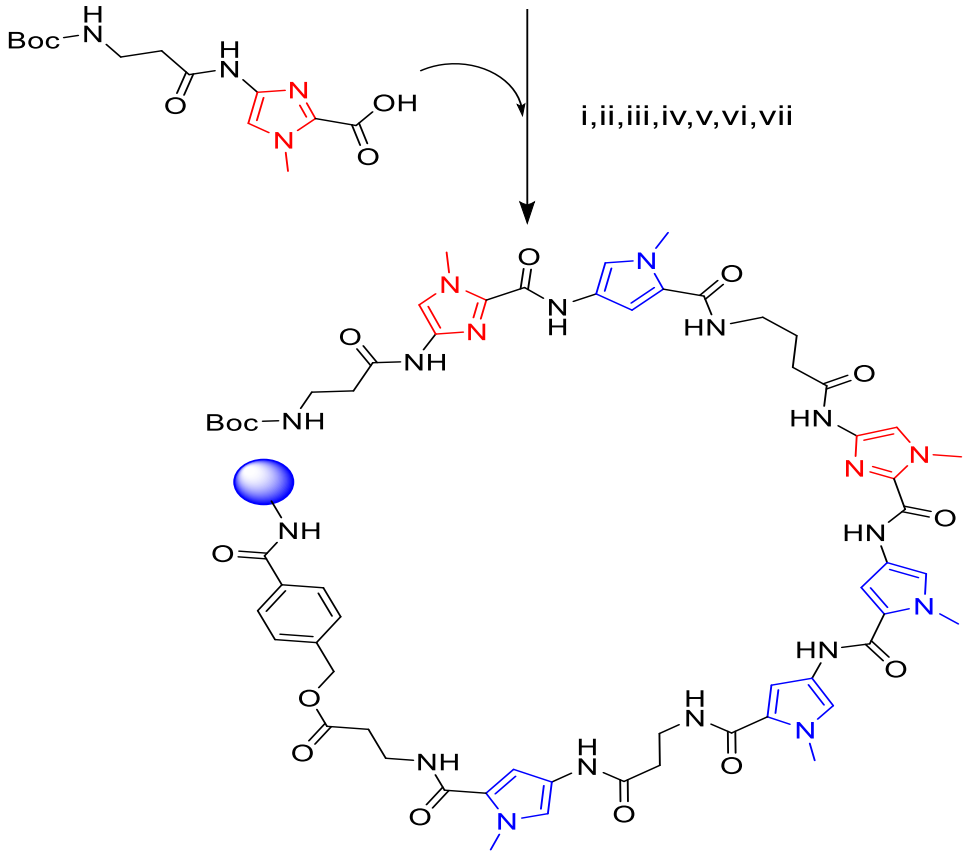
After completing the first sequence, the programmed system transferred a solution of Boc-Py-Py-COOH (500 mg, 1.38 mmol) in DMF (6.5 mL), was well dissolved by vortex and loaded in AA#4 to TVA and MVA for sequence coupling followed up by capping. After sequence 2 plus cap was completed, a solution of Boc-γ-Im-COOH (450 mg, 1.38 mmol) in DMF (6.5 mL) and DIPEA (1mL) was well dissolved by vortex; loaded in AA#5 transferred to TVA then MVA before transferred to RV for coupling reaction for 180 min followed by capping with acetic anhydride for 17 min.

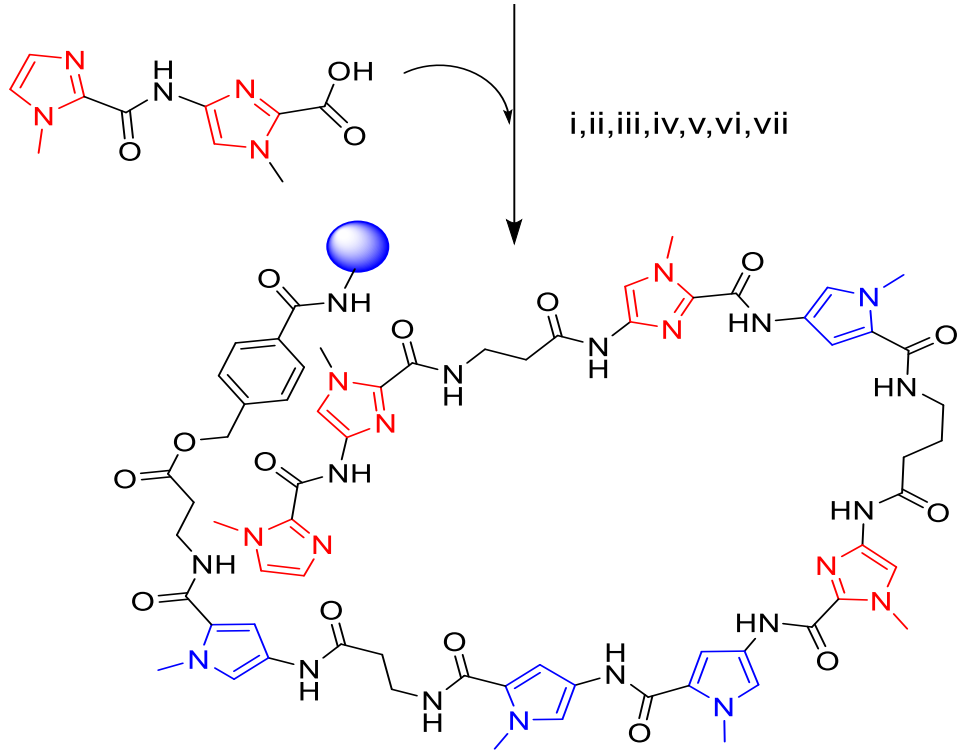
After sequence 3 plus cap was completed, a solution of Boc-Py-Obt (657 mg, 1.84 mmol) in 6.5mL DMF was well dissolved by vortex; loaded in AA#6, transferred to TVA then MVA before transferred to RV for coupling reaction for 180 min followed by capping with acetic anhydride for 17 min. After sequence 4 was finished, the programed system was initiated and followed by transferring a well-dissolved by vortex solution of Boc-β-Im-COOH (432 mg, 1.38 mmol) in DMF (6.5 mL) and DIPEA (1mL) from AA#6 to TVA then MVA before transferred to RV for coupling reaction for 180 min following by capping with acetic anhydride for 17 min. After sequence 5 was finished, the programed system was initiated and followed by transferring a well-dissolved by vortex solution of dimer desIm-Im-COOH (345 mg, 1.38 mmol) in DMF (6.5 mL) and DIPEA (1mL) from AA#7 to TVA then MVA before transferred to RV for coupling reaction for 180 min following by capping with acetic anhydride for 17 min. **(Scheme 2.14)**

The sequence was finished after 30 h. 0,1 mg of resin was cleaved with 100 μL Ta solution for 2 h at 45^{o}C, then the heated solution was collected and filtered before being analyzed with a Bruker LC-MS instrument. The exact mass was determined to be $[M+H]^{+}$, 1460.73 m/z, which confirmed the presence of the desired product. All the resin was carefully transferred from the RV into a beaker using a small spatula, and the resin was washed three-time with DCM and MeOH, then dried under vacuum for 30 min until reaching dryness. The resin was then collected in a 20 mL scintillation vial and stored in the freezer before cleavage the next day.


Boc
HN
O
O
O
N
H
Boc
H
N
H
N
OH
O
O
N
i,ii,iii,iv,v,vi,vii
Boc
NH
NH
HN
O
N
H
N
O
O
O
Boc
NH
N
H
N
OH
O
N
i,ii,iii,iv,v,vi,vii
Boc
NH
HN
O
N
NH
O
O
N
H
N
O
O
HN
O
O
N


Boc
OH
i,ii,iii,iv,v,vi,vii
Boc
i,ii,iii,iv,v,vi,vii
Boc


i,ii,iii,iv,v,vi,vii


i,ii,iii,iv,v,vi,vii

Scheme 2.14 *Scheme of synthesis polyamide **1131** resin with solid-phase synthesis method. It started from a commercially acquired Boc-β-PAM resin, and each sequence contains multiple steps from (i) to (viii). (i) wash the resin with DCM and DMF, (ii) deprotection with 60% TFA/ DCM/ 0.5 M Indole, (iii) 25% piperidine/ DMF; (iv) building blocks with the structures shown on the scheme, (v) 0.5M PyBOP/ DMF as coupling reagents, (vi) DIPEA, (vii) DMSO, (viii) capping step with acetic anhydride for 17 min.*

Polyamide **1131** resin was synthesized with 12 consecutive programs. The total coupling time was 30 h **(Table 2.5)**.

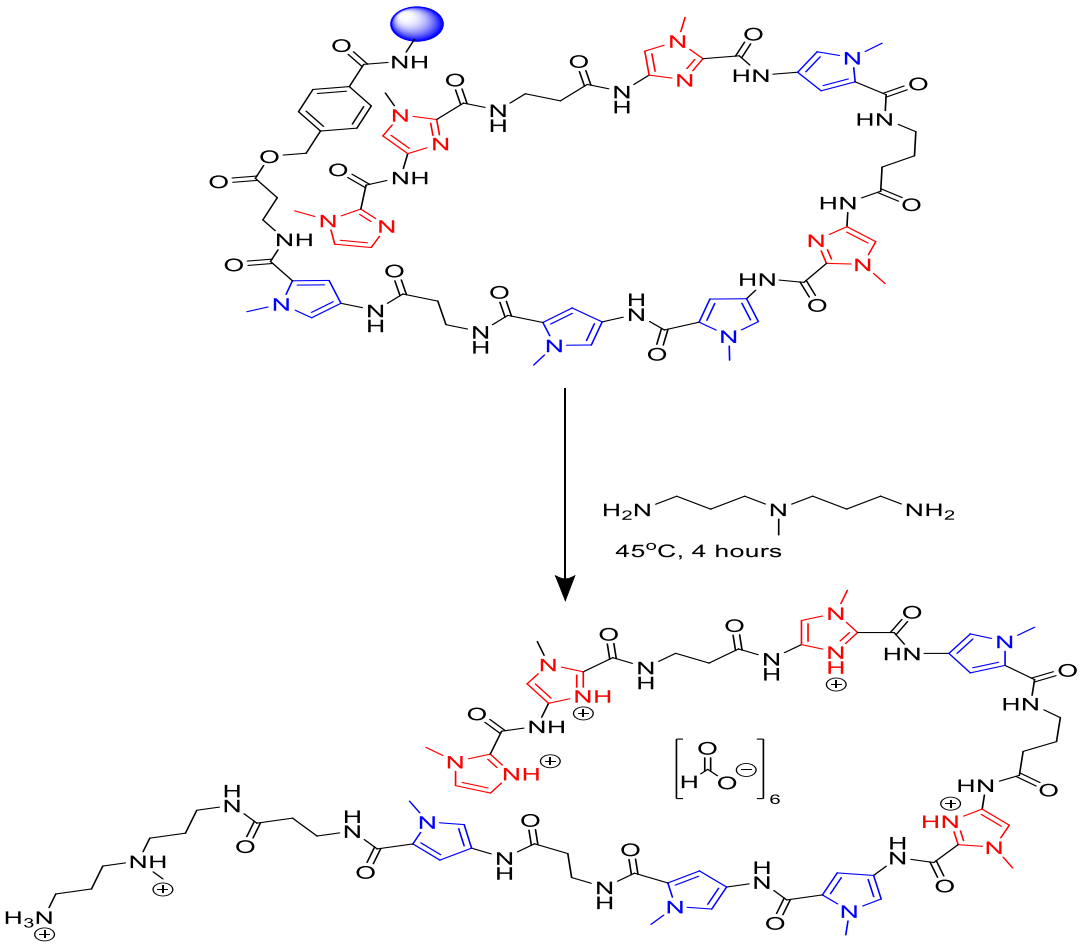
Table 2.5. Programmed sequence of **1131** resin synthesis.

Building block(s) and step(s)	Additional Solvent mixture	Applied program
1-Boc-β-Py-COOH	(6.5 mL) DMF	Add-DIPEA-180-min-DIPEA-Flows.CFN
2-Capping	None	Cap-17-min-continues.CFN
3-Boc-Py-Py-COOH	(6.5 mL) DMF	Add-DIPEA-180-min-DIPEA-Flows.CFN
4-Capping	None	Cap-17-min-continues.CFN
5-Boc-γ-Im-COOH	(6.5 mL) DMF + DIPEA (1mL)	Add-DIPEA-180-min-DIPEA-Flows.CFN
6-Capping	None	Cap-17-min-continues.CFN
7-Boc-Py-OBt	(6.5 mL) DMF	Add-DIPEA-180-min-DIPEA-Flows.CFN
8-Capping	None	Cap-17-min-continues.CFN
9-Boc β-Im-COOH	(6.5 mL) DMF + DIPEA (1mL)	Add-DIPEA-180-min-DIPEA-Flows.CFN
10-Capping	None	Cap-17-min-continues.CFN
11-desIm-Im-COOH	(6.5 mL) DMF + DIPEA (1mL)	Add-DIPEA-180-min-DIPEA-Flows.CFN
12-Capping	None	Cap-17-min-continues.CFN

Cleavage of PAM resin using Ta as a nucleophile.

The next day, 275 mg resin was cleaved with aminolysis in 2000 μL of Ta at 45 ºC in a heating device Isotemp Fisher Scientific for 4 h, and the separation of polyamides solution from

the PAM resin support was performed in a 20 µM polyethylene frit placed inside a disposable polypropylene syringe by washing with MeOH and H_2O. The polyamide solution was concentrated in the rotor evaporator before purification with preparative HPLC to afford an off-white solid of **1131** (34 mg, 0.63 mmol, 37% yield) (**Scheme 2.15**).



Scheme 2.15 Cleavage of PA**1131** from PAM resin using Ta as the nucleophile.

Purification of compound

The polyamides solution was diluted with a mixture of 200 µL DMSO and H_2O/0.2% HCOOH (200 µL), then purified by preparative HPLC using a Phenomenex Luna 250x30 mm, 5 µM, 100 Å, C18 column maintained at 25 ºC. The organic phase was 0.2% HCOOH in MilliQ water and 100%HPLC grade MeOH. The applied gradient was 10% MeOH for 8 min, followed by a ramp rate to 90% Methanol over 35.6 min at a 20 mL/ min flow rate. All collected fractions were analyzed and selected for 95% or higher purity with analytical HPLC using ThermoFisher - UHPLC; analysis was processed and analyzed with the program Chromeleon 3.0. High-purity

fractions were collected for rotary evaporation to remove MeOH from the solution before the lyophilization process.

Lyophilization of compound

Compound **1131** was diluted with a 1:1 mixture of ACN (5 mL) and 0.2% HCOOH in H_2O (5 mL). The samples were frozen before the lyophilization process with ground dry ice until an even coated of frozen material results at an acute angle in the vial. The frozen sample was covered with porous Kimwipe paper and secured by a rubber band. The sample was settled in a lyophilizer glass vessel and connected to a vacuum port. The frozen polyamide solutions sublimed as an off-white solid powder after 48 h. The lyophilizer vacuum was set at 300 torr within 30 min prior to the lyophilization process, and the inside temperature was set at -83 to -85°C.

Characterization of the compound.

Analytical HPLC characterization was performed with a C12 Phenomenex Juniper Proteo column maintained at 40°C. The mobile phase consisted of 0.1% formic acid in water (A) and HPLC grade ACN (B). The applied gradient consisted of 5% B for 0.75 min, followed by a ramp to 60% B over 6.5 min at 2.0 mL/min. The retention was 3.17 min. (**Figure 2.8**)

Characterization result of PA 1131

PA-**1131,** desImImβImPyγImPyPyPyPyβTa $(HCOO^-)_6$

^{1}H-NMR (600 MHz, DMSO-d_6) δ= 10.74 (s, 1H), 10.37 (s, 1H), 9.97- 9.88 (t, J= 18.97 Hz, 3H), 8.37 (s, 9H), 8.22 (s, 1H), 8.04 (s, 1H), 7.96 (s, 1H), 7.67 (s, 1H), 7.57 (s, 1H), 7.49 (s, 1H), 7.44 (s, 1H), 7.43 (s, 1H), 7.41(s, 2H), 7.37 (s, 1H), 7.29 (s, 1H), 7.19 (s, 1H), 7.10 (s, 1H), 7.03 (s, 1H), 6.93 (s, 1H), 6.85 (s, 2H), 6.62 (s, 1H), 3.96 (s, 3H), 3.93- 3.91 (d, J= 8.17 Hz, 2H), 3.89 (s, 1H), 3.87 (s, 2H), 3.77 (s, 2H), 3.75 (s, 2H), 3.51 (s, 1H), 3.48 (br, 2H), 3.03 (s, 3H), 2.88 (s, 3H), 2.84-2.80 (t, J= 7.87 Hz, 2H), 2.77-2.76 (d, J= 4.92 Hz, 7H), 2.47 (s, 4H), 2.32- 2.30 (d, J= 7.87 Hz, 4H), 2.25 (br, 4H), 2.07 (s, 8H), 1.66-1.63 (q, J= 21.65 Hz, 8H), 1.51-1.47 (q, J= 23.62 Hz, 2H).

^{13}C-HNMR (151 MHz, DMSO-d_6) δ= 170.4, 169.9, 168.7, 165.6, 165.5, 165.5, 165.5, 165.5, 165.5, 165.5, 161.2, 158.3, 155.8, 137.5, 134.6, 134.2, 133.7, 127.6, 127.0, 126.9, 126.1, 115.4, 118.8, 117.6, 113.8, 113.9, 104.4, 103.3, 67.3, 54.5, 54.3, 53.9, 53.9, 53.9, 53.4, 41.4, 41.2, 41.3, 41.2, 40.0, 39.9, 39.8, 39.6, 39.5, 39.4, 39. 2, 39.1, 37.5, 37.4, 37.3, 37.1, 36.7, 36.6, 36.1, 35.9, 35.9, 35.6, 35.5, 35.1, 34.9, 34.8, 34.6, 32.9, 32.8, 32.7, 26.7, 24.6, 26.5, 24.6, 24.6, 24.5, 24.4.

HRMS (ESI) was calculated for **1131**, $[M+H]^+$ found 1460.73 m/z.

HPLC purity: 93%

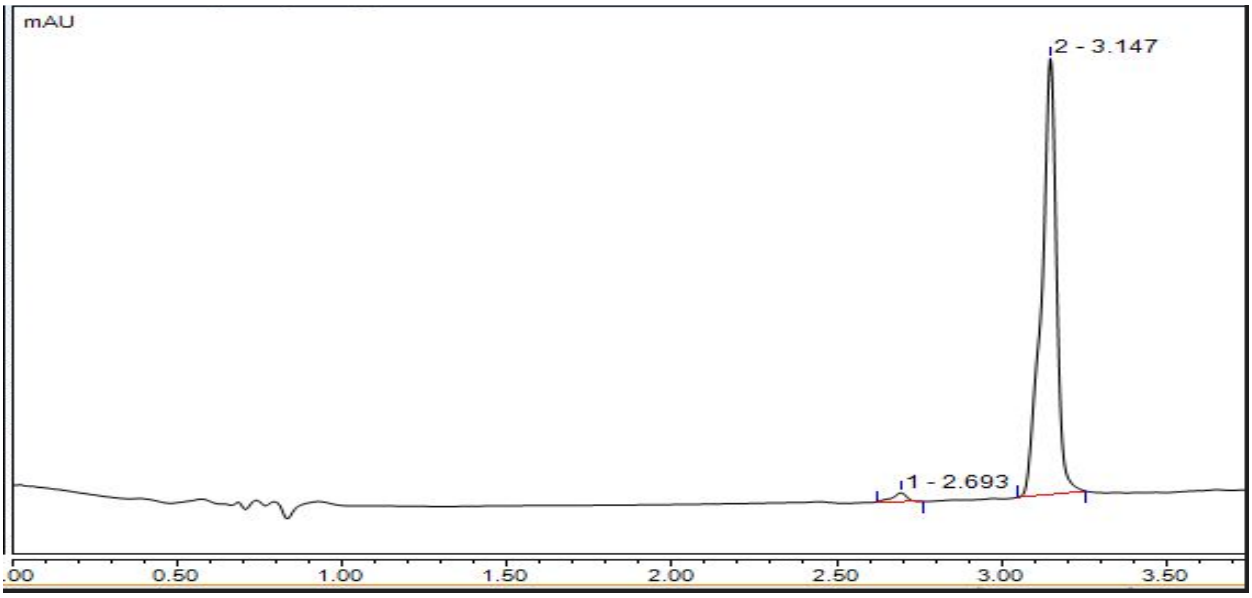


Figure 2.8 Analytical HPLC purity of compound **1131** with λ monitored. Compound purity 93%, retention time 3.147 min.

2.2.12 Synthesis of desIm-Im-Py-Im-β-γ-Im-Py-Py-Py-Py-β-Ta

Synthesis of 1132

Synthesis of polyamides **1132** resin started by initiating the reaction sequences program with the CSBio program. The CSBio reaction vessel was set at a temperature of 35°C throughout the whole process. Boc-β-alanine-PAM resin (1000 mg) was loaded into the RV. PyBOP (benzotriazole-1-yl-oxytripyrrolelidinophosphonium hexafluorophosphate (11 g) was dissolved in DMF (40 mL) and filled in the R4 solvent bottle.

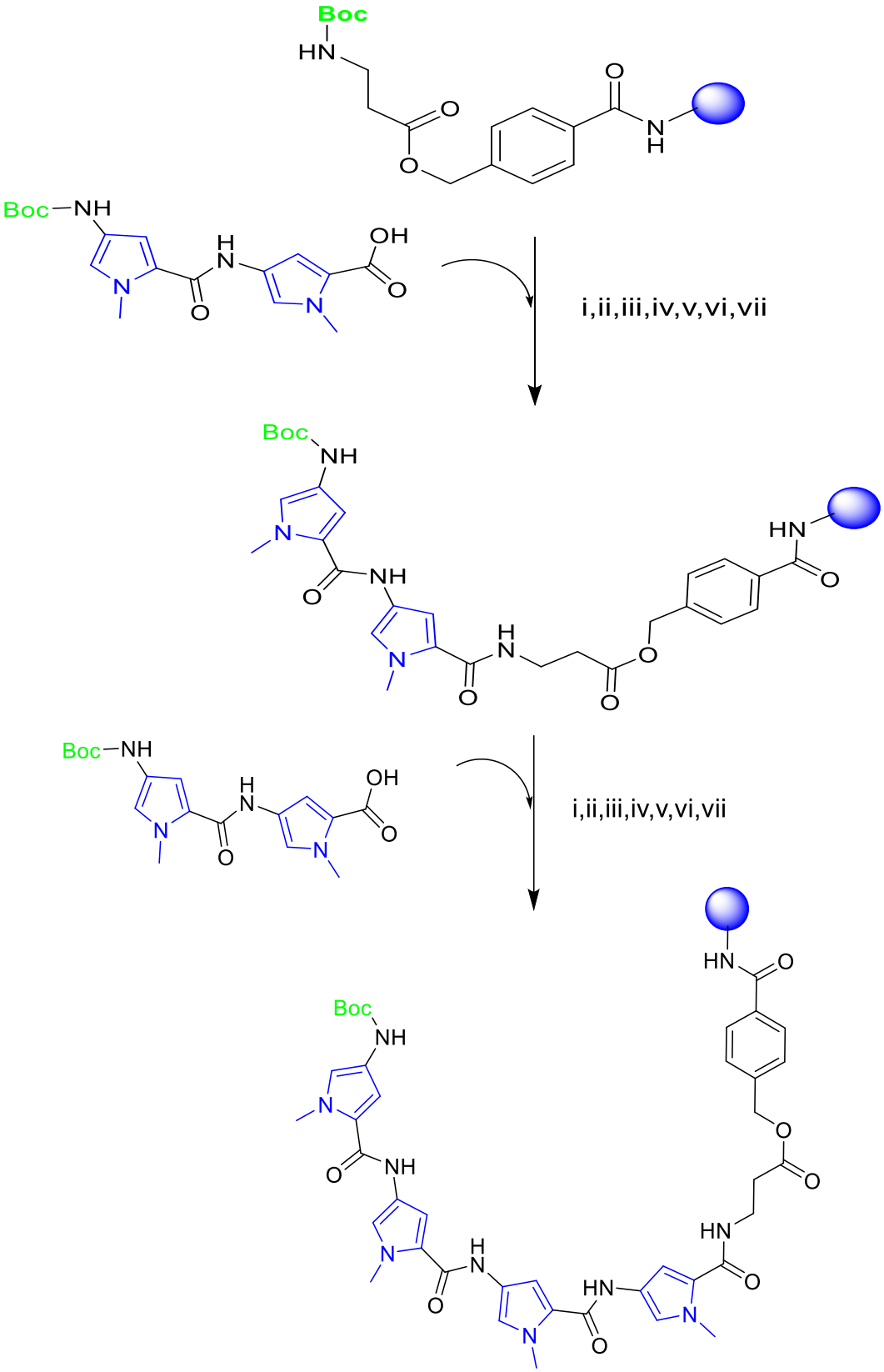
After the program started the first wash and deprotection reaction for 1h, a solution of Boc-Py-Py-COOH (502 mg, 1.38 mmol) in DMF (6.5 mL) was stirred to dissolve and transferred into AA#3 fully. The coupling program directed the synthesizer to transfer the solution from AA#3 to TVA and then MVA to begin the activation before transferring to RV for 180 min, followed by capping with acetic anhydride for 17 min to remove impurities such as unreacted amine. After completing the first sequence, the program proceeded to "cap" to remove any nonreacted amine before moving on to the next sequence. Each sequence took around 6 h to complete.

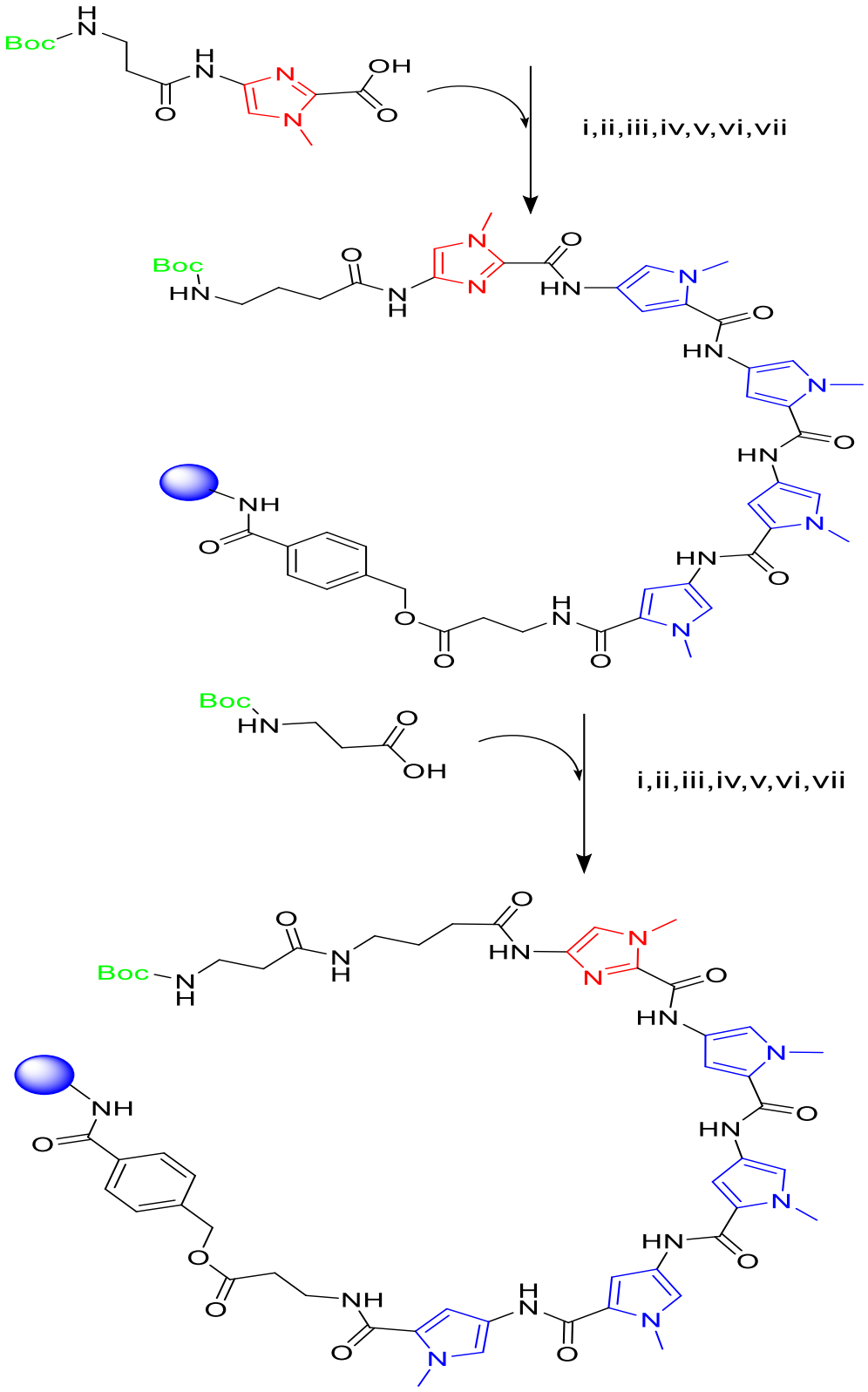
After completing the first sequence, the programmed system transferred a solution of Boc-Py-Py-COOH (499 mg, 1.38 mmol) in DMF (6.5 mL), was well dissolved by vortex and loaded in AA#4 to TVA and MVA for sequence coupling followed up by capping. After sequence 3 plus cap was completed, a solution of Boc-γ-Im-COOH (453 mg, 1.38 mmol) in DMF and DIPEA (1 mL) was well dissolved by vortex; loaded in AA#5 transferred to TVA then MVA before transferred to RV for coupling reaction for 180 min followed by capping with acetic anhydride for 17 min.

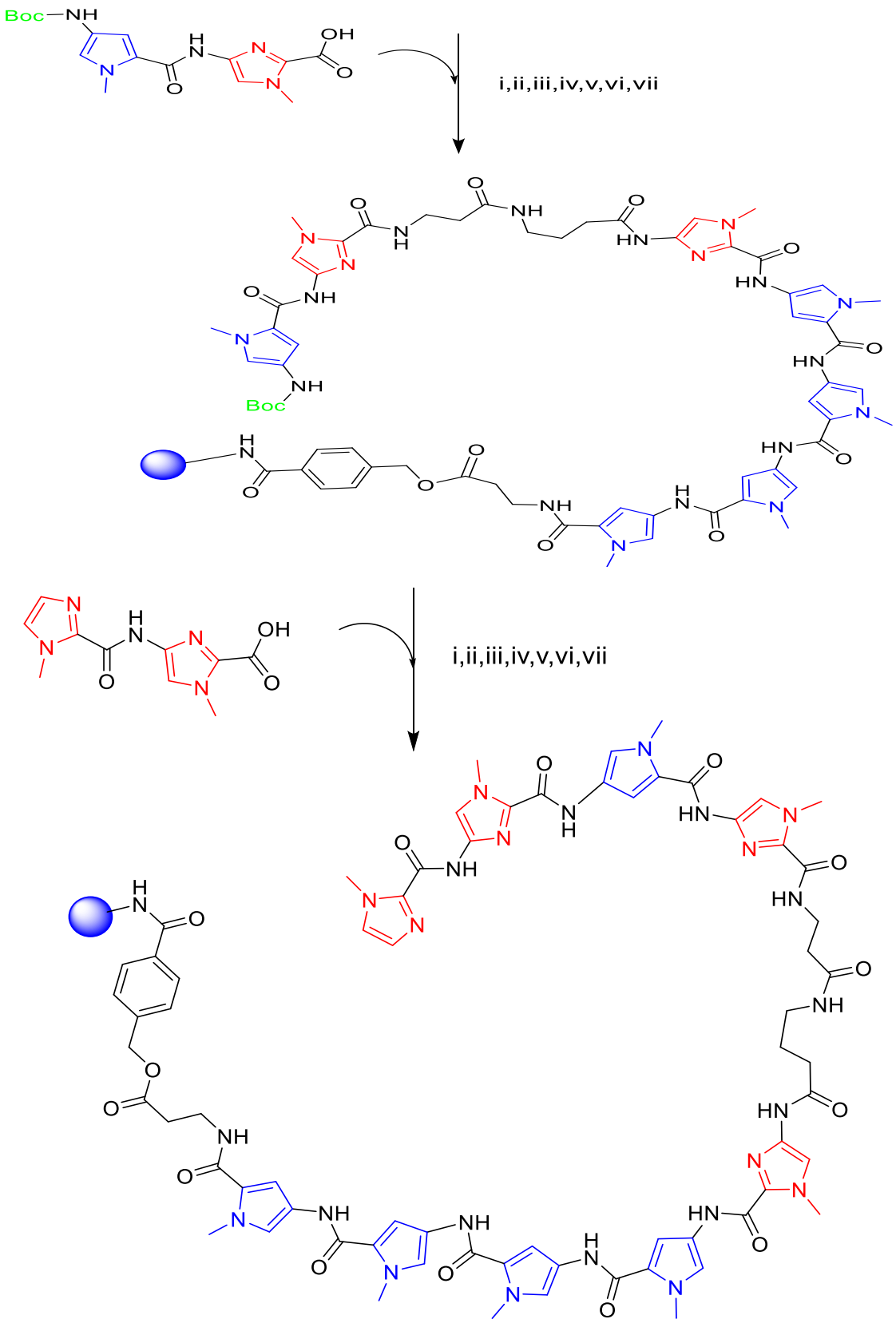
After the third sequence plus cap was completed, a solution of Boc-β-COOH (350 mg, 1.84 mmol) in DMF (6.5 mL) was well dissolved by vortex; loaded in AA#6, transferred to TVA then MVA before transferred to RV for coupling reaction for 180 min followed by capping with acetic anhydride for 17 min.

After the fourth sequence was finished, the programed system was initiated and followed by transferring a well dissolved by vortex solution of Boc-Py-Im-COOH (505 mg, 1.38 mmol) in DMF (6.5 mL) and DIPEA (1mL) from AA#6 to TVA then MVA before transferred to RV for coupling reaction for 180 min following by capping with acetic anhydride for 17 min. After the fifth sequence finished, the programed system initiated and followed by transferring a well-dissolved by vortex solution of dimer desIm-Im-COOH (345 mg, 1.38 mmol) and DIPEA (1ml) in DMF (6.5 mL) from AA#7 to TVA then MVA before transferred to RV for coupling reaction for 180 min following by capping with acetic anhydride for 17 min. (**Scheme 2.16**)

The sequence was finished after 30 h. 0,1 mg of resin was cleaved with 100 μL Ta solution for 2h at 45^{o}C, and the heated solution was collected and filtered before being analyzed with a Bruker LC-MS instrument. The exact mass was determined to be $[M+H]^+$ as 1460.73 m/z, which confirmed the presence of the desired product. All the resin was carefully transferred from the RV into a beaker using a small spatula, and the resin was washed three-time with DCM and MeOH, then dried under vacuum for 30 min until reaching dryness. The resin was then collected in a 20mL scintillation vial and stored in the freezer before cleavage the next day.


Boc
HN
O
O
O
N
H
Boc—NH
H
N
OH
O
N
O
N
i,ii,iii,iv,v,vi,vii
Boc
NH
N
HN
O
NH
O
N
H
N
O
O
O
Boc—NH
H
N
OH
O
N
O
N
i,ii,iii,iv,v,vi,vii
HN
O
Boc
NH
N
O
NH
O
N
H
N
O
HN
O
N
HN
O
O
O


Boc
OH
i,ii,iii,iv,v,vi,vii
Boc
NH
Boc
OH
i,ii,iii,iv,v,vi,vii
Boc
NH


Boc
NH
OH
i,ii,iii,iv,v,vi,vii
Boc
i,ii,iii,iv,v,vi,vii

Scheme 2.16 *Scheme of synthesis polyamide **1132** applying a solid-phase synthesis method. It started from a commercially acquired Boc-β-PAM resin, and each sequence contains multiple steps from (i) to (viii). (i) wash the resin with DCM and DMF, (ii) deprotection with 60% TFA/ DCM/ 0.5 M Indole, (iii) 25% piperidine/DMF; (iv) building blocks with the structures shown on the scheme, (v) 0.5M PyBOP/DMF as coupling reagents, (vi) DIPEA, (vii) DMSO, (viii) capping step with acetic anhydride for 17 min.*

Polyamide **1132** resin was synthesized with 12 consecutive programs. The total coupling time was 30 h (**Table 2.6**).

Table 2.6. Programmed sequence of **1132** resin synthesis.

Building block(s) and step(s)	Additional Solvent mixture	Applied program
1-Boc-Py-Py-COOH	(6.5 mL) DMF	Add-DIPEA-180-min-DIPEA-Flows.CFN
2-Capping	None	Cap-17-min-continues.CFN
3-Boc-Py-Py-COOH	(6.5 mL) DMF	Add-DIPEA-180-min-DIPEA-Flows.CFN
4-Capping	None	Cap-17-min-continues.CFN
5-Boc-γ-Im-COOH	(6.5 mL) DMF+DIPEA (1mL)	Add-DIPEA-180-min-DIPEA-Flows.CFN
6-Capping	None	Cap-17-min-continues.CFN
7-Boc-β-COOH	(6.5 mL) DMF	Add-DIPEA-180-min-DIPEA-Flows.CFN
8-Capping	None	Cap-17-min-continues.CFN
9-Boc Py-Im-COOH	(6.5 mL) DMF+DIPEA (1mL)	Add-DIPEA-180-min-DIPEA-Flows.CFN
10-Capping	None	Cap-17-min-continues.CFN
11-desIm-Im-COOH	(6.5 mL) DMF+DIPEA (1mL)	Add-DIPEA-180-min-DIPEA-Flows.CFN
12-Capping	None	Cap-17-min-continues.CFN

Cleavage of PAM resin using Ta as a nucleophile.

The next day, 334 mg resin was cleaved with aminolysis in 2000 µL of Ta at 45ºC in a heating device Isotemp Fisher Scientific for 4 h, and the separation of polyamides solution from the PAM resin support was performed in a 20 µM polyethylene frit placed inside a disposable polypropylene syringe by washing with MeOH and H_2O. The polyamide solution was concentrated in the rotor evaporator before purification with preparative HPLC to afford the final mass of **1132** (26 mg, 0.768 mmol, 23% yield) **(Scheme 2.17)**

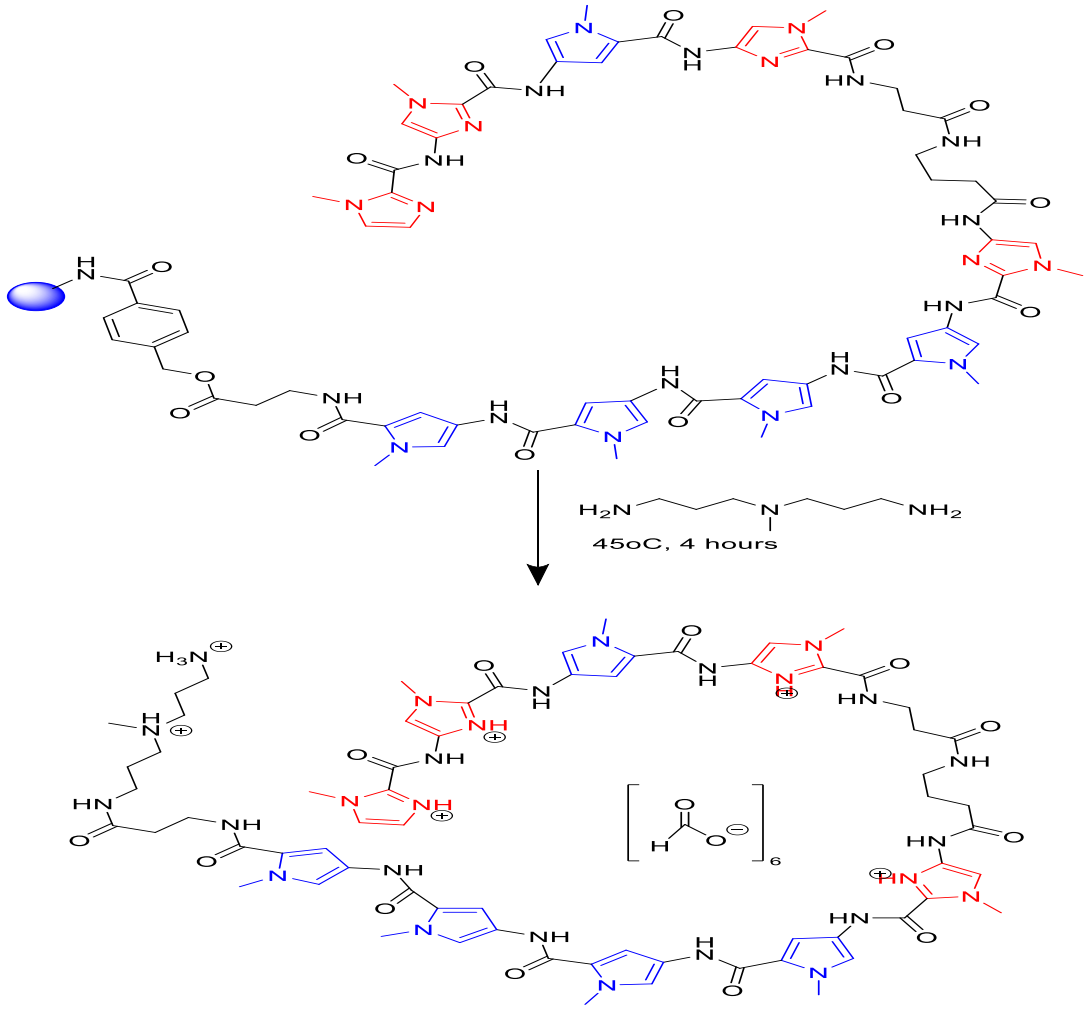


Figure 2.17 Cleavage of PA**1132** from PAM resin using Ta as the nucleophile**.**

Purification of compound

The polyamides solution was diluted with a mixture of DMSO (200 µL) and H_2O/0.2% HCOOH (200 µL), then purified by preparative HPLC using a Phenomenex Luna 250x30 mm, 5 µM, 100 Å C18 column maintained at 25ºC. The organic phase was 0.2% HCOOH in MilliQ water

and 100% HPLC grade MeOH. The applied gradient was 10% MeOH for 8 min, followed by a ramp rate to 90% Methanol over 35.6 min at a 20 mL/min flow rate. All collected fractions were analyzed and selected for 95% or higher purity with analytical HPLC using ThermoFisher - UHPLC; analysis was processed and analyzed with the program Chromeleon 3.0. High-purity fractions were collected for rotary evaporation to remove MeOH before the lyophilization.

Lyophilization of compound

Compound **1132** was diluted with a 1:1 mixture of ACN (5 mL) and 0.2% HCOOH in H_2O (5 mL). The samples were frozen before the lyophilization process with ground dry ice until an even coated of frozen material results at an acute angle in the vial. The frozen sample was covered with porous Kimwipe paper and secured by a rubber band. The sample was settled in a lyophilizer glass vessel and connected to a vacuum port. The frozen polyamide solution was sublimed as an off-white solid powder after 48 h. The lyophilizer vacuum was set at 300 torr within 30 min prior to the lyophilization process, and the inside temperature was set at -83 to -85^{o}C

Characterization of the compound.

Analytical HPLC characterization was performed with a C12 Phenomenex Juniper Proteo column maintained at 40^{o}C. The mobile phase consisted of 0.1% formic acid in water (A) and HPLC grade ACN (B). The applied gradient consisted of 5%B for 0.75 min, followed by a ramp to 60% B over 6.5 min at 2.0 mL/min. The retention was 3.030 min.

PA1132, desImImPyImβγImPyPyPyPyβTa $(HCOO^-)_6$

^{1}H-NMR (600 MHz, DMSO-d_6) δ= 10.38 (s, 1H),10.28 (s, 1H), 9.98 (s, 2H), 9.04 (s, 1H), 8.350 (9H), 8.01 (s, 2H), 7.56 (s, 1H),7.49 (s,1H), 7.428 (s, 1H), 7.37 (s, 1H), 7.25 (s, 1H), 7.23 (s, 1H), 7.18 (s, 1H), 7.15 (s, 1H), 6.83 (s, 1H), 4.03-4.01 (q, J= 20.6 Hz, 3H), 3.98 (s, 2H), 3.97 (s, 1H), 3.92 (s, 1H), 3.90 (s, 1H), 3.83 (s,1H), 3.78 (s, 1H), 3.71 (s, 1H), 3.55 (s, 1H), 3.44- 3.41 (q, J= 6.76 Hz, 2H), 3.35- 3.32 (q, J= 19.7 Hz, 2H), 3.04 (s, 2H), 2.77 (s, 4H), 2.73 (s, 2H), 2.63 (s, 2H), 2.48 (s, 7H), 2.26- 2.23 (q, J=20.8 Hz, 3H), 2.19 (s, 1H), 2.09 (s, 12H), 1.67 (s, 4H), 1.42 (s, 5H), 1.16 (s, 4H), 1.21 (s, 1H).

^{13}C-NMR (151 MHz, DMSO-d_6) δ=171.4, 170.6, 170.5, 170.5, 169.9, 165.4, 161.3, 161.2, 158.7, 158.5, 158.5, 158.5, 158.4, 155.8, 155.6, 152.8, 137.8, 136.2, 136.1, 137.7, 134.4, 133.7, 127.7, 127.7, 127.6, 127.1, 127.0, 127.0, 126.9, 126.9, 126.1, 123.1, 122.8, 122.8, 122.3, 122.2, 122.2, 122.1, 121.3, 121.2, 119.5, 118.5, 118.4, 117.9, 116.5, 114.3, 114.0, 113.9, 113.9, 113.9, 105.8,

104.6, 104.8, 104.7, 104.2, 104.2, 103.4, 59.6, 54.5, 54.3, 53.9, 41.3, 41.2, 39.9, 39.8, 39.7, 39.5, 39.4, 39.2, 39.0, 37.2, 35.9, 35.1, 34.9, 34.9, 34.8, 34.0, 28.2, 25.3, 24.5, 14.1.

HRMS (ESI) was calculated for **1132**, $C_{67}H_{85}N_{27}O_{12}$, $[M+H]^+$,1460.70 m/z, found 1460.73 m/z

HPLC purity:92%

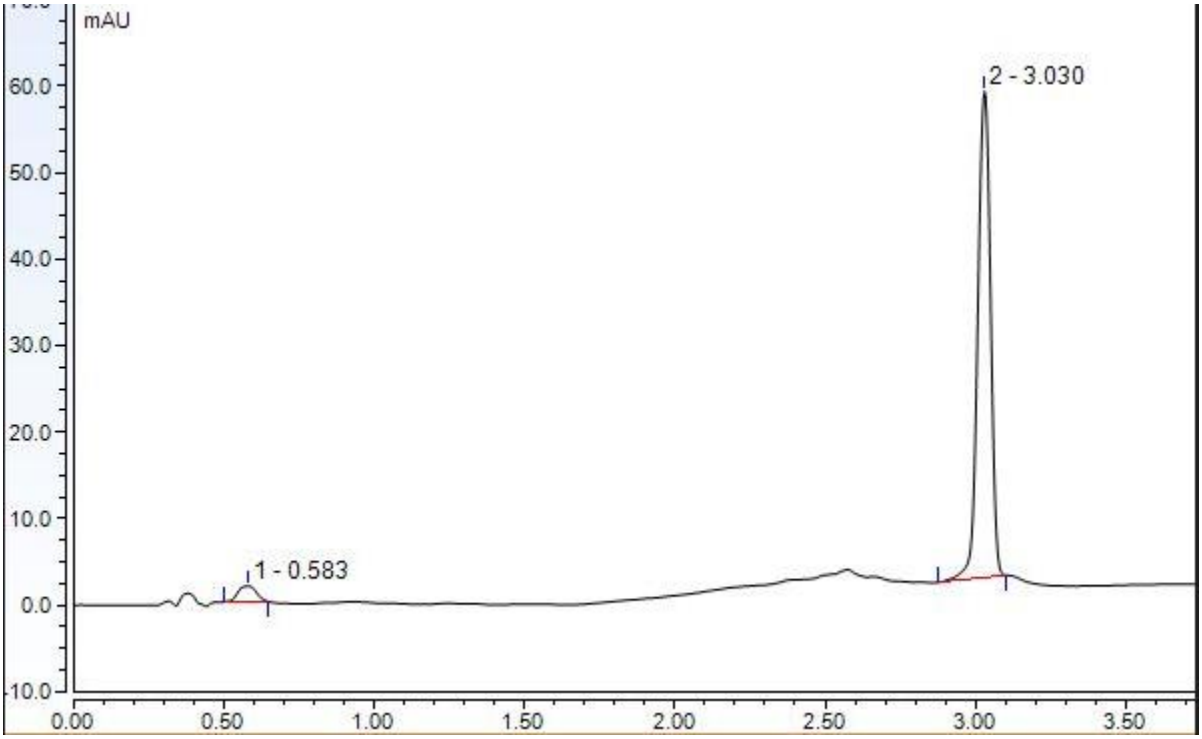


Figure 2.9 analytical HPLC purity of compound **1132** with λ monitored. Compound purity 92%, retention time 3.030 min.

2.3 Results.

As described in this chapter, all polyamides were synthesized with the solid-phase peptide synthesis method. Some building blocks include Boc-Py-Obt, Boc-Py-Im-COOH, NO_2-Im-COOMe, Boc-γ-COOH, Boc-β-COOH, Boc-Py-Py-COOH are commercially available and inexpensive, so most of them can be utilized directly for coupling synthesis. The other building blocks, such as desIm-Im-COOH, Boc-γ-Im-COOH, and Boc-β-Im-COOH, are expensive and require careful synthesis and workup to afford high purity without using column chromatography.

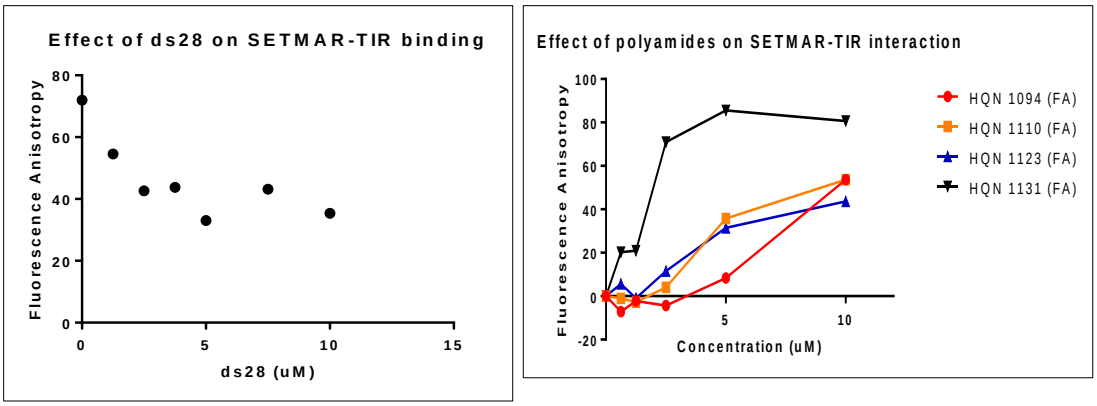


Figure 2.10 *The result of the Fluorescence Anisotropy study on SETMAR-TIR interaction. A) the effect of SETMAR-TIR binding at various concentrations of unlabeled TIR-DNA. B) The effect of increasing concentration of Polyamides on SETMAR-TIR interactions*.

Our collaborator performed the following study of competitive binding using Fluorescence Anisotropy (FA) to study the effect of Polyamides on SETMAR-TIR binding. We sent compounds **1094**, **1110**, **1123**, and **1131** to test for their effects. As shown in **Figure 2.10** B, at lower concentrations, PAs appeared to increase the binding of SETMAR to the labeled TIR-DNA.

The presence of **1131** appeared to increase the FA values more significantly than the other PAs. It was unclear whether the presence of the polyamides increased the binding effects of SETMAR-TIR or if the binding of polyamides caused the binding effect to be increased. Our collaborator also showed us the original values of the competitive binding study before subtracting the signal of the labeled TIR-DNA (**Figure 2.11)**, and it gave us a better look at the effect of polyamides on SETMAR-TIR binding.

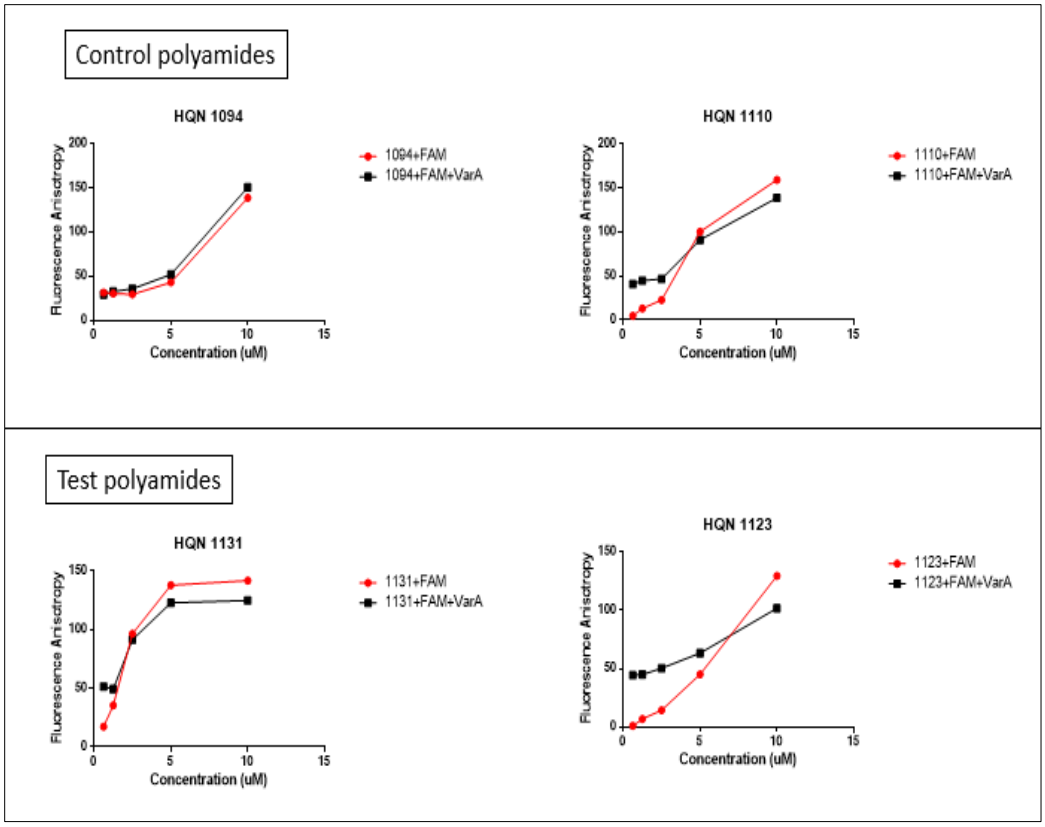


Figure 2.11: *The effect of polyamides with the variant of SETMAR(VarA) and labeled DNA(FAM).*

Compared to the result between a single mismatch PA**1094** to the test polyamide **1131**, we observed that the addition of a SETMAR variant (VarA) at a higher concentration, PA131 had no change to the response signal of FA while there is a rapid increase with the mismatch **PA1094**. The more rigid polyamides like **1110** and **1123** have a similar effect in increasing the binding of SETMAR-TIR. The presence of continuous cyclic rings appeared to cause a rapid increment in FA responses. Our collaborator also tests the binding effects of polyamides at much higher concentrations to seek more information. **(Figure 2.12**)

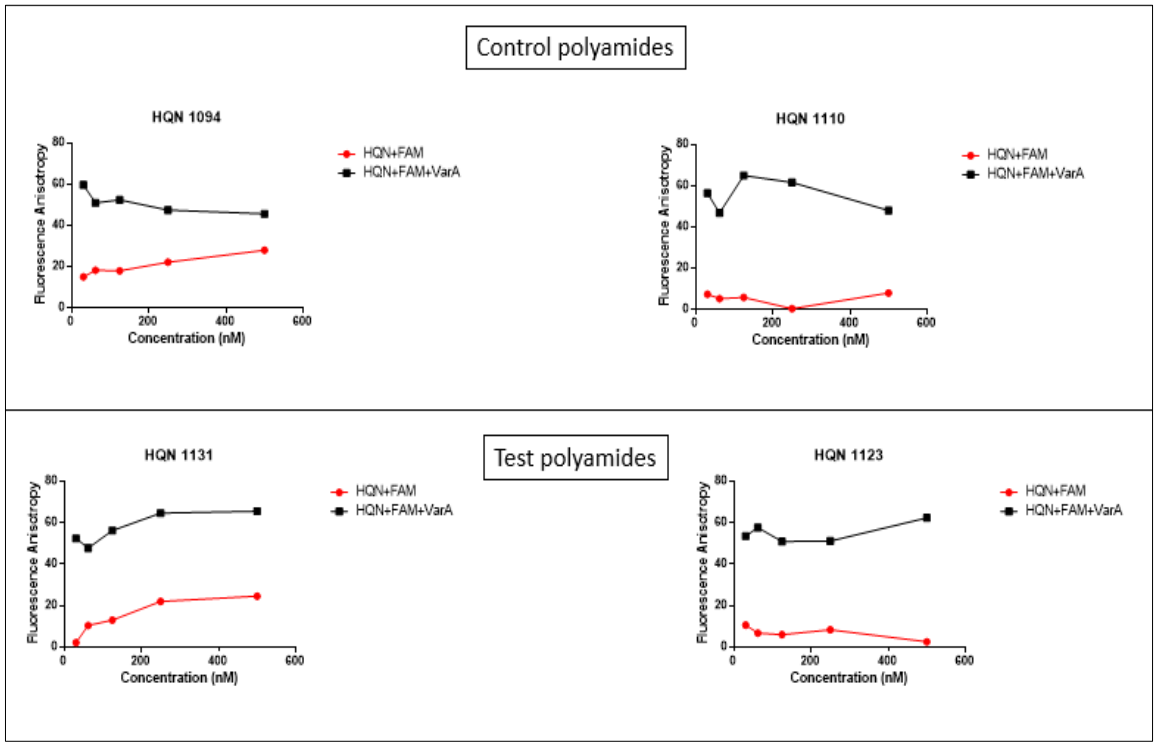


Figure 2.12 *The effects of polyamides toward SETMAR-TIR DNA binding at higher Polyamides concentration.*

At higher concentrations (**Figure 2.12**), we were able to confirm that the presence of polyamides alone can increase the FA signals, while at lower concentrations (**Figure 2.11)**, we do not see much effect on SETMAR-TIR binding. At higher concentrations, the effect of PA **1131** does not seem to change the FA signals. The mismatched polyamides **1094** and **1110** slightly decreased the FA signals at very high concentrations. This showed us that the presence of our polyamides alone could increase the SETMAR-TIR binding instead of decreasing it, and at the higher concentration, there is not much difference in the effects of polyamides on SETMAR-TIR binding.

We reached out to the assistance of our second collaborator, Dr. Wilson, and his researchers to study the binding affinity of our polyamides on the SETMAR-TIR binding. They applied Surface Plasmon Resonance (SPR) to allow us to study our polyamides' binding affinity, which should give us some hints about our research approach. At first, they discovered that they could not read any signal with PA**1123**. The rigid structure of the **1123** aggregate appeared to result in poor signal from the SPR. However, we received some new discoveries after another SPR study with PA**1131** and its mismatch counterpart **1094**. (**Figure 2.13**)

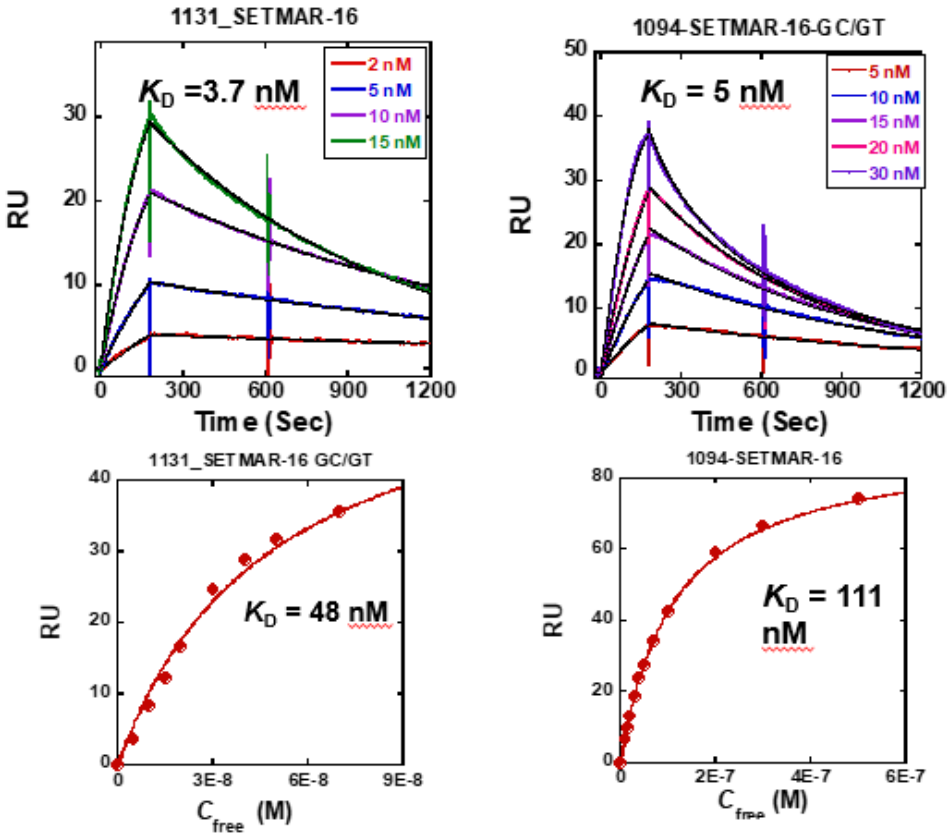


Figure 2.13: *Binding affinity study of PA131 and PA**1094***.

As shown in **Figure 2.13**, both polyamides were shown to have very good results. The polyamides with a single β-alanine substitution in their structures appear to keep them from aggregation due to the continuous heterocyclic rings. PA**1131** and PA**1094** were shown to have a good binding affinity (**Figure 2.13** top graphs). We observed rapid associations and a decent dissociation rate from both polyamide binding. It is noticeable that the matched sequence polyamide **1131** has a better binding affinity than its mismatched counterpart PA**1094.**

Additionally, the bottom figures showed that the equilibrium analysis of PA**1131** shows that the graph has not reached equilibrium, and this indicates that polyamide **1131** binds tightly to the TIR of DNA, and the complex has not been completely released. Future compounds can be designed based on the initial results. The additional study of PA **1132** binding affinity provided us with further information about the PA: TIR binding; despite having some aggregation in binding with either a short sequence of SETMAR or a more extended sequence, **1132** seems to have a similar binding affinity with them. However, the compound showed a much stronger binding affinity to a mismatch sequence (**Figure 2.14**), showing that the linkage β-alanine might be

responsible for the compound's affinity. This information can provide us to modify and study more about the binding of SETMAR TIR.

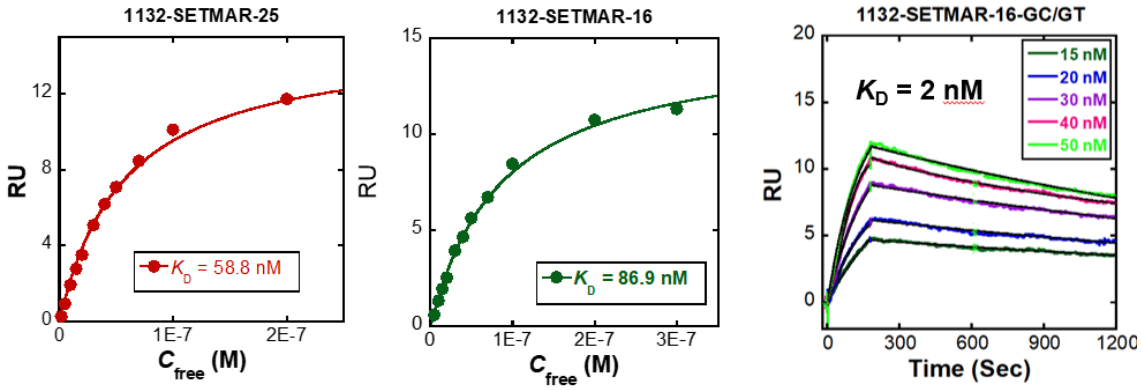


Figure 2.14 SPR study results of PA**1132.**

2.4 Conclusion

In summary, we designed and synthesized polyamides sequence to recognize and bind to the conserved target TIR DNA of the SETMAR gene. We used the solid phase peptide synthesis method to successfully synthesize the polyamides by linking them with amide bonds starting with a Boc-β-alanine-PAM resin.[72]

The result from DNA competitive binding showed us that the presence of polyamides significantly increases the binding of SETMAR-TIR DNA. At higher concentrations, that effect was not that much different than at lower concentrations. The results from SPR to study for binding affinity showed that sequences with rigid structures like **1110** and **1123** aggregated and heavily affected the binding affinity of the polyamides. However, a single β-alanine substitution and the location of β-alanine within the sequence can relax the sequence and increase the binding affinity.[68,73]

Therefore, substituting more than one β-alanine in the polyamide sequences and introducing cationic charges to the N-terminus of the designed sequence is needed to investigate more about SETMAR-Polyamide binding capabilities. In addition, new sequences should have a more flexible structure in order to avoid the rigidity of the sequence.

2.5 Bibliography

1. Kopka ML, Yoon C, Goodsell D, Pjura P, Dickerson RE. Binding of an Antitumour Drug to DNA Netropsin and CGCGAATTCGCG. Published online 1985.

2. Kopka ML, Yoon C, Goodsell D, Pjura P, Dickerson RE. The molecular origin of DNA-drug specificity in netropsin and distamycin. *Proceedings of the National Academy of Sciences of the United States of America*. 1985;82(5):1376-1.38. doi:10.1073/pnas.82.5.1376

3. Alniss HY. Thermodynamics of DNA Minor Groove Binders. *Journal of Medicinal Chemistry*. 2019;62(2):385-402. doi:10.1021/acs.jmedchem.8b00233

4. Peter B. Dervan. Molecular Recognition of DNA by small molecule. 2001;9(July):2-57.

5. Kawamoto Y, Bando T, Sugiyama H. Sequence-specific DNA binding Pyrrole–imidazole polyamides and their applications. *Bioorganic and Medicinal Chemistry*. 2018;26(8):1393-1411. doi:10.1016/j.bmc.2018.01.026

6. Swalley SE, Baird EE, Dervan PB. Effects of γ-turn and β-tail amino acids on sequence-specific recognition of DNA by hairpin polyamides. *Journal of the American Chemical Society*. 1999;121(6):1113-1120. doi:10.1021/ja9830905

7. Liu B, Wang S, Aston K, et al. β-Alanine and N-terminal cationic substituents affect polyamide-DNA binding. *Organic and Biomolecular Chemistry*. 2017;15(46):9880-9888. doi:10.1039/c7ob02513k

8. Chen Q, Bates AM, Hanquier JN, et al. Structural and genome-wide analyses suggest that transposon-derived protein SETMAR alters transcription and splicing. *Journal of Biological Chemistry*. 2022;298(5):101894. doi:10.1016/j.jbc.2022.101894

9. Cordaux R, Udit S, Batzer MA, Feschotte C. Birth of a chimeric primate gene by capture of the transposase gene from a mobile element. *Proceedings of the National Academy of Sciences of the United States of America*. 2006;103(21):8101-8106. doi:10.1073/pnas.0601161103

10. Tellier M, Chalmers R. The roles of the human SETMAR (Metnase) protein in illegitimate DNA recombination and non-homologous end joining repair. *DNA Repair*. 2019;80(March):26-35. doi:10.1016/j.dnarep.2019.06.006

11. Dussaussois-Montagne A, Jaillet J, Babin L, et al. SETMAR isoforms in glioblastoma: A matter of protein stability. *Oncotarget*. 2017;8(6):9835-9848. doi:10.18632/oncotarget.14218

12. Chen Q, Georgiadis M. Crystallization of and selenomethionine phasing strategy for a SETMAR-DNA complex. *Acta Crystallographica Section:F Structural Biology Communications*. 2016;72:713-719. doi:10.1107/S2053230X16012723

13. Swalley SE, Baird EE, Dervan PB. Effects of γ-turn and β-tail amino acids on sequence-specific recognition of DNA by hairpin polyamides. *Journal of the American Chemical Society*. 1999;121(6):1113-1120. doi:10.1021/ja9830905

14. Koeller KJ. DNA Binding Polyamides and the Importance of DNA Recognition in their use as Gene-Specific and Antiviral Agents. *Medicinal Chemistry*. 2014;04(02):338-344. doi:10.4172/2161-0444.1000162

15. Wang S, Aston K, Koeller KJ, et al. Modulation of DNA-polyamide interaction by β-alanine substitutions: A study of positional effects on binding affinity, kinetics and thermodynamics. *Organic and Biomolecular Chemistry*. 2014;12(38):7523-7536. doi:10.1039/c4ob01456a

16. Jaramillo D, Liu Q, Aldrich-Wright J, Tor Y. Synthesis of N-Methylpyrrole and N-Methylimidazole Amino Acids Suitable for Solid-Phase Synthesis. *The Journal of Organic Chemistry*. 2004;69(23):8151-8153. doi:10.1021/jo048686r

17. Baird EE, Dervan PB. Solid Phase Synthesis of Polyamides Containing Imidazole and Pyrrole Amino Acids. 1996;7863(9):6141-6146.

18. Rucker VC, Melander C, Dervan PB. Influence of beta -Alanine on Hairpin Polyamide Orientation in the DNA Minor Groove. 2003;86.

Chapter 3: Substitution of β-Alanine to study and improve SETMAR DNA binding affinity.

3.1 Introduction

Pyrrole-Imidazole Polyamides (PIPs) are nanomolecular, and they are well known for their feature of binding to the specific minor groove of DNA. PIPs also permeate through the cell and be taken up by living cells.[1,74] PIPs consist of N-methylpyrrole (Py) polymers and N-methylimidazole (Im) linking through various amide linkages. They are analogs of the naturally derived compounds netropsin and distamycin A, known for binding to the minor groove of A/T-rich DNA.[2,19,75]

Polyamides possess the features allowing them to recognize and form hydrogen bonding with DNA base pairs within a double strand. Dervan and his coworkers reported the polyamide pairing rules where Py/Py recognized A/T or T/A base pairs. Additionally, Py/β and β/Py can recognize A/T and T/A as well. While Py/Im pair distinguishes C/G from G/C. Additionally, it's been reported that substituting β-alanine linkage can reduce the rigidity and increase the binding affinity of Polyamide-DNA binding. Hairpin polyamides have multiple potentials in biomedical applications since they can recognize the DNA sequence and bind to two opposite strands of DNA at the minor groove.[1,21,72,73]

SETMAR is a chimeric gene created by the fusion of the SET methyltransferase domain and the Homosapien mariner one transposase (Hsmar1).[45,61] SETMAR gene is located on chromosome 3p26.1, and it is reportedly linked to a variety of diseases.[60] Although its mechanism is still unclear and SETMAR no longer have active transpose functionality, many reports discovered the presence of SETMAR mRNA in various type of cancer such as glioblastoma, gliomas, colorectal cancer, non-Hodgkin's lymphoma, and acute leukemia, etc.[54,62]

The TIRs(Terminal Inverted Repeats) sequence of Hsmar1 has been conserved since 40 million years ago in the evolution history of primates, and the discovery of SETMAR-TIR binding location by Georgiadis and her colleagues.[65,66] It's been shown that the potential of introducing a DNA minor groove compound binding to that specific SETMAR-TIR binding location in the minor groove can investigate and suppress the overexpression of the SETMAR gene to study and prevent cancer development.

Most modern PIPs are prepared with a solid-phase peptide synthesizer; this technique allows researchers to synthesize polyamide from a short chain of 5 to 6 building blocks to a more extensive sequence of possibly longer than 20 building blocks. It was first developed by the Nobel Prize winner Bruce Merrifield in the early 1960s, with the proposal of the use of polystyrene-base solid support to synthesize peptides from the C-terminus to the N-terminus direction.[76] Boc-β-PAM resin (**Figure 3.1**) is one of the most frequent methods we utilized in our group to design and

synthesize polyamides due to its accessibility and a milder condition in Boc deprotection methodology. In addition, PAM resin can be readily cleaved by using either Dp or Ta as nucleophiles for the completion of polyamide synthesis.[70,77,78]

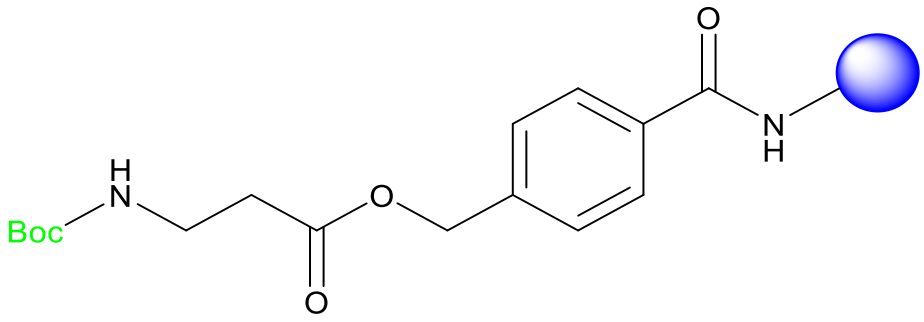


Figure 3.1 Structure of Boc-β-Alanine PAM resin.

In this chapter, we substitute pyrrole for β-alanine in the polyamide sequences. Our goal is to modify the polyamide sequence to increase the binding affinity and decrease the rigidity of the sequence by substituting pyrrole for β-alanine at some specific locations, as shown in **Table 3.1.** We will have a β-alanine after two or three heterocyclic rings of imidazole or pyrrole. All the sequences are synthesized starting with a β-alanine-PAM resin and coupling the building blocks linking by amide bonds.

Table 3.1 Polyamides designed to investigate improving the binding affinity of polyamides.

Compound	N-terminus	1	2	3	4	5	6	7	8	9	10	11	C-Terminus
1134	desIm	Im	Py	Im	Py	γ	Im	β	Py	Py	Py	β	Ta
1161	desIm	Im	β	Im	Py	γ	Im	Py	β	Py	Py	β	Ta
1162	desIm	Im	β	Im	Py	γ	Im	β	Py	Py	Py	β	Ta
1166	desIm	Im	β	Im	Py	γ	Im	Py	Py	β	Py	β	Ta
1171	desIm	Im	β	Im	β	γ	Im	Py	β	Py	Py	β	Ta

3.2 Methods

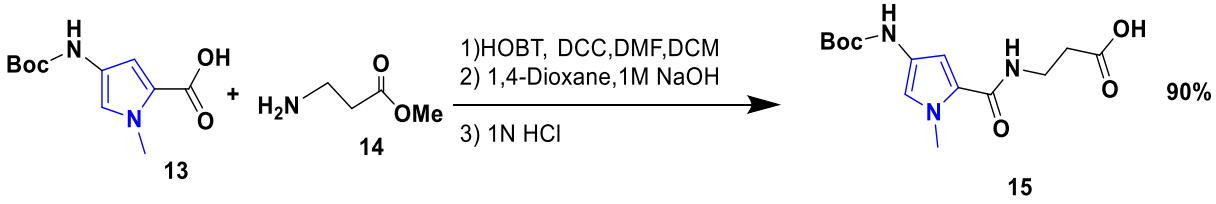
3.2.1 Material and Instrumentation

Boc- β-alanine-PAM resin was ordered from Peptides International (Louisville, KY), 3-(dimethylamino)-1-propylamine, N-methyl-imidazole, N-methyl-pyrrole, piperidine, indole, ammonium formate, 20% Pd/C, dimethyl sulfoxide, HCl, and NaOH were purchased from Sigma Aldrich (St. Louis, MO). 1-methyl-4-nitro-imidazole-2-carboxylate (NO_2-Im-COOMe) was ordered

from AC Pharmachem, Inc (Worcester, MA). 4-tert-butoxycarbonylaminobutyric acid (Boc-γ-COOH), 1-ethyl-4-amino-β-alanine-carboxylate (H_2N-β-COOEt), 4-tert-butoxy-β-alanine-carboxylic acid (Boc-β-COOH), trifluoroacetic acid (TFA), formic acid (FA), hexafluorophosphate azabenzotriazole tetramethyl uronium (HATU), benzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate (PyBOP), hydroxybenzotriazole (HOBT), dimethyl formamide (DMF), and methanol (MeOH) was acquired from Oakwood Chemical. Triethylamine, N, N-diisopropylethylamine (DIPEA), formic acid, and acetic anhydride were purchased from Fisher Scientific.

^{1}H-NMR and ^{13}C-NMR spectra were recorded in DMSO-d_6 at 300mHz, and 600MHz respectively. ^{1}H-NMR is referenced to DMSO at 2.5 ppm, and ^{13}C-NMR in DMSO at 39.52 ppm. Coupling constant J is reported in Hz. The PIPs are synthesized using the automated solid-phase peptide synthesizer device CS336S/CS336X. High-resolution mass spectrometry was carried out on a Bruker maXis II TOF-MS spectrometer. HPLC was carried out in the ThermoFisher U-HPLC, and the data were processed with the program Chromeleon7. Purifications of Polyamides were carried out in ThermoFisher preparative HPLC.

3.2.2 Synthesis of dimer building block Boc-Py-β-COOH (15, Scheme 3.1)



Scheme 3.1 Synthesis of Boc-Py-β-COOH **15.**

In an RBF, Boc-Py-COOH **13 (**5.0 g, 20.78 mmol) was dissolved in DMF (104 mL), followed by HOBT (4.38 g, 31.7mmol) and DCC (6.54 g, 31.7 mmol) in DCM (66 mL), stirred well under inert nitrogen gas for 20 min. After 30 min at room temperature, the mixture formed DCU and turned white solution, let the reaction ran overnight at room temperature. After 12 h, add H_2N-β-COOMe (3.27 g, 31.7 mmol) in DCM (20mL), and the reaction was stirred well under Nitrogen gas. The reaction's completion was confirmed by TLC and HPLC, and the reaction was completed after 18 h at room temperature. The mixture was filtered to remove the solid DCU; the mixture was washed with DCM and followed by solvent removal by vacuum evaporation to afford an oily solution. The oil mixture was immediately dissolved in 2,4-dioxane and stirred under N_2 for 30 min before adding 1N NaOH. The reaction was heated in an oil bath at a controlled temperature of 45^{o}C for 1 to 2 h, and the reaction was checked for completion with analytical HPLC. The reaction was removed from heat and settled at room temperature for 10 min, and 1 N HCl was added dropwise

to reduce the solution pH to around 4.35 to afford off-white to off-gray solid **15 (**5.85 g, 20.78 mmol, 90% yield).

3.2.3 Synthesis of desIm-Im-Py-Im-Py-γ-Im-β-Py-Py-Py-β-Ta (1134)

Synthesis of 1134 resin

Synthesis of polyamides **1134** resin started by initiating the reaction sequences program with the CSBio program. The CSBio reaction vessel was set at a temperature of 35°C throughout the whole process. Boc-β-alanine-PAM resin (1000 mg) was loaded into the RV. PyBOP (11 g) was dissolved in DMF (40 mL) and filled in the R4 solvent bottle.

After the program started the first wash and deprotection reaction for 1 h, a solution of dimer Boc-Py-Py-COOH (503 mg, 1.38 mmol) in DMF (6.5 mL) was stirred to fully dissolved and transferred into AA#3. The coupling program directed the synthesizer to transfer the solution from AA#3 to TVA and then MVA to begin the activation before transferring to RV for coupling for 180 min, followed by capping with acetic anhydride for 17 min to remove impurities such as an unreacted amine. After completing the first sequence, the program proceeded to "cap" to remove any nonreacted amine before moving on to the next sequence. Each sequence took around 6 h to complete.

After completing the first sequence, the programmed system transferred a solution of Boc-β-PyCOOH (420 mg, 1.38 mmol) in DMF (6.5 mL), was well dissolved by vortex and loaded in AA#4 to TVA and MVA for sequence coupling followed up by capping. After sequence 3 plus cap was completed, a solution of Boc-γ-Im-COOH (450 mg, 1.38 mmol) in DMF (6.5 mL) and DIPEA (1mL) was well dissolved by vortex; loaded in AA#5 transferred to TVA, then MVA before transferred to RV for coupling reaction for 180 min followed by capping with acetic anhydride for 17 min.

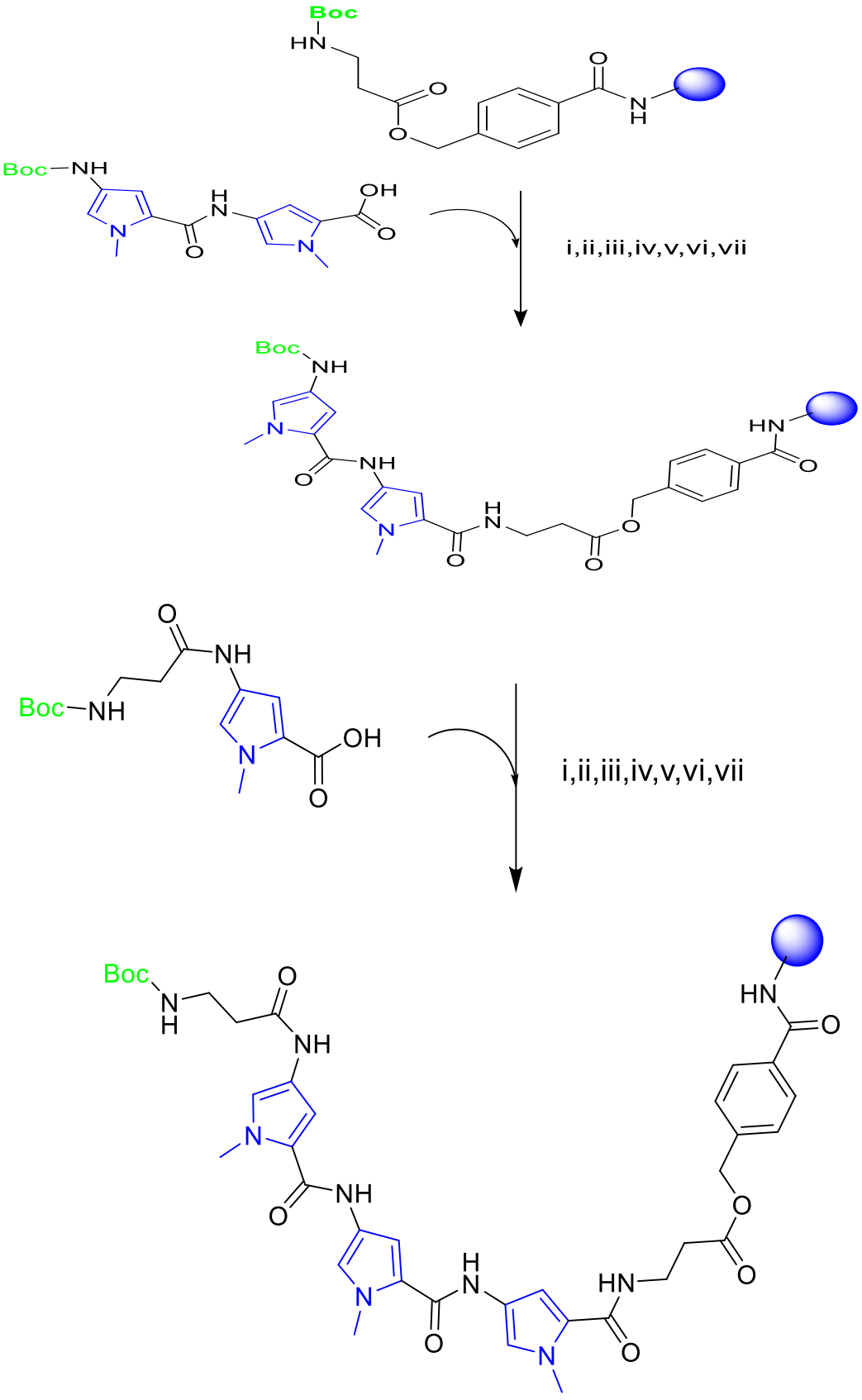
After sequence 4 plus cap was completed, a solution of Boc-Py-OBt (660 mg, 1.84 mmol) in DMF (6.5 mL) was well dissolved by vortex; loaded in AA#6 transferred to TVA then MVA before transferred to RV for coupling reaction for 180 min followed by capping with acetic anhydride for 17 min. After sequence 5 was finished, the programed system was initiated and followed by transferring a well dissolved by vortex solution of Boc-Py-Im-COOH (502 mg, 1.38 mmol) in DMF (6.5 mL) and DIPEA (1mL) from AA#6 to TVA then MVA before being transferred to RV for coupling reaction for 180 min following by capping with acetic anhydride for 17 min.

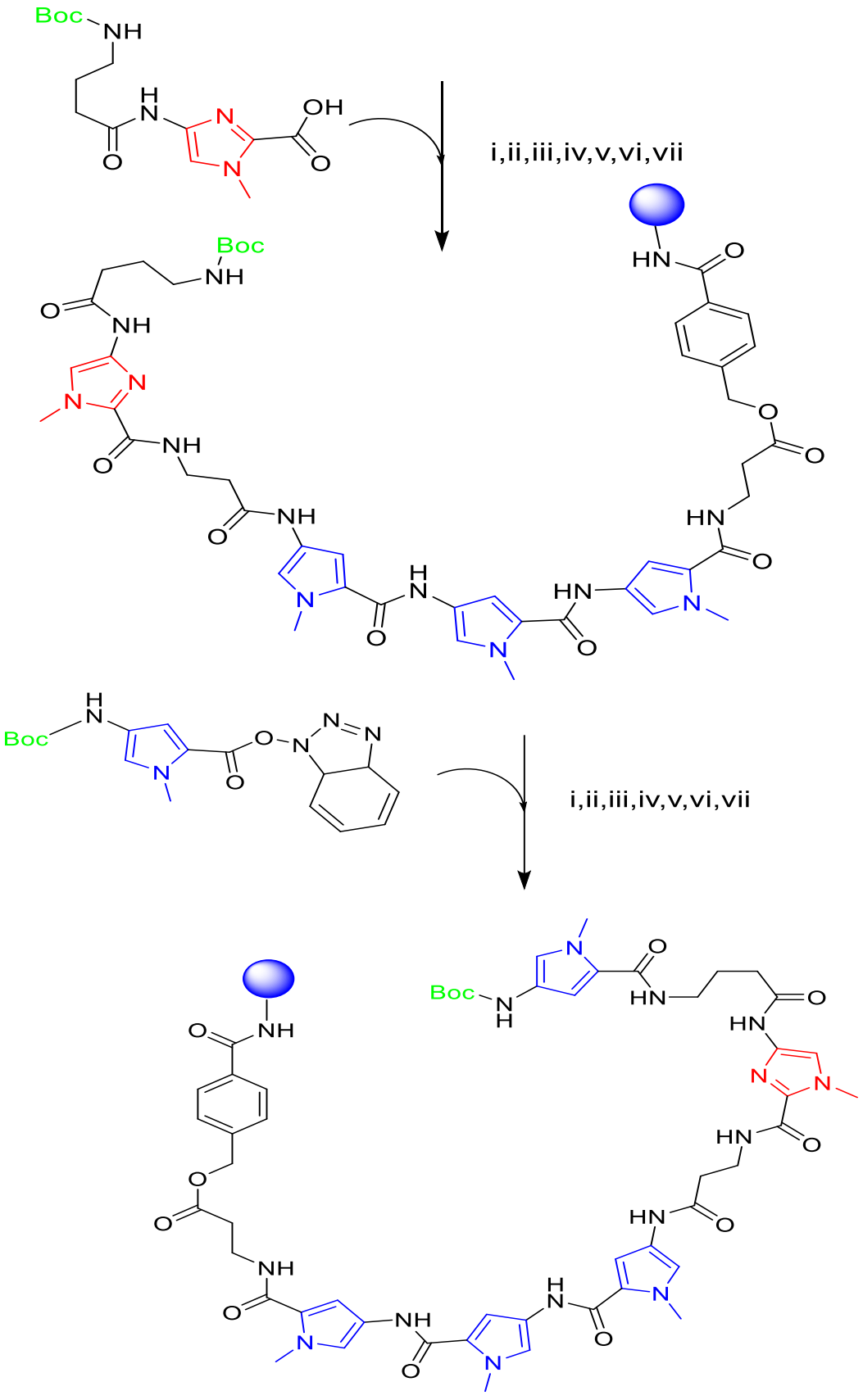
After sequence 6 was finished, the programed system was initiated and followed by transferring a well-dissolved by vortex solution of dimer desIm-Im-COOH (345 mg, 1.38 mmol) and DIPEA (1ml) in DMF (6.5 mL) from AA#8 to TVA then MVA before transferred to RV for coupling reaction for 180 min following by capping with acetic anhydride for 17 min. (**Scheme 3.2**)

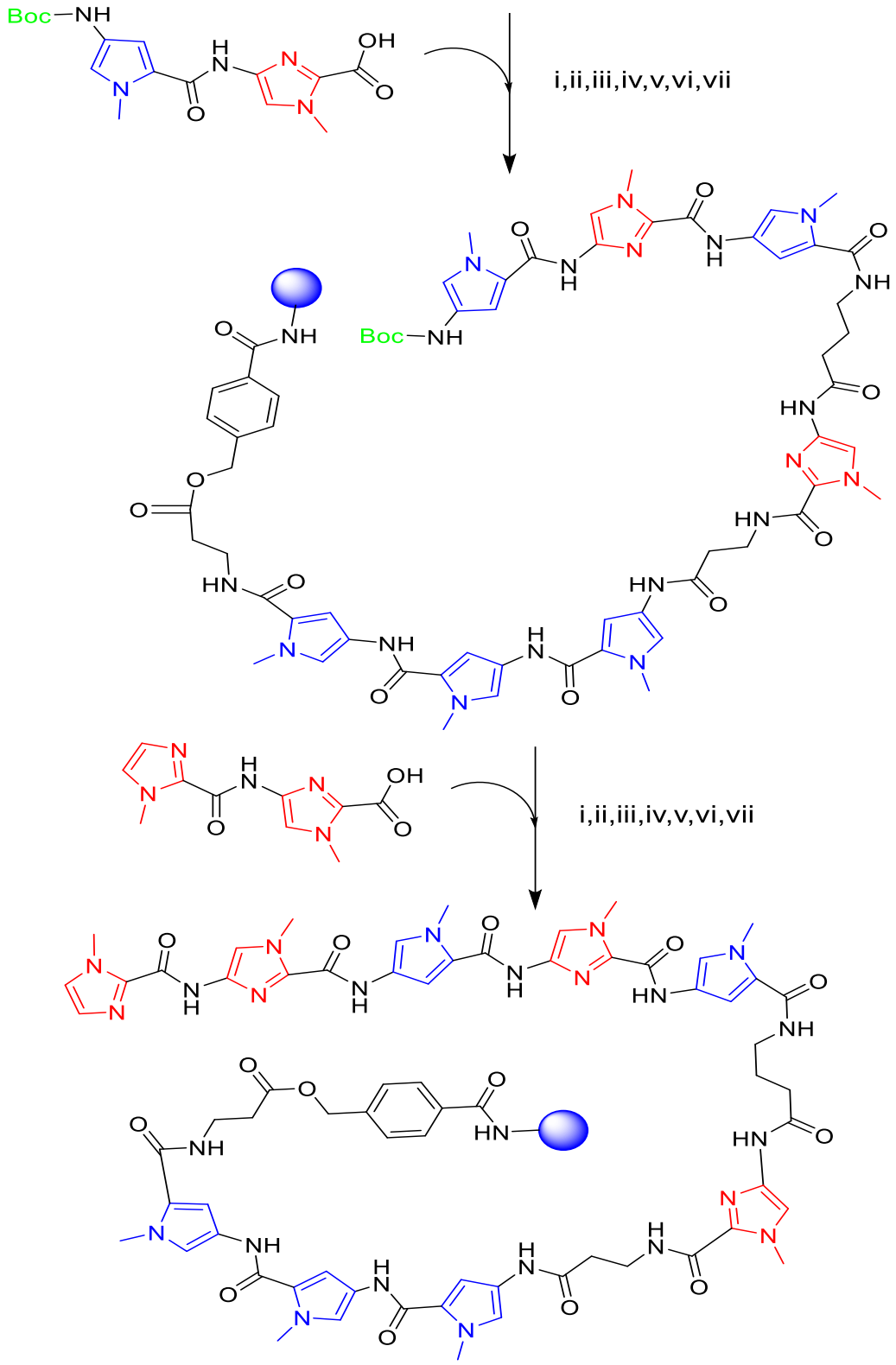
The sequence was finished after 30 h. 0,1 mg of resin was cleaved with 100 µL Ta solution for 2 h at 45°C, then the heated solution was collected and filtered before being analyzed with a Bruker LC-MS instrument. The exact mass was determined to be $[M+H]^+$ at 1460.62 m/z, confirming the mass of the desired product. All the resin was carefully transferred from the RV into a beaker using a small spatula, and the resin was washed three-time with DCM and MeOH, then dried under vacuum for 30 min until reaching dryness. The resin was then collected in a 20mL scintillation vial and stored in the freezer before cleavage in the next step.

Table 3.2. Programmed sequences of **1134** resin synthesis

Building block(s) and step(s)	Additional Solvent mixture	Applied program
1. Boc-Py-Py-COOH	(6.5 mL) DMF	Add-DIPEA-180-min-DIPEA-Flows.CFN
2. Capping	None	Cap-17-min-continues.CFN
3. Boc-β-Py-COOH	(6.5 mL) DMF	Add-DIPEA-180-min-DIPEA-Flows.CFN
4. Capping	None	Cap-17-min-continues.CFN
5. Boc-γ-Im-COOH	(6.5 mL) DMF + DIPEA (1mL)	Add-DIPEA-180-min-DIPEA-Flows.CFN
6. Capping	None	Cap-17-min-continues.CFN
7. Boc-Py-OBt	(6.5 mL) DMF	Add-DIPEA-180-min-DIPEA-Flows.CFN
8. Capping	None	Cap-17-min-continues.CFN
9. Boc Py-Im-COOH	(6.5 mL) DMF + DIPEA (1mL)	Add-DIPEA-180-min-DIPEA-Flows.CFN
10. Capping	None	Cap-17-min-continues.CFN
11. desIm-Im-COOH	(6.5 mL) DMF + DIPEA (1mL)	Add-DIPEA-180-min-DIPEA-Flows.CFN
12. Capping	None	Cap-17-min-continues.CFN


Boc
HN
O
O
O
N
H
Boc
NH
N
H
N
O
O
OH
O
i,ii,iii,iv,v,vi,vii
Boc
NH
N
O
NH
N
H
O
O
HN
O
Boc
NH
O
NH
N
OH
O
i,ii,iii,iv,v,vi,vii
Boc
N
H
O
NH
N
O
NH
N
H
N
O
O
HN
O
O
HN
O


Boc
NH
H
N
O
N
OH
O
N
i,ii,iii,iv,v,vi,vii
HN
O
O
O
HN
O
HN
N
Boc
NH
O
NH
N
N
O
NH
O
NH
N
H
N
O
HN
N
Boc
H
N
N
O
O
N
N
i,ii,iii,iv,v,vi,vii
N
Boc
N
H
O
HN
O
HN
N
N
HN
O
HN
O
O
NH
O
O
NH
O
NH
N
H
N
O
N
NH
O
N

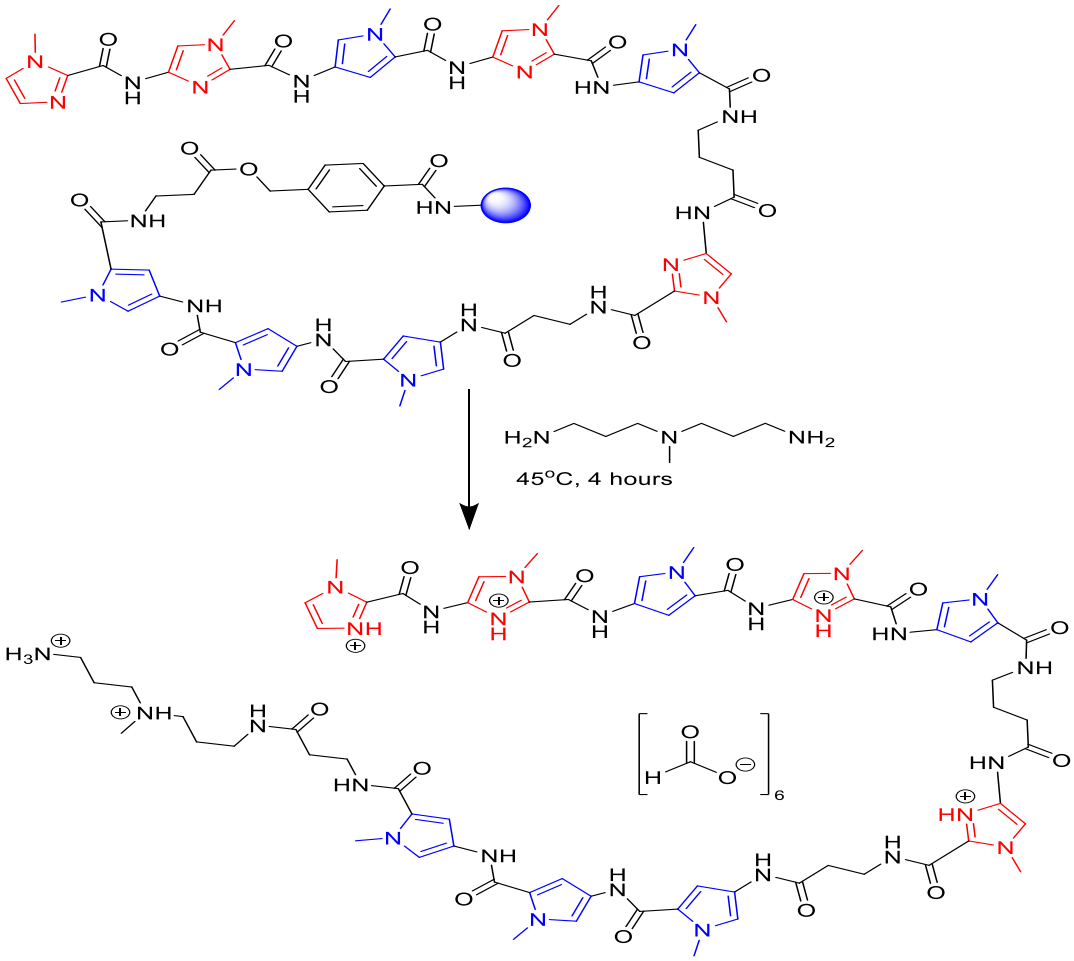


Scheme 3.2: *Scheme of synthesis polyamide **1134** resin using solid-phase synthesis method. Started from a commercially acquired Boc-β-PAM resin, and each sequence contains multiple steps from (i) to(viii). (i) wash the resin with DCM and DMF, (ii) deprotection with 60% TFA/DCM/0.5 M*

Indole,(iii) 25% piperidine/DMF; (iv) building blocks with the structures shown on the scheme,(v) 0.5M PyBOP/DMF as coupling reagents, (vi) DIPEA, (vii) DMSO, (viii) capping step with acetic anhydride for 17 min.

Cleavage of PAM resin using Ta as a nucleophile.

The next day, 350 g resin was cleaved with aminolysis in 2000 µL of Ta at 45ºC in a heating device Isotemp Fisher Scientific for 4 h, and the separation of polyamides solution from the PAM resin support was performed in a 20 µM polyethylene frit placed inside a disposable polypropylene syringe by washing with MeOH and H_2O. The polyamide solution was concentrated in a rotary evaporator to afford off-white solid **1134** (22 mg, 0.0805 mmol, 18.7% yield) (**Scheme 3.3**).



Scheme 3.3 Cleavage PA **1134** from PAM resin via Ta as the nucleophile.

Purification of compound

The polyamides solution was diluted with a mixture of DMSO (200 µL) and H_2O/0.2% HCOOH (200 µL), then purified by preparative HPLC using a Phenomenex Luna 250x30 mm, 5 µM, 100Å, C18 column maintained at 25ºC. The organic phase was 0.2% HCOOH in MilliQ water

and 100% HPLC grade MeOH. The applied gradient was 10% MeOH for 8 min, followed by a ramp rate to 90% Methanol over 35.6 min at a 20 mL/min flow rate. All collected fractions were analyzed and selected for 95% or higher purity with analytical HPLC using ThermoFisher U-HPLC; analysis was processed and analyzed with the program Chromeleon 3.0. High-purity fractions were collected for rotary evaporation to remove MeOH from the solution before lyophilization.

Lyophilization of compound

Compound **1134** was diluted with a 1:1 mixture of ACN (5 mL): and 0.2% HCOOH in H_2O (5 mL). The samples were frozen before the lyophilization process with ground dry ice until an even coated of frozen material results at an acute angle in the vial. The frozen sample was covered with porous Kimwipe paper and secured by a rubber band. The sample was settled in a lyophilizer glass vessel and connected to a vacuum port. The frozen polyamide solutions sublimed as an off-white solid powder after 48 h. The lyophilizer vacuum was set at 300 torr within 30 min prior to the lyophilization process, and the inside temperature was set at -83 to -85^{o}C.

Characterization of the compound.

Analytical HPLC characterization was performed with a C12 Phenomenex Juniper Proteo column maintained at 40^{o}C. The mobile phase consisted of 0.1% formic acid in water (A) and HPLC grade ACN (B). The applied gradient consisted of 5% B for 0.75 min, followed by a ramp to 60% B over 6.5 min at 2.0 mL/min. The retention was 3.050 min.

PA 1134, desIm-Im-Py-Im-Py-γ-Im-β-Py-Py-Py-β-Ta $(HCOO^-)_6$

^{1}H-NMR (600 MHz, DMSO-d_6) δ= 10.355 (s, 1H), 10.32-10.25 (d, J= 22.9 Hz, 2H), 10.04 (s, 1H), 9.97- 9.90 (t, J= 43.1 Hz, 3H), 8.37 (s, 4H), 8.05 (s, 1H),8.01-7.97 (d, J= 20.12 Hz, 2H), 7.90 (s,1H), 7.57- 7.53 (d, J=21.55 Hz, 2H), 7.44 (s, 1H), 7.39(s, 2H), 7.05 (s, 1H), 7.20-7.18 (d, J= 12.94 Hz, 2H), 7.16-7.13 (d, J= 20.21 Hz, 2H), 6.96 (s, 1H), 6.86 (s, 1H), 6.82 (s, 1H), 3.99 (s, 3H), 3.94 (s, 2H), 3.89 (s, 2H), 3.83 (s, 2H), 3.808 (s, 2H), 3.77 (s, 2H), 3.72 (s, 1H), 3.53 -3.49 (d, J=30.1Hz,, 3H), 3.17 (s, 1H), 3.03 (s, 1H), 2.886 (s, 1H), 2.77 (s, 3H), 2.48 (s, 9H), 2.31- 2.30 (d, J= 7.19Hz, 6H), 2.24 (s, 4H), 2.125 (s, 2H), 2.07 (s, 5H), 1.75-1.73 (t, J= 17.2 Hz, 2H), 1.66 (s, 5H), 1.50-1.48 (t, J=10.06 Hz, 2H), 1.410 (s, 2H), 1.205 (s, 1H), 1.15-1.13 (t, J= 8.6 Hz, 2H).

^{13}C-NMR (151 MHz, DMSO-d_6) δ= 169.89, 167.98, 165.64, 165.62, 161.25, 161.17, 158.72, 158.48, 158.42, 155.78, 155.73, 155.61, 137.81, 136.15, 136.05, 134.72, 134.37, 134.17, 127.71, 127.10, 127.10, 123.33, 122.81, 122.77, 122.23, 122.18, 121.82, 121.39, 119.51, 118.48, 118.20,

117.92, 117.88, 114.70, 114.07, 113.60, 105.74, 104.74, 104.20, 103.96, 54.52, 54.36, 53.97, 41.32, 41.28, 41.24, 40.07, 39.95, 39.81, 39.67, 39.53, 39.39, 39.25, 39.11, 38.14, 37.53, 37.15, 36.62, 36.22,36.02, 35.93, 35.54, 35.48, 35.17, 35.11, 34.91, 34.78, 34.01, 33.32, 32.84, 29.28, 28.19, 26.70, 25.50, 25.33, 24.72, 24.63, 24.50, 24.46.

HRMS (ESI) was calculated for **1134,** $C_{67}H_{85}N_{27}O_{12,}$ $[M+H]^+$, 1460.69 m/z, and found to be 1460.73 m/z.

HPLC purity: 99%

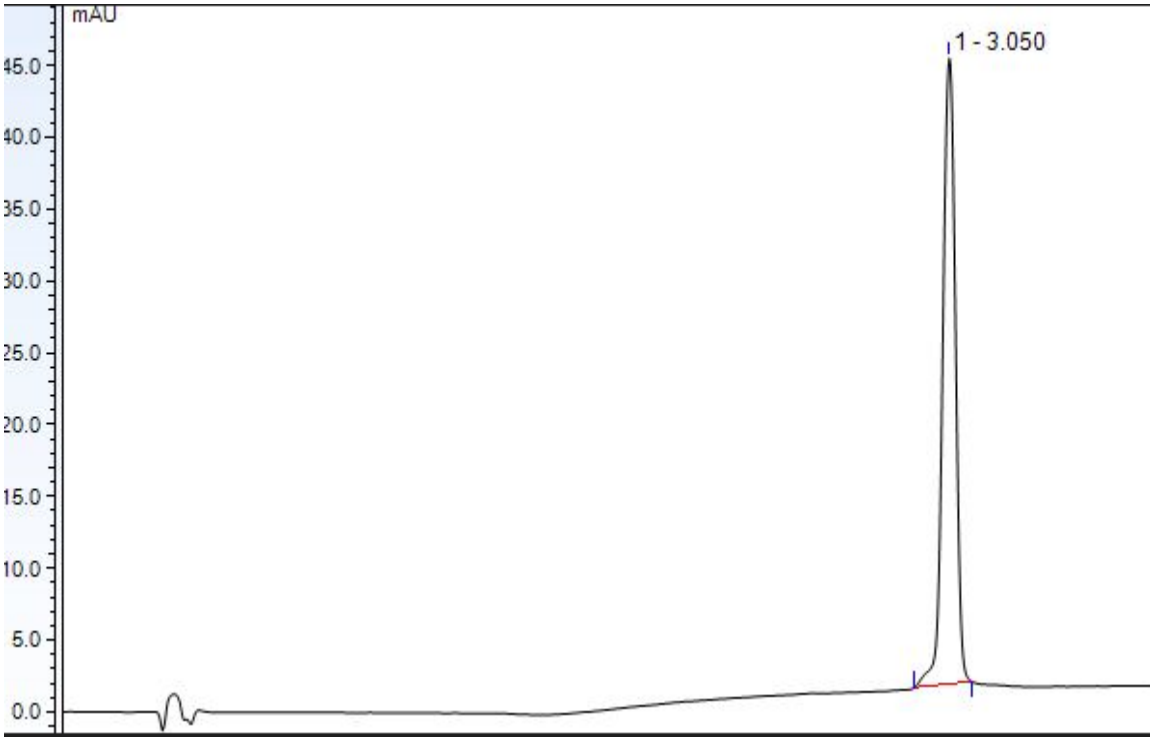


Figure 3.2. A*nalytical HPLC purity of compound **1134** with λ monitored. Compound purity 98%, retention time 3.050 min*.

3.2.4 Synthesis of desIm-Im-β-Im-Py-γ-Im-Py-β-Py-Py-β-Ta (1161)

Synthesis of 1161 resin

Synthesis of polyamides **1161** resin started by initiating the reaction sequences program with the CSBio program. The CSBio reaction vessel was set at a temperature of 35ºC throughout the whole process. Boc-β-alanine-PAM resin (1000 mg) was loaded into the RV. PyBOP (11 g) was dissolved in DMF (40 mL) and filled in the R4 solvent bottle.

After the program started the first wash and deprotection reaction for 1 h, a solution of dimer Boc-Py-Py-COOH (506 mg, 1.38 mmol) in DMF (6.5 mL) was stirred to fully dissolved and

transferred into AA#3. The coupling program directed the synthesizer to transfer the solution from AA#3 to TVA and then MVA to begin the activation before transferring to RV for coupling for 180 min, followed by capping with acetic anhydride for 17 min to remove impurities such as an unreacted amine. After completing the first sequence, the program proceeded to "cap" to remove any nonreacted amine before moving on to the following sequence. Each sequence took around 6 h to complete.

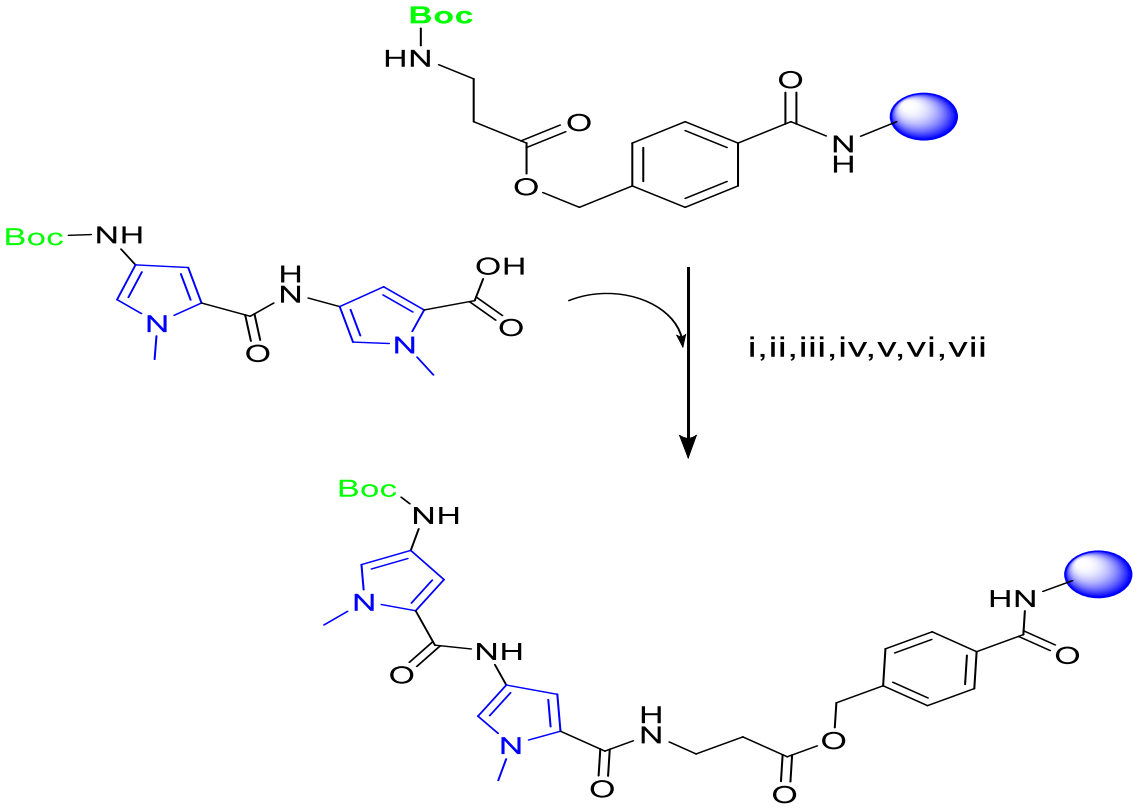
After completing the first sequence, the programmed system transferred a solution of dimer Boc-Py-β-COOH (433 mg, 1.38 mmol) in DMF (6.5 mL), was well dissolved by vortex and loaded in AA#4 to TVA and MVA for sequence coupling followed up by capping. After sequence 3 plus cap was completed, a solution of Boc-γ-Im-COOH (450 mg, 1.38 mmol) in DMF (6.5 mL) and DIPEA (1mL) was well dissolved by vortex; loaded in AA#5 transferred to TVA, then MVA before transferred to RV for coupling reaction for 180 min followed by capping with acetic anhydride for 17 min. After sequence 4 plus cap was completed, a solution of Boc-Py-OBt (660 mg, 1840 mmol) in DMF(6.5mL) was well dissolved by vortex; loaded in AA#6, transferred to TVA then MVA before transferred to RV for coupling reaction for 180 min followed by capping with acetic anhydride for 17 min.

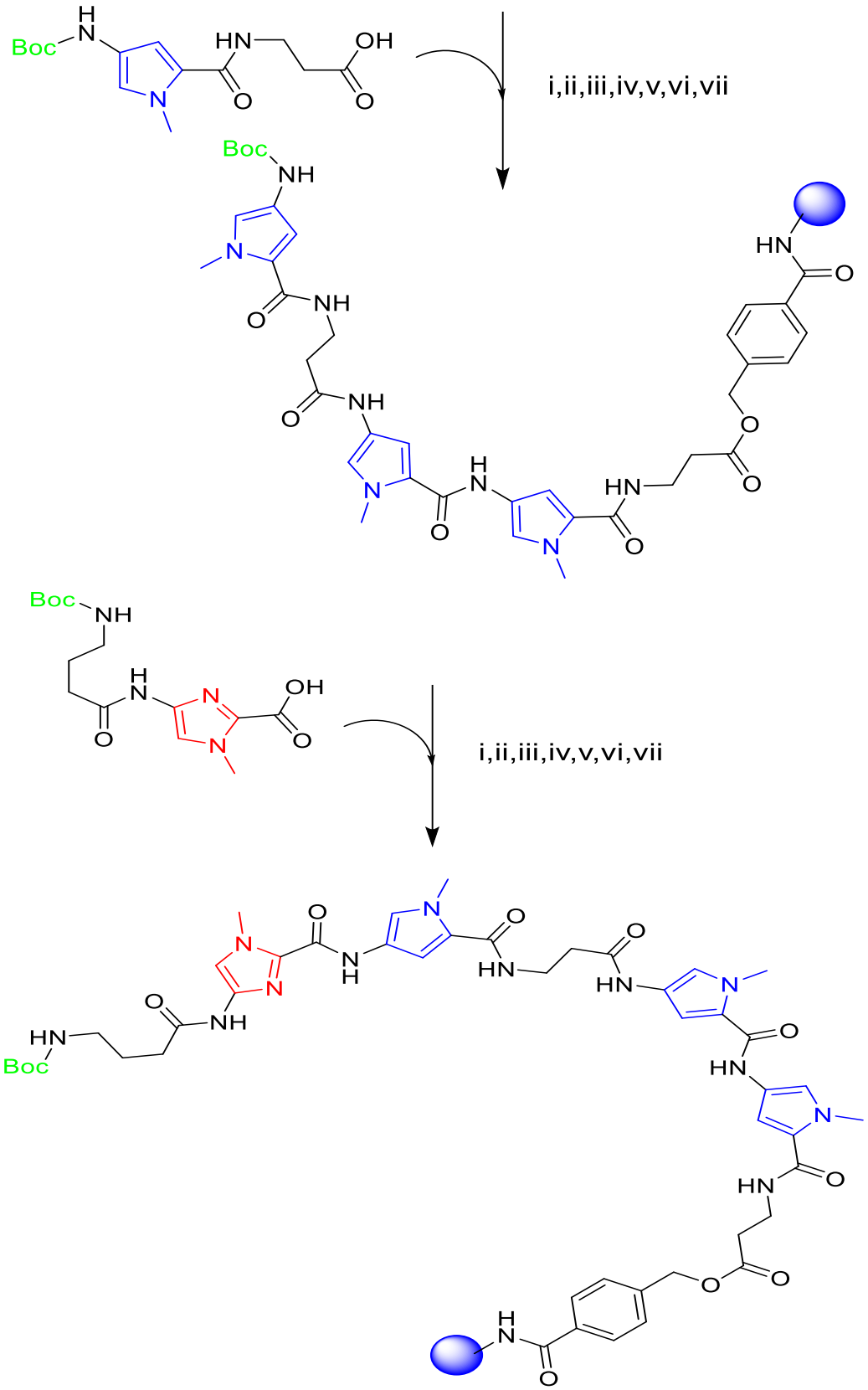
After sequence 5 was finished, the programed system was initiated and followed by transferring a well dissolved by vortex solution of Boc-β-Im-COOH (433 mg, 1.38 mmol) in DMF (6.5 mL) and DIPEA (1mL) from AA#6 to TVA then MVA before being transferred to RV for coupling reaction for 180 min following by capping with acetic anhydride for 17 min. After sequence 6 was finished, the programed system was initiated and followed by transferring a well-dissolved by vortex solution of dimer desIm-Im-COOH (346mg, 1.38 mmol) and DIPEA (1ml) in DMF (6.5 mL) from AA#7 to TVA then MVA before transferred to RV for coupling reaction for 180 min following by capping with acetic anhydride for 17 min.

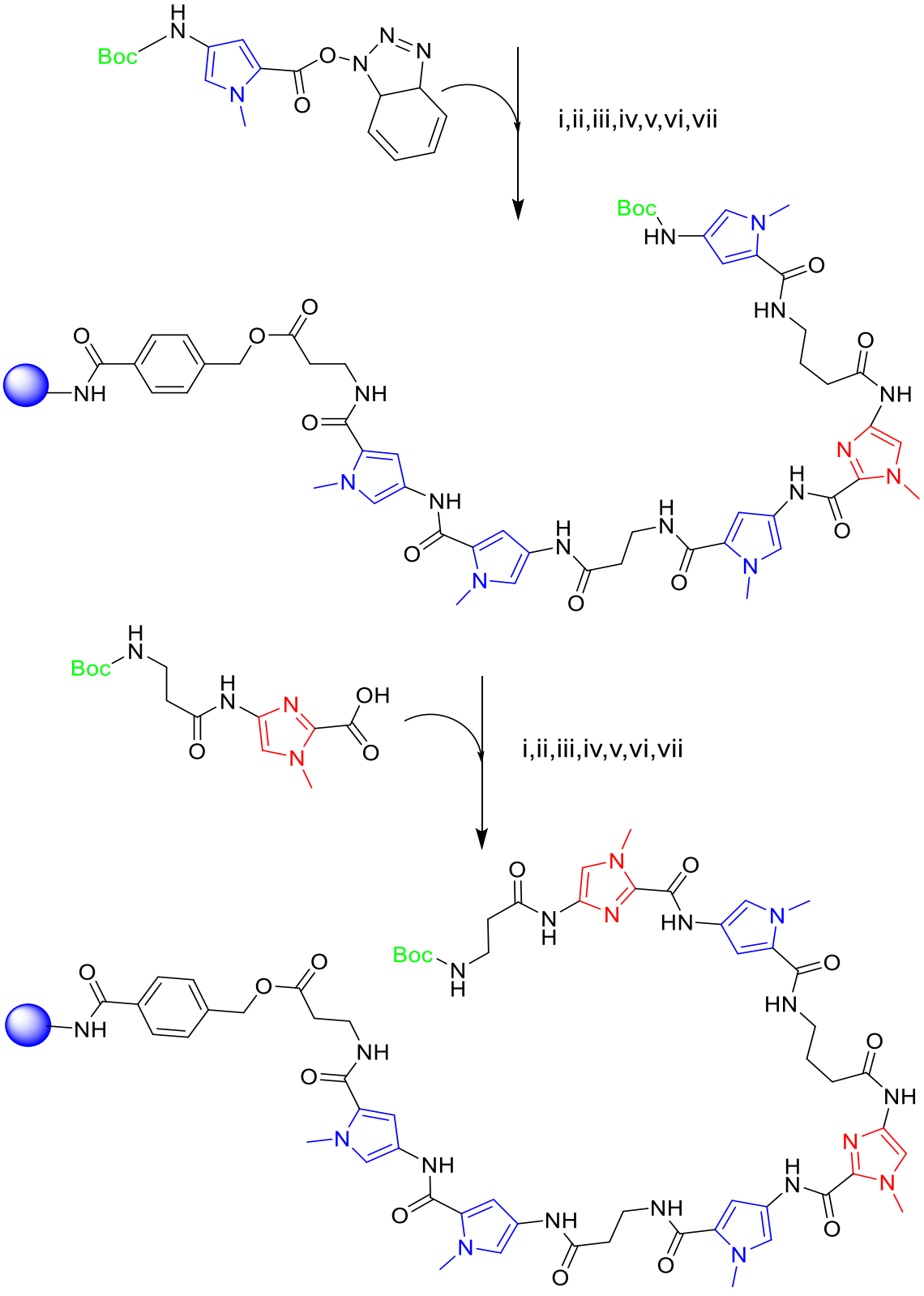
The sequence was finished after 30 h. 0,1 mg of resin was cleaved with 100 μL Ta solution for 2 h at 45ºC, and then the heated solution was collected and filtered before being analyzed with a Bruker LC-MS instrument. The exact mass was determined to be $[M+H]^+$ 1409.73 m/z, which confirmed the presence of the desired product. All the resin was carefully transferred from the RV into a beaker using a small spatula, and the resin was washed three-time with DCM and MeOH, then dried under vacuum for 30 min until reaching dryness. The resin was then collected in a 20mL scintillation vial and stored in the freezer before cleavage the next day.

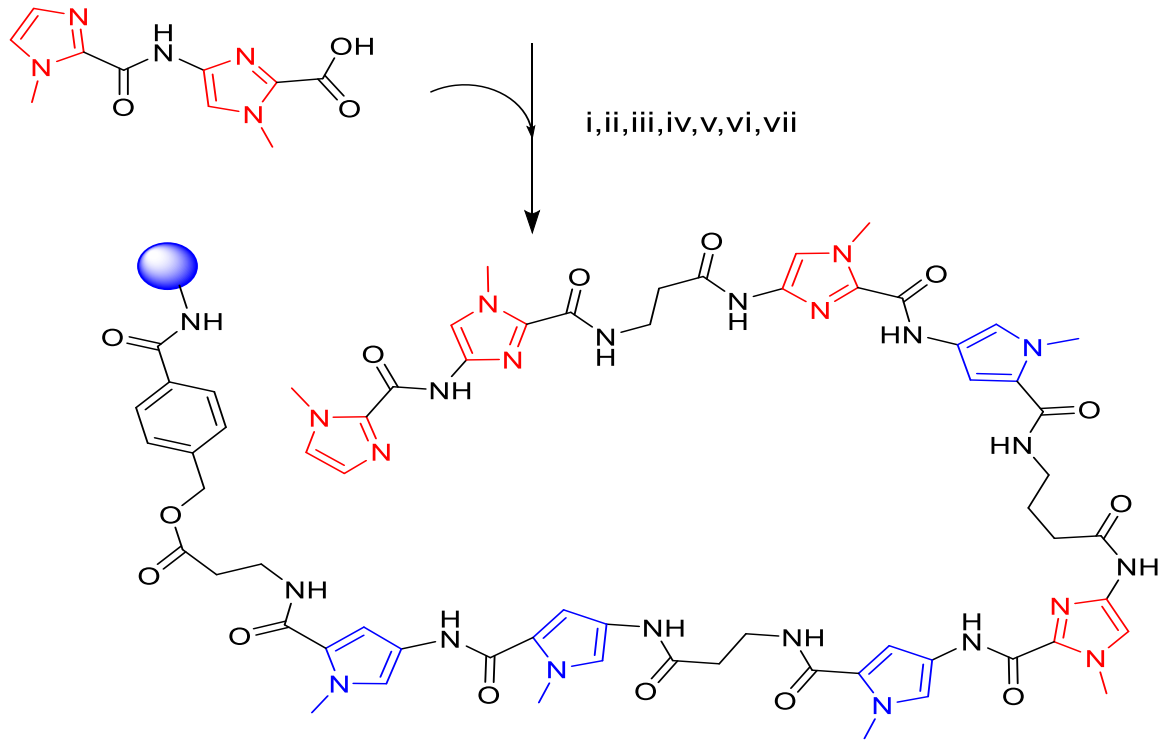
Table 3.3. Programmed sequence of **1161** resin synthesis

Building block(s) and step(s)	Additional Solvent mixture	Applied program
1-Boc-Py-Py-COOH	(6.5 mL) DMF	Add-DIPEA-180-min-DIPEA-Flows.CFN
2-Capping	None	Cap-17-min-continues.CFN
3-Boc-Py-β-COOH	(6.5 mL) DMF	Add-DIPEA-180-min-DIPEA-Flows.CFN
4-Capping	None	Cap-17-min-continues.CFN
5-Boc-γ-Im-COOH	(6.5 mL) DMF+DIPEA (1mL)	Add-DIPEA-180-min-DIPEA-Flows.CFN
6-Capping	None	Cap-17-min-continues.CFN
7-Boc-Py-OBt	(6.5 mL) DMF	Add-DIPEA-180-min-DIPEA-Flows.CFN
8-Capping	None	Cap-17-min-continues.CFN
9-Boc β-Im-COOH	(6.5 mL) DMF+DIPEA (1mL)	Add-DIPEA-180-min-DIPEA-Flows.CFN
10-Capping	None	Cap-17-min-continues.CFN
11-desIm-Im-COOH	(6.5 mL) DMF+DIPEA (1mL)	Add-DIPEA-180-min-DIPEA-Flows.CFN
12-Capping	None	Cap-17-min-continues.CFN




i,ii,iii,iv,v,vi,vii
i,ii,iii,iv,v,vi,vii

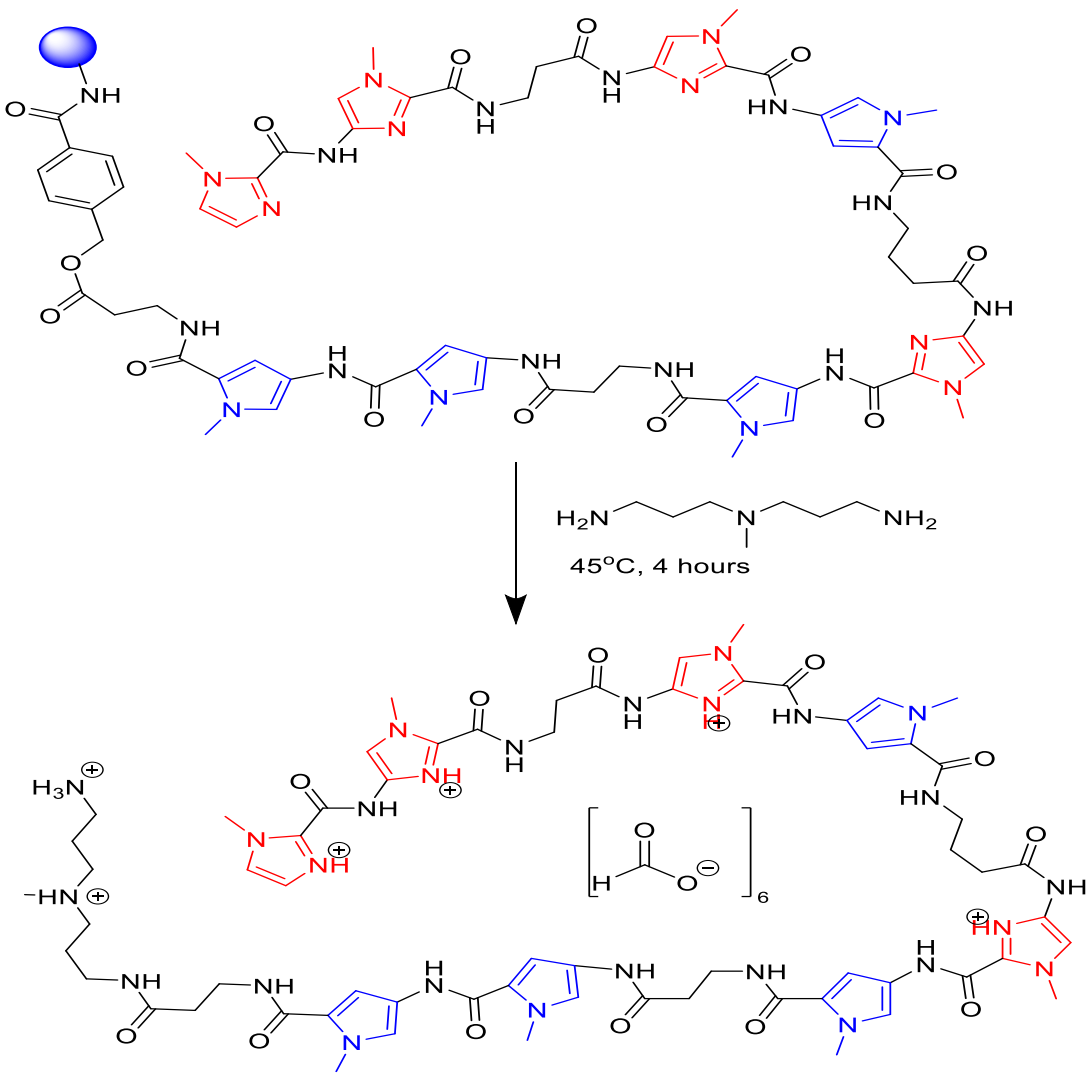

Boc
i,ii,iii,iv,v,vi,vii
Boc
i,ii,iii,iv,v,vi,vii



Scheme 3.4 Scheme of synthesis polyamide **1161** resin by utilizing solid-phase synthesis method. It started from a commercially acquired Boc-β-PAM resin, and each sequence contains multiple steps from (i) to (viii). (i) wash the resin with DCM and DMF, (ii) deprotection with 60% TFA/ DCM/ 0.5 M Indole, (iii) 25% piperidine/ DMF; (iv) building blocks with the structures shown on the scheme,(v) 0.5M PyBOP/ DMF as coupling reagents, (vi) DIPEA, (vii) DMSO, (viii) capping step with acetic anhydride for 17 min.

Cleavage of PAM resin using Ta as a nucleophile.

The next day, 270 mg resin was cleaved with aminolysis in 2000 µL of Ta at 45°C in a heating device Isotemp Fisher Scientific for 4 h, and the separation of polyamides solution from the PAM resin support was performed in a 20 µM polyethylene frit placed inside a disposable polypropylene syringe by washing with MeOH and H_2O. The polyamide solution was concentrated in a rotary evaporator to afford off-white solid **1161** (35 mg, 62.1 mmol, 38.6% yield) off-white solid. **(Scheme 3.5**).



Scheme 3.5 *Cleavage PA**1161** from PAM resin using Ta as the nucleophile.*

Purification of compound

The polyamide solution was diluted with a mixture of DMSO (200 µL) and H_2O/0.2% HCOOH (200 µL), then it was purified by preparative HPLC using a Phenomenex Luna 250x30 mm, 5 µM, 100Å C18 column maintained at 25 °C. The organic phase was 0.2% HCOOH in MilliQ water and 100% HPLC grade MeOH. The applied gradient was 10% MeOH for 8 min, followed by a ramp rate to 90% Methanol over 35.6 min at a 20 mL/min flow rate. All collected fractions were analyzed and selected for 95% or higher purity with analytical HPLC using ThermoFisher-UHPLC; analysis was processed and analyzed with the program Chromeleon 3.0. High-purity fractions were collected for rotary evaporation to remove MeOH before the lyophilization.

Lyophilization of the compound.

The compound **1161** was diluted with a 1:1 mixture of ACN (5 mL) and 0.2% HCOOH in H_2O (5 mL) after the rotary evaporator. The samples were frozen before the lyophilization process with ground dry ice until an even coated of frozen material results at an acute angle in the vial. The frozen sample was covered with porous Kimwipe paper and secured by a rubber band. The sample was settled in a lyophilizer glass vessel and connected to a vacuum port. The frozen polyamide solutions sublimed as an off-white solid powder after 48 h. The lyophilizer vacuum was set at 300 torr within 30 min prior to the lyophilization process, and the inside temperature was set at -83 to -85°C.

Characterization of the compound.

Analytical HPLC characterization was performed with a C12 Phenomenex Juniper Proteo column maintained at 40°C. The mobile phase consisted of 0.1% formic acid in water(A) and HPLC grade ACN (B). The applied gradient consisted of 5%B for 0.75 min, followed by a ramp to 60% B over 6.5 min at 2.0 mL/min. The retention was 2.527 min.

1161: desIm-Im-β-Im-Py-γ-Im-Py-β-Py-Py-β-Ta $(HCOO^-)_6$

^{1}H-NMR (600 MHz, DMSO-d_6) δ= 10.32 (s, 1H), 10.22 (s,1H), 9.88- 9.81 (t, J= 31.9 Hz, 4H), 9.73 (s,1H), 8.198 (s, 2H), 8.13 (s, 1H), 8.05-8.04 (d, J=5.71Hz, 2H), 8.01-7.98 (t, J=14.2 Hz, 4H), 7.50 (s, 1H), 7.44 (s, 1H), 7.41 (s, 2H), 7.20-7.14 (d, J= 11.4 Hz, 2H), 7.13 (d, 1H), 7.03 (s, 1H), 6.94-6.90 (d, J=16.2 Hz, 2H), 6.85-6.81(d, J=17.1 Hz, 2H), 3.97 (s, 3H), 3.93-3.90 (t, J=17.1 Hz,10H), 3.79- 3.77 (d, J= 11.4 Hz, 12H), 3.57 (s, 1H), 3.50- 3.48 (q, J= 7.62 Hz, 3H), 3.44- 3.39 (q, J= 20.9 Hz, 4H), 3.38- 3.72 (q, J= 22.8 Hz, 4H), 3.19-3.15 (q, J= 19.1 Hz, 3H), 3.11- 3.05 (q, J= 20.0 Hz, 4H), 3.02 (s, 2H), 2.95- 2.92(d, J= 6.7 Hz, 2H), 2.84- 2.81(t, J= 15.2Hz, 8H), 2.64 (s, 2H), 2.61-2.57 (t, J= 12.4 Hz, 3H), 2.54 (s, 1H), 2.34- 2.30 (t, J= 16.2 Hz, 3H), 1.84 (s, 3H), 1.74 (s, 2H), 1.43(s, 1H), 1.2(s, 1H), 1.16-1.13 (t, J=16.1 Hz, 2H).

^{13}C-NMR (151 MHz, DMSO-d_6) δ= 174.0, 173.1, 171.7, 170.9, 164.4, 164.3, 164.3, 161.5, 161.41, 161.3, 159.0, 158.9, 158.8, 140.9, 139.2, 139.0, 137.8, 137.3, 137.2, 137.1, 130.7, 130.1, 126.4, 126.2 ,125.9, 125.8, 125.2, 125.0, 124.2, 124.2, 121.3, 121.2, 121.1, 121.1, 121.1, 117.2, 117.1, 117.0, 107.4, 107.4, 107.1, 56.5, 56.0, 55.4, 48.4, 48.8, 41.2, 39.8, 39.5, 39.2, 39.2, 39.1, 39.1, 39.0, 38.9, 38.8, 38.7, 38.6, 38.5, 38.2, 38.2, 38.0, 38.0, 37.9, 35.9, 35.5, 28.6, 27.4, 25.9, 25.2.

HRMS (ESI) was calculated for **1161**, $[M+H]^+$ 1409.68 m/z, found to be 1409.73 m/z.

HPLC purity of 1161:97%

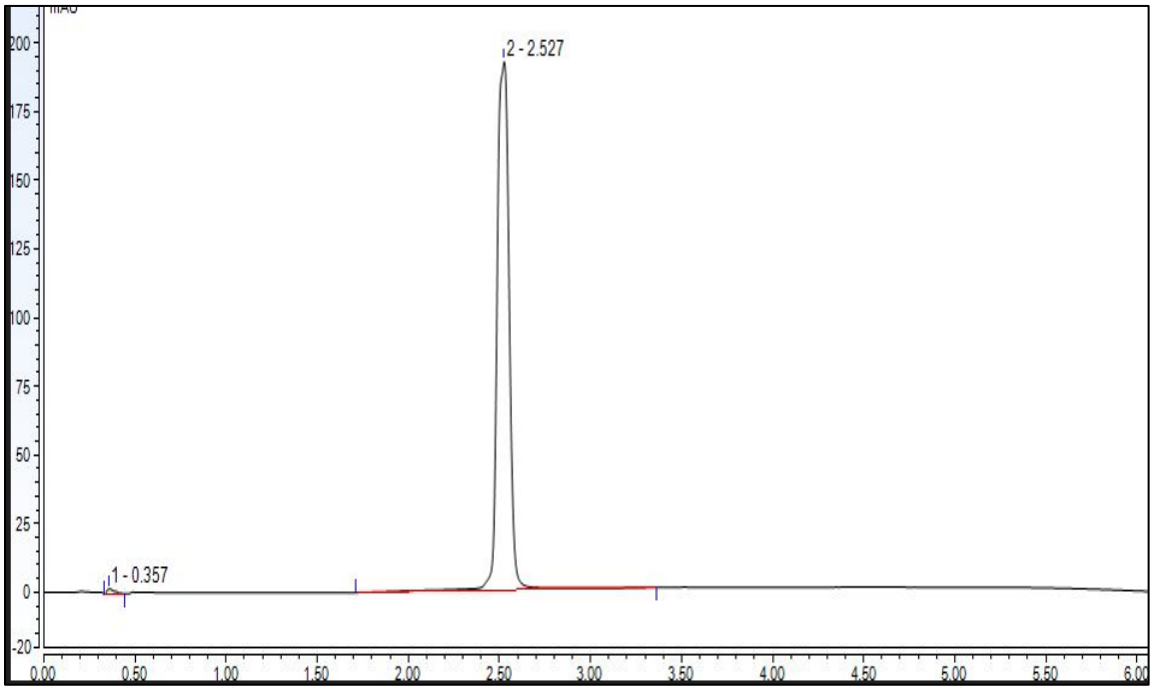


Figure 3.3 Analytical HPLC purity of compound **1161** with λ monitored. Compound purity 97%, retention time 2.527 min.

3.2.5 Synthesis of desIm-Im-β-Im-Py-γ-Im-β-Py-Py-Py-β-Ta (1162)

Synthesis of 1162.

Synthesis of polyamides **1162** resin started by initiating the reaction sequences program with the CSBio program. (**Table 3.4)** The CSBio reaction vessel was set at a temperature of 35°C throughout the process. Boc-β-alanine-PAM resin (1000 mg) was loaded into the RV. PyBOP (11 g) was dissolved in DMF (40 mL) and filled in the R4 solvent bottle. After the program started the first wash and deprotection reaction for 1 h, a solution of Boc-Py-Py-COOH (500 mg, 1.38 mmol) in DMF (6.5 mL) was stirred to fully dissolved and transferred into AA#3.

The coupling program directed the synthesizer to transfer the solution from AA#3 to TVA and then MVA to begin the activation before transferring to RV for coupling for 180 min, followed by capping with acetic anhydride for 17 min to remove impurities such as an unreacted amine. After completing the first sequence, the "capping" program removed any nonreacted amine before moving on to the next sequence. Each sequence took around 6 h to complete.

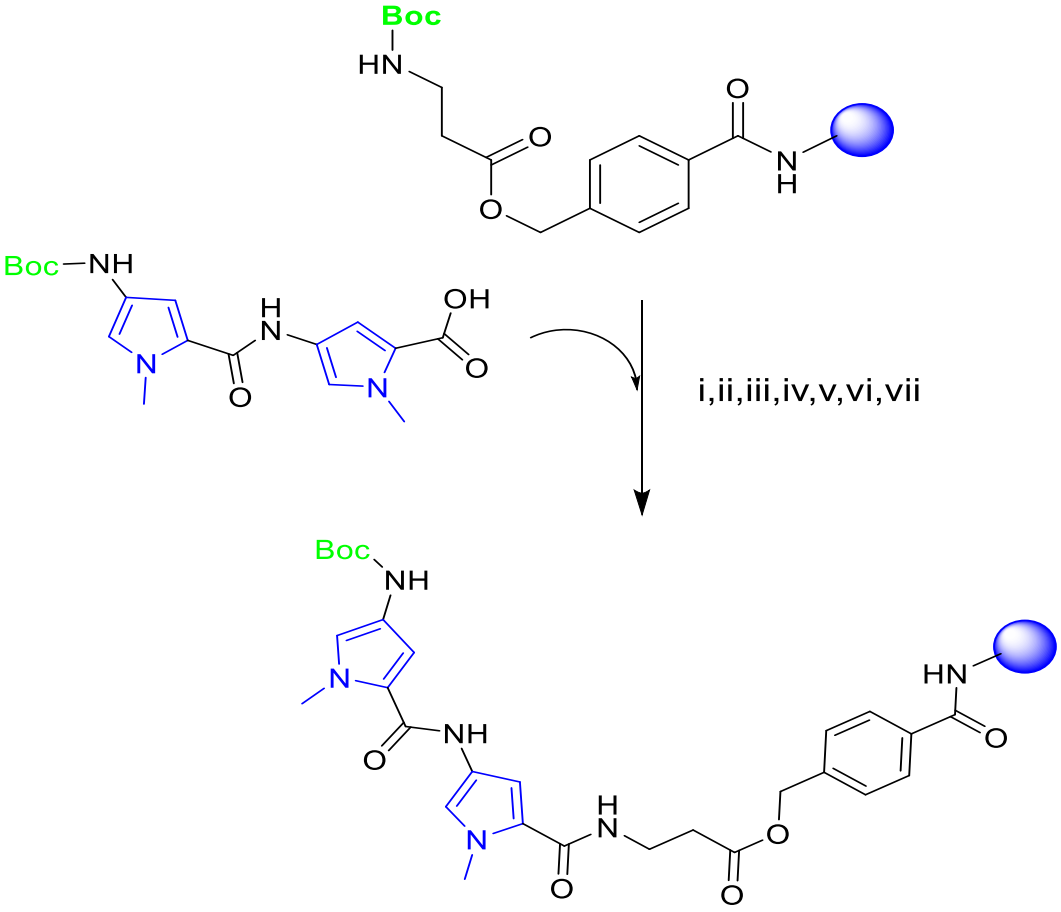
Table 3.4. Programmed sequence of **1162** resin synthesis.

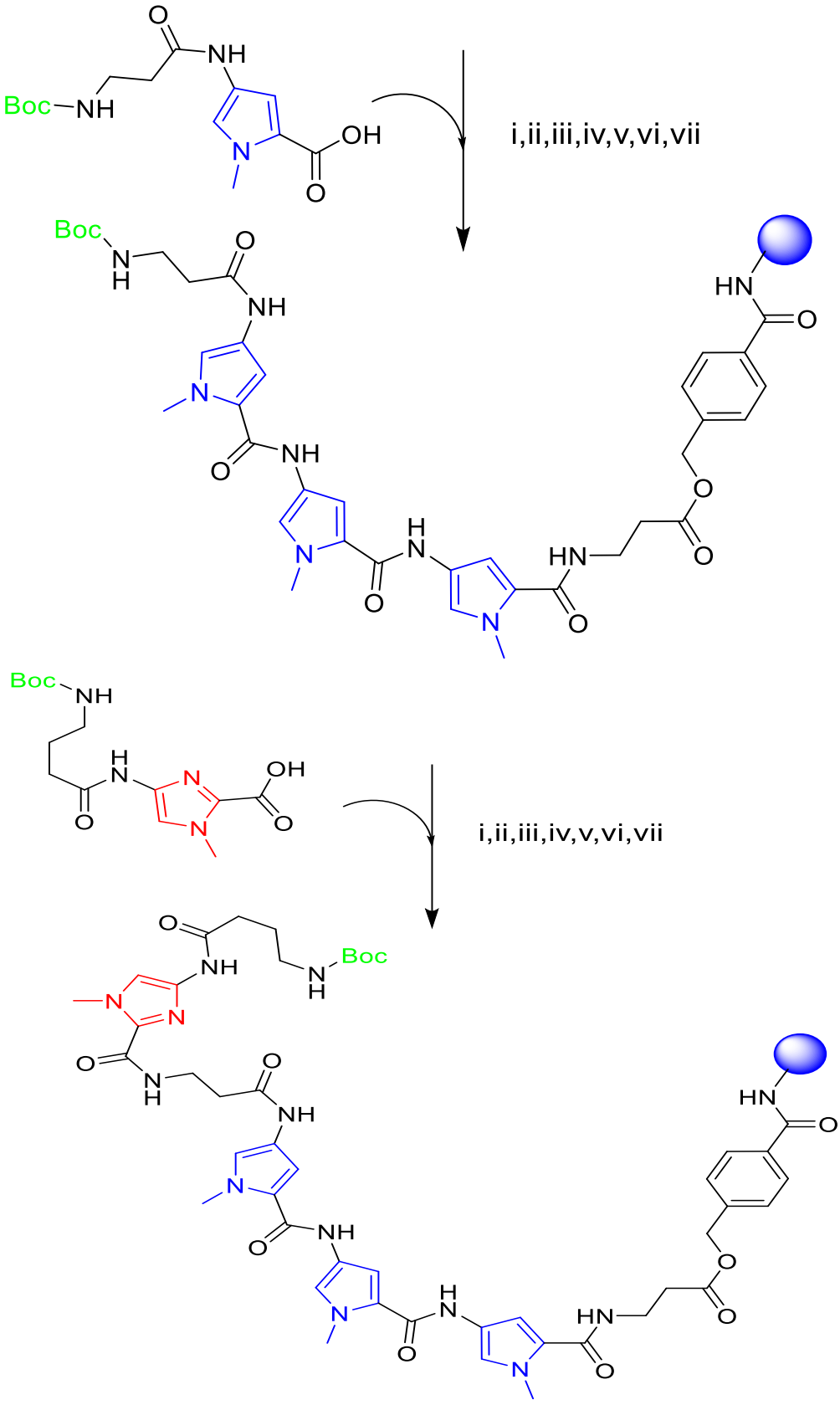
Building block(s) and step(s)	Additional Solvent mixture	Applied program
1-Boc-Py-Py-COOH	(6.5 mL) DMF	Add-DIPEA-180-min-DIPEA-Flows.CFN
2-Capping	None	Cap-17-min-continues.CFN
3-Boc-β-Py-COOH	(6.5 mL) DMF	Add-DIPEA-180-min-DIPEA-Flows.CFN
4-Capping	None	Cap-17-min-continues.CFN
5-Boc-γ-Im-COOH	(6.5 mL) DMF+DIPEA (1mL)	Add-DIPEA-180-min-DIPEA-Flows.CFN
6-Capping	None	Cap-17-min-continues.CFN
7-Boc-Py-OBt	(6.5 mL) DMF	Add-DIPEA-180-min-DIPEA-Flows.CFN
8-Capping	None	Cap-17-min-continues.CFN
9-Boc β-Im-COOH	(6.5 mL) DMF+DIPEA (1mL)	Add-DIPEA-180-min-DIPEA-Flows.CFN
10-Capping	None	Cap-17-min-continues.CFN
11-desIm-Im-COOH	(6.5 mL) DMF+DIPEA (1mL)	Add-DIPEA-180-min-DIPEA-Flows.CFN
12-Capping	None	Cap-17-min-continues.CFN

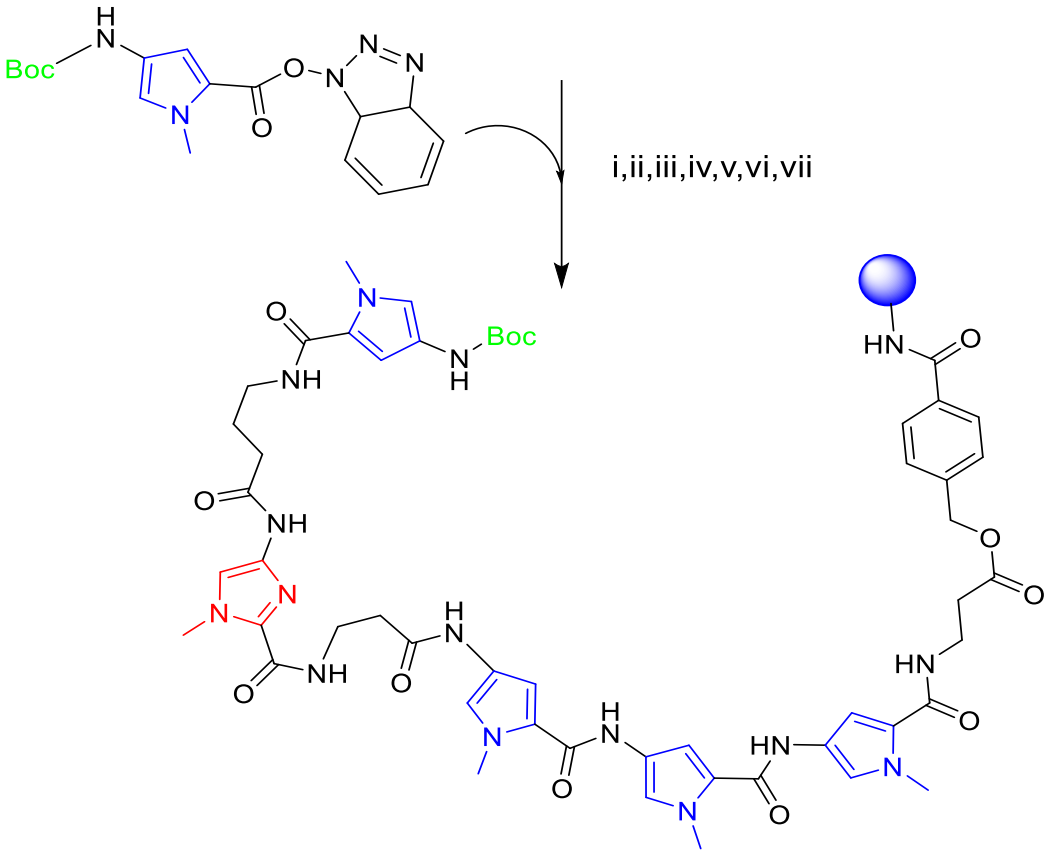
After completing the first sequence, the programmed system transferred a solution of Boc-β-Py-COOH (432 g, 1.38 mmol) in DMF (6.5 mL), was well dissolved by vortex and loaded in AA#4 to TVA and MVA for sequence coupling followed up by capping. After sequence 3 plus cap was completed, a solution of Boc-γ-Im-COOH (452 g, 1.38 mmol) and 1mL in DMF (6.5 mL) was well dissolved by vortex; loaded in AA#5 transferred to TVA then MVA before transferred to RV for coupling reaction for 180 min followed by capping with acetic anhydride for 17 min. After sequence 4 plus cap was completed, a solution of Boc-Py-Obt (660 mg, 1.38 mmol) in DMF (6.5 mL) was well dissolved by vortex; loaded in AA#6 transferred to TVA then MVA before transferred to RV for coupling reaction for 180 min followed by capping with acetic anhydride for 17 min.

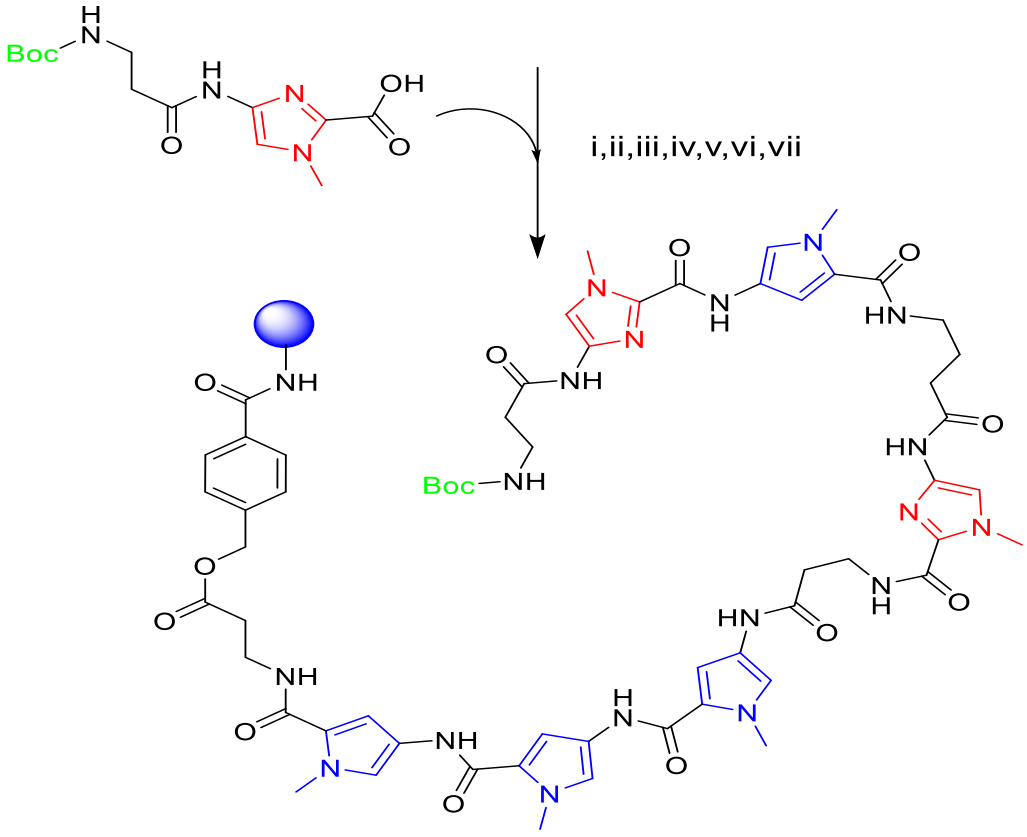
After sequence 5 was finished, the programed system was initiated and followed by transferring a well dissolved by vortex solution of Boc-β-Im-COOH (432 mg, 1.38 mmol) and DIPEA (1mL) in DMF(6.5mL) from AA#6 to TVA, then MVA before being transferred to RV for coupling reaction for 180 min following by capping with acetic anhydride for 17 min. After sequence 6 was finished, the programed system was initiated and followed by transferring a well dissolved by vortex solution of dimer desIm-Im-COOH (345 mg, 1.38 mmol) and DIPEA (1ml) in DMF (6.5 mL) from AA#8 to TVA then MVA before transferred to RV for coupling reaction for 180 min

following by capping with acetic anhydride for 17 min. (**Scheme 2.6**) The sequence was finished after 32 h total (The sequence paused due to sensor issues). 0,1 mg of resin was cleaved with 100 µL Ta solution for 2 H at 45°C; then, the heated solution was collected and filtered before being analyzed with a Bruker LC-MS instrument. The exact mass was determined to be $[M+H]^+$ at 1409.53 m/z, confirming the mass of the desired product. All the resin was carefully transferred from the RV into a beaker using a small spatula, and the resin was washed three-time with DCM and MeOH, then dried under vacuum for 30 min until reaching dryness. The resin was then collected in a 20 mL scintillation vial and stored in the freezer before cleavage in the next day.




Boc
NH
O
NH
N
OH
O
i,ii,iii,iv,v,vi,vii
Boc
N
H
O
NH
N
O
NH
N
H
N
O
HN
O
O
HN
O
N
HN
O
Boc
NH
H
N
O
N
N
OH
O
i,ii,iii,iv,v,vi,vii
O
NH
N
H
Boc
N
N
O
N
H
O
NH
N
O
NH
N
H
N
O
HN
O
O
HN
O
N
HN
O


Boc
i,ii,iii,iv,v,vi,vii


Boc
i,ii,iii,iv,v,vi,vii
Boc

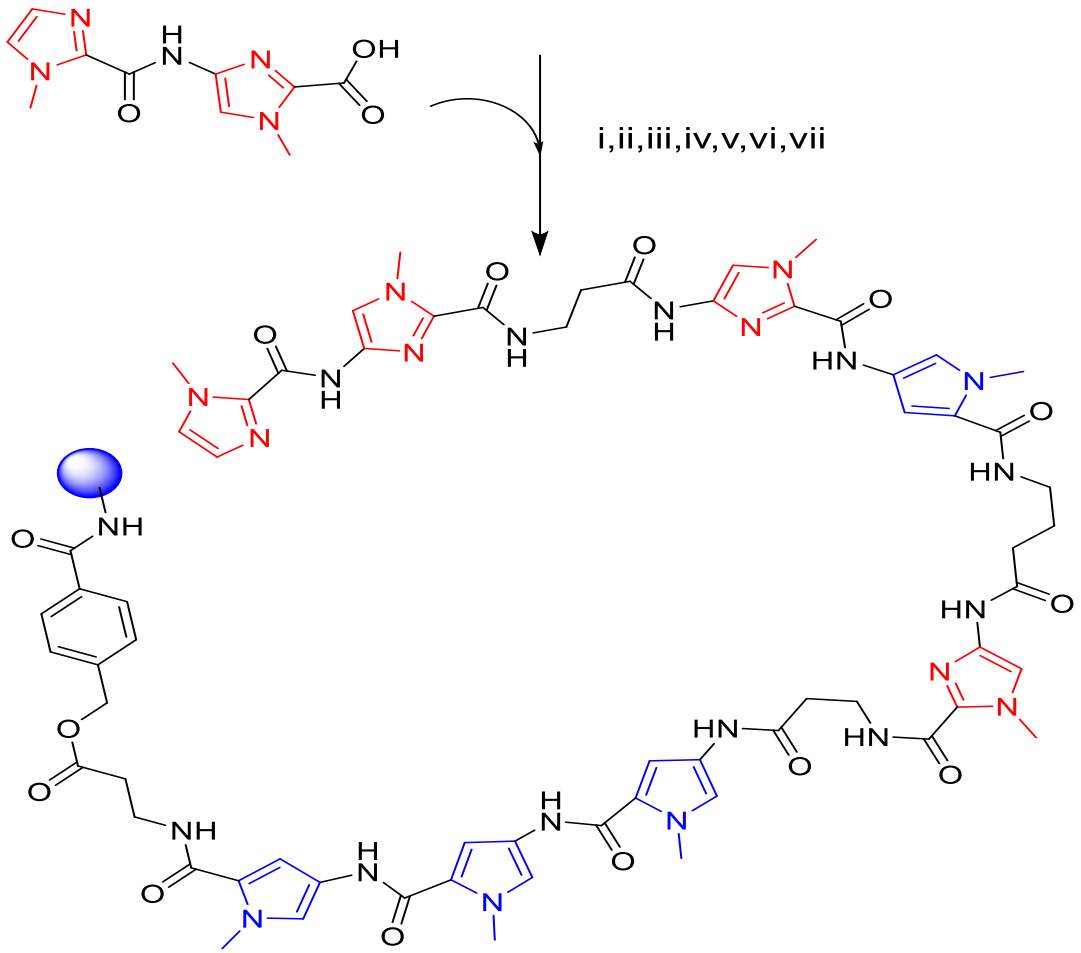
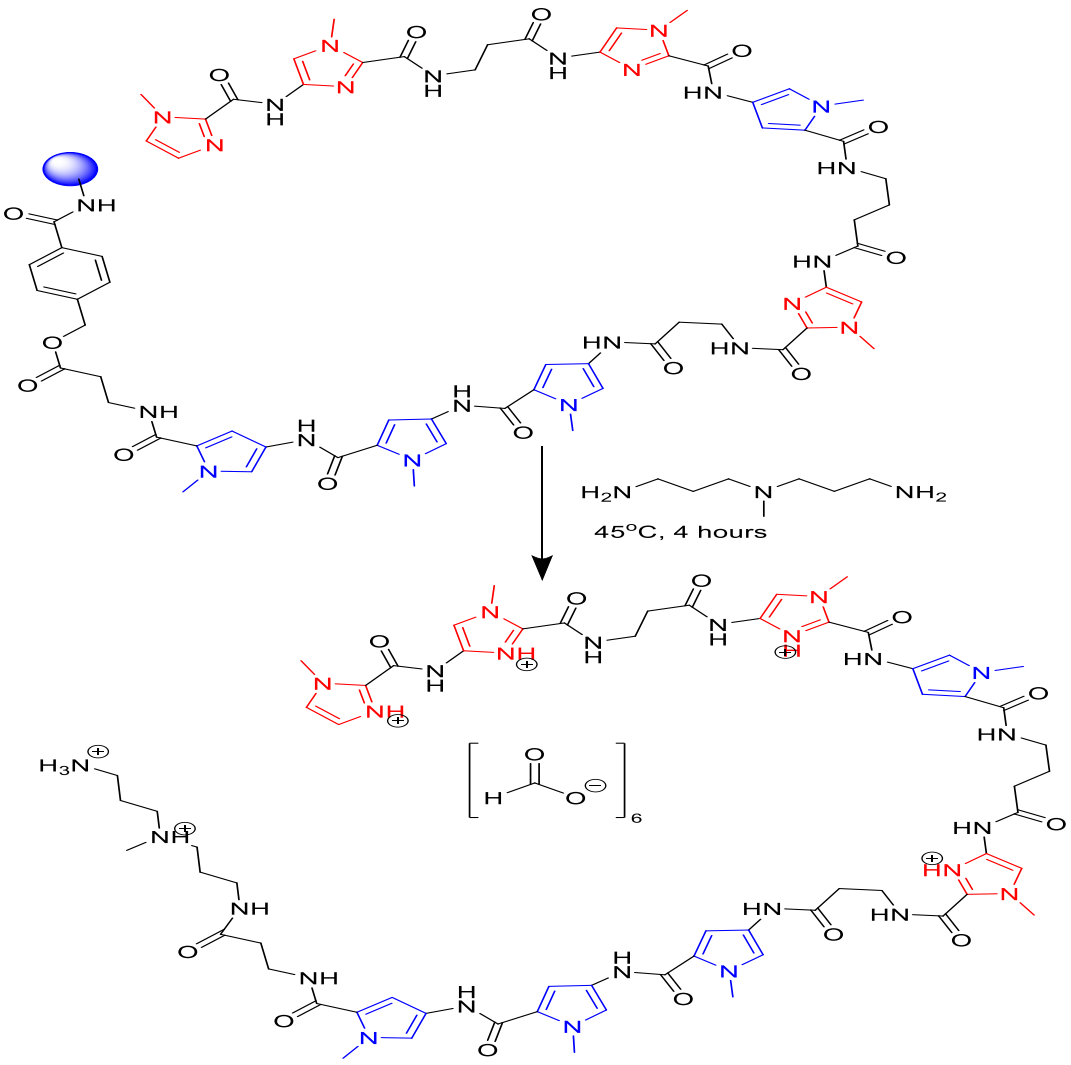


Figure 3.9: *Scheme of synthesis polyamide **1162** resin with solid-phase synthesis method. It started from a commercially acquired Boc-β-PAM resin, and each sequence contains multiple steps from (i) to (viii). (i) wash the resin with DCM and DMF, (ii) deprotection with 60% TFA/ DCM/ 0.5 M Indole, (iii) 25% piperidine/DMF; (iv) building blocks with the structures shown on the scheme,(v) 0.5M PyBOP/ DMF as coupling reagents, (vi) DIPEA, (vii) DMSO, (viii) capping step with acetic anhydride for 17 min.*

Cleavage of PAM resin using Ta as a nucleophile.

The day after, 315 mg resin was cleaved with aminolysis in 2000 µL of Ta at 45ºC in a heating device Isotemp Fisher Scientific for 4 h, and the separation of polyamides solution from the PAM resin support was performed in a 20 µM polyethylene frit placed inside a disposable polypropylene syringe by washing with MeOH and H_2O. The polyamide solution was concentrated in a rotary evaporator to afford off-white solid **1162** (27 mg,72.45 mmol, 25.53% yield). **(Scheme 3.7**).



Scheme 3.7 Cleavage of PA**1162** from PAM resin via Ta as the nucleophile.

Purification of the compound.

The polyamides solution was diluted with a mixture of DMSO (200 µL) and H_2O/0.2% HCOOH (200 µL), then purified by preparative HPLC using a Phenomenex Luna 250x30 mm, 5 µM, 100Å C18 column maintained at 25 °C. The organic phase was 0.2% HCOOH in MilliQ water and 100% HPLC grade MeOH. The applied gradient was 10% MeOH for 8 min, followed by a ramp rate to 90% Methanol over 35.6 min at a 20 mL/min flow rate. All collected fractions were analyzed and selected for 95% or higher purity with analytical HPLC using ThermoFisher - UHPLC; analysis was processed and analyzed with the program Chromeleon 3.0. High-purity fractions were collected for rotary evaporation to remove MeOH from the solution before lyophilization.

Lyophilization of the compound.

Compound **1162** was diluted with a 1:1 mixture of ACN (5 mL): and 0.2% HCOOH in H_2O (5 mL). The samples were frozen before the lyophilization process with ground dry ice until an even coated of frozen material results at an acute angle in the vial. The frozen sample was covered with porous Kimwipe paper and secured by a rubber band. The sample was settled in a lyophilizer glass vessel and connected to a vacuum port. The frozen polyamides sublimed as an off-white solid powder after 48 h. The lyophilizer vacuum was set at 300 torr within 30 min prior to the lyophilization process, and the inside temperature was set at -83 to -85ºC.

Characterization of the compound.

Analytical HPLC characterization was performed with a C12 Phenomenex Juniper Proteo column maintained at 40ºC. The mobile phase consisted of 0.1% formic acid in water(A) and HPLC grade ACN (B). The applied gradient consisted of 5% B for 0.75 min, followed by a ramp to 60% B over 6.5 min at 2.0 mL/min. The retention was 2.563 min.

1162, desIm-Im-β-Im-Py-γ-Im-β-Py-Py-Py-β-Ta $(HCOO^-)_6$.

^{1}H-NMR (600 MHz, DMSO-d_6) δ= 10.32 (s, 1H), 10.29 (s, 1H), 9.952(s, 1H), 9.89 (s, 1H), 9.86-9.83 (d, J= 18.7 Hz, 2H), 9.75 (s, 1H), 8.20 (s, 1H), 8.08 (s, 1H), 7.98 (br, 6H), 7.90 (s, 1H), 7.50 (s, 1H), 7.42- 7.40 (d, J= 17.1 Hz, 2H), 7.38 (s, 1H), 7.20 (s, 1H), 7.17-7.13 (t, J=17.12 Hz, 2H), 7.03 (d, 1H), 6.92 (s, 1H), 6.85 (s, 2H), 3.96 (s, 3H), 3.92- 3.91 (d, J=10.35 Hz, 3H), 3.89 (s, 3H), 3.81 (s, 3H), 3.76- 3.75 (d, J= 9.05 Hz, 6H), 3.49 (s, 3H), 3.36 (s, 3H), 3.15- 3.13 (d, J=4.50 Hz, 4H), 3.09 (s, 2H), 2.94 (s, 1H), 2.86 (br, 6H) 2.73 (s, 4H), 2.71 (s, 4H), 2.60-2.57 (t, J= 15.4 Hz, 2H), 2.48 (s, 6H), 2.34- 2.31 (t, J= 14.7 Hz, 2H), 2.29* 2.27(t, J=14.2 Hz, 2H), 2.52 (br, 2H), 1.93-1.91 (t, J= 12.8 Hz, 4H), 1.21 (s, 1H), 1.15 (s, 1H), 1.75 (s, 3H).

^{13}C-NMR (151 MHz, DMSO-d_6) δ= 174.12, 172.97, 171.70, 171.08, 164.38, 161.62, 161.52, 161.41,158.92,158.82, 140.93, 139.23, 138.95,137.76, 137.26, 137.16, 136.62, 130.69, 130.10, 126.41, 125.92, 125.91, 125.87, 125.27, 125.20, 124.90, 124.18, 121.60, 121.31, 121.06, 119.20, 117.21, 117.20, 117.01, 116.70, 107.43, 107.36, 56.38, 55.3606, 55.3605, 55.3603, 55.20, 48.43, 43.05, 42.91, 42.77, 42.49, 42.35, 42.2, 41.23, 39.29, 39.23, 39.17, 39.16, 39.09, 39.03, 38.69, 38.52, 38.20, 38.17, 38.05, 38.04, 37.99, 35.92, 35.53, 28.58, 27.06, 26.73, 24.89.

HRMS(ESI) calculated for **1162**, $C_{64}H_{84}N_{26}O_{12}$, $[M+H]^+$,1409.68 m/z, found 1409.53 m/z.

HPLC purity:99%

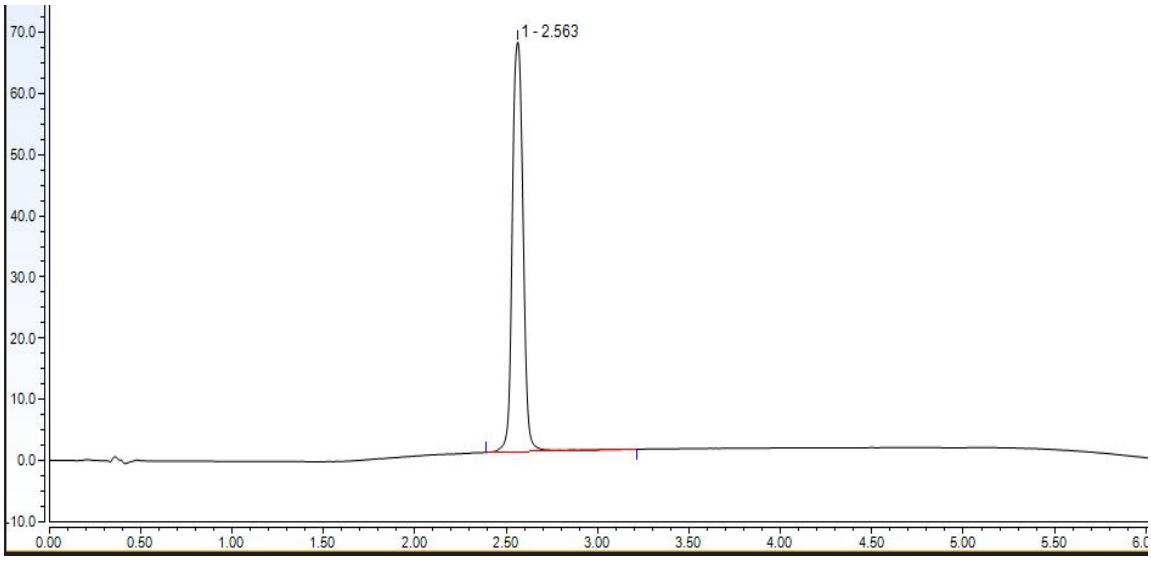


Figure 3.4 Analytical HPLC purity of compound **1162** with λ monitored. Compound purity 99%, retention time 2.563 min.

3.2.6 Synthesis of desIm-Im-β-Im-Py-γ-Im-Py-Py-β-Py-β-Ta

Synthesis of 1166.

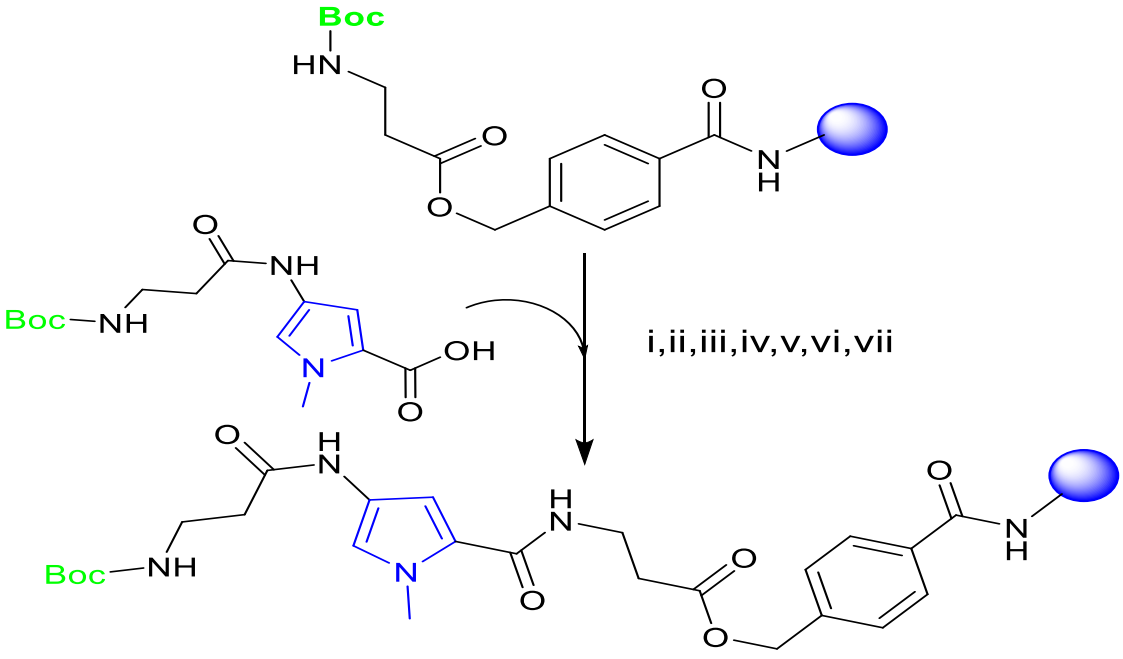
Synthesis of polyamides **1166** resin started by initiating the reaction sequences program with the CSBio program (**Table 3.5)**. The CSBio reaction vessel was set at a temperature of 35ºC throughout the whole process. Boc-β-alanine-PAM resin (1000 mg) was loaded into the RV. Py-BOP (11 g) was dissolved in DMF (40 mL) and filled in the R4 solvent bottle.

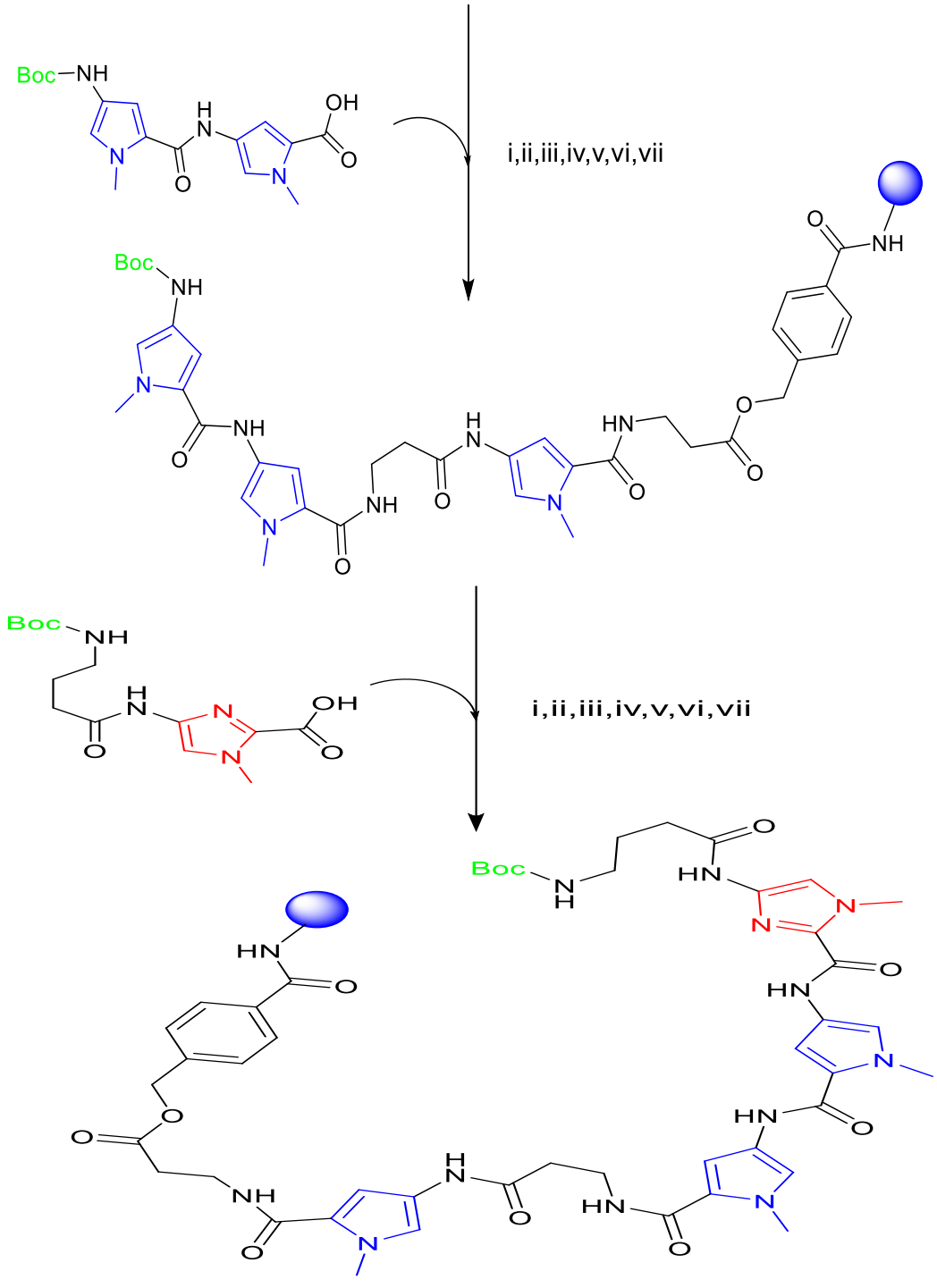
After the program started the first wash and deprotection reaction for 1 h, a solution of a monomer Boc-β-Py-COOH (430 mg, 1.38 mmol) in DMF (6.5 mL) was stirred to fully dissolved and transferred into AA#3. The coupling program directed the synthesizer to transfer the solution from AA#3 to TVA and then MVA to begin the activation before transferring to RV for coupling for 180 min, followed by capping with acetic anhydride for 17 min to remove impurities such as an unreacted amine. After completing the first sequence, the program proceeded to "cap" to remove any nonreacted amine before moving on to the next sequence. Each sequence took around 6 h to complete.

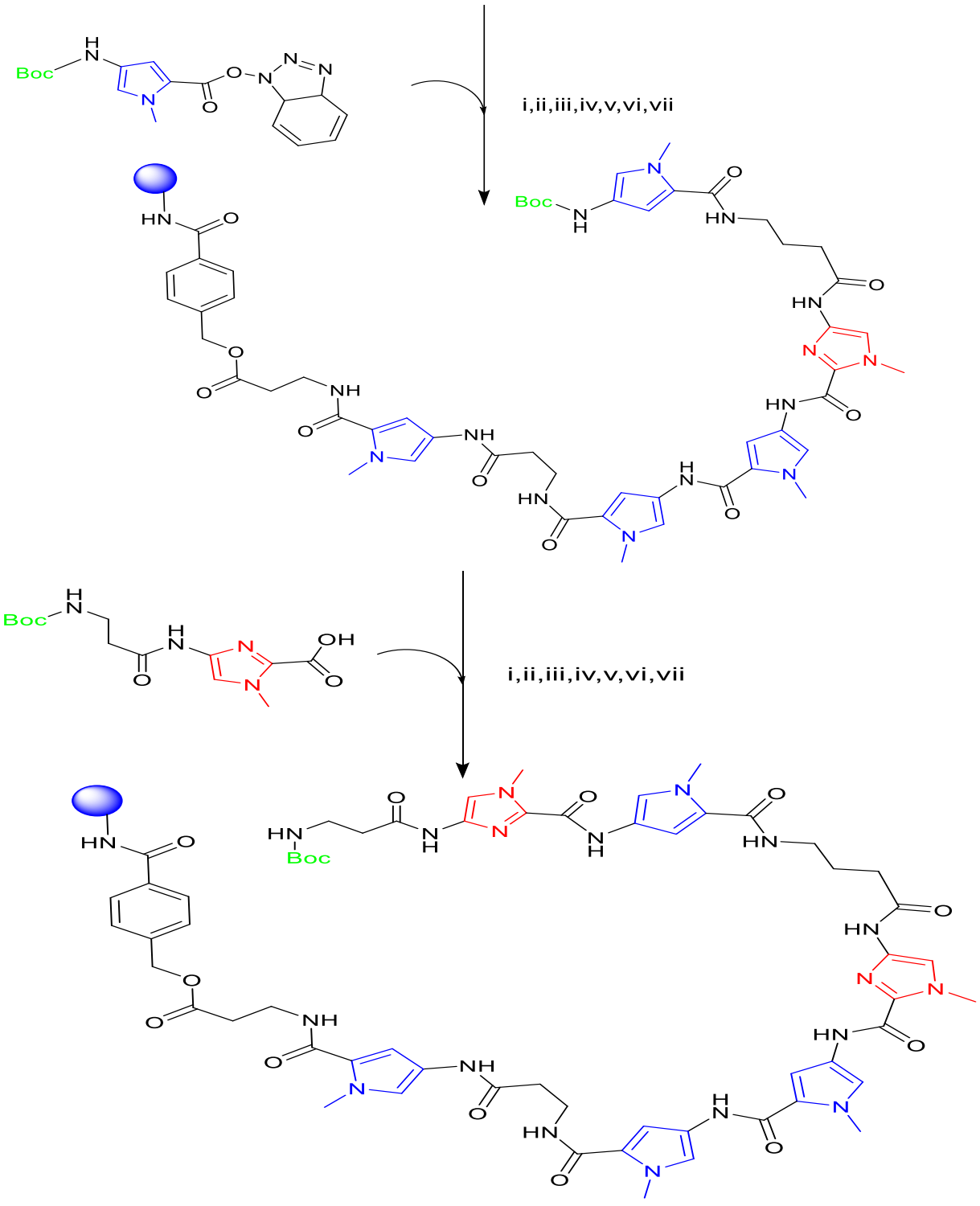
After completing the first sequence, the programmed system transferred a solution of Boc-Py-Py-COOH (503 mg, 1.38 mmol) in DMF (6.5 mL), was well dissolved by vortex and loaded in AA#4 to TVA and MVA for sequence coupling followed up by capping. After sequence 3 plus cap was completed, a solution of Boc-γ-Im-COOH (450 mg, 1.38 mmol) and DIPEA (1 mL) in DMF

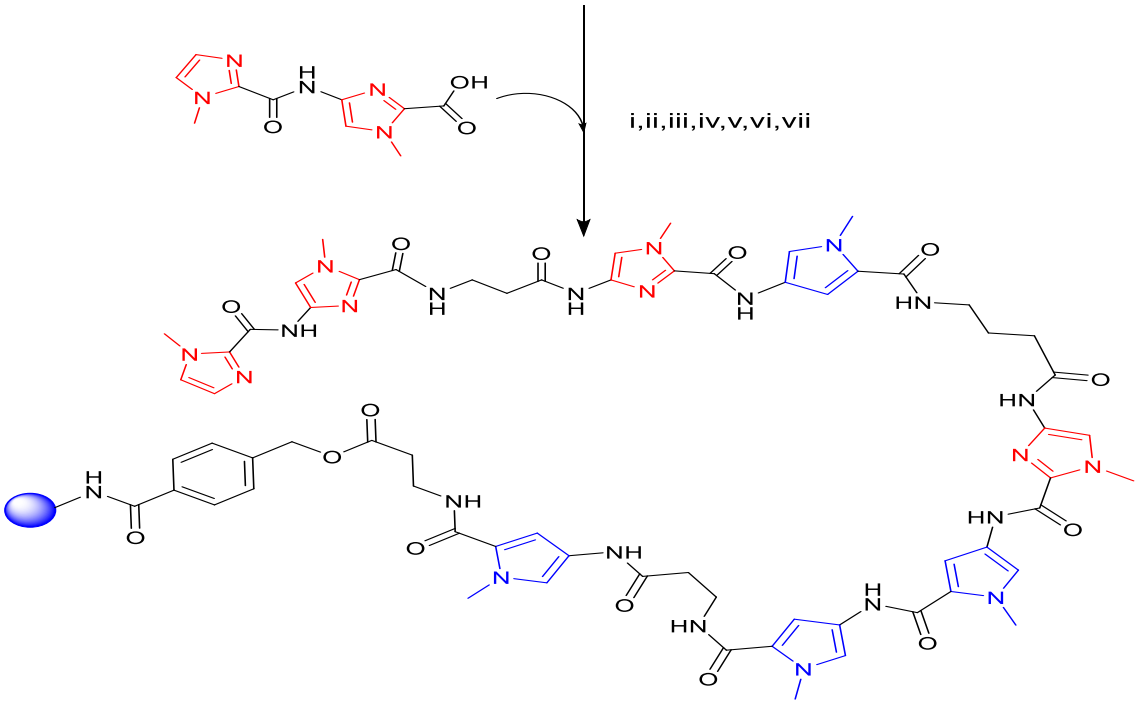
(6.5 mL) was well dissolved by vortex; loaded in AA#5 transferred to TVA then MVA before transferred to RV for coupling reaction for 180 min followed by capping with acetic anhydride for 17 min. After sequence 4 plus cap was completed, a solution of Boc-Py-OBt (659 mg, 1.840 mmol) in DMF (6.5 mL) was well dissolved by vortex; loaded in AA#6, transferred to TVA then MVA before transferred to RV for coupling reaction for 180 min followed by capping with acetic anhydride for 17 min. After sequence 5 was finished, the programed system was initiated and followed by transferring a well dissolved by vortex solution of Boc-β-Im-COOH (432 mg, 1.38 mmol) and DIPEA (1mL) in DMF (6.5 mL) from AA#6 to TVA, then MVA before being transferred to RV for coupling reaction for 180 min following by capping with acetic anhydride for 17 min.

After sequence 6 was finished, the programed system was initiated and followed by transferring a well dissolved by vortex solution of dimer desIm-Im-COOH (345 mg, 1.38 mmol) and DIPEA (1ml) in DMF (6.5 mL) from AA#8 to TVA then MVA before transferred to RV for coupling reaction for 180 min following by capping with acetic anhydride for 17 min. The sequence was finished after 35 h. (**Scheme 3.8**) 0,1 mg of resin was cleaved with 100 µL Ta solution for 2 H at 45°C, and then the heated solution was collected and filtered before being analyzed with a Bruker LC-MS instrument. The exact mass was determined to be $[M+H]^+$ as 1408.73 m/z, which confirmed the presence of the desired product. All the resin was carefully transferred from the RV into a beaker using a small spatula, and the resin was washed three-time with DCM and MeOH, then dried under vacuum for 30 min until reaching dryness. The resin was then collected in a 20mL scintillation vial and stored in the freezer before cleavage in the next day.




Boc
NH
OH
i,ii,iii,iv,v,vi,vii
Boc
NH
i,ii,iii,iv,v,vi,vii
Boc


i,ii,iii,iv,v,vi,vii
i,ii,iii,iv,v,vi,vii



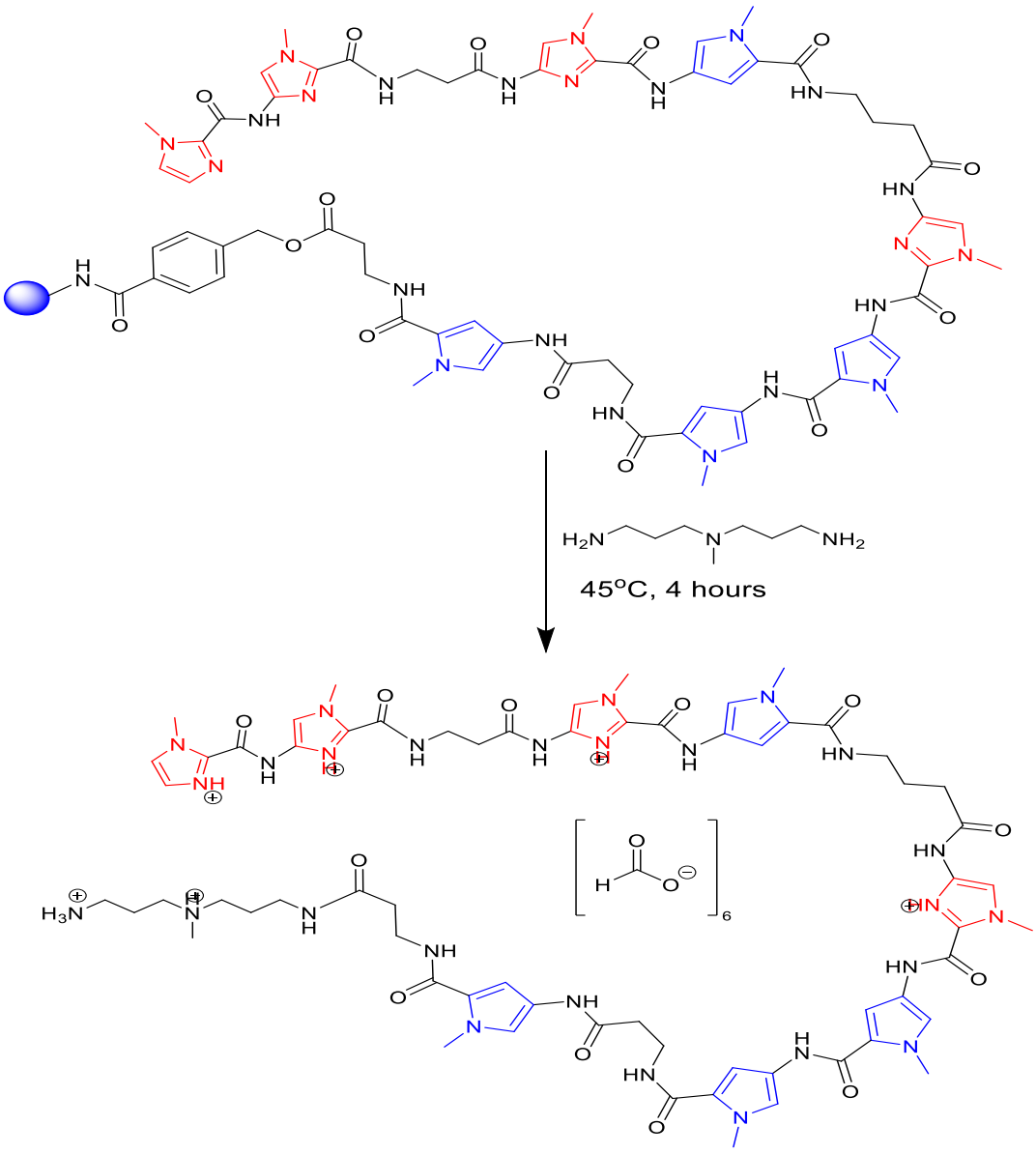
Scheme 3.8: *Scheme of synthesis polyamide **1166** resin with solid phase synthesis. It started from a commercially acquired Boc-β-PAM resin, and each sequence contains multiple steps from (i) to (viii). (i) wash the resin with DCM and DMF, (ii) deprotection with 60% TFA/ DCM/ 0.5 M Indole,(iii) 25% piperidine/ DMF; (iv) building blocks with the structures shown on the scheme, (v) 0.5M PyBOP/ DMF as coupling reagents, (vi) DIPEA, (vii) DMSO, (viii) capping step with acetic anhydride for 17 min.*

Table 3.5. Programmed sequence of **1166** resin synthesis

Building block(s) and step(s)	Additional Solvent mixture	Applied program
1-Boc-β-Py-COOH	(6.5 mL) DMF	Add-DIPEA-180-min-DIPEA-Flows.CFN
2-Capping	None	Cap-17-min-continues.CFN
3-Boc-Py-Py-COOH	(6.5 mL) DMF	Add-DIPEA-180-min-DIPEA-Flows.CFN
4-Capping	None	Cap-17-min-continues.CFN
5-Boc-γ-Im-COOH	(6.5 mL) DMF+DIPEA (1mL)	Add-DIPEA-180-min-DIPEA-Flows.CFN
6-Capping	None	Cap-17-min-continues.CFN
7-Boc-Py-OBt	(6.5 mL) DMF	Add-DIPEA-180-min-DIPEA-Flows.CFN
8-Capping	None	Cap-17-min-continues.CFN
9-Boc β-Im-COOH	(6.5 mL) DMF+DIPEA (1mL)	Add-DIPEA-180-min-DIPEA-Flows.CFN
10-Capping	None	Cap-17-min-continues.CFN
11-desIm-Im-COOH	(6.5 mL) DMF+DIPEA (1mL)	Add-DIPEA-180-min-DIPEA-Flows.CFN
12-Capping	None	Cap-17-min-continues.CFN

Cleavage of PAM resin using Ta as a nucleophile

The next day, 395 mg resin was cleaved with aminolysis in 2000 µL of Ta at 45°C in a heating device Isotemp Fisher Scientific for 4 h, and the separation of polyamides solution from the PAM resin support was performed in a 20 µM polyethylene frit placed inside a disposable polypropylene syringe by washing with MeOH and H_2O. A rotary evaporator concentrated the polyamide solution to afford a white solid **1166** (40 mg, 0.0908 mmol, 31.2% yield) (**Scheme 3.9**).



Scheme 3.9 Cleavage PA**1166** using Ta as the nucleophile.

Purification of the compound.

The polyamides solution was diluted with a mixture of DMSO (200 μL) and H_2O/0.2% HCOOH (200 μL), then purified by preparative HPLC using a Phenomenex Luna 250x30 mm, 5 μM, 100 Å C18 column maintained at 25 ºC. The organic phase was 0.2% HCOOH in MilliQ water and 100% HPLC grade MeOH. The applied gradient was 10% MeOH for 8 min, followed by a ramp to 90% Methanol over 35.6 min at a 20 mL/min flow rate. All collected fractions were analyzed and selected for 95% or higher purity with analytical HPLC using ThermoFisher-UHPLC; analysis was processed and analyzed with the program Chromeleon 3.0. High-purity fractions were collected for rotary evaporation to remove MeOH before the lyophilization.

Lyophilization of the compound.

Compound **1166** was diluted with a 1:1 mixture of ACN (5 mL): and 0.2% HCOOH in H_2O (5 mL). The samples were frozen before the lyophilization process with ground dry ice until an even coated of frozen material results at an acute angle in the vial. The frozen sample was covered with porous Kimwipe paper and secured by a rubber band. The sample was settled in a lyophilizer glass vessel and connected to a vacuum port. The frozen polyamide solution was sublimed as an off-white solid powder after 48 h. The lyophilizer vacuum was set at 300 torr within 30 min prior to the lyophilization process, and the inside temperature was set at -83 to -85ºC.

Characterization of the compound.

Analytical HPLC characterization was performed with a C12 Phenomenex Juniper Proteo column maintained at 40ºC. The mobile phase consisted of 0.1% formic acid in water (A) and HPLC grade ACN (B). The applied gradient consisted of 5% B for 0.75 min, followed by a ramp to 60% B over 6.5 min at 2.0 mL/min. The retention was 2.530 min.

1166, desIm-Im-β-Im-Py-γ-Im-Py-Py-β-Py-β-Ta $(HCOO^-)_6$

^{1}H-NMR (600 MHz, DMSO-d_6) δ= 10.34 (s, 1H), 10.24 (s, 1H), 9.95 (s, 1H), 9.85 (s, 1H), 9.78 (s, 1H), 8.21- 8.18(t, J=12.77 Hz, 1H), 8.13-8.11 (d, 5.95 Hz, 1H), 8.04- 7.99 (d, J= 17.7 Hz, 4H), 7.50 (s, 1H), 7.44-7.40 (t, J= 17.7 Hz, 3H), 7.24 (s, 1H), 7.19 (s, 1H), 7.16 (s, 1H), 7.02 (s, 1H), 6.94 (s, 1H), 6.83 (s, 1H), 6.64 (s, 3H), 3.96 (s, 3H), 3.93 (s, 3H), 3.92 (s, 3H), 3.82 (s, 3H), 3.78 (s, 3H), 3.75 (s, 3H), 3.57 (s, 3H), 3.18- 3.15 (d, J= 8.23 Hz, 4H), 3.09 (s, 4H), 2.95- 2.92 (t, J= 13.7 Hz, 4H), 2.83 (br, 6H), 2.63 (s, 4H), 2.59-5.57 (t, J=13.07 Hz, 3H), 2.53 (s, 2H), 2.51-2.49 (s, J= 11.9 Hz, 2H), 2.37-2.28 (q, J=20.38 Hz, 3H), 1.841(br, 5H), 1.730(s, 5H), 1.23- 1.17 (t, J= 22.07, 2H), 1.03- 0.97 (q, J= 35.55 Hz, 1H).

^{13}C-NMR (151 MHz, DMSO-d_6) δ= 174.0, 173.1, 171.4, 170.9, 166.2, 164.4, 164.3, 161.6, 161.5, 161.4, 161.4, 161.4, 158.9, 158.9, 158.8, 140.9, 139.2, 138.9, 137.8, 137.2, 137.2, 137.1, 130.7, 130.1, 126.4, 126.2, 125.9, 125.2, 124.9, 124.3, 124.2, 121.7, 121.3, 121.1, 120.8, 119.3, 117.2, 117.1, 117.0, 107.9, 106.6, 56.5, 56.1, 56.0, 55.9, 55.9, 55.9, 55.4, 50.6, 50.6, 48.4, 43.5, 43.1, 43.1, 43.1, 43.0, 42.8, 42.7, 42.6, 42.5, 42.3, 42.2, 41.2, 39.7, 39.4, 39.3, 39.2, 39.1, 39.0, 38.9, 38.8, 38.7, 38.6, 38.5, 38.2, 38.2, 38.0, 37.9, 36.4, 35.9, 35.5, 28.6, 28.4, 27.6, 27.3, 27.3, 27.3, 26.7, 25.6, 25.2, 25.1.

HRMS (ESI) was calculated for **1166**, $C_{64}H_{84}N_{26}O_{12}$ $[M+H]^+$ as 1408.68 m/z, found to be 1408.73 m/z.

HPLC purity: 93%

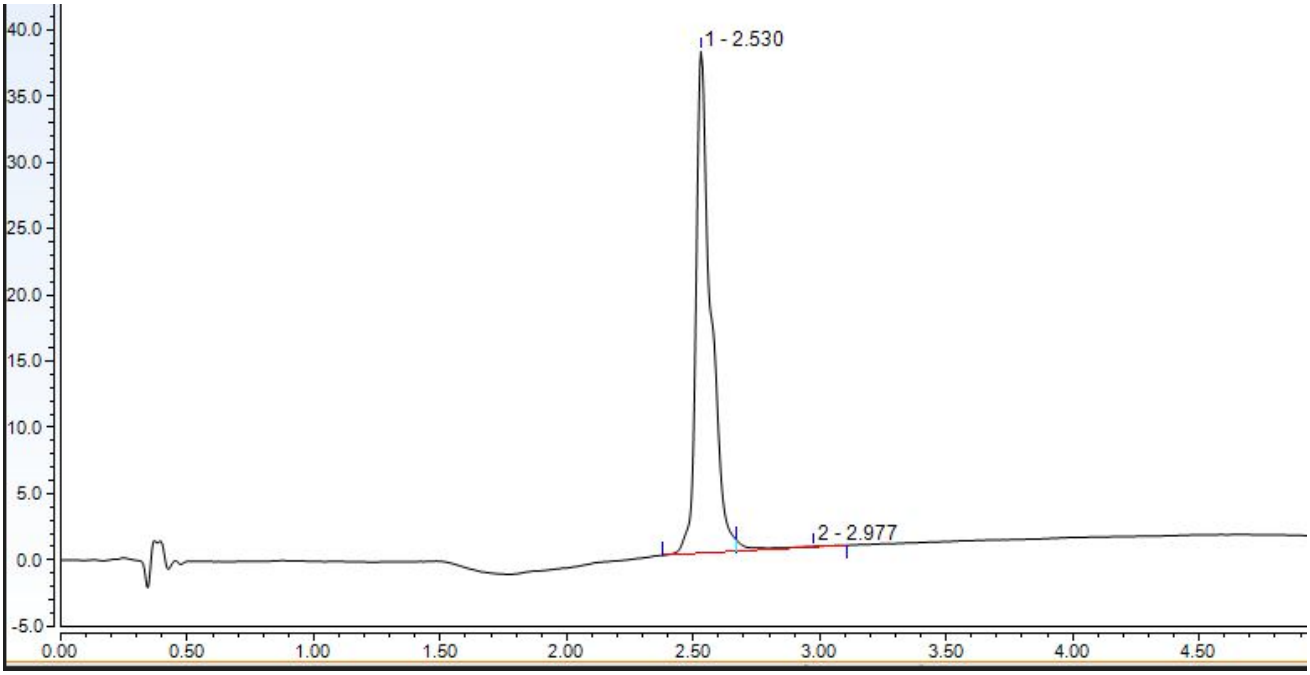


Figure 3.5 Analytical HPLC purity of compound **1166** with λ monitored. Compound purity 93%, retention time 2.530 min.

3.2.7 Synthesis of desIm-Im-β-Im-β-γ-Im-Py-β-Py-Py-β-Ta (1171)

Synthesis of 1171.

Synthesis of polyamides **1171** resin started by initiating the reaction sequences program with the CSBio program (**Table 3.6**). The CSBio reaction vessel was set at a temperature of 35ºC throughout the process. Boc-β-alanine-PAM resin (1000 mg) was loaded into the RV. PyBOP (11 g) was dissolved in DMF (40 mL) and filled in the R4 solvent bottle. After the program started the

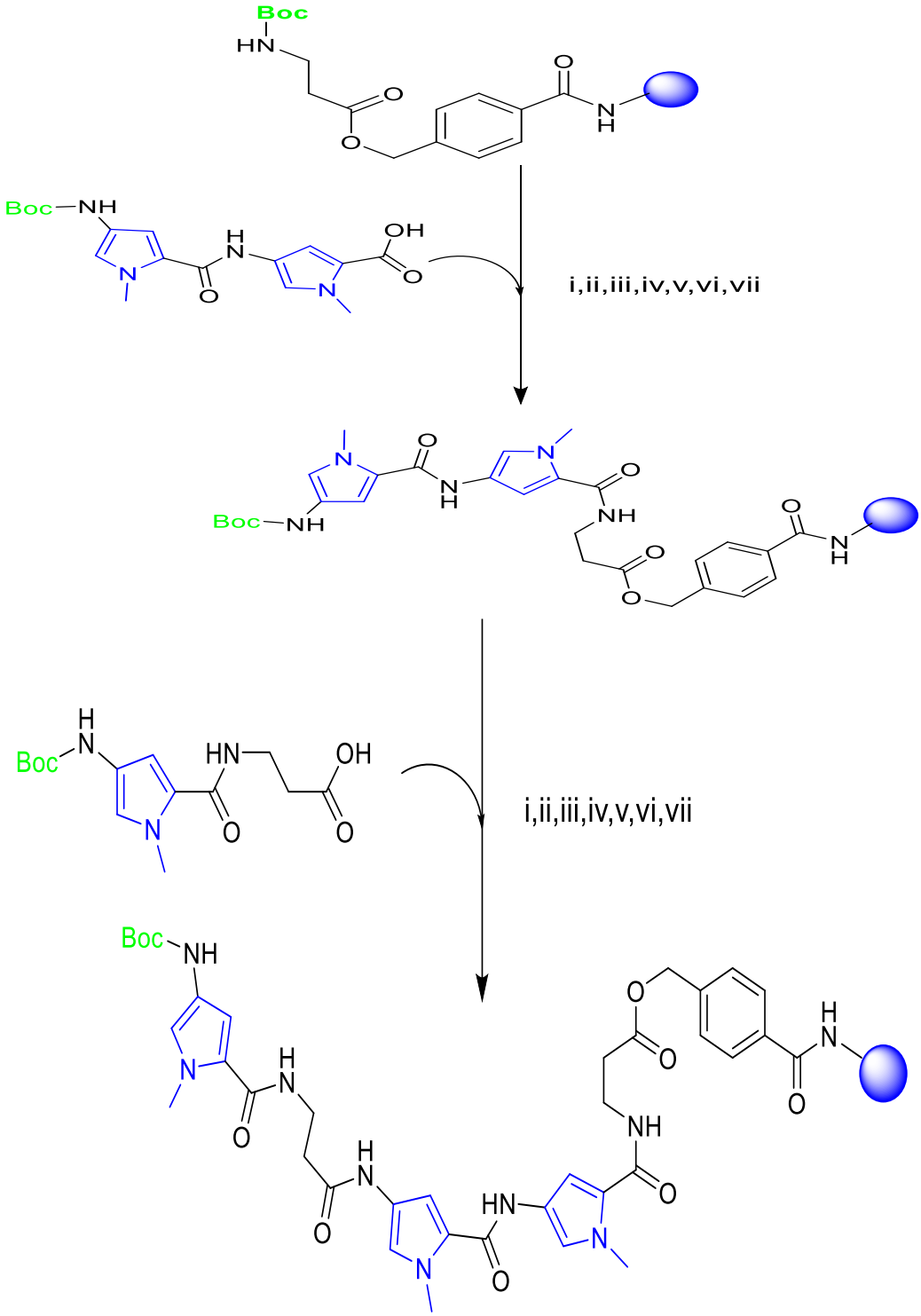
first wash and deprotection reaction for 1 h, a solution of a monomer Boc-β-Py-COOH (431 mg, 1.38 mmol) in DMF (6.5 mL) was stirred to fully dissolved and transferred into AA#3. The coupling program directed the synthesizer to transfer the solution from AA#3 to TVA and then MVA to begin the activation before transferring to RV for coupling for 180 min, followed by capping with acetic anhydride for 17 min to remove impurities such as unreacted amines. After completing the first sequence, the program proceeded to "cap" to remove any nonreacted amine before moving on to the next sequence. Each sequence took around 6 h to complete.

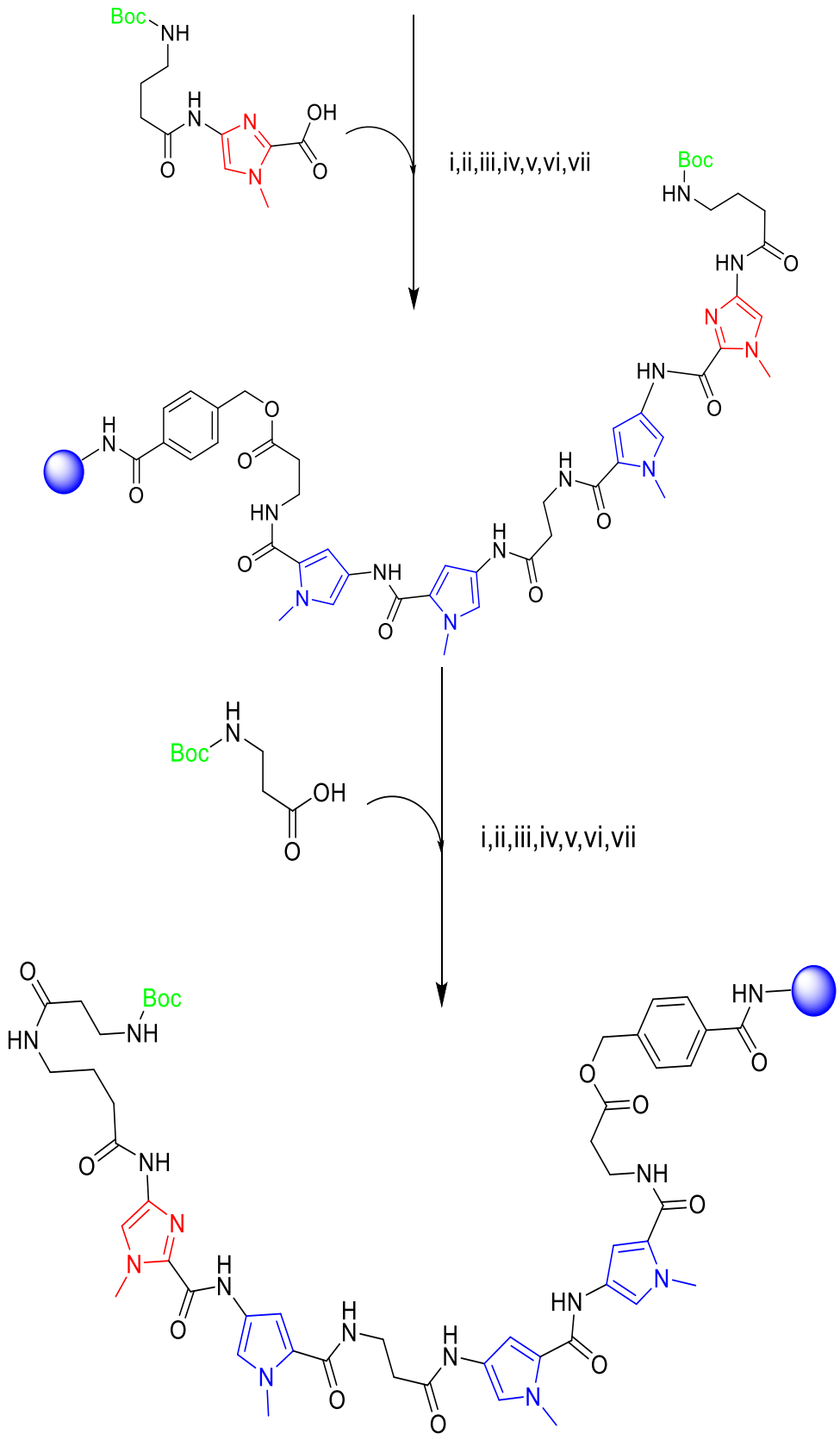
After completing the first sequence, the programmed system transferred a solution of Boc-Py-Py-COOH (500 mg, 1.38 mmol) in DMF (6.5 mL), was well dissolved by vortex and loaded in AA#4 to TVA and MVA for sequence coupling followed up by capping. After sequence 2 plus cap was completed, a solution of Boc-γ-Im-COOH (452 mg, 1.38 mmol) and DIPEA (1mL) in 6.5mL DMF was well dissolved by vortex; loaded in AA#5 transferred to TVA, then MVA before transferred to RV for coupling reaction for 180 min followed by capping with acetic anhydride for 17 min. After sequence 3 plus cap was completed, a solution of Boc-Py-Obt (659 mg, 1840 mmol) in DMF (6.5 mL) was well dissolved by vortex; loaded in AA#6 transferred to TVA then MVA before transferred to RV for coupling reaction for 180 min followed by capping with acetic anhydride for 17 min. After sequence 4 was finished, the programed system was initiated and followed by transferring a well dissolved by vortex solution of Boc-β-Im-COOH (430 mg, 1.38 mmol) and DIPEA (1 mL) in DMF (6.5 mL) from AA#6 to TVA, then MVA before transferred to RV for coupling reaction for 180 min following by capping with acetic anhydride for 17 min. After sequence 5 was finished, the programed system was initiated and followed by transferring a well-dissolved by vortex solution of dimer desIm-Im-COOH (343 mg, 1.38 mmol) and DIPEA (1 mL) in DMF (6.5 mL) from AA#8 to TVA then MVA before transferred to RV for coupling reaction for 180 min following by capping with acetic anhydride for 17 min. The sequence was finished after 30 h (**Scheme 3.10**). 0,1 mg of resin was cleaved with 100 μL Ta solution for 2 h at 45^{o}C; then, the heated solution was collected and filtered before being analyzed with a Bruker LC-MS instrument.

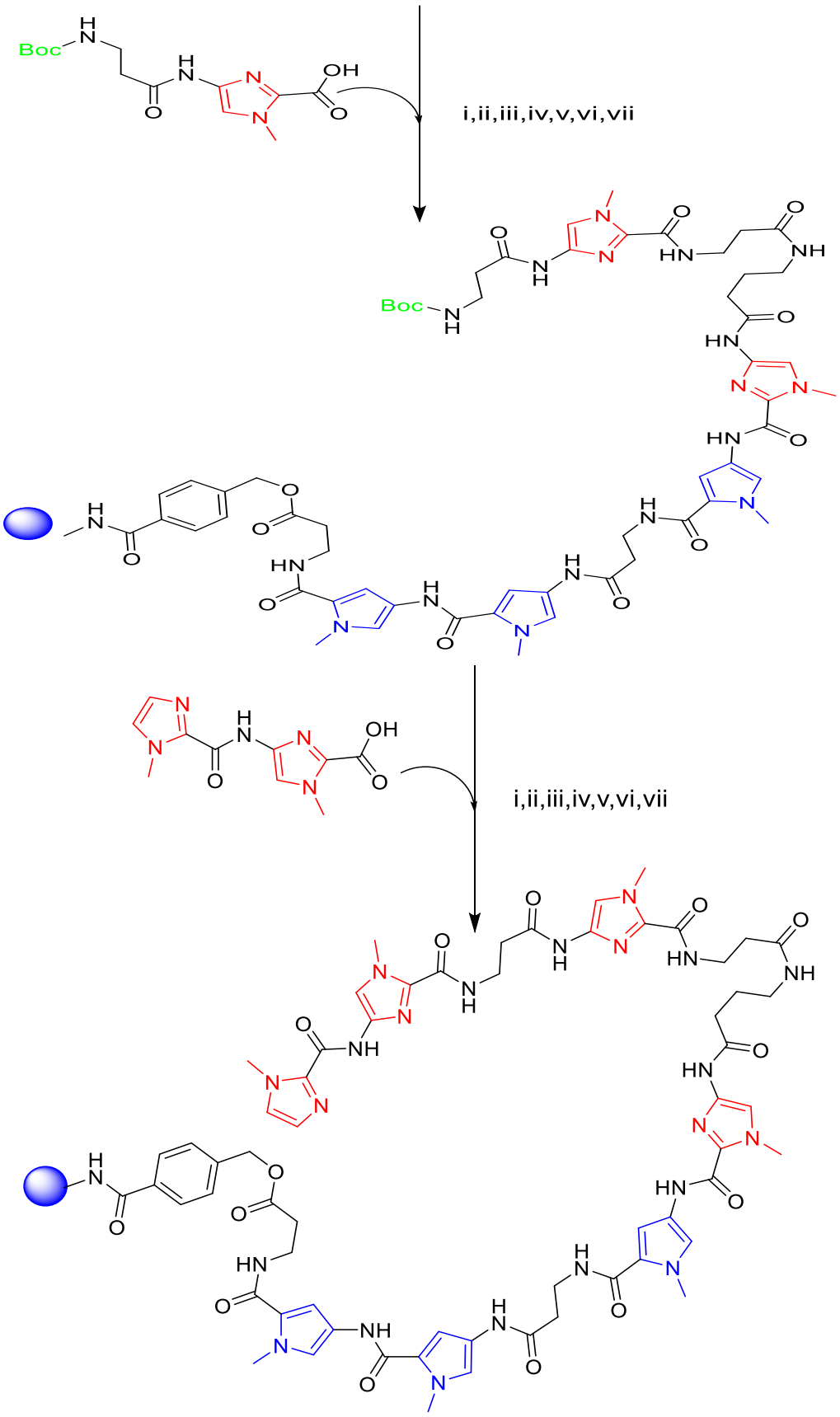
The exact mass was determined to be $[M+H]^+$ as 1358.76 m/z, which confirmed the presence of the desired product. All the resin was carefully transferred from the RV into a beaker using a small spatula, and the resin was washed three times with DCM and MeOH, then dried under vacuum for 30 min until reaching dryness. The resin was then collected in a 20 mL scintillation vial and stored in the freezer before cleavage the next day.

Table 3.6. Programmed sequence of **1171** resin synthesis

Building block(s) and step(s)	Additional Solvent mixture	Applied program
1-Boc-Py-Py-COOH	(6.5 mL) DMF	Add-DIPEA-180-min-DIPEA-Flows.CFN
2-Capping	None	Cap-17-min-continues.CFN
3-Boc-Py-β-COOH	(6.5 mL) DMF	Add-DIPEA-180-min-DIPEA-Flows.CFN
4-Capping	None	Cap-17-min-continues.CFN
5-Boc-γ-Im-COOH	(6.5 mL) DMF + DIPEA (1mL)	Add-DIPEA-180-min-DIPEA-Flows.CFN
6-Capping	None	Cap-17-min-continues.CFN
7-Boc-β-COOH	(6.5 mL) DMF	Add-DIPEA-180-min-DIPEA-Flows.CFN
8-Capping	None	Cap-17-min-continues.CFN
9-Boc β-Im-COOH	(6.5 mL) DMF + DIPEA (1mL)	Add-DIPEA-180-min-DIPEA-Flows.CFN
10-Capping	None	Cap-17-min-continues.CFN
11-desIm-Im-COOH	(6.5 mL) DMF + DIPEA (1mL)	Add-DIPEA-180-min-DIPEA-Flows.CFN
12-Capping	None	Cap-17-min-continues.CFN


Boc
HN
O
O
O
N
H
Boc—NH
N
H
N
O
O
OH
O
i,ii,iii,iv,v,vi,vii
Boc—NH
N
O
N
H
N
O
NH
O
O
O
N
H
H
N
Boc
HN
OH
N
O
O
i,ii,iii,iv,v,vi,vii
Boc
NH
N
H
N
O
O
H
N
O
NH
O
HN
N
O
O
O
N
H
O

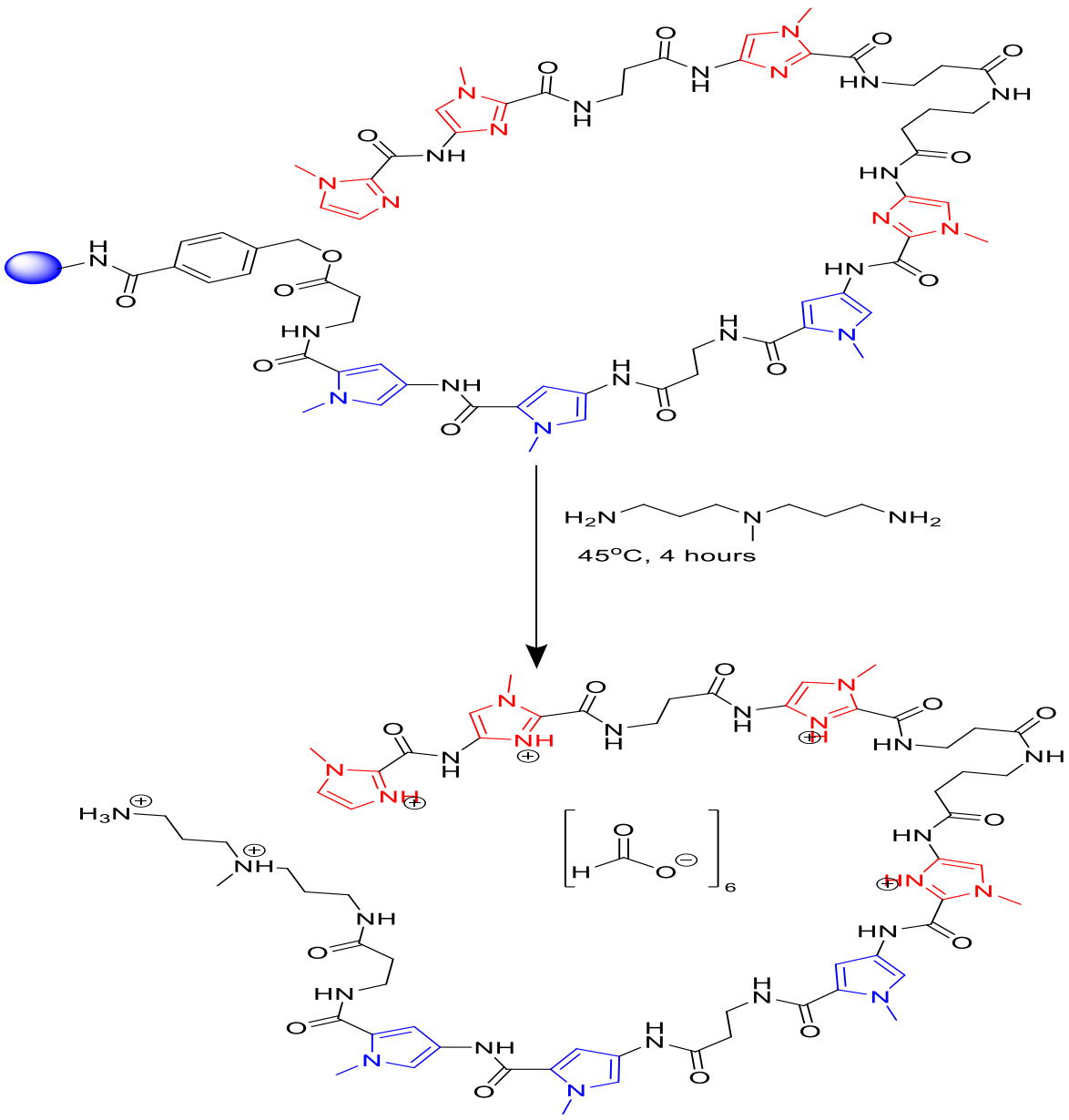

Boc
NH
H
N
N
OH
O
O
N
i,ii,iii,iv,v,vi,vii
Boc
HN
HN
O
N
N
HN
O
H
N
O
O
O
HN
O
N
NH
H
N
O
N
O
Boc
H
N
OH
O
i,ii,iii,iv,v,vi,vii
O
Boc
HN
NH
O
NH
N
N
H
N
O
N
H
N
O
N
HN
O
HN
O
O
O
NH
O
N
HN
O


i,ii,iii,iv,v,vi,vii
i,ii,iii,iv,v,vi,vii

Scheme 3.10 *Scheme of synthesis polyamide* **1171** *resin with solid-phase synthesis method. It started from a commercially acquired Boc-β-PAM resin, and each sequence contains multiple steps from (i) to (viii). (i) wash the resin with DCM and DMF, (ii) deprotection with 60% TFA/ DCM/ 0.5 M Indole, (iii) 25% piperidine/DMF; (iv) building blocks with the structures shown on the scheme,(v) 0.5M PyBOP/ DMF as coupling reagents, (vi) DIPEA, (vii) DMSO, (viii) capping step with acetic anhydride for 17 min.*

Cleavage of PAM resin using Ta as a nucleophile.

The next day, 365 mg resin was cleaved with Ta (2000 μL) via aminolysis at 45°C in a heating device Isotemp Fisher Scientific for 4 h, and the separation of polyamides solution from the PAM resin support was performed in a 20 μM polyethylene frit placed inside a disposable polypropylene syringe by washing with MeOH and H_2O. The polyamide solution was concentrated in a rotary evaporator to afford off-white **1171**(31 mg, 0.836 mmol, 27.1% yield) (**Scheme 3.11**).



Scheme 3.11 Cleavage of PIPs **1171** from PAM resin using Ta as the nucleophile.

Purification of the compound.

The polyamide solution was diluted with a mixture of DMSO (200 µL) and H_2O/0.2% HCOOH (200 µL), then it was purified by preparative HPLC using a Phenomenex Luna 250x30 mm, 5 µM, 100 Å C18 column maintained at 25 °C. The organic phase was 0.2% HCOOH in MilliQ water and 100% HPLC grade MeOH. The applied gradient was 10% MeOH for 8 min, followed by a ramp to 90% Methanol over 35.6 min at a 20 mL/min flow rate. All collected fractions were analyzed and selected for 95% or higher purity with analytical HPLC using ThermoFisher-UHPLC;

analysis was processed and analyzed with the program Chromeleon 3.0. High-purity fractions were collected for rotary evaporation to remove MeOH before the lyophilization.

Lyophilization of compound

Compound **1171** was diluted with a 1:1 mixture of ACN (5 mL): and 0.2% HCOOH in H_2O (5mL). The samples were frozen before the lyophilization process with ground dry ice until an even coated of frozen material results at an acute angle in the vial. The frozen sample was covered with porous Kimwipe paper and secured by a rubber band. The sample was settled in a lyophilizer glass vessel and connected to a vacuum port. The frozen polyamide solution was sublimed into an off-white solid powder after 48 h. The lyophilizer vacuum was set at 300 torr within 30 min prior to the lyophilization process, and the inside temperature was set at -83 to -85ºC.

Characterization of the compound.

Analytical HPLC characterization was performed with a C12 Phenomenex Juniper Proteo column maintained at 40ºC. The mobile phase consisted of 0.1% formic acid in water (A) and HPLC grade ACN (B). The applied gradient consisted of 5% B for 0.75 min, followed by a ramp to 60% B over 6.5 min at 2.0 mL/ min. The retention was 3.17 min.

PA **1171**, dIm-Im-β-Im- β-γ-Im-Py-β-Py-Py-β-Ta $(HCOO^-)_6$

^{1}H-NMR (600 MHz, DMSO-d_6) δ= 10.48 (s, 1H), 10.28 (s, 1H), 10.04 (s, 1H), 9.91 (s, 1H), 8.41 (s, 9H), 8.21- 8.18 (t, J= 12.92 Hz, 1H), 8.14- 8.10 (t, J=12.92, 1H), 8.01-7.96 (d, J= 11.07, 4H), 7.49 (s, 1H), 7.41-7.37 (t, J=5.54 Hz, 3H), 7.21 (s, 1H), 7.16 (s, 1H), 7.04 (s, 1H), 6.90 (s, 1H), 3.84 (s, 1H), 3.96 (s, 3H), 3.93 (s, 3H), 3.90 (s, 3H), 3.89 (s, 3H), 3.79 (s, 4H), 3.76 (s, 2H), 3.46-3.44 (q, J= 16.61 Hz, 2H), 3.41- 3.37 (q, J= 16.61 Hz, 2H), 3.03 (br, 3H), 2.76 (br, 5H), 2.56-5.23 (t, J=12.46 Hz, 3H), 2.47 (br, 3H), 2.23 (br ,2H), 2.07 (s, 4H), 2.05 (s, 2H), 1.66 (br, 9H), 1.49-1.47(t, J= 13.36 Hz, 1H), 1.21 (s, 1H), 1.16 (s, 2H).

^{13}C-NMR (151 MHz, DMSO-d_6) δ= 173.7, 173.6, 172.9, 171.6, 170.9, 169.3, 164.3, 164.3, 164.5, 161.5, 161.4, 161.4, 161.3, 158.9, 158.8, 140.9, 139.2, 139.0, 137.8, 137.2, 137.0, 136.7, 130.7, 130.7, 130.1, 130.1, 126.2, 125.9, 125.2, 125.1, 24.2, 121.3, 121.2, 116.9, 116.9, 116.9, 116.9, 116.7, 116.7, 116.7, 107.4, 107.2, 107.0, 57.6, 57.5, 57.1, 44.5, 44.4, 44.2, 43.0, 42.2, 42.3, 42.5, 42.6, 42.7, 42.9, 40.6, 40.2, 39.7, 39.2, 38.6, 38.6, 38.2, 38.2, 38.1, 38.0, 37.9, 37.9, 37.8, 35.8, 29.8, 28.4, 27.9.

HRMS (ESI) was calculated for **1171**, $C_{61}H_{83}N_{25}O_{12}$ with predicted $[M+H]^+$, 1358.68 m/z, found 1358.76 m/z.

HPLC purity of 1171: 94%

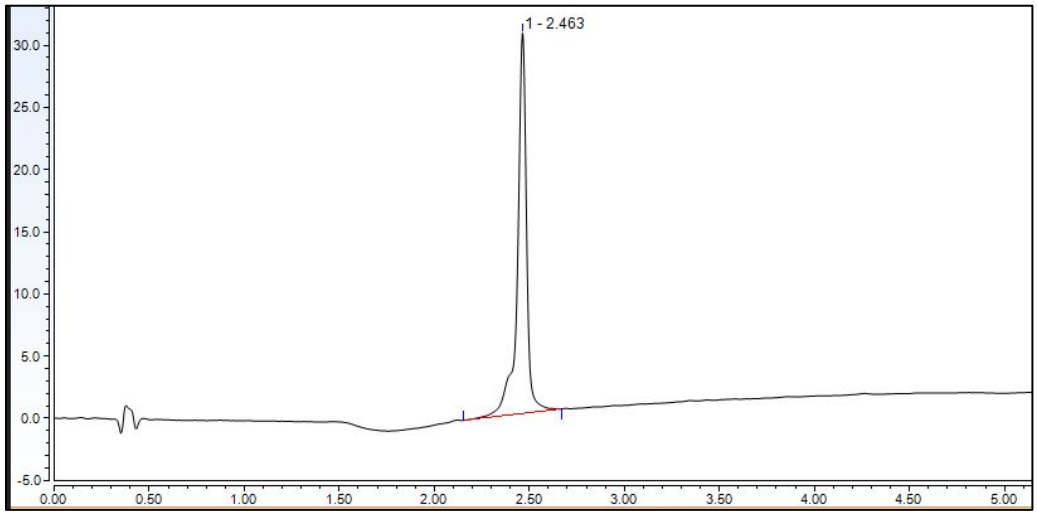


Figure 3.6 Analytical HPLC purity of compound **1171** with λ monitored. Compound purity 94%, retention time 2.463 min.

3.3. Results and Discussion.

In this chapter, we have redesigned and modified our polyamides based on the result of the previous project. We included PA**1143** to compare its binding affinity with the new PIPs to observe the role of β-alanine in the binding complex. As expected, PA **1134** became aggregated and unable to bind to the SETMAR DNA sequence, which appeared to happen to compound **1162** as well.

Both **1134** and **1162** will require further investigation in the modification of the compounds and the SPR study model for the study of the binding affinities. However, the binding affinity of the other PIPs gave us interesting and puzzling results. While **1161** has been shown to bind strongly with the SETMAR sequence and even the mismatch SETMAR sequence, it's been noted to us by our collaborators that the compound has very low to almost no selectivity in binding (**Figure 3.7**).

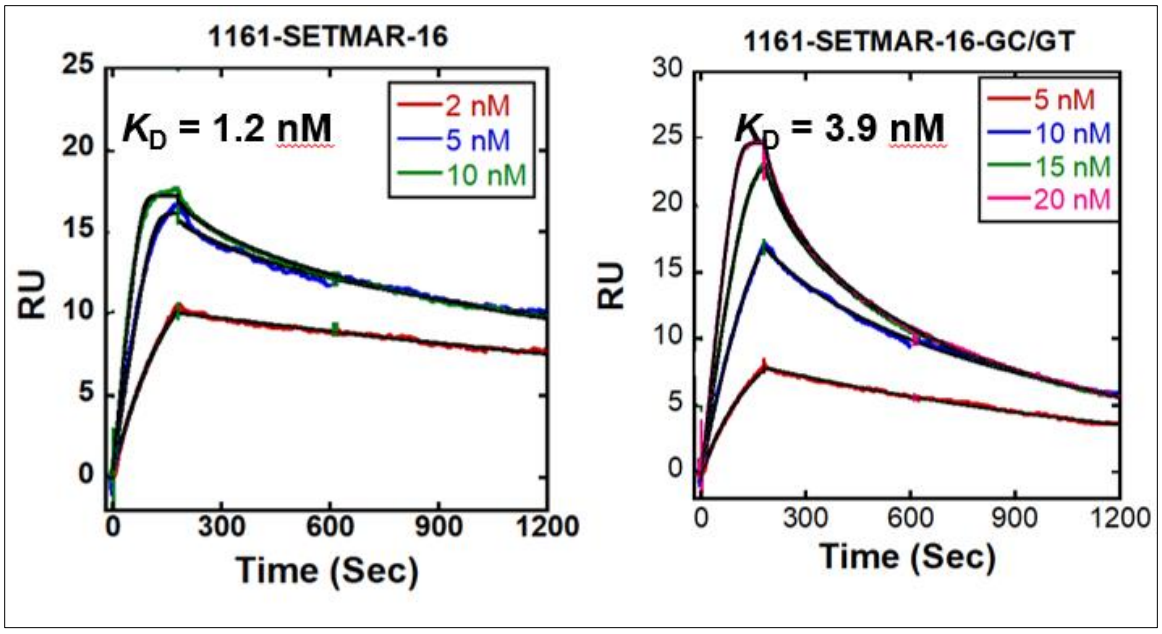


Figure 3.7: Binding affinity study of PA **1161.**

PA **1166** seems to have the most interesting result; despite the aggregation when introduced with a longer sequence of SETMAR, **1166** is shown to bind strongly with the promoter of the shorter SETMAR-16 sequence. The structure of PA **1166** ensures the compound has a much lesser rigidity by substituting β-alanines for two pyrrole molecules in the PIP sequence. Additionally, the SPR result was reported to have a high selectivity toward binding the promoter sequence of SETMAR (**Figure 3.8).**

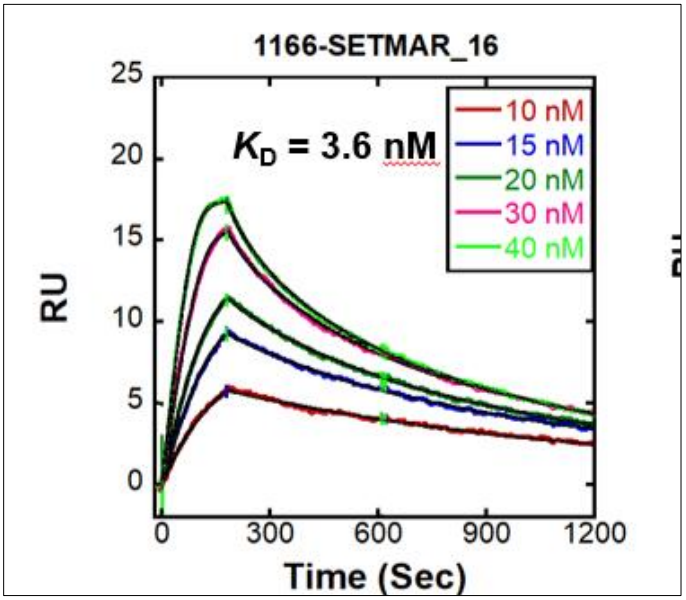


Figure 3.8: Binding affinity study of **1166.**

In this project, we also attempted to increase the flexibility of the PIPs by designing a polyamide compound with two β-alanine in between an imidazole. Hence the structure of PA **1171:** desIm-Im-β-Im-β-γ-Im-Py-β-Py-Py-β-Ta, as we aimed to increase the flexibility and increase the binding affinity and selectivity of the designed PIPs. As expected, the compound was highly flexible, and the compound provided clean sensorgram reading of the SPR study. However, this PIP has no sequence selectivity. It was assumed that more than three β-alanines in the sequence might have caused this problem. (**Figure 3.9**)

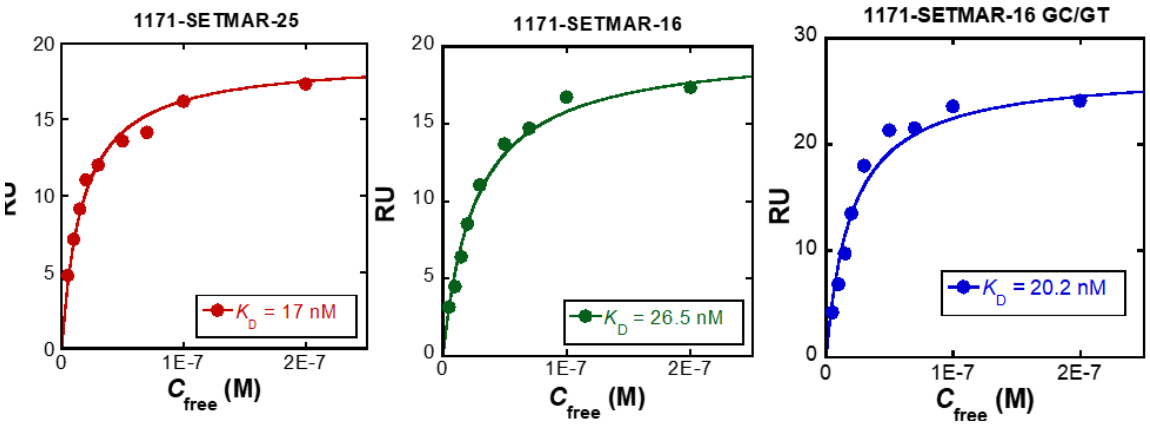


Figure 3.9. Binding affinity study of PA **1171.**

3.4. Conclusion.

In this chapter, we synthesized several compounds with one or two β-alanines substituting for pyrroles to investigate the binding affinity of PIPs and the TIR DNA of the SETMAR gene. In addition, we also discovered that the relaxation of the sequence rigidity increases the yield of the solid phase synthesis post-isolation and purification, as mentioned in our lab's previous works.[79,80]. The binding affinity study results provided us with a deeper understanding of SETMAR TIR DNA: polyamide binding.

We learned that reducing the rigidity of the PIP sequence can increase the binding affinity and selectivity with SETMAR binding; despite some aggregation of the polyamide with longer SETMAR DNA sequences. For every two to three aromatic rings of imidazoles or pyrroles, a β-alanine linkage can reduce the rigidity can increase the binding affinity and selectivity of the SETMAR promoter sequence. However, the presence of more than three β-alanines linkages can be too flexible for binding selectivity. Further studies will be needed to improve and discover SETMAR TIR binding.

3.5. Bibliography

1. Wemmer DE, Dervan PB. Targeting the minor groove of DNA. *Curr Opin Struct Biol*. 1997;7(3):355-361. doi:https://doi.org/10.1016/S0959-440X(97)80051-6

2. Kopka ML, Yoon C, Goodsell D, Pjura P, Dickerson RE. The molecular origin of DNA-drug specificity in netropsin and distamycin. *Proc Natl Acad Sci U S A*. 1985;82(5):1376-1380. doi:10.1073/pnas.82.5.1376

3. Cabeen M, Aaron G, Christine H, et al. *DNA Structure & Chemistry*. Harvard University; 2020. https://projects.iq.harvard.edu/lifesciences1abookv1

4. Blackledge MS, Melander C. Programmable DNA-binding small molecules. *Bioorganic Med Chem*. 2013;21(20):6101-6114. doi:10.1016/j.bmc.2013.04.023

5. White S, Szewczyk JW, Turner JM, Baird EE, Dervan PB. Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands. *Nature*. 1998;391(6666):468-471. doi:10.1038/35106

6. Bailly C, Chaires JB. Sequence-Specific DNA Minor Groove Binders. Design and Synthesis of Netropsin and Distamycin Analogues. *Bioconjug Chem*. 1998;9(5):513-538. doi:10.1021/bc980008m

7. Coll M, Frederick CA, Wang AH, Rich A. A bifurcated hydrogen-bonded conformation in the d(A.T) base pairs of the DNA dodecamer d(CGCAAATTTGCG) and its complex with distamycin. *Proc Natl Acad Sci U S A*. 1987;84(23):8385-8389. doi:10.1073/pnas.84.23.8385

8. Fechter EJ, Olenyuk B, Dervan PB. Sequence-specific fluorescence detection of DNA by polyamide thiazole orange conjugates. *J Am Chem Soc*. 2005;127(47):16685-16691. doi:10.1021/ja054650k

9. Kopka ML, Yoon C, Goodsell D, Pjura P, Dickerson RE. Binding of an Antitumour Drug to DNA Netropsin and CGCGAATTCGCG. Published online 1985.

10. Alniss HY. Thermodynamics of DNA Minor Groove Binders. *J Med Chem*. 2019;62(2):385-402. doi:10.1021/acs.jmedchem.8b00233

11. Peter B. Dervan. Molecular Recognition of DNA by small molecule. 2001;9(July):2-57.

12. J.William L, Krowicki K, Bhat G. Molecular recognition between oligopeptidesand nucleic acids: novel imidazole-containing oligopeptides related tonetropsin that exhibit altered DNA sequence specificity. *Biochemistry*. 1986;25:7408–7416. doi:10.1002/(SICI)1099-1409(200006/07)4:4<393::AID-JPP227>3.3.CO;2-2

13. Design of Sequence-Specific DNA-Binding Molecules Author (s): Peter B . Dervan Published by : American Association for the Advancement of Science Stable URL : http://www.jstor.org/stable/1696414. *Am Assoc Adv Sci*. 1986;232(4749):464-471.

14. Mrksich M, Wade WS, Dwyer TJ, Geierstanger BH, Wemmer DE, Dervan PB. Antiparallel side-by-side dimeric motif for sequence-specific recognition in the minor groove of DNA by the designed peptide 1-methylimidazole-2-carboxamide netropsin. *Proc Natl Acad Sci U S A*. 1992;89(16):7586-7590. doi:10.1073/pnas.89.16.7586

15. Wade WS, Mrksich M, Dervan PB. Design of Peptides That Bind in the Minor Groove of DNA at 5′-(A,T)G(A,T)C(A,T)-3′ Sequences by a Dimeric Side-by-Side Motif. *J Am Chem Soc*. 1992;114(23):8783-8794. doi:10.1021/ja00049a006

16. Koeller KJ. DNA Binding Polyamides and the Importance of DNA Recognition in their use as Gene-Specific and Antiviral Agents. *Med Chem (Los Angeles)*. 2014;04(02):338-344. doi:10.4172/2161-0444.1000162

17. Castaneda CH, Scuderi MJ, Edwards TG, et al. Improved antiviral activity of a polyamide against high-risk human papillomavirus: Via N-terminal guanidinium substitution. *Medchemcomm*. 2016;7(11):2076-2082. doi:10.1039/c6md00371k

18. Song Y, Niederschulte J, Bales KN, Park AH, Bashkin JK, Dupureur CM. DNA binding thermodynamics and site stoichiometry as a function of polyamide size. *Biochimie*. 2019;165:170-178. doi:10.1016/j.biochi.2019.07.021

19. Trauger JW, Baird EE, Dervan PB. Extended hairpin polyamide motif for sequence-specific recognition in the minor groove of DNA. *Chem Biol*. 1996;3(5):369-377. doi:10.1016/S1074-5521(96)90120-9

20. White S, Baird EE, Dervan PB. On the pairing rules for recognition in the minor groove of DNA by pyrrole-imidazole polyamides. *Chem Biol*. 1997;4(8):569-578. doi:10.1016/S1074-5521(97)90243-X

21. Han YW, Kashiwazaki G, Morinaga H, et al. Effect of single pyrrole replacement with β-alanine on DNA binding affinity and sequence specificity of hairpin pyrrole/imidazole polyamides targeting 5'-GCGC-3'. *Bioorg Med Chem*. 2013;21(17):5436-5441. doi:10.1016/j.bmc.2013.06.005

22. Kawamoto Y, Bando T, Sugiyama H. Sequence-specific DNA binding Pyrrole–imidazole polyamides and their applications. *Bioorganic Med Chem*. 2018;26(8):1393-1411. doi:10.1016/j.bmc.2018.01.026

23. Neely L, Trauger JW, Baird EE, Dervan PB, Gottesfeld JM. Importance of minor groove binding zinc fingers within the transcription factor IIIA-DNA complex. *J Mol Biol*. 1997;274(4):439-445. doi:10.1006/jmbi.1997.1411

24. Dickinson LA, Gulizia RJ, Trauger JW, et al. Inhibition of RNA polymerase II transcription in human cells by synthetic DNA-binding ligands. *Proc Natl Acad Sci U S A*. 1998;95(22):12890-12895. doi:10.1073/pnas.95.22.12890

25. Olenyuk BZ, Zhang GJ, Klco JM, Nickols NG, Kaelin WG, Dervan PB. Inhibition of vascular endothelial growth factor with a sequence-specific hypoxia response element antagonist. *Proc Natl Acad Sci U S A*. 2004;101(48):16768-16773. doi:10.1073/pnas.0407617101

26. Mysore VS, Szablowski J, Dervan PB, Frost PJ. A DNA-binding Molecule Targeting the Adaptive Hypoxic Response in Multiple Myeloma Has Potent Antitumor Activity. *Mol Cancer Res*. 2016;14(3):253-266. doi:10.1158/1541-7786.MCR-15-0361

27. Yao EH, Fukuda N, Ueno T, et al. A pyrrole–imidazole polyamide targeting transforming growth factor-β1 inhibits restenosis and preserves endothelialization in the injured artery. *Cardiovasc Res*. 2008;81(4):797-804. doi:10.1093/cvr/cvn355

28. Martínez TF, Phillips JW, Karanja KK, et al. Replication stress by Py–Im polyamides induces a non-canonical ATR-dependent checkpoint response. *Nucleic Acids Res*. 2014;42(18):11546-11559. doi:10.1093/nar/gku866

29. Schaal TD, Mallet WG, McMinn DL, et al. Inhibition of human papilloma virus E2 DNA binding protein by covalently linked polyamides. *Nucleic Acids Res*. 2003;31(4):1282-

1291. doi:10.1093/nar/gkg206

30. Edwards TG, Vidmar TJ, Koeller K, Bashkin JK, Fisher C. DNA Damage Repair Genes Controlling Human Papillomavirus (HPV) Episome Levels under Conditions of Stability and Extreme Instability. *PLoS One*. 2013;8(10). doi:10.1371/journal.pone.0075406

31. Oyoshi T, Kawakami W, Narita A, Bando T, Sugiyama H. Inhibition of transcription at a coding sequence by alkylating polyamide. *J Am Chem Soc*. 2003;125(16):4752 – 4754. doi:10.1021/ja029196o

32. Shinohara K ichi, Sasaki S, Minoshima M, Bando T, Sugiyama H. Alkylation of template strand of coding region causes effective gene silencing. *Nucleic Acids Res*. 2006;34(4):1189-1195. doi:10.1093/nar/gkl005

33. Jespersen C, Soragni E, James Chou C, Arora PS, Dervan PB, Gottesfeld JM. Chromatin structure determines accessibility of a hairpin polyamide–chlorambucil conjugate at histone H4 genes in pancreatic cancer cells. *Bioorg Med Chem Lett*. 2012;22(12):4068-4071. doi:https://doi.org/10.1016/j.bmcl.2012.04.090

34. Lin J, Hiraoka K, Watanabe T, et al. Identification of binding targets of a pyrrole-imidazole polyamide KR12 in the LS180 colorectal cancer genome. *PLoS One*. 2016;11(10):1-19. doi:10.1371/journal.pone.0165581

35. Wang YD, Dziegielewski J, Wurtz NR, Dziegielewska B, Dervan PB, Beerman TA. DNA crosslinking and biological activity of a hairpin polyamide-chlorambucil conjugate. *Nucleic Acids Res*. 2003;31(4):1208 – 1215. doi:10.1093/nar/gkg215

36. Wicker T, Sabot F, Hua-Van A, et al. Reply: A unified classification system for eukaryotic transposable elements should reflect their phylogeny. *Nat Rev Genet*. 2009;10(4):276. doi:10.1038/nrg2165-c4

37. Ravindran S. Barbara McClintock and the discovery of jumping genes. *Proc Natl Acad Sci U S A*. 2012;109(50):20198-20199. doi:10.1073/pnas.1219372109

38. McCLINTOCK B. The origin and behavior of mutable loci in maize. *Proc Natl Acad Sci U S A*. 1950;36(6):344-355. doi:10.1073/pnas.36.6.344

39. Finnegan DJ. Eukaryotic transposable elements and genome evolution. *Trends Genet*. 1989;5(C):103-107. doi:10.1016/0168-9525(89)90039-5

40. Robertson HM. The Tcl-mariner superfamily of transposons in animals. *J Insect Physiol*. 1995;41(2):99-105. doi:https://doi.org/10.1016/0022-1910(94)00082-R

41. Sinzelle L, Izsvák Z, Ivics Z. Molecular domestication of transposable elements: From detrimental parasites to useful host genes. *Cell Mol Life Sci*. 2009;66(6):1073-1093. doi:10.1007/s00018-009-8376-3

42. Ågren JA, Clark AG. Selfish genetic elements. *PLOS Genet*. 2018;14(11):e1007700. doi:10.1371/journal.pgen.1007700

43. Feschotte C, Pritham EJ. DNA transposons and the evolution of eukaryotic genomes. *Annu Rev Genet*. 2007;41(35):331-368. doi:10.1146/annurev.genet.40.110405.090448

44. Pace JK, Feschotte C. The evolutionary history of human DNA transposons: Evidence for intense activity in the primate lineage. *Genome Res*. 2007;17(4):422-432. doi:10.1101/gr.5826307

45. Cordaux R, Udit S, Batzer MA, Feschotte C. Birth of a chimeric primate gene by capture of the transposase gene from a mobile element. *Proc Natl Acad Sci U S A*. 2006;103(21):8101-8106. doi:10.1073/pnas.0601161103

46. Ayarpadikannan S, Kim HS. The Impact of Transposable Elements in Genome Evolution and Genetic Instability and Their Implications in Various Diseases. *Genomics Inform*. 2014;12(3):98. doi:10.5808/gi.2014.12.3.98

47. Robertson HM. The mariner transposable element is widespread in insects. *Nature*. 1993;362(6417):241-245. doi:10.1038/362241a0

48. Plasterk RH. The Tc1/mariner transposon family. *Curr Top Microbiol Immunol*. 1996;204:125-143. doi:10.1007/978-3-642-79795-8_6

49. Lampe DJ, Akerley BJ, Rubin EJ, Mekalanos JJ, Robertson HM. Hyperactive transposase mutants of the Himar1 mariner transposon. *Proc Natl Acad Sci U S A*. 1999;96(20):11428-11433. doi:10.1073/pnas.96.20.11428

50. Smit AF, Riggs AD. Tiggers and DNA transposon fossils in the human genome. *Proc Natl Acad Sci U S A*. 1996;93(4):1443-1448. doi:10.1073/pnas.93.4.1443

51. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature*. 2001;409(6822):860-921. doi:10.1038/35057062

52. Robertson HM, Zumpano KL. Molecular evolution of an ancient mariner transposon, Hsmar1, in the human genome. *Gene*. 1997;205(1-2):203-217. doi:10.1016/s0378-1119(97)00472-1

53. Robertson HM, Martos R. Molecular evolution of the second ancient human mariner transposon, Hsmar2, illustrates patterns of neutral evolution in the human genome lineage. *Gene*. 1997;205(1-2):219-228. doi:10.1016/s0378-1119(97)00471-x

54. Lié O, Renault S, Augé-Gouillou C. SETMAR, a case of primate co-opted genes: towards new perspectives. *Mob DNA*. 2022;13(1):1-14. doi:10.1186/s13100-022-00267-1

55. Qian C, Zhou MM. SET domain protein lysine methyltransferases: Structure, specificity and catalysis. *Cell Mol Life Sci*. 2006;63(23):2755-2763. doi:10.1007/s00018-006-6274-5

56. Kuzmichev A, Nishioka K, Erdjument-Bromage H, Tempst P, Reinberg D. Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. *Genes Dev*. 2002;16(22):2893-2905. doi:10.1101/gad.1035902

57. Liu D, Bischerour J, Siddique A, Buisine N, Bigot Y, Chalmers R. The Human SETMAR Protein Preserves Most of the Activities of the Ancestral Hsmar1 Transposase . *Mol Cell Biol*. 2007;27(3):1125-1132. doi:10.1128/mcb.01899-06

58. Kim HS, Kim SK, Hromas R, Lee SH. The set domain is essential for metnase functions in replication restart and the 5' end of ss-overhang cleavage. *PLoS One*. 2015;10(10):1-16. doi:10.1371/journal.pone.0139418

59. Chen Q. A ROLE FOR SETMAR IN GENE REGULATION : INSIGHTS FROM STRUCTURAL ANALYSIS OF THE DNA-BINDING DOMAIN IN COMPLEX WITH DNA. 2016;(August).

60. Higgins JJ, Pucilowska J, Lombardi RQ, Rooney JP. Candidate genes for recessive non-syndromic mental retardation on chromosome 3p (MRT2A)*. *Clin Genet*.

2004;65(6):496-500. doi:https://doi.org/10.1111/j.0009-9163.2004.00267.x

61. Kaur E, Nair J, Ghorai A, et al. Inhibition of SETMAR-H3K36me2-NHEJ repair axis in residual disease cells prevents glioblastoma recurrence. *Neuro Oncol*. 2020;22(12):1785-1796. doi:10.1093/neuonc/noaa128

62. Dussaussois-Montagne A, Jaillet J, Babin L, et al. SETMAR isoforms in glioblastoma: A matter of protein stability. *Oncotarget*. 2017;8(6):9835-9848. doi:10.18632/oncotarget.14218

63. Lee SH, Oshige M, Durant ST, et al. The SET domain protein Metnase mediates foreign DNA integration and links integration to nonhomologous end-joining repair. *Proc Natl Acad Sci U S A*. 2005;102(50):18075-18080. doi:10.1073/pnas.0503676102

64. Wray J, Williamson EA, Chester S, et al. The transposase domain protein Metnase/SETMAR suppresses chromosomal translocations. *Cancer Genet Cytogenet*. 2010;200(2):184-190. doi:10.1016/j.cancergencyto.2010.04.011

65. Chen Q, Bates AM, Hanquier JN, et al. Structural and genome-wide analyses suggest that transposon-derived protein SETMAR alters transcription and splicing. *J Biol Chem*. 2022;298(5):101894. doi:10.1016/j.jbc.2022.101894

66. Chen Q, Georgiadis M. Crystallization of and selenomethionine phasing strategy for a SETMAR-DNA complex. *Acta Crystallogr Sect Struct Biol Commun*. 2016;72:713-719. doi:10.1107/S2053230X16012723

67. Swalley SE, Baird EE, Dervan PB. Effects of γ-turn and β-tail amino acids on sequence-specific recognition of DNA by hairpin polyamides. *J Am Chem Soc*. 1999;121(6):1113-1120. doi:10.1021/ja9830905

68. Liu B, Wang S, Aston K, et al. β-Alanine and N-terminal cationic substituents affect polyamide-DNA binding. *Org Biomol Chem*. 2017;15(46):9880-9888. doi:10.1039/c7ob02513k

69. Tellier M, Chalmers R. The roles of the human SETMAR (Metnase) protein in illegitimate DNA recombination and non-homologous end joining repair. *DNA Repair (Amst)*. 2019;80(March):26-35. doi:10.1016/j.dnarep.2019.06.006

70. Wang S, Aston K, Koeller KJ, et al. Modulation of DNA-polyamide interaction by β-alanine substitutions: A study of positional effects on binding affinity, kinetics and thermodynamics. *Org Biomol Chem*. 2014;12(38):7523-7536. doi:10.1039/c4ob01456a

71. Jaramillo D, Liu Q, Aldrich-Wright J, Tor Y. Synthesis of N-Methylpyrrole and N-Methylimidazole Amino Acids Suitable for Solid-Phase Synthesis. *J Org Chem*. 2004;69(23):8151-8153. doi:10.1021/jo048686r

72. Baird EE, Dervan PB. Solid Phase Synthesis of Polyamides Containing Imidazole and Pyrrole Amino Acids. 1996;7863(9):6141-6146.

73. Rucker VC, Melander C, Dervan PB. Influence of beta -Alanine on Hairpin Polyamide Orientation in the DNA Minor Groove. 2003;86.

74. Chenoweth DM, Harki DA, Dervan PB. Solution-Phase Synthesis of Pyrrole - Imidazole Polyamides. Published online 2009:7175-7181.

75. Reeves R, Nissen MS. The A·T-DNA-binding domain of mammalian high mobility group

I chromosomal proteins. A novel peptide motif for recognizing DNA structure. *J Biol Chem*. 1990;265(15):8573-8582. doi:10.1016/s0021-9258(19)38926-4

76. Fields GB. Introduction to peptide synthesis. *Curr Protoc protein Sci*. 2002;Chapter 18:Unit 18.1. doi:10.1002/0471140864.ps1801s26

77. Greenberg WA, Baird EE, Dervan PB. A comparison of H-pin and hairpin polyamide motifs for the recognition of the minor groove of DNA. *Chem - A Eur J*. 1998;4(5):796-805. doi:10.1002/(SICI)1521-3765(19980515)4:5<796::AID-CHEM796>3.0.CO;2-G

78. Xiao J, Yuan G, Huang W, Chan ASC, Lee KLD. A convenient method for the synthesis of DNA-recognizing polyamides in solution. *J Org Chem*. 2000;65(18):5506-5513. doi:10.1021/jo000135n

79. Liu B, Wang S, Aston K, et al. β-Alanine and N-terminal cationic substituents affect polyamide–DNA binding. *Org Biomol Chem*. 2017;15(46):9880-9888. doi:10.1039/c7ob02513k

80. Gumpper RH, Li W, Castañeda CH, Scuderi MJ, Bashkin JK, Luo M. A Polyamide Inhibits Replication of Vesicular Stomatitis Virus by Targeting RNA in the Nucleocapsid. *J Virol*. 2018;92(8). doi:10.1128/jvi.00146-18

81. Khan GS, Shah A, Zia-Ur-Rehman, Barker D. Chemistry of DNA minor groove binding agents. *J Photochem Photobiol B Biol*. 2012;115:105-118. doi:10.1016/j.jphotobiol.2012.07.003

82. Cho J, Parks ME, Dervan PB. Cyclic polyamides for recognition in the minor groove of DNA. *Proc Natl Acad Sci U S A*. 1995;92(22):10389-10392. doi:10.1073/pnas.92.22.10389

83. Tellier M. Structure, Activity, and Function of SETMAR Protein Lysine Methyltransferase. *Life*. 2021;11(12). doi:10.3390/life11121342

84. Tellier M, Chalmers R. Human SETMAR is a DNA sequence-specific histone-methylase with a broad effect on the transcriptome. *Nucleic Acids Res*. 2019;47(1):122-133. doi:10.1093/nar/gky937

Chapter 4: Insertion of the Cationic groups into the N-terminal of the Novel Polyamides Sequence.

4.1. Introduction

Pyrrole-Imidazole polyamides (PIPs) are heterocyclic polymers that consist of aromatic heterocycles N-methylpyrrole(Py) and N-methylimidazole(Im) linked by amide bonds.[13] Many useful biomedical applications of polyamides include inhibiting protein-DNA interaction, gene expression, and anti-cancer drugs.[5,25] PIPs are derived from the natural DNA minor groove binding compounds such as netropsin and distamycin. Similar to Netropsin and Distamycin, polyamides can fit through the tight space of the DNA minor groove, recognize the specific sequence of dsDNA, and form hydrogen bonds with the DNA base pair in the minor groove.[2,81]

Since the late 1980s, the works of Dickerson, Lown, and Dervan have introduced us to the binding rule of polyamides, allowing more polyamide sequences to be synthesized to bind the desired DNA sequence.[2,9,12,13] DNA base pairs allow position for Hydrogen donors and acceptors for the building blocks. As mentioned by the "binding rules," A, T, and C can be recognized by pyrrole indifferently and β-alanine; meanwhile, G is selectively recognized by imidazole.[1,82] In addition, substituting β-alanine for pyrrole has been shown to improve the binding affinity and the binding kinetics of polyamide sequences.[17,70,79] Reports showed that the cationic charge of the polyamide sequences improves the PIP's solubility and affinity to the anionic DNA sequences.

SETMAR is a chimeric fusion of a SET methyltransferase domain and a mariner transposon domain, and this chimeric Transposable element has lost its active transposon activity since the rise of the primate transposon around 40 million years ago.[45,57,60] Despite having an inactive transposon function, SETMAR still has an unknown role that involves the development of cancer cells.[64,83,84] With the discovery of SETMAR-TIR specific binding location in the minor groove at the 19bp conserved sequence 5'-GGTGCAAAAGTAATTGCGG-3' by Georgiadis and colleagues.[65,66] This provides an opportunity to design DNA minor groove compounds to study the mechanism of the SETMAR function and prevent the overexpression of SETMAR.[62]

In this study, we modified the polyamides to have an additional cationic charge at the N-terminus, and all the polyamides are synthesized with the solid phase peptide synthesis method, starting with a Boc-β-PAM resin linking the building blocks with amide bonds. We introduced additional charges, such as Guan and TMG groups, to the N-terminal of the sequence.[68] Due to the presence of two nitrogens on the imidazole heterocycle combining with the weak nucleophile NH_2 group making a stronger electron-withdrawing effect of the imidazole, and it is very unstable coupling to NH_2-Im-Im dimers in the polyamides sequences.[78] To solve this problem, we add an additional Boc-Pyrrole prior to the Imidazole-imidazole dimer to ensure the success of the coupling

reactions and ensure the sequence stabilizes for the addition of the Guan and TMG group (**Table 4.1**).

Table 4.1: Designed polyamides sequences used for this study.

Compound	N-terminus	1	2	3	4	5	6	7	8	9	10	11	12	C-terminus
1176	TMG-Py	**Im**	**Im**	**β**	**Im**	Py	**γ**	**Im**	Py	**β**	Py	Py	**β**	Ta
1178	Guan-Py	**Im**	**Im**	**β**	**Im**	Py	**γ**	**Im**	Py	**β**	Py	Py	**β**	Ta
1181	TMG-Py	**Im**	**Im**	**β**	**Im**	Py	**γ**	**Im**	**β**	Py	Py	Py	**β**	Ta
1182	Guan-Py	**Im**	**Im**	**β**	**Im**	Py	**γ**	**Im**	**β**	Py	Py	Py	**β**	Ta

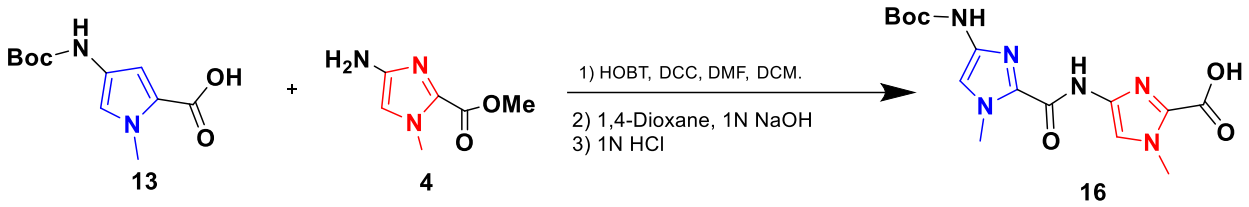
4.2 Method

4.2.1 Material and Instrumentation

Boc- β-alanine-PAM resin was ordered from Peptides International (Louisville, KY), 3-(dimethylamino)-1-propylamine, N-methyl-imidazole, N-methyl-pyrrole, piperidine, indole, ammonium formate, 20% Pd/C, dimethyl sulfoxide, HCl, and NaOH were purchased from Sigma Aldrich (St. Louis, MO). 1-methyl-4-nitro-imidazole-2-carboxylate (NO_2-Im-COOMe) was ordered from AC Pharmachem, Inc (Worcester, MA). 4-tert-butoxycarbonylaminobutyric acid (Boc-γ-COOH), 1-ethyl-4-amino-β-alanine-carboxylate (H_2N-β-COOEt),4-tert-butoxy-β-alanine-carboxylic acid (Boc-β-COOH), trifluoroacetic acid (TFA), formic acid (FA), hexafluorophosphate azabenzotriazole tetramethyl uronium (HATU), benzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate (PyBOP), hydroxybenzotriazole (HOBT), dimethyl formamide (DMF), and methanol (MeOH) were acquired from Oakwood Chemical. Triethylamine, N, N-diisopropylethylamine (DIPEA), formic acid, and acetic anhydride were purchased from Fisher Scientific.

^{1}H-NMR and ^{13}C-NMR spectra were recorded in DMSO-d_6 at 300 MHz and 600 MHz, respectively. ^{1}H-NMR is referenced to DMSO at 2.5 ppm, and ^{13}C-NMR in DMSO at 39.5 ppm. Coupling constants J are reported in Hz. High-resolution mass spectrometry was carried out on a Bruker Maxis II spectrometer. Analytical HPLC data was analyzed by Agilent HPLC. Purifications of Polyamides were carried out in ThermoFisher preparative HPLC.

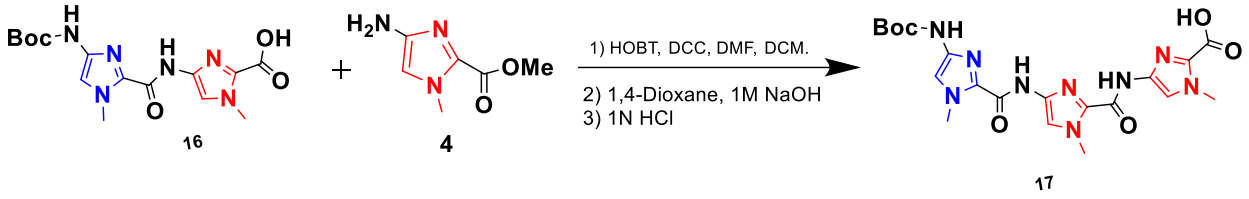
4.2.2. Synthesis of Boc-Py-Im-COOH 16 (Scheme 4.1)



Scheme 4.1. Synthesis of Boc-Py-Im-COOH **16**.

In a 250 mL round bottom flask, Boc-Py-COOH **13 (**8 g, 33.3 mmol) was dissolved in DMF (16 mL), followed by HOBT (6.7 g, 49.9 mmol) and DCC (10.3 g, 49.9 mmol) in DCM (40 mL), stir well under nitrogen gas for 20 min. After 30 min at room temperature, the mixture formed DCU and turned white solution, let the reaction ran overnight at room temperature. After 12 h, add H_2N-Im-COOMe **4** (7.75 g, 49.9 mmol) in DCM (20 mL), and the reaction was stirred well under Nitrogen gas. The reaction was monitored by TLC and HPLC, and the reaction was completed after 18 h at room temperature. The mixture was filtered to remove the solid DCU, and the DCM solvent was removed by rotary evaporator. EtOAc was added to the flask in vacuo, and then the mixture was washed with H_2O, aq $NaHCO_3$, 4% HCl, and brine. The organic solvent was dried with anhydrous Na_2SO_4 and gravity filtered to collect a red solution. The solution's solvent is removed by rotor evaporator to afford an oil solution. The oil mixture was immediately dissolved in 1,4-Dioxane (40 mL) and stirred until dissolved under N_2 for 30 min before adding 1N NaOH (40 mL). The reaction was heated in an oil bath with a temperature controlled at 45°C for 1 to 2 h and checked for completion with analytical HPLC. The reaction was removed from heat and settled at room temperature for 10 min, reducing the pH to around 4.25 to afford off red-brow solid **16 (**8.76 g, 33.3 mmol, 69.7% yield).

4.2.2 Synthesis of dimer building blocks Boc-Py-Im-Im-COOH. (17, Scheme 4.2)



Scheme 4.2. Synthesis of Boc-Py-Im-Im-COOH 17.

In a 250 mL round bottom flask, Boc-Py-Im-COOH **16 (**5 g, 13.7 mmol) was dissolved in DMF (13mL), followed by HOBT (2.77 g, 20.55 mmol), and DCC (4.24 g, 20.55 mmol) in 40mL DCM, stir well with a stir bar until dissolved nitrogen gas for 30 min. After 5 minutes, there was

still undissolved solid; the RBF was ultrasonicated until all solids had dissolved after 10 minutes. After 30 min at room temperature, the mixture formed DCU and turned white solution, let the reaction ran overnight at room temperature.

After 12 h, add H_2N-Im-COOMe **4 (**3.2 g, 20.55 mmol) in DCM (20 mL), and the reaction was stirred well under Nitrogen gas. The reaction was monitored by TLC and HPLC, and the reaction was completed after 18 h at room temperature. The mixture was filtered to remove the solid DCU, and the DCM solvent was removed in a rotary evaporator. EtOAc was added to the flask, and then the mixture was washed with H_2O, aq $NaHCO_3$, 4% HCL, and brine. The organic solvent was dried with anhydrous Na_2SO_4 and gravity filtered to collect an orange-red solution. The solution's solvent is removed by rotor evaporator to afford an oil solution. The oil mixture was immediately dissolved in 2,4-dioxane (40 mL) and stirred under N_2 for 30 min before adding 1N NaOH (40 mL). The reaction was heated in an oil bath with a temperature controlled at 45 ^oC for 1 to 2 h and checked for completion with analytical HPLC. The reaction was removed from heat and settled at room temperature for 10 min, reducing the pH to around 3.75 to afford a solid of light pink color **17** (2.97 g, 13.7 mmol, 44.6% yield).

4.2.4 Synthesis of Boc-Py-Im-Im-β-Im-Py-γ-Im-Py-β-Py-Py-β-PAM resin.

Synthesis of 1174.

Synthesis of polyamides **1174** resin started by initiating the reaction sequences program with the CSBio program (**Table 4.2**). The CSBio reaction vessel was set at a temperature of 35^oC throughout the whole process. Boc-β-alanine-PAM resin (1000 mg) was loaded into the RV. Py-BOP (11 g) was dissolved in DMF (40 mL) and filled in the R4 solvent bottle. After the program started the first wash and deprotection reaction for 1 h, a solution of a monomer Boc-Py-Py-COOH (505 mg, 1.38 mmol) in DMF (6.5 mL) was stirred to fully dissolved and transferred into **AA#3**.

The coupling program directed the synthesizer to transfer the solution from **AA#3** to TVA and then MVA to begin the activation before transferring to RV for 180 min, followed by capping with acetic anhydride for 17 min to remove impurities such as an unreacted amine. After completing the first sequence, the program proceeded to "cap" to remove any nonreacted amine before moving on to the next sequence. Each sequence took around 6 h to complete.

After completing the first sequence, the programmed system transferred a solution of Boc-Py-β-COOH (433 mg, 1.38 mmol) in DMF (6.5 mL), was well dissolved by vortex and loaded in **AA#4** to TVA and MVA for sequence coupling followed up by capping. After sequence 2 plus cap was completed, a solution of Boc-γ-Im-COOH (452mg, 1.38 mmol) and DIPEA (1mL) in DMF

(6.5mL) was well dissolved by vortex; loaded in **AA#5** transferred to TVA then MVA before transferred to RV for coupling reaction for 180 min followed by capping with acetic anhydride for 17 min.

After sequence 3 plus cap was completed, a solution of Boc-Py-OBt (656 mg, 1.84 mmol) in DMF (6.5 mL) was well dissolved by vortex; loaded in **AA#6,** transferred to TVA then MVA before transferred to RV for coupling reaction for 180 min followed by capping with acetic anhydride for 17 min. After sequence 4 was finished, the programed system was initiated and followed by transferring a well-dissolved by vortex solution of Boc-β-Im-COOH (432 mg, 1.38 mmol) in DMF (6.5 mL) and DIPEA (1mL) from **AA#7** to TVA, then MVA before transferred to RV for coupling reaction for 180 min following by capping with acetic anhydride for 17 min.

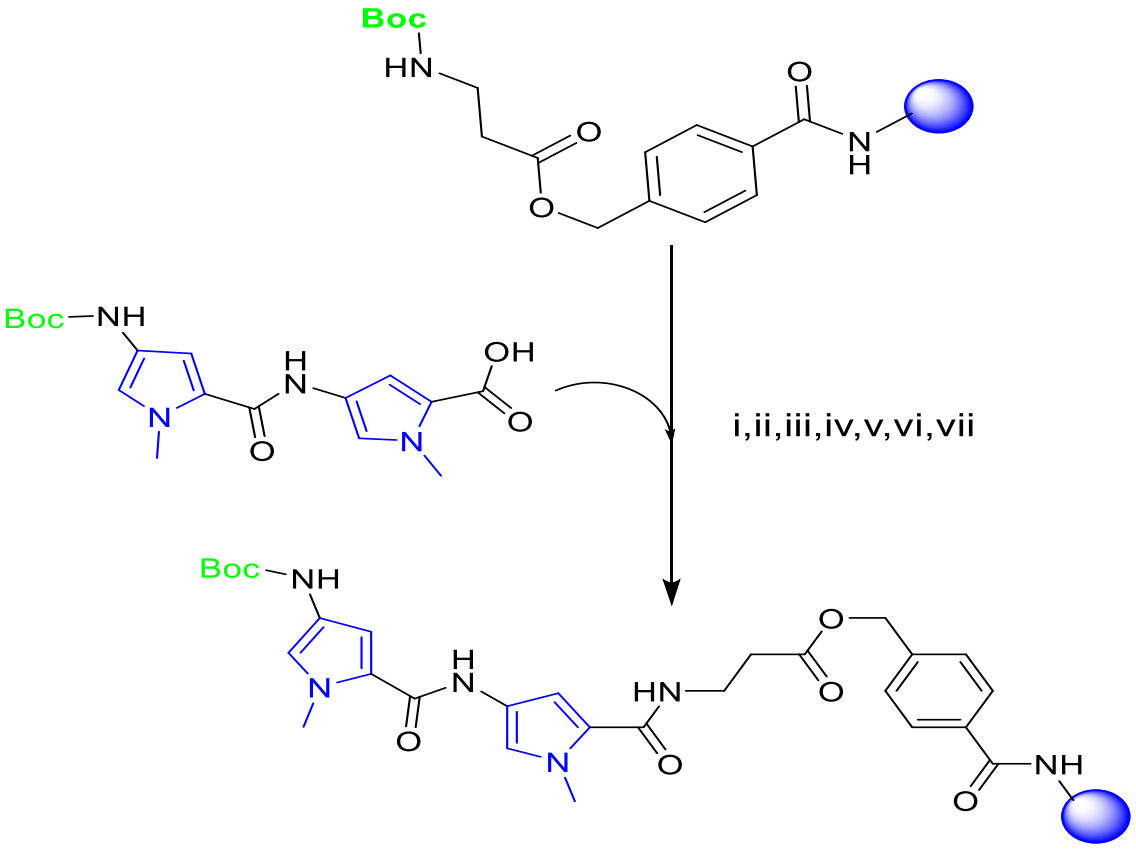
After sequence 5 was finished, the programed system was initiated and followed by transferring a well-dissolved by vortex solution of trimer Boc-Py-Im-Im-COOH (672 mg, 1.38 mmol) in DMF(6.5 mL) and DIPEA (1mL) from **AA#8** to TVA, then MVA before transferred to RV for coupling reaction for 180 min following by capping with acetic anhydride for 17 min (**Figure 4.3).**

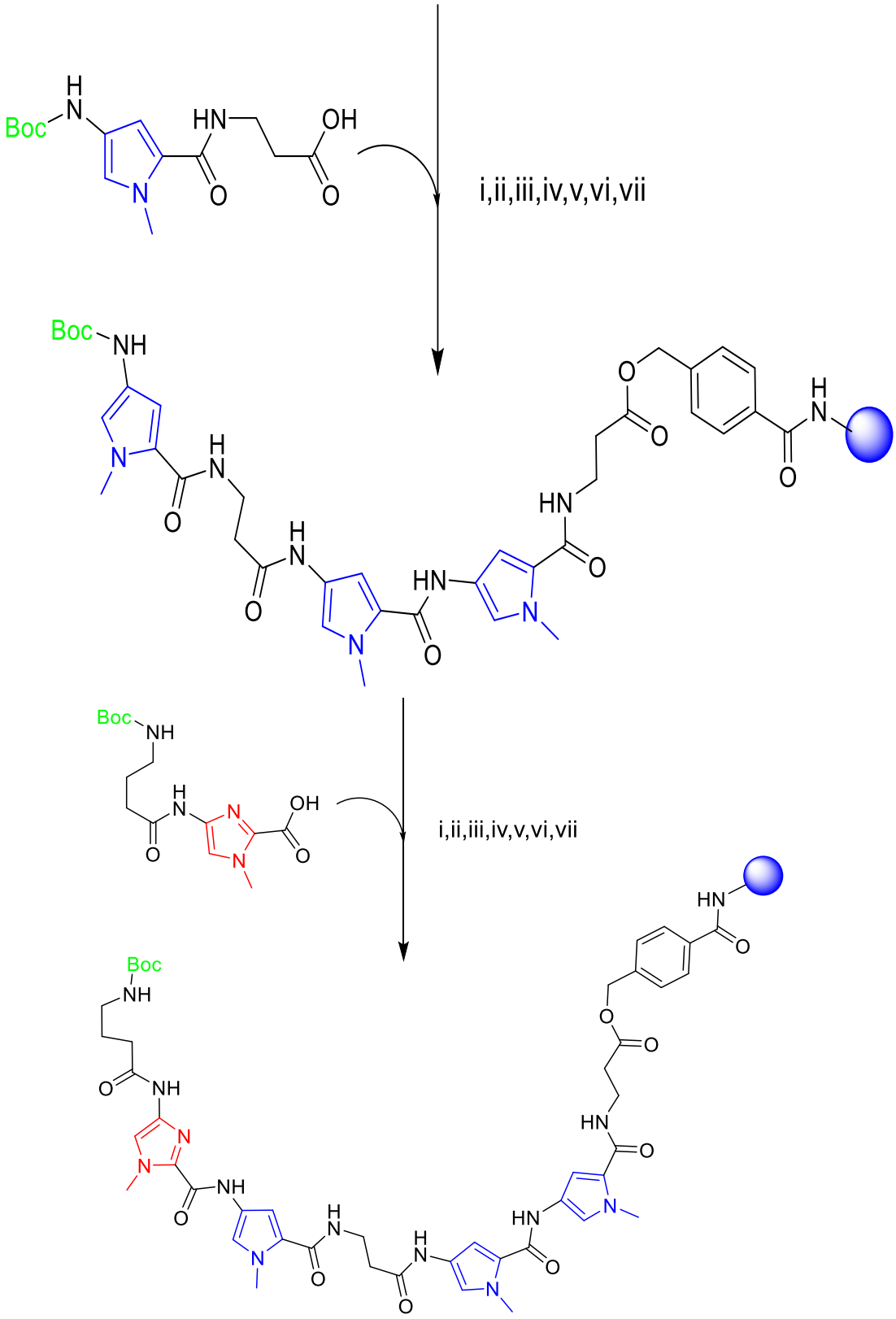
The sequence was finished after 30 h. 0,1 mg of resin was cleaved with 100 µL Ta solution for 2 h at 45^{o}C, and the heated solution was collected and filtered before being analyzed with a Bruker LC-MS instrument. The exact mass was determined to be $[M+H]^+$ as 1646.82 m/z, confirming the coupling reaction's completion. All the resin was carefully transferred from the RV into a beaker using a small spatula, and the resin was washed three times with DCM and MeOH, then dried under vacuum for 30 min until reaching dryness. The resin was then collected in a 20mL scintillation vial and stored in the freezer.

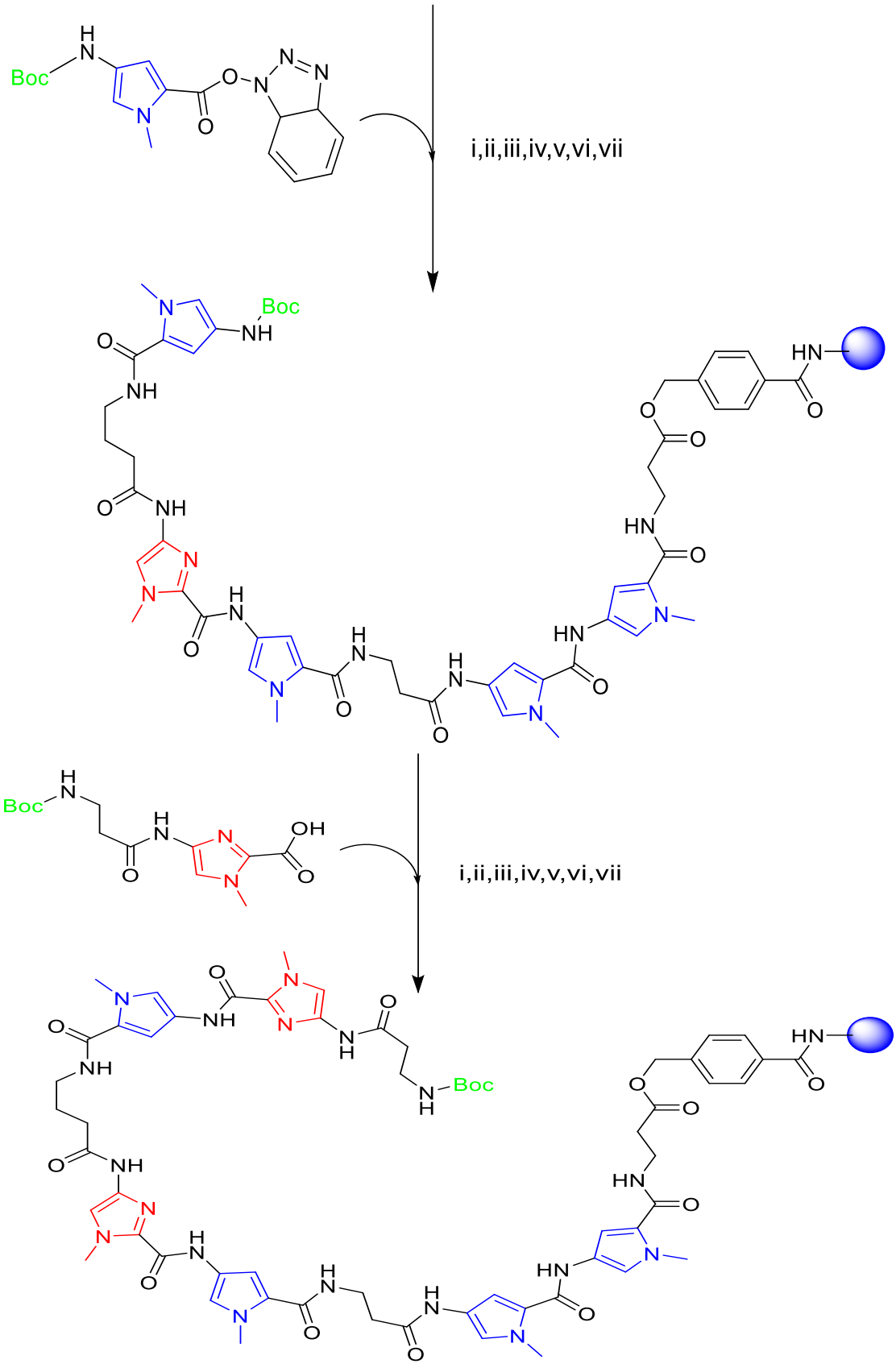
Table 4.2 Program sequence of **1174** resin synthesis.

Building block(s) and step(s)	Additional Solvent mixture	Applied program
1-Boc-Py-Py-COOH	(6.5 mL) DMF	Add-DIPEA-180-min-DIPEA-Flows.CFN
2-Capping	None	Cap-17-min-continues.CFN
3-Boc-Py-β-COOH	(6.5 mL) DMF	Add-DIPEA-180-min-DIPEA-Flows.CFN
4-Capping	None	Cap-17-min-continues.CFN
5-Boc-γ-Im-COOH	(6.5 mL) DMF + DIPEA (1mL)	Add-DIPEA-180-min-DIPEA-Flows.CFN

6-Capping	None	Cap-17-min-continues.CFN
7-Boc-Py-OBt	(6.5 mL) DMF	Add-DIPEA-180-min-DIPEA-Flows.CFN
8-Capping	None	Cap-17-min-continues.CFN
9-Boc β-Im-COOH	(6.5 mL) DMF + DIPEA (1mL)	Add-DIPEA-180-min-DIPEA-Flows.CFN
10-Capping	None	Cap-17-min-continues.CFN
11-Boc-Py-Im-Im-COOH	(6.5 mL) DMF + DIPEA (1mL)	Add-DIPEA-180-min-DIPEA-Flows.CFN
12-Capping	None	Cap-17-min-continues.CFN




Boc
H
N
HN
OH
O
N
O
i,ii,iii,iv,v,vi,vii
Boc
NH
N
H
N
O
H
N
O
HN
O
O
HN
O
N
O
H
N
O
Boc
NH
H
N
N
OH
O
N
O
i,ii,iii,iv,v,vi,vii
Boc
NH
O
NH
N
N
O
NH
H
N
O
H
N
O
HN
O
N
HN
O
O
HN
O
O
O
HN
O


Boc
i,ii,iii,iv,v,vi,vii
Boc
i,ii,iii,iv,v,vi,vii

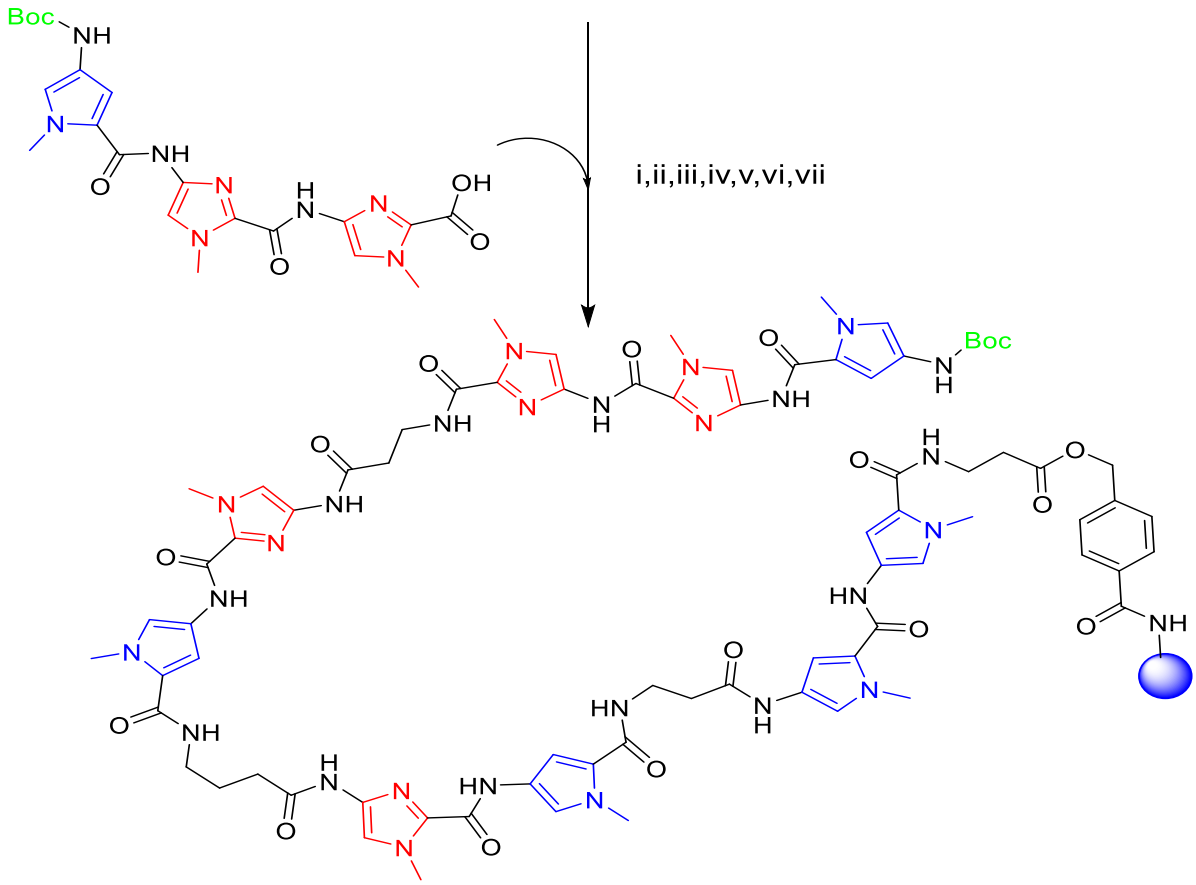


Figure 4.3: *Scheme of synthesis polyamide* **1174** resin. *It started from a commercially acquired Boc-β-PAM resin, and each sequence contains multiple steps from (i) to (viii). (i) wash the resin with DCM and DMF, (ii) deprotection with 60% TFA/ DCM/ 0.5 M Indole, (iii) 25% piperidine/ DMF; (iv) building blocks with the structures shown on the scheme,(v) 0.5M PyBOP/ DMF as coupling reagents, (vi) DIPEA, (vii) DMSO, (viii) capping step with acetic anhydride for 17 min.*

4.2.5. Synthesis of Guan-Py-Im-Im-β-Im-Py-γ-Im-Py-β-Py-Py-β-PAM resin.

Synthesis of polyamides **1176** resin started by initiating the reaction sequences program with the CSBio program (**Table 4.3**). The CSBio reaction vessel was set at a temperature of 35°C throughout the whole process. Resin from the previous step, **1174** (0.512 mg), was loaded into the RV. The first step was washing the resin with DMF, followed by capping to remove any extra impurities.

A solution of Bis-Boc-Pyrazolocarboxamidine (1.09 g, 3.48 mmol) in DIPEA (1 mL) and of DMF (6.5 mL) was added into the RV 3h after step 3 began. After 26 h, the program started the

deprotection step using TFA to deprotect the Boc group from Boc_2Guan to introduce the Guan group in the N-terminus of the polyamide **1176**. After the program finished, 0.1 mg of resin was cleaved with Ta as a nucleophile for four h at 45°C, and the sample was analyzed with a Bruker LC-MS instrument. The exact mass was determined to be $[M+H]^+$ as 1588.78 m/z, confirming the presence of the desired product.

All the resin was carefully transferred from the RV into a beaker using a small spatula, and the resin was washed three times with DCM and MeOH, then dried under vacuum for 30 min until reaching dryness. The resin was then collected in a 20 mL scintillation vial and stored in the freezer before cleavage the next day.

Synthesis of 1176.

Table 4.3 Program sequence of **1176** synthesis.

Building block(s) and step(s)	Additional Solvent mix-ture	Applied program
1 Wash	None	DMF washes. CFN
2-Capping	None	Cap_17min_continue. CFN
3-Boc_2Guan	(6.5 mL) DMF + DIPEA (1mL)	1 gram-pause-24h-DMF-Flows-Benedit. CFN
4-Deprotection	None	DMF-DCM-TFA-DCM-Flows. CFN

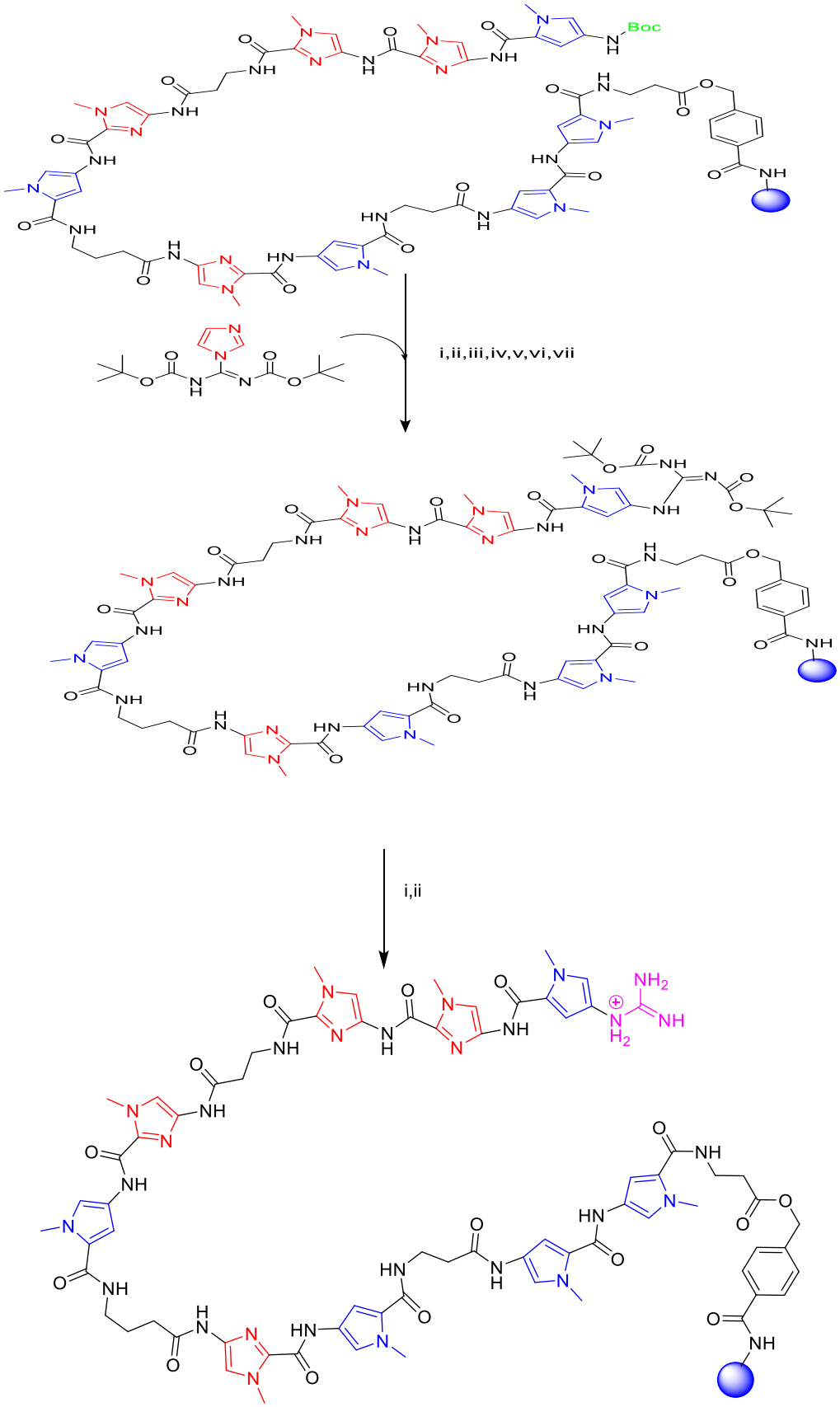

Boc
i,ii,iii,iv,v,vi,vii
i,ii

Figure 4.4: *Scheme of synthesis polyamide* **1176** resin *synthesis. It started from a previously synthesized* ***1174*** *resin, and each sequence contains multiple steps from (i) to (viii). (i) wash the resin with DCM and DMF, (ii) deprotection with 60% TFA/ DCM /0.5 M Indole, (iii) 25% piperidine /DMF; (iv) building blocks with the structures shown on the scheme,(v) 0.5M PyBOP /DMF as coupling reagents, (vi) DIPEA, (vii) DMSO, (viii) capping step with acetic anhydride for 17 min.*

Cleavage of PAM resin using Ta as a nucleophile.

The next day, 453 mg resin was cleaved with aminolysis in 2000 µL of Ta at 45°C in a heating device Isotemp Fisher Scientific for 4 h, and the separation of polyamides solution from the PAM resin support was performed in a 20 µM polyethylene frit placed inside a disposable polypropylene syringe by washing with MeOH and H_2O. The polyamide solution was concentrated by a rotary evaporator to afford 37 mg (0.104 mmol, 22.35% yield) of yellow solid (**Figure 4.5**).

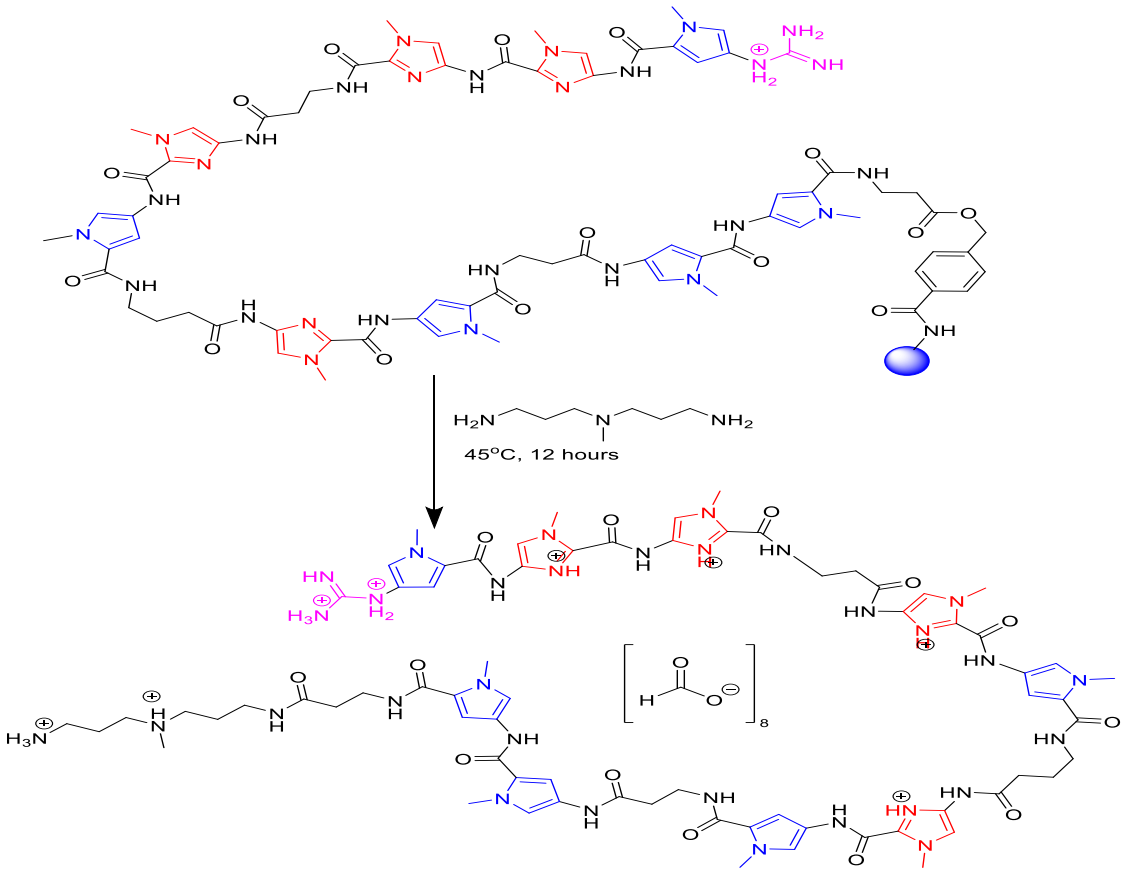


Figure 4.5 Cleavage of PA **1176** using Ta as the nucleophile.

Purification of the compound.

The polyamide solution was diluted with a mixture of DMSO (200 μL) and H_2O/ 0.2% HCOOH (200 μL), then purified by preparative HPLC using a Phenomenex Luna 250x30 mm, 5 μM, 100 Å C18 column maintained at 25 °C. The organic phase was 0.2% HCOOH in MilliQ water and 100% HPLC grade MeOH. The applied gradient was 10% MeOH for 8 min, followed by a ramp rate to 90% Methanol over 35.6 min at a 20 mL/min flow rate. All collected fractions were analyzed and selected for 95% or higher purity with analytical HPLC using ThermoFisher-UHPLC; analyses were processed and analyzed with the program Chromeleon 3.0. High-purity fractions were collected for rotary evaporation to remove MeOH before the lyophilization.

Lyophilization of the compound.

Compound **1176** was diluted with a 1:1 mixture of ACN (5 mL): and 0.2% HCOOH in H_2O (5mL). The samples were frozen before the lyophilization process with ground dry ice until an even coated of frozen material results at an acute angle in the vial. The frozen sample was covered with porous Kimwipe paper and secured by a rubber band. The sample was settled in a lyophilizer glass vessel and connected to a vacuum port. The frozen polyamide solution was sublimed as an off-white solid powder after 48 h. The lyophilizer vacuum was set at 300 torr within 30 min prior to the lyophilization process, and the inside temperature was set at -83 to -85 °C.

Characterization of compound

Analytical HPLC characterization was performed with a C12 Phenomenex Juniper Proteo column maintained at 40°C. The mobile phase consisted of 0.1% formic acid in water (A) and HPLC grade ACN (B). The applied gradient consisted of 5%B for 0.75 min, followed by a ramp to 60% B over 6.5 min at 2.0 mL/ min. The retention was 2.518 min.

1176, Guan-Py-Im-Im-β-Im-Py-γ-Im-Py-β-Py-Py-β-Ta $(HCOO^-)_8$

^{1}H-NMR (600 MHz, DMSO-d_6) δ= 8.42 (br, 8H), 8.18 (s,1H), 8.13 (s, 3H), 8.03 (s, 3H), 7.59 (s, 1H), 7.50 (s, 1H), 7.42 (s, 2H), 7.18 (s, 1H), 7.16-7.12 (d, J= 14.0 Hz, 4H), 7.00 (s. 2H), 6.95-6.93 (d, J= 15.73 Hz, 4H), 6.84 (s, 2H), 6.79(s, 2H), 3.96 (s, 2H), 3.93- 3.92 (d, J= 4.35 Hz, 3H), 3.91 (s, 2H), 3.83(s, 2H), 3.78- 3.77 (d, J= 3.28 Hz, 3H), 3.76- 3.75 (d, J= 3.98 Hz,4H), 3.53- 3.52 (t, J= 10.22 Hz, 2H), 3.48- 3.46 (t, J= 11.78 Hz, 2H), 3.34-3.30 (q, J= 20.43 Hz, 3H), 3.17- 3.15 (t, J= 14.15 Hz, 3H), 3.03- 3.00 (q, J= 19.65 Hz, 2H), 2.75- 2.73 (t, J= 12.57 Hz, 8H), 2.60- 2.57 (T, J= 14.15 Hz, 3H), 2.47- 2.46 (q, J= 7.34 Hz, 4H), 2.31-2.29 (t, J= 12.86 Hz, 10H), 2.06 (s, 2H), 2.04 (s, 1H), 1.77-1.72 (q, J= 27.51 Hz, 4H), 1.66-1.62 (q, J= 27.35 Hz, 11H), 1.49-1.47 (q, J= 20.64 Hz, 3H) .

$_{13}$C-HNMR (151 MHz, DMSO-d_6) δ= 173.1, 172.4, 171.8, 169.5, 164.3, 164.3, 161.5, 161.3, 161.3, 160.6, 158.8, 158.8, 158.3, 139.6, 139.2, 138.9, 137.7, 137.3, 137.2, 137.1, 135.9, 126.9, 126.4, 126.4, 125.9, 125.8, 125.2, 124.2, 121.2, 121.2, 121.1, 121.0, 117.9, 117.1, 117.0, 116.4, 113.1, 107.4, 107.2, 107.0, 57.6, 57.5, 57.1, 44.5,44.5, 44.4, 43.1, 43.0, 42.9, 42.7, 42.6, 42.5, 42.3, 42.2, 41.3, 41.2, 40.6, 40.5, 40.2, 40.1, 39.7, 39.2, 39.1, 39.0, 38.6, 38.3, 38.1, 38.0, 37.9, 7.9, 36.4, 35.9, 29.8, 28.8, 28.7, 27.9.

HRMS (ESI) was calculated for **1176**, $C_{71}H_{93}N_{31}O_{13}$, $[M+H]^+$,1588.77 m/z, found 1588.78 m/z

HPLC purity 97.7%

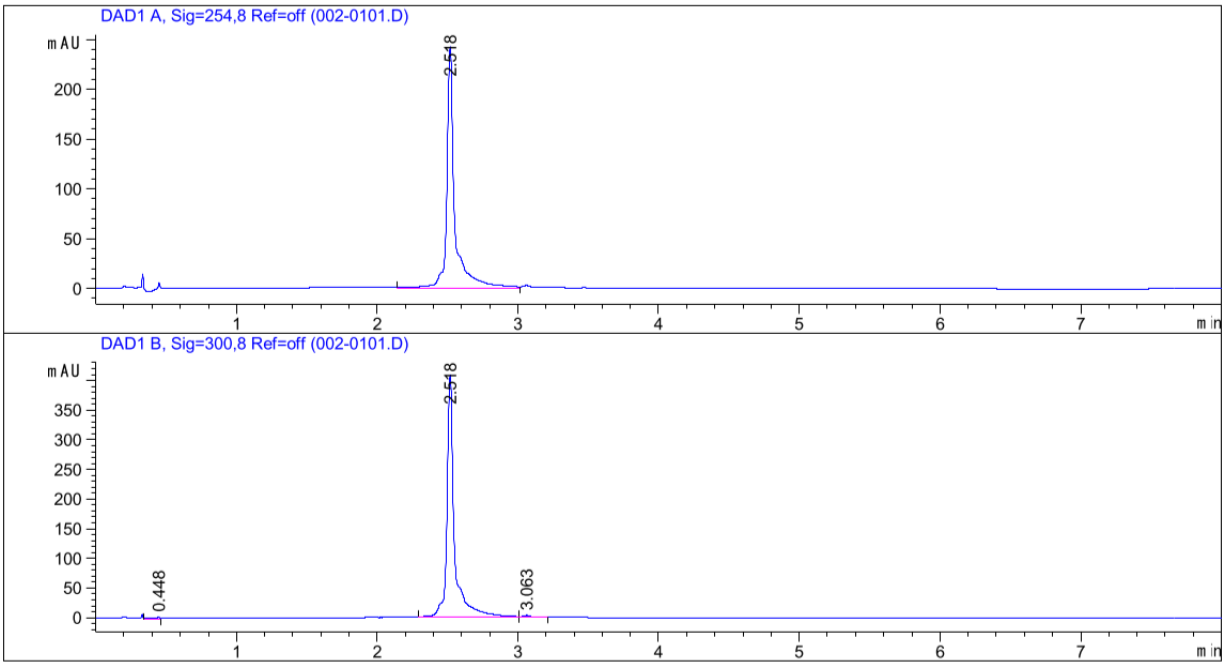


Figure 4.6: Analytical HPLC purity of compound **1176** with λ monitored at 300nm. Compound purity 97%, retention time 2.518 min.

4.2.6 Synthesis of TMG-Py-Im-Im-β-Im-Py-γ-Im-Py-β-Py-Py-β-Ta

Synthesis of 1178.

Synthesis of polyamides **1178** resin started by initiating the reaction sequences program with the CSBio program (**Table 4.4**). The CSBio reaction vessel was set at a temperature of 35°C throughout the whole process. The previously synthesized **1174** resin (0.520 mg) was loaded into the RV. The first step was washing the resin with DMF, followed by capping to remove any extra impurities.

The third program sequence is coupling HATU (hexafluorophosphate azabenzotriazole tetramethyl uronium) to the PIP synthesis sequence; after 3 h since the sequence began, HATU (380 mg,1.38 mmol). After the program finished, 0.1 mg of resin was cleaved with Ta as a nucleophile for 4 h at 45ºC; the sample was analyzed with a Bruker LC-MS instrument. The exact mass was determined to be $[M+H]^+$ as 1588.78 m/z, confirming the coupling reaction's completion. All the resin was carefully transferred from the RV into a beaker using a small spatula, and the resin was washed three times with DCM and MeOH, then dried under vacuum for 30 min until reaching dryness. The resin was then collected in a 20 mL scintillation vial and stored in the freezer before cleavage in the next day.

Table 4.4 Programmed sequence of **1178** PAM resin synthesis.

Building block(s) and step(s)	Additional Solvent mixture	Applied program
1-Wash resin	None	DMF washes. CFN
2-Capping	None	Cap-17-min-continues. CFN
3-HATU solution	(6.5 mL) DMF	AddTMG_DIEA_180_DMSO_DIEA_flows. CFN

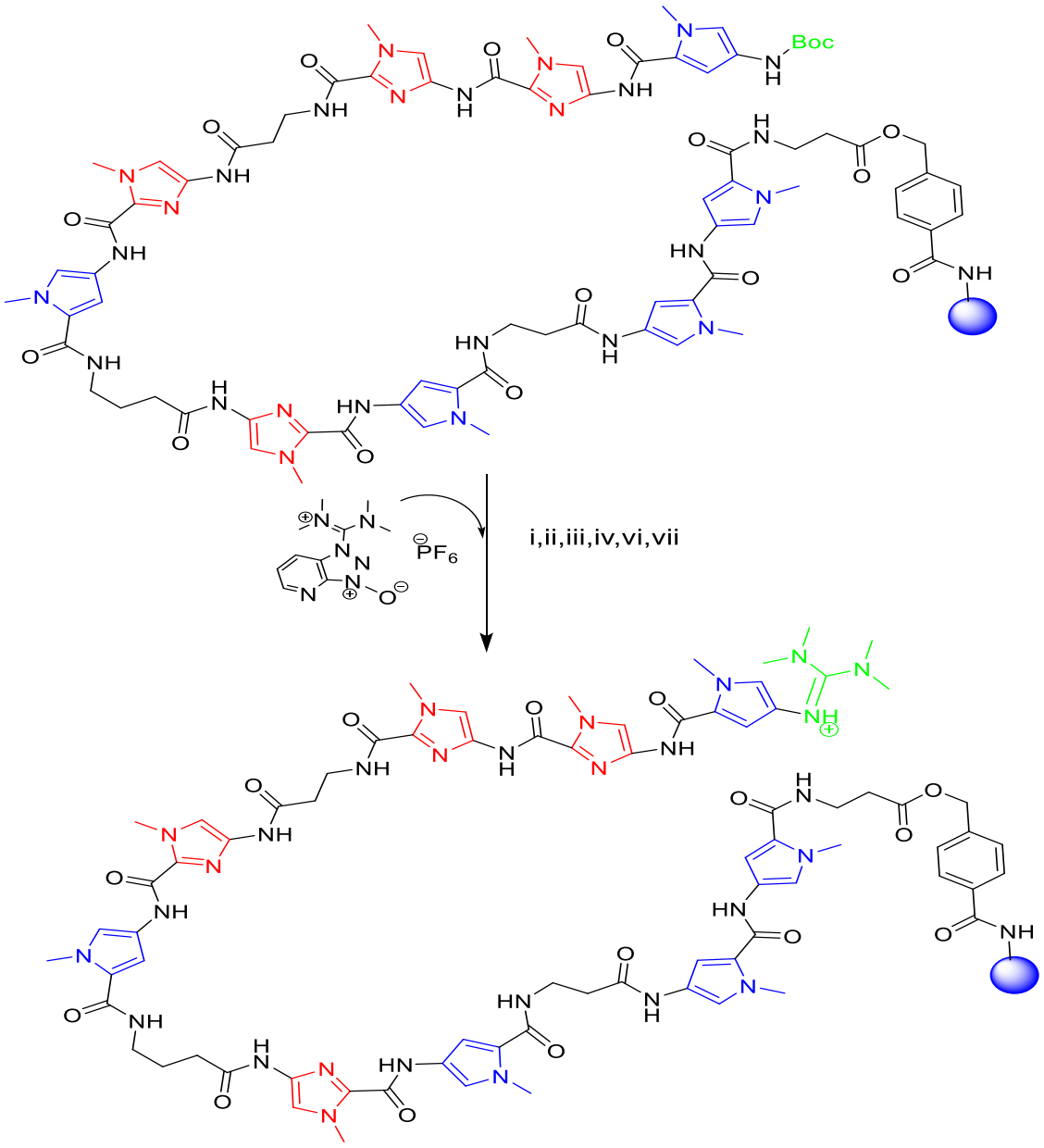


Figure 4.7. Synthesis of **1178** PAM resin synthesis.

Cleavage of PAM resin using Ta as a nucleophile.

The next day, 386 mg resin was cleaved with aminolysis in 2000 µL of Ta at 45ºC in a heating device Isotemp Fisher Scientific for 12 h, and the separation of polyamides solution from the PAM resin support was performed in a 20 µM polyethylene frit placed inside a disposable polypropylene syringe by washing with MeOH and H_2O. A rotary evaporator concentrated the

polyamides solution to afford an off-white solid **1178** (37 mg, 0.088 mmol, 25.33% yield) (**Figure 4.8**).

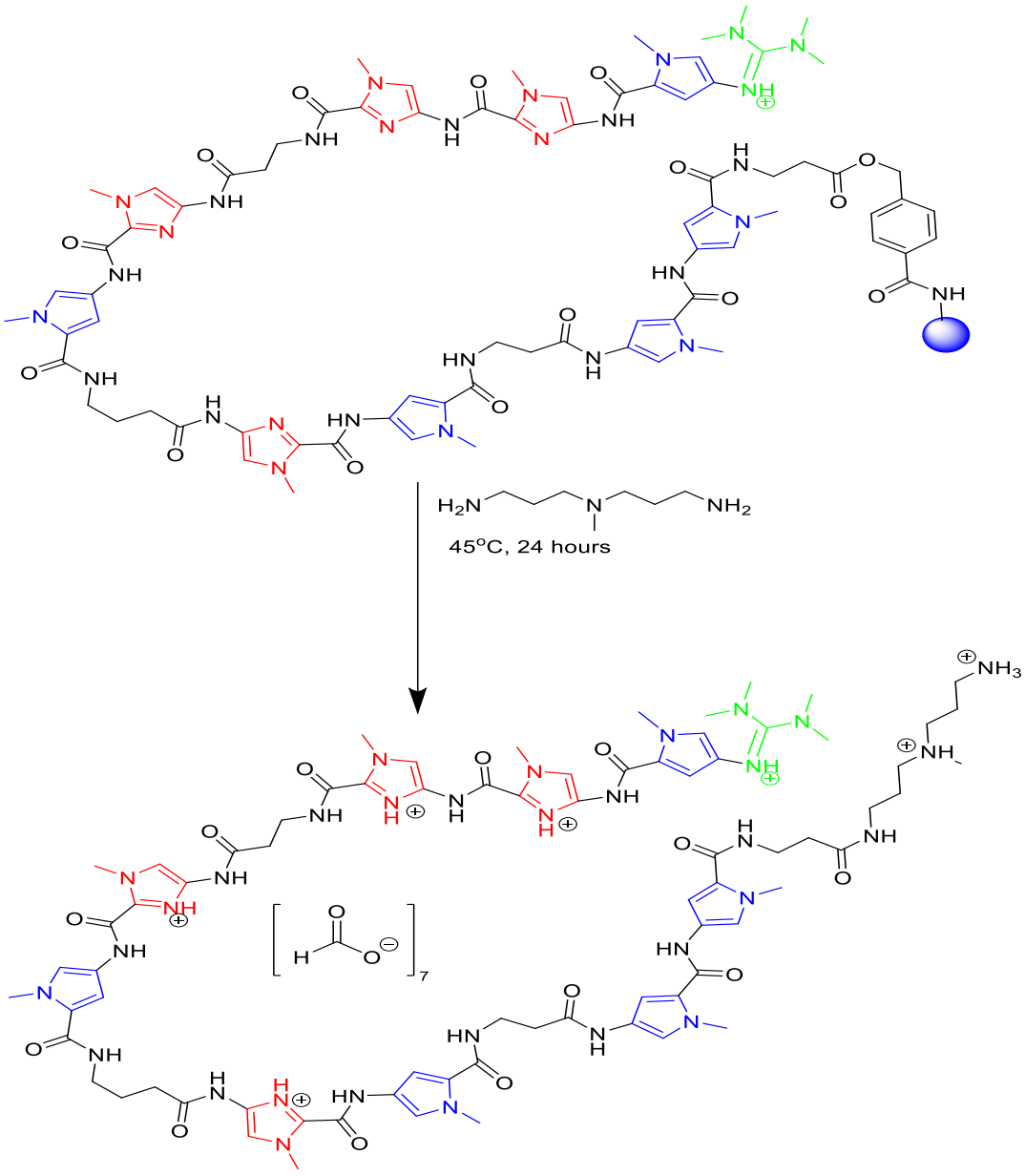


Figure 4.8 Cleavage of **1178** from PAM resin using Ta as the nucleophile.

Purification of the compound.

The polyamide solution was diluted with a mixture of DMSO (200 µL) and H_2O/ 0.2% HCOOH (200 µL), then purified by preparative HPLC using a Phenomenex Luna 250x30 mm, 5 µM, 100 Å C18 column maintained at 25 °C. The organic phase was 0.2% HCOOH in MilliQ water and 100% HPLC grade MeOH. The applied gradient was 10% MeOH for 8 min, followed by a ramp to 90% Methanol over 35.6 min at a 20 mL/min flow rate. All collected fractions were

analyzed and selected for 95% or higher purity with analytical HPLC using ThermoFisher-UHPLC; analyses were processed and analyzed with the program Chromeleon 3.0. High-purity fractions were collected for rotary evaporation to remove MeOH before the lyophilization.

Lyophilization of the compound.

Compound **1178** was diluted with a 1:1 mixture of ACN (5 mL): and 0.2% HCOOH in H_2O (5 mL). The samples were frozen before the lyophilization process with ground dry ice until an even coated of frozen material results at an acute angle in the vial. The frozen sample was covered with porous Kimwipe paper and secured by a rubber band. The sample was settled in a lyophilizer glass vessel and connected to a vacuum port. The frozen polyamide solution was sublimed as an off-white solid powder after 48 h. The lyophilizer vacuum was set at 300 torr within 30 min prior to the lyophilization process, and the inside temperature was set at -83 to -85 ^{o}C.

Characterization of compound

Analytical HPLC characterization was performed with a C12 Phenomenex Juniper Proteo column maintained at 40^{o}C. The mobile phase consisted of 0.1% formic acid in water(A) and HPLC grade ACN (B). The applied gradient consisted of 5% B for 0.75 min, followed by a ramp to 60% B over 6.5 min at 2.0 mL/min. The retention was 2.669 min.

PA **1178**, TMG-Py-Im-Im-β-Im-Py-γ-Im-Py-β-Py-Py-β-Ta $(HCOO^{-})_7$

^{1}H-NMR (600MHz, DMSO-d_6) δ= 8.41 (s, 8H), 8.15(s, 1H), 8.08- 8.05 (d, J= 16.73 Hz, 2H), 8.00-7.96 (d, J= 22.81 Hz, 2H), 7.60 (s, 1H), 7.50 (s, 1H), 7.43-7.40 (d, J= 15.21 Hz, 2H), 7.20 (s, 2H), 7.16- 7.13 (d, J= 12.79 Hz, 2H), 6.95- 6.92 (d, J= 11.29 Hz, 2H), 6.88 (s, 1H), 6.82 (s, 1H), 6.79-6.76 (d, J= 16.56 Hz, 2H), 3.97 (s, 3H), 3.93-3.92 (d, J= 6.77 Hz, 4H), 3.91(s, 3H), 3.84 (s, 3H), 3.78- 3.77 (d, J= 5.27 Hz, 5H), 3.76- 3.75 (d, J= 5.27 Hz, 5H), 3.55 (s, 1H), 3.50- 3.47 (q, J= 19.57 Hz, 3H), 3.33- 3.30 (q, J= 21.83 Hz, 3H), 3.18- 3.14 (q, J= 22.58 Hz, 3H), 2.76-2.73 (t, J= 16.56 Hz, 6H), 2.60- 2.57 (t, J= 15.81 Hz, 2H), 2.48 (s, 15H), 2.32- 2.27 (q, J= 23.33 Hz, 9H), 2.23- 2.20 (q, J= 20.32 Hz, 4H), 2.07(s, 4H), 2.05 (s, 3H), 1.76- 1.73 (t, J= 15.05 Hz, 3H), 1.64- 1.63 (d, J= 9.03 Hz, 3H), 1.49- 1.47 (t, J= 15.81 Hz, 3H), 1.21 (s, 1H),

^{13}C-NMR (151 MHz, DMSO-d_6) δ= 173.5, 173.1, 172.4, 171.9,168.9, 168.9, 168.8, 168.8, 168.7, 168.7, 164.3, 164.2, 161.5, 161.3, 161.2, 160.9, 158.8, 158.2, 139.6, 138.9, 137.7, 137.3, 137.2, 137.1, 135.9, 126.6, 126.6, 126.4, 126.2, 126.1, 125.9, 125.8, 125.2, 125.2, 124.2, 124.2, 123.4, 121.9, 121.2, 121.2, 121.2, 121.0, 121.0, 120.9, 118.1, 118.0, 117.3, 117.2, 117.0, 116.4, 109.3, 107.4, 107.2, 107.0, 57.6, 57.5, 57.1, 44.5.44.4, 43.1, 43.0, 42.9, 42.7,42.6, 42.5, 42.3, 42.2, 41.3,

41.3, 41.2, 40.7, 40.3, 39.7, 39.6, 39.2, 39.1, 38.6, 38.6, 38.3, 38.1, 38.0, 37.9, 37.9, 36.4, 35.9, 29.8, 28.8, 28.7, 27.9.

HRMS (ESI) was calculated for 1178,$C_{75}H_{101}N_{31}O_{13}$, $[M+H]^+$, 1645.84 found 1645.93.

HPLC purity 99.5%

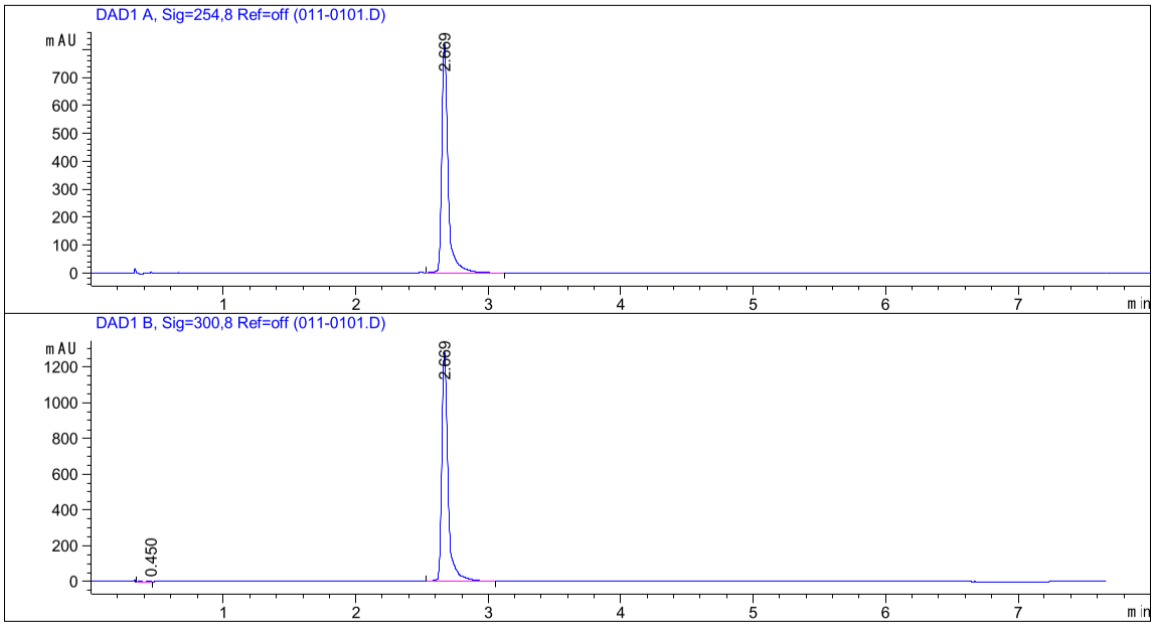


Figure 4.9. Analytical HPLC purity of compound **1178** with λ monitored at 300nm. Compound purity is 99.5%, retention time 3.357 min.

4.2.7 Synthesis of Boc-Py-Im-Im-β-Im-Py-γ-Im-β-Py-Py-Py-β-PAM resin

Synthesis of 1180.

Synthesis of polyamides **1180** was started by initiating the reaction sequences program with the CSBio program (**Table 4.5**).

Table 4.5 Programmed sequence of **1180** PAM resin synthesis.

Building block(s) and step(s)	Additional Solvent mixture	Applied program
1-Boc-Py-Py-COOH	(6.5 mL) DMF	Add-DIPEA-180-min-DIPEA-Flows.CFN
2-Capping	None	Cap-17-min-continues.CFN

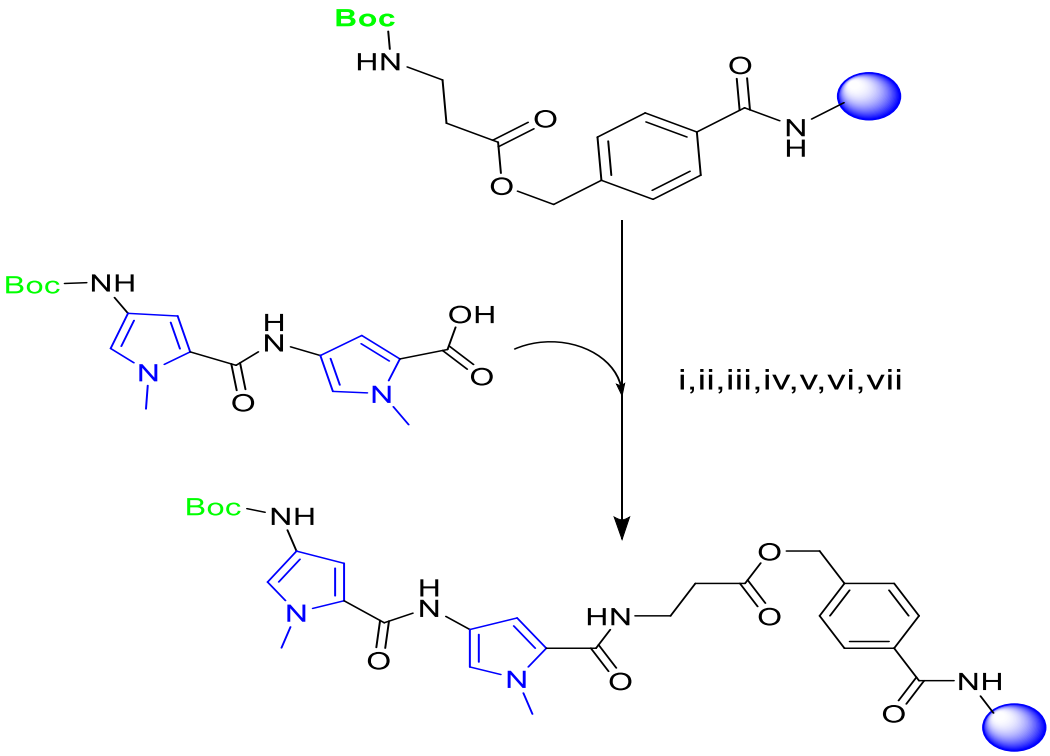
3-Boc-β-Py-COOH	(6.5 mL) DMF	Add-DIPEA-180-min-DIPEA-Flows.CFN
4-Capping	None	Cap-17-min-continues.CFN
5-Boc-γ-Im-COOH	(6.5 mL) DMF + DIPEA (1mL)	Add-DIPEA-180-min-DIPEA-Flows.CFN
6-Capping	None	Cap-17-min-continues.CFN
7-Boc-Py-OBt	(6.5 mL) DMF	Add-DIPEA-180-min-DIPEA-Flows.CFN
8-Capping	None	Cap-17-min-continues.CFN
9-Boc β-Im-COOH	(6.5 mL) DMF + DIPEA (1mL)	Add-DIPEA-180-min-DIPEA-Flows.CFN
10-Capping	None	Cap-17-min-continues.CFN
11-Boc-Py-Im-Im-COOH	(6.5 mL) DMF + DIPEA (1mL)	Add-DIPEA-180-min-DIPEA-Flows.CFN
12-Capping	None	Cap-17-min-continues.CFN

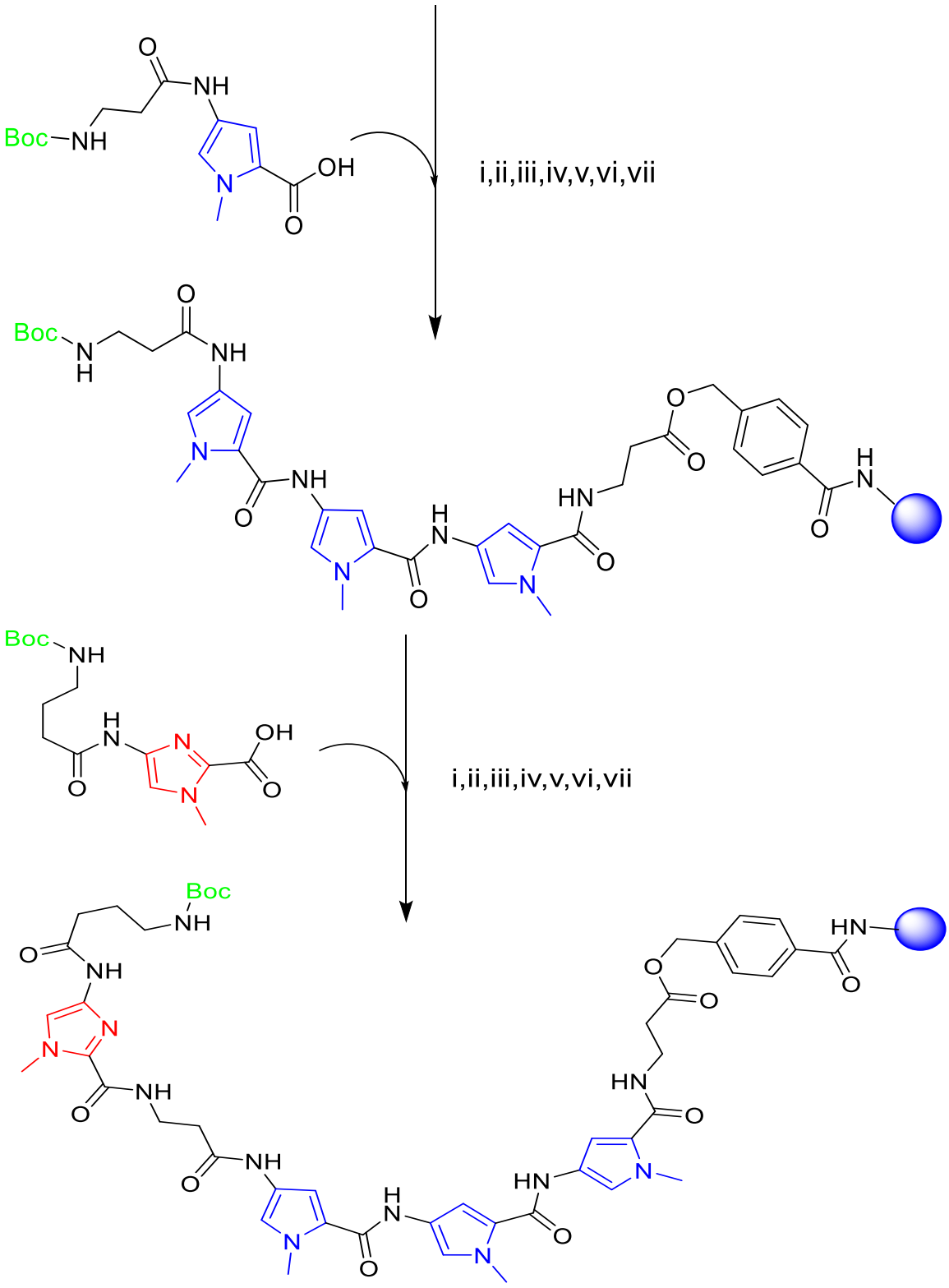
The CSBio reaction vessel was set at a temperature of 35^{o}C throughout the whole process. Boc-β-alanine-PAM resin (1000 mg) was loaded into the RV. PyBOP (11 g) was dissolved in DMF (40 mL) and filled in the R4 solvent bottle. After the program started the first wash and deprotection reaction for 1 h, a solution of Boc-Py-Py-COOH (503 mg, 1.38 mmol) in DMF (6.5 mL) was stirred to fully dissolved and transferred into AA#3. The coupling program directed the synthesizer to transfer the solution from AA#3 to TVA and then MVA to begin the activation before transferring to RV for 180 min, followed by capping with acetic anhydride for 17 min to remove impurities such as an unreacted amine. After completing the first sequence, the program proceeded to "cap" to remove any nonreacted amine before moving on to the next sequence. Each sequence took around 6 h to complete.

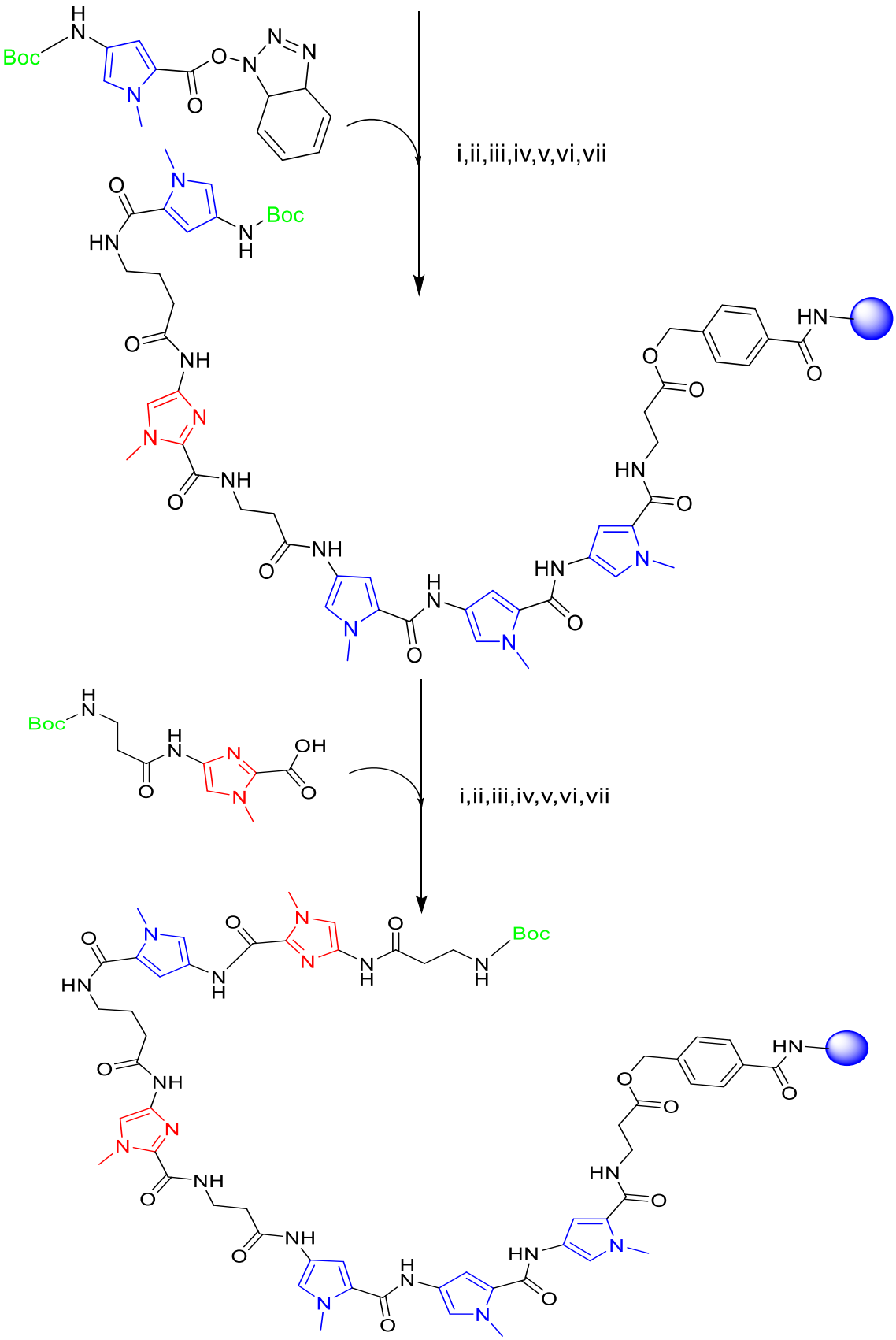
After completing the first sequence, the programmed system transferred a solution of Boc-β-Py-COOH (431 mg, 1.38 mmol) in DMF (6.5 mL), was well dissolved by vortex and loaded in AA#4 to TVA and MVA for sequence coupling followed up by capping. After sequence 3 plus cap was completed, a solution of Boc-γ-Im-COOH (452 mg, 1.38 mmol) and DIPEA (1 mL) in DMF (6.5 mL) was well dissolved by vortex; loaded in AA#5 transferred to TVA then MVA before transferred to RV for coupling reaction for 180 min followed by capping with acetic anhydride for 17 min.

After sequence 4 plus cap was completed, a solution of Boc-Py-Obt (658 mg, 1.84 mmol) in DMF was well dissolved by vortex; loaded in AA#6, transferred to TVA, then MVA before transferred to RV for coupling reaction for 180 min followed by capping with acetic anhydride for 17 min. After sequence 5 was finished, the programed system was initiated and followed by transferring a well-dissolved by vortex solution of Boc-β-Im-COOH (432 mg, 13.8 mmol) and DIPEA (1 mL) in DMF (6.5 mL) from AA#6 to TVA then MVA before transferred to RV for. After sequence 6 was finished, the programed system was initiated and followed by transferring a well-dissolved vortex solution of BocPy-Im-Im-COOH (671 mg, 1.38 mmol) and DIPEA (1 ml) in DMF (6.5 mL) from AA#8 to TVA then MVA before transferred to RV for coupling reaction for 180 min following by capping with acetic anhydride for 17 min. The sequence was finished after 31 h (**Figure 4.10**) 0,1 mg of resin was cleaved with 100 µL Ta solution for 2 h at 45ºC, and the heated solution was collected and filtered before being analyzed with a Bruker LC-MS instrument.

The exact mass was determined to be $[M+H]^+$ 1645.79 m/z, which confirmed the presence of the desired product. All the resin was carefully transferred from the RV into a beaker using a small spatula, and the resin was washed three times with DCM and MeOH, then dried under vacuum for 30 min until reaching dryness. The resin was then collected in a 20 mL scintillation vial and stored in the freezer.




Boc
NH
O
NH
OH
N
O
i,ii,iii,iv,v,vi,vii
Boc
N
H
O
NH
N
O
NH
H
N
O
N
HN
O
O
H
N
O
Boc
NH
H
N
O
N
N
OH
O
i,ii,iii,iv,v,vi,vii
Boc
NH
O
NH
N
N
O
NH
O
NH
N
H
N
O
N
HN
O
O
O
HN
O
HN
N
O
HN
O


i,ii,iii,iv,v,vi,vii
i,ii,iii,iv,v,vi,vii

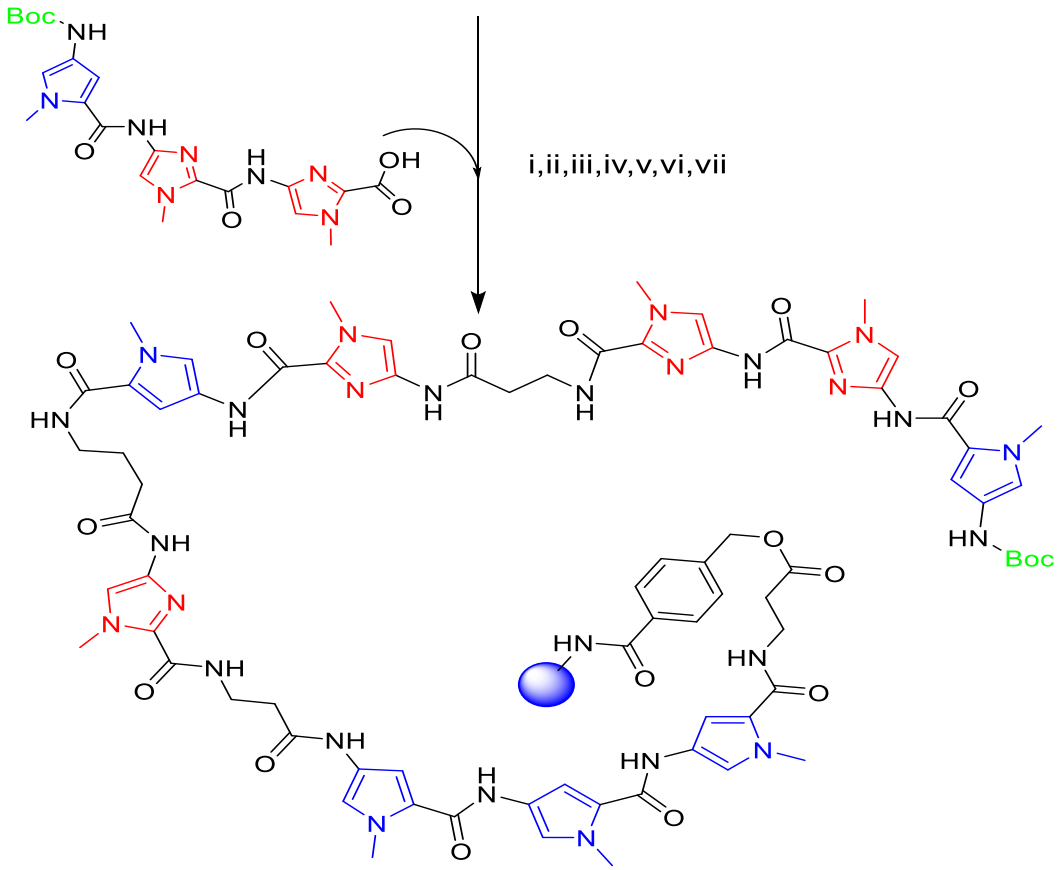


Figure 4.10: *Scheme of synthesis polyamide **1180** PAM resin synthesis. Started from a commercially acquired Boc-β-PAM resin, each sequence containing multiple steps from (i) to (viii). (i) wash the resin with DCM and DMF, (ii) deprotection with 60% TFA/ DCM /0.5 M Indole, (iii) 25% piperidine/ DMF; (iv) building blocks with the structures shown on the scheme, (v) 0.5M PyBOP/ DMF as coupling reagents, (vi) DIPEA, (vii) DMSO, (viii) capping step with acetic anhydride for 17 min.*

4.2.8 Synthesis of 1181.

Synthesis of polyamides **1181** resin started by initiating the reaction sequences program with the CSBio program (**Table 4.3**).

Table 4.6 Programmed sequence of **1181** PAM resin synthesis.

Building block(s) and step(s)	Additional Solvent mix-ture	Applied program
1 Wash	None	DMF washes. CFN
2-Capping	None	Cap_17min_continue.CFN

3-Boc_2Guan	(6.5 mL) DMF+DIPEA (1mL)	1 gram-pause-24h-DMF-Flows-Ben-edit.CFN
4-Deprotection	None	DMF-DCM-TFA-DCM-Flows. CFN

The CSBio reaction vessel was set at a temperature of 35°C throughout the whole process. Resin from the previous sequence, **1180** resin (516 mg), was loaded into the RV. The first step was washing the resin with DMF, followed by capping to remove any extra impurities. A solution of Bis-Boc-Pyrazolocarboxamidine (1.09g, 3.48 mmol) in DIPEA (1mL) and 6.5 mL of DMF was added into the RV 3h after step 3 began (**Table 4.6)**.

After 26 h, the program started the deprotection step using TFA to deprotect the Boc group from Boc_2Guan to introduce the Guan group in the N-terminus of the polyamide **1181**. (**Figure 4.11**) After the program finished, 0.1 mg of resin was cleaved with Ta as the nucleophile for 4h at 45°C; the sample was analyzed with a Bruker LC-MS instrument. The exact mass was determined to be $[M+H]^+$ as 1588.79 m/z, confirming the coupling reaction's completion. All the resin was carefully transferred from the RV into a beaker using a small spatula, and the resin was washed three times with DCM and MeOH, then dried under vacuum for 30 min until reaching dryness. The resin was then collected in a 20mL scintillation vial and stored in the freezer before cleavage in the next day.

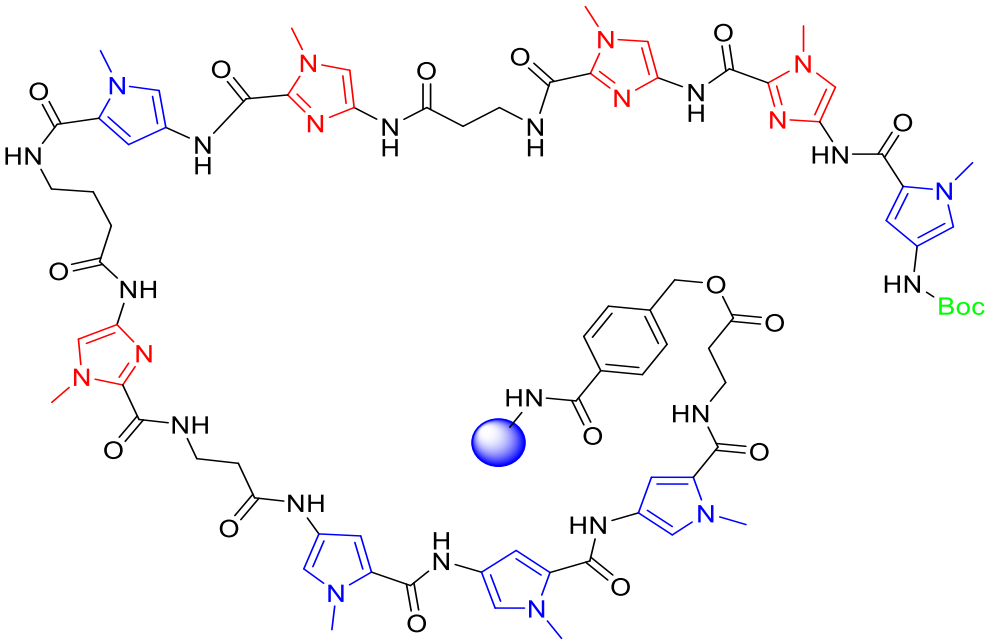


Figure 4.11. figure and description continue in the next pages.

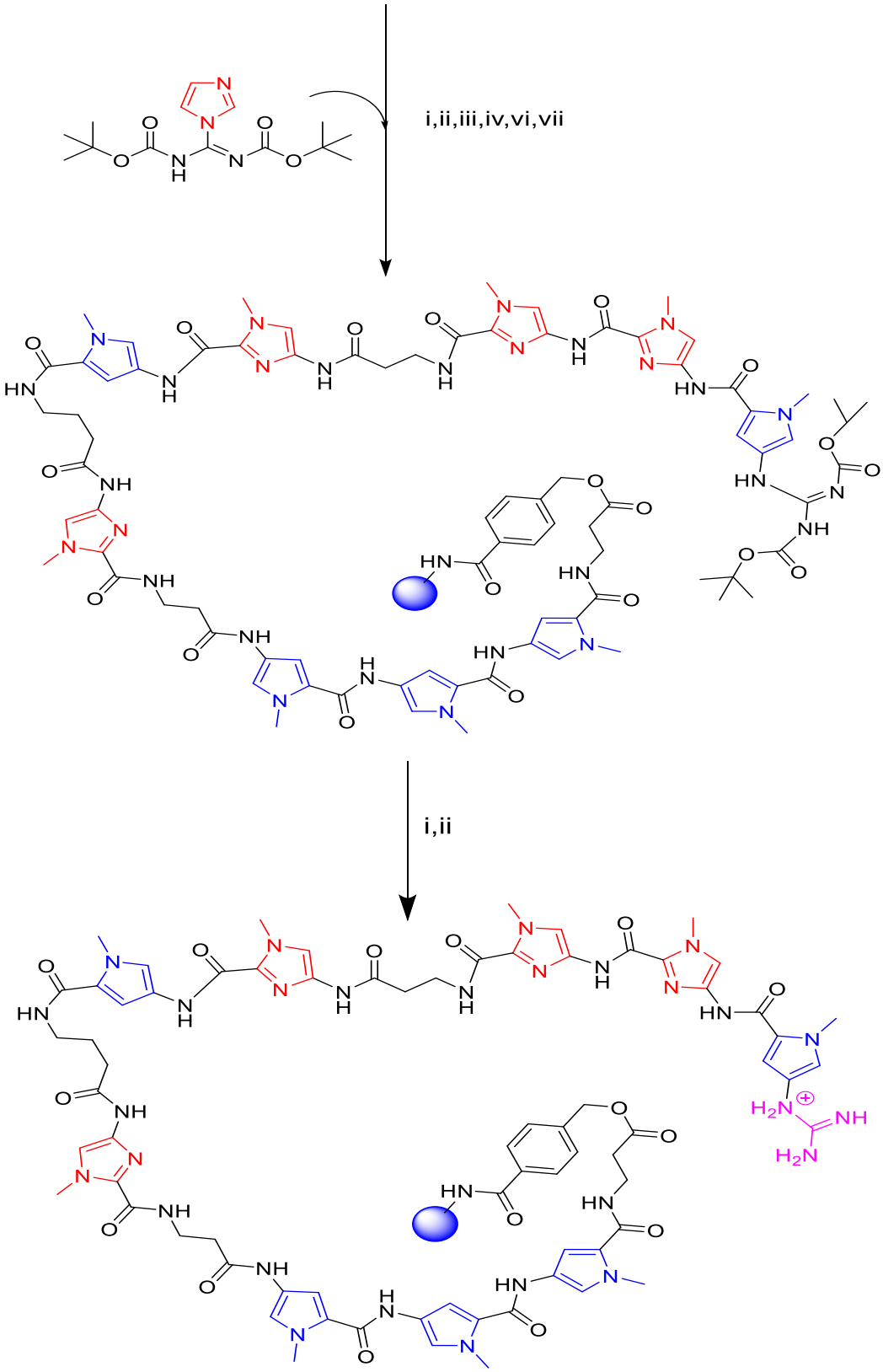

i,ii,iii,iv,vi,vii
i,ii

Figure 4.11: *Scheme of synthesis polyamide **1181** PAM resin synthesis. Started from a previously synthesized resin **1180**, each sequence containing multiple steps from (i) to (viii). (i) wash the resin with DCM and DMF, (ii) deprotection with 60% TFA/ DCM/ 0.5 M Indole,(iii) 25% piperidine/ DMF; (iv) building blocks with the structures shown on the scheme, (v) 0.5M PyBOP/ DMF as coupling reagents, (vi) DIPEA, (vii) DMSO, (viii) capping step with acetic anhydride for 17 min included during the procedure.*

Cleavage of PAM resin using Ta as a nucleophile.

The next day, 417 mg resin was cleaved with aminolysis in 2000 μL of Ta at 45ºC in a heating device Isotemp Fisher Scientific for 4 h, and the separation of polyamides solution from the PAM resin support was performed in a 20 μM polyethylene frit placed inside a disposable polypropylene syringe by washing with MeOH. The polyamide solution was concentrated by the rotary evaporator to afford off-white solid **1181** (47 mg, 0.096 mmol, 30.9% yield) (**Figure 4.10**).

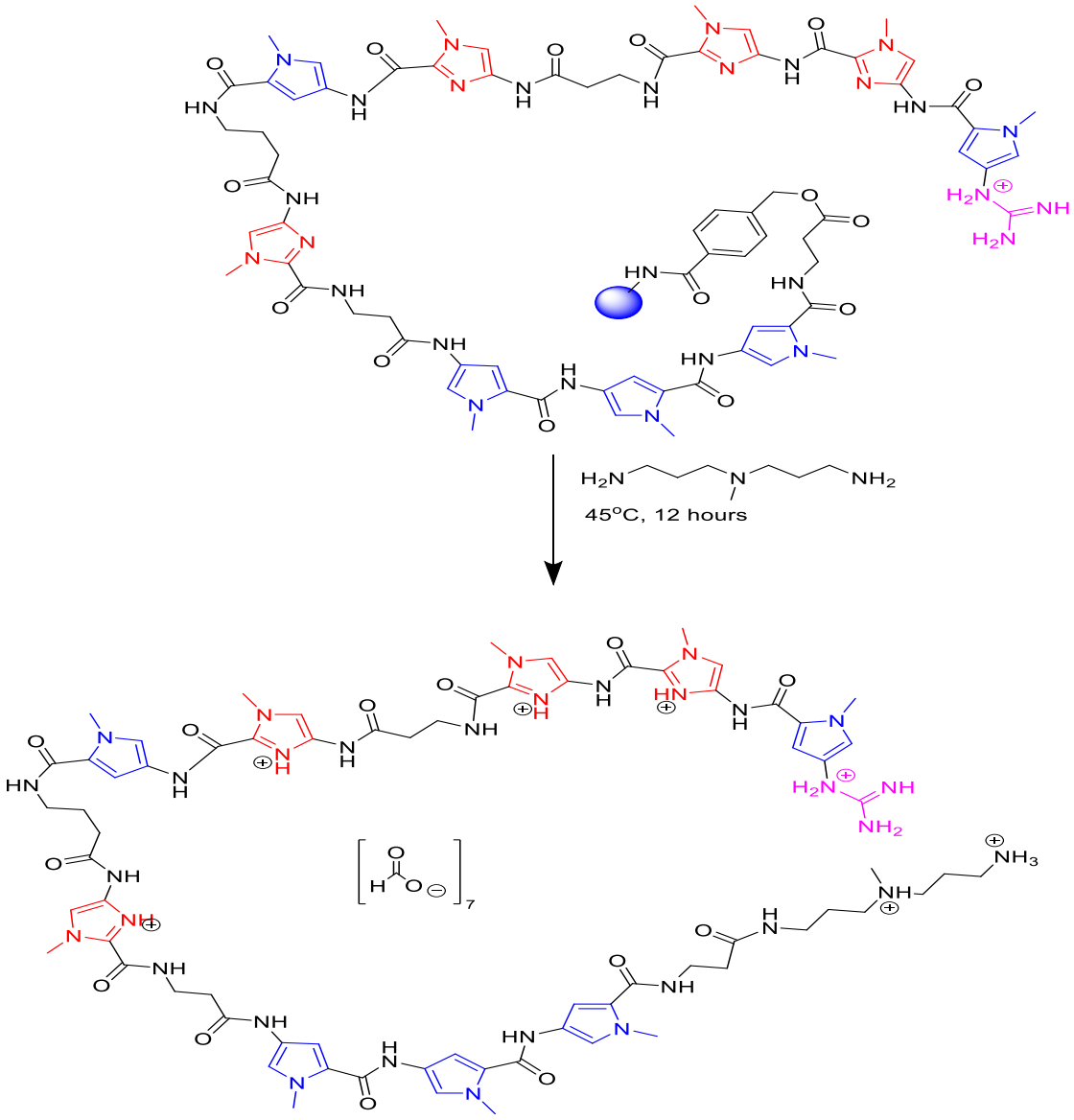


Figure 4.12 *Cleavage of **1181** from PAM resin using Ta as the nucleophile*.

Purification of the compound.

The polyamides solution was diluted with a mixture of DMSO (200 µL) and H_2O/0.2% HCOOH (200 µL), then purified by preparative HPLC using a Phenomenex Luna 250x30 mm, 5 µM, 100Å C18 column maintained at 25°C. The organic phase was 0.2% HCOOH in MilliQ water and 100% HPLC grade MeOH. The applied gradient was 10% MeOH for 8 min, followed by a ramp rate to 90% Methanol over 35.6 min at a 20 mL/min flow rate. All collected fractions were analyzed and selected for 95% or higher purity with analytical HPLC using ThermoFisher-UHPLC; analysis was processed and analyzed with the program Chromeleon 3.0. High-purity fractions were collected for rotary evaporation to remove MeOH before the lyophilization.

Lyophilization of the compound.

Compound 1181 was diluted with a 1:1 mixture of ACN (5 mL): and 0.2% HCOOH in H_2O (5 mL). The samples were frozen before the lyophilization process with ground dry ice until an even coated of frozen material results at an acute angle in the vial. The frozen sample was covered with porous Kimwipe paper and secured by a rubber band. The sample was settled in a lyophilizer glass vessel and connected to a vacuum port. The frozen polyamide solutions sublimed as an off-white solid powder after 48 h. The lyophilizer vacuum was set at 300 torr within 30 min prior to the lyophilization process, and the inside temperature was set at -83 to -85ºC.

Characterization of compound

Analytical HPLC characterization was performed with a C12 Phenomenex Juniper Proteo column maintained at 40ºC. The mobile phase consisted of 0.1% formic acid in water (A) and HPLC grade ACN (B). The applied gradient consisted of 5% B for 0.75 min, followed by the ramp to 60% B over 6.5 min at 1.2 mL/min. The retention was 2.533 min.

1181, Guan-Py-Im-Im-β-Im-Py- γ-Im-β-Py-Py-Py-β-Ta $(HCOO^-)_7$

^{1}H-NMR (600 MHz, DMSO d_6) δ= 8.46 (s,7H), 8.22 (s, 1H), 8.08(S,1H), 8.05 (s,1H), 8.01 (s, 1H), 7.95 (s, 1H), 7.64 (s, 1H), 7.54(s, 1H), 7.46 (s, 1H), 7.41 (s, 1H), 7.25 (s, 1H), 7.20 (s, 3H), 7.04-7.02 (s, J= 10.01 Hz, 2H), 7.01 (s, 1H), 6.95 (s, 1H), 6.88 (s, 1H), 6.85 (s, 1H), 4.01 (s, 3H), 3.96-3.95 (d, J= 6.20 Hz, 6H), 3.92 (s, 3H), 3.87 (s, 3H), 3.84 (s, 6H), 3.80- 3.79 (s, J= 9.29 Hz, 6H), 3.52-3.51 (d, J= 5.68 Hz, 4H), 3.38-3.36 (d, J= 5.25 Hz, 3H), 3.17-3.16 (d, J= 6.42 Hz, 2H), 3.06-3.05 (d, J= 5.83 Hz, 2H), 2.79 (s, 6H), 2.63 - 2.60 (t, J= 13.41 Hz, 2H), 2.55- 2.53 (t, J= 13.41 Hz, 2H), 2.37- 2.31 (q, J= 30.91 Hz, 10H), 2.26- 2.24 (t, J=14.0 Hz, 3H), 2.11 (s, 4H), 2.09 (s, 3H), 1.78 -1.75 (q, J= 19.24 Hz, 2H), 1.67- 1.66 (d, J= 6.99 Hz, 6H), 1.53- 1.51 (t, J= 13.41 Hz, 2H), 1.24 (s, 1H), 1.19 (s, 2H).

^{13}C-HNMR (151 MHz, DMSO d_6) δ= 173.6, 172.9, 171.8, 171.1, 169.5, 169.4, 169.4, 169.4, 169.3, 169.3, 164.3, 164.2, 161.6, 161.5, 161.3, 161.3, 160.5, 158.8, 158.3, 139.6, 139.2, 138.9, 137.7, 137.3, 136.6, 135.9, 127.1, 126.5, 126.3, 125.9, 125.8, 125.3, 125.2, 124.9, 124.2, 121.6, 121.3, 121.1, 121.0, 120.9, 118.0, 117.1, 116.7, 116.5, 116.4, 113.2, 107.8, 107.4, 107.3, 107.0, 57.6, 57.5, 57.1, 44.5, 44.4, 43.1, 43.0, 42.9, 42.7, 42.6, 42.5, 42.3, 42.2, 41.2, 40.7, 40.3, 39.7, 39.6, 39.2, 39.2, 39.1, 39.1, 38.6, 38.6, 38.2, 38.1, 38.0, 38.0, 37.9, 35.9, 29.8, 28.6, 27.9.

HRMS (ESI) was calculated for **1181**, $C_{71}H_{93}N_{31}O_{13}$ $[M+H]^+$ 1588.77 found, 1588.79 m/z

HPLC purity 95.7%

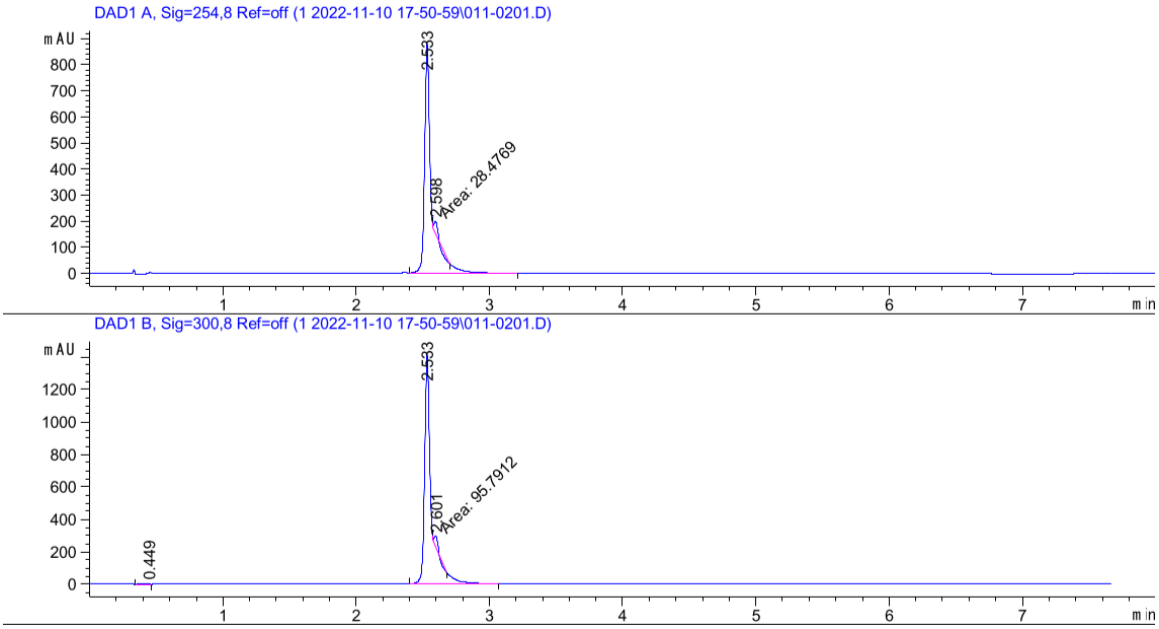


Figure 4.13. Analytical HPLC purity of compound **1181** with λ monitored at 300nm. Compound purity 95.7%, retention time 2.533 min.

4.2.7 Synthesis of TMG-Py-Im-Im-β-Im-Py-γ-Im-β-Py-Py-Py-β-Ta

Synthesis of 1182

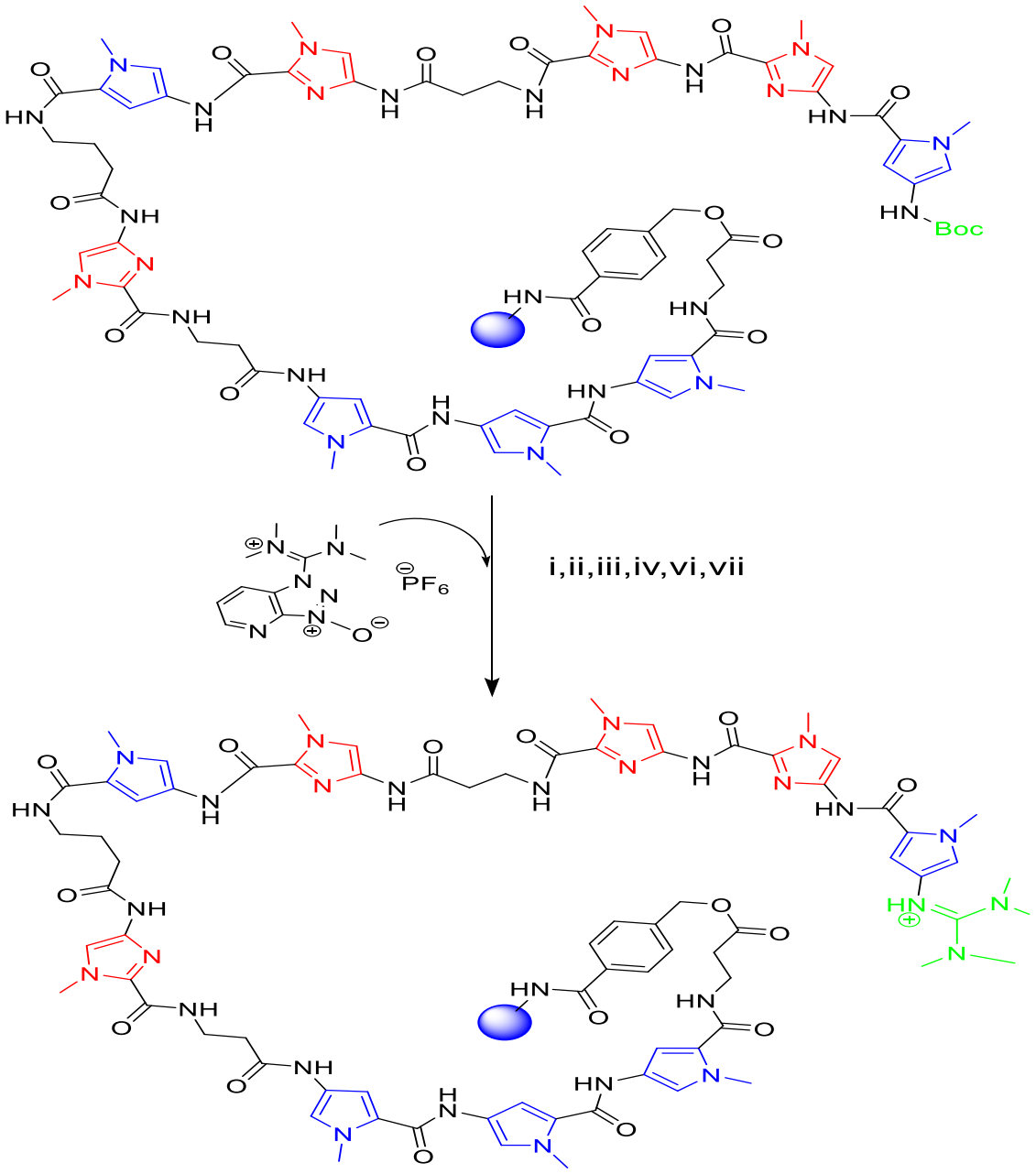


Figure 4.14: *Scheme of synthesis polyamide **1182** PAM resin synthesis. Started from a commercially acquired Boc-β-PAM resin, and each sequence contain multiple step from (i) to(viii). (i) wash resin with DCM and DMF, (ii) deprotection with 60% TFA/DCM/0.5 M Indole,(iii) 25% piperidine/DMF; (iv) building blocks with the structures shown on the scheme,(v) 0.5M PyBOP/DMF as coupling reagents, (vi) DIPEA, (vii) DMSO, (viii) capping step with acetic anhydride for 17 min.*

Synthesis of polyamides **1182** resin started by initiating the reaction sequences program with the CSBio program (**Table 4.7**). The CSBio reaction vessel was set at a temperature of 35°C

throughout the whole process. The previously synthesized **1180** resin (0.562 mg) was loaded into the RV. The first step was washing the resin with DMF, followed by capping to remove any extra impurities. The third program sequence is coupling HATU (Hexafluorophosphate Azabenzotriazole Tetramethyl Uronium) to the PIP sequence; after 3h since the sequence began, HATU (380 mg,1.38 mmol). After the program finished, 0.1 mg of resin was cleaved with Ta as the nucleophile for 4h at 45ºC; the sample was analyzed with a Bruker LC-MS instrument. The exact mass was determined to be $[M+H]^+$ as 1588.78 m/z, confirming the coupling reaction's completion. All the resin was carefully transferred from the RV into a beaker using a small spatula, and the resin was washed three times with DCM and MeOH, then dried under vacuum for 30 min until reaching dryness. The resin was then collected in a 20 mL scintillation vial and stored in the freezer before cleavage the next day.

Table 4.7 Program sequence of **1182** PAM resin synthesis.

Building block(s) and step(s)	Additional Solvent mixture	Applied program
1-Wash resin	None	DMF washes.CFN
2-Capping	None	Cap-17-min-continues.CFN
3-HATU solution	(6.5 mL) DMF	AddTMG_DIEA_180_DMSO_DIEA_flows.CFN

Cleavage of PAM resin using Ta as a nucleophile.

The next day, 441 mg resin was cleaved with aminolysis in 2000 µL of Ta at 45ºC in a heating device Isotemp Fisher Scientific for 12 h, and the separation of polyamides solution from the PAM resin support was performed in a 20 µM polyethylene frit placed inside a disposable polypropylene syringe by washing with MeOH and H_2O. The polyamide solution was concentrated by rotary evaporator to afford off-white solid of **1182** (57 mg, 0.101 mmol, 34.2% yield) (**Figure 4.13**).

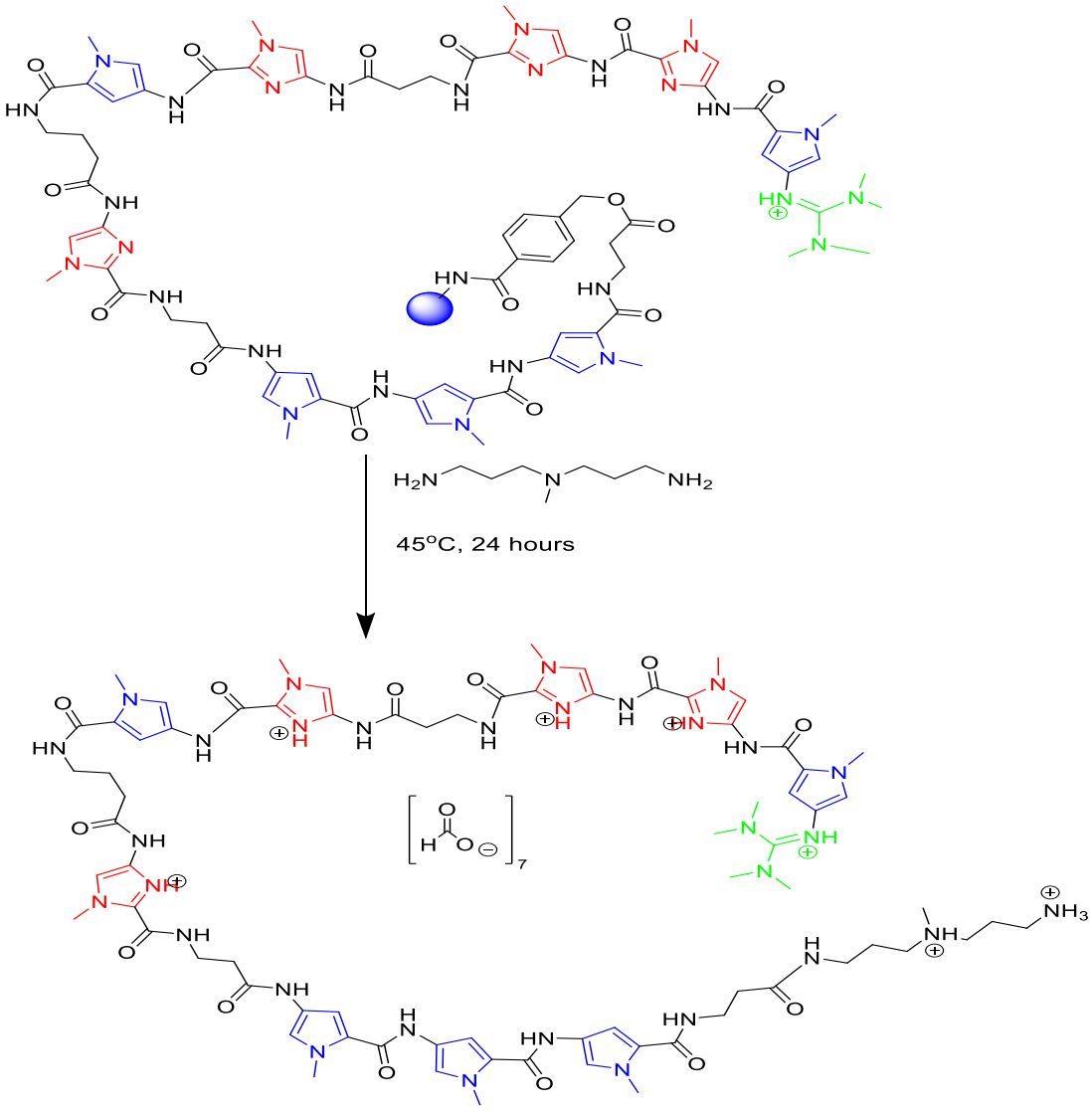


Figure 4.15 *Cleavage of **1182** from PAM resin using Ta as the nucleophile*.

Purification of the compound.

The polyamides solution was diluted with a mixture of DMSO (200 µL) and H_2O/ 0.2% HCOOH (200 µL), then purified by preparative HPLC using a Phenomenex Luna 250x30 mm, 5 µM, 100Å C18 column maintained at 25 °C. The organic phase was 0.2% HCOOH in MilliQ water and 100% HPLC grade MeOH. The applied gradient was 10% MeOH for 8 min, followed by a ramp to 90% Methanol over 35.6 min at a 20 mL/min flow rate. All collected fractions were analyzed and selected for 95% or higher purity with analytical HPLC using ThermoFisher-UHPLC; analysis was processed and analyzed with the program Chromeleon 3.0. High-purity fractions were collected for rotary evaporation to remove MeOH before lyophilization.

Lyophilization of compound

Compound 1182 was diluted with a 1:1 mixture of ACN (5 mL) and 0.2% HCOOH in H_2O (5 mL). The samples were frozen before the lyophilization process with ground dry ice until an even coated of frozen material results at an acute angle in the vial. The frozen sample was covered with porous Kimwipe paper and secured by a rubber band. The sample was settled in a lyophilizer glass vessel and connected to a vacuum port. The frozen polyamide solution sublimed as an off-white solid powder after 48 h. The lyophilizer vacuum was set at 300 torr within 30 min prior to the lyophilization process, and the inside temperature was set at -83 to -85ºC.

Characterization of compound

Analytical HPLC characterization was performed with a C12 Phenomenex Juniper Proteo column maintained at 40ºC. The mobile phase consisted of 0.1% formic acid in water (A) and HPLC grade ACN (B). The applied gradient consisted of 5% B for 0.75 min, followed by a ramp to 60% B over 6.5 min at 2.0 mL/min. The retention was 2.684 min.

1182 with ionized data

^{1}H-NMR (600MHz, DMSO d_6) δ= 8.47 (s, 5H), 8.18 (s, 1H), 8.08 (s, 1H), 8.05 (s, 1H), 8.01 (s, 1H), 7.95 (s, 1H), 7.65 (s, 1H), 7.54 (s, 1H), 7.46 (s, 1H), 7.41 (s, 1H), 7.25 (s, 1H), 7.20- 7. 19(d, J= 7.12 Hz, 4H), 7.05(s, 1H), 6.96 (s, 1H), 6.89 (s, 2H), 6.85 (s, 1H), 6.80 (s, 1H), 3.99 (s, 3H), 3.96-3.96 (d, J= 4.25 Hz, 3H), 3.92 (s, 3H), 3.87 (s, 3H), 3.84- 3.83 (d, J= 2.73 Hz, 6H), 3.79- 3.78 (d, J= 7.78 Hz, 6H), 3.52- 3.51(d, J= 5.45 Hz, 4H), 3.38- 3.34 (q, J= 22.19 Hz, 3H), 3.18- 3.15 (q, J= 20.63 Hz, 3H), 3.07-3.03 (q, J= 22.18 Hz, 3H), 2.92 (br, 8H), 2.79 (s, 6H), 2.64-2.61 (t, J= 13.99 Hz, 3H), 2.56- 2.52 (t, J= 13.57 Hz, 3H), 2.33-2.32 (d, J= 7.31 Hz, 10H), 2.26-2.24 (t, J= 12.53 Hz, 3H), 2.09 (s, 3H), 2.08 (s, 3H) 1.77-1.74 (t, J= 13.57 Hz, 2H), 1.66 (br, 6H), 1.54- 1.50 (t, J= 14.62 Hz, 3H), 1.24 (s, 1H).

^{13}C-NMR (151 MHz, DMSO d_6) δ= 170.5, 169.9, 168.9, 168.0, 165.8, 165.7, 165.7, 161.3,161.2, 158.5, 158.4, 158.2, 158.2, 157.9, 155.7, 155.2, 136.6, 136.2, 135.8, 134.6, 134.2, 134.1, 133.5, 132.8, 123.3, 123.3, 123.1, 122.7, 122.6, 122.2, 122.1, 121.8, 121.7, 121.1, 120.4, 118.5, 118.4, 118.2, 118.0, 117.9, 117.6, 115.0, 114.2, 113.6, 113.6, 113.3, 113.3, 106.2, 104.7, 104.3, 104.2, 103.9, 54.6, 54.4, 54.0, 41.4, 41.3, 40.1, 39.9, 39.8, 39.6, 39.5, 39.4, 39.2, 39.1, 38.1, 37.6, 37.2, 36.6, 36.5, 36.1, 36.1, 36.1, 35.9, 35.5, 35.5, 35.2, 35.1, 35.0, 34.9, 34.8, 32.8, 30.6, 26.8, 25.5, 24.9.

HRMS (ESI) was calculated for **1182**, $C_{75}H_{101}N_{31}O_{13}$, $[M+H]^+$ 1644.82 found 1644. 78.

HPLC purity. 99.1%

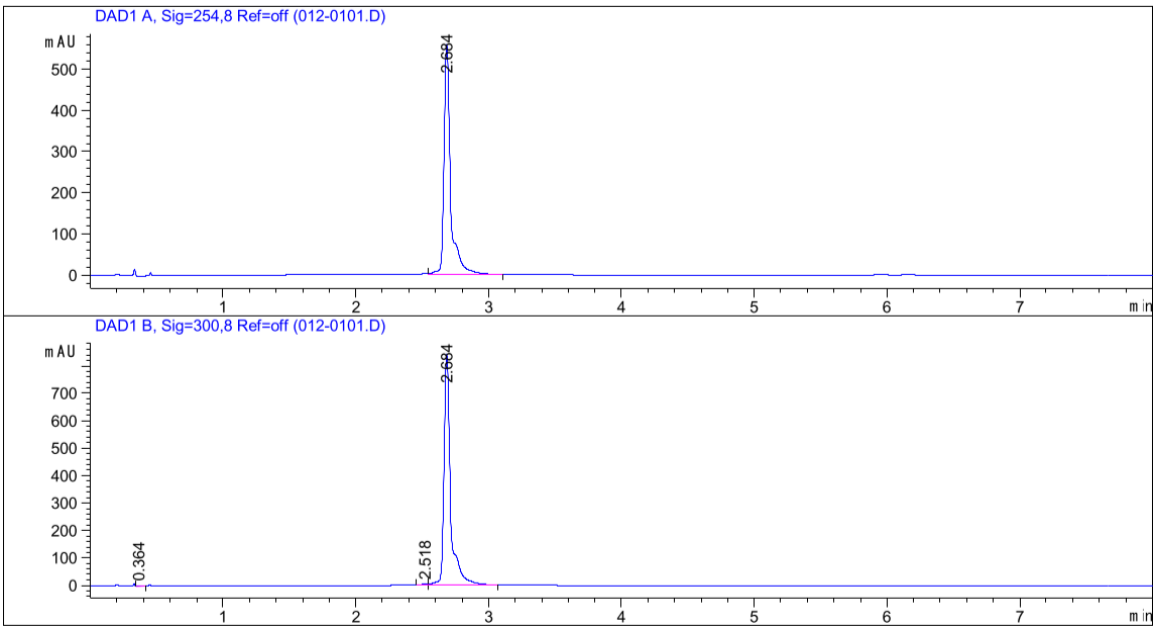


Figure 4.16. Analytical HPLC purity of compound **1182** with λ monitored at 300nm. Compound purity 99.1%, retention time 2.68 min.

4.3 Conclusion

We successfully modified and introduced the additional cationic groups to the PIP sequences in this chapter. Based on the previous results of our compounds, we expected to learn more about the binding affinity and selectivity of SETMAR TIR: PIPs binding complexes. The additional charges to the PIP sequence can further improve the selectivity of the compounds and hopefully avoid the aggregation of the compounds. We also improved the success of our PIP's coupling and avoided the strong electron-withdrawing effect of the N-terminus imidazoles by adding a pyrrole into the N-terminus of our polyamides.

4.4 Bibliography

1. Wemmer DE, Dervan PB. Targeting the minor groove of DNA. *Curr Opin Struct Biol*. 1997;7(3):355-361. doi:https://doi.org/10.1016/S0959-440X(97)80051-6

2. Kopka ML, Yoon C, Goodsell D, Pjura P, Dickerson RE. The molecular origin of DNA-drug specificity in netropsin and distamycin. *Proc Natl Acad Sci U S A*. 1985;82(5):1376-1380. doi:10.1073/pnas.82.5.1376

3. Cabeen M, Aaron G, Christine H, et al. *DNA Structure & Chemistry*. Harvard University; 2020. https://projects.iq.harvard.edu/lifesciences1abookv1

4. Blackledge MS, Melander C. Programmable DNA-binding small molecules. *Bioorganic Med Chem*. 2013;21(20):6101-6114. doi:10.1016/j.bmc.2013.04.023

5. White S, Szewczyk JW, Turner JM, Baird EE, Dervan PB. Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands. *Nature*. 1998;391(6666):468-471. doi:10.1038/35106

6. Bailly C, Chaires JB. Sequence-Specific DNA Minor Groove Binders. Design and Synthesis of Netropsin and Distamycin Analogues. *Bioconjug Chem*. 1998;9(5):513-538. doi:10.1021/bc980008m

7. Coll M, Frederick CA, Wang AH, Rich A. A bifurcated hydrogen-bonded conformation in the d(A.T) base pairs of the DNA dodecamer d(CGCAAATTTGCG) and its complex with distamycin. *Proc Natl Acad Sci U S A*. 1987;84(23):8385-8389. doi:10.1073/pnas.84.23.8385

8. Fechter EJ, Olenyuk B, Dervan PB. Sequence-specific fluorescence detection of DNA by polyamide thiazole orange conjugates. *J Am Chem Soc*. 2005;127(47):16685-16691. doi:10.1021/ja054650k

9. Kopka ML, Yoon C, Goodsell D, Pjura P, Dickerson RE. Binding of an Antitumour Drug to DNA Netropsin and CGCGAATTCGCG. Published online 1985.

10. Alniss HY. Thermodynamics of DNA Minor Groove Binders. *J Med Chem*. 2019;62(2):385-402. doi:10.1021/acs.jmedchem.8b00233

11. Peter B. Dervan. Molecular Recognition of DNA by small molecule. 2001;9(July):2-57.

12. J.William L, Krowicki K, Bhat G. Molecular recognition between oligopeptidesand nucleic acids: novel imidazole-containing oligopeptides related tonetropsin that exhibit altered DNA sequence specificity. *Biochemistry*. 1986;25:7408–7416. doi:10.1002/(SICI)1099-1409(200006/07)4:4<393::AID-JPP227>3.3.CO;2-2

13. Design of Sequence-Specific DNA-Binding Molecules Author (s): Peter B . Dervan Published by : American Association for the Advancement of Science Stable URL : http://www.jstor.org/stable/1696414. *Am Assoc Adv Sci*. 1986;232(4749):464-471.

14. Mrksich M, Wade WS, Dwyer TJ, Geierstanger BH, Wemmer DE, Dervan PB. Antiparallel side-by-side dimeric motif for sequence-specific recognition in the minor groove of DNA by the designed peptide 1-methylimidazole-2-carboxamide netropsin. *Proc Natl Acad Sci U S A*. 1992;89(16):7586-7590. doi:10.1073/pnas.89.16.7586

15. Wade WS, Mrksich M, Dervan PB. Design of Peptides That Bind in the Minor Groove of DNA at 5′-(A,T)G(A,T)C(A,T)-3′ Sequences by a Dimeric Side-by-Side Motif. *J Am Chem Soc*. 1992;114(23):8783-8794. doi:10.1021/ja00049a006

16. Koeller KJ. DNA Binding Polyamides and the Importance of DNA Recognition in their use as Gene-Specific and Antiviral Agents. *Med Chem (Los Angeles)*. 2014;04(02):338-344. doi:10.4172/2161-0444.1000162

17. Castaneda CH, Scuderi MJ, Edwards TG, et al. Improved antiviral activity of a polyamide against high-risk human papillomavirus: Via N-terminal guanidinium substitution. *Medchemcomm*. 2016;7(11):2076-2082. doi:10.1039/c6md00371k

18. Song Y, Niederschulte J, Bales KN, Park AH, Bashkin JK, Dupureur CM. DNA binding thermodynamics and site stoichiometry as a function of polyamide size. *Biochimie*. 2019;165:170-178. doi:10.1016/j.biochi.2019.07.021

19. Trauger JW, Baird EE, Dervan PB. Extended hairpin polyamide motif for sequence-

specific recognition in the minor groove of DNA. *Chem Biol*. 1996;3(5):369-377. doi:10.1016/S1074-5521(96)90120-9

20. White S, Baird EE, Dervan PB. On the pairing rules for recognition in the minor groove of DNA by pyrrole-imidazole polyamides. *Chem Biol*. 1997;4(8):569-578. doi:10.1016/S1074-5521(97)90243-X

21. Han YW, Kashiwazaki G, Morinaga H, et al. Effect of single pyrrole replacement with β-alanine on DNA binding affinity and sequence specificity of hairpin pyrrole/imidazole polyamides targeting 5'-GCGC-3'. *Bioorg Med Chem*. 2013;21(17):5436-5441. doi:10.1016/j.bmc.2013.06.005

22. Kawamoto Y, Bando T, Sugiyama H. Sequence-specific DNA binding Pyrrole–imidazole polyamides and their applications. *Bioorganic Med Chem*. 2018;26(8):1393-1411. doi:10.1016/j.bmc.2018.01.026

23. Neely L, Trauger JW, Baird EE, Dervan PB, Gottesfeld JM. Importance of minor groove binding zinc fingers within the transcription factor IIIA-DNA complex. *J Mol Biol*. 1997;274(4):439-445. doi:10.1006/jmbi.1997.1411

24. Dickinson LA, Gulizia RJ, Trauger JW, et al. Inhibition of RNA polymerase II transcription in human cells by synthetic DNA-binding ligands. *Proc Natl Acad Sci U S A*. 1998;95(22):12890-12895. doi:10.1073/pnas.95.22.12890

25. Olenyuk BZ, Zhang GJ, Klco JM, Nickols NG, Kaelin WG, Dervan PB. Inhibition of vascular endothelial growth factor with a sequence-specific hypoxia response element antagonist. *Proc Natl Acad Sci U S A*. 2004;101(48):16768-16773. doi:10.1073/pnas.0407617101

26. Mysore VS, Szablowski J, Dervan PB, Frost PJ. A DNA-binding Molecule Targeting the Adaptive Hypoxic Response in Multiple Myeloma Has Potent Antitumor Activity. *Mol Cancer Res*. 2016;14(3):253-266. doi:10.1158/1541-7786.MCR-15-0361

27. Yao EH, Fukuda N, Ueno T, et al. A pyrrole–imidazole polyamide targeting transforming growth factor-β1 inhibits restenosis and preserves endothelialization in the injured artery. *Cardiovasc Res*. 2008;81(4):797-804. doi:10.1093/cvr/cvn355

28. Martínez TF, Phillips JW, Karanja KK, et al. Replication stress by Py–Im polyamides induces a non-canonical ATR-dependent checkpoint response. *Nucleic Acids Res*. 2014;42(18):11546-11559. doi:10.1093/nar/gku866

29. Schaal TD, Mallet WG, McMinn DL, et al. Inhibition of human papilloma virus E2 DNA binding protein by covalently linked polyamides. *Nucleic Acids Res*. 2003;31(4):1282-1291. doi:10.1093/nar/gkg206

30. Edwards TG, Vidmar TJ, Koeller K, Bashkin JK, Fisher C. DNA Damage Repair Genes Controlling Human Papillomavirus (HPV) Episome Levels under Conditions of Stability and Extreme Instability. *PLoS One*. 2013;8(10). doi:10.1371/journal.pone.0075406

31. Oyoshi T, Kawakami W, Narita A, Bando T, Sugiyama H. Inhibition of transcription at a coding sequence by alkylating polyamide. *J Am Chem Soc*. 2003;125(16):4752 – 4754. doi:10.1021/ja029196o

32. Shinohara K ichi, Sasaki S, Minoshima M, Bando T, Sugiyama H. Alkylation of template strand of coding region causes effective gene silencing. *Nucleic Acids Res*. 2006;34(4):1189-1195. doi:10.1093/nar/gkl005

33. Jespersen C, Soragni E, James Chou C, Arora PS, Dervan PB, Gottesfeld JM. Chromatin structure determines accessibility of a hairpin polyamide–chlorambucil conjugate at histone H4 genes in pancreatic cancer cells. *Bioorg Med Chem Lett*. 2012;22(12):4068-4071. doi:https://doi.org/10.1016/j.bmcl.2012.04.090

34. Lin J, Hiraoka K, Watanabe T, et al. Identification of binding targets of a pyrrole-imidazole polyamide KR12 in the LS180 colorectal cancer genome. *PLoS One*. 2016;11(10):1-19. doi:10.1371/journal.pone.0165581

35. Wang YD, Dziegielewski J, Wurtz NR, Dziegielewska B, Dervan PB, Beerman TA. DNA crosslinking and biological activity of a hairpin polyamide-chlorambucil conjugate. *Nucleic Acids Res*. 2003;31(4):1208 – 1215. doi:10.1093/nar/gkg215

36. Wicker T, Sabot F, Hua-Van A, et al. Reply: A unified classification system for eukaryotic transposable elements should reflect their phylogeny. *Nat Rev Genet*. 2009;10(4):276. doi:10.1038/nrg2165-c4

37. Ravindran S. Barbara McClintock and the discovery of jumping genes. *Proc Natl Acad Sci U S A*. 2012;109(50):20198-20199. doi:10.1073/pnas.1219372109

38. McCLINTOCK B. The origin and behavior of mutable loci in maize. *Proc Natl Acad Sci U S A*. 1950;36(6):344-355. doi:10.1073/pnas.36.6.344

39. Finnegan DJ. Eukaryotic transposable elements and genome evolution. *Trends Genet*. 1989;5(C):103-107. doi:10.1016/0168-9525(89)90039-5

40. Robertson HM. The Tcl-mariner superfamily of transposons in animals. *J Insect Physiol*. 1995;41(2):99-105. doi:https://doi.org/10.1016/0022-1910(94)00082-R

41. Sinzelle L, Izsvák Z, Ivics Z. Molecular domestication of transposable elements: From detrimental parasites to useful host genes. *Cell Mol Life Sci*. 2009;66(6):1073-1093. doi:10.1007/s00018-009-8376-3

42. Ågren JA, Clark AG. Selfish genetic elements. *PLOS Genet*. 2018;14(11):e1007700. doi:10.1371/journal.pgen.1007700

43. Feschotte C, Pritham EJ. DNA transposons and the evolution of eukaryotic genomes. *Annu Rev Genet*. 2007;41(35):331-368. doi:10.1146/annurev.genet.40.110405.090448

44. Pace JK, Feschotte C. The evolutionary history of human DNA transposons: Evidence for intense activity in the primate lineage. *Genome Res*. 2007;17(4):422-432. doi:10.1101/gr.5826307

45. Cordaux R, Udit S, Batzer MA, Feschotte C. Birth of a chimeric primate gene by capture of the transposase gene from a mobile element. *Proc Natl Acad Sci U S A*. 2006;103(21):8101-8106. doi:10.1073/pnas.0601161103

46. Ayarpadikannan S, Kim HS. The Impact of Transposable Elements in Genome Evolution and Genetic Instability and Their Implications in Various Diseases. *Genomics Inform*. 2014;12(3):98. doi:10.5808/gi.2014.12.3.98

47. Robertson HM. The mariner transposable element is widespread in insects. *Nature*. 1993;362(6417):241-245. doi:10.1038/362241a0

48. Plasterk RH. The Tc1/mariner transposon family. *Curr Top Microbiol Immunol*. 1996;204:125-143. doi:10.1007/978-3-642-79795-8_6

49. Lampe DJ, Akerley BJ, Rubin EJ, Mekalanos JJ, Robertson HM. Hyperactive transposase mutants of the Himar1 mariner transposon. *Proc Natl Acad Sci U S A*. 1999;96(20):11428-11433. doi:10.1073/pnas.96.20.11428

50. Smit AF, Riggs AD. Tiggers and DNA transposon fossils in the human genome. *Proc Natl Acad Sci U S A*. 1996;93(4):1443-1448. doi:10.1073/pnas.93.4.1443

51. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature*. 2001;409(6822):860-921. doi:10.1038/35057062

52. Robertson HM, Zumpano KL. Molecular evolution of an ancient mariner transposon, Hsmar1, in the human genome. *Gene*. 1997;205(1-2):203-217. doi:10.1016/s0378-1119(97)00472-1

53. Robertson HM, Martos R. Molecular evolution of the second ancient human mariner transposon, Hsmar2, illustrates patterns of neutral evolution in the human genome lineage. *Gene*. 1997;205(1-2):219-228. doi:10.1016/s0378-1119(97)00471-x

54. Lié O, Renault S, Augé-Gouillou C. SETMAR, a case of primate co-opted genes: towards new perspectives. *Mob DNA*. 2022;13(1):1-14. doi:10.1186/s13100-022-00267-1

55. Qian C, Zhou MM. SET domain protein lysine methyltransferases: Structure, specificity and catalysis. *Cell Mol Life Sci*. 2006;63(23):2755-2763. doi:10.1007/s00018-006-6274-5

56. Kuzmichev A, Nishioka K, Erdjument-Bromage H, Tempst P, Reinberg D. Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. *Genes Dev*. 2002;16(22):2893-2905. doi:10.1101/gad.1035902

57. Liu D, Bischerour J, Siddique A, Buisine N, Bigot Y, Chalmers R. The Human SETMAR Protein Preserves Most of the Activities of the Ancestral Hsmar1 Transposase . *Mol Cell Biol*. 2007;27(3):1125-1132. doi:10.1128/mcb.01899-06

58. Kim HS, Kim SK, Hromas R, Lee SH. The set domain is essential for metnase functions in replication restart and the 5' end of ss-overhang cleavage. *PLoS One*. 2015;10(10):1-16. doi:10.1371/journal.pone.0139418

59. Chen Q. A ROLE FOR SETMAR IN GENE REGULATION : INSIGHTS FROM STRUCTURAL ANALYSIS OF THE DNA-BINDING DOMAIN IN COMPLEX WITH DNA. 2016;(August).

60. Higgins JJ, Pucilowska J, Lombardi RQ, Rooney JP. Candidate genes for recessive non-syndromic mental retardation on chromosome 3p (MRT2A)*. *Clin Genet*. 2004;65(6):496-500. doi:https://doi.org/10.1111/j.0009-9163.2004.00267.x

61. Kaur E, Nair J, Ghorai A, et al. Inhibition of SETMAR-H3K36me2-NHEJ repair axis in residual disease cells prevents glioblastoma recurrence. *Neuro Oncol*. 2020;22(12):1785-1796. doi:10.1093/neuonc/noaa128

62. Dussaussois-Montagne A, Jaillet J, Babin L, et al. SETMAR isoforms in glioblastoma: A matter of protein stability. *Oncotarget*. 2017;8(6):9835-9848. doi:10.18632/oncotarget.14218

63. Lee SH, Oshige M, Durant ST, et al. The SET domain protein Metnase mediates foreign DNA integration and links integration to nonhomologous end-joining repair. *Proc Natl Acad Sci U S A*. 2005;102(50):18075-18080. doi:10.1073/pnas.0503676102

64. Wray J, Williamson EA, Chester S, et al. The transposase domain protein Metnase/SETMAR suppresses chromosomal translocations. *Cancer Genet Cytogenet*. 2010;200(2):184-190. doi:10.1016/j.cancergencyto.2010.04.011

65. Chen Q, Bates AM, Hanquier JN, et al. Structural and genome-wide analyses suggest that transposon-derived protein SETMAR alters transcription and splicing. *J Biol Chem*. 2022;298(5):101894. doi:10.1016/j.jbc.2022.101894

66. Chen Q, Georgiadis M. Crystallization of and selenomethionine phasing strategy for a SETMAR-DNA complex. *Acta Crystallogr Sect Struct Biol Commun*. 2016;72:713-719. doi:10.1107/S2053230X16012723

67. Swalley SE, Baird EE, Dervan PB. Effects of γ-turn and β-tail amino acids on sequence-specific recognition of DNA by hairpin polyamides. *J Am Chem Soc*. 1999;121(6):1113-1120. doi:10.1021/ja9830905

68. Liu B, Wang S, Aston K, et al. β-Alanine and N-terminal cationic substituents affect polyamide-DNA binding. *Org Biomol Chem*. 2017;15(46):9880-9888. doi:10.1039/c7ob02513k

69. Tellier M, Chalmers R. The roles of the human SETMAR (Metnase) protein in illegitimate DNA recombination and non-homologous end joining repair. *DNA Repair (Amst)*. 2019;80(March):26-35. doi:10.1016/j.dnarep.2019.06.006

70. Wang S, Aston K, Koeller KJ, et al. Modulation of DNA-polyamide interaction by β-alanine substitutions: A study of positional effects on binding affinity, kinetics and thermodynamics. *Org Biomol Chem*. 2014;12(38):7523-7536. doi:10.1039/c4ob01456a

71. Jaramillo D, Liu Q, Aldrich-Wright J, Tor Y. Synthesis of N-Methylpyrrole and N-Methylimidazole Amino Acids Suitable for Solid-Phase Synthesis. *J Org Chem*. 2004;69(23):8151-8153. doi:10.1021/jo048686r

72. Baird EE, Dervan PB. Solid Phase Synthesis of Polyamides Containing Imidazole and Pyrrole Amino Acids. 1996;7863(9):6141-6146.

73. Rucker VC, Melander C, Dervan PB. Influence of beta -Alanine on Hairpin Polyamide Orientation in the DNA Minor Groove. 2003;86.

74. Chenoweth DM, Harki DA, Dervan PB. Solution-Phase Synthesis of Pyrrole - Imidazole Polyamides. Published online 2009:7175-7181.

75. Reeves R, Nissen MS. The A·T-DNA-binding domain of mammalian high mobility group I chromosomal proteins. A novel peptide motif for recognizing DNA structure. *J Biol Chem*. 1990;265(15):8573-8582. doi:10.1016/s0021-9258(19)38926-4

76. Fields GB. Introduction to peptide synthesis. *Curr Protoc protein Sci*. 2002;Chapter 18:Unit 18.1. doi:10.1002/0471140864.ps1801s26

77. Greenberg WA, Baird EE, Dervan PB. A comparison of H-pin and hairpin polyamide motifs for the recognition of the minor groove of DNA. *Chem - A Eur J*. 1998;4(5):796-805. doi:10.1002/(SICI)1521-3765(19980515)4:5<796::AID-CHEM796>3.0.CO;2-G

78. Xiao J, Yuan G, Huang W, Chan ASC, Lee KLD. A convenient method for the synthesis of DNA-recognizing polyamides in solution. *J Org Chem*. 2000;65(18):5506-5513. doi:10.1021/jo000135n

79. Liu B, Wang S, Aston K, et al. β-Alanine and N-terminal cationic substituents affect polyamide–DNA binding. *Org Biomol Chem*. 2017;15(46):9880-9888. doi:10.1039/c7ob02513k

80. Gumpper RH, Li W, Castañeda CH, Scuderi MJ, Bashkin JK, Luo M. A Polyamide Inhibits Replication of Vesicular Stomatitis Virus by Targeting RNA in the Nucleocapsid. *J Virol*. 2018;92(8). doi:10.1128/jvi.00146-18

81. Khan GS, Shah A, Zia-Ur-Rehman, Barker D. Chemistry of DNA minor groove binding agents. *J Photochem Photobiol B Biol*. 2012;115:105-118. doi:10.1016/j.jphotobiol.2012.07.003

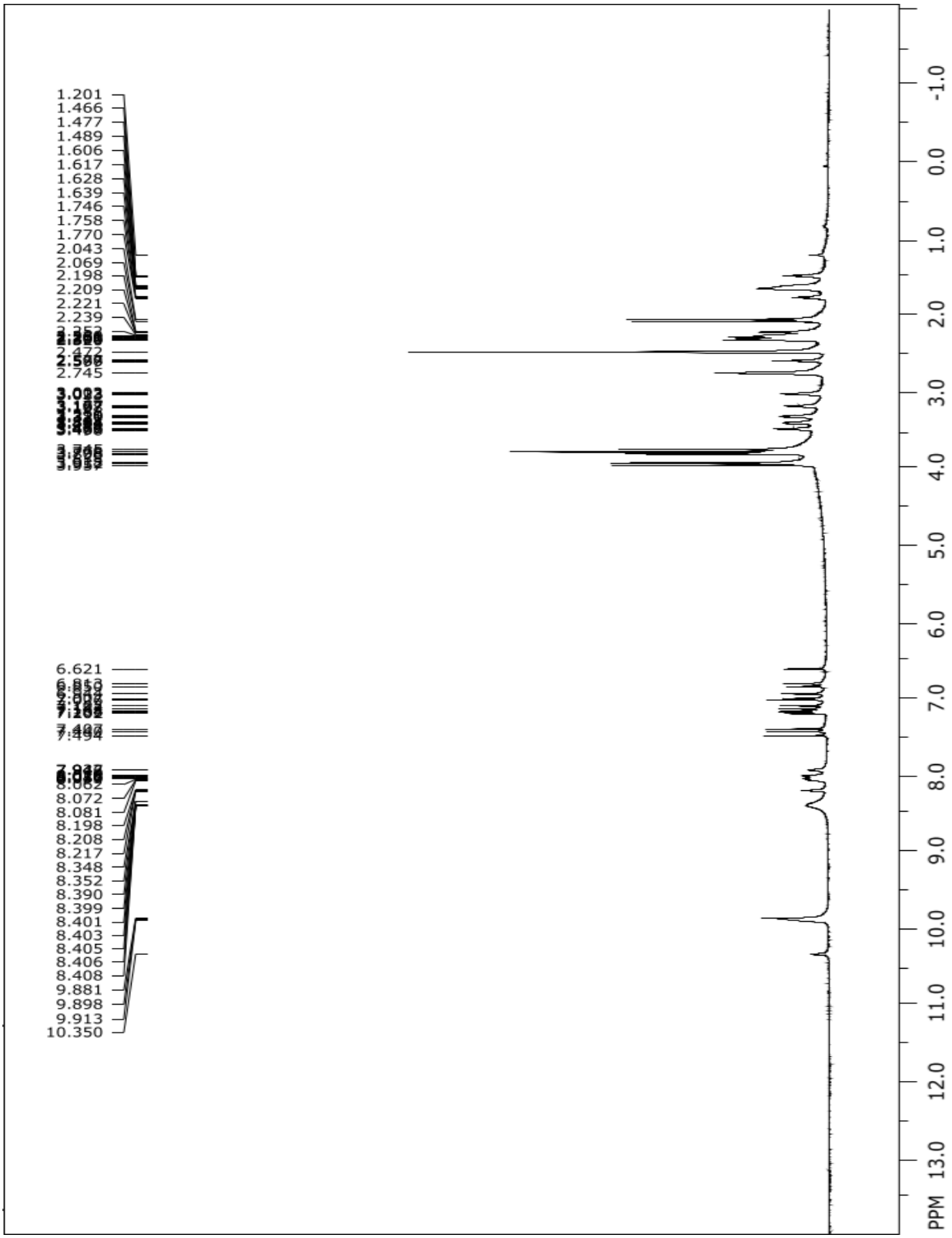
82. Cho J, Parks ME, Dervan PB. Cyclic polyamides for recognition in the minor groove of DNA. *Proc Natl Acad Sci U S A*. 1995;92(22):10389-10392. doi:10.1073/pnas.92.22.10389

83. Tellier M. Structure, Activity, and Function of SETMAR Protein Lysine Methyltransferase. *Life*. 2021;11(12). doi:10.3390/life11121342

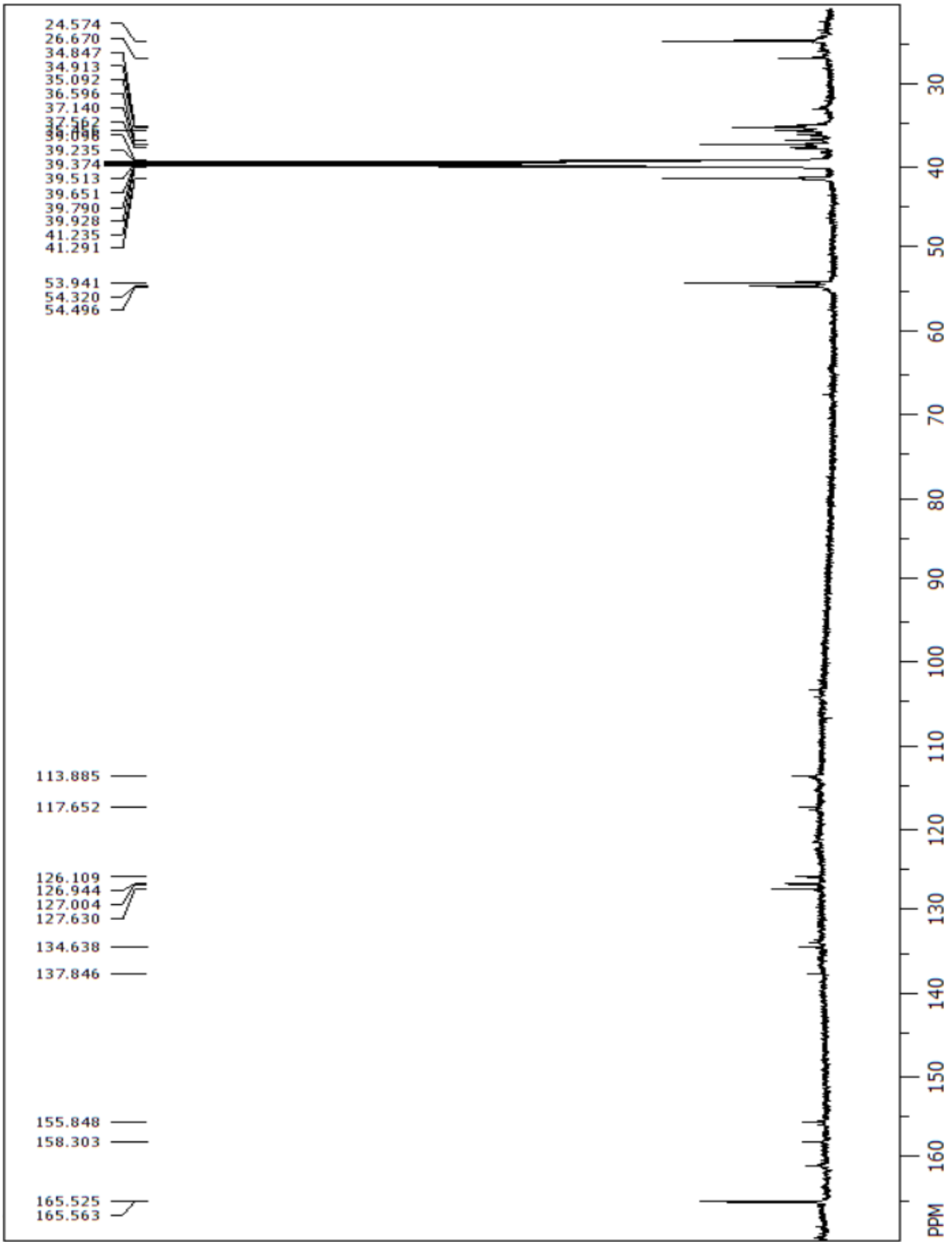
84. Tellier M, Chalmers R. Human SETMAR is a DNA sequence-specific histone-methylase with a broad effect on the transcriptome. *Nucleic Acids Res*. 2019;47(1):122-133. doi:10.1093/nar/gky937

Appendix

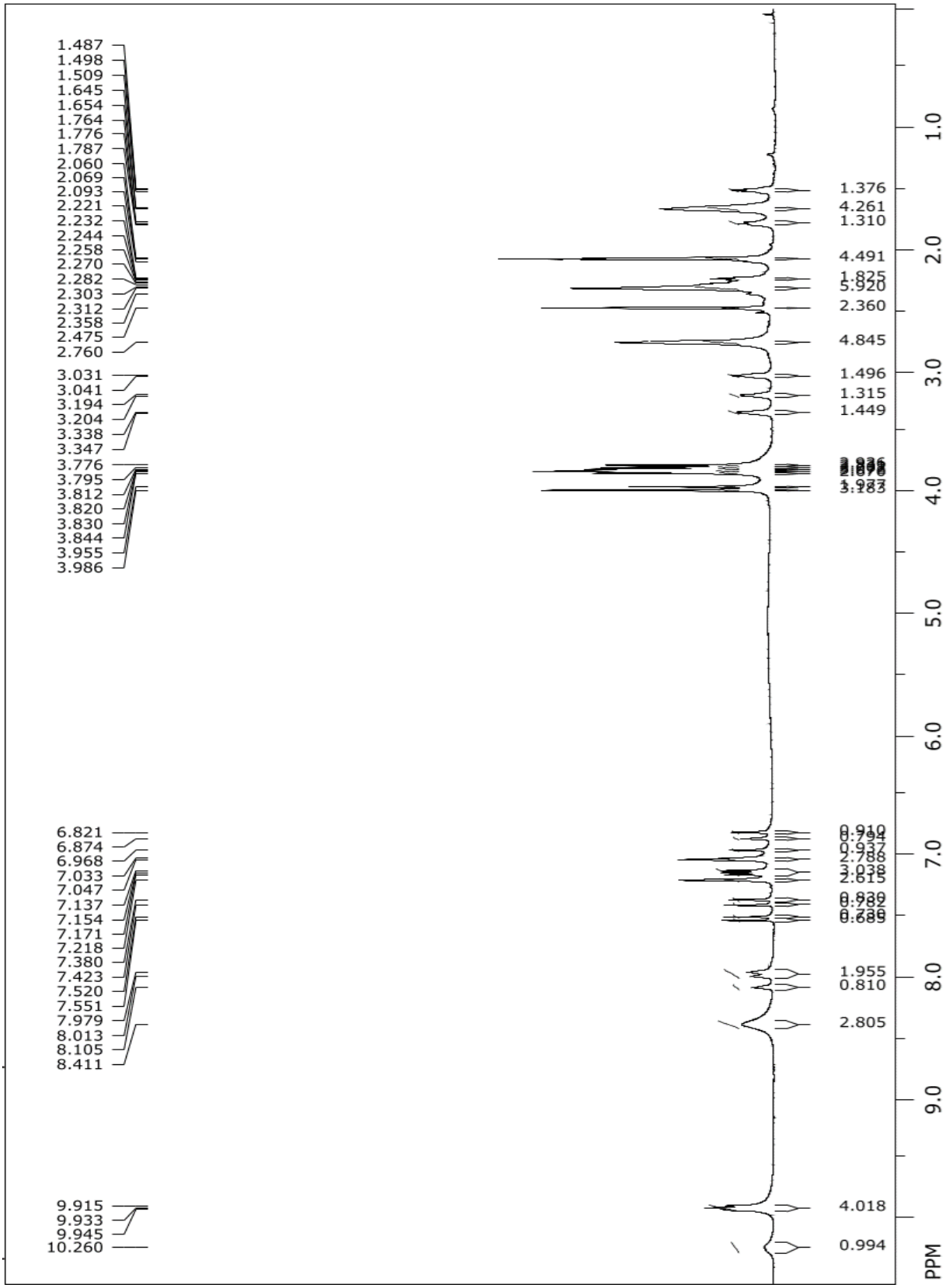
1094 H-NMR


1.201
1.466
1.477
1.489
1.606
1.617
1.628
1.639
1.746
1.758
1.770
2.043
2.069
2.198
2.209
2.221
2.239
2.472
2.745
6.621
7.494
8.062
8.072
8.081
8.198
8.208
8.217
8.348
8.352
8.390
8.399
8.401
8.403
8.405
8.406
8.408
9.881
9.898
9.913
10.350
PPM
13.0
12.0
11.0
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
-1.0

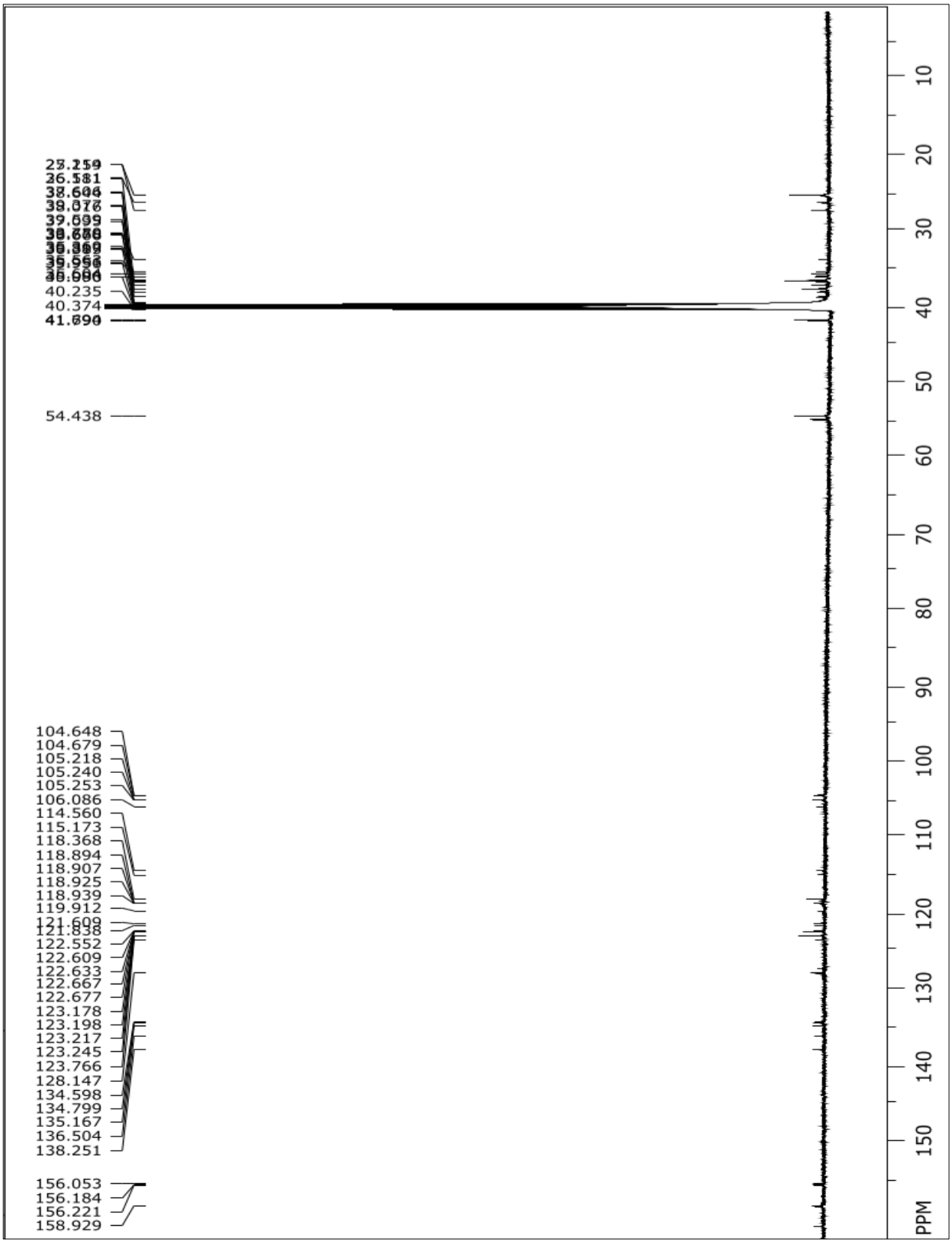
1094 C-NMR:


24.574
26.670
34.847
34.913
35.092
36.596
37.140
37.562
39.235
39.374
39.513
39.651
39.790
39.928
41.235
41.291
53.941
54.320
54.496
113.885
117.652
126.109
126.944
127.004
127.630
134.638
137.846
155.848
158.303
165.525
165.563
30
40
50
60
70
80
90
100
110
120
130
140
150
160
PPM

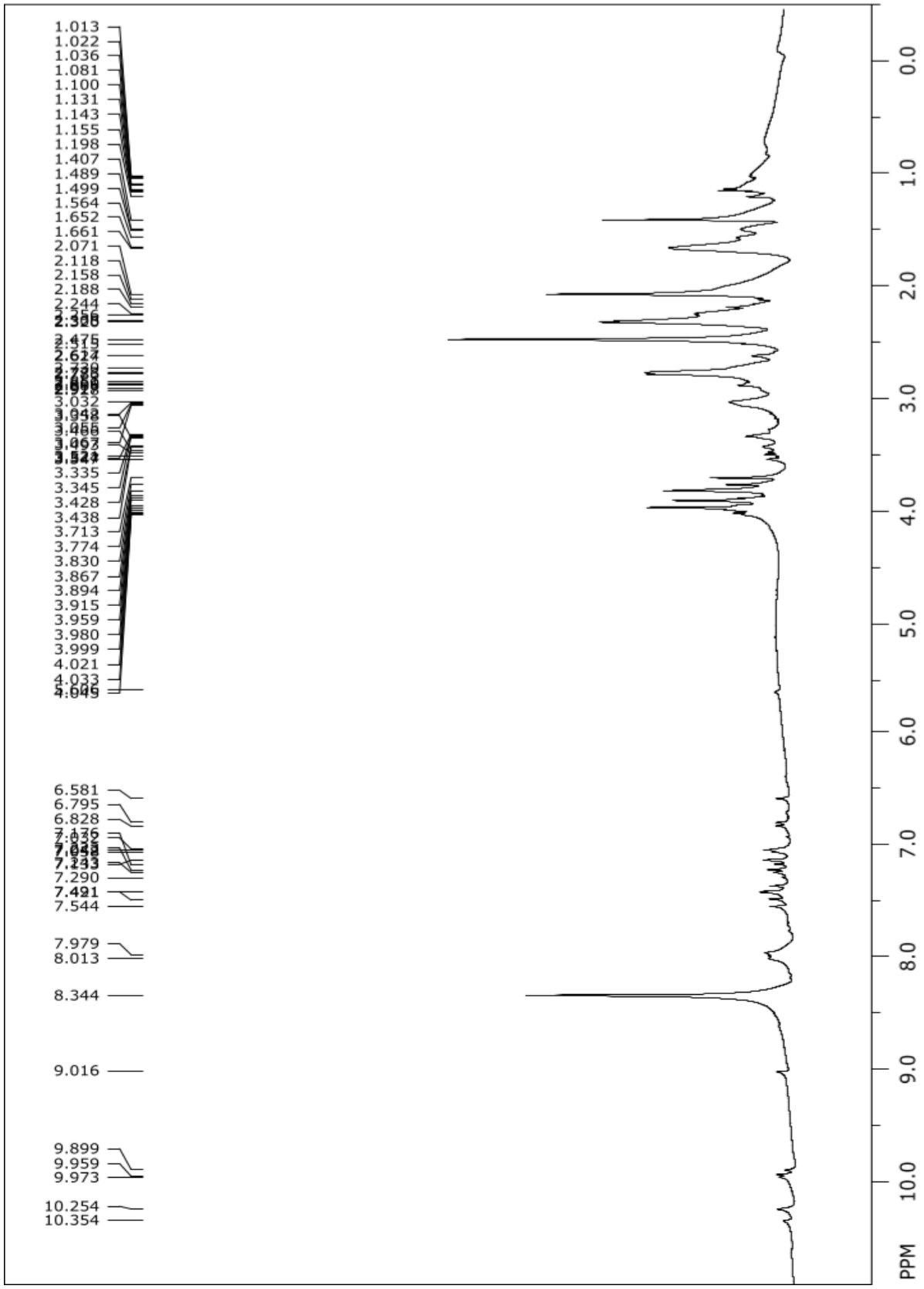
1110 H-NMR:


1.487
1.498
1.509
1.645
1.654
1.764
1.776
1.787
2.060
2.069
2.093
2.221
2.232
2.244
2.258
2.270
2.282
2.303
2.312
2.358
2.475
2.760
3.031
3.041
3.194
3.204
3.338
3.347
3.776
3.795
3.812
3.820
3.830
3.844
3.955
3.986
6.821
6.874
6.968
7.033
7.047
7.137
7.154
7.171
7.218
7.380
7.423
7.520
7.551
7.979
8.013
8.105
8.411
9.915
9.933
9.945
10.260
1.376
4.261
1.310
4.491
1.825
5.920
2.360
4.845
1.496
1.315
1.449
3.183
0.910
0.937
2.788
3.038
2.615
1.955
0.810
2.805
4.018
0.994
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
PPM

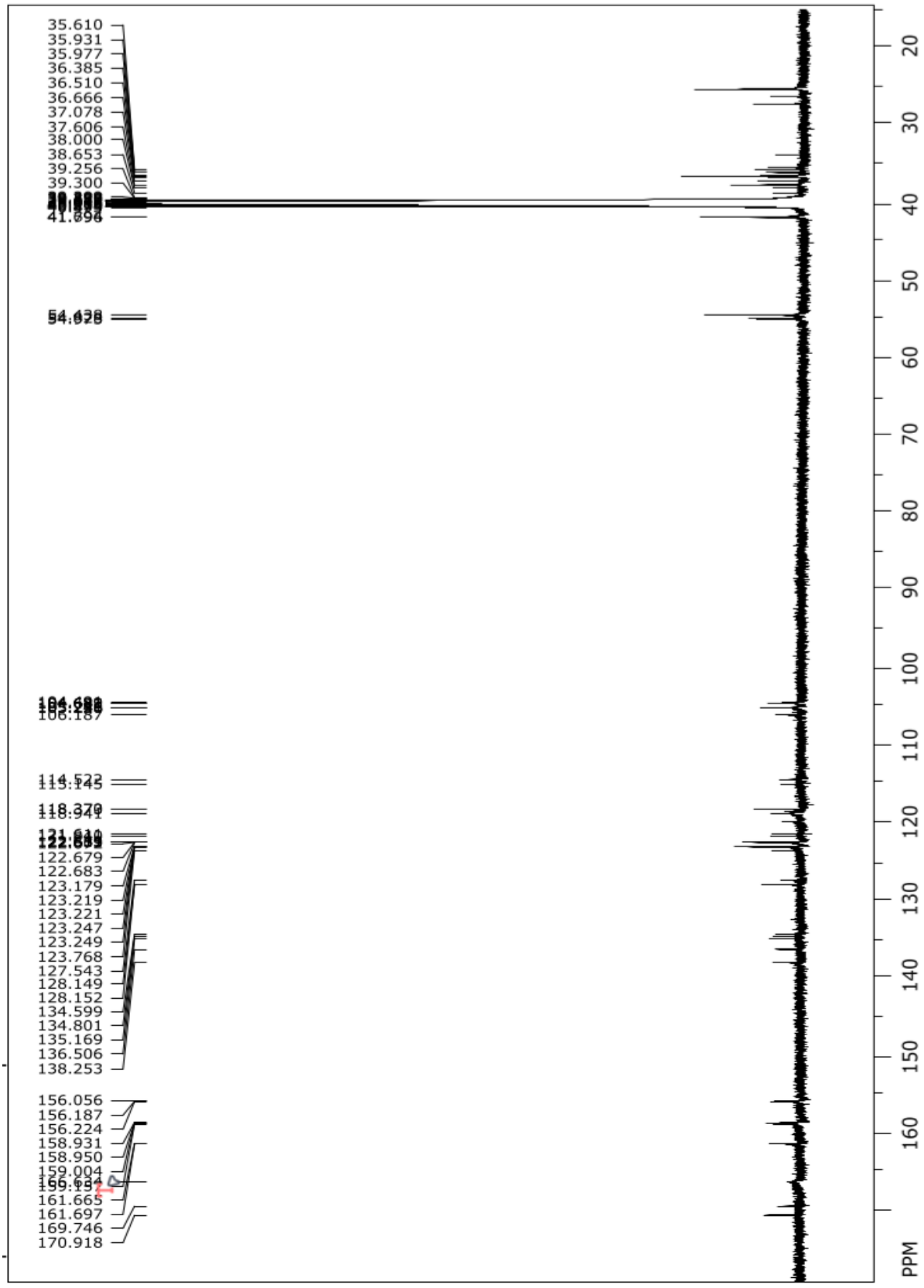
1110 C-NMR:


40.235
40.374
54.438
104.648
104.679
105.218
105.240
105.253
106.086
114.560
115.173
118.368
118.894
118.907
118.925
118.939
119.912
121.609
121.838
122.552
122.609
122.633
122.667
122.677
123.178
123.198
123.217
123.245
123.766
128.147
134.598
134.799
135.167
136.504
138.251
156.053
156.184
156.221
158.929
PPM

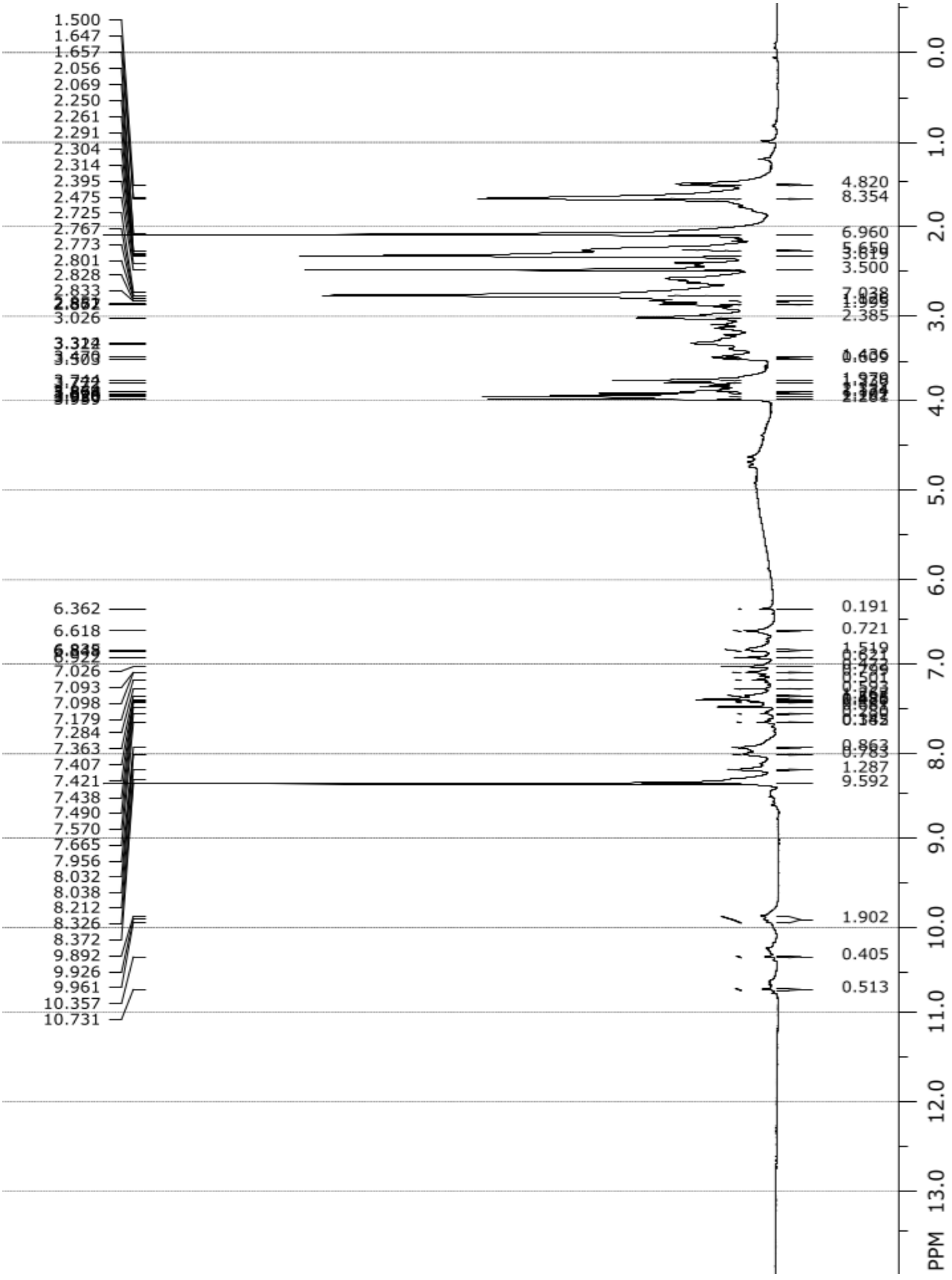
1123 H-NMR:


1.013
1.022
1.036
1.081
1.100
1.131
1.143
1.155
1.198
1.407
1.489
1.499
1.564
1.652
1.661
2.071
2.118
2.158
2.188
2.244
3.032
3.335
3.345
3.428
3.438
3.713
3.774
3.830
3.867
3.894
3.915
3.959
3.980
3.999
4.021
4.033
6.581
6.795
6.828
7.290
7.544
7.979
8.013
8.344
9.016
9.899
9.959
9.973
10.254
10.354
0.0
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
10.0
PPM

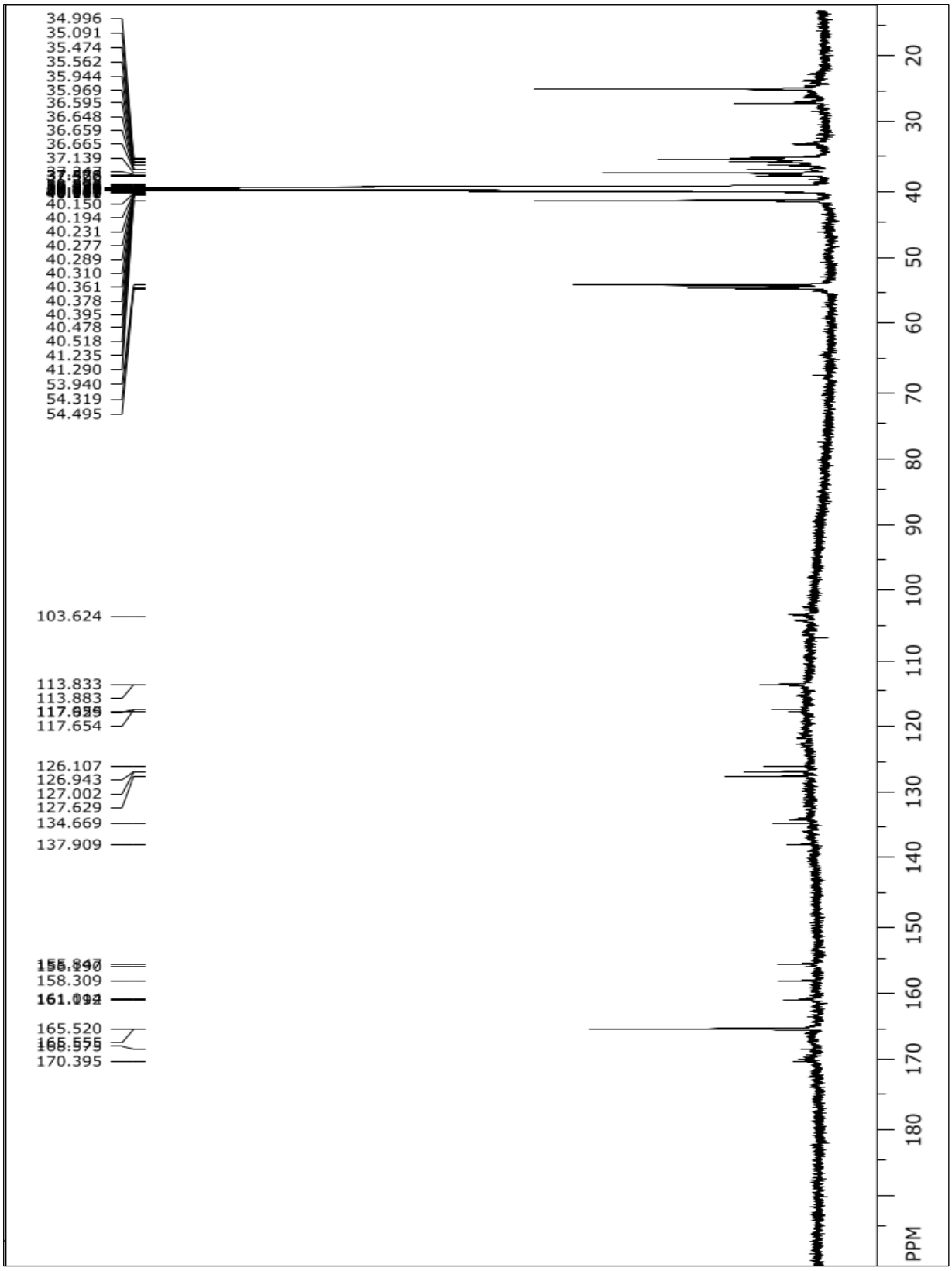
1123 C-NMR:


35.610
35.931
35.977
36.385
36.510
36.666
37.078
37.606
38.000
38.653
39.256
39.300
41.098
54.438
54.628
106.187
114.522
115.145
118.370
118.941
122.679
122.683
123.179
123.219
123.221
123.247
123.249
123.768
127.543
128.149
128.152
134.599
134.801
135.169
136.506
138.253
156.056
156.187
156.224
158.931
158.950
159.004
161.665
161.697
169.746
170.918
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
PPM

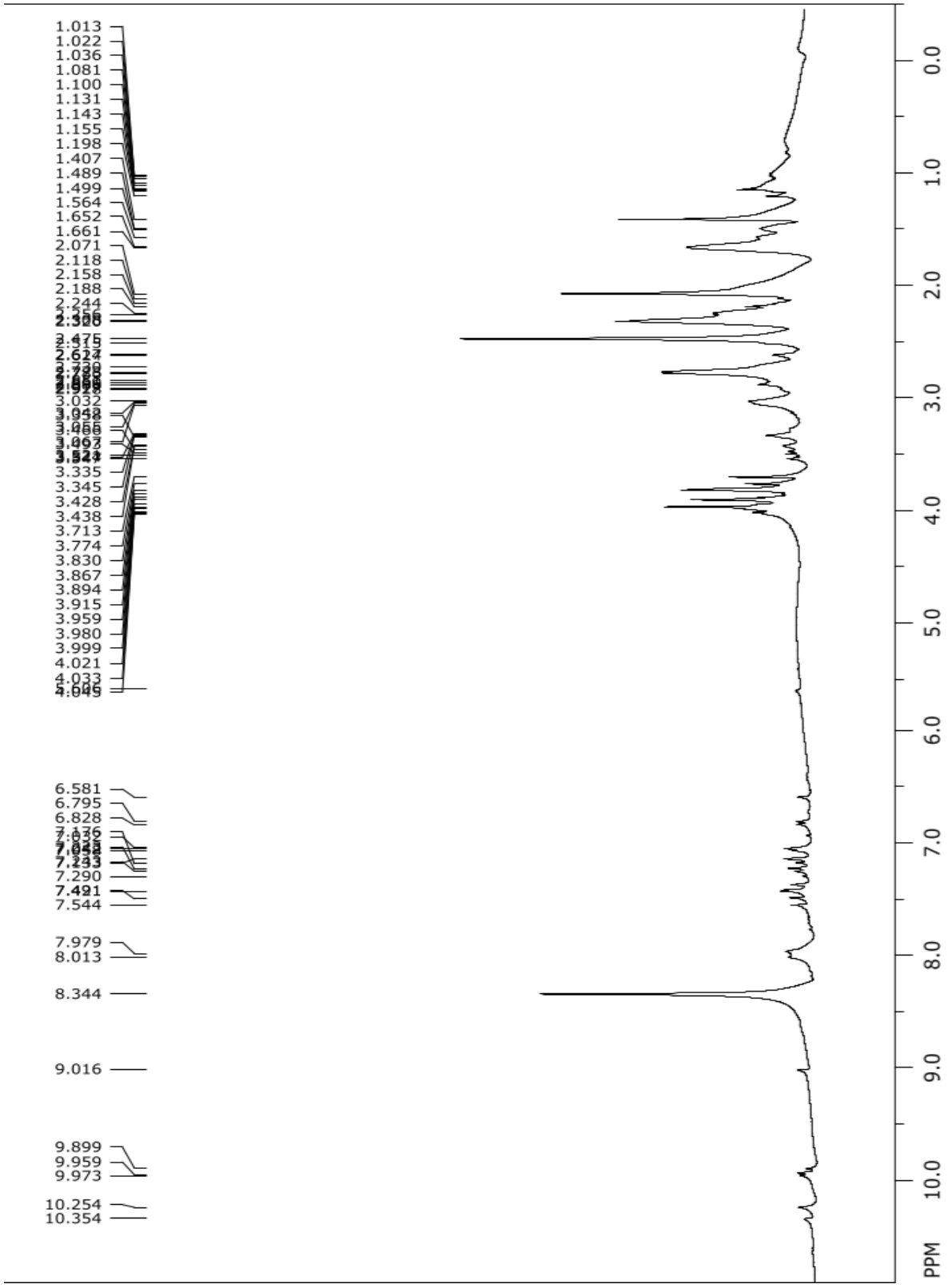
1131H-NMR:


1.500
1.647
1.657
2.056
2.069
2.250
2.261
2.291
2.304
2.314
2.395
2.475
2.725
2.767
2.773
2.801
2.828
2.833
3.026
6.362
6.618
7.026
7.093
7.098
7.179
7.284
7.363
7.407
7.421
7.438
7.490
7.570
7.665
7.956
8.032
8.038
8.212
8.326
8.372
9.892
9.926
9.961
10.357
10.731
4.820
8.354
6.960
3.500
2.385
0.191
0.721
1.519
0.863
1.287
9.592
1.902
0.405
0.513
PPM

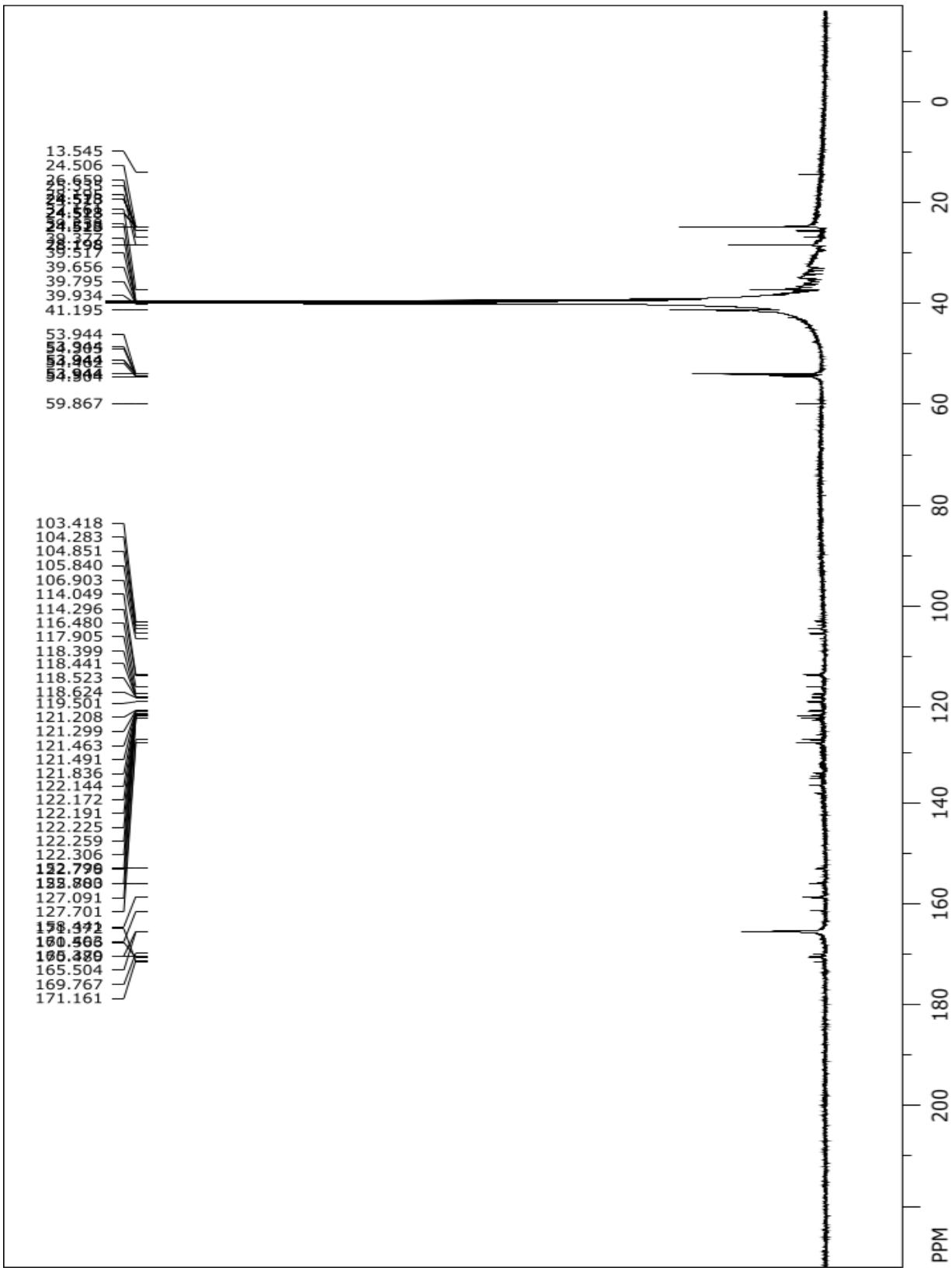
1131C-NMR:


34.996
35.091
35.474
35.562
35.944
35.969
36.595
36.648
36.659
36.665
37.139
40.150
40.194
40.231
40.277
40.289
40.310
40.361
40.378
40.395
40.478
40.518
41.235
41.290
53.940
54.319
54.495
103.624
113.833
113.883
117.654
126.107
126.943
127.002
127.629
134.669
137.909
158.309
165.520
170.395
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
PPM

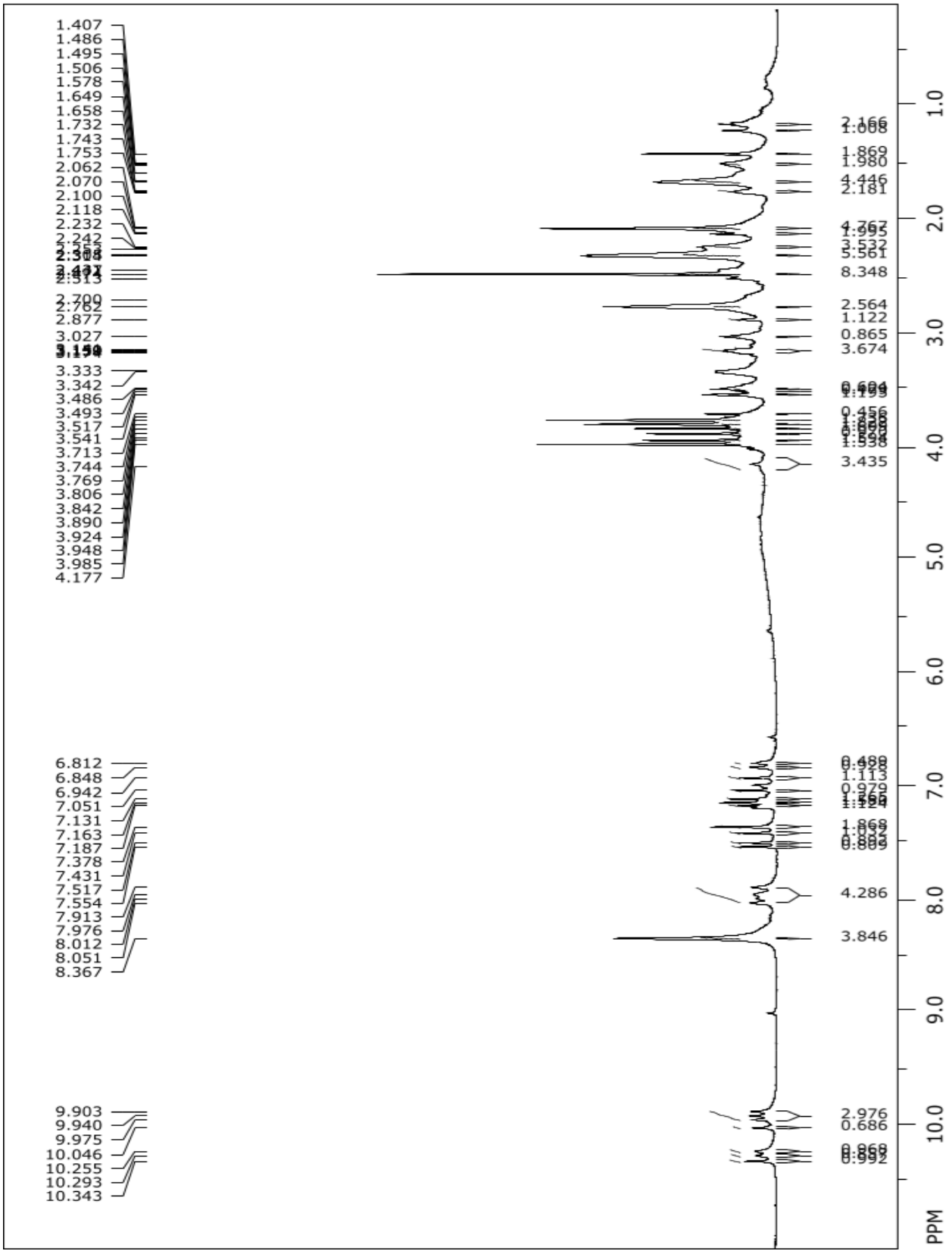
1132 H-NMR:


1.013
1.022
1.036
1.081
1.100
1.131
1.143
1.155
1.198
1.407
1.489
1.499
1.564
1.652
1.661
2.071
2.118
2.158
2.188
2.244
2.300
3.032
3.335
3.345
3.428
3.438
3.713
3.774
3.830
3.867
3.894
3.915
3.959
3.980
3.999
4.021
4.033
6.581
6.795
6.828
7.290
7.491
7.544
7.979
8.013
8.344
9.016
9.899
9.959
9.973
10.254
10.354
0.0
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
10.0
PPM

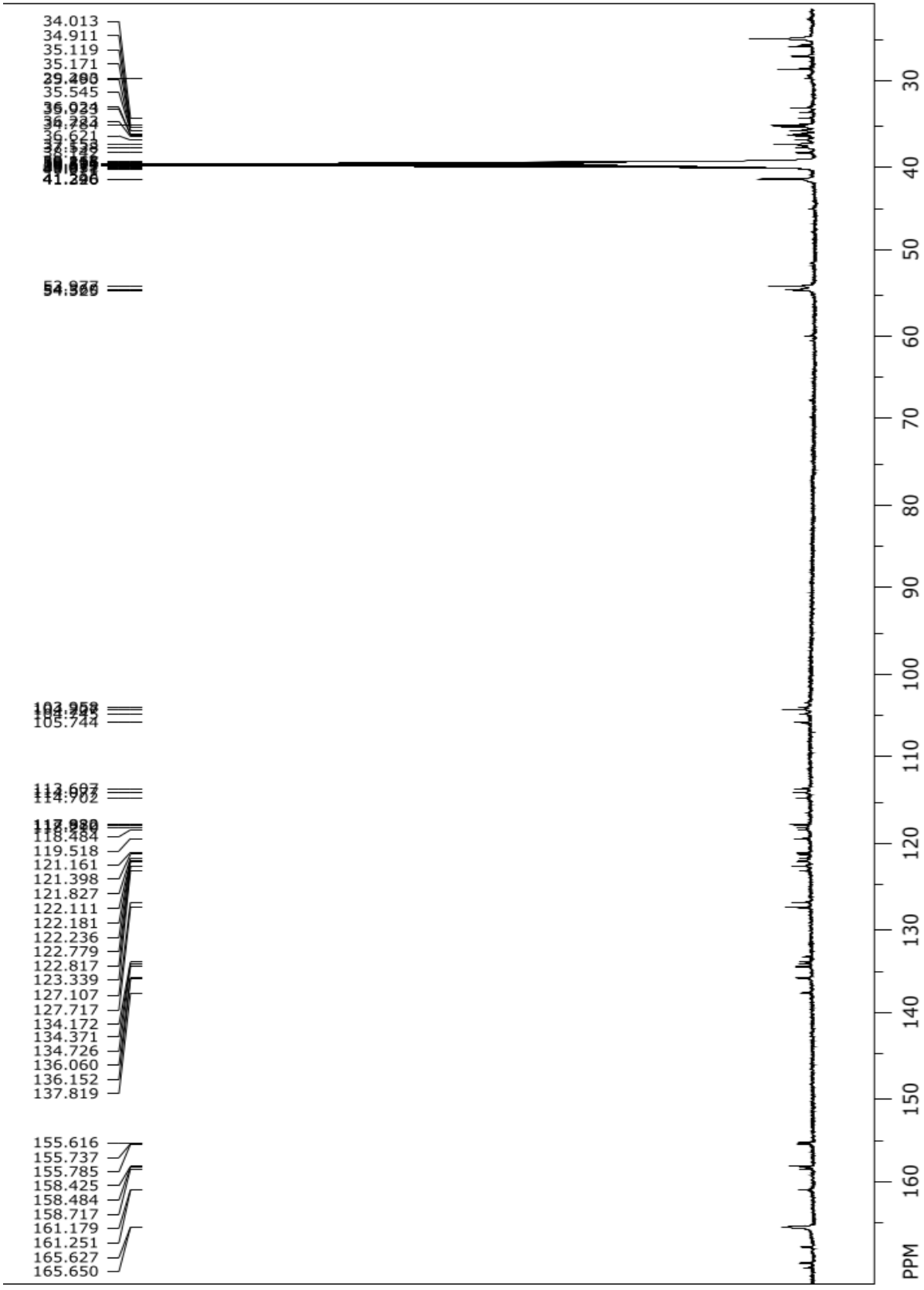
1132C-NMR:


13.545
24.506
39.517
39.656
39.795
39.934
41.195
53.944
59.867
103.418
104.283
104.851
105.840
106.903
114.049
114.296
116.480
117.905
118.399
118.441
118.523
118.624
119.501
121.208
121.299
121.463
121.491
121.836
122.144
122.172
122.191
122.225
122.259
122.306
127.091
127.701
165.504
169.767
171.161
0
20
40
60
80
100
120
140
160
180
200
PPM

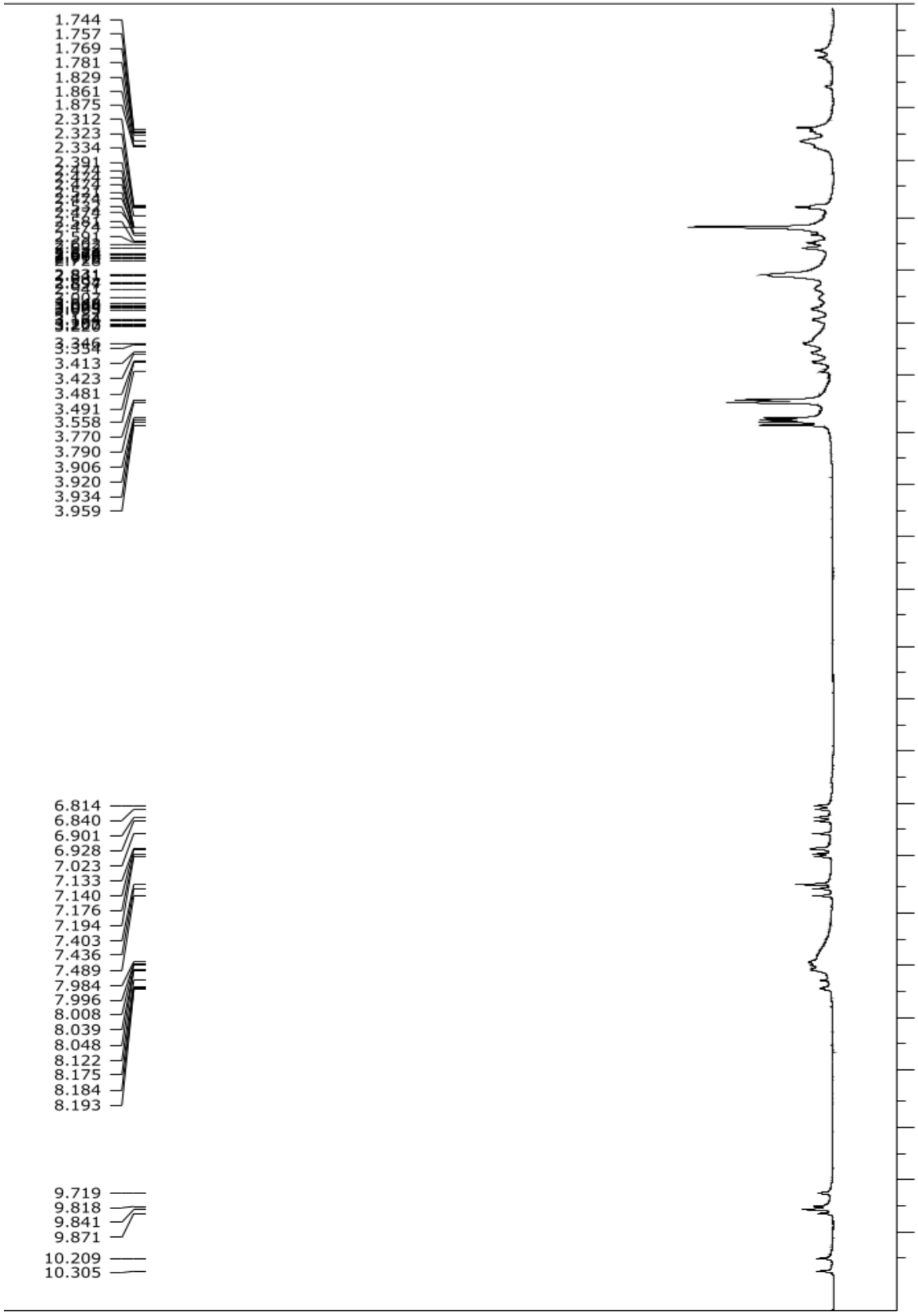
1134 H-NMR:


1.407
1.486
1.495
1.506
1.578
1.649
1.658
1.732
1.743
1.753
2.062
2.070
2.100
2.118
2.232
2.242
2.700
2.762
2.877
3.027
3.333
3.342
3.486
3.493
3.517
3.541
3.713
3.744
3.769
3.806
3.842
3.890
3.924
3.948
3.985
4.177
6.812
6.848
6.942
7.051
7.131
7.163
7.187
7.378
7.431
7.517
7.554
7.913
7.976
8.012
8.051
8.367
9.903
9.940
9.975
10.046
10.255
10.293
10.343
2.166
1.869
1.980
4.446
2.181
4.767
3.532
5.561
8.348
2.564
1.122
0.865
3.674
0.456
3.435
1.113
0.979
4.286
3.846
2.976
0.686
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
10.0
PPM

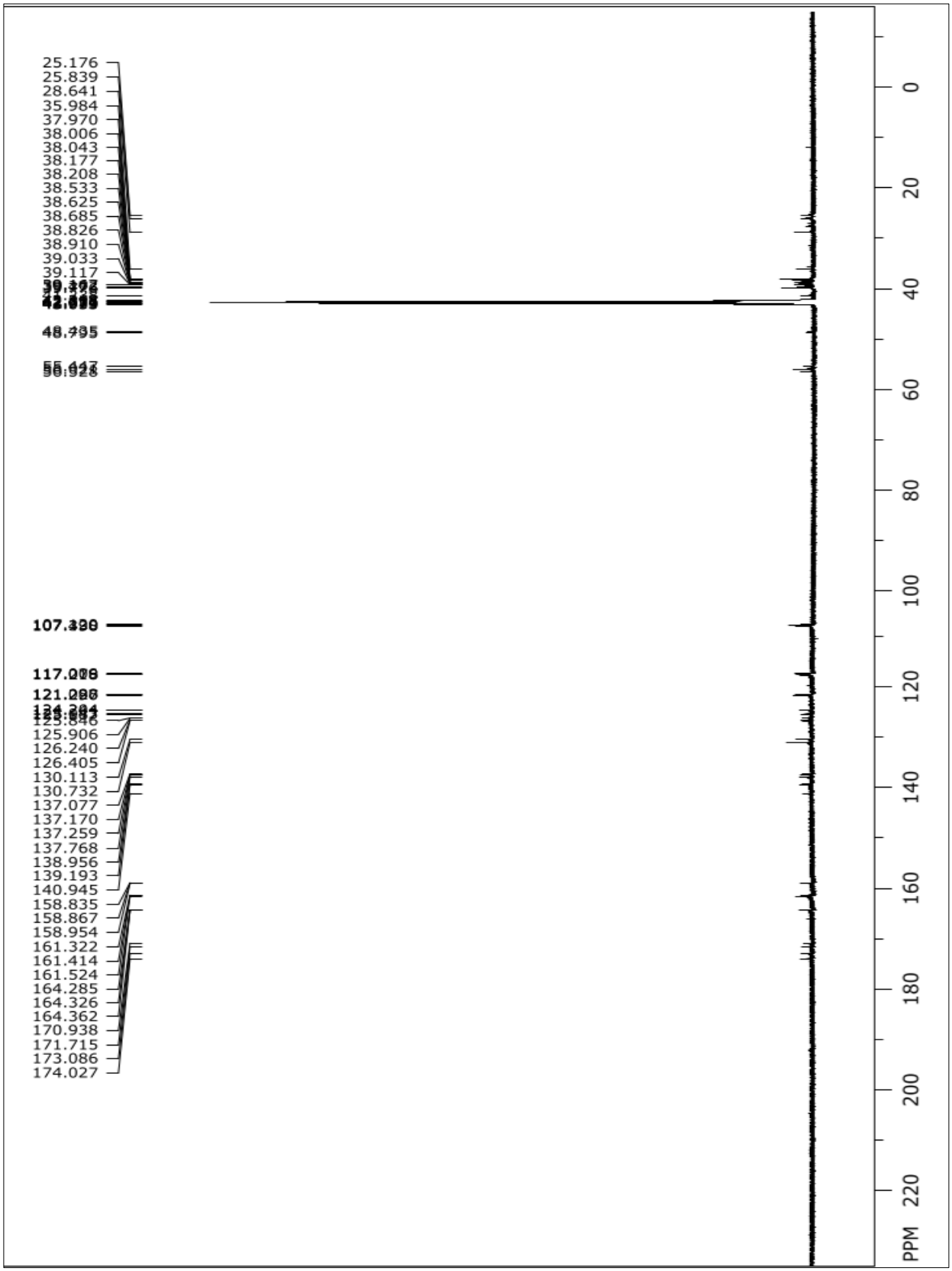
1134 C-NMR:


34.013
34.911
35.119
35.171
35.545
41.296
105.744
113.607
114.702
118.484
119.518
121.161
121.398
121.827
122.111
122.181
122.236
122.779
122.817
123.339
127.107
127.717
134.172
134.371
134.726
136.060
136.152
137.819
155.616
155.737
155.785
158.425
158.484
158.717
161.179
161.251
165.627
165.650
30
40
50
60
70
80
90
100
110
120
130
140
150
160
PPM

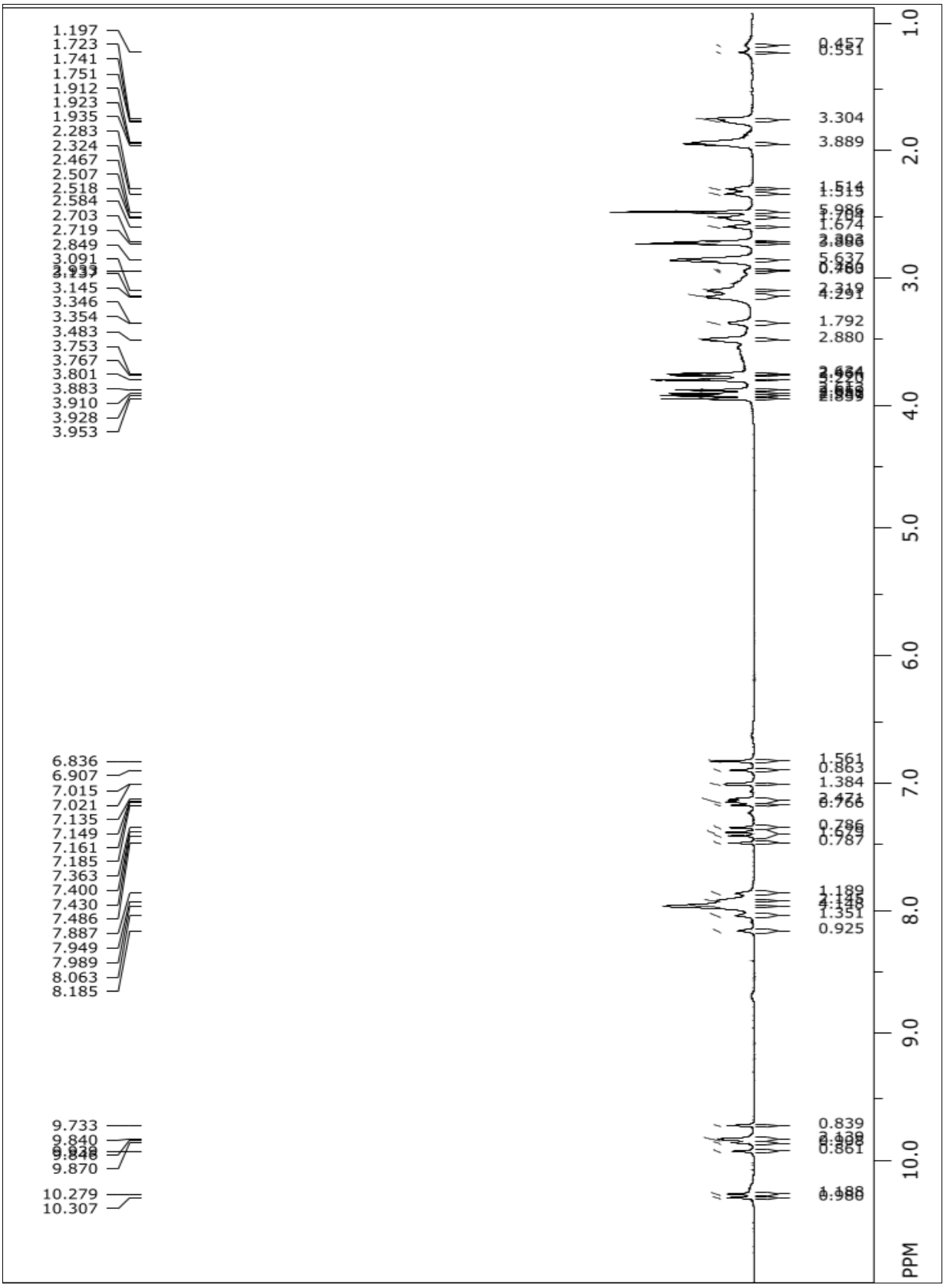
1161 H-NMR:


1.744
1.757
1.769
1.781
1.829
1.861
1.875
2.312
2.323
2.334
2.391
3.346
3.354
3.413
3.423
3.481
3.491
3.558
3.770
3.790
3.906
3.920
3.934
3.959
6.814
6.840
6.901
6.928
7.023
7.133
7.140
7.176
7.194
7.403
7.436
7.489
7.984
7.996
8.008
8.039
8.048
8.122
8.175
8.184
8.193
9.719
9.818
9.841
9.871
10.209
10.305
PPM 10.0 9.6 9.2 8.8 8.4 8.0 7.6 7.2 6.8 6.4 6.0 5.6 5.2 4.8 4.4 4.0 3.6 3.2 2.8 2.4 2.0 1.6 1.2

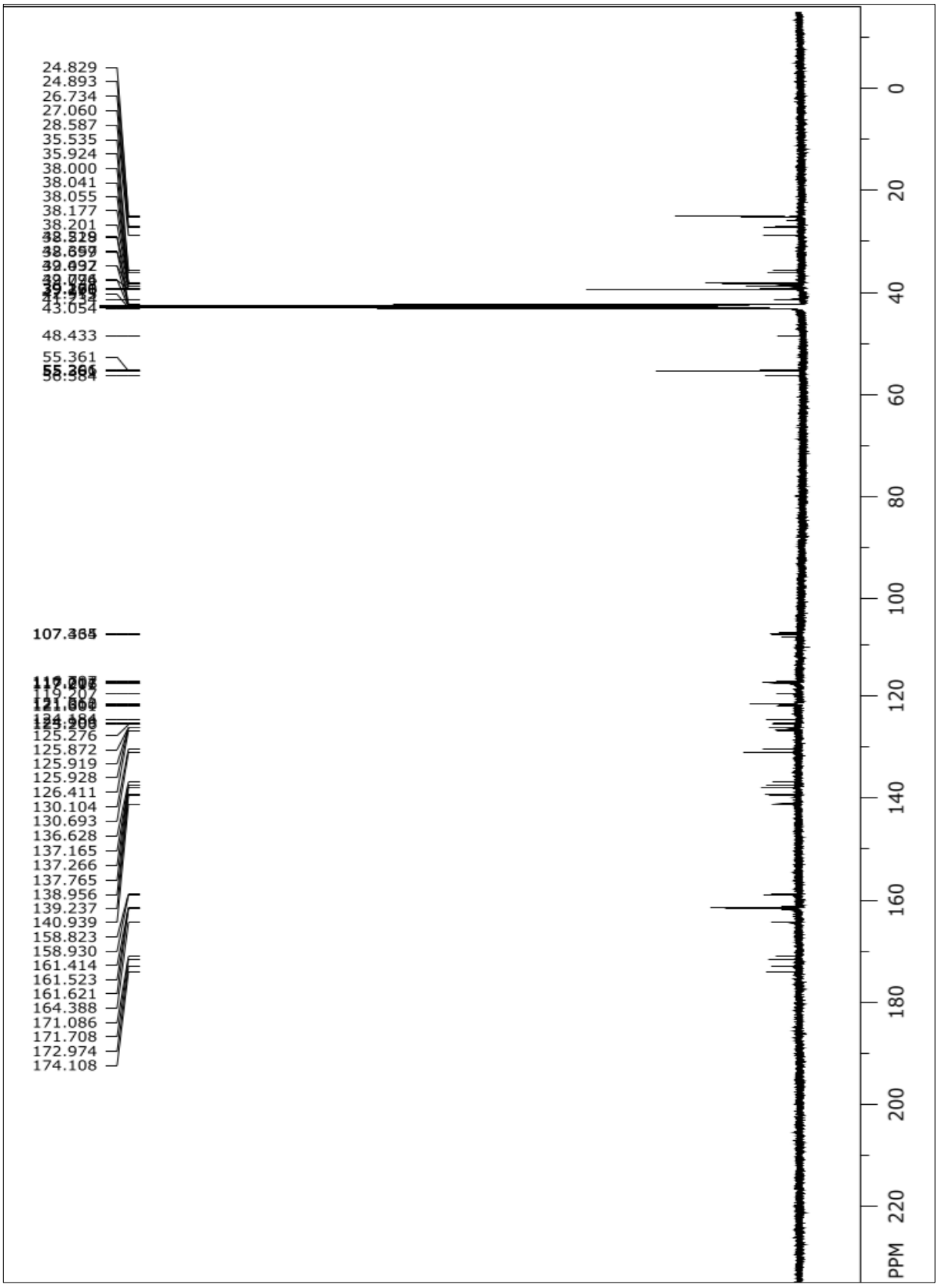
1161 C-NMR:


25.176
25.839
28.641
35.984
37.970
38.006
38.043
38.177
38.208
38.533
38.625
38.685
38.826
38.910
39.033
39.117
48.795
55.447
107.390
117.008
121.088
125.906
126.240
126.405
130.113
130.732
137.077
137.170
137.259
137.768
138.956
139.193
140.945
158.835
158.867
158.954
161.322
161.414
161.524
164.285
164.326
164.362
170.938
171.715
173.086
174.027
0
20
40
60
80
100
120
140
160
180
200
220
PPM

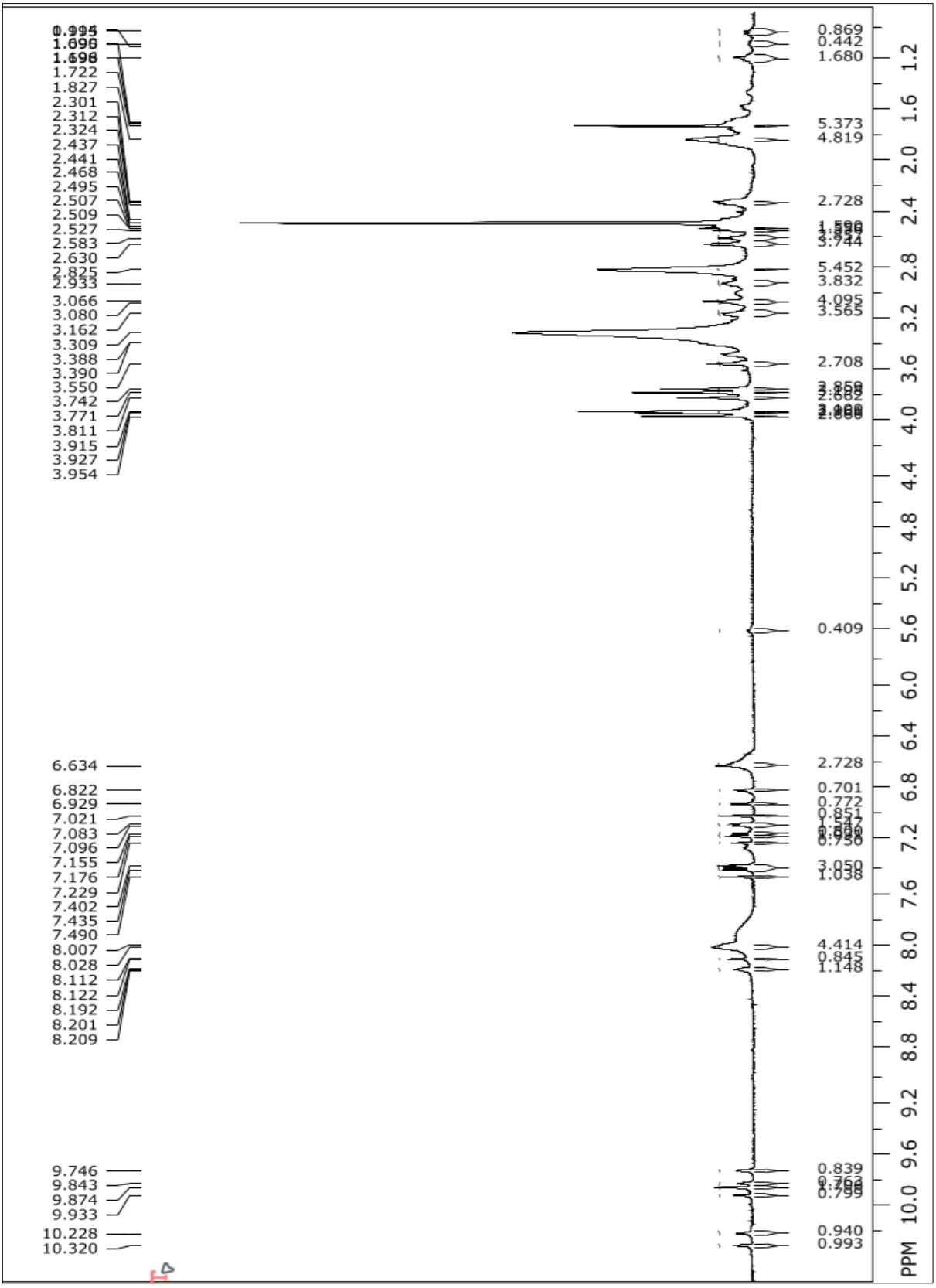
1162 H-NMR:


1.197
1.723
1.741
1.751
1.912
1.923
1.935
2.283
2.324
2.467
2.507
2.518
2.584
2.703
2.719
2.849
3.091
3.145
3.346
3.354
3.483
3.753
3.767
3.801
3.883
3.910
3.928
3.953
6.836
6.907
7.015
7.021
7.135
7.149
7.161
7.185
7.363
7.400
7.430
7.486
7.887
7.949
7.989
8.063
8.185
9.733
9.840
9.870
10.279
10.307
0.457
0.551
3.304
3.889
1.514
1.515
5.986
1.674
5.637
4.291
1.792
2.880
1.561
0.863
1.384
0.786
0.787
1.189
1.351
0.925
0.839
0.861
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
10.0
PPM

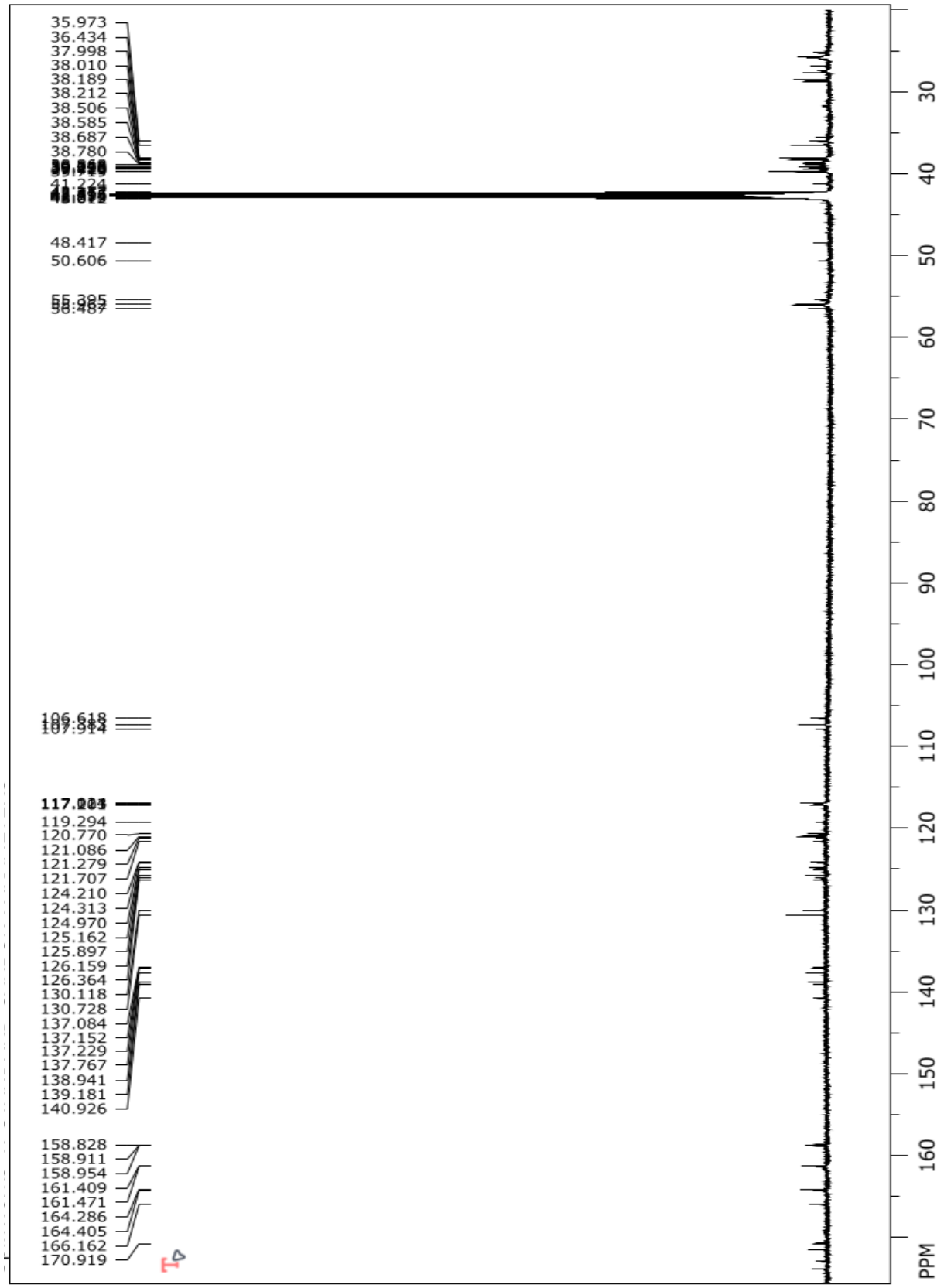
1162 C-NMR:


24.829
24.893
26.734
27.060
28.587
35.535
35.924
38.000
38.041
38.055
38.177
38.201
43.054
48.433
55.361
107.365
125.276
125.872
125.919
125.928
126.411
130.104
130.693
136.628
137.165
137.266
137.765
138.956
139.237
140.939
158.823
158.930
161.414
161.523
161.621
164.388
171.086
171.708
172.974
174.108
0
20
40
60
80
100
120
140
160
180
200
220
PPM

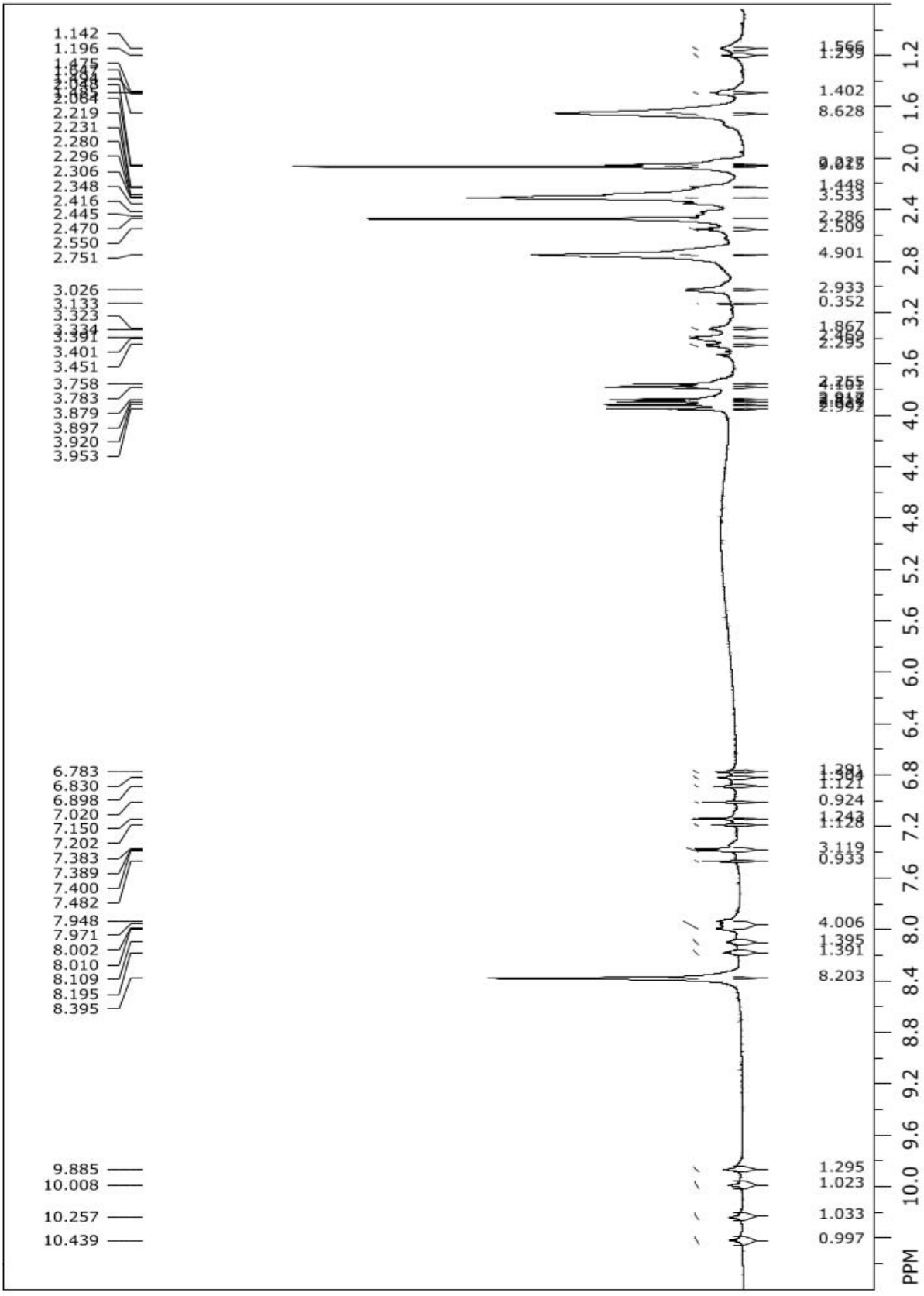
1166H-NMR:


1.722
1.827
2.301
2.312
2.324
2.437
2.441
2.468
2.495
2.507
2.509
2.527
2.583
2.630
2.825
2.933
3.066
3.080
3.162
3.309
3.388
3.390
3.550
3.742
3.771
3.811
3.915
3.927
3.954
6.634
6.822
6.929
7.021
7.083
7.096
7.155
7.176
7.229
7.402
7.435
7.490
8.007
8.028
8.112
8.122
8.192
8.201
8.209
9.746
9.843
9.874
9.933
10.228
10.320
0.869
0.442
1.680
5.373
4.819
2.728
3.744
5.452
3.832
4.095
3.565
2.708
0.409
2.728
0.701
0.772
0.851
0.750
3.050
1.038
4.414
0.845
1.148
0.839
0.799
0.940
0.993
PPM

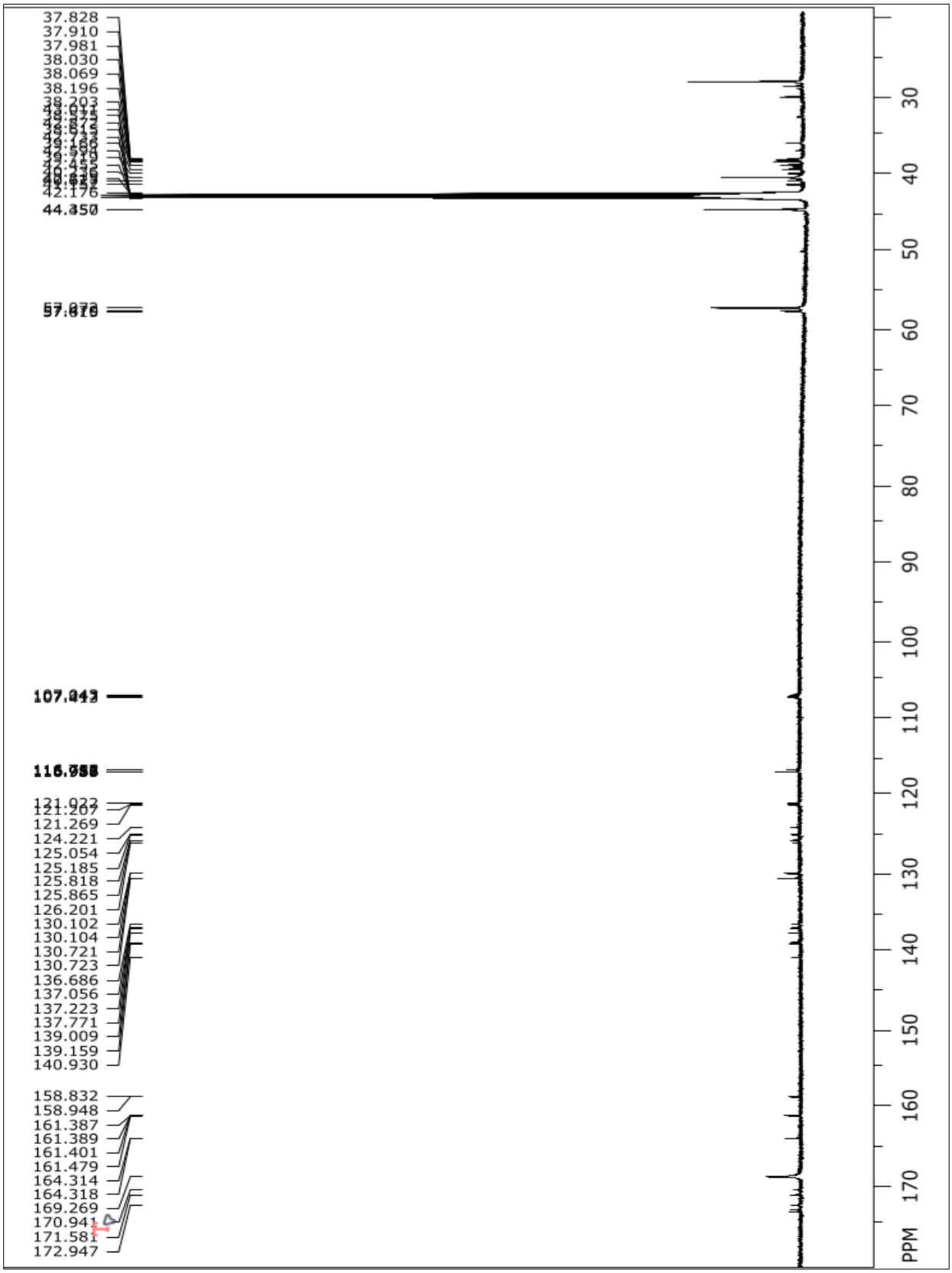
1166C-NMR:


35.973
36.434
37.998
38.010
38.189
38.212
38.506
38.585
38.687
38.780
41.224
48.417
50.606
106.618
107.914
119.294
120.770
121.086
121.279
121.707
124.210
124.313
124.970
125.162
125.897
126.159
126.364
130.118
130.728
137.084
137.152
137.229
137.767
138.941
139.181
140.926
158.828
158.911
158.954
161.409
161.471
164.286
164.405
166.162
170.919
30
40
50
60
70
80
90
100
110
120
130
140
150
160
PPM

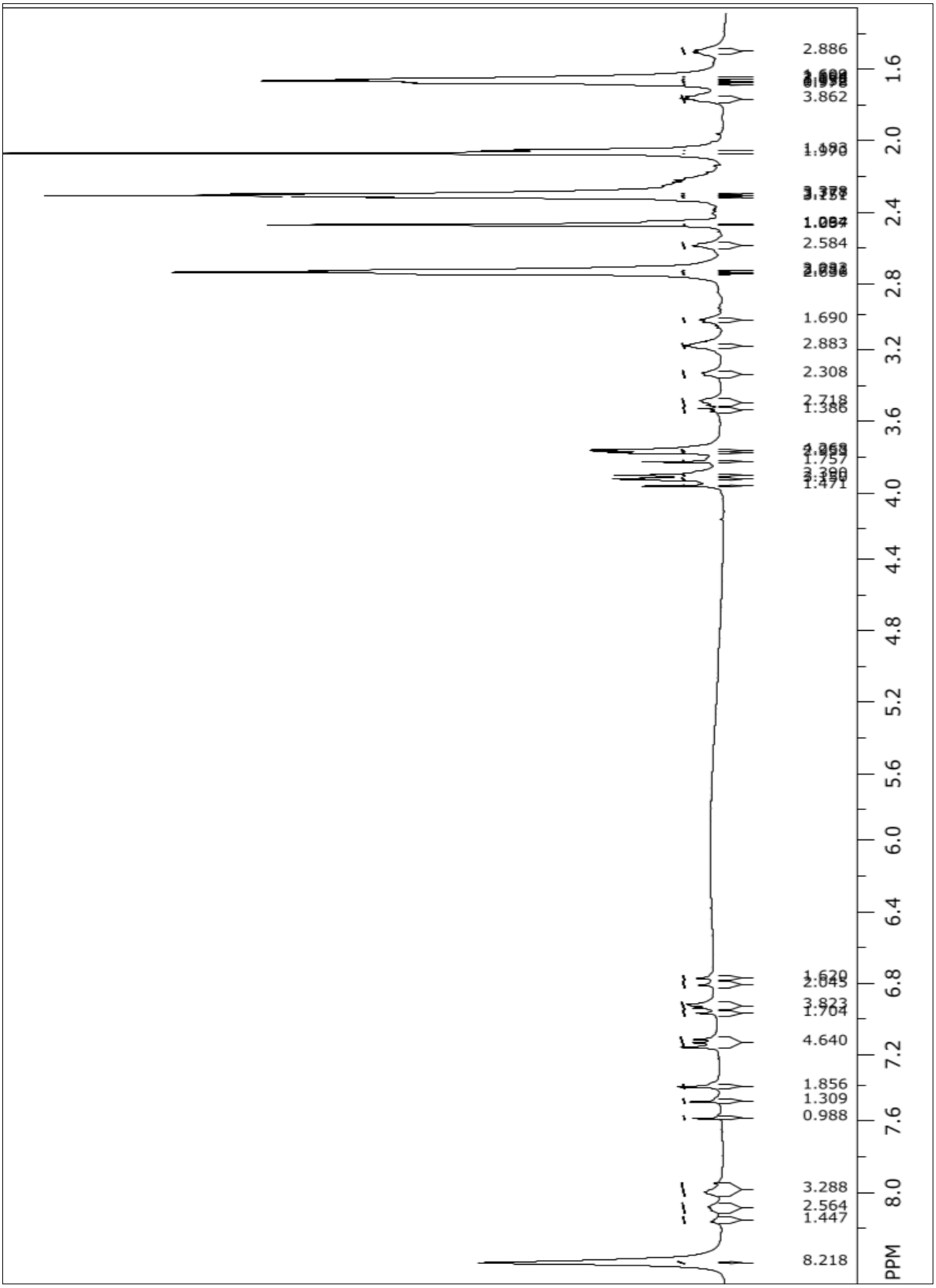
1171 H-NMR:


1.142
1.196
2.219
2.231
2.280
2.296
2.306
2.348
2.416
2.445
2.470
2.550
2.751
3.026
3.133
3.323
3.401
3.451
3.758
3.783
3.879
3.897
3.920
3.953
6.783
6.830
6.898
7.020
7.150
7.202
7.383
7.389
7.400
7.482
7.948
7.971
8.002
8.010
8.109
8.195
8.395
9.885
10.008
10.257
10.439
1.566
1.402
8.628
1.448
3.533
2.286
2.509
4.901
2.933
0.352
1.867
1.121
0.924
3.119
0.933
4.006
1.395
1.391
8.203
1.295
1.023
1.033
0.997
PPM

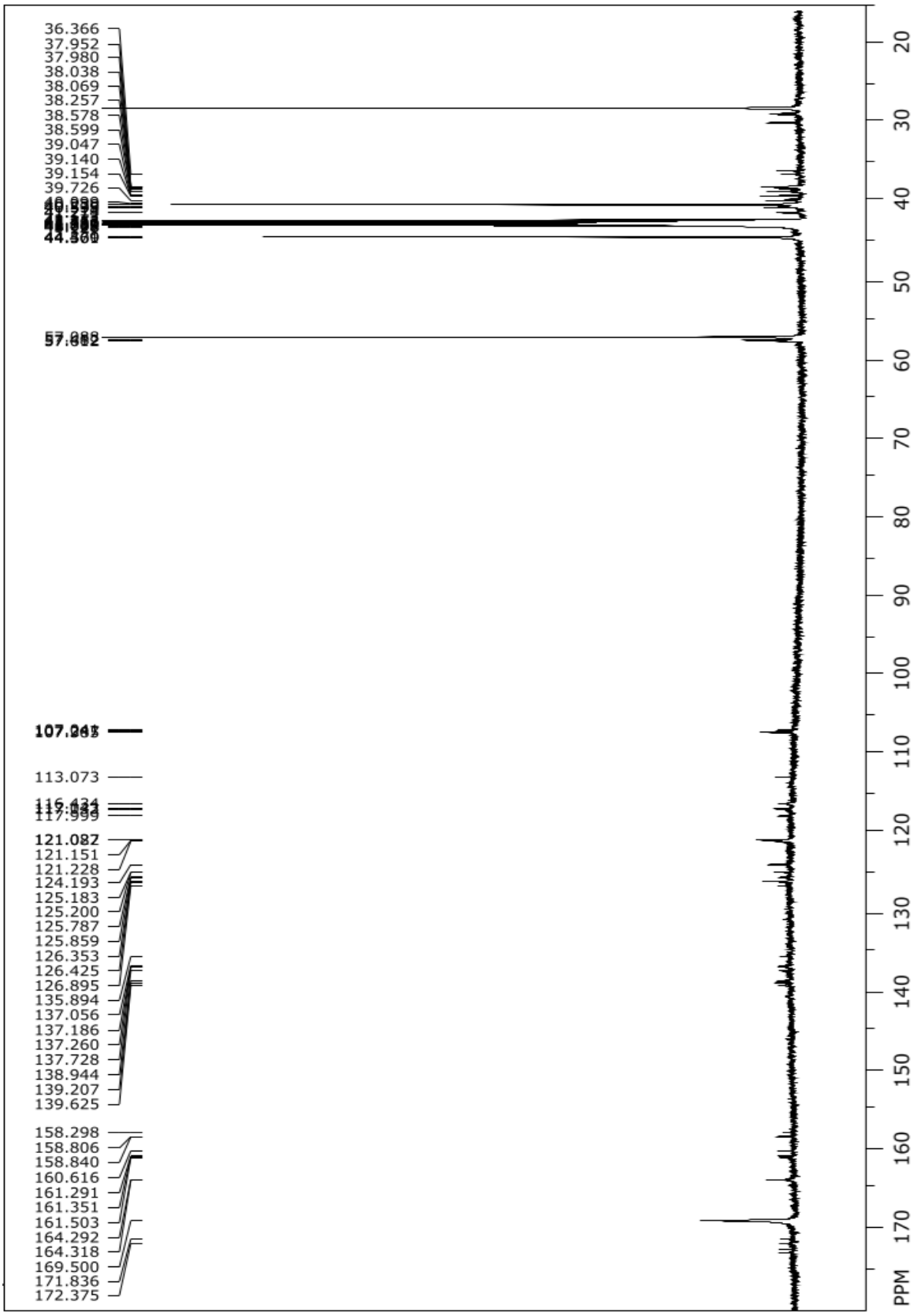
1171C-NMR:


37.828
37.910
37.981
38.030
38.069
38.196
44.350
107.043
116.988
121.022
121.207
121.269
124.221
125.054
125.185
125.818
125.865
126.201
130.102
130.104
130.721
130.723
136.686
137.056
137.223
137.771
139.009
139.159
140.930
158.832
158.948
161.387
161.389
161.401
161.479
164.314
164.318
169.269
170.941
171.581
172.947
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
PPM

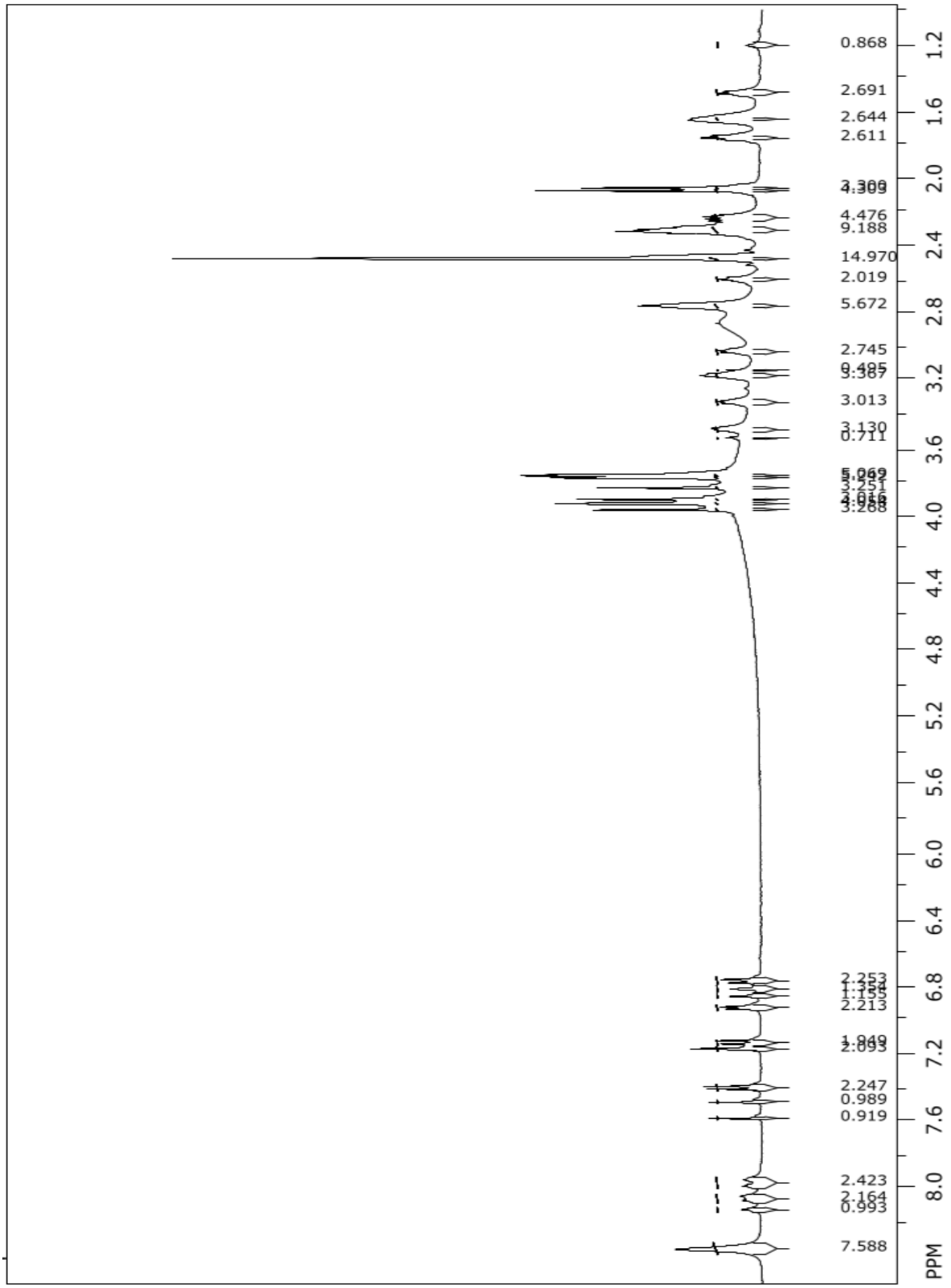
1176's H-NMR:


2.886
3.862
2.584
1.690
2.883
2.308
1.386
1.471
1.620
2.045
3.823
1.704
4.640
1.856
1.309
0.988
3.288
2.564
1.447
8.218
PPM

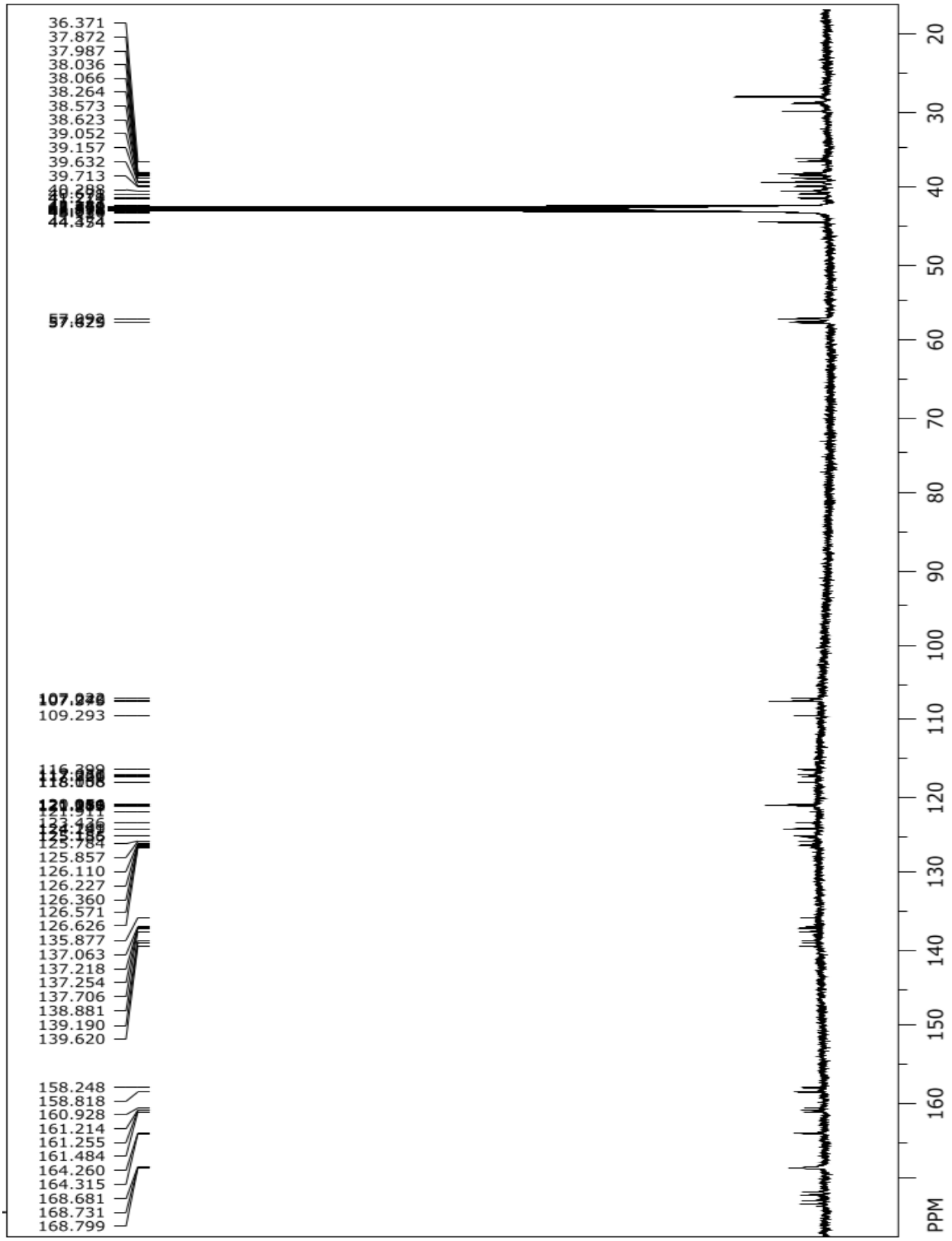
1176's C-NMR:


36.366
37.952
37.980
38.038
38.069
38.257
38.578
38.599
39.047
39.140
39.154
39.726
113.073
121.087
121.151
121.228
124.193
125.183
125.200
125.787
125.859
126.353
126.425
126.895
135.894
137.056
137.186
137.260
137.728
138.944
139.207
139.625
158.298
158.806
158.840
160.616
161.291
161.351
161.503
164.292
164.318
169.500
171.836
172.375
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
PPM

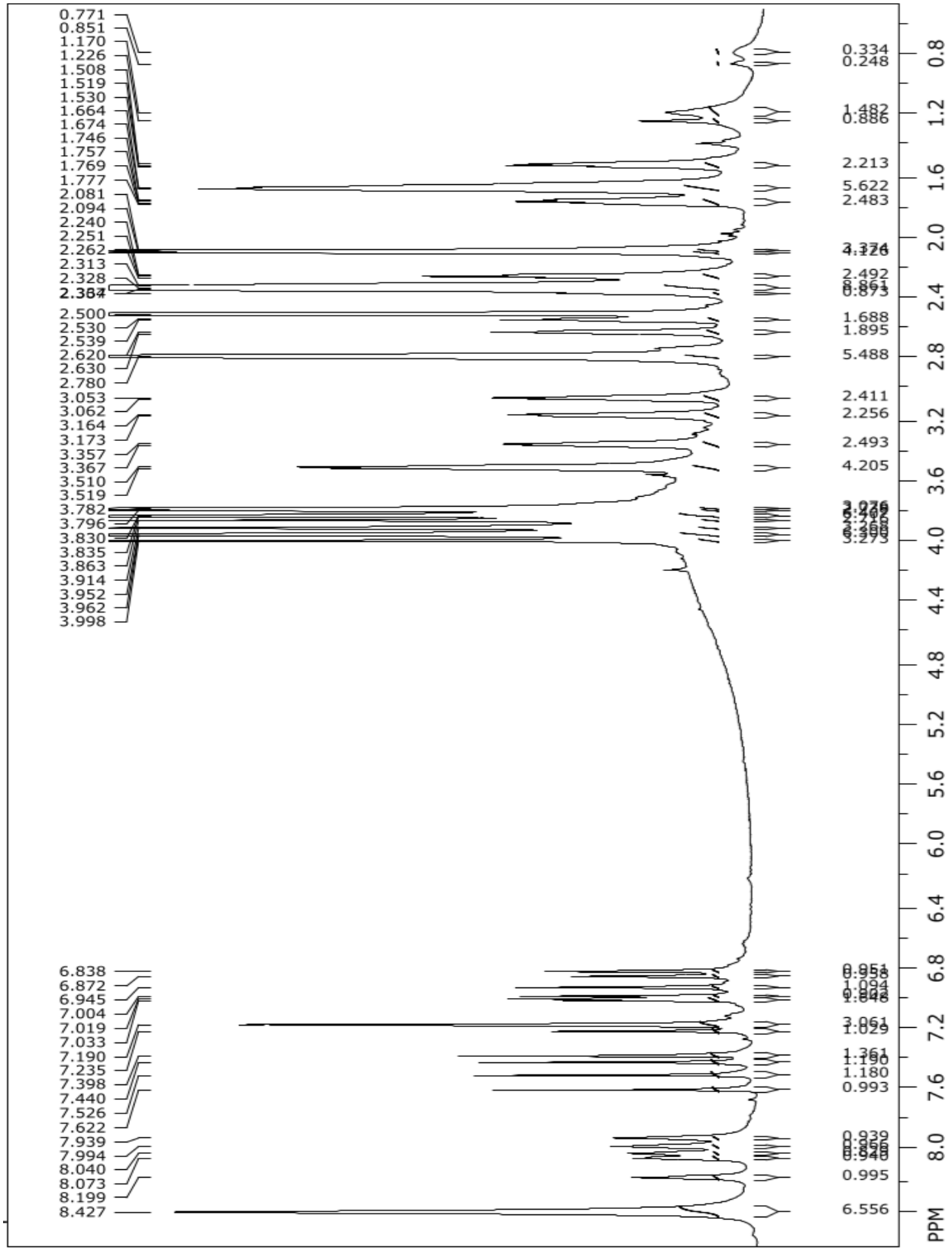
1178's H-NMR:


0.868
2.691
2.644
2.611
4.476
9.188
14.970
2.019
5.672
2.745
3.013
0.711
2.253
2.213
2.247
0.989
0.919
2.423
0.993
7.588
1.2
1.6
2.0
2.4
2.8
3.2
3.6
4.0
4.4
4.8
5.2
5.6
6.0
6.4
6.8
7.2
7.6
8.0
PPM

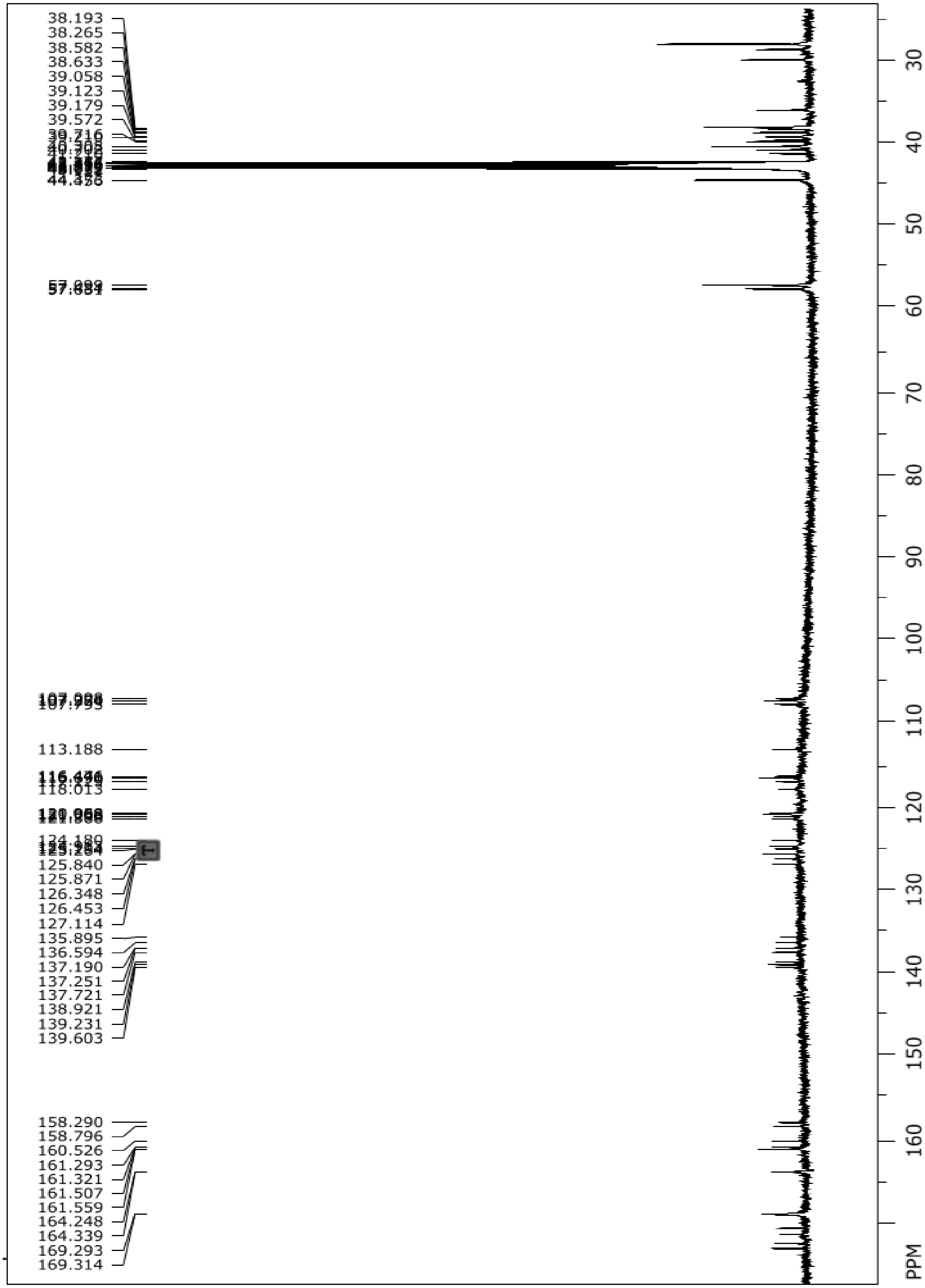
1178's C-NMR:


36.371
37.872
37.987
38.036
38.066
38.264
38.573
38.623
39.052
39.157
39.632
39.713
44.454
57.029
57.629
107.238
109.293
125.784
125.857
126.110
126.227
126.360
126.571
126.626
135.877
137.063
137.218
137.254
137.706
138.881
139.190
139.620
158.248
158.818
160.928
161.214
161.255
161.484
164.260
164.315
168.681
168.731
168.799
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
PPM

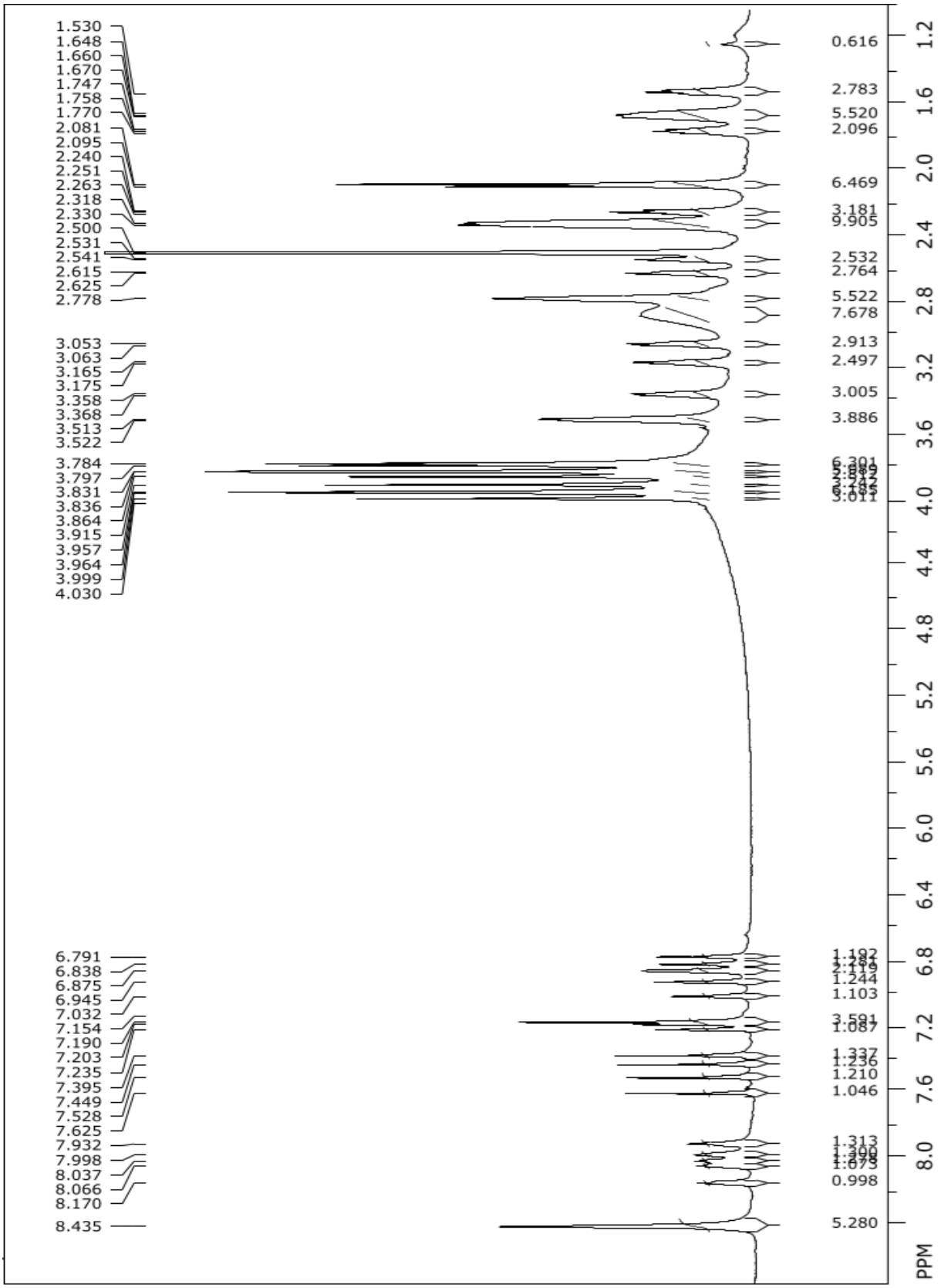
1181's H-NMR:


0.771
0.851
1.170
1.226
1.508
1.519
1.530
1.664
1.674
1.746
1.757
1.769
1.777
2.081
2.094
2.240
2.251
2.262
2.313
2.328
2.384
2.500
2.530
2.539
2.620
2.630
2.780
3.053
3.062
3.164
3.173
3.357
3.367
3.510
3.519
3.782
3.796
3.830
3.835
3.863
3.914
3.952
3.962
3.998
6.838
6.872
6.945
7.004
7.019
7.033
7.190
7.235
7.398
7.440
7.526
7.622
7.939
7.994
8.040
8.073
8.199
8.427
0.334
0.248
1.482
0.886
2.213
5.622
2.483
2.492
0.873
1.688
1.895
5.488
2.411
2.256
2.493
4.205
3.273
1.029
1.180
0.993
0.995
6.556
PPM

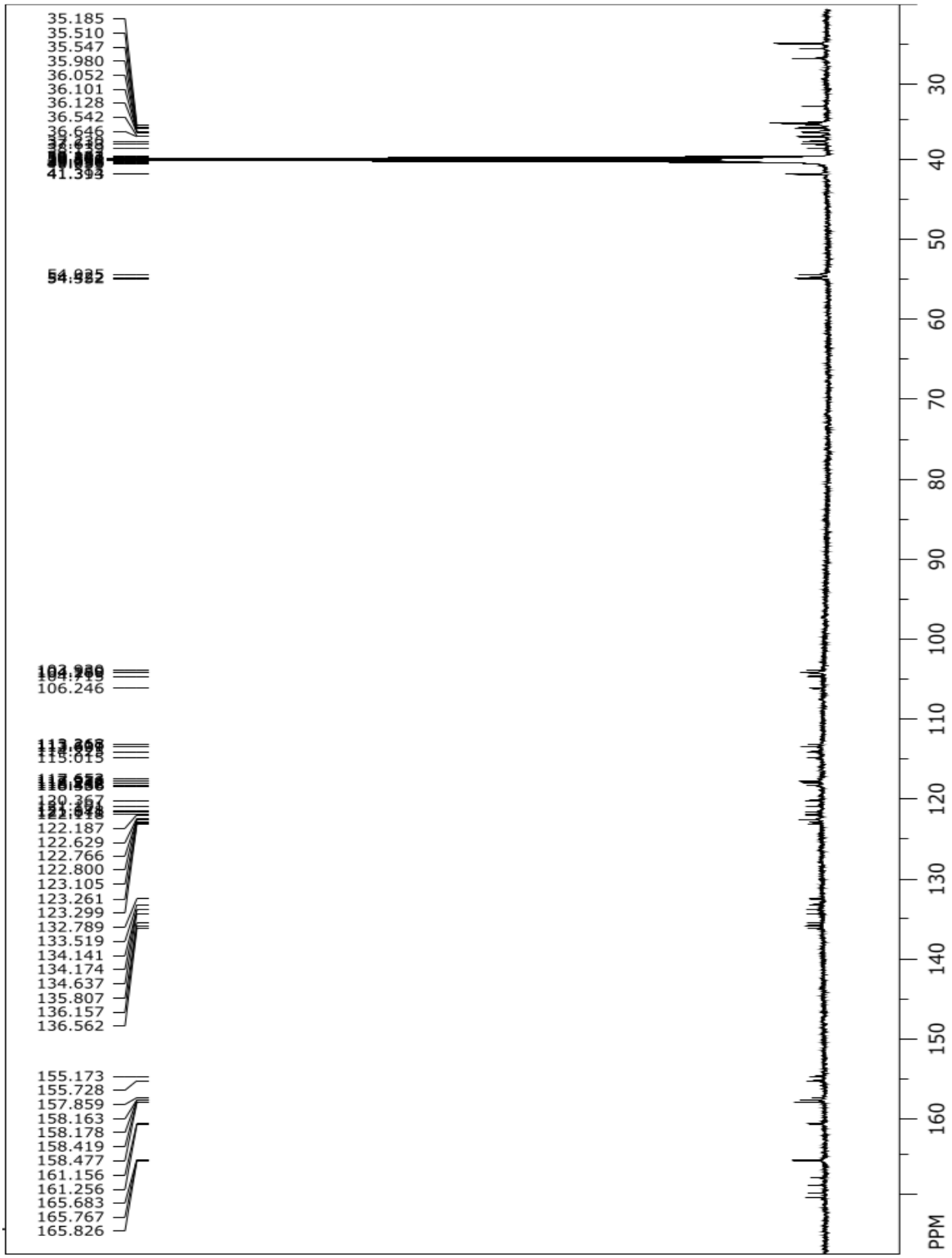
1181's C-NMR:


38.193
38.265
38.582
38.633
39.058
39.123
39.179
39.572
39.716
57.631
113.188
118.013
124.180
125.840
125.871
126.348
126.453
127.114
135.895
136.594
137.190
137.251
137.721
138.921
139.231
139.603
158.290
158.796
160.526
161.293
161.321
161.507
161.559
164.248
164.339
169.293
169.314
PPM

1182 H-NMR:


1.530
1.648
1.660
1.670
1.747
1.758
1.770
2.081
2.095
2.240
2.251
2.263
2.318
2.330
2.500
2.531
2.541
2.615
2.625
2.778
3.053
3.063
3.165
3.175
3.358
3.368
3.513
3.522
3.784
3.797
3.831
3.836
3.864
3.915
3.957
3.964
3.999
4.030
6.791
6.838
6.875
6.945
7.032
7.154
7.190
7.203
7.235
7.395
7.449
7.528
7.625
7.932
7.998
8.037
8.066
8.170
8.435
0.616
2.783
5.520
2.096
6.469
3.181
9.905
2.532
2.764
5.522
7.678
2.913
2.497
3.005
3.886
6.301
3.011
1.192
1.244
1.103
3.591
1.087
1.337
1.210
1.046
1.313
0.998
5.280
1.2
1.6
2.0
2.4
2.8
3.2
3.6
4.0
4.4
4.8
5.2
5.6
6.0
6.4
6.8
7.2
7.6
8.0
PPM

1182 C-NMR:


35.185
35.510
35.547
35.980
36.052
36.101
36.128
36.542
36.646
41.393
54.935
54.982
106.246
115.015
120.367
122.187
122.629
122.766
122.800
123.105
123.261
123.299
132.789
133.519
134.141
134.174
134.637
135.807
136.157
136.562
155.173
155.728
157.859
158.163
158.178
158.419
158.477
161.156
161.256
165.683
165.767
165.826
30
40
50
60
70
80
90
100
110
120
130
140
150
160
PPM