Engineering of a Solution Phase Efficient Pathway for the Development of Nucleobase Containing Chimeric Compounds

M.Sc. Thesis

By

Sudarshan Majee

2003131026

DEPARTMENT OF CHEMISTRY

INDIAN INSTITUTE OF TECHNOLOGY INDORE

2021 - 2022

Engineering of a Solution Phase Efficient Pathway for the Development of Nucleobase Containing Chimeric Compounds

A Thesis

Submitted in partial fulfilment

of requirements for the

awards of the degree

of

Master of Science

by

Sudarshan Majee

2003131026

INDIAN INSTITUTE OF TECHNOLOGY INDORE

CANDIDATE DECLARATION

I hereby certify that the work is shown in the thesis entitled **"Engineering of a Solution Phase Efficient Pathway for the Development of Nucleobase Containing Chimeric Compounds"** in the partial fulfilment of the requirements for the award of the degree of **Master of Science** and submitted to the Department of Chemistry, IIT Indore, is an authentic record of my work carried out during the time period 03/08/2021 to 21/05/2022 under the supervision of **Prof. Apurba K. Das**, Professor, Department of Chemistry, Indian Institute of Technology Indore.

I have not submitted the matter presented in the thesis for the award of any other degree by this or any other institute.

Sudarshan

Sudarshan Majee

This is to certify that the above statement made by the candidate is correct to the best of my/our knowledge.

Prof. Apurba K. Das

Sudarshan Majee has successfully given his M.Sc. oral examination held on the date 25.05.2022.

Signature of supervisor of M.Sc. thesis	Convener, DPGC
Prof. Apurba K. Das	**Dr. Tushar Kanti Mukherjee**
Date: 27.05.2022	Date: 27.05.2022
Signature of PSPC Member	Signature of PSPC Member
Dr. Tushar Kanti Mukherjee	**Dr. Selvakumar Sermadurai**
Date: 27.05.2022	Date: 27.05.2022

ACKNOWLEDGEMENT

I would like to thank my project supervisor **Prof. Apurba K. Das** (Professor, Department of Chemistry, IIT Indore), for allowing me to work in the field of my interest with his research group. His valuable guidance, advice and motivational support throughout the project work help me to complete my project work.

I would also like to thank **Dr. Selvakumar Sermadurai** and **Dr. Tushar Kanti Mukherjee** for their valuable suggestion and guidance.

I wish to express my gratitude to Prof. Suhas S. Joshi, Director, IIT Indore, for his continuous encouragement, help, and support in every aspect.

I would like to thank all faculty members of the Department of Chemistry for their motivational inspiration.

I would like to thank all lab members for their valuable suggestions, and help without their support would not be possible.

I would like to thank the Department of Chemistry IIT Indore for working in the laboratory.

I am also thankful to SIC, IIT Indore and its members for their technical help and support.

I would like to express my gratitude to all of my friends who assisted me directly or indirectly during my post-graduation program; in particular, I would like to thank Sajal, Saroj, Shalini, and Ankita. I have learnt a lot from them, both in chemistry and outside it.

I would also like to convey my deepest gratitude to my family and friends for their kind cooperation, understanding, and valuable support.

Sudarshan Majee

M.Sc. 2nd year

Roll No. 2003131026

The Department of Chemistry IIT Indore

DEDICATED

TO

MY FAMILY

ABSTRACT

Nucleopeptides are an important class of molecules having diverse applications in therapeutics, diagnostics and biomaterials development. The nucleopeptides have significant importance due to their ability to target complementary DNA and RNA strands. In addition, these compounds have interesting self-assembly properties and are used as self-replicating materials in prebiotic chemistry and constituents of drug delivery systems. Therefore, we have focused on synthesising purine functionalised amino acids derivatives. All the intermediates are purified using column chromatography and characterised by mass spectrometry and NMR spectroscopy.

NOMENCLATURES

α	Alpha
J	Coupling constant
˚C	Degree Celsius
δ	Delta (chemical shift)
d	Doublet
γ	Gamma
g	Gram
Hz	Hertz
K	Kelvin
mg	Milligram
mL	Milliliter
m	Multiplet
nm	Nanometer
ppm	Parts per million
s	Singlet

ACRONYMS

DMSO	Dimethyl sulfoxide
$(Boc)_2O$	Di-*tert*-butyl decarbonate
EtOAc	Ethyl acetate
MeOH	Methanol
NMR	Nuclear Magnetic Resonance
DMF	*N, N*-dimethylformamide
K_2CO_3	Potassium carbonate
NaOH	Sodium hydroxide
Na_2SO_4	Sodium sulfate
Boc	*tert*-butoxycarbonyl
THF	Tetrahydrofuran
TBAI	Tetra-n-butylammonium iodide
TLC	Thin Layer Chromatography

LIST OF FIGURES

TABLE OF CONTENT

Chapter 6

Chapter 1

1.1 INTRODUCTION

Chimeric compounds may often develop by the conjugation of two biologically active compounds. This idea describes the formation of a novel bifunctional molecule in which all the individual constituents have some specific pharmacological activity [1]. Therefore, the integration of bioactive nucleobases with amino acids/ small peptides fabricates an entirely new class of materials, called nucleopeptide which holds considerable biological importance [2]. Willardiine, a natural nucleobases uracil containing amino acid derivative which plays a role of an agonist of AMPA receptor, plays a crucial role in neurological disorders [3]. In another example, N6-threonylcarbamoyl adenosine was found to play a vital role in maintaining decoding accuracy during the protein synthesis process [4]. Xu et al. successfully developed short nucleopeptides based hydrogelators upon connecting the nucleobases with di, tripeptides, which self-assemble in water to form supramolecular hydrogels under physiological conditions [5].

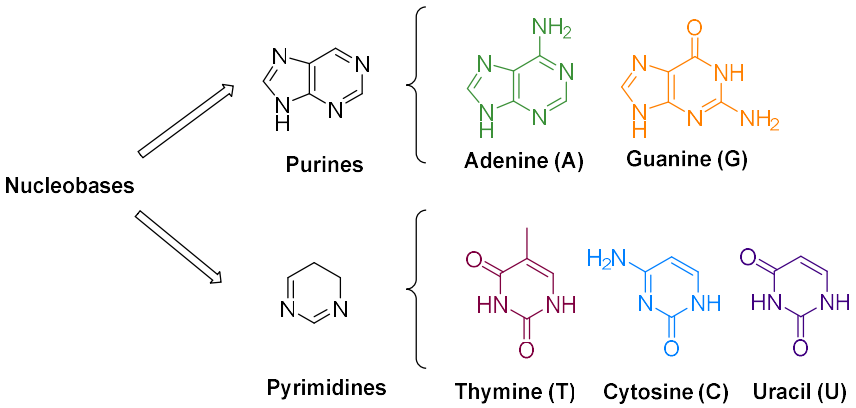


Figure 1. Purine and pyrimidine-based nucleobases.

Now, nucleobases can be classified into two groups, one purine (adenine, guanine) and another pyrimidine (cytosine, thymine and uracil) based on their chemical structures. Purine bases offer more H-bonding donor and

acceptor sites than pyrimidine bases, promoting significant non-covalent interactions to generate purine functionalised supramolecular self-assembled analogues. Guanine (G) and its derivatives are well known to form different self-assembled architectures such as G-quartets, G-ribbon and helical structures under suitable conditions. In general, G develops self-assembled G-quartets structures in the presence of suitable metal ions through the intermolecular H-bonding and ion-dipole interactions. Further, these quartets stack on top of one another through the π-π stacking interactions to form G-quadruplex structures [6]. On the other hand, G-ribbon and helical structures are formed without metal ions [7,8].

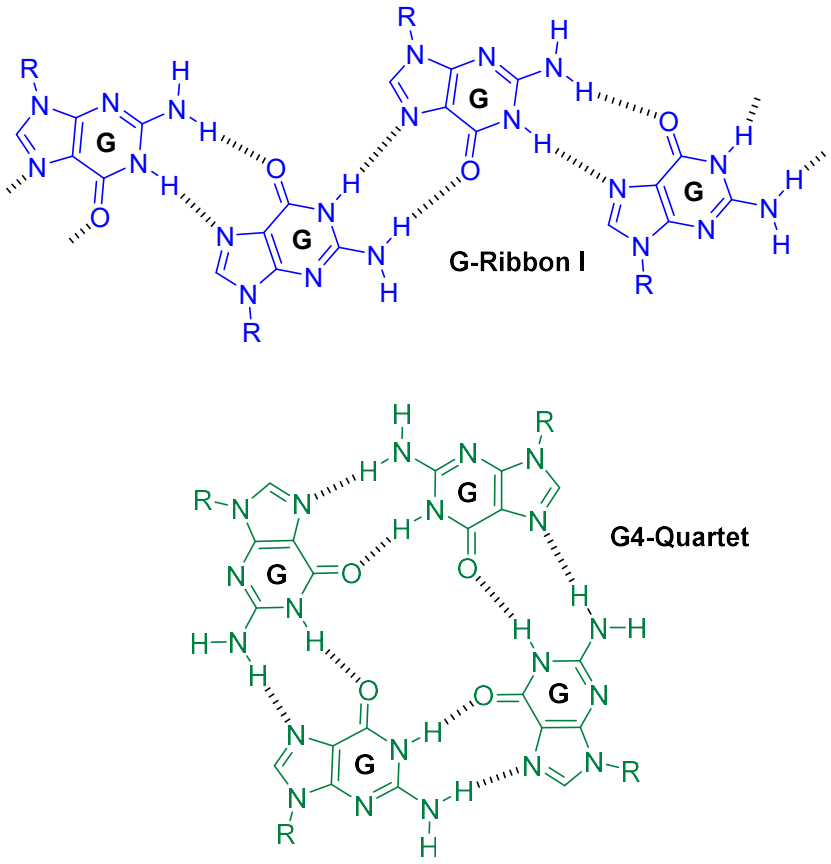


Figure 2. G-ribbon-I and the G4-quartet are examples of different self-assembly motifs formed from Guanine-containing molecules.

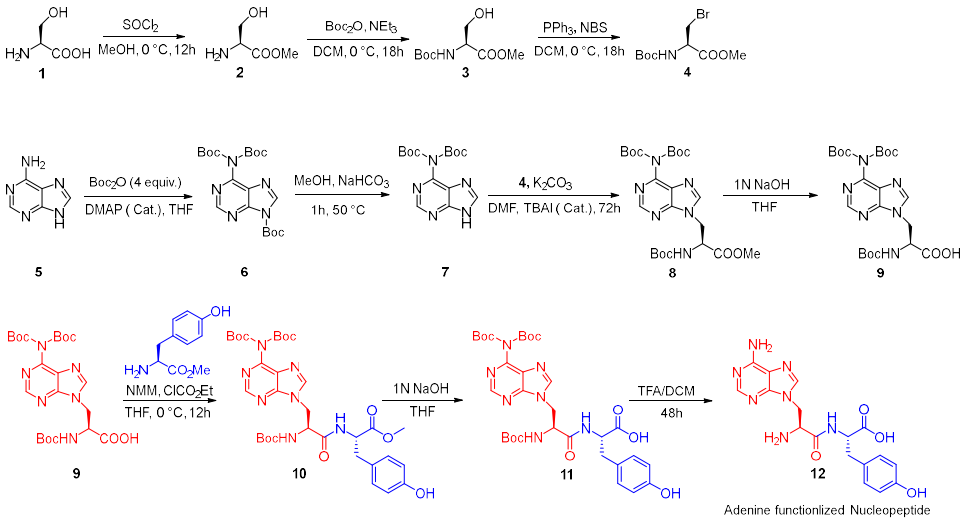
Therefore, guanine and its functionalised derivatives hold special attention because of their unique self-assemblies, the self-assembled structures offering significant importance in biological ion transportation, molecular engineering [9], and molecular electronics [10].

L-Tyrosine plays a vital role in binding protein-nucleic acid, like in the case of herpes simplex virus DNA binding proteins recognising DNA [11]. In this process, the interaction of the L-tyrosine residues on the protein surface with nucleic acid is very dominant. Aromatic stacking pairing of tyrosine-adenine-tyrosine plays a vital role in the N-glycosidase activity of gelonin, derived from the plant *Gelonium multiorum*, a ribosome-inactivating protein, proposing the importance of this bonding in the RNA binding by this protein [12]. The development of covalent tyrosine-nucleobase adducts apart from aromatic non-covalent nucleobase-tyrosine bonds has shown that this type of by-product plays a crucial role in DNA protein crosslinking works under DNA exposure which helps in ionising radiation [13].

Because residues of L-tyrosine in proteins play a role in binding to nucleic acids and also RNA can interact with that free L-tyrosine, we aimed to develop an entirely new efficient synthetic pathway to synthesise nucleobases (adenine and guanine) conjugating tyrosine-based nucleopeptides. The *N-9* substituted analogues of natural purine nucleosides have been synthesised differently in which the nucleobases are present at the side chain of the peptide backbone.

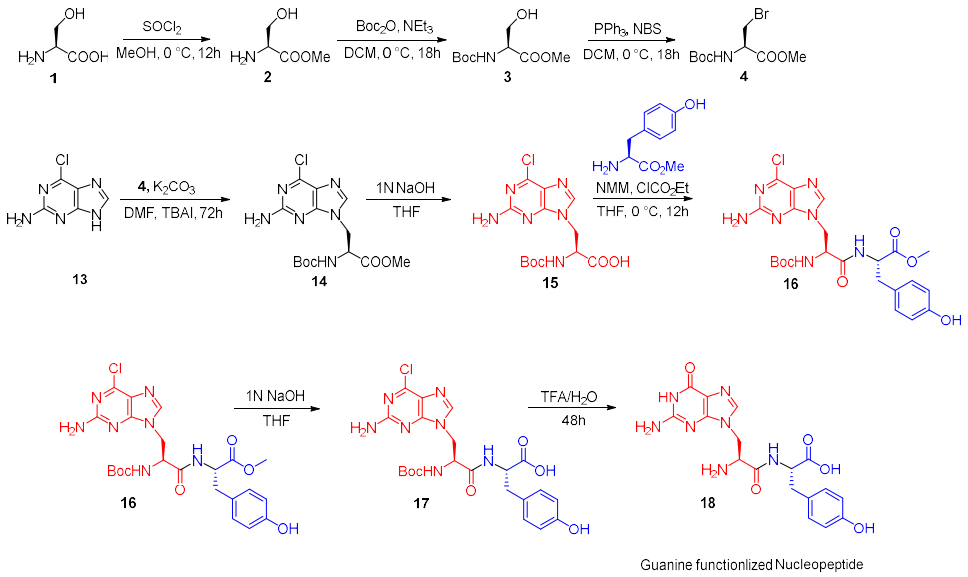
1.2 Reaction Scheme

Scheme 1: Overall Synthesis of Compound (12)



L-serine (1) was initially converted into serine methyl ester (**2**), followed by *N*-termini amino group protection using Boc-anhydride. The Boc-protected ester (**3**) was then converted into the bromo alanine derivative (**4**) using NBS as the brominating agent. On the other hand, adenine (**5**) was dissolved in THF, and excess Boc-anhydride was added in the presence of DMAP as a catalytic amount to synthesise Tris-Boc protected adenine. The Tris-Boc adenine (**6**) was converted into Bis-Boc-adenine (**7**) under mild basic conditions. Then the alanyl PNA monomer (**8**) was synthesised by substitution reaction on the Bromo-alanine derivative (**4**) using Bis-Boc-adenine (**7**) as the nucleophile. Then hydrolysis of compound (**8**) took place in the presence of a base to form compound (**9**); later, tyrosine methyl ester was coupled with compound (**9**) to form adenine based peptide methyl ester (**10**) in the presence of coupling reagent N-Methyl morpholine and Ethyl chloroformate followed by hydrolysis to get compound (**11**) finally, deprotection was taken place in the presence of acidic medium to get our final compound (**12**). The intermediates and the final product were characterised using mass spectrometry and ^{1}H and ^{13}C NMR spectroscopy.

Scheme 2: Overall Synthesis of Compound (18)



On the contrary, a substitution reaction has taken place on bromo alanine derivative (**4**) using 2-amino-6 chloropurine (**13**) as a nucleophile and TBAI as a catalyst giving compound (**14**) as a product. Later ester hydrolysis of compound (**14**) is carried out with 1N NaOH, resulting in compound (**15**). Successively coupling reaction has occurred between tyrosine methyl ester and compound (**15**) using coupling reagents N-methyl morpholine and Ethyl chloroformate to get compound (**16**). After hydrolysis was carried out using the base for transforming methyl ester into acid, i.e., compound (**17**). Removing the chlorine from 2-amino-6-chloropurine by acid-catalyzed hydrolysis pathway is an efficient route for forming guanine residue from its 2-amino-6-chloropurine precursor. Hence, compound (**17**) was treated with TFA/H_2O to acquire our desired product (**18**). The intermediates and the final product were characterised using mass spectrometry and ^{1}H and ^{13}C NMR spectroscopy.

Chapter 2

2.1: Experimental Section

2.1.1 Materials

The solvents and reagents were purchased from commercially available sources. Alfa Aesar, Sigma Aldrich-India, Merck, Spectro-chem and TCI are some available sources.

Adenine and potassium carbonate (K_2CO_3) was obtained from Alfa Aesar. Sodium hydroxide, Ethyl chloroformate, Thionyl Chloride and diethyl ether were purchased from Spectro-chem, whereas DMF, THF, methanol, 1,4-dioxane, were purchased from Merck. 1N NaOH, Boc-Anhydride, triphenylphosphine, DMAP, *N*-Methyl morpholine, L-serine and Trifluoroacetic acid (TFA) were obtained from SRL, and NBS was obtained from TCI. For moisture sensitive reactions, the dry solvent has been used in the presence of N_2 or Ar gas. After completing the reaction, column chromatography was done for further purification using silica (100-200 mesh) as an immobile phase and hexane, ethyl acetate or toluene as a mobile phase.

2.1.2 General

TLC monitored the course of reactions. All ^{1}H and ^{13}C NMR spectra were set down on Bruker Avance (500 MHz) instrument at 25 ℃. Using ESI positive mode, mass spectra were set down on the Bruker instrument. The NMR spectra of all intermediates and final compounds were analysed using MestReNova software. The NMR samples were prepared in DMSO-d_6 and $CDCl_3$. Solvent. The chemical shift was expressed in the form of ppm (δ) relative to surplus solvents protons as internal standards ($CHCl_3$: δ 7.26, DMSO: δ = 2.50 For ^{1}H NMR; and $CHCl_3$: δ 77.00, DMSO: δ 39.50 for ^{13}C NMR).

Chapter 3

3. Results and Discussions

3.1 Synthesis of (*S*)-2-(*N*-Boc-amino)-3-bromopropinoic acid methyl ester (4) as crucial intermediate:

The preparation of short nucleodipeptides can be divided into the synthesis of nucleobase functionalised amino acid derivatives and the coupling of second aromatic acid (L-tyrosine) with the nucleobase-amino acid monomer (**Scheme 1** and **2**). In the first case, we have chosen the amino acid L-serine to synthesise our alanyl nucelo amino acids (**9** and **14**). In this approach, the (*S*)-2-(*N*-Boc-amino)-3-bromopropinoic acid methyl ester (**4**) is the key primary intermediate to substitute the two nucleobases (A, G), resulting in the corresponding alanyl nucleo amino acids. This key intermediate (4) was prepared from L-serine in three steps, as displayed in **scheme 1**. At first, the free carboxylic group of L-serine is protected to its methyl ester (**2**). The L-serine methyl ester was then reacted with Boc-anhydride in the presence of triethylamine and DCM as a solvent to synthesise *N*-Boc-serine methyl ester (**3**) with a yield of 90%. Compound 3 was then transformed to the required bromo derivative **4** according to the literature reported procedure for converting the primary hydroxyl group to their bromo derivative by introducing NBS as brominating agents in the presence of PPh_3. Thus in a precooled solution of NBS and PPh_3 in DCM, compound **3** was added in the presence of the catalytic amount of pyridine to afford compound **4** in a 70% yield.

3.2 Derivatization of purine nucleobase:

In this step, we focus on synthesising bis-Boc protected purine nucleobases to prevent or minimise the side reaction during peptide coupling step and use of toxic high boiling point solvent such as DMF. Further, this bis-Boc protected method has improved the poor solubility of purine nucleobases in most organic solvents. The boc-protected purine

nucleobases were prepared by following Garner's way with a little modification [14,]. This pathway provides an efficient approach to the gram-scale synthesis of the corresponding Bis-boc protected nucleobase (**7**) without further purification steps. In this step, to a suspension of adenine (**5**) in dry THF, an excess amount of Boc-anhydride was added in the presence of a catalytic amount of DMAP to produce tris-Boc-protected adenine (**6**). The progress of the reaction was monitored with TLC as the reaction mixture gradually turned into a clear solution. This tris-Boc-protected adenine (**7**) can be easily transformed into corresponding bis-Boc-adenine (**8**) by treating aqueous $NaHCO_3$ in MeOH at a reflux condition. The reaction was finished within 2h in quantitative yield. However, we observed the significant formation of adenine with a longer reaction time under reflux conditions in a basic medium.

In the case of guanine derivatisation, due to the poor solubility and regioselectivity of guanine, it is replaced with a 2-amino-6-chloropurine (13) analogue. However, we used this 2-amino-6-chloropurine for the next alkylation step without transforming it into its bis-Boc derivative.

3.3 Synthesis of Purine amino acids monomer (9 & 14):

The exocyclic amino group of adenine was protected with Boc, as mentioned above, to synthesise the adenine amino acid monomer. The alkylation of bis-Boc-adenine (**8**) with (S)-2-(N-Boc-amino)-3-bromopropinoic acid methyl ester (**4**) was performed in the presence of K_2CO_3 in DMF. This substitution step was accomplished by the first formation of anion in K_2CO_3 (as a base) in dry DMF, followed by substituting primary bromide with the nucleobase (Scheme **1** and **2**) to produce **9**. The nucleophilic substitution was successfully performed with a moderate yield of 40%. This substation reaction was optimised in the presence of a TBAI catalyst (Table 1). As shown in Table 1, the highest result (40%) was obtained in 0.10 Equiv. Of TBAI with the reaction time of 72 h. It is noteworthy that a very low yield was obtained without TBAI.

Further, the progress of the reaction was not improved with an excess amount of TBAI.

Similarly, the 2-amino-6-chloropurine (**13**) was suspended in the presence of K_2CO_3, and anhydrous DMF, followed by the addition of **4** to produce the corresponding purine nucleo amino acids (**14**).

3.4 Development of nulceodipeptides (12 & 18):

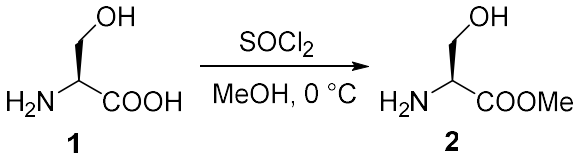
In this step, the nucleo amino acids monomers (i.e. adenine monomer **9** and 2-amino-6-chloro purine monomer **15**) were coupled with L-tyrosine methyl ester to formulate the nucleodipeptides (Scheme 1 & 2). Therefore, the methyl ester groups in compounds **8** and **14** were successfully deprotected into free carboxylic acid groups by simply ester hydrolysis methods. This ester hydrolysis method was carried out under basic conditions (1N NaOH) at 0 °C, and corresponding acid-free nucleo amino acid derivatives (**9** & **15**) were produced in 65 to 70% yields. The solution phase peptide coupling of these compounds was performed using ethyl chloroformate as peptide coupling reagents. In detail, the acid-free nucleo amino acid derivatives (9 & 15) were cooled in dry THF under the non-nucleophilic base *N*-methyl morpholine (NMM). Then after stirring at ice cold-conditions for 10 mins, ethyl chloroformate was dropwise added to allow the formation of mixed anhydride and a solution of L-tyrosine methyl ester in THF was added to it, resulting in production of nucleodipeptides (**10**, **16**). The hydrolysis of resulting methyl ester protected nucleopeptides (**10**, **16**) was performed with 1N NaOH in THF at continuous ice-cold conditions, and the reaction proceeded quickly (3-4 h) to obtain the corresponding *C*-terminal (carboxyl) free nucleodipeptides **11** and **17** as off-white solid in quantitative yields (≥75%). The N-Boc groups in 11 were finally deprotected by using a 1:1 solution of TFA and DCM. Therefore, both *N*-and *C*- terminal free nucleodipeptide (adenine functionalised nucleodiepeptide **12**) was isolated after lyophilisation in yield 68%.

Similarly, in the case of 2-amino-6-chloropurine containing nucleopeptide, the corresponding guanine base was generated in the presence of TFA/H_2O conditions. The chlorine group in 2-amino-6-chloropurine (**17**) can be easily converted to a 6-oxo derivative under acid-catalyzed hydrolysis. Further, upon using TFA, the N-Boc group of amino acids was also deprotected. The corresponding TFA salts from the free amino acids were isolated by precipitation with Et_2O. Finally, the solid predicated were dried under vacuum to obtain guanine functionalised nucleodipeptide (**18**).

Chapter 4

4: Synthesis of Compounds

4.1. Synthesis of Compound **2**



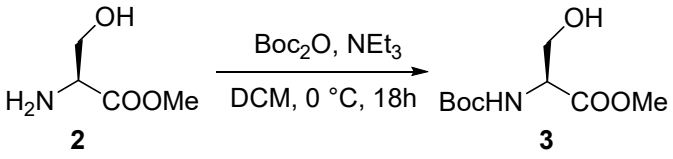
1 **2**

In a 250 mL RB flask, MeOH (30 mL) was added and kept to stir under ice-cold conditions. After cooling for 5 min, $SOCl_2$ (8.14 mL, 113.22 mmol) was added to it under an ice-cold condition. Then L-serine (6 g, 56.61 mmol) was added to the RB and allowed to stir for 12 h. Then the solution was evaporated in a rotavapor, and the residue was co-evaporated with Et_2O. The white solid was collected and resulted in L-serine methyl ester.

Yield = 91% (5.47 g). ^{1}H NMR (500 MHz, DMSO-d_6): δ = 8.60 (s, 2H), 4.10 (s, 1H), 3.83 (s, 2H), 3.74 (s, 3H), 3.39 (s, 1H); ^{13}C NMR (125 MHz, DMSO-d_6, TMS): δ = 169.19, 59.89, 54.85, 53.23 ppm.

ESI-MS m/z: $[M+Na]^+$ calcd for $C_4H_9NO_3$ 142.0475; found 142.0473.

4.2. Synthesis of Compound **3**



2 **3**

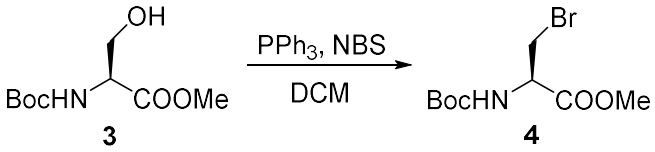
Compound **2** was taken in a 100 mL RB flask and dissolved in DCM (25 mL). After that, Boc-anhydride (3 mL, 12.8 mmol) was added under ice-cold conditions followed by the addition of compound **2** (2.0 g, 12.8 mmol) was added and then Et_3N (1.94 g, 19.2 mmol) was added dropwise over 10 min. The reaction mixture was kept on stirring condition for 18 h. The reaction mixture was then diluted with DCM and later washed with

sat aq NH_4Cl (2×15 mL), sat aq $NaHCO_3$ (2×15 mL) and brine (2×15 mL). The organic layer was collected and dried over Na_2SO_4. The organic parts were concentrated under rotavapor to get colourless oil as Boc-protected amino acids.

Yield = 90% (3.63 g). ^{1}H NMR (500 MHz, $CDCl_3$): δ = 5.62 (d, 1H, J = 6.5 Hz), 4.30 (s, 1H), 3.88 (s, 2H), 3.81-3.78 (m, 1H), 3.70 (s, 3H), 1.38 (s, 9H). ^{13}C NMR (125 MHz, $CDCl_3$, TMS): δ = 171.57, 155.86, 85.22, 69.10, 62.69, 52.56, 28.25ppm.

ESI-MS m/z: $[M+Na]^+$ calcd for $C_9H_{17}NO_5$ 242.0999; found 242.1149

4.3. Synthesis of Compound **4**



3 **4**

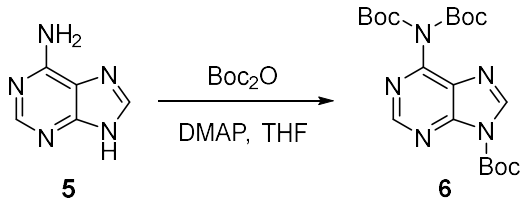
Solution of triphenylphosphine (3.35 g, 12.77 mmol) was dropwise added in a suspension of NBS (2.43 g, 13.68 mmol) in anhydrous DCM with ambient temperature. The evolved reddish-brown mixture was stirred for a further 5 min. Then an anhydrous solution of **(3)** (1 g, 4.55 mmol) was put into anhydrous DCM; the solution was stirred for 12h.

Eventually, the reaction mixture was concentrated in rotavapor then the residue was co-evaporated with toluene. Then the dark brown oily residue was titrated with diethyl ether. The purification of the product was carried out using column chromatography using EtOAc- hexane.

Yield = 67% (0.67 g). ^{1}H NMR (500 MHz, $CDCl_3$): δ = 5.40 (d, 1H, J = 6.75 Hz), 4.74 (m, 1H, J = 7.85 Hz), 3.78 (s, 3H), 3.71-3.68 (m, 2H), 1.38 (s, 9H); ^{13}C NMR (125 MHz, $CDCl_3$, TMS): δ = 171.57, 155.86, 85.22, 69.10, 62.69, 52.56, 28.25ppm.

ESI-MS m/z: $[M+Na]^+$ calcd for $C_9H_{17}NO_5Na$ 242.0999; found 242.1149.

4.4. Synthesis of Compound **6**



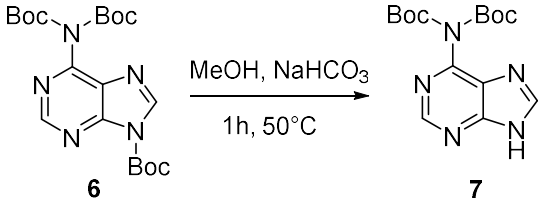
5 6

In an Ar-flushed RB flask, adenine (1 g, 7.4 mmol) and DMAP (0.362 g, 2.96 mmol) were suspended in 20 mL of dry THF. The reaction mixture was then left to stir for 10 mins. After that, Boc-anhydride (6.46 g, 29.60 mmol) was slowly mixed under the Ar atmosphere. Finally, the reaction mixture was kept in stirring condition for 12 h. Then the excess amount of THF was evaporated, and the residue was dissolved in excess ethyl acetate. The ethyl acetate part was washed three times with HCL (1N) and brine solution. The organic portion was dehydrated over Na_2SO_4 and concentrated in rota vapor to yield a pale-yellow oil.

Yield = 83% (0.827 g). ^{1}H NMR (500 MHz, $CDCl_3$): δ = 8.94 (s, 1H), 8.45 (s, 1H), 1.65 (s, 9H), 1.36 (s, 18H); ^{13}C NMR (125 MHz, $CDCl_3$): δ = 154.04, 152.44, 151.19, 150.01, 145.62, 143.22, 129.57, 87.48, 83.94, 27.90, 27.71 ppm.

ESI-MS m/z: $[M+Na]^+$ calcd for $C_{20}H_{29}O_6N_5Na$ 458.2010; found 458.2112.

4.5. Synthesis of Compound **7**



6 7

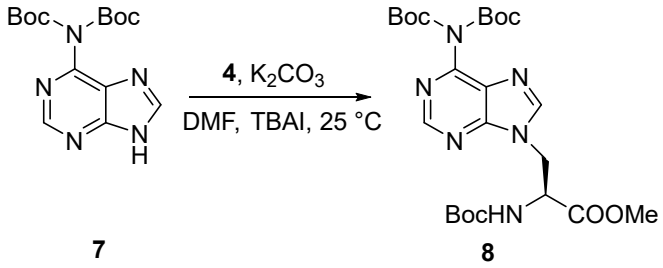
Initially, compound **6** (1.0 g, 3.3 mmol) was added to MeOH (30 mL) and dissolved, followed by 15 mL of saturated $NaHCO_3$ solution was added. The resulting turbid solution was then allowed to reflux at 50 °C for 1 hour, and TLC tracked the course of the reaction. After the reaction, the solvents were evaporated, and the aqueous layers were extracted with the

DCM. The DCM parts were collected, dried over Na_2SO_4 filtered and concentrated under vacuum to give **7** a solid white powder.

Yield: 82% (0.821 g). ^{1}H NMR (500 MHz, $CDCl_3$): δ = 8.78 (s, 1H), 8.33 (s, 1H), 1.46 (s, 18H); ^{13}C NMR (125 MHz, $CDCl_3$, TMS): δ = 151.88, 150.21, 144.01, 84.91, 27.76 ppm.

ESI-MS m/z: $[M+Na]^+$ calcd for $C_{15}H_{21}N_5O_4Na$ 358.1486; found 358.1562.

4.6. Synthesis of Compound **8**



7 **8**

In the compound 8 synthesis, TBAI was used as a catalyst for enhancing the efficiency of the product. When the reaction was conducted at room temp and with no catalyst, only the trace amount of the product was isolated. As the equivalent of TBAI was increased, the yield was also increased. A yield of 25% was obtained when TBAI was raised to 0.075 equivalents. Even on increasing the amount of catalyst from 0.10 Equiv. to 0.20 equiv., there was no further increment in the yield. Therefore, the catalyst amount was fixed as 0.10 equivalents.

To a solution of **7** (1.08 g, 3.22 mmol) in dry DMF, anhydrous K_2CO_3 (0.445 g, 3.2 mmol) was added under an inert (N_2) atmosphere. The reaction mixture was then left to stir for 30 mins. Then, compound **4** (0.908 g, 3.22 mmol) was added followed by the addition of tetra-n-butylammonium iodide (TBAI) (0.059 g, 0.161 mmol). Then stirring was carried out for 72h. Then, the dilution of the reaction was carried out

Table 1. Optimisation table of the reaction.

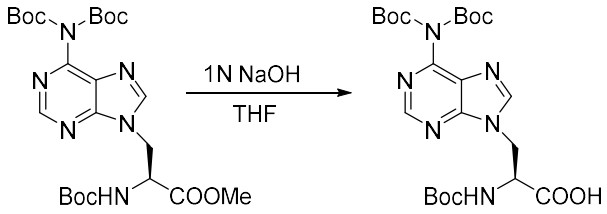
S.No.	Catalyst	Amount (Eq.)	Time (h)	Yield (%)
1	TBAI	0.0	48	10
2	TBAI	0.0	72	12
3	TBAI	0.05	72	20
4	TBAI	0.075	72	25
5	**TBAI**	**0.10**	**72**	**40**
6	TBAI	0.15	72	40
7	TBAI	0.20	72	40
8	TBAI	0.10	42	35
9	TBAI	0.10	96	40
10	TBAI	0.10	120	40

Using excess ethyl acetate and the organic layer was washed with 1N HCl (3×22 mL) and brine (2×12 mL). The organic part was dehydrated over sodium sulphate and concentrated in rota vapor to get the crude. The crude was then purified using column chromatography using 60% ethyl acetate/hexane as eluent.

Yield: 40% (0.378 g). ^{1}H NMR (500 MHz, $CDCl_3$): δ = 8.84 (s, 1H), 5.36 (s, 1H), 4.76 (m, 2H), 3.76 (s, 3H), 1.46 (s, 18H), 1.25 (s, 9H); ^{13}C NMR (125 MHz, $CDCl_3$): δ = 168.5, 152.0, 151.86, 151.39, 151.03, 150.57, 135.65, 125.40, 67.31, 52.95, 52.45, 30.81, 30.19 ppm.

ESI-MS m/z: $[M+K]^+$ calcd for $C_{24}H_{36}N_6O_8K$ 575.2226; found 575.2074.

4.7. Synthesis of Compound **9**



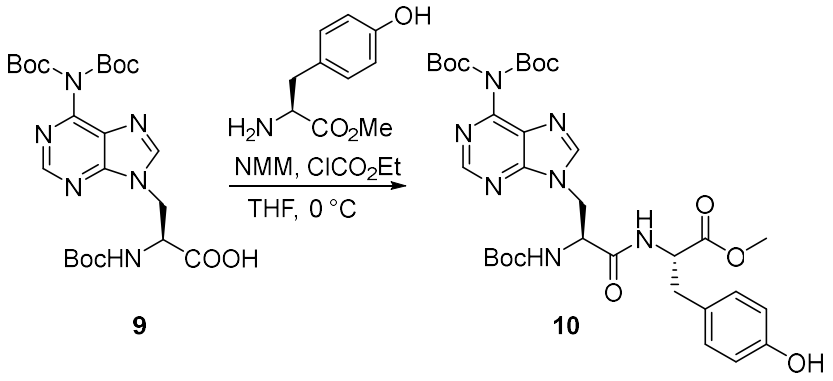
8 **9**

Compound **8** (0.375 g, mmol) was taken in a round bottom flask and dissolved in THF under ice-cold conditions. After cooling, 1 N NaOH was portion-wise added till the reaction was completed. TLC monitored the course of the reaction. After completion of the reaction, THF was evaporated, and the reaction mixture was diluted with distilled water. The water part was washed with diethyl ether. Then the pH of that aqueous part was adjusted to 2 by adding $KHSO_4$. The aqueous portion was then drawn out with ethyl acetate (3×30 mL). Then collected the organic layer and dried it with Na_2SO_4. The ethyl acetate part was evaporated to yield compound **9**.

Yield: 75% (0.450 g). ^{1}H NMR (500 MHz, $CDCl_3$): δ = 8.82 (s, 1H), 8.24(s, 1H), 4.91 (s, 1H, J = 3.59 Hz), 4.71 (s, 2H), 3.63 (s, 1H), 1.54 (s, 9H), 1.43 (s, 18H); ^{13}C NMR (125 MHz, $CDCl_3$): δ = 171.97, 155.38, 153.93, 152.36, 151.70, 150.40, 150.11, 145.97, 127.88, 80.66, 54.02, 45.39, 27.91 ppm.

ESI-MS m/z: $[M-H]^-$ calcd for $C_{23}H_{35}N_6O_8$ 521.2354; found 521.1962.

4.8. Synthesis of Compound **10**



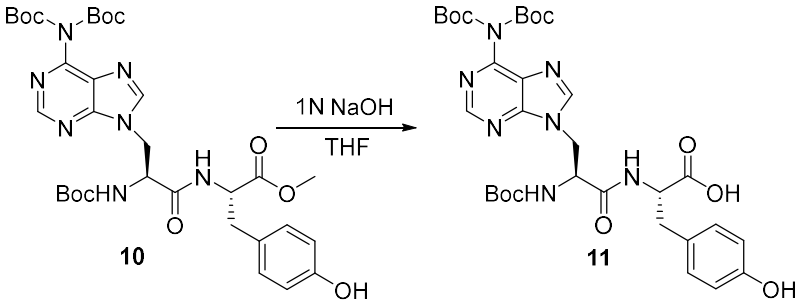
9 **10**

In a 100 mL r.b. flask, compound **9** (0.300 g, 0.57 mmol) was dissolved in THF under an inert atmosphere along with ice-cold conditions. The solution was left to stir for a duration of 10 mins, and ethyl chloroformate (0.062 g, 0.57 mmol) was added to the reaction mixture, followed by *N*-methyl morpholine (NMM) (0.057 g, 0.57 mmol). The reaction was then allowed to stir for 30 mins under ice-cold conditions. Then L-tyrosine methyl ester (0.134 g, 0.69 mmol) was added, and the reaction was left for 12 h. THF was evaporated in the rota vapor, and the residue was solvated in ethyl acetate. The ethyl acetate layer was then washed with 1N HCl (3×20 mL) and brine (2×15 mL) and again with 1N NaOH (3×1530 mL) and brine (2×10 mL). The organic portion was collected and dried using Na_2SO_4 and concentrated *in vacuo* to give the crude product. The purification of the crude was carried out using column chromatography using 60% ethyl acetate-hexane. Compound **10** was obtained as an off-white solid.

Yield: 70% (0.560 g)

ESI-MS m/z: $[M+Na]^+$ calcd for $C_{33}H_{46}N_7O_{10}$ 722.3120; found 722.3050.

4.9. Synthesis of Compound **11**



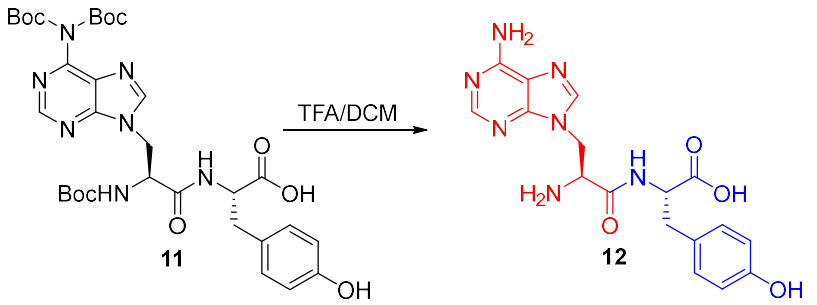
Compound **10** (0.375 g) was taken in an R.B. flask and dissolved in THF under ice-cold conditions. After cooling the reaction mixture, 1N NaOH was portion-wise added till the reaction was complete. TLC monitored the course of the reaction. Then THF was evaporated in the rota, and dilution of the reaction mixture was done with distilled water. The water part was

washed with diethyl ether. Then the pH of that aqueous part was adjusted to 2 by adding $KHSO_4$. The aqueous layer was then extracted by ethyl acetate (3×30 mL). The organic layer was collected and dried over Na_2SO_4. The ethyl acetate part was evaporated in rota vapor to afford the desired compound **11**.

Yield: 75% (0.510 g). ^{1}H NMR (500 MHz, DMSO-d_6): δ = 9.15 (d, 1H, J = 12.3), 8.51 (s, 1H), 8.36 (s, 1H), 7.31 (s, 1H), 7.19 (d, 2H, J = 7.95), 6.96 (dd, 2H), 4.98 (s, 2H), 4.29 (s,1H), 4.06 (d, 3H, J = 4.65), 3.44 (dd, 1H), 3.18 (dd, 1H), 1.76 (s, 18H), 1.58 (s, 9H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 168.8, 155.6, 155.4, 152.2, 152.1, 151.6, 151.3, 150.8, 135.9, 130.3, 129.9,128.4, 125.6, 115.7, 84.8, 84.4, 53.6, 53.2, 52.7, 50.9, 34.4,30.45 ppm.

ESI-MS m/z: $[M+H]^+$ calcd for $C_{32}H_{44}N_7O_{10}$ 586.2455 ; found 586.2544.

4.10. Synthesis of Compound **12**



Adenine functionlized Nucleopeptide

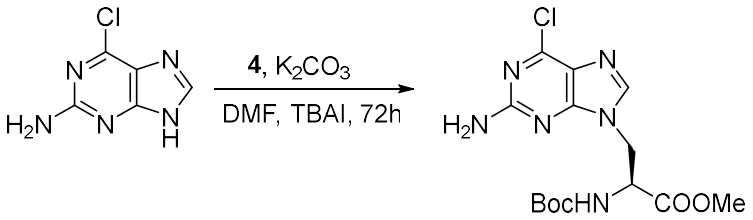
The compound was added in DCM, and TFA (1 mL) was added to it. Then stir the reaction mixture for 24h. After the completion of the reaction, TFA was evaporated under a vacuum, and the residue was diluted with water. The aqueous part was washed with diethyl ether (2×10 mL) followed by adjustment of pH to 9 by addition of NH_4OH. Then obtained solution is lyophilised.

Yield: 68% (0.220 g). ^{1}H NMR (500 MHz, DMSO-d_6): δ = 8.31 (s, 1H), 8.14 (s, 1H), 8.00 (s, 1H), 7.88 (s, 1H), 7.17 (s, 2H), 6.92 (dd, 2H), 6.63

(d, 2H, *J* = 5Hz), 4.37 (d, 1H, *J* = 3Hz), 4.09 (s, 1H), 3.50 (s, 1H), 2.87 (s, 2H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 175.9, 172.7, 158.3, 158.1, 155.9, 152.3, 149.8, 141.3, 130.2, 127.4, 118.4, 115.0, 114.9, 61.9, 60.4, 53.9, 36.2 ppm.

ESI-MS m/z: $[M+H]^+$ calcd for $C_{17}H_{20}N_7O_4$ 386.1571 ; found 386.1697

4.11. Synthesis of Compound **14**



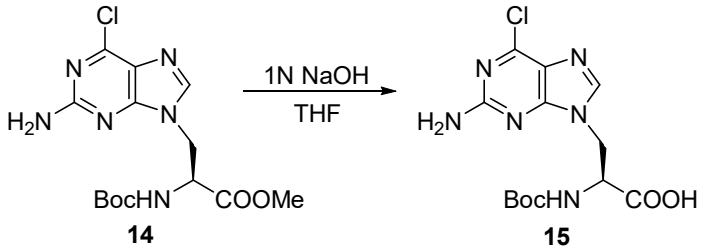
13 **14**

To the solution of **13 (**0.135 g, 0.80 mmol) in dry DMF, anhydrous K_2CO_3 (0.111 g, 0.81 mmol). Then the reaction is stirred for 30 mins. Then compound **4** (0.150 g, 0.53 mmol) was put into the reaction followed by the introduction of tetra-n-butylammonium iodide (TBAI) (0.019 g, 0.053 mmol). Then stirring was carried out for 72h. After completion of the reaction, the dilution was done with excess ethyl acetate and washed by using 1N HCl (3×15 mL) and brine (2×15 mL). The organic portions were collected and dried over Na_2SO_4 and concentrated in a vacuum to give a pale-yellow crude. The crude was then purified using column chromatography using 40% ethyl acetate-hexane as eluent.

Yield: 65% (0.550 g). ^{1}H NMR (500 MHz, DMSO-d_6): δ = 7.88 (s, 1H), 7.32 (d, 1H, *J* = 5Hz), 6.77 (s, 2H), 4.46 (d, 2H), 4.24 (dd, 1H), 3.66 (s, 3H), 1.30 (s, 9H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 169.9, 159.7, 155.1, 153.8, 149.4, 142.9, 123.3, 52.4, 52.1, 27.9 ppm.

ESI-MS m/z: $[M+H]^+$ calcd for $C_{14}H_{20}N_6O_4Cl$ calcd. 371.1229, found 371.1279.

4.12. Synthesis of Compound **15**



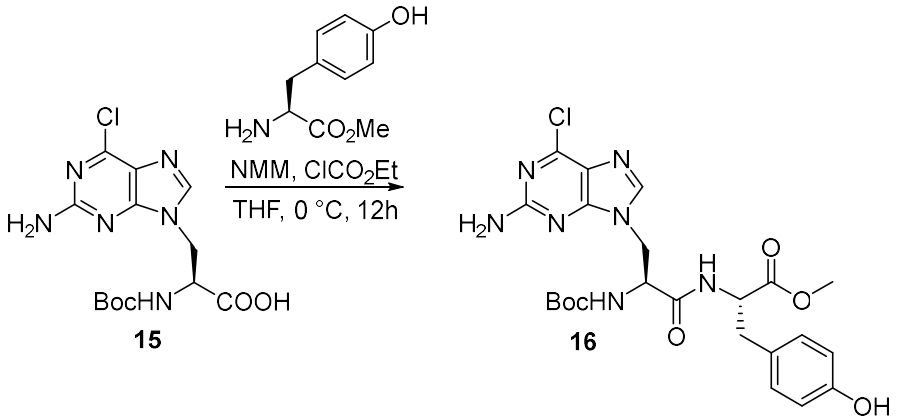
14 **15**

Compound 14 (0.250 g) was dissolved in THF at an ice-cold condition in an R.B flask. After cooling the reaction mixture, 1N NaOH was portion-wise added till the reaction was complete. TLC monitored the course of the reaction. Then THF was evaporated in the rota, and the reaction mixture was diluted with distilled water. The water part is washed with diethyl ether. Then the pH of that aqueous part was adjusted to 2 by adding $KHSO_4$. The aqueous layer was then extracted by ethyl acetate (3×30 mL). The organic layer was then dried over Na_2SO_4 and evaporated in the rota vapor to yield compound **15**.

Yield: 70% (0.250 g). ^{1}H NMR (500 MHz, DMSO-d_6): δ = 12.84 (s,1H), 7.93 (s, 1H), 7.29 (d, 1H, J = 8.2), 6.94 (s, 2H), 4.45 (dd, 1H), 4.30 (m, 2H), 1.28 (s, 9H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 171.3, 159.8, 155.3, 154.2, 149.3, 143.4, 123.2, 78.6, 52.42, 28.02.

ESI-MS m/z: $[M+H]^+$ calcd for $C_{13}H_{18}N_6O_4Cl$ calcd. 357.1073, found 357.1090.

4.13. Synthesis of Compound **16**



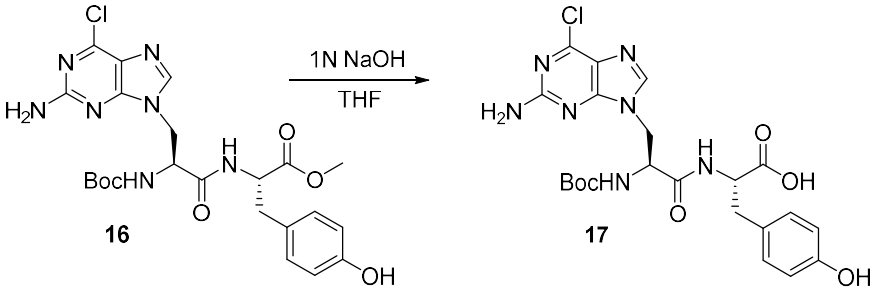
15 **16**

In a 100 ml R.B. flask, compound **15** (0.050 g, 0.14 mmol) was dissolved in THF under an inert atmosphere and ice-cold conditions. The reaction mixture was kept on stirring for 10 mins, and then Ethyl Chloroformate (0.018 g, 0.21 mmol, 16µL) was added to it, followed by the addition of N-methyl morpholine (NMM) (0.014 g, 0.14 mmol, 16µL) in an interval of 10mins. After stirring the reaction mixture for half an hour, Tyrosine methyl ester (0.041 g, 0.21 mmol) was added and stirred the reaction for 12h. After the conclusion of the reaction, THF was evaporated in the rota vapor and the residue was dissolved in ethyl acetate. The ethyl acetate part was washed with 1N HCl (3×15 mL) and brine (2×10 mL) and again with 1N NaOH (3×15 mL) and brine (2×10 mL). The organic portions were dried over Na_2SO_4 and concentrated in rota vapor to give compound **16**.

Yield: 60% (0.650 g). ^{1}H NMR (500 MHz, DMSO-d_6): δ = 9.65 (s, 1H), 8.07 (dd, 1H), 7.96 (d, 1H, J = 6.95), 7.58 (dd, 1H), 7.45 (m, 1H), 7.40 (s,1H), 7.25 (dd, 2H), 6.94 (s, 2H), 4.48 (m, 2H), 4.01 (m, 1H), 3.65 (s, 3H), 1.23 (s, 9H).

ESI-MS m/z: $[M+H]^+$ calcd for $C_{23}H_{29}N_7O_6Cl$ calcd. 534.1862 found 534.1599.

4.14. Synthesis of Compound **17**

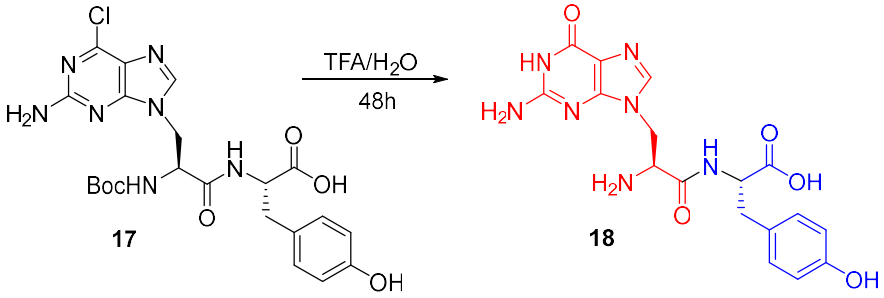


Under ice-cold conditions, compound 16 (0.150 g) was initially taken in an R.B. flask and dissolved in THF. After cooling the solution, NaOH (1 N) was put into the reaction mixture. TLC monitored the course of the reaction. After completion of the reaction, THF was evaporated in the rota vapor, and the response was diluted with distilled water. Ether wash was

carried out, and the water component was collected, pH of the water part was adjusted to mild acidic by adding $KHSO_4$ and later extracted with Ethyl Acetate. Sodium sulphate is used to dry the organic layer and evaporate in the rota to afford the desired compound **17**.

Yield: 75% (0.345 g). ^{1}H NMR (500 MHz, DMSO-d_6): δ = 7.94 (s, 1H), 7.40 (m, 1H), 7.31 (d, 1H, J = 7.5), 6.95 (s, 2H), 6.88 (s, 2H), 6.55 (s, 2H), 4.46 (d, 1H), 4.39 (m, 1H), 4.22 (dd, 2H), 3.51(s, 2H), 1.36 (s, 9H); ESI-MS m/z: $[M+H]^+$ calcd for $C_{22}H_{26}N_7O_6ClNa$ calcd. 542.1512 found 542.2495.

4.15. Synthesis of Compound **18**



Under nitrogen atmosphere (0.10 g, 1.6 mmol) of compound (17) was put in 1 mL 50% TFA/water and the reaction mixture was kept on stirring for 24h. Then solvents were evaporated in rotavapor, and the crude product was precipitated by the addition of diethyl ether. The white powder was collected by filtration.

Yield: 60% (0.345 g). ^{1}H NMR (500 MHz, DMSO-d_6): δ = 10.85 (s, 1H), 8.62 (s, 4H), 7.66 (s, 1H), 7.10 (dd, 1H), 4.47 (s, 2H), 4.42 (d, 2H, J = 5.15), 3.40 (s, 1H), 3.17 (s, 1H).

^{13}C NMR (125 MHz, DMSO-d_6): δ = 168.3, 167.95, 156.6, 153.7, 153.0, 151.1, 137.7, 128.5, 116.5, 110.4, 108.6.

Chapter 5

5.1. Characterization

^{1}H and ^{13}C NMR spectral data

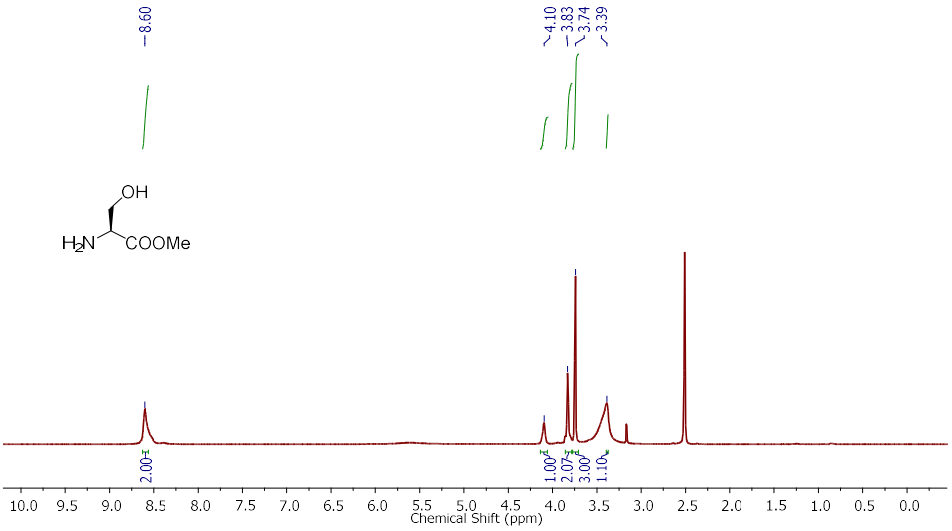


Figure 3. ^{1}H NMR (500 MHz, DMSO) spectrum of compound **2**.

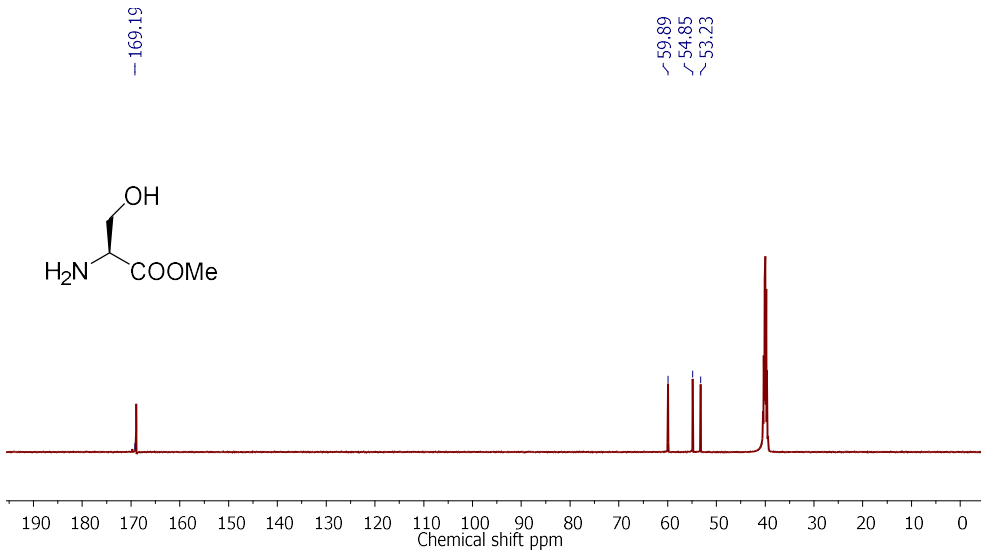


Figure 4. ^{13}C NMR (125 MHz, DMSO) spectrum of compound **2**.

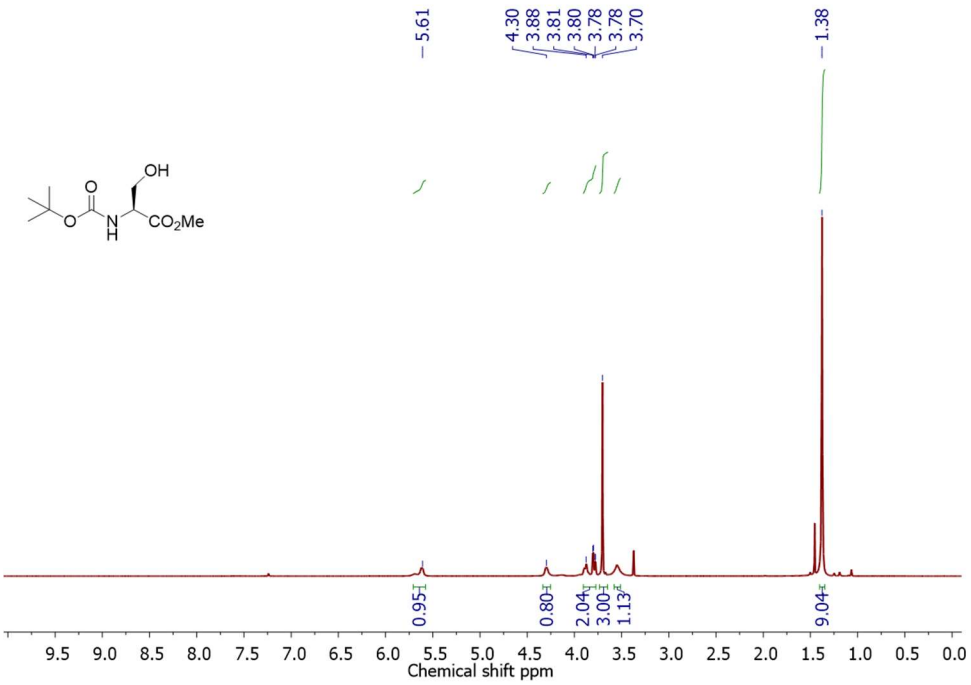


Figure 5. ^{1}H NMR (500 MHz, $CDCl_3$) spectrum of compound **3**.

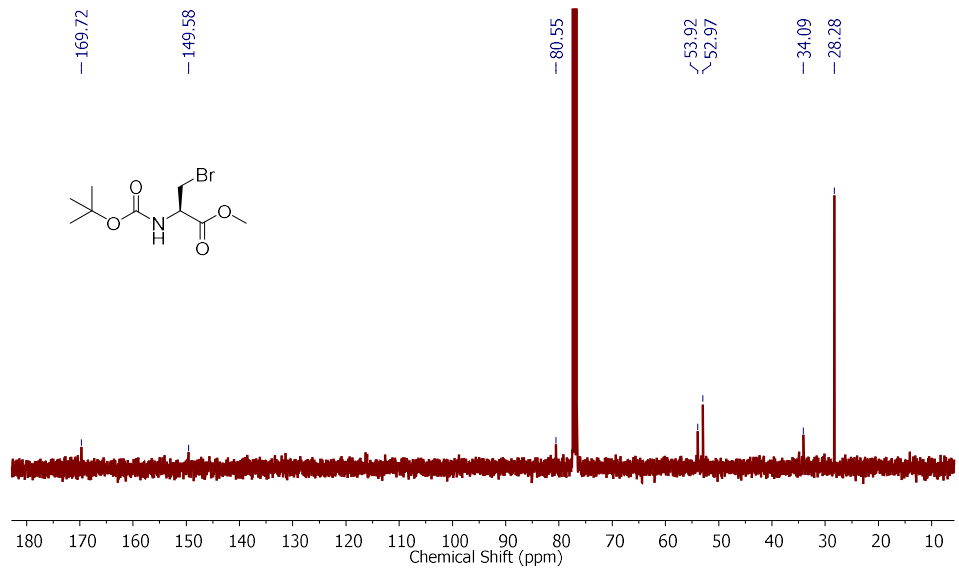


Figure 6.^{13}C NMR (125 MHz, $CDCl_3$) spectrum of compound **3**.

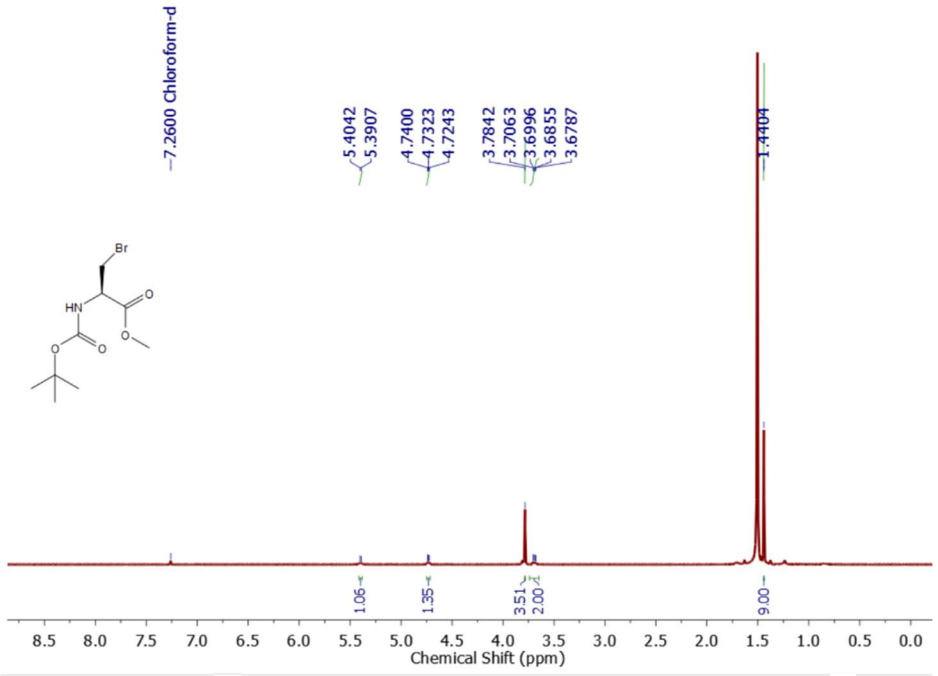


Figure 7. ^{1}H NMR (500 MHz, $CDCl_3$) spectrum of compound **4**.

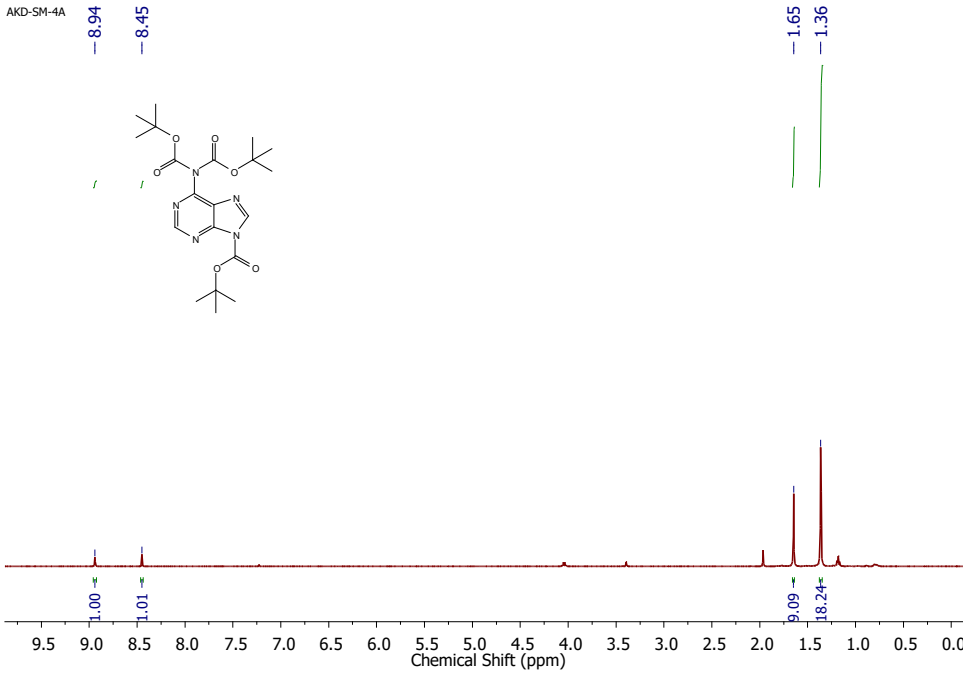


Figure 8. ^{1}H NMR (500 MHz, $CDCl_3$) spectrum of compound **6**.

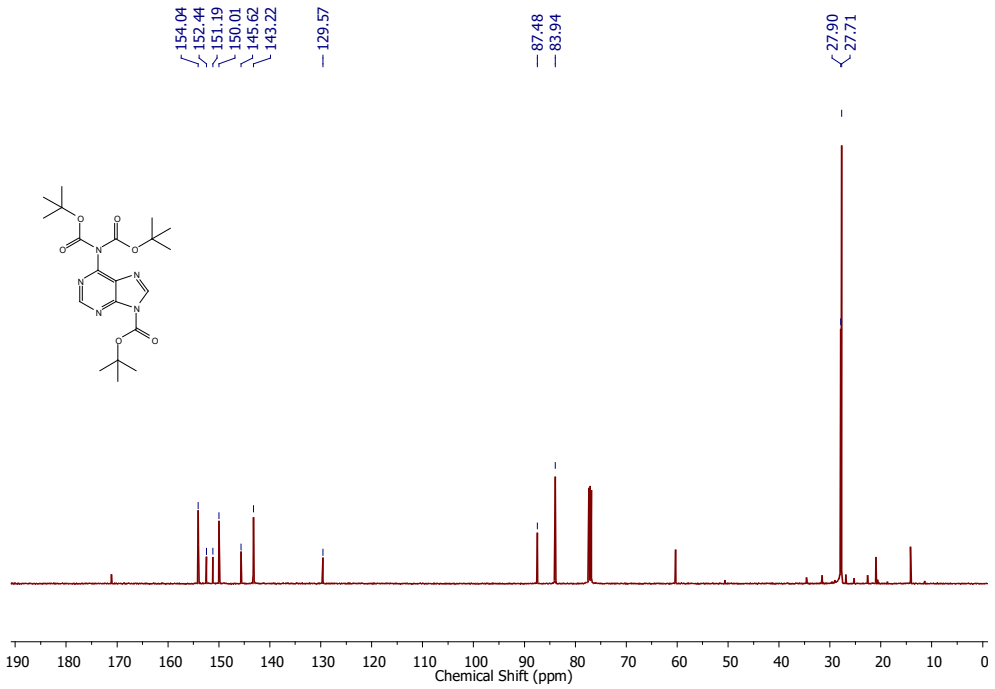


Figure 9. ^{13}C NMR (125 MHz, $CDCl_3$) spectrum of compound **6.**

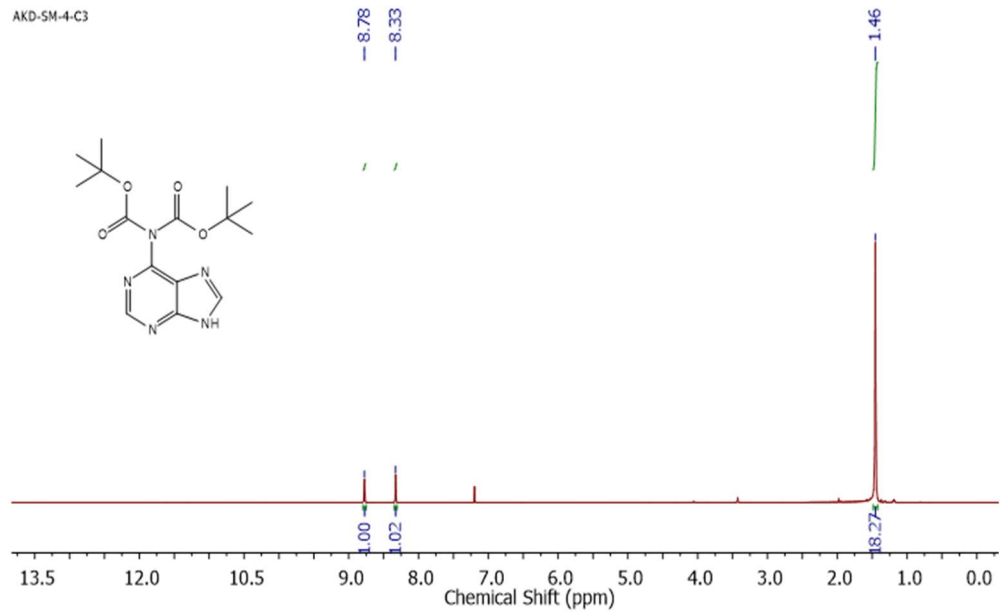


Figure 10. ^{1}H NMR (500 MHz, $CDCl_3$) spectrum of compound **7**.

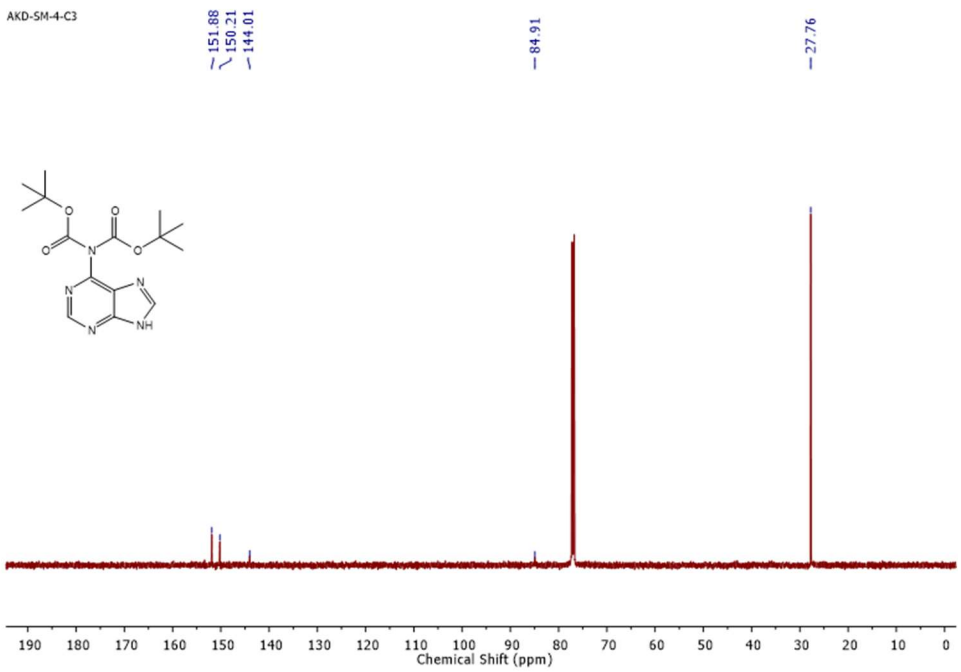


Figure 11. ^{13}C NMR (125 MHz, $CDCl_3$) spectrum of compound **7.**

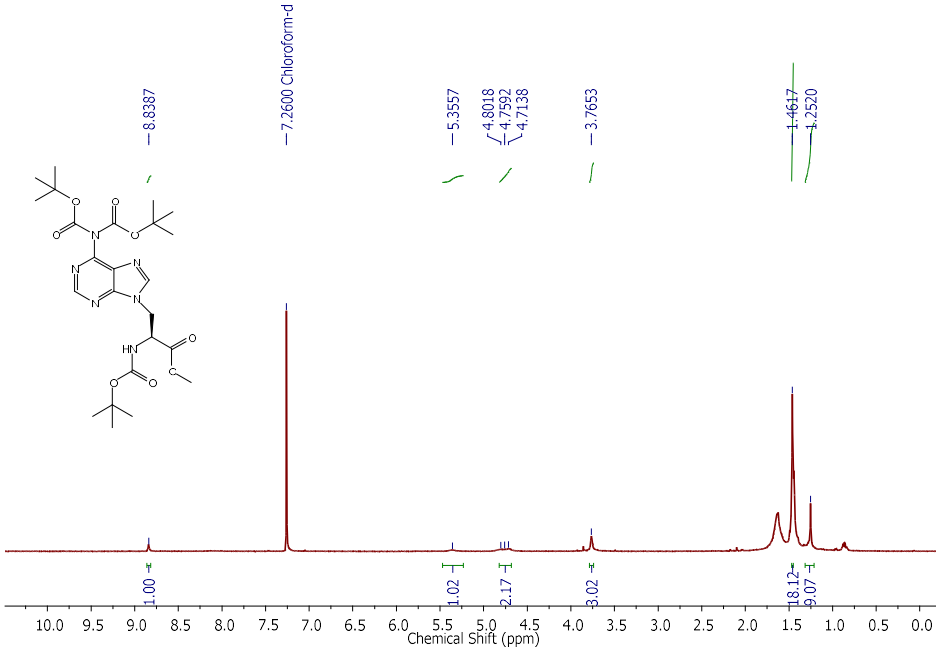


Figure 12. ^{1}H NMR (500 MHz, $CDCl_3$) spectrum of compound **8**.

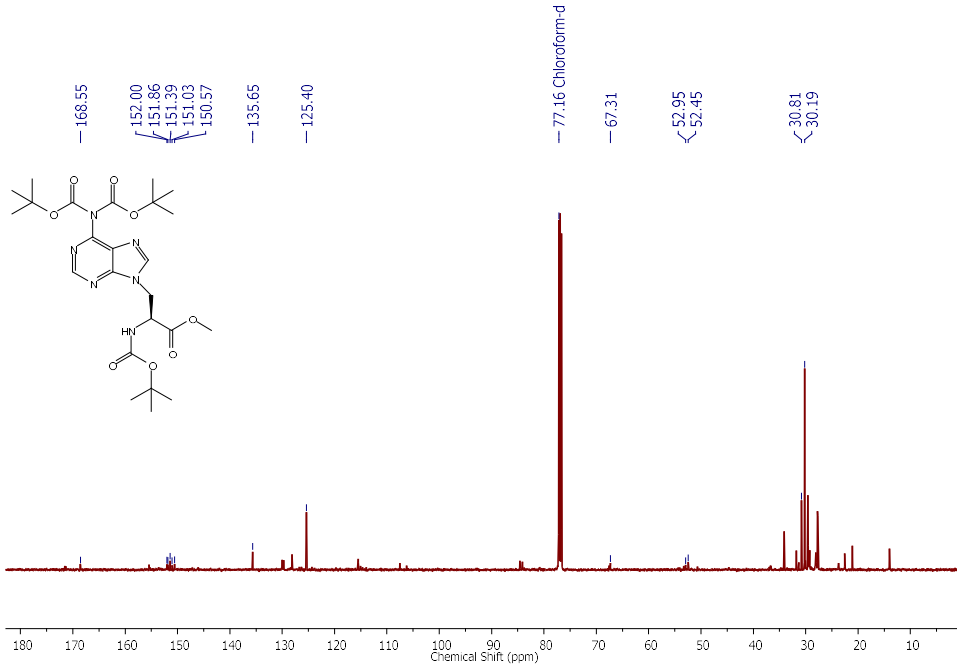


Figure 13. ^{13}C NMR (125 MHz, $CDCl_3$) spectrum of compound **8.**

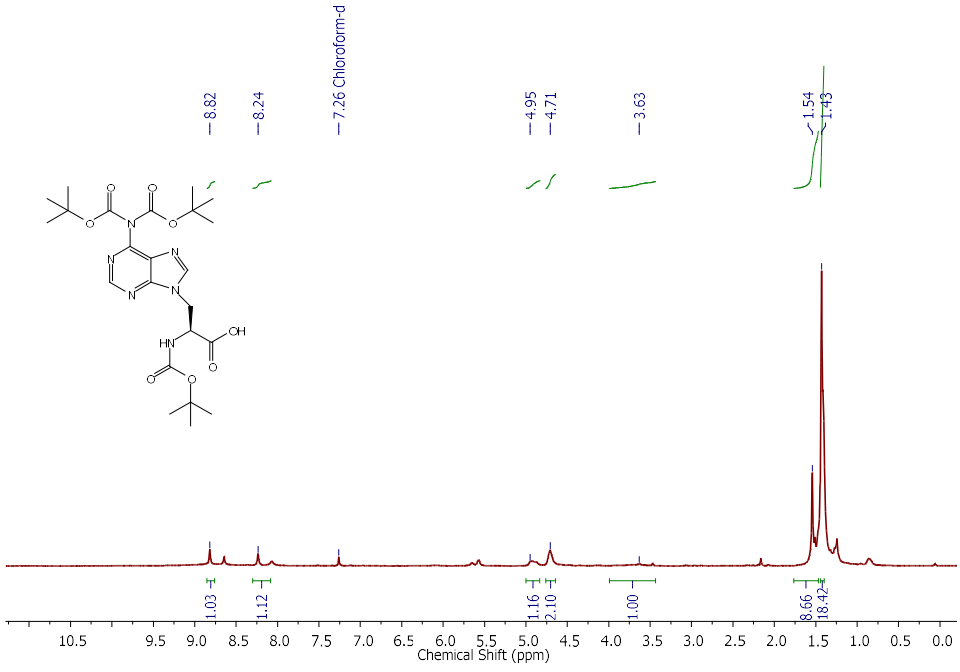


Figure 14. ^{1}H NMR (500 MHz, $CDCl_3$) spectrum of compound **9**.

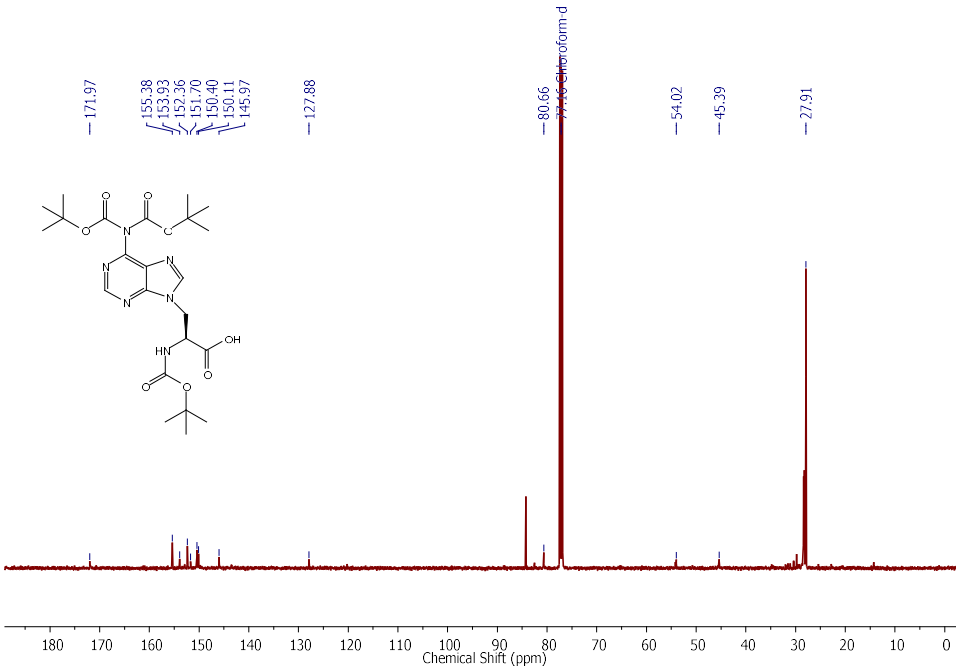


Figure 15. ^{13}C NMR (125 MHz, $CDCl_3$) spectrum of compound **9.**

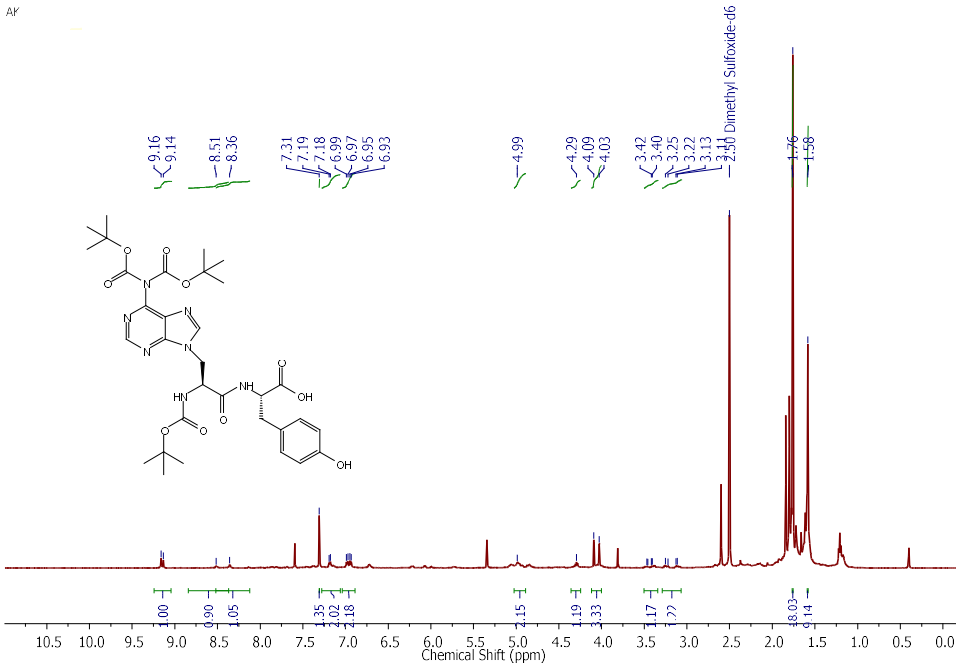


Figure 16. ^{1}H NMR (500 MHz, $CDCl_3$) spectrum of compound **11**.

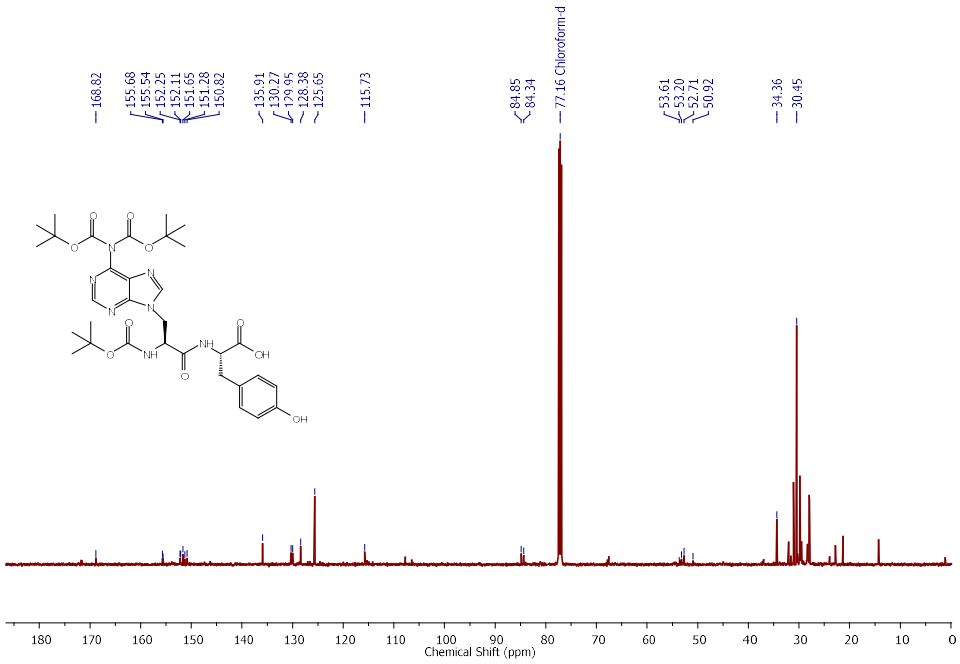


Figure 17. ^{13}C NMR (125 MHz, $CDCl_3$) spectrum of compound **11.**

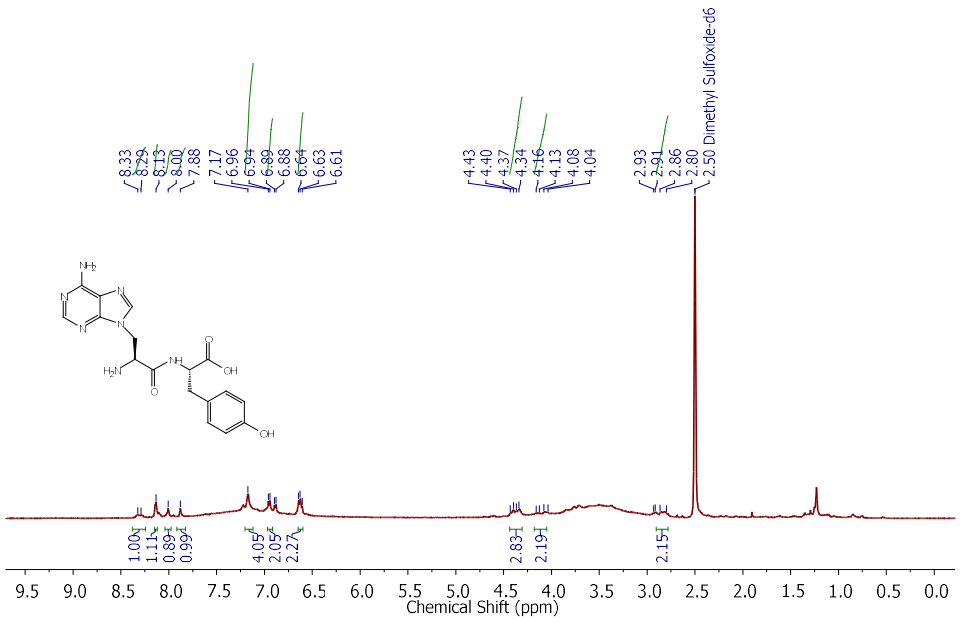


Figure 18. ^{1}H NMR (500 MHz, DMSO-d_6) spectrum of compound **12**.

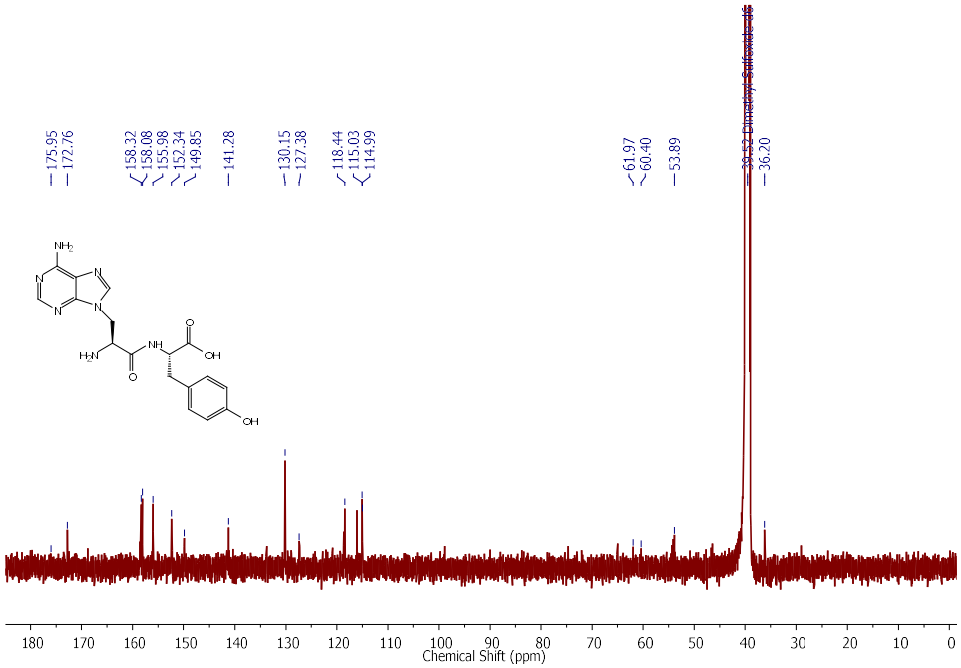


Figure 19. ^{13}C NMR (125 MHz, DMSO-d_6) spectrum of compound **12.**

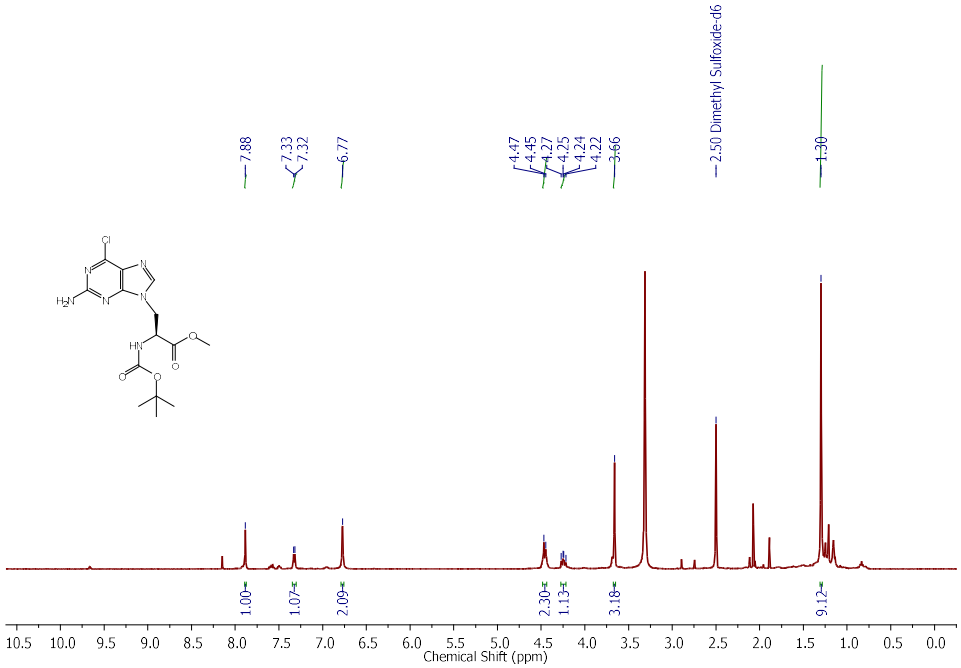


Figure 20. ^{1}H NMR (500 MHz, DMSO-d_6) spectrum of compound **14**.

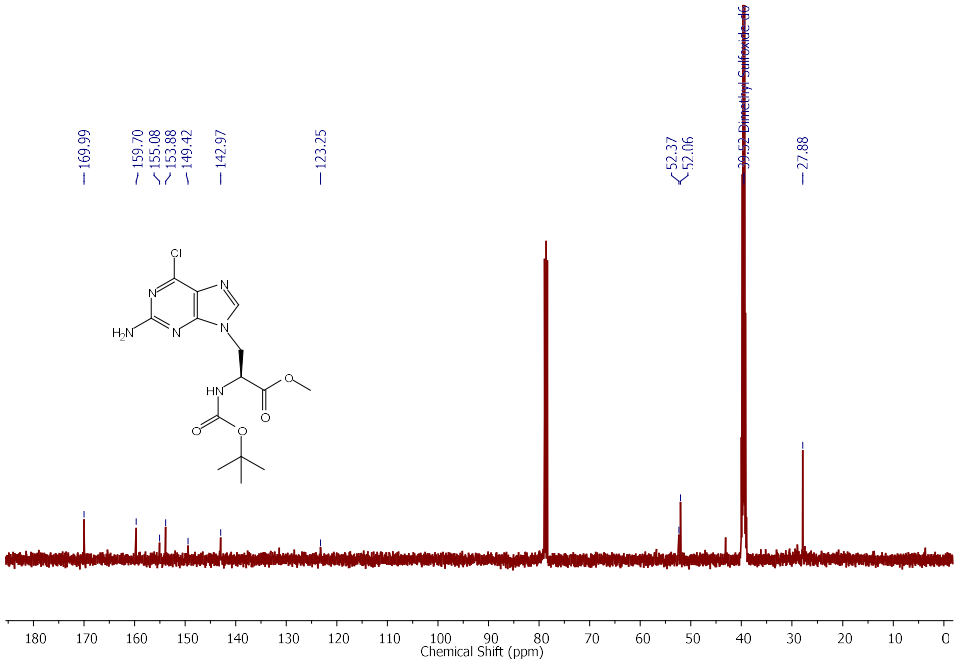


Figure 21. ^{13}C NMR (125 MHz, DMSO-d_6) spectrum of compound **14.**

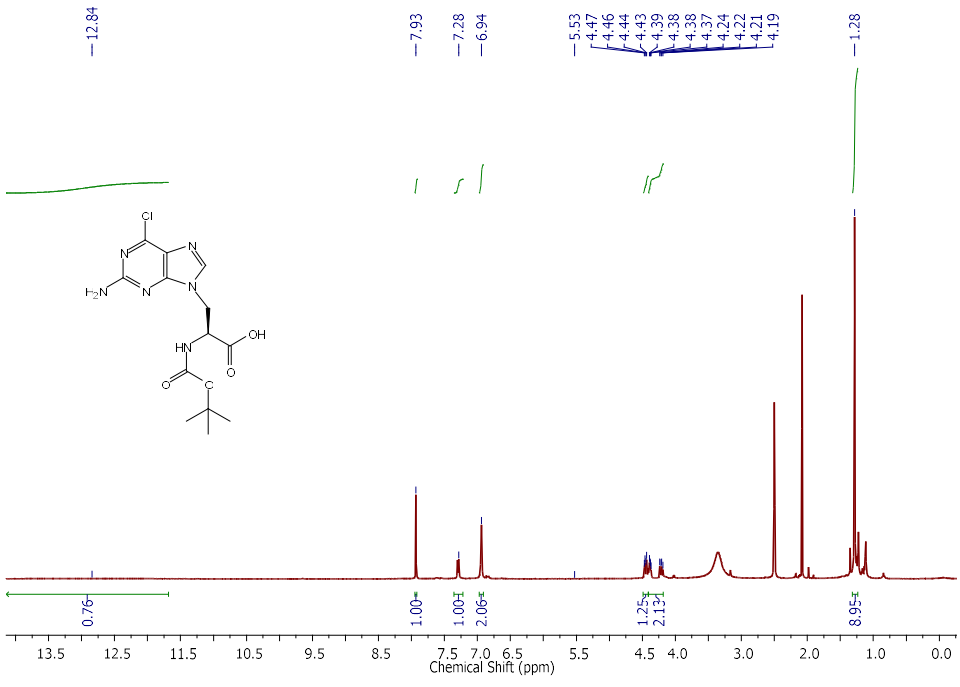


Figure 22. ^{1}H NMR (500 MHz, DMSO-d_6) spectrum of compound **15**.

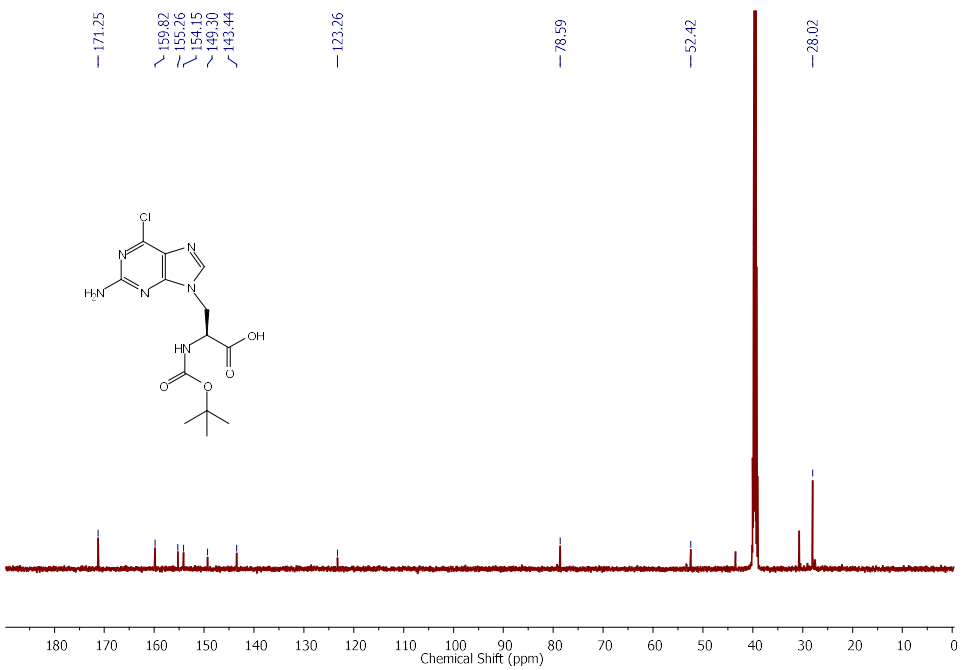


Figure 23. ^{13}C NMR (125 MHz, DMSO-d_6) spectrum of compound **15.**

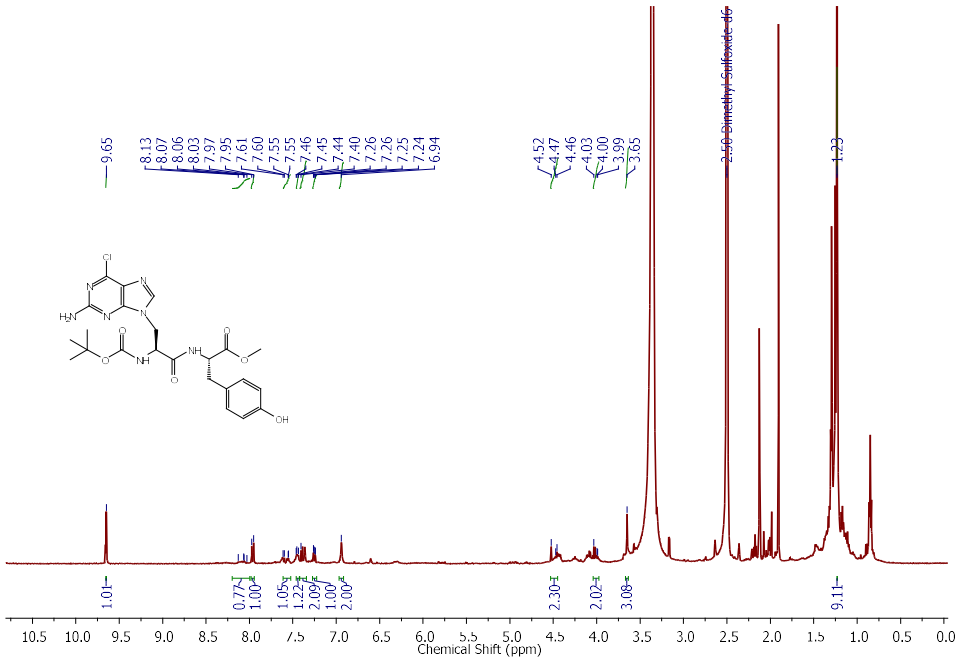


Figure 24. ^{1}H NMR (500 MHz, DMSO-d_6) spectrum of compound **16**.

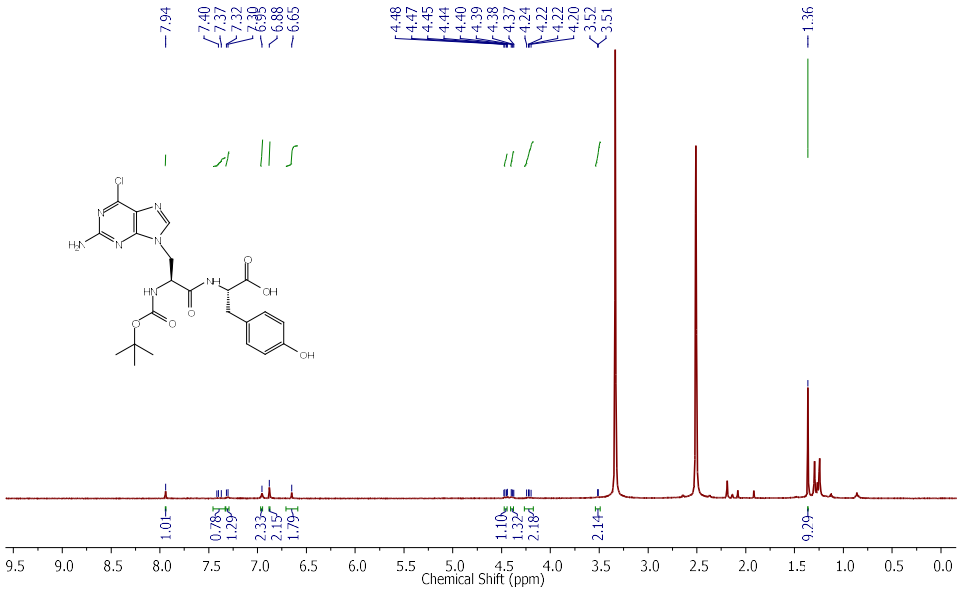


Figure 25. ^{1}H NMR (500 MHz, DMSO-d_6) spectrum of compound **17**.

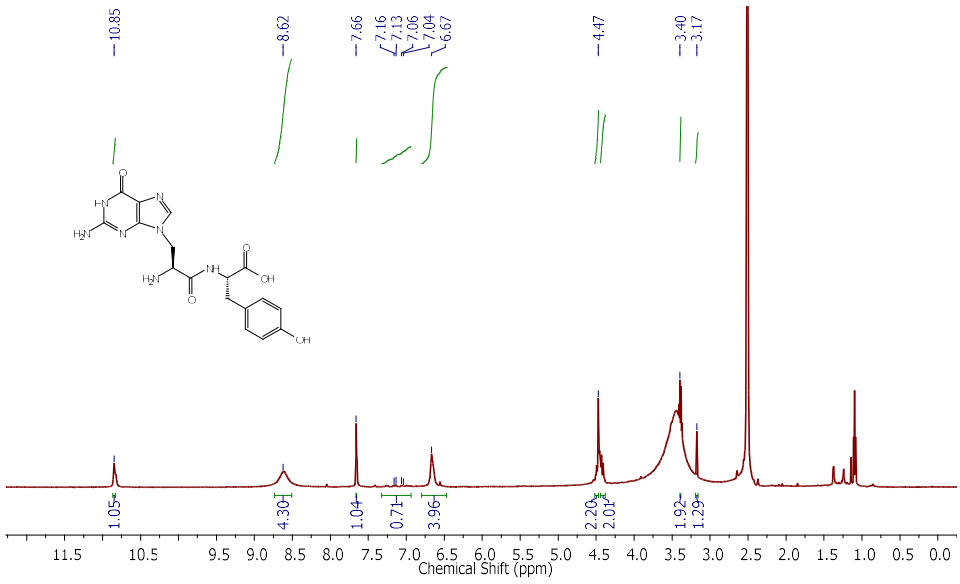


Figure 26. ^{1}H NMR (500 MHz, DMSO-d_6) spectrum of compound **18**.

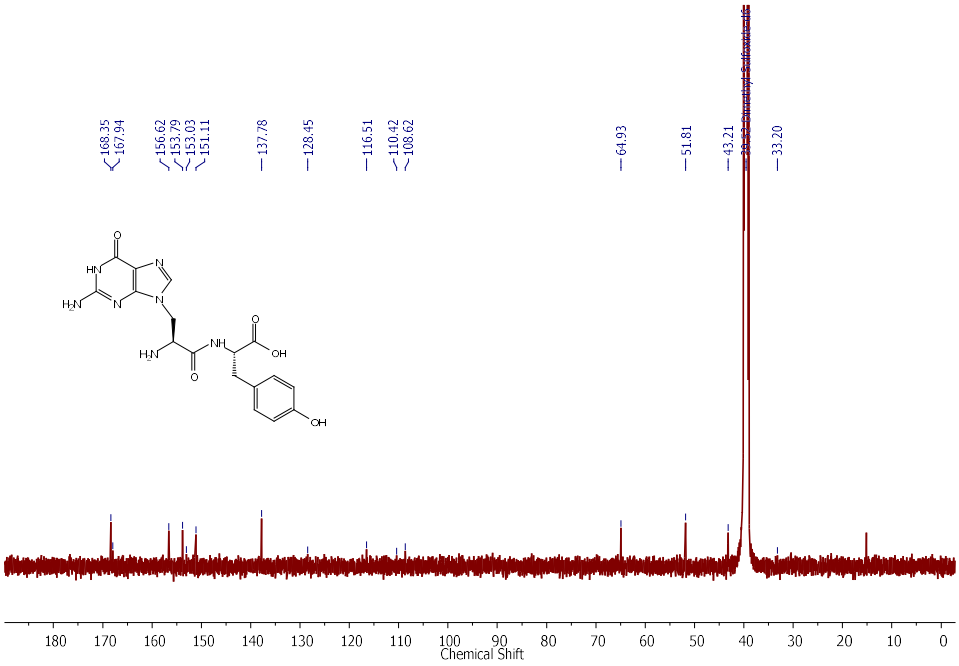


Figure 27. ^{13}C NMR (125 MHz, DMSO-d_6) spectrum of compound **18.**

Mass spectral data

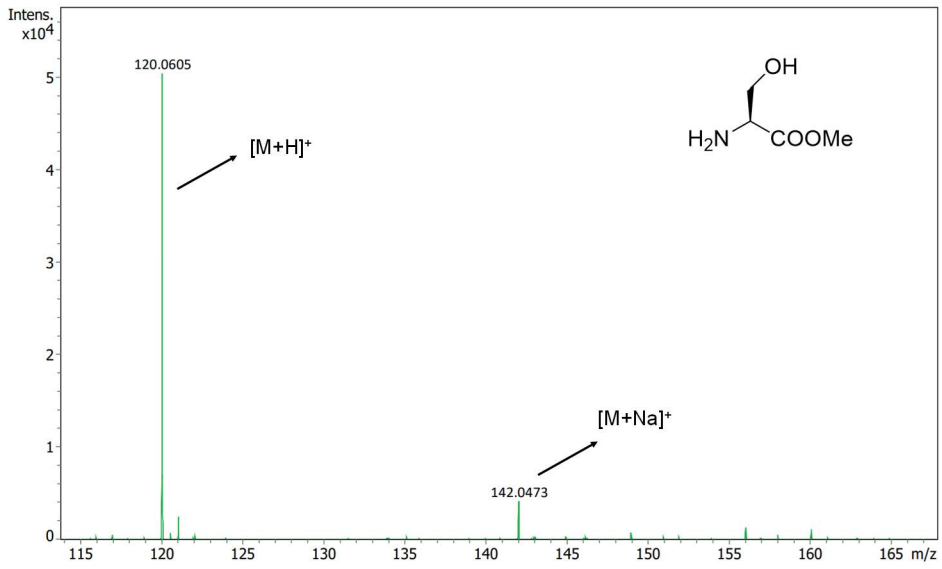


Figure 28. Mass spectrum of compound **2**; (ESI-MS) m/z: $[M+Na]^+$ calculated for $C_4H_9NO_3Na$ 142.0475; found 142.0473.

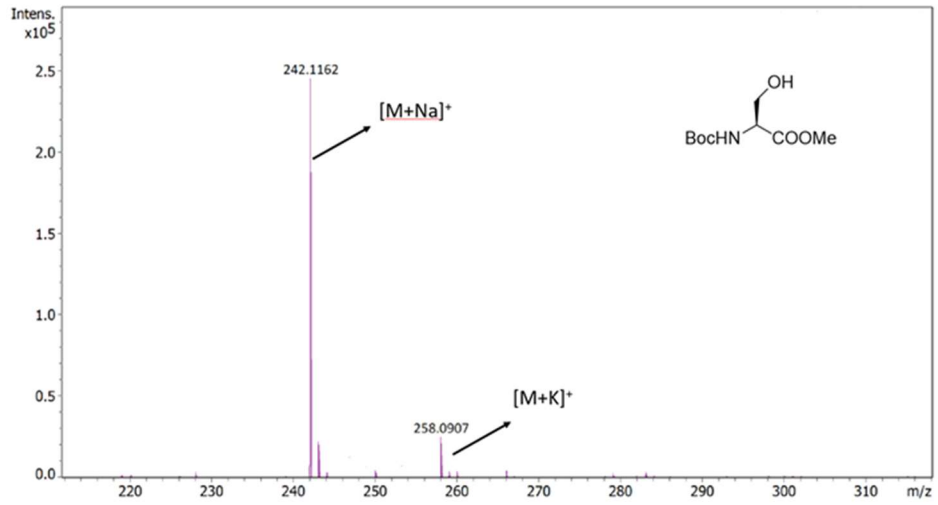


Figure 29. Mass spectrum of compound **3**; (ESI-MS) m/z: $[M+Na]^+$ calcd for $C_9H_{17}NO_5Na$ 242.0999; found 242.1149

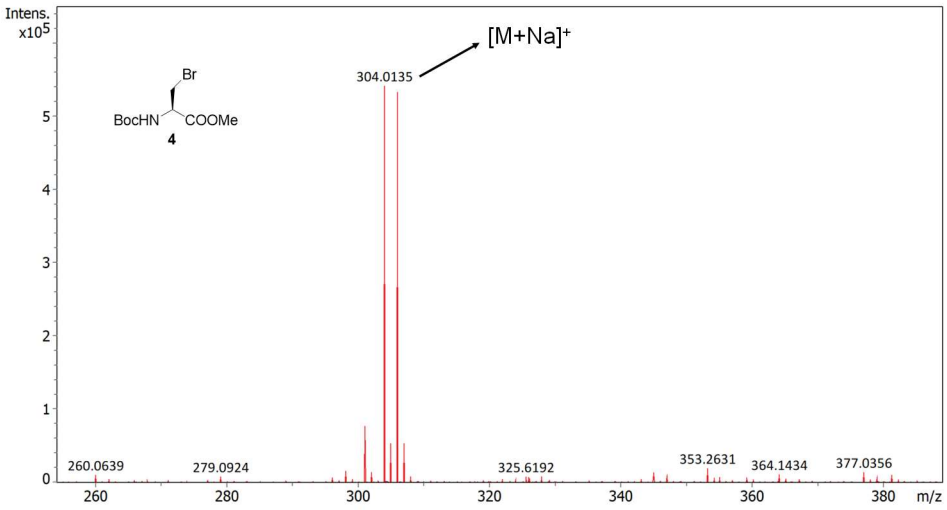


Figure 30. Mass spectrum of compound **4;** ESI-MS, m/z: $[M+Na]^+$ calcd. for $C_9H_{16}BrNO_4Na$ 304.0155; found 304.0135.

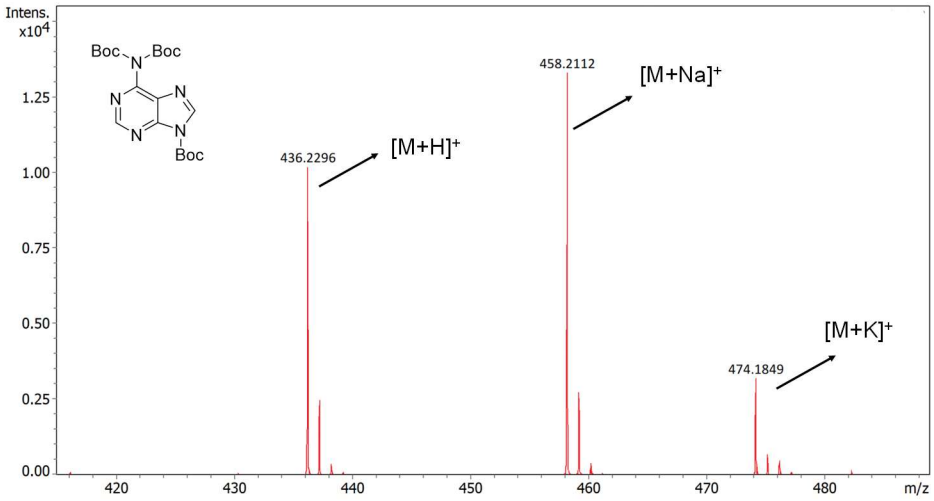


Figure 31. Mass spectrum of compound **6**. ESI-MS m/z: $[M+Na]^+$ calcd for $C_{20}H_{29}O_6N_5Na$ 458.2010; found 458.2112

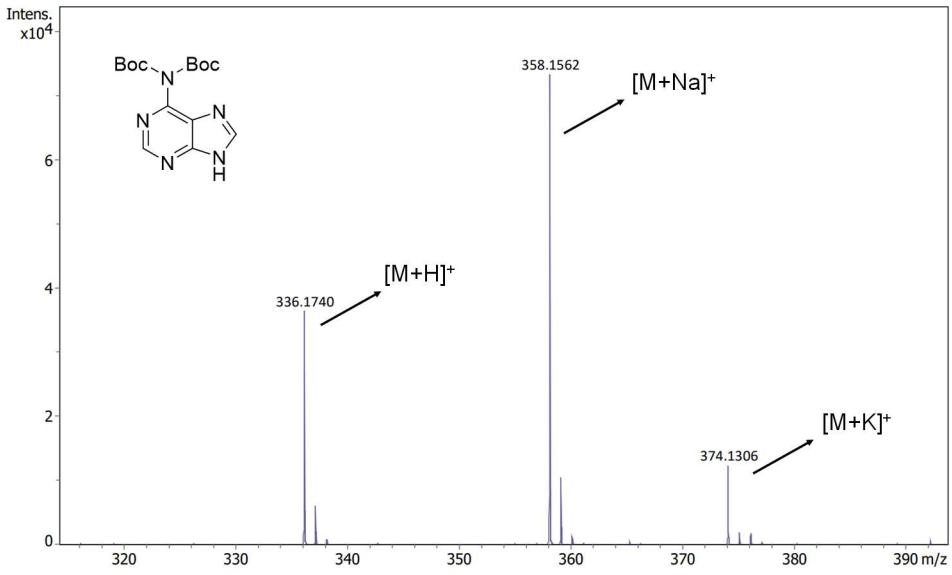


Figure 32. Mass spectrum of compound **7**; (ESI-MS) m/z: $[M+Na]^+$ calcd for $C_{15}H_{21}N_5O_4Na$ 358.1486; found 358.1562

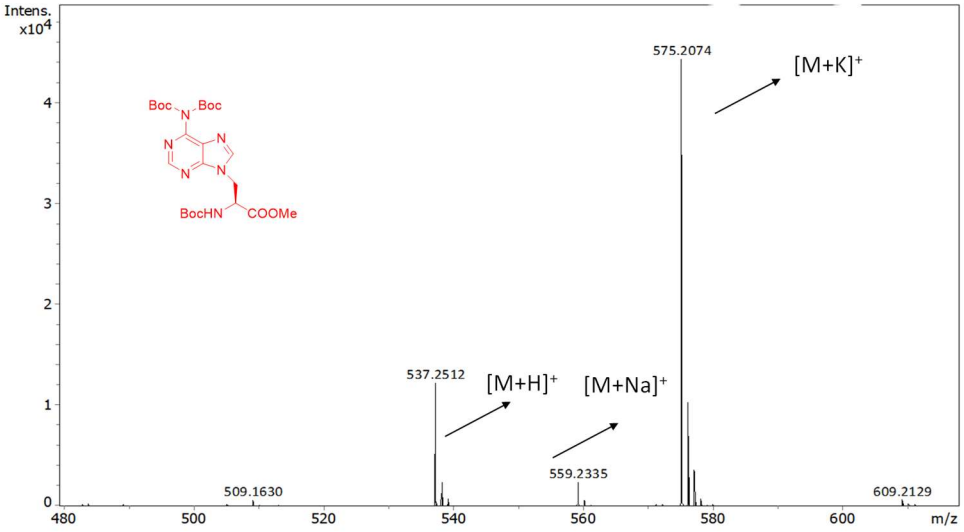


Figure 33. Mass spectrum of compound **8** ; (ESI-MS) m/z: $[M+K]^+$ calcd for $C_{24}H_{36}N_6O_8K$ 575.2226; found 575.2074.

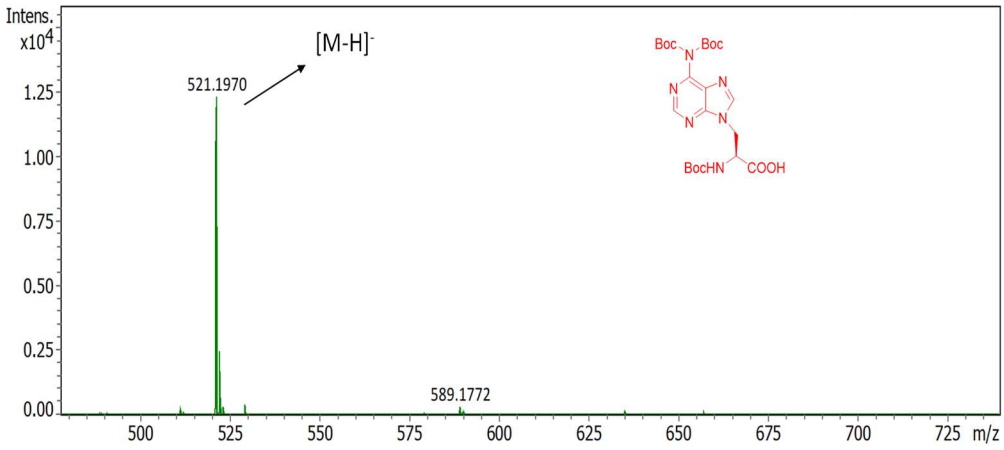


Figure 34. Mass spectrum of compound **9** ; (ESI-MS) m/z: $[M-H]^-$ calcd for $C_{23}H_{35}N_6O_8$ 521.2354; found 521.1962.

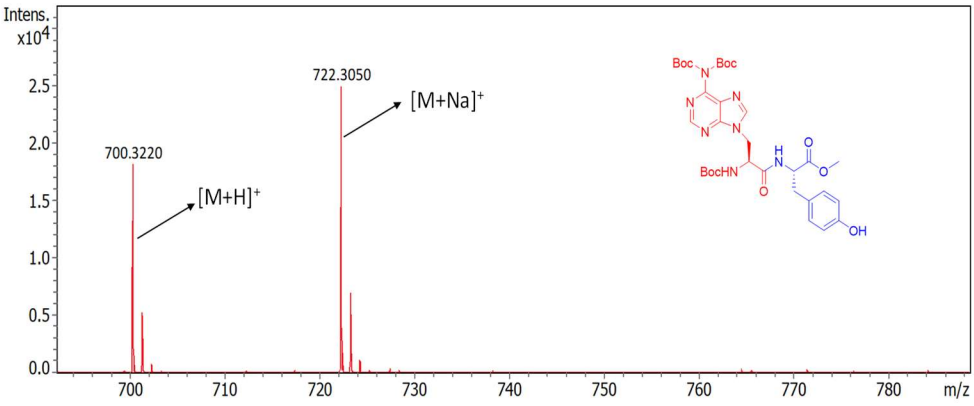


Figure 35. Mass spectrum of compound **10** ; (ESI-MS) m/z: $[M+Na]^+$ calcd for $C_{33}H_{46}N_7O_{10}Na$ 722.3120; found 722.3050.

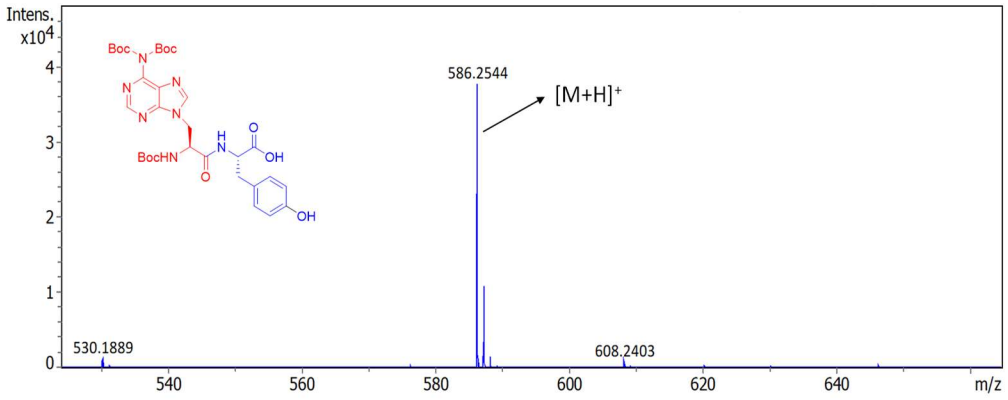


Figure 36. Mass spectrum of compound **11** ; (ESI-MS) m/z: $[M+H]^+$ calcd for $C_{32}H_{44}N_7O_{10}$ 586.2455 ; found 586.2544.

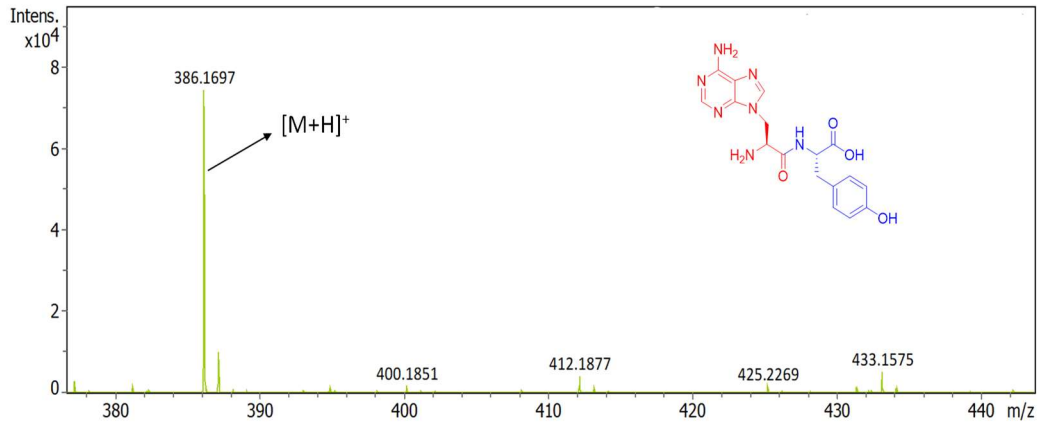


Figure 37. Mass spectrum of compound **12 ;** (ESI-MS) m/z: $[M+H]^+$ calcd for $C_{17}H_{20}N_7O_4$ 386.1571 ; found 386.1697.

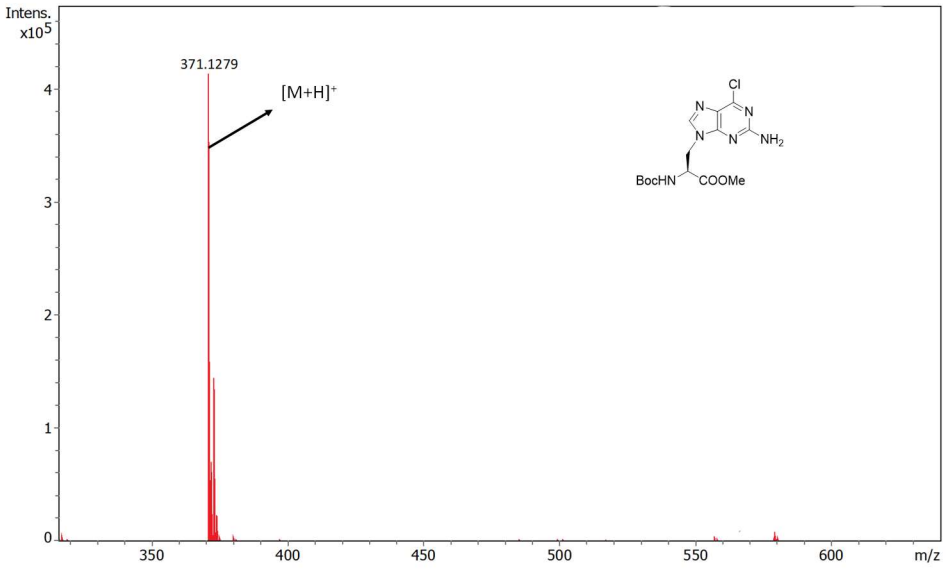


Figure 38. Mass spectrum of compound **14.** (ESI-MS) m/z: $[M+H]^+$ calcd for $C_{14}H_{20}N_6O_4Cl$ calcd. 371.1229, found 371.1279.

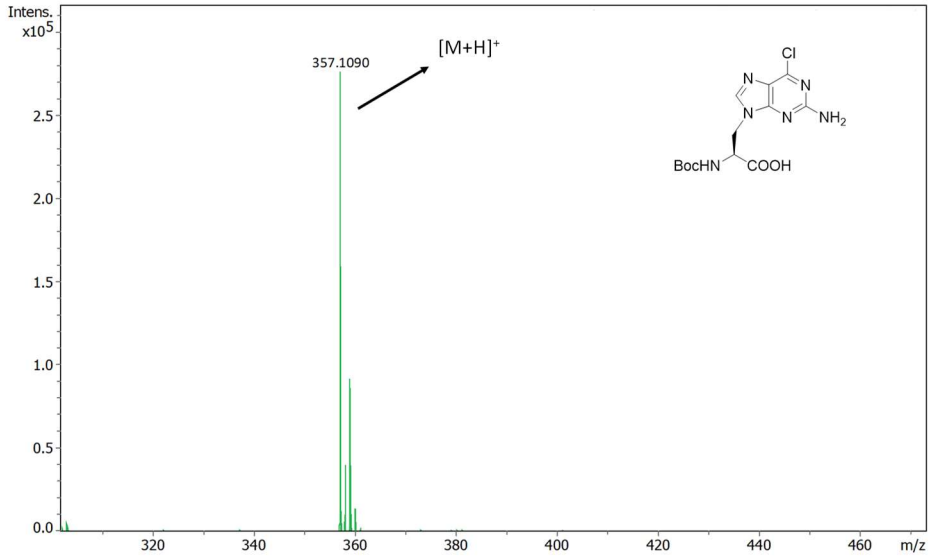


Figure 39. Mass spectrum of compound **15.** (ESI-MS) m/z: $[M+H]^+$ calcd for $C_{13}H_{18}N_6O_4Cl$ calcd. 357.1073, found 357.1090.

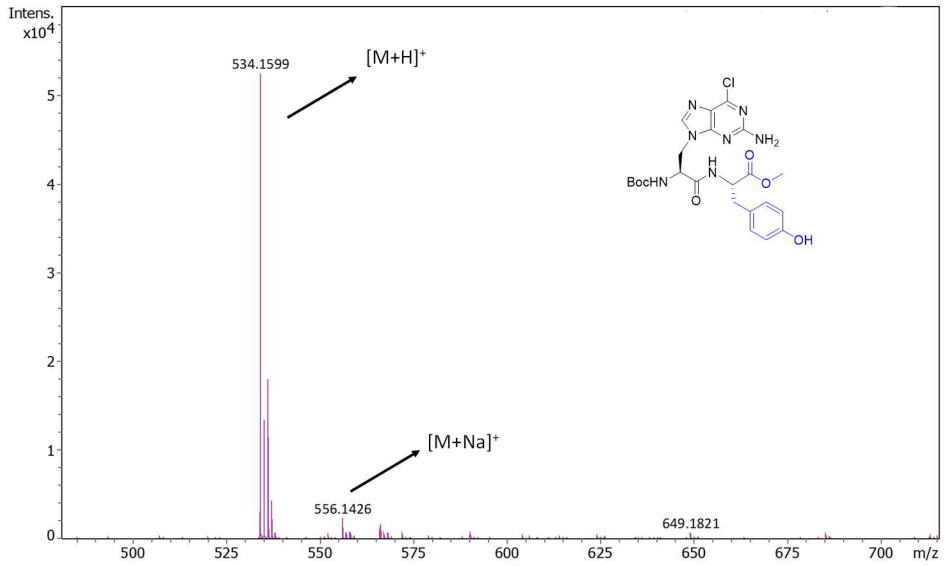


Figure 40. Mass spectrum of compound **15.** (ESI-MS) m/z: $[M+H]^+$ calcd for $C_{23}H_{29}N_7O_6Cl$ calcd. 534.1862 found 534.1599.

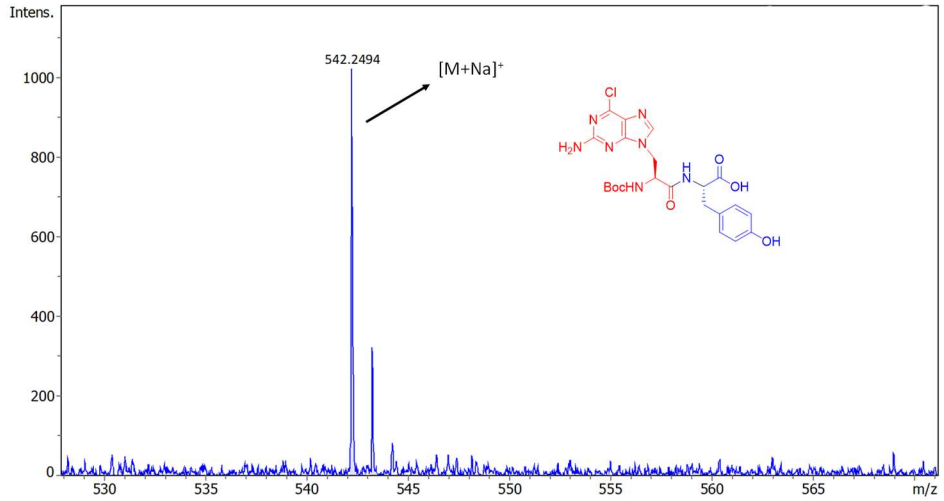


Figure 41. Mass spectrum of compound **16.** (ESI-MS) m/z: $[M+Na]^+$ calcd for $C_{22}H_{26}N_7O_6ClNa$ calcd. 542.1512 found 542.2495.

Chapter 6

6.1 Conclusion

We have successfully synthesised two purine nucleobase (adenine and guanine) containing amino acid monomers in an efficient solution-phase pathway via the crucial intermediate (S)-2-(N-Boc-amino)-3-bromopropinoic acid methyl ester by nucleophilic substitution of the bromo group with the respective nucleobase derivative. Apart from the 2-amino-6-chloropurine derivative, the exocyclic amino group of adenine was protected to avoid unnecessary side reactions and improve solubility for better chemical reactions. The conventional mixed carbonic anhydride method for peptide synthesis was introduced to couple the nucleo amino acids with tyrosine methyl ester, and the resulting dipeptides were purified using column chromatography and characterised by mass and NMR spectroscopy method. Finally, the *C*-termini methyl ester and *N*-Boc of nucleodipeptides were cleaved by ester hydrolysis and TFA, respectively to obtain free nucleodipeptides. We have also observed an efficient pathway to synthesise short nucleopeptides of high purity using standard solution-phase peptide synthesis methods.

6.2 Future Plans

The current work has synthesised the purine nucleobase functionalised nucleodipeptides (**12** and **18**). As the present final compounds consist of both bioactive nucleobase and peptide units, such nucleopeptide will show a wide range of biological activities, more relevant to their ability to interact with different nucleic acids selectively. Therefore, these novel nucleopeptides will be further used to investigate their binding with RNA and DNA targets. CD techniques will also analyse the interaction of these nucleopeptides with different nucleic acids. Moreover, all these future findings will encourage the researchers to design such novel nucleobase containing chimeric compounds for future biomedicine.

6.3 References

[1] Kuan S. L., Bergamini F. R. G., Weil T. (2018), Functional protein nanostructures: a chemical toolbox, Chem. Soc. Rev, 47, 9069 (DOI: 10.1039/C8CS00590G)

[2] Roviello G. N., (2018), Novel insights into nucleoamino acids: biomolecular recognition and aggregation studies of a thymine-conjugated L-phenyl alanine, Amino Acids, 50, 933-941 (DOI: 10.1007/s00726-018-2562-2)

[3] Roviello G. N., Roviello V., Autiero I., Saviano M. (2016), Solid-phase synthesis of TyrT, thymine–tyrosine conjugate with poly(A) RNA-binding ability, RSC Adv, 6, 78515 (DOI: 10.1039/C6RA00294C)

[4] Miyauchi K., Kimura S., Suzuki T. (2012), A cyclic form of N6-threonyl carbamoyl adenosine as a widely distributed tRNA hypermodification, Nature Chemical Biology, 9(2), 105–111 (DOI: 10.1038/nchembio.1137)

[5] Li X., Kuang Yi., Lin H. C., Gao Y., Shi J., Xu B. (2011) Supramolecular Nanofibers and Hydrogels of Nucleopeptides, Angew. Chem. IntEd., 50, 9365-9369 (DOI: 10.1016/j.matt.2019.09.015)

[6] Davis J. T. (2004), G-Quartets 40 Years Later: From 5′-GMP to Molecular Biology and Supramolecular Chemistry, Angew. Chem. IntEd., 43, 668-698 (DOI: 10.1002/anie.200300589)

[7] Meng L., Liu K., Mo S., Mao Y., Yi T. (2013), From G-quartets to G-ribbon gel by concentration and sonication control, Org. Biomol. Chem., 11, 1525 (DOI: 10.1039/C3OB27204D)

[8] Wu G., Kwan I. C. M. (2009), Helical Structure of Disodium 5′-Guanosine Monophosphate Self-Assembly in Neutral Solution, J. Am. Chem. Soc., 131, 9, 3180–3182 (DOI: 10.1021/ja809258y)

[9] Kaucher M. S., Harrell W. A., Davis J. T. A (2006), Unimolecular G-Quadruplex that Functions as a Synthetic Transmembrane Na^+ Transporter, J. Am. Chem. Soc. 128, 1, 38–39 (DOI: 10.1021/ja056888e)

[10] Liddar H., Li J., Neogia A. (2008), Self-assembled deoxyguanosine based molecular electronic device on GaN substrates, Appl. Phys. Lett., 92, 013309 (DOI:10.1063/1.2828405)

[11] Ruyechan W. T., Olson J. W. (1992), Surface lysine and tyrosine residues are required for interaction of the major herpes simplex virus type 1 DNA-binding protein with single-stranded DNA, J. Virol., 66, 6273 – 6279 (DOI: 10.1128/JVI.66.11.6273-6279.1992)

[12] Katiyar S. P., Bakkiyaraj D., Pandian S. K. (2011), Role of aromatic stack pairing at the catalytic site of gelonin protein. Biochem. Biophys. Res. Commun., 410, 75-80 (DOI: 10.1007/978-1-4939-7756-7_8)

[13] Charlton T. S., Ingelse B. A., Black D. S., Craig D. C., Mason K. E., Duncan M. W. (1999), A covalent thymine–tyrosine adducts involved in DNA–protein crosslinks, Free Radical Biol. Med., *27*, 254 – 261 (DOI: 10.1016/s0891-5849(99)00030-1)

[14] Dey S., Garner P., (2000), Synthesis of tert-Butoxycarbonyl (Boc)-Protected Purines J. Org. Chem., 65, 7697 (DOI: 10.1021/jo000983i)