Synthesis of Unnatural Amino Acids and Their Spectroscopic Applications

A
Dissertation Submitted to the
Indian Institute of Technology Guwahati
As Partial Fulfillment for the Award of the Degree of

Doctor of Philosophy in Chemistry

By
Suranjan De
(Roll No. 126122035)

Under The Supervision of

Prof. Subhendu Sekhar Bag

Department of Chemistry
Indian Institute of Technology Guwahati
Guwahati 781039, India
25th May 2018

Dedicated to

Late Damoder De (Uncle),

Late Chanchala De & Late Gauri Pal (Grandmother),

My Parents,

&

Dr. Mirjana Živanov

INDIAN INSTITUTE OF TECHNOLOGY, GUWAHATI
Department of Chemistry

Declaration

*I do hereby declare that the research work embodied in this thesis entitled "**Synthesis of Unnatural Amino Acids and Their Spectroscopic Applications**" has been carried out by me under the supervision of **Prof. Subhendu Sekhar Bag** in the Department of Chemistry, Indian Institute of Technology Guwahati, India.*

In keeping with the general practice of reporting scientific observations, due acknowledgements have been made wherever the work described is based on the findings of other investigators.

IIT Guwahati *Suranjan De*
August, 2017.

Dr. Subhendu Sekhar Bag, Ph.D.
Professor,
Department of Chemistry
Indian Institute of Technology
Guwahati -781039
Assam, INDIA

Ph: +91-361-258-2324(O)
Ph: +91-361-258-4324 (R)
Fax: +91-361-258-2349

E-mail: ssbag75@iitg.ernet.in
ssbag75@yahoo.co.in

CERTIFICATE

This is to certify that the research work presented in this thesis entitled "***Synthesis of Unnatural Amino Acids and Their Spectroscopic Applications***" is an authentic record of the results obtained from the research work carried out by Mr. Suranjan De under my supervision in the Department of Chemistry, Indian Institute of technology Guwahati, India. This work is original and has not been submitted elsewhere for a degree or award.

IIT Guwahati
August, 2017

Prof. Subhendu Sekhar Bag
(Thesis Supervisor)

ACKNOWLODGEMENT

I take this opportunity to express a deep sense of sincere gratitude to my supervisor **Prof. Subhendu Sekhar Bag** for his guidance, suggestions during the entire course of my research work. I would like to thank my Doctoral Committee members, **Prof. Anil Kumar Saikia** (chairman), **Dr. Chandan Das** (member) **& Dr. Chandan Mukherjee** (member) for their intellectual input, encouragement, valuable suggestions and comments during the course of my research work. I would also like to thank **Prof. Arun Chattopadhyay** and **Prof. G. Krishnamoorthy** for their valuable suggestion and help.

I would like to express my heartiest regards to my lab seniors Dr. Rajen Kundu, Dr. Mohan Kasula, Dr. Manoj K.Pradhan, Dr. Subhasish Jana, and lab mates Suman, Hiranya, Afsana, Ananya, Sayantan, Krishanu for their cooperation, support and pleasant company during my research work. I would also like to acknowledge summer trainee Kamini Jakhar for her help and pleasant company in the laboratory during her summer project work. Sincere thanks goes to my other lab mates Shilaj Roy, Soumen Saha, Nilotpal Singha, Minati Das, Ila Verma for their cooperation. Thanks to all of my friends, juniors and seniors whom I meet during my last five years for their help. I want to express my sincere thanks to Mr. Chandan Borgohain, Mr. Deep Manoranjan, and Mr. Asim Malakar for their great effort in CIF instrumental section and smooth data collection system. I am grateful to our technical and official staffs for their help and support.

Finally, I owe success to my parents (Mr. Dipankar De and Mrs. Manika De), my brother (Mr. Suhhas Ranjan De) and Mrs. Basana Rani Mondal who have been a constant source of inspiration, support and love to carry out my research work.

I would like to acknowledge the Department of Chemistry, IIT Guwahati for giving me the opportunity and fellowship to carry out my research work.

Suranjan De

Present Address

C/O: Prof. Subhendu Sekhar Bag
Department of Chemistry
Indian Institute of Technology Guwahati
Guwahati – 781039, Assam, India
Phone: +91 361 2582324

Permanent Address

Village – Karakberia
P. O. – K-Sadhuadal
Dist. – Bankura, P.S-Joypur
Pin – 722154
West Bengal, India
E mail: suranjan@iitg.ernet.in
de.suranjan@gmail.com

Area of Interest

Design & Synthesis of Side Chain Modified Unnatural Amino Acids for Peptidomimetic; Protein Conformational and Activity Study Using High Field NMR; Protein Nucleic Acids (PNAs) and Their Therapeutic Application

Education:

2012-2018	**Ph. D. [2018]** IIT Guwahati. ***Title****: "Synthesis of Unnatural Amino Acids and Their Spectroscopic Applications"*
2010-2012	**Research Trainee [under DBT-BIF],** West Bengal University of Technology, Salt Lake-1, India. ***Title****: "MALDI-Mass Spectrometric Fragmentation of Short Peptide Molecules".*
2008-2010	**Master of Science (*Organic Chemistry*)** West Bengal State University**.** West Bengal, India. ***Title- "****Extraction of Piperine from Black Pepper and Its Characterisation Using NMR Spectroscopy".*
2005-2008	**Bachelor of Science *(Chemistry Hons.),*** *University of Burdwan, India.*

Honors/Awards

- **RSC ChemChomm Best Poster Prize and Book Boucher in FICS 2017.**
- **Qualified GATE–2012** (Graduate Aptitude Test in Engineering) 99.21 percentile; organized by Ministry of Human Resource Development (MHRD), Government of India. All India Rank-119.
- **Merit-Cum-Means Scholarship from State Government of India** (2008-2010).

List of Publications

1. ●Isothiocyanyl Alanine as a Synthetic Intermediate for the Synthesis of Thioureayl Alanines and Subsequent Aminotetrazolyl Alanines● Bag, Subhendu Sekhar*; **De, Suranjan, *J. Org.Chem*, 2017,** *82*, 12276**.**

2. ●Multipurpose Isothiocyanyl Alanine/Lysine: Use as Solvatochromic IR Probes and in Site Specific Labeling/Ligation of Short Peptides● Bag, Subhendu Sekhar*; **De, Suranjan *Bioorganic Med. Chem. Lett*. 2018,** *18*, 1404**.**

3. ●Triazolo-β-Aza-ε-Amino Acid and Its Aromatic Analogue as Novel Scaffolds for β-turn Peptidomimetics● Bag, Subhendu Sekhar*; Jana, Subhashis; Yashmeen, Afsana; **De, Suranjan**● *Chemical Communication*, **2015**, *51*, 5242.

4. ●Pyrenylthioureayl Alanine as a Selective Switch-on Fluorescent Sensor for Hg^{2+} and Cu^{2+} Ions● Bag, Subhendu Sekhar*; **De, Suranjan (Under Review)**

List of Conferences/Symposiums/Workshop

1. International Conference on Sophisticated Instruments in Modern Research (**ICSIMR 2017**), IIT Guwahati

2. XIIth J-NOST Conference for Research Scholars (**JNOST-2016**), CSIR-CDRI, Lucknow.

3. **FICS-2016**, Department of Chemistry, IIT Guwahati

4. National Conference on Recent Advance in Cancer Biology and Therapeutice-2014, Department of Biotechnology, IIT Guwahati

5. UGC-Sponsored National Level Seminar in Frontier Area of Chemistry-A Modern Perspective-2010. Department of Chemistry, Ramakrishna Mission Vidyamandira, Belur, India.

6. National Work shop on Basics of Statistics & Databases in Bioinformatics-2011, West Bengal University of Technology, Salt Lake, India

Abstract

Amino acids have important biological roles in the living system such as protein synthesis, protein stability, and function. They play a very important role for the structural requirements for the bioactivity of several peptides and proteins. The physical and chemical property of a protein depends on the nature of amino acids. Proteins play a vital role in all living organisms to maintain the cell structures, properties, and functions. In nature, only nine canonical amino acids with a limited number of functional groups are available for protein modification. So, the lack of significant side chain modified amino acids with novel functionality has prompted the design of new side chain modified unnatural amino acids (UNAAs) for research in protein chemistry. To have more complex job done by proteins, their modification with designer unnatural amino acids is essential. UNAAs are mostly used as enzyme inhibitors in cells or more in general as therapeutic agents. A major limitation of current drug development strategies lies on the incapability of traditional pharmaceuticals to target the extended protein interfaces. Therefore, modified amino acids are good choice to overcome this situation and can open wide scope of applications. Also, UNAAs with modified reactive side chain play a crucial role in site-specific labeling/ligation of peptides under physiological condition which can be useful for imaging of protein, monitor protein-protein interaction, protein function and dynamics. Unnatural amino acid (UNAAs) incorporation has become the common practice in the field of protein engineering. UNAAs can be introduced into recombinant proteins via residue-specific or site-specific incorporation methods for its application in therapeutic domain. Therefore, synthetically modified UNAAs (with or without fluorophoric functional groups) have been a challenging field of research for the past decade.

Therefore, with the advancement of knowledge and strategies of chemistry and biology, side chain modified UNAAs would continue to be the lead in assisting researchers to gain insights into fundamental questions concerning life's essential biological functions involving protein interactions, recognition, and synthesis. Therefore, the thesis entitled **"Synthesis of Unnatural Amino Acids and Their Spectroscopic Applications"** is based on research work towards: the (a) synthesis of aromatic triazolyl, sulfur containing acyclic unnatural amino acids; (b) synthesis of new class of thioureayl and tetrazolyl amino acids; (c) application of isothiocyanyl amino acids as solvatochromic IR responsive probes and site specific labeling/ligation of short peptides; (d) application of thioureayl amino acids in metal ion sensing. Thus, a variety of novel unnatural side chain modified amino acids, peptides have been synthesised and their conformation and physicochemical/photophysical properties have been investigated in detail.

The thesis contains a total of **5 Chapters** including one Introduction Chapter (**Chapter-1**). Each chapter contains their individual introduction, experimental and reference sections. In short, **Chapter-1** is a critical review of applications of UNAAs in genetic incorporation, site specific protein modification and fluorophoric unnatural amino acids in sensing. **Chapter-2** elaborates the synthesis of triazolyl, thiocyanyl, and isothiocyanyl unnatural amino acids *via* "click'' chemistry, nucleophilic substitution reaction and *via* decomposition of dithiocarbamic acids salts respectively. Our design concept is focused on possible site specific incorporation of triazolyl amino acids, isothiocyanyl and thiocyanyl amino acids as a new class of side chain modified sulfur-containing amino acids. **Chapter-3** describes the application of isothiocyanyl alanine (**^{NCS}Ala)** in the synthesis of thioureayl alanines and then to aminotetrazolyl alanine as other new classes of unnatural amino acids. Our design concept exploited the electrophilicity of –NCS functionality. **Chapter-4** deals with isothiocyanyl alanine/lysine as solvatochromic IR responsive probes for possible probing of local structures and electrostatic microenvironment of a short peptide containing **^{NCS}Ala/ ^{NCS}Lys** with the help of distinct infrared absorption property of –NCS functionality. The electrophilicity of **^{NCS}Ala** and **^{NCS}Lys** in a short tri-/hexa-peptides have further been exploited to label covalently with chromophoric/fluorophoric amine and detail conformational study has been carried out in order to monitor conformational changes. **Chapter-5** describes application of fluorescent pyrenylthioureayl alanine amino acid in sensing of Hg^{2+} and Cu^{2+} ions in semi aqueous medium.

CHAPTER-1: *Synthesis and Applications of Side Chain Modified Unnatural Amino Acids: A Review*

This chapter highlights some of the applications of side chain modified unnatural amino acids (UNAAs) in genetic incorporation for the expansion of genetic code, site-selective protein modification as a key role for the labeling of protein, use as IR sensitive spectral probes to investigate the structure, dynamics, localisation, and biomolecular interactions. The fluorescent unnatural amino acids in protein can be utilised for studying conformation/cell imaging/ microenvironment sensitivity and metal ion sensing.

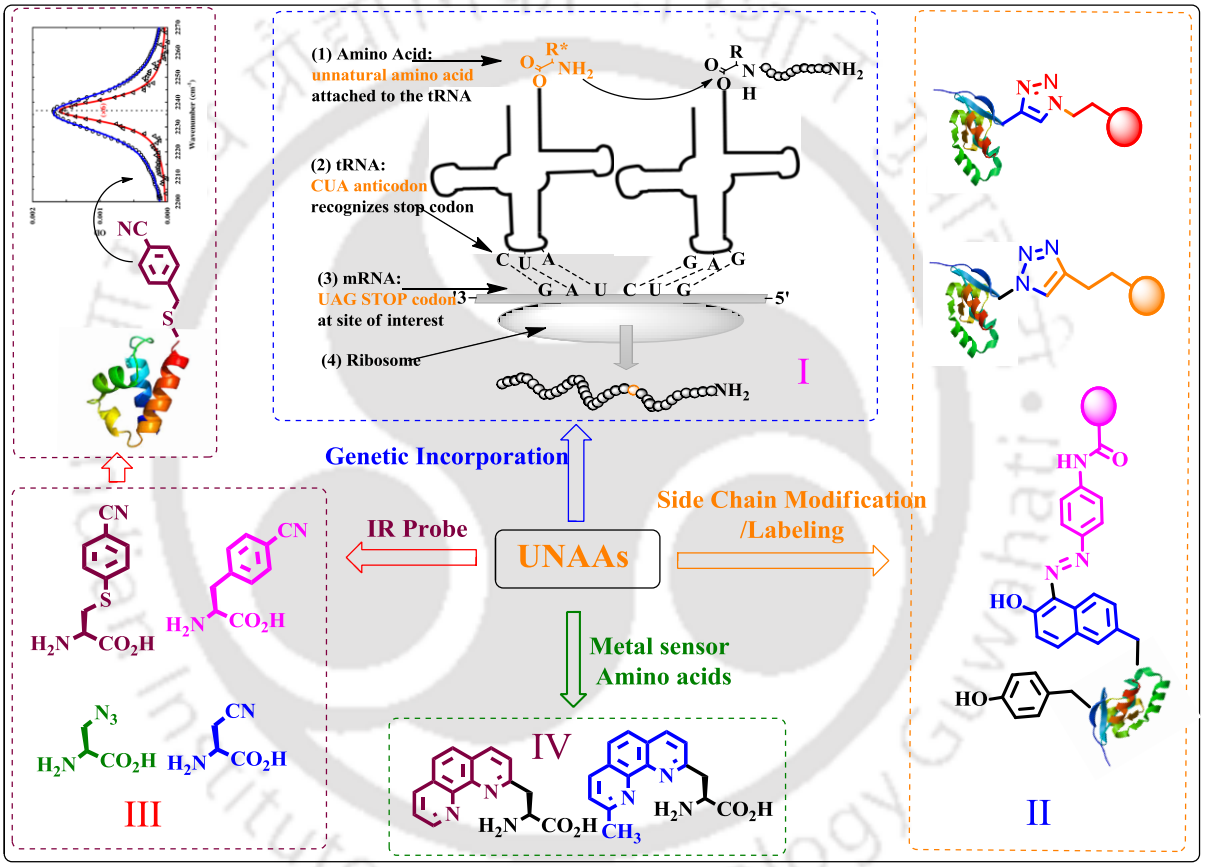


Figure A1: Schematic application of unnatural amino acids.

In nature, only nine canonical amino acids with a limited number of functional groups available for protein modification. So, the lack of significant side chain modified amino acids with novel functionality has prompted the design of new side chain modified UNAAs for research in protein chemistry. Thus, much effort have been devoted to develop site-specific modification method in amino acids, several side chain modified peptides/proteins for investigating folding mechanism, labeling of a target protein, IR probe and studying protein-protein/protein-drug interactions.

CHAPTER-2: *Synthesis of Triazolyl, Thiocyanyl and Isothiocyanyl Unnatural Amino Acids*

As a part of our ongoing research efforts in the synthesis of side chain modified amino acids we introduced the new functionality in peptides with triazolyl and sulfur containing groups (-SCN/-NCS). This chapter describes the design and synthesis of triazolyl, thiocyanyl and isothiocyanyl unnatural amino acids **(Figure A2)**. The side chain cyclisation is observed during the synthesis of isothiocyanyl alanine/serine and the five membered side products were isolated from this reaction. All detail experimental results are presented in this chapter.

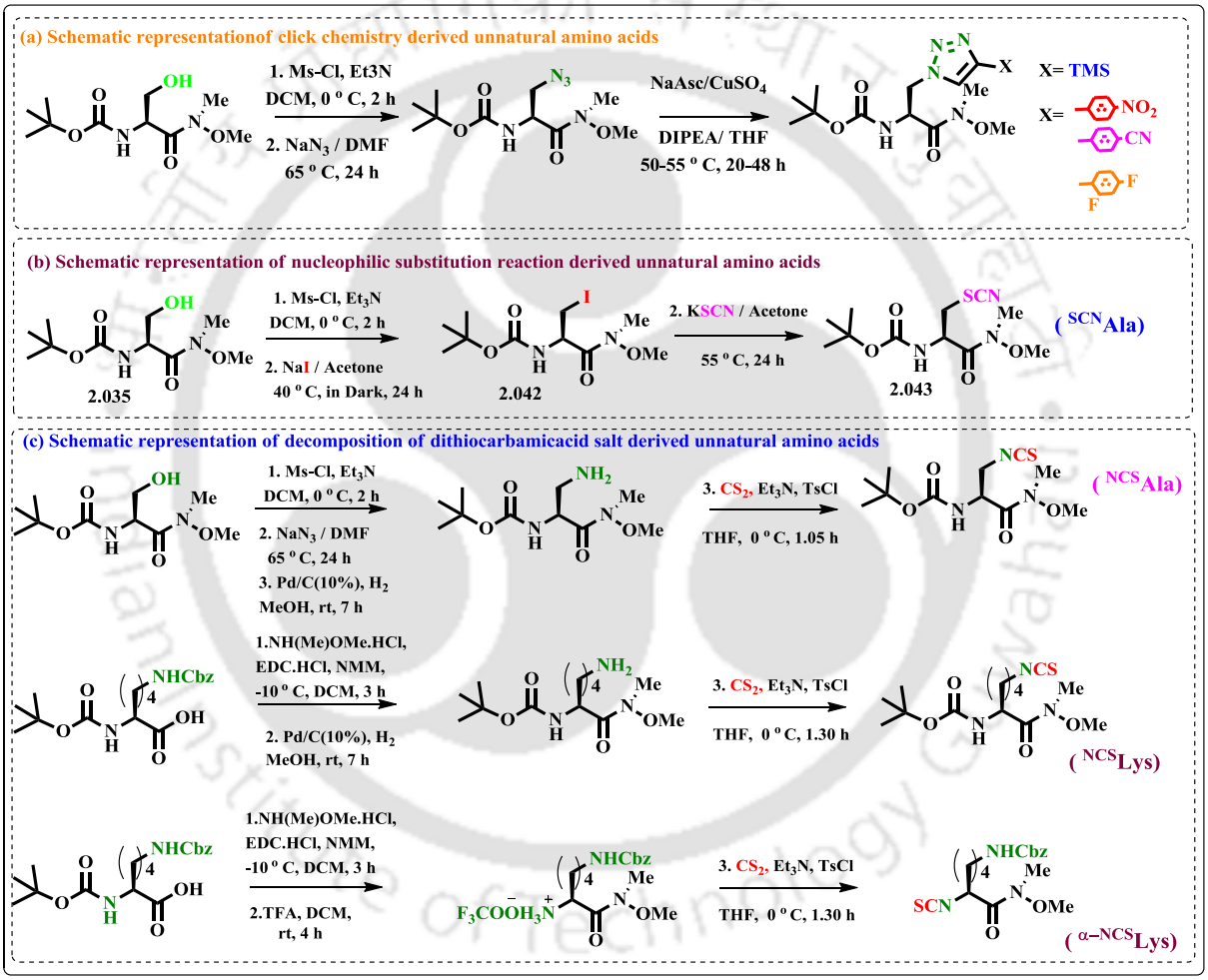


Figure A2: Schematic of synthesis of (a) triazolyl unnatural amino acids (b) thiocyanyl alanine and (c) isothiocyanyl alanine and lysine.

CHAPTER-3: *Application of ^{NCS}Ala in the Synthesis of Thioureayl Alanines and then Aminotetrazolyl Alanines as another Class of New Unnatural Amino Acids*

After successful synthesis of isothiocyanyl alanine (^{NCS}**Ala**) in **Chapter 2**, we exploited electrophilicity of -NCS functional group for the synthesis of another new class of unnatural amino acids. This chapter describes the design and synthesis of thioureayl alanines (^{TU}**Ala)** which is further utilised to achieve aminotetrazolyl alanines (^{TzA}**Ala) (Figure A3)**.Thus, a variety of aliphatic/aromatic substituted thioureayl alanines and aromatic aminotetrazolyl alanines have been successfully synthesised. The photophysical properties of the as synthesised fluorescent unnatural amino acids have been studied. All detail experimental results are presented in this chapter.

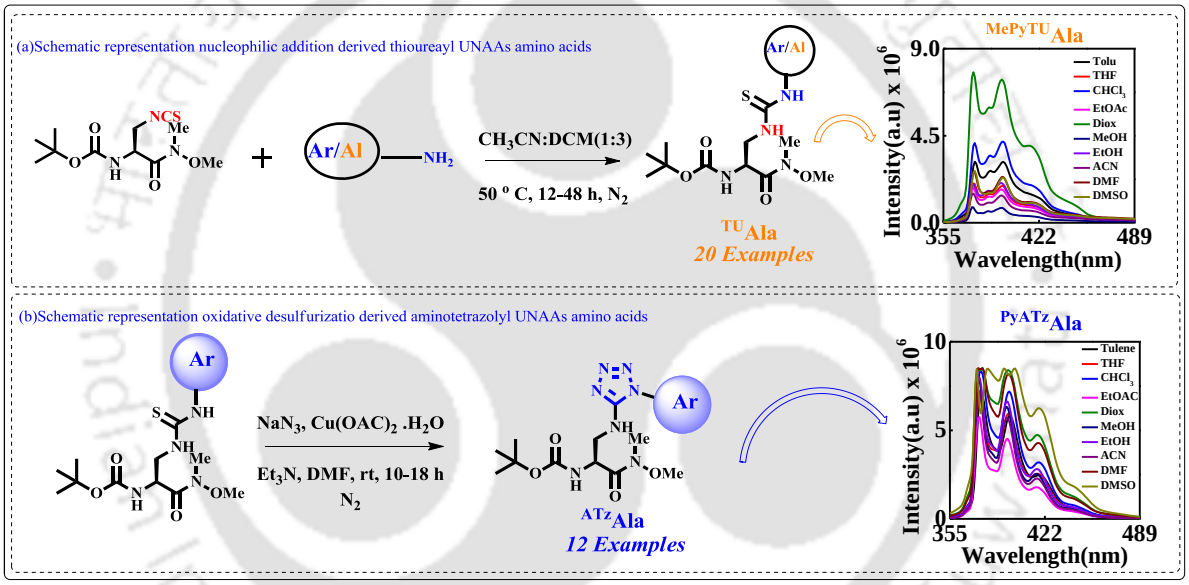


Figure A3: (a) Schematic of formation of thioureayl alanines and (b) aminotetrazolyl alanines UNAAs**.**

CHAPTER-4: *Isothiocyanyl Alanine / Lysine as Solvatochromic IR Responsive Probes and Use in Site Specific Labeling/Ligation of Short Peptides*

This chapter describes the application of isothiocyanyl unnatural amino acids in peptide. After successful demonstration of the synthetic utility of isothiocyanyl alanine (^{NCS}**Ala**), here we have utilised the side chain –NCS (^{NCS}**Ala**/ ^{NCS}**Lys**) as novel chromophoric probe for determining the electrostatic microenvironment of a short peptide by exploring the distinct infrared (IR) absorption property of NCS group (**Figure A4)**. We exploited ^{NCS}**Ala**/^{NCS}**Lys** as a solvatochromic IR probe in protein for monitoring protein conformation and folding. The electrophilicity of isothiocyanyl group (-NCS) of ^{NCS}**Ala**/ ^{NCS}**Lys** in a tri-/hexa-peptides has been exploited to label covalently with chromophoric/fluorophoric amines to give the corresponding thioureayl derivatives. All detail experimental results are presented in this chapter.

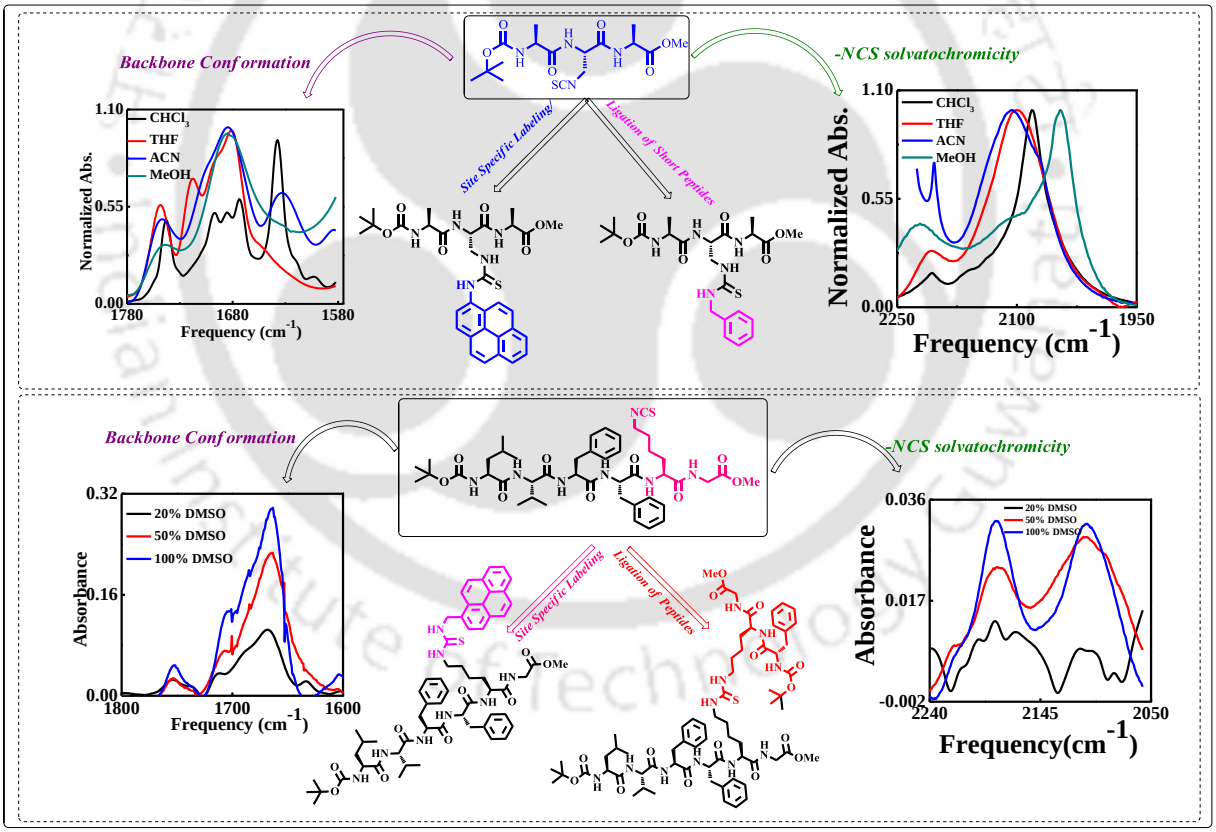


Figure A4: Schematic application of ^{NCS}**Ala**/^{NCS}**Lys** in peptides.

CHAPTER-5: Studies on the Switch-on Fluorescence Sensing for Hg^{2+} and Cu^{2+} Ions by Pyrenylthioureayl Alanine Amino Acid

This chapter describes application of fluorescent pyrenylthioureayl alanine amino acid (its synthesis has been described in **chapter 3)** in sensing of Hg^{2+} and Cu^{2+} ions in semi aqueous medium. The probe is effectively sensing Hg^{2+}ion and Cu^{2+} ion in presence or absence of other interfering metal ions (**Figure A5)**. The study revealed the sensing efficiency of **^{PyTU}Ala** for Hg^{2+} ion is three times more than Cu^{2+} ion. This is evident from the following studies. The quantum yield and life time of the probe **^{PyTU}Ala** has increased much more in presence of Hg^{2+} ions. The binding affinity of Hg^{2+} ion with the probe is $3.93x10^{10}$ (M^{-1}) whereas for Cu^{2+} ion it is comparatively less and which has been further confirmed by the optical and NMR titration. Binding stoichiometry of the probe is 2:1 (Probe: Metal) for both metal ions which has been further verified by fluorescence studies, NMR, ESI-MS and DFT optimisation. All detail experimental results are presented in this chapter.

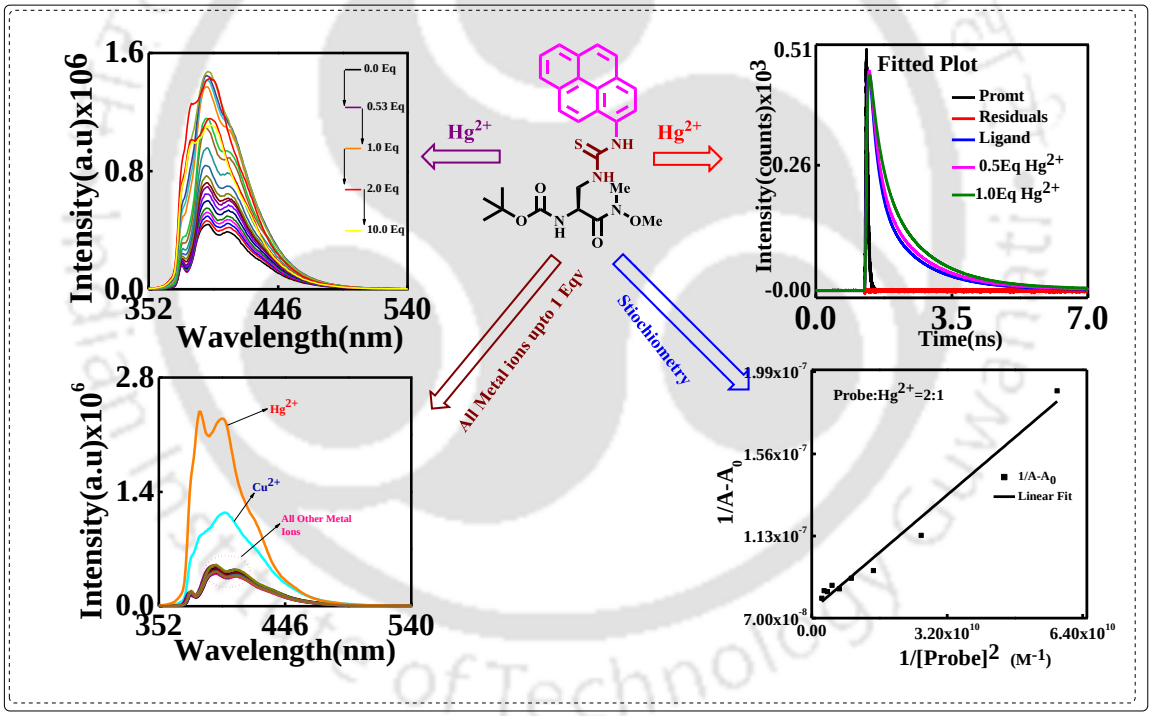


Figure A5: Schematic application on pyrenylthioureayl alanine (**^{PyTu}Ala**) in metal sensing.

List of Abbreviations

Ala (A)	Alanine
AA	Amino acid
Abs	Absorbance
Ac	Acyl
ACN	Acetonitrile
Aha	Azidohomoalanine
Ag	Silver
Ar	Aryl
Arg (R)	Arginine
Asp (N)	Asparagine
B3LYP	Becke, three-parameter, Lee-Yang-Parr
BODIPY	Boron Bipyrromethene
BOC	Di-tert-butyl dicarbonate
Bn	Benzyl
Cbz	Carboxy benzyl
BTNP	Benzothiazole–naphthalene
Ca	Calcium
Cd	Cadmium
CD	Circular dichrosim
CHEF	chelation-enhanced fluorescence
Co	Cobalt
CR	Click reagen
CS_2	Carbon di sulphide
Cu	Copper
$CuSO_4$	Copper sulfate
CT	Charge transfer
CuAAC	Copper-catalysed azide-alkyne cycloaddition
CuI	Copper iodide
Cyhex	Cyclohexane
$CHCl_3$	Chloroform
$CuSO_4$	Copper sulfate
CH_3CN	Acetonitrile
$CDCl_3$	Deuterated chloroform
COSY	Homonuclear correlation spectroscopy
CN	Cyano
Cys (C)	Cysteine
DCM	Dichloromethane
DFT	Density functional theory
Diox/Dx	1,4-Dioxane
DIPEA	N,N-Diisopropyl ethyl amine
DMAP	N, N-Dimethylamino pyridine
4-DMAP	4-(N,N-Dimethylamino)-phthalimide

DMF	Dimethyl formamide
DMSO	Dimethyl sulfoxide
DMSO-d6	Deuterated dimethyl sulfoxide
DNA	Deoxyribonucleic acid
E. coli	Escherichia coli
EDC.HCl	1-(3-Dimethylaminopropyl)-3-ethylcarbodiimide-hydrochloride
EPR	Electron paramagnetic resonance
Eq/Equiv.	Equivalent
Et_3N	Triethylamine
EtOAc	Ethyl acetate
EtOH	Ethanol
F_{CN}	p-Cyanophenylalanine
Fe	Iron
FESEM	Field emission scanning electron microscope
FID	Free Induction decay
Fl	Fluorescence
Fmoc	Fluorenyl methyloxy carbonyl
FRET	Fluorescence resonance energy transfer
FSD	Fourier self-deconvolution
Glu(E)	Glutamic acid
Gln (Q)	Glutamine
Gly (G)	Glycine
GPCRs	G-protein coupled receptors
h	hour
HEPES	4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid
HEDM	High energy density materials
Hg	Mercury
His (H)	Histidine
HIV	Human immunodeficiency virus
HRMS	High resolution mass spectroscopy
HOMO	Highest occupied molecular orbital
HOBT	1-hydroxy-benzotriazole
HSQC	The heteronuclear single quantum coherence or heteronuclear single quantum correlation
ICT	Intramolecular charge transfer
Ile (I)	Isoleucine
K	Kelvin
IR	Infrared spectroscopy
KCN	Potassium cyanide
K_2CO_3	Potassium carbonate
KOH	Potassium hydroxide
KSCN	Potassium thiocyanate
KSI	Ketosteroid isomerase
Leu (L)	Leucine

LiOH	Lithium hydroxide
LUMO	Lowest unoccupied molecular orbital
Lys (k)	Lysine
Max	Maxima
MD	Molecular dynamics
MeOH	Methanol
Met (M)	Methionine
Mg	Magnesium
mM	Mili molar
m.p.	Melting point
MALDI-ToF	Matrix assisted laser desorption ionisation-time of flight
Mn	Manganese
Na	Sodium
Na_2CO_3	Sodium carbonate
$NaHCO_3$	Sodium bicarbonate
NaI	Sodium iodide
NaN_3	Sodium azide
NCS	Isothiocyanate
Ni	Nickel
NMR	Nuclear magnetic resonance
NOESY	Nuclear overhauser effect spectroscopy
nM	Nano molar
nm	Nanometer
ORTEP	Oak ridge thermal ellipsoid plot
Pb	Lead
PET	Photoinduced electron transfer
PEITC	Phenethyl isothiocyanate
Pd	Palladium
$PdCl_2(PPh_3)_2$	Bis(triphenylphosphine)palladium(II) dichloride
Ph	Phenyl
Phe (F)	Phenyl alanine
PNA	Protein nucleic acid
PPh_3	Triphenyl phosphine
p-$PheN_3$	para-azido-L-phenylalanine
Pro (P)	Proline
ppm	Parts per million
PXRD	Powder X-ray diffraction
PyBOP	(Benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate
Py	Pyrene
RNA	Ribonucleic acid
rt	Room temperature
SCN	Thiocyanate
Ser (S)	Serine

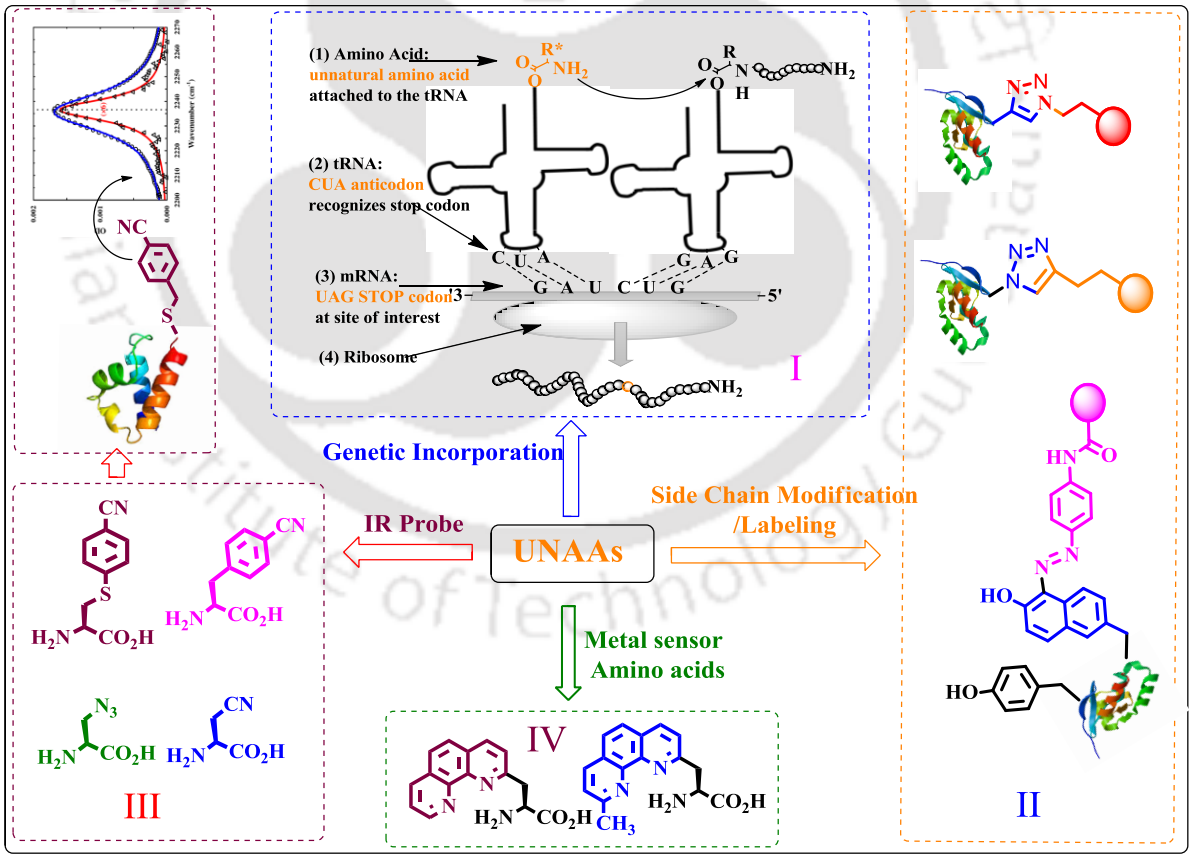
SPPS	Solid phase peptide synthesis
T	Temperature
TBAF	Tetrabutylammonium fluoride
TBAB	Tetrabutylammonium bromide
TBS	*Tert*-butyldimethylsilyl
TDDFT	Time dependent density functional theory
TEM	Transmission electron microscopy
TEA	Triethyl amine
TFA	Trifluoroacetic acid
TFE	Trifluoroethanol
THF	Tetrahydrofuran
TLC	Thin layer chromatography
TMS	Trimethylsilyl
TOCSY	Total correlated spectroscopy
Tolu	Toluene
Trp (W)	Tryptophan
TsCl	Tosyl chloride
TU	Thiourea
Tyr (T)	Tyrosine
Tz	Tetrazole
UNAAs	Unnatural amino acids
UV	Ultra violate
Val (V)	Valine
VT	Variable temperature
Zn	Zinc
μM	Micro molar
Wt/H_2O	Water
Φ	Quantum yield
ε	Molar extinction co-efficient
τ	Decay time
Å	Angstrom (10^{-8}cm)
$\tilde{\nu}$	Wave Number
λ	Wavelength
λ_{max}^{abs}	Absorption maxima
λ_{max}^{fl}	Fluorescence maxima
Δf	Solvent polarity parameter
μ_e	Excited state dipole moment
μ_g	Ground state dipole moment

Synthesized Amino acids	
^{PhTU}**Ala**	Phenylthioureaylalanine
^{MeBTU}**Ala**	*p*-methylphenyl thioureayl alanine
^{EtBTU}**Ala**	*p*-Ethylphenyl thioureayl alanine
^{BuBTU}**Ala**	*p*-Butylphenyl thioureayl alanine
^{DMBTU}**Ala**	*m,p*-Dimethylphenyl thioureayl alanine
^{MOBTU}**Ala**	*p*-Methoxyphenyl thioureayl alanine
^{DMOBTU}**Ala**	*m, m*-Dimethoxyphenyl thioureayl alanine
^{MePyTU}**Ala**	Pyrennylmethylthioureayl alanine
^{NapTU}**Ala**	Naphthyl thioureayl alanine
^{PyTU}**Ala**	Pyrennyl thioureayl alanine
^{HBTU}**Ala**	*p*-Hydroxyphenyl thioureayl alanine
^{ClBTU}**Ala**	*p*-Chlorophenyl thioureayl alanine
^{CNBTU}**Ala**	*p*-Cyanophenyl thioureayl alanine
^{AcBTU}**Ala**	*p*-Acetylphenyl thioureayl alanine
^{BuTU}**Ala**	n-Butyl thioureayl alanine
^{HepTU}**Ala**	n-Heptyl thioureayl alanine
^{EtOHTU}**Ala**	Ethanol thioureayl alanine
^{CyHTU}**Ala**	Cyclohexyl thioureayl alanine
^{PhTzA}**Ala**	1-Phenyltetrazolyl-5-amino alanine
^{MeBTzA}**Ala**	1- (*p*-methylphenyltetrazolyl)-5-amino alanine
^{EtBTzA}**Ala**	1-(*p*-ethylphenyltetrazolyl)-5-amino alanine
^{BuBTzA}**Ala**	1-(*p*-nButylphenyl tetrazolyl)-5-amino alanine
^{DMBTzA}**Ala**	1-(*m,p*-Dimethyphenyl tetrazolyl)-5-amino alanine
^{HBTzA}**Ala**	1-(*p*-Hydroxyphenyltetrazolyl)-5-amino alanine
^{MOBTzA}**Ala**	1-(*p*-Methoxyphenyl tetrazolyl)-5-amino alanine
^{NapTzA}**Ala**	1-Napthyl tetrazolyl-5-amino alanine
^{PyTzA}**Ala**	1-Pyrene tetrazolyl-5-amino alanine
^{ClBTzA}**Ala**	1-(*p*-Chlorophenyl tetrazolyl)-5-amino alanine
^{CNBTzA}**Ala**	1-(*p*-Cyano phenyl tetrazolyl)-5-amino alanine
^{AcBTzA}**Ala**	1-(*p*-Acetylphenyl tetrazolyl)-5-amino alanine
β-AlaN_3/^{N3}Ala	*β*-azidoalanine/Azido alanine
^{NCS}**Ala/Ita**	Isothiocyanyl Alanine
^{NCS}**Lys/Itl**	Isothiocyanyl Lysine
^{SCN}**Ala/ β-Ala$_{SCN}$**	Thiocyanyl Alanine/beta- thiocyanyl Alanine
^{NH2}**Ala**	Amino alanine
^{TU}**Ala**	Thioureayl alanines
^{Tz}**Ala**	Aminotetrazolyl alanines

NMR Data	
δ	Chemical shift in NMR
s	singlet
d	doublet
t	triplet
q	quartet
h	heptete
m	multiplet
bs	broad singlet
dd	double doublet
dt	doublet of triplet
ddd	doublet of doublet of doublet
J	coupling constant in Hz

Chapter-1

Synthesis and Applications of Side Chain Modified Unnatural Amino Acids: A Review



1.1. Introduction

Amino acids (AAs) are responsible for the majority of functional attributes of all living organisms. Amino acids are building blocks of proteins and play an important role in metabolism.[1a] The chemical properties of the amino acids of proteins regulate the biological property of the protein. Most amino acids are responsible for forming tissues, organs, muscles, skin and hair in mammalians.[1b] About 500 amino acids are known in the literature which includes both natural and synthetic. There are 22 proteinogenic amino acids out of which only 20 amino acids appear in the genetic code.[1] They are known as standard or canonical amino acids. The other two amino acids are known as non-standard or non-canonical amino acids. They are selenocysteine[1e] **1.001** (**Figure 1.1**) and pyrrolysine[1e]**1.002** (**Figure 1.1**). The selenocysteine & pyrrolysine are encoded *via* variant codons. As for example, pyrrolysine and selenocysteine are encoded by the stop codon UAG and UGA respectively.[1e,2a] In nature there are another two sulfur-containing amino acids homocysteine **1.003** (**Figure 1.1**) and taurine **1.004** (**Figure 1.1**) (though taurine is not technically an amino acid) which have not been incorporated into proteins.[2c,d]

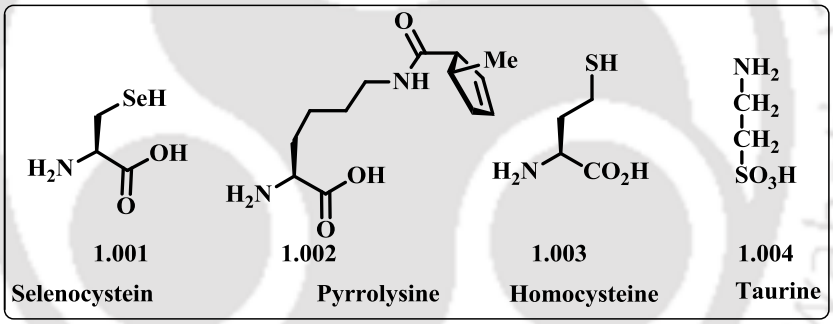


Figure 1.1. Structure of 21st and 22nd amino acids and sulfur containing amino acids.[1e,2c]

The 20 naturally occurring amino acids bear a wide variety of chemical versatility. In living systems, proteins with compact folded structure and self-assembly shows specific biochemical functions as a result of their particular conformation.[2b] The diversity of protein structures and functions are being widely expanded by the formation of some unnatural amino acids (UNAAs) containing modified side chains.[1d,3] Therefore, advanced knowledge to incorporate unnatural amino acids (UNAAs) specifically into proteins represents a very powerful tool for protein engineering.[1d] Unnatural amino acids are non-proteinogenic amino acids that can occur naturally but they are most often chemically synthesised. They are not only used as units to synthesize unnatural peptidic and non peptidic-compounds but also as conformational constraints and molecular scaffolds in order to fulfill several requirements concerning various biological phenomena. In addition, UNAAs are being widely utilised in the generation of pharmacologically active drug candidates.[3b] Thus, chemical diversity of UNAAs can be used to create proteins with new functions.[1c-d,3] As for example replacement of biscysteine based

peptide bond by a dithiol amino acid is an important strategy to design thiopeptide based inhibitor. Avoiding of disulfide bond formation and achieving better inhibiting activity are the major advantages of such strategy.[3]

The research involving the synthesis of UNAA has been a great focus since the last century. Since Theodor Curtius and Emil Fisher synthesised the first simple peptides a century ago, research in developments of methodologies for peptide synthesis has undergone dramatic development.[4] Different technologies are now available for the production of peptides and proteins including (a) the extraction from natural sources, [5b] (b) the production by recombinant DNA technology,[5b] (c) the production in cell-free expression systems,[5b] (d) the production in transgenic animals , plants , chemical synthesis and enzyme technology using proteolytic enzymes.[6a,b] However, the biochemical methods involving the site-specific incorporation of UNAAs into proteins has widely expanded the scope of design of proteins with diverse structure and function.[7] After the revolutionary works of P.G.Schultz's and Chamberlin's group, remarkable progress has been made in exploring these amino acids, in finding novel methods for tRNA aminoacylation (even *in vitro*) and in expanding genetic codes.[19] Aminoacylation of tRNA (Amber and frame shift suppressor tRNAs) with UNAAs have been introduced by focused evolution of aminoacyl-tRNA synthetases (aaRS) or some ribozymes.[19b-m] Codons have been extended to include different unnatural base pairs. Multiple incorporation of different UNAAs has been achieved by variation in codon for each tRNA.[19b-h] The combination of these innovative methods have opened the opportunity for synthesising unnatural mutant proteins in both cell-free systems and in living cells.[8] However the chemical route is often a superior option for the synthesis of medium size peptides comprising of molecules of pharmaceutical importance.[9b] It is also a fundamental tool for understanding the structure-function relationship in proteins/peptides, the discovery of novel therapeutic and diagnostic agents and the production of synthetic vaccines.[9] Recently it has been applied for the design of synthetic biocatalysts, [2c] which has turned out to be an important area of research.

In last few decades many research groups were involved in the design of side chain modified amino acids for peptidomimetics which can be classified into three categories such as- (a) with small side chain functionality[16] (**1.005-1.008**, **Figure 1.2**) (b) large side chain functionality[16] (**1.009-1.012**, **Figure 1.3**) and (c) with linker/ reactive functional group[16,8] (**1.013-1.016**, **Figure 1.4).** These can be utilised for further site specific reactions for protein modification or labeling, for studying protein conformational changes and dynamics with minimal or no perturbation.

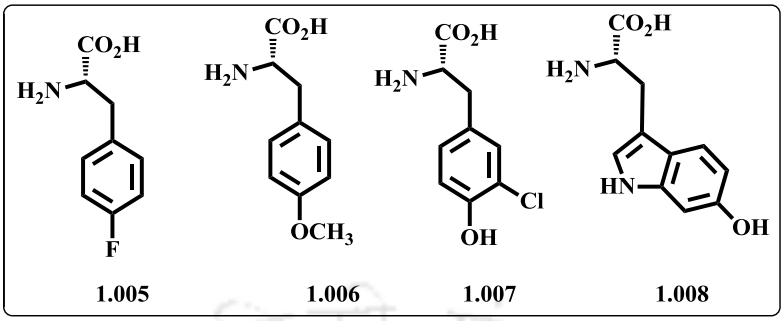


Figure 1.2. Structure of small side chain modified UNAAs.[16]

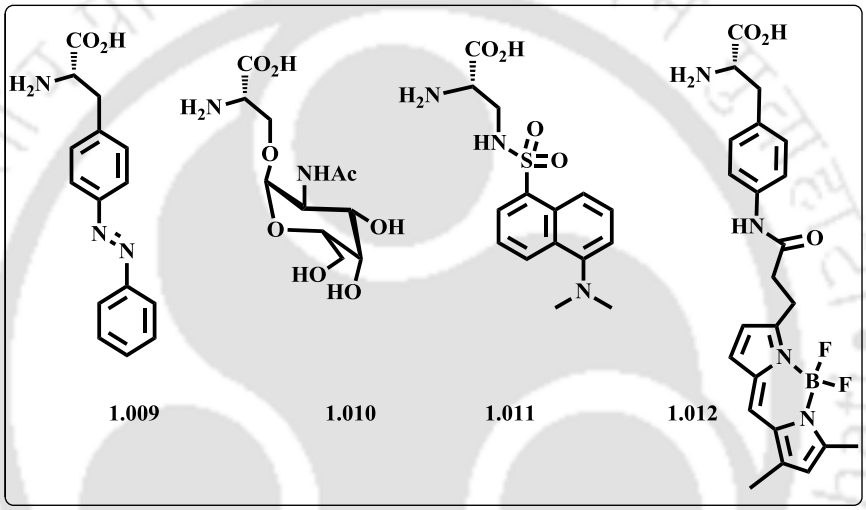


Figure 1.3. Structure of large side chain modified UNAAs.[8a]

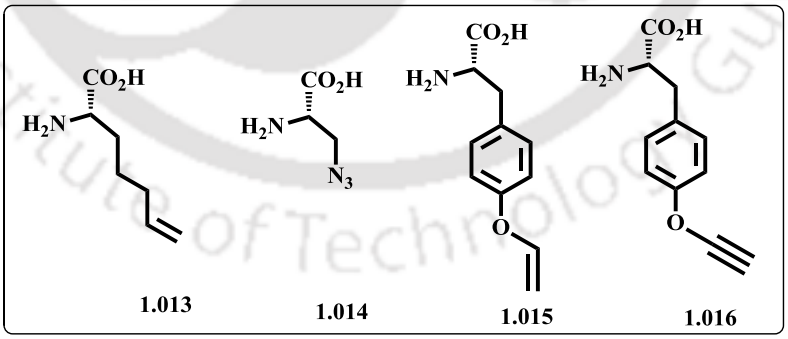


Figure 1.4. Structure of side chain modified UNAAs with linker functional group.[16,8a]

Side chains of amino acids in the peptides can be replaced with analogues of amino acids that have functional properties similar to natural amino acid side chains.[3b] As for example, Heinis *et al*., have synthesised small (**1.017, Figure 1.5**) and large side chain modified dithiol amino acids (**1.018-1.020, Figure 1.5**).They have shown that these amino acids played an important structural role such as it can mimic cysteine residue with respect to functions and size. They reported these thiol groups in dithio amino acids (DTAAs) are capable of forming disulfide link similar to naturally existing cysteines. They applied solid-phase peptide synthesis (SPPS) method to incorporate these dithiol groups containing amino acids effectively into peptides.[10a]

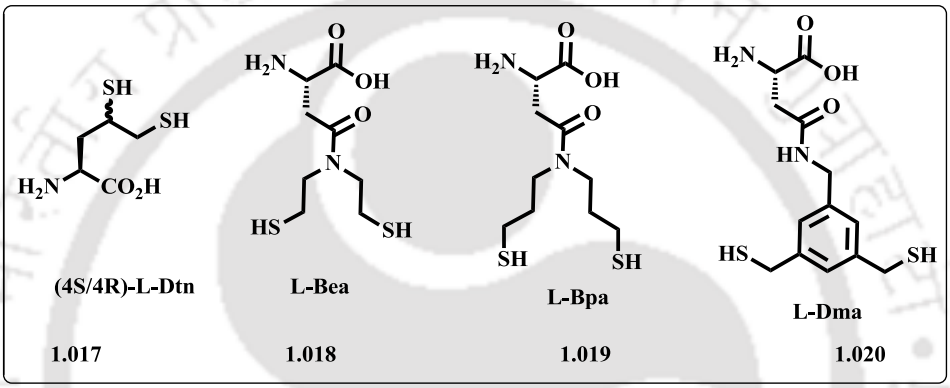


Figure 1.5. Structure of modified dithiol unnatural amino acids.

Peptidomimetics have benefits over natural peptides in terms of stability and bioavailability and therefore, peptidomimetics have better prospective in drug discovery. Peptidomimetics involves main- or side-chain modifications of the parent peptide.[3b] As for example, Nowick *et al*., reported UNAA, "Hao" (**1.022, Figure 1.6**) which is a tripeptide mimicshowing the hydrogen-bonding ability like tripeptide β–strand (**1.021**, **Figure 1.6**). Peptides containing "Hao" and natural amino acids show hydrogen-bonding surfaces that are complementary to the hydrogen-bonding edges of natural protein β-sheets.[10b]

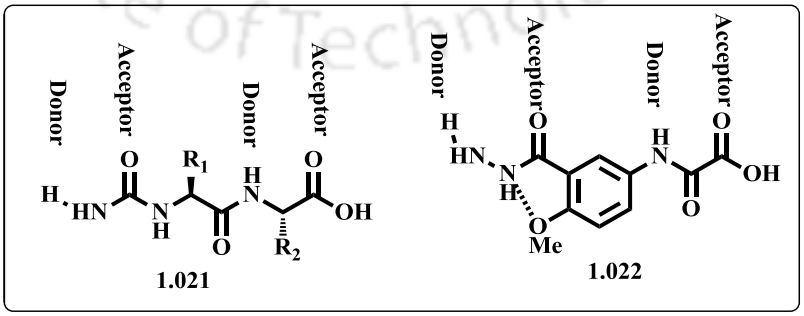


Figure 1.6. Structure of UNAA containing β-strand mimic.

Protein solubility plays a significant role in relation to study diseases that depends on protein misfolding.[11] As for example, protein misfolding and stability in the cell is associated with the kind of neurodegenerative diseases such as Alzheimer's, Huntington's, and Parkinson's disease.[11a-b] Solubility and folding efficacy of a protein can be altered by sequence alterations, site specific modification with different unnatural amino acids residue or modification in side chains.[11c] Several multiple chromophores labeled peptides/proteins have also been developed for investigating folding mechanism,[12] detection of a target protein, sensing microenvironment,[12e,f] sensing metal ions, enzyme activity detection[12e] and studying protein-protein/protein-drug interactions utilising various fluorescence phenomena such as FRET, excimer, and exciplex emission.[52] Fluorescent amino acids can also be used for *in vitro* and cellular imaging of protein localisation, biomolecular interactions and sensing toxic metal ions etc.[16h] The fluorescent proteins are valuable tools in cell biology for monitoring molecular localisation and activities of proteins and gene expression in live cells.[16h]As a result, there is a need to generate modified amino acids to monitor folding and unfolding events on fast timescales in living cells by utilising the fluorescent property of the AA. Only a few naturally occurring AA's have been reported to have florescent properties, therefore, incorporation of fluorescent active UNAA can be utilised for such applications. [12] Additionally, use of a spectroscopic labeled unnatural amino acids to monitor *in vivo* biomolecular distribution enables the construction of highly selective reporting/ imaging agents.[17] Selectively incorporating labeled amino acids can be useful for monitoring sensitive biological processes, such as enzyme catalysis,[13b] product release, [13b] and protein folding.[13b].

This chapter focuses on the critical survey of unnatural amino acids towards the genetic incorporation, the position specific peptide side chain modification, and applications of fluorescent UNAAs in sensing protein micro environment as well as sensing metal ions.

1.2. Importance of Unnatural Amino Acids

Significance behind the use of unnatural amino acids:

(a) Site-specific incorporation of unnatural amino acids would enhance the functional properties of protein. Therefore, protein engineering with unnatural amino acids (UNAAs) is growing at a fast rate to study the structure, function, and dynamics as well as to study the protein-protein/protein-DNA or other inter-biomolecular interactions.[14a]

(b) Optical control of K+ channel in mammalian cells is done by using photo-releasable UNAAs.[14b]

(c) Fluorescent UNAAs can be used for *in vitro* and *in vivo* cellular imaging which could enable to visualize and study the protein localisation, biomolecular interactions and conformational changes in a particular protein system.[14c]

(d) The unnatural D-amino acids are being utilised in treating bacterial infections of the skin and the bacteria that causes the food poisoning.[14d]

(e) UNAAs can also be used as a probe for the structural and functional study of proteins with the help of biophysical techniques such as circular dichroism(CD), nuclear magnetic resonance(NMR), electron paramagnetic resonance(EPR), and infrared (IR) spectroscopy.[14e,13d]

(f) UNAAs can also find applications in designing peptidomimetics drug candidates, often in stabilising a particular conformation of a protein *via* side chain hydrophobic interaction.[14f]

(g) Metal-binding unnatural amino acids (UNAAs) are used in protein research for various applications such as protein sensors, for metalloenzyme engineering, and metal ion detection.[14g]

Therefore, looking at the vast applications and importance, research on the development of newer unnatural amino acids is growing at a fast pace. There are three major focuses on the development of unnatural amino acids which are: (a) unnatural amino acids for genetic incorporation; (b) side chain modified unnatural amino acids and their wide applications; and (c) different fluorescent unnatural amino acids for monitoring protein conformation, sensing microenvironment, cell imaging and sensing metal ions.

1.3. Unnatural Amino Acids for Genetic Incorporation

Protein engineering with UNAAs to develop a group of new protein building blocks finds several applications, like monitoring protein function, *in vivo* probe labeling, capturing transient protein-protein interaction *via* crosslinking etc.[14h] However, incorporation of amino acids into proteins through genetic engineering was limited to the 20 natural amino acids. Later, the pioneering discovery of unnatural amino acids for incorporation into peptide and proteins by Offord [15c], Kaiser[15a-b] and Schultz *et al*., towards expanding the genetic code.[16] The genetic code includes 61 codons for designating 20 naturally occurring amino acids and three codons as stop signals. To incorporate unnatural amino acids into the current genetic code system, it is necessary to generate new codons that are specific to them. In 1989, an introduction of unnatural amino acids into the specific position of proteins through amber suppression method developed by P.G. Schultz was found to be an important discovery to expand protein function.[16]

There are three well-known methods reported by different research group for the genetic incorporation of unnatural amino acids. These are: (a) three base (stop codon) codon method [16b-m]: in this method the amber suppressor tRNAs are aminoacylated with preferred unnatural amino acids through chemical aminoacylation (i.e. an enzymatic ligation of an aminoacyl-dinucleotide with a tRNA lacking the terminal dinucleotide unit) **(Figure 1.7-1.8**).(b) four-base blank codon method [18]: it has overcome the limitation of amber codon suppression method and helps to encode multiple distinct UNAAs into proteins by the use of new blank codons. Taira *et al*., [18] have shown that this method in a eukaryotic cell-free method opens the probability of a four base codon mediated incorporation of such unnatural amino acids residue into proteins of living eukaryotic cells; (c) five-base codon method [19]: developed be Hohsaka *et al*., [19] is another novel strategy for unnatural mutagenesis as well as a novel insight into the mechanism of frameshift suppression. There are 15-five base codons CGGN1N2, where N1 and N2 indicate one of four nucleotides. These codons have been successfully decoded by aminoacyl-tRNAs containing the complementary five-base anticodons.

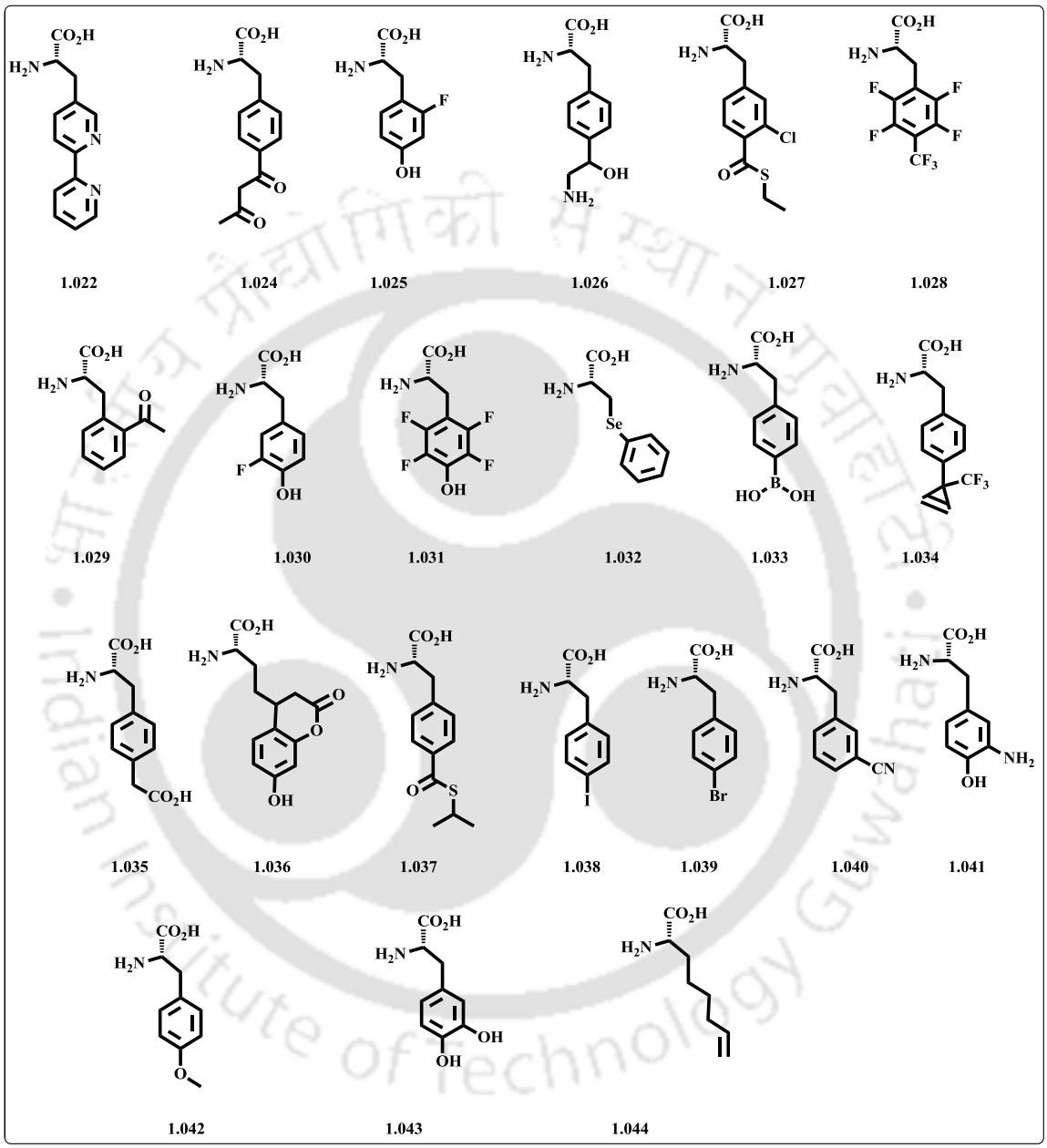


Figure 1.7. Chemical structure of genetically incorporated unnatural amino acids.[16]

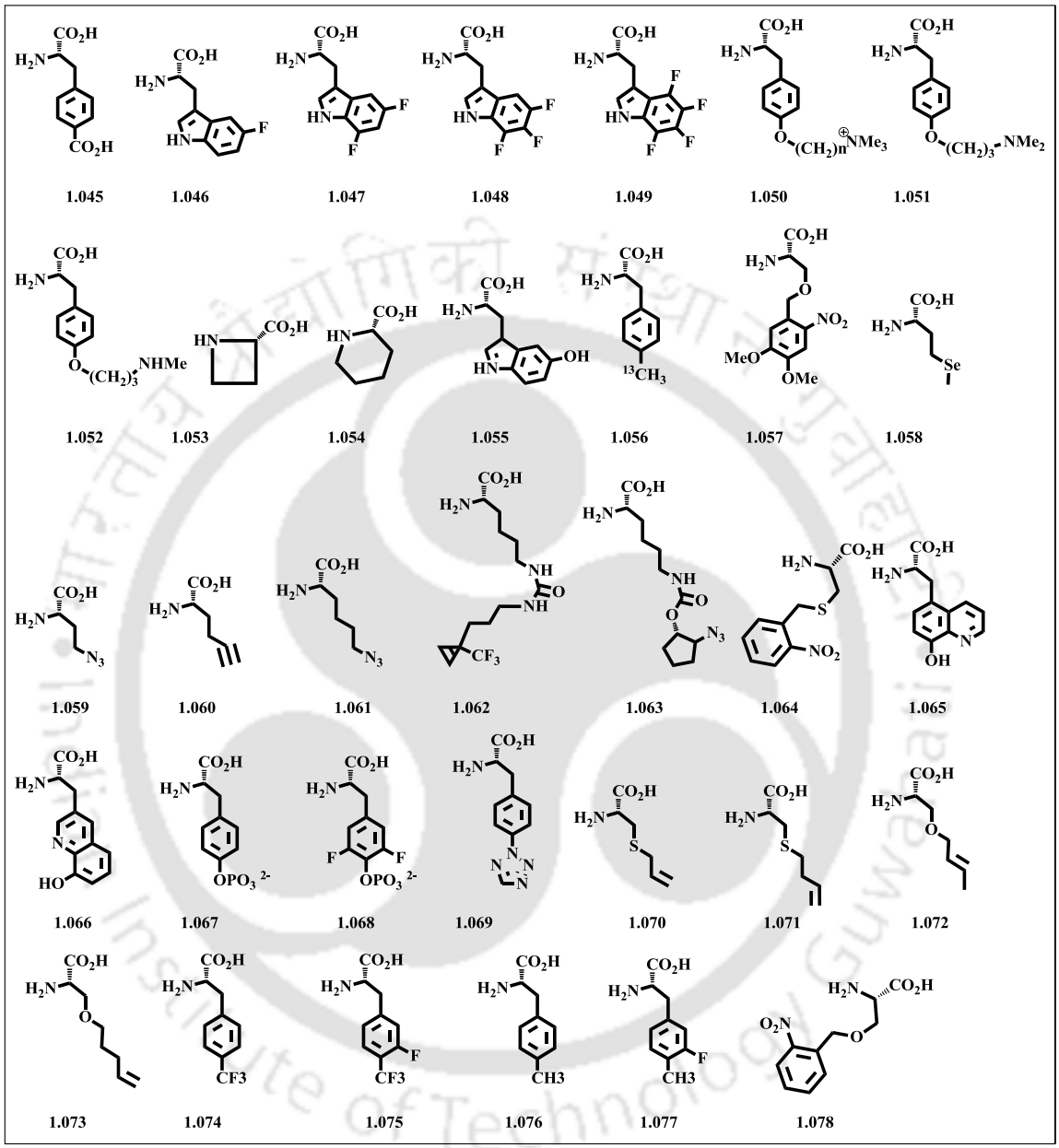


Figure 1.8. Chemical structure of genetically incorporated unnatural amino acids.[16]

1.4. Unnatural Amino Acids as Solvatochromic Probes

The solvatochromic probes have been recognised as ideal probes to study the polarity of microenvironments within a biomacromolecule and molecular moiety.[20] Therefore, solvatochromic fluorophores are important for investigation of protein folding and dynamics.[20e] The IR-spectroscopic study with labeled unnatural amino acids is useful for monitoring protein conformation, folding, and binding.[20] The development of the biological and chemical methods to prepare IR-probe-labeled biomolecules and biopolymers, a variety of vibrational spectroscopic methods has been newly intended and developed, which can be effectively used to monitor the shift, fluctuation, and heterogeneity of IR-probe frequency in real time.[20e] Thus, often fluorophores, such as dye sensitised molecules, and other fluorescent molecule have been used for this reason. However, these have a large molecular size, compared to the usual size of amino acid side chains in proteins.[20] So, the attachment of an external fluorophore to a protein can cause unfavorable perturbation to the natural structure.[20] On the other hand functional groups such as -NO, -CO, -SCN, or -CN are much smaller and they can be introduced into an amino acid side chain (**1.079-1.082, Figure 1.9**) without much structural alterations in amino acids. Hence, these can be used as solvatochromic IR sensitive probe in peptides/protein to investigate the biomolecular interactions. Getahun *et al*.,[20a] Waegele *et al*.,[20b] Jo *et al.*, [20c] and Ye *et al*., [20d] have used different side chain modified amino acids as IR responsive probe for monitoring protein structure and dynamics.

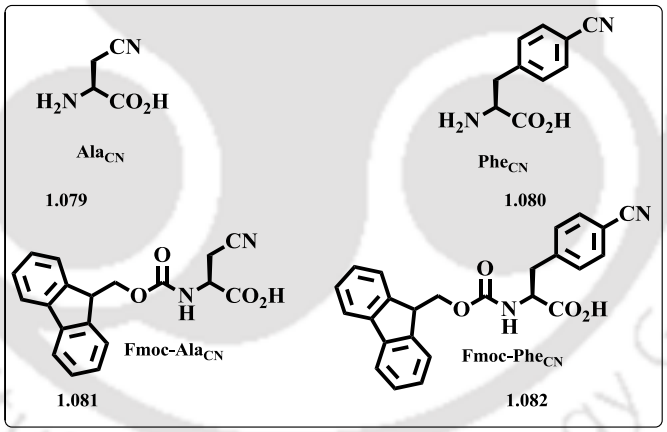


Figure 1.9. Nitrile-derivatised amino acids used for IR responsive probe.[22]

1.5. Side Chain Modification of an Amino Acid in Peptides/Proteins

Nature uses the covalent modification [21a] of proteins to control their different function in a particular system.[21d] So, protein modification *via* chemical route is an important and attractive research area. The complementary use of both genetic and chemical methods has provided a large scope of preparation of modified protein with diverse properties and functions.[21b] These chemical modifications are called posttranslational modifications (PTMs)[21] and they occur

after the synthesis of protein in ribosome. PTMs mostly involve incorporation of new functional groups into target proteins and provides extra benefit to functional property in peptide/protein which plays a crucial role in biological processes.[21c] Undoubtedly, many Intra- and extracellular reactions controlled by the specific chemical change inside the system determines the specific change in both protein structure and function.[21d] As for example, in order to prevent degradation of proteins, they are being regularly modified, leading to intense variation in their properties. In general, PTMs bring benefit to the chemical sensitivity of the protein reaction which helps in particular functional denouement.[21d] These modifications include acylation, methylation, phosphorylation, sulfonation, farnesylation, ubiquitination, and glycosylation which play a crucial role in important cellular processes including trafficking, differentiation, migration, and signaling.[22] Additionally, the bio-orthogonal modification of unnatural moieties/amino acids makes the site-selective modification of proteins as a key tool for investigation of *in vitro* and *in vivo* in the biological system.[23,24] One can decide which residue to target in order to modify the protein to express the desired property/function in the protein of interest. As for example, the circulation half-life of a therapeutic protein may be enhanced by the addition of polyethylene glycol (PEG).[25] In addition, site-selective modification of protein is crucial for the preparation of the next generation of biopharmaceuticals including antibody-drug conjugates (ADCs) for cancer therapy and glycoproteins for vaccination.[27] Such approaches would enable clear molecular examination of modified protein structure on therapeutic function.[26a,e] On the other hand, incorporation of reactive side chain containing UNAAs play important role for peptide/ protein functionalisation. As for example, when azido and ethynyl functional group containing amino acids (**1.059-1.061**, **Figure 1.8**) have been incorporated into recombinant proteins in *E. coli*, it has been reported that those amino acids are capable of crosslinking *via* click chemistry and offer stability to the alpha helical structure of a leucine zipper protein at high temperatures.[28] In addition, incorporation of UNAAs in a larger protein helps to form crosslink at the internal part of the protein. This type of experiment provides fruitful information that this method of "protein stapling" can be expanded to the larger proteins.[29a,b]

There are two approaches: (a) transition metal-free and (b) transition metal-mediated approaches. There are also two ways of modification towards the formation of unnatural amino acids in side chain modification (c) residue/amino acid/site-selective transformations that preferentially modify one amino acid residue over the others (e.g., cysteine versus lysine) and can be considered examples of chemoselective reactions; and (d) regioselective transformation favorably modify only one particular set of the same amino acid, especially when more than one is present in the same molecule (e.g., solvent exposed lysine versus internal lysine).

1.5.1. Transition Metal-Free Approaches

1.5.1.1. Modern Methods for Targeting Natural Amino Acids in Proteins

1.5.1.1.1. Lysine and N-Terminus Side Chain Modification in Proteins

Lysines (Lys) have a primary amine group and most of the time its side chains are protonated at physiological condition but they can act as nucleophiles. Weil and co-workers reported a site and regioselective Lysine modification method by using commercially available N-hydroxysulfosuccinimide (NHS) to form fluorescent probe (**1.083, Figure 1.10a**).[30a] It has been found that some of the sulfonated derivatives of NHS esters have better water solubility of the probe which can be used for *in vitro* labeling of a protein under physiological conditions.[30b] Likewise, isothiocyanates react with primary amines of Lysine and form thiourea (**1.085, Figure 1.10b**) but the minimum pH for an isothiocyanate reaction is 9 to 9.5 which is higher than the formation of NHS esters (pH 8 to 9) which might be unfavorable for the modification of alkaline-sensitive protein due to its high sensitivity of nucleophilic addition reaction.[21d] Not only the difficulties in synthesis of NHS esters and isothiocyanates but also amino group of lysine as a nucleophile can permanently open 1H-benzo[d][1,3]oxazine-2,4-dione leading to orthoaminebenzamide formation inside cells (**1.086, Figure 1.10c**). [21d] In 2006, Hooker *et al.*,[31a] have modified proteins using an oxidative approach with $NaIO_4$ and a dialkyl acyl phenylenediamine (**1.087, Figure 1.10c**).

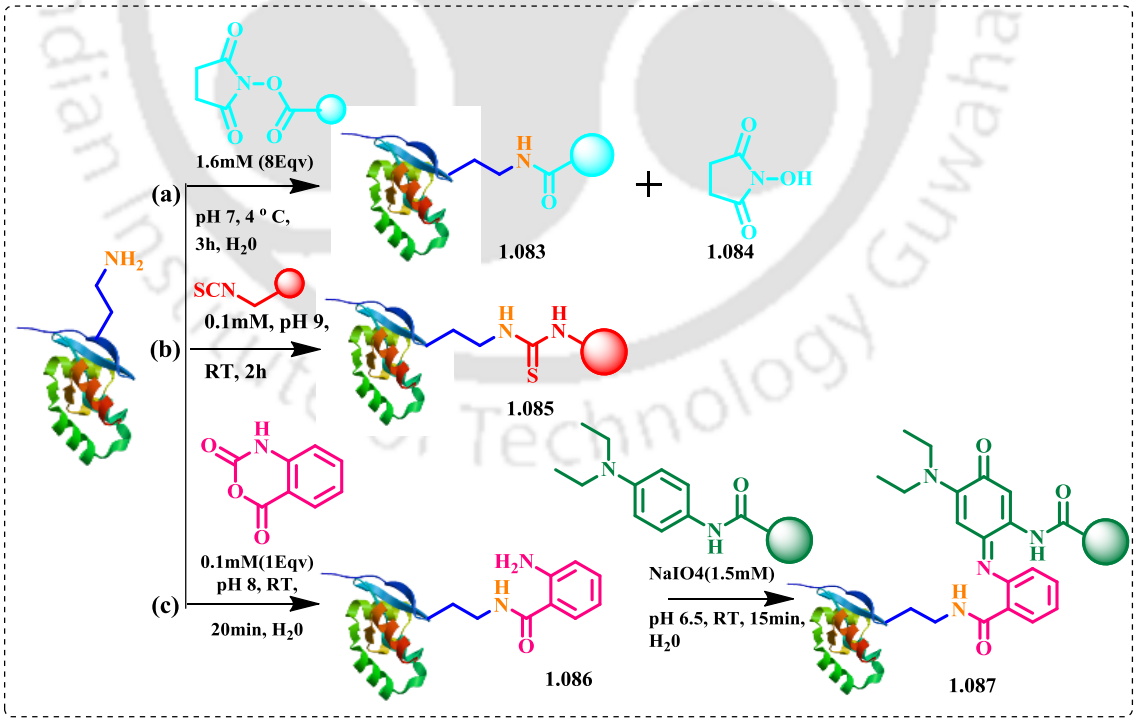


Figure 1.10. Site selective method for lysine side chain modification.

In 1979, Jentoft and Dearborn,[31b] showed that the reductive amination of an aldehyde by forming imine or Schiff base forms quickly and reversibly with both part of lysine and N-terminus. In next sequence, the imine got reduced into the secondary amine using $NaBH_3CN$ (**1.089, Figure 1.11d**) for further modification. In 1999, Tanaka *et al*.,[31c] reported another efficient way which can overcome the reversibility of the imine conversion, *via* a 6-π-aza electrocyclisation (**1.090, Figure 1.11e**). The irreversible conversation of the corresponding zwitterions (**1.091, Figure 1.11e**) has been shown by Tanaka *et al*.[31d,e] Recent site-selective methods for the modification of lysine and the N-terminal position using imino-2-methoxyethyl reagents (IME) (**1.092, Figure 1.11f**), reversible lysine N-terminal modification *via* stable Iminobornates (**1.093, Figure 1.11g**) are reported by different research group.[31f]

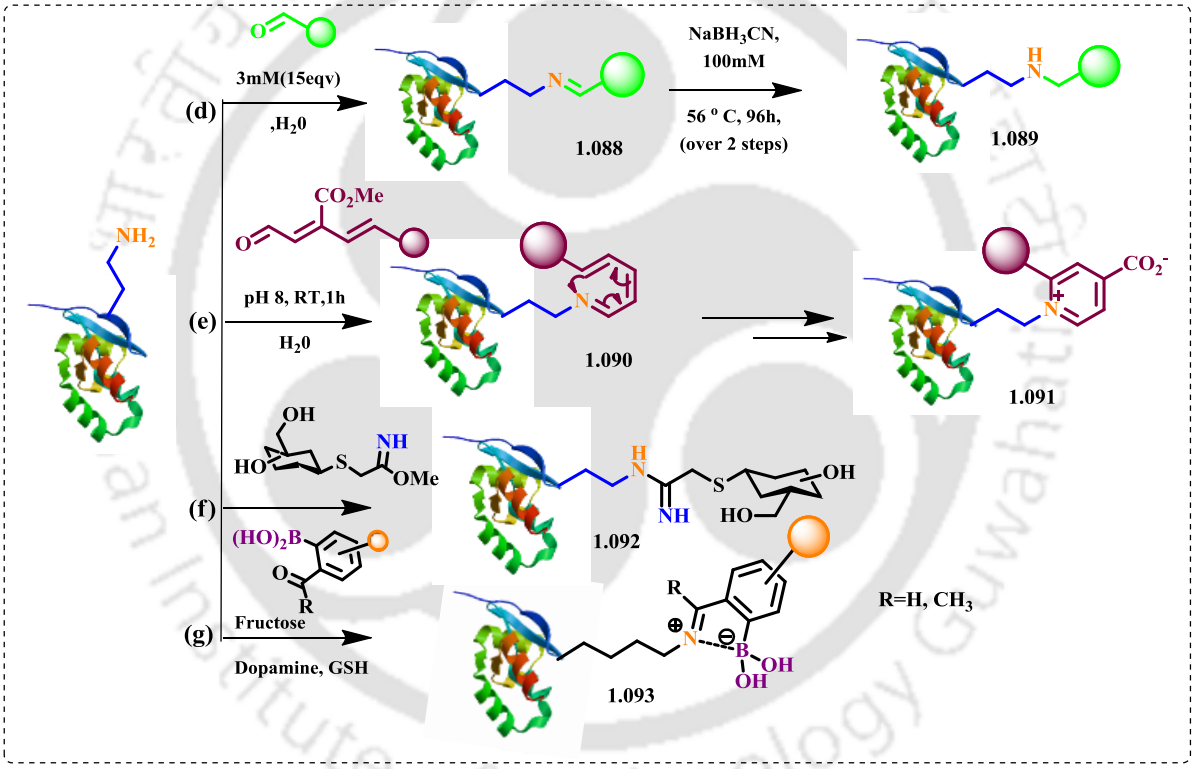


Figure 1.11. Site selective method for lysine side chain modification.

1.5.1.1.2. Tyrosine Side Chain Modification in Proteins

We know that the phenol groups present on tyrosine, can act as nucleophile *via* addition to several activated bonds. The common process reported conversion of phenol to a diazonium salt. In 2004, Hooker *et al.*,[32a] selectively modified ortho position of the phenol group (**1.094, Figure 1.12a**) into diazonium derivative. However, the substrate was very inadequate to aryl diazonium salts containing electron-withdrawing groups in the para position. Fascinatingly, aniline can be formed by the reduction of the azo bond and converted to a stable benzoxazine with an acrylamide containing different functional group (**1.096, Figure 1.11a**).[21d]

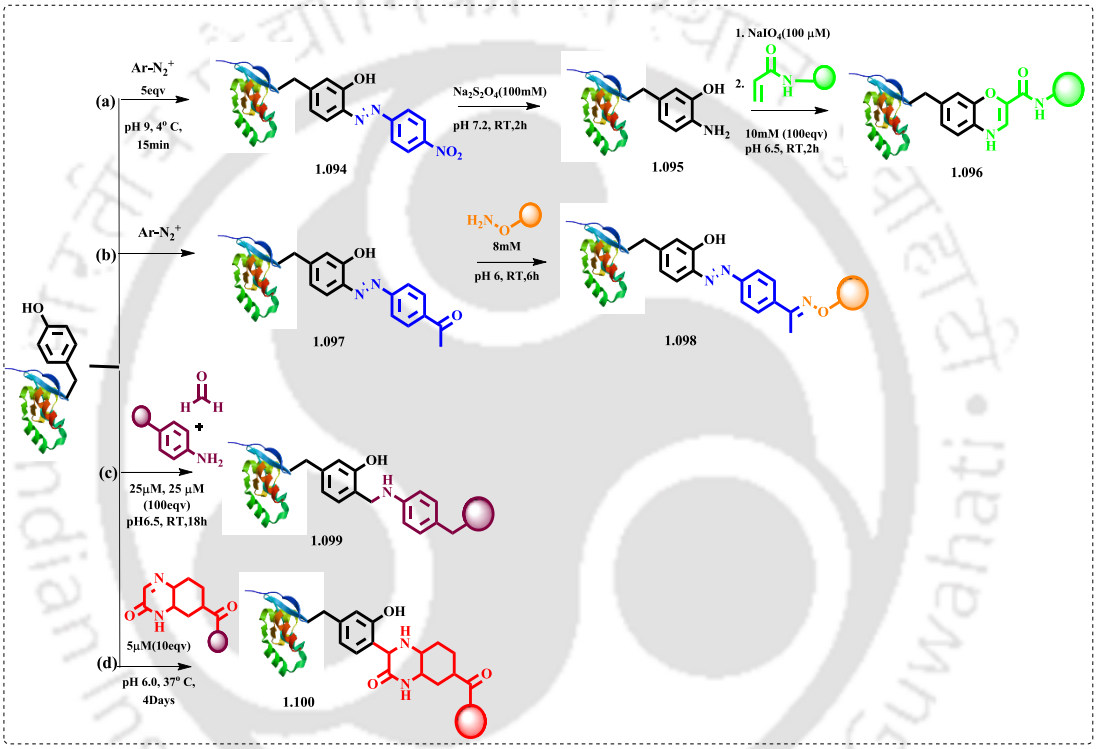


Figure 1.12. Site selective method for tyrosine side chain modification.

In the next year 2005, Schlick *et al.*,[32b] have reported another process in which a ketone-contained aryldiazonium upon condensation with alkoxy amine formed oxime ethers (**1.097-1.098, Figure 1.12b**). Joshi *et al.*,[32c] have shown another method based on using a Mannich reaction which requires an aldehyde and an amine as supplementary reactants. It allows for tyrosine residues to be modified with a high selectivity (**1.099, Figure 1.12c**). McFarland *et al.*,[32d] have reported that the Mannich reaction condition doesn't control the selective modification of tyrosine residue. Recently in 2009, Guo *et al.*,[32e] have prepared special kind of fluorophore containing imines which can react well in the labeling purpose of proteins (**1.100, Figure 1.12d**).

1.5.1.1.3. Cysteine Side Chain Modification in Proteins

The thiol moiety shows more nucleophilic character than the primary amine mainly at a pH below 9 and the amine is protonated under this condition. Therefore, cysteine favorably reacts much faster than lysine.[21d] So one can use this strategy towards the selective modification of cysteine in presence of lysine residue. In 2008, Bernardes *et al*.,[33a] have shown that oxidative elimination of cysteine upon reaction with O-mesitylenesulfonylhydroxylamine (MSH) leads to the formation of dehydroalanine (**1.101, Figure 1.13a**) under basic pH. In the next year 2009, Chalker *et al*., [33b] have reported that these modified proteins undergo subsequent reaction by the addition of thiols leading to thioether analogs (**1.102, Figure 1.13a**). In 2008, Kim *et al*.,[33c] have shown that Maleimides react selectively and stoichiometrically with thiols (**1.103, Figure 1.13b**). Similarly, an alkylation can be done using iodoacetamide (**1.104, Figure 1.13c**). In 1959, Gundlach *et al*., [33d] reported in some particular cases that these reactive species can undergo reaction with methionine residues. Formation of disulfide is selective for cysteine residues, Rademann *et al*., found that the covalent attachment of the probe through a disulfide bond (**1.105, Figure 1.13d**) brings in equilibrium in a system. However, the disulfide link is very unstable in presence of mercaptoethanol or dithiothreitol.[33f]

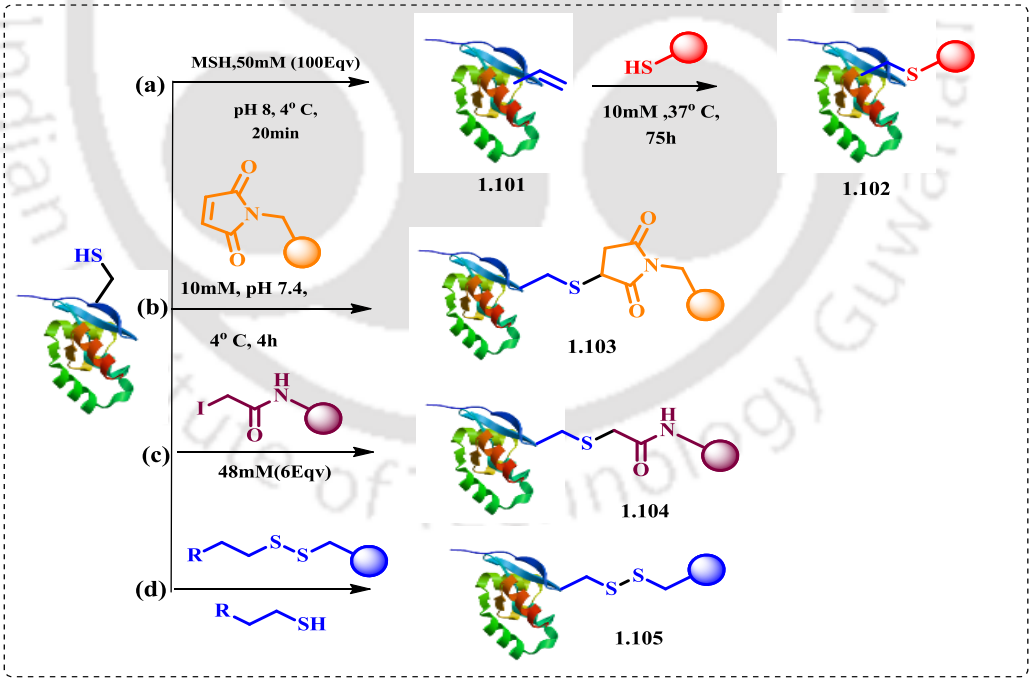


Figure 1.13. Site selective method for cysteine side chain modification.

In 2009, Chalker *et al*.,[33e] have successfully shown, S-Allyl cysteine can also be formed by applying one pot treatment with MSH with 2-propen-1-thiol (**1.106**, **Figure 1.14f)**. At the same time, Crich *et al.*,[33g,h], and Chalker *et al.*,[33a,b] have reported the rearrangement and reduction of allylic disulfides and allylic diselenides for the modification of proteins (**1.106, Figure 1.14e,g**). Thiols can also be added to electron-poor triple bonds such as alkenoic amides (**1.107, Figure 1.14h**) which can take part in a second addition reaction for protein release which has been shown by Shiu *et al*.[33j] Boons, Popik, and co-workers [33i] have demonstrated the selective and reversible photochemical derivatisation of cysteine residues on proteins using 3-(hydroxymethyl)-2-naphthol derivatives (NQMPs) (**1.108, Figure 1.14i**). Barbas and co-workers expanded this methodology as a rapid thiol-selective protein modification strategy in a recombinant HSA and a fusion maltose-binding-HA peptide protein (MBP-C-HA), generating fluorophore and PEGylated conjugates superior in stability to maleimide-conjugated proteins in human plasma (**1.109, Figure 1.14j**).[33k] Recently, Abbas, Xing, and Loh have introduced allenamides for cys-selective protein in aqueous media (**1.110, Figure 1.14k**).[33]

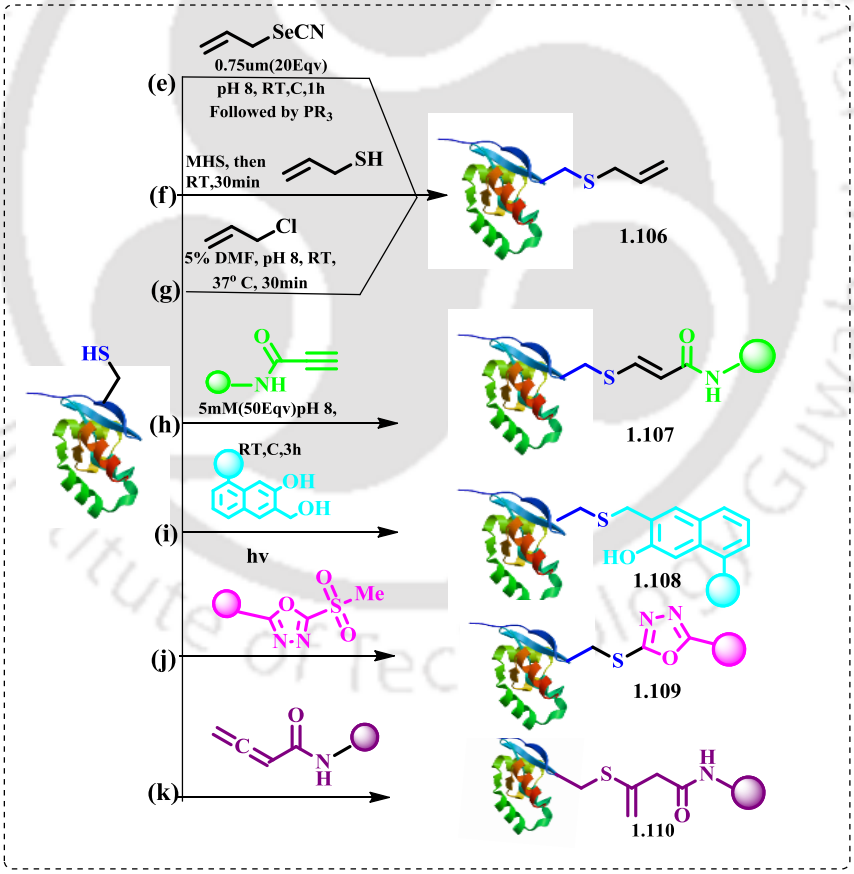


Figure 1.14. Site selective method for cysteine side chain modification.

1.5.1.1.4. Arginine Side Chain Modification in Proteins

Arginine has side chain amine group which can be modified selectively. In 1999, Oya *et al*.,[34] shown the formation of a pyrimidine derivative (**1.111-1.112, Figure 1.15**) at a particular concentration in a phosphate buffer (pH 7.4). The reaction has been applied on a simple model, Na-acetyl-L-arginine, then applied to bovine serum albumin to show the selectivity towards the modification of arginine.[34]

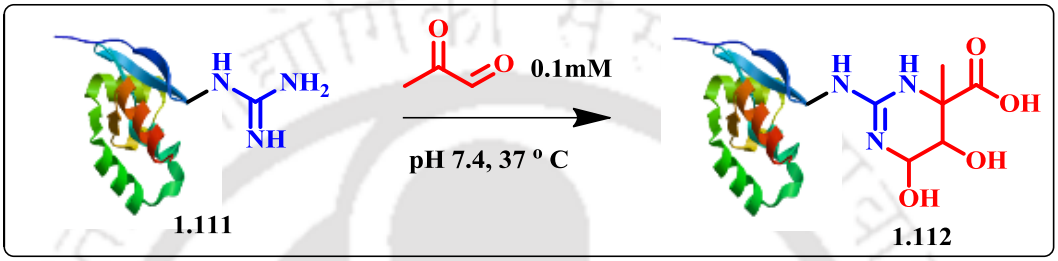


Figure 1.15. Site selective method for arginine side chain modification.

1.5.1.1.5. Histidine Side Chain Modification in Proteins

Histidine has an imidazole side chain (which is partially protonated) as a positively charged amino acid at physiological pH. Epoxides have been used commonly to modify proteins and these simultaneously react with various nucleophiles leading to unsystematic protein surface modification. So, it needs selectivity. In 2003, Chen *et al*.,[35a] have reported that selective histidine modification can be done (**1.113-1.114, Figure 1.16**) in a smiliar fashion. Scheck *et al*.,[35b] has found that tryptophan and histidine give Pictet-Spengler adducts with high yields whereas lysine and glutamine failed to give similar adduct. In 2007, this method was used for the N-terminal selective modification of antibodies by Scheck and Francis.[35c]

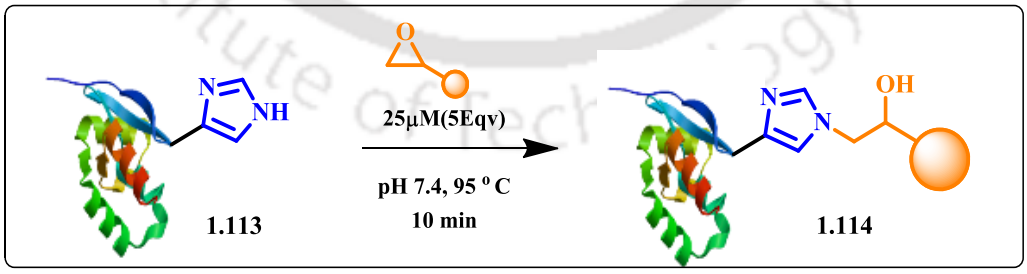


Fig 1.16. Site selective method for Histidine side chain modification.

1.5.1.1.6. Glutamate and Aspartate Side Chain Modification in Proteins

Both glutamate and aspartate has the side chain carboxylic group, so modification can be done selectively in the side chain carboxylic groups or the C-terminal position of a protein. This process is composed of standard peptide coupling reactions making amide bonds with this terminal functionality and using water soluble carbodiimide.[36a] After a few years later in 2005, Schlick *et al*., [36b] position specifically modified glutamate residues on the protein coat of a tobacco mosaic virus by using solution phase peptide coupling method in phosphate or HEPES buffer (pH7.4) at room temperature (**1.115-1.116, Figure 1.17**).

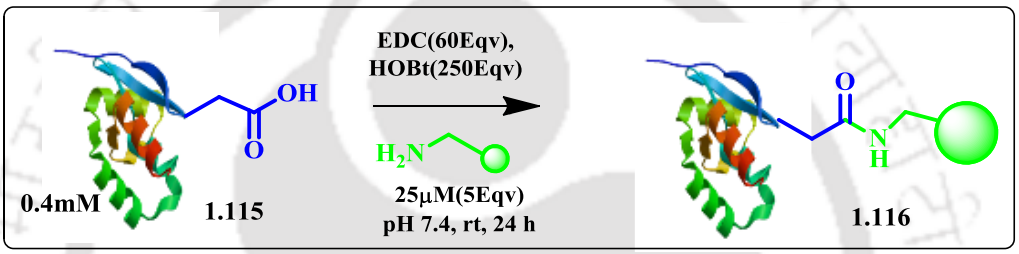


Figure 1.17. Site selective method for glutamate/aspartate side chain modification.

1.5.1.1.7. Disulfides Bond Modification in Proteins

The modification of native disulfide bonds through bis-alkylation protocols (**Figure 1.18**) has attracted growing interest in current years due to novel therapeutic proteins in the global biotechnology market.[21c] This endeavor has stimulated the development of new methods for improving the properties and safety of these protein-based pharmaceuticals.[21c] Among these technologies, the work by Shaunak and coworkers [37b] was found to be mainly suitable (**1.117-1.118, Figure 1.18**). Disulfide modification allows th e site-selective conjugation of poly (ethylene glycol) (PEG) or other polymers to proteins from which optimal pharmacokinetics are achieved, hence ensuring protein purity, maintenance of tertiary structure, and stability.[37] In the same way, Haddleton, Caddick, and co-workers employed dibromo[38]and dithiophenolmaleimides[39] and Baker and coworkers used aryloxymaleimides[40] as bis-alkylating reagents for the *in situ* bridging of disulfides allowing the incorporation of PEG chains in a very efficient manner (**1.119-1.120, Figure 1.18)** at pH 6.2. It has been found that this thiomaleamic acid linker is stable under physiological conditions, it can be cleaved easily at lysosomal pH 4.5 at 37 ° C, assisting the preparation of homogeneous antibody−drug conjugates (ADCs) at disulfide sites.[41]

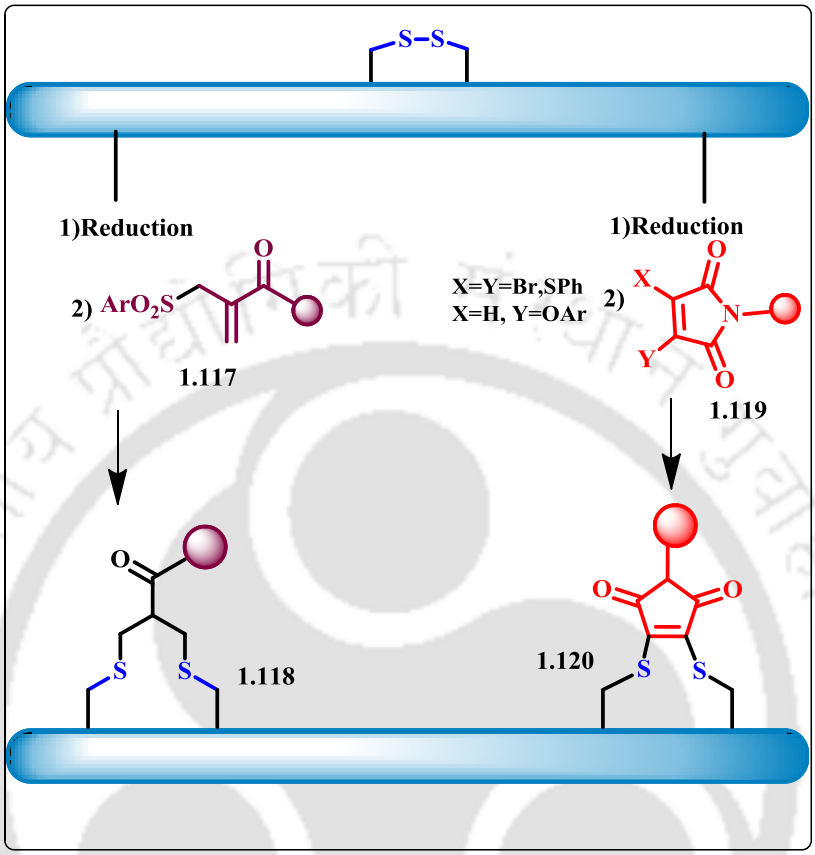


Figure 1.18. Site-selective methods for the modification of exposed disulfides.[21c]

1.5.1.1.8. Photochemically Induced Bioconjugation in Proteins

Photochemically induced bioconjugation in proteins is another promising area to produce highly reactive reactant species like radicals, nitrenes, and carbenes using particular wavelengths of light.[42a, 21d] It can be generated by the decomposition of benzophenone [42a,b] (**1.121**, **Figure 1.19a)**, aryl azide [42a,b] (**1.124**, **Figure 1.19b)**, and diazirine [42c-d] (**1.129,1.133**, **Figure 1.19c-d)**. As they are highly reactive, these species undergo rapid reaction with any types of molecule without considering its nature and these also get hydrolysed easily by the solvent. On the other hand, these species can react with any types of molecule, including DNA, oligosaccharides and proteins. This method is very useful for molecular biology study. However, photo labeling is neither site nor residue specific most of the time.[21d]

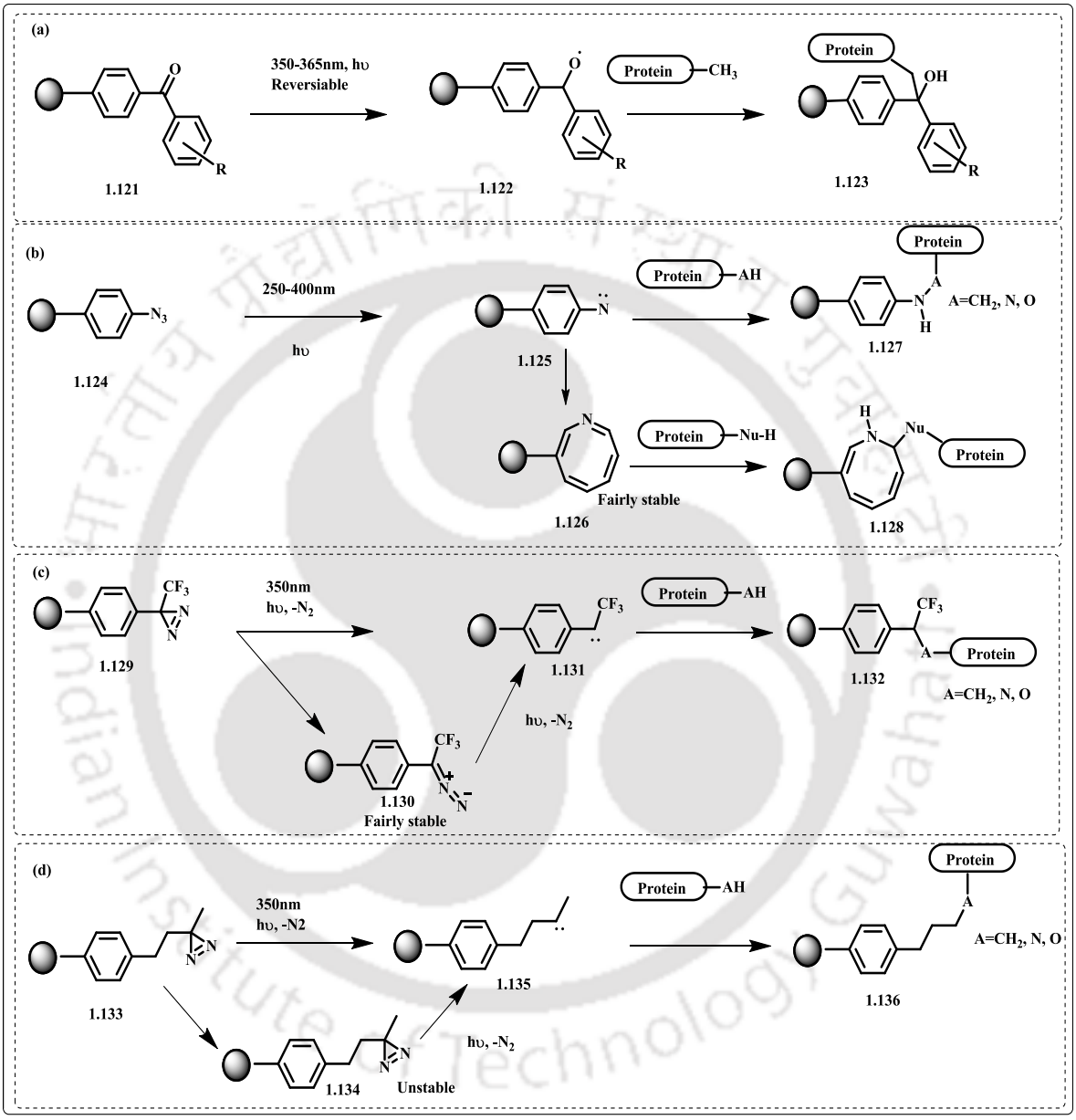


Figure 1.19. Protein photo labeling.[21d]

1.5.2. Transition Metal-Mediated Approaches

1.5.2.1. Modern Methods for Targeting Natural Amino Acids in Proteins

The incorporation of UNAAs combined with chemical protein synthesis allows access to chemically modified proteins which are difficult to obtain using standard molecular biology methods. As for example, the unnatural amino acid containing alkene/alkyne, azido, or carbonyl groups serves as the effective tool for the further bioorthogonal modifications.[43] To overcome the problem associated with amino acid side chain modification, transition metal complexes have been smartly used by Antos *et al.*; Francis *et al.*,[44a] and van Maarseveen *et al.*,[44b] to bring selective labeling of proteins. In 2005, McFarland and Francis *et al.*,[44d] modified lysine side chain (**1.139, Figure 1.20a**) using Iridium (II) complex. In 2006, Tilley and Francis *et al.*, successfully labeled tyrosine (**1.142, Figure 1.20b**) in chymotrypsinogen using Palladium (II) Complex.[44e] In 2004, tryptophan was irreversibly modified (**1.145, Figure 1.20c)** by diazoacetate derivatives in the presence of Rhodium (II) complexes by Antos and Francis *et al.*[44c]

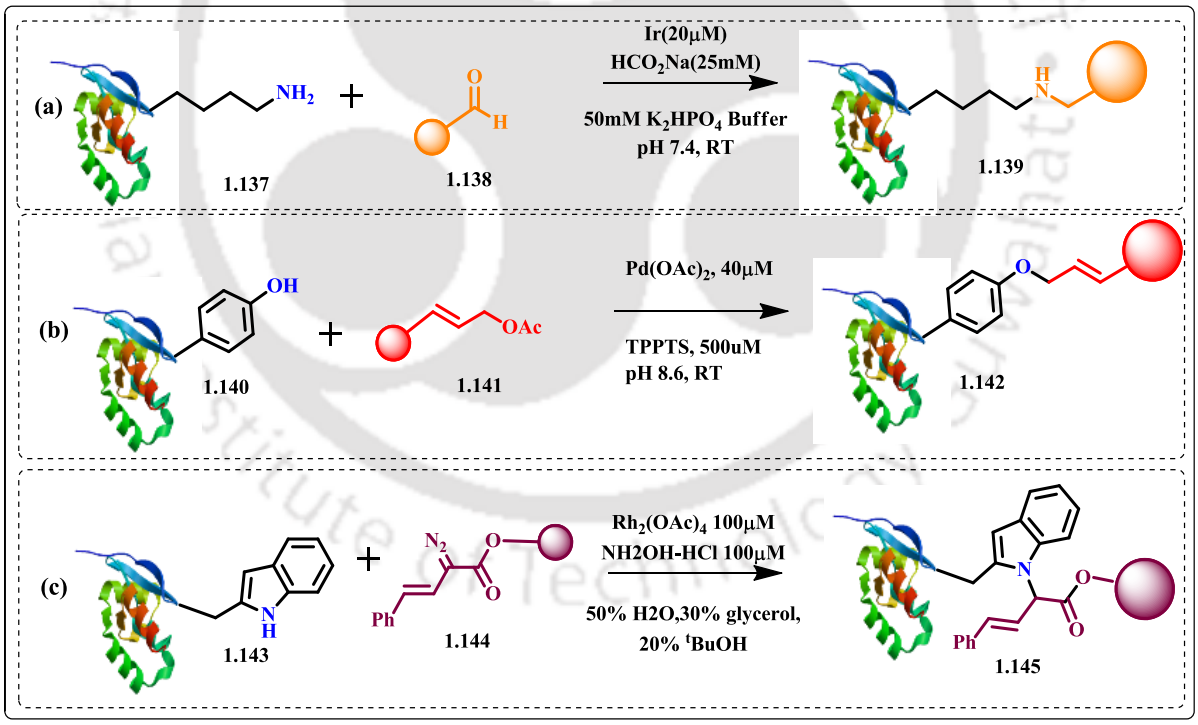


Figure 1.20. Transition metal-catalyzed bioconjugation.

1.6. Fluorescent Unnatural Amino Acids

Unnatural amino acids with polyaromatic side chains are useful fluorescent probe in sensing protein's microenvironment and can be widely used for protein engineering.[48] Position specific incorporation of such types of amino acids into proteins is thus important for monitoring the structures, dynamics, and inter-biomolecular interactions.[45] Polyaromatic side chains amino acids are highly electron rich species and thus can act as electron donor for electron transfer in protein.[48] The highly sensitive fluorescence probe is useful for visualising intracellular events and understanding molecular interactions inside a cell such as elucidating protein's structures [45], functions and dynamics,[45h-i] receptor-ligand binding,[45c-d] monitoring enzyme activity,[45b,h-i] and proteomics[45m]. Also position specific incorporation of fluorescent unnatural amino acids into protein for microenvironment sensitivity [46,47] is a promising area of protein research. Therefore, one can design and generate fluorescent unnatural amino acids (FUAAs) for genetic encoding or to generate labeled proteins/peptide for studying conformational or diverse functional realm.[48] In 1999, M. Sisido *et al*.,[8,18,] have efficiently incorporated various fluorescent unnatural amino acids with modified aromatic side chain (**1.145-1.165, Figure 1.21**).[48]

In order to improve the concept of environment-sensitive fluorescent amino acid building blocks, fluorophores systems were developed, with the goal of increasing the fluorescence changes upon binding in target systems.[49a-b] The environment-sensitive amino acid building blocks based on the polar fluorophoric groups have also been exploited for probing protein interactions and protein unfolding.[50] A brief review of literature of UNAAs bearing fluorescent heterocyclic moieties have been described in this section as fluorescent and/or colorimetric sensor for protein conformation monitoring, sensing microenvironment, cell imaging and chemosensors for metal cations.

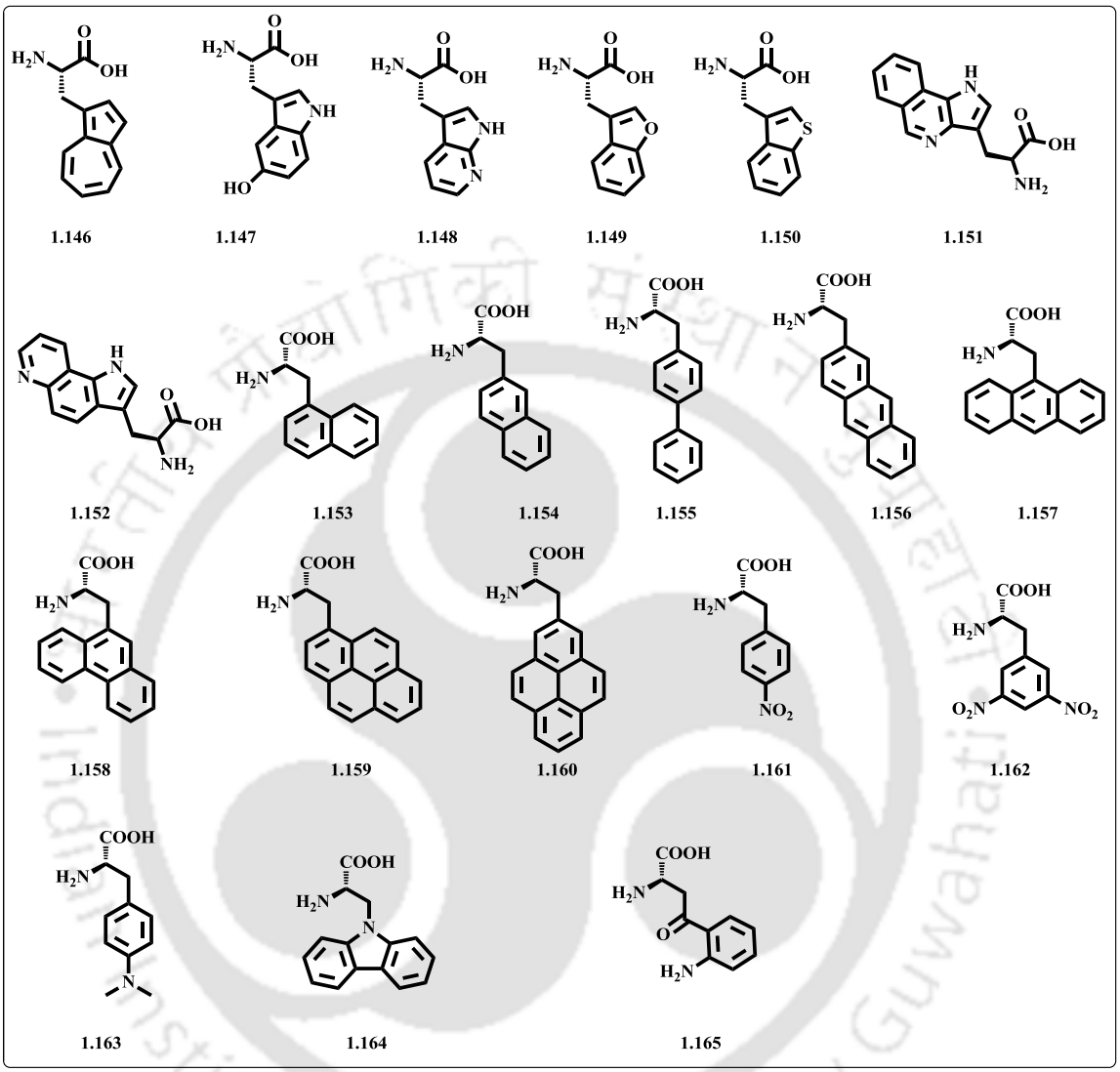


Figure 1.21. Various fluorescent unnatural amino acids (FUAAs).[.48]

1.6.1 Applications of Fluorescent Unnatural Amino Acids: Protein conformation/ Cell imaging / Microenvironment Sensitivity

Petersson *et al*., [51] have utilised coumaryl amino acids as a fluorescent signaling unit to study the kinetics of protease enzymes (**Figure 1.22)** which shows "turn-on" fluorescent protease substrates. Utilising this strategy, they tested a variety of serine, cysteine, carboxyl, and metallo-proteases, including trypsin, chymotrypsin, pepsin, thermolysin, papain, and calpain. This technique can be used for real time monitoring of protease activity in crude preparations of virtually any protease.[51]

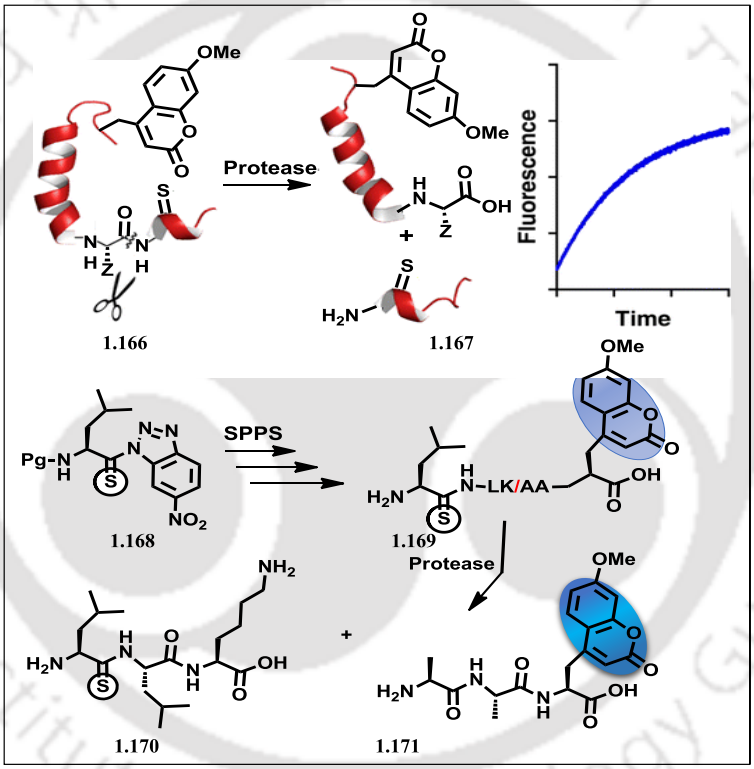


Figure 1.22. Fluorescence turn-on pathway to monitor protease activity.

Petersson *et al*.,[52] have also synthesised cyanotryptophans and incorporated into dihydrofolate reductase (DHFR) protein (**1.172-1.174, Figure 1.23**) into two different positions of DHFR. The 6-cyanotryptophan (6-CNTrp) **1.172** was found to form an effective forster resonance energy transfer (FRET) pair with L-(7-hydroxycoumarin-4-yl)ethylglycine (HCO) at position-17. Thus, they demonstrated that a UNAAs could be used as a FRET partner for monitoring protein−nucleic acid interactions. They also demonstrated that 6-CNTrp could be utilised as a fluorescence donor for the study of protein conformational changes.[52]

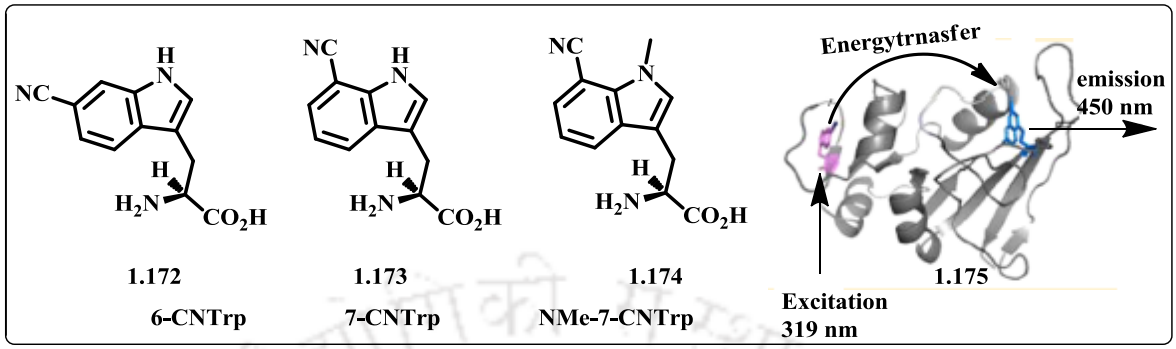


Figure 1.23. Incorporated cyanotryptophane analogue into dihydrofolate reductase protein.

Another approach to generate structurally novel fluorescent UNAAs is the π-conjugation in tyrosine/phenylalanine ring. These UNAAs showed improved photophysical properties than the corresponding tyrosine/phenylalanine amino acids. Tsao *et al.* [53] have incorporated the 2-naphthol analogue of tyrosine, 2-amino-3-(6-hydroxy-2-naphthyl) propanoic acid (NpOH) genetically into proteins (**1.176, Figure 1.24**) of *Escherichia coli*. This results provide an alternative way for biocompatible chemical transformation of protein and provides protein conformational information.[53]

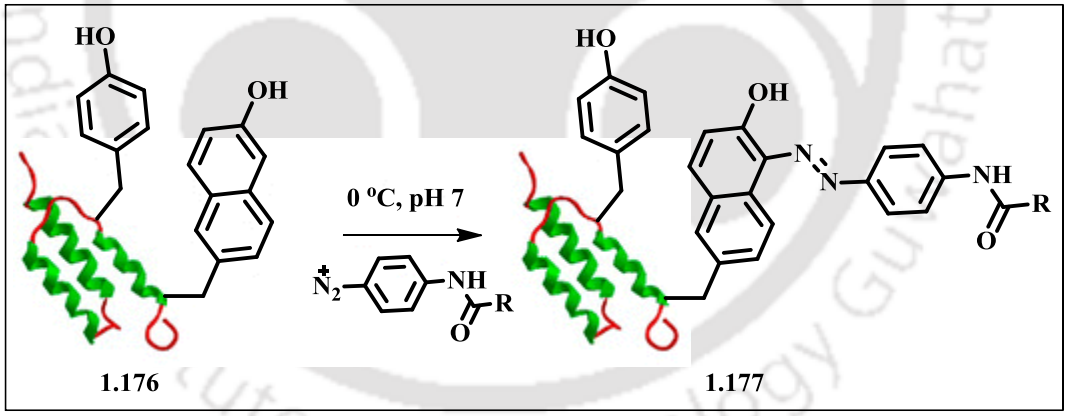


Figure 1.24. Incorporation of 2-naphthol analogue of tyrosine into *Escherichia coli* proteins.

Two fluorescent amino acids, including the novel fluorescent species 4-biphenyl-l-phenylalanine (**1.178, Figure 1.25**), and l-(7-hydroxycoumarin-4-yl)ethylglycine (**1.182, Figure 1.25**) have been incorporated at positions-17 and 115, respectively, of dihydrofolate reductase (DHFR) to facilitate the monitor of conformational switching linked with inhibitor binding *via* Förster resonance energy transfer (FRET).[54]

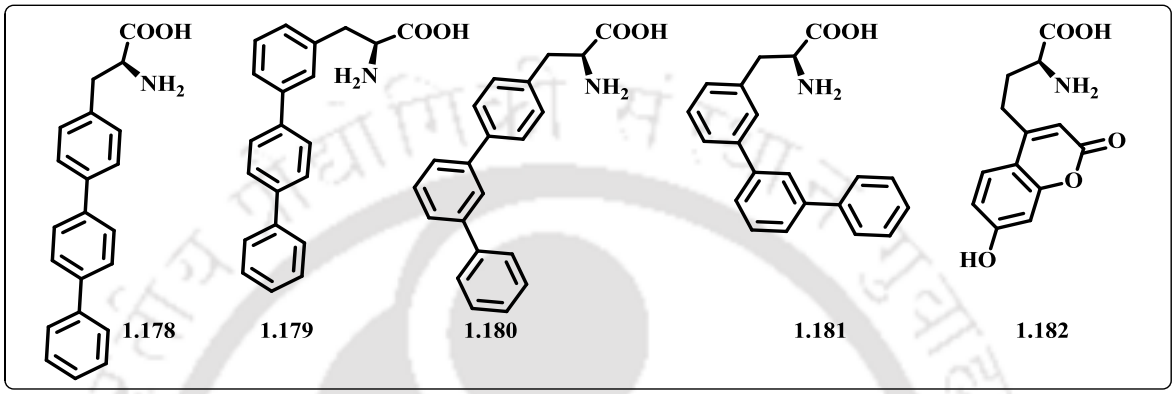


Figure 1.25. Structure of biphenyl-phenylalanines and coumaryl derivative of UNAAs.

Recently, Sidney M. Hecht *et al.* [55] have synthesised oxazole and acridon-2-ylalanine based fluorescent amino acids (**1.183-1.184, Figure 1.26**) which are efficient in FRET and demonstrated the potential utility of the fluorophores as reporter groups (FRET probe) to monitor protein conformational changes and molecular interactions.[55]

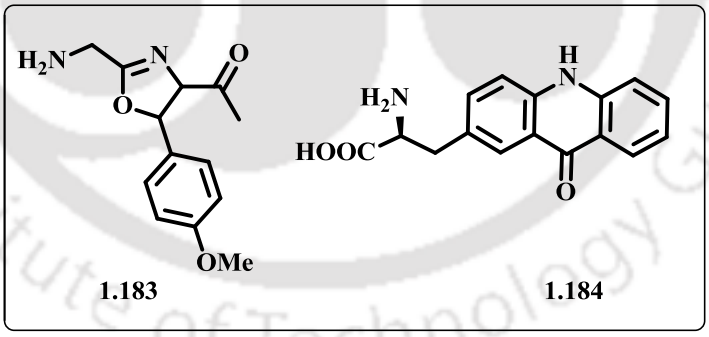


Figure 1.26. Structure of oxazole and acridon-2-ylalanine derivative.

In 2004, Garbay *et al*., have developed direct synthetic method to prepare optically pure coumarine labelled amino acids (**1.187-1.188**, **Figure 1.27**) from inexpensive chiral materials.[56] They have investigated the cellular localisation of coumaryl amino acid labeled peptides through confocal fluorescence microscopy.

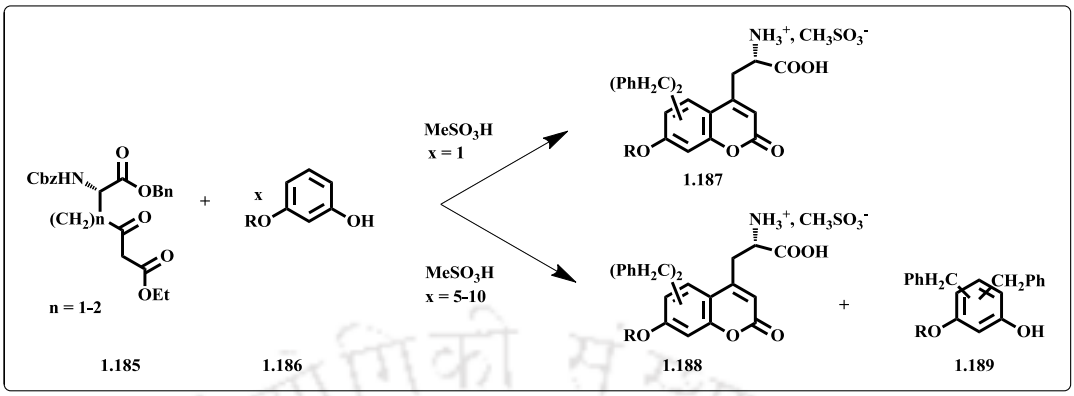


Figure 1.27. General synthesis of coumaryl amino acids.

The α-amino acid, 6-(2-dimethylaminonaphthoyl) alanine (DANA) (**1.191**, **Figure 1.28**) containing highly solvatochromic fluorophore 6-propionyl-2-(dimethylamino) naphthalene (PRODAN), (**1.190**, **Figure 1.28**) was synthesised in 2002.[57] These amino acid have been utilised in peptides for monitoring peptide–protein/protein–protein functions and interactions, where binding ability often helps substrates to come in contact with more hydrophobic surfaces.

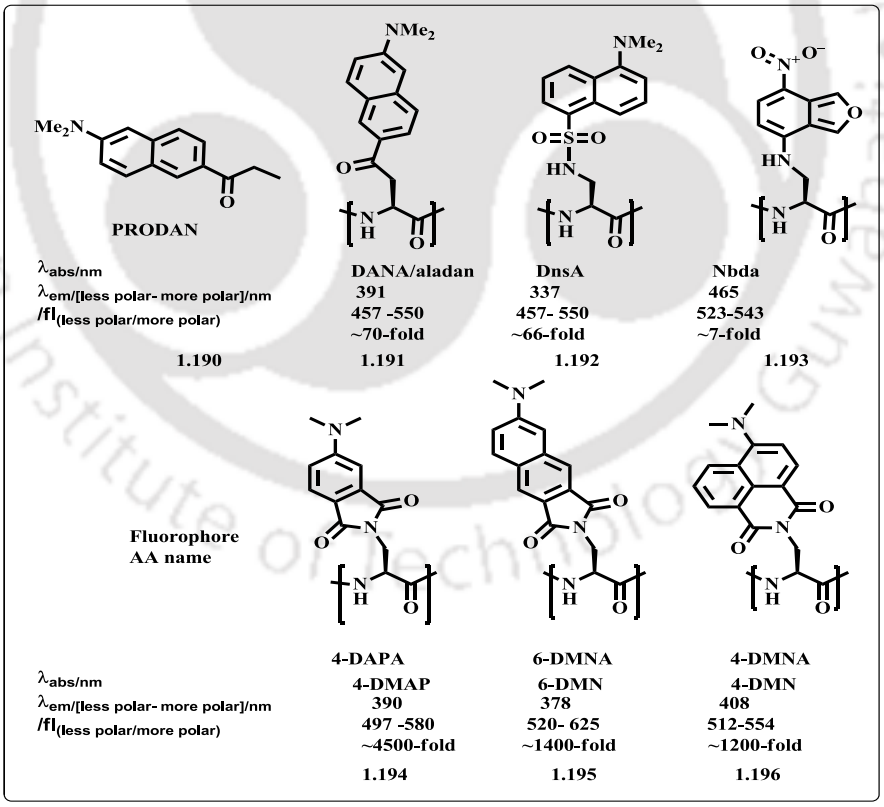


Figure 1.28. Some environment-sensitive fluorophores and their solvatochromatic properties

Impariali *et al*., [58] have synthesised the fluorescent amino acid, 4-(*N*,*N*-dimethylamino)-phthalimide propionic acid (4-DAPA) containing 4-(*N*,*N*-dimethylamino)-phthalimide (4-DMAP, **1.194, Figure 1.28)** and it was found to have the advantage of being a mimic of tyrosine due to closer size. They have also developed a series of novel fluorogenic probes containing solvatochromic amino acids 6-*N*,*N*-dimethylamino-2-3-naphthalimidoalanine (6-DMNA, **1.195, Figure 1.28**) and 4-*N*,*N*-dimethylaminophthalimidoalanine (4-DAPA). Upon binding to class II MHC proteins these fluorophores show large changes in emission spectra, quantum yield and fluorescence lifetime. They have also prepared the same series of peptide with the amino acids 4-DMAP, 6-DMN, BADAN,[58] dansyl and NBD (**Figure 1.29**). The microenvironment sensitivity of these fluorophores makes them perfect tools for monitoring structural dynamics in proteins.

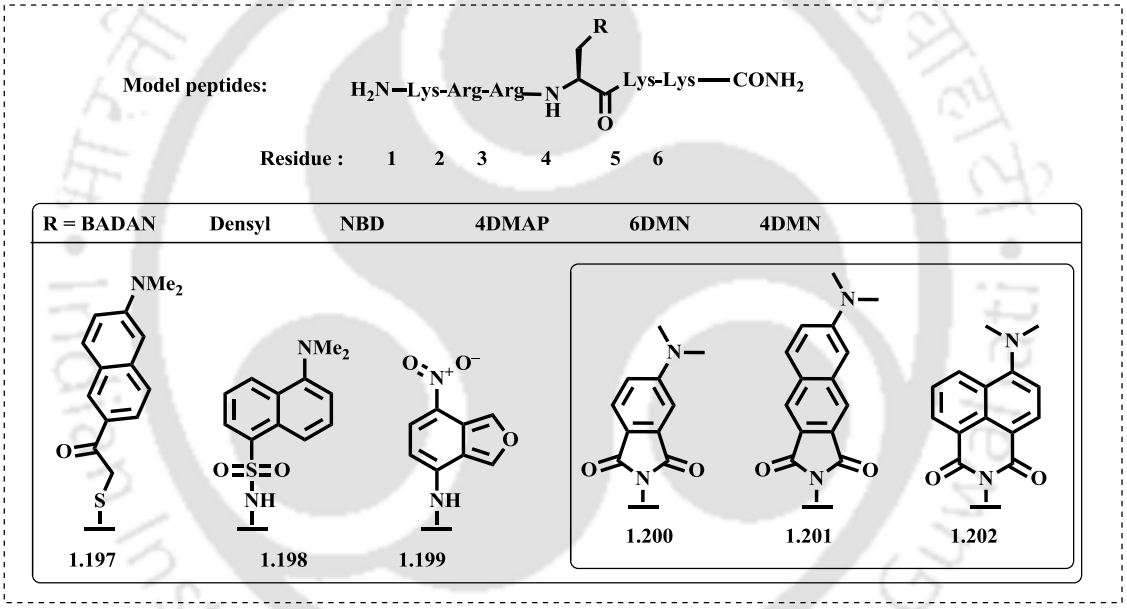


Figure 1.29. Solvatochromic amino acids incorporated model peptide series.

Wang *et al*.; have synthesised styryl conjugated tyrosine analogues, a new class of fluorescent UNAAs with excellent optical tunability in emission spectra (**1.203-1.204**, **Figure 1.30**). These novel fluorescent UNAAs consist of stilbene and meta-phenylene vinylene units as fluorophores have wide range of emission color from blue to near IR. They have used these UNAAs to prepare cell-penetrating peptide by using in solid-phase peptide synthesis (SPPS) and demonstrated fluorescent peptide for cell imaging.[59]

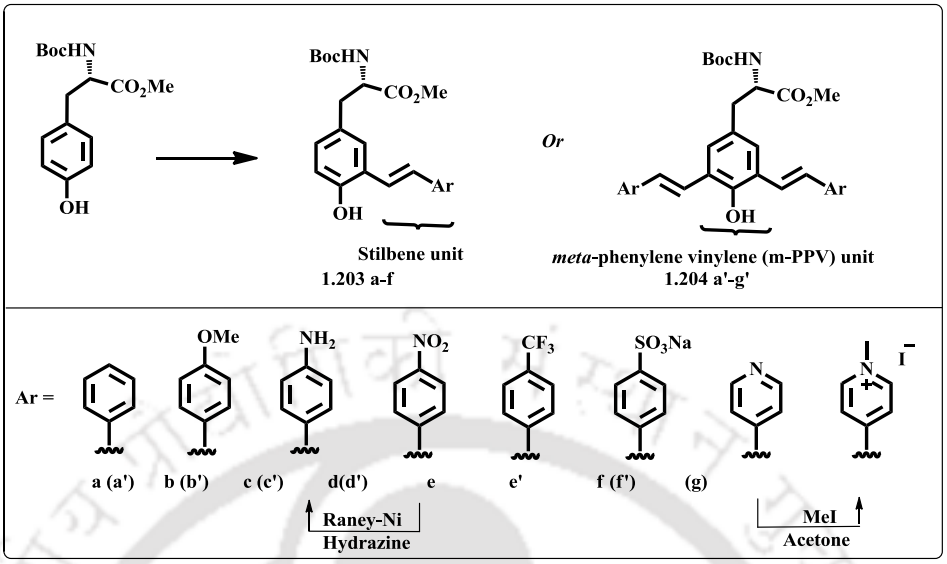


Figure 1.30. Fluorescent UNAAs with extending the π-conjugation of aromatic side chain.

In 2008, Katritzky *et al*., have reported [60a,b] the syntheses of coumarin-labeled amino acids **1.205–1.208 (Figure 1.31)** and dipeptides **1.209 (Figure 1.31)** which afford enantiomerically pure fluorescent building blocks suitable for utilising benzotriazole activated fluorogenic agents as coupling reagents.

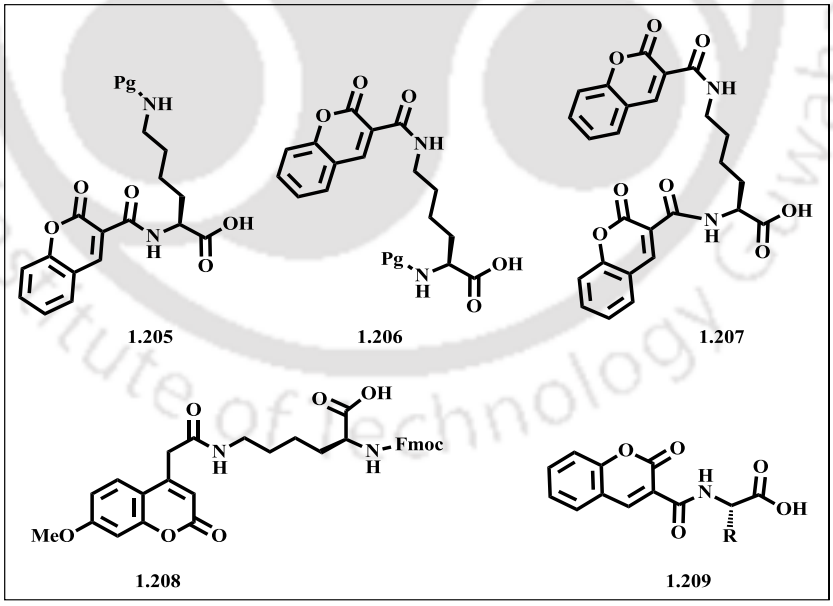


Figure 1.31. Some example of coumarin-labeled amino acids.

In 2008, Katritzky *et al.*,[60c-d] have also designed and synthesised a protected monosaccharide containing Fmoc-L-lysine scaffold based on fluorescent amino acid building blocks **1.210–1.213** (**Figure 1.32**).[60c-d] After deprotection of **1.210–1.123** provide water soluble organic fluorophores **1.124** which could be useful for peptide labeling at the C-terminus in inverse solid phase peptide synthesis (SPPS).

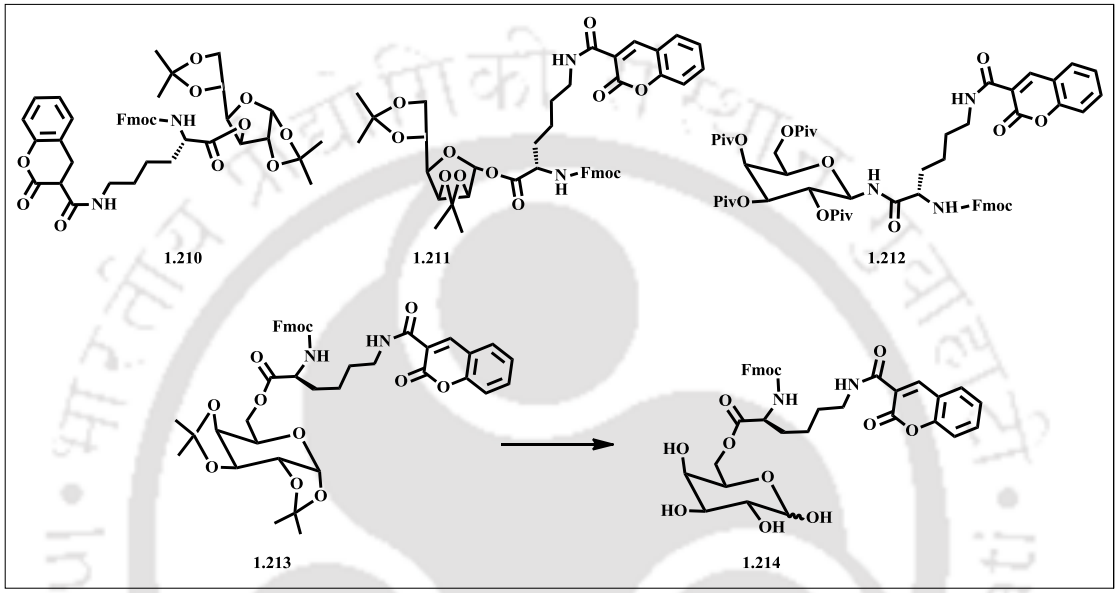


Figure 1.32. Monosaccharide containing Fmoc-L-lysine scaffold based fluorescent amino acid.

Other polarity-sensitive fluorescent unnatural amino acids L-Anap (**1.218**, **Figure 1.33**) has been reported and genetically encoded by Wang *et. al*., to image protein modifications and activities in mammalian cells.[61]

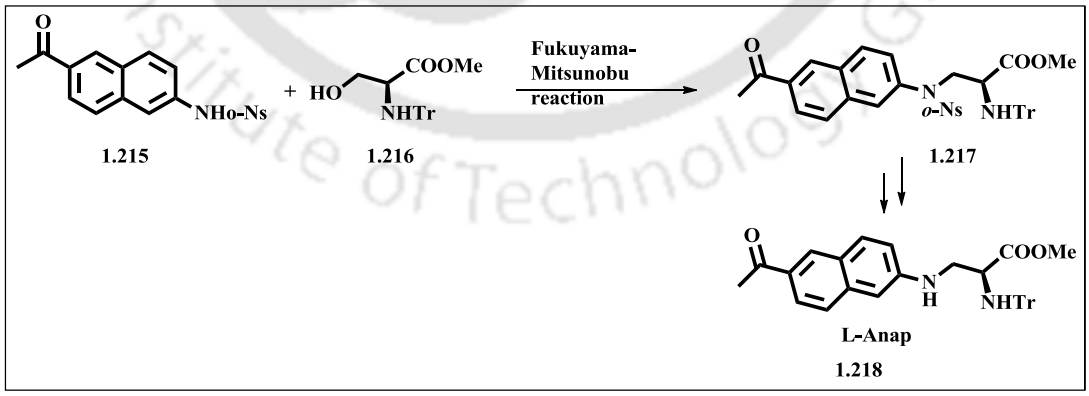


Figure 1.33. Polarity-sensitive fluorescent unnatural amino acids L-Anap.

Sisido *et al*., [62] have developed a modern technology for screening fluorescent peptides containing multiple-labeled fluorescent amino acids based on their protein binding specificity. Thus, they synthesised 8-mer peptides modified with several fluorescent unnatural amino acids (**1.219-1.224**, **Figure 1.34**). Utilising a combination of fluorescence analysis with gel filtration method they were able to quantify the binding peptides simultaneously.[62]

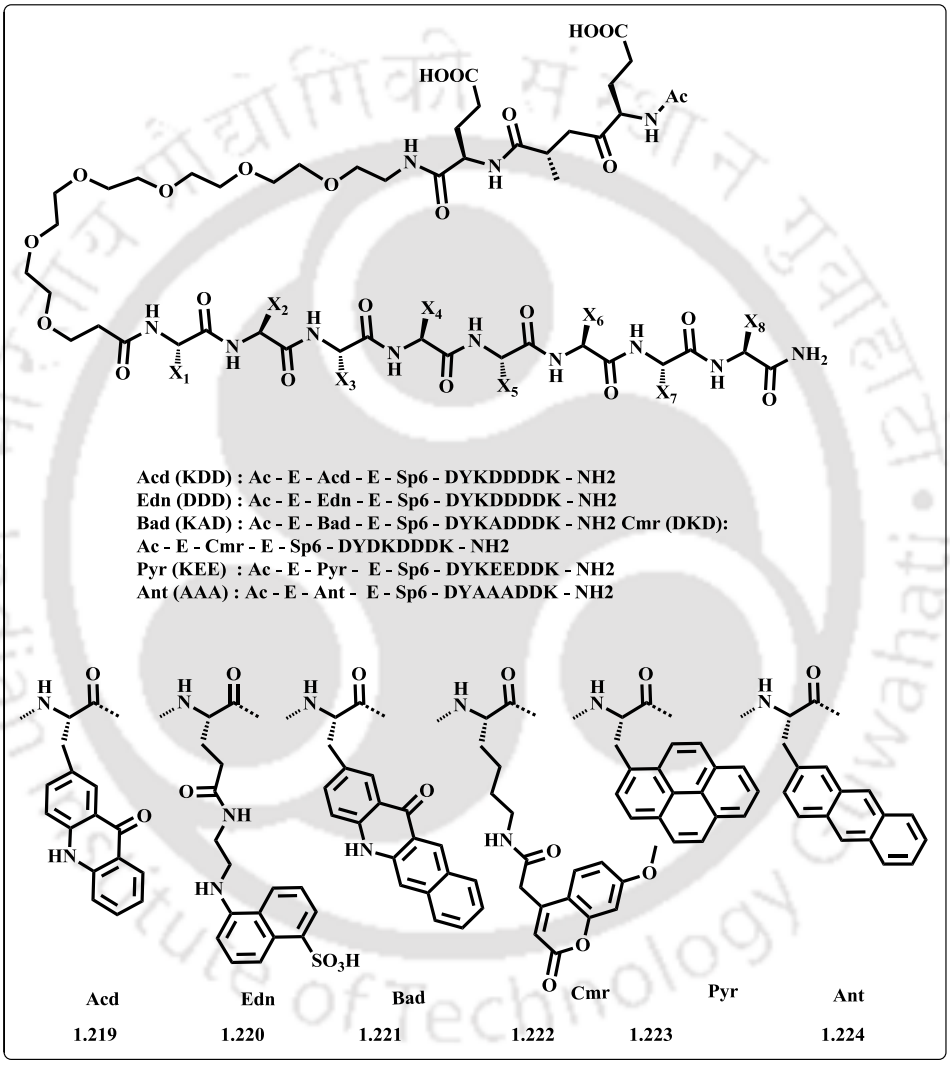


Figure 1.34. Structure of peptides modified with fluorescent amino acids.

1.6. 2. Application of Unnatural Amino Acids in Sensing Metal Ions

Metal-binding unnatural amino acids (UNAAs) are used in protein research for various applications such as protein sensors, for metalloproteins,[63] and for protein NMR spectroscopy.[64] Recent research in the *de novo* design of metalloproteins are based on the collection of polypeptides with defined structural and potentially functional properties.[65,66] The formation of peptidyl motifs designed to sense Zn(II) in solution highlights the significance of tuning the metal binding affinity of such concepts.[67] The ability to modify metal binding affinity is very crucial in order to prepare such peptide-based sensors over a wide range of concentration of metal cations. Additionally, it is essential for the peptide to bind selectively the metal cation of interest to show activity of such peptides in sensory applications.[68] So, the construction of these devices involves the use of novel multidentate metal-binding residues.[69-74] By the utilisation of such modified amino acid residues in small peptidyl unit, one can enhance the metal binding affinities. On the other hand nature uses the protein architecture within metalloprotein systems elegantly for the metal-binding with the help of side chain functionalities or adjustment in protein structure.[68]

As we know amino acids and peptides both have the ability to complex with metal ions through nitrogen, oxygen and sulfur donor atoms at the main side chain, one can design amino acids and peptides that coordinate with metals. Such work has potential applications for the study of protein-protein interactions mediated by metals, protein binding to nanoparticles and metal surfaces, and the development of selective chemosensors for metals to use *in vivo* and *in vitro***.**

In 1996, Imperiali *et al*.,[68] described the R-amino acids containing dominant bidentate ligands (bipyridyl and phenanthrolyl groups). As for examples 6-, 5-, and 4-substituted (*S*)-2-amino-3-(2,2′-bipyridyl)propanoic acids (**1.225, 6Bpa; 1.226, 5Bpa; 1.227**, **4Bpa**, **Figure 1.35**), (*S*)-2-amino-3-(1,10-phenanthrol-2-yl)-propanoic acid (**Fen, 1.228**, **Figure 1.35**), and (*S*)-2-amino-3-(9-methyl-1,10-phenanthrol-2-yl)propanoic acid (**1.229, Neo**) show differential response in presence of various metal ions due to the special orientation of the ligand ring system (2,2-bipyridyl or 1,10-phenanthrolyl). These bidentate unnatural side chain residues with variable metal-binding abilities can be incorporated into small peptides for further studies.[68]

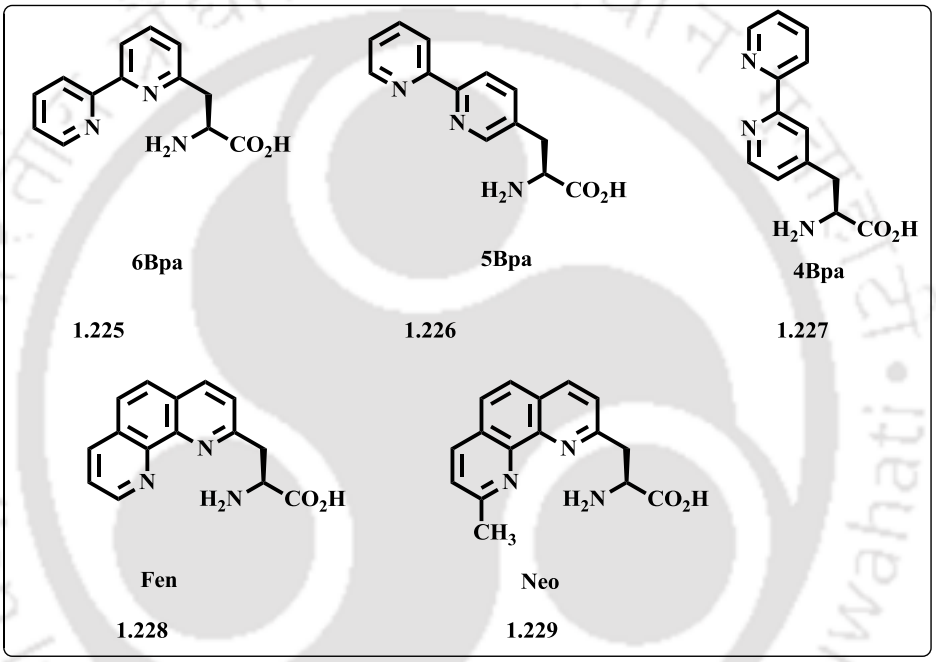


Figure 1.35. Structure of bipyridyl and phenanthrolyl fluorescent unnatural amino acids.

In the same year 1996, Imperiali *et al*., also synthesised three new amino acids (**1.230, 1.231 and 1.232, Figure 1.36)** which showed good metal-coordination property and fluorescence quenching properties respectively. They have synthesised three amino acid based peptide building blocks *via* solid phase synthesis (SPPS) that showed metal-dependent photo induced electron transfer (PET). Fluorescence quenching was observed upon titration with Zn^{+2} and Co^{+2}.[76]

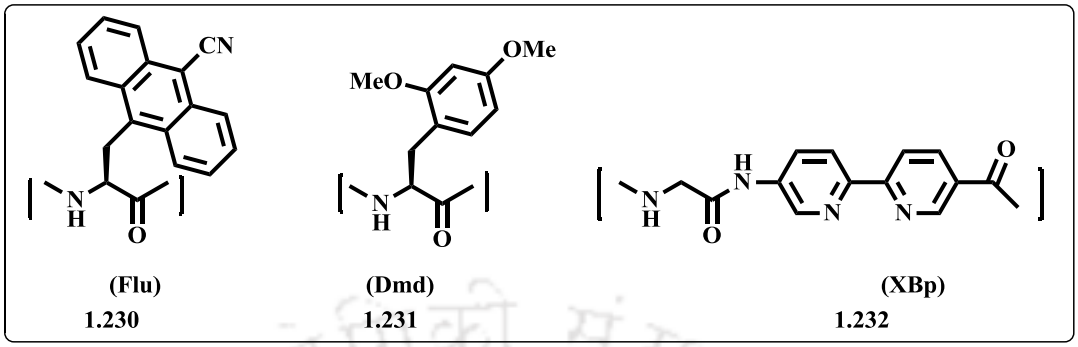


Figure 1.36. Structure of donor/acceptor fluorescent unnatural amino acids.

On the other hand, they incorporated those UNAAs **(1.230-1.232, Figure 1.36**) in small peptide unit and have shown the conformational changes (**1.233-1.234**, **Figure 1.37**) in presence of complexing metal ions which was monitored by using photo electron transfer (PET).[76a-d] The designed peptide based sensors are shown in **Table 1.**

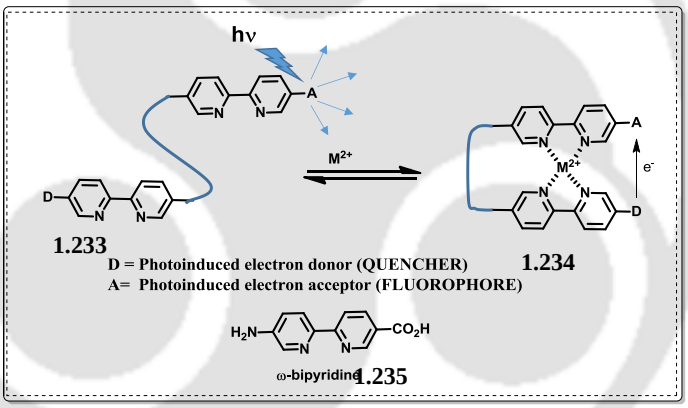


Figure 1.37. Schematic presentation of photo electron transfer (PET).

Peptide	Sequences
01	AcFluXBpGlyGlyGlyXBpDmdGlyNH2
02	AcFluXBpGlyGlyGlyXBpDmdThrArgNH2
03	AcArgFluXBpGlySerSerXBpDmdThrArgNH2

Table 1. Sequence of peptide building block modified with chelating ligand

In 1998, Imperiali *et al*., reported the synthesis of (*S*)-2-amino-*N*R-9-fluorenyl methoxycarbonyl-3-(oxine-2-yl) propionic acid (Fmoc-2Oxn-OH, **1.236, Figure 1.39**) and (*S*)-2-amino-*N*R-9fluorenylmethoxycarbonyl-3-(oxine-5-yl) propionic acid (Fmoc-5Oxn-OH, **1.237, Figure 1.39**). They have utilised these amino acids in short peptides which were found to capable of sensing sub-micromolar levels of Zn(II).[77]

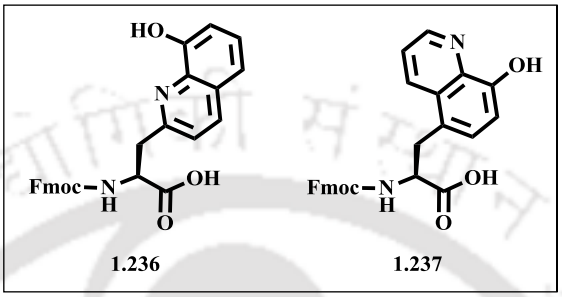


Figure 1.39. Example Zn (II) metal ion sensing fluorescent unnatural amino acids.

In 2011, Lee *et al*., have reported dansyl fluorophore containing UNAAs which showed hypersensitive turn on response to Hg^{2+} ions in aqueous solution. The amino acids **1.238-1.240 (Figure 1.40)** showed sensitivity towards Hg^{2+} ions whereas **1.241-1.243** did not show any response to various metal ions. Furthermore, amino acids **1.240-1.241** were able to detect Hg^{2+} ions in aqueous solution without interferences by any other metal ions. Thus, probes were efficient for detecting the maximum allowable level (2 ppb) of mercury ions in drinking water. It has also been shown that the chain length between two dansyl sulfonamide parts is not enough for discrimination of Hg^{2+} ions but plays vital role in binding affinity for Hg^{2+} ions.[78]

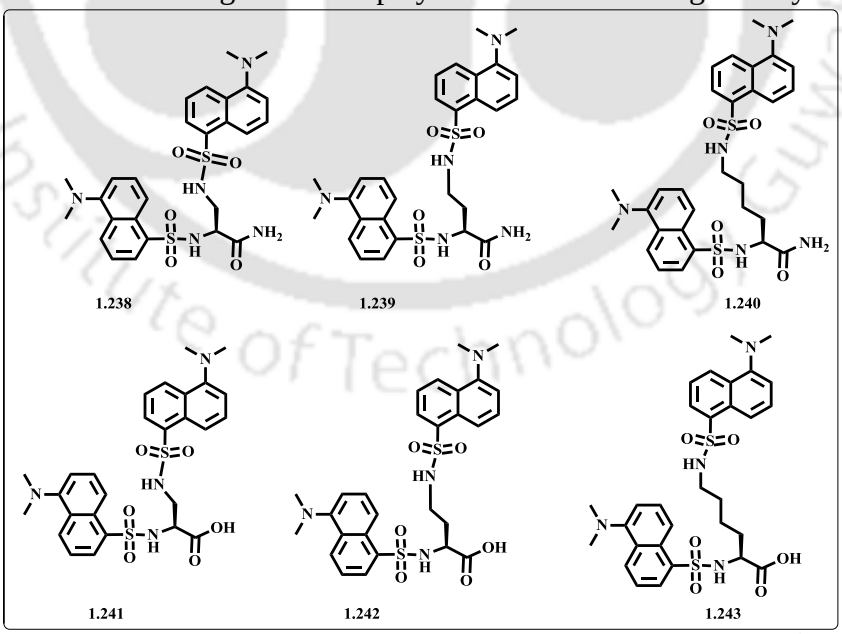


Figure 1.40. Dansyl alanine amino acids derivatives utilised for Hg^{2+} detection.

1.7. Summary and Future Prospects

Amino acids are essential components of life system, protein functions and protein stability. So, design and synthesis of unnatural amino acids is very important area in expansion of genetic code. Much research has been done to develop protocols for the genetic incorporation of such modified amino acids into a protein system in *Eukaryotic cells*. Currently many laboratories are dedicated in the generation of new orthogonal aaRS/tRNA pairs for the efficient encoding of additional unnatural amino acids. However incorporation of such amino acids in complex system needs more research and vast knowledge for the generation of new aaRS/tRNA pairs which may also allow the *in vivo* synthesis of biopolymers with unnatural backbones. The incorporation of unnatural amino acid into protein not only generates protein based therapeutics but also it may provide stability in protein structure which is important in protein based drug delivery.

The incorporation of UNAAs by post translational modification and chemical protein modifications are an important area in protein research. Protein conformational variations and dynamics can also be observed at particular position of interest with negligible or no perturbation of sequence by applying these technique. In addition, site-selective protein modification will play a very key role for the preparation of the next generation of biopharmaceuticals, including ADCs for cancer therapy and glycoproteins for vaccination. Such methods will lead to develop concepts and also will open new molecular division in protein as therapeutic function with better safety and effectiveness.

On the other hand the design of fluorescent unnatural amino acids (FUAA) is very attractive area in current ongoing research. These novel fluorescent amino acids and peptides can be generated from synthetic management at the side chain of naturally occurring amino acids. These amino acids or peptide could possess improved physicochemical or photophysical properties. They can not only be utilised for monitoring protein's microenvironment, conformation, and dynamics of proteins, protein-DNA interactions but also can be used for the cellular imaging and dynamics. Fluorescent unnatural amino acids are also very important tool for sensing metal ions like Hg (II), Zn (II) .Therefore, with the help of advance knowledge and strategies in chemistry and biology, many side chain modified UNAAs and FUAAs can be generated for the application in this direction. Thus, such research activity will help all researchers to understand the fundamental questions about biological process utilising UNNAs.

1.8. References

1. (a) Wu, G. *Amino Acids*. **2009**, *37*, 17. (b) http://www.encyclopedia.com/science-and-technology/biochemistry/biochemistry/protein (c) Wang, A.; Nairn, W.N.; Marelli, M.; Grabstein, K. **2012**. Protein Engineering with Non-Natural Amino Acids, Protein Engineering, Prof. Pravin Kaumaya (Ed.), ISBN: 978-953-51-0037-9 (d) Zhang, W.H.; Otting, G.; Jackson, C. *J. Curr Opin Struct Biol*. **2013**, *23*, 581.(e) Zhang,Y.; Gladyshev, V, N. *Nucleic Acids Res*. **2007,** 35,4952.
2. (a)Tarfah, I.; Hassan, M.A.; Al-Hazimi b.; Ayman El-Faham b,c. *J. Saudi Chem. Soc*.**2012**, *16*, 97. (b)Wright, G.; Carver, A.; Cottom, D.; Reeves, D.; Scott, A.; Simons, P.; Wilmut, I.; Garner, I.; Colman, A. *Biotechnol*. **1991**, *9*, 830. (c) Tobin, P.H.; Richards, D.H.; Callender, R.A.; Wilson,C.J. *Curr Drug Meta*b. **2014**, *15*, 743. (c) Ripps, H.; Shen, W. *Mol Vis*. **2012**, *18*, 2673. (d) Brosnan, J.T.; Brosnan, M.E. *J. Nutr*. **2006**, *136*, 1636.
3. (a) van Kasteren, S.I.; Krame, H.B.; Jensen, H.H.; Campbell, S.J.; Kirkpatrick, J.; Oldham, N.J.; Anthony, D.C.; Davis, B.G. *Nature.* **2007**, *446*, 1105. (b) Gokhale, A.S.; Satyanarayanajois,S. *Immunotherapy* , 2014, 6, 755.
4. Chandrudu, S.; Simerska, P.; Toth, I.; *Molecules.* **2013**, *18*, 4373
5. (a)Bröcker, M. J.; Ho, J. M. L.; Church, G.M.; Söll, D.; Donoghue, P. O. *Angew. Chem. Int., Ed*. **2014**, *53*, 319. (b)Federico, K.; Geoffrey, C.; Wieslaw, k. *Trends Biotechnol*. **2005,** *23*, 150.
6. (a)Tarfah, I.; Hassan, M.A.; Al-Hazimi b.; Ayman El-Faham b,c. J. Saudi *Chem. Soc*.**2012**, *16*, 97–116. (b)Wright, G.; Carver, A.; Cottom, D.; Reeves, D.; Scott, A.; Simons, P.; Wilmut, I.; Garner, I.; Colman, A. *Biotechnol*. **1991**, *9*, 830. (c) Tobin, P.H.; Richards,D.H.;, Callender,R.A.; Wilson,C.J. *Curr Drug Meta*b. **2014**, *15*, 743.
7. Wang, N.; Ju, T.; Niu, W.; Guo, J.; *ACS Synth. Biol*., **2015**, *4*, 207.
8. (a) Takahiro, H.; Masahiko, Sisido. *Current Opin. Chem. Biol.* **2002**, *6*, 809. (b)Young, T.S.; Schultz P.G. *J. Biol. Chem.***2010,** *285*, 11039.
9. (a) Guzmán,F.; Barberis, S.; Illanes, A. *Electronic Journal of Biotechnology* , **2007**, *10* . (b) Uhlig, T.; Kyprianoua, T.; Martinelli, F.G.; Oppici, C.A.; Heiligers , D.; Hills, D.; Calvo , X.E.; Verhaert, P. *Eupa open proteomics* .**2014**, *4*, 58.
10. (a) Chen, S.; Gopalakrishnan,R.; Schaer,T.; Marger,F.; Hovius,R.; Daniel Bertrand, D.; Pojer,F.; Heinis,C. *Nat. Chem*. **2014**, *6*, 1009. (b) Nowick, J.S.; Chung, D.M.; Maitra,K.; Maitra, S.; Stigers,K.D.; Sun.Y. *J. Am. Chem. Soc.* **2000,** *122*, 7654.
11. (a) Dobson, C. M. Trends Biochem. Sci. **1999**, *24*, 329. (b) Ross, C.A.; Poirier, M.A. *Nat Med.* **2004**, *10*, S10. (c) Ignatova, Z. Microb Cell Fact. **2005**, *4*, 23.
12. (a)Thirumurugan, P.; Matosiuk, D.; Jozwiak, K. *Chem. Rev.* **2013**, *113*, 4905 (b) Neuweiler, H.; Schultz, A.; Vaiana, A. C.; Smith, J. C.; Kaul, S.; Wolfrum, J.; Sauer, M.

Angew. Chem., Int. Ed., **2002**, *41*, 4769. (c) Kohn, J. E.; Plaxco, K. W. *Proc. Natl. Acad. Sci. U.S.A*. **2005**, *102*, 10841. (d) Chen, S. X.; Wang, L.; Fahmi, N. E.; Benkovic, S. J.; Hecht, S. M. *J. Am. Chem. Soc*. **2012**, *134*, 18883. (e) Fujii, A.; Sekiguchi, Y.; Matsumura, H.; Inoue, T.; Chung, W. S.; Hirota, S.; Matsuo, T. *Bioconjugate Chem.* **2015**, *26*, 537. (f) Tamura, T.; Hamachi,T. *ACS Chem. Biol*. **2014**, *9*, 2708.

13. (a) Skrynnikov, N. R.; Mulder, F. A. A.; Hon, B.; Dahlquist, F. W.; Kay, L. E. *J. Am. Chem. Soc.* **2001**, *123*, 4556. (b) Cellitti, S.E.; Jones, D.H.; Lagpacan, L.; Hao, X.; Zhang, Q.; Hu,H.; Brittain, S.M.; Brinker,A.; Caldwell,J.; Bursulaya,B.; Spraggon,G.; Brock,A.; Ryu,Y.; Uno,T.; Schultz, P.G.; Geierstanger, B.H. *J. Am. Chem. Soc*. **2008**, *130*, 9268.(c) Li, L.; Bergen,M.V.; Mandelkow, E.-M.; Mandelkow, E. *J. Biol. Chem*. **2002** , *277*, 41390. (d) Jones, D.H.; Cellitti, S.E.; Hao, X. *J Biomol NMR*. **2010**, *46*, 89.
14. (a) Ge, H; *Nucleic Acids Res*. **2000**, *28*, 2.(b) Kang, J.-Y.; Kawaguchi, D.; Coin, I.; Xiang, Z.; O'Leary, D. D. M.; Slesinger, P. A.; Wang, L. *Neuro.* **2013**, *80*, 2. (c) Truong, K.; Ikura, M. *Cur Opin Struc Biol***. 2001**,*11*, 573. (d) Kolodkin-Ga, I.; Romero, D.; Cao, S.; Clardy, J.; Kolter, R.; Losick, R.; *Science* , **2010**, *328*, 627 (e) Goodman, M.; Kossoy, A. *J. Am. Chem. Soc*. 1966, 88, 5010. Wang,Y.; Paletta, J. T.; Berg, K.; Reinhart, E.; Rajca, S.; Rajca, A. *Org. Lett*. **2014**, *16*, 5298. (f) Mojsoska, B.; Jenssen, H. *Pharmaceuticals*. **2015**, *8*, 366. (g) Liu, X.; Li, J.; Hu, C.; Zhou, Q.; Zhang, W.; Hu, M.; Zhou, J.; Wang, J. *Angew. Chem. Int. Ed.* **2013**, *52*, 4805. (h) Ngounou Wetie, A.G.; Sokolowska, I.; Woods, A.G.; Roy, U.; Loo, J.A.; Darie, C.C. Proteomics. 2013 13, 1.
15. (a) Kaiser, E.T. *Acc Chem Res*, **1989**, *22*, 47. (b) Kaiser, E.T.; Lawrence D.S.; Rokita, S.E. *Rev Biochem*, **1985**, *54*, 565. (c) Kaiser, E.T.; Lawrence, D.S. *Science*, **1984**, *226*, 505. (d) Offord, R.E. Protein Eniz. **1987**, *3*, 151.
16. (a) Chen, P. R. *J. Am. Chem. Soc*. **2011**, *133*, 20581. (b) Wu, N.; Deiters, A.; Cropp, T.A.; King, D.; Schultz, P.G.; *J. Am. Chem. Soc*. **2004**, *126*, 14306.(c) *Nature Methods*. **2007**, *4*, 239. (d) Wang, L.; Xie, J.; Schultz P.G. *Annu Rev Biophys Biomol Struct*. **2006**, *35*, 225. (e) Wang,L.; Schultz, P.G. *Angew. Chem. Int., Ed.* **2005**, *44*, 34. (f) Chin, J.W.; Cropp, A.T.; Anderson, J.C.; Mukherji, M.; Zhang, Z.; Schultz,P.G. Science. **2003**, *301*, 964. (g) Deiters, A.; Groff, D.; Ryu, Y.; Xie, J.; Schultz, P. G. Angew. Chem., Int. Ed. 2006, 45, 2728. (h) Summerer, D.; Chen, S.; Wu, N.; Deiters, A.; Chin, J. W.; Schultz, P. G. *Natl. Acad. Sci. U.S.A* .**2006**, *103*, 9785. (i)Tsao, M. L.; Summerer, D.; Ryu, Y.; Schultz, P. G. *J. Am. Chem. Soc*. **2006**,*128*, 4572. (j) Deiters, A., Geierstanger, B. H., and Schultz, P. G. *ChemBioChem*.**2005**, *6*, 55. (k) Xie, J.; Wang, L.; Wu, N.; Brock, A.; Spraggon, G.; Schultz, P. G. *Nat. Biotechnol*. **2004**, 22, 1297.(l)Alfonta, L.; Zhang, Z.; Uryu, S.; Loo, J. A.; Schultz, P. G. *J. Am. Chem. Soc*. **2003**, 125, 14662. **(**m**)** Suydam, I. T.; Boxer, S. G. *Biochemistry*.**2003**, 42, 12050. (n) Connor, R.E.; Tirrell, D.A. *Polymer Reviews*. **2007**, *47*, 9.(o) Lang, K.; Chin, J.W**.** *Chem. Rev***. 2014,** *114*, 4764. (d) Wang, Q.; Parrish, A.R.; Wang, L. Chemistry & Biology.**2009**, *16*, 323.(p) Neumann, H.; Wang,

K.; Davis, L.; Garcia-Alai, M.; Chin,J.W. Nature. **2010**, *464*, 18. (Q) Liu, D. R.; Schultz, P.G.; *Proc. Natl. Acad. Sci. USA.* **1999**, 96, 4780.

17. Boutureira,O.; Bernardes, G.- J. L. *Chem. Rev*. **2015**, *115*, 2174.
18. Taira, H.; Fukushima, M.; Hohsaka, T.; Sisido, M. *J. Biosci. Bioeng*. **2005**, 99, 473.
19. Hohsaka,T.; Ashizuka,Y.; Murakami,H.; Sisido,M. *Nucleic Acids. Res*. **2001**, *29*, 3646.
20. (a)Getahun, Z.; Huang, C.-Y.; Wang, T.; Leo´, B.D.; DeGrado, W.F.; Gai, F. *J. Am. Chem. Soc*. **2003**, *125*, 405. (b) Waegele M.M.; Tucker, M.J.; Gai, F. *Chem Phys Lett*. **2009**, *478*, 249. (c) Jo, H.; Culik, R. M.; Korendovych, I. V.; DeGrado, W. F.; Gai, F. *Biochemistry* **2010**, *49*, 10354. (d)Ye, S.; Zaitseva, E.; Caltabiano, G.; Schertler, G. F. X.; Sakmar, T. P.; Deupi, X.; Vogel, R. *Nature.* **2010**, *464*, 1386. (e) Kim ,H.; Cho,M. *Chem. Rev*. **2013**, *113*, 5817.
21. (a)Young,T.S.; Schultz, P. G. J. *Biol. Chem. **2010**, 285*, 11039. (b) Stephanopoulos, N.; Francis, M. B. *Nat. Chem. Biol*. **2011**, *7*, 876. (c) Boutureira,O.; Bernardes, G.- J. L. *Chem. Rev*. **2015**, *115*, 2174. (d) Basle, E.; Joubert, N.; Pucheault, M. Chemistry *&* *Biology*. **2010**, *17*, 213.
22. Walsh, C. T.; Garneau-Tsodikova, S.; Gregory J. Gatto, J. *Angew.Chem., Int. Ed*. **2005**, *44*, 7342.
23. Kee, J.-M.; Muir, T. W. *ACS Chem. Biol*. **2011**, *7*, 44.
24. Bernardes, G. J. L.; Chalker, J. M.; Davis, B. G. Chemical Protein Modification; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2010.
25. Harris, J. M.; Chess, R. B. *Nature Reviews*. **2003**, *2*, 214.
26. (a) Sun, S. B.; Schultz, P. G.; Kim, C. H. *Chem. Bio. Chem*. **2014**, *15*, 1721. (b) Pelegri-O'Day, E. M.; Lin, E.-W.; Maynard, H. D. *J. Am. Chem. Soc.* **2014**, *136*, 14323. (c) Chari, R. V. J.; Miller, M. L.; Widdison, W. C. *Angew. Chem., Int. Ed*. **2014**, *53*, 3796. (d) Schumacher, D.; Hackenberger, C. P. R. *Curr. Opin. Chem. Biol*. **2014**, *22*, 62. (e) Steiner, M.; Hartmann, I.; Perrino, E.; Casi, G.; Brighton, S.; Jelesarov, I.; Bernardes, G. J. L.; Neri, D. *Chem. Sci*. **2013**, *4*, 297.
27. Shen, B.-Q.; Xu, K.; Liu, L.; Raab, H.; Bhakta, S.; Kenrick, M.; Parsons-Reponte, K. L.; Tien, J.; Yu, S.-F.; Mai, E.; Li, D.; Tibbitts, J.; Baudys, J.; Saad, O. M.; Scales, S. J.; McDonald, P. J.; Hass, P. E.;Eigenbrot, C.; Nguyen, T.; Solis, W. A.; Fuji, R. N.; Flagella, K. M.;Patel, D.; Spencer, S. D.; Khawli, L. A.; Ebens, A.; Wong, W. L.; Vandlen, R.; Kaur, S.; Sliwkowski, M. X.; Scheller, R. H.; Polakis, P.; Junutula, J. R. *Nat. Biotechnol*. **2012**, *30*, 184. (e) Steiner, M.; Hartmann, I.; Perrino, E.; Casi, G.; Brighton, S.; Jelesarov, I.; Bernardes, G. J. L.; Neri, D. *Chem. Sci*. **2013**, *4*, 297. (a) Sun, S. B.; Schultz, P. G.; Kim, C. H. *Chem. Bio. Chem*. **2014**, *15*, 1721.
28. Padmanabhan, S.; Baldwin, R. L. *Protein Sci*. **1994**, *3*, 1992.

29. (a)Yeo, D.J.; Warriner, Stuart L.; Wilson, A. *J. Chem. Commun*. **2013**, *49***,** 9131.(b) Grison,C.M.; Burslem, G.M.; Miles, J.A.; Pilsl, L. K. A.; Yeo, D.J.; Imaniab, Z.; Warriner,S.L.; Webb, M.E.; Wilson,A.J. *Chem. Sci*. **2017**,
30. (a)Chen, X.; Muthoosamy, K.; Pfisterer, A.; Neumann, B.; Weil, T. *Bioconjugate Chem*. **2012**, *23*, 500. (b) Klykov,O.; Weller,M.G.; *Anal. Methods*. **2015**, *7*, 6443.
31. (a)Hooker, J.M.; Esser-Kahn, A.P.; and Francis, M.B. *J. Am. Chem. Soc.* **2006**, *128*, 15558. (b)Jentoft, N.; Dearborn, D.G. *J. Biol. Chem*. **1979,** *254*, 4359. (c)Tanaka, K.; Fukase, K.; Katsumura, S. *Synlett*. **2011**, 2115.(d)Tanaka, K.; Kamatani, M.; Mori, H.; Fujii, S.; Ikeda, K.; Hisada, M.; Itagaki, Y.; Katsumura, S. *Tetrahedron*.**1999**, *55*, 1657. (e) Tanaka, Y.; Bond, M.R.; Kohler, J. *J. Mol. Biosyst*. **200**8, *4*, 473. (f) Robinson, M. A.; Charlton, S. T.; Garnier, P.; Wang, X.-T.;Davis, S. S.; Perkins, A. C.; Frier, M.; Duncan, R.; Savage, T. J.; Wyatt, D. A.; Watson, S. A.; Davis, B. G. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 14527.
32. (a)Hooker, J.M.; Kovacs, E.W.; Francis, M.B. *J. Am. Chem. Soc*. **2004**, *126*, 3718. (b) Schlick, T.L.; Ding, Z.; Kovacs, E.W.; Francis, M.B. *J. Am. Chem. Soc*. **2005**, *127*, 3718. (c) Joshi, N.S.; Whitaker, L.R.; Francis, M.B. *J. Am. Chem.Soc*. **2004**, *126*, 15942. (d)McFarland, J.M.; Joshi, N.S.; Francis, M.B. *J. Am. Chem. Soc*.**2008** *130*, 7639. (e) Guo, H.-M.; Minakawa, M.; Ueno, L.; Tanaka, F. *Bioorg. Med. Chem. Lett*. **2009**, *19*, 1210.
33. (a)Bernardes, G.J.; Chalker, J.M.; Errey, J.C.; Davis, B.G. *J. Am. Chem. Soc.* **2008**, *130*, 5052. (b) Chalker, J.M.; Bernardes, G.J.; Lin, Y.A.; Davis, B.G. *Chem. Asian J*. **2009**, *4*, 630. (c)Kim, Y.; Ho, S.O.; Gassman, N.R.; Korlann, Y.; Landorf, E.V.; Collart, F.R.; Weiss, S. *Bioconjug. Chem*. **2008**, *19*, 786. (d) Gundlach, H.G.; Moore, S.; Stein, W.H. *J. Biol. Chem*. **1959**, *234*, 1761. (e) Chalker, J.M.; Lin, Y.A.; Boutureira, O.; Davis, B.G. *Chem. Commun*. 2009, 3714. (f) Rademann, J. *Angew. Chem. Int. Ed. Engl*. **2004**, *43*, 4554. (g) Crich, D.; Krishnamurthy, V.; Hutton, T.K. *J. Am. Chem. Soc.* **2006**, *128*, 2544. (h) Crich, D.; Krishnamurthy, V.; Brebion, F.; Karatholuvhu, M.; Subramanian, V.; Hutton, T.K. *J. Am. Chem. Soc*. **2007**, *129*, 10282. (i) Arumugam, S.; Guo, J.; Mbua, N. E.; Friscourt, F.; Lin, N.; Nekongo, E.; Boons, G.-J.; Popik, V. V. Chem. Sci. 2014, 5, 1591. (j) Shiu, H.Y., Chan, T.C., Ho, C.M., Liu, Y., Wong, M.K., and Che, C.M. *Chemistry*.**2009**, *15*, 3839. (k) Toda, N.; Asano, S.; Barbas, C. F. *Angew. Chem., Int. Ed.* **2013**, *52*, 12592. (l) Abbas, A.; Xing, B.; Loh, T.-P. *Angew. Chem., Int. Ed*. **2014**, *53*, 7491.
34. (a) Oya, T.; Hattori, N.; Mizuno, Y.; Miyata, S.; Maeda, S.; Osawa, T.; Uchida, K. *J. Biol. Chem*. **1999**, *274*, 18492.
35. (a) Chen, G.; Heim, A.; Riether, D.; Yee, D.; Milgrom, Y.; Gawinowicz, M.A.; Sames, D. *J. Am. Chem. Soc*. **2003**, *125*, 8130. (b) Scheck, R.A.; Dedeo, M.T.; Iavarone, A.T.;

Francis, M.B. *J. Am. Chem. Soc*. **2008**, *130*, 11762.(c) Scheck, R.A.; Francis, M.B. *ACS Chem. Biol*. **2007**, *2*, 247.

36. (a) Gilles, M.A.; Hudson, A.Q.; Borders, C.L. *Anal. Biochem*. **1990**, *184*, 244. (b) Schlick, T.L., Ding, Z., Kovacs, E.W.; Francis, M.B. *J. Am. Chem. Soc*. **2005**, 127, 3718.
37. (a)Brocchini, S.; Godwin, A.; Balan, S.; Choi, J.-W.; Zloh, M.; Shaunak, S. *Adv. Drug Delivery Rev*. **2008**, *60*, 3.(b) Brocchini, S.; Godwin, A.; Balan, S.; Choi, J.-W.; Zloh, M.; Shaunak, S. *Adv. Drug Delivery Rev*. **2008**, *60*, 3.
38. (a) Bryden, F.; Maruani, A.; Savoie, H.; Chudasama, V.; Smith, M. E. B.; Caddick, S.; Boyle, R. W. *Bioconjugate Chem*. **2014**, *25*, 611. (b) Jones, M. W.; Strickland, R. A.; Schumacher, F. F.; Caddick, S.; Baker, J. R.; Gibson, M. I.; Haddleton, D. M. *J. Am. Chem. Soc*. **2012**, *134*, 1847.
39. (a) Jones, M. W.; Strickland, R. A.; Schumacher, F. F.; Caddick, S.; Baker, J. R.; Gibson, M. I.; Haddleton, D. M. *Chem. Commun*. **2012**, *48*, 4064. (b) Schumacher, F. F.; Nunes, J. P. M.; Maruani, A.; Chudasama, V.; Smith, M. E. B.; Chester, K. A.; Baker, J. R.; Caddick, S. *Org. Biomol. Chem*. **2014**, *12*, 7261.
40. Marculescu, C.; Kossen, H.; Morgan, R. E.; Mayer, P.; Fletcher,S. A.; Tolner, B.; Chester, K. A.; Jones, L. H.; Baker, J. R. *Chem.Commun*. **2014**, *50*, 7139.
41. Castañeda, L.; Maruani, A.; Schumacher, F. F.; Miranda, E.;Chudasama, V.; Chester, K. A.; Baker, J. R.; Smith, M. E. B.; Caddick, S. *Chem. Commun*. **2013**, *49*, 8187.
42. (a) Brunner, J. *Annu. Rev. Biochem*. **1993**, *62*, 483. (b)Knowles, J.R. *Acc. Chem. Res*. **1972**, *5*, 155. (c) Singh, A., Thornton, E.R., Westheimer, F.H. *J. Biol. Chem*. **1962**, *237*, 3006.(d) Hashimoto, M.; Hatanaka, Y. *Eur. J. Org. Chem*. **2008**, 2513.
43. Van Hest, J. C. M.; van Delft, F. L. *ChemBioChem*, **2011**, *12*, 1309.
44. (a)Antos, J.M.; Francis, M.B. *Curr. Opin. Chem. Biol***. 2006**, *10*, 253. (b) Antos, J.M.; Francis, M.B. *J. Am. Chem. Soc*. **2004**, *126*, 10256. (c) van Maarseveen, J.H.; Reek, J.N.; Back, J.W. *Angew. Chem. Int. Ed. Engl*. **2006**, *45*, 1841. (d) McFarland, J.M.; Francis, M.B. J. Am. Chem. Soc. **2005**, *127*, 13490.(e) Tilley, S.D.; Francis, M.B. *J. Am. Chem. Soc*.**2006**, *128*, 1080.
45. (a) Liu, Q.; Wang, J.; Boyd, B. J. *Talanta* **2015,** *136*, 114. (b) Pazos, E.; Vázquez, O.; Mascareñas, J. L.; Vázquez, M. E. *Chem. Soc. Rev*. **2009,** *38*, 3348. (c) Tung, C. -H. *Peptide Sci*. **2004**, *76*, 391. (d) Kenworthy, A. K. *Methods* **2001**, *24*, 289. (e) Hoppe, A.; Christensen, K.; Swanson, J. A. *Biophys. J*. **2002**, *83*, 3652. (f) Ozawa, T.; Umezawa, Y. *Supramol. Chem*. **2002**, *14*, 271. (g) Packard, B. Z.; Toptygin, D. D.; Komoriya, A.; Brand, L. *Natl. Acad. Sci. USA* **1996**, *93*, 11640. (h) Bains, G.; Patel, A. B.; Narayanaswami, V. *Biochemistry* **2011**, *16*, 7909. (i) Matsumoto, C.; Hamasaki, K.; Mihara, H.; Ueno, A. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 1857. (k) Chafekar, S. M. *BBA Mol. Bas. Dis*. **2008**, *1782*, 523.

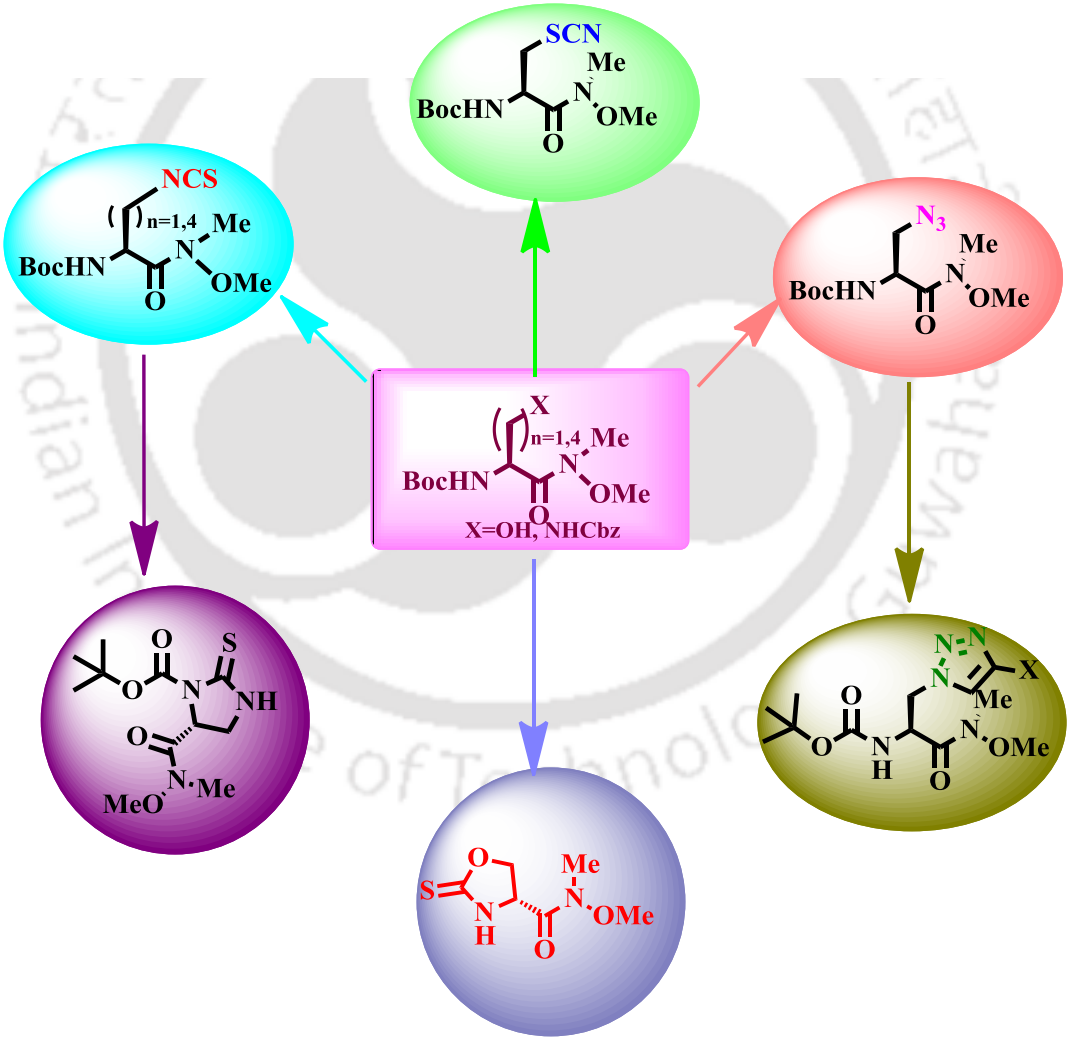
46. (a) Martin, L. J.; Imperiali, B. *Methods Mol. Biol.* **2015**, *1248*, 201. (b) Talukder, P.; Chen, S. C.; Liu, T.; Baldwin, E. A.; Benkovic, S. J.; Hecht, S. M. *Bioorg. Med. Chem. Lett.* **2014**, *22*, 5924.
47. (a) Lu, C. H.; Li, J.; Zhang, X. L.; Zheng, A. X.; Yang, H. H.; Chen, Xi.; Chen, G. N. *Anal. Chem.* **2011**, *83*, 7276. (b) Oh, K. J.; Cash, K. J.; Plaxco, K. W. *J. Am. Chem. Soc.* **2006**, *128*, 14018. (c) Marmé, N.; Knemeyer, J. P.; Wolfrun, J.; Sauer, M. *Angew.Chem., Int. Ed. Engl*. **2004**, *43*, 3798.
48. Hohsaka, T.; Kajihara, D.; Ashizuka, Y.; Murakami, H.; Sisido, M. *J. Am. Chem. Soc*. **1999**, *121*, 34.
49. (a) Grabchev, I.; Chovelon, J. -M.; Qian, X.; *J. Photochem. Photobiol. A* **2003**, *158*, 37. (b) Martin, E.; Weigand, R.; Pardo, A. *J. Lumin.* **1996**, *68*, 157.
50. (a) Turcatti, G.*;* Zoffmann, S.; Lowe III, J. A.*;* Drozda, S. E.; Chassaing, G.; Schwartz, T. W.; A. Chollet, *J. Biol. Chem*. **1997**, *272*, 21167.*(b)* Goguen, B. N.; Loving, G. S.; Imperiali, B. *Bioorg. Med. Chem. Lett*. **2011**, *21*, 5058.
51. Goldberg, J. M.; Chen, X.; Meinhardt, N.; Greenbaum, D. C.; Petersson, E. J. *J. Am. Chem. Soc.* **2014**, *136*, 2086.
52. Talukder, P.; Chen, S.; Roy, B.; Yakovchuk, P.; Spiering, M. M.; Alam, M. P.; Madathil, M. M.; Bhattacharya, C.; Benkovic, S. J. Hecht, S. M. *Biochemistry* **2015**, *54*, 7457.
53. Chen, S.; Tsao, M.-L. *Bioconjugate Chem.* **2013**, *24*, 1645.
54. (a) Chen, S.; Fahmi, N. E.; Bhattacharya, C.; Wang, L.; Jin, Y.; Benkovic, S. J.; Hecht, S. M. *Biochemistry* **2013**, *52*, 8580. (b) Chen, S.; Fahmi, N. E.; Wang, L.; Bhattacharya, C.; Benkovic, S. J.; Hecht, S. M. *J. Am. Chem. Soc.* **2013**, *135*, 12924.
55. (a) Maini, R.; Dedkova, L. M.; Paul, R.; Madathil, M. M.; Chowdhury, S. R.; Chen. S.; Hecht, S. M. *J. Am. Chem. Soc.* **2015**, *137*, 11206. (b) Chowdhury, S. R.; Chauhan, P. S.; Dedkova, L. M.; Bai, X.; Chen, S.; Talukder, P.; Hecht, S. M. *Biochemistry* **2016**, *55* 2427.
56. Brun, M. P.; Bischoff, L.; Garbay, C. *Angew. Chem. Int. Ed.* **2004**, *43*, 3432
57. Nitz, M.; Mezo, A. R.; Ali, M. H.; Imperiali, B. *Chem. Commun*. **2002**, 1912.
58. Loving, G.; Imperiali, B. *J. Am. Chem. Soc*. **2008**, *130*, 13630. (b) Zhang, J.; Wallar, B. J.; Popescu, C. V.; Renner, D. B.; Thomas, D. D.; Lipscomb, J. D. *Biochemistry.* **2006**, *45*, 2913. (c) Zuhlke, R. D.; Pitt, G. S.; Deisseroth, K.; Tsien, R. W.; Reuter, H. *Nature* **1999**, *399*, 159. (d) Torok, K.; Cowley, D. J.; Brandmeier, B. D.; Howell, S.; Aitken, A.; Trentham, D. R. *Biochemistry* **1998**, *37*, 6188. (e) Wang, Q. Z.; Lawrence, D. S. *J. Am. Chem. Soc.* **2005**, *127*, 7684.
59. Cheruku, P.; Huang, J.–H.; Yen, H.–J.; Iyer, R. S.; Rector, K. D.; Martinez, J. S.; Wang H.–L. *Chem. Sci.* **2015**, ***6***, 1150.
60. (a) Katritzky, A. R.; Narindoshvili, T.; Angrish, P.; *Synthesis* **2008**, *13*, 2013. (b) Katritzky, A. R.; Yoshioka, M.; Narindoshvili, T.; Chung, A.; Khashab, N. M.; Johnson,

J. V. Org. Biomol. Chem. **2008**, *6*, 4582. (c) Katritzky, A. R.; Cusido, J.; Narindoshvili, T.; *Bioconj. Chem*. **2008**, *19*, 1471. (d) Johansson, A.; Akerblom, E.; Ersmark, K.; Lindeberg, G.; Hallberg, A. *J. Comb. Chem*. **2000**, *2*, 496.

61. Xiang, Z.; Wang, L. J. Org. Chem., **2011**, *76* , 6367
62. (a) Sisido, M. *Prog. Polym. Sci.* **1992**, *17*, 699. (b) Sisido, M. *Ad*v. *Photochem.* **1997**, *22*, 197. (c) Hohsaka, T.; Ashizuka, Y.; Murakami, H.; Sisido, M. *J. Am. Chem. Soc.* **1996**, *118*, 9778. (d) Hohsaka, T.; Kajihara, D.; Ashizuka, Y.; Murakami, H.; Sisido, M. *J. Am. Chem. Soc*. **1999**, *121*, 34.
63. Ishida,H.; Kyakuno,M.; Oishi,S. Biopolymers. 2004, 76, 69.
64. (a) Skrynnikov, N. R.; Mulder, F. A. A.; Hon, B.; Dahlquist, F. W.; Kay, L. E. *J. Am. Chem. Soc.* **2001**, *123*, 4556. (b) Cellitti,S.E.; Jones, D.H.; Lagpacan,L.; Hao, X.; Zhang, Q.; Hu,H.; Brittain, S.M.; Brinker,A.; Caldwell,J.; Bursulaya,B.; Spraggon,G.; Brock,A.; Ryu,Y.; Uno,T.; Schultz, P.G.; Geierstanger, B.H. *J. Am. Chem. Soc*. **2008**, *130*, 9268.
65. Rabanal, F.; DeGrado, W. F.; Dutton, P. L. *J. Am. Chem. Soc.* **1996**, *118*, 473 474.
66. Regan, L. *Trends Biochem. Sci.* **1995**, *20*, 280.
67. (a) Walkup, G. K.; Imperiali, B. *J. Am. Chem. Soc.* **1996**, *118*, 3053. (b) Godwin, H. A.; Berg, J. M. *J. Am. Chem. Soc.* **1996**, *118*, 6514.
68. Cheng, R.P.; Fisher, S.L.; Imperiali, B. *J. Am. Chem. Soc*. **1996**, *118*, 11349.
69. Kazmierski, W. M. *Int. J. Pept. Protein Res.* **1995**, *45*, 241.
70. Gilbertson, S. R.; Chen, G.; McLoughlin, M. *J. Am. Chem. Soc.* **1994**, *116*, 4481.
71. (a) Imperiali, B.; Prins, T. J.; Fisher, S. L. *J. Org. Chem.* **1993**, *58*, 1613. (b) Imperiali, B.; Fisher, S. L. *J. Org. Chem.* **1992**, *57*, 757. (c) Imperiali, B.; Fisher, S. L. *J. Am. Chem. Soc.* **1991**, *113*, 8527.
72. Rana, T. M.; Ban, M.; Hearst, J. E. *Tetrahedron Lett.* **1992**, *33*, 4521.
73. Wilson, S. R.; Yasmin, A.; Wu, Y. *J. Org. Chem.* **1992**, *57*, 6941.
74. Ruan, F.; Chen, Y.; Hopkins, P. B. *J. Am. Chem. Soc.* **1990**, *112*, 9403.
75. Torrado, A.; Imperiali, B. *J. Org. Chem.* **1996,** *61*, 8940.
76. (a) Imperiali, B.; Fisher, S. L. *J. Org. Chem.* **1993**, *58*, 1613. (b) Imperiali, B.; Kapoor, T. M. *Tetrahedron* **1993**, *49*, 3501. (c) Walkup, G. K.; Imperiali, B. *J. Am. Chem. Soc.* **1996**, *118*, 3053. (d) Godwin, H. A.; Berg, J. M. *J. Am. Chem. Soc.* **1996**, *118*, 6514
77. Walkup, G. K.; Imperiali, B. *J. Org. Chem.* **1998,** *63*, 6727.
78. (a) Lohani, C. R.; Kim, J. M.; Lee, K.-H. *Tetrahedron.* **2011**, *67*, 4130. (b) Lohani, C.; Neupane, R. L. N.; Kim, J.–M.; Lee, K. -H. *Sens. Actuators.* B. **2012**, *161*, 1088.

Chapter-2

Synthesis of Triazolyl, Thiocyanyl and Isothiocyanyl Unnatural Amino Acids



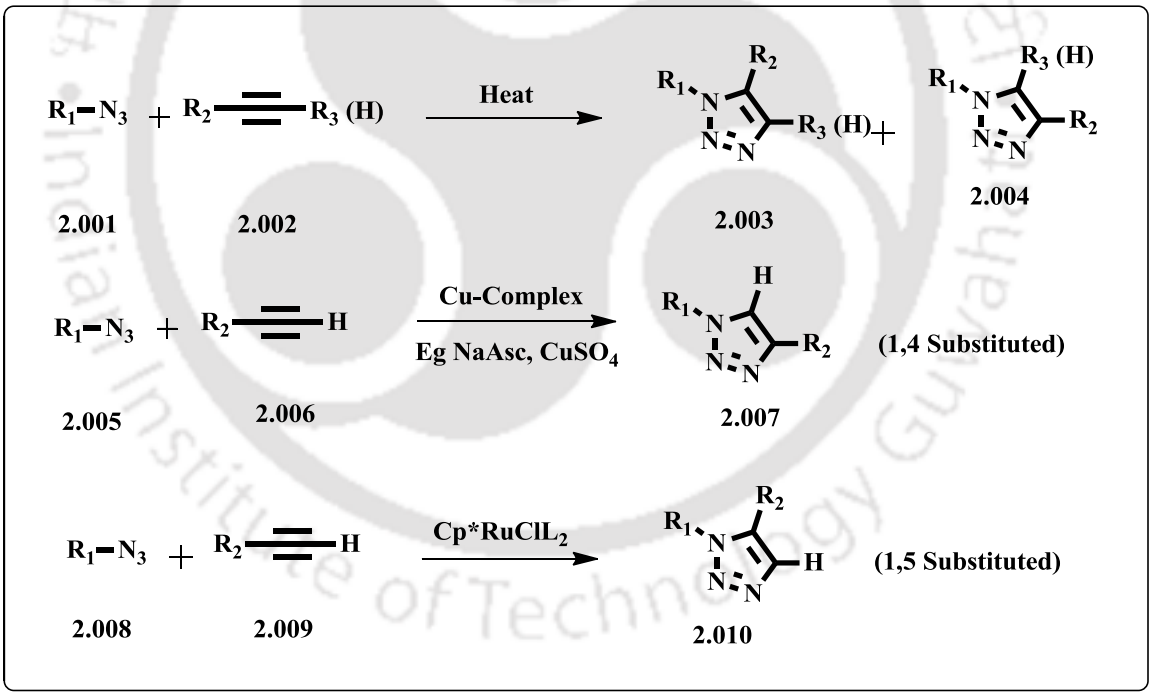
2.1. Introduction

The diversity in protein structures and their functions are being widely expanded by the introduction of UNAAs.[1f] Therefore, synthesising unnatural UNAAs is an attractive area of research. Further, small synthetic UNAAs can be incorporated into the peptides to mimic the secondary structural features of peptides.[1b] Triazoles based UNAAs are an important class of side chain modified amino acids, which can mimic the naturally occurring residue in terms of size, reactivity, stability and physiochemical activity. The triazolyl moiety is commonly found in many biologically active molecules and is responsible for physiological activities such as anti-HIV, antibiotics, antiviral and antibacterial agents.[2] In recent years, click chemistry has got much attention among the scientific community for the preparation of triazole based useful drug molecules and pharmaceutical leads.[3,4] The click chemistry is one of the dominant and common reaction which provides a wide variety of chemical diversity[4] and has extensive applications such as chemical ligation,[5] labeling of proteins, lipids [7] and nucleic acids[8] *in vitro* / *vivo*. [5, 6]

The sulfur-containing amino acids (SAAs) can play an important role in chelation of neurotoxin drugs such as chlorella and cilantro [1c,d]. The SAAs chelates physiologically important metal ions such as copper (Cu), lead (Pb) for inter tissue transport and their removal from the body.[1d] The SAAs involves in disulfide bond formation and consequently helps to fold regular secondary structures such as helices, sheets, and turns.[1b] As for example, thiocyanates and isothiocyanates can mimic the natural sulfur-containing residue (SAAs) in terms of stability and reactivity and also provide a labile functional group in chemical synthesis of biologically active sulfur-containing heterocycles.[27a-1] It has been earlier reported that the isothiocyanate and thiocyanate moieties are found in many natural products [27a-l,]; few having the anti-cancer activity.[28] It has been found, compared to thiocyanates, isothiocyanates readily take part in nucleophilic addition and cycloaddition reactions due to the strong electrophilic character of the latter.[27a-o] So, the presence of isothiocyanate group in the side chain of the amino acids would bring functional tunability, which would be an interesting topic.

2.2. General Synthesis of Triazoles

Among the variety of heterocyclic molecule, 1,2,3-triazole heterocycles got much attention in recent time because of its wide applicability in the various fields ranging from chemistry, biology to material science.[9,10] The most classical way to the synthesis of 1,2,3-triazoles involves thermal 1,3-dipolar cycloaddition of azides with alkynes with no selectivity (**2.03-2.004, Scheme 2.01**), as initially proposed by Huisgen.[10d] Cu(I)-catalyzed version of the Huisgen 1,3-dipolar cycloaddition reaction became popular to generate 1,4-disubstituted 1,2,3–triazoles (**2.007, Scheme 2.01**) which is referred as "click reaction" defined by K. B. Sharpless in 2001.[10] Similarly 1,5-substitued triazoles (**2.010, Scheme 2.01**) are easily prepared *via* Ru-catalyzed click reactions.[10c] The advantages of this reaction is based on the fact that it can easily generate a triazole unit with complete regiospecificity. **Scheme 2.01** shows the methods for the generation of 1,2,3-triazole *via* click reactions. The triazoles are rotationally restricted small molecular scaffold having high metabolic stability, biocompatibility, and easy associability *via* H-bonding with biomolecules.[11]



Scheme 2.01. Methods of 1,2,3-triazole synthesis.

2.2.1. Applications of Triazoles

Triazole is a distinctive structural unit, associated with several biological activities. This triazole acts as an antimicrobial, anti-inflammatory, antidepressant, anti-malarial, analgesic, and anticancer agents.[12a] Freddy.*et al*.,[12b] have synthesised a 3-[4-substituted-phenyl-5-thioxo-4,5-dihydro-1*H*-[1,2,4]-triazole-3-yl-methoxy)-phenyl]-2-phenyl-3*H*-quinazoline-4-one **(2.011, Figure 2.01)** which are screened for antifungal activity. Sztanke *et al*.,[12c] have synthesised a series of 3-(un)substituted-7-aryl-5H-6,7-dihydroimidazo-[2,1-c][1,2,4]-triazoles compounds and their derivatives which are **(2.012, Figure 2.01)** screened for antimicrobial and antifungal activities. Sung *et al*.,[12d] have synthesised [(4-Amino 5-Disubstituted-4-H-[1,2,4]-triazole-3-yl)thio] alkanoic acid based molecules (**2.013, Figure 2.01)** and tested anti-inflammatory activity. Demirbas *et al*.,[12e] have prepared a series of compounds 4-amino-3-substituted-5-oxo-4,5-dihydro-[1,2,4]-triazole-1-yl-aceticacid-2,4-dichlorobenzylidene hydrazide derivatives (**2.015, Figure 2.01**) and tested their antitumor activity. A series of 5-[(Biphenyl-4-yloxy)methyl]-4-n-substituents-3-mercapto-(4H)-[1,2,4]-triazole has been synthesised by Kumar *et .al*.[12f] The derivative such as 5-[(Biphenyl-4-yloxy)methyl]-4-fluorophenyl-3-mercapto-(4H)-[1,2,4]-triazole (**2.014, Figure 2.01)** screened for the analgesic activity.

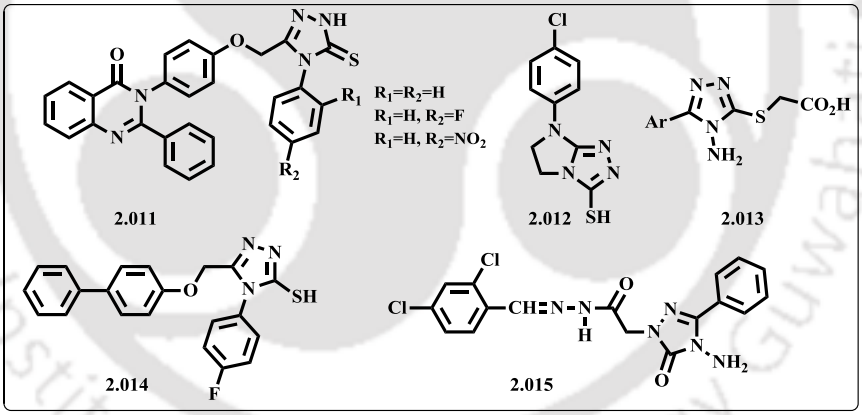


Figure 2.01. Biologically active triazole molecules.

In 2011, Zerrouki and his co-workers [13] described a productive strategy for the synthesis of a series of tri (ethylene glycol)-based oligo triazoles (**2.016, Figure 2.02**). They have substituted all backbone amide bonds of a peptide nucleic acid (PNA) oligomer by [1,2,3] triazole moieties.[13] Esslinger *et al*., have first time synthesised and reported triazole based unnatural amino acids (**2.017, Figure 2.02**) by using click reaction.[14a] They have used those amino acids to probe binding interactions between ligand and neutral amino acid transport proteins. They found that the triazole amino acids are basically unreactive at the SN1 transporter protein.[14a] Our group have synthesised donor/acceptor based triazolyl unnatural amino acids and

effectively incorporated into tripeptide moiety (**2.018, Figure 2.02**) which mimics *β*-turn conformation of natural *β*-turn analogue.[14b]

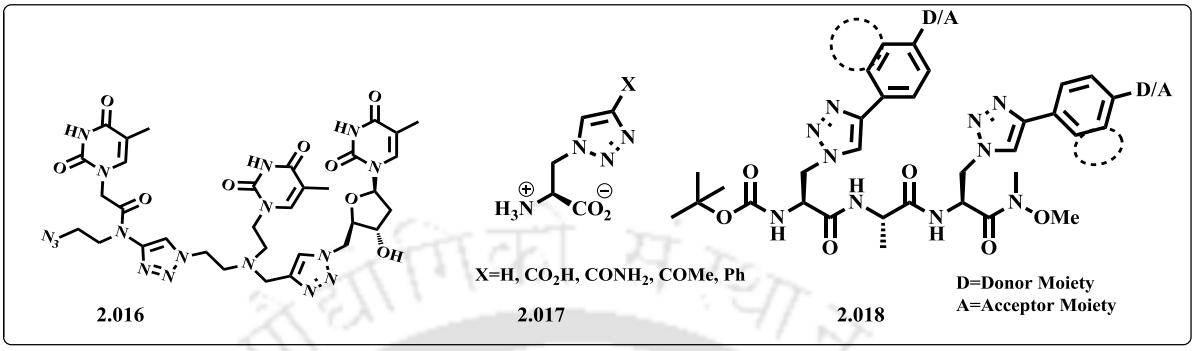


Figure 2.02. Oligotriazole, triazole amino acid and triazole containing peptide.

2.3. General Synthesis of Isothiocyanates and Thiocyanates

Various synthetic methods have been reported to convert readily available amines into the isothiocyanate analogue (**Figure-2.03).** Most of the syntheses of isothiocyanate are based on the decomposition of dithiocarbamic acid salts into isothiocyanates. The required dithiocarbamic acid salts can be formed *in situ* by treatment of an amine with carbon disulfide (CS_2) in the presence of triethylamine (Et_3N).[15] The schematic formation of isothiocyanate is given below

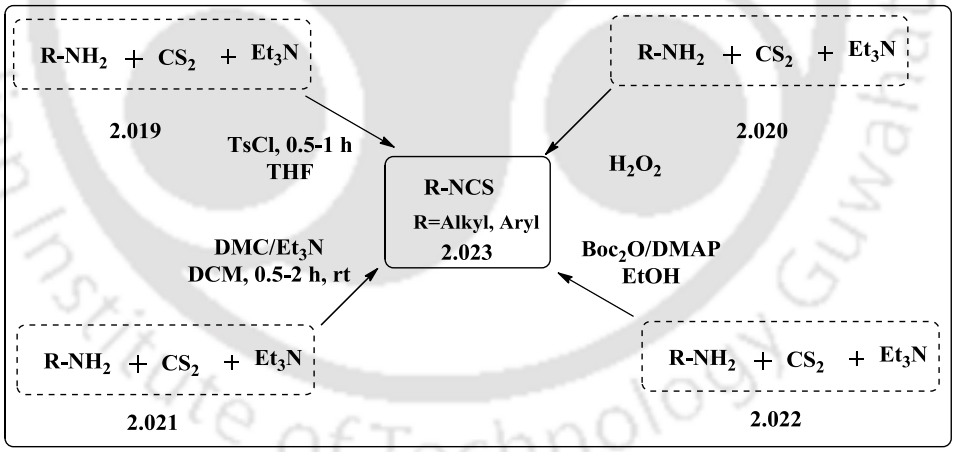


Figure 2.03. Schematic representation of isothiocyanate synthesis.

Various synthetic methods have been reported to prepare thiocyanate analogues (**Figure-2.04)**. Most of them are based on (a) nucleophilic substitution of alkyl halides or tosylates[16a], alcohols, thiols, carboxylic acids, silyl ethers, silyl carboxylates [16b] sulfonate esters [16c] and/or (b) cross coupling reaction of aryl boronic acids.[16d]

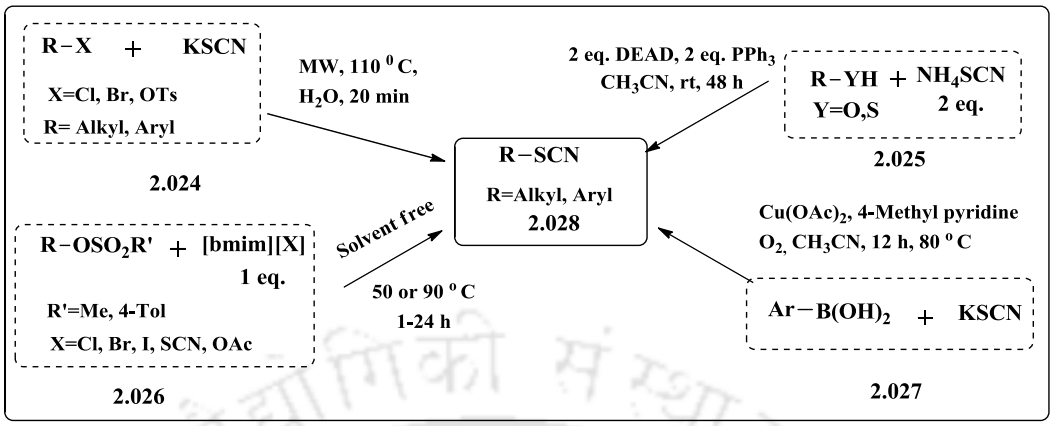


Figure 2.04. Schematic representation of thiocyanate synthesis.

2.3.1. Applications of Isothiocyanates and Thiocyanates

Isothiocyanates occur widely in nature and are of interest in food science and medicine. As for example, phenethyl isothiocyanate (PEITC) shows apoptosis-resistant to some currently used chemotherapeutic drugs.[17a, 17b] Certain isothiocyanates effectively binds to the mutated p53 proteins found in many types of tumors, causing an increase in the rate of cell death.[17c] T. Kim *et al*., [18] have synthesised with various sulforaphane analogs (**2.029-2.030, Figure 2.05**) from aromatic and aliphatic amines and tested on cisplatin-treated cultured LLC-PK1 kidney cell line. Among different analogs, **2.029** show a potent effect on kidney cell protection assay. Lamani *et al*., [19] have synthesised 1,3,4-oxadiazole containing peptidomimetics **(2.032, Figure 2.05**) by coupling N-protected amino acid hydrazides with amino acid-derived isothiocyanate esters (**2.031, Figure 2.05**).

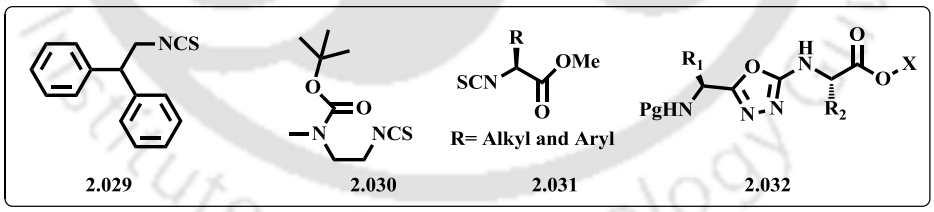


Figure 2.05. Synthesised isothiocyanates and oxadiazole compound.

Thiocyanate plays very crucial role in the biosynthesis of hypothiocyanite by a lactoperoxidase.[20a-c] Hence the lack of thiocyanate [21] or reduced thiocyanate [22] in the human body, (e.g., cystic fibrosis) is often responsible for destroying the defense system in human body.[23][24] Thiocyanate is an effective inhibitor of the thyroid sodium-iodide symporter. [25] Iodine is an important element of thyroxin and thiocyanates decreases iodide transport into the thyroid follicular cell which helps to maintain thyroxine level in the thyroid gland.

2.4. Background

From the literature report it is clear that in the fast moving journey of expanding the genetic code, several unnatural amino acids (UNAAs) have been synthesised and incorporated in protein's framework to bring new functionalities into proteins. UNAAs always create an opportunity and provide an effective way for site-selective modification of proteins in broader and more suitable way. However, modification of such residues can be limited by methods for their installation and the chemistry available for reaction. Therefore, the interest in the design and application of new unnatural amino acids is increasing substantially. Moreover, the synthesis of unnatural amino acids with a new side chain functional group is highly attractive and core area of research in amino acids and peptides. The sulfur-containing amino acids not only play an important role in disulfide bond formation in peptide but also it acts in connective tissue for their flexibility. Sometimes disulfide bond creates the problem because it aggregates the protein and it happens by oxidation of cysteine residues to sulfonic acids which leads to excessive disulfide bonding, protein misfolding, and aggregation.[1a] So, the side chain modified sulfur amino acids can mimic the cysteine residue in terms of stability and physiochemical properties like oxidation, reduction. So, synthesizing modified sulfur containing UNAAs is an important area of research in protein field. Therefore, there is not essentially a need for new chemistry for transforming proteins but better methodology for unnatural amino acids is essential.

Towards this journey, a very few new triazole and sulfur containing amino acids have been synthesised to encoded genetically and site-specifically into proteins, which has already been described in **Chapter 1**. Therefore, there is a great demand to develop new side chain modified unnatural amino acids with reactive or unreactive functional groups. Recently, our research group is involved in the expansion and synthesis of side chain modified UNAAs. Herein, we present our views on this fast-developing field, with a focus on the synthesis of side chain modified unnatural amino acids.

2.5. Objective

With this above literature background of unnatural amino acids, we found that very few sulfur containing and triazole based amino acids have been reported for the purpose of modification in peptide and expansion of genetic code. As a part of our current research efforts on side chain modified amino acids synthesis *via* azide-alkyne cycloaddition reaction, we thought that it would be worthwhile to synthesise triazole based amino acids with different functional groups which can be used for various purposes in peptide chemistry. Small side chain which can be useful IR spectroscopy study as well has non-perturbing properties can of be immense interest in the full of peptide protein research. Such small side chain may be –NCS which is poorly explored in the current research. As of now, many reports are available on the synthesis of isothiocyanate and thiocyanate containing compounds with novel functionality. However, isothiocyanyl and thiocyanyl amino acids have not been reported yet. Isothiocyanates are a very important class of chromophores known for a long time for their analytical use in the determination of the primary structure of peptides and proteins.[12] In recent time, these chromophores have attracted much research attention for their stability, synthetic easiness of thioureas, thioamides and other precursors of various sulfur containing biologically active heterocycles.[13] So, the exploration of isothiocyanyl and thiocyanyl amino acid chemistry would generate a new class of unnatural amino acid derivatives.

Therefore, with this aim and idea we framed our objectives as below:

(a) Design and synthesis of side chain modified triazole amino acids using click chemistry, starting from easily available L-serine (**Figure 2.06**).

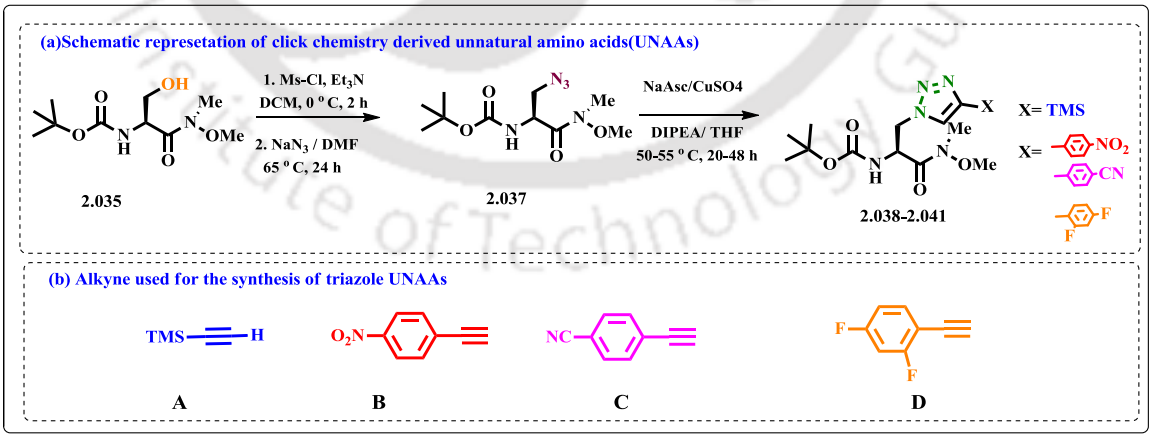


Figure 2.06. (a) Schematic of representation of synthesis of triazolyl unnatural amino acids (UNAAs) and (b) the alkynes used in this study.

(b) Several methodologies exist for the synthesis of thiocyanates from the corresponding alkyl halides or tosylates, alcohols, thiols, carboxylic acids, silyl ethers, silyl carboxylates and sulfonate esters. However, we adopted an easy and convenient literature reported methodology which relied on nucleophilic substitution at alkyl halide in presence of potassium thiocyanate (KSCN) (**Figure 2.07**) to prepare thiocyanyl alanine (^{SCN}**Ala**).

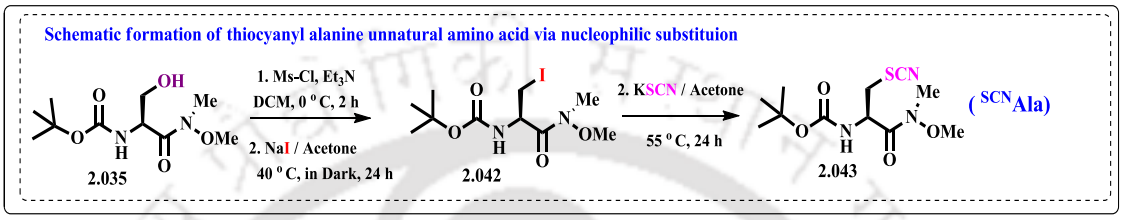


Figure 2.07. Schematic presentation of formation of thiocyanyl alanine (^{SCN}**Ala**).

(c) Many methodologies were reported for the synthesis of isothiocyanates from the corresponding amines. Here we used a convenient literature reported methodology which is based on TsCl-promoted decomposition of dithiocarbamic acid salts into isothiocyanates to yield isothiocyanyl alanine (^{NCS}**Ala = Ita**) and isothiocyanyl lysine (^{NCS}**Lys = Itl**) (**Figure 2.08)**.

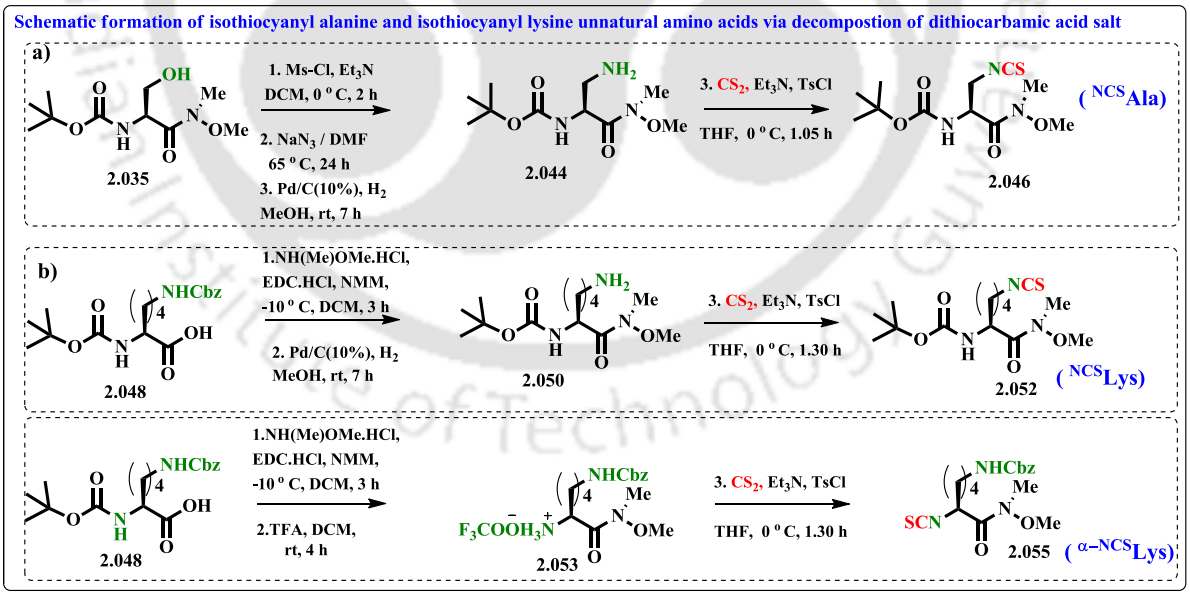
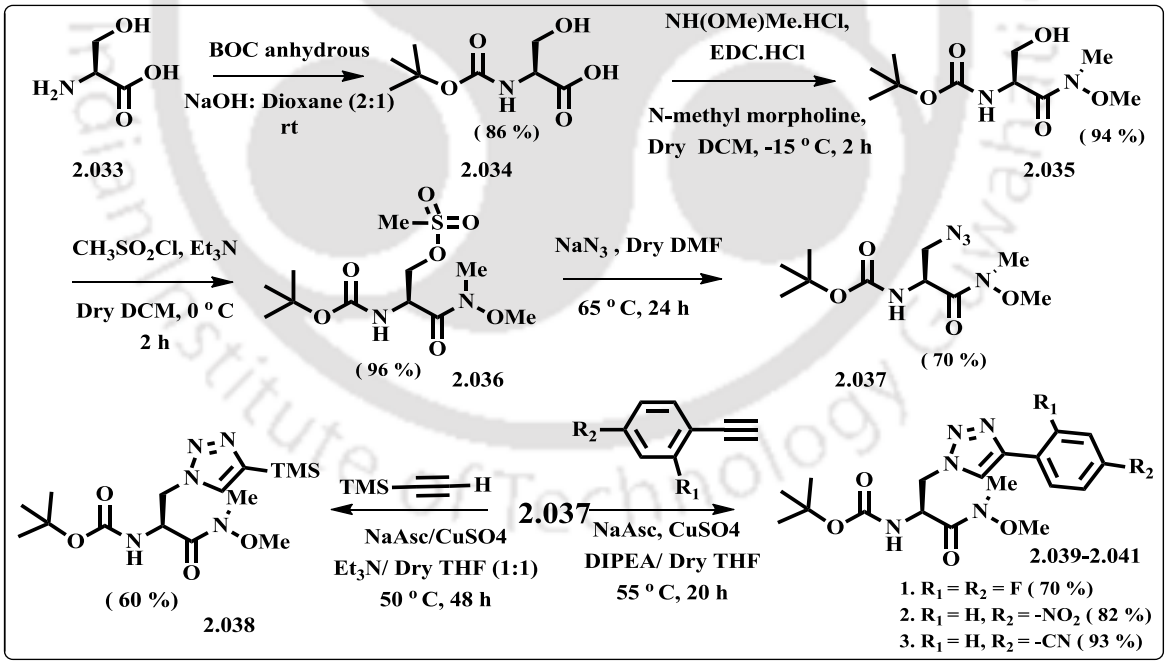


Figure 2.08. Schematic of formation of (a) isothiocyanyl alanine (^{NCS}**Ala**) and (b) isothiocyanyl lysine (^{NCS}**Lys**) (c) α- Isothiocyanyl Lysine ($^{\alpha-NCS}$**Lys**).

2.6. Results and Discussion

2.6.1. Synthesis of Triazolyl Amino Acids

The synthesis of unnatural triazolyl amino acids was achieved *via* a 1,3-dipolar cycloaddition reaction between an aromatic alkyne and serine azide which is shown in **Scheme 2.07**. The synthesis started with natural L-serine (**2.033, Scheme 2.02**) which was protected first with Boc-anhydride to afford N-Boc-serine (**2.034)**. Afterward, the C-terminus was protected as weinreb amide (**2.035**) with N-methoxy methylamine hydrochloride following a peptide coupling protocol. The N, C-diprotected serine **2.035** was then allowed to react with mesyl chloride in presence of triethylamine in dry DCM at 0 ^{o} C for 2 hours to afford N, C-diprotected *O*-mesyl serine (**2.036**). Upon treatment with sodium azide in dry DMF at 65 ^{o} C, mesyl was converted to N, C-diprotected azidoalanine (**2.037, Scheme 2.02**). Finally, Cu(I) catalysed 1,3-azide-alkyne cycloaddition reaction was carried out between the azidoalanine and various aromatic and aliphatic alkynes (**A-D, Figure 2.06**) at 50-55 ^{o} C in THF to produce the desired triazolyl amino acids (**2.038-2.041, Figure 2.09**) in good to excellent yield (60-93%). Formation of all the compounds was confirmed by NMR, IR and ESI-HR mass spectrometry.



Scheme 2.02. Schematic of synthesis of triazolyl amino acid derivatives.

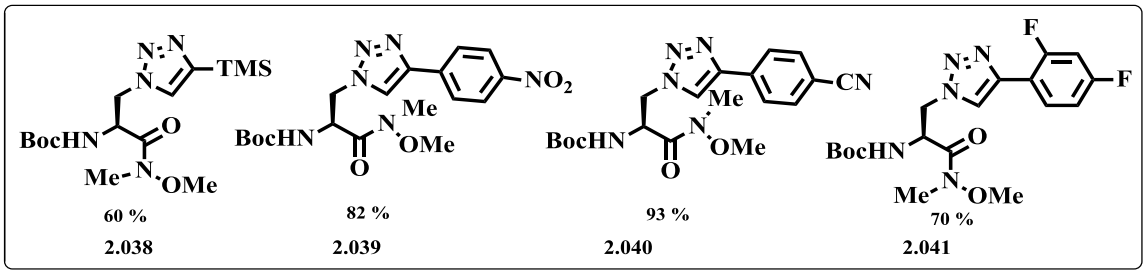
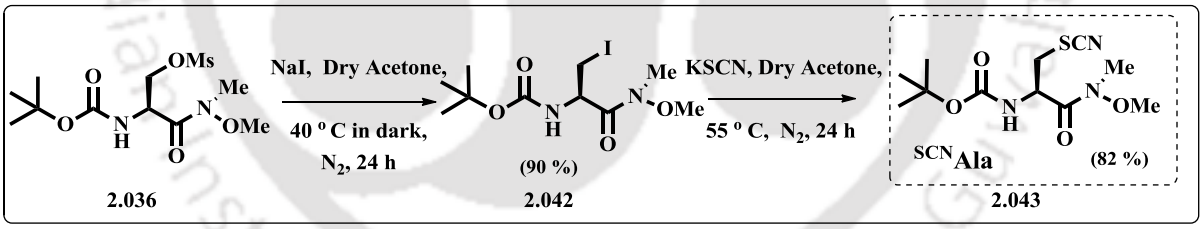


Figure 2.09. Chemical structures of synthesised triazolyl amino acids.

2.6.2. Synthesis of Thiocyanyl Alanine (^{SCN}Ala)

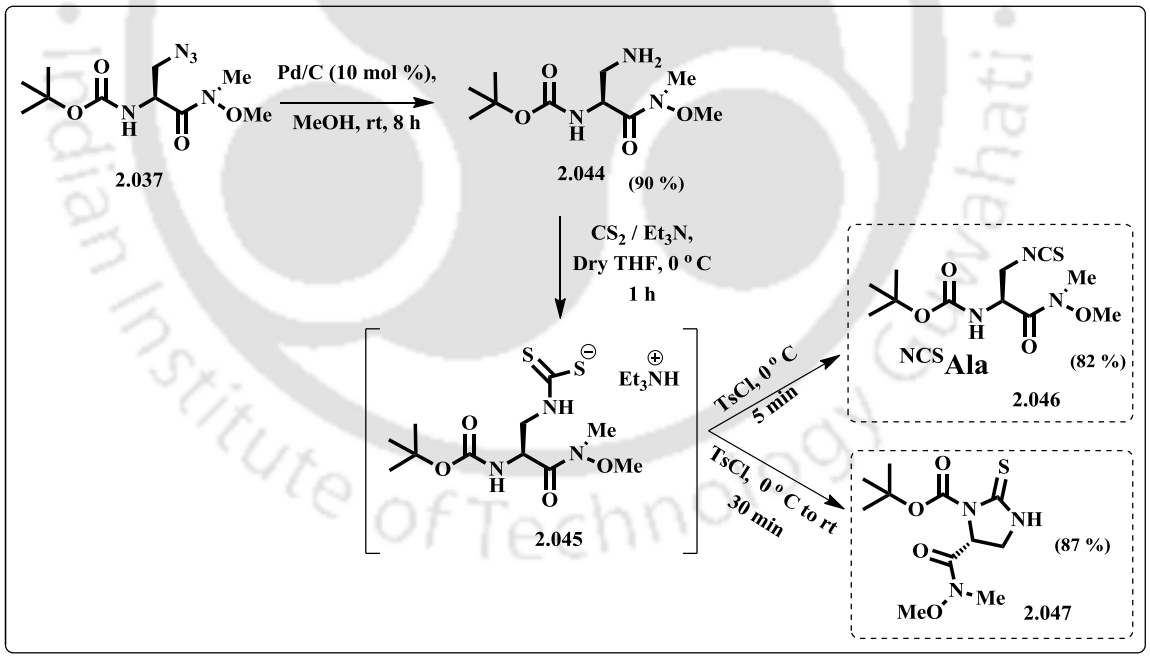
The synthesis of thiocyanyl alanine was achieved *via* nucleophilic substitution of N, C-diprotected iodoalanine (**2.042, Scheme 2.03**). The N, C-diprotected iodoalanine was obtained from N, C-diprotected O-mesyl serine (**2.036**) with the treatment of sodium iodide (NaI) in dry acetone at 40 ^{o} C in dark under N_2 atmosphere for 24 hours. A nucleophilic substitution reaction carried out between the N, C-diprotected iodoalanine with potassium thiocyanate (KSCN) at 55 ^{o} C in dry acetone under N_2 atmosphere for 24 hours afforded the desired thiocyanyl alanine (**^{SCN}Ala , 2.043**) in 81 % yield. Formation of all the compounds was characterised by NMR, IR and ESI-HRMS



Scheme 2.03. Schematic presentation of formation of thiocyanyl alanine (**^{SCN}Ala**).

2.6.3. Synthesis of Isothiocyanyl Alanine (^{NCS}Ala)

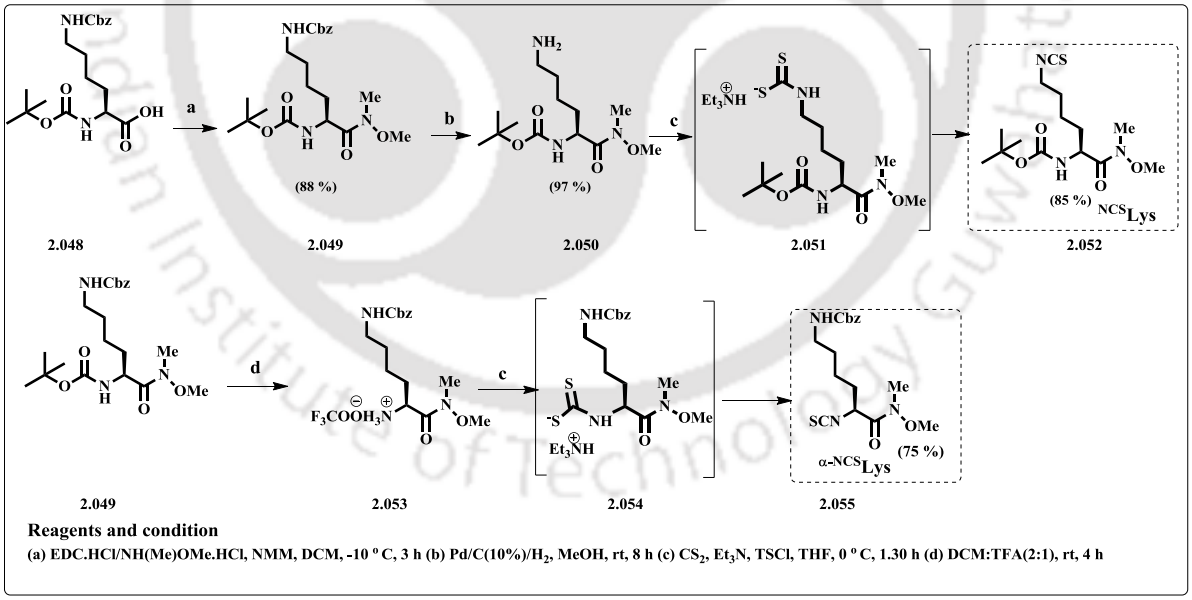
The synthesis of isothiocyanyl alanine (**2.046**) was achieved *via* decomposition of dithiocarbamic acid salt (**2.045, Scheme 2.04**) of corresponding aminoalanine (**2.044**). The N, C-diprotected aminoalanine (**2.044, Scheme 2.04**) was generated upon treatment of N, C-diprotected azidoalanine (**2.037, Scheme 2.02**) with Pd/C (10%) under hydrogen atmosphere in dry methanol for 7 hours. The required dithiocarbamic acid salt was generated *in situ* via the treatment of corresponding amine **2.044** derivatives with carbon disulfide (CS_2) in the presence of triethylamine (Et_3N) in dry THF at 0 ^{o} C for 1 hour. Finally the isothiocyante was formed upon treatment of dithiocarbamic acid salt with tosyl chloride (TsCl) at 0 ^{o} C for 5 minutes under N_2 atmosphere with very good yields. We also observed that during formation of isothiocyanyl alanine, one side product was forming and its percentage of conversion was highly dependent on time. When reaction time was 5 minutes, the acyclic NCS compound (**2.046, Scheme 2.04**) was forming exclusively in 82% isolated yield. On the other hand, prolonged reaction time for 30 minutes offered the corresponding cyclic amino acid derivative (**2.047**) exclusively with 87 % yield. The formation of the cyclic compound **2.047** was confirmed by NMR and X-ray single crystal analysis (**Figure 2.10**).



Scheme 2.04. Schematic presentation of synthesis of isothiocyanyl alanine (^{NCS}**Ala**) and its cyclic derivative.

2.6.4. Synthesis of Isothiocyanyl Lysine (^{NCS}Lys) and α- Isothiocyanyl Lysine ($^{α–NCS}$Lys)

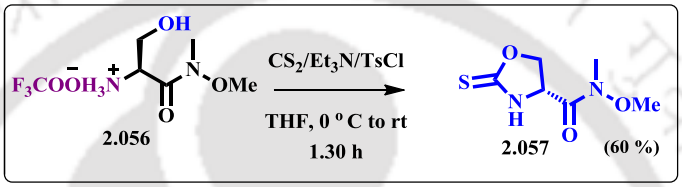
The synthesis of N,C-diprotected isothiocyanyl lysine **2.052** (**^{NCS}Lys, Scheme 2.05**) and α- isothiocyanyl lysine **2.055** (**$^{α–NCS}$Lys, Scheme 2.05**) was achieved via decomposition of dithiocarbamic acid salt of corresponding N, C-diprotected lysine respectively. At first C-terminus of Boc-lysine(Cbz)-CO_2H (**2.048, Scheme 2.05**) was protected as weinreb amide to form $^{\epsilon}$N,$^{\alpha}$N,C-triprotected lysine(**2.049**) with N-methoxy methylamine hydrochloride following a peptide coupling protocol**.** Afterward corresponding $^{\alpha}$N,C-diprotected lysine (**2.050**) was generated *via* selective deprotection of $^{\epsilon}$N-carboxybenzyl (Cbz) group of **2.049** in presence of Pd/C(10%) under hydrogen atmosphere at room temperature in dry methanol. The $^{\epsilon}$N, C-diprotected lysine (**2.053**) was generated *via* selective deprotection of α-N-Boc upon treatment with trifluoroacetic acid in dry DCM for 4 hours. The required dithiocarbamic acid salts (**2.051** & **2.054**) were generated *in situ* by treatment of corresponding amine derivatives (**2.050** and/or **2.053**) with carbon disulfide (CS_2) in presence of triethylamine (Et_3N) in dry THF for 1 hour. Finally, isothiocyanates (**2.052** & **2.055, Scheme 2.05**) were obtained upon treatment of dithiocarbamic acid salts with tosyl chloride (TsCl) at 0 ^{o} C for 30 minutes under N_2 atmosphere with good yields of 85% and 75% respectively. Here we did not observe any cyclic derivative of **^{NCS}Lys**. Formation of all compounds was confirmed by NMR, IR and ESI-HRMS



Scheme 2.05. Synthesis of isothiocyanyl lysine (**^{NCS}Lys**) and α- isothiocyanyl lysine (**$^{α-NCS}$Lys**).

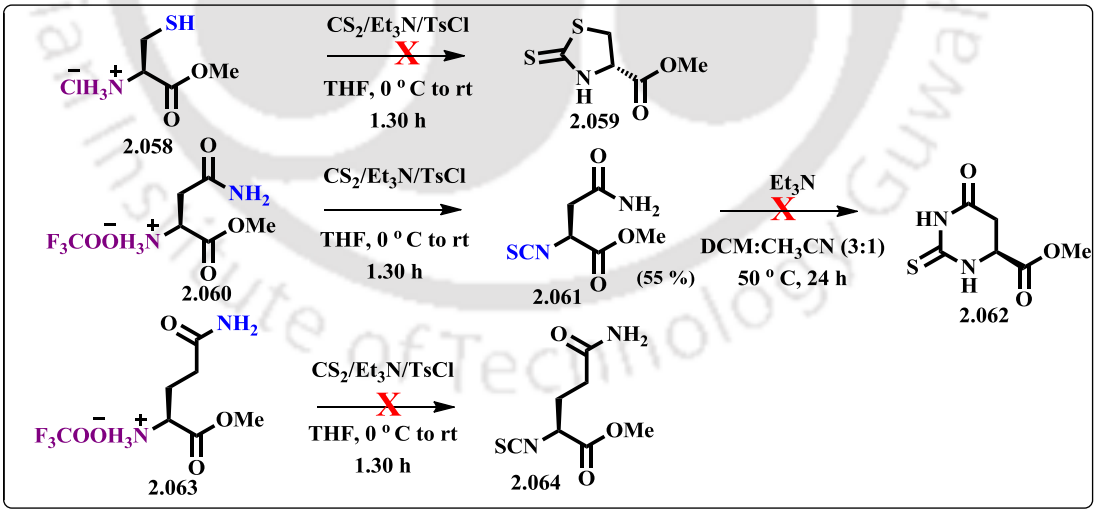
2.6.5. Studies on the Intramolecular Cyclisation of Isothiocyanyl Amino Acid Derivatives and Other Isothiocyanyl Compounds

Our observation on the formation of cyclic amino acid derivative was discussed in **Scheme 2.04** which led us to further investigate the intramolecular cyclisation process in other isothiocyanyl amino acid residues. For that purpose, we have chosen different C-protected amino acids containing nucleophilic functional group at side chain. We employed same reaction condition (**Scheme 2.06)** for all the cases towards the formation of isothiocyanyl moiety and we found that under the similar reaction conditions C-protected serine(**2.056**) was converted to similar kind of cyclic amino acid derivatives (**2.057, Scheme 2.06)**.

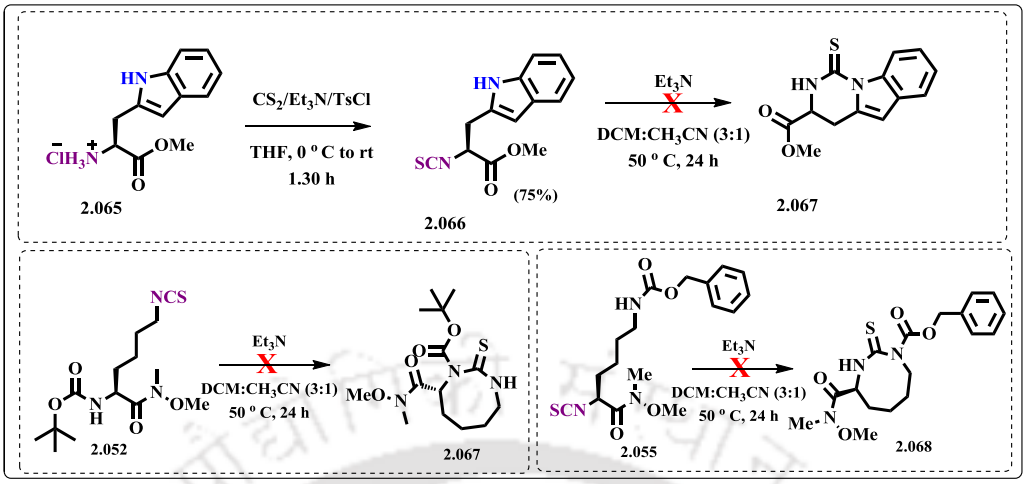


Scheme 2.06. Intramolecular cyclisation of C-protected serine under the condition of isothiocyanlation.

The other amino acids tested afforded only acyclic derivatives (**2.052, 2.055 and 2.061, 2.066**). We also tried to study the cyclisation of some of the purified acyclic -NCS derivatives (**2.052, 2.055, 2.061 and 2.066**) by treating with Et_3N in DCM: CH_3CN(3:1) at 50 ^{o} C for 24 hours (**Scheme 2.07-2.08)**. However, no cyclic product was observed in all the cases.

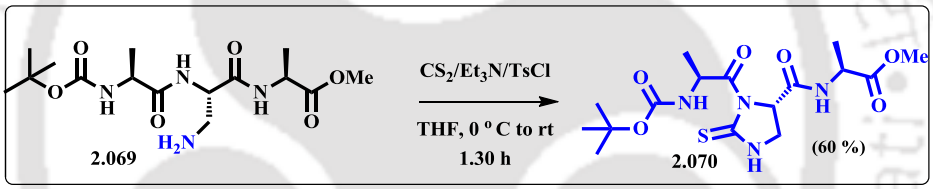


Scheme 2.07. Intramolecular cyclisation of various amino acids under the condition of isothiocyanlation.



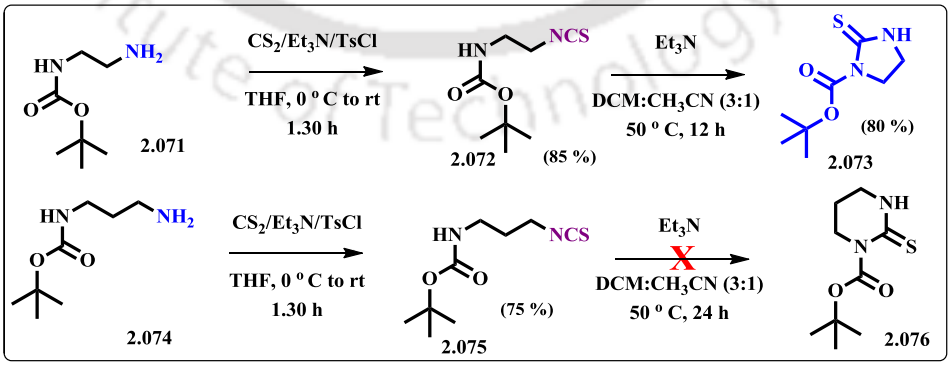
Scheme 2.08. Intramolecular cyclisation of various amino acids under the condition of isothiocyanlation.

Next, we were curious to study the cyclisation in peptide. Thus, we have synthesised amine containing tripeptide (**2.069**) which under similar condition afforded corresponding cyclic derivative **2.070 (Scheme 2.09)**.



Scheme 2.09. Intramolecular cyclisation in tripeptide under the condition of isothiocyanlation.

We also chose two aliphatic diamines, Boc-ethylenediamine (**2.071**) and Boc-diaminopropane diamine (**2.074**) for the same reaction, while compound **2.071** and **2.074** under the similar reaction condition afforded only the acyclic NCS derivatives **(2.072, 2.075, Scheme 2.10)**. The NCS derivative of Boc-ethylenediamine **2.072** was yielded cyclic derivative (**2.073, Scheme 2.10**) when treated with Et_3N in DCM:CH_3CN (3:1) at 50 °C for 12 hours.



Scheme 2.10. Intramolecular cyclisation in diamino derivatives under the condition of isothiocyanlation.

2.7. Characterisations of Few Unnatural Amino Acids

2.7.1. ^{1}H-NMR Characterisations of Few Unnatural Amino Acids

All the synthesised amino acids are in fully protected form and were characterised by NMR spectroscopy. As a representative example the structural assignments, we have chosen three protected amino acids **2.038**, **2.043** and **2.046 (Figure 2.09).** For amino acid **2.038** (**^{TzTMS}Ala**), the triazolyl hydrogen appeared as a singlet at the characteristic position of δ 7.49. α-NH of Boc group appeared at δ 5.52 as a doublet with J = 6.4Hz. The C-α hydrogen of alanyl unit resonated as a double δ5.02 with J = 13.4, 5.6Hz. The β-CH_2 hydrogens of alanyl unit appeared as a doublet at δ4.66 with J = 4.9Hz. Methyl hydrogens of -NMe and -OMe group also resonated as a singlet at δ3.17 and 3.69 respectively. The ^{t}Bu hydrogens of Boc- protected group appeared as a singlet at δ1.36. Trimethyl group of trimethylsilyal (TMS) appeared at δ0.25.

For amino acid **2.043** (**^{SCN}Ala**), α-NH of Boc group appeared at δ 5.58 as a doublet with J = 4.5 Hz. The C-α hydrogen of alanyl unit resonated as a double δ4.99 with J = 4.0 Hz. The β-CH_2 hydrogens of alanyl unit appeared as a double doublet at δ3.38 with J = 13.7, 4.0 Hz and the doublet at δ3.29 J = 4.8 Hz. Methyl hydrogens of -NMe and -OMe group also resonated as a singlet at δ3.25 and 3.79 respectively. The ^{t}Bu hydrogens of Boc- protected group appeared as a singlet at δ1.45. For amino acid **2.046** (**^{NCS}Ala**), α-NH of Boc group appeared at δ5.54 as a doublet with J = 7.5 Hz. The C-α hydrogen of alanyl unit resonated as a double δ4.88 with J = 3.5 Hz. The β-CH_2 hydrogens of alanyl unit appeared as a double doublet at δ3.84 with J = 14.3, 4.4 Hz. Methyl hydrogens of -NMe and -OMe group also resonated as a singlet at δ3.25 and 3.76 respectively. The ^{t}Bu hydrogens of Boc- protected group appeared as a singlet at δ1.45.

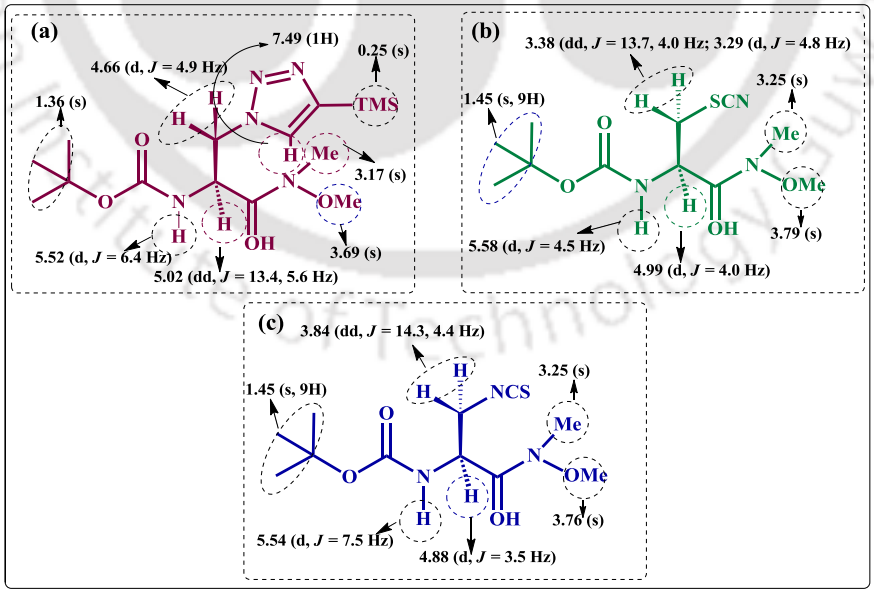


Figure 2.09. ^{1}H-NMR Characterisation of (a) **2.038** (**^{TzTMS}Ala**) (b) **2.043** (**^{SCN}Ala**) (c) **2.046** (**^{NCS}Ala**).

2.7.2. The Single Crystal X-Ray Structure of Cyclic Derivative of ^{NCS}Ala

The structure of cyclic ^{NCS}**Ala** derivative **2.047** was confirmed from X-ray single crystal analysis (**Figure 2.10, Table 2.1**).

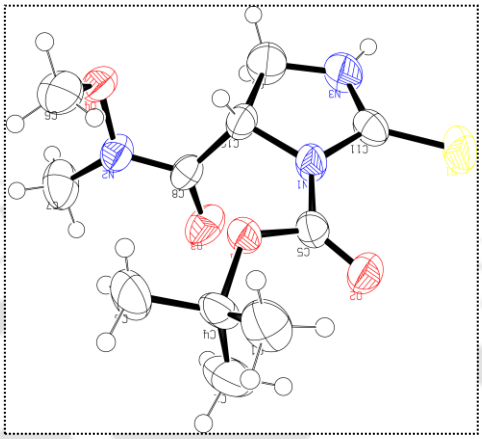


Figure 2.10. ORTEP diagram cyclic derivative **2.047** of ^{NCS}**Ala.**

Table 2.1. Summary table of crystal parameter of cyclic derivative of ^{NCS}**Ala**

Compound	**2.047**
Z	**4**
Molecular Formula	**$C_{11}H_{19}N_3O_4S$**
Temperature	**296K**
Space Group	**P 21 21 21 (orthorhombic)**
a (Å)	**8.7239(9)**
b (Å)	**10.2860 (11)**
c (Å)	**16.5805 (18)**
α, β, γ (Degree)	**α = β = γ = 90 °**
M_r	**289.35**
V (Å^3)	**1487.8**
μ (mm^{-1})	**0.231**
F(000)	**616.0**
R(reflection)	**0.0340**
wR2(reflection)	**0.0890**

2.8. Conclusion

We have successfully synthesised triazole based unnatural amino acid derivatives *via* click reaction with good yields. We have also synthesised unexplored isothiocyanyl alanine (**^{NCS}Ala**), isothiocyanyl lysine (**^{NCS}Lys/$^{\alpha\text{-}NCS}$Lys**) and thiocyanyl alanine (**^{SCN}Ala**) from natural amino acids serine and lysine. Furthermore, we explored the intramolecular cyclisation of some of the -NCS containing amino acids and peptide. The –NCS containing amino acids could be useful in designing side chain modified peptides for introducing the fluorophore. Impact of side chain –NCS on the conformation of a designed peptide could also be investigated further. Moreover, **^{NCS}Ala/^{NCS}Lya** amino acids could be utilised as site-specific and residue specific incorporation of **^{NCS}Ala** or **^{NCS}Lys** into peptides can help to investigate the effect of modification in terms of protein structure, stability and labeling experiments.

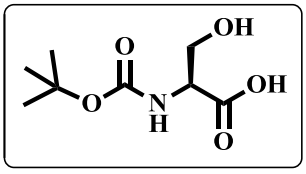
2.9. Experimental Section

2.9.1 General Experimental

All the reactions were carried out under the nitrogen atmosphere using oven-dried round bottom flasks. Reagents were purchased at the highest commercial quality and used without further purification unless otherwise stated. Reactions were monitored by thin-layer chromatography (TLC) carried out on a 0.25 mm silica gel 60F–254 and silica gel-G (1:4) and visualised under UV light at 254 nm. Further visualisation was achieved by iodine vapor adsorbed on silica gel depending on the product type. Organic extracts were dried over anhydrous sodium sulfate. Solvents were removed in a rotary evaporator under reduced pressure. Column chromatography was performed on silica gel 60–120 mesh using a mixture of hexane and ethyl acetate as mobile solvent. The isolated compounds were characterised by ^{1}H, ^{13}C NMR and IR spectroscopic techniques and Mass spectrometry. All the NMR spectra were recorded at ambient temperature on Bruker Ascend TM Aeon 600MHz spectrometer where ^{1}H frequency was 600MHz and ^{13}C frequency was 150MHz. NMR spectra for all the samples were measured in either in $CDCl_3$ or in d_6-DMSO. The chemical shift values were reported in ppm downfield from tetramethylsilane, using chloroform-d (δ = 7.26 for ^{1}H NMR, δ = 77.23 for ^{13}C NMR) or using deuterated dimethyl sulfoxide-d_6 (δ = 2.50 for ^{1}H-NMR, δ = 39.50 for ^{13}C NMR). ^{1}H-NMR coupling constant(s) J is represented in Hertz (Hz). All the NMR-FID has processed in MestReNova v6.0.2 software. High-resolution mass spectra (HRMS) were recorded on a Water system Mass spectrometer in positive mode using electrospray ionisation-time of flight (ESI-TOF) and/or Atmospheric pressure chemical ionisation-Time of flight (APCI-TOF) reflection experiments. IR spectra were recorded on KBr plate in Perkin Elmer spectrometer and reported in frequency of absorption (cm^{-1}).

2.9.2. Synthesis and Characterisations

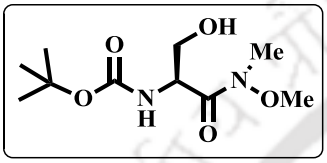
(S) N-Boc-serine (2.034): A solution of L-Serine **2.033** (6000mg, 57.09 mmol) in 1(N) aqueous NaOH (80 mL) and dioxane (40 mL) at 0 ^{o}C was treated with ditert butyl dicarbonate



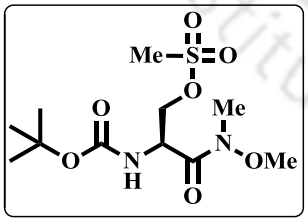
(17ml, 74.22 mmol) and the mixture was allowed to warm to room temperature and stirred for 24 hours. The dioxane was then evaporated and the aqueous layer was washed with ethyl acetate to remove ditert butyl dicarbonate. Then aqueous layer cooled to 0 ^{o}C and acidified with a 1(M) H_2SO_4 solution (pH 2-3) and extracted with ethyl acetate at least three times (50 ml). Then organic layer was treated with brine solution and dried over Na_2SO_4. Evaporated under vacuum to give viscous liquid and on standing it forms white solid. A white solid; 10gm,Yield 86 %; m.p. 88-91 ^{o}C. IR (KBr) $\bar{\upsilon}$ 3382, 3008, 2973, 2929, 2654, 1722, 1700,

1625, 1527, 1460, 1438, 1397, 1371 cm^{-1}. ^1H NMR ($CDCl_3$; 600 MHz) δ 6.43 (1H, s), 5.92 (1H, s), 5.60(1H, s), 4.34-4.20 (1H, m), 4.03-3.83 (2H, m), 1.43 (9H, s); ^{13}C NMR ($CDCl_3$; 150 MHz) δ 174, 156.5, 80.8, 63.1, 55.7, 28.5 +ESI-HRMS (m/z) calculated for $C_8H_{15}NO_5Na$ $[M+Na]^+$ 228.0848, found 228.0843.

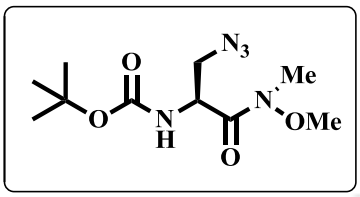
(S) N,C-diprotected serine (2.035): At first compound-**2.034** (7000mg, 34.13 mmol) was loaded in 100 ml single Neck round bottom flask (R.B) and then dry DCM (70 mL) was added to it. After that N, O dimethyl hydroxyl amine hydrochloride (3460mg, 35.49 mmol) was added to the stirred solution. Then the solution was treated with N-methyl morpholine (4.7ml, 36.86 mmol) under N_2 at -10 ^{o} C and to this resulting solution EDC.HCl (7078mg, 36.86mmol was added five portions over 30 minutes (36.86 mmol). After stirring at -10 ^{o} C for 2 h the reaction mixture was treated with 1(M) aqueous HCl solution and extracted with 200 mL DCM. After that organic layer was treated with saturated $NaHCO_3$ and brine solution and dried over Na_2SO_4. After Evaporation of DCM to give light greenish solid and the residue was purified by column chromatography on silica gel (60-120) using hexane/ethyl acetate (R_f = 0.4 in 1:2 hexane/ethyl acetate). A white solid; 8000mg,Yield 94 %; m.p. 112-116 ^{o} C. IR (KBr) ῡ 3473, 3357, 3003, 2978, 2944, 1703, 1647, 1538, 1467, 1448, 1394 cm^{-1}. ^1H NMR (600 MHz; $CDCl_3$) δ 5.64 (1H, d, J = 7.1 Hz), 4.78 (1H, s), 3.80 (2H, d, J = 2.6 Hz), 3.76 (3H, s), 3.21 (3H, s), 2.90 (1H, s), 1.43 (9H, s); ^{13}C NMR (150 MHz; $CDCl_3$) δ 171.1, 156.0, 80.2, 63.8, 61.7, 52.6, 32.3, 28.5. +ESI-HRMS (m/z) calculated for $C_{10}H_{20}N_2O_5Na$ $[M+Na]^+$ 271.1270, found 271.1277.



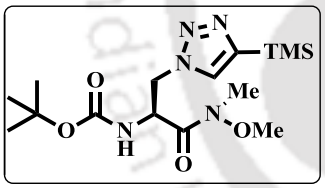
(S)N,C-diprotected O-mesyl serine (2.036): At First compound-**2.035** (4000mg, 16.12mmol) was loaded in 50 mL single neck round bottom flask. Then 50 mL dry DCM was added it and cool it at 0 ^{o} C in an ice bath. Then Et_3N (2.8ml, 19.67 mmol) was added to it.After that methyl sulphonyl chloride (1.6ml, 19.67 mmol) was added to the reaction mixture dropwise under N_2 atmosphere. The whole reaction mixture was stirred under N_2 atmosphere for 2 h in an ice bath. After completion of reaction, the mixture was extracted with DCM and washed with brine and saturated $NaHCO_3$ solution. After evaporation of DCM to give yellowish liquid and the residue was purified by column chromatography on silica gel (60-120) using hexane/ethyl acetate (R_f = 0.4 in 1:1 hexane/ethyl acetate). Light green gummy compound; 5100mg, Yield 96 %. IR (KBr) ῡ 3423, 2978, 2939, 2817, 1713, 1665, 1513, 1459, 1392, 1358 cm^{-1}.^1H NMR (600 MHz; $CDCl_3$) δ 5.57(1H, d, J = 7.7 Hz), 4.91 (1H, s), 4.45 – 4.38 (2H, m), 3.74 (3H, s), 3.21 (3H, s), 3.00 (3H, s), 1.41 (9H, s); ^{13}C NMR (150 MHz; $CDCl_3$) δ 168.4, 155.2, 80.4, 68.4, 61.8, 50.6, 37.8, 32.4, 28.4. +ESI-HRMS (m/z) calculated for $C_{11}H_{23}N_2O_7SNa$ $[M+Na]^+$ 349.1040, found 349.1072.



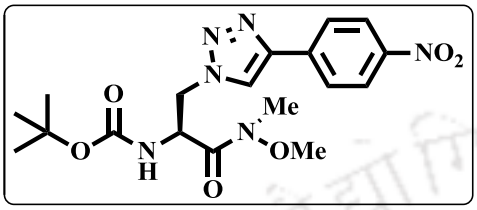
(S)N,C-diprotected azido alanine (2.037): At first compound-**2.036** (5100mg,15.63mmol) was loaded in 100 mL single neck round bottom flask, dry DMF (15 mL) was added to it. Then degassed under high vacuum for 10 minutes and N_2 atmosphere was created. After that NaN_3 (3050mg, 46.91 mmol) was added to it and again degassed in high vacuum. Then reaction mixture was refluxed at 65 ^{o}C for 24 h under N_2 atmosphere. After completion of the reaction the reaction mixture was extracted with ethyl acetate and washed with brine solution and evaporation of ethyl acetate to give light yellowish liquid and residue was purified by column chromatography on silica gel (60-120) using hexane/ethyl acetate (R_f = 0.4 in 4:1, hexane/ethyl acetate). Light green liquid; 3200mg, Yield 75 %. IR (KBr) ῡ 3422, 3058, 2980, 2939, 2105, 1714, 1666, 1499, 1455, 1392 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 5.49 (1H, d, J = 7.8 Hz), 4.84 (1H, d, J = 3.4 Hz), 3.76 (3H, s), 3.55 (2H, ddd, J = 41.1, 12.4, 4.7 Hz), 3.23 (3H, s), 1.44 (9H, s); ^{13}C NMR (150 MHz; $CDCl_3$) δ 169.6, 155.1, 80.0, 61.6, 52.3, 50.6, 32.2, 28.2, 28.2. ESI-HRMS (m/z) calculated for $C_{10}H_{19}N_5O_4Na$ $[M+Na]^+$ 296.1335, found 296.0961.



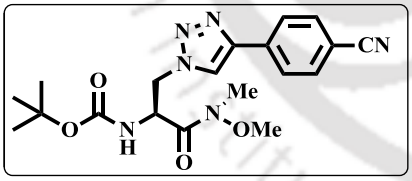
(S)Trimethylsilyl triazolyl alanine (2.038): At first compound-**2.037** 1500mg (5.5 mmol) was loaded in 100 mL round bottom flask and dry t-BuOH (8ml) and Et_3N (2ml) added to it. Then TMS acetylene (13.73 mmol) followed by sodium ascorbate (3.3 mmol) and $CuSO_4$, $5H_2O$ (0.55 mmol) were added to reaction mixture and reaction was stirred 50 ^{o}C for 48 h. After complication of reaction, the reaction mixture was dried under high vacuum. Residue was extracted with ethyl acetate and washed with saturated ammonium chloride and brine solution. Then residue was purified on silica gel (60-120) using hexane/ethyl acetate (R_f = 0.4 in 1:3, hexane/ethyl acetate). A white semi solid compound; 1245mg, Yield 60 %. IR (KBr) ῡ 3325, 3129, 3009, 2962, 2930, 2856, 2695, 1688, 1658, 1490, 1434, 1392, 1366, 1329, 1307, 1250, 1196, 1169 cm^{-1}.^{1}H NMR (600 MHz; $CDCl_3$) δ 7.49 (1H, s), 5.52 (1H, d, J = 6.4 Hz), 5.02 (1H, dd, J = 13.4, 5.6 Hz), 4.66 (2H,d, J = 4.9 Hz), 3.69 (3H, s), 3.17 (3H, s), 1.36 (9H, s), 0.25 (9H, s); ^{13}C NMR (150 MHz; $CDCl_3$) δ 169.1, 155.1, 146.5, 129.9, 80.3, 61.8, 51.2, 50.4, 32.5, 28.3, -1.0. +ESI-HRMS (m/z) calculated for $C_{15}H_{30}N_5O_4Si$ $[M+H]^+$ 372.2062, found 372.2062.



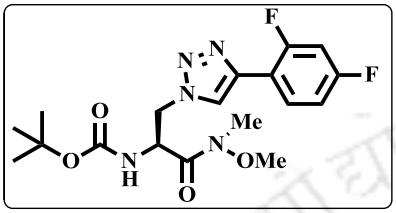
(S)p-Nitrophenyl triazolyl alanine (2.039): At first compound-**2.037** 1000mg (3.66 mmol) was loaded in 50 mL round bottom flask. Then Dry THF (25ml) added to it. The degassed 10 minutes in high vacuum. Then 1-ethynyl-4-nitrobenzene (4.02 mmol) and DIPEA (4.39 mmol) were added to it. Then sodium ascorbate (2.19 mmol) and $CuSO_4$, 5 H_2O (0.366 mmol) in 1ml water were added to reaction mixture and reaction mixture was degassed for 5 minutes, followed by refluxed at 55 ^{o} C for 20 h. After complication of reaction, the reaction mixture was dried under high vacuum. Residue was extracted with ethyl acetate and washed with saturated ammonium chloride and brine solution. Then residue was purified on silica gel (60-120) using hexane/ethyl acetate (R_f = 0.4 in 1:1, hexane/ethyl acetate). A yellowish solid compound; 1275mg, Yield 82 %; m.p. 163-164 ^{o} C. IR (KBr) $\bar{\upsilon}$ 3448, 3414, 3150, 3128, 3106, 2984, 2934, 2445, 2292, 1690, 1667, 1608, 1514, 1461, 1432, 1389, 1351 cm^{-1}. ^{1}H NMR ($CDCl_3$; 600 MHz) δ 8.18 – 8.13 (2H, m), 8.01 (1H, s), 7.91 (2H, d, J = 8.8 Hz), 5.74 (1H, d, J = 4.7 Hz), 5.08 (1H, d, J = 3.4 Hz), 4.78 – 4.66 (2H, m), 3.72 (3H, s), 3.16 (3H, s), 1.29 (9H, s); ^{13}C NMR ($CDCl_3$; 150 MHz) δ 168.7, 155.1, 147.1, 145.4, 136.9, 126.1, 124.2, 122.5, 80.4, 62.0, 51.2, 51.1, 32.4, 28.2. +ESI-HRMS (m/z) calculated for $C_{18}H_{25}N_6O_6$ $[M+H]^+$ 421.1757, found 421.1859.



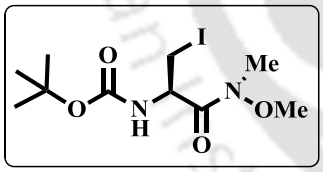
(S)p-Cyanophenyl triazolyl alanine (2.040): At fist compound-**2.037** 100mg (3.66mmol) was loaded in 50 mL round bottom flask. Then Dry THF (25ml) added to it and degassed 10 minutes in high vacuum. Then 4-ethynylbenzonitrile (4.02mmol) and DIPEA (4.39mmol) were added to it. Then sodium ascorbate (2.2mmol) and $CuSO_4$, $5H_2O$ (0.36mmol) in 1ml water were added to reaction mixture and reaction mixture was degassed for 5 minutes. Then whole reaction was refluxed 55 ^{o} C for 20 h. After complication of reaction, the reaction mixture was dried under high vacuum. Residue was extracted with ethyl acetate and washed with saturated ammonium chloride and brine solution. Then residue was purified on silica gel (60-120) using hexane /ethyl acetate (R_f = 0.4 in 1:1, hexane/ethyl acetate). A white solid compound; 1375mg, Yield 93 %; m.p. 143-144 ^{o} C. IR (KBr) $\bar{\upsilon}$ 3373, 3122, 3122, 3104, 2978, 2937, 2222, 1698, 1673, 1655, 1613, 1514, 1478, 1458 cm^{-1}.^{1}H NMR ($CDCl_3$; 600 MHz) δ 7.90 (3H, d, J = 8.1 Hz), 7.66 (2H, d, J = 8.1 Hz), 5.60 (1H, d, J = 4.2 Hz), 5.08 (1H, d, J = 4.1 Hz), 4.74 (2H, d, J = 4.4 Hz), 3.76 (3H, s), 3.21 (3H, s), 1.37 (9H, s).; ^{13}C NMR ($CDCl_3$; 150 MHz) δ 168.7, 155.2, 146.0, 135.1, 132.8, 126.2, 122.0, 119.0, 111.5, 80.7, 62.0, 51.43-51.30, 32.6, 28.3. +ESI-HRMS (m/z) calculated for $C_{19}H_{25}N_6O_4$ $[M+H]^+$ 401.1932, found 401.1953.



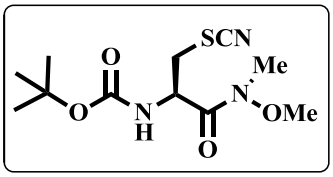
(S) o,p-Difluorophenyl triazolyl alanine (2.041): At first compound-**2.037,** 220mg (0.806 mmol) was loaded in 50 mL round bottom flask. Then Dry THF (10ml) added to it. Then 1-ethynyl-2, 4-difluorobenzene (0.887 mmol) and DIPEA were added to it. After that sodium ascorbate (0.484 mmol) and $CuSO_4$, 5 H_2O (0.0806 mmol) in 1ml water were added to reaction mixture and reaction mixture was degassed 5 minutes followed by refluxed at 55 ^{o}C for 20 h. After complication of reaction, the reaction mixture was dried under high vacuum. Residue was extracted with ethyl acetate and washed with saturated ammonium chloride and brine solution. Then residue was purified on silica gel (60-120) using hexane/ethyl acetate (R_f = 0.4 in 1:1, hexane/ethyl acetate). A light greenish semi solid compound; 235mg, Yield 70 %. IR (KBr) ῡ 3380, 3162, 3081, 2981, 2935, 2870, 2853, 1681, 1632, 1602, 1561, 1493, 1463, 1443, 1421, 1387 cm^{-1}. ^{1}H NMR ($CDCl_3$; 600 MHz) δ 8.22-8.18 (1H, q, J =8.4 Hz), 7.84 (1H, s), 6.93 (1H, t, J =16.2 Hz), 6.84 (1H, t, J = 18.6 Hz), 5.65-5.66 (1H, d, J =7.2 Hz), 5.04 (1H, s), 4.73-4.73 (2H, d, J = 3.6 Hz), 3.740 (3H, s), 3.199 (3H, s), 1.367 (9H, s); ^{13}C NMR ($CDCl_3$; 150 MHz) δ 168.8, 163.4-163.3, 161.7-161.6, 160.1-160.0, 158.4-158.4, 155.1, 140.5, 128.9-128.8, 123.4-123.3, 115.1-115.1, 112.1-111.9, 104.2-103.9, 80.5, 62.0, 51.4, 51.0, 32.5, 28.3. +ESI-HRMS (m/z) calculated for $C_{18}H_{24}F_2N_5O_4$ $[M+H]^+$ 412.1791, found 412.1799.



(S)N,C-diprotected iodoalanine (2.042): At first compound-2.036 500mg (1.5 mmol) was loaded in 100 mL R.B. Then 10 minutes degassed in high vacuum. After that dry acetone (5 mL) was added to it and degassed in high vacuum for 5 minutes. Then dry NaI (7.5 mmol) was added then to it. Again reaction mixture was degassed in high vacuum for 5 minutes. Finally the reaction mixture was stirred at 30 ^{o} C for 24 h in dark under nitrogen atmosphere. Then reaction mixture was dried under high vacuum and 100 ml $CHCl_3$ was added to the reaction mixture, followed by filtered through what Mann filter paper and organic phase was washed with saturated sodium thiosulphate solution. After evaporation of $CHCl_3$, residue (R_f = 0.4 in 2:1, hexane/ethyl acetate) was purified on silica gel (60-120) using hexane/ethyl acetate. Amber color liquid; Yield 90 %. ^{1}H NMR (600 MHz; $CDCl_3$) δ 5.45 (1H, d, J = 8.2 Hz), 4.77 (1H, d, J = 3.2 Hz), 3.75 (3H, s), 3.47 (1H, dd, J = 10.1, 5.1 Hz), 3.37 (1H, dd, J = 9.7, 5.6 Hz), 3.22 (3H, s), 1.42 (9H, s); ^{13}C NMR ($CDCl_3$; 150 MHz) δ 169.0, 155.0, 80.3, 61.9, 51.2, 28.4, 5.9. +ESI-HRMS (m/z) calculated for $C_{10}H_{20}IN_2O_4$ $[M+H]^+$ 359.0468, found-359.1791

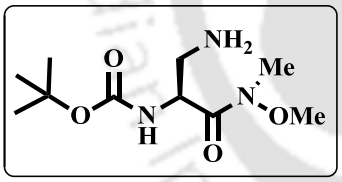


(S) N,C-diprotected thiocyanyl alanine (2.043, ^{SCN}Ala): At first compound-**2.042** 1200mg (3.35mmol) was loaded in 100 mL R.B. Then 10 minutes degassed in high vacuum.



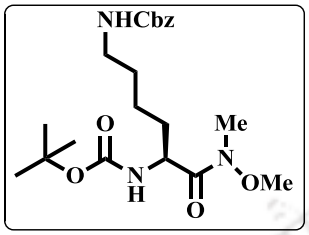
After that dry acetone (10 mL) was added to it and again degassed in high vacuum for 5 minutes. Then preheated molecular sieve was added to it. Then KSCN (16.75 mmol added then) was to it. Again reaction mixture was degassed in high vacuum for 5 minutes. Finally the reaction mixture was stirred at 55 ^{o} C for 24 h. Then reaction mixture was dried under high vacuum. Then ethyl acetate was added to the reaction mixture and filtered through what Mann filter paper and organic phase was washed with saturated brine solution. Then residue (R_f = 0.4 in 2:1, hexane/ethyl acetate) was purified on silica gel (60-120) using hexane/ethyl acetate. A yellowish solid compound; 800mg, Yield 82 %; m.p. 119-123 ^{o} C. IR (KBr) ῡ 3347, 2990, 2979, 2925, 2853, 2151, 1715, 1663, 1529, 1458, 1413, 1391 cm^{-1}.^{1}H NMR (600 MHz; $CDCl_3$) ^{1}H NMR (600 MHz, $CDCl_3$) δ 5.58 (1H, d, J = 4.5 Hz), 4.99 (1H, d, J = 4.0 Hz), 3.79 (3H, s), 3.38 (1H, dd, J = 13.7, 4.0 Hz), 3.29 (1H, d, J = 4.8 Hz), 3.25 (3H, s), 1.45 (9H, s); ^{13}C NMR (150 MHz; $CDCl_3$) δ 168.8, 155.1, 112.0, 80.8, 62.1, 50.6, 36.9, 32.5, 28.4. +ESI-HRMS (m/z) calculated for $C_{11}H_{20}N_3O_4S$ $[M+H]^+$ 290.109, found-290.1118.

(S) N,C-diprotected aminoalanine (2.044): At first compound-**2.037** (1700 mg, 6.22 mmol) was loaded in a two neck R.B and dry methanol was added to it. Evacuation unde high vacuum pump for 10 minutes and degassing by N_2 gas, Pd/C (132 mg, 1.24 mmol) was added

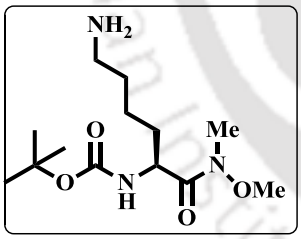


and H_2 atmosphere was created by a balloon. The reaction mixture was stirred at room temperature for 7 h after which the reaction mixture was filtered through Whatmann filter paper, solvent was evaporated under high vacuum to afford white semi solid product (R_f = 0.4 in 1:3, $CHCl_3$/MeOH) which was directly used for the next step without further purification. White semi solid compound; 1350mg, Yield 88 %; IR (KBr) ῡ 3400, 3342, 3002, 2978, 2932, 1704, 1646, 1530, 1455, 1394, 1365, 1287, 1251 cm^{-1}.^{1}H NMR (600 MHz; $CDCl_3$) δ 5.50(1H, d, J = 7.1 Hz,), 4.69 (1H, s), 3.76 (s, 3H), 3.20 (3H, s), 3.04 – 2.98 (1H, m), 2.85 (1H, dd, J = 13.3, 5.8 Hz), 1.75 (2H, s), 1.42 (9H, s); ^{13}C NMR (150 MHz; $CDCl_3$) δ 171.6, 155.9, 79.9, 61.8, 53.2, 44.3, 32.2, 28.5. +ESI-HRMS (m/z) calculated for ($C_{10}H_{22}N_3O_4$) $(M+H)^+$ 248.1610, found-248.1606.

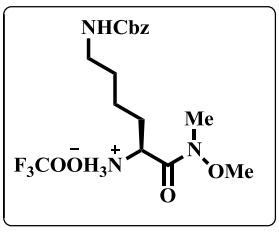
(S) $^{\omega}$N, $^{\alpha}$N, C-triprotected lysine (2.049): At first compound-**2.048** (1700mg, 4.47 mmol) was loaded in 100 ml single Neck round bottom flask and then dry DCM (30 mL) was added to it. After that N, O dimethyl hydroxyl amine hydrochloride (456mg, 4.64mmol) was added to the stirred solution. Then the solution was treated with N-methyl morpholine (0.750ml, 4.8mmol) under N_2 at -10 ^{o}C and to this resulting solution EDC. HCl (928mg, 4.8mmol was added five portion over 30 minutes .After stirring at -10 ^{o}C for 3 h. After that reaction mixture was treated with 1(M) aqueous HCl solution and extracted with 200 mL DCM. After that organic layer was treated with saturated $NaHCO_3$ and brine solution and dried over Na_2SO_4. After evaporation of DCM to give light greenish liquid and the residue (R_f = 0.4 in 1:2, hexane/ethyl acetate) was purified by column chromatography on silica gel (60-120) using hexane/ethyl acetate. Transparent liquid compound; 1670mg, Yield 88 %. IR (KBr) $\bar{\upsilon}$ 3337, 3064, 3003, 2938, 2867, 2825, 1718, 1538, 1454, 1390 cm^{-1}.^{1}H NMR (600 MHz; $CDCl_3$) δ 7.34 (4H, d, *J* = 4.3 Hz), 7.30 (1H, q, *J* = 4.6 Hz), 5.24 (1H, d, *J* = 9.2 Hz), 5.07 (2H, s), 5.00 – 4.89 (1H, m), 4.64 (1H, q, *J* = 7.8, 7.1 Hz), 3.75 (3H, s), 3.18 (5H, s), 1.93 – 1.82 (1H, m), 1.69 (1H, p, *J* = 6.3 Hz), 1.51 (3H, ddp, *J* = 20.6, 13.5, 6.8 Hz), 1.41 (10H, s). +ESI-HRMS (m/z) calculated for $C_{21}H_{34}N_3O_6$ $[M+H]^+$ 424.2448, found 424.2449.



(S) $^{\alpha}$N,C-diprotected lysine (2.050): At first Compound **2.049** (1500mg, 3.55mmol) was loaded in preheated dried single neck 50ml round bottom flask and dry methanol was added to it. After that it was degassed under high vacuum and N_2 atmosphere was created. Then under N_2 atmosphere 10%-Pd/C (76mg, 0.70mmol) was added and again degassed under high vacuum and H_2 atmosphere was created by balloon. The reaction mixture was stirred at room temperature (rt) for 8 h. After completion, the reaction mixture was filtered through Whatmann filter paper and evaporation of methanol in high vacuum gave white semi solid product and compound was directly used in next step of reaction without further purification. Colorless semi solid; 100mg, Yield 97 %.



***(S)*$^{\omega}$N,C-diprotected lysine *trifluoroacetate salt (2.053)*:** At first compound-**2.049** (**5**00mg, (0.90mmol) was loaded into 50 ml single neck round bottom flask and DCM (2ml) and TFA (1ml) were added. It was allowed to stir at room temperature for about 4 h. Then the RM was dried at high vacuum and to the obtained compound, chloroform (30ml, 3 times) was added and evaporated to obtain brownish gummy compound. The compound was directly used for next reaction without any purification.

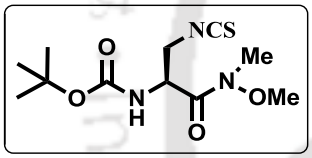


Weight of the salt = 630mg.

General synthesis procedure for Isothiocyanyl derivative:

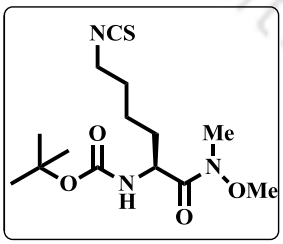
At first amino, the compound was loaded in dry 50 mL round bottom flask under nitrogen atmosphere and dry THF was added to it. Then reaction mixture was cooled to 0 ^{o} C and Et_3N 3.3 equivalent was added to it, after 5 minute CS_2 2 equivalent was added to it and stirred for 1 hour at 0 ^{o} C. After that at 0 ^{o} C tosyl chloride 1.2 equivalents was added to it and whole reaction mixture was stirred at 0 ^{o} C for 5 minutes to 30 minutes. Afterward, reaction mixture was dried under high vacuum and ethyl acetate was added to the reaction mixture and organic layer was treated with 1(N) HCl followed by saturated $NaHCO_3$ and brine solution and dried over Na_2SO_4. After Evaporation of ethyl acetate, the residue was purified by column chromatography on silica gel (60-120) using hexane/ethyl acetate.

(S)N,C-diprotected isothiocyanyl alanine (2.046, **^{NCS}Ala*)***: Using general synthesis for isothacyanyl and starting from 830mg (3.35mmol) of compound **2.044**. The compound **2.046**



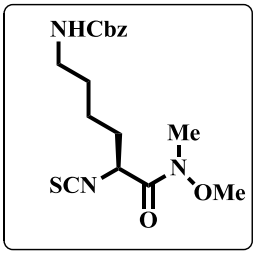
was obtained (R_f = 0.4 in 2:1, hexane/ethyl acetate) within 5 minutes. A light greenish solid compound; 800mg, Yield 82 % ; m.p. 80-84 ^{0} C. IR (KBr) ῡ 3343, 3009, 2980, 2937, 2223, 2120, 1697, 1658, 1526, 1462, 1430, 1391 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 5.54 (1H, d, J = 7.5 Hz), 4.88 (1H, d, J = 3.5 Hz), 3.84 (1H, dd, J = 14.3, 4.4 Hz), 3.76 (4H, s), 3.25 (3H, s), 1.45 (9H, s); ^{13}C NMR (150 MHz; $CDCl_3$) δ 168.8, 155.1, 134.1, 80.7 62.0, 50.6, 47.1, 32.5, 28.5. +ESI-HRMS (m/z) calculated for $C_{11}H_{20}N_3O_4SNa$ $[M+Na]^+$ 312.0994, found= 312.0998.

(S)N,C-diprotected isothiocyanyl lysine (2.052, **^{NCS}Lys*)***: Using general synthesis for isothiocyanyl and start from 900mg (3.11mmol) of **2.050**, the desired compound **2.052** was

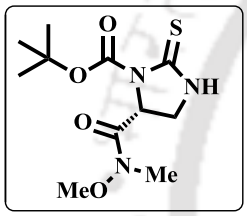


obtained (R_f = 0.4 in 1:1, hexane/ethyl acetate). A light green gummy liquid; 880mg, Yield 85 %. IR (KBr) ῡ 3434, 2975, 2983, 2861, 2148, 2105, 1709, 1660 1503, 1456, 1389, 1366 cm^{-1}.^{1}H NMR (600 MHz; $CDCl_3$) δ 5.21 (1H, d, J = 8.8 Hz), 4.62 (1H ,d, J = 4.0 Hz), 3.73 (3H, s), 3.47 (2H, t, J = 6.7 Hz), 3.16 (3H, s), 1.75–1.58 (3H, m), 1.56 – 1.40 (3H, m), 1.38 (9H, s); ^{13}C NMR (150 MHz; $CDCl_3$) δ 172.8, 155.6, 130.0, 79.7, 61.7, 49.9, 44.8, 32.2, 32.1, 29.5, 28.4, 22.4. +ESI-HRMS (m/z) calculated for $C_{14}H_{26}N_3O_4S$ $[M+H]^+$ 332.1644, found 332.1643.

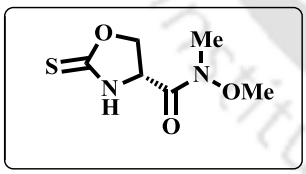
(S)N,C-diprotected α–isothiocyanyl lysine (2.055, $^{\alpha\text{-NCS}}$Lys): Using general synthesis for isothiocyanyl and start from 630mg (1.48 mmol) of **2.053**, the desired compound **2.055** was obtained (R_f = 0.4 in 1:1, hexane/ethyl acetate). A light green gummy liquid; 410mg, Yield 75 %. IR (KBr) ῡ 3447, 3063, 3033, 2863, 2072, 1718, 1537, 1389, 1455, 1389, 1323 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 7.28 (5H, d, J = 18.9 Hz), 5.21 (1H, s), 5.03 (2H, s), 4.50 (1H, s), 3.65 (3H, s), 3.15 (5H, d, J = 2.2 Hz), 1.81 (2H, d, J = 3.3 Hz), 1.53–1.31 (4H, m); ^{13}C NMR (150 MHz; $CDCl_3$) δ 168.0, 156.4, 136.7, 136.5, 128.4, 127.9, 127.9, 66.4, 61.6, 57.4, 40.4, 32.5, 32.4, 29.0, 22.7.+ESI-HRMS (m/z) calculated for $C_{17}H_{24}N_3O_4S$ $[M+H]^+$ 366.1488, found 366.1512.



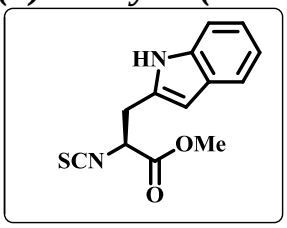
(R)-tert-butyl5-(methoxy(methyl)carbamoyl)-2-thioxoimidazolidine-1-carboxylate (2.047) : Using general synthesis for isothacyanyl and starting from 830mg (3.36 mmol) of compound **2.044**, the compound **2.047** was obtained (R_f = 0.4, 1:5, hexane/ethyl acetate) within 30 minutes. On standing it forms white crystal. A white solid compound; 850mg, Yield 87 %; m.p. 160-164 ^{o} C. IR (KBr) ῡ 3356, 2975, 2938, 1749, 1669, 1510, 1477, 1438, 1368, 1342, 1396 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$+DMSO-d_6) δ 7.27 (1H, s), 5.16 (1H, dd, J = 11.0, 5.1 Hz), 3.83 – 3.74 (1H, m), 3.64 (3H, s), 3.33 (1H, dd, J = 11.0, 5.2 Hz), 3.12 (3H, s), 1.37 (9H, s).^{13}C NMR (150 MHz; $CDCl_3$+ DMSO-d_6) δ 180.9, 169.7, 149.7, 83.7, 61.4, 58.7, 44.5, 32.5, 27.8.+ESI-HRMS (m/z) calculated for $C_{11}H_{20}N_3O_4S$ $[M+H]^+$ 290.1175, found 290.1188.



(R)-N-methoxy-N-methyl-2-thioxooxazolidine-4-carboxamide (2.057):
Using general synthesis for isothiocyanyl and start from 800mg (3.05 mmol) of **2.056**, the desired compound **2.057** was obtained (R_f = 0.4 in 1:5, hexane/ethyl acetate). A light green gummy liquid; 350mg, Yield 60 %. IR (KBr) ῡ 3242, 2966, 2996, 2922, 2852, 1760, 1666, 1516, 1394, 1324 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 8.45 (1H, s), 4.94 (1H, dd, J = 9.6, 7.2 Hz), 4.86 (1H, t, J = 9.7 Hz,), 4.69 – 4.63 (1H, m), 3.71 (3H, s), 3.22 (3H, s); ^{13}C NMR (150 MHz; $CDCl_3$) δ 189.9, 168.4, 72.2, 61.7, 57.1, 32.6. +ESI-HRMS (m/z) calculated for $C_6H_{11}N_2O_3S$ $[M+H]^+$ 191.0490, found 191.0490.

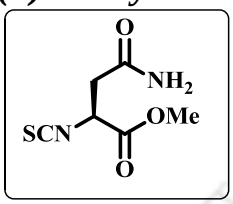


(S)-methyl 3-(1H-indol-2-yl)-2-isothiocyanatopropanoate(2.066):
Using general synthesis for isothiocyanyl and start from 500mg (1.96 mmol) of **2.065**, the desired compound **2.066** was obtained (R_f = 0.4 in 1:2, hexane/ethyl acetate). A light yellow semi solid; 385mg, Yield 75 %. IR (KBr) ῡ 3416, 1953, 3055, 2953, 2925, 2853, 2226, 2076, 1745, 1619, 1546, 1488, 1457 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 8.21 (1H,

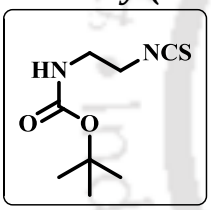


s), 7.58 (1H, d, *J* = 7.6 Hz), 7.37 (1H, d, *J* = 7.9 Hz), 7.23 (1H, t, *J* = 7.2 Hz), 7.17 (1H, t, *J* = 7.2 Hz), 7.13 (1H, s), 4.63 – 4.55 (1H, m), 3.76 (3H, s), 3.50 – 3.41 (1H, m), 3.36 (1H, dd, *J* = 14.5, 7.8 Hz); ^{13}C NMR (150 MHz; $CDCl_3$) δ 168.9, 137.3, 136.2, 126.9, 124.2, 122.4, 119.9, 118.4, 111.6, 109.1, 60.4, 53.3, 30.1. +ESI-HRMS (m/z) calculated for $C_{13}H_{13}N_2O_2S$ $[M+H]^+$ 261.0698, found 261.0687.

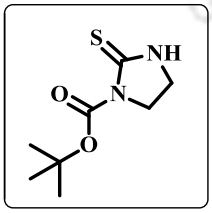
(S)-methyl4-amino-2-isothiocyanato-4-oxobutanoate(2.061): Using general synthesis for isothiocyanyl and start from 500mg (1.34 mmol) of **2.060**, the desired compound **2.061** was obtained (R_f = 0.4 in 1:3, hexane/ethyl acetate). A semi solid; 140mg, Yield 55 %. IR (KBr) $\bar{\upsilon}$ 3434, 3192, 2954, 2061, 1957, 1742, 1668, 1610, 1668, 1610, 1511, 1411, 1313 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 5.97 (1H, s), 5.86 (1H, s), 4.78 (1H, d, *J* = 4.6 Hz), 3.84 (3H, s), 2.93 – 2.75 (2H, m); ^{13}C NMR (150 MHz; $CDCl_3$) δ 170.1, 168.5, 139.4, 55.7, 53.7, 38.7. +ESI-HRMS (m/z) calculated for $C_6H_9N_2O_3S$ $[M+H]^+$ 189.0345, Found 189.0452.



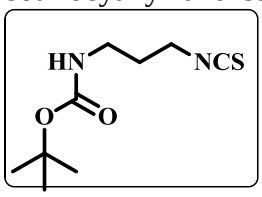
Tert-butyl(2-isothiocyanatoethyl)carbamate(2.072): Using general synthesis for isothiocyanyl and start from 240mg (1.5 mmol) of **2.071**, the desired compound**2.071** was obtained (R_f = 0.4 in 2:1, hexane/ethyl acetate). A liquid; 260mg, Yield 85 %. IR (KBr) $\bar{\upsilon}$ 3207, 2977, 2931, 2896, 2196, 2111, 1740, 1700, 1560, 1522, 1486, 1476, 1459, 1395, 1367, 1346, 1278 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 4.87 (1H, s), 3.47 (2H, t, *J* = 5.1 Hz,), 3.20 (2H, d, *J* = 5.4 Hz), 1.27 (9H, s); ^{13}C NMR (150 MHz; $CDCl_3$) δ 155.8, 132.1, 80.2, 77.4, 77.23, 77.0, 45.5, 40.6, 28.4. +ESI-HRMS (m/z) calculated for $C_8H_{15}N_2O_2S$ $[M+H]^+$ 203.0854, found 203.0825.



Tert-butyl 2-thioxoimidazolidine-1-carboxylate (2.073): At first compound **2.072** 60 mg, (0.30 mmol) was loaded into 25ml R.B. Then DCM: CH_3CN and Et_3N (0.90 mmol) were added to it. After that reaction mixture was refluxed at 50 ^{o} C for 12 h. Next reaction mixture was dried under vacuum and residue was purified by column chromatography on silica gel (60-120) using hexane/ethyl acetate. The desired compound **2.072** was obtained (R_f = 0.4 in 1:2, hexane/ethyl acetate). A semi solid compound; 49mg, Yield 80 %. IR (KBr) $\bar{\upsilon}$ 3370, 2969, 2915, 1724, 1641, 1506, 1477, 1385, 1369, 1340, 1306, 1275, 1233 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 7.72 (1H, s), 4.08 – 4.01 (2H, m), 3.56 (2H, t, *J* = 8.8 Hz), 1.51 (9H, s); ^{13}C NMR (150 MHz; $CDCl_3$) δ 181.2, 150.4, 83.5, 47.7, 41.1, 28.1. +ESI-HRMS (m/z) calculated for $C_8H_{15}N_2O_2S$ $[M+H]^+$ 203.0854 Found 203.0907.

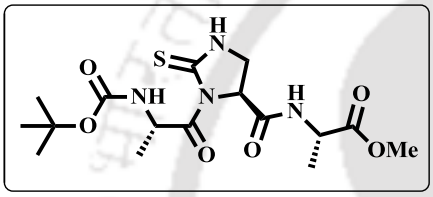


Tert-butyl (3-isothiocyanatopropyl)carbamate (2.075): Using general synthesis for isothiocyanyl and start from 604mg (3.5 mmol) of **2.074**, the desired compound **2.075** was



obtained (R_f = 0.4 in 6:1, hexane/ethyl acetate). A light green gummy liquid; 570mg, Yield 75 %. IR (KBr) ῡ 3375, 3010, 2999, 2985, 2973, 2940, 2198, 2150, 2090, 1674, 1517, 1473 cm^{-1}.^{1}H NMR (600 MHz; $CDCl_3$) δ 4.82 (1H, s), 3.56 (2H, t, J = 6.5 Hz,), 3.19 (2H, d, J = 6.0 Hz), 1.89 – 1.77 (2H, m), 1.39 (9H, s); ^{13}C NMR (150 MHz; $CDCl_3$) δ 156.1, 130.4, 79.6, 42.7, 37.7, 30.4, 28.4, 28.4. +ESI-HRMS (m/z) calculated for $C_4H_9N_2S$ [(M+H)-Boc]$^+$ 117.0486, found 117.0453.

(6S, 9S)-tert-butyl 6-(((S)-1-methoxy-1-oxopropan-2-yl)carbamoyl)-9-methyl-8-oxo3-thioxo-1,2,4,7-tetrazonane-1-carboxylate (2.070): Using general synthesis for isothiocyanyl and start from 100 mg (0.24 mmol) of **2.069** the desired compound **2.070** was obtained (R_f = 0.4 in 1:2, hexane/ethyl acetate). A gummy liquid; 58mg, Yield 60 %. IR (KBr) ῡ 3327, 2979, 2926, 2854, 2106, 1740, 1681, 1523, 1454, 1388, 1367.78, 1216, 1165, 1047 cm^{-1}.^{1}H NMR (600 MHz; $CDCl_3$) δ 7.73 (1H, d, J = 25.9 Hz,), 6.91 (1H, s), 6.49 – 6.37 (1H, m), 5.24 (1H, s), 5.16 (1H, t, J = 6.6 Hz), 4.54 (1H, p, J = 7.2 Hz), 3.83 (2H , d, J = 7.4 Hz), 3.74 (3H, s), 1.47 (3H, s), 1.43 – 1.39 (12H, m); ^{13}C NMR (150 MHz; $CDCl_3$) δ 179.5, 176.6, 173.2, 168.1, 155.4, 80.1, 60.8, 52.8, 49.0, 48.6, 44.6, 28.5, 19.3, 18.1. +ESI-HRMS (m/z) calculated for $C_{16}H_{27}N_4O_6S^+$ [M+H]$^+$ 403.1651, found 403.1604.



2.9.3. ^{1}H and ^{13}C NMR Spectra of Selected Amino acids and cyclic Derivatives

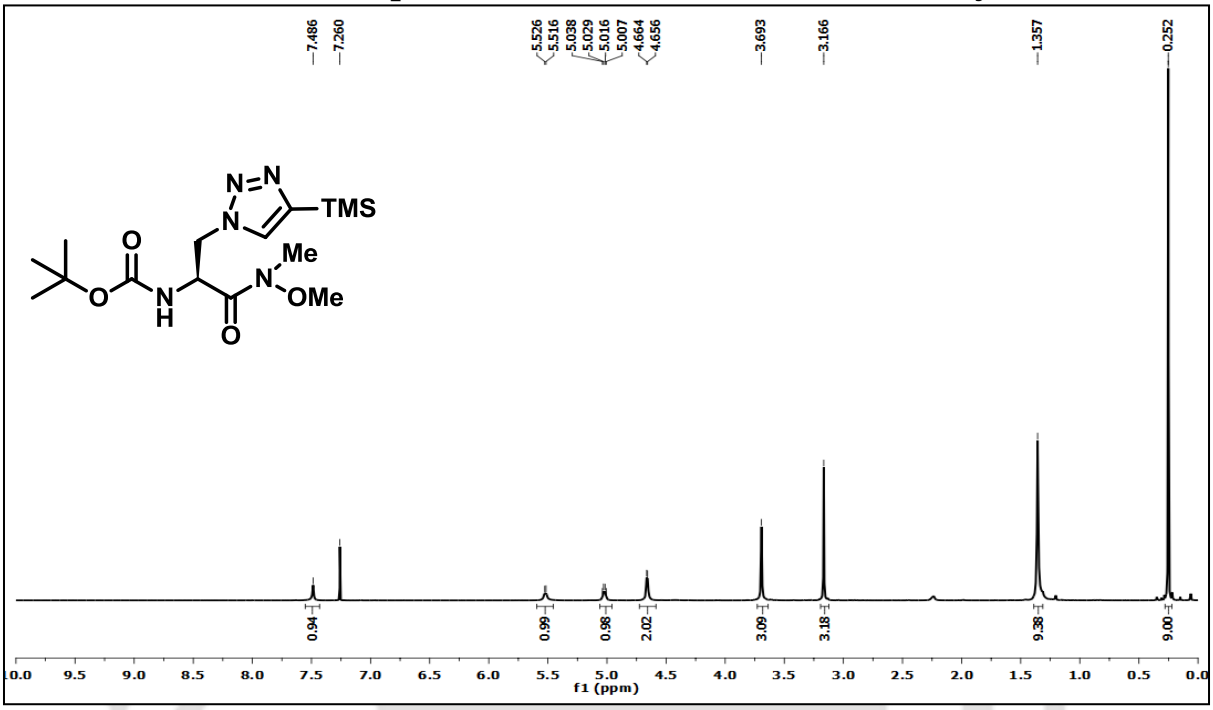


Figure 2.11. ^{1}H NMR spectra of synthesised compound **2.038.**

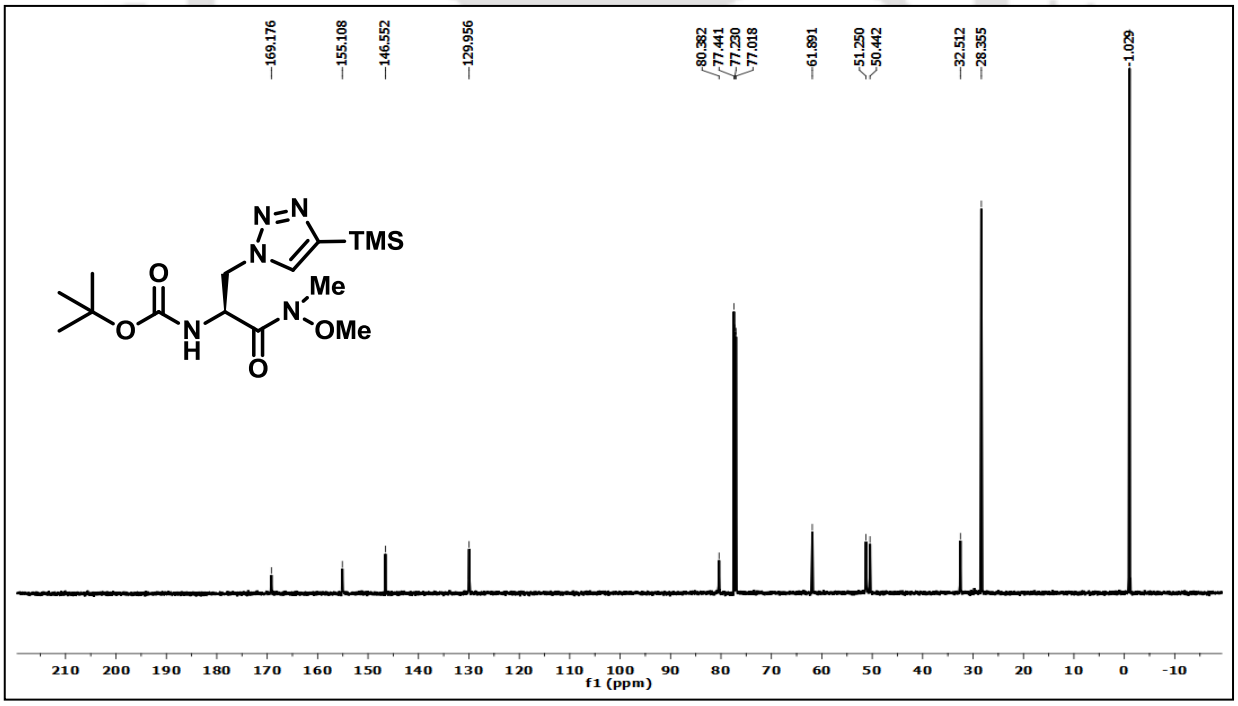


Figure 2.12. ^{13}C NMR spectra of synthesised compound **2.038.**

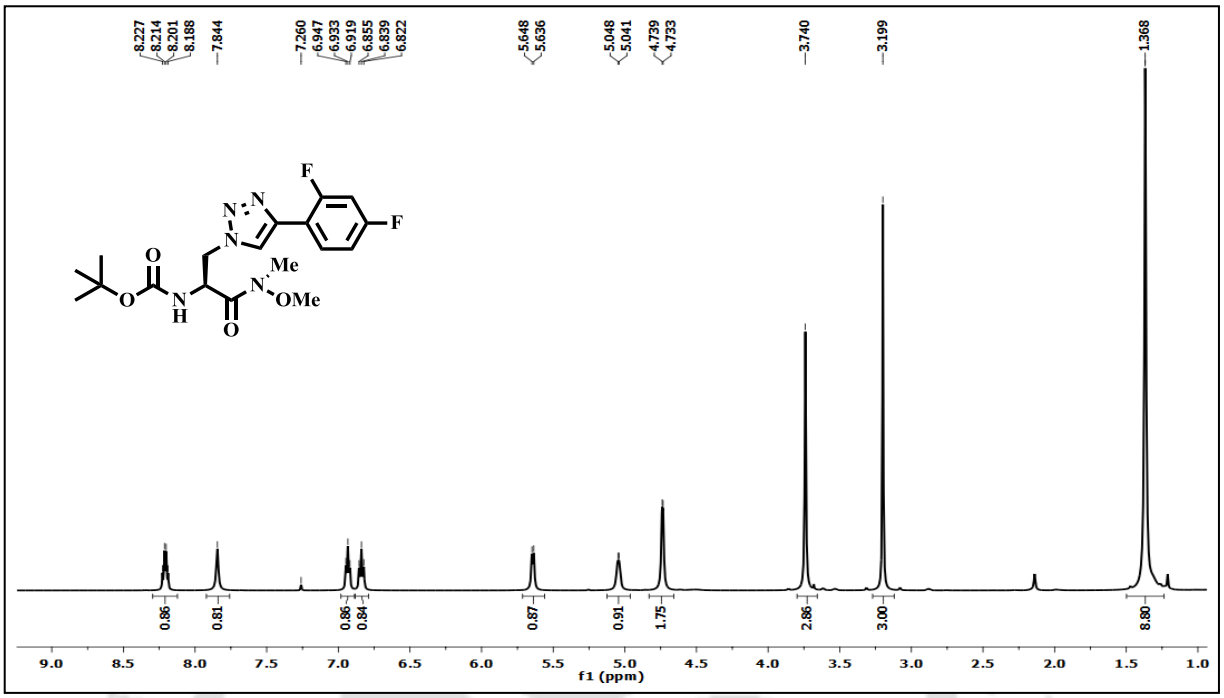


Figure 2.13. ^{1}H NMR spectra of synthesised compound **2.041.**

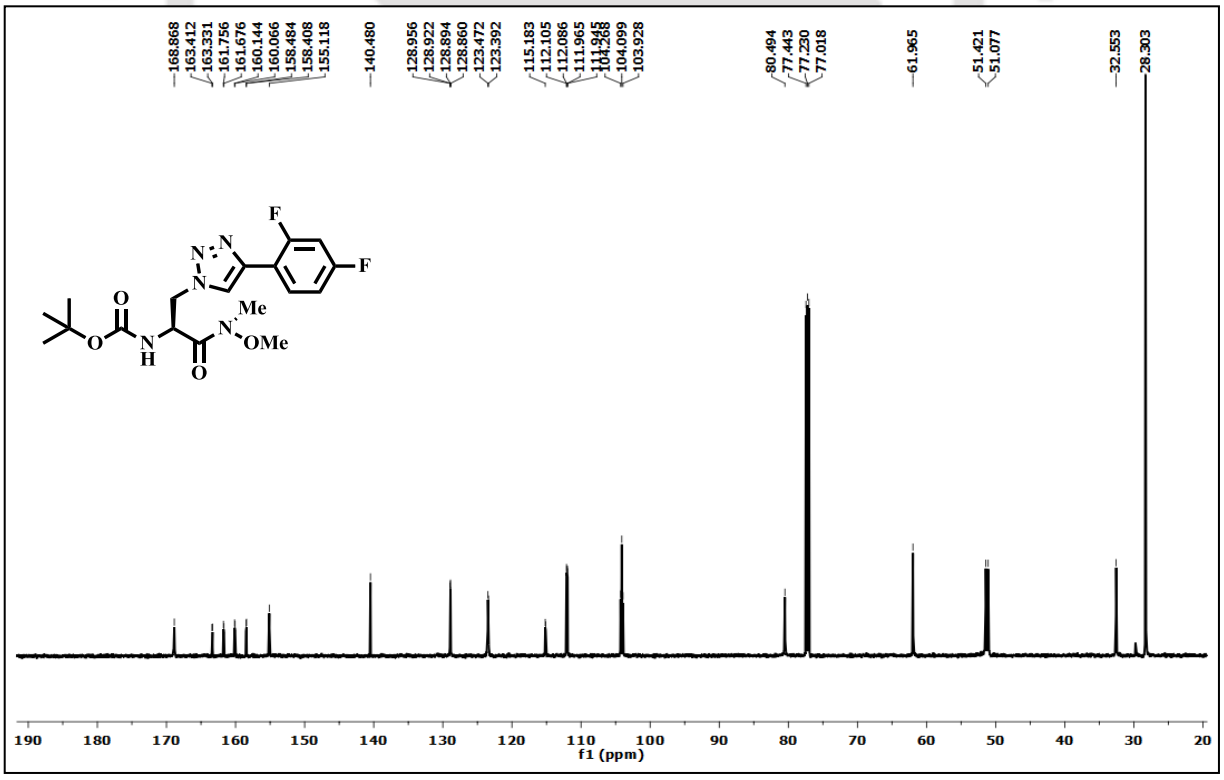


Figure 2.14. ^{13}C NMR spectra of synthesised compound **2.041.**

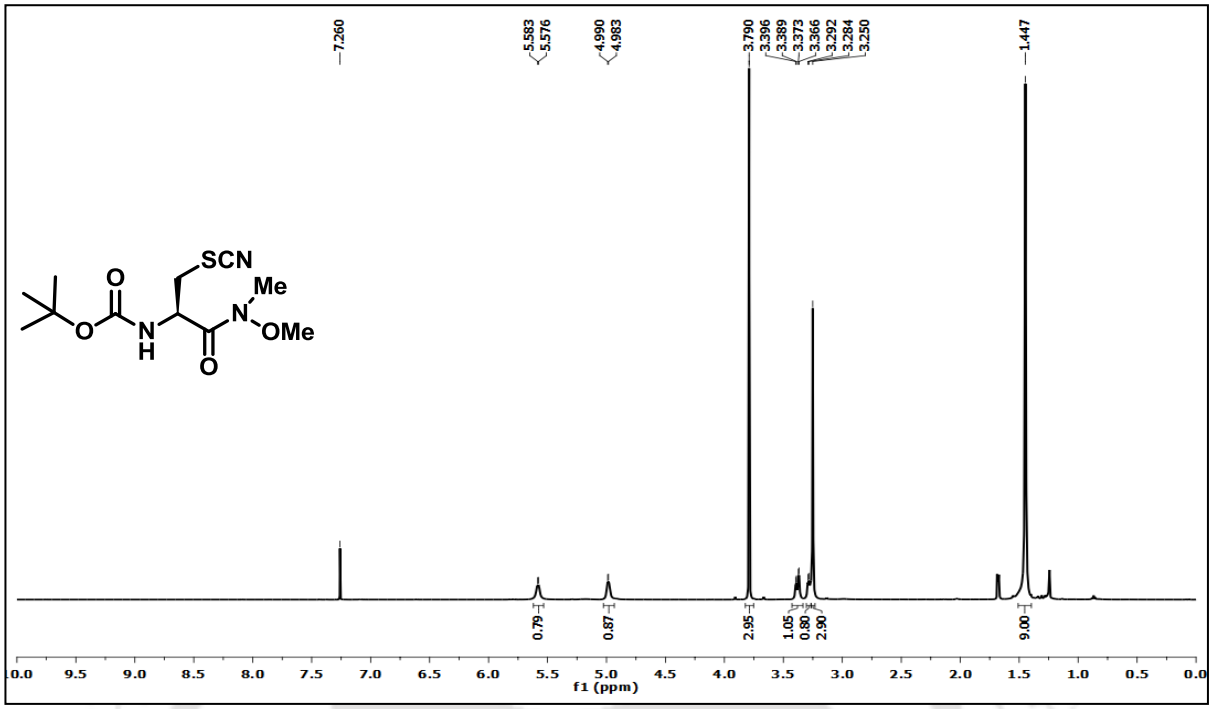


Figure 2.15. ^{1}H NMR spectra of synthesised compound **2.043.**

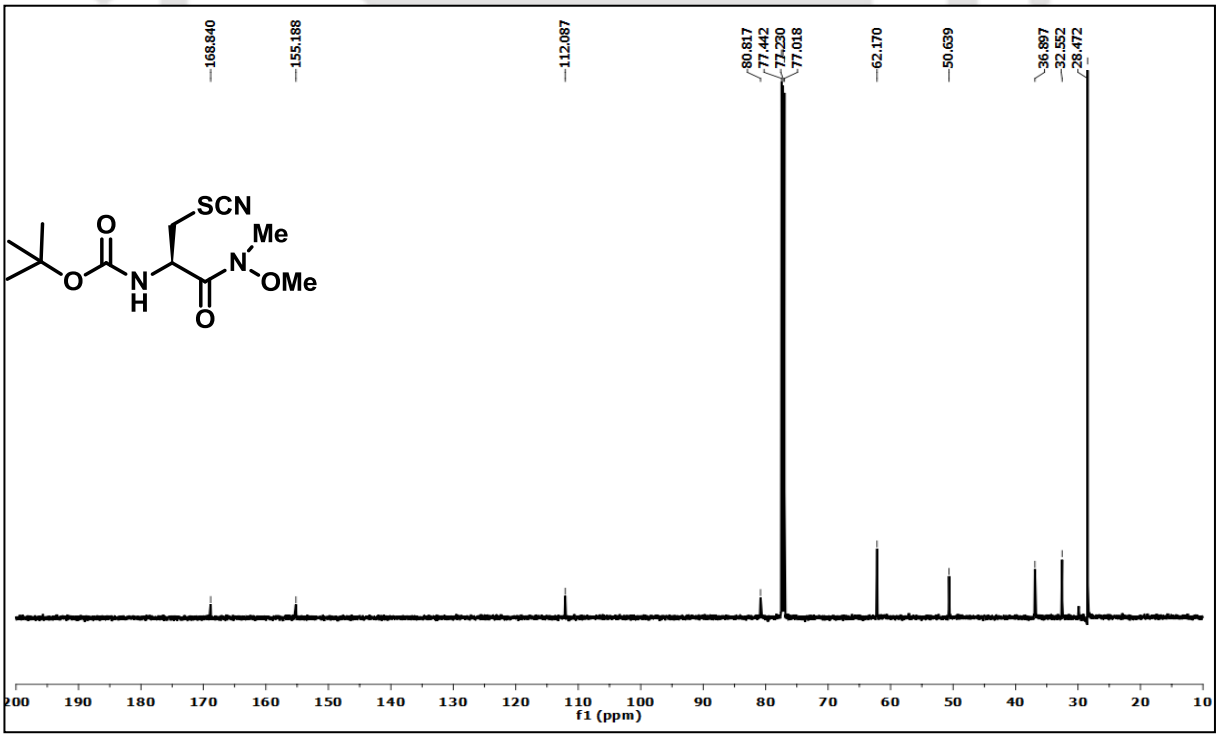


Figure 2.16. ^{13}C NMR spectra of synthesised compound **2.043.**

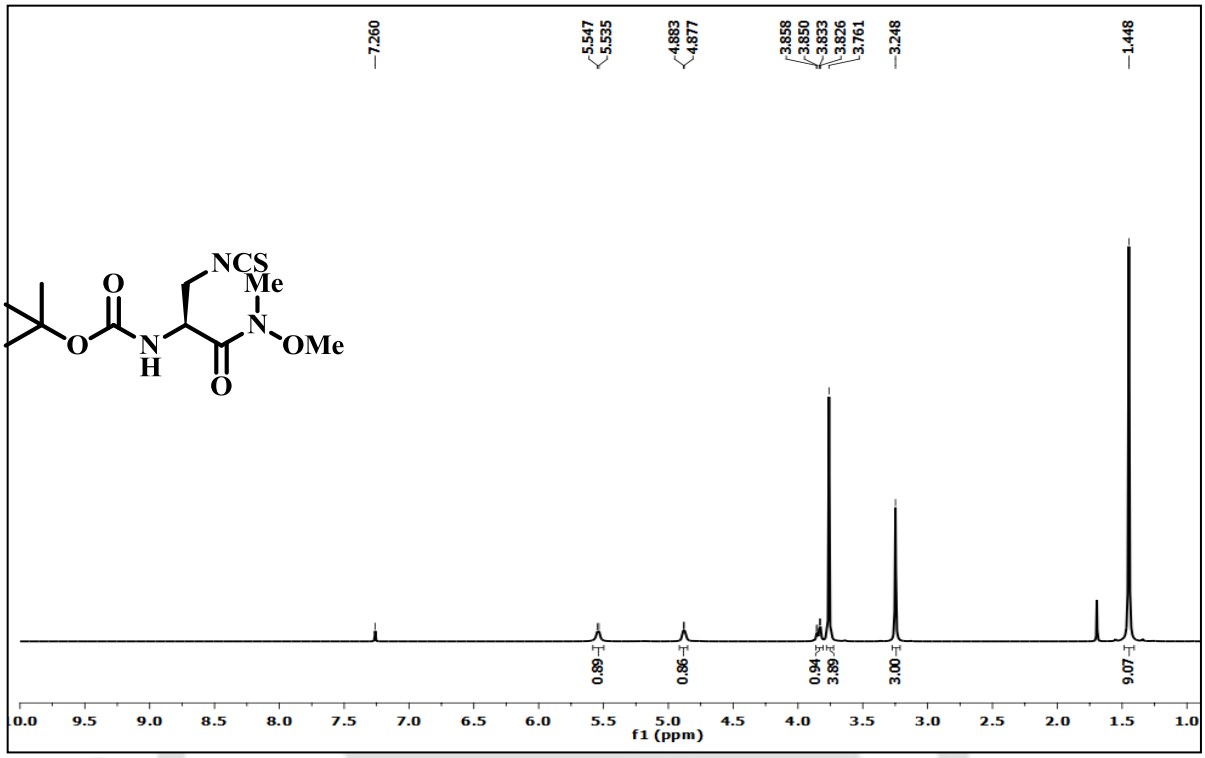


Figure 2.17. ^{1}H NMR spectra of synthesised compound **2.046.**

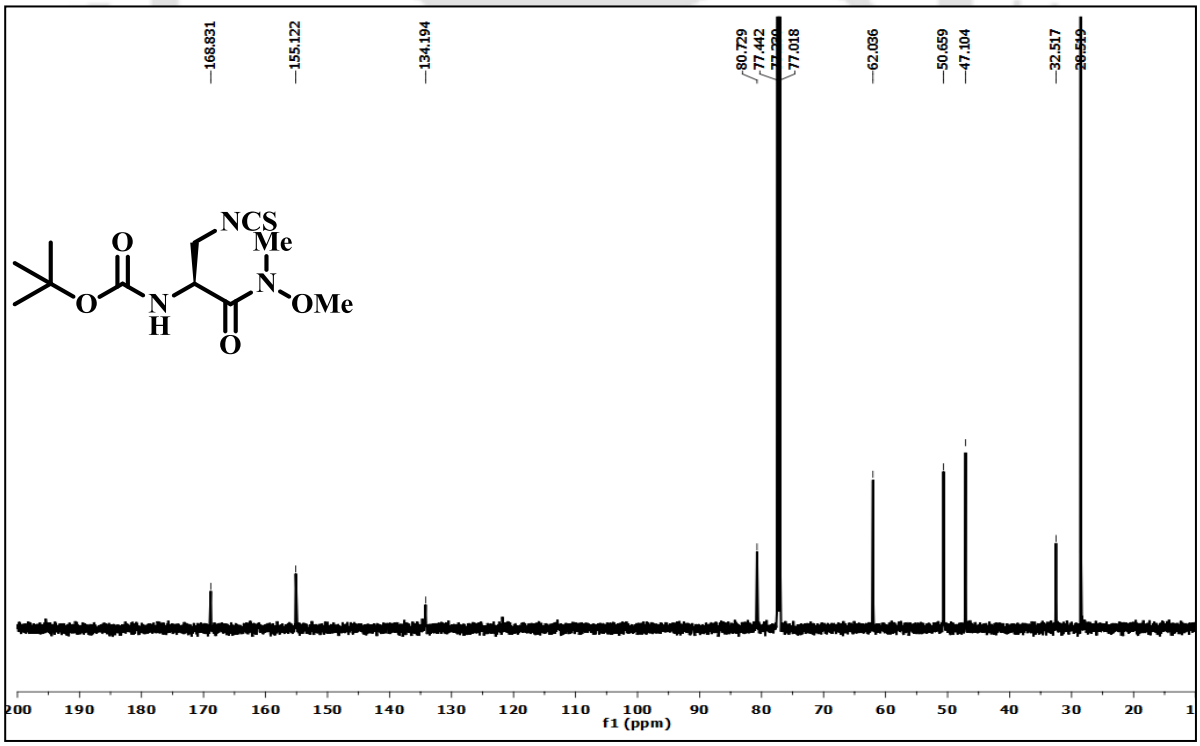


Figure 2.18. ^{13}C NMR spectra of synthesised compound **2.046.**

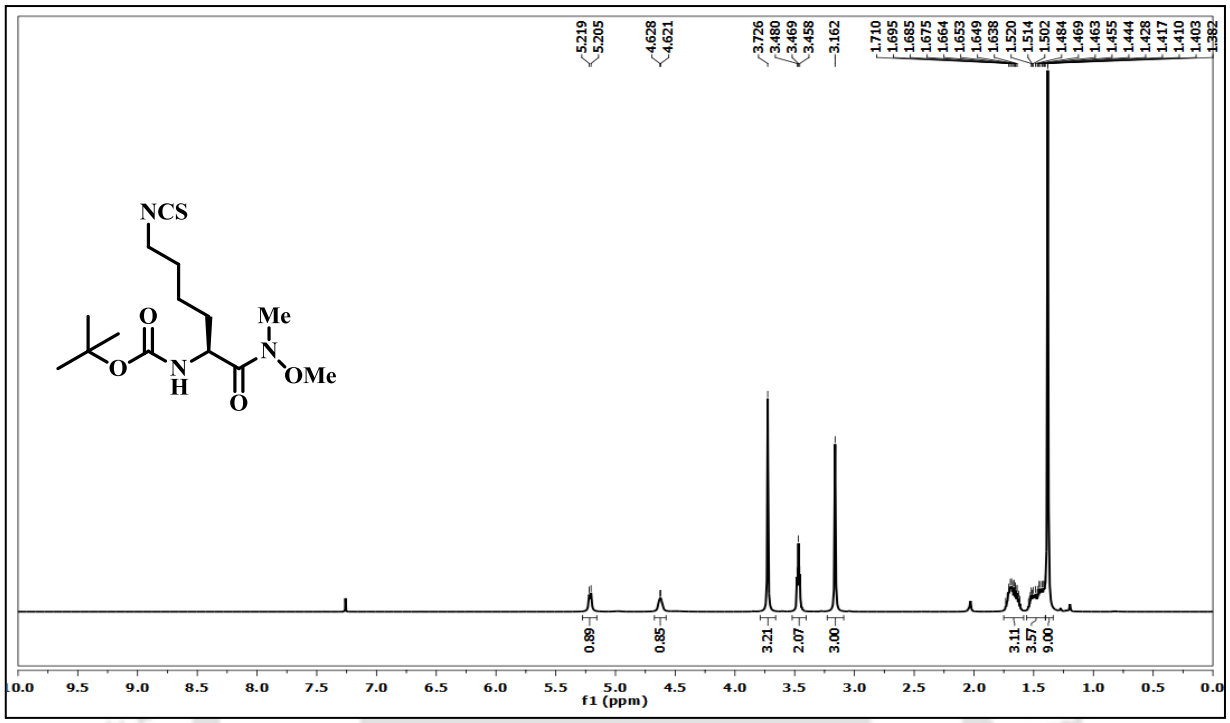


Figure 2.19. ^{1}H NMR spectra of synthesised compound **2.052.**

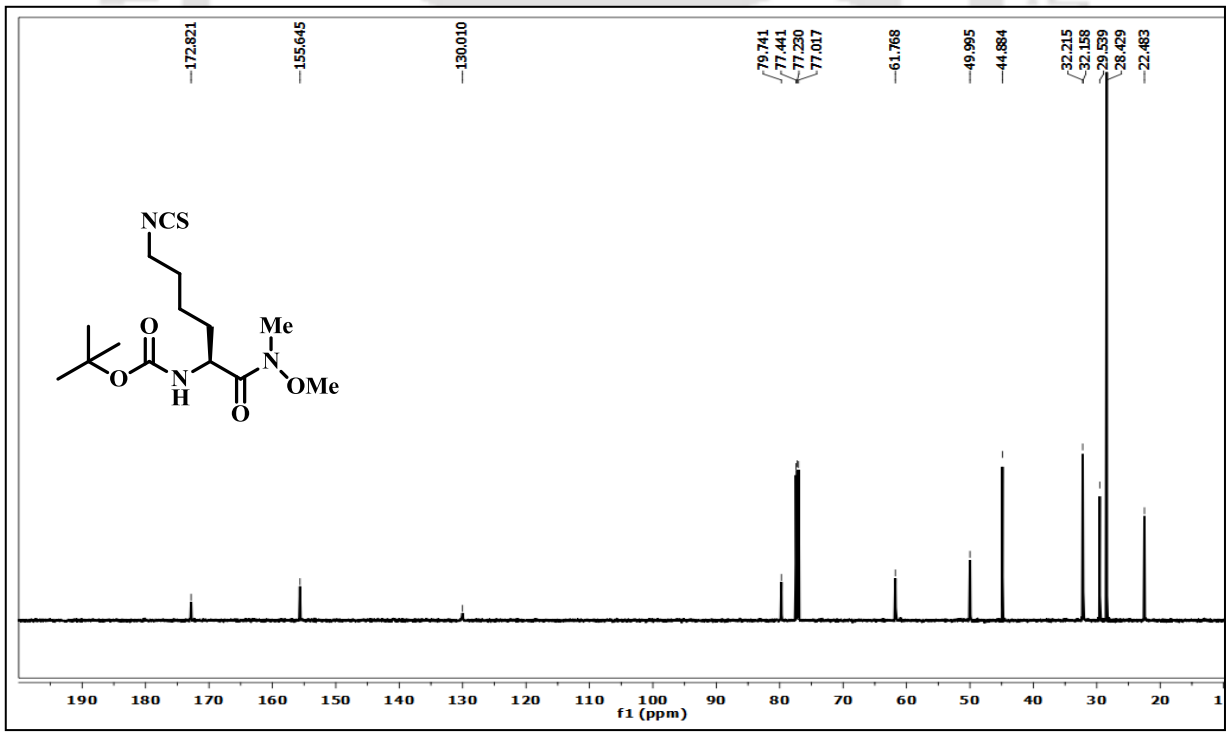


Figure 2.20. ^{13}C NMR spectra of synthesised compound **2.052.**

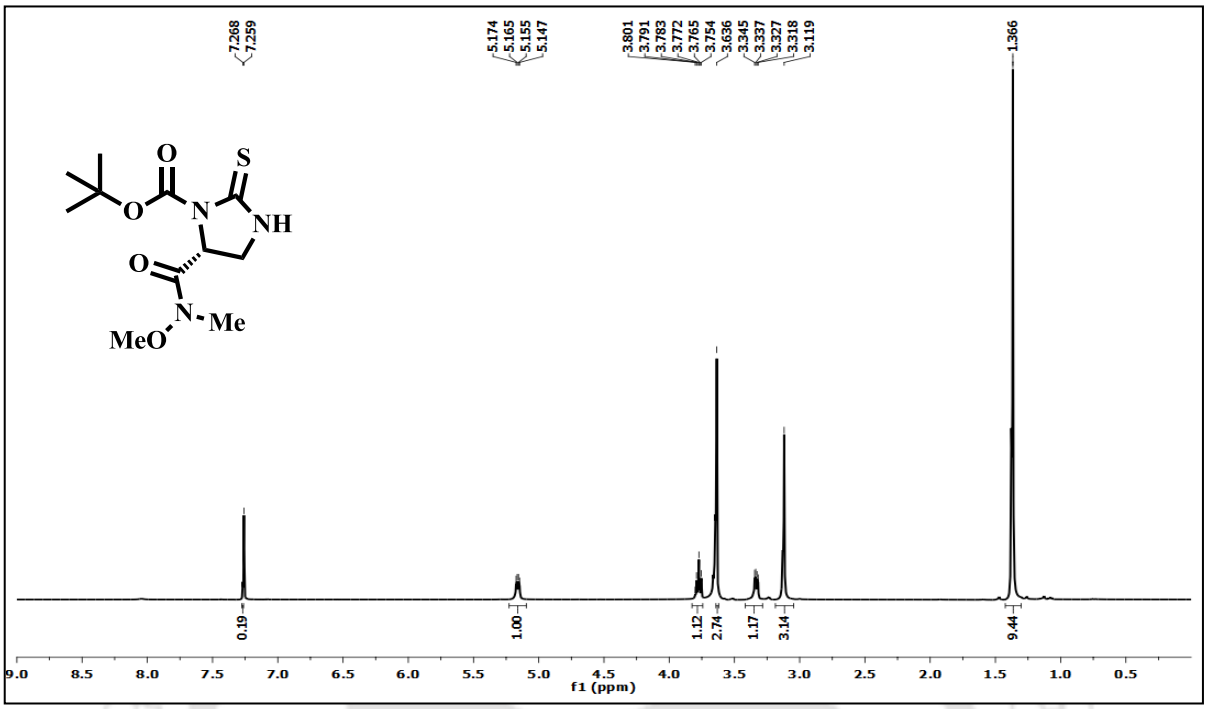


Figure 2.21. ^{1}H NMR spectra of synthesised compound **2.047.**

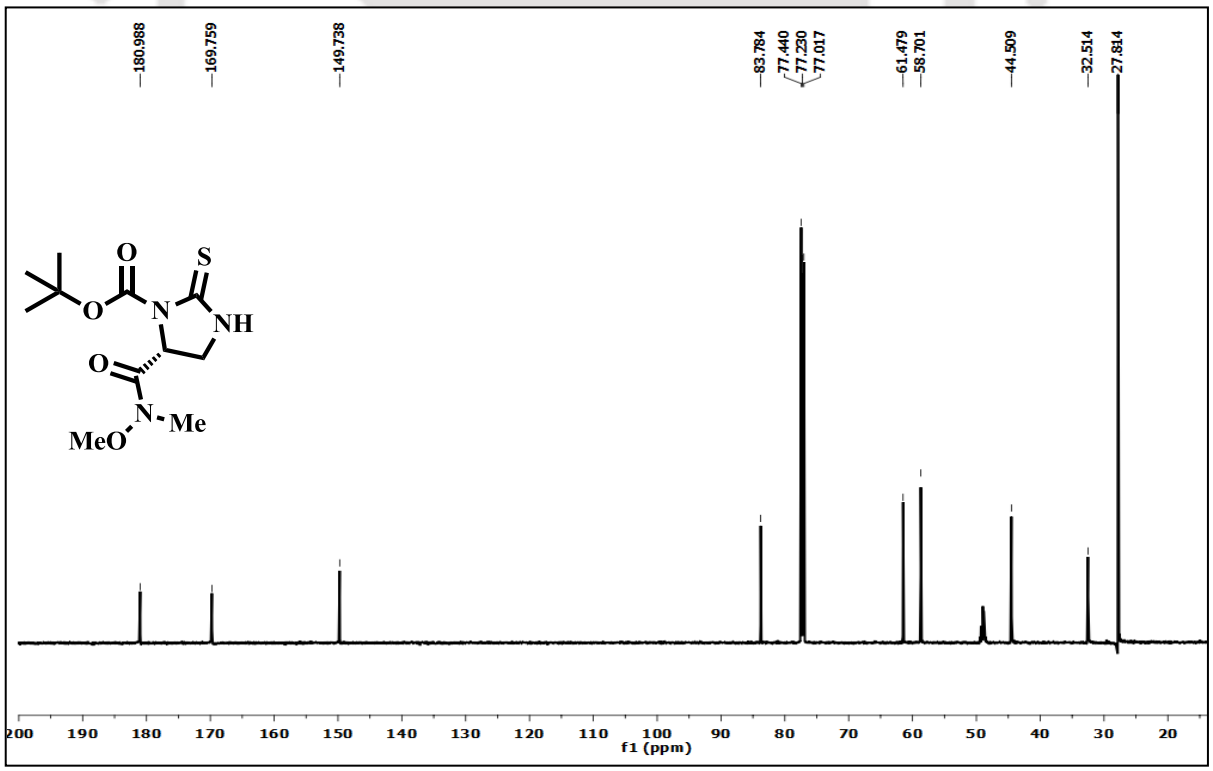


Figure 2.22. ^{13}C NMR spectra of synthesised compound **2.047.**

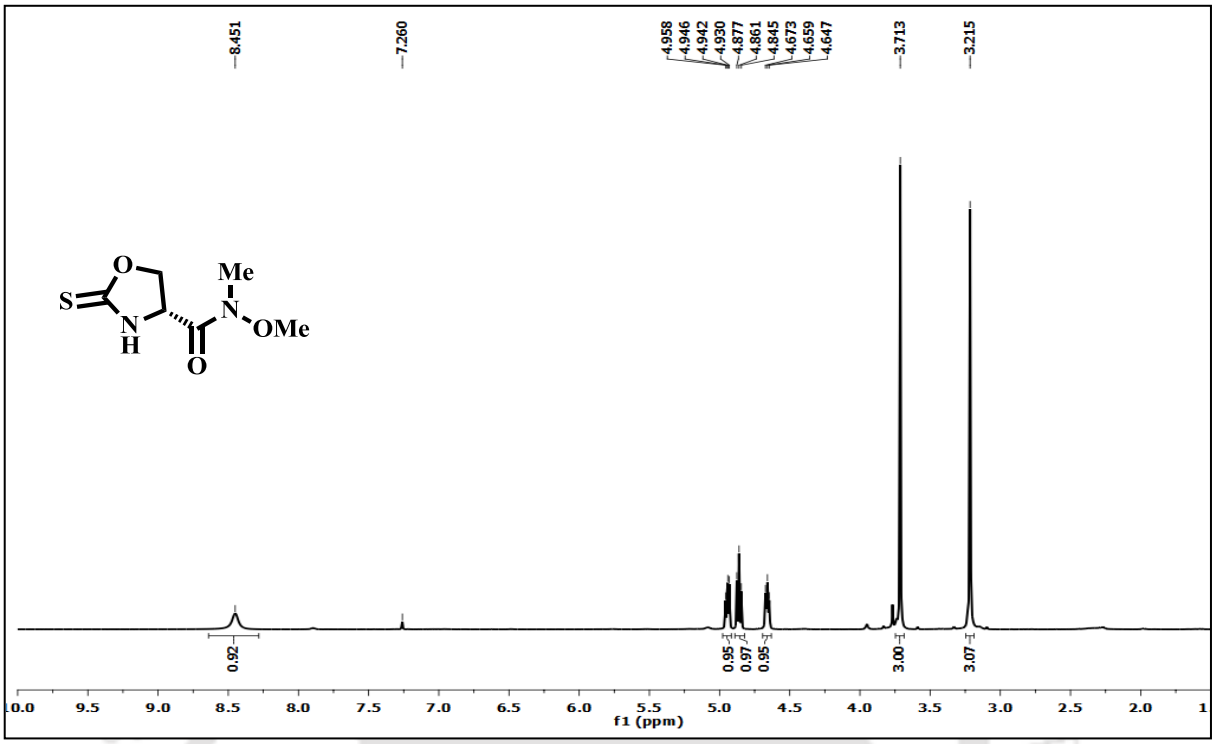


Figure 2.23. ^{1}H NMR spectra of synthesised compound **2.057.**

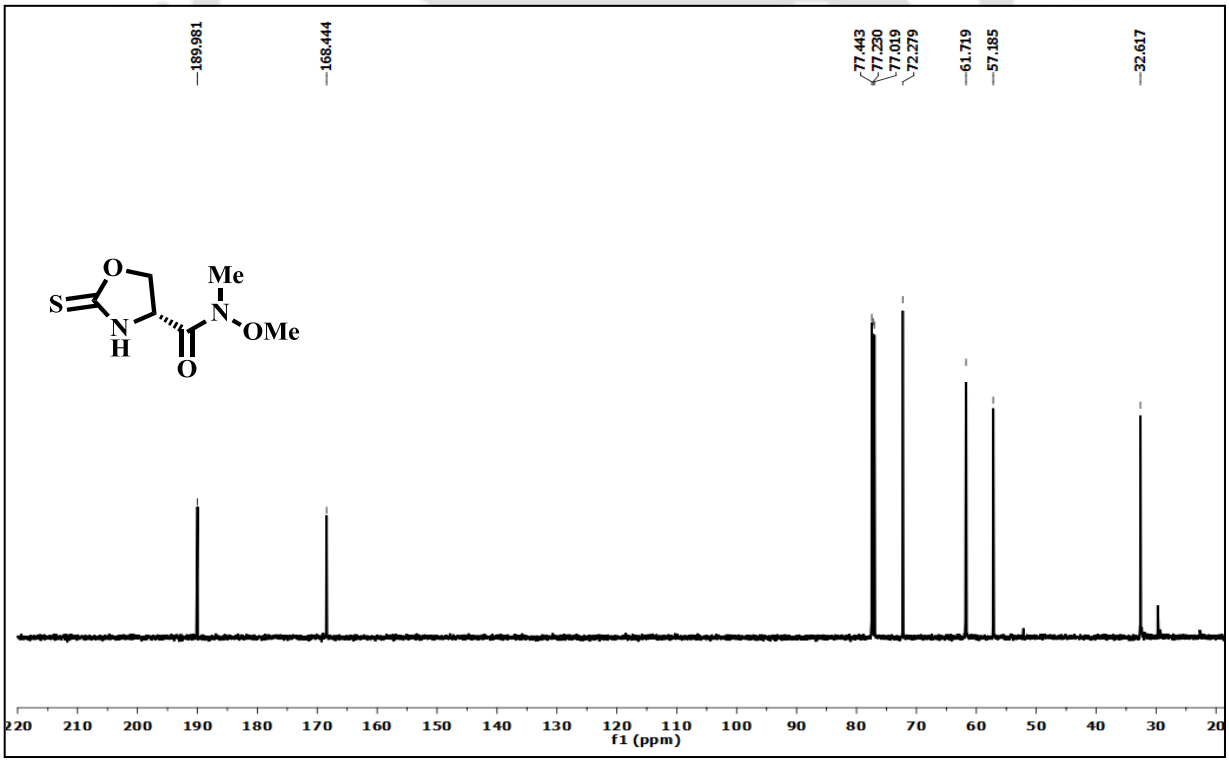


Figure 2.24. ^{13}C NMR spectra of synthesised compound **2.057.**

2.10. References

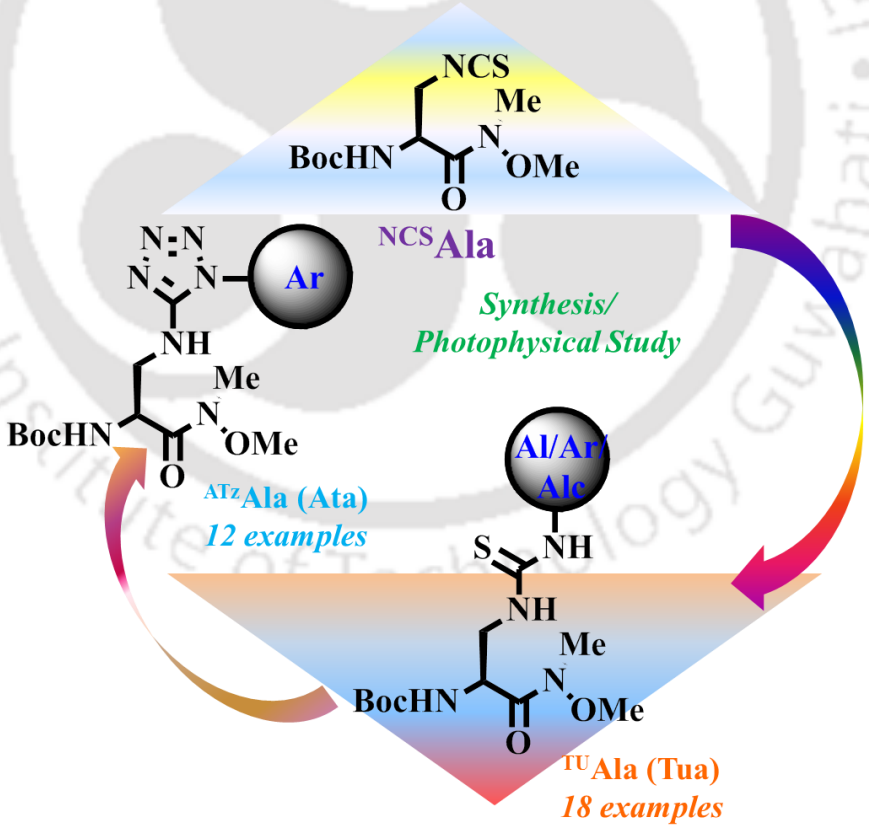
1. (a) Cumming, R.C.; Schubert, D. *FASEB J*. **2005**, *19*, 2060. (b) Konda, M.; Kauffmann, B.; Rasalea, D.B.; Das, A. K. *Org. Biomol. Chem*. **2016**, *14*, 4089 (c) Flora, J.S.S.; Pachauri, V. *Int. J. Environ. Res. Public Health.* **2010**, *7*, 2745. (d) Sears, M.E. *Scientific World J*. **2013**, 21984. (e) Williams, R.M.; Sinclair, P.J.; Zhai, D.; Chen, D. *J. Am. Chem. Soc*. **1988** *110*, 1551.(f) Mohite, A.R.; Bhat, R.G. *J. Org. Chem.* **2012**, *77*, 5423.
2. Tron, G. C.; Pirali, T.; Billington, R.A.; Canonico, P.L.; Sorba, G.; Genazzani, A.A. *Med. Res. Rev*. **2008**, *28*, 278.
3. (a) Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem., Int. Ed*. **2001**, *40*, 2004. (b) Wu, P.; Feldman, A. K.; Nugent, A.K.; Hawker, C. J.; Scheel, A.; Voit, B.; Pyun, J.; Fréchet, J. M. J.; Sharpless, K.B.; Fokin, V. V. *Angew. Chem., Int. Ed*. **2004**, *43*, 3928. (c) Finn, M. G.; Fokin, V.V. *Chem. Rev*. **2010**, *39*, 1231.
4. Kolb, H. C.; Sharpless, K. B. *Drug Discovery Today* **2003**, *8*, 1128. (b) Moses, J.E.; Moorhouse, A.D. *Chem. Rev*. **2007**, *36*, 1249.
5. (a) Bock, V. D.; Hiemstra, H.; van Maarseveen, J. H. *Eur. J. Org. Chem*. **2005**, 51. (b) Briehn, C. A.; Schiedel, M.-S.; Bonsen, E. M.; Schuhmann, W.; Bauerle, P. *Angew. Chem., Int. Ed*. **2001**, *40*, 4680. (c) Sawa, M.; Hsu, T.-L.; Itoh, T.; Sugiyama, M.; Hanson, S. R.; Vogt, P. K.; Wong, C.-H. *Proc. Natl. Acad. Sci. U. S. A*. **2006**, *103*, 12371.
6. Evans, C. E.; Lovell, P. A. *Chem. Commun*. **2009**, 2305.
7. Dirks, A. J.; Cornelissen, J. J. L.M.; Nolte, R. J. M. *Bioconjugate Chem*. **2009**, *20*, 1129.
8. Guo, H. M.; Minakawa, M.; Tanaka, F. *J. Org. Chem*. **2008**, *73*, 3964.
9. (a) Agalave, S. G.; Maujan, S.R.; Pore, V.S. *Chem. Asian. J*. **2011**, *6*, 2696. (b) Meldal, M.; Tornøe, C.W. *Chem. Rev*. **2008**, *108*, 2952.
10. (a) Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem., Int. Ed*. **2001**, *40*, 2004. (b) Wu, P.; Feldman, A.K.; Nugent, A.K.; Hawker, C.J.; Scheel, A.; Voit, B.; Pyun, J.; Frechet, J. M. J.; Sharpless, K. B.; Fokin, V. V. *Angew. Chem., Int. Ed*. **2004**, *43*, 3928. (c) Kann,N.; Johansson, J.R.; Beke-Somfai, T. *Org. Biomol. Chem*. **2015**, *13*, 2776. (d) Huisgen, R. *J. Org. Chem*. **1976**, *41*, 403
11. (a) Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem., Int. Ed*. **2001**, *40*, 2004.(b) Wu, P.; Feldman, A. K.; Nugent, A. K.; Hawker, C. J.; Scheel, A.; Voit, B.; Pyun, J.; Fréchet, J. M. J.; Sharpless, K. B.; Fokin, V.V. *Angew. Chem., Int. Ed*. **2004**, *43*, 3928. (c) Finn, M.G.; Fokin, V. V. *Chem. Rev*. **2010**, *39*, 1231.
12. (a) Singhal, N.; Sharma,P.K.; Dudhe,R.; Kumar, R. *J. Chem. Pharm. Res*. **2011**, *3*, 126. (b) Havaldar, H.F.; Patil. R.A. *E-J. of Chem*. **2008**, *5*, 347. (c) Sztanke,K.; Tuzimski,T.; Rzymowska, J.; Pasternak, K.; Szerszen. M. K *Eur. J. Med. Chem*. **2008**, 43, 404. (d) Sung, K.; Lee**.**A.R. *J Heterocyclic Chem***. 1992**, 29, 1101. (e) Demirbas. N.; Karaoglub,S.A.; Demirbasa,A.; Sancaka, K. *Eur. J of Med Chem*. **2004**, 39,793. (f) Kumar, H.; Javed, A.S.; Khan, A.S.; Amir. M. *Eur J Med Chem*. **2008**, 43, 2688.
13. (a) Vergnaud, J.; Faugeras, P.-A.; Chaleix, V.; Champavier, Y.; Zerrouki, R.; *Tetrahedron Lett*. **2011**, *52*, 6185.
14. (a) Gajewski, M.; Seaver,B.; Esslinger, C.S. *Bioorg Med Chem Lett*. **2007**, 17, 4163. (b) Bag, S. S.; Jana, S.; Yashmeen, A.; Senthilkumar, K; Bag, R. *Chem Comm*. **2014**, *50*, 433.
15. (a) Hodgkins, J. E.; Ettlinger, M. G. *J. Org. Chem*. **1956**, *21* , 404. (b) Hodgkins, J. E.; Reeves, W. P. *J. Org. Chem.* **1964**, *29*, 3098. (c)Wong,R.; Dolman,J.S. *J. Org. Chem.* **2007**, *72*, 3969. (d) Kim ,T.; Kim,Y.-J.; Han, I.-H.; Lee,D.; Hama,J.; Kang, K.S.; Jae Wook Lee,

J.W. *Bioorg. Med. Chem. Lett*. **2015**, *25*, 62**. (e)** Lamani,R.S.; Nagendra, G.; Sureshbabu, V.V. *Tetrahedron Lett*. **2010**, *51*, 4705**.** (f) Isobe, T.; Fukuda, K.; Tokunaga,T.; Seki,H.;Yamaguchi, K. Ishikawa, T. *J. Org. Chem.* **2000,** *65*, 7774.

16. (a) Ju, Y.; Kumar, D.; Varma, R.S. *J. Org. Chem.* **2006**, *71*, 6697. (b) Iranpoor, N.; Firouzabadi, H.; Akhlaghinia, B. ; Azadi, R. *Synthesis*, **2004**, 9296 (c) Liu, Y.; Xu, Y.; Jung, S. H.; Chae, J. Synlett. **2012**, 23, 2663. (d) Sun, N.; Zhang, H.; Mo, W.; Hu, B. ; Shen, Z.; Hu, X. *Synlett*. **2013**, *24*, 1443.
17. (a)Tsimberidou, A.M.; Keating M. J. *Cancer*. **2009**, *115*, 2824. (b) Madhunapantula, S.V.; Robertson, G. P. *Enzyme Res*. **2011**, 327923. (c) Wang,X.; Di Pasqua, A.J.; Govind,S.; McCracken,E.; Hong, C.; Mi,L.; Mao,Y.; Yu-Chieh Wu, J.; Tomita, Y.; Woodrick, J.C.; Fine, R.L.; Chung, F.-L. *J. Med. Chem.*, **2011**, *54*, 809.
18. Kim, T.; Kim, Y.J.; Han, I.-M.; Lee , D.; Hama, J.; Kang , K.S.; Lee, J.W. *Bioorg. Med. Chem. Lett.* **2015**, *25*, 62.
19. Lamani, R.S.; Nagendra, G.; Sureshbabu, V.V. *Tetrahedron Lett*. **2010**, *51*, 4705.
20. (a) Conner, G. E.; Wijkstrom-Frei, C.; Randell, S. H.; Fernandez, V. E.; Salathe, M. *FEBS Letters*. **2007**, *581***,** 271. (b) White, W.E.; Pruitt, K.M.; Mansson-Rahemtulla, B. *Antimicrobial Agents and Chemotherapy*. **1983**, *23*, 267. (c) Thomas, E. L.; Aune, T. M. Infection and Immunity. **1978**, *20*, 456.
21. Childers, M.; Eckel, G.; Himmel, A.; Caldwell, *J. Medical Hypotheses*. **2007**, *68*, 101.
22. Minarowski, Ł.; Sands, D.; Minarowska, A.; Karwowska, A.; Sulewska, A.; Gacko, M.; Chyczewska, E. Folia *Histochemica et Cytobiologica*. **2008**, *46*, 245.
23. Moskwa, P.; Lorentzen, D.; Excoffon, K. J.; Zabner, J.; McCray, P. B. Jr.; Nauseef, W. M.; Dupuy, C.; Bánfi, B. *Am J Respir Crit Care Med*. **2007**, *175*, 174.
24. Xu, Y.; Szép, S.; Lu, Z. *Pnas.* **2009**, *106*, 20515.
25. Braverman L. E.; He X.; Pino S. J. Clin Endocrinol Metab. **2005**, *90*, 700.
26. Spicer, C.D.; Davis, B.G. *Nat. Comm*. **2014**, *5*, 4740.
27. (a) Mukerjee, A. K.; Ashare, R. *Chem. Rev.* **1991**, 91, 1. (b) Wu, Y.-J.; Zhang, Y. *Tetrahedron Lett.* **2008**, *49*, 2869. (c) Seitz, O. *ChemBioChem.* **2000**, *1*, 215. (d) Gunther, W.; Kunz, H. *Angew. Chem., Int. Ed. Engl.* **1990**, *29*, 1050. (e) Rodriguez-Lucena, D.; Benito, J.; Ortiz Mellet, C.; Garcia Fernandez, J. M. *Chem. Commun.* **2007**, 831. (f) Gomez-Garcia, M.; Benito, J. M.; Rodriguez-Lucena, D.; Yu, J.-X.; Chmurski, K.; Ortiz Mellet, C.; Gutirrez-Gallego, R.; Maestre, A.; Defaye, J.; Garcia Fernandez, J. M. *J. Am. Chem. Soc.* **2005**, *127*, 7970. (g) Garcia Fernandez, J. M.; Ortiz Mellet, C. *Adv. Carbohydr. Chem. Biochem.* **2000**, *55*, 35. (h) Falconer, R. A.; Toth, I. *Bioorg. Med. Chem*. **2007**, *15*, 7012. (i) Wittstock, U.; Agerbirk, N.; Stauber, E. J.; Olsen, C. E.; Hippler, M.; Mitchell-Olds, T.; Gershenzon, J.; Vogel, H. *PNAS* **2004**, *101*, 4859. (j) Michalski, O.; Ciez, D. J. *Mol. Struct.* **2013**, *1037*, 225. (k) Halpern, B.; Patton, W.; Crabbé, P. *J. Chem. Soc. B* **1969**, 1143 (l) Bones, A. M.; Rossiter, J. T. *Physiol. Plant.* **1996**, *96*, 194. (m) Cejpek, K.; Valusek, J.; Velısek, J. *J. Agric. Food Chem.* **2000**, *48*, 3560. (n) Ciez, D. *Tetrahedron.* **2007**, *63*, 4510. (o) Kalinina, S.; Gliemann, H.; Lopez-Garcıac, M.; Petershans, A.; Auernheimer, J.; Schimmel, T.; Bruns, M.; Schambony, A.; Kessler, H.; Wedlich, D. *Biomaterials.* **2008**, *29*, 3004.
28. (a) Zhang, Y.; Li, T.; Gonzalez, V. *Mol. Cancer Ther*. **2003**, *2*, 1045. (b) Xiao, D.; Vogel, V.; Singh, S. V. *Mol. Cancer Ther*. **2006**, *5*, 2931.

Chapter-3

Application of ^{NCS}Ala in the Synthesis of Thioureayl Alanines and then Aminotetrazolyl Alanines as another Class of New Unnatural Amino Acids



3.1. Introduction

The key role of protein and amino acids in the structure and function of living organism has encouraged extensive studies of amino acids. Modified amino acids with enhanced biological activity, proteolytic stability and bioavailability are of increasing interest in protein design and engineering.[1a] In the last few years, several efforts have been devoted to the synthesis of side chain modified unnatural amino acids. This modification can be achieved by oxidation, halogenation, C-H activation, electrophilic addition and nucleophilic substitution reactions.

Isothiocyanate as modified side chain can bring the functional tunability in the amino acids as was described in the C**hapter 2**. The well-known functional conversion of isothiocyanates to thioureas is attractive in this regard. Thiourea is an organosulfur compound which is structurally analogous to urea. However, the nature of urea and thiourea vary considerably due to its high coordination ability and larger cross-sectional area.[1c] Thiourea constitutes an important class of compounds that possess widespread applications in medicinal chemistry[1b-e] such as fungicides, herbicides,[2] rodenticides[3]and phenoloxidase enzymatic inhibitors.[4] It also acts as valuable building blocks for the synthesis of amides, guanidines and varieties of heterocycles including tetrazoles.[5a-f] It is also used for anion recognition/sensing[5g-i] and in organocatalysis.[6] The synthesis of thiourea-peptides is thus an attractive area of research, since the thiourea residues are fairly rigid[1b] and may play an important role in protein secondary structure and it's folding.

The tetrazole group acts as bioisosteric to the carboxyl group [7f] with enhanced biological functions.[7g] Aminotetrazoles is another class of widely used heterocyclic compound in the current research area. It has the maximum number of nitrogens among the heterocyclic compounds and displays remarkably high thermal stabilities.[7a-e] It has a wide range of applications such as in high energy density materials (HEDM),[8] as useful ligands in coordination chemistry[9] and in bioorganic chemistry.[10] In particular, 5-aminotetrazoles function as anti-allergic and anti-asthmatic,[12]antiviral and anti-inflammatory,[13] antineoplastic[14] and cognition disorder activities.[15a] On the other hand 1, 5-disubstituted tetrazoles are effective bioisosters for cis–amide bonds in peptidomimetics.[7h,kj]The importance for the development of new aminotetrazole based amino acids molecules using different starting material and different reaction methodology would be thus an active area of research. To meet the challenges in side chain modifications, we describe the few reported synthesis of new side chain modified thioureayl and aminotetrazolyl amino acids.

3.2. General Synthetic Procedure for Thioureas

There are several reports in the literature describing *in situ* generation of substituted thiourea.[17] It can be generated by condensation of the primary or secondary amine with isothiocyanate [17a], thiophosgene [17b] or its derivatives[17c]. Many isothiocyanates are difficult to synthesize and are very unstable. Hence, several routes have been reported by different research groups for the preparation of both symmetrical [18a-f]**(Figure 3.01a**) and unsymmetrical thioureas [19a-h] **(Figure 3.01b**) without using isothiocyanates directly. Prabhu and co-workers reported an efficient method for the preparation of symmetrical and unsymmetrical thiourea derivatives by simple condensation between amines and carbon disulfide in aqueous medium (**3.003, Figure 3.01a**).[20a]

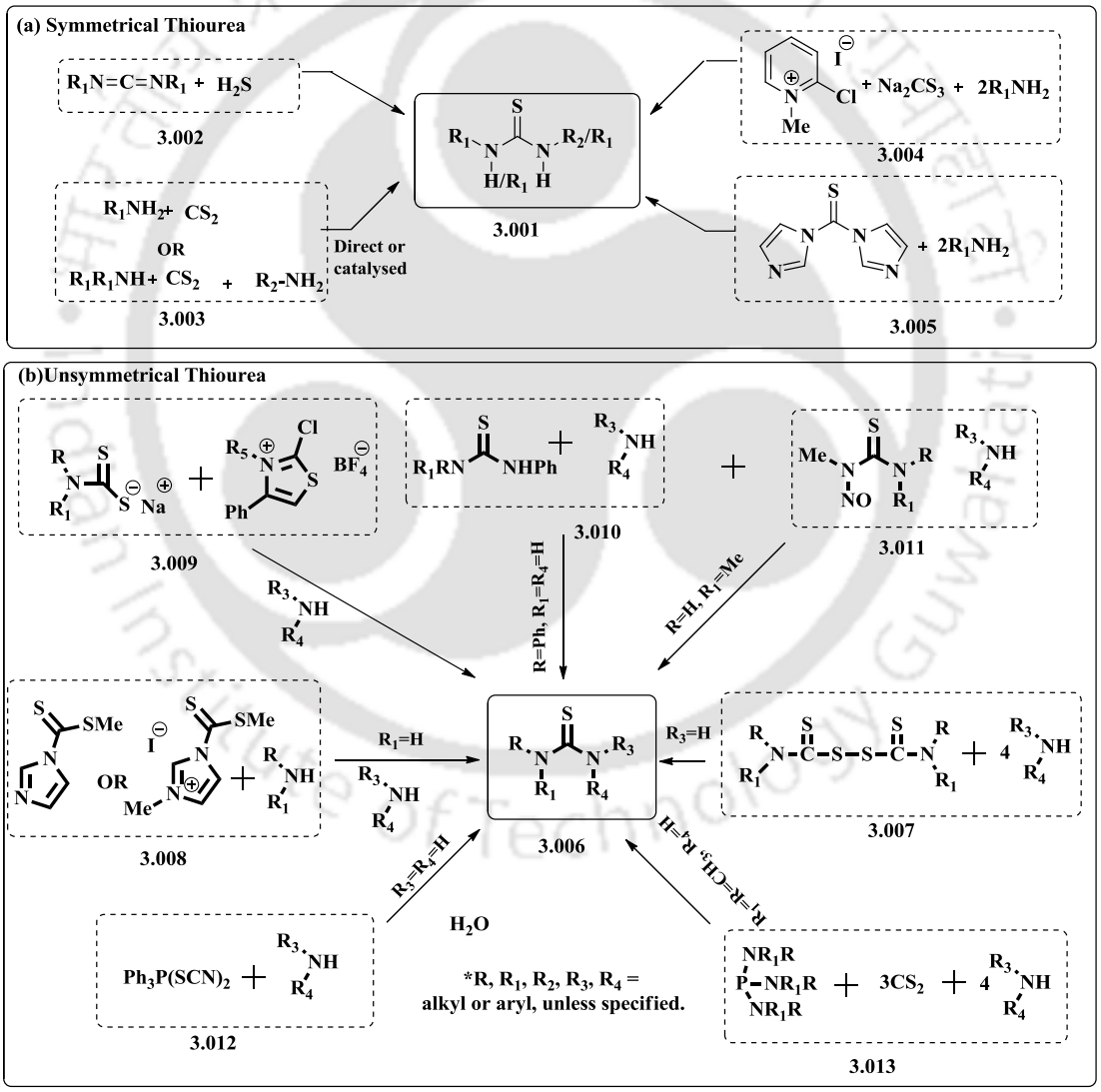


Figure 3.01. Reported procedures for the synthesis of (a) symmetrical and (b) unsymmetrical thiourea.

3.2.1. Applications of Thioureas

3.2.1.1. Thioureas in Biology and Medicinal Chemistry

In 1998, Dallaire *et al*.,[21a] have synthesised a new thiourea based building block **(3.014, Figure 3.02)** where from new analogues of peptide nucleic acids can be obtained. During the past several years, peptide nucleic acids (PNAs) have appeared to be useful mimics of DNA with promising applications in diagnostics and in the pharmaceutical areas (e.g. antisense-antigene-therapeutic agents). They have incorporated modifications into the polyamide backbone of PNAs and explored its effects on the binding properties with DNA/RNA and on the transportation phenomena across the cell membranes.[21a] In 2010 Liu *et al*., [21b] and their co-workers have synthesised and studied pseudo-peptide thioureas containing α-aminophosphonate moiety (**3.015, Figure 3.02)** as potential antitumor agents.[21b] In 2012, Klein *et al*., [21c] have synthesised a new class of bis(thiourea) hydrazide pseudopeptides **(3.016, Figure 3.02**). This class of compounds is designed to display "amphifinity", i.e. association with a peptide strand on one face of the scaffold but not on the other, and hence could potentially inhibit β-sheet aggregation.[21c]

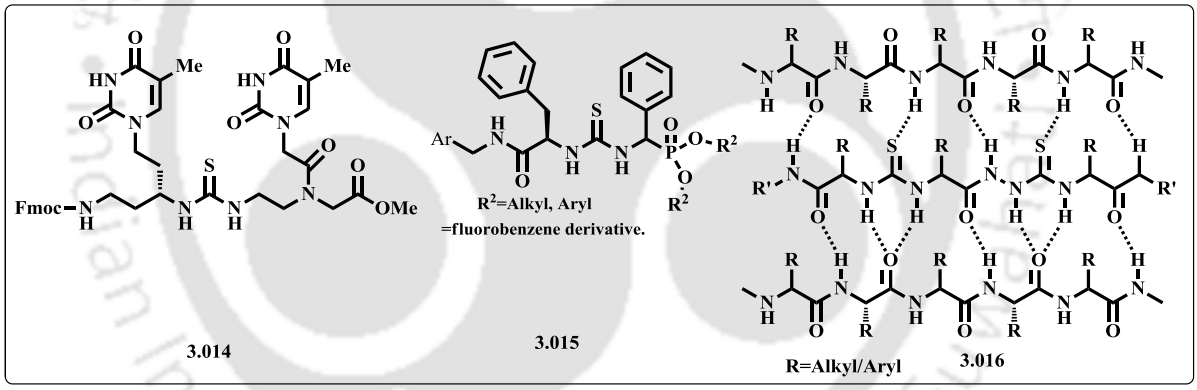


Figure 3.02. Biologically important thiourea derivatives.

In 2012, Gowda *et al*., [21d] have synthesised thiouriedo derivatives of peptide conjugated heterocycles (**3.017, Figure 3.03).** These compounds were evaluated for their ability to inhibit the growth of a panel of microorganisms. All the synthesised compounds in this series have been found to be display an excellent antimicrobial activity.[21d] Wang *et al*., [21g] have synthesised novel dehydroabietic acid derivatives conjugated with acyl-thiourea peptide moiety (**3.018, Figure 3.03)** and evaluated their anti-tumor activity. They also showed that compound **3.020** exhibited the best antitumor activity against the HeLa cell line and even displayed more potent inhibitory activity than commercial antitumor drug 5-fluorouracil (5-FU).[21g] In 2016, Ghasemi and his co-workers [21h] have produced thiourea-peptoids (**3.019, Figure 3.03**) in moderate to good yields from a four component Ugi reaction. As small peptides are an

important group of compounds in medicinal chemistry, they modified peptide moiety using unnatural amino acid which shows more diversity compared to natural analogues.[21h]

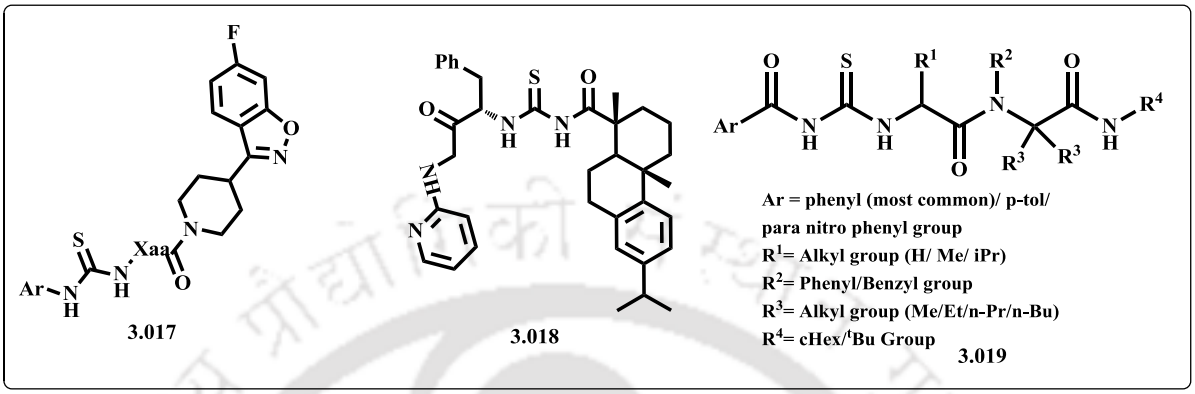


Figure 3.03. Medically useful thiourea derivatives.

3.2.1.2. Thioureas as Sensor

In 2013, Jolliffe and co-workers have synthesised tripodal thiourea (**3.020, Figure 3.04**) based on a cyclic peptide scaffold and studied their anion binding properties. The synthesised product has high affinity towards sulfate binding in aqueous solution. The complex is mainly supramolecular and uses hydrogen bonding as complexation technique. Thus this provides many potential applications such as receptor binding and discriminating between anions in aqueous media. [21e] In the year 2015, Costero *et al.*, [21f] have made a thiourea-modified fluorescein derivative (**3.021, Figure 3.04)** by the reaction of fluorescein isothiocyanate with 2-(2-aminoethoxy)ethan-1-ol. This simple fluorescein derivative works as a sensor and is able to discriminate among linear aliphatic ω-amino acids with different chain lengths.[21f]

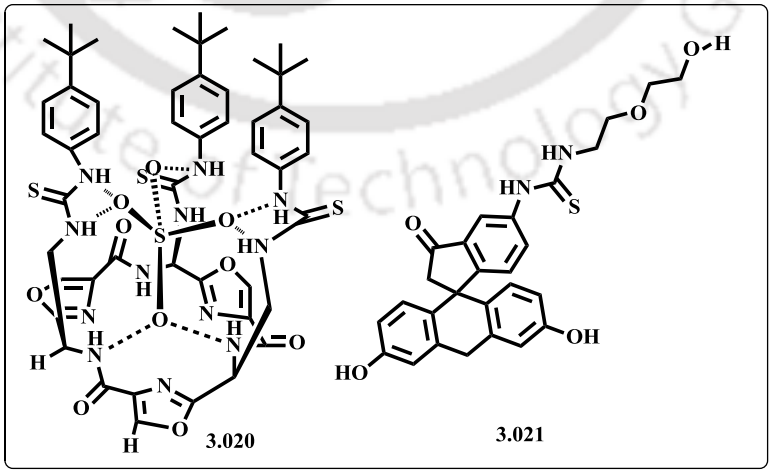


Figure 3.04. Reported various thiourea compounds for sensor.

3.2.1.2. Thioureas as Catalyst

In 2005, Takemoto *et al*., have demonstrated some enantioselective reactions catalyzed by thiourea derivatives **(3.022, Figure 3.05)** as general acid catalysts as well as diastereoselective reactions considering H-bond ability of thiourea derivatives.[21i] In 2006, Akiyama *et al*., have reviewed recent Progress in chiral Brønsted Acid Catalysis (**3.023, Figure 3.05**) where they have shown monofunctional and bifunctional thiourea catalysts in asymmetric synthesis [21j]**.** In 2015, Fu and co-workers have reported an important pathway for the direct enantioselective amination of α-ketoester which is catalyzed by tertiary amine thiourea **(3.024-3.027, Figure 3.05**). These organic catalysts were used to carry out the reaction with 93% enantioselectivity and 90% yield. [21k]

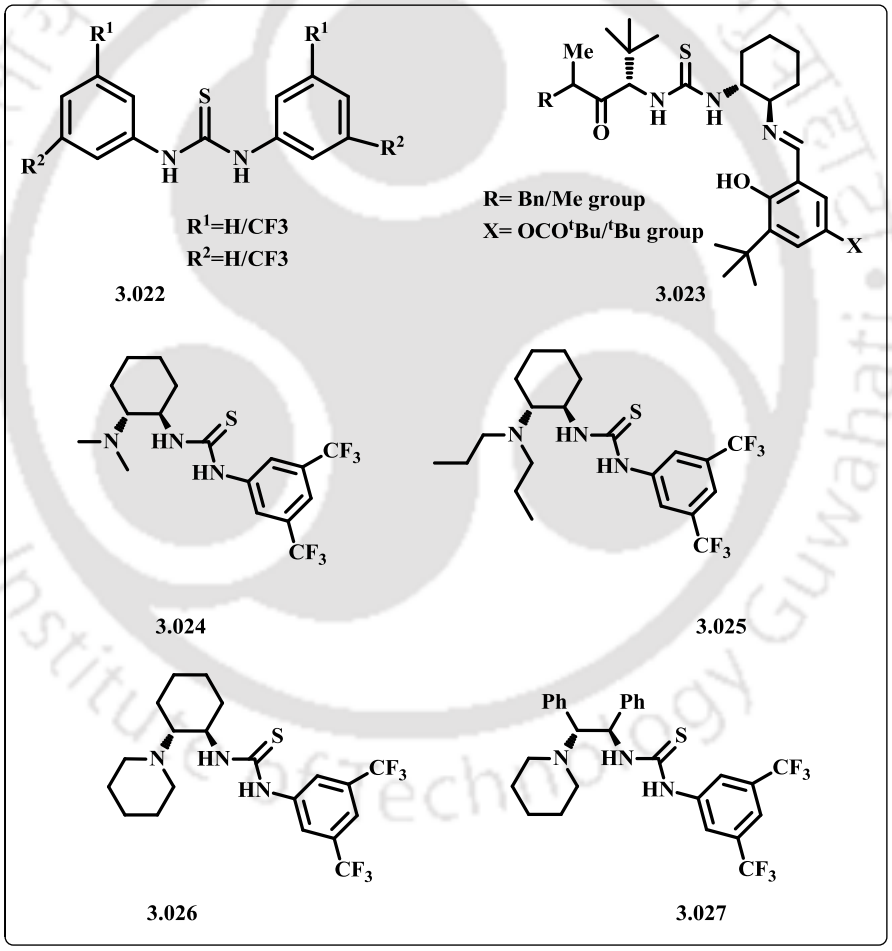


Figure 3.05. Reported various thiourea catalysts for enantioselective reactions.

3.3. General Synthetic Procedure for Aminotetrazoles

There are several reports in the literature describing the *in situ* generation of substituted aminotetrazoles. There are two ways: (a) reactions of aminoguanidine derivatives with sodium nitrite **(3.028, Figure 3.06)**; [22a-c] (b) synthesis of 5-aminotetrazole with the reaction of azide anion to (i) carbodiimides (**3.029, Figure 3.06)** [22c,23] (ii) cyanamides (**3.030, Figure 3.06)** [24], (iii) nucleophilic substitution of chlorine by N_3^- in R-chloroformamidines (**Figure 3.031)** [25a-d] (iv) sulfite anion by N_3^- in aminoiminomethanesulfonic acids (**3.032, Figure 3.06)** [25e] and (v) nucleophilic addition of N_3^- to thiourea (**3.033, Figure 3.06)** in the presence of copper, mercury [25f] or lead salts.[22a]

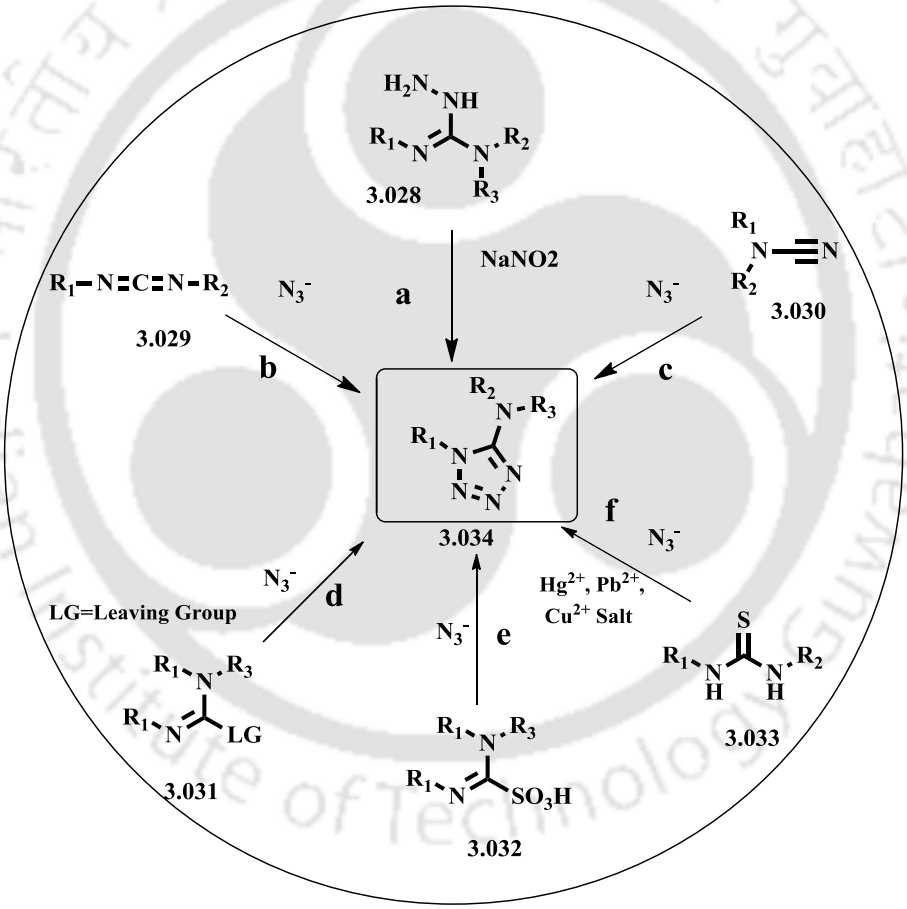


Figure 3.06. Reported procedure for the synthesis of aminotetrazole.

3.3.1. Application of Aminotetrazoles

In 1969, Morley *et al*., have reported the synthesis of an analog of C-terminal tetrapeptide amide sequence of gastrin wherein aspartic acid in position 16 was replaced by p-(5-tetrazolyl)-L-alanyl residue **(3.035, Figure 3.07**).[26a] This tetrazole analog of the physiologically active tetrapeptide amide was found to be as potent as its aspartyl relative. However, all the other synthetic analogues synthesised by modification at aspartyl position showed less activity. In 1993, Ornstei *et al*.,[26b] have reported the synthesis and pharmacological characterisation of two novel series of acidic amino acids (o-tetrazolesubstituted acyclic α–amino acids), **3.036** and 4-(tetrazolylalkyl) piperazine-2-carboxylic acids) **3.037-3.038** as potential N-methyl-D-aspartate (NMDA) antagonists. They also synthesised the piperazine tetrazoles which showed a moderate affinity for the NMDA receptor.[26b] Sureshbabu *et al*., have synthesised tetrazole analogues of amino acids starting from N$^{\alpha}$-Fmoc amino acid.[26c] They have synthesised aminotetrazole analogues of aspartic and glutamic acid and successfully incorporated these amino acids into peptides (**3.039, Figure 3.07)** for further study.[26c]

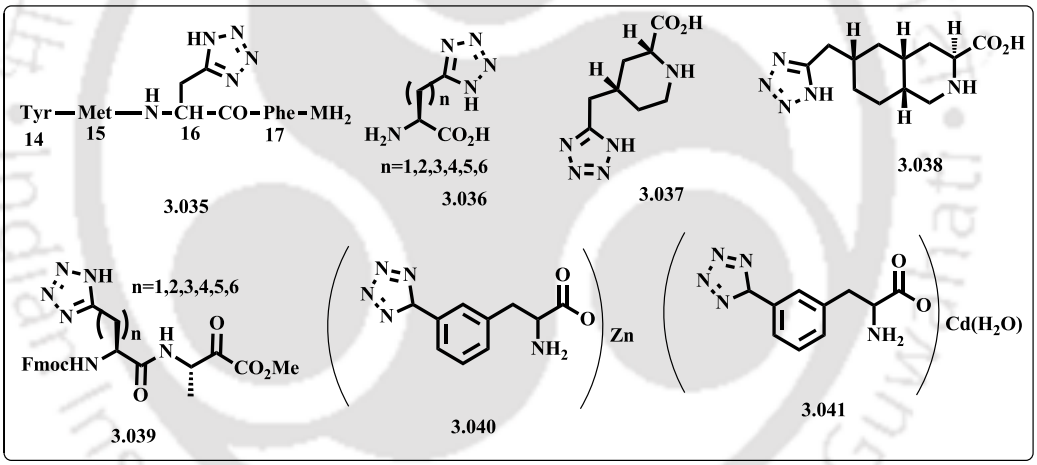


Figure 3.07. Some reported aminotetrazoles with different applications.

Few years later, Nachman *et al*., have reported the conformation of a C-terminal pentapeptide, modified with tetrazole moiety in active-core of insect kinin analog which mimic a cis peptide bond and favorably induce the formation of type VI β-turn.[26d] McManus and Herbst *et. al*.,[24b,c] proved that 5-substituted tetrazoles can be used as isosteric replacements of a carboxyl group and 1,5-disubstituted tetrazoles can be used as isosteres of the cis-amide bond of peptides. In 2003, Qu *et al*., have synthesised 3D homochiral networks containing mono [(*S*)-5-(3-tetrazoyl)-phenylalaninato] zinc(II) **(3.040, Figure 3.07)** and mono [(*S*)-5-(3-tetrazoyl)-phenylalaninato] cadmium(II) monoaqua(II) (**3.041, Figure 3.07)**.
The structure of these coordination polymers gives new insight into Sharpless reaction of chiral 5-substituted 1H-tetrazole and homochiral supramolecular array constructions.[26e]

3.3.2. Biologically Active Tetrazoles

The best well-known example of tetrazole-based antihypertensive medicine is losartan (**3.042, Figure 3.08**) and its analogues and the compounds may prove useful for the treatment of diabetes. Ram *et al*., [27a-b] have reported that the compounds **3.043** and **3.044** act as promising drug candidate against cerebral ischemia, schizophrenia. Ornstein *et al*., have synthesised a compound **3.045** by 6-substitution of decahydroisoquinoline-3-carboxylic acids [28a] and found that the compound behaved as antagonist of AMPA–2-amino-3-(3-hydroxy-5methylisoxazol-4-yl) propionic acid and kainase receptors. Another class of glutamate receptor agonists is the derivatives of isoxazoles **3.046** which has been reported by Ornstein and Dominguez *et al*.[28b-c]

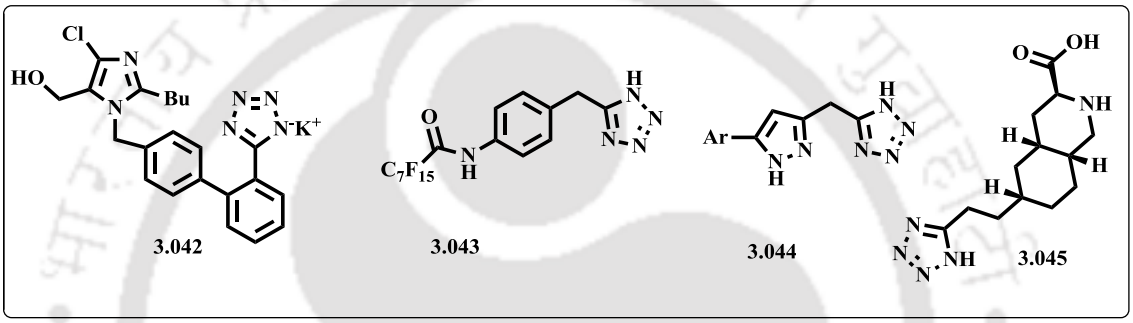


Figure 3.08. Examples of biologically active aminotetrazoles.

1,5-disubstituted tetrazole (**3.047-3.048, Figure 3.09**) fragments were studied actively and it has been found that such compounds are anti-inflammatory agents. On the other hand, they also exhibit weak antiulcer and analgesic activity [29a,b]. Bayer *et al*., have synthesised derivatives of 3'-(5-amino1,2,3,4-tetrazol-4-yl)-3'-deoxythymidines **3.049**, which show activity against the human immune deficiency virus. [30a]

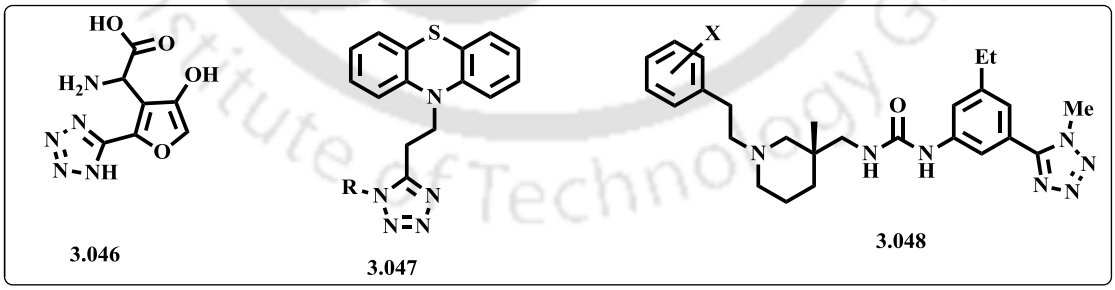


Figure 3.09. Examples of biologically active aminotetrazoles.

Another tetrazole-based drug found in cephalosporin class [30b-c] which has been reported by Powers and lee *et al*. Cephalosporin has almost similar structure to penicillinin but these types of antibiotics have low toxicity, for example: latamoxef (**3.050, Figure 3.10).** Habich *et al*., have synthesised very interesting terazole containing peptide molecule **3.051** with various

substituents at position-1 of the tetrazole and reported that such products can be used as antihypertensive agents.[31a] Another class of 2,5-disubstituted tetrazole derivative 9H-xanthene-9-carboxylic acid **3.052**, has been reported by Huwyler *et al*., where tetrazole has been replaced by oxadiazole ring.[31b] Such compounds may find use as glutamate receptor modulators. Similar kind of compound **3.053** has been studied by Chang *et al*., which showed antiviral activity. [31c]

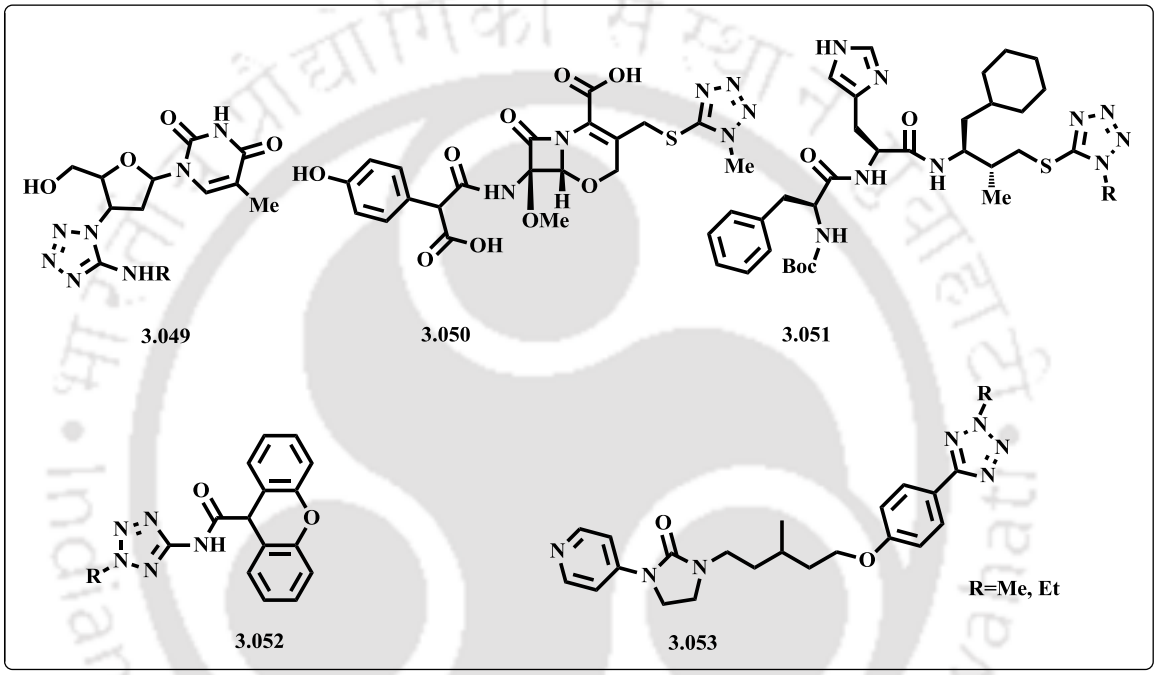


Figure 3.10. Examples of biologically active aminotetrazoles.

3.4. Background

The interest in the synthesis and application of novel unnatural amino acids (UNAAs) is growing universally in recent years. Moreover, the synthesis of unnatural amino acids with novel functionality is highly challenging and an emerging area of research. Side chain modified UNAAs are very important and an attractive tool for the development of less perturbing amino acids. The complementary use of both genetic and chemical methods has provided a wide scope of preparation of protein construct with either natural or synthetically modified amino acid residues.[11b,c] In addition, side chain modified amino acids can take part in site-selective protein modification which is very important for the preparation of the next generation of biopharmaceuticals. Furthermore, side chain modified UNAAs with a fluorophoric group applied for generation of unnatural fluorescent amino acids has been explored previously. Tetrazoles and thioureas are important class of scaffolds in peptidomimetic chemistry. However, application of these scaffolds in the UNAAs synthesis is less explored. Therefore, generation of side chain modification with the above scaffolds would be an attractive area of research in amino acids and protein field.

Recently, our research group is involved in the synthesis of side chain modified UNAAs. In the present chapter, we report the well-known chemistry of –NCS which has been exploited to generate thioureayl alanines (^{TU}**Ala = Tua**) as a new class of UNAAs which is further converted to aminotetrazolyl alanines (^{ATz}**Ala = Ata**) as another class of new UNAAs. Herein, we also report and study the photophysical property of pyrenylthioureayl and aminotetrazolyl UNAAs as highly emissive unnatural amino acids.

3.5. Objective

Several synthetic methodologies have appeared in the literature for the generation of substituted thioureido and tetrazoles including 5-amino tetrazoles. However, there is no report of unnatural thioureayl and aminotetrazolyl amino acids. Hence, we are the first to report thioureayl and 5-aminotetrazolyl as novel class of unnatural amino acids. Considering the importance of UNAAs in biological system, we made an effort to the synthesis of new thioureayl and aminotetrazolyl unnatural amino acids. These unnatural amino acids can be used for various applications like site-specific modification in protein, genetic incorporation and as sensor.

Utilising the above concept we outlined our objectives as follows:

(a) Design and synthesis of thioureayl alanines (^{TU}**Ala**) *via* nucleophilic addition of aromatic and aliphatic amines **(Figure 3.11)** to isothiocyanyl alanine (**Ita**).

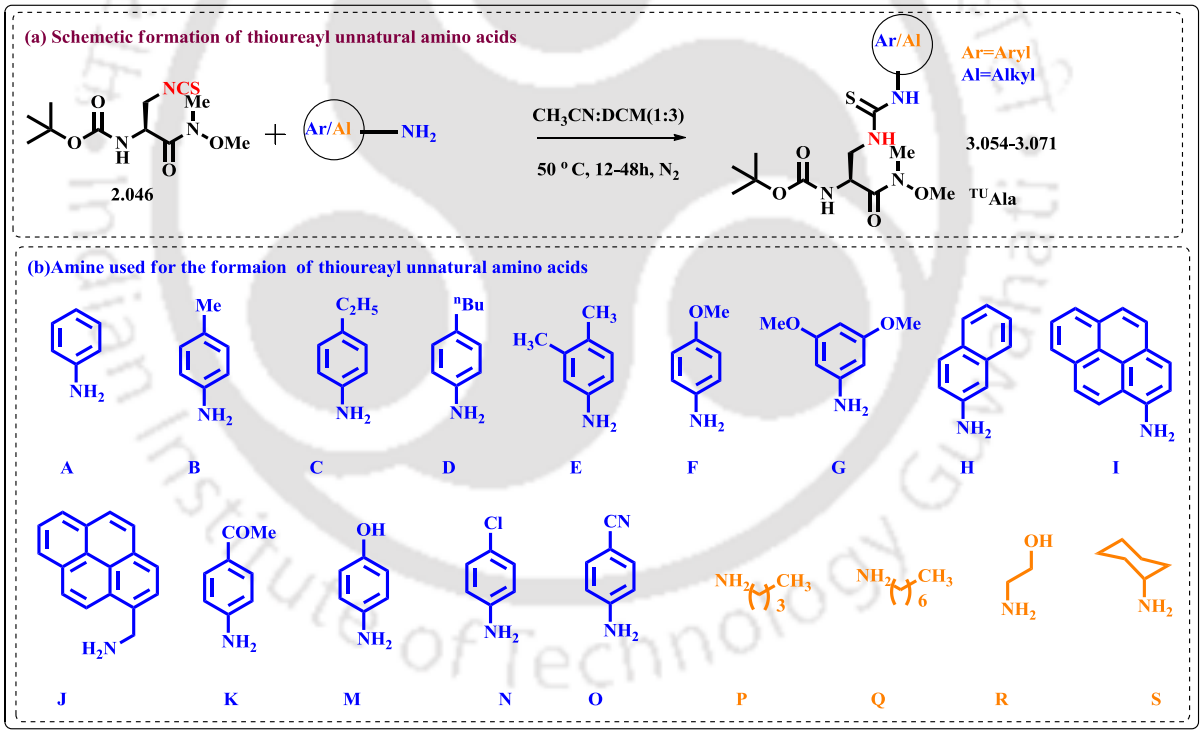


Figure 3.11. Schematic of formation of thioureayl alanine amino acids.

(b) Study of $Cu(OAc)_2$ mediated desulfurisation of thioureayl alanines **(Figure 3.12)** for the formation of carbodiimide intermediates followed by their electrocyclisation with azide to generate aminotetrazolyl unnatural amino acids.

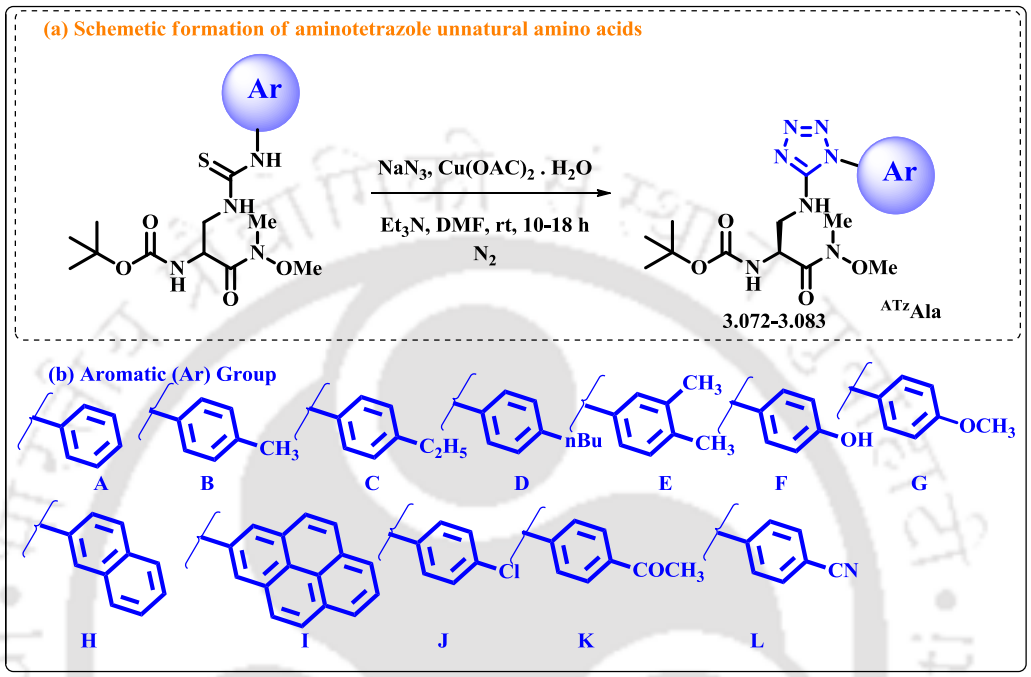


Figure 3.12. Schematic of formation of 5-aminotetrazolyl alanine amino acids.

(c) Study of photophysical properties of methylpyrenylthioureayl alanine (^{MePyTU}**Ala**), pyrenylthioureayl alanine (^{PyTU}**Ala**) **(Figure 3.13)**, and pyrenylaminotetrazolyl alanine (^{PyTzA}**Ala**) unnatural amino acids by UV-visible and fluorescence spectroscopy.

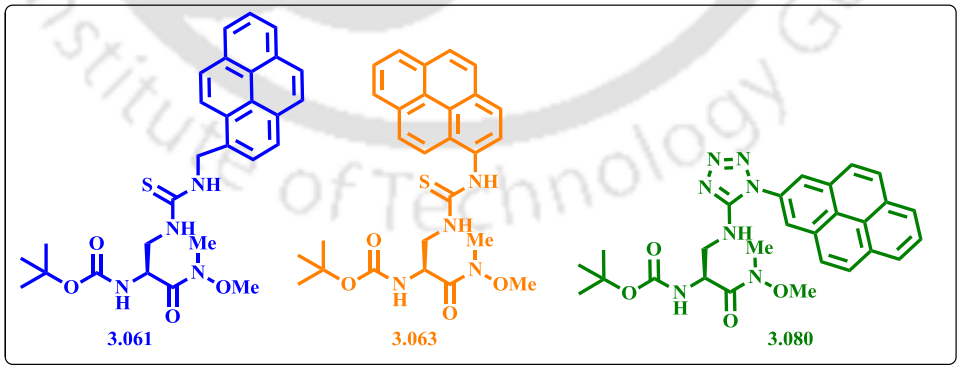
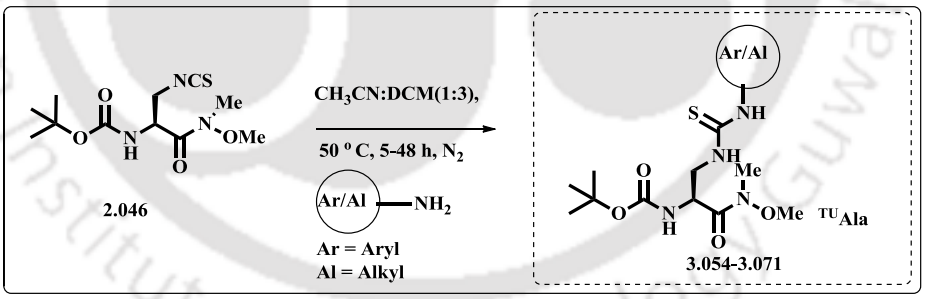


Figure 3.13. Synthesised fluorescent pyrenylthioureayl and pyrenylaminotetrazolyl amino acids.

3.6. Results and Discussion

3.6.1. Synthesis of Unnatural Thioureayl Alanines

We have already discussed the synthesis of **^{NCS}Ala (2.046)** in **Chapter-2**. Here, the electrophilicity of –NCS functionality of **^{NCS}Ala** was first used to achieve the synthesis of thioureayl unreported side-chain chiral thioureayl alanines (**^{TU}Ala = Tua**) **(Scheme 3.01)**. Therefore, we carried out the simple condensation reaction of **^{NCS}Ala** with various amines (**Figure 3.11, A-S)** in CH_3CN: DCM (1:3) solvent at 50 ^{o} C for about 12-48 hours (**Scheme 3.01**). Yields of the products also varied. All the aromatic amines irrespective of the nature of the substituents, except for *p*-acetyl aniline, afforded thioureayl alanines in very good to excellent yield of 81-98% (**Figure 3.14**). The *p*-acetyl aniline afforded moderate yield of thioureayl alanines (**3.075**, 65%) which might be because of the formation of other undetectable side product. In case of aliphatic amines, the yields were good to moderate. The amines with short carbon chain (ethanolamine-**R, Figure 3.11**) or compact amine (cyclohexylamine-**S, Figure 3.11**) gave very good yield (70-78%) while the long chain amines (butyl-**P, Figure 3.11** or heptylamine-**Q, Figure 3.11**) yielded moderately (50-60%) which might be attributed to the steric effect. A correlation based on basicity/nucleophilicity is at this stage very difficult to access. However, in a nut shell, the novel thioureayl amino acids are the first report which can be easily utilised for catalytic transformation which is our future research focus.[32] Formation of all compounds was confirmed by NMR, ESI-HRMS and IR.



Scheme 3.01. Synthetic scheme for thioureayl unnatural amino acids.

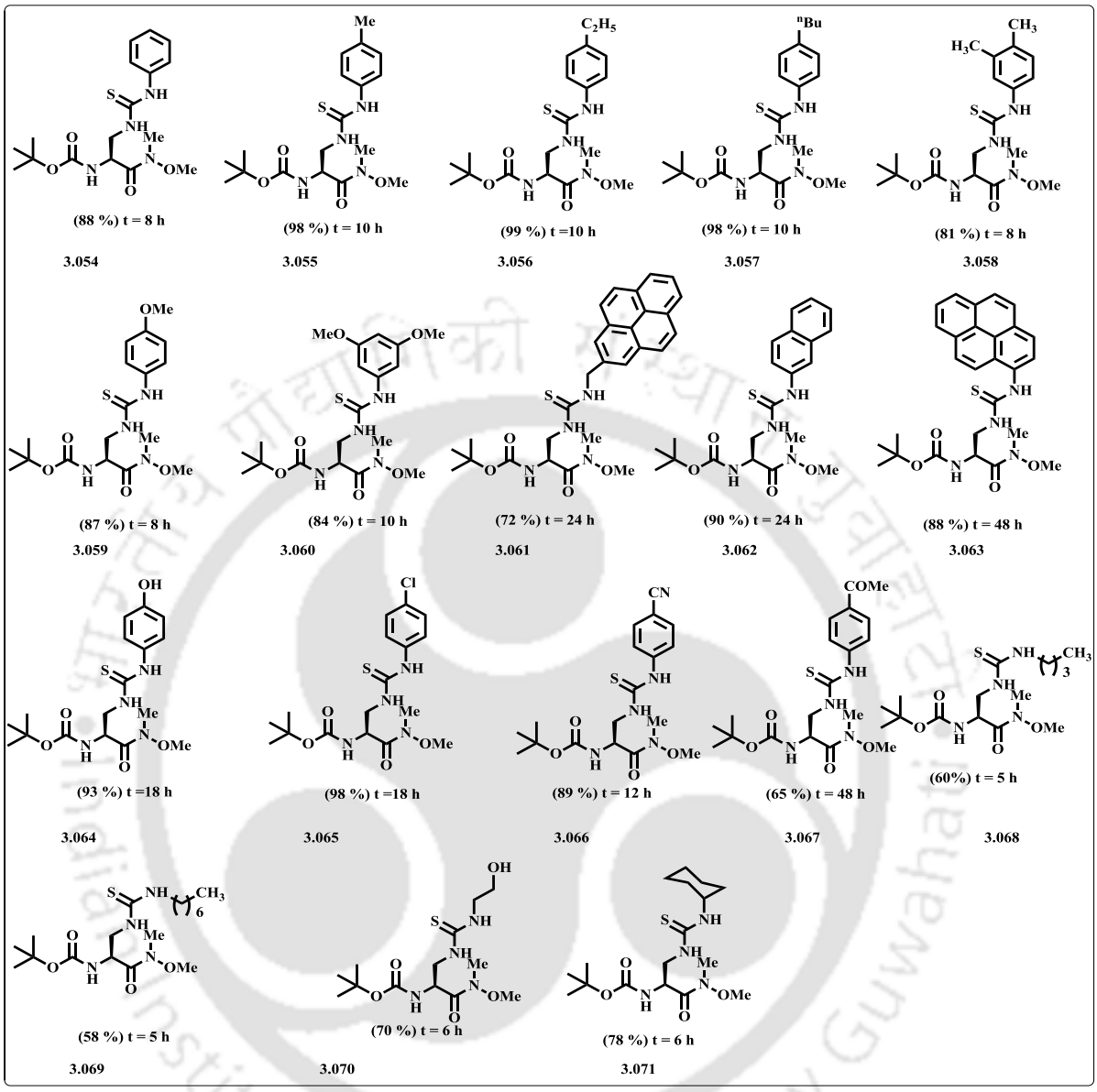
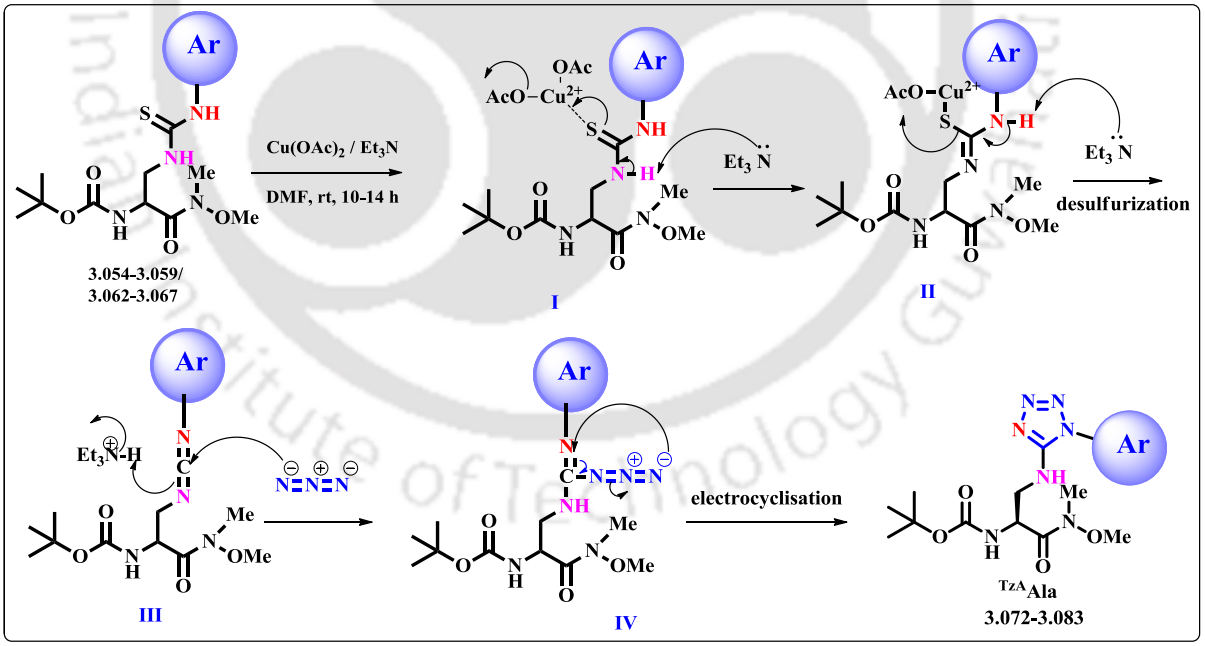


Figure 3.14. Structures of the synthesised N,C-diprotected thioureayl alanines.

3.6.2. Synthesis of Unnatural Aminotetrazolyl Alanines (^{TzA}Ala)

After successful synthesis of several thioureayl alanines (**^{TU}Ala = Tua)**, we thought that the intermediate generated from an oxidative desulfurisation of various **^{TU}Ala** could be trapped by external nucleophiles such as sodium azide which would ultimately lead to the generation of another novel class of 5-aminotetrazolyl alanines **^{TzA}Ala** (**Ata**) *via* electrocyclisation. Among the various available strategies, we adopted $Cu(OAc)_2$ mediated desulfurisation of thioureayl (**Scheme 3.02)** alanines for the formation of carbodiimide intermediates followed by their electrocyclisation with azide to generate aminotetrazolyl alanines. We are the first to report this

novel class of unnatural amino acids in good to excellent yields **(Figure 3.15).** The synthesis was carried out by stirring a solution of thioureayl alanines in DMF at room temperature with NaN_3 (3 equiv.) in presence of $Cu(OAc)_2/Et_3N$ for 12 hours. We synthesised various 5-aminotetrazolyl alanines in good to excellent yields (45%-85%). Thus, 1-phenyl-5-aminotetrazolyl alanine (**3.072**) from phenylthioureayl alanine (**3.054**) was obtained in moderate yield. High yields of 1-aryl-5-aminotetrazolyl alanines were observed in cases where aryl group of arylthioureayl alanines were substituted with –Me, -Et, -nBu, dimethyl, -OH and -Cl substituents. However, strong electron donating (-OMe) or strong electron withdrawing (-COMe/CN) groups reduced the yield from moderate to low (61%-45%). The regiochemistry with respect to the tetrazole moiety is the reflection of the difference in the pK_as of the amino functionalities onto the thioureayl alanines.[33] Thus, amine with less basicity, in this case the aromatic amines, became the part of the ring nitrogen and the other amine of the *β*-amino alanyl moiety with more basicity was the part of the exocyclic nitrogen with respect to the tetrazole moiety. This may be explained by considering the protonation of more basic 'N' of carbodiimide intermediate **III** (**Scheme 3.02)**, followed by nucleophilic attack by azide to give azidoguanidyl alanine **IV** (**Scheme 3.02)**, which underwent electrocyclisation giving products with the observed regiochemistry at the tetrazole unit. Formation of all compounds was confirmed by NMR, ESI-HRMS and IR.



Scheme 3.02. The plausible mechanism of regiochemical outcome for the synthesis of aminotetrazolyl alanines (**^{TzA}Ala**) from thioureayl alanines (**^{TU}Ala**).

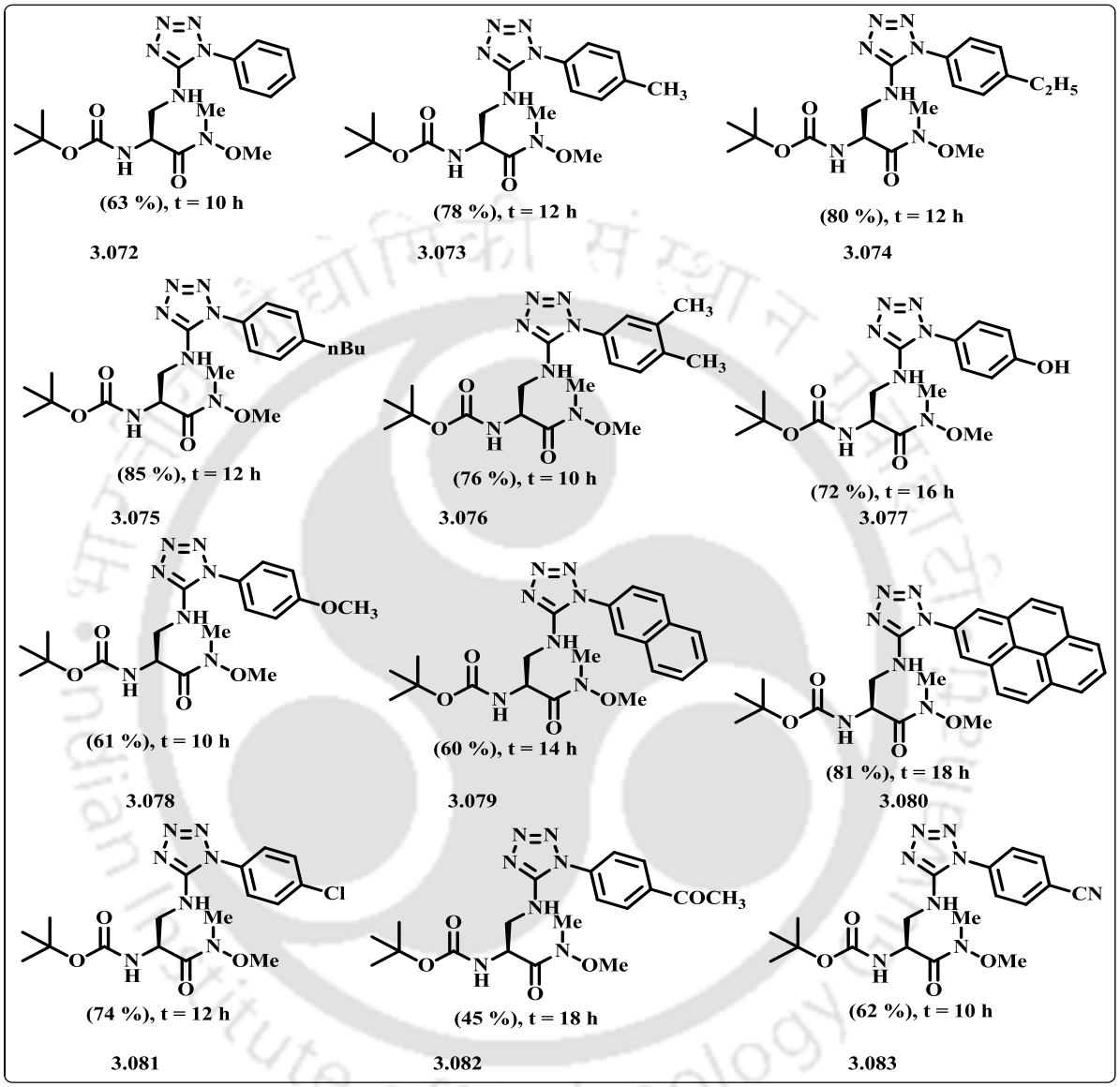


Figure 3.15. Structures of the synthesised N,C-diprotected 5-aminotetrazolyl alanines.

3.6.3. ^{1}H-NMR Characterisation of Representative Thioureayl Alanine and Aminotetrazolyl Alanine Unnatural Amino Acids

All the amino acids in the fully protected form were characterised mainly by NMR spectroscopy. For amino acid **3.057 (Figure 3.16a)**, the hydrogen of aromatic unit appeared at δ 7.16 and 7.12 as a doublet with coupling constant J = 7.6Hz and 8.1Hz respectively. One NH (NH_c) appeared as a broad singlet at δ 8.303, another NH (-NH_b) appeared as a doublet with J= 5.7Hz at δ 6.68 and BOC-NH_a appeared as a singlet at δ 5.62. The α–CH hydrogen of alanyl unit was found as a broad singlet (bs) at δ 4.81ppm and β–CH_2 hydrogens appeared as a multiplet at δ 4.03 to 3.77. Three protons of N-Me and OMe group appeared as a singlet at δ 3.14 and 3.73 respectively. N-butyl group hydrogen appeared at δ 2.56 as a triplet with J=7.7 Hz, δ 1.54 as a double doublet with J=15.1 Hz, 7.8 Hz and δ 0.88 as a double doublet with J=7.6 Hz, 7.1 Hz. The ^{t}Bu hydrogens of Boc- group and two hydrogen of n-Bu group appeared as a singlet at δ 1.33.

For amino acid **3.075 (Figure 3.16b)**, the hydrogen of aromatic unit appeared at δ 7.36 and 7.30 as a double with coupling constant J = 7.9 Hz and 7.8 Hz respectively. One NH (-NH_b) appeared as a broad singlet at δ 5.36 and another NH (-NH_a) appeared as a doublet at δ 5.72 with J=6.7 Hz. The α–CH hydrogen of alanyl unit appeared as a broad singlet (bs) at δ 4.87. The β–CH_2 hydrogens of alanyl unit appeared as a broad singlet at δ 3.85 and 3.58. Methyl hydrogens of N-Me and OMe group appeared as a singlet at δ 3.14 and 3.76 respectively. N-butyl group hydrogen appeared at δ 2.63 as a triplet with J=7.8 Hz, at δ 1.58 as a doublet of triplet with J=15.3 Hz, 7.7 Hz and at δ 0.90 as a triplet with J=7.4 Hz. The ^{t}Bu hydrogens of Boc group and two hydrogens of aromatic ^{t}Bu group appeared as a singlet at δ 1.32.

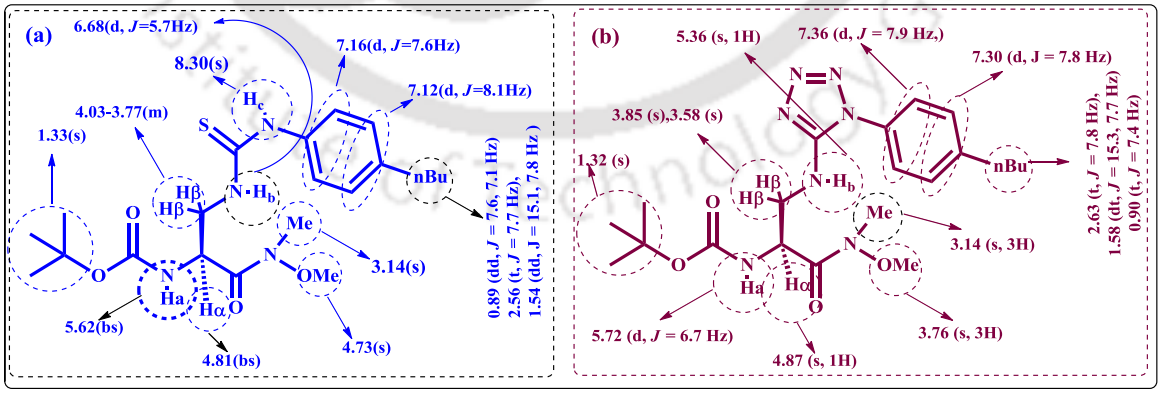


Figure 3.16. The chemical shift assignment for unnatural amino acids **3.057** and **3.075.**

3.6.4 Study of Photophysical Properties

Finally, we want to highlight the preliminary photophysical property of the fluorescent amino acids **(3.061, 3.063, and 3.080**) in various organic solvents to check the solvatochromic nature. The UV–visible spectra of all the compounds (10 μm) were measured using a UV-visible Spectrophotometer with a cell of 1 cm path length at 25 ^{o} C and 1 nm slit width. All the sample solutions were prepared before an hour of the measurement. Fluorescence emission spectra were recorded in a fluorescence spectrophotometer using a cell of 1 cm path length and 3nm excitation/emission slit width at 25 ^{o} C. The excitation wavelength for recording the emission spectra was set at the maximum wavelength of absorbance (λ_{abs}) in each case.

The UV-visible spectra of pyrenylthioureayl alanine **3.063** (**^{PyTU}Ala**), showed slightly distorted shapes of absorption bands compared to the pyrene in all solvents. The intensity of absorption followed an irregular trend with a minute shift (2-4 nm) of absorption wavelength as the polarity of solvent increases from toluene (329, 346 nm) to MeOH (326, 341 nm) (**Table 3.1)**. However, in THF and DMF, structureless broad absorption centered at 361 nm appeared in the spectra (**Figure 3.17a**). All these observations reflected the nominal solvatochromicity and spectral modulation by the thiourea moiety.[34a-b] Whereas, in fluorescence an irregular trend of emission had appeared at short wavelength range 390-398 nm and 410-425 nm in long wavelength range. The emission was drastically quenched in polar solvents like DMSO, ACN, MeOH etc (**Figure 3.17b**).[34c-d]

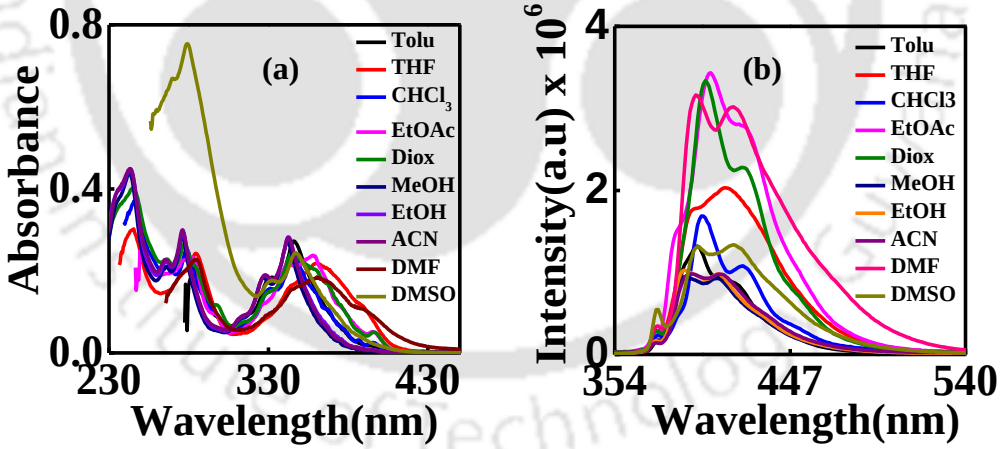


Figure 3.17. (a) UV-visible and (b) fluorescence spectra of amino acid **3.063** (^{PyTU}Ala) in various organic solvents (concentration of amino acids was 10 μM).

Similarly, we have recorded the UV-visible spectra of pyrenylmethylthioureayl alanine **3.061** (**^{MePyTU}Ala**). Thus, the UV-visible spectra showed characteristic pyrenyl absorbance bands at 347, 331, 317 nm in toluene (**Table 3.1)**. As the solvent polarity increases, it showed blue shifted solvatochromism along with increased absorbance at 342, 326, 311 nm in MeOH (**Figure 3.18a**). Upon excitation at 345 nm and increasing the solvent polarity from dioxane to

MeOH it showed emission at 376, 396, and 420 nm (**Table 3.1)** with decreased intensity which may be due to increased non-radiative decay *via* solvent-solute or H-bonding interaction (**Figure 3.18b**). [34c-d]

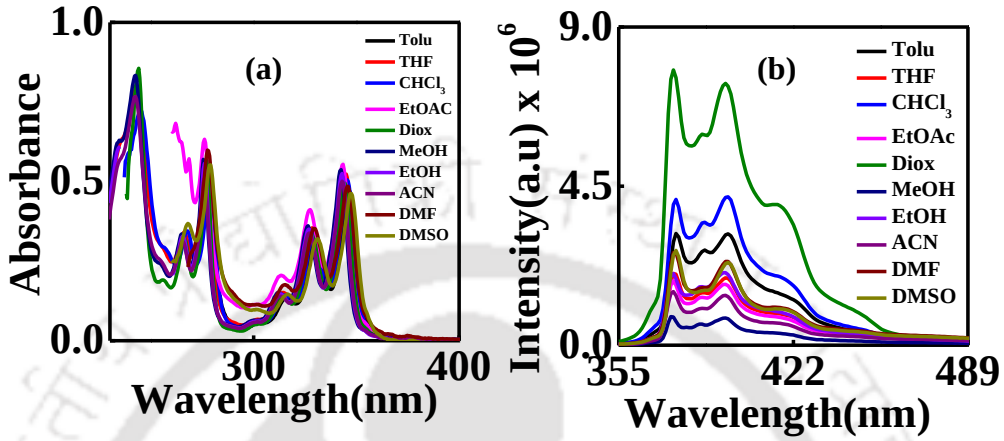


Figure 3.18. (a) UV-visible and (b) fluorescence spectra of amino acid of amino acid **3.061** (**^{MePyTU}Ala**) in various organic solvents (concentration of amino acids was 10 μM).

On the other hand, the corresponding tetrazole derivative i.e. pyrenylaminotetrazolyl alanine **3.080** (**^{PyTzA}Ala**) showed both the absorption and emission characteristics of pyrene. In the UV visible spectra a blue shifted solvatochromicity with increased absorption was observed for all three absorptions as the solvent polarity was increased from toluene (316, 331, 346 nm) to MeOH (311, 327, 342 nm) (**Figure 3.19a**). Upon excitation at 345 nm, the three emission bands appearing at around 375, 395 and 415 nm experienced the enhancement of emission with a blue shift (**Table 3.1)** as the solvent polarity was increased from toluene to MeOH except for ethyl acetate which had a quenching effect (**Figure 3.19b**).

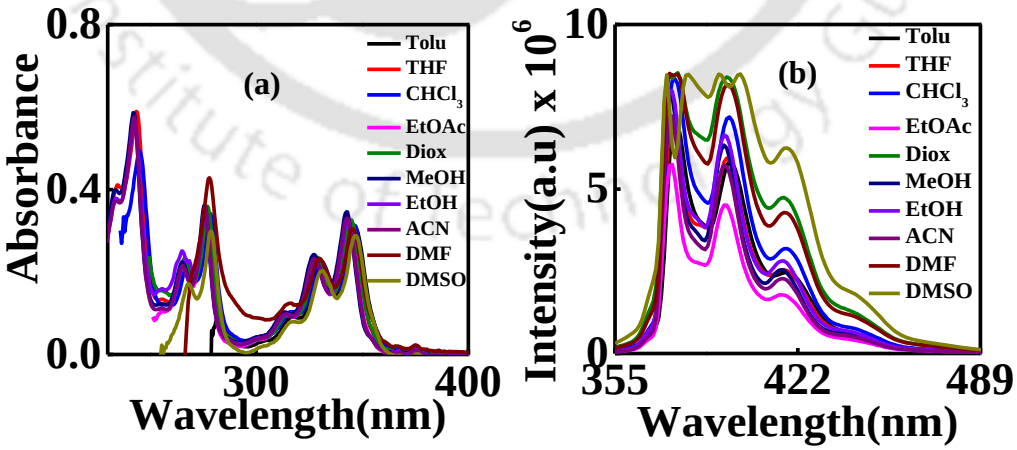


Figure 3.19. (a) UV-visible and (b) fluorescence spectra of amino acid **3.080** (**^{PyTzA}Ala**) in various organic solvents (concentration of amino acids was 10 μM).

Table 3.1. Summary of photophysical properties of amino acids **3.063** (^{PyTU}**Ala**), **3.061** (^{MePyTU}**Ala**) and **3.080** (^{PyTzA}**Ala**)

Entry→	**3.063 (^{PyTU}Ala)**									
Solvent→	Tolu	THF	$CHCl_3$	EtOAc	Diox	MeOH	EtOH	ACN	DMF	DMSO
Δf	0.013	0.210	0.148	0.201	0.021	0.309	0.290	0.305	0.275	0.265
(λ_{abs}) nm	346 329 315 282	360 343 285	345 329 279	358 342 281	345 329 277	341 326 275	342 327 276	342 327 276	389 361 345	347 331 276
(ε_{max} X 10^3) M^{-1} cm^{-1}	27.2	16.7	23.6	21.1	24.1	22.02	24.4	24.0	15.6	24.0
(λ_{fl}) nm	415 398	411 394	422 401	422 405	422 401	408 390	409 391	411 394	416 396	417 398
Entry→	**3.061 (^{MePyTU}Ala)**									
Δf	0.013	0.210	0.148	0.201	0.021	0.309	0.290	0.305	0.275	0.265
(λ_{abs})nm	347 329 315	344 329 315 277	346 328 314 279	343 327 313 275	345 327 313 277	343 326 311 276	343 326 311 276	343 327 312 276	347 328 313 277	347 331 316 278
(ε_{max} X 10^3) M^{-1} cm^{-1}	46.7	42.0	46.0	30.7	32.9	18.6	22.7	23.7	45.0	46.1
(λ_{fl})nm	421 397 387 377	418 395 386 376	421 397 387 377	419, 396 386 375	416 396 376 387	419 394 375 387	420 395 385 375	419 396 375	419 397 376	419 396 376
Entry→	**3.080 (^{PyTzA}Ala)**									
Δf	0.013	0.210	0.148	0.201	0.021	0.309	0.290	0.305	0.275	0.265
(λ_{abs})nm	346 330 316 283	344 330 314 277	345 330 315 278	344 328 313 276	345 329 315 276	343 327 312 275	343 327 313 275	343 328 312 275	345 329 315 278	346 331 315 278
(ε_{max} X 10^3) M^{-1} cm^{-1}	30.3	30.7	31.3	26.64	30.9	24.2	25.6	25.1	30.2	28.4
(λ_{fl})nm	420 397 377 445	443 417 396 376	441 418 397 377	440 416 396 375	443 417 396 375	440 416 395 375	442 415 395 375	441 416 395 374	443 417 396 377	444 418 400 381

***Δf values is the Solvent polarity function where $\Delta f = [\{(D-1)/(2D+1)\}-\{(n^2-1)/(2n^2+1)\}]$**

3.7. Conclusion

In conclusion, we have successfully demonstrated the multifaceted use of unexplored unnatural amino acids, isothiocyanyl alanine ($^{\mathbf{NCS}}$**Ala**, **Ita**). A variety of aliphatic and aromatics substituted thioureayl alanines from isothiocyanyl alanine and subsequently the aromatic substituted aminotetrazolyl alanines have been successfully synthesised in good to excellent yields. The new thioureayl alanine amino acids are the first report which can be easily utilised for catalytic transformation which is our future research focus. The regiochemistry of 5-aminotetrazolyl alanines $^{\mathbf{ATz}}$**Ala** (**Ata**), the third class of new unnatural amino acids, with respect to the tetrazole moiety has been explained on the basis of the difference in the pK_as of the amino functionalities in the thioureayl alanines. Three of the pyrene labeled amino acids from two new class possessed good photophysical properties which could find applications in future. The thiourea moiety did not show any modulation of fluorescence of a probe probably due to an increase in distance from the probe core. The result is exemplified in the fluorescence of pyrenylmethylthioureayl alanine ($^{\mathbf{MePyTU}}$**Ala**) which exhibited both the absorption and the fluorescence property, characteristic of pyrene and similar to what was observed for pyrenylaminotetrazolyl alanine. The UV-visible, fluorescence property of the thioureayl amino acids can be utilised for the metal sensing study which is described in **Chapter 5**.

3.8. Experimental section

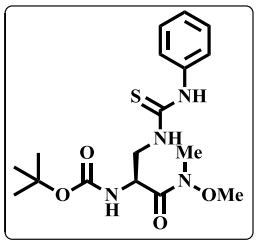
3.8.1. General Experimental

All the reactions were carried out under nitrogen atmosphere using oven-dried round bottom flasks. Reagents were purchased at the highest commercial quality and used without further purification, unless otherwise state. Reactions were monitored by thin layer chromatography (TLC) carried out on a 0.25 mm silica gel 60F–254 and silica gel-G (1:4) and visualized under UV light at 254 nm. Further visualisation was achieved by iodine vapor adsorbed on silica gel depending on the product type. Organic extracts were dried over anhydrous sodium sulfate. Solvents were removed in a rotary evaporator under reduced pressure. Column chromatography was performed on silica gel 60–120mesh using a mixture of hexane and ethyl acetate as mobile solvent. The isolated compounds were characterised by ^{1}H, ^{13}C NMR, 2D-NMR and IR spectroscopic techniques and Mass spectrometry. All the NMR spectra were recorded at ambient temperature on Bruker Ascend TM Aeon 600MHz spectrometer where ^{1}H frequency was 600MHz and ^{13}C frequency was 150MHz. NMR spectra for all the samples were measured in either in $CDCl_3$ or in d_6-DMSO. The chemical shift values were reported in ppm downfield from tetramethylsilane, using chloroform-d (δ= 7.26 for ^{1}H NMR, δ=77.23 for ^{13}C NMR) or using deuterated dimethyl sulfoxide-d_6 (δ= 2.50 for ^{1}H NMR, δ=39.50 for ^{13}C NMR). ^{1}H NMR coupling constant(s) *J* is represented in Hertz (Hz). All the NMR-FID has processed in MestReNova v6.0.2 software. High resolution mass spectra (HRMS) were recorded on a Water system Mass spectrometer in positive mode using electrospray ionisation-time of flight (ESI-TOF) and/or Atmospheric pressure chemical ionisation-Time of flight (APCI-TOF) reflection experiments. IR spectra were recorded in KBr plate in Perkin Elmer spectrometer and reported in frequency of absorption (cm^{-1}).

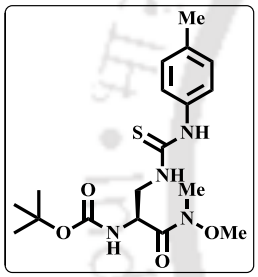
General procedure for the synthesis of thioureayl alanines (3.054-3.072, ^{TU}Ala = tua) from isothiocynyl alanine (2.046, ^{NCS}Ala): In 5 mL anhydrous CH_2Cl_2/CH_3CN (3:1) taken in a round bottom flask, **^{NCS}Ala** (200 mg, 0.69 mmol) was added under nitogen atmosphere. Degassing by N_2 gas for 5 minutes, various aliphatic/aromatic amine(s) (0.75 mmol) was added to it and the reaction mixture was heated at 50 ^{o} C for about 8-12 h (for almost all aromatic amines except for napthylamine, pyrenylamine, acetyl aniline, aminophenol, chloroaniline the time needed are 24 h, 48 h, 48 h, 18 h, 18 h, respectively and for aliphatic or cyclohexyl amines the time needed are 5-6 h) under N_2 atmosphere. After completion of the starting material as indicated by TLC, the reaction mixture was dried under high vacuum and directly packed into a silica gel (60-120) column. The residue was purified using hexane/ethyl acetate.

3.8.2. Synthesis and Characterisations

(S)-Phenylthioureaylalanine (3.054, ^{PhTU}Ala): Using the general procedure and starting from 0.69 mmol ***^{NCS}Ala***, the desired compound **3.054** was obtained as semi solid compound (R_f = 0.4 in 1:1, hexane/ethyl acetate). A semi solid; 231mg, Yield 87 %. IR (KBr) ῡ 3483, 3376, 3310, 2999, 2980, 2931, 1679, 1683, 1597.16, 1523, 1496, 1352, 1319, 1299, 1256, 1201, 1162, 1066, 1024, 991 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 8.13(1H, d, J = 42.3 Hz), 7.41 (2H, d, J = 6.8 Hz), 7.29 – 7.26 (3H, m), 6.79 (1H, d, J = 4.2 Hz), 5.63 (1H, t, J = 7.5 Hz), 4.85 (1H, d, J = 4.8 Hz), 4.07 – 4.06 (1H, m,), 3.85 – 3.70 (1H, m), 3.77 (3H, s), 3.19 (3H, s), 1.37 (9H, s); ^{13}C NMR (150 MHz; $CDCl_3$) δ 181.2, 170.2, 155.8, 136.1, 130.1, 127.2, 125.4, 80.3, 62.0, 50.5, 48.0, 32.5, 29.8, 28.4. +APCI-HRMS (m/z) calculated for $C_{17}H_{27}N_4O_4S$ $[M+H]^+$ 383.1753, found 383.1790.

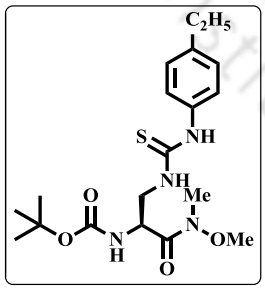


(S)-p-methylphenyl thioureayl alanine (3.055, ^{MeBTU}Ala)*:** Using the general procedure and starting from 0.36 mmol ***^{NCS}Ala, the desired compound **3.055** was obtained (R_f = 0.4 in 1:1, hexane/ethyl acetate). A semi solid compound; 140mg, Yield 98 %. IR (KBr) ῡ 3440, 2975, 2924, 2853, 1707, 1693, 1648, 1535, 1515, 1385, 1367, 1310, 1282, 1252, 1165, 1055, 1010, 988 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 8.27 (1H, s), 7.16 (2H, d, J = 7.8 Hz), 7.10 (2H, d, J = 7.9 Hz), 6.67 (1H, s), 5.63 (1H, d, J = 12.6 Hz), 4.81 (1H, s), 4.01 (1H , dt, J = 13.1, 4.5 Hz), 3.73 (4H, s), 3.14 (3H, s), 2.30 (3H, s), 1.34 (9H, s,); ^{13}C NMR (150 MHz; $CDCl_3$) δ 181.2, 170.2, 155.7, 137.1, 133.6, 130.5, 125.5, 80.1, 61.9, 50.5, 47.7, 32.4, 28.3, 21.1. +APCI-HRMS (m/z) calculated for $C_{18}H_{29}N_4O_4S$ $[M+H]^+$ 397.1910, found 397.1906.

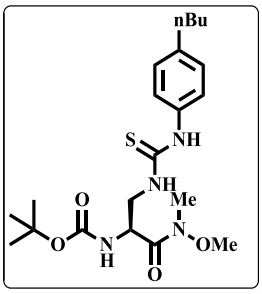


.

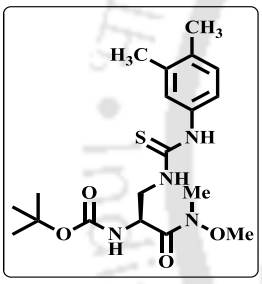
(S)-p-Ethylphenyl thioureayl alanine (3.056, ^{EtBTU}Ala)*:** Using the general procedure and starting from 0.38 mmol ***^{NCS}Ala, the desired compound **3.056** was obtained (R_f = 0.4 in 1:1, hexane/ethyl acetate). A white solid compound; 155mg, Yield 99 %; m.p. 147-151 ^{o}C. IR (KBr) ῡ 3442, 2969, 2931, 2850, 1701, 1651, 1533, 1515, 1390, 1367, 1314, 1253, 1166, 1055, 1022, 991 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 7.91 (1H, s), 7.22 (2H, d, J = 8.2 Hz), 7.14 (2H, d, J = 8.2 Hz), 6.66 (1H, s), 5.56 (1H, d, J = 5.4 Hz,), 4.83 (1H, s), 4.09 – 3.99 (1H, m), 3.85 – 3.78 (1H, m), 3.76 (3H, s), 3.18 (3H, s), 2.64 (2H, q, J = 7.6 Hz), 1.36 (9H, s), 1.23 (3H, t, J =7.6 Hz); ^{13}C NMR (150 MHz; $CDCl_3$) δ 181.4, 170.2, 155.8, 143.7, 133.6, 129.5, 125.7, 80.3, 62.0, 50.6, 48.1, 32.5, 28.5, 28.4, 15.5. +APCI-HRMS (m/z) calculated for $C_{19}H_{31}N_4O_4S$ $[M+H]^+$ 411.2066, found 411.2073.



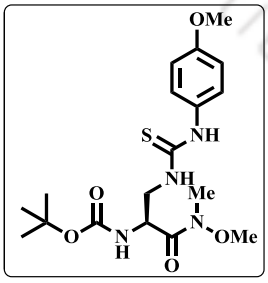
(S)-p-Butylphenyl thioureayl alanine (3.057, ^{BuBTU}Ala): Using the general procedure and starting from 0.41 mmol ***^{NCS}Ala***, the desired compound **3.057** was obtained (R_f = 0.4 in 1:1, hexane/ethyl acetate). A semi solid compound; 175mg, Yield 97 %. IR(KBr) ῡ 3363, 3308, 3208, 3150, 2962, 2929, 2857, 2104, 1676, 1649, 1614, 1536, 1521, 1445, 1372, 1317,1339, 1287, 1255, 1197, 1167, 1065, 1023, 989 cm^{-1}. ^{1}H NMR (600MHz; $CDCl_3$) δ 8.30 (1H s,), 7.16 (2H d, J = 7.6 Hz,), 7.12 (2H, d, J = 8.1 Hz), 6.68 (1H, s), 5.62 (1H, d, J = 5.7 Hz), 4.81 (1H, s), 4.05 – 3.97 (1H , m), 3.82 – 3.76 (1H, m), 3.73(3H, s), 3.14 (3H, s), 2.56 (2H, t, J = 7.7 Hz), 1.54 (2H, dd, J = 15.1, 7.8 Hz), 1.33 (12H, s), 0.89 (3H, dd, J = 7.6, 7.1 Hz); ^{13}C NMR (150 MHz; $CDCl_3$) δ 181.1, 170.2, 155.7,142.1, 133.7, 129.9, 125.3, 80.1, 61.9, 50.5, 47.6, 35.2, 33.5, 32.4, 28.3, 22.4, 14.0. +ESI-HRMS (m/z) calculated for $C_{21}H_{35}N_4O_4S$ $[M+H]^+$ 439.2379, found 439.2378.



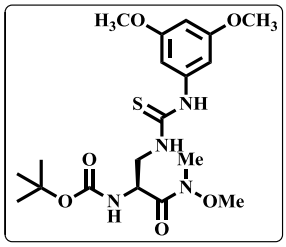
(S)-m,p-Dimethylphenyl thioureayl alanine (3.058,^{DMBTU}Ala): Using the general procedure and starting from 0.55 mmol ***^{NCS}Ala***, the desired compound **3.058** was obtained (R_f = 0.4 in 1:1, hexane/ethyl acetate). A solid compound; 170mg, Yield 75 %; m.p. 128-132 ^{o}C. IR (KBr) ῡ 3415, 3019, 2978, 2917, 2853, 1706, 1653, 1538, 1503, 1392, 1368, 1215, 1164, 1052, 1018, 991 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 8.25 (1H, s), 7.10 (1H, d, J = 7.5 Hz), 6.97 – 6.93 (2H, m), 6.65 (1H, s), 5.61 (1H, d, J = 8.1 Hz), 4.81 (1H, s), 3.99 (1H, dd, J = 8.3, 4.8 Hz,), 3.85 – 3.77 (1H, m,), 3.73 (4H, s), 3.14 (3H, s), 2.21(1H, s), 2.10 (3H, s), 1.33 (9H, s); ^{13}C NMR (150 MHz; $CDCl_3$) δ 181.0, 170.3, 155.6, 138.4, 135.8, 133.7, 130.9, 126.5, 122.8, 80.0, 61.9, 50.5, 47.4, 32.4, 28.2, 19.8, 19.4. +APCI-HRMS (m/z) calculated for $C_{19}H_{31}N_4O_4S$ $[M+H]^+$ 411.2066, found 411.2068.



(S)-p-Methoxyphenyl thioureayl alanine (3.059, ^{MOBTU}Ala): Using the general procedure and starting from 0.39 mmol ***^{NCS}Ala***, the desired compound **3.059** was obtained (R_f = 0.4 in 1:1, hexane/ethyl acetate). A white solid compound; 130mg, Yield 87 %; m.p. 115- 119 ^{o}C. IR (KBr) ῡ 3483, 1973, 2931, 2850, 1707, 1653, 1540, 1511, 1392, 1367, 1338, 1298, 1247, 1167, 1030, 985 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 7.71 (1H, s), 7.17 (2H, d, J = 8.5 Hz), 6.92 (2H, d, J = 8.8 Hz), 6.52 (1H, s), 5.54 (1H , d, J = 4.4 Hz), 4.81 (1H, s), 4.05 (1H, dd, J = 10.2, 5.4 Hz), 3.81 (3H, s), 3.77 (4H, s), 3.19 (3H, s), 1.37 (9H, s); ^{13}C NMR (150 MHz; $CDCl_3$) δ 181.9, 170.2, 159.1, 155.9, 128.5, 128.0, 115.3, 80.3, 62.0, 55.6, 50.7, 48.2, 32.5, 28.4. +ESI-HRMS (m/z) calculated for $C_{18}H_{29}N_4O_5S$ $[M+H]^+$ 413.1859, found 413.1843.

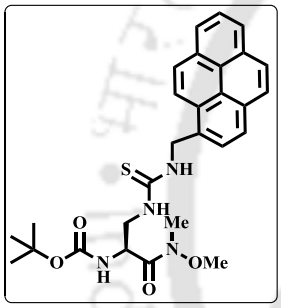


(S)-m, m-Dimethoxyphenyl thioureayl alanine (3.060, ^{DMOBTU}Ala)*:* Using the general procedure and starting from 0.36 mmol ***^{NCS}Ala***, the desired compound **3.060** was obtained (R_f = 0.4 in 1:1, hexane/ethyl acetate). A white solid compound; 135mg, Yield 84 %;



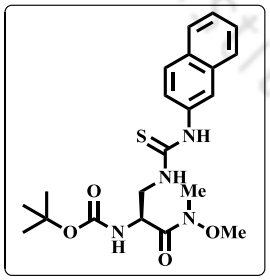
m.p. 139-143 ^{o} C. IR (KBr) ῡ 3471, 3439, 2961, 2917, 2856, 1707, 1643, 1603, 1538, 1451, 1390, 1365, 1332, 1268, 1204, 1160, 1055, 991 cm^{-1}.^{1}H NMR (600 MHz; $CDCl_3$) δ 8.37 (1H, dd, J = 28.0, 14.0 Hz), 6.91 (1H, s), 6.38 (2H, s), 6.29 (1H, s), 5.64 (1H, s), 4.83 (1H, d, J = 4.4 Hz,), 4.00 (1H, dd, J = 11.9, 5.0 Hz), 3.91 – 3.80 (1H, m), 3.74 (9H, s), 3.14 (3H, s), 1.33 (9H, s); ^{13}C NMR (150 MHz; $CDCl_3$) δ 180.7, 170.3, 161.7, 155.7, 137.9, 102.8, 99.1, 80.2, 61.9, 60.5, 55.5, 50.5, 32.4, 28.3. +APCI-HRMS (m/z) calculated for $C_{19}H_{31}N_4O_6S$ $[M+H]^+$ 443.1964, found 443.1976.

(S)-Pyrennylmethylthioureayl alanine (3.061, ^{MePyTU}Ala)*:* Using the general procedure and starting from 0.69 mmol ***^{NCS}Ala***, the desired compound **3.061** was obtained (R_f =



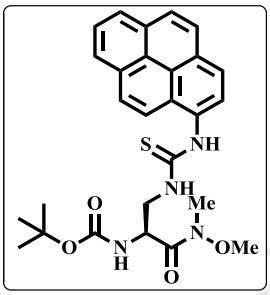
0.4 in 2:1, hexane/ethyl acetate). A semi solid compound; 260mg, Yield 72 %; m.p. 177-181 ^{o} C. IR (KBr) ῡ 3369, 3298, 3044, 2977, 2930, 1685, 1650, 1519, 1455, 1433, 1382, 1367, 1338, 1289, 1266, 1206, 1166, 1045, 1022 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 8.12 (3H, d, J = 2.9 Hz), 8.06 – 8.00 (2H, m), 7.96 (3H, d, J = 12.1 Hz), 7.88 (1H, d, J = 6.7 Hz), 7.01 (1H, s), 6.95 (1H, s), 5.86 (1H, s), 5.22 (2H, m), 4.71 (1H, s), 3.81 (2H, s), 3.46 (3H, s), 2.67 (3H, s), 1.32 (9H, s); ^{13}C NMR (150 MHz; $CDCl_3$) δ 182.5, 170.5, 155.9, 131.2, 130.7, 129.13, 128.2, 127.5, 127.3, 126.1, 125.4, 124.9, 124.7, 124.6, 123.0, 80.5, 61.6, 50.5, 47.1, 32.0, 29.8, 28.3. +ESI-HRMS (m/z) calculated for $C_{28}H_{33}N_4O_4S$ $[M+H]^+$ 521.2223, found 521.2232.

(S)-Naphthyl thioureayl alanine (3.062, ^{NapTU}Ala)*:* Using the general procedure and starting from 0.41 mmol ***^{NCS}Ala***, the desired compound **3.062** was obtained (R_f = 0.4 in 1:1,



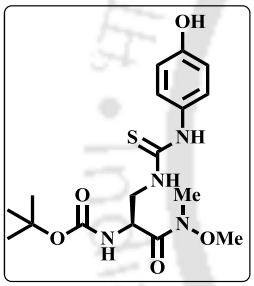
hexane/ethyl acetate). A semi solid compound; 160mg, Yield 90 %. IR (KBr) ῡ 3409, 2976, 2925, 2853, 1708, 1653, 1600, 1533, 1392, 1367, 1338, 1271, 1164, 1053, 1020, 990 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 8.15 (1H, s), 8.03 (1H, d, J = 8.3 Hz), 7.89 (2H, dd, J = 11.9, 8.2 Hz), 7.58 – 7.20 (4H, m), 6.51 (1H, s), 5.40 (1H, d, J = 7.4 Hz), 4.71 – 4.69 (1H, m), 3.94 (1H, dd, J = 11.4, 5.0 Hz), 3.80 – 3.76 (1H, m), 3.65 (3H, s), 3.06 (3H, s), 1.26 (9H, s); ^{13}C NMR (150 MHz; $CDCl_3$) δ 182.0, 170.1, 155.6, 134.8, 131.7, 130.1, 129.1, 128.6, 127.6, 127.2, 125.9, 125.6, 122.7, 80.2, 61.9, 50.3, 48.1, 32.4, 28.3. +ESI-HRMS (m/z) calculated for $C_{21}H_{29}N_4O_4S$ $[M+H]^+$ 433.1910, found 433.1905.

(S)-Pyrennyl thioureayl alanine (3.063, ^{PyTU}Ala): Using the general procedure and starting from 0.51 mmol ***^{NCS}Ala***, the desired compound **3.063** was obtained (R_f = 0.4 in 1:1, hexane/ethyl acetate). A light brown solid compound; 195mg, Yield 75 % ; m.p. 183-187 ^{o}C.



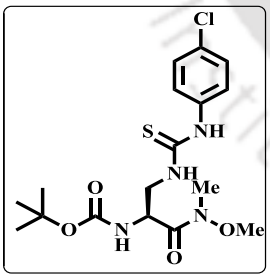
IR (KBr) ῡ 3380, 3354, 3208, 3035, 2967, 2932, 2806,1691, 1671, 1602, 1524, 1460, 1438, 1338, 1365, 1305, 1278, 1246, 1196, 1164, 1114, 1053, 1020, 993, 957 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 8.49 (1H, s), 8.21 – 8.13 (6H, m), 8.10 (1H, d, *J* = 8.9 Hz), 8.05– 8.01 (2H, m), 7.94 (1H, d, *J* = 8.0 Hz), 6.54 (1H, s), 5.43 (1H, d, *J* = 7.1 Hz,), 4.65(1H, d, *J* = 4.0 Hz,), 3.96-3.79 (2H, m,), 3.56 (3H, s), 2.98 (3H, s), 1.01 (9H, s); ^{13}C NMR(150 MHz; $CDCl_3$) δ 182.3, 170.1, 155.6, 131.3, 131.1, 131.0, 129.2, 128.7, 128.4, 127.9,127.1, 126.7, 126.0, 126.0, 125.6, 125.5, 124.5, 121.8, 80.0, 61.8, 50.4, 48.1, 32.3, 28.0. +APCI HRMS (m/z) calculated for ($C_{27}H_{31}N_4O_4S$) $[M+H]^+$ 507.2066, found 507.2063.

(S)-p-Hydroxyphenyl thioureayl alanine (3.064, ^{HBTU}Ala): Using the general procedure and starting from 0.39 mmol ***^{NCS}Ala***, the desired compound **3.064** was obtained (R_f = 0.4 in 1:2, hexane/ethyl acetate). A semi solid compound; 145mg, Yield 93 %. IR (KBr) ῡ 3471,



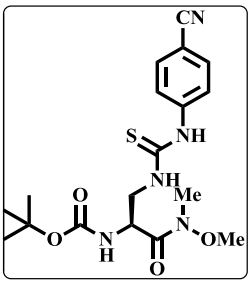
3439.68, 2975, 2920, 2850, 1690, 1637, 1538, 1510, 1449.19, 1390, 1371, 1338, 1249, 1157, 1096, 1055, 980 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 8.04 (1H, s), 7.02 (2H,d, *J* = 7.6 Hz), 6.79 (2H, d, *J* = 8.2 Hz), 6.53 (1H, s), 5.78 (1H, s), 4.87 (1H, s), 3.97 – 3.86 (2H, m), 3.74 (3H, s), 3.16 (3H, s), 1.39 (9H, s); ^{13}C NMR (150 MHz; $CDCl_3$) δ 181.7, 170.5, 156.2, 128.0, 127.8, 116.9, 80.7, 62.0, 50.6, 47.1, 32.5, 28.5.+APCI-HRMS(m/z) calculated for $C_{17}H_{27}N_4O_5S$ $[M+H]^+$ 399.1702, found 399.1715.

(S)-p-Chlorophenyl thioureayl alanine (3.065, ^{ClBTU}Ala): Using the general procedure and starting from 0.34 mmol ***^{NCS}Ala***, the desired compound **3.065** was obtained (R_f =

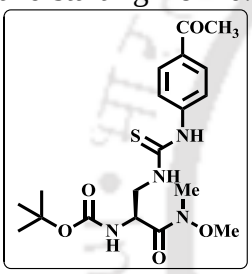


0.4 in 1:1, hexane/ethyl acetate). A solid compound; 139mg,Yield 98 %; m.p. 115-119 ^{o}C. IR (KBr) ῡ 3440, 2975, 2932, 2850, 1690, 1648, 1538, 1492, 1449, 1392, 1367, 1304, 1338, 1251, 1164, 1055, 1090, 1014, 990 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 8.75 (1H d, *J* = 11.4 Hz,), 7.29 (2H, d, *J* = 4.4 Hz), 7.21 (2H, d, *J* = 8.3 Hz), 6.94 (1H, s), 5.80 (1H, s), 4.82 (1H, s), 4.03 (1H, s), 3.74 (4H, s), 3.12 (3H, s), 1.33 (9H, s); ^{13}C NMR (150 MHz; $CDCl_3$) δ 181.1, 170.2, 155.9, 135.4, 132.1, 129.8, 126.4, 80.4, 61.9, 50.5, 47.6, 32.4, 28.3.+APCI-HRMS (m/z) calculated for $C_{17}H_{26}ClN_4O_4S$ $[M+H]^+$ 417.1363, found 417.1371.

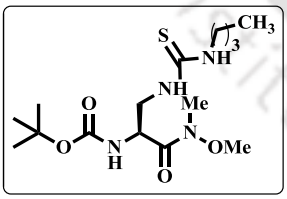
(S)-p-Cyanophenyl thioureayl alanine (3.066, ^{CNBTU}Ala): Using the general procedure and starting from 0.36 mmol ***^{NCS}Ala***, the desired compound **3.066** was obtained (R_f = 0.4 in 1:1, hexane/ethyl acetate). A white solid compound;130mg, Yield 88 %; m.p. 186-190 ^{o}C. IR (KBr) ῡ 3435, 3323, 3207, 3156, 3137, 2978, 2929, 2856, 2221, 1675, 1651, 1627, 1606, 1526, 1445, 1372, 1331, 1283, 1253, 1200, 1168, 1066, 1025, 989 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 8.93(1H, s), 7.60(2H, d, J = 7.2 Hz), 7.53 (2H, s), 7.39 (1H, s), 5.90 (1H, s), 4.85 (1H, s), 4.08-3.87(2H, m), 3.79 (3H, s), 3.18 (3H, s), 1.36 (9H, s); ^{13}C NMR (150 MHz; $CDCl_3$) δ 180.9, 170.3, 156.2, 141.9, 133.4, 123.4, 118.7, 108.4, 80.9, 62.0, 50.6, 47.5, 32.6, 28.4.+APCI-HRMS (m/z) calculated for $C_{18}H_{26}N_5O_4S$ $[M+H]^+$ 408.1706, found 408.1716.



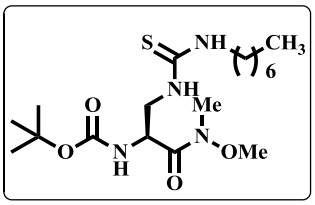
(S)-p-Acetylphenyl thioureayl alanine (3.067, ^{AcBTU}Ala): Using the general procedure and starting from 0.36 mmol ***^{NCS}Ala***, the desired compound **3.067** was obtained (R_f = 0.4 in 1:1, hexane/ethyl acetate). A gummy compound; 100mg, Yield 65 %. IR (KBr) ῡ 3449, 2975, 2924, 2850, 1701, 1647, 1603, 1537, 1449, 1367, 1338, 1270, 1165, 1049, 1201, 991 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 8.48 (1H, s), 7.97 (2H, d, J = 7.2 Hz), 7.41 (2H, s), 7.22 (1H, s), 5.72 (1H, s), 4.87 (1H, s), 4.17 – 4.03 (1H, m), 3.89 – 3.81 (1H, m), 3.79 (3H, s), 3.21 (3H, s), 2.57 (3H, s), 1.36 (9H, s).^{13}C NMR (150 MHz; $CDCl_3$) δ 197.03, 180.9, 170.4, 156.0, 141.7, 134.1, 130.0, 123.0, 80.6, 62.0, 50.6, 32.5, 28.4, 26.6.+APCI-HRMS (m/z) calculated for $C_{19}H_{29}N_4O_5S$ $[M+H]^+$ 425.1859, found 425.1869.



(S)-n-Butyl thioureayl alanine (3.068, ^{BuTU}Ala): Using the general procedure and starting from 0.34 mmol ***^{NCS}Ala***, the desired compound **3.068** was obtained (R_f = 0.4 in 1:2, hexane/ethyl acetate).A semi solid compound; 75mg, Yield 60 %. IR (KBr) ῡ 3374, 3308, 2959, 2929, 2856, 1684, 1647, 1549, 1524, 1484, 1434, 1387, 1341, 1283, 1252, 1212, 1170, 1062, 1022, 989 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 6.85 (1H, s), 6.54 (1H, s), 5.89 (1H, s), 4.79 (1H, s), 3.94 (1H, s), 3.7 6(4H, s), 3.18 (3H, s), 1.52 (2H, s), 1.39 (10H, s), 1.33 (2H, s), 0.88 (3H, s); ^{13}C NMR (150 MHz; $CDCl_3$) δ 182.0, 170.4, 156.4, 80.5, 61.9, 50.7, 43.9, 32.5, 30.9, 29.7, 28.4, 20.1, 13. +ESI-HRMS (m/z) calculated for $C_{15}H_{31}N_4O_4S$ $[M+H]^+$ 363.2066, found 363.2066.

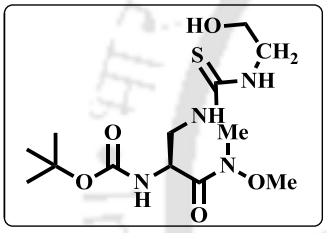


(S)-n-Heptyl thioureayl alanine (3.069, ^{HepTU}Ala): Using the general procedure and starting from 0.34 mmol ***^{NCS}Ala***, the desired compound **3.069** was obtained (R_f = 0.4 in 1:2, hexane/ethyl acetate). A white solid compound; 80mg, Yield 58 %; m.p. 112-116 ^{o}C. IR (KBr)



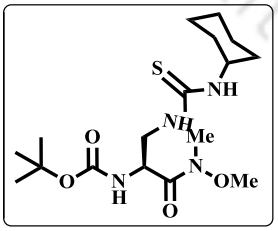
ῡ 3350, 2927, 2958, 2855, 2104, 1710, 1661, 1501, 1461, 1390, 1368, 1265, 1168, 1054, 991 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 6.88 (1H, s), 6.66 (1H, s), 6.55(1H, s) 5.89 (1H, s), 4.78 (1H, s), 3.93 (1H, s), 3.75 (4H, s), 3.17 (3H, s), 1.53 (2H, s), 1.38 (10H, s), 1.24 – 1.21(9H, m), 0.83 – 0.82 (3H, m); ^{13}C NMR (150 MHz; $CDCl_3$) δ 182.1, 170.4, 156.3, 80.5, 61.9, 50.7, 32.4, 31.8, 29.0, 28.9, 28.4, 28.4, 26.9, 22.6, 22.6, 14.1, 14.1. +ESI-HRMS (m/z) calculated for $C_{18}H_{37}N_4O_4S$ $[M+H]^+$ 405.2537, found 405.2547.

(S)-Ethanol thioureayl alanine (3.070, ^{EtOHTU}Ala): Using the general procedure and starting from 0.29 mmol ***^{NCS}Ala***, the desired compound **3.070** was obtained (R_f = 0.4 in 1:4, hexane/ethyl acetate). A semi solid compound; 71mg, Yield 70 %.



IR (KBr) ῡ 3475, 2979, 2932, 2874, 2017, 1673, 1566, 1528, 1385, 1301, 1253, 1207, 1162, 1054, 910 cm^{-1}.^{1}H NMR (600 MHz; $CDCl_3$) δ 7.21 (1H, s), 6.16 (1H, s), 4.81 (1H, s), 3.93 (2H, s), 3.80 (3H, s), 3.74 (3H, s), 3.67–3.39 (2H, m), 3.20 (3H, s), 1.41 (9H, s).^{13}C NMR (150 MHz; $CDCl_3$) δ 183.6, 171.0, 156.2, 80.6, 61.9, 61.6, 50.9, 46.8, 32.7, 29.8, 28.5.+ESI-HRMS (m/z) calculated for $C_{13}H_{27}N_4O_5S$ $[M+H]^+$ 351.1702, found 351.1702.

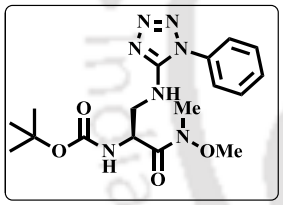
(S)-Cyclohexyl thioureayl alanine (3.071, ^{CyHTU}Ala): Using the general procedure and starting from 0.41 mmol ***^{NCS}Ala***, the desired compound **3.071** was obtained (R_f = 0.4 in 1:2, hexane/ethyl acetate). A white solid compound; 125mg, Yield 78 % ; m.p. 165-169 ^{o}C. IR



(KBr) ῡ 3365, 3305, 3081, 2980, 2932, 2853, 1686, 1650, 1545, 1451, 1382, 1367, 1341, 1316, 1288, 1256, 1214, 1170, 1056, 1021, 988 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 6.87 (1H s,), 6.35 (1H s,), 5.82 (1H s,), 4.79 (1H, s), 3.95 (1H, s), 3.76(4H, s,), 3.62 (1H, s), 3.19 (3H, s,), 2.02–1.94 (3H, m,), 1.71 (2H, t), 1.60 (1H ,d, J=12.6Hz,), 1.41(9H, s), 1.38 – 1.28 (2H, m), 1.23 – 1.13 (3H, m); ^{13}C NMR (150 MHz; $CDCl_3$) δ 180.9, 170.2, 156.6, 80.6, 62.0, 52.8, 50.7, 32.8, 32.5, 28.4, 25.5, 24.9, 24.8. +ESI-HRMS (m/z) calculated for $C_{17}H_{33}N_4O_4S$ $[M+H]^+$ 389.2223, found 389.2229.

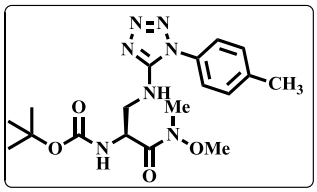
General procedure for the synthesis of 5-aminotetrazolyl alanines (3.072-3.083, ^{TzA}Ala = Tza) from aromatic thioureayl alanines (3.054-3.059, 3.062-3.067, TUAla = tua): In a flame dried round bottom flask various thioureayl alanines **(3.054-3.059, 3.062-3.067)** were dissolved in 5 ml dry DMF under N_2 atmosphere. Evacuation under high vacuum pump for 5 minutes and degassing by N_2 gas 3 equivalent of NaN_3, 0.5 equivalent of Cu $(OAC)_2$, H_2O and 3 equivalent of Et_3N were added to reaction mixture while degassing was continued. Then the reaction mixture was degassed for another 2 minutes and stirred for 8-16 hour at room temperature under N_2 atmosphere. After completion of the reaction indicated by TLC (hexane/ethyl, 1:1-1:2; R_f = 0.4), the residue was extracted with ethyl acetate and washed with saturated ammonium chloride and brine solution. The organic extract was dried over Na_2SO_4, evaporated to dryness in a rotary evaporator and purified on silica gel (60-120) column chromatography using using (hexane/ethyl, 1:1-1:2) as eluent to afford desired 5-aminotetrazolyl alanines (**3.080-3.091**) as solid or semi-solid material with good to excellent yield (60-85%). All the compounds were characterised by ^{1}H, ^{13}C, IR spectroscopic techniques and mass spectrometry.

(S)-1-Phenyltetrazolyl-5-amino alanine (3.072, ^{PhTzA}Ala): Using the general procedure and starting from 0.16 mmol of ^{PhTU}***Ala (3.054)***, the desired compound **3.072** was obtained (R_f = 0.4 in 1:2, hexane/ethyl acetate). A white solid compound; 40mg, Yield 63 %;



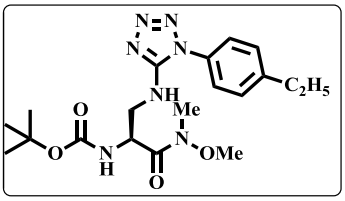
m.p. 130-133 ^{o} C. IR (KBr) $\bar{\upsilon}$ 3379, 2975, 2926, 2854, 1709, 1658, 1609, 1509, 1459, 1390, 1366, 1321, 1302, 1251, 1165, 1088, 1063, 1020, 999 cm^{-1}.^{1}H NMR (600 MHz, $CDCl_3$) δ 7.57–7.49 (5H, m), 5.68 (1H, d, J = 5.8 Hz), 5.38 (1H, s), 4.90 (1H, s), 3.92 (1H, d, J = 10.7 Hz), 3.80 (3H, s), 3.62 (1H, d, J = 4.8 Hz), 3.20 (3H, s), 1.36 (9H, s); ^{13}C NMR (150 MHz; $CDCl_3$) δ 169.8, 156.3, 154.7, 133.4, 130.3, 129.8, 124.1, 80.6 , 62.1, 51.0, 47.7, 32.5, 28.4. +ESI-HRMS (m/z) calculated for $C_{17}H_{26}N_7O_4$ $(M+H)^+$ 392.2046, found 392.2049.

(S)-1-(p-methylphenyltetrazolyl)-5-amino alanine (3.073, ^{MeBTzA}Ala): Using the general procedure and starting from 0.34 mmol of ^{TolTU}***Ala (3.055)***, the desired compound **3.073** was obtained (R_f = 0.4 in 1:2, hexane/ethyl acetate). A solid compound; 105mg,Yield 78 %;



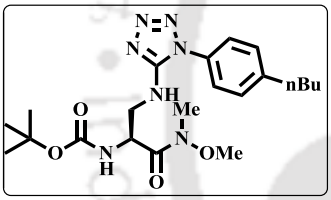
m.p. 102- 106 ^{o} C. IR (KBr) $\bar{\upsilon}$ 3417, 3050, 2956, 2925, 2853, 1705, 1657, 1608, 1519, 1459, 1391, 1367, 1302, 1264, 1166, 1088, 1049, 1018, 990 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 7.36 (2H, d, J = 8.2 Hz), 7.33 (2H, d, J = 8.2 Hz), 5.68 (1H, d, J = 7.6 Hz), 5.32 (1H, s), 4.89 (1H, d, J = 2.1 Hz), 3.88 – 3.86 (1H, m), 3.78 (3H, s), 3.60 (1H, s), 3.17 (3H, s), 2.41 (3H, s), 1.35 (9H, s); ^{13}C NMR (150 MHz; $CDCl_3$) δ 169.9, 156.2, 154.7, 140.2, 130.8, 130.7, 124.0, 80.5, 62.0, 50.9, 47.5, 32.4, 28.3, 21.4.+ESI-HRMS (m/z) calculated for $C_{18}H_{28}N_7O_4$ $[M+H]^+$ 406.2203, found 406.2203.

(S)-1-(p-ethylphenyltetrazolyl)-5-amino alanine (3.074, ^{EtBTzA}Ala): Using the general procedure and starting from 0.31 mmol of ^{EtBTU}***Ala (3.056)***, the desired compound **3.074** was obtained (R_f = 0.4 in 1:2, hexane/ethyl acetate). A semi solid compound; 105mg,



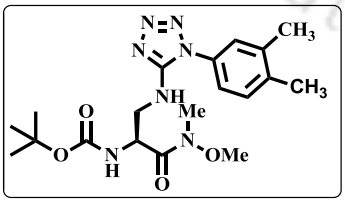
Yield 80 %. IR (KBr) ῡ 3347, 2969, 2926, 2854, 1710, 1658, 1604, 1519, 1458, 1431, 1390, 1366, 1296, 1251, 1166, 1094, 1018, 1056, 1018, 1056, 990 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 7.39 (2H, d, J = 8.2 Hz), 7.35 (2H, d, J = 8.1 Hz), 5.69 (1H, d, J = 7.4 Hz), 5.34 (1H, s), 4.89 (1H, s), 3.89 – 3.87 (1H, m), 3.78 (3H, s), 3.59 (1H, s), 3.17 (3H, s), 2.70 (2H, q, J = 7.6 Hz), 1.34 (9H, s), 1.24 (3H, dd, J = 13.8, 6.1 Hz,).^{13}C NMR (150 MHz; $CDCl_3$) δ 169.8, 156.2, 154.7, 146.3, 130.8, 129.6, 124.1, 80.4, 62.0, 51.0, 47.5, 32.4, 28.7, 28.3, 15.5. +ESI-HRMS (m/z) calculated for $C_{19}H_{30}N_7O_4$ $[M+H]^+$ 420.2359, found 420.2358.

(S)-1-(p-nButylphenyl tetrazolyl)-5-amino alanine (3.075, ^{BuBTzA}Ala): Using the general procedure and starting from 0.32 mmol of ^{BuBTU}***Ala (3.057)***, the desired compound **3.075** was obtained (R_f = 0.4 in 1:2, hexane/ethyl acetate). A semi solid compound;



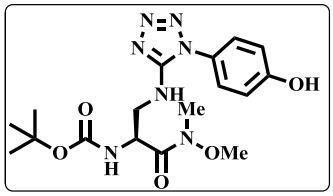
122mg,Yield 85 %. IR (KBr) ῡ 3298, 3052, 2977, 2952, 2928, 2857, 1690, 1653, 1604, 1578, 1538, 1519, 1458, 1432, 1394, 1365, 1290, 1255, 1172, 1091, 1050, 1026, 992 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 7.36 (2H, d, J = 7.9 Hz), 7.30 (2H, d, J = 7.8 Hz), 5.72 (1H, d, J = 6.7 Hz), 5.36 (1H, s), 4.87 (1H, s), 3.85 (1H, s), 3.76 (3H, s), 3.58 (1H, s), 3.14 (3H, s), 2.63 (2H, t, J = 7.8 Hz), 1.58 (2H, dt, J = 15.3, 7.7 Hz), 1.35 – 1.32 (11H, m), 0.90 (3H, t, J = 7.4 Hz); ^{13}C NMR (150 MHz; $CDCl_3$) δ 169.8, 156.2, 154.7, 145.0, 130.8, 130.1, 123.9, 80.3, 61.9, 50.9, 47.4, 35.3, 33.4, 32.4, 29.7, 28.3, 22.3, 14.0. +ESI-HRMS (m/z) calculated for $C_{21}H_{34}N_7O_4$ $[M+H]^+$ 448.2672, found 448.2670.

(S)-1-(m,p-Dimethyphenyl tetrazolyl)-5-amino alanine (3.076, ^{DMBTzA}Ala): Using the general procedure and starting from 0.28 mmol of ^{DMBTU}***Ala (3.058)***, the desired compound **3.076** was obtained (R_f = 0.4 in 1:1, hexane/ethyl acetate). A semi solid compound; 90mg, Yield

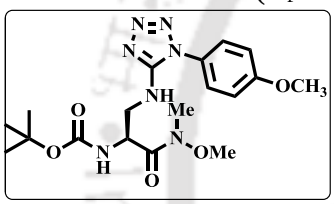


76 %. IR (KBr) ῡ 3354, 2975, 2925, 2854, 1710, 1658, 1604, 1510, 1456.99, 1390.56, 1366.94, 1302, 1277, 1251, 1166, 1094, 1054, 1022, 992 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 7.26 (2H, t, J = 8.8 Hz), 7.18 (1H, dd, J = 7.8, 1.5 Hz), 5.68 (1H, d, J = 7.1 Hz), 5.34 (1H, s), 4.88–4.87 (1H, m), 3.88–3.86 (1H, m), 3.78 (3H, s), 3.61–3.60 (1H, m), 3.17 (3H, s), 2.31 (3H, s), 2.30 (3H, s), 1.34 (9H, s); ^{13}C NMR (150 MHz, $CDCl_3$) δ 169.9, 156.2, 154.7, 139.1 ,138.8, 131.14, 130.8, 125.2 , 121.2, 80.4, 62.0, 50.9, 47.4, 32.4 , 28.3, 19.9, 19.7 +ESI-HRMS (m/z) calculated for $C_{19}H_{30}N_7O_4$ $[M+H]^+$ 420.2359, found 420.2360.

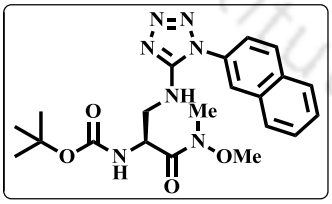
(S)-1-(p-Hydroxyphenyltetrazolyl)-5-amino alanine (3.077, ^{HBTzA}Ala): Using the general procedure and starting from 0.34 mmol of ^{HBTU}***Ala (3.064)***, the desired compound **3.077** was obtained (R_f = 0.4 in 1:2, hexane/ethyl acetate). A solid compound;100mg, Yield 72 %; m.p. 170-174 ^{o}C. IR (KBr) ῡ 3353, 3270, 2967, 2924, 2853, 1698, 1674, 1659, 1617, 1517, 1458, 1390, 1367, 1315, 1279, 1228, 1168, 1105, 1063, 1019, 967 cm^{-1}. ^{1}H NMR (600MHz; $CDCl_3$+DMSO-d_6) δ 9.42 (1H, s), 7.17 (2H, d, *J* = 8.5 Hz), 6.88 (2H, d, *J* = 8.5 Hz), 5.91 (1H, d, *J* = 7.9 Hz), 5.55 (1H, s), 4.78 (1H, s), 3.71 (3H, s), 3.62 (2H, dd, *J* = 24.3, 18.1 Hz), 3.06 (3H, s), 1.28 (9H, s); ^{13}C NMR (150MHz; $CDCl_3$+$DMSOd_6$) δ 170.3, 158.7, 156.1, 155.0, 126.2, 124.4, 116.7, 80.4, 61.9, 50.8, 46.2, 32.3, 28.2. +ESI-HRMS (m/z) calculated for $C_{17}H_{26}N_7O_5$ $[M+H]^+$ 408.1995, found 408.1998.



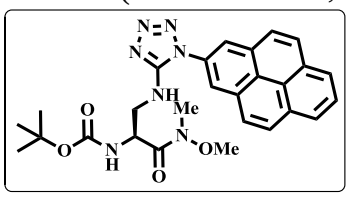
(S)-1-(p-Methoxyphenyl tetrazolyl)-5-amino alanine (3.078, ^{MOBTzA}Ala): Using the general procedure and starting from 0.26 mmol of ^{MOBTU}***Ala (3.059)***, the desired compound **3.078** was obtained (R_f = 0.4 in 1:2, hexane/ethyl acetate). A semi solid compound; 67mg, Yield 61 %. IR (KBr) ῡ 3381, 3006, 2975, 2925, 2853, 1709, 1658, 1606, 1582, 1519, 1461, 1435, 1391, 1367, 1303, 1254, 1171, 1093, 1024, 1043, 985 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 7.39 (2H ,d, *J* = 8.5 Hz), 7.02 (2H, d, *J* = 8.6 Hz), 5.69 (1H, d, *J* = 6.7 Hz), 5.24 (1H, s), 4.88 (1H, s), 3.84 (4H, s), 3.78 (3H, s), 3.60 (1H, s), 3.17 (3H, s), 1.35 (9H, s); ^{13}C NMR (150 MHz; $CDCl_3$) δ 169.8, 160.6, 156.2, 154.9, 125.9, 115.3, 80.5, 62.0, 55.8, 51.0, 47.4, 32.4, 28.3. +ESI-HRMS (m/z) calculated for $C_{18}H_{28}N_7O_5$ $[M+H]^+$ 422.2153, found 422.2153.



(S)-1-Napthyl tetrazolyl-5-amino alanine (3.079, ^{NapTzA}Ala): Using the general procedure and starting from 0.35 mmol of ^{NapTU}***Ala (3.062)***, the desired compound **3.079** was obtained (R_f = 0.4 in 1:2, hexane/ethyl acetate). A solid compound; 93 mg ,Yield 60 %; m.p. 190- 195 ^{o}C. IR (KBr) ῡ 3311, 3056, 3000, 2975, 2924, 2853, 1711, 1640, 1548, 1548, 1490, 1425, 1394, 1366, 1276, 1301, 1230, 1166, 1104, 1075, 1047, 1020, 948 cm^{-1}.^{1}H NMR (600 MHz; $CDCl_3$) δ 8.52 (1H, d, *J* = 7.8 Hz), 7.85 (1H, d, *J* = 8.1 Hz), 7.62 (1H, d, *J* = 7.5 Hz), 7.54 (2H , dd, *J* = 14.3, 7.9 Hz), 7.47 (1H ,t, *J* = 7.2 Hz), 6.23 (1H ,s), 6.09 (1H, s), 4.99 (1H, s), 3.92 (2H, t, *J* = 18.8 Hz), 3.79 (3H, s), 3.12 (3H, s), 1.42 (9H, s); ^{13}C NMR (150 MHz, $CDCl_3$) δ 170.53, 168.0, 155.9, 148.0, 132.3, 128.1, 127.0, 126.0 , 125.6, 124.0, 122.1, 118.8, 80.2, 61.9, 51.2, 47.4 , 32.3 , 28.5.

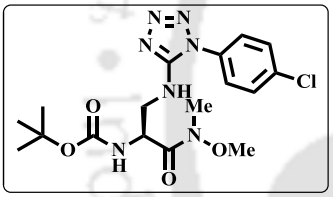


(S)-1-Pyrene tetrazolyl-5-amino alanine (3.080, ^{PyTzA}Ala): Using the general procedure and starting from 0.30 mmol of ***^{PyTU}Ala (3.063)***, the desired compound **3.080** was obtained (Rf = 0.4 in 1:2, hexane/ethyl acetate). A semi solid compound; 126mg, Yield 81 %.



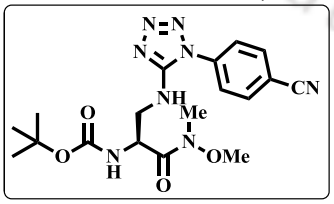
IR (KBr) ῡ 3417, 2953, 2853, 1707, 1644, 1607, 1462, 1390, 1366, 1166, 1086, 1020, 849 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 8.31 – 8.23 (3H, m), 8.19 (2H, dd, *J* = 11.0, 4.8 Hz), 8.14 – 8.05 (2H, m), 8.00 (1H, d, *J* = 7.9 Hz), 7.70 (1H, d, *J* = 9.1 Hz), 5.56 (1H, d, *J* = 6.4 Hz), 5.00 (1H, s), 4.83 (1H, s), 3.82 (1H, s), 3.75 (3H, s), 3.71 (1H, s), 3.13 (3H, s), 1.15 (9H, s); ^{13}C NMR (150 MHz; $CDCl_3$) δ 170.0, 156.3, 155.9, 133.1, 131.1, 130.7 , 130.3 , 129.6 , 127.4 , 127.0, 126.8, 126.6, 125.4 , 125.1, 124.6, 124.2, 121.0, 80.3, 62.0, 50.9, 46.9, 32.5, 28.2. +ESI-HRMS (m/z) calculated for $C_{27}H_{30}N_7O_4$ $(M+H)^+$ 516.2359, found 516.2350.

(S)-1-(p-Chlorophenyl tetrazolyl)-5-amino alanine (3.081, ^{ClBTzA}Ala): Using the general procedure and starting from 0.32 mmol of ***^{ClBTU}Ala (3.065)***, the desired compound **3.089** was obtained (R_f = 0.4 in 1:2, hexane/ethyl acetate). A semi solid compound; 100mg, Yield 74 %.



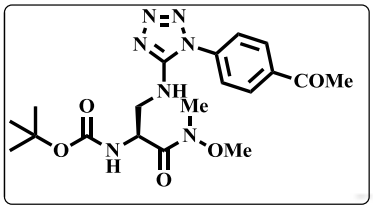
IR (KBr) ῡ 3334, 2978, 2926, 2853, 1709, 1659, 1608, 1499, 1391, 1366, 1251, 1166, 1092, 1013, 990 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 7.51 (2H, d, *J* = 8.4 Hz), 7.47 (2H, d, *J* = 8.4 Hz), 5.75 (1H, d, *J* = 6.3 Hz), 5.55 (1H, s), 4.87 (1H, s), 3.90–3.88 (1H, m), 3.78 (3H, s), 3.59–3.58 (1H, m), 3.17 (3H, s), 1.34 (9H, s); ^{13}C NMR (150 MHz; $CDCl_3$) δ 169.7, 156.4, 154.4, 135.7, 131.9, 130.4, 125.3, 80.6, 62.0, 51.0, 47.8, 32.4, 28.3. +ESI-HRMS (m/z) calculated for $C_{17}H_{25}ClN_7O_4$ $(M+H)^+$ 426.1657, found 426.1657.

(S)-1-(p-Cyano phenyl tetrazolyl)-5-amino alanine (3.083, ^{CNBTzA}Ala): Using the general procedure and starting from 0.21 mmol of ***^{CNBTU}Ala (3.066)***, the desired compound **3.082** was obtained (R_f = 0.4 in 1:2, hexane/ethyl acetate). A semi solid compound; 54mg, Yield



62 %. IR (KBr) ῡ 3378, 2961, 2925, 2854, 2232, 1708, 1659, 1603, 1568, 1514, 1459, 1426, 1392, 1366, 1280, 1251, 1166, 1094, 1018, 990 cm^{-1}.^{1}H NMR (600 MHz; $CDCl_3$) δ 7.85 (2H, d, *J* = 8.2 Hz), 7.74 (2H, d, *J* = 8.2 Hz), 5.97 (1H, s), 5.82 (1H, s), 4.88 (1H, s), 3.94 (1H, s), 3.79 (3H, s), 3.58 (1H, s), 3.20 (3H, s), 1.34 (9H, s); ^{13}C NMR (150 MHz; $CDCl_3$) δ 169.4, 156.7, 154.3, 137.1, 134.2, 123.9, 117.6, 113.2, 80.8, 62.0, 51.0, 48.2, 32.5, 28.3. +ESI-HRMS (m/z) calculated for $C_{18}H_{25}N_8O_4$ $[M+H]^+$ 417.1999, found 417.1995.

(S)-1-(p-Acetylphenyl tetrazolyl)-5-amino alanine (3.082, ^{AcBTzA}Ala)*:* Using the general procedure and starting from 0.35 mmol of ^{AcBTU}***Ala (3.067)***, the desired compound **3.083** was obtained (R_f = 0.4 in 1:2, hexane/ethyl acetate). A greenish white semi solid



compound; 45mg, Yield 45 %. IR (KBr) ῡ 3377, 2961, 2925, 2854, 1687, 1657, 1602, 1574, 1515, 1463, 1426, 1392, 1364, 1266, 1166, 1094, 1017, 985 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 8.15 (2H, d, *J* = 8.5 Hz), 7.68 (2H, d, *J* = 8.2 Hz), 5.73 (1H, d, *J* = 5.8 Hz), 5.67 (1H, s), 4.90 (1H, s), 3.97–3.96 (1H, m), 3.81 (3H, s), 3.62 (1H, d, *J* = 6.3 Hz), 3.22 (3H, s), 2.66 (3H, s), 1.36 (9H, s); ^{13}C NMR (150 MHz; $CDCl_3$) δ 196.6, 169.5, 156.6, 154.5, 130.4, 123.4, 80.8, 62.1, 51.0, 48.2, 32.5, 28.4, 26.9.+ESI-HRMS (m/z) calculated for $C_{19}H_{28}N_7O_5$ $[M+H]^+$ 434.2152, found 434.2154.

3.8.3. Photophysical Studies of the Amino Acids

UV-visible measurements: All the UV–visible spectra of the amino acids (10 μM) were measured in different solvents using a UV-visible spectrophotometer with a cell of 1 cm path length. The measurements were carried out in absorbance mode. The absorbance values of the sample solutions were measured in the wavelength regime of 200–600 nm. All the sample solutions were prepared just before doing the experiment.

Fluorescence experiments: All the sample solutions were prepared as described in UV-visible measurement experiments. Fluorescence spectra were obtained using a fluorescence spectrophotometer at 25 °C using 1 cm path length cell. The excitation wavelengths for all the cases were set at the absorbance maxima of each sample in each solvents and emission spectra were measured in the wavelength region of 352–600 nm with an integration time of 0.2 sec and with slit width 3 nm. All the sample solutions were prepared just before doing the experiment. Total volume of 1.0 ml from a stock solution of 3 ml of 10 μM concentration for each case was used for fluorescence experiment in 1 ml cell. Fluorescence emissions were collected exciting the samples at the wave length corresponding to their absorption maxima.

3.8.4. ^{1}H and ^{13}C NMR Spectra Few Selected Amino Acids

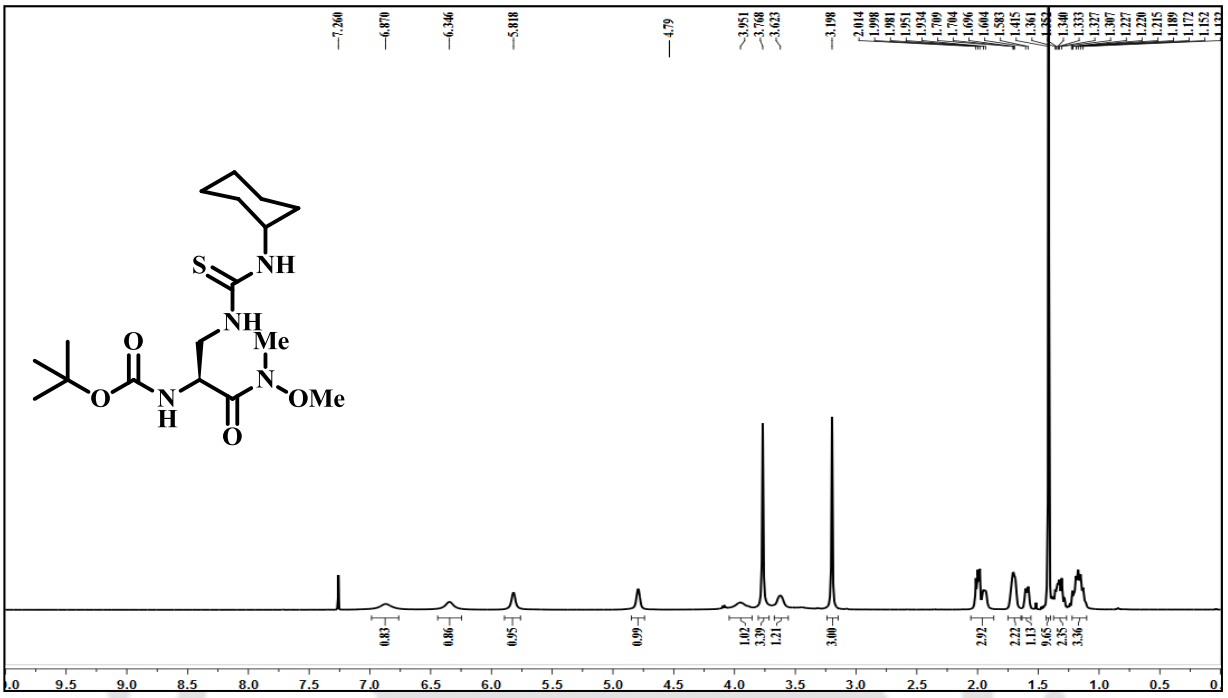


Figure 3.20. ^{1}H NMR spectra of synthesised compound **3.071.**

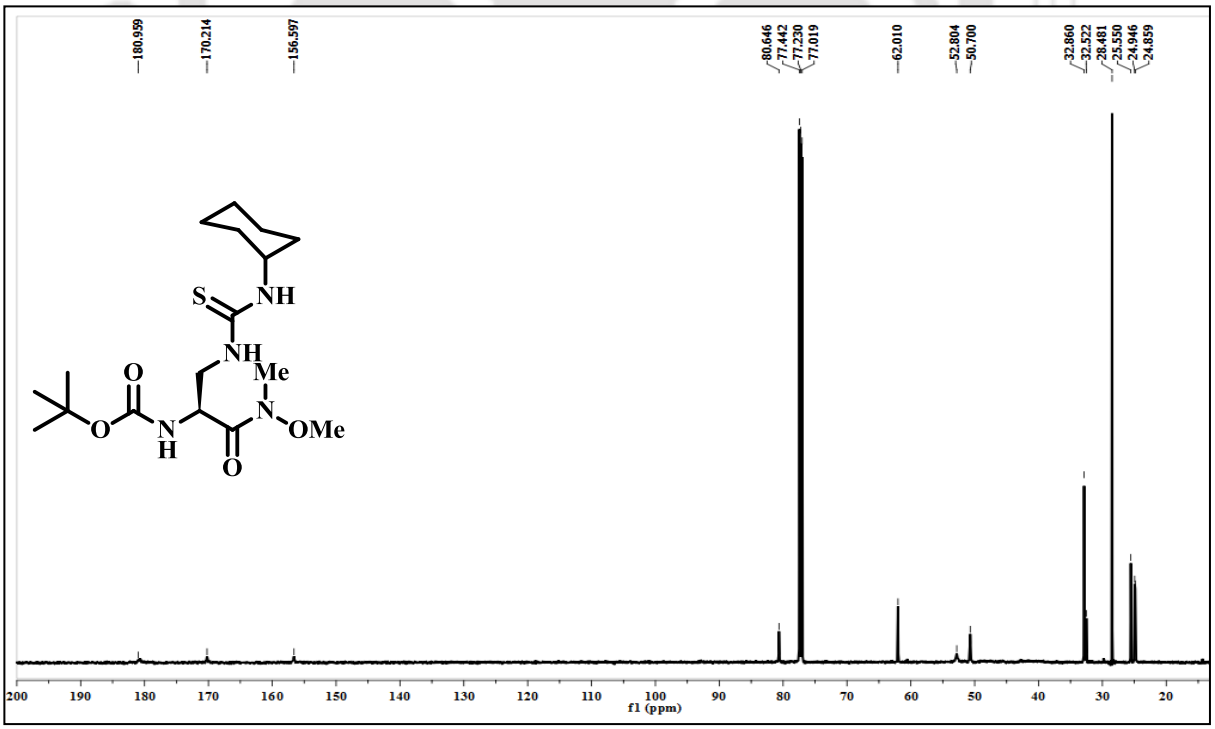


Figure 3.21. ^{13}C NMR spectra of synthesised compound **3.071.**

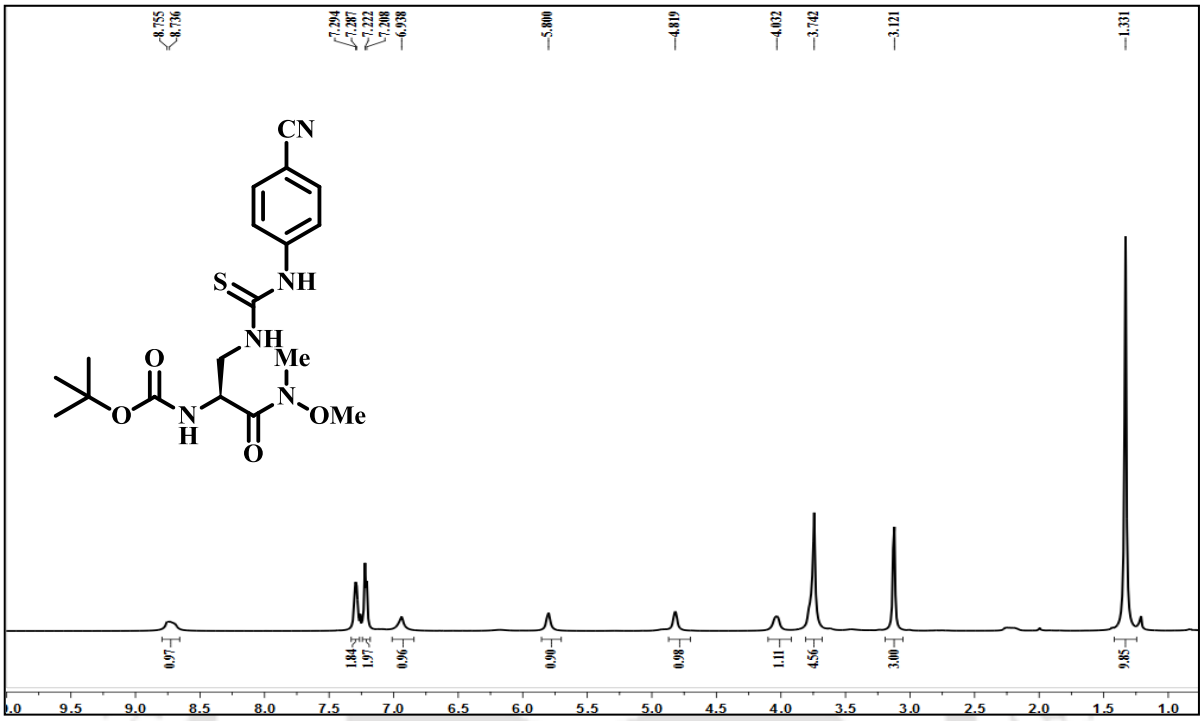


Figure 3.22. ^{1}H NMR spectra of synthesised compound **3.066.**

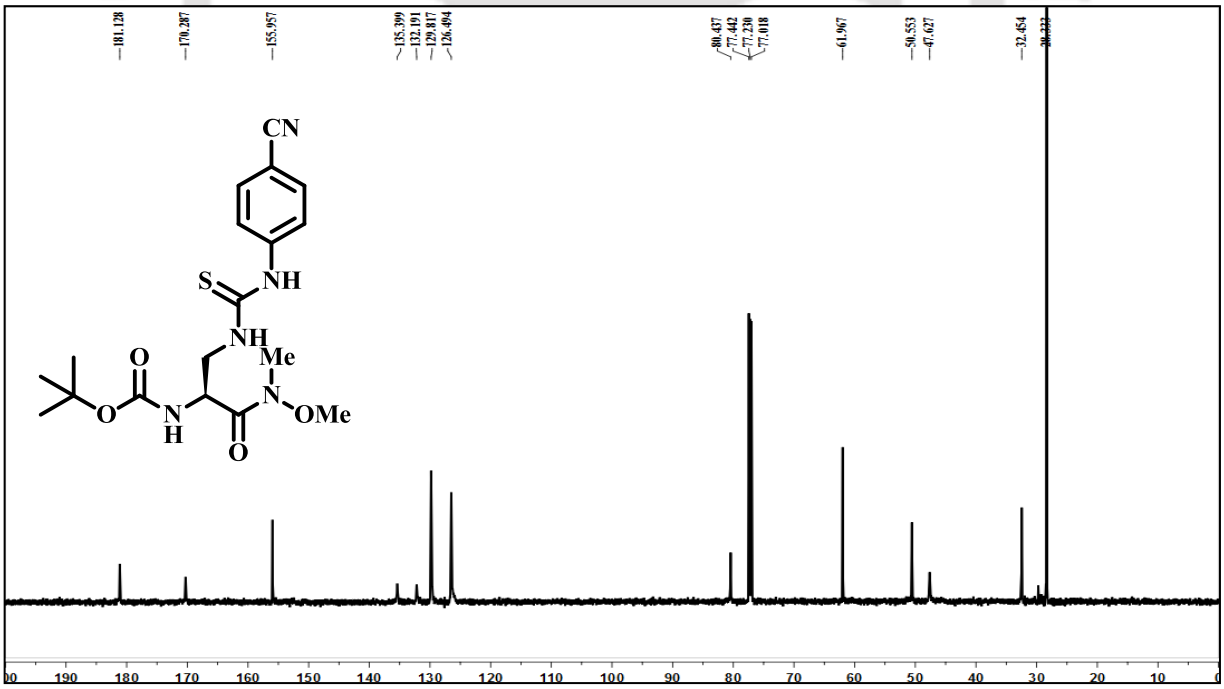


Figure 3.23. ^{13}C NMR spectra of synthesised compound **3.066.**

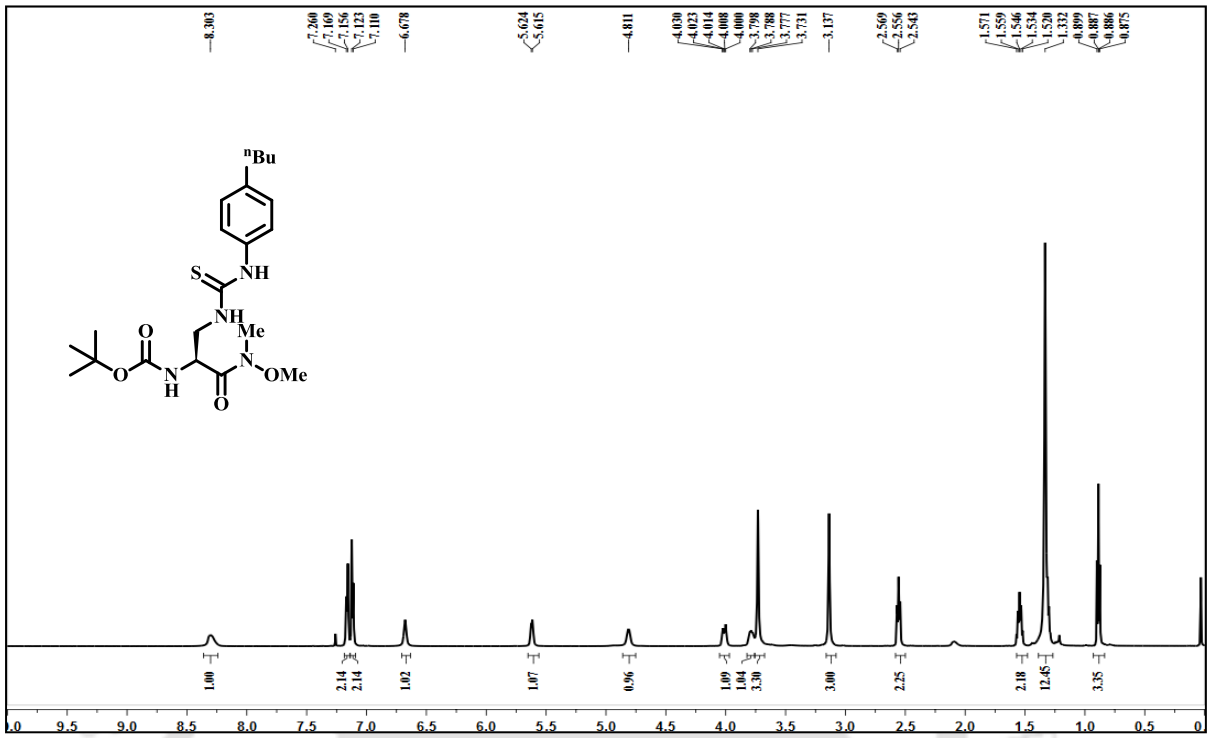


Figure 3.24. ^{1}H NMR spectra of synthesised compound **3.057**

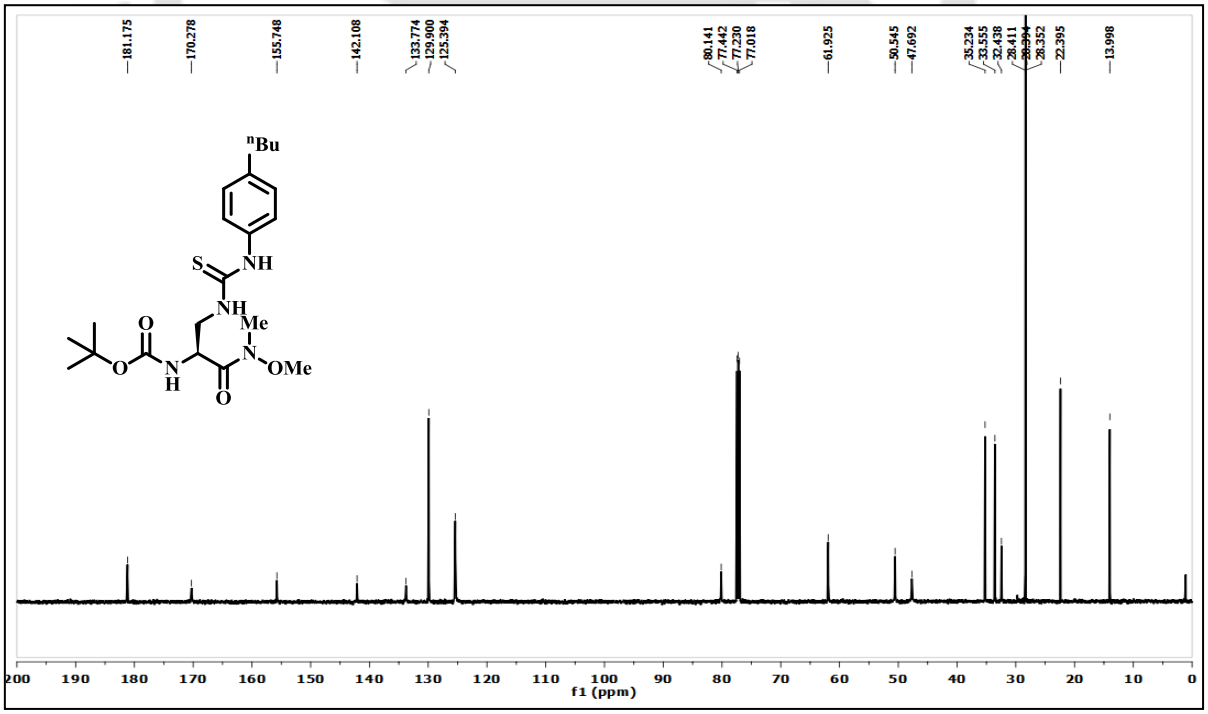


Figure 3.25. ^{13}C NMR spectra of synthesised compound **3.057**

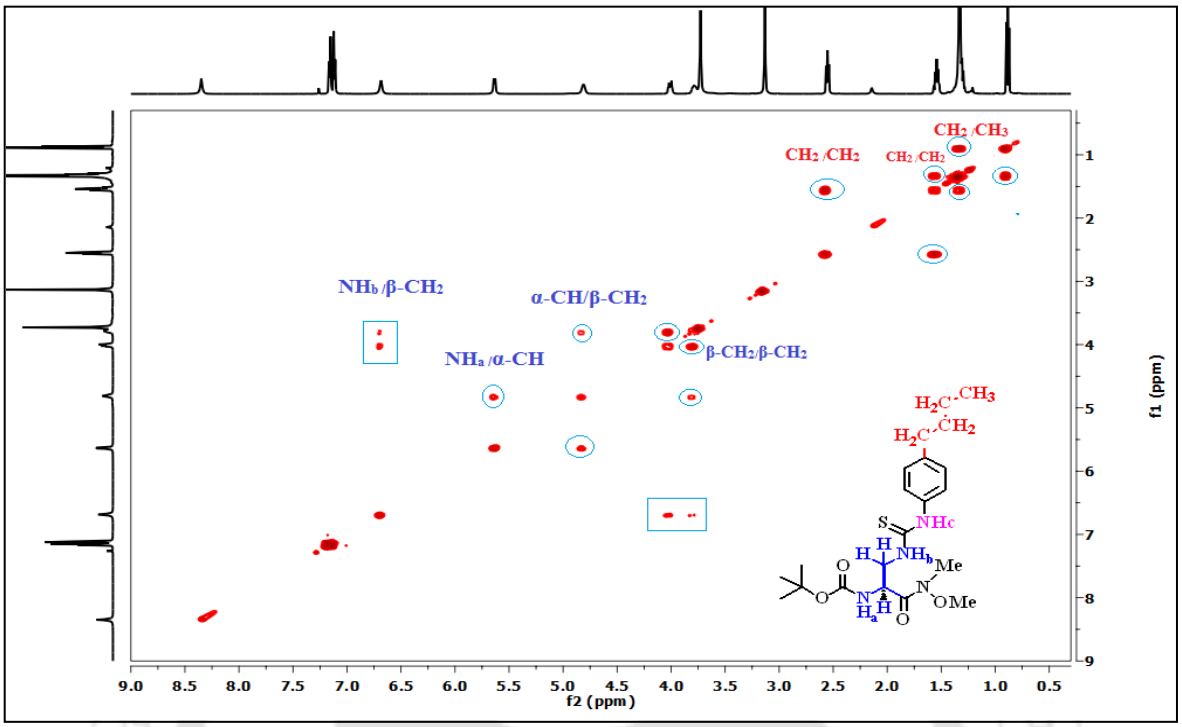


Figure 3.26. COSY (^{1}H-^{1}H) NMR spectra of synthesised compound **3.057**

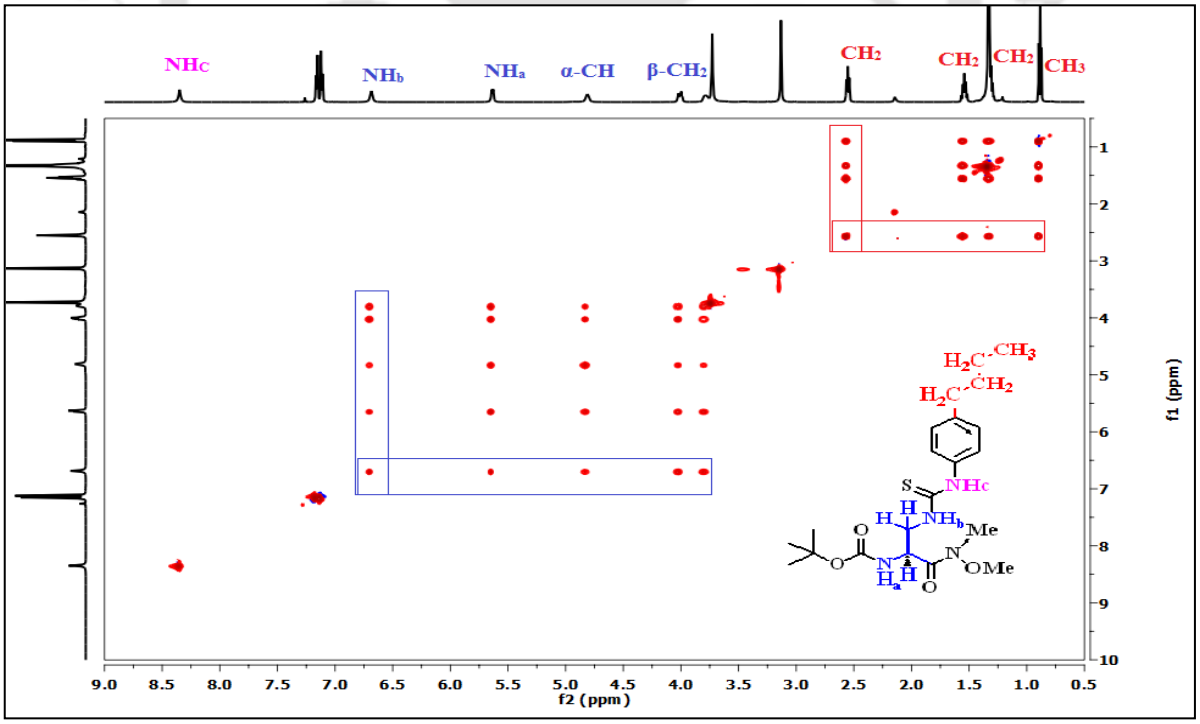


Figure 3.27. TOCSY (^{1}H-^{1}H) NMR spectra of synthesised compound **3.057.**

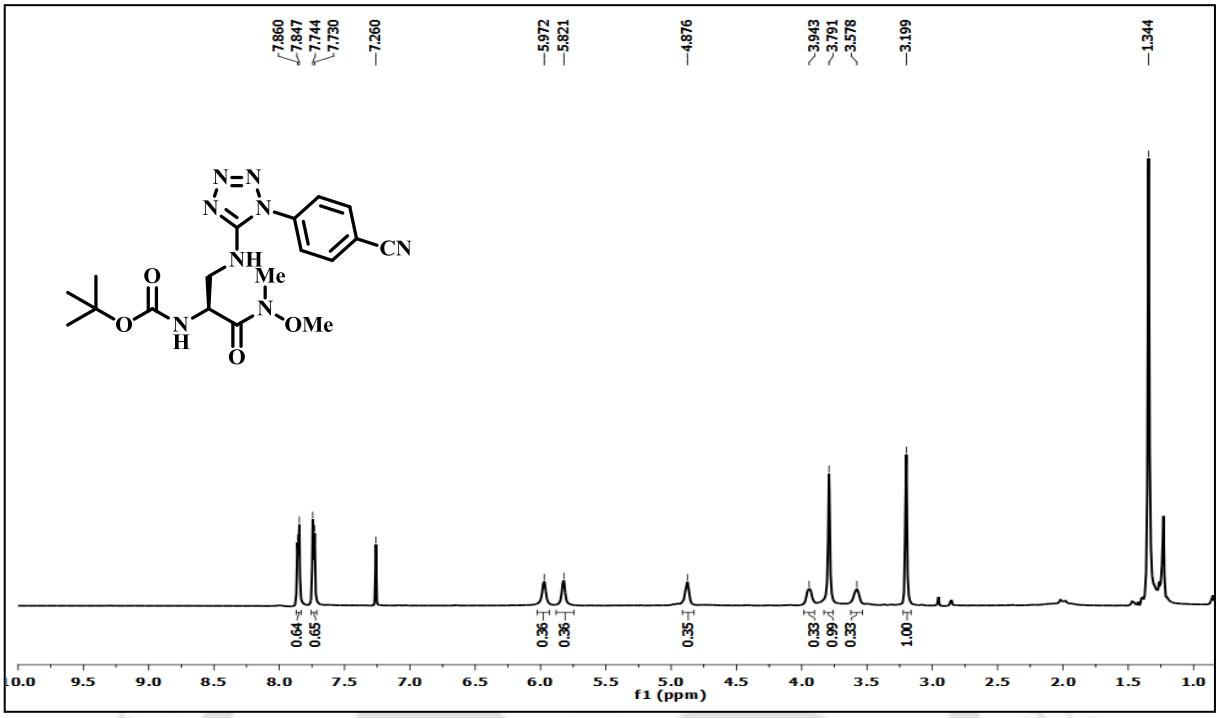


Figure 3.28. ^{1}H NMR spectra of synthesised compound **3.083.**

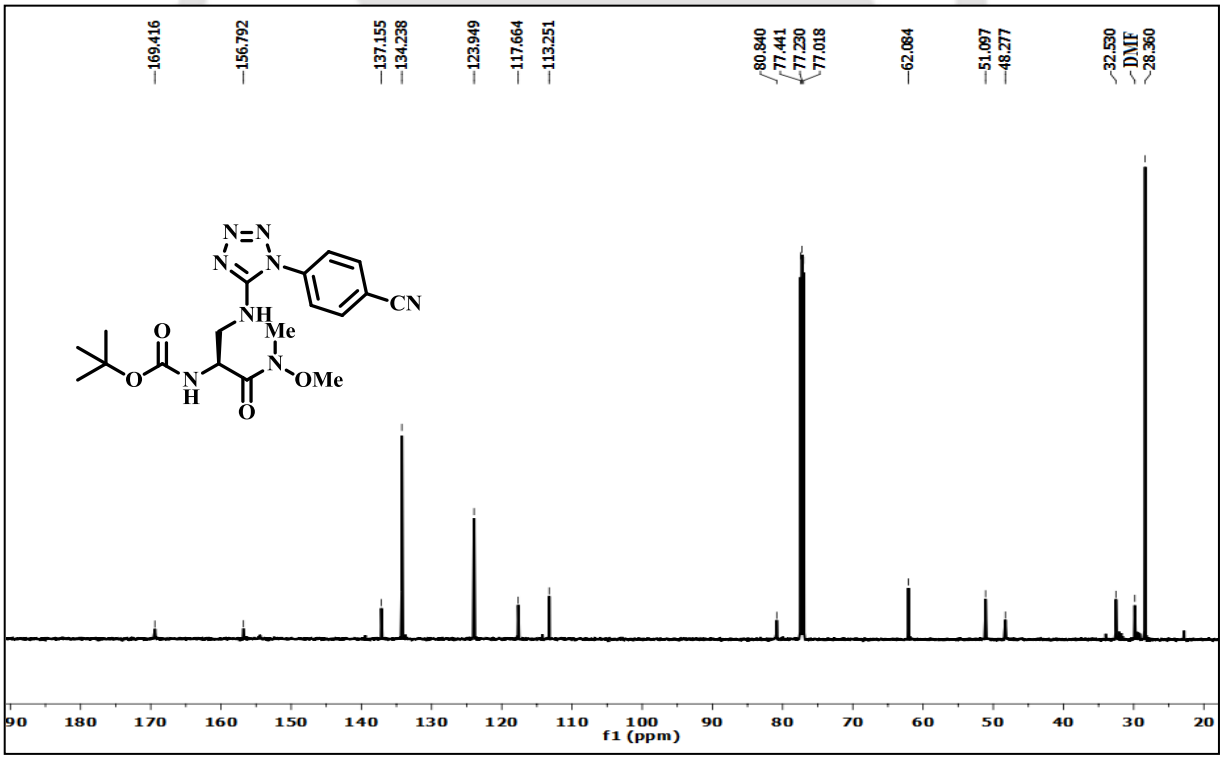


Figure 3.29. ^{13}C NMR spectra of synthesised compound **3.083.**

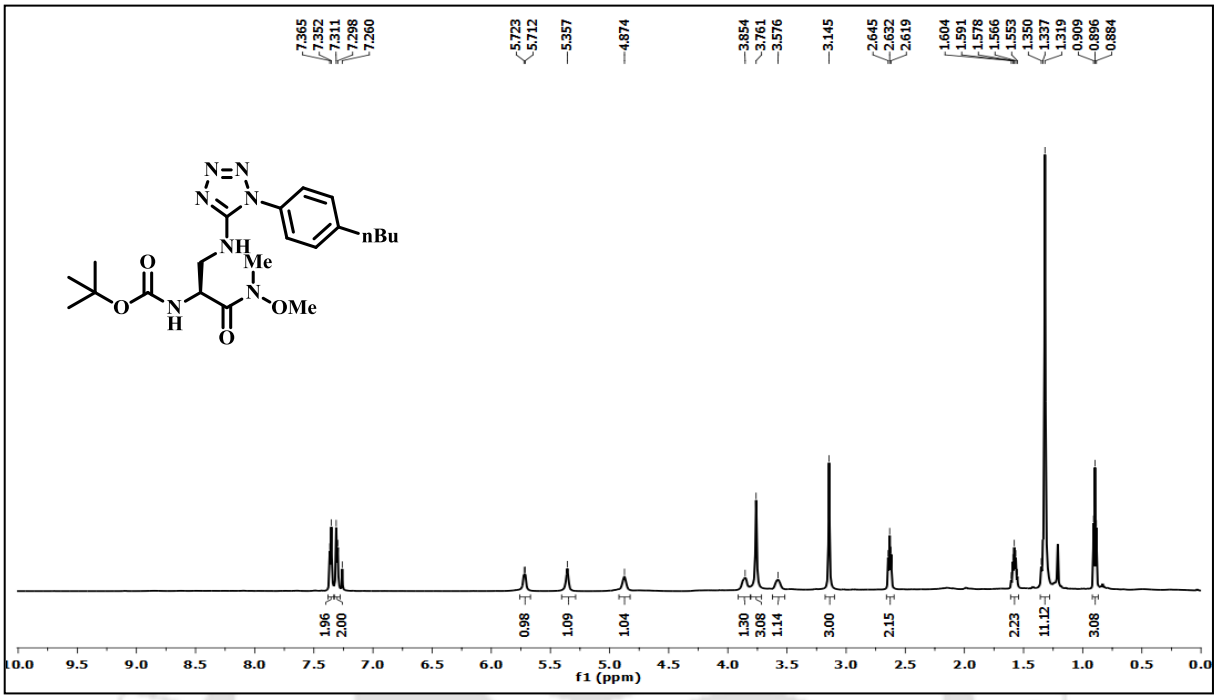


Figure 3.30. ^{1}H NMR spectra of synthesised compound **3.075**

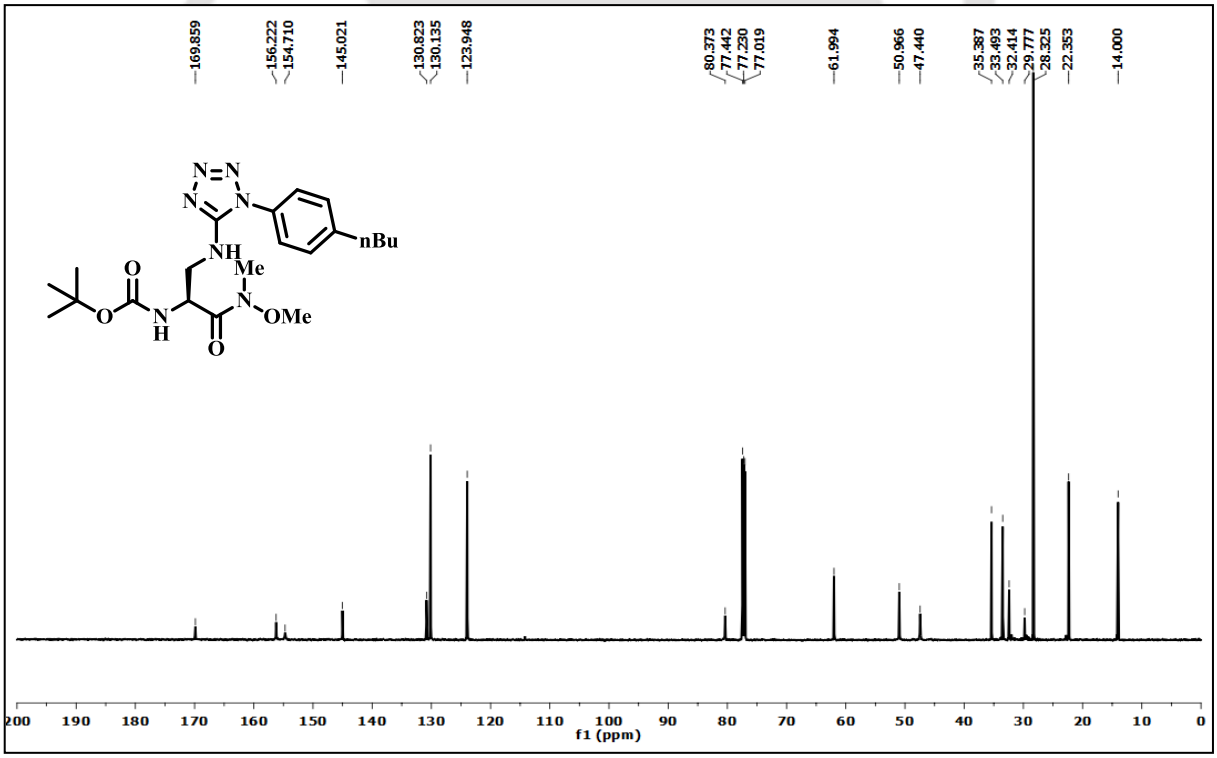


Figure 3.31. ^{13}C NMR spectra of synthesised compound **3.075.**

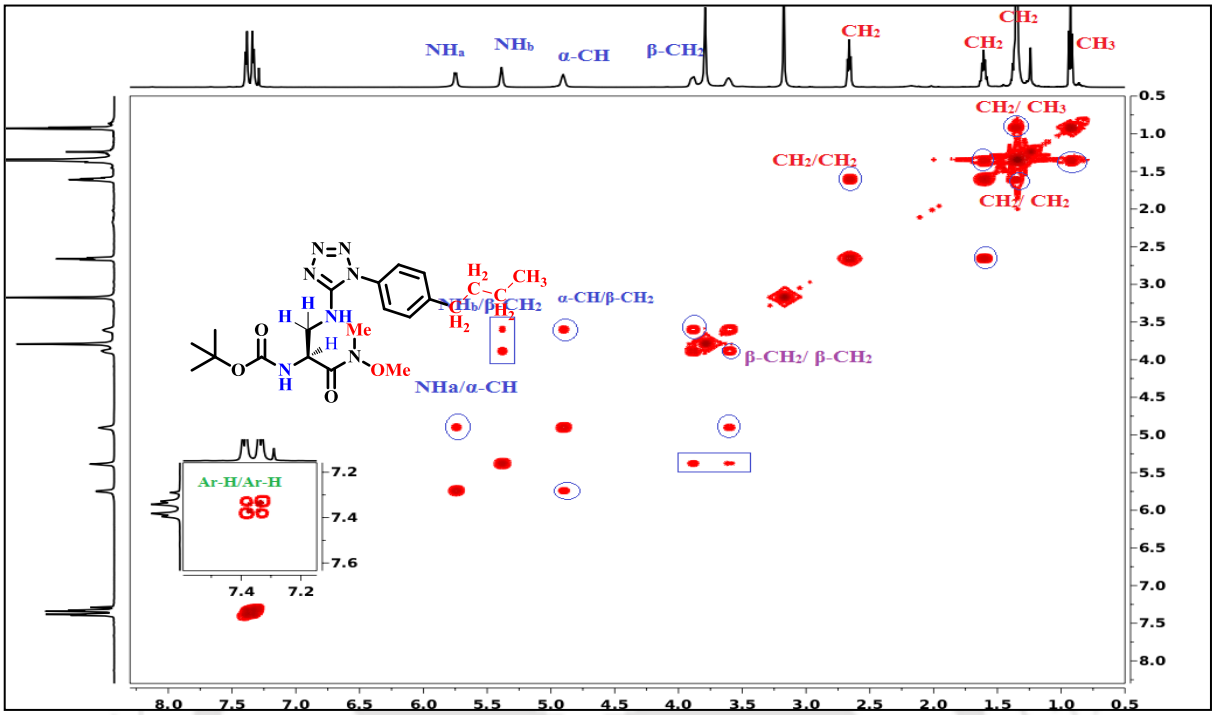


Figure 3.32. COSY (^{1}H-^{1}H) NMR spectra of synthesised compound **3.075**

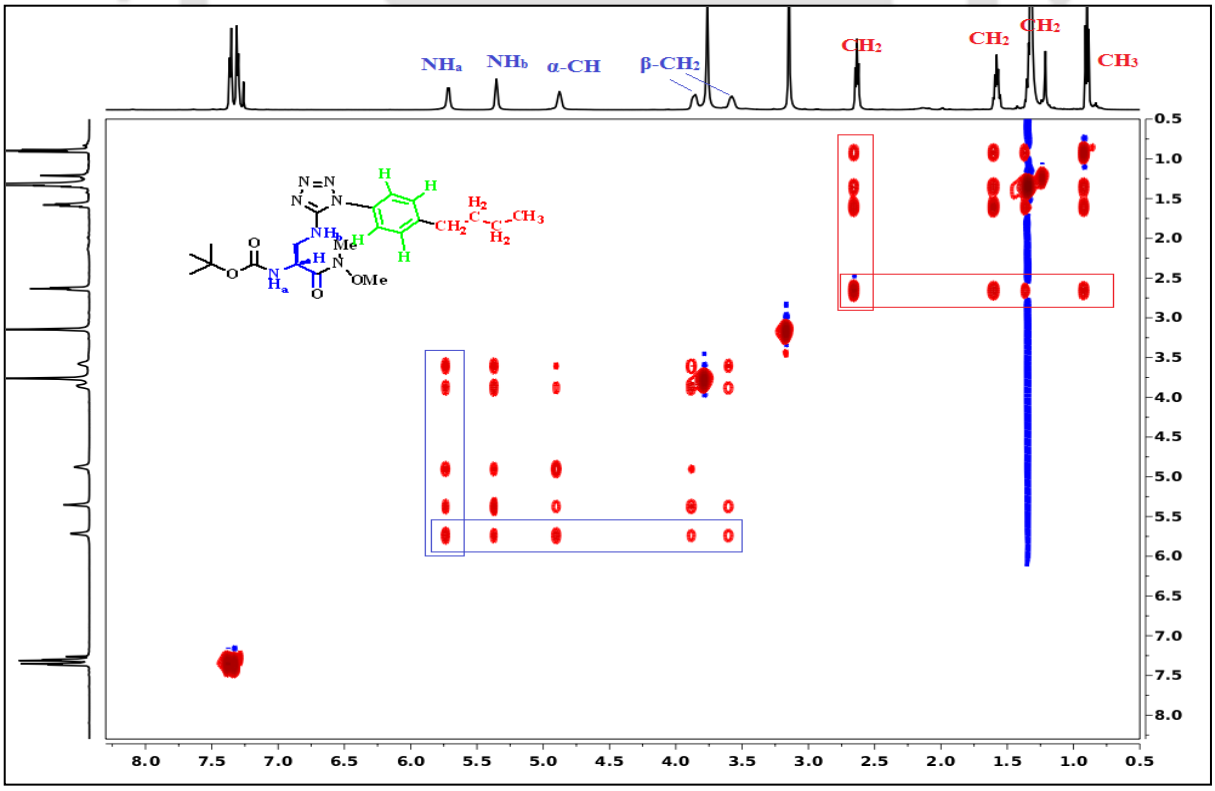


Figure 3.33. TOCSY (^{1}H-^{1}H) NMR spectra of synthesised compound **3.075**

3.9. References

1. (a) Mohite, A.R.; Bhat, R.G. *J. Org. Chem.* **2012**, *77*, 5423. (b) Ghasemia, E.; Shahvelayatib,A.S.; Yavaria, I. *Phosphorus, Sulfur, And Silicon.* **2016**, *191*,746. (c) Harris, K.D.M. *J. Chin. Chem. Soc.* **1999,** *46*,5. (d) Batey, R.A.; Powell, D.A. *Org. Lett*. **2000**, *2*, 3237. (d) Schroeder, D. C. *Chem. Rev.* **1955**, *55*, 181. (e) Struga, M.; Kossakowski, J.;Kedzierska, E.; Fidecka, S.; Stefańska, *J. Chem. Pharm. Bull.* **2007**, *55*, 796. (f) Chayah, M.; Camacho, M.E.; Carrión, M.D.; Gallo, M.; Romero,M.; Duarte, J. *Med. Chem. Commun.* **2016**, *7*, 667.
2. Walpole, C.; Ko, S. Y.; Brown, M.; Beattie, D.; Campbell, E.; Dickenson, F.; Ewan, S.; Hughes, G. A.; Lemaira, M.; Lerpiniere, J.; Patel,S.; Urban, L. *J. Med. Chem*. **1998**, *41*, 3159.
3. (a) Schroeder, D.C. *Chem. Rev*. **1955**, 55, 181. (b) Sarkis, G.Y.; Faisal, E.D. *J. Heterocycl. Chem.* **1985**, *22*, 137.
4. Makhsumov, A. G.; Safaev, A.S.; Abidova, S. V. Katal Pererab. Uglevodordn. Syrya **1968**, *2*, 101; *Chem. Abstr*. **1969**, *71*, 101668.
5. (a) Levallet, C.; Lerpiniere, J.; Ko, S. Y. *Tetrahedron.* **1997**, *53*, 5291. (b) Kasmi, S.; Hamelin, J.; Benhaoua, H. *Tetrahedron Lett*. **1998**, *39*, 8093. (c) Kearney, P. C.; Fernandez, M.; Flygare, J.A. *J. Org. Chem*. **1998**, *63*, 196. (d) Boga, C.; Forlani, L.; Silvestroni, C.; Corradi, A.B.; Sgarabotto, P. J. *Chem. Soc., Perkin Trans*. **1999**,1, 1363. (e) Kidwai, M.; Venkataramanan, R.; Dave, B. *Green Chem*. **2001**, *3*, 278. (f) Du, W.; Curran, D.P. *Org. Lett*. **2003**, *5*, 1765. (g) Liu, W.-X.; Jiang, Y.-B. *J. Org. Chem*. **2008**, *73*, 1124.(h) Hargrove, A.E.; Nieto, S.; Zhang, T.; Sessler, J.L.; Anslyn, E. V. *Chem.Rev*. **2011**, *111*, 6603. (i) Chen, H.-L.; Guo, Z.-F.; Lu, Z.-l. *Org.Lett*. **2012**, *14*, 5070.
6. (a) Taylor, M. S.; Jacobsen, E. N. *Angew. Chem., Int. Ed.* **2006**, *45*, 1520. (b) Connon, S. *J. Chem.; Eur*. J. **2006,** *12*, 5418. (c) Takemoto, Y. *Org. Biomol. Chem*. **2005**, *3*, 4299. (d) Schreiner, P. R. *Chem. Soc. Rev*. **2003**, *32*, 289. (e) McCooey, S. H.; Connon, S. J. *Angew. Chem., Int. Ed*. **2005**, *44*, 6367. (f) Wang, J.; Li, H.; Yu, X.; Zu, L.; Wang, W. *Org. Lett*. **2005**, *7*, 4293. (g) Hoashi, Y.; Okino, T.; Takemoto, Y. *Angew. Chem., Int. Ed*. **2005**, *44*, 4032.
7. (a)Levchik, S. V.; Balabanovich, A. I.; Ivashkevich, O. A.; Lesnikovich,A. I.; Gaponik, P. N.; Costa, L. *Thermochim. Acta* **1992**, *207*, 115; (b) Levchik, S.V.; Balabanovich, A.I.; Ivashkevich, O.A.; Lesnikovich, A.I.; Gaponik, P.N.; Costa, L. *Thermochim. Acta* **1993**, *255*, 53. (c) Lesnikovich, A. I.; Ivashkevich, O. A.; Levchik, S.V.; Balabanovich, A.I.; Gaponik, P.N.; Kulak, A. A. *Thermochim. Acta.* **2002**, *388*, 233. (d) Gao, A.; Oyumi, Y.; Brill, T. B. *Combust. Flame* **1991**, *83*, 345. (e)Levchik, S.V.; Balabanovich, A.I.; Ivashkevich, O.A.; Gaponik,P. N.; Costa, L. *Polym. Degrad. Stability* **1995**, *47*, 333. (f)Butler RN. Tetrazoles. In Comprehensive Heterocyclic Chemistry II, Katritzky AR, Rees CW, Scriven EFV (eds). Pergamon Press: Oxford, **1996**, 621. (g) Duncia J.V.; Pierce M.E.; Santella J.B. *J. Org. Chem*. **1991**, *56*, 2395. (h) Zabrocki , J.; Smith G.D.; Dunbar J.B.; Iijima H.; Marsall G.R.; *J. Am. Chem. Soc*. **1988**, *110*, 5875.(i) Zabrocki ,

J.; Dunbar J.B.; Marsall K.W.; Toth M. V.; Marsall G.R. *J. Org. Chem.* **1992**, *57*, 202. (j) Valle, G.; Crisma, M.; Yu K.-L.; Toniolo ,C.; Mishra, R.K.; Johnson R. L.; *Chem. Commun*. **1988**, *53*, 2863. (k) May B.C.H.; Abell A.D. *Tetrahedron Lett*. **2001**, *42*, 5641.

8. (a) Singh, R. P.; Verma, R. D.; Meshri, D. T.; Shreeve, J. M. *Angew.Chem., Int. Ed.* **2006**, *45*, 3584. (b) Steinhauser, G.; Klapo¨tke, T. M.*Angew. Chem., Int. Ed.* **2008**, *47*, 3330. (c) Singh, R. P.; Gao, H.; Meshri, D. T.; Shreeve, J. M. In *High Energy Density Materials*; Klapo¨tke,T. M., Ed.; Springer: Berlin, Heidelberg, 2007; pp 35. (d) Klapo¨tke,T. M. In *High Energy Density Materials*; Klapo¨tke, T. M., Ed.; Springer:Berlin, Heidelberg, **2007**, 85.

9. (a) Tappan, B. C.; Huynh, M. H.; Hiskey, M. A.; Chavez, D. E.; Luther, E. P.; Mang, J. T.; Son, S. F. *J. Am. Chem. Soc.* **2006**, *128*, 6589. (b) Huynh, M. H.; Hiskey, M. A.; Meyer, T. J.; Wetzler, M. *Proc.Natl. Acad. Sci.* **2006**, *103*, 5409. (c) Hiskey, M. A.; Chavez, D. E.;Naud, D. L. 2001, US 6214139; *Chem. Abstr*. **2001**, *134*, 282929. (d) Hammer, A.; Holl, G.; Kaiser, M.; Klapo¨tke, T. M.; Mayer, P.; No¨th, H.;Piotrowski; Warchhold, H. M. *Eur. J. Inorg. Chem.* **2002**, 834.

10. (a)Sureshbabu, V. V.; Venkataramanarao, R.; Naik, S. A.; Chennakrishnareddy,G. *Tetrahedron Lett.* **2007**, *48*, 7038. (b) Purchase,C. F., II; White, A. D.; Anderson, M. K.; Bocan, T. M. A.; Bousley, R. F.;Hamelehle, K. L.; Homan, R.; Krause, B. R.; Lee, P.; Mueller, S. B.; Speyer,C.; Stanfield, R. L.; Reindel, J. F. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 1753. (c) Yamazaki, K.; Hasegawa, H.; Umekawa, Ueki, K.; Y.; Ohashi,N.; Kanaoka, M. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 1275. (d) Vieira,E.; Huwyler, J.; Jolidon, S.; Knoflach, F.; Mutelb, V.; Wichmann, J. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 4628.

11. (a)Wittenberger, S. J. *Org. Prep. Proced. Int.* **1994**, *26*, 499. (b) Butler, R. N. In *Comprehensive Heterocyclic Chemistry II*; Katritzky, A. R., Rees, C. W., Scriven, E. F. V., Eds.; Pergamon: Oxford, 1996;Vol. 4, p 674.

12. (a) Yoshinaga, J.; Shogaki, T.; Kakita, T.; Ozeki, H.; Sugimoto,N.; Kato, Y. Eur. Patent EP 262873, 1988; *Chem. Abstr*. **1988**, *109*, 54782. (b) Cetenko, W. A.; Connor, D. T.; Mullican, M. D.; Sorenson,R. J. Eur. Patent EP 249236, 1987; *Chem. Abstr*. **1988**, *108*, 150482.(c) Connor, D. T.; Unangst, P. C.; Weikert, R. J. Eur. Patent EP 279466,1988; *Chem. Abstr*. **1988**, *109*, 231035. (d) Ford, R. E.; Knowles, P.;Lunt, E.; Marshall, S. M.; Penrose, A. J.; Ramsden, C. A.; Summers,A. J. H.; Walker, J. L.; Wright, D. E. *J. Med. Chem*. **1986**, *29*, 538. (e)Peet, N. P.; Baugh, L. E.; Sundler, S.; Lewis, J. E.; Matthews, E. H.; Olberding, E. L.; Shah, D. N. *J. Med. Chem.* **1986**, *29*, 2403.

13. Girijavallabhan, V. M.; Pinto, P. A.; Ganguly, A. K.; Versace, R.W. Eur. Patent EP 274867, 1988; *Chem. Abstr*. **1989**, *110*, 23890.

14. (a) Akimoto, H.; Ootsu, K.; Itoh, F. Eur. Patent EP 530537, **1993**; *Chem. Abstr*. **1993**, *119*, 226417. (b) Taveras, A. G.; Mallams, A. K.; Afonso, A. Int. Patent WO 9811093, 1998; *Chem. Abstr*. **1998**, *128*, 230253.

15. (a) Mitch, C. H.; Quimby, S. J. Int. Patent WO 9851312, 1998, *Chem. Abstr*. **1998**, *130*, 13997. (b) Castro Pineiro, J. L.; Chambers,M. S.; Hobbs, S. C.; Matassa, V. G. Int. Patent WO 9319052, 1993;*Chem. Abstr*. **1994**, *120*, 134543.

16. Andrus, A.; Partridge, B.; Heck, J. V.; Christensen, B. G. *Tetrahedron Lett.* **1984**, *25*, 911.

17. (a) Katritzky, A. R.; Ledoux, S.; Witek, R. M.; Nair, S. K. *J. Org. Chem*. **2004**, *69*, 2976. (b) For a review, see: Sharma, S. *Synthesis* **1978**, 803. (c) Staab, H. A. *Angew. Chem., Int. Ed. Engl.* **1962**, *1*, 351. (d) Yamazaki, N.; Tomioka, T.; Higashi, F. *Synthesis* **1975**, 384. (e) Ranu, B. C.; Dey, S. S.; Bag, S. ARKIVOC. **2003**, *10*, 14. (f) Ballabeni, M.; Ballini, R.; Bigi, F.; Maggi, R.; Parrini, M.; Predieri, G.; Sartori, G.J. Org. Chem. **1999**, *64*, 1029. (g) Schroeder, D. C. Chem. Rev. 1955, 55, 181.

18. (a) Allen, C. F. H.; Edens, C. O.; vanAllan, J. In *Organic Synthesis; Wiley: New York*, **1955**; Vol. 3, p 394. (b)Yamazaki, N.; Higashi, F.; Iguchi, T. *Tetrahedron Lett*. **1974**, 1191. (c) Ballabeni, M.; Ballini, R.; Bigi, F.; Maggi, R.; Parrini, M.;Predieri, G.; Sartori, G. *J. Org. Chem*. **1999**, *64*, 1029. (d)Kurzer, F.; Douraghi-Zadeh, K. *Chem. Rev*. **1967**, *67*, 107.(e) Takikawa, Y.; Inoue, N.; Sato, R.; Takizawa, S. *Chem. Lett*.**1982**, 641. (f) Staab, H. A. *Angew. Chem., Int. Ed. Engl*. **1962**, *1*, 351. (b) Staab, H. A.; Walther, G. *Leibigs Ann. Chem*. **1962**, *657*, *98*.

19. (a) Sugimoto, H.; Makino, I.; Hirai, K. *J. Org. Chem*. **1988**, *53*, 2263.(b) Ramadas, K.; Srinivasan, N.; Janarthanan, N. *Tetrahedron Lett*.**1993**, *34*, 6447.(c) Xian, M.; Zhu, X.; Li, Q.; Cheng, J.-P. *Tetrahedron Lett*. **1999**, *40*, 1957. (d) Ramadas, K.; Srinivasan, N. *Synth. Commun*. **1995**, *25*, 3381. (e) Yamazaki, N.; Tomioka, T.; Higashi, F. *Synthesis* **1975**, 384. (f)Tamura, Y.; Adachi, M.; Kawasaki, T.; Kita, Y. *TetrahedronLett*. **1978**, 1753. (g) Tamura, Y.; Kawasaki, T.; Adachi, M.; Kita, Y. *Chem. Pharm. Bull*. **1979**, *27*, 1636 (h) Alan R. Katritzky,A.R.; Ledoux,S.; Witek,R.M.; Nair, S.K. *J. Org. Chem*. **2004**, *69*, 2976.

20. (a)Maddani, M.R.; Prabhu, K.R. *J. Org. Chem.* **2010**, *75*, 2327.(b) Nguyen, T.B.; Ermolenko, L.; AlMourabit, A. *Synthesis*. **2014**, *46*, 3172.

21. (a)Dallaire, C.;Arya, P. Tetrahedron Lett. **1998**, *39*, 5129.(b)Liu, J. Z.; Song, B. A.; Fan, H. T.; Bhadury, P. S.; Wan, W. T.; Yang, S.; Xu, W.; Wu, J.; Jin, L. H.; Wei, X.; Hu, D. Y.; Zeng, S. Eur. J. Med. Chem. **2010**, *45*, 5108.(c)Klein, J. J.; Hecht, S. Org. Lett. **2012**, *14*, 1, 330. (d)Suhas, R.; Chandrashekar, S.; Gowda, D. C. Eur. J. Med. Chem. **2012**, *48*, 179.(e)Dungan, V. J.; Ngo, H. T.; Young, P. G.; Jolliffe, K. A. Chem. Commun. **2013**, *49*, 264.(f)Martí, A.; Costero, A. M.; Gaviña, P.; Parra, M. Eur. J. Org. Chem., **2015**, 6597.(g)Jin, L.; Qu, H. E.; Huang, X. C.; Pan, Y. M.; Liang, D.; Chen, Z. F.; Wang, H. S.; Zhang, Y. Int. J. Mol. Sci. **2015**, *16*, 14571.(h)Ghasemi, E.; Shahvelayati, A. S.; Yavari, I. *Phosphorus Sulfur Silicon Relat. Elem*. **2016**, *00*, 1.(i)Takemoto, Y. *Org. Biomol. Chem*. **2005**, *3*, 4299.(j)Akiyama, T.; Itoh, J.; Fuchibe, K. Adv. Synth. Catal. **2006**, *348*, 999.(k)Fu, J. Y.; Wang, Q. L.; Gui, Y. Y.; Wang, L. X. Tetrahedron Lett. **2015**, *56*, 4220.

22. (a) Finnegan, W. G.; Henry, R. A.; Lieber, E. *J. Org. Chem.* **1953**, *18*, 779. (b) Jensen, K. A.; Holm, A.; Rachlin, S. *Acta Chem. Scand.* **1966**, *20*, 2795. (c) Percival, D. F.; Herbst, R. M. *J. Org. Chem.* **1957**, *22*, 925. (d) Ford, R. E.; Knowles, P.; Lunt, E.; Marshall, S. M.; Penrose, A. J.; Ramsden, C. A.; Summers, A. J. H.; Walker, J. L.; Wright, D. E. *J. Med. Chem*. **1986**, *29*, 538.

23. Ding, Y.-X.; Weber, W. P. *Synthesis* **1987**, 823.

24. (a) Moderhack, D.; Goos, K.-H.; Preu, L. *Chem. Ber.* **1990**, *123*, 1575. (b) Garbrecht, W. L.; Herbst, R. M. *J. Org. Chem.* **1953**, *18*, 1014.(c) Herbst, R. M.; Roberts, C. W.; Harvill, E. J. *J. Org. Chem.* **1951**,*16*, 139. (d) Marchalin, M.; Martvon, A. *Collect. Czech. Chem. Commun.* **1980**, *45*, 2329

25. (a) Ried, W.; Erle, H.-E. *Liebigs Ann. Chem.* **1982**, 201. (b) Ried, W.; Dietschmann, H.; Erle, H.-E. *Synthesis* **1980**, 619. (c) Ku¨ hle, E.; Anders, B.; Klauke, E.; Tarnow, H.; Zumach, G. *Angew. Chem., Int. Ed. Engl.* **1969**, *8*, 20. (d) Ku¨ hle, E.; Anders, B.; Zumach, G. *Angew.Chem., Int. Ed. Engl.* **1967**, *6*, 649. (e) Miller, A. E.; Feenay, D. J.; Ma, Y.; Zarcone, L.; Aziz, M. A.; Magnuson, E. *Synth. Commun.* **1990**, *20*,217. (f) Batey, R. A.; Powell, D. A. *Org. Lett.* **2000**, *2*, 3237.

26. (a) Morley, J.S. *J. Chem. Soc.* C. **1969**, 809.(b)Crnstein, P.L.; Arnold,M.B.; Evrard, D.; Leander, J.D.; Lodge, D.; Darryte , D. *Bioorg. Med. Chem. Lett*. **1993**, 3, 43.(c)Sureshbabu, V.V.; Venkataramanarao, R.; Naik, S.A.; Chennakrishnareddy,G. Tetrahedron Lett.**2007**, *48* , 7038. (d)Nachmana, R.J.; Zabrockia, J.; Olczakb, J.; Williamsc, H.J.; Guillermo Moynac, G.; Scottc, A.I.; Coaste,G.M.; *Peptides*. **2002**, *23*, 709.(e) Qu, Z.-R.; Zhao,H.; Wang, X-S.; Li, Y.-H.; Song,Y.-M.; Liu,Y.-J.; Ye, Q.; Xiong, R.-G.; Abrahams, B.F.; Xue, Z.-L.; You, and X.-Z. *Inorg Chem*. **2003**, *42*, 7710.

27. (a) Sharon, A.; Pratar, R.; Tiwari, P.; Srivastava, A.; Maulik, P.R.; Ram, V.J. *Bioorg. Med. Chem. Lett.* **2005**, 15, 2115. (b)Brauner-Osborn, H.; Egebjerg, J.; Nielsen, E. O.; Madsen, U.; Krogsgaard-Larsen, P. *J. Med. Chem.* **2000**, *43*, 2609.

28. (a) P. L. Ornstein, M. B. Arnold, N. K. Allen, T. Bleisch, P. S. Borromeo, C. W. Lugar, J. D. Leander,D. Lodge, and D. Schoepp, *J. Med. Chem.***1996,** *39*, 2219. (b) M. J. O'Neill, M.J.; Bond, A.; Ornstein, P. L.; Ward, M.A.; Hichs, C.A.; Hoo, K.; Bleakman, D.; Lodge, D. *Neuropharmacology*. **1998**. *37*, 1211. (c) Dominguez, E.; Iyengar, S.; Shannon, H. E .; Bleakman, D.; Alt, A.; Arnold, B. M.;Bell, M.G.; T. J. Bleisch, T.J.; Buckmaster, J. L.; Castano, A. M.; DelPrado, M.; Escribano, A.; Filla, S. A. ; Ho, K. H.; Hudziak, K. J.; Jones, C. K.; Martinez-Perez, J. A. ; Mateo, A.; Mathes, B. M.; Mattiuz, E. L.; Ogden, A. M. L.; Simmons, R. M. A.; Stack, D. R.; Stratford, R. E.; Winter, M. A.; Zhipei Wu, Ornstein, P. L. *J. Med.Chem.* **2005**, *48*, 4200.

29. (a) Rajasekaran. A.; Thampi, P. P. *Eur. J. Med. Chem*. **2004**, *39*, 273.(b) Batt, D. G. ; Houghton, G. C. ; Roderick, J. ; Santella, J. B.; Wacker, D. A.; Welch, P. K.; Orlovsky, Y. I. ; Wadman, E. A. ; Trzaskov, J. M.; Davies, Decicco, C. P.; Carter, P. H. *Bioorg. Med. Chem. Lett.* **2005,** *15*, 787.

30. (a) Habich, D. *Synthesis*, **1992**, 358.(b) Powers, R. A.; Shoichet, B. K. *J. Med. Chem.* **2002,** *45*, 3222. (c) Lee, P.-Y. ; Chang, W.-N. ; Lu, C.-H. ; Lin, M.-W. ; Cheng, B.-C.; Chien, C.-C.; Chang, C.-J.; Chang, H.-W.; *Antimicrob. Agents Chemother.* **2003,** *51*, 957.

31. (a) Ashton, W. T. ; Cantone, C. L.; Meurer, L. C. ; Tolman, R. L.; Greenlee, W. J.; Patchett, A. A.; Lynch, R. J.; Schorn, T. W.; Strouse, J. F.; Sieg, P. K. S. J. Med. Chem. **1992**, *36*, 2103.(b)Vieira, E.; Huwyler, J.; Jolidon, S.; Knoflach, F.; Mutel, V.; Wichmann, J.; *Bioorg. Med. Chem. Lett.* **2005**, *15*, 4628. (c) Chang, C.-S. ; Lin, Y.-T.; Lee, C.-C.; Lee, Y.-C.; Tai, C.-L.; Tseng, S.-N.; Chern, J.-H. *J. Med. Chem.* ***2005***, *48*, 3522.

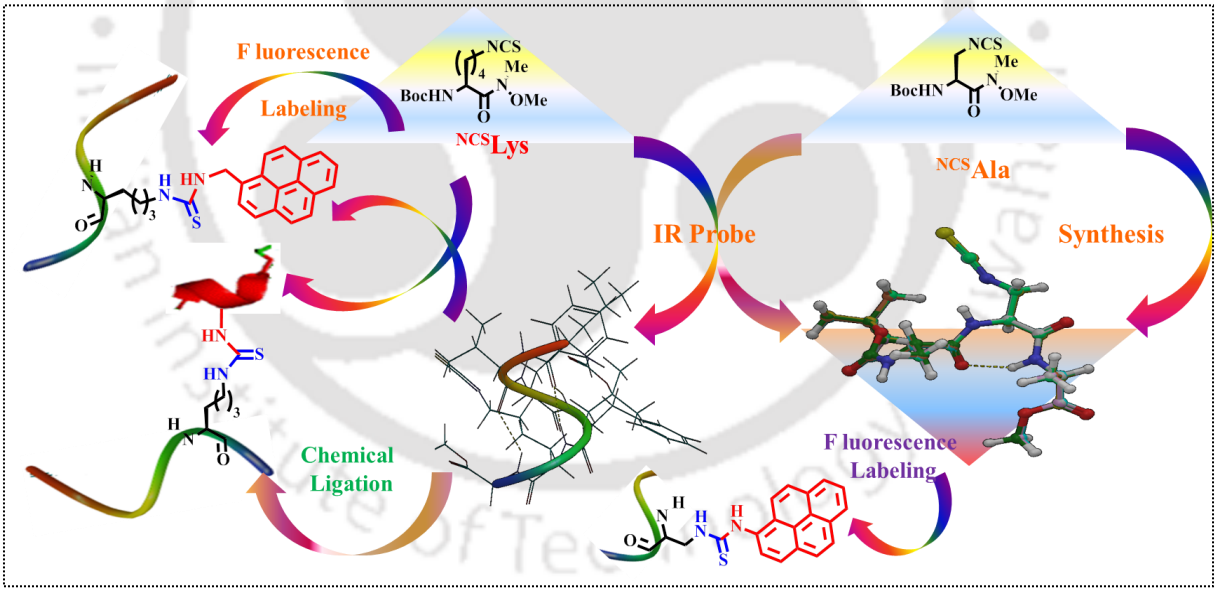
32. (a) Frings, M.; Thomé, I.; Bolm, C. Beilstein *J. Org. Chem*. **2012**, *8*, 1443. (b) Volla, C. M. R.; Atodiresei, I.; Rueping, M. *Chem. Rev*. **2014**, *114*, 2390.

33. Mahdavi, M.; Asadi, M.; Khoshbakht, M.; Saeedi, M.;Bayat, M.; Foroumadi, A.; Shafiee, A. *Helv. Chim. Acta* **2016**, *99*, 378.

34. a) Klonis, N; Sawyer, W. H. *Photochem. Photobiol.* **2003**, *77*, 502. (b) Shah, N. M.; Mehta, D. H. J. *Ind. Chem. Soc.* **1954**, *31*, 784. (c) Grabowksi, Z. R.; Rotkiewicz, K.; Rettig, W. *Chem. Rev.* **2003**, *103*, 3899. (d) Badger, G. M.; Walker, I. S. *J. Chem. Soc.* **1956**, 122.

Chapter-4

Isothiocyanyl Alanine / Lysine as Solvatochromic IR Responsive Probes and Use in Site Specific Labeling/Ligation of Short Peptides



4.1. Introduction

Side chain modified unnatural amino acids (UNAAs) have attracted current research activity as a novel spectral probe of peptides/proteins. As for example, fluorescent UNAAs [1-3] (**4.001-4.002, Figure 4.01**), labeled EPR active UNAAs[4] (**4.003-4.004, Figure 4.01**), photoaffinity labels[5](**4.005-4.006, Figure 4.01**) have been reported in recent years. As a part of our research, we also have contributed to the generation of fluorescent triazolyl UNAAs as peptide labels for conformational study and addressing fundamental photophysical aspects in predefined peptide conformations.[6] However, large sizes of the labels often caused problems of structural perturbation. Therefore, simple and small functional groups as side chain labels would be more beneficial leading to minimal or no structural perturbation of the peptide/protein secondary structure. The intrinsic stretching vibrational motion of diatomic functional groups as chromophores, thus, have attracted much recent interest as infrared (IR) spectral probes of peptides/proteins.[7-8]

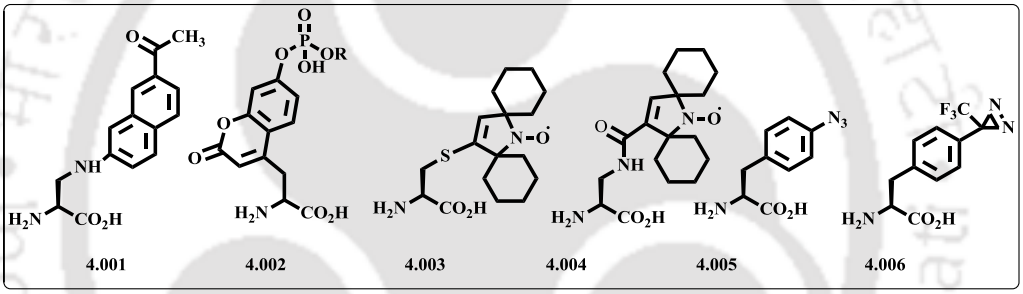


Figure 4.01. Some fluorescent, EPR and photoaffinity labeled UNAAs.

However, choosing an appropriate IR probe for a specific application is a challenging task. In spite of the great success of –CN as IR probe, [9a-d, 10a,15] it often suffers from shortcomings such as small dipole strength of –CN leading to weaker IR intensity than – N_3 and interference by Fermi resonance.[9c-d] On the other hand, because of large dipole strength $–N_3$ has got popularity as IR probe.[9a,10a-b] However, broad bandwidth, markedly short vibrational lifetime and overlapped stretching with accidental Fermi resonance bands are its drawback.[10c-4f] Furthermore, many of the IR-chromophores, such as, -SCN, has been introduced *via* derivatisation of cysteine which could be a problem when present in multiples in a protein.[11] Therefore, novel side chain which does not rely on post-synthetic modification and possesses stronger IR intensity would be more attractive research. However, choosing an appropriate IR probe for a specific application is a challenging task since the probe's geometry in solution and vibrational properties need to be considered. Despite their analytical use,[12] isothiocyanates have attracted much research attention in current synthetic thrust.[13]

Moreover, the stronger IR transition dipole strength of -NCS compared to reported >C=O and comparable to -CN/N_3 group, would be a promising IR probe for studying peptide's/protein's microenvironment. As a part of our continuous research efforts, the design

of unnatural amino acids [6] and the limitations of the currently used IR probes, have led us to think –NCS, for the first time, as a new family of possible high sensitive and intensive IR stretch probe.

4.2. Solvalotochromic IR Responsive Probes of Proteins

4.2.1. The –CN as Solvalotochromic IR Responsive Probe

In 2003, Getahun *et al*.,[14] have reported, nitrile-derivatised amino acids (**1.079-1.080, Chapter-1**) and they used them as local environment sensors in model peptide MLCK579-595 (**4.007, Figure 4.02**). It is mainly advantageous to employ such a probe to investigate the processes where a specific amino acid undergoes large environmental changes, as for example, from polar to hydrophobic or vice versa. They also reported infrared spectroscopy in combination with the modern chemical as well as a biological process may be used to study a variety of structural and dynamic properties, such as side-chain packing, side-chain interaction, peptide-protein binding, and protein-protein interaction. In 2010, Taskent-Sezgin *et al*.,[18] have incorporated also *p*-cyanophenylalanine in the model peptide (**4.010, Figure 4.02**) and monitored IR of modified peptide (**4.011, Figure 4.02.**). They explained by IR that the Phe_{CN} fluorescence was quenched by interactions with a Tyrosine side chain in the native state where –CN group was exposed to solvent. C. Schultz *et al*.,[15] have incorporated *p*-cyanophenylalanine by the genetically as infrared probe in bacteria. They also reported this unnatural amino acid is useful for site-specific probe of protein folding, conformational changes, and biomolecular interactions based on IR experiments.

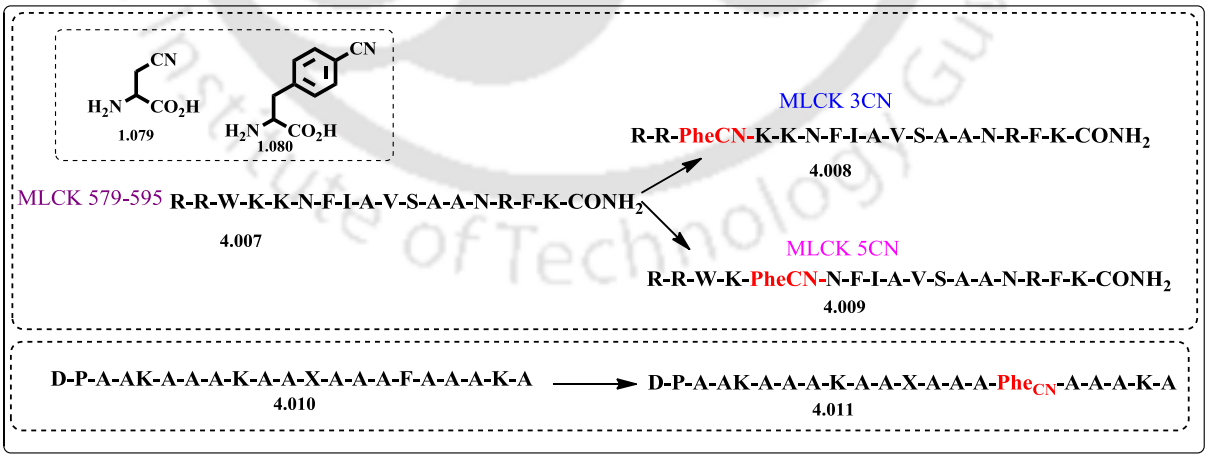


Figure 4.02. Structure of *p*-cyanophenylalanine (**1.080**), *β*-cyanoalanine (**1.079**), some natural peptide analogues **(4.007, 4.010**) and modified peptide analogues (**4.008-4.009, 4.011**) with *p*-cyanophenylalanine.

Waegele *et al.*,[17] have reported that the C≡N stretching vibration of 5-cyanotryptophan (**4.012, Figure 4.03**) is a useful infrared probe of the local environment in a membrane protein. They investigated how the C≡N stretching band of the free amino acid changes with the solvent in two ^{CN}**Trp** containing mutants (**4.014-4.015, Figure 4.03**) of a cationic antimicrobial peptide indolicidine **4.013** (ILPWKWPWWPWRR-NH_2) after modification once at the 9th and 11th position with 5-cyanotryptophan. They found on binding to the model membrane the IR bandwidth of –CN vibrational mode in indolicine changes, thus it acts as a very sensitive infrared probe of local hydration status of the peptides.

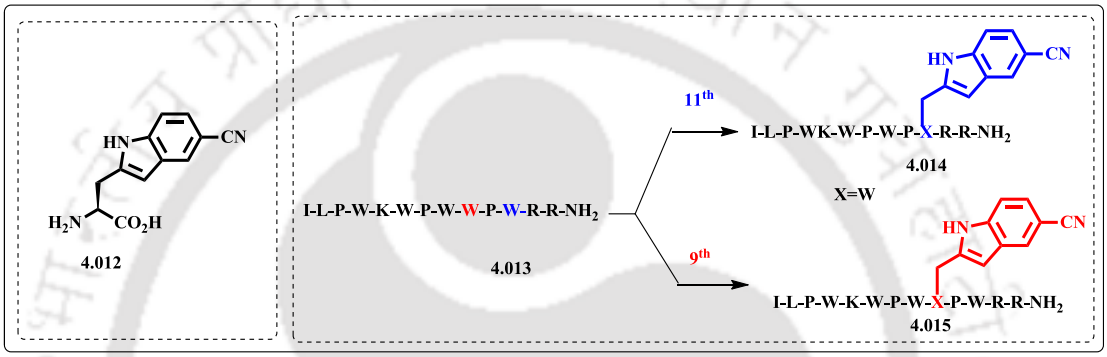


Figure 4.03. Structure of 5-cyanotryptophan (**4.012**), iodolicine (**4.013**) and 5-cyanotryptophan containing indolicidine (**4.014-4.015**).

Jo *et al.*,[11a] have prepared several aromatic nitrile cysteine (**4.016-4.017, Figure 4.04**) and successfully incorporated site specifically into peptides/proteins *via* either thiol alkylation or arylation reaction (**4.018-4.019, Figure 4.04**) of cysteine and studied IR. They found on addition of Ca^{2+} the bandwidth of the IR transition is decreased from ~19 cm^{-1} to ~11 cm^{-1}. Furthermore, the peak position (**4.020, Figure 4.04**) of this nitrile band (ν = 2236.0 cm^{-1}) remains in the same position, on addition of $Ca^{2+,}$ which signifies that the labeled cysteine residue is exposed to solvent in CaM structures obtained in both the absence and presence of Ca^{2+}.

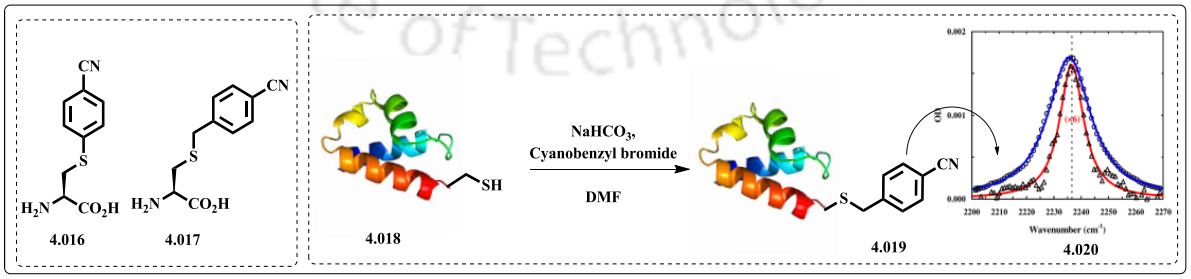


Figure 4.04. Structure of aromatic nitrile cysteine (**4.016-4.017**), schematic cysteine alkylation (**4.018-4.019**) and IR spectra (**4.020**) of modified peptide.

4.2.2. The $-N_3$ as Solvalotochromic IR Responsive Probe

In 2008, Oh *et al.*,[19] have established that *β*-azidoalanine (*β*-AlaN_3) (**1.014, Chapter-1**) has positive benefits over nitrile-derivatized amino acids when used as an IR probe as dipole strength of the CN and SCN stretch mode is almost 19 and 5 times lower than azide stretch mode. They also revealed the efficacy of *β*-azidoalanine (*β*-AlaN_3) in place of alanine as an IR probe of the local electrostatic environment in proteins and they have successfully investigated the aggregation of A*β*(16-22) amyloid peptide(**4.021, Figure 4.05**) containing a single *β*-azidoalanine at position twenty-one. They found that azido peak frequency of such amyloid peptides in both water and DMSO is 2104 cm^{-1} which signifies that the azido group in the aggregate is encircled by a local electrostatic environment similar to DMSO.

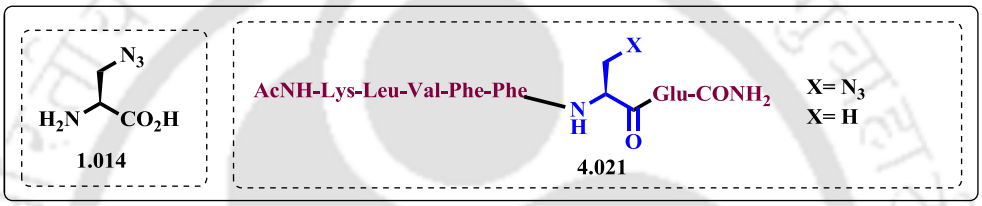


Figure 4.05. Azidoalanine (**1.014**), A*β*(16-22) amyloid peptide (X=H) and A*β*(16-22) amyloid peptide containing azidoalanine .

In 2009, Ye *et al.*,[20] have reported that the position specific incorporation of an IR-sensitive amino acid, para-azido-L-phenylalanine (*p*-PheN_3) (**4.005, Figure 4.06**), into the G protein–coupled receptor rhodopsin incorporated at positions M253 and L72 (**4.023, Figure 4.06**), which are part of the solvent-exposed G-protein-binding pocket in the active conformation of rhodopsin using amber codon suppression technology. The antisymmetric stretch vibration of the azido group absorbs at ~2,100 cm^{-1} in a clear spectral window which is sensitive to its electrostatic environment. They used FTIR spectroscopy to study the azido probe and showed that the electrostatic environments of specific interhelical networks change during receptor activation.

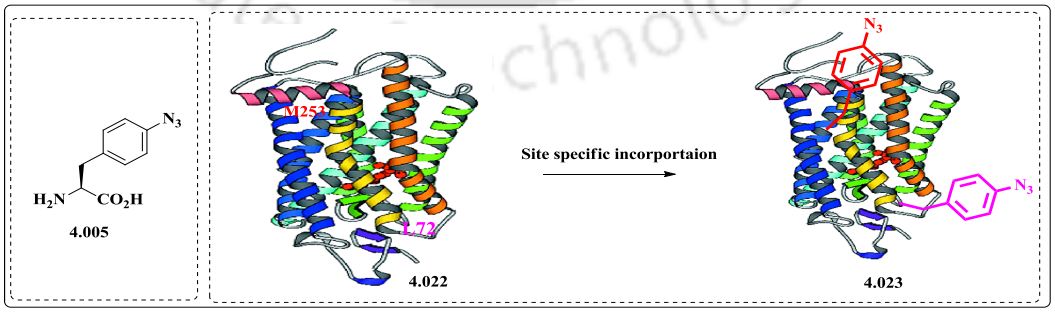


Figure 4.06. Structure of *p*-azido-L-phenylalanine, G protein–coupled receptor rhodopsin **4.022** and G protein–coupled receptor rhodopsin containing *p*-azido-L-phenylalanine **4.023**.

In the next year 2010, Taskent-Sezgin *et al.*,[21] have chosen N-terminal domain of the ribosomal protein L9 (NTL9) **4.025** (**Figure 4.07)** and targeted methionine1(Met1) by expression (NTL9-Met1Aha) and isoleucine-4(Ile4) **4.026** (NTL9*-Ile4Aha) by chemical synthesis as sites for incorporation of the azidohomoalanine **(**Aha**)** substitutions which acts as a good IR sensitive probe of protein for monitoring protein folding, protein structure, and electrostatic effects . They explained that the probe can be easily incorporated into proteins in high yield in a position-specific manner using simple, readily available auxotrophic expression systems as well as by solid-phase peptide synthesis.

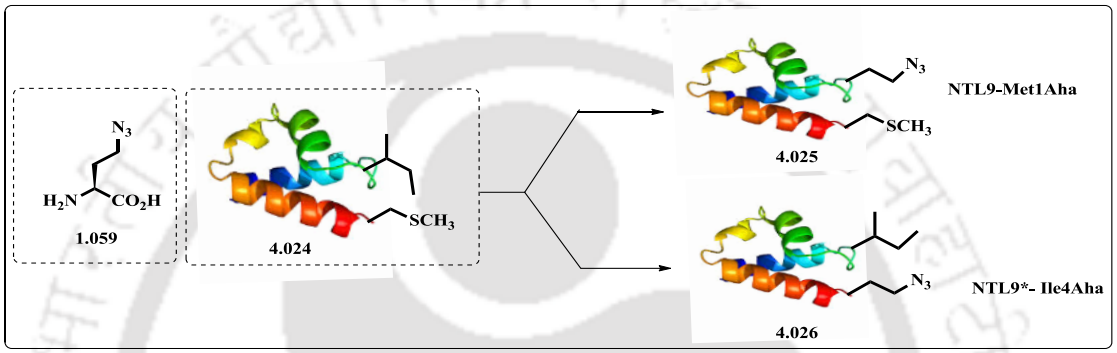


Figure 4.07. Structure of azidohomoalanine (**1.059**), ribosomal protein L9 (**4.024**) and modified ribosomal protein L9 containing azidohomoalanine (**4.025-4.026**).

4.2.3. The –SCN as Solvalotochromic IR Responsive Probe

The SCN group has relatively large Stark tuning rate of the nitrile stretching vibrations and it can be used to monitor the local electric field in chemical as well as biological systems. As for example, by positioning beta-thiocyanatoalanine (***β*-^{SCN}Ala**) (**4.027, Figure 4.08**) into the active site pocket of the enzyme ketosteroid isomerase (KSI) (**4.028-4.029, Figure 4.08**) and using the nitrile stretching mode as a vibrational stark probe, Boxer, Herschlag, and co-workers [22] investigated the changes in the local electrostatic field associated with substrate binding and steroid isomerisation induced charge localisation.

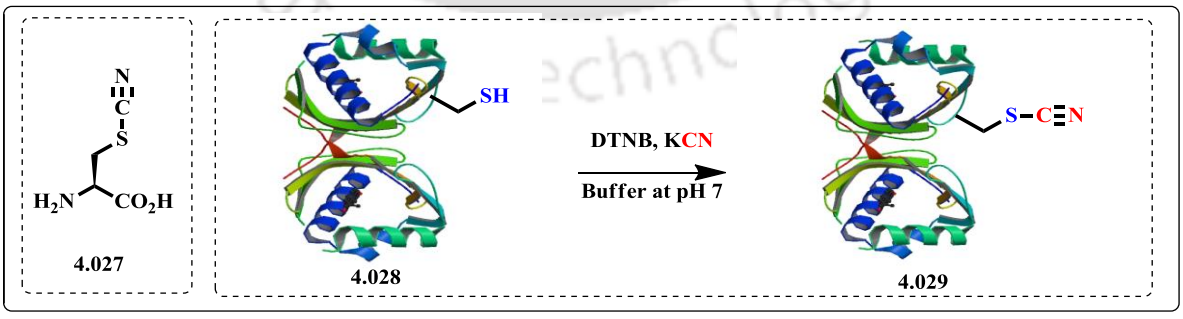


Figure 4.08. Beta-thiocyanatalanine (**4.027**), enzyme ketosteroid isomerase (**4.028**) and active site pocket of the enzyme ketosteroid isomerase contained thiocyanatoalanine (**4.029**).

In 2010, Edelstein *et al.*,[23] have synthesised thiocyanatoalanine, or cyanylated cysteine and showed that solvent exposed cysteine residues in proteins find wide use in large protein systems (Ac-AAAAKAAAAKAAAAKAAAAKAAAAK-NH_2) as a probe of site-specific structure and dynamics. The SCN stretching vibration of this artificial side chain provides an isolated infrared chromophore. The cyanylated cysteine side chain in peptides (**4.031-4.033, Figure 4.09**) displayed to destabilize, but not completely disrupt, the helical structure of the folded peptide when replaced for alanine.

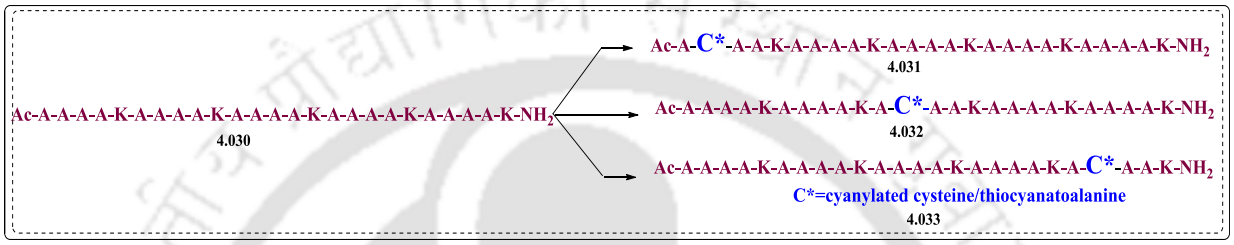


Figure 4.09. Structure of natural peptide analogues (**4.030**) and thiocyanatoalanine containing peptide analogues (**4.031-4.033**).

In the same year McMahon *et al.*,[24] have used *β*-thiocyanatoalanine into peptides *via* post-translation modification at cysteine residue and studied cyanylated cysteine in two model peptides (Ac-KWKLFKKIGAVLKVL-NH_2 **4.034**; Ac-NPVVHFFKNIVTPRT-NH_2; **4.035**) (**Figure 4.10**) at selected sites and have shown binding to the membrane interfaces of the human myelin basic protein and the antimicrobial peptide CM15, infrared spectra of both systems in buffer and exposed to dodecylphosphocholine micelles indicate that the CN stretching absorption band of cyanylated cysteine can clearly differentiate between membrane burial and solvent exposure of the artificial side chain. Later Fafarman *et al.*,[25] have also developed a straightforward and general method to introduce the thiocyanate nitrile stretch as a site-specific electric field probe for proteins.

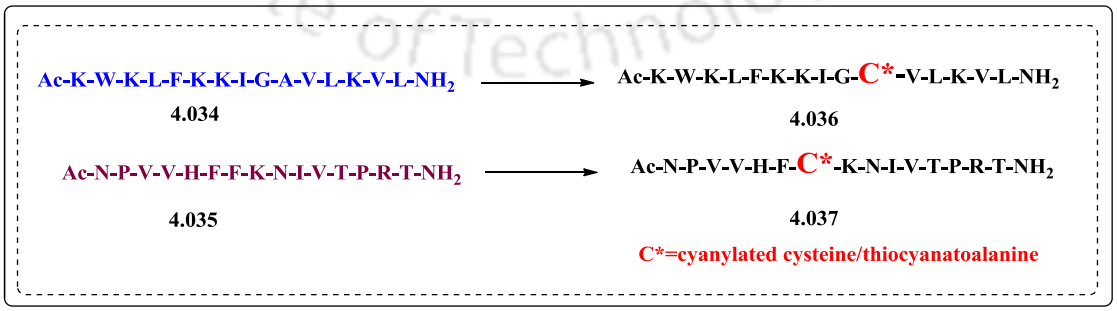


Figure 4.10. Structure of natural peptide analogues (**4.034, 4.035**) and cyanylated cysteine containing peptide analogues (**4.036, 4.037**).

4.3. Background

Large number of unnatural amino acids (UNAAs) with various novel functional roles has been introduced [1] and many of them have been incorporated into proteins *via* modern synthetic and biological methods.[1-2] Several of such functional amino acids have been utilised to investigate the structure, dynamics, localisation, and biomolecular interactions both *in vitro* and *in vivo*.[26] Many side chain modified UNAAs have also been reported as novel spectral probes of peptides/proteins. Since IR probe as label has been used in a variety of lipid systems for monitoring the protein-lipid interactions of membrane, it has been found that some dye molecules, and other fluorescent molecules are used as probes most of the times for this purpose. However, large sizes of the side chain labels often caused problems of unfavorable structural perturbation. The advantage lies in the fact that the IR absorption bands can be tuned by judicious design of side chain functional group to avoid any spectral overlap with vibrational bands in native amino acids. Therefore, several of such chromophores, such as CO, [27a] nitrobenzoxadiazole,[27b] –CN [28a-b] , -SCN[28b-c] , -NCO, -PH, and –SiH [7] with distinct IR-absorption in the well-known open window of peptide (2100–2300 cm^{-1}) have been exploited to extract information of the micro environment of peptides/proteins.[3f, 8, 29-30]

Therefore, the limitations of the currently used IR probes led us to think, -NCS as a new family of possible high sensitive and intensive IR stretch probe which expectedly would be utilised more efficiently for the study of the structure and dynamics of peptides/proteins. Furthermore, the installation of such small side chains on the backbone conformation of peptides/proteins is much less studied. Recently, our research group is involved in the synthesis of side chain modified UNAAs. Herein, we present ^{NCS}**Ala** & ^{NCS}**Lys** unnatural amino acids as infrared (IR) responsive probe in short peptides.

4.4. Objective

The study of backbone conformations of small/large peptides from the knowledge of amide-I IR absorptions has been a subject of investigation in recent time.[8,27,31] The applications of various novel IR chromophores have been reported as IR reporters for the structure and dynamics of the local environment.[7,28] However, choosing an appropriate IR probe for a specific application is a challenging task, so in this chapter we demonstrated the application of –NCS as a novel chromophore. Herein, the –NCS is used for possible probing of local structures and electrostatic microenvironment of a short peptide containing ^{NCS}**Ala/** ^{NCS}**Lys**. From literature survey, it is clear that very few studies are available for solvatochromic IR sensitive probe for monitoring protein folding and electrostatics. Hence, utilisation of -NCS is not only acts as fluorescent labeling /ligation group in peptides but also as solvatochromic IR probe in ^{NCS}**Ala/**^{NCS}**Lys** containing peptides to monitor the protein conformation. Therefore, with this aim and idea we framed our objectives as below:

(a) Design and synthesis of natural and side chain modified tripeptide/pentapeptide/hexapeptide by incorporating natural and ^{NCS}**Ala /**^{NCS}**Lys** amino acids respectively in peptide (**Figure 4.11**) *via* well know solution phase peptide coupling protocol.

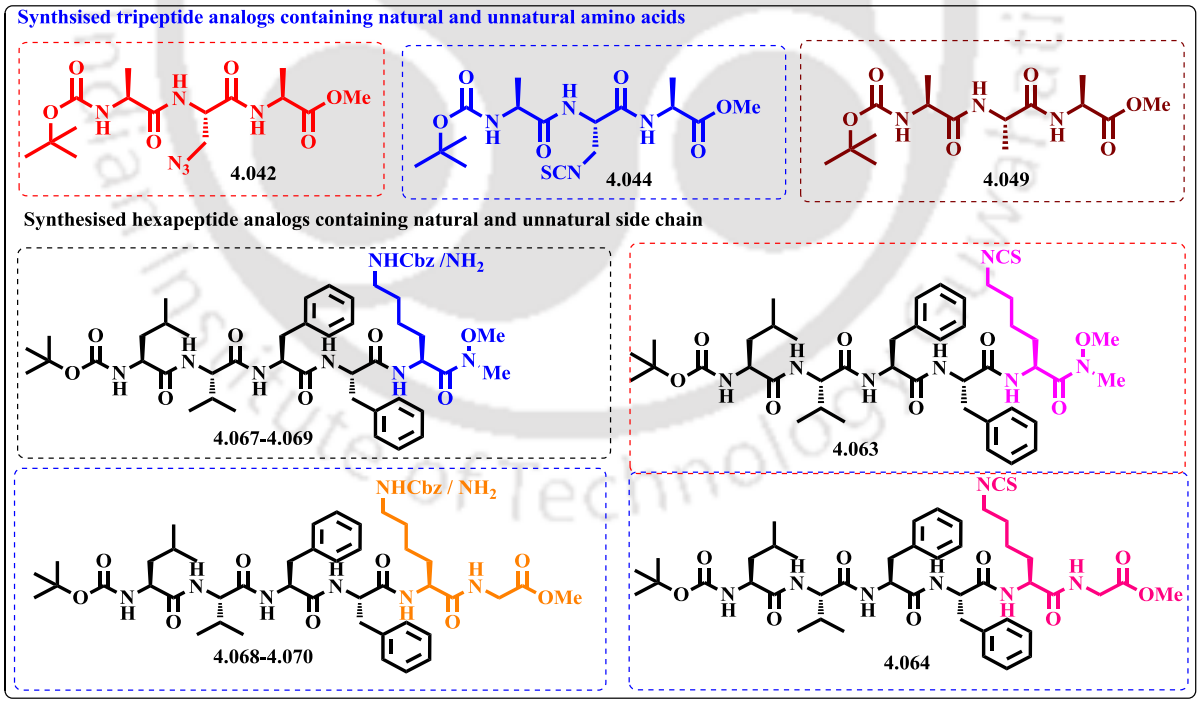


Figure 4.11. Synthesised natural and side chain modified tripeptide/pentapeptide/hexapeptide analogues.

(b) The electrophilicity of –NCS in ^{NCS}**Ala**/ ^{NCS}**Lys** has been utilised for site-specific fluorescent labeling/ligation reaction in tripeptide/hexapeptide (**Figure 4.12**) under physiological condition.

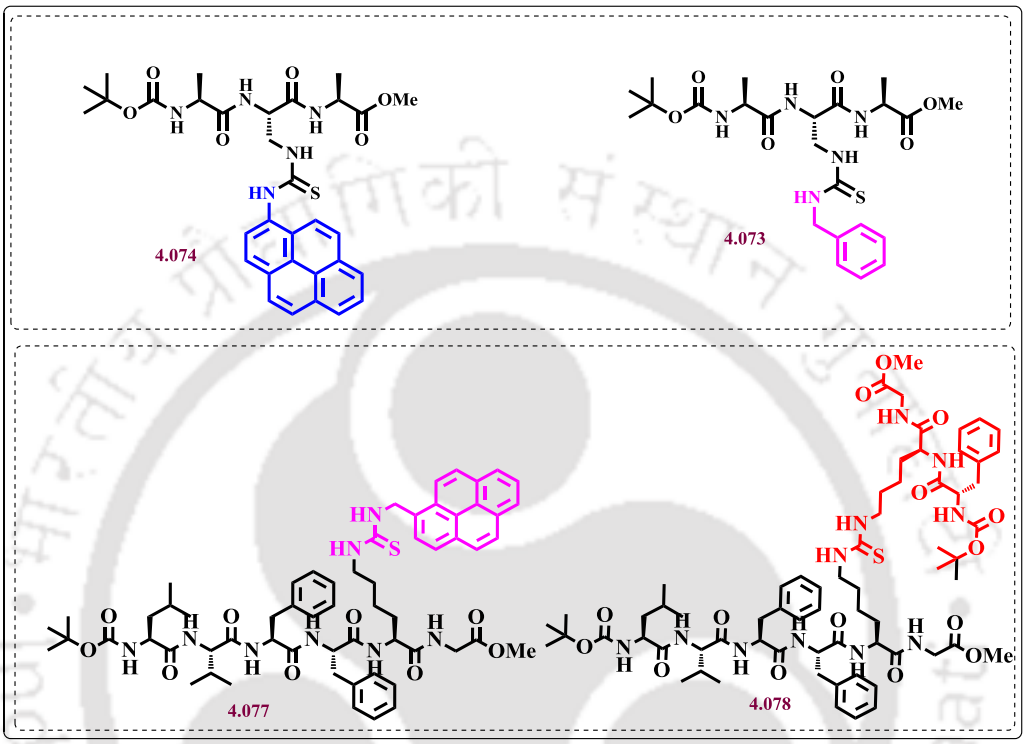


Figure 4.12. Synthesised labeled tripeptides/ hexapeptide and ligated hexapeptide.

(c) Study of solvatochromic IR sensitivity of ^{NCS}**Ala**/ ^{NCS}**Lys** of in small peptide/protein

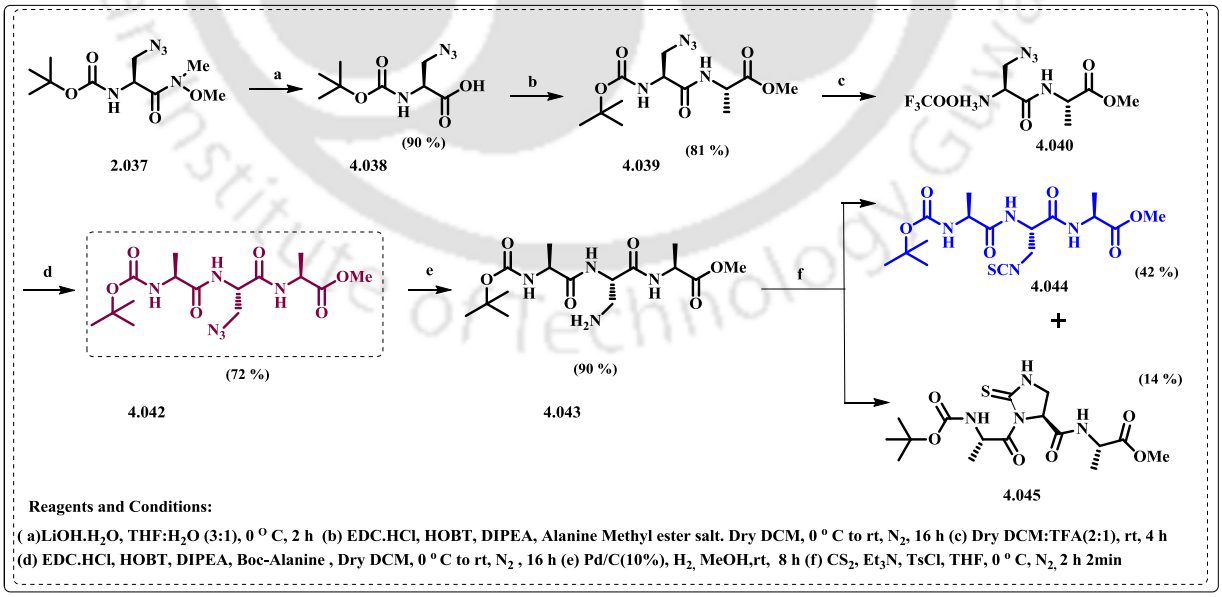
(d) We also examined the effect of –NCS side-chain conformation of short tripeptide, pentapeptide and hexapeptide using NMR, CD, and powder x-ray diffraction.

(e) Study the photophysical properties of the fluorescence-labeled pyrenylthioureayl hexapeptide.

4.5. Results and Discussion

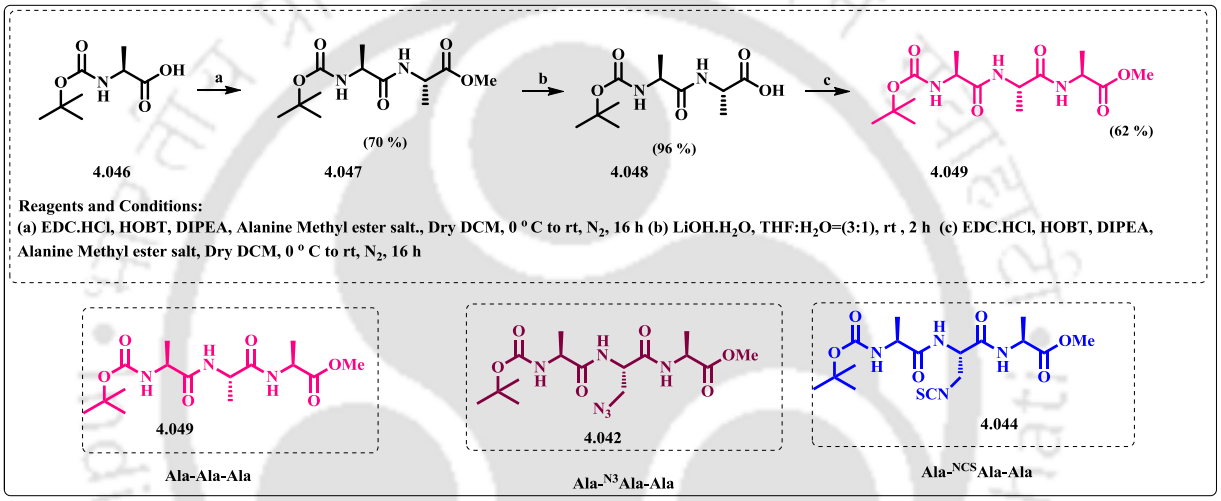
4.5.1. Synthesis of Tripeptide Containing ^{NCS}Ala

To get quick access of the tripeptide containing **^{NCS}Ala**, we adopted a strategy which is depicted in **Scheme 4.01**. We have synthesised our designed tripeptides (**Scheme 4.01**) (a) Ala-**^{N3}Ala**-Ala (**4.047**) (b) Ala **^{NCS}Ala**-Ala (**4.048**) and (c) Ala-Ala-Ala (**4.054**) decorated with our synthesised unnatural amino acids **^{NCS}Ala** and **$^{N}{}_{3}$Ala** (**2.037**, **described in Chapter 2**). Firstly, we have prepared BocNH-**^{N3}Ala**-CO_2H (**4.038**) starting from N,C diprotected serine azide (**2.037**) upon treatment with LiOH in water/THF (3:1) at 0 ^{o} C for 2 hours followed by acidification with dilute HCl (pH=2-3). Secondly, $^{N}{}_{3}$Ala-Ala dipeptide was prepared (**4.039**) by EDC.HCl and HOBT mediated solution phase peptide coupling protocol between, BocNH-^{N3}Ala-CO_2H (**4.046**) and HCl-salt of alanine methyl ester. Next, the BOC of $^{N}{}_{3}$Ala-Ala dipeptide was deprotected to corresponding TFA salt of dipeptide (**4.040**) upon treatment with TFA in dry DCM for 4 hours which was further used for EDC.HCl and HOBT mediated solution phase peptide coupling with Boc-Alanine to afford our targeted azido tripeptides (**4.042**). After that, H_2/Pd-C mediated reduction of azido functionality of tripeptide afforded the tripeptide **4.043** containing amino alanine (^{NH2}Ala). Finally, CS_2 and Et_3N were added to a solution of amino alanyl tripeptide **4.043** in THF and stirred for 4 min at 0 ^{o} C followed by addition of TsCl. Within 2 minutes the reaction was completed with the generation of desired tripeptide **4.044** containing **^{NCS}Ala** in 42% and a cyclic tripeptidyl thioamide **4.045** in 14% isolated yield. All the intermediates and the final peptides were characterised by IR, NMR and mass spectrometry.



Scheme 4.01. Schematic of formation of different tripeptide analogues.

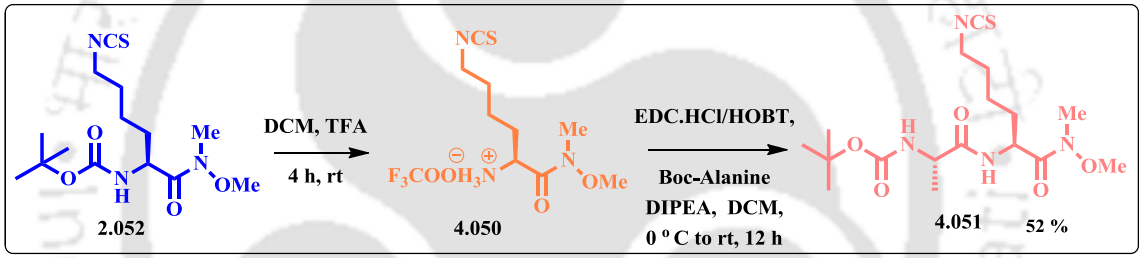
Similarly, Ala-Ala-Ala (**4.049, Scheme 4.02**) was synthesised *via* solution phase peptide coupling method. Firstly, Ala-Ala dipeptide (**4.047**) was prepared by EDC.HCl and HOBT mediated solution phase peptide coupling protocol between Boc-Ala-CO_2H (**4.046**) and HCl-salt of alanine methyl ester. Next, the LiOH mediated hydrolysis of ester group of dipeptide **4.047** was used to prepare dipeptide acid **4.048** which was further used for EDC.HCl and HOBT mediated solution phase peptide coupling with HCl salt of alanine ester to form desired tripeptide Ala-Ala-Ala (**4.049**). All the intermediates and the final peptides were characterised by IR, NMR and mass spectrometry.



Scheme 4.02. Schematic of formation of different tripeptide analogues.

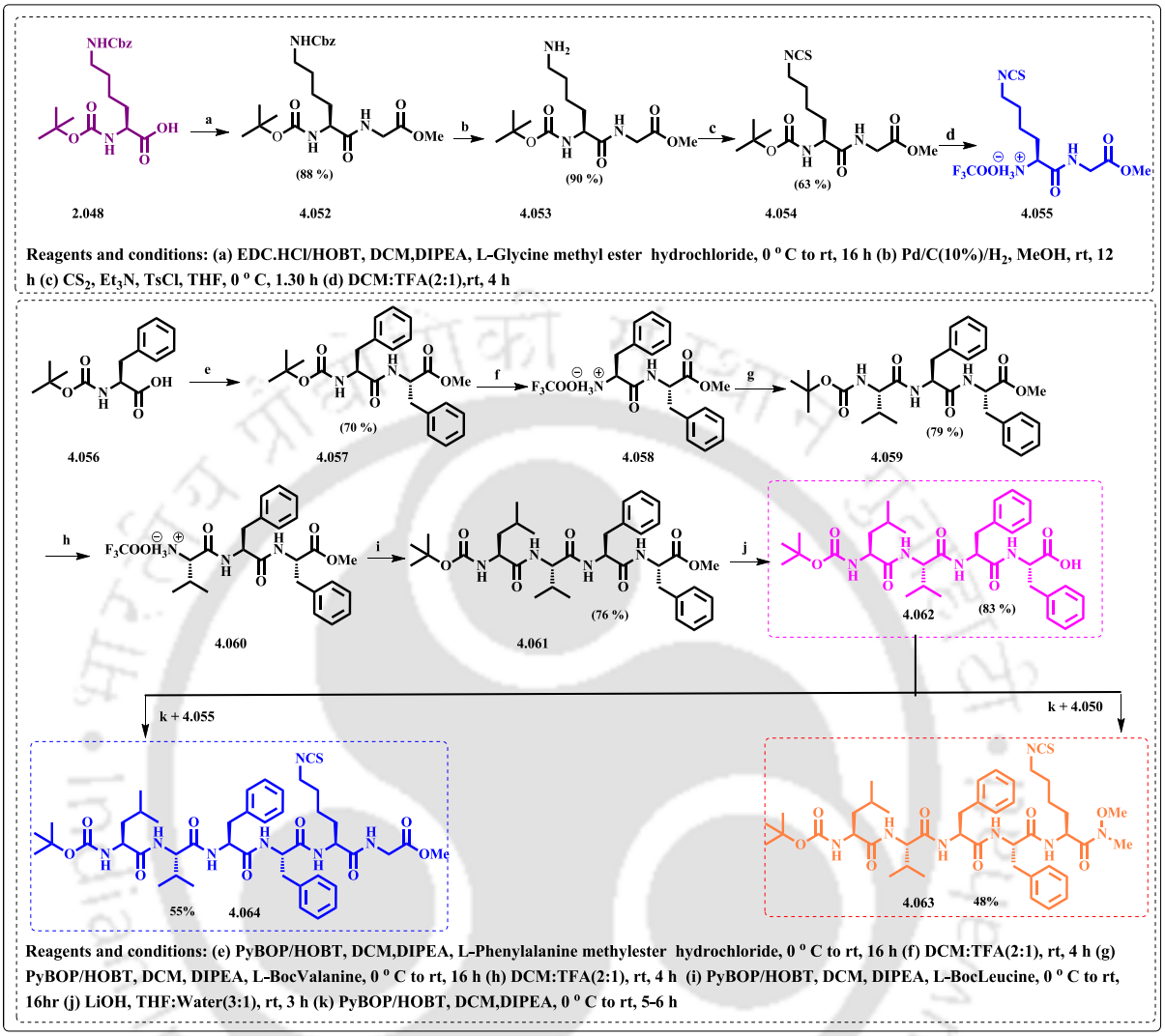
4.5.2. Synthesis of Hexapeptides with ^{NCS}Lys

The high reactivity of ^{NCS}**Ala** led us to synthesise an alternative one i.e. ^{NCS}**Lys** and its incorporation into a short peptide. Thus, we synthesised ^{NCS}**Lys (2.052)** in **Chapter-2** without having any other side product in contrast to ^{NCS}**Ala** where side product was formed. Therefore, compared to the ^{NCS}**Ala,** ^{NCS}**Lys** was found to be less reactive, more efficient and easily incorporable into short peptide without any difficulty. The ^{NCS}**Lys** smoothly underwent reaction with Boc-alanine under EDC/HOBt coupling condition as was evident from a test reaction to give dipeptide **4.051** (**Scheme 4.03**). Thus, the ^{NCS}**Lys** could be incorporated into a peptide at any stage *via* normal peptide coupling protocol allowing peptide/protein functionalisation very easy. Therefore, -NCS functionalisation of **Lys** does not rely on post-synthetic modification, i.e. ^{NCS}**Lys** directly can be utilised for protein/peptide modification which is highly beneficial without getting interfered by any other natural amino acids.



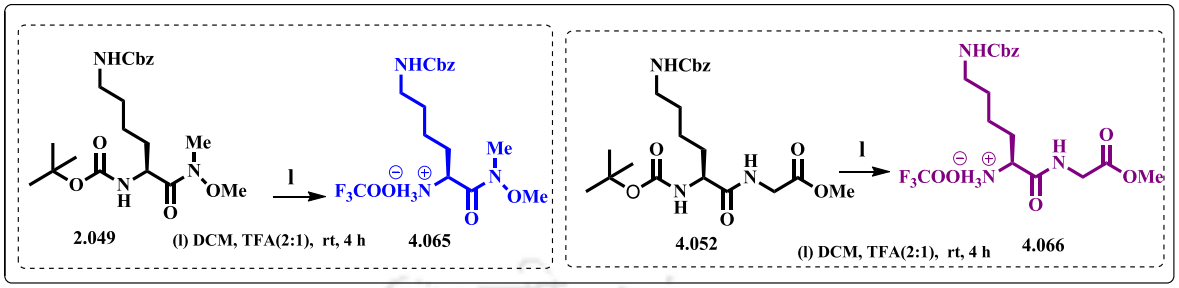
Scheme 4.03. Schematic of formation of dipeptide for peptide coupling compatibility test of ^{NCS}**Lys**.

After getting ^{NCS}**Lys** in the pure form we thought that it would be worthwhile to incorporate it into comparatively large peptide backbone. Thus, we chose a hexapeptide sequence which is a modified sequence of Soto's β-sheet breaker peptide. We ultimately synthesised ^{NCS}**Lys** containing pentapeptide **4.063** BocNH-Leu-Val-Phe-Phe-^{NCS}**Lys**-CO(NMeOMe)] and hexapeptide **4.064** [BocNH-Leu-Val-Phe-Phe-^{NCS}**Lys**-Gly-CO_2Me] following scheme **4.06** *via* a peptide coupling protocol (**Scheme 4.04**). Thus, starting from BocNH-Phe (**4.056**) the tetrapeptide **4.062** was synthesised through the generation of compound **4.057** to **4.061** in a very good yield. A final round of peptide coupling using PyBOP and HOBt mediated solution phase coupling between tetrapeptide **4.062** and TFA salt of ^{NCS}**Lys** (**4.050**) and TFA salt of dipeptide ^{NCS}Lys-Gly (**4.055**) afforded the final targeted pentapeptide **4.063** and hexapeptide **4.064** respectively (**Scheme 4.04**), in a moderate yield. All the intermediates and the final peptides were characterised by IR, NMR and mass spectrometry.

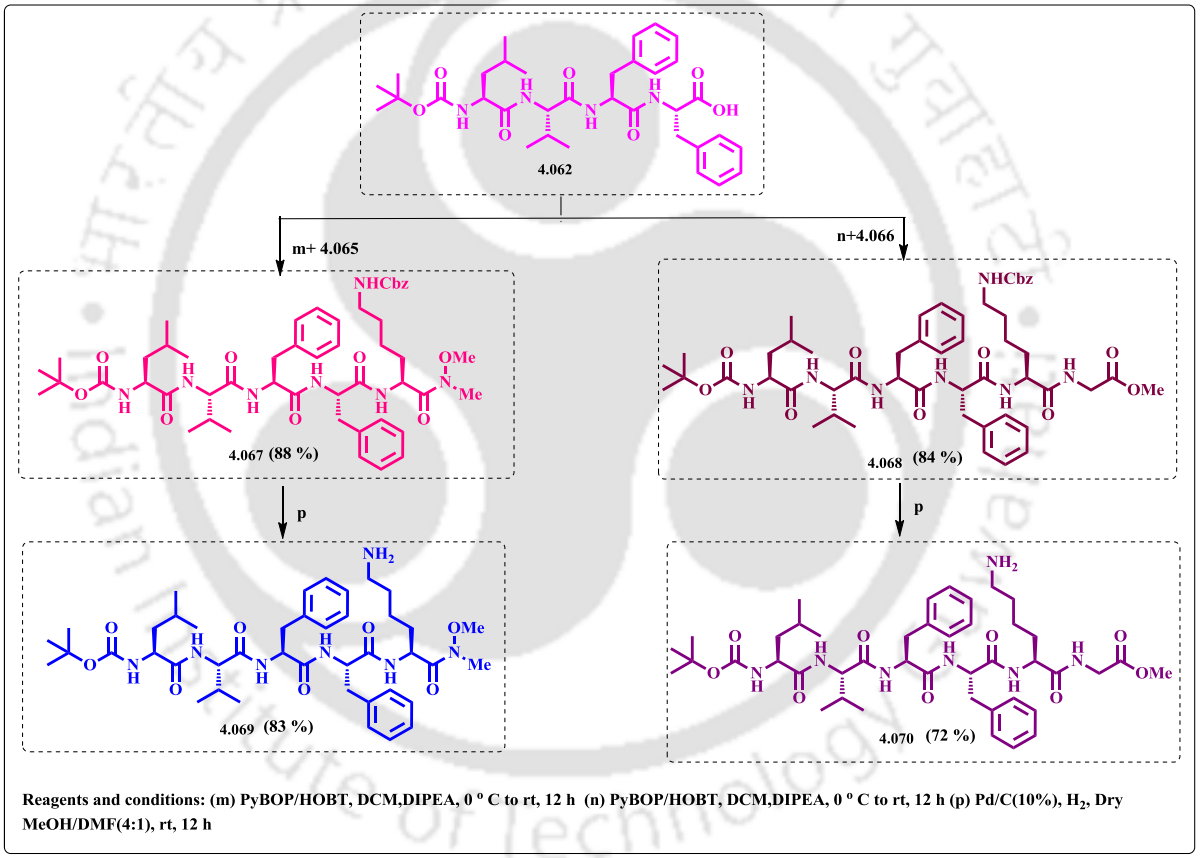


Scheme 4.04. Schematic of formation of ^{NCS}**Lys** containing pentapeptide and hexapeptide.

Similarly in order to prepare natural penta (**4.067**) and hexapeptide (**4.068**) analogues we have first synthesised two corresponding TFA salt of the amino acid (**4.065, Scheme 4.05**) and dipeptide (**4.066, Scheme 4.05**) by treating corresponding Boc-protected amino acid (**2.049, Chapter-2**) and dipeptide (**4.052, Scheme 4.04**) with TFA in dry DCM for 4 hours. Next we performed PyBOP and HOBT mediated solution phase peptide coupling protocol between tetrapeptide (**4.062**) and TFA-salts of **4.065** and **4.066** to afford natural Cbz-protected penta (**4.067, Scheme 4.06**) and hexapeptide (**4.068, Scheme 4.06**) respectively. The corresponding side chain amino pentapeptide (**4.069, Scheme 4.06**) and hexapeptide (**4.070, Scheme 4.06**) were prepared by deprotecting Cbz group using Pd/C under H_2 in dry methanol/DMF (4:1) for 12 hours. All the intermediates and final peptides were purified by silica-gel column chromatography and characterised by IR, NMR and mass spectrometry.



Scheme 4.05. Schematic of formation of TFA salt of amino acid and dipeptide.



Scheme 4.06. Schematic of formation of ^{NHCbz}**Lys** and ^{NH2}**Lys** containing penta and hexapeptide.

4.5.3. ^{1}H-NMR Characterisation of Some Selected Peptides

The synthesized three tripeptides and hexapeptide in the fully protected form were characterised mainly by NMR spectroscopy. As representative example the structural assignments of peptides **4.044** and **4.064** are shown in **Figure 3.13** and **3.14**, respectively. For peptide **4.044** (Boc-NH-Ala-^{NCS}**Ala**-Ala-CO_2Me; **Figure 4.13**) -NH of the N terminal alanine and C-terminal alanine appeared as doublet at δ 7.10 and δ 8.20 with coupling value J = 6.9Hz and 8.1Hz respectively. Another –NH of ^{NCS}**Ala** unit appeared as doublet at δ 8.49 with coupling value J =7.0Hz. The C-α hydrogen of ^{NCS}**Ala** unit appeared at δ 4.01–3.93 as a multiplet. The C-α hydrogen of N-terminal alanine appeared as a double doublet at δ 4.58 with coupling value J = 13.2, 6.4 Hz and C-terminal alanine appeared as pentate at δ 4.27 with coupling value J=7.1Hz. The β-CH_2 hydrogens of ^{NCS}**Ala** unit appeared as a doublet with coupling constants J = 6.2Hz. The methyl hydrogens of OMe groups resonated as singlet at δ 3.62. The methyl hydrogens of alanyl unit appeared as a doublet at δ 1.19 and δ 1.29 with coupling constant J = 7.3 Hz and 7.2Hz respectively. The ^{t}Bu-hydrogens of Boc-protected group appeared as a singlet at δ 1.37.

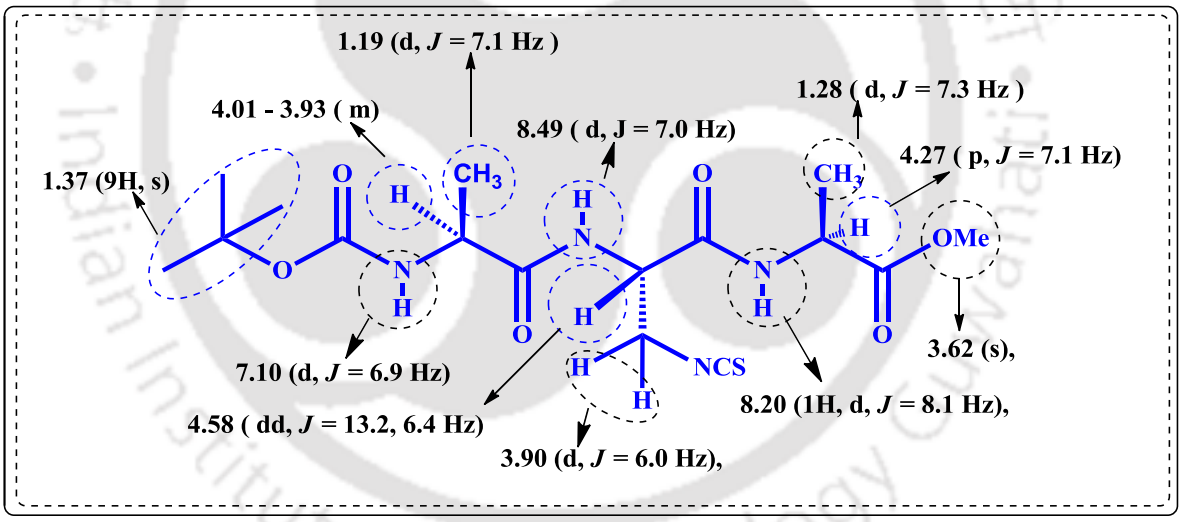


Figure 4.13. ^{1}H-NMR of assignment of ^{NCS}**Ala** tripeptide-**4.044.**

For peptide **4.064** (BocNH-Leu-Val-Phe-Phe-^{NCS}**Lys**-Gly-CO_2Me; **Figure 4.14**). -NH of N-terminal leucine appeared as doublet at δ 7.05 with coupling value J = 8.6 Hz; -NH of valine appeared as doublet at δ 7.45 with coupling value J = 9.0 Hz; two –NH of phenylalanine appeared as multiplet at δ 8.12-8.02; -NH of lysine appeared as doublet at δ 8.14 with coupling value J = 7.9Hz; -NH of glycine appeared as doublet at δ 8.35 with coupling value J = 5.7Hz. The α–hydrogen of two phenylalanine appeared as multiplet at δ 4.61–4.47; α-hydrogen of lysine appeared as double doublet at δ 4.33 with coupling value 13.8, 8.1 Hz; α-hydrogen of valine appeared as double doublet at δ 4.14 with coupling value J = 15.0, 6.9 Hz; α-hydrogen of

leucine and glycine are overlapped and appeared as multiplet at δ 3.96-3.78. δ and ω-hydrogen of lysine overlapped and appeared at δ 3.63 as quintet with coupling value J =6.7 Hz. The all β-CH_2 hydrogens of phenyl alanine appeared at δ 3.02-2.71 as multiplet and double doublet. The β hydrogen of valine appeared as multiplet at δ 1.90–1.79; β-hydrogen of lysine appeared as doublet of triplet at δ 1.64 with coupling value J = 13.8, 7.5 Hz; β-hydrogen of leucine appeared at δ 1.60–1.45 as multiplet; γ-hydrogens of lysine and ^{t}Bu-hydrogens of Boc are overlapped and appeared at 1.43–1.24 as multiplet. Methyl hydrogens of leucine appeared at δ 0.82 as doublet of doublet with coupling value J = 19.5Hz, 6.4Hz. Methyl hydrogens of valine appeared at δ 0.78-0.64 as multiplet.

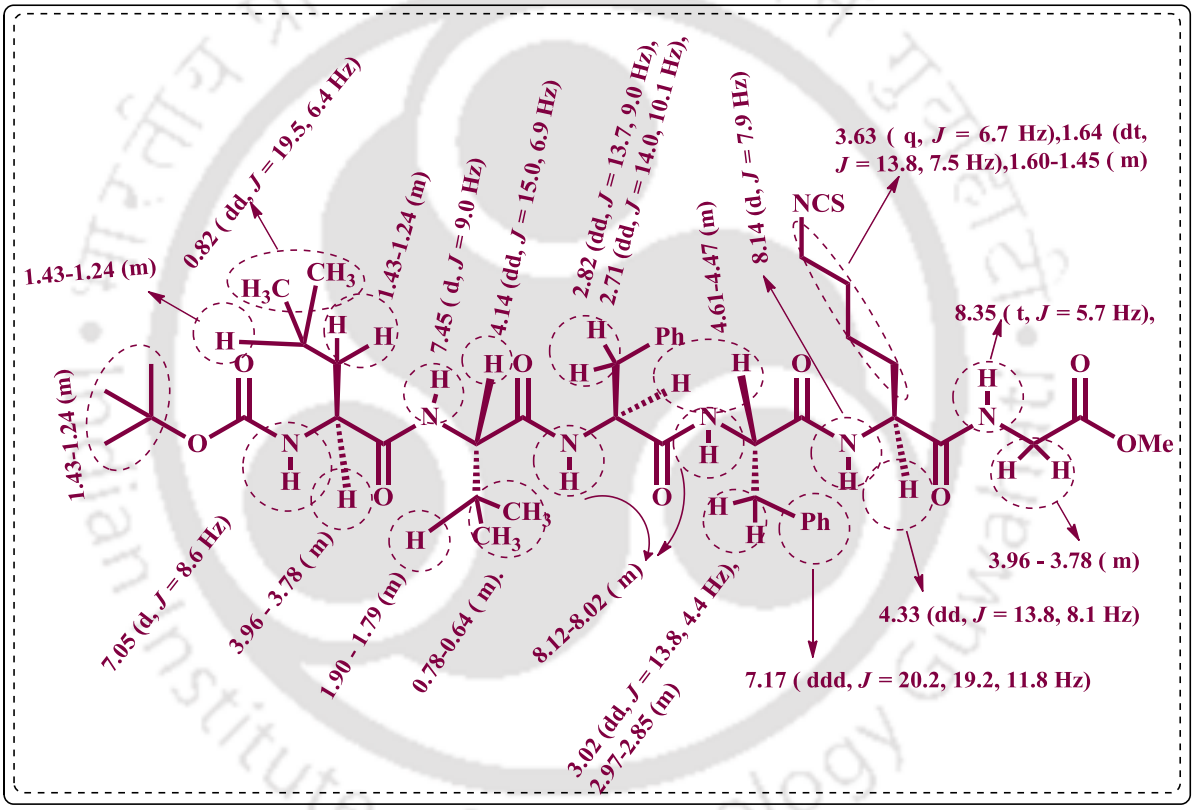


Figure 4.14. ^{1}H-NMR of assignment of ^{NCS}**Lys** Hexapeptide-**4.064.**

4.5.4. Conformational Analysis of Various Synthesised Peptides *via* CD Spectroscopy

After getting all the peptides in pure form, we next carried out the spectroscopic studies on the backbone and side-chain conformation of ^{NCS}**Ala** containing tripeptide **4.044** with the help of NMR, amide-I IR absorption, and circular dichroism (CD) spectroscopy. The conformation based on CD spectroscopic study was compared with that of azidoalanyltripeptide **4.042** and natural tripeptide Ala-Ala-Ala (**4.049**). We also studied the conformation of ^{NCS}**Lys** containing pentapeptide **4.063** [BocNH-Leu-Val-Phe-Phe-^{NCS}**Lys**-CO(NMeOMe)] and hexapeptide **4.064** [BocNH-Leu-Val-Phe-Phe-^{NCS}**Lys**-Gly-CO_2Me] based on CD spectra analysis and compared with their corresponding natural analogue pentapeptide **4.069** [BocNH-Leu-Val-Phe-Phe-**Lys**-CO(NMeOMe)] and hexapeptide **4.070** [BocNH-Leu-Val- Phe-Phe-**Lys**-Gly-CO_2Me].

The CD spectra revealed that the ^{NCS}**Ala** containing tripeptide **4.044** mostly retained its C_7-turn conformation[32] in both ACN and MeOH solvents [(+)195,-199 nm and (-) 223 nm] while the natural trialanyl tripeptide **4.042** was found to adopt a switchable P_{II} conformation[32] [(-) 195-197 nm and (+) 216 nm, (-) 231 nm] upon changing the solvent polarity. The tripeptide **4.042** containing ^{N3}**Ala** behaved in a similar way as that of natural trialanine **4.049**. The blue shifted solvatochromicity (10 nm) of –NCS functionality was also reflected in the CD spectra[32] indicating a strong H-bonding and electrostatic interaction (**Figure 4.15a**). Therefore, -NCS side group had not much effect on the perturbation of the backbone conformation which behaved in a similar way as the natural analogues adopted the conformation.

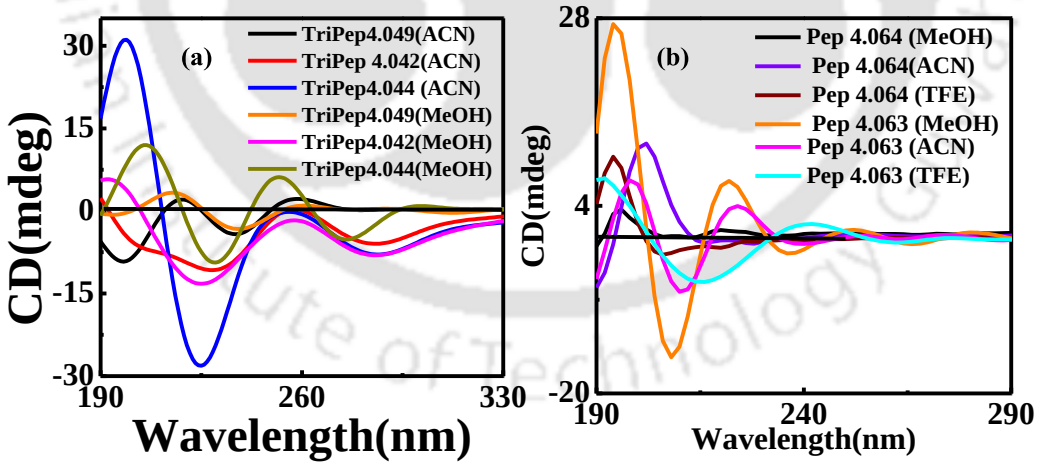


Figure 4.15. CD spectra of (a) tripeptides Ala-Ala-Ala (**4.049**), Ala-^{N3}**Ala**-Ala (**4.042**) and Ala-^{NCS}**Ala**-Ala (**4.044**) and (b) ^{NCS}**Lys** containing pentapeptide 4.063 [BocNH-Leu-Val-Phe-Phe-^{NCS}**Lys**-CO(NMeOMe)] and hexapeptide **4.064** [BocNH-Leu-Val-Phe-Phe-^{NCS}**Lys**-Gly-CO_2Me] in various solvents [concentration of peptides was 60 μM].

On the other hand the CD spectra of ^{NCS}**Lys** containing pentapeptide **4.063** exhibited strong, sharp band at (+) 198-193 and broadband at (-) 210-218 in all three solvents (ACN, MeOH and TFE) (**Figure 4.15b**) similar to its native pentapeptide **4.067/4.069** indicating a mixed β-sheet and α-helical like conformation (**Figure 4.16c-d**). and -NCS side group had no perturbing effect on the conformation. On the other hand, ^{NCS}**Lys** containing hexapeptide **4.064** was found to adopt predominantly α-helical sheet and turn conformation in ACN and MeOH (**Figure 4.15b**). Again the natural analogous hexapeptide **4.068/4.070** also adopted similar conformation (**Figure 4.16a-b**) indicating that the side chain –NCS had no effect on the backbone conformation of the peptides containing ^{NCS}**Lys**. Furthermore, both the peptides containing ^{NCS}**Lys** maintained relatively rigid conformations in all solvents (**Figure 4.15b**) studied even when the polarity was changed. Therefore ^{NCS}**Lys** could act as a very good spectroscopic probe of protein/peptide which would have negligible perturbation effect on the backbone conformation of peptide/protein.

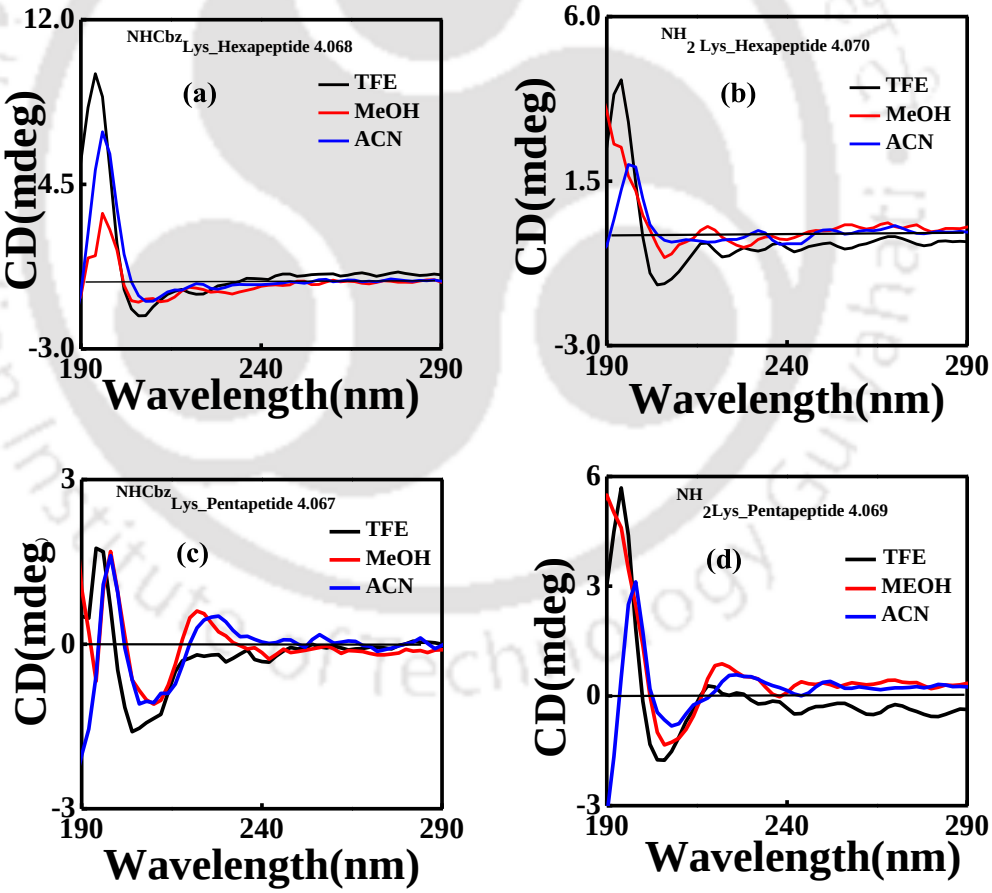


Figure 4.16. CD spectra of natural hexa (a-b) and pentapeptides (c-d) containing ^{CbzNH}**Lys** or ^{NH2}**Lys.**

4.5.5. 2D-NMR Conformational Analysis of Peptides

Next, we carried out ^{1}H-^{1}H NOESY experiment on ^{NCS}**Ala** containing tripeptide **4.044** in d_6-DMSO at 25 ° C. The NOESY spectra of peptide **4.044** (**Figure 4.17**) revealed the presence of cross peak between Hα(i)/NH (i) for all three NHs; all three NH/H_β (of ^{NCS}**Ala**); NH(i, of C-terminal Ala)/NH (i + 1 of ^{NCS}**Ala** and i+2 of N-terminal Ala) indicating a folded C_7 like conformation.[33] Moreover, NH(i)/H_β interactions of each residue and all H_α/H_β (of ^{NCS}**Ala**) are indicative of peptide folded conformation. It was also clear from the NOESY cross peak intensity that the $H^{\beta1}$/ $H^{\beta2}$ and carbamate-NH are very much closer and forming an H-bonded 7-membered ring like C_7 conformation.[33] The presence of long-range backbone interactions accompanying ^{NCS}Ala and side chain interactions are shown in figure **4.19a.**

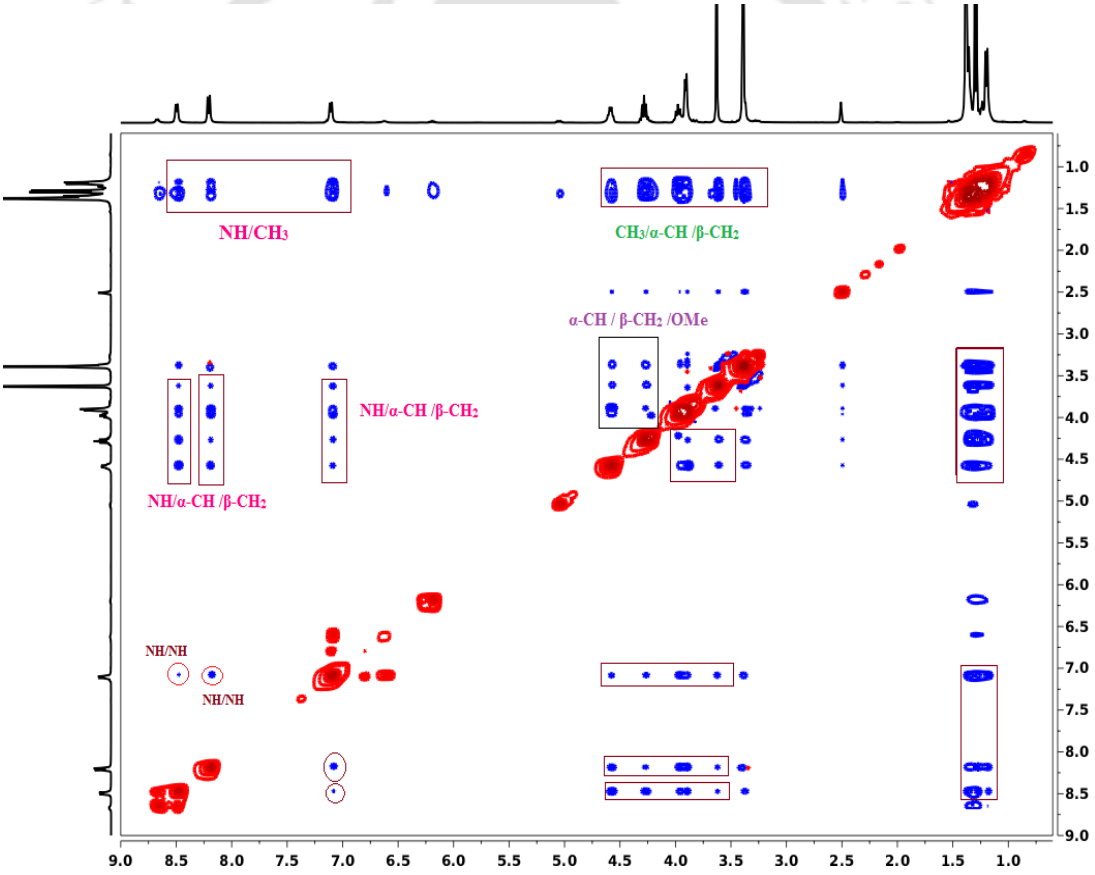


Figure 4.17. The ^{1}H-^{1}H NOESY spectra of tripeptide **4.044.**

On the other hand, the NOESY spectra of peptide **4.064 (Figure 4.18)** revealed the presence of a pattern of consecutive, medium-to-strong H_α(i)/NH(i); H_α(i)/NH(i+1) [Val-Phe]; H_α(i)/NH(i+1) [Phe-Phe]; H_α(i)/NH(i+4) [Gly-Val) NOE connectivites accompanied by NH(i)/NH(i+1) [Leu-Val]; NH(i)/NH(i+1) (Phe-^{NCS}Lys); NH(i)/NH(i+1) [^{NCS}Lys- Gly]; NH(i)/NH(i+2) [Phe-^{NCS}Lys]; NH(i)/NH(i+3)[Val-^{NCS}Lys]; H_α(i)/H_β(i+1) [Leu-Val]; H_β (i)/NH (i+1) [Phe-Phe]; H_β (i)/NH (i+1) [Phe-^{NCS}Lys]; H_β (i)/NH (i+4) [Leu- ^{NCS}Lys] interactions. All these interactions suggested a folded mixed conformation characteristic of α-helix and β-strand. The presence of long range backbone interactions accompanying ^{NCS}Lys and side chain interactions such as H_β, H_γ(i)/H_β, H_ω(i+4), indicated the peptide fold encompasses residues 1 and 5 (**Figure 4.19b**).

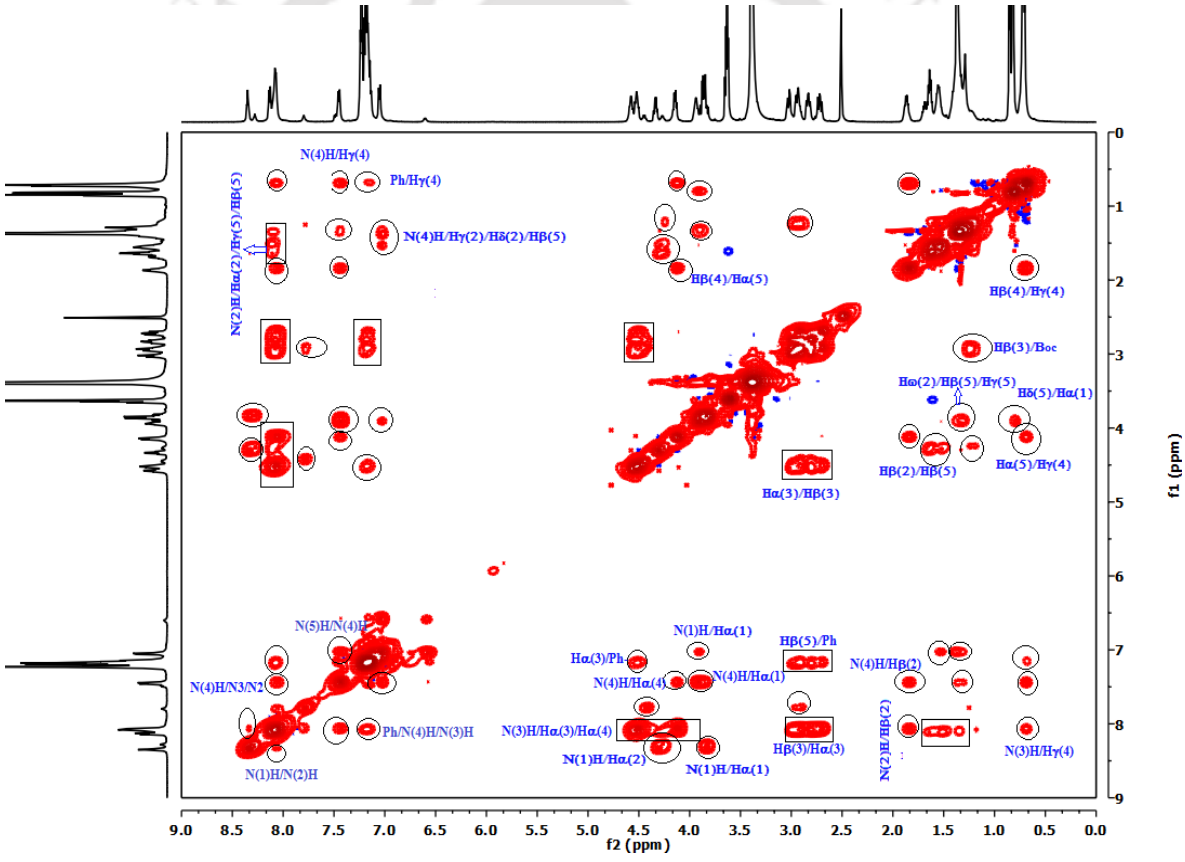


Figure 4.18. The ^{1}H-^{1}H NOESY spectra of hexapeptide **4.064.**

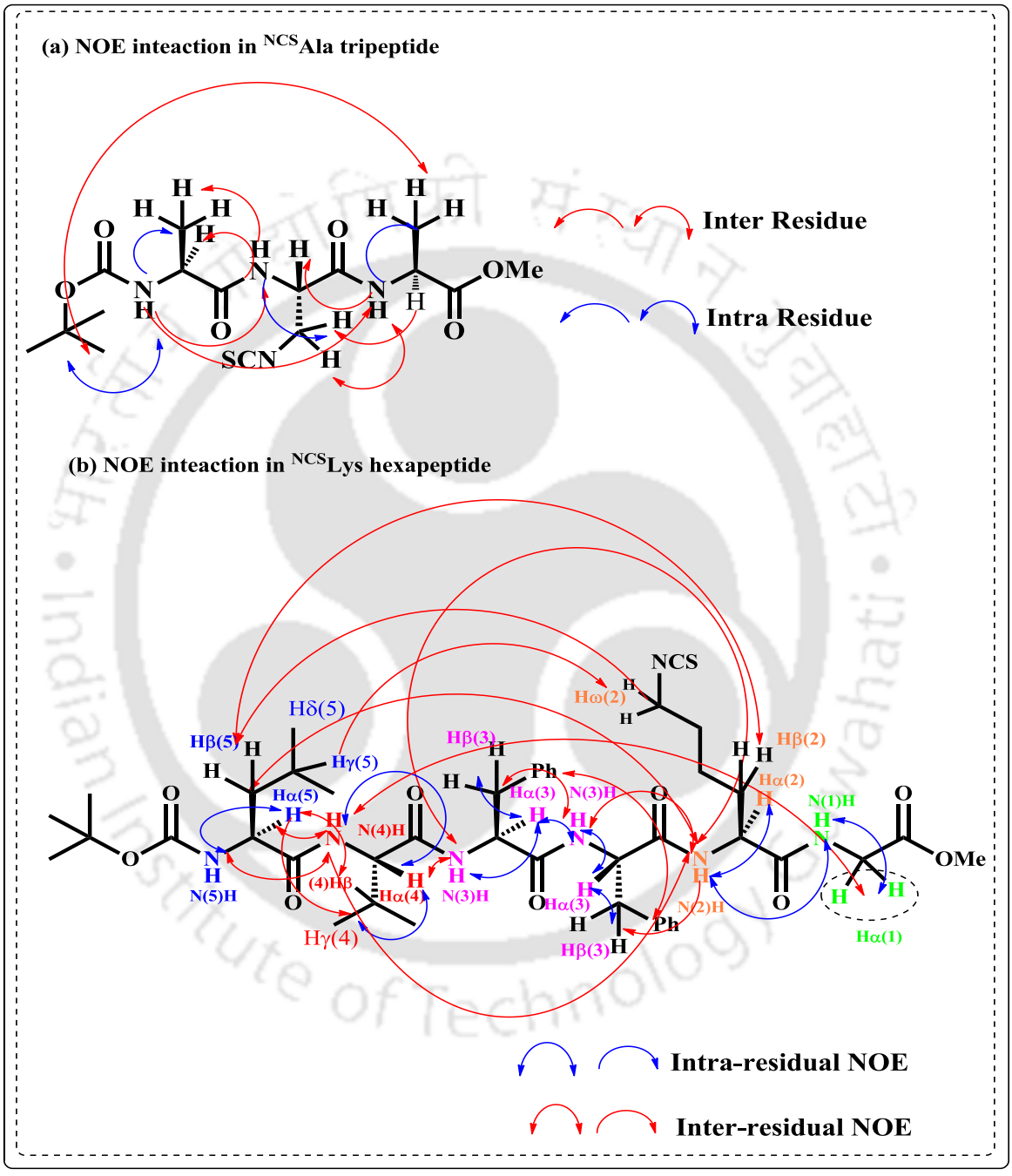


Figure 4.19. 2D-NOE interactions in (a) Boc-Ala-^{NCS}**Ala**-Ala-CO_2Me and (b) BocNH-Leu-Val-Phe-Phe-^{NCS}**Lys**-Gly-CO_2Me from NOESY spectra

4.5.6. Probing the Intramolecular H-Bonding: Study of Variable Temperature ^{1}H-NMR

Inspection of the temperature coefficients of chemical shifts (Δδ/ΔT) in d_6-DMSO from a variable temperature (VT) ^{1}H NMR (**Figure 4.20**) of hexapeptide **4.064** provided some useful information about the H-bonding abilities of the amide protons of each residue. The measured Δδ/ΔT values were spread over the range − 2.8 to − 6.6 ppb/K indicating the involvement of the amide-NH for strong to moderate intramolecular H-bonding ability. The amide-NH groups of residue 2 (Val), 3/4 (Phe/Phe) showed lowest to moderate Δδ/ΔT value of − 2.8 ppb/K and -5.0 ppb/K, respectively, indicating their strong intramolecular H-bonding status. The other residues showed moderate H-bonding ability as was evident from moderate Δδ/ΔT values ranging from -6.4 (^{NCS}**Lys**-NH) to -6.5 (Gly-NH) ppb/K (**Table 4.1**).

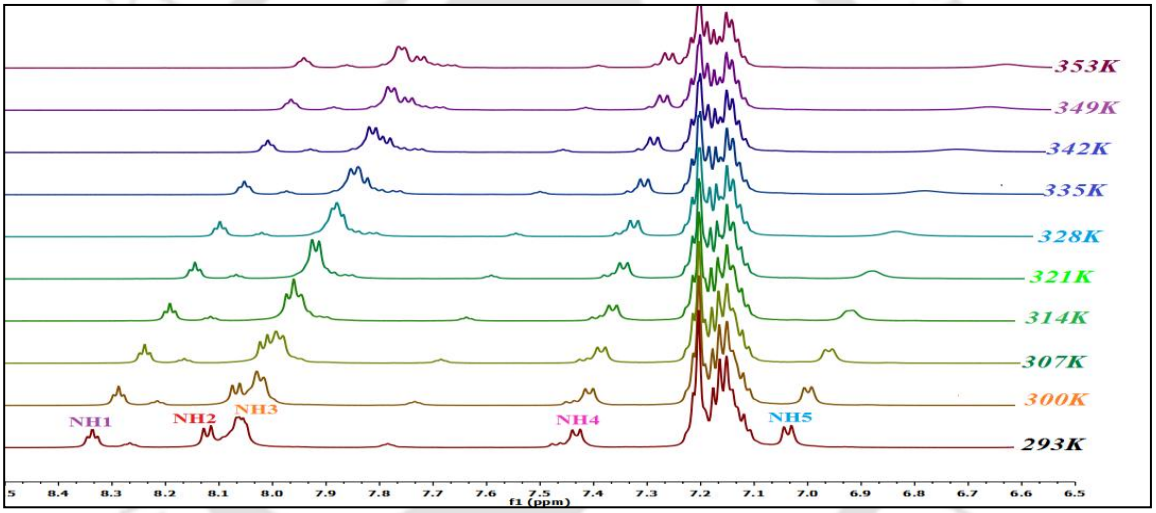


Figure 4.20. Variable temperature NMR of hexapeptide **4.064.**

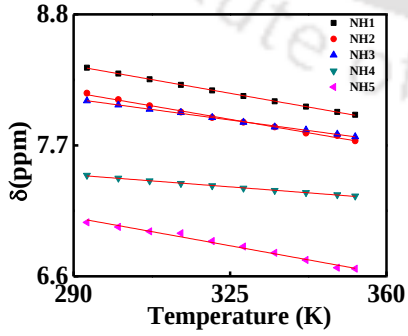
(a) Plot δ vs T	(b) Δδ vs Δ T in ppb/K				
	NH1 (Gly)	NH2 (Lys)	NH3 (Phe/Phe)	NH4 (val)	NH5 (Leu)
	-6.5	-6.4	-5.0	-2.8	-6.6

Table 4.1. (a) Plot of δ vs. T. (b) Values of Δδ/ΔT in ppb/K

4.5.7. IR Spectral Analysis of Backbone Conformation of Peptides

The IR absorption of the amide I vibration appeared at 1750 cm^{-1} and a broad structured band at 1718-1690 cm^{-1} in both ACN and MeOH for peptide **4.044** indicating a *β*-turn conformation with some contribution from *β*-strand (**Figure 4.21a**).[32f-g] For hexapeptide **4.064**, because of its solubility, we could record IR in DMSO and DMSO-dioxane mixed solvent for hexapeptide **4.064**. Thus, in pure 20% DMSO in dioxane the amide-II and I bands appeared at 1526 and 1668 (s), 1651 (w) cm^{-1}, respectively. Increasing the polarity by increasing DMSO content, the bands remained almost unperturbed (**Figure 4.21b**) indicating that the side chain functional group –NCS did not perturb the α-helical sheet-like conformation. The solution IR of the natural hexapeptide containing ^{CbzNH}**Lys 4.068** (**Figure 4.21c-e**) also possess similar IR bands reflecting that there is no conformational perturbation caused by the –NCS side chain of ^{NCS}**Lys** in the hexapeptide **4.064.**

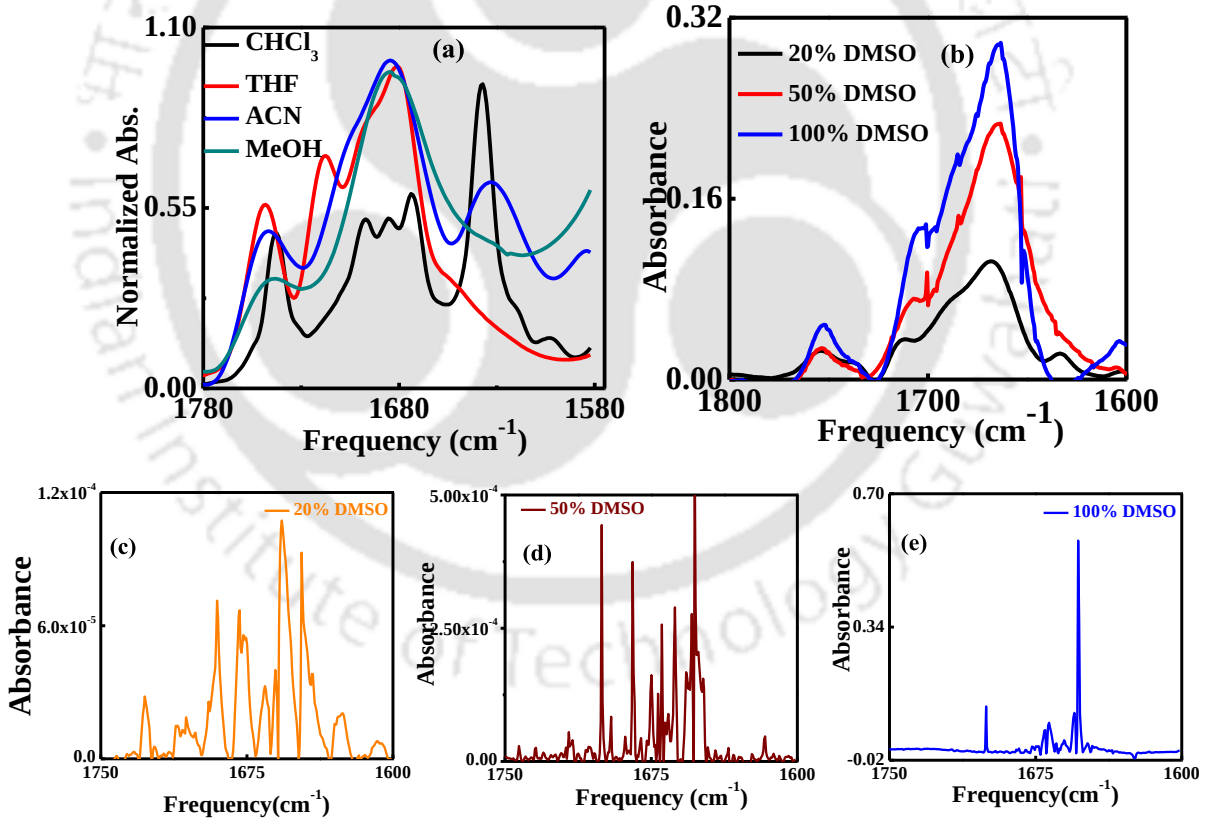


Figure 4.21. Normalized IR spectra of ^{NCS}**Ala** containing tripeptide **4.044** (a) ^{NCS}**Lys 4.064** (b) and ^{NHCbz}**Lys 4.068** containing hexapeptide (c-e) in various solvents. The (a-d) >C=O absorption regions.

4.5.8. Solvatochromic IR Absorptions of –NCS Functionality in Peptides

Next, the effect of solvent polarity sensitive IR absorption of –NCS at 2100-2200 cm^{-1} was studied in the short and long peptides **4.044** and **4.064** containing ^{NCS}**Ala** and ^{NCS}**Lys**, respectively (**Figure 4.22**). Thus, tripeptide **4.044** in the least polar solvent, $CHCl_3$, showed -NCS absorption at 2081cm^{-1}. The $\bar{\upsilon}_{NCS}$ gradually increased (blue-shift) in aprotic solvents with increasing solvent polarity and appeared at 2099 and 2103 cm^{-1} in THF and ACN, respectively. Thus, a 22 cm^{-1} blue shift in $\bar{\upsilon}_{SCN}$ was observed for -NCS as the solvation environment was changed from $CHCl_3$ to ACN (**Figure 4.22a**). Therefore, the shift in IR absorption of –NCS reflected its very high sensitivity toward solvent polarity and thus can be utilised as solvatochromic IR probe for studying protein structure and dynamics. When dissolved in a polar protic solvent like MeOH, $\bar{\upsilon}_{NCS}$ become very lower (2044.8 cm^{-1}) than that in ACN (2103 cm^{-1}) similar to the case of >C=O which showed a red-shift in water compared to that observed in DMSO.[34] The large red-shift of -NCS in MeOH can be rationalised by considering the possible electrostatic and H-bonding interaction between MeOH and –NCS functionality that established a large dipole near the –NCS IR probe. Moreover, the significant splitting of the vibrational bands of –NCS was observed in H-bonding solvent such as MeOH. Thus, along with the intense band weak but prominent other bands appeared at frequencies scattered around 2116 and 2221 cm^{-1} (**Figure 4.22a**).

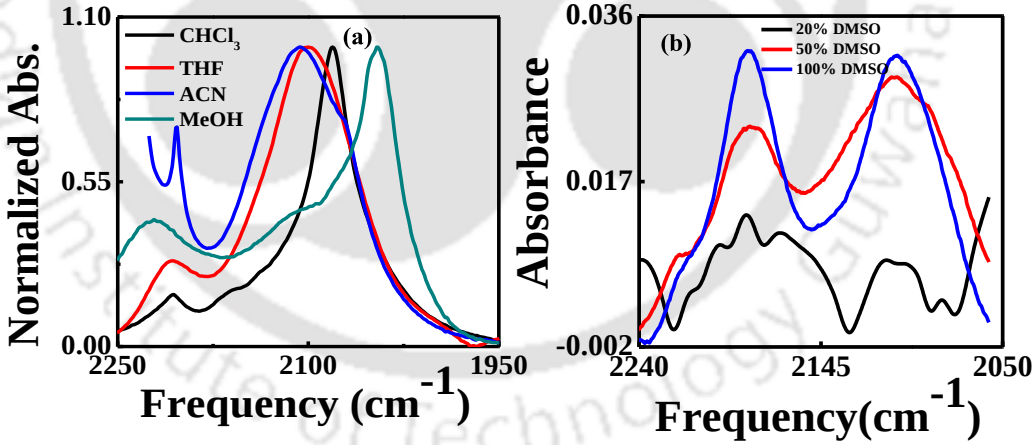


Figure: 4.22. (a) Normalised IR spectra of ^{NCS}**Ala** containing tripeptide **4.044** and (b) ^{NCS}**Lys** containing hexapeptide **4.064** in various solvents. The (a-b)–NCS absorption regions.

Next, we have studied solvent polarity sensitive IR response of –NCS functionality of ^{NCS}**Lys** in hexapeptide **4.064** in DMSO and DMSO-dioxane mixed solvent system. Thus, in pure DMSO, it exhibited IR absorption at 2106.8 and 2183 cm^{-1}. In 50% dioxane in DMSO bands appeared at 2105.7 and 2181.7 cm^{-1}. Both the bands either in pure DMSO or in 1:1 mixture of dioxane/DMSO appeared to be strong and broad with almost equal in intensity

(**Figure 4.22b**). In the least polar solvent 20%, DMSO in dioxane, a medium intense doublet at 2097-2104 cm^{-1} and a quintet like broadband span over 2145-2213 cm^{-1} was the result indicating the solvatochromic nature of -NCS functional group of ^{NCS}**Lys** in hexapeptide **4.064**.

The DFT-calculated IR absorptions were found to be highly sensitive to solvent polarity for ^{NCS}**Ala** containing tripeptide **4.044,** free amino acid, ^{NCS}**Lys** and its hexapeptide **4.064 (Figure 4.23a-f**) supporting the experimental observation and reflecting H-bonding interaction in a polar protic solvent like methanol.

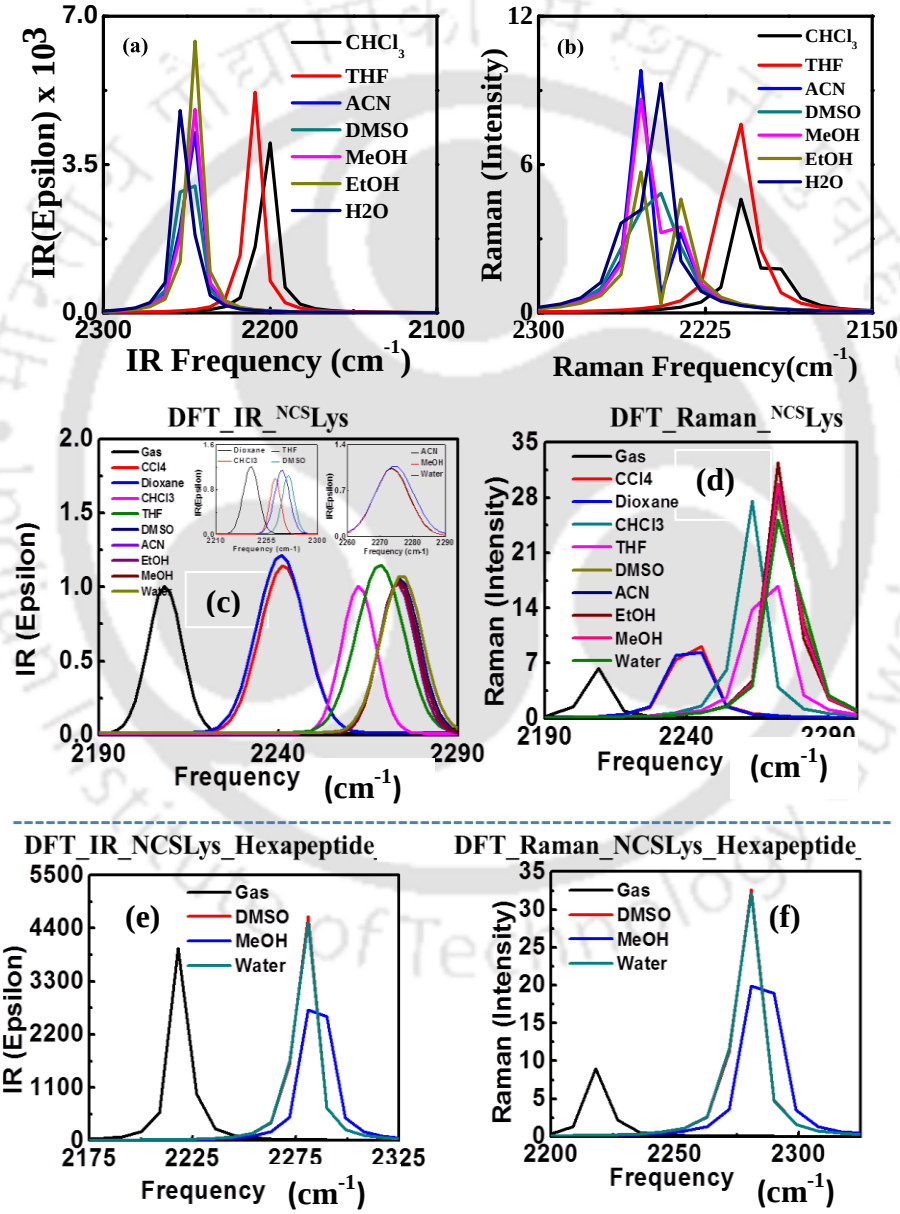


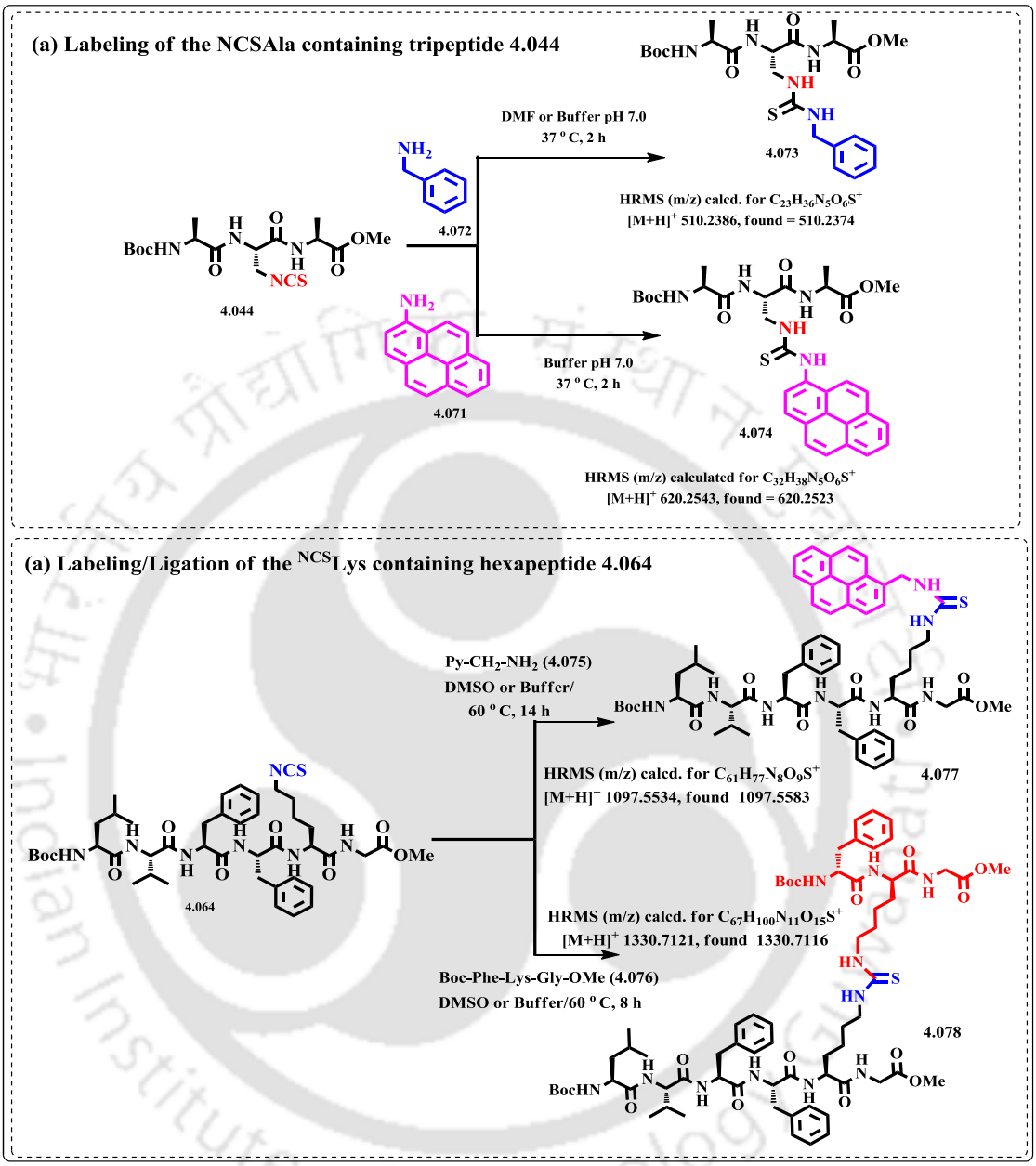
Figure 4.23. Gaussian frequency analysis of ^{NCS}**Ala** containing tripeptide **4.044** (a-b) and ^{NCS}**Lys** (c-d) and ^{NCS}**Lys** containing hexapeptide **4.064** (e-f) in various solvents.

Therefore, the unexplored –NCS can be used as a probe of the local environment in peptide/protein because it's central frequency is sensitive to the polarity and hydrogen bonding ability of local environment. In our analysis, we always observe a blue shift as the solvent polarity increases in an aprotic solvent for both ^{NCS}**Ala** and ^{NCS}**Lys** containing peptides. In particular case of ^{NSC}**Ala** containing tripeptide, we also found that the stretching frequency of -NCS exhibited high sensitivity to H-bonding interactions in a polar protic solvent like MeOH which can be exploited further in detail to utilise it as a specific reporter of H-bonding interactions. The IR intensity was also found to be high and comparable to that of reported -CN or $–N_3$ groups.[35] Further, the -NCS functionality have almost three times stronger IR transition dipole strengths than that of a typical >C=O stretching mode which would show promise to use –NCS as useful IR probe for peptide structures

4.5.9. Fluorescent Labeling/Ligation of the ^{NCS}Ala Containing Tripeptide 4.044 and ^{NCS}Lys Containing Hexapeptide 4.064

Finally, we exploited the electrophilicity of –NCS of reactive tripeptide **4.044** and ^{NCS}**Lys** containing hexapeptide **4.064** to label covalently with chromophoric/fluorophoric amines such as benzyl/pyrenyl or pyrenylmethyl amines as nucleophiles (**Scheme 4.07)**.[36] As a test case, we carried out the labeling both in organic solvent and in a neutral aqueous buffer at room temperature with two separate nucleophiles-namely benzylamines and pyrenylamines for labeling the ^{NCS}**Ala** containing tripeptide **4.044**. Thus, a solution of ^{NCS}**Ala** containing tripeptide in DMF was allowed to react with benzylamine at an incubation temperature of 37 ^{o}C for about 1 h after which the product (benzylthioureayl alanine containing tripeptide, **4.073** formed was isolated (71% yield) and characterised by NMR, IR and Mass spectrometry (**Scheme 4.07a)**. Next, the same reaction was carried out under physiological condition i.e. by incubating the peptide **4.044** (1.2×10^{-5} mM) in 50 mM sodium phosphate buffer of pH 7.0 at 37 ^{o}C with an equivalent amount of amine nucleophiles, benzylamine and/or pyrenylamine for 2 hours.

The long hexapeptide containing ^{NCS}**Lys** (**4.064**) was next labeled with a fluorophore, pyrenylmethylamine (**4.075**) as well as with non-fluorescent tripeptidyl amine **4.076** to showcase the utility of -NCS containing amino acid as a possible labeling/ligation site in a long peptide/protein (**Scheme 4.07b)**. The reaction of ^{NCS}**Lys** containing hexapeptide (**4.064**) with fluorescent pyrenylmethylamine **4.075** and tripeptidy amine **4.076** underwent smoothly at 60 ^{o}C either in DMSO or in buffer to afford pyrenylmethylthioureayl lysine containing fluorescent hexapeptide **4.077** and thioureayl linked bispeptide **4.078** (**Scheme 4.07b**). Thus, ^{NCS}**Lys** can be used for ligation type reaction to afford ligated bis-peptide/protein.[37]



Scheme 4.07. (a) Labeling of ^{NCS}**Ala** containing tripeptide and (b) Labeling/ligation of ^{NCS}**Lys** containing hexapeptide.

Therefore, we demonstrated the efficient, specific, chromogenic labeling of ^{NCS}**Ala** containing tripeptide **4.044** and ^{NCS}**Lys** containing hexapeptide **4.064** exploiting the –NCS functionality. Thus, the unnatural amino acid, ^{NCS}**Ala**/^{NCS}**Lys** if can be incorporated into proteins site-specifically then the generated protein can easily be labeled site-specifically for various biotechnological applications. Furthermore, ^{NCS}**Lys** can be an alternative in generation of thiourea ligated peptide/protein.[13]

4.5.10. The Study of Morphology of Modified Peptide and Natural Analogues

Out of our curiosity and as the ^{NCS}**Lys** containing hexapeptide **4.064** has Phe-Phe dipeptide unit which is the core recognition motif of the Alzheimer's *β*-amyloid peptide[40], we thought that it would be worthwhile to study the morphology of it. We envisaged that depending on the interaction among the hydrophobic side chains and the orientation with the Phe-Phe core the peptide might self-assemble into nanostructures. Thus, the formation of supramolecular structures through molecular self-assembly was primarily accessed by field emission scanning electron microscopy (FESEM) and then transmission electron microscopy (TEM) (with uranyl acetate as the stain) in MeOH with a sample concentration of 10 μM. Our preliminary results showed that the hexapeptide molecules form nanocluster consisting of nanoflakes. It is interesting to note that the nanoclusters were grown with highly uniform size distribution and the average clusters size was about 200 nm (**Figure 4.24a**). From FESEM image it is clear that the spherical in shape nanoparticles of average size 200-400 nm were linked together *via* a nanovesicle of 200nm long forming a helical garland like a network. The nanospherical particle size ranges from 100-200 nm of ^{NCS}**Lys** and ^{NH2}**Lys** hexapeptide was evident from TEM (**Figure 4.24b-c)**.

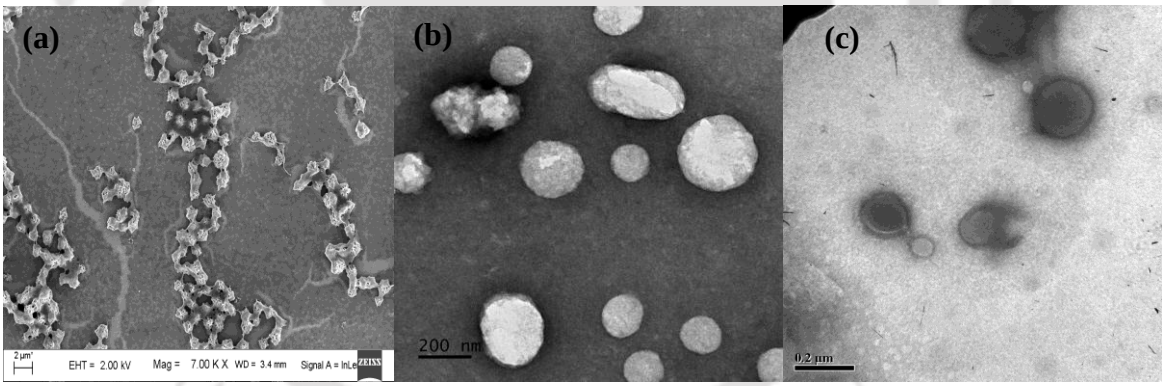


Figure 4.24. (a) FESEM image of ^{NCS}**Lys** hexapeptide **4.064** (b) TEM image of ^{NCS}**Lys** hexapeptide **4.064** (c) TEM image of ^{NH_2}**Lys** hexapeptide **4.070.**

The crystalline nature of the peptide nanoflakes was confirmed from the X-ray diffraction measurement on the powder sample of the hexapeptide showing three diffraction peaks (**Figure 4.25a-b**). Our results suggested that the initial formation of aggregates was possibly promoted by hydrophobic interactions of as well as backbone conformation and hydrogen-bonding interactions. More details morphological study is our future target.

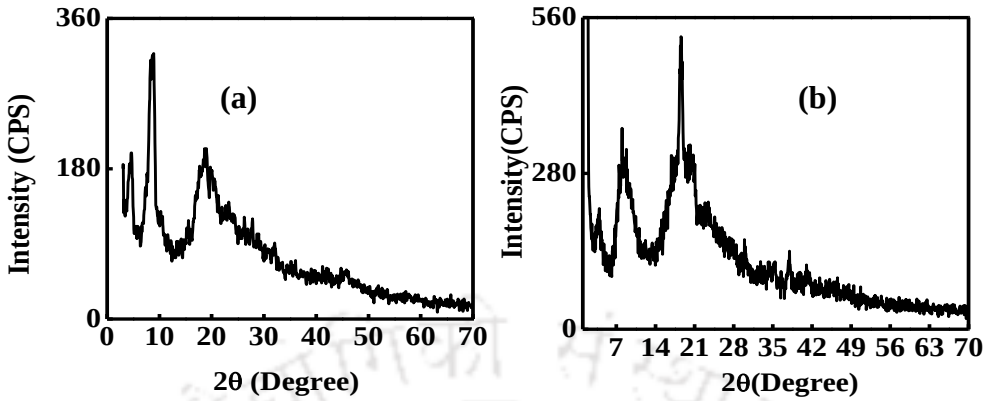


Figure 4.25. (a) PXRD of hexapeptide **4.064** and (b) PXRD of hexapeptide **4.068**.

4.5.11. Study of Photophysical Properties of Pyrenylmethylthioureayl Lysine Labeled Hexapeptide

We also recorded the fluorescence of the pyrenylmethylthioureayl lysine labeled hexapeptide **4.077** to test the photophysical outcome and we observed that in the peptide the chromophore retained its pyrene like absorption with good solvatochromicity and emission property. Thus, the UV-visible spectra showed absorptions at 316, 330 and 346 in less polar solvent $CHCl_3$. In aprotic polar solvent, ACN, it showed a hypsochromic shift with hypochromic effect (311, 327 and 343 nm) while the entire bands experienced hypsochromic shift with hyperchromic effect in polar protic solvent MeOH (312, 326 and 342 nm). When excited at 345 nm, the long wavelength absorption band of the chromophore, the peptide exhibited characteristic pyrenyl emission bands at 377, 397, 417 nm with almost similar intensity in all solvents except in DMSO wherein an enhanced emission was observed (**Figure 4.26a-b**).

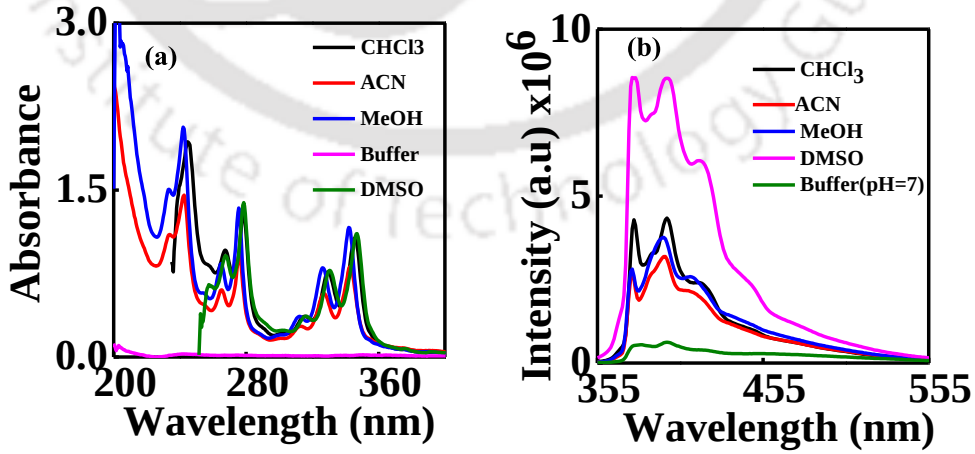


Figure 4.26. (a)UV-visible and (b) fluorescence emission spectra of fluorescent pyrenylmethylthioureayl lysine amino acid containing hexapeptide **4.077** in various organic solvents (conc.10 μM).

4.5.12. DFT and Macromodel Optimised Geometries of ^{NCS}Ala Containing Tripeptide and ^{NCS}Lys Containing Hexapeptide

We also carried out a DFT optimisation of the tripeptide containing **^{NCS}Ala** at B3LYP/6-31G** level of theory in water using CPCM solvent model in G09 program package-45 and found the ϕ, *χ1* and ψ angles as -87, 61 and 63.8 degrees, respectively, in relation to the central side chain modified amino acid **^{NCS}Ala**. All these results indicated that the backbone conformation of the peptide **4.044** was close to C_7, γ-turn conformation both in d6-DMSO from experiment and in water from theory.[32f,39] The DFT optimisation at B3LYP/6-31G** level of theory in water as solvent (CPCM solvent model) using G09 program package-45 which also supported an H-bonded (2.05 Å) 7-membered ring conformation (**Figure 4.27a**) in relation to **^{NCS}Ala** of the peptide **4.044**.

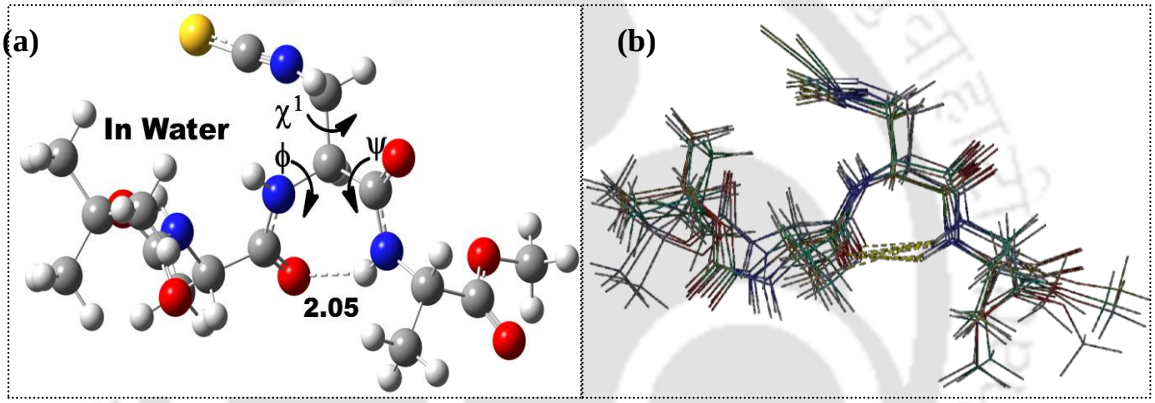


Figure 4.27. (a) DFToptimised structure and (b) Clustering of structures (within 21 kJ/mole global minima) obtained from molecular dynamics simulation for the tripeptide Ala-**^{NCS}Ala**-Ala (**4.044**)

For hexapeptide **4.064**, the ^{3}J (HN, H_α) was came out to be 7.9 Hz and it showed helical sheet-like conformation which was also supported from Gaussian optimised geometry (**Figure 4.28**). Next, we carried out MD simulations using Schrodinger Macromodel software in water for the peptide **4.044** and **4.064** using an OPLS 2005 force field and taking global minimum conformer as the starting structure for the peptide.[38] The MD simulation of the tripeptide **4.044** (**Figure 4.27b**) containing **^{NCS}Ala** amino acid in the middle of the backbone fully supported the H-bonded C_7-turn conformation with backbone H-bond involving >CO at i and NH at i+2 supporting the experimental as well as DFT calculated optimised structure. MD simulation (**Figure 4.29**) for **^{NCS}Lys** amino acid containing hexapeptide **4.064** showed H-bonded helical sheet-like conformation with backbone and side chains torsion ϕ, ψ, *χ1* and *χ2* angles as -85, -19, 79 and -174 degrees, respectively, in relation to the side chain modified amino acid **^{NCS}Lys** supporting the other experimental observations.

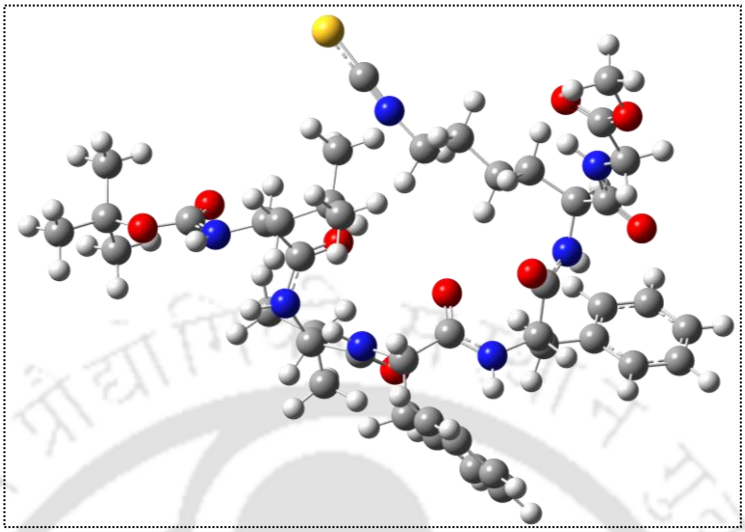


Figure 4.28. DFT optimised structure for the hexapeptide **4.064** containing ^{NCS}**Lys**.

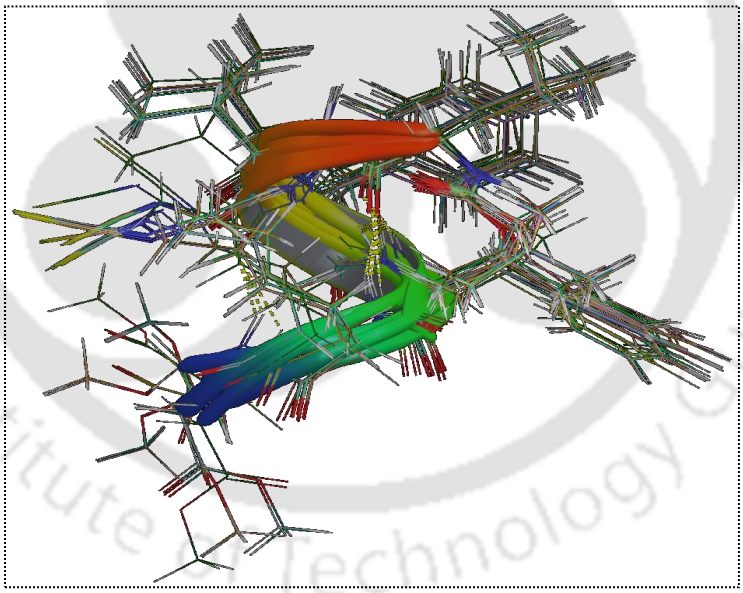
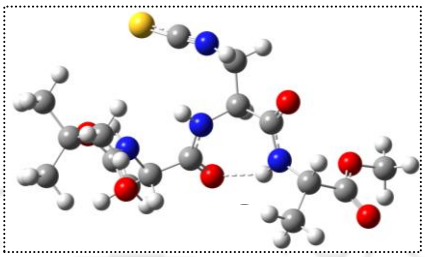


Figure 4.29. Clustering of structures (within 21 kJ/mole global minima) obtained from molecular dynamics simulation for the hexapeptide **4.064** containing ^{NCS}**Lys**.

4.5.13. DFT Calculation

4.5.13.1. Geometry in water as representative example and cartesian coordinates of 4.044

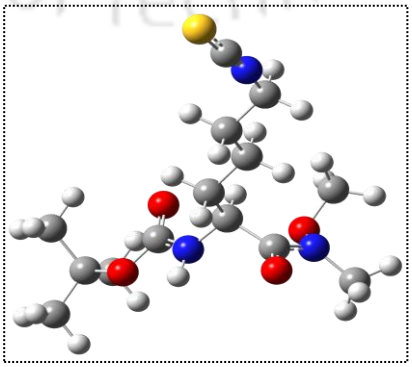


E(RB3LYP) = -1694.04271134 au; Imaginary Freq. = 0; Dipole Moment = 2.8157 D

Center Number	Atomic Number	Atomic Type	Coordinates (Angstroms) X	Y	Z
1	6	0	1.224533	-0.114078	1.862275
2	6	0	-0.138735	0.327985	1.307071
3	8	0	-1.148318	0.331690	2.024335
4	7	0	-0.162617	0.701200	0.008778
5	6	0	-1.347220	1.273703	-0.616855
6	6	0	-0.957989	2.239474	-1.745978
7	6	0	-2.304623	0.190229	-1.170501
8	7	0	-2.818952	-0.611412	-0.210654
9	8	0	-2.552337	0.093578	-2.373361
10	6	0	-3.743848	-1.692267	-0.497212
11	6	0	-5.184702	-1.386171	-0.063823
12	8	0	-6.039634	-2.244122	0.048376
13	6	0	1.488422	0.566975	3.211609
14	6	0	-3.267811	-3.022797	0.097294
15	8	0	-5.404504	-0.080332	0.134803
16	6	0	-6.754431	0.283658	0.497201
17	7	0	-0.090874	3.277476	-1.277797
18	6	0	0.785781	3.706496	-0.604035
19	16	0	1.958464	4.389534	0.252073
20	7	0	2.320376	0.103209	0.923735
21	8	0	4.211872	-0.679201	0.003035
22	6	0	5.095245	-1.672485	-0.655494
23	6	0	4.466290	-2.151763	-1.966953

24	6	0	6.355279	-0.849217	-0.927463
25	6	0	5.397356	-2.817329	0.313346
26	6	0	2.962829	-1.005133	0.403606
27	8	0	2.419544	-2.100152	0.309325
28	1	0	1.156662	-1.199618	1.997562
29	1	0	0.699028	0.611454	-0.516359
30	1	0	-1.882015	1.824275	0.164365
31	1	0	-0.462077	1.703142	-2.560726
32	1	0	-1.862021	2.687715	-2.161146
33	1	0	-2.563589	-0.410703	0.754842
34	1	0	-3.781234	-1.773653	-1.589473
35	1	0	0.637367	0.408877	3.875585
36	1	0	2.385777	0.148108	3.672421
37	1	0	1.629112	1.645954	3.088157
38	1	0	-2.305894	-3.300294	-0.338974
39	1	0	-3.148396	-2.945750	1.182673
40	1	0	-3.996093	-3.807647	-0.112390
41	1	0	-7.450194	-0.002834	-0.293826
42	1	0	-7.042298	-0.206204	1.429362
43	1	0	-6.741215	1.364946	0.621837
44	1	0	2.869909	0.945349	1.040614
45	1	0	5.178270	-2.795873	-2.491625
46	1	0	3.551936	-2.717785	-1.787498
47	1	0	4.237954	-1.299733	-2.614404
48	1	0	7.106400	-1.471931	-1.421194
49	1	0	6.779059	-0.471087	0.007110
50	1	0	6.129230	0.001448	-1.576754
51	1	0	6.136557	-3.486481	-0.136219
52	1	0	4.499950	-3.390064	0.543440
53	1	0	5.815610	-2.422835	1.244399

4.5.13.2. Geometry in water and cartesian coordinates of ^{NCS}Lys

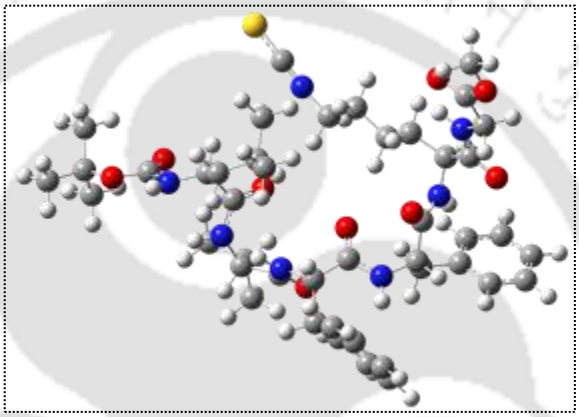


E(RB3LYP) = -1411.91998053 a.u.; Imaginary Freq = 0; Dipole Moment = 8.0098 Debye

Center Number	Atomic Number	Atomic Type	Coordinates (Angstroms) X	Y	Z
1	7	0	-1.722440	-0.069116	-0.691344
2	1	0	-1.826709	-0.208322	-1.686928
3	6	0	-0.537086	0.627038	-0.204436
4	1	0	-0.691742	0.775198	0.862618
5	6	0	0.747112	-0.181982	-0.449265
6	6	0	-0.487018	1.974134	-0.946276
7	1	0	0.593890	-1.177127	-0.016401
8	1	0	0.878577	-0.310502	-1.529363
9	8	0	-0.068530	2.043681	-2.103028
10	7	0	-0.915711	3.088790	-0.279265
11	6	0	-1.196044	4.345413	-0.959835
12	1	0	-2.249426	4.399272	-1.254777
13	1	0	-0.960476	5.182470	-0.299775
14	1	0	-0.563952	4.392051	-1.844683
15	8	0	-1.669810	2.900918	0.890898
16	6	0	-0.953634	3.386212	2.040176
17	1	0	-0.721124	4.451182	1.943023
18	1	0	-1.633482	3.235879	2.880373
19	1	0	-0.031125	2.818788	2.197902
20	8	0	-3.502983	-1.385700	-0.599435
21	6	0	-4.567167	-2.192329	0.031515
22	6	0	-3.949750	-3.361107	0.804955
23	1	0	-3.375696	-3.006550	1.661229
24	1	0	-4.745305	-4.021810	1.162827
25	1	0	-3.292430	-3.943658	0.152036
26	6	0	-5.442540	-1.301315	0.916940
27	1	0	-5.831045	-0.456342	0.340105
28	1	0	-6.293896	-1.880010	1.288223
29	1	0	-4.878576	-0.920337	1.768500
30	6	0	-5.356820	-2.699996	-1.177054
31	1	0	-4.717325	-3.296839	-1.833854
32	1	0	-6.188619	-3.325959	-0.842288
33	1	0	-5.762991	-1.863673	-1.753312
34	8	0	-2.476247	-0.755195	1.361529
35	6	0	-2.566649	-0.748142	0.139136
36	6	0	2.001618	0.459930	0.156137
37	6	0	3.254579	-0.393504	-0.073929
38	1	0	1.853958	0.608854	1.234163

39	1	0	2.156110	1.454421	-0.281289
40	6	0	4.498990	0.257521	0.547573
41	1	0	3.119126	-1.389058	0.363724
42	1	0	3.424822	-0.531929	-1.147588
43	1	0	4.674569	1.248342	0.114526
44	1	0	4.367958	0.390061	1.627126
45	7	0	5.674127	-0.526407	0.341284
46	6	0	6.537054	-1.286148	0.092663
47	16	0	7.751925	-2.300552	-0.227700

4.5.13.3.Geometry in water and cartesian coordinates of 4.064



E(RB3LYP) = -3173.37472359 a.u.; Imaginary Freq = 0; Dipole Moment = 11.6465 Debye

---

Center Number	Atomic Number	Atomic Type	Coordinates (Angstroms) X	Y	Z
1	7	0	1.223094	-2.409822	-0.996204
2	6	0	-0.153668	-2.614260	-1.428187
3	6	0	-1.125127	-1.927181	-0.429212
4	6	0	-0.403202	-4.119086	-1.734541
5	6	0	-1.616835	-4.450912	-2.583671
6	7	0	-2.369919	-2.453493	-0.339205
7	8	0	-0.790337	-0.920457	0.193264
8	6	0	1.683726	-2.800215	0.217329
9	6	0	-2.588541	-5.351000	-2.119904
10	6	0	-3.695934	-5.683404	-2.906208
11	6	0	-3.847165	-5.121125	-4.173368
12	6	0	-2.882455	-4.229101	-4.651336
13	6	0	-1.779156	-3.899093	-3.865217
14	6	0	-3.378862	-1.931083	0.576783
15	6	0	-3.590633	-0.418263	0.307655
16	7	0	-3.843336	0.362936	1.382812

17	8	0	-3.587743	0.019741	-0.842563
18	6	0	-3.056779	-2.371702	2.030965
19	6	0	-4.181064	-2.247711	3.043612
20	6	0	-3.984778	-1.548880	4.244553
21	6	0	-5.004190	-1.451874	5.196551
22	6	0	-6.238537	-2.057305	4.962814
23	6	0	-6.445317	-2.762914	3.773620
24	6	0	-5.427142	-2.857590	2.825671
25	8	0	0.950300	-3.312680	1.064158
26	6	0	3.214504	-2.649697	0.426245
27	6	0	3.628606	-2.674125	1.926543
28	7	0	3.747501	-1.525230	-0.349741
29	6	0	3.461555	-0.219789	-0.100537
30	8	0	2.651098	0.146977	0.751117
31	6	0	4.155703	0.795124	-1.031615
32	6	0	3.496579	-4.095411	2.500775
33	6	0	5.066839	-2.175455	2.128029
34	7	0	5.496386	0.366148	-1.418741
35	6	0	6.556217	0.550870	-0.568518
36	8	0	7.711577	0.322756	-1.222302
37	6	0	3.327518	1.030638	-2.314809
38	6	0	1.917783	1.623174	-2.126520
39	6	0	1.967289	3.057814	-1.581253
40	6	0	1.160632	1.578911	-3.462673
41	8	0	6.426913	0.859752	0.610365
42	6	0	9.020973	0.401729	-0.537803
43	6	0	10.001074	0.061804	-1.662192
44	6	0	9.253648	1.824921	-0.023921
45	6	0	9.089063	-0.644533	0.577598
46	6	0	-4.079132	1.798616	1.272027
47	6	0	-5.297790	2.053674	0.351313
48	6	0	-2.797356	2.591695	0.945847
49	7	0	-5.233368	3.127480	-0.468871
50	6	0	-6.330680	3.467075	-1.346217
51	1	0	-4.386351	3.665951	-0.589460
52	6	0	-1.618069	2.221419	1.857238
53	8	0	-6.304111	1.348495	0.426439
54	6	0	-0.439224	3.189954	1.700885
55	6	0	0.805136	2.678263	2.443774
56	6	0	-5.916814	4.620905	-2.238642
57	8	0	-4.820935	5.148804	-2.211536

58	7	0	1.869968	3.631686	2.421580
59	6	0	2.655819	4.505891	2.396844
60	16	0	3.761355	5.684311	2.367612
61	8	0	-6.912956	4.974872	-3.054713
62	6	0	-6.634976	6.065821	-3.961567
63	1	0	1.876462	-2.002476	-1.647805
64	1	0	-0.259652	-2.048047	-2.359903
65	1	0	-0.438347	-4.668108	-0.790143
66	1	0	0.493555	-4.462564	-2.262295
67	1	0	-2.618608	-3.257269	-0.897631
68	1	0	-2.472283	-5.802118	-1.137570
69	1	0	-4.434699	-6.382950	-2.526544
70	1	0	-4.706040	-5.376480	-4.786586
71	1	0	-2.988355	-3.792347	-5.640033
72	1	0	-1.032120	-3.213984	-4.257517
73	1	0	-4.321342	-2.399520	0.273838
74	1	0	-3.906505	-0.052371	2.299816
75	1	0	-2.172547	-1.830042	2.377337
76	1	0	-2.764849	-3.425631	1.960090
77	1	0	-3.021059	-1.085112	4.441302
78	1	0	-4.828521	-0.906248	6.118915
79	1	0	-7.032679	-1.984201	5.699622
80	1	0	-7.401179	-3.243135	3.585771
81	1	0	-5.601863	-3.420550	1.912635
82	1	0	3.650921	-3.539883	-0.045052
83	1	0	2.948155	-2.009353	2.469048
84	1	0	4.543183	-1.699596	-0.949775
85	1	0	4.228002	1.717079	-0.448698
86	1	0	2.483453	-4.480283	2.387525
87	1	0	3.753206	-4.095865	3.565413
88	1	0	4.189253	-4.780239	1.995080
89	1	0	5.204517	-1.134820	1.828368
90	1	0	5.337449	-2.258045	3.185969
91	1	0	5.778660	-2.786760	1.559305
92	1	0	5.712845	0.285340	-2.403202
93	1	0	3.908229	1.706268	-2.957029
94	1	0	3.253699	0.076548	-2.852870
95	1	0	1.370280	1.000843	-1.407884
96	1	0	0.958304	3.475969	-1.502826
97	1	0	2.423698	3.104854	-0.588390
98	1	0	2.543865	3.709363	-2.249258
99	1	0	0.147491	1.978471	-3.349749

100	1	0	1.078447	0.555013	-3.843467
101	1	0	1.672289	2.179262	-4.224772
102	1	0	11.025790	0.084954	-1.281205
103	1	0	9.802291	-0.937491	-2.059946
104	1	0	9.918740	0.784593	-2.478915
105	1	0	10.268720	1.905186	0.376768
106	1	0	8.544481	2.080100	0.763832
107	1	0	9.152478	2.545751	-0.841035
108	1	0	8.866500	-1.639497	0.179941
109	1	0	8.384180	-0.416190	1.377393
110	1	0	10.100042	-0.664115	0.995836
111	1	0	-4.420067	2.112462	2.265706
112	1	0	-2.513319	2.421202	-0.097858
113	1	0	-3.019317	3.660081	1.061641
114	1	0	-7.225744	3.755508	-0.782250
115	1	0	-6.620086	2.615315	-1.973390
116	1	0	-1.945030	2.218742	2.905821
117	1	0	-1.296424	1.201769	1.621114
118	1	0	-0.188848	3.316482	0.641895
119	1	0	-0.708301	4.179091	2.088526
120	1	0	0.569358	2.467928	3.493292
121	1	0	1.179669	1.755065	1.988819
122	1	0	-7.548390	6.204779	-4.536689
123	1	0	-5.802742	5.808001	-4.619118
124	1	0	-6.392792	6.970790	-3.401334

4.6. Conclusion

The ^{NCS}**Ala** has been incorporated into a tripeptide sequence while ^{NCS}**Lys** in a hexapeptide sequence *via* a peptide coupling protocol. While ^{NCS}**Ala** containing tripeptide adopted a C_7-turn conformation, a predominant helical sheet-like conformation was evident for the hexapeptide containing ^{NCS}**Lys**. It was further observed that the side chain –NCS does not perturb the backbone conformation of the hexapeptide which was evident from IR, CD, NMR and powder XRD analysis of both the modified and native peptide analogues. The –NCS functionality is also shown to have the solvatochromic IR sensitivity which further can be utilised as IR sensitive spectral probe to investigate the structure, dynamics, localisation, and biomolecular interactions. In particular, the strong IR transition dipole strength of -NCS compared to reported >C=O and comparable to -CN/N_3 group, would show promise as useful IR probes for peptide structures. Furthermore, the high sensitivity to H-bonding interactions in protic polar solvent could allow the –NCS as a highly specific reporter for the same in a biomolecular system. However, a rigorous theoretical calculation is needed to elucidate the molecular origin of H-bonding and IR stretching shift. The electrophilicity of isothiocyanyl alanine (^{NCS}**Ala = Ita**) and ^{NCS}**Lys** in a short tri-/hexa-peptides has further been exploited to label covalently with chromophoric/fluorophoric amines such as benzyl/pyrenylamines/pyrenylmethylamines specific ligation site in a hexapeptide, thus could be utilised as alternative to cysteine for chemical ligation of peptides/proteins to offer thiourea linked bis-peptides/proteins. Therefore, ^{NCS}**Ala**/^{NCS}**Lys** amino acids would be beneficial over the available nine canonical amino acids for site-specific labeling and in ligation of peptide or protein if incorporated enzymatically or studying peptides conformation and dynamics which is our current research focus.

4.7. Experimental section

4.7.1. General Experimental

Reactions were carried out under the nitrogen atmosphere and few in air using oven-dried round bottom flasks. Reagents were purchased at the highest commercial quality and used without further purification unless otherwise stated. Reactions were monitored by thin-layer chromatography (TLC) carried out on a 0.25 mm silica gel 60F–254 and silica gel-G (1:4) and visualized under UV light at 254 nm. Further visualisation was achieved by iodine vapour adsorbed on silica gel depending on the product type. Organic extracts were dried over anhydrous sodium sulphate. Solvents were removed in a rotary evaporator under reduced pressure. Column chromatography was performed on silica gel 60–120 mesh using a mixture of hexane and ethyl acetate as a mobile solvent, an isolated compounds were characterised by 1D-NMR (^{1}H , ^{13}C) and 2D-NMR (COSY, TOCSY, NOESY), HRMS, and IR spectroscopic technique. All NMR spectra were recorded at ambient temperature on Bruker Ascend ^{TM} Aeon 600 MHz and Bruker Avance II, 400 MHz spectrometers where ^{1}H frequency were 600MHz, 400MHz and ^{13}C frequency were 150 MHz and 100MHz. NMR spectra for the samples were measured in deuterochloroform ($CDCl_3$) and deuterated dimethyl sulfoxide-d_6 (DMSO-d_6). The chemical shifts value are quoted in δ units, parts per million (ppm) up field from the signal of internal TMS. The chemical shift values were reported in ppm downfield from tetramethylsilane, using chloroform-d (δ = 7.26 for ^{1}H NMR, δ = 77.23 for ^{13}C NMR), using deuterated dimethyl sulfoxide-d_6 (δ = 2.50 for ^{1}H NMR, δ = 39.50 for ^{13}C NMR). ^{1}H NMR data is represented as follows: Chemical shift, multiplicity (s = singlet, d = doublet, dd = double doublet, ddd = Doublet of doublet of doublets, dt= Doublet of triplet, t = triplet, q = quartet, p = Pentate, h=heptate, m = multiplet), integration and coupling constant(s) *J* in Hertz (Hz). High-resolution mass spectra (HRMS) were recorded on a Mass spectrometer in positive mode using electrospray ionisation-time of flight (ESI-TOF) reflection experiments. IR spectra of the dry compound were recorded in KBr plate in Perkin Elmer spectrometer and reported in frequency of absorption (cm^{-1}). Solution phase IR spectra recorded on 32mm drilled and undrilled KBr plate from Sigma-Aldrich. All the NMR-FID was processed in MestReNova v6.0.2. Software.

4.7.2. Synthesis and Characterisations

General solution phase peptide coupling procedure: To a solution N-protected amino acids/peptides in dry DCM or 3:1 mixture of dry DCM and DMF, 1-[3-dimethyl amino propyl]-3-ehtylcarbo-diimide hydrochloride (EDC.HCl) (1.4 eqv) and HOBT (1.2 eqv) or PyBOP (1.2 eqiv) and HOBt (1.4 eqiv) were added and the reaction mixture was stirred for 1 h at 0 ^{o} C. Then the amine salt of wienreb amide or methyl ester protected corresponding amino acids or peptide (1.1 eqiv) were added followed by diisopropylethylamine (DIPEA) (2.4-3.0 eqiv). The reaction mixture was stirred for another 6-20 h at 0 ^{o} C to room temperature. Afterward, the reaction mixture was extracted with ethyl acetate three times in presence of aqueous 1N HCl (30ml) and saturated $NaHCO_3$ solution (50 ml each). Then the organic layer was washed with brine solution and dried over Na_2SO_4. After evaporation of organic layer, the residue was purified by silica-gel column chromatography (60-120) using hexane/ethylacetate as eluent to afford the pure product.

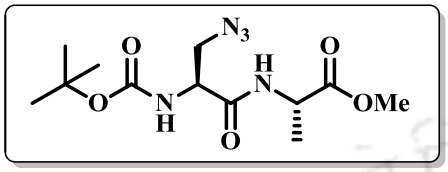
General procedure for tert-butoxycarbonyl (Boc) Deprotection: To solution BOC-protected amino acids/peptides in dry DCM, TFA was added in 1:2 ratios. It was allowed to stir at room temperature for about 4 hours. Then the reaction mixture was dried under high vacuum. In order to obtain excess TFA free compound, the chloroform (20 ml, 3 times) and toluene (20 ml, 2 times) were added sequentially to the reaction mixture and co-evaporated to obtain a gummy compound. The compound was directly used for next reaction without any purification.

General procedure for conversion of azidoalanine (N_3) to aminoalanine (NH_2) or carboxybenzyl (Cbz) deprotection: To a solution of N, C-terminal protected amino acids/peptides in dry MeOH, Pd/C (10 moles %) 0.3 equivalent was added under nitrogen atmosphere. Then the reaction mixture was degassed for 5 min under high vacuum and hydrogen atmosphere was created. Afterward, reaction was stirred in the H_2 atmosphere for 7 hours for azide to amine conversion and 12 hours for Cbz deprotection at room temperature. After completion of the reaction, the reaction mixture was dried under vacuum and directly used in next step without any further purification.

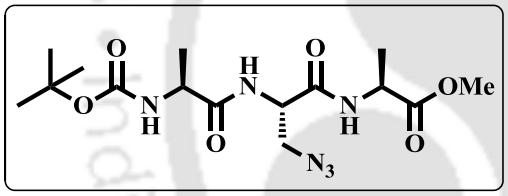
(S)N,C-diprotected isothiocyanyl lysine (2.052): Synthesised and described in Chapter 2

(S)$^{\omega}$N, $^{\alpha}$N, C-triprotected lysine (2.049): Synthesised and described in Chapter 2.

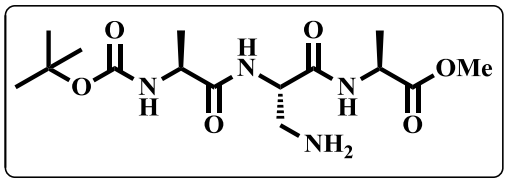
(S,S) N, C-diprotected azidoalanyl dipeptide (4.039): Following the general procedure of peptide coupling protocol, starting from 1000 mg (4.3 mmol) of tert-butoxycarbonyl-protected azido alanine acid **4.038** and HCl salt of alanyl methylester 900mg (6.45mmol) which were synthesised in **chapter-2**, the dipeptide **4.039** was obtained. A semi solid compound; 1100mg, Yield 81 %. IR (KBr) ῡ 3468, 3320, 2105, 1751, 1690, 1656, 152, 1449, 1391, 1292, 1215, 1166, 1055, 986, 928, 729 cm^{-1}. ^1H NMR (600 MHz; $CDCl_3$) δ 7.01 (1H, s), 5.45 (1H, s), 4.55 (1H,dd, J = 14.3, 7.1 Hz), 4.32 (1H, s), 3.78 (1H, dd, J = 12.3, 5.3 Hz), 3.73 (3H, s), 3.52 (1H, dd, J = 11.7, 4.7 Hz), 1.43 (9H, s), 1.40 (3H,dd, J = 7.1, 2.2 Hz); ^{13}C NMR (150 MHz; CDCl3) δ 173.08, 169.23, 155.5, 80.9, 53.8, 52.7, 52.4, 48.4, 28.3, 18.3.+ESI-HRMS (m/z) calculated for $C_{12}H_{22}N_5O_5$ $[M+H]^+$ 316.1621, found 316.1611.



(S, S, S) N,C-diprotected azidoalanyl tripeptide tripeptide (4.042): Following the general procedure of peptide coupling protocol, starting from 4.49 mmol of tert-butoxycarbonyl-protected alanine acid and TFA-salt of azidoalanyl dipeptide-**4.040** (6.29 mmol), the tripeptide **4.042** was obtained. A white solid compound ;1250mg, Yield 72 %; m.p. 155-159 ^{o} C. IR (KBr) ῡ 3321, 3276, 2984, 2933, 2108, 1740, 1701, 1673, 1693, 1522, 1451, 1393, 1368, 1272, 1252, 1222, 1167, 1086, 1051, 1029 cm^{-1}. ^1H NMR (600 MHz; $CDCl_3$) δ 7.19 (1H, d, J = 4.0 Hz), 7.09 (1H, d, J = 7.6 Hz), 5.09 (1H, d, J = 5.0 Hz), 4.63 (1H, dt, J = 8.0, 5.0 Hz), 4.55 (1H, p, J = 7.2 Hz), 4.18 (1H, s), 3.91–3.82 (1H, m), 3.74 (3H, s), 3.54 (1H, dd, J = 12.4, 5.3 Hz), 1.44 (9H, s), 1.42 (3H, d, J = 7.2 Hz), 1.39 (3H, d, J = 7.1 Hz); ^{13}C NMR (150 MHz, $CDCl_3$) δ 173.0, 172.9, 168.7, 155.9, 80.9, 52.7, 52.5, 52.0, 50.8, 48.6, 28.4, 18.0 .+ESI-HRMS (m/z) calculated for $C_{15}H_{27}N_6O_6$ $[M+H]^+$ calculated 387.1992, found 387.1990.

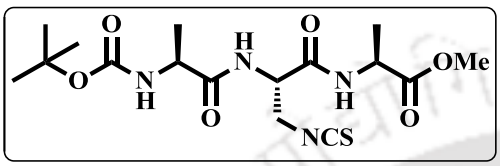


(S,S,S) N,C-diprotected aminoalanyl tripeptide (4.043): Using the general procedure for the synthesis of amino alanine, starting from azidoalanyl tripeptide-**4.042** (422 mg, 1.09 mmol), and the desired aminoalanyl tripeptide **4.043** containing amino alanine was obtained. A gummy material; 355mg, Yield 90 %. IR (KBr) ῡ 3444, 2981, 2928, 2038, 1670, 1533, 1455, 1368, 1323, 1299, 1251, 1167 cm^{-1}. ^1H NMR (400 MHz; DMSO-d_6) δ 8.52–8.27 (1H, m), 8.00–7.79 (1H, m), 7.13–6.89 (1H, m), 4.70–4.50 (1H, m), 4.26 (2H, ddd, J = 28.5, 14.3, 7.3 Hz), 3.90 (3H, s), 3.67–3.53 (2H, m), 3.31 – 2.90 (2H, m), 1.36 (9H, s), 1.31–1.22 (3H, m), 1.21–1.12 (3H, m). ^{13}C NMR (150 MHz; $CDCl_3$) δ 174.9, 173.6, 169.6, 156.3,

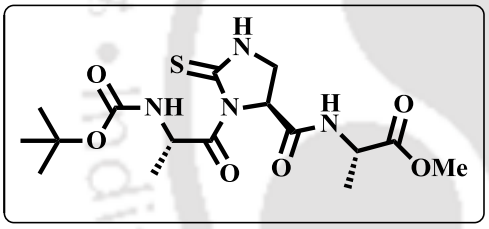


80.3, 52.7, 51.8, 50.86 48.7, 41.5, 28.4, 18.0, 17.1. +ESI-HRMS (m/z) calculated for $C_{15}H_{29}N_4O_6$ $[M+H]^+$ 361.2087, found 361.2071.

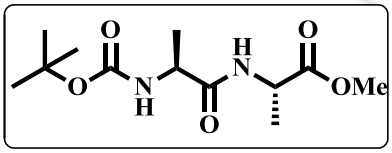
(S.S,S) N,C-diprotected isothiocyanylalanine tripeptide(4.044): Using the procedure for the synthesis of ^{NCS}**Ala** (**Chapter-2**), starting from aminoalanyl tripeptide **4.043** (355 mg, 0.88mmol), the desired tripeptide, BocNH-Ala-^{NCS}**Ala**-Ala-CO2Me (**4.044**) was obtained (R_f = 0.5 in 1:1, in hexane/ethyl acetate). A white solid material; 150mg, Yield 42 %; m.p. 153-157 ^{o} C. IR (KBr) ῡ 3322, 3271, 3006, 2983, 2943, 2213, 2106, 1742, 1700, 1675, 1644, 1521, 1452, 1393, 1369, 1300, 1253, 1216, 1167, 1109, 1080, 1049 cm^{-1}. ^{1}H NMR (400 MHz; DMSO-d_6) δ 8.49 (1H, d, J = 7.0 Hz), 8.20 (1H, d, J = 8.1 Hz), 7.10 (1H, d, J = 6.9 Hz), 4.58 (1H, dd, J = 13.2, 6.4 Hz), 4.27 (1H, p, J = 7.1 Hz), 4.01 – 3.93 (1H, m), 3.90 (2H, d, J = 6.0 Hz), 3.62 (3H, s), 1.37 (9H, s), 1.28 (3H, d, J = 7.3 Hz), 1.19 (3H, d, J = 7.1 Hz); ^{13}C NMR (100 MHz; DMSO-d_6) δ 173.2, 172.5, 167.9, 155.2, 129.5, 78.3, 52.1, 51.6, 50.0, 47.8, 46.2, 28.2, 17.9, 16.8. +ESI-HRMS (m/z) calculated for $C_{16}H_{27}N_4O_6S$ $[M+H]^+$ 403.1651, found 403.1622.



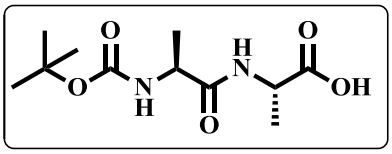
As a side product we also obtained (R_f = 0.4 in 1:2, hexane/ethyl acetate) the cyclic thioureayl peptide **4.045.** A light greenish liquid compound; Yield 14 %. IR (KBr) ῡ 3327, 2979, 2926, 2854, 2106, 1740, 1681, 1523, 1454, 1388, 1367, 1216, 1165, 1047 cm^{-1}.^{1}H NMR (600 MHz; $CDCl_3$) δ 7.73 (1H, d, J = 25.9 Hz), 6.91 (1H, s), 6.49 – 6.37 (1H, m), 5.24 (1H, s), 5.16 (1H, t, J = 6.6 Hz), 4.54 (1H, p, J = 7.2 Hz), 3.83 (2H, d, J = 7.4 Hz), 3.74 (3H, s), 1.47 (3H, s), 1.43 – 1.39 (12H, m). ^{13}C NMR (150 MHz; $CDCl_3$) δ 179.5, 176.6, 173.2, 168.1, 155.4, 80.1, 60.8, 52.8, 49.0, 48.6, 44.6, 28.5, 19.3, 18.1. +ESI-HRMS (m/z) calculated for $C_{16}H_{27}N_4O_6S$ $[M+H]^+$ 403.1651, found 403.1604.



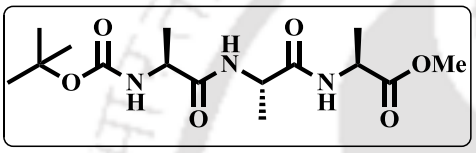
(S,S)N,C-diprotected Ala-Ala dipeptide (4.047): Using EDC.HCl/HOBt mediated solution phase peptide coupling protocol , starting from 500mg (2.64 mmol) Boc-alanine-**4.046** and methylester alanine salt the desired dipeptide-**4.047** was obtained (R_f= 0.5 in 1:1, hexane/ethyl acetate). A Semi solid compound; 510mg, Yield 70 %. IR (KBr) ῡ 3323, 3090, 2984, 2939, 2878, 2850, 2789, 1745, 1683, 1655, 1553, 1524, 1455, 1392, 1371, 1324, 1252, 1222, 1110, 1167, 1071 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 6.78 (1H, s), 5.13 (1H, d, J = 4.3 Hz), 4.54 (1H, p, J = 7.2 Hz), 4.18 (1H, s), 3.72 (3H, s), 1.42 (9H, s), 1.38 (3H, d, J = 7.2 Hz), 1.34 (3H, d, J = 7.1 Hz).^{13}C NMR (150 MHz; $CDCl_3$) δ 173.4, 172.4, 155.6, 80.2, 52.6, 50.1, 48.1, 28.4, 18.4. ESI-HRMS (m/z) calculated for $C_{12}H_{23}N_2O_5$ $[M+H]^+$ 275.1607, found 275.1606.



(S,S)C-diprotected Ala-Ala dipeptide acid (4.048): At first **compound-4.047** 400mg (1.45mmol) was loaded in 100 mL single neck R.B and THF/Water (3:1) was added to it. After that $LiOH.H_2O$ 92mg (2.18mmol) was added to the stirred solution and it was allowed to stir for about 2 hours at room temperature. The reaction mixture was dried under high vacuum, 5-6 ml water was added, and it was acidified with about 1N HCl until the pH 2-3.Then product was extracted with Ethyl acetate and was dried over Na_2SO_4. After Evaporation of ethyl acetate, we got semi solid compound- 4.053 and it was used directly for the next step without purification. A semi solid compound; 365mg, Yield 96 %.

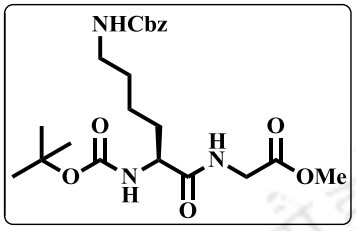


(S,S,S) N, C-diprotected trialanyl tripeptide (4.049): Following the general procedure of peptide coupling protocol, starting from 1.4 mmol of tert-butoxycarbonylprotected alanyl alanine dipeptidic acid **4.048** and TFA-salt of alanine methyl ester dipeptide *(*1.54 mmol), the *N*, *C*-diprotected trialanyl tripeptide **4.049** was obtained (R_f = 0.4 in 1:3, hexane/ethyl acetate) as solid compound. A white solid; 300mg, Yield 62 % ; m.p. 184-188 ^{o} C. IR (KBr) ῡ 3279, 3080, 2980, 2933, 1746, 1697, 1640, 1524, 1451, 1391, 1367, 1336, 1217, 1168, 1052, 1027, 957, 862 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 7.48 (1H, t, J = 8.9 Hz), 7.39 (1H, d, J = 5.5 Hz), 5.62 (1H, d, J = 6.0 Hz), 4.56 (1H, d, J = 5.9 Hz), 4.46 (1H, p, J = 7.2 Hz), 4.21 (1H, s), 3.64 (3H, s), 1.33 (9H, s), 1.31 – 1.23 (9H, m). ^{13}C NMR (150 MHz; $CDCl_3$) δ 173.1, 173.0, 172.1, 155.5, 79.7, 52.3, 50.0, 48.7, 48.0, 28.3, 28.3, 18.8, 18.5, 17.8. +ESI-HRMS (m/z) calculated for $C_{15}H_{28}N_3O_6$ $[M+H]^+$ 346.1978, found 346.1963.



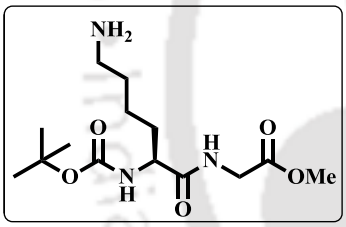
(S,S)N,C-diprotected Ala-^{NCS}Lys dipeptide (4.051):Using EDC.HCl/HOBt mediated solution phase peptide coupling protocol, starting from 350mg (1.85 mmol) Boc-alanine-**4.046** and 600mg (1.85mmol) isothiocyanyl lysineine TFAsalt-**4.050** , the desired dipeptide was obtained (R_f=0.3 in 1:1, hexane/ethyl acetate) in 6 hours. The residue was purified by column chromatography on silica gel (60-120) by using hexane/ethyl acetate. A colourless liquid compound; 390mg, Yield 52 %. IR (KBr) ῡ 3318, 2977, 2935, 2869, 2184, 2107, 1651, 1513, 1453 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 7.14 (1H, d, J = 3.1 Hz), 5.29 (1H, d, J = 3.7 Hz), 4.88 (1H, d, J = 3.2 Hz), 4.12 (1H, s), 3.67 (3H, s), 3.40 (2H, t, J = 6.6 Hz), 3.11 (3H, s), 1.73–1.61 (2H, m), 1.54 (2H, tdd, J = 13.8, 11.8, 6.3 Hz), 1.39 –1.28 (11H, m), 1.22 (3H, d, J = 7.0 Hz); ^{13}C NMR (150 MHz; $CDCl_3$) δ 172.7, 171.9, 155.2, 129.7, 79.5, 61.5, 49.9, 48.4, 44.6, 31.9, 31.4, 29.4, 29.2, 28.2, 22.1, 18.5. +ESI-HRMS (m/z) calculated for $C_{17}H_{31}N_4O_5S$ $[M+H]^+$ 402.2015, found-402.2017.

(S)$^{\omega}$N,$^{\alpha}$N,C-triprotected Lys-Gly dipeptide (4.059): Using EDC.HCl/HOBt mediated solution phase peptide coupling protocol, starting from 900 mg (2.36mmol) Boc-lysine(Cbz)CO_2H **2.048** (**Chapter-2**) and glycine methylester salt 320mg (2.60mmol) the desired dipeptide-**4.052** was obtained as gummy liquid (R_f = 0.35 in 1:10, hexane/ethyl acetate).

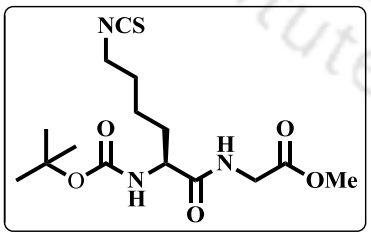


A gummy liquid compound; 940mg, Yield 88 %. IR (KBr) ῡ 3331, 3067, 3028, 2928, 2856, 1746, 1702, 1528, 1455, 1438, 1391, 1367, 1250, 1213, 1169, 1025 cm^{-1}.^{1}H NMR (600 MHz; $CDCl_3$) δ 7.28 (4H, d, *J* = 4.0 Hz), 7.26 – 7.21 (1H, m), 5.54 (1H, s), 5.37 (1H, s), 5.02 (2H, s), 4.15 (1H, d, *J* = 5.7 Hz), 4.00 (1H, dd, *J* = 18.0, 4.6 Hz), 3.87 (1H, d, *J* = 17.8 Hz), 3.62 (3H, s), 3.11 (2H, d, *J* = 4.6 Hz), 1.81 – 1.68 (1H, m), 1.60 (1H, dd, *J* = 13.7, 6.9 Hz), 1.52 – 1.42 (2H, m), 1.37 (11H, s).^{13}C NMR (150 MHz; $CDCl_3$) δ 172.9, 170.2, 156.6, 155.8, 136.6, 128.4, 128.0, 128.0, 79.8, 66.5, 54.1, 52.2, 41.0, 40.4, 32.1, 29.3, 28.3, 22.3. +ESI-HRMS (m/z) calculated for $C_{22}H_{34}N_3O_7$ $[M+H]^+$ 452.2397, found 452.2390.

(S)$^{\alpha}$N,C-diprotected Lys-Gly dipeptide (4.053):Using general protocol for carboxy benzyl (Cbz) deprotection, starting from 350mg (0.77mmol) dipeptide-4.059, the desired amino dipeptide-**4.053** was obtained as semi solid material. The compound was directly used for next reaction without any purification. A semi solid compound; 220 mg, Yield = 90 %.

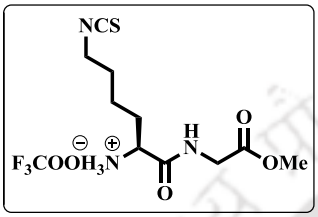


(S)$^{\alpha}$N,C-diprotected ^{NCS}Lys-Gly dipeptide (4.054): The Compound-**4.053** (190mg, 0.59mmol) was loaded in dry 50 mL single neck round bottom flask under nitogen atmosphere and dry THF was added to it. Then reaction mixture was cooled to 0 ^{o}C and Et_3N (0.248ml, 1.77mmol) was added to it, after 5 minute CS_2 (0.054ml, 0.89mmol) was added to it and stirred for 1 hour at 0 ^{o}C under N_2 atmosphere. After that at 0 ^{o}C tosyl chloride (124mg, 0.64mmol) was added to it and whole reaction mixture was stirred at 0 ^{o}C for 20 minute under N_2 atmosphere. Then reaction mixture was dried under high vacuum and ethyl acetate was added to the reaction mixture and organic layer was treated with 1(N) HCl followed by saturated $NaHCO_3$ and brine solution and dried over Na_2SO_4. After Evaporation of ethyl acetate, residue (R_f =0.3 in, 1:10 hexane/ethyl acetate) was purified by column chromatography on silica gel (60-120) using hexane/ethyl acetate. A gummy liquid Compound; 135mg, Yield 63 %. IR (KBr) ῡ 3318, 2984, 2953, 2926, 2853, 2185, 2107, 1745, 1664, 1523, 1455, 1438, 1391, 1367, 1248, 1210, 1170 cm^{-1}.^{1}H NMR (600 MHz; $CDCl_3$) δ 7.11 (1H, s), 5.35 (1H, s), 4.19 (1H, d, *J* = 6.0 Hz), 4.07–



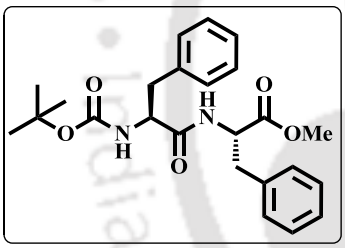
3.90 (2H, m), 3.69 (3H, s), 3.49 (2H, t, *J* = 6.0 Hz), 1.86–1.76 (1H, m), 1.72–1.56 (3H, m), 1.53–1.44 (2H, m), 1.39 (9H, s); ^{13}C NMR (150 MHz; $CDCl_3$) δ 172.5, 170.2, 155.8, 130.0, 80.2, 53.9, 52.4, 44.9, 41.1, 31.9, 29.6, 28.3, 22.6. +ESI-HRMS (m/z) calculated for $C_{15}H_{26}N_3O_5S$ $[M+H]^+$ 360.1593 Found 360.1579.

(S) C-protected ^{NCS}Lys-Gly dipeptide trifluoroacetate salt (4.055): using general protocol for Boc deprotection, starting from 250mg dipeptide-**4.054** (250mg, 0.69mmol) the



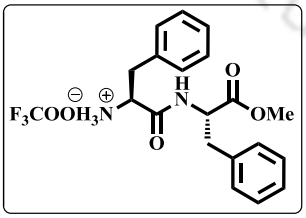
TFA salt of dipeptide **4.055** was obtained as brown semi solid material. The compound was directly used for next reaction without any purification. A brown gummy compound; 334mg.

(S,S)N,C-diprotected Phe-Phe dipeptide (4.057): Using EDC.HCl/HOBt mediated solution phase peptide coupling protocol, starting from 700mg (3.02 mmol) Boc-phenylalanine-**4.056** and phenylalanine methylester salt 715mg (3.33mmol) the desired dipeptide-**4.057** was obtained (R_f =0.5 in 4:1, hexane/ethyl acetate).White solid compound;



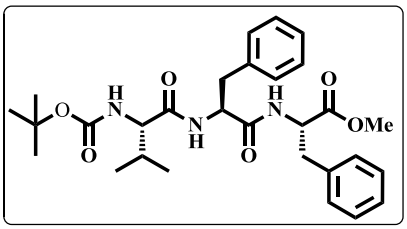
900mg, Yield 70 %. IR (KBr) ῡ 3331, 3087, 3063, 3033, 2986, 2927, 2859, 1941, 1744, 1698, 1666, 1521, 1495, 1445, 1388, 1366, 1349, 1299, 1282 1249, 1223, 1165, 1117 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 7.32–7.15 (8H, m), 6.99 (2H, d, *J* = 7.0 Hz), 6.42 (1H, d, *J* = 5.9 Hz), 5.04 (1H, s), 4.78 (1H, d, *J* = 5.5 Hz), 4.36 (1H, s), 3.65 (3H, s), 3.04 (4H, qd, *J* = 14.0, 6.2 Hz), 1.39 (9H, s); ^{13}C NMR (150 MHz; $CDCl_3$) δ 171.5, 170.9, 155.4, 136.6, 135.8, 129.4, 129.3, 128.7, 128.6, 127.2, 127.0, 80.2, 55.7, 53.4, 52.3, 38.3, 38.0, 28.3. +ESI-HRMS (m/z) calculated for $C_{24}H_{31}N_2O_5$ $[M+H]^+$ 427.2233, found-427.2227.

(S,S)C-protected Phe-Phe dipeptide trifluoroacetate salt (4.058): using general

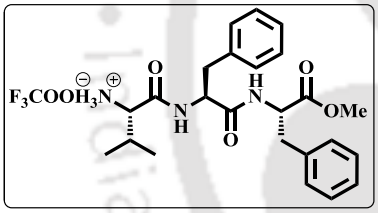


protocol for Boc deprotection, starting from 1100mg (2.58mmol) dipeptide-**4.057** the TFA salt of dipeptide-**4.058** was obtained as light brown liquid. The compound was directly used for next reaction without any purification. Weight of the compound = 1280mg.

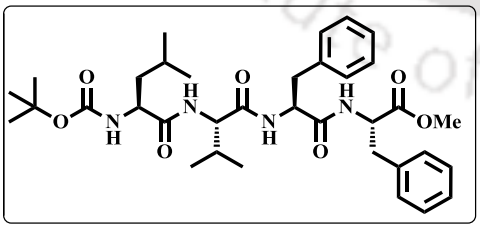
(S,S,S)N,C-diprotected Val-Phe-Phe tripeptide (4.059): Using EDC.HCl/HOBt mediated solution phase peptide coupling protocol , starting from 650mg (2.99 mmol) Boc-valine and dipeptide TFAsalt-**4.058** 715mg (3.33mmol) the desired tripeptide-**4.059** was obtained (R_f =0.35 in 1:2, hexane/ethyl acetate). A white solid compound; 1250mg, Yield 79 % ; m.p. 156-160 ^{0} C. IR (KBr) ῡ 3314, 3278, 3064, 3030, 3001, 2968, 2934, 2871, 1728, 1688, 1664, 1643, 1543, 1530, 1495, 1453, 1438, 1389, 1366, 1348, 1301, 1281, 1235 cm^{-1}. ^{1}H NMR (600 MHz, $CDCl_3$) δ 7.26–7.07 (9H, m), 6.88 (2H, d, J = 6.7 Hz), 6.45 (1H, s), 6.15 (1H, s), 4.86 (1H, s), 4.66 (1H, dd, J = 13.4, 6.3 Hz), 4.56 (1H, dd, J = 13.9, 7.0 Hz), 3.86 – 3.78 (1H, m), 3.59 (3H, s), 3.03 (1H, dd, J = 13.5, 4.9 Hz), 2.97 (1H, dd, J = 13.9, 6.1 Hz), 2.91 (2H, ddd, J = 13.8, 6.8, 4.2 Hz), 2.00 (1H, dt, J = 13.4, 6.9 Hz), 1.70 – 1.62 (2H, m), 1.37 (9H, s), 0.80 (3H, d, J = 6.8 Hz), 0.71 (3H, d, J = 6.6 Hz); ^{13}C NMR (150 MHz; $CDCl_3$) δ 171.6, 171.3, 170.2, 156.0, 136.4, 135.7, 129.5, 129.3, 128.8, 128.7, 127.3, 127.2, 80.2, 60.1, 54.3, 53.6, 52.4, 38.3, 38.0, 30.8, 28.4, 19.4, 17.6. +ESI-HRMS (m/z) calculated for $C_{29}H_{40}N_3O_6$ $[M+H]^+$ 526.2917, found-526.2838



(S,S,S)C-protected Val-Phe-Phe tripeptide trifluoroacetate salt (4.060): using general protocol for Boc deprotection, starting from 300mg tripeptide-**4.059** (0.57mmol) the TFA salt of tripeptide was obtained as semi solid material. The compound was directly used for next reaction without any purification. Weight of the compound = 384mg.

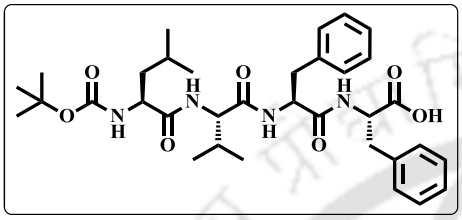


(S,S,S,S)N,C-diprotected Leu-Val-Phe-Phe tetrapeptide(4.061): Using PyBOP/ HOBt mediated solution phase peptide coupling protocol , starting from 170mg (0.73 mmol) Boc-leucine and tripeptide TFAsalt-**4.059** 384mg (0.73mmol) the desired tetrapeptide-**4.061** was obtained (R_f = 0.4 in 20:1, chloroform/methanol). A white Solid Compound; 360mg, Yield 76 %; m.p. 186-190 ^{o} C. IR (KBr) ῡ 3283, 3066, 3031, 2959, 2930, 2871, 1744, 1692, 1639, 1547, 1524, 1499, 1454, 1391, 1336, 1320, 1266, 1235, 1171, 1118 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 7.21 (8H, ddd, J = 28.9, 15.7, 7.6 Hz), 7.03 (3H, d, J = 6.4 Hz), 6.89–6.77 (1H, m), 6.78 –6.64 (1H, m), 5.17–5.04 (1H, m), 4.76 (2H, dd, J = 13.7, 6.8 Hz), 4.29 (1H, s), 4.10 (1H, d, J = 4.2 Hz), 3.64 (3H, d, J = 1.9 Hz), 3.07 (2H, dd, J = 12.8, 4.9 Hz), 3.05–2.91 (2H, m), 2.10 (1H, s), 1.92–1.79 (2H, m), 1.71–1.64 (1H, m), 1.59 (1H, d, J = 4.5 Hz), 1.54–1.46 (1H, m), 1.42 (9H, s), 0.92 (6H, dd, J = 11.7, 6.5 Hz), 0.85 (3H, d, J = 6.8 Hz), 0.78 (3H, s);^{13}C NMR (150 MHz; $CDCl_3$) δ 173.2, 171.5, 171.1, 170.7,

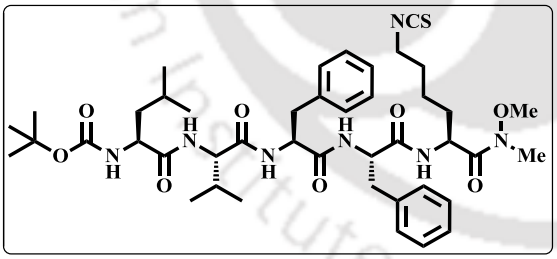


156.0, 136.7, 136.2, 129.4, 129.3, 128.6, 128.6, 127.1, 126.9, 80.0, 58.6, 54.1, 53.6, 53.6, 52.3, 41.1, 38.5, 38.2, 31.1, 28.5, 24.9, 23.2, 22.2, 19.4, 18.1.+ESI-HRMS (m/z) calculated for $C_{35}H_{51}N_4O_7$ $[M+H]^+$ 639.3758, Found-639.3779.

(S,S,S,S)N-protected Leu-Val-Phe-Phe tetrapeptide acid (4.062): At first compound-**4.061** (350mg, 0.54 mmol) was loaded in 50 mL single neck R.B and THF/Water (3:1) was added to it. After that $LiOH.H_2O$ (35mg, 0.82mmol) was added to the stirred solution at 0 ^{0} C and it was allowed to stir for about 4 hours at room temperature. The reaction mixture was dried under high vacuum, 2-3 ml water was added, and it was acidified with about 1N HCl until the pH 2-3.Then product was extracted with ethyl acetate and was dried with Na_2SO_4. After evaporation of ethyl acetate could afford solid compound and it was directly used it for next step without purification. A white solid; 280mg, Yield 83%.

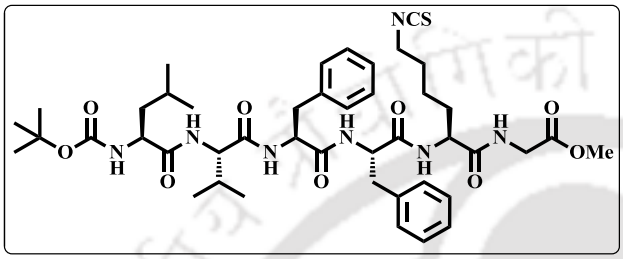


(S,S,S,S,S)N,C-diprotected Leu-Val-Phe-Phe-^{NCS}Lys pentapeptide (4.063): Using PyBOP/HOBt mediated solution phase peptide coupling protocol, starting from 200mg (0.32 mmol) Boc-tetrapeptide acid-**4.062** and 116mg (0.35 mmol) isothiocyanyllysine TFA salt-**4.050**, the desired pentapeptide-**4.063** was obtained (R_f= 0.4 in, 1:20 hexane/ethyl acetate) in 5 hours. A white solid; 130mg, Yield 48 %; m.p. 201-205 ^{o} C. IR (KBr) $\bar{\upsilon}$ 3293, 3065, 3030, 2928, 2186, 2103, 1710, 1688, 1649, 1523, 1454, 1390, 1366, 1236, 1168 cm^{-1}.^{1}H NMR (600 MHz; $CDCl_3$) δ 7.61–7.34 (2H, m), 7.16 (10H, ddd, J = 33.2, 20.2, 7.2 Hz), 6.88 (1H, d, J = 9.8 Hz), 5.35 (1H, s), 4.98 (2H, s), 4.79 (1H, s), 4.30 (1H, d, J = 14.7 Hz), 4.11 (1H, s), 3.71 (3H, s), 3.48 (2H, dt, J = 12.4, 6.2 Hz), 3.20 (4H, s), 3.11–3.01 (1H, m), 2.97 (1H, dd, J = 13.7, 8.6 Hz), 2.82 – 2.70 (1H, m), 2.12 – 1.99 (1H, m), 1.88 (2H, d, J = 24.6 Hz), 1.81–1.61 (6H, m), 1.57 (3H, d, J = 5.9 Hz), 1.46 (2H, d, J = 7.9 Hz), 1.42 (9H, s), 1.25 (1H, s), 0.93 (6H, dd, J = 16.1, 6.6 Hz), 0.85 (3H, d, J = 6.8 Hz), 0.78 (3H, s); ^{13}C NMR (150 MHz; $CDCl_3$) δ 173.5, 171.7, 171.2, 171.1, 170.8, 156.3, 137.1, 137.0, 130.0, 129.5, 129.2, 128.5, 128.4, 126.8, 126.7, 80.3, 61.8, 59.0, 54.3, 54.0, 48.4, 44.9, 41.3, 38.5, 32.2, 32.0, 29.8, 29.7, 28.5, 24.9, 23.2, 22.5, 22.1, 19.3, 18.3.+ESI-HRMS (m/z) calculated for $C_{43}H_{64}N_7O_8S$ $[M+H]^+$ 838.4537, found-838.4542.



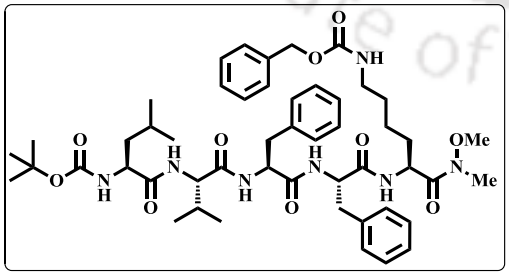
(S,S, S,S,S)N,C-diprotected Leu-Val-Phe-Phe-^{NCS}Lys-Gly hexapeptide (4.064):

Using PyBOP/HOBt mediated solution phase peptide coupling protocol, starting from 330mg (0.52 mmol) Boc-tetrapeptide acid-**4.062** and 187mg (0.52 mmol) dipeptide TFA salt-**4.055** the desired hexapeptide-**4.064** was obtained in 6 hours. A white gummy solid was formed and after washing that solid with, EtOAc, $CHCl_3$ and diethyl ether a solid residue was formed. A white solid; 250mg, Yield 55 %; m.p. 249-253 ^{o} C. IR (KBr) ῡ 3414, 3281, 3086, 3028, 2957, 2925, 2873, 2854, 2184, 2105, 1751, 1693, 1632, 1544, 1454, 1392, 1367, 1285, 1214 cm^{-1}. ^{1}H NMR (600 MHz; DMSO-d_6) δ 8.35 (1H, t, J = 5.7 Hz), 8.14 (1H, d, J = 7.9 Hz), 8.12–8.02 (2H, m), 7.45 (1H, d, J = 9.0 Hz), 7.17 (10H, ddd, J = 20.2, 19.2, 11.8 Hz), 7.05 (1H, d, J = 8.6 Hz), 4.61–4.47 (2H, m), 4.33 (1H, dd, J = 13.8, 8.1 Hz), 4.14 (1H, dd, J = 15.0, 6.9 Hz), 3.96 – 3.78 (3H, m), 3.63 (4H, q, J = 6.7 Hz), 3.02 (1H, dd, J = 13.8, 4.4 Hz), 2.97–2.85 (2H, m), 2.82 (1H, dd, J = 13.7, 9.0 Hz), 2.71 (1H, dd, J = 14.0, 10.1 Hz), 1.90 – 1.79 (1H, m), 1.64 (3H, dt, J = 13.8, 7.5 Hz), 1.60–1.45 (2H, m), 1.43–1.24 (13H, m), 0.82 (6H, dd, J = 19.5, 6.4 Hz), 0.78–0.64 (6H, m); ^{13}C NMR (150 MHz, DMSO-d_6) δ 172.2, 171.8, 170.9, 170.6, 170.5, 170.1, 155.3, 137.6, 137.5, 129.2, 129.03, 128.07, 128.00, 126.2, 126.1, 78.1, 57.0, 53.7, 53.5, 53.0, 51.8, 51.7, 44.6, 40.5, 37.4, 31.4, 31.0, 28.8, 28.2, 24.2, 23.0, 22.1, 21.5, 19.1, 17.7.+ESI-HRMS (m/z) calculated for $C_{44}H_{64}N_7O_9S$ $[M+H]^+$ 866.4486; $C_{44}H_{63}N_7NaO_9S$ $[M+Na]^+$ 888.4306, found-866.4535, 888.4319.



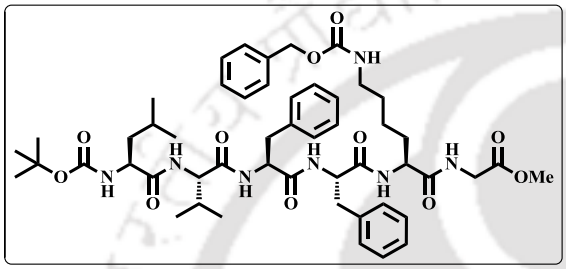
S,S, S,S,S)N,C-diprotected Leu-Val-Phe-Phe-^{NHCbz}Lys pentapeptide (4.067):

Using EDC.HCl/HOBt mediated general solution phase peptide coupling protocol, starting from 210mg (0.33 mmol) Boc-tetrapeptide-**4.062** and 211mg (0.49mmol) Cbz-lysine TFA salt-**4.065**, the desired pentapeptide **4.067** was obtained (R_f=0.4 in 20:1 $CHCl_3$/Methanol) in 16 hours. A white solid; 270mg, Yield 88 %; m.p.148-152 ^{o} C. IR (KBr) ῡ 3289, 3065, 2960, 2871, 1703, 1636, 1528, 1454, 1391, 1366, 1241, 1169 cm^{-1}. ^{1}H NMR (600 MHz; $CDCl_3$) δ 7.754(1H, s), 7.37–7.27 (5H, m), 7.14 (10H, ddd, J = 45.6, 23.7, 7.4 Hz), 6.82 (1H, s),5.83(1H, s), 5.49 (1H, s), 5.27 (1H, s), 5.13 (1H, d, J = 8.0 Hz), 5.11 – 5.01 (2H, m), 5.02 – 4.82 (2H, m), 4.83 –4.64 (1H, m), 4.20 (1H, s), 4.07 (1H, s), 3.71 (3H, s), 3.19 (3H, s), 3.14 (3H, dt, J = 12.9, 7.1 Hz), 3.05 (1H, d, J = 14.7 Hz), 2.95 (1H, t, J = 11.4 Hz), 2.74 (1H, d, J = 12.5 Hz), 2.11 – 2.02 (1H, m), 1.74 (1H, p, J = 6.9, 6.3 Hz), 1.64 (2H, h, J = 7.8, 7.4 Hz), 1.51 (5H, dp, J = 21.8,

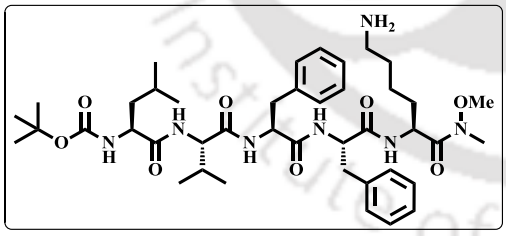


7.4, 6.9 Hz), 1.41 (11H, s), 0.90 (6H, dd, *J* = 14.9, 6.5 Hz), 0.78 (6H, dd, *J* = 42.7, 6.8 Hz). +APCI-HRMS (m/z) calculated for $C_{50}H_{72}N_7O_{10}$ $[M+H]^+$ 930.5341, found-930.5343.

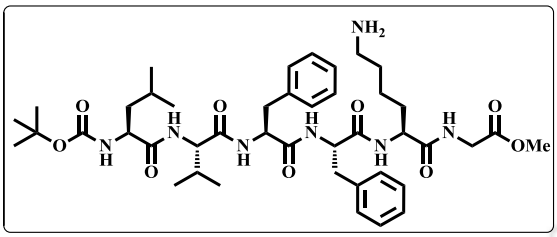
(S,S,S,S,S)N,C-diprotectedLeu-Val-Phe-Phe-^{NHCbz}Lys-Gly hexapeptide(4.068): Using EDC.HCl/HOBt mediated general solution phase peptide coupling protocol, starting from 206mg (0.32 mmol) Boc-tetrapeptide-**4.062** and 223mg (0.48mmol) dipeptide TFA salt-**4.066**, the desired hexaeptide-**4.068** was obtained in 16 hours. A white gummy solid was formed and after washing that solid with, EtOAc and diethyl ether a solid residue was formed. A White solid; 260mg, Yield 84 %; m.p. 233-237 ^{o} C. IR (KBr) ῡ 3397, 3278, 3089, 3064, 3028, 2958, 2931, 2867, 1749, 1713, 1693, 1629, 1546, 1452, 1438, 1404, 1388 cm^{-1}. ^{1}H NMR (600 MHz; DMSO-d_6) δ 8.25 (1H, t, *J* = 5.9 Hz), 8.06 (3H, dq, *J* = 11.0, 7.3, 6.7 Hz), 7.43 (1H, d, *J* = 8.9 Hz), 7.32 (4H, ddt, *J* = 19.5, 13.9, 7.2 Hz), 7.26 – 7.09 (10H, m), 7.03 (1H, d, *J* = 8.5 Hz), 4.99 (2H, s), 4.57 (1H, td, *J* = 8.3, 4.7 Hz), 4.51 (1H, td, *J* = 8.8, 4.3 Hz), 4.26 (1H, td, *J* = 8.1, 5.4 Hz), 4.13 (1H, dd, *J* = 9.0, 6.3 Hz), 3.93 (1H, td, *J* = 9.2, 5.0 Hz), 3.83 (2H, qd, *J* = 17.4, 5.9 Hz), 3.61 (3H, s), 3.06 – 2.89 (4H, m), 2.81 (1H, dd, *J* = 14.1, 8.8 Hz), 2.71 (1H, dd, *J* = 14.2, 9.6 Hz), 1.90 – 1.80 (1H, m), 1.64 (1H, dt, *J* = 13.5, 6.8 Hz), 1.52 (2H, dp, *J* = 13.8, 8.5, 6.3 Hz), 1.40 (3H, q, *J* = 7.4 Hz), 1.36 (9H, s), 1.27 (3Hq, *J* = 8.1, 7.4 Hz), 0.82 (6H, dd, *J* = 19.2, 6.6 Hz), 0.71 (6H, q, *J* = 7.9, 7.3 Hz). +APCI-HRMS (m/z) calculated for $C_{51}H_{72}N_7O_{11}$ $[M+H]^+$ 958.5290, found-958.5290.



(S,S,S,S,S)N,C-diprotected Leu-Val-Phe-Phe-^{NH2}Lys hexapeptide (4.069): Using general protocol for carboxy benzyl (Cbz) deprotection, starting from 50mg (0.053mmol) Cbz-pentapeptide-**4.067**, the desired amino pentapeptide-**4.069** was obtained as a solid material. The formation of compound was confirmed by HRMS. A solid compound; 35mg, Yield 83%.+ESI-HRMS (m/z) calculated for $C_{42}H_{66}N_7O_8$ $[M+H]^+$ 796.4973, found 796.4971.

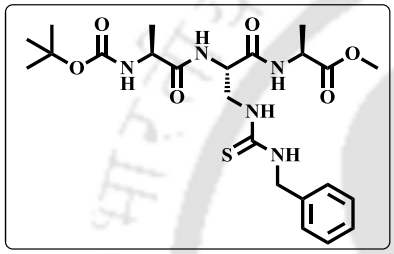


(S,S, S,S,S)N,C-diprotected Leu-Val-Phe-Phe-^{NH2}Lys-Gly hexapeptide (4.070):

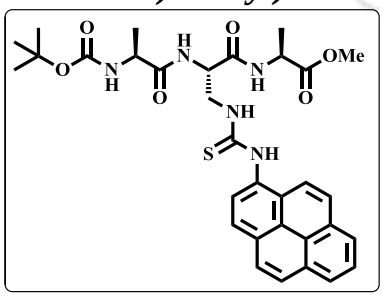


Using general protocol for carboxy benzyl (Cbz) deprotection, starting from 50mg (0.052mmol) Cbz-hexapeptide-**4.068**, the desired amino hexapeptide-**4.070** was obtained as a solid material in 20 hours. The formation of compound was confirmed by HRMS. A white soilid; 30mg, Yield 72%. +ESI-HRMS (m/z) calculated for $C_{43}H_{66}N_7O_9$ $[M+H]^+$ 824.4922, found 824.4925.

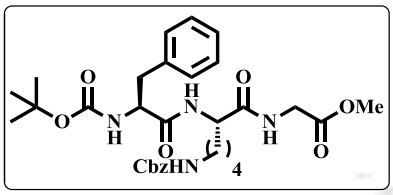
(6S,9S,12S)-Methyl9-((3-benzylthioureido)methyl)-2,2,6,12-tetramethyl-4, 7,10-trioxo-3-oxa-5,8,11-triazatridecan-13-oate (4.073): In a 10 mL single neck round bottom flask,compound-**4.044** (5mg, 0.012mmol) was loaded and 1ml sodium phosphate buffer PH = 7.0 1ml and 30ul DMF was added to it. After that benzyl amine (2mg, 0.012mmol) was added to it and reaction mixture was stirred for 2 h under N_2 atmosphere. After completion of reaction, reaction mixture was dried under high vacuum and formation of thiourea derivative was confirmed by NMR and HRMS. Yield 71 %. IR (KBr) ῡ 3346, 2975, 2923, 1740, 1655, 1536, 1536, 1449, 1371, 1342, 1246, 1168, 1072, 1025 cm^{-1}. ^{1}H NMR (600 MHz, $CDCl_3$) δ 7.29 –7.24 (3H, m), 7.22 –7.18 (2H m,), 7.10 (1H, d, J = 4.9 Hz), 6.50 (1H, d, J = 13.3 Hz), 6.36 (1H, s), 4.99–4.95 (2H, m), 4.84 (2H, d, J = 5.1 Hz), 4.72 (1H, d, J = 13.2 Hz), 4.69–4.63 (1H, m), 4.13 (1H, d, J = 8.3 Hz), 3.53 (3H, s), 3.01 (1H ,dd, J = 11.2, 10.4 Hz), 1.49 (3H, d, J = 7.5 Hz), 1.35 (9H, s), 1.27 (1H ,d, J = 7.1 Hz). ^{13}C NMR (150 MHz, $CDCl_3$) δ 181.2, 173.6, 171.2, 168.9, 155.4, 137.9, 128.9, 128.2, 127.8, 80.5, , 57.9, 53.0, 52.8, 50.4, 48.4, 46.4, 28.5, 18.9, 15.1. +APCI-HRMS (m/z) calculated for $C_{23}H_{36}N_5O_6S$ [M+H] + 510.2386, found 510.2385.



(6S,9S,12S)-Methyl2,2,6,12-tetramethyl-4,7,10-trioxo-9-((3-(pyren-1-yl) thioureido)methyl)-3-oxa-5,8,11triazatridecan-13-oate (4.074):In a 10 mL round bottom flask compound-**4.044** (5mg, 0.012mmol) was loaded and 1ml sodium phosphate buffer pH = 7.0 1ml and 30µl DMF was added to it. 1-amino pyrene (3mg, 0.012mmol) was added to it and reaction mixture was stirred for 2 h under N_2 atmosphere. After completion of reaction, reaction mixture was dried under high vacuum and formation of thiourea derivative was confirmed by ESI-HRMS.+APCI-HRMS (m/z) calculated for $C_{32}H_{38}N_5O_6S$ $[M+H]^+$620.2543, found 620.2523

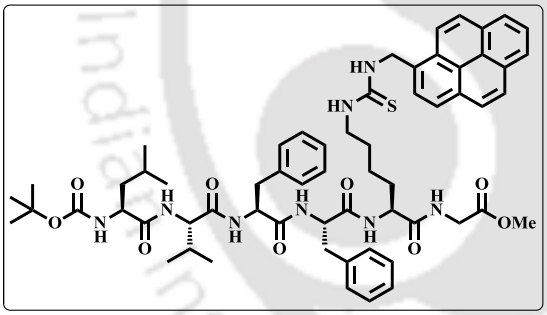


(6S,9S)-Methyl6-benzyl-9-(4-(((benzyloxy)carbonyl)amino)butyl)-2,2-dimethyl-4,7,10-trioxo-3-oxa-5,8,11-triazatridecan-13-oate : Using EDC.HCl/HOBt mediated solution phase peptide coupling protocol, starting from 600mg (2.26mmol) Boc-alanine and



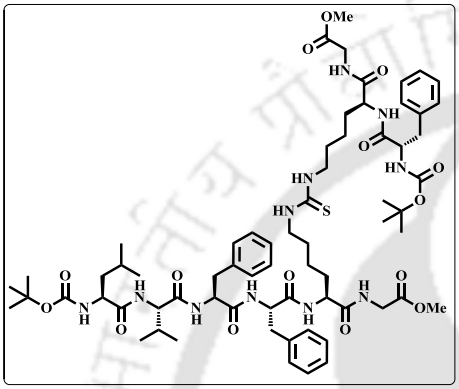
327mg (2.26mmol) dipeptide TFAsalt-**4.075,** the desired dipeptide was obtained (R_f= 0.3 in, 1:20 $CHCl_3$/Methanol) in 16 hours. The residue was purified by chromatography on silica gel (60-120) by using $CHCl_3$/Methanol. White solid Compound; 1000mg, Yield 73 %. ^{1}H NMR (600 MHz; $CDCl_3$) δ 7.39–7.17 (10H, m), 6.75 (2H, dt, *J* = 12.0, 6.7 Hz), 5.13 (4H, dd, *J* = 35.0, 6.0 Hz), 4.47 (1H, td, *J* = 7.9, 5.1 Hz), 4.40 (1H, q, *J* = 7.0 Hz), 3.95 (2H, qd, *J* = 18.1, 5.7 Hz), 3.26–3.08 (3H, m), 3.03 (1H, dd, *J* = 13.9, 7.4 Hz), 1.89 (3H, s), 1.73–1.62 (1H, m), 1.51 (2H, h, *J* = 7.0 Hz), 1.40 (9H, s), 1.36–1.26 (3H, m).+ESI-HRMS calculated for $C_{31}H_{43}N_4O_8$ [M+H]+599.3081, found 599.3093.

(6S,9S,12S,15S,18S)-Methyl12,15-dibenzyl-6-isobutyl-9-isopropyl-2,2-dimethyl 4, 7, 10, 13, 16, 19-hexaoxo-18-(4-(3-(pyren-1 ylmethyl) thioureido) butyl)-3-oxa-5,8,11,14, 17, 20-hexaazadocosan-22-oate (4.077): In a 10 mL single neck round



bottom flask, compound-**4.064** (20mg, 0.02 mmol) was loaded and 1ml DMSO was added to it. After that pyrene methyl amine (5mg, 0.02mmol) was added to it and reaction mixture was stirred for 14 h at 60 ^{0} C under N_2 atmosphere. After completion of reaction, reaction mixture was dried under high vacuum and formation of thiourea derivative was confirmed by ^{1}HNMR and ESI-HRMS. ^{1}H NMR (600 MHz; DMSO-d_6) δ 8.38 (1H, s), 8.34 – 8.22 (2H, m), 8.18 – 8.04 (4H, m), 8.02 (1H, d, *J* = 6.5 Hz), 7.80 (1H, s), 7.52 – 7.39 (1H, m), 7.29 – 7.07 (10H, m), 7.05 (1H, d, *J* = 8.0 Hz), 6.60 (2H, s), 5.37 (1H, s,), 4.57 (1H, s), 4.50 (1H, s), 4.43 (1H,d, *J* = 6.7 Hz), 4.31 – 4.20 (1H,m), 4.12 (2H, d, *J* = 7.3 Hz), 3.93 (2H, d, *J* = 4.6 Hz), 3.89 –3.76 (3H, m), 3.60 (3H, d, *J* = 4.4 Hz), 3.07 – 2.95 (2H, m), 2.95 – 2.86 (2H, m), 2.81 (1H, dd, *J* = 12.9, 8.5 Hz), 2.71 (1H, dd, *J* = 24.2, 14.5 Hz), 1.84 (1H, d, *J* = 6.3 Hz), 1.63 (1H, s), 1.58–1.44 (2H, m), 1.43–1.13 (14H, m), 0.82 (6H, dd, *J* = 19.6, 4.7 Hz),0.70 (6H, d, *J*=5.0Hz).+ESI-HRMS (m/z) calculated for $C_{61}H_{77}N_8O_9S$ $[M+H]^+$1097.5534, found 1097.5583.

(5S,17S)-Dimethyl5-((S)-2-((tert-butoxycarbonyl)amino)-3 phenylpropanamido)-17-((6S,9S,12S,15S)-12,15-dibenzyl-6-isobutyl-9-isopropyl-2,2-dimethyl-4,7,10,13-tetraoxo-3-oxa-5,8,11,14-tetraazahexadecanamido)-4,18-dioxo-11-thioxo-3,10,12,19-tetraazahenicosane-1,21-dioate (4.078): In a 10 mL single neck round bottom flask, compound-**4.064** (20mg, 0.02 mmol) was loaded and 1ml DMSO was added to it. After that tripeptide amine (9mg, 0.02mmol) was added to it and reaction mixture was stirred for 8 h at 60 ^{0} C under N_2 atmosphere. After completion of reaction, reaction mixture was dried under high vacuum and formation of thiourea derivative was confirmed by ^{1}HNMR, TOCSY and ESI-HRMS.^{1}H NMR (600 MHz; DMSO-d_6) δ 8.37 (1H, s), 8.29 (1H, d, J = 5.7 Hz), 8.09 (3H,d, J = 13.5 Hz), 7.95 (1H,d, J = 7.1 Hz), 7.80 (1H, s), 7.44 (1H, d, J = 7.7 Hz), 7.37 (1H, s), 7.30 – 7.08 (13H, m), 7.05 (1H, d, J = 7.6 Hz), 6.96 (1H, d, J = 7.9 Hz), 4.57 (1H, s), 4.51 (1H, dd, J = 8.0, 4.0 Hz), 4.28 (1H, dd, J = 16.4, 7.6 Hz), 4.15 (2H, dd, J = 23.2, 6.9 Hz), 3.93 (1H, d, J = 5.1 Hz), 3.83 (3H, dt, J = 16.0, 10.1 Hz), 3.61 (4H, s), 3.00 (2H, dd, J = 21.7, 13.5 Hz), 2.95 – 2.86 (2H, m), 2.85 – 2.76 (1H, m), 2.76 – 2.66 (2H, m), 1.85 (1H, d, J = 6.0 Hz), 1.73 – 1.59 (2H, m), 1.54 (3H, s), 1.45 (3H, d, J = 6.1 Hz), 1.35 (9H, s), 1.28 (9H, s), 1.21 (2H, s), 0.82 (6H, dd, J = 19.4, 5.7 Hz), 0.70 (6H, d, J = 5.3 Hz).+ESI-HRMS (m/z) calculated for $C_{67}H_{100}N_{11}O_{15}S$ $[M+H]^{+}$1330.7121, found 1330.7116.



4.7.3. ^{1}H, ^{13}C and TOCSY NMR Spectra of Few Selected Peptides

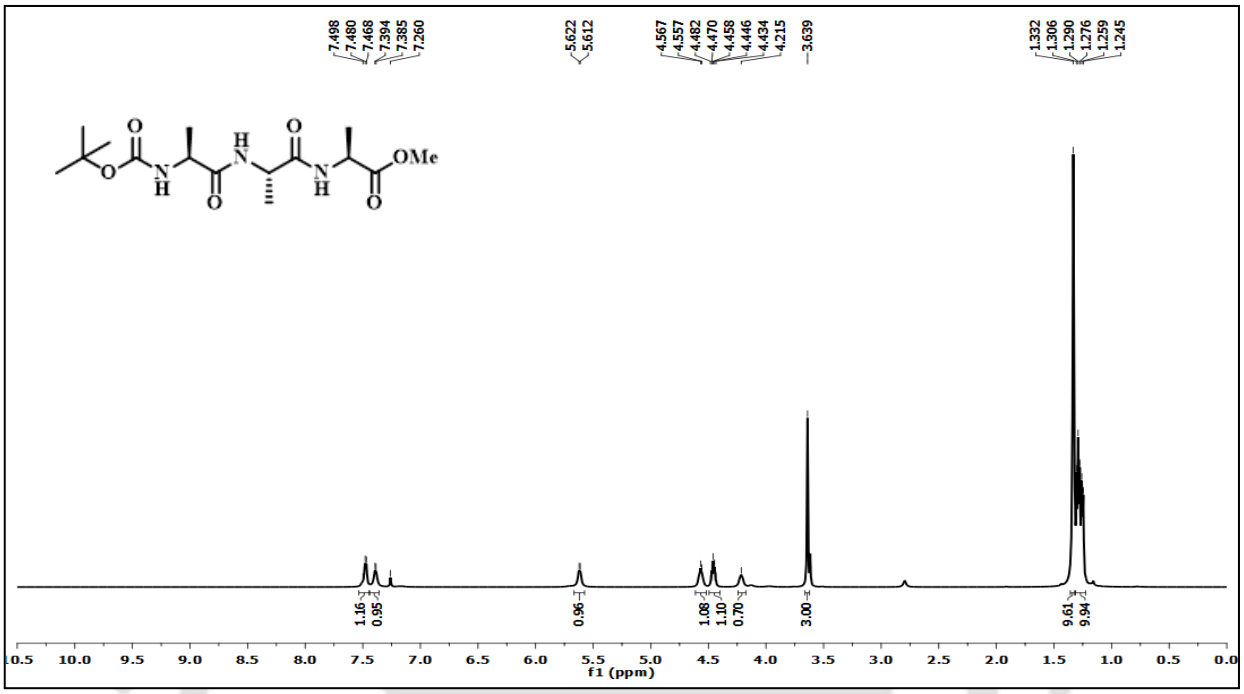


Figure 4.30. ^{1}H NMR spectra of synthesised compound **4.049.**

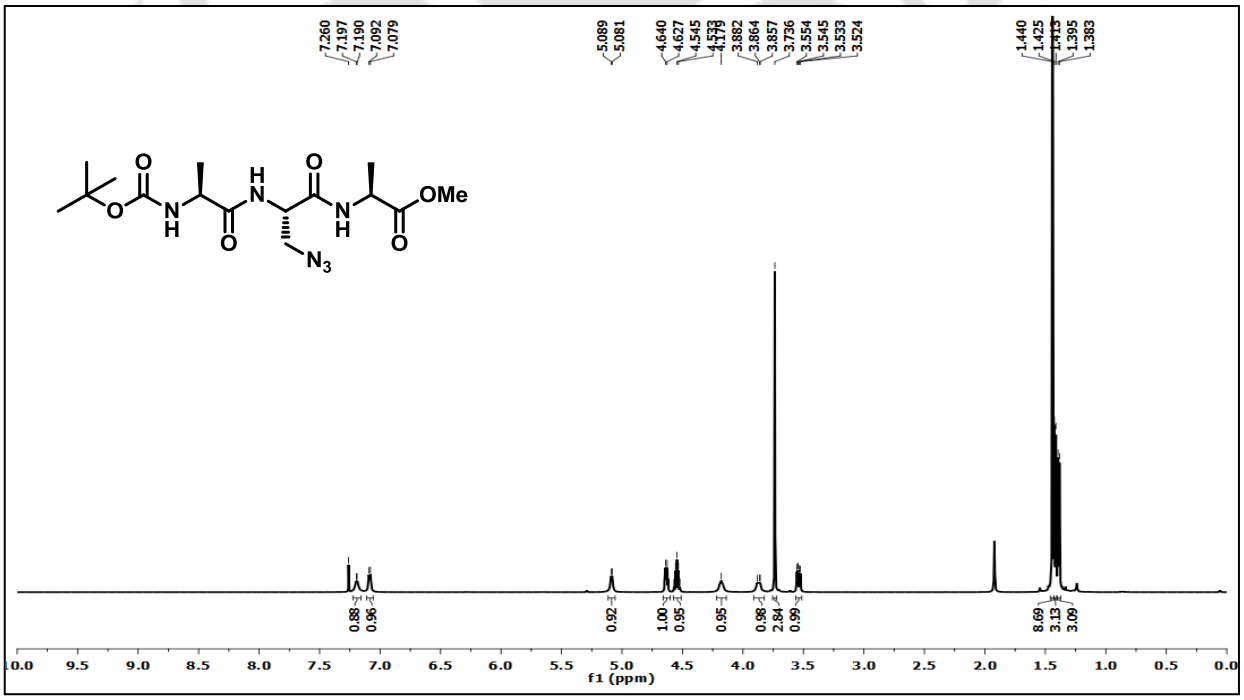


Figure 4.31. ^{1}H NMR spectra of synthesised compound **4.042.**

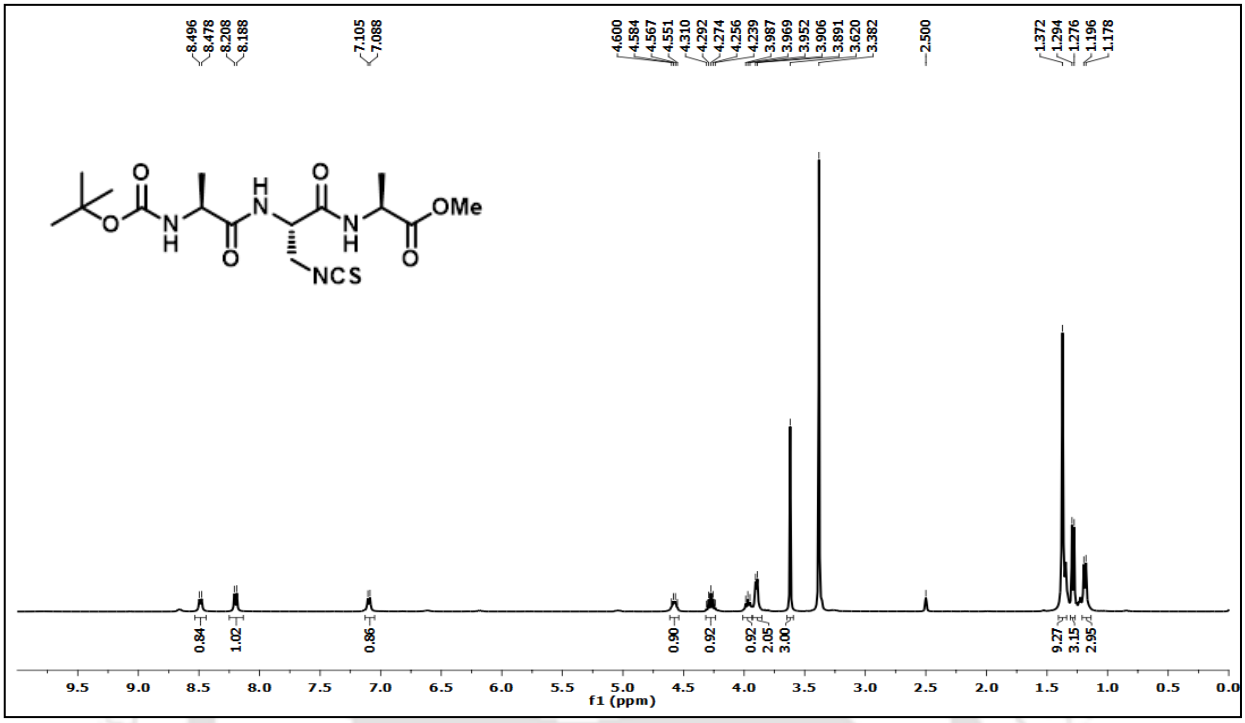


Figure 4.32. ^{1}H NMR spectra of synthesised compound **4.044.**

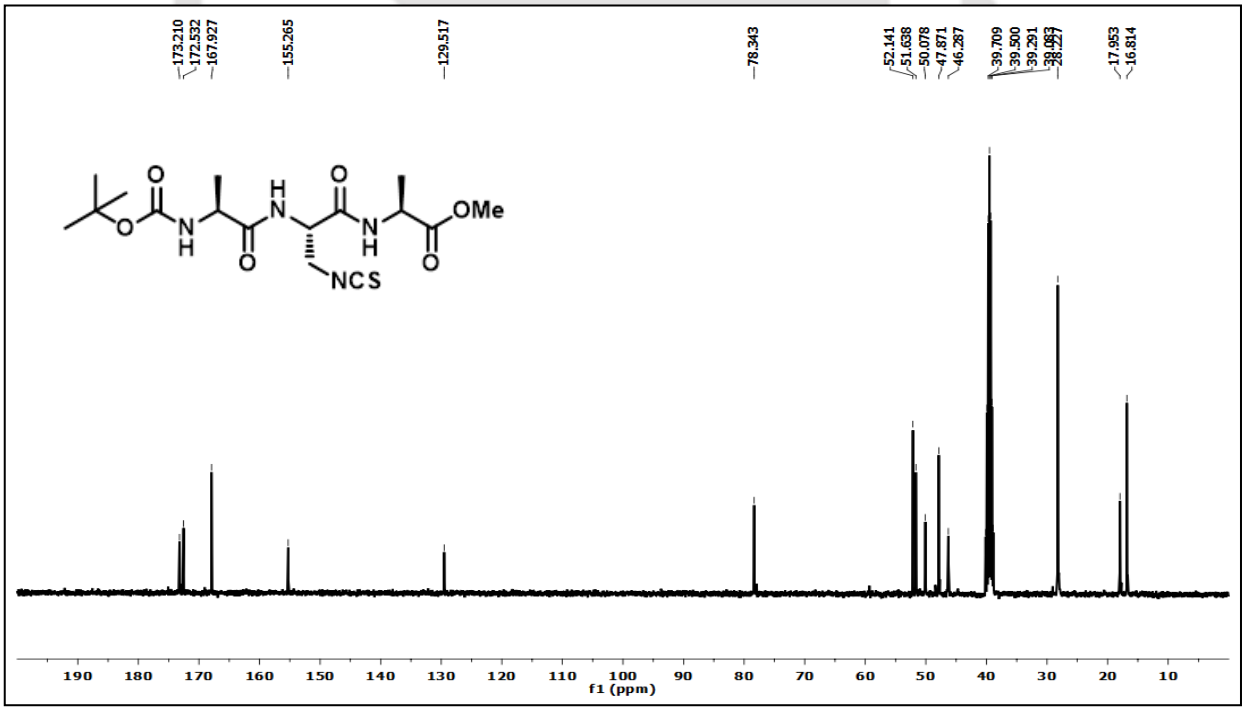


Figure 4.33. ^{13}C NMR spectra of synthesised compound **4.044.**

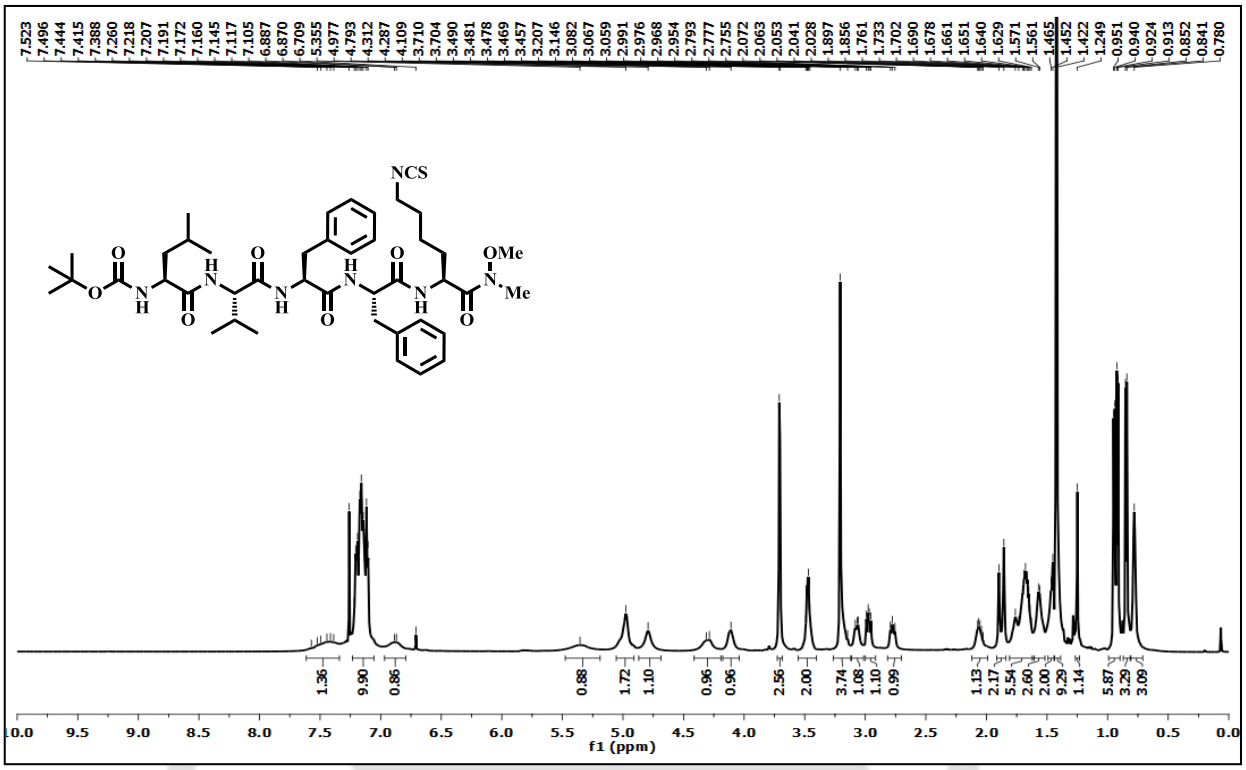


Figure 4.34. ^{1}H NMR spectra of synthesised compound **4.063.**

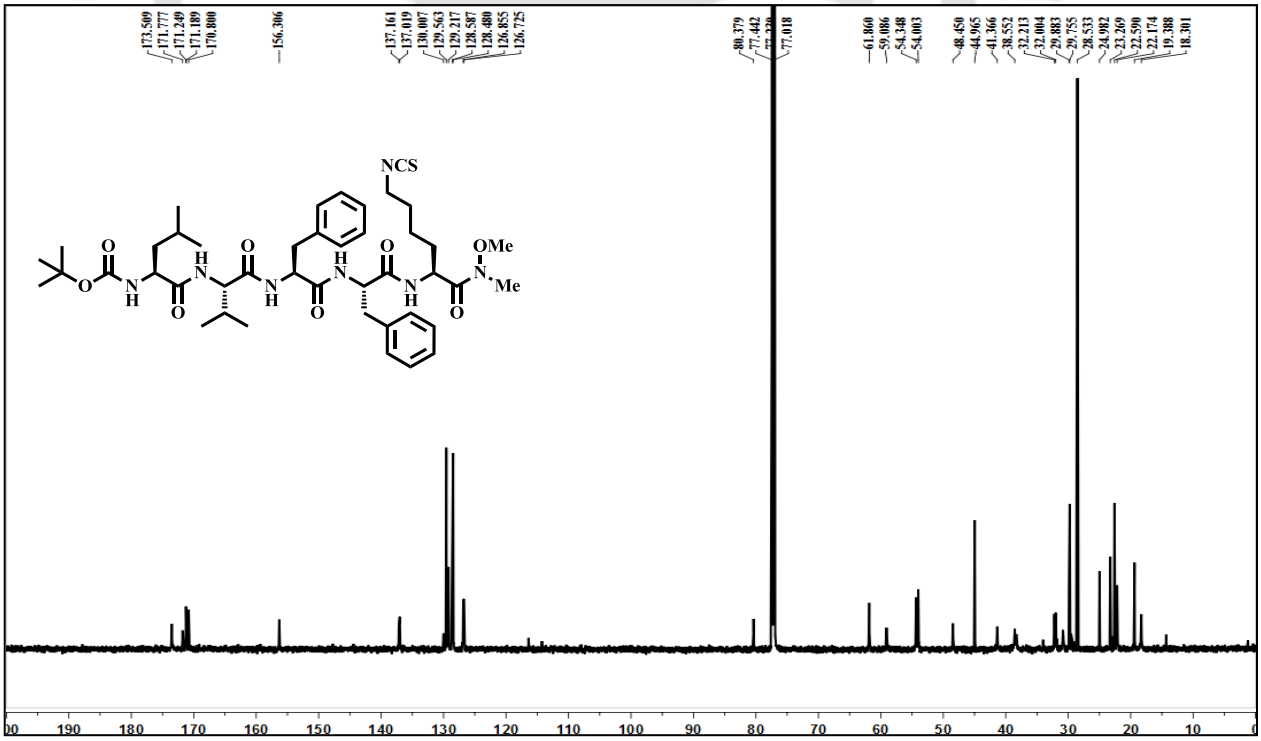


Figure 4.35. ^{13}C NMR spectra of synthesised compound **4.063.**

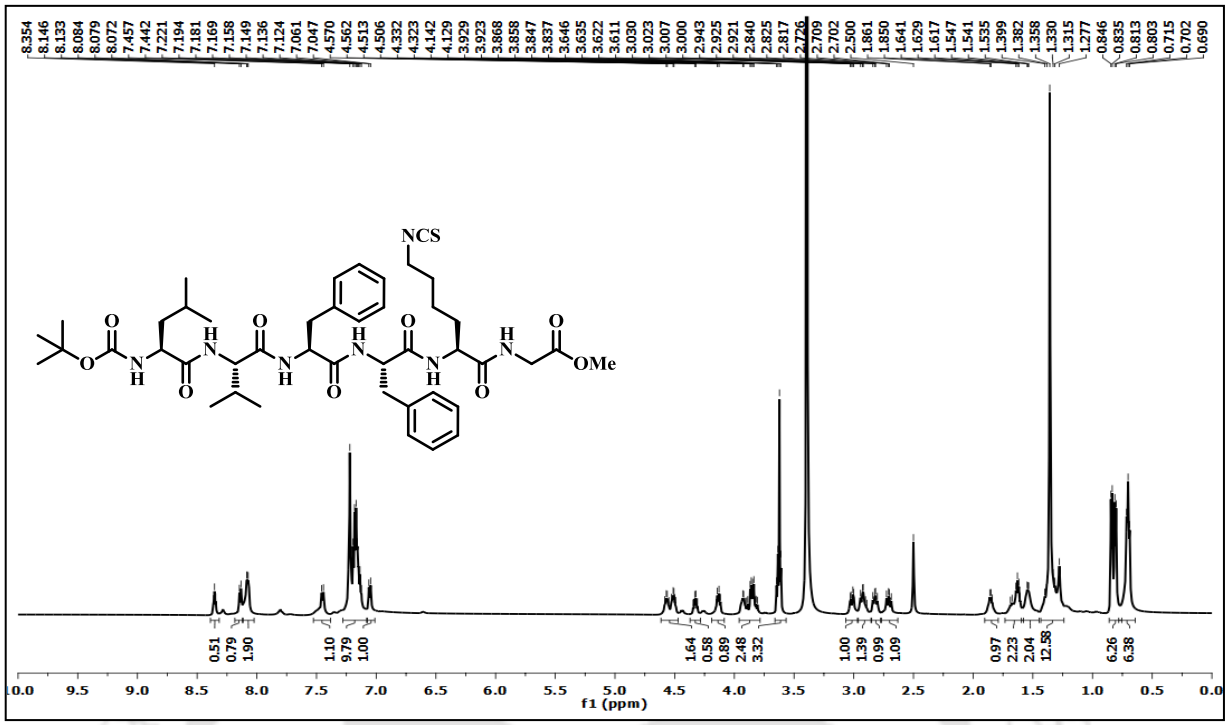


Figure 4.36. ^{1}H NMR spectra of synthesized compound **4.064.**

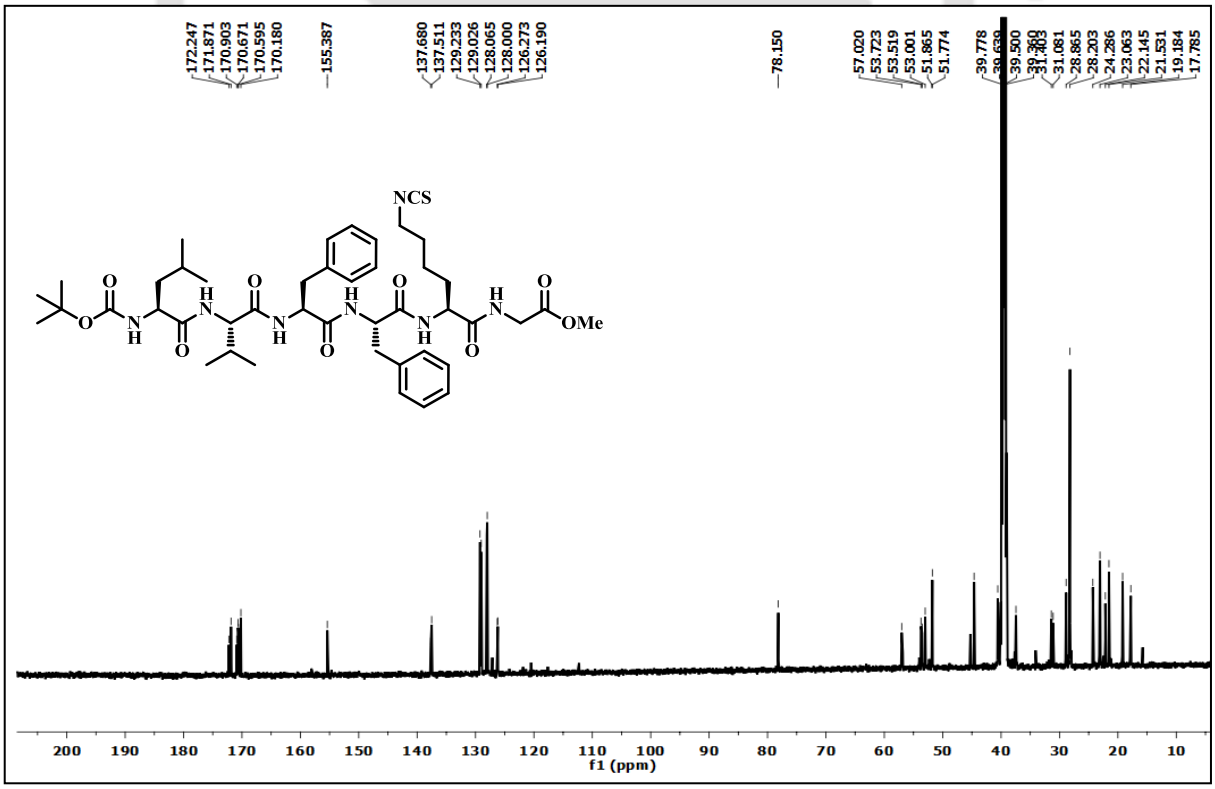


Figure 4.37. ^{13}C NMR spectra of synthesised compound **4.064.**

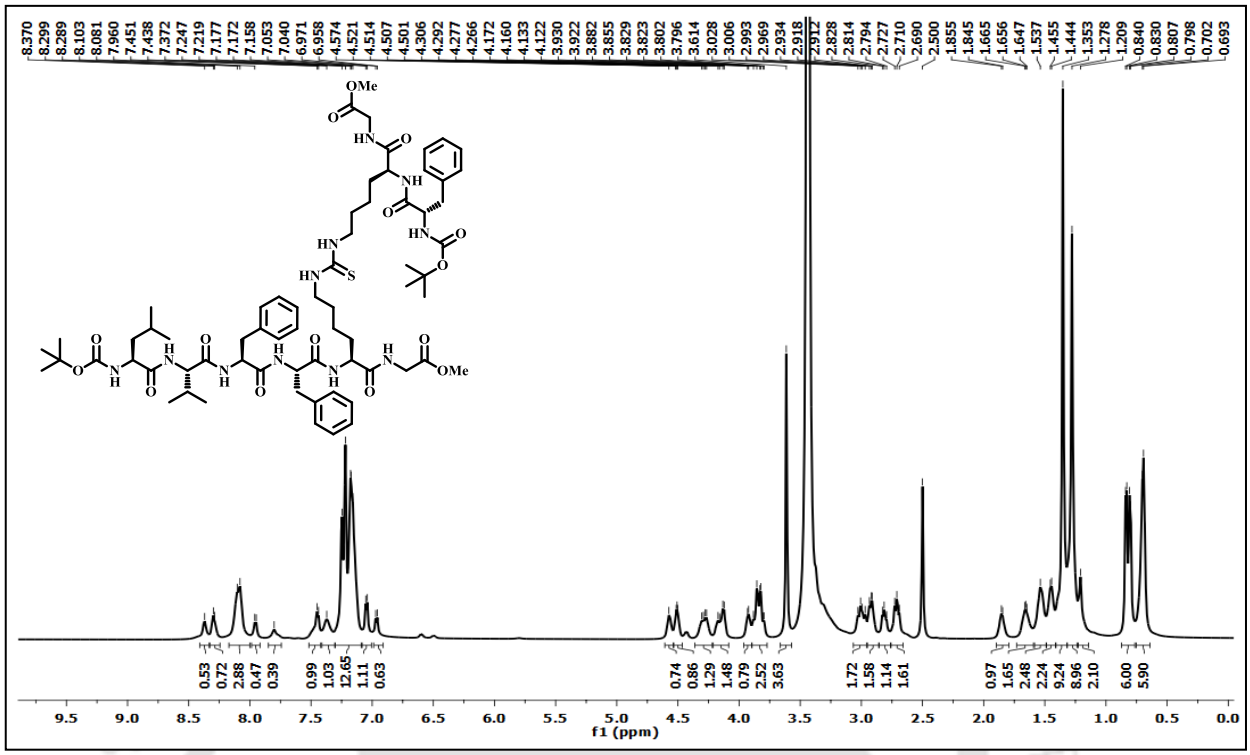


Figure 4.38. ^{1}H NMR spectra of synthesised compound **4.078**

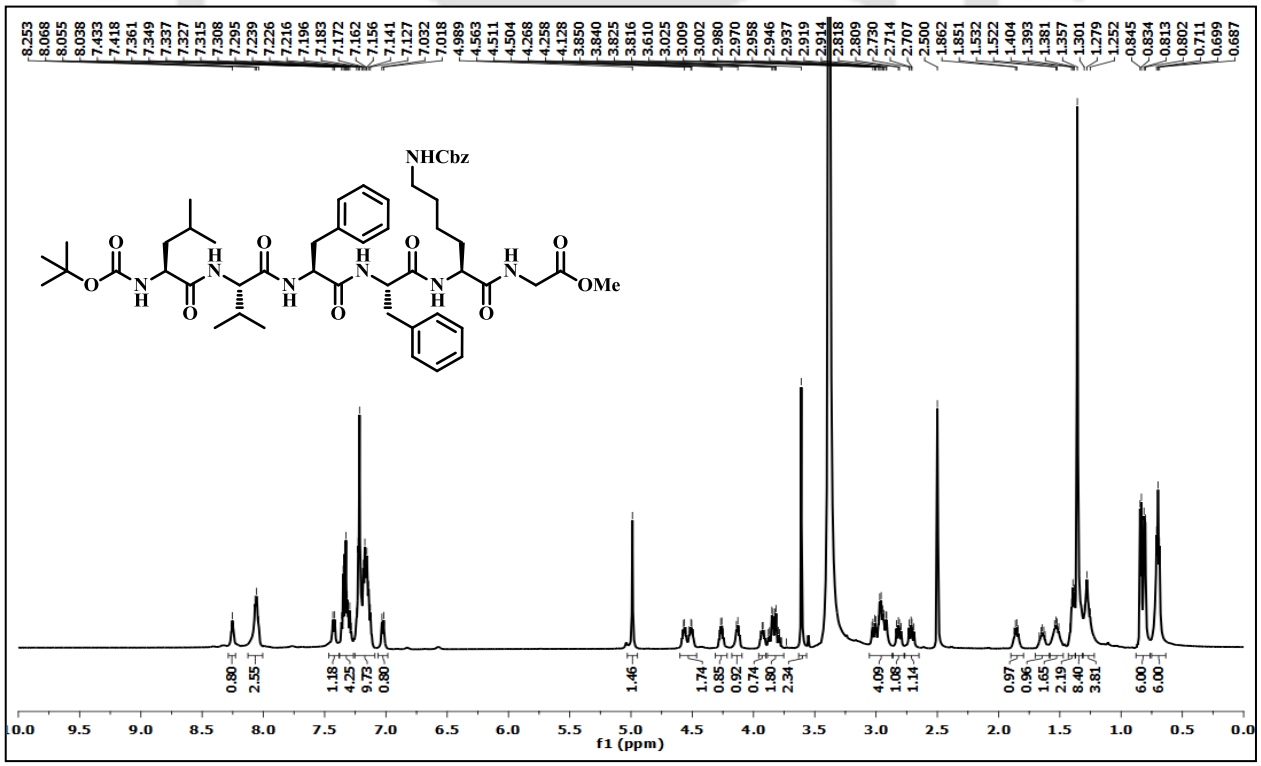


Figure 4.39. ^{1}H NMR spectra of synthesised compound **4.068.**

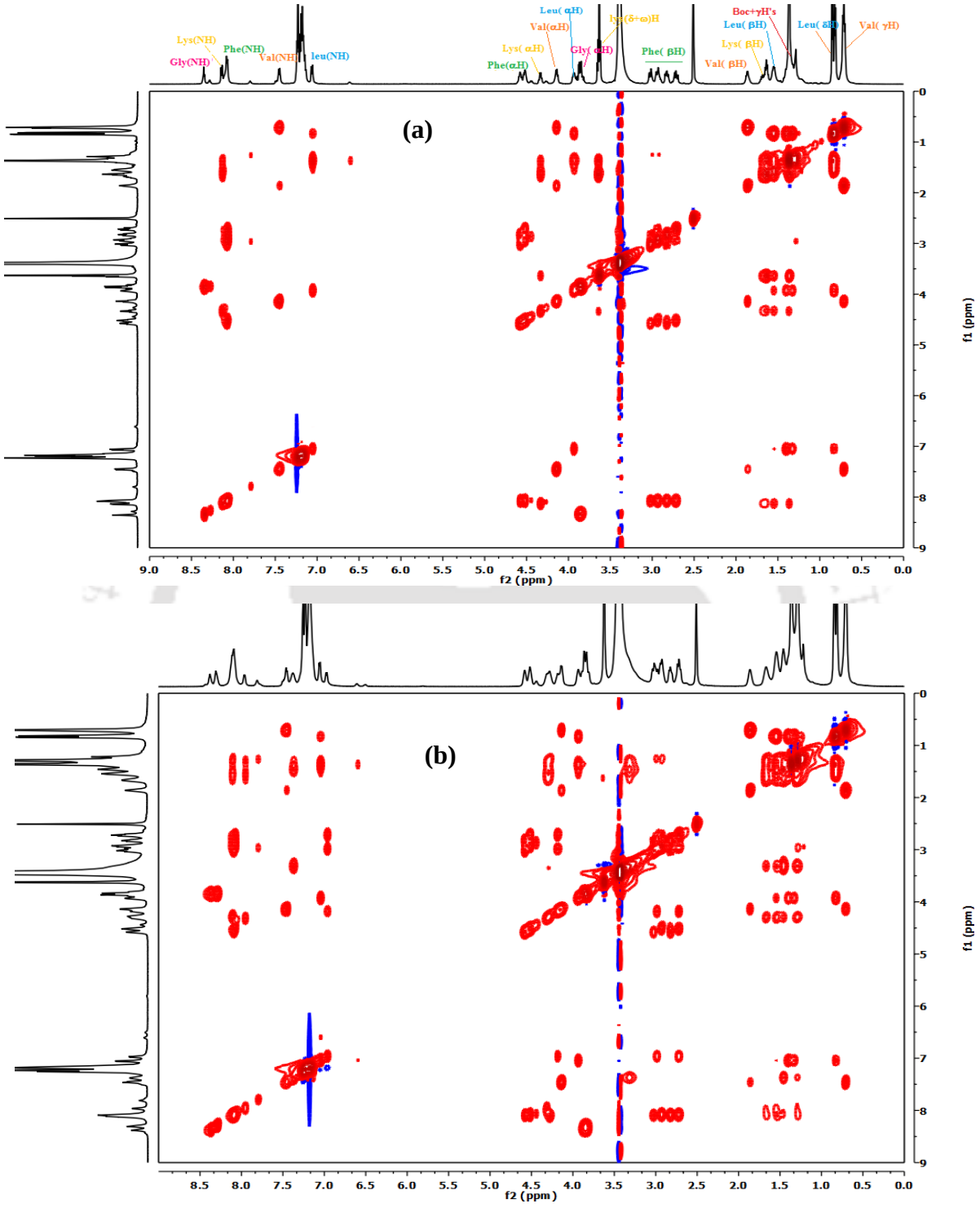


Figure 4.40. TOCSY (^{1}H-^{1}H) NMR spectra of synthesised (a) compound **4.064** and (b) compound **4.078.**

4.8. References

1. Hong, S. H.; Kwon, Y. C.; Jewett, M. C. *Front. Chem.* **2014**, *2, 34.* (b) Wissner, R. F.; Batjargal, S.; Fadzen, C. M.; Petersson, J. E. *J. Am. Chem. Soc.* **2013**, *135,* 6529. (c) Chalker, J. M.; Bernardes, G. J. L.; Davis, B. G. *Acc. Chem. Res.* **2011**, *44*, 730. (d) Lee, H. S.; Guo, J.; Lemke, E. A.; Dimla, R. D.; Schultz, P. G. *J. Am. Chem. Soc.* **2009**, *131*, 12921. (e) Liu, C. C.; Schultz, P. G. *Annu. Rev. Biochem.* **2010**, *79*, 413.(e) Ge, J.; Li, L.; Yao, S.Q. *Chem. Commun*. **2011**, **47**, 10939.
2. (c) Bains, G.; Patel, A. B.; Narayanaswami, V. *Molecules* **2011**, *16*, 7909. (d) Marmé, N.; Knemeyer, J. P.; Sauer, M.; Wolfrun, J. *Bioconjugate Chem.* **2003**, *14*, 1133. (e) Bains, G.; Patel, A. B.; Narayanaswami, V. *Biochemistry* **2011**, *16*, 7909. Chen, S.; Tsao, M. –L. *Biocon. Chem.* **2013**, *24*, 1645. (f) Sainlos, M.; Tigaret, C.; Poujol, C.; Olivier, N. B.; Bard, L.; Breillat, C.; Thiolon, K.; Choquet, D.; Imperiali, B. *Nat. Chem. Biol*. **2011**, *7*, 81. (g) Loving, G. S.; Sainlos, M.; Imperiali, B. *Trends Biotechnol*. **2010**, *28*, 73.
3. (a) Sinkeldam, R.W.; Greco, J. N.; Tor. Y. *Chem. Rev*. **2010**, *110*, 2579. (b) Lee, H. S.; Guo, J.; Lemke, E. A.; Dimla, R. D.; Schultz, P. G. *J. Am. Chem. Soc.* **2009**, *131*, 12921. (c) Shen, B.; Xiang, Z.; Miller, B.; Louie, G.; Wang, W.; Noel, J. P.; Gage, F. 60 H.; Wang, L. *Stem Cells* **2011**, *29*, 1231. (d) Sainlos, M.; Tigaret, C.; Poujol, C.; Olivier, N. B.; Bard, L.; Breillat, C.; Thiolon, K.; Choquet, D.; Imperiali, B. *Nat. Chem. Biol*. **2011**, *7*, 81. (e) Loving, G. S.; Sainlos, M.; Imperiali, B. *Trends Biotechnol*. **2010**, *28*, 73. (f) Gosavi, P. M.; Korendovych, I. V. Curr. Opin. Chem.Biol. **2016**, *34*, 103.
4. (b) Fielding, A. J.; Concilio, M. G.; Heaven, G.; Hollas, M. A. *Molecules* **2014**, *19*, 16998. (c) Jagtap, A. P.; Krstic, I.; Kunjir, N. C.; Hänsel, R.; Prisner, T. F.; Sigurdsson, S. T. *Free Radic. Res.* **2015**, *49*, 78. (d) Fleissner, M. R.; Brustad, E. M.; Kálai, T.; Altenbach, C.; Cascio, D.; Peters, F. B.; Hideg, K.; Peuker, S.; Schultz, P. G.; Hubbell, W. L. *Proc. Natl. Acad. Sci. U S A.* **2009**, *106*, 21637. (e) Schmidt, M.J.; Borbas,J.; Drescher, M.; Summerer, D. *J. Am. Chem. Soc*. **2014**, *136*, 1238. (f) Klare, J. P. *Biol Chem.* **2013**,*394*, 1281 and references therein.
5. (a) Smith, E.; Collins, I.; *Future Med. Chem.* **2015**, *7*, 159. (c) Wright, K.; Moretto, A.; Crisma, M.; Wakselman, M.; Mazaleyrat, J. P.; Formaggio, F.; Toniolo, C. *Org. Biomol. Chem.* **2010**, *8*, 3281. (d) Bush, J. T.; Walport, L. J.; McGouran, J. F.; Leung, I. K. H.; Berridge, G.; van Berkel, S. S.; Basak, A.; Kessler, B. M.; Schofield, C. J. *Chem. Sci.* **2013**, *4*, 4115.(e) Lin, S.; Zhang, Z.; Xu, H.; Li, L.; Chen, S.; Li,J.; Hao,Z.; Peng R.; Chen, P.R. *J. Am. Chem. Soc*. **2011**, *133*, 20581.
6. (a) Bag, S. S.; Jana, S.; Pradhan, M. K. *Bioorg. Med. Chem.* **2016**, *24*, 3579. (b) Bag, S. S.; Jana, S.; Pradhan, M. K. Pal, S. *RSC Adv*. **2016**, *6*, 72654. (c) Bag, S. S.; Jana, S.; Yashmeen, A.; De, S. *Chem. Commun.* **2015**, *51*, 5242. (d) Bag, S. S.; Jana, S.; Yashmeen, Senthilkumar, K.; Bag, R. *Chem. Commun.* **2014**, *50*, 433.
7. (a) Zheng, M. L.; Zheng, D. C.; Wang, J. *J. Phys. Chem.B* **2010,** *114*, 2327. (b) Watson, M. D.; Gai, X. S.; Gillies, A. T.; Brewer, S. H.; Fenlon, E. E. *J. Phys. Chem.* 61*B* **2008**, *112*, 13188. (c) Tucker, M. J.; Kim, Y. S.; Hochstrasser, R. M. *Chem. Phys.Lett.* **2009**, *470*, 80.

8. (a) Ma, J.; Pazos, I.M.; Zhang, W.; Culik, R.M.; Gai, F. *Annu. Rev. Phys. Chem.* **2015**, *66*, 357 and references therein. (b) Koziol, K. L., Johnson, P. J.; Stucki-Buchli, B.; Waldauer, S. A.; Hamm, P. *Curr. Opin. Struct. Biol.* **2015**, *34*, 1. (c) Hunt, N. T. *Chem. Soc. Rev.* **2009**, *38*, 1837. (g) Kim, H.; Cho, M. *Chem. Rev.* **2013**, *113*, 5817.
9. (a) Waegele, M. M.; Culik, R.M.; Gai, F. *J. Phys. Chem. Lett.* **2011**, *2*, 2598. (b) Bagchi, S.; Boxer, S. G.; Fayer, M. D. *J. Phys. Chem. B* **2012**, *116*, 4034. (c) Weeks, C. L.; Polishchuk, A.; Getahun, Z.; DeGrado, W.F.; Spiro, T.G. *J. Raman Spectrosc.* **2008**, *39*, 1606. (d) Nyquist, R. A. *Appl. Spectrosc.* **1990**, *44*, 1405. (f) Reimers, J. R.; Hall, L. E. *J. Am. Chem. Soc.* **1999**,*121*, 3730.
10. (a) Gai, X. S.; Coutifaris, B. A.; Brewer, S. H.; Fenlon, E. E. *Phys. Chem.Chem. Phys.* **2011**, *13*, 5926. (b) Wolfshorndl, M. P.; Baskin, R.; Dhawan, I.; Londergan, C. H. *J. Phys. Chem. B* **2012**, *116*, 1172. (c) Nydegger, M.W.; Dutta, S.; Cheatum, C. M. *J. Chem. Phys.* **2010**, *133*, 134506. (d) Dutta, S.; Rock, W.; Cook, R. J.; Kohen, A.; Cheatum, C. M. *J. Chem. Phys.* **2011**, *135*, 055106. (e) Lipkin, J. S.; Song, R.; Fenlon, E. E.; Brewer, S. H. *J. Phys. Chem. Lett.* **2011**, *2*, 1672. (f) Lieber, E.; Rao, C. N. R.; Thomas, A. E.; Oftedahl, E.; Minnis, R.; Nambury, C. V. N. *Spectrochim. Acta.* **1963**, *19*, 1135.
11. (a) Jo, H.; Culik, R. M.; Korendovych, I. V.; DeGrado, W. F.; Gai, F. *Biochemistry* **2010**, *49*, 10354. (b) Alfieri, K. N.; Vienneau, A. R.; Londergan, C. H. *Biochemistry*62**2011**, *50*, 11097. (c) Gonzalez, J. D.; Levonyak, N. S.; Schneider, S. C.; Smith, M. J.; Cremeens, M. E. *J. Mol. Struct.* **2014**, *1056-1057*, 7. (d) Stafford, A. J.; Ensign, D.L.; Webb, L. J. *J. Phys. Chem. B* **2010**, *114*, 15331. (e) Fafarman, A. T.; Boxer, S. G. *J.Phys. Chem. B* **2010**, *114*, 13536.
12. (a) Edman, P. *Acta Chem. Scand.* **1956**, *10*, 761. (b) Edman, P. Sequence determination. In *Protein Sequence Determination*; Needleman, S. B., Ed.; Springer- Verlag: BRD, 1970; pp 211−265. (c) Cohen, S. A.; Strydom, D. J. *Anal. Biochem.* **1988**, *174*, 1.
13. (a) Mukerjee, A. K.; Ashare, R. *Chem. Rev.* **1991**, 91, 1 and references therein. (b) Wu, Y.-J.; Zhang, Y. *Tetrahedron Lett.* **2008**, *49*, 2869. (c) Seitz, O. *ChemBioChem* **2000**, *1*, 215. (d) Gunther, W.; Kunz, H. *Angew. Chem., Int. Ed. Engl.* **1990**, *29*, 1050. (e) Rodriguez-Lucena, D.; Benito, J.; Ortiz Mellet, C.; Garcia Fernandez, J.M. *Chem. Commun.* **2007**, 831. (f) Gomez-Garcia, M.; Benito, J. M.; Rodriguez-Lucena, D.; Yu, J.-X.; Chmurski, K.; Ortiz Mellet, C.; Gutirrez-Gallego, R.; Maestre, A.; Defaye, J.; Garcia Fernandez, J. M. *J. Am. Chem. Soc.* **2005**, *127*, 7970. (g) Garcia Fernandez, J. M.; Ortiz Mellet, C. *Adv. Carbohydr. Chem. Biochem.* **2000**, *55*, 35. (h) Falconer, R. A.; Toth, I. Bioorg. Med. Chem. **2007**, *15*, 7012.
14. Getahun, Z.; Huang, C. Y.; Wang, T.; DeLeon, B.; DeGrado, W.F.; Gai, F. *J. Am. Chem. Soc*. **2003**, *125*, 405.
15. Schultz, K.C.; Supekova, L.;Ryu,Y.;Xie,J.; Perera,R.; Schultz, P.G. *J. Am. Chem. Soc.* **2006**, *128*, 13984.
16. Jo, H.; Culik, R. M.; Korendovych, I. V.; DeGrado, W. F.; Gai, F. *Biochemistry.* **2010**, *49*, 10354.
17. Waegele M.M.; Tucker, M.J.; Gai, F. *Chem Phys Lett*. **2009**, *478*, 249.

18. (a) Taskent-Sezgin, H.; Chung, J.; Patsalo, V.; Miyake-Stoner, S. J.; Miller, A. M.; Brewer, S. H.; Mehl, R. A.; Green, D. F.; Raleigh, D. P.; Carrico, I. *Biochemistry* **2009**, *48*, 9040.
19. Oh, K.I.; Lee, J.-H.; Joo, C.; Han, H.; Cho, M. *J. Phys. Chem. B.* **2008**, *112*, 10352.
20. Ye, S.; Zaitseva, E.; Caltabiano, G.; Schertler, G. F. X.; Sakmar, T. P.; Deupi, X.; Vogel, R. *Nature.* **2010**, *464*, 1386.
21. Taskent-Sezgin, H.; Chung, J.; Banerjee, P. S.; Nagarajan, S.; Dyer, R. B.; Carrico, I.; Raleigh, D. P. *Angew. Chem., Int. Ed*. **2010**, *49*, 7473.
22. Fafarman, A.T.; Sigala, P.A.; Herschlag, D.; Boxer, S.G. *J. Am. Chem. Soc.* **2010**, *132*, 12811.
23. Edelstein, L.; Stetz, M.A.; McMahon, H. A.; Londergan, C.H. *J. Phys. Chem. B.* **2010**, *114*, 4931.
24. McMahon, H.A.; Alfieri, K.N.; Clark, C. A.A.; Londergan, C.H. *J. Phys. Chem. Lett.* **2010**, *1*, 850.
25. (a) Fafarman,A.T.; Webb, L. J.;Chuang, J. I.; Boxer, S.G. *J. Am. Chem. Soc*. **2006**, *128*, 13356. (b) Fafarman, A. T.; Boxer, S. G. *J. Phys. Chem*. B **2010**, *114*, 13536.
26. (a) Jotterand, N.; Pearce, D. A.; Imperiali, B. *J. Org. Chem.* **2001,** *66*, 3224. (b) Ben-Efraim, I.; Strahilevitz, J.; Bach, D.; Shai, Y. *Biochemistry* **1994**, *33*, 6966. (b) Marmé, N.; Knemeyer, J. P.; Sauer, M.; Wolfrun, J. *Bioconjugate Chem.* **2003**, *14*, 1133.
27. (a) Park, E. S.; Boxer, S. G. *J. Phys. Chem. B* **2002**, *106*, 5800. (b) Woys, A. M.; Mukherjee, S. S.; Skoff, D. R.; Moran, S. D.; Zanni, M. T. *J. Phys. Chem. B* **2013**, *117*, 5009. (c) Peran, I.; Oudenhoven, T.; Woys, A. M.; Watson, M. D.; Zhang, T. O.; Carrico, I.; Zanni, M. T.; Raleigh, D. P. *J. Phys. Chem. B* **2014**, *118*, 7946. (d) Dekkers, D. W.C.; Comfurius, P.; Schroit, A. J.; Bevers, E.M.; Zwaal, R.F. A. *Biochemistry* **1998**, *37*, 14833.
28. (a) Lindquist, B. A.; Furse, K. E.; Corcelli, S. A. *Phys. Chem. Chem. Phys.* **2009**, *11*, 8119. (b) Taskent-Sezgin, H.; Chung, J.; Patsalo, V.; Miyake-Stoner, S. J.; Miller, A. M.; Brewer, S. H.; Mehl, R. A.; Green, D. F.; Raleigh, D. P.; Carrico, I. *Biochemistry* **2009**, *48*, 9040. (c) Zimmermann, J.; Thielges, M.C.; Seo, Y. J.; Dawson, P. E.; Romesberg, F.E. *Angew. Chem. Int. Ed.* **2011**, *50*, 8333. (d) Suydam, I. T.; Snow, C.D.; Pande, V.S.; Boxer, S. G. *Science* **2006**, *313*, 200. (e) Fafarman, A. T.; Boxer, S.G. *J. Phys. Chem. B **2010**, 114, 13536.* (f) Sigala, P. A.; Fafarman, A. T.; Bogard, P. E.; Boxer, S. G.; Herschlag, D. *J. Am. Chem. Soc. **2007**, 129, 12104.* (g) van Wilderen, L. J. G. W.; Kern-Michler, D.; Muller-Werkmeister, H. M.; Bredenbeck, J. *Phys. Chem. Chem. Phys.* **2014**, *16*, 19643. (h) Adhikary, R.; Zimmermann, J.; Dawson, P. E.; Romesberg, F. E. *Anal. Chem.* **2015**, *87*, 11561.
29. (a) Getahun, Z.; Huang, C. Y.; Wang, T.; León, B. D.; DeGrado, W. F.; Gai, F. *J. Am. Chem. Soc.* **2003**, *125, 405.* (b) Fafarman, A. T.; Boxer, S. G. *J. Phys. Chem. B* **2010**, *114, 13536.* (c) Getahun, Z.; Huang, C. Y.; Wang, T.; De León, B.; DeGrado, W. F.; Gai, F. *J. Am. Chem. Soc.* **2003**, *125*, *405*. (d) Waegele, M. M.; Culik, R. M.; Gai, F. *J. Phys. Chem. Lett.* **2011**, *2*, 2598. (e) Bazewicz, C.G.; Lipkin, J. S.; Smith, E. E.; Liskov, M. T.; Brewer, S. H. *J. Phys. Chem. B* **2012**, *116*, 10824. (f) Woys, A. M.; Mukherjee, S. S.; Skoff, D. R.; Moran, S. D.; Zanni, M. T. *T J. Phys. Chem. B* **2013**, *117*, 5009. (g) Ghosh, A.; Tucker, M. J.; Gai, F. *J. Phys. Chem. B* **2014**, *118*, 7799. (h) Ding, B.;

Panahi, A.; Ho, J.-J. Laaser, J. E.; Brooks, III, C. L.; Zanni, M. T.; Chen, Z. *J. Am. Chem. Soc.* **2015**, *137*, 10190. (i) Shrestha, R.; Cardenas, A. E.; Elber, R.; Webb, L. J. *J. Phys. Chem. B* **2015**, *119*, 2869. (j) Deb, P.; Haldar, T.; Kashid, S. M.; Banerjee, S.; Chakrabarty, S.; Bagchi, S. *J. Phys. Chem. B* **2016**, *120*, 4034. (k) Gao, Y.; Zou, Y.; Ma, Y.; Wang, D.; Sun, Y.; Ma, G. *Langmuir* **2016**, *32*, 937.

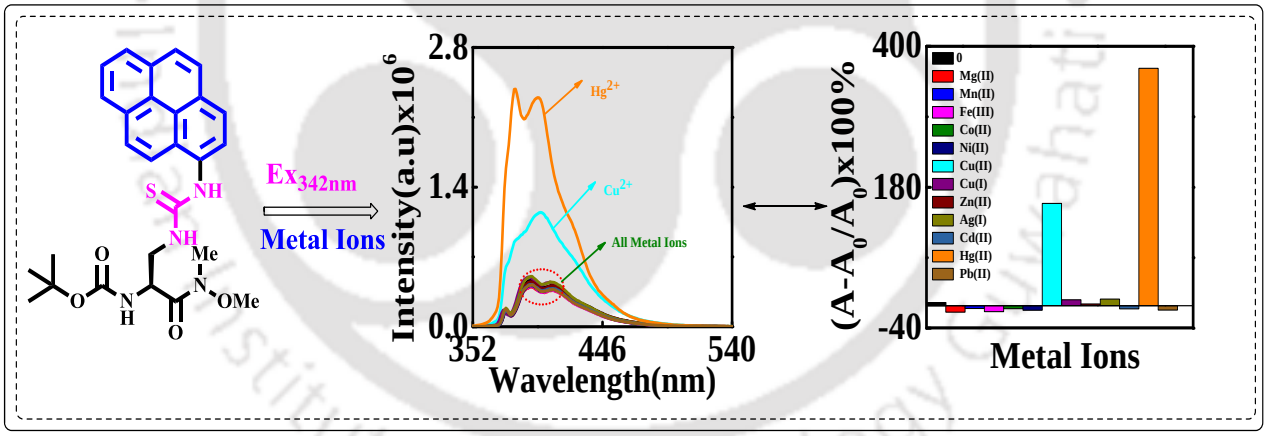
30. (a) van Wilderen, L. J. G.; Kern-Michler, W. D.; Müller-Werkmeister, H. M.; Bredenbeck, J. *Phys. Chem. Chem. Phys.* **2014**, *16*, 19643. (b) Peran, I.; Oudenhoven, T.; Woys, A. M.; Watson, M. D.; Zhang, T.O.; Carrico, I.; Zanni, M. T.; Raleigh D. P. *J. Phys. Chem. B* **2014**, *118*, 7946. (c) Shrestha, R.; Cardenas, A. E.; Elber, R.; Webb, L.J. *J. Phys. Chem. B* **2015**, *119*, 2869.
31. (a) Schweitzer-Stenner, R.; Eker, F.; Huang, Q.; Griebenow, K. *J. Am. Chem. Soc.* **2001**, *123*, 9628. (b) Eker, F.; Cao, X.; Nafie, L.; Schweitzer-Stenner, R. *J. Am. Chem. Soc.* **2002**, *124*, 14330. (c) Woutersen, S.; Pfister, R.; Hamm, P.; Mu, Y.; Kosov, D.; Stock,G. *J. Chem. Phys.* **2002**, *117*, 6833. (d) Eker, F.; Griebenow, K.; Schweitzer-Stenner, R. *J. Am. Chem. Soc.* **2003**, *125*, 8178. (e) Toal, S.; Meral, D.; Verbaro, D.; Urbanc, B.; Schweitzer-Stenner, R. *J. Phys. Chem. B* **2013**, *117*, 3689. (f) Feng, Y.; Huang, J.; Kim, S.; Shim, J. H.; MacKerell, Jr., A. D.; Ge, N.-H.; *J. Phys. Chem. B* **2016**, 120, 5325 and references therein. (g) Bastida, A.; Zuniga, J.;Requena, A.; Miguel, B.; Candela, M. E.; Soler, M. A. *J. Phys. Chem. B* **2016**, *120*,348.
32. (a) Toal, S.; Meral, D.; Verbaro, D.; Urbanc, B.; Schweitzer-Stenner, R. *J. Phys. Chem. B* **2013**, *117*, 3689. (b) Liu, Z.; Chen, K.; Ng, A.; Shi,Z.; Woody, R.W.; Kallenbach, N. R.; *J. Am. Chem. Soc.* **2004**, *126*, 15141. (c) Eker, F.; Griebenow, K.; Schweitzer-Stenner, R. *J. Am. Chem. Soc.* **2003**, *125*, 8178. (d) Madison, V.; Kopple, K. D. *J. Am. Chem. Soc.* **1980**, *102*, 4855. (e) Rao, B. N. N.; Kumar, A.; Balaram, H.; Ravit, A.; Balaram, P. *J. Am. Chem. Soc.* **1983**, *105*, 7423. (f) Oh, K.-I ; Kim, W.; Joo, C.; Yoo, D.-G.; Han, H.; Hwang, G.-S.; Cho, M. *J. Phys. Chem. B* **2010**, *114*, 1302 (g) Han, C.; Wang, J.; *ChemPhysChem*, **2012**, *13*, 1522. (h) Shi, Z.; Chen, K.; Liu, Z.; Kallenbach, N. R. *Chem. Rev.* **2006**, *106*, 1877.
33. (a) Madison, V.; Kopple, K.D. *J. Am. Chem. Soc.* **1980**, *102*, 4855. (b) Narasinga Rao, B. N.; Kumar, A.; Balaram, H.; Ravit, A.; Balaram, P. *J. Am.Chem. Soc.* **1983**, *105*, 7423. (c) Oh, K.-I.; Kim, W.; Joo, C.; Yoo, D.-G.; Han, H.; Hwang, G.-S.; Cho, M. *J. Phys. Chem. B* **2010**, *114*, 13021. (d) Kong, J.; Yu, S. *Acta. Biochim. Biophys. Sin.* **2007**, *39*, 549.(e) Feher-Voelger, A.; Borges-Gonzlez, J.; Carrillo, R.; Morales, E.Q.; Martn, J.G.-P.T. *Chem. Eur. J.* **2014**, *20*, 1.(f) Angell, Y.; Burgess, K. *J. Org. Chem.* **2005**, *70*, 9595. (g) Chouhan, G.; James, K. *Org. Lett.* **2013**, *15*, 6. (h) Chakraborty, T. K.; Rao, K. S.; Kiran, M. U.; Jagadeesh, B. *Tetrahedron Lett.* **2009**, *50*, 4350.
34. (a) Ma, J.; Pazos, I.M.; Zhang, W.; Culik, R.M.; Gai, F. *Annu. Rev. Phys. Chem.* **2015**, *66*, 357 (b) Gao, Y.; Zou, Y.; Ma,Y.; Wang, D.; Sun,Y.; Ma, G. *Langmuir* **2016**, *32*, 937.
35. (a) Waegele, M. M.; Culik, R. M.; Gai, F. *J. Phys. Chem. Lett.* **2011**, *2*, 2598. (b) Bagchi, S.; Boxer, S. G.; Fayer, M.D. *J. Phys. Chem. B.* **2012**, *116*, 4034. (c) Weeks, C.L.; Polishchuk, A.; Getahun, Z.; DeGrado, W. F.; Spiro, T. G. *J. Raman Spectrosc.* **2008**, *39*, 1606. (d) Schultz, K. C.; Supekova, L.; Ryu, Y.; Xie, J.; Perera, R.; Schultz, P. G. *J. Am. Chem. Soc.* **2006**, *128*, 13984. (e) Gai, X. S.; Coutifaris, B. A.; Brewer, S. H.;

Fenlon, E. E. *Phys. Chem.Chem. Phys.* **2011**, *13*, 5926. (f) Nyquist, R. A. *Appl. Spectrosc.* **1990**, *44*, 1405. (g) Reimers, J. R.; Hall, L.E.; *J. Am. Chem. Soc.* **1999**, *121*, 3730.

36. (a)Yang, M.; Li, P. R. J.; Chen, *Chem. Soc. Rev.* **2014**, *43*, 6511. (b) Grammel, M.; Hang, H. C.; *Nat. Chem. Biol.* **2013**, *9*, 475. (c) Ni, Z.; Zhou, L.; Li, X.; Zhang, J.; Dong, S.; *PLoS One* **2015**, *10*, e0141918. (d) Dumas, A.; Lercher, L.; Spicer, C. D.; Davis, B. G. *Chem. Sci.* **2015**, ***6***, 50. (e) Gunnoo, S. B.; Madder, A. *Org. Biomol. Chem.* **2016**, *14*, 8002. (f) Yin, J.; Liu, F.; Li, X.; Walsh, C. T.; *J Am Chem. Soc.* **2004**, *126*, 7754. (g) Chen, I.; Howarth, M.; Lin, W.; Ting, A. Y.; *Nat Methods* **2005**, *2*, 99.
37. (a) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. *Science* **1994**, *266*, 776. (b) Kent, S. B.; *Chem. Soc. Rev.* **2009**, *38*, 338. (c) Zhang, Y.; Xu, C.; Kam, H. Y.; Lee, C. L.; Li, X. *Proc. Natl. Acad. Sci. USA.* **2013**, *17*, 6657.
38. MacroModel, Version 9.0 Schrodinger, LLC, New York, NY, **2009**. (b) Maestro,version 9.0, Schrödinger, LLC, New York, NY, **2009**. (c) Lee, H.-J.; Park, H.-M.;Lee, K.-B. *Biophys. Chem*. **2007**, *125*, 117. (d) Feher-Voelger, A.; Borges-Gonzlez, J.; Carrillo, R.; Morales, E. Q.; Gonzlez-Platas, J.; Martn, T. *Chem. Eur. J.* **2014**, *20*, 1. (e) Shivakumar, D.; Williams, J.; Wu, Y.; Damm, W.; Shelley, J.; Sherman, W. *J. Chem. Theory Comput.* **2010**, *6*, 1509. (f) Gnanakaran, S.; García, A. *Proteins: Struc.Func. Bioinform.* **2005**, *59*, 773. (g) Wang, L.; Berne, B. J.; Friesne, R**.** A. *PNAS.* **2012**, *109*, 1937.
39. (a)Madison, V.; Schellman, J. *Biopolymers* **1972**, *11*, 1041. (b) Lee, K.-K.; Kim, E.; Joo, C.; Song, J.; Han, H.; Cho, M. *J. Phys. Chem. B* **2008**, *112*, 16782. (c) Shi, Z.; Chen, K.; Liu, Z.; Kallenbach, N. R. *Chem. Re*v. **2006**, *106*, 1877.
40. de Groot, N.S.; Parella, T.; Aviles, F.X.; Vendrell, J.; Ventura, S.; *Biophys J.* **2007,** *92*, 1732. (b) Adler-Abramovich, L.; Gazit, E. *Chem. Soc. Rev*. **2014**, *43*, 6881.

Chapter-5

Studies on the Switch-on Fluorescence Sensing for Hg^{2+} and Cu^{2+} Ions by Pyrenylthioureayl Alanine Amino Acid



5.1. Introduction

In medicinal chemistry, fluorescent sensors are very important because of their significance to fulfill the necessity of *in vivo* probes, such as mapping the spatial and temporal distribution of the biological analytes. They have other advantages including extremely high sensitivity, moderately low cost and easy accessibility, as well as several modes of detection. Fluorescent chemosensors that display an increase in the fluorescence intensity due to the effect of a coordinated metal ion is known as a chelation-enhanced fluorescence (CHEF) effect (**Figure 5.01)**.[1a-d] Such type of sensors have a lone pair with suitable energy, which can reduce the fluorescence by an asset of the photoinduced electron transfer (PET) effect (**Figure 5.01**).[19]

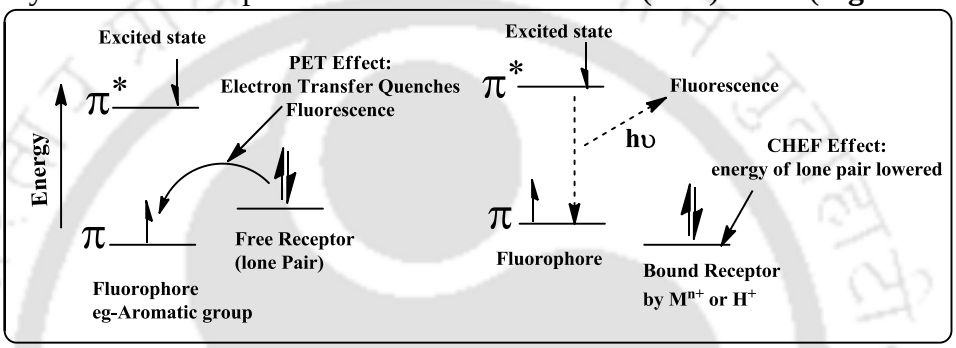


Figure 5.01. PET and CHEF effect on flurosence[1a].

Research in the field of fluorescence sensing of heavy metal as well as transition metal ions is rapidly increasing because of their adverse effects on biological systems and on environment.[2] Among the various heavy metal ions-based pollutants, the mercury ion (Hg^{2+}) is in the frontline for its adverse impact on the environment and human health.[3] On the other hand, the transition metal ion copper (Cu^{2+}) is important for some physiological processes.[3a-b] However, an excess amount of Cu^{2+} is highly toxic to organisms.[4c-d] Therefore design of fluorescent chemosensors for the sensing of both Hg^{2+} and Cu^{2+} is highly important for the protection of both environment and biological systems. A tremendous number of literatures exists which report the design of various types of chemosensors for the detection of either Cu^{2+} or Hg^{2+} ion.[5] Many efforts have been put forth in recent times to design the fluorescence sensors specific for the sensing of both Hg^{2+} and Cu^{2+} ion detection.[6] Few reports of chemosensors also exist which relies the sensing *via* change in both color and fluorescence.[7] However, many of them suffer from shortcomings such as fluorescence quenching by paramagnetic Cu^{2+} and Hg^{2+} ions mostly due to electron transfer mechanism [8] and spin–orbit coupling[8] respectively. Moreover, many organic chemosensors for Hg^{2+} have poor water solubility and suffers from surrounding effects and by the interference of other metal ions with low limit of detection.[9] Therefore, the development of fluorescence switch-on chemosensors for the detection of Cu^{2+} and/or Hg^{2+} is highly desirable.

5.2. Some Fluorescence Switch-on Sensors for Cu^{2+} Ion

For most of the reported Cu^{2+} fluorescent sensors, it is found that the binding to the Cu^{2+} ion causes "Turn off" fluorescence due to its paramagnetic nature [8]. As a result of various research efforts, a large number of fluorescence turn-off sensors for Cu^{2+} ion have been developed and reported. There exist a few reports of sensors where the binding of the Cu^{2+} ion affects an increase in the fluorescence intensity of the probe. Here, in this section, a few such switch-on fluorescent sensors for Cu^{2+} ion are presented in very brief.

Yoon *et al*., [10a] have reported a fluorescence "Turn-On" chemodosimeter (**5.001, Figure 5.02**) for Cu^{2+} ion which effectively recognized Cu^{2+} ion *via* a selective hydrolysis of the acetyl group. Qian *et al*., [10b-c] have reported a ratiometric and selective fluorescent sensor (**5.002-5.003, Figure 5.02**) which showed a large red-shift in emission in presence of Cu^{2+} ion which in turn acts as an effective sensor for a Cu^{2+} ion. A new pyrene-based probe, 1,4-bis(1-pyrenyl)-2,3-diaza-1,3-butadiene (**5.004, Figure 5.02**), which selectively senses Hg^{2+} and Cu^{2+} through the enhancement of the fluorescence with the red shift of the excimer emission has been developed by Molina *et al*.[10d] Later on, they reported another fluorescent probe, 2-aza-1,3-butadiene derivatives[10e] containing a pyrene unit (**5.005, Figure 5.02**) which detects Cu^{2+} *via* an enhancement of fluorescence intensity in an aqueous environment.

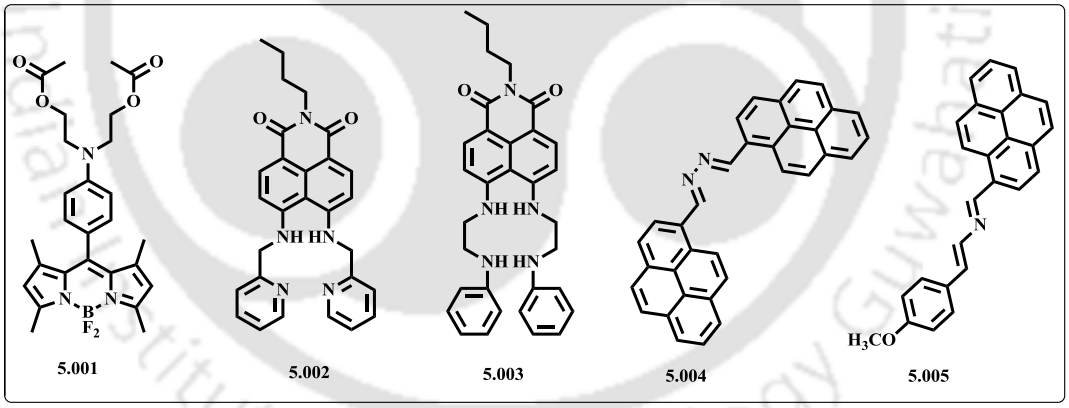


Figure 5.02. Examples of some turn-on probe for the detection of Cu^{2+} ion.

Shiraishi *et al*.,[10f] have reported a new rhodamine derivative containing an ethylenediamine-*N*,*N*-diacetic acid moiety and used it in Cu^{2+} ion sensing (**5.006, Figure 5.03**). The compound **5.006** showed strong green fluorescence in presence of Cu^{2+} ion, while with other metal ions it showed very weak orange fluorescence. Kim *et al*.,[10J] have reported a mono pyrenylalkylamine derivative (**5.007, Figure 5.03**) which exhibited an excimer emission at 455 nm along with a weak monomer emission at 375 nm in presence of Cu^{2+} ion in 1:1 (V/v) H_2O/CH_3CN.

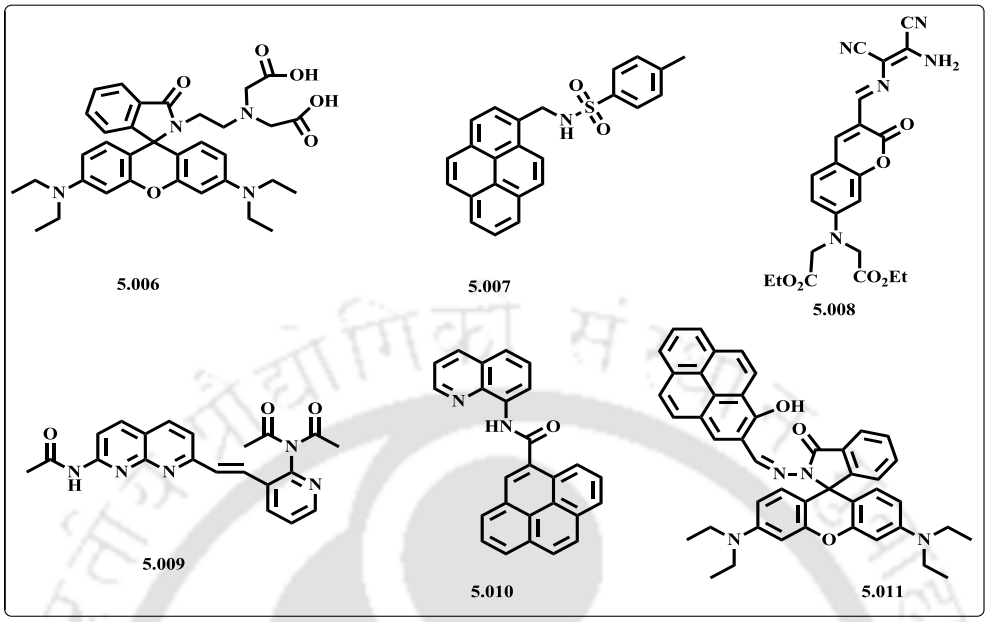


Figure 5.03. Examples of some turn-on probe for the detection of Cu^{2+} ion

A coumarin-based colorimetric chemosensor has also been developed by Wang *et al*., [10g] (**5.008, Figure 5.03**), which exhibited good sensitivity and selectivity for the Cu^{2+} ion over other tested cations (transition metal, alkali and alkaline earth metal cations) in aqueous solution. A 1,8-naphthyridine-based fluorescent chemodosimeter (**5.009, Figure 5.03**), for the selective detection of Zn^{2+} and Cu^{2+} ion, have been reported by Wei *et al*.[10h] The compound **5.009** exhibited a dual emissive behaviour in presence of Zn^{2+} ion and a fluorescent "on-off" behaviour upon addition of Cu^{2+} ion. Ham *et al*.,[10K] have reported a sensor containing a pyrene derivative **5.010** for the Cu^{2+} ion. In CH_3CN, compound **5.010** exhibited a Cu^{2+} ion induced strong static excimer emission at 460 nm, along with a weak monomer emission at 388 nm. Yoon *et al*., [10i] reported a rhodamine-based (**5.011, Figure 5.03**) ratiometric and fluorescent "on-off" chemosensor for the Cu^{2+} ion. The compound **5.011** showed a selective and chelation enhanced ratiometric fluorescence change and colorimetric change with Cu^{2+} ion over the other metal ions examined.

5.3. Some Fluorescence Switch-on Sensors for Hg^{2+} Ion

Although Hg^{2+} ion acts as a fluorescence quencher for many of the fluorophore [8], there exists a few reports of sensors where the binding of the Hg^{2+} ion results in an increase in the fluorescence intensity of the probe. Here, in this section, a few such switch-on fluorescent sensors for Hg^{2+} ion are presented in very brief.

Peng *et al*.,[11a] have designed and synthesised chemodosimeter (**5.012, Figure 5.04**), which is obtained from rhodamine B and it has the same structure to that of isopropenyl acetate. This probe effectively detects Hg^{2+} ion in ethanol–H_2O solutions. In 2012, Das *et al*., [11b] have prepared rhodamine 6G based sensor and its different derivatives (**5.013-5.016, Figure 5.04)** to achieve different fluorescence emission spectrum. These kinds of probe effectively detect Hg^{2+} ion in CH_3CN–aqueous HEPES buffer at pH 7.

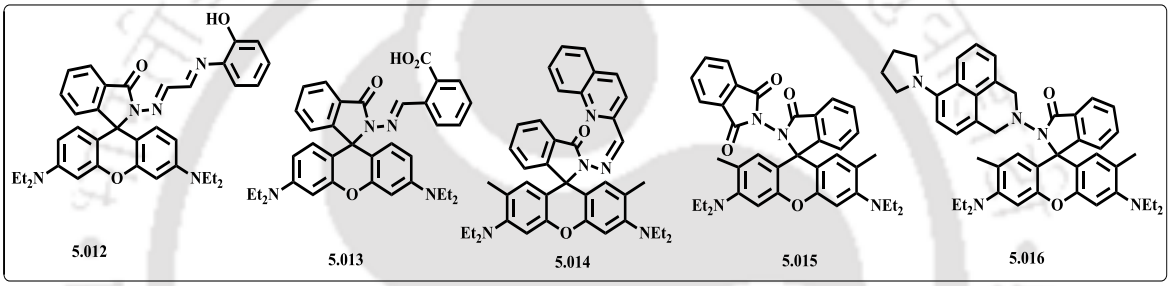


Figure 5.04. Examples of some turn on probe for the detection of Hg^{2+} ion.

Chellappa *et al*., [11c] have prepared a new pyrene-based ‘‘turn on’’ fluorescent sensor (**5.017, Figure 5.05**) which showed significantly enhanced fluorescence intensity in the presence of Hg^{2+} ions and a high selectivity towards Hg^{2+} ions over a wide range of various metal ion in an aqueous medium. In 2014, N.Wanichacheva *et al*.,[11d] have synthesised non-fluorescent 2-[3-(2-aminoethylsulfanyl)propylsulfanyl]ethanamine (**5.018, Figure 5.05**) which is covalently attached to one moiety of rhodamine-6G. It showed turn on fluorescence selectivity and sensitivity to Hg^{2+} ion over other various competitive metal ions.

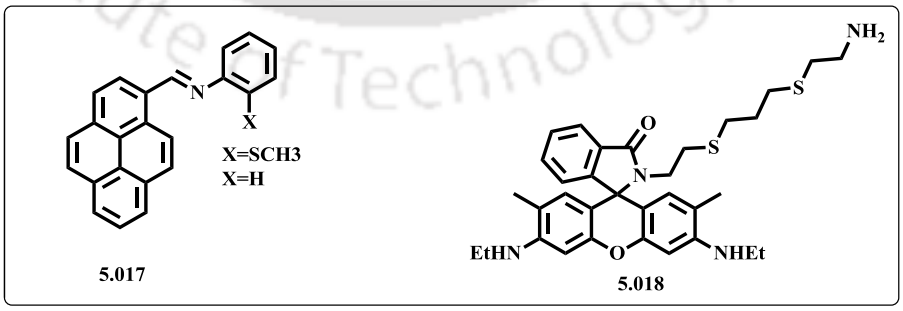


Figure 5.05. Examples of some turn on probe for the detection of Hg^{2+} ion.

Ye *et al*., [11e] have reported rhodamine-5-ferrocenyl-1,3,4 thiadiazol as the fluorescent sensor (**5.019, Figure 5.06**) for Hg^{2+} in living cells. The probe showed high sensitivity and selectivity for Cu^{2+} and Hg^{2+} over several common alkali, alkaline earth, and transition metal ions. The probe was effectively used for cell imaging in HeLa cells in presence of Hg^{2+} ion. Wu *et al*.,[11f] have prepared Schiff-base colorimetric sensor (**5.020, Figure 5.06**). Not only it showed a fluorescence 'off-on' mode in the presence of Hg^{2+} ion with high selectivity and sensitivity but also change in color is visualized by the naked eye in presence of Hg^{2+} ion. In 2012, Chen *et al*.,[11g] have synthesised new fluorophore, a cyclen moiety (**5.021,Figure 5.06**) as the ionophore which showed a turn-on fluorescent sensor and displayed an effective and selective sensing ligand for Zn^{2+}, Cd^{2+}, and Hg^{2+} ions under the physiological condition with the help of cysteine as an auxiliary reagent. In 2014,[11h] De la Cruz-Guzman *et al*., have synthesised a rhodamine organosilane derivative (Rh-UTES) (**5.022, Figure 5.06**) which shows selective turn-on fluorescent solid-sensor for Hg (II) detection.

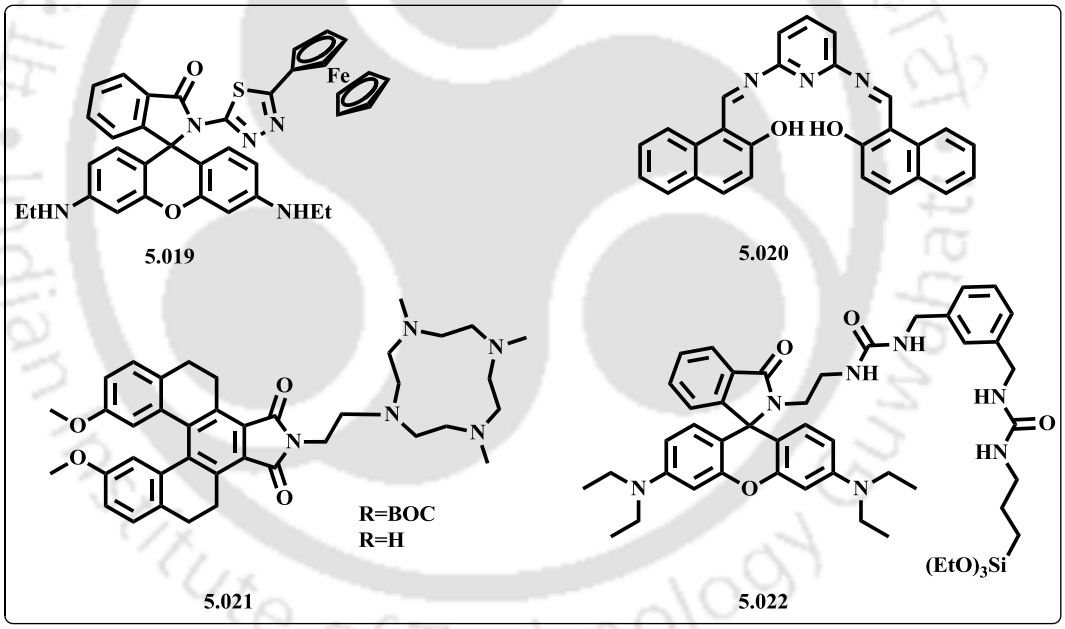


Figure 5.06. Examples of some turn-on probe for the detection of Hg^{2+} ion.

5.4. Thiourea-based Fluorescence Switch-on Metal Sensors

Dwivedi *et al*.,[12a] have synthesised a thiourea-based benzothiazole–naphthalene (BTNP) (**5.023, Figure 5.07**) probe for the detection of Hg^{2+} and Cu^{2+} transition-metal ions and it shows higher sensitivity compared to other transition metal ions. They have demonstrated BTNP probe act as an effective fluorescent ‘‘On–Off’’ probe in an aqueous medium. Velmathi *et al*., [12b] have synthesised a simple thiourea linked with amino Pyrene (**5.024, Figure 5.07**). It not only acts as a fluorescence sensor but also naked eye sensor for the detection of Cu^{2+} and Hg^{2+} ions. The complexes showed good binding constants for Cu^{2+} and Hg^{2+} with stoichiometry ratio 1:1 for both Hg^{2+} and Cu^{2+}.

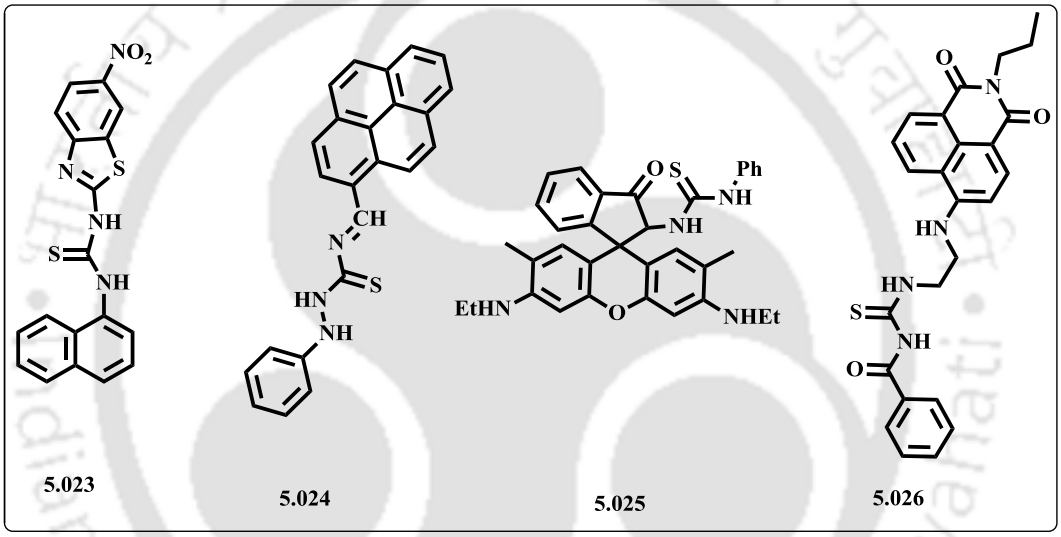


Figure 5.07. Thiourea-based switch-on metal fluorescence sensor for Hg^{2+} and Cu^{2+} ion.

Yoon *et al*., [12c-e] have developed a thiohydrazide derivative of rhodamine B ligand (**5.025, Figure 5.07**) which showed 26 fold increment of fluorescence intensity upon addition of 1 equivalent 1mM solution of Hg^{2+} to a solution of the ligand in water–methanol (80/20 V/v). It is also able to detect Hg^{2+} ion below 2 ppb. Liu *et al*., [12f] have reported 1,8-Naphthalimidederived N-acylthioamide, **5.026**. It undergoes a sulfur-to-nitrogen exchange reaction promoted by Hg^{2+} ion and forms the guanylated compound which emits fluorescence at a different wavelength. Thus, this conversion allows a ratiometric detection of Hg^{2+} ions.

Finney *et al*., [12g] have reported that thiourea-naphthalimide conjugates (**5.027, Figure 5.08**) with extended binding areas act as metal-responsive fluorescent chemosensors in aqueous media. They have found that it is very useful turn-on response to Zn^{2+}, Cd^{2+} and Hg^{2+} ions at nanomolar (nM) concentration. They have also shown that the ligand is a good optical imaging agent for the Hg^{2+} in live cells. Zhu *et al*.,[12h] have synthesised N, N′ bis(1-phenylthiouredoethyl)-1,6,7,12-tetrakis(4-methylphenoxy)perylene-3,4,9,10-tetracarboxylic (**5.028**) and studied with the focus on the photophysical, chemodosimetric mechanism, as well as fluorogenic behaviours toward various metal cations. This new "turn-on" fluorescent chemodosimeter **5.028** shows high sensing for the Hg^{2+} ion over other metal cations.

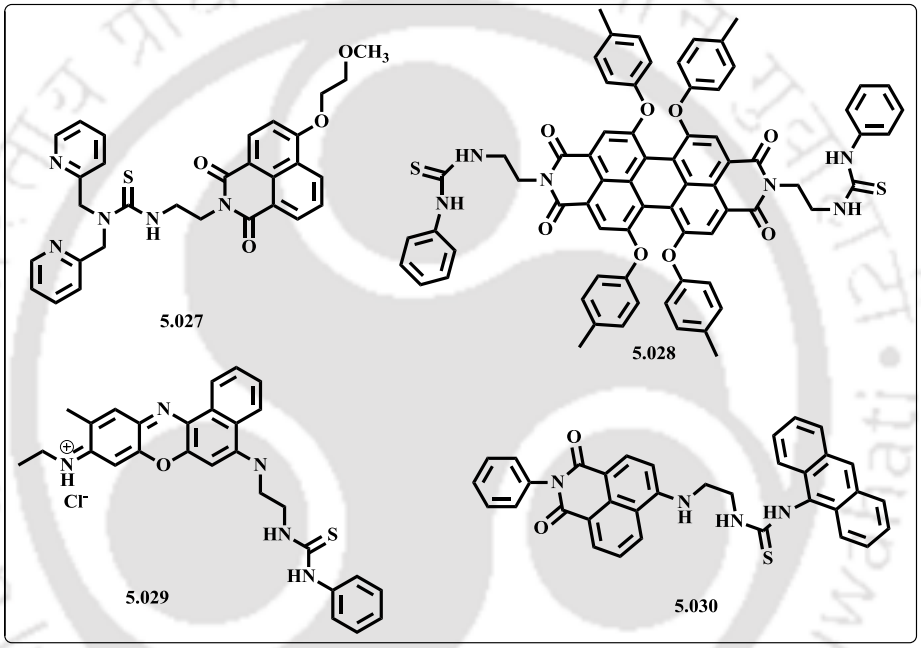


Figure 5.08. Thiourea-based switch-on metal fluorescence sensor for Hg^{2+} ion.

Nile-blue-based chemodosimeter, [12i] **5.029,** was prepared for detection of Hg^{2+} ions in 100% aqueous solution. On addition of Hg^{2+} to a solution of **5.029**, the sensor system shows significant blue shifts in its absorption and emission spectra due to desulfurisation-based cyclisation reaction. The sensor is efficient for the Hg^{2+} ion detection at pH 2 to 9. Its higher sensitivity is sufficient to detect the Hg^{2+} ion (less than 2ppb) in drinking water.[12i] Misra *et al*., have reported the thiourea-based fluorescent molecular dyad (ANTU) **5.030**, which bears anthracene and naphthalimide fluorophore. After addition of Hg^{2+} to the ANTU solution, it shows significant fluorescence intensity change which has detection limit of 4 ppb for Hg^{2+} in aqueous-ethanol medium.[12j]

Qian *et al*., have developed BODIPY-rhodamine (BODIPY=Boron Bipyrromethene) FRET-based ratiometric probe **5.031** (**Figure 5.09**) which can selectively detect lower amounts of Hg^{2+} ions on the ppb level under normal biological conditions. It displays a clear Hg^{2+} mediated change in the intensity ratio of the two well-separated and comparably strong emission bands of BODIPY and rhodamine. The substantial changes in the fluorescence color can be monitored by naked eye. [13a] Zheng *et al*., have reported similar FRET sensor **5.032**, based on intramolecular energy transfer (ICT) from fluorescein to a rhodamine B moiety. On addition of Hg^{2+} to a solution of **5.032**, the sensor system shows dual fluorescence peaks, positioned at 520 and 591 nm. There is a 65-fold increase in the ratiometric value (I_{591}/I_{520}), up to addition of 1.0 equivalent of Hg^{2+}. [13b] Another FRET-based sensor **5.033** (**Figure 5.09**), containing rhodamine B and a naphthalimide moiety, was made for the detecting of Hg^{2+} over a wide pH range by Lie *et al*. When Hg^{2+} is added, it helps desulfurisation of the thiocarbonyl part which leads to ring opening at rhodamine B part, forming imidazole derivative which helps in the FRET process. They also reported that the compound showed color change from colorless to red which can be noticed by the naked eye.[13c]

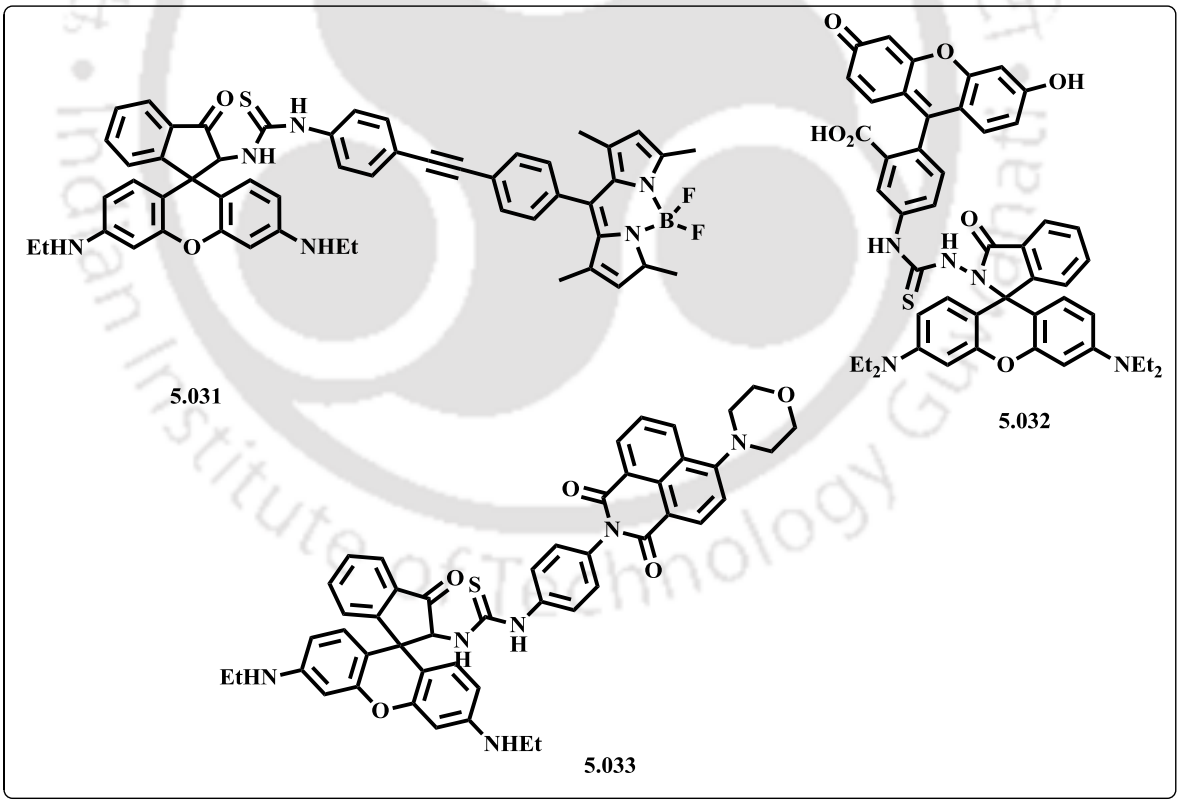


Figure 5.09. FRET based sensor for Hg^{2+} detection.

5.5. Background

The unnatural amino acids based sensor systems have been developed for many years. From **Chapter-1** discussion it is clear that fluorescent amino acids whose emission spectra and quantum yields are very sensitive to solvent polarity, are widely used as useful probes for investigating chemical, biochemical, and biological phenomena including metal ion sensing.[19e] Fluorescent chemosensors for the detection of environmentally as well as biologically hazardous metal ions and some physiological essential trace elements are highly desirable. [19] Transition and heavy-metal ions play vital roles in living systems but their high level concentration is harmful.[19d] Thus, the design and synthesis of environmental friendly chemosensors for the detection of essential trace elements and the ecological hazardous elements are of high importance. A fruitful design of a fluorescence chemosensor involves attachment of a metal binding unit with a fluorophore through a spacer. Binding of a target metal display noticeable change in the photophysical response of the fluorophore [14] further, the selectivity of a metal sensor is solely dependent on the nature of the coordinating ligands in the receptor unit and the interaction thereof. Therefore, many of the metal sensors containing N/O atoms [14] and few containing S-donor [15] in the metal binding site have been reported. Recently, the potentialities of mixed-donor-S containing ligands are considered for the design of metal sensors. [16] In particular, thiophilic nature of Hg^{2+} and Cu^{2+} attracted very much for the synthesis of sensors containing S as ligating atom. In this respect, in recent time, few thiourea based sensor have been reported as fruitful as sensors for Hg^{2+} and Cu^{2+} ions and other heavy metals.[17] However, newer design of more effective thiourea-based sensor would be interesting and is still in high demand.

Recently, our research group is involved in the development and application of pyrenyl thioureanyl amino acids as fluorescent chemosensor in metal ion detection. In the present chapter, we reported the interesting photophysical properties of amino pyrene linked thiourea used as fluorimetric chemosensor for biologically important transition metals i.e Hg^{2+}/Cu^{2+}.

5.6. Objective

With this above literature background, it is clear that many thiourea based metal sensor systems have been developed for Hg^{2+}/Cu^{2+} ion detection. However, there is no report of thiourea based unnatural amino acid as a metal chemosensory system, as smart fluorescent probe for the Hg^{2+}/ Cu^{2+} ion over other metal ions. Keeping the idea in mind and our continuing research efforts toward the design of unnatural fluorescent amino acids [18a-d] and sensors for biologically and environmentally important heavy and transition metal ions,[18e-f] we thought that it would be worthwhile to explore fluorescent pyrenylthioureayl alanine amino acid (**3.063**,**Chapter-3**, ^{PyTU}**Ala**) as a possible sensor for heavy and transition metal ions (**Figure 5.10**). We envisaged that thiourea-moiety would allow strong coordination with Hg^{2+} and/or Cu^{2+}due to their thiophilic nature and hence would generate a detectable fluorescence signal from pyrenyl unit.

Therefore, in this particular, we want to report the pyrenylthioureayl alanine amino acid (**3.063**, ^{PyTU}**Ala**) as a potential sensor for the detection of heavy and transition metal–ions. Thus, we investigated the binding interaction of our newly synthesized thioureayl amino acid for various heavy and transition metal–ions by spectroscopic methods. Among the various metal ions, the probe showed a selective affinity for Hg^{2+} ion in the partial aqueous medium. However, the sensor also showed affinity for Cu^{2+}ion in the absence of interfering metal ions. In both the cases the sensing events were indicated by an enhancement of fluorescence and a 1:2 metal-ligand complexation possibly via the coordination with S atom of thiourea unit.

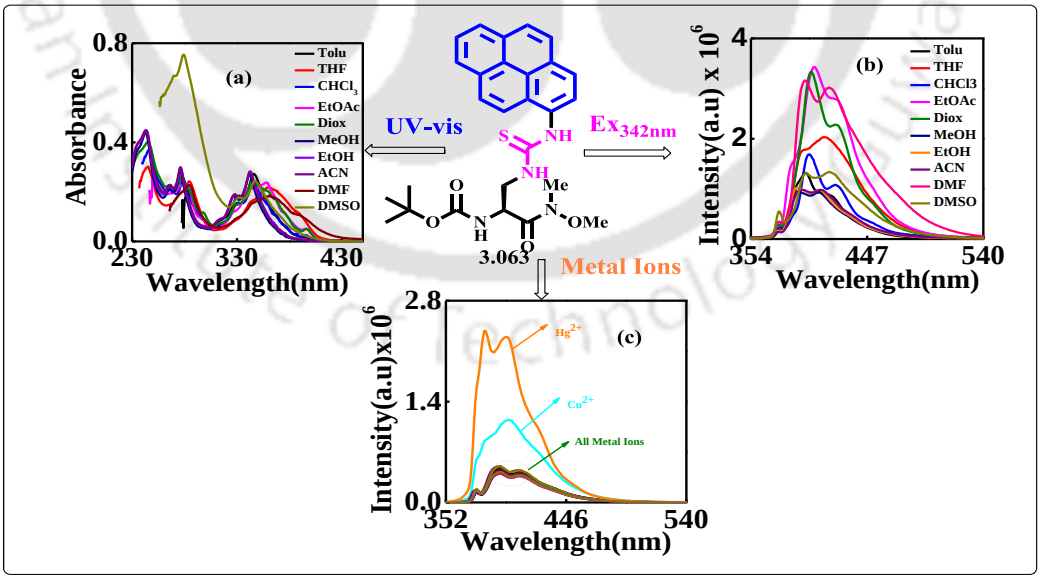


Figure 5.10. (a) UV-visible spectra of probe **3.063** (10 μM) (b) Fluorescence emission spectra of the probe **3.063** (10 μM) in different solvents(c) Fluorescence emission spectra of the probe in presence of 1 equivalent (10 uM) concentration of various metal ions .

5.7. Results and Discussion

The synthesis of the probe pyrenylthioureayl alanine (**3.063**, **^{PyTU}Ala**) was achieved by following our reported protocol discussed in **Chapter 3** with good yield (75 %).

5.7.1. ^{1}H NMR Characterisation of Pyrenylthioureayl Alanine

The 9 hydrogens of pyrene unit appeared at δ 8.21-8.13 ppm as a multiplet, δ 8.10 as a doublet with coupling constant (*J)* = 8.9 Hz, δ 8.01-8.05 as a multiplet and δ 7.94 as a doublet with *J* = 8.0 Hz. NHγ appeared as a broad singlet at δ 8.49, -NH_{β} appeared at δ 6.54 as a singlet and -NH_{α} appeared at δ 5.43 as a doublet with *J* = 7.1 Hz. The α–CH hydrogen of alanyl unit found as a doublet at δ 4.65 ppm with *J* = 4.0Hz and β–CH_2 hydrogens of alanyl unit appeared as a multiplet at δ 3.96-3.79. Methyl hydrogens of N-Me and OMe groups appeared as a singlet at δ 2.98 and 3.56 respectively. The ^{t}Bu hydrogens of Boc- found as a singlet at δ 1.01.

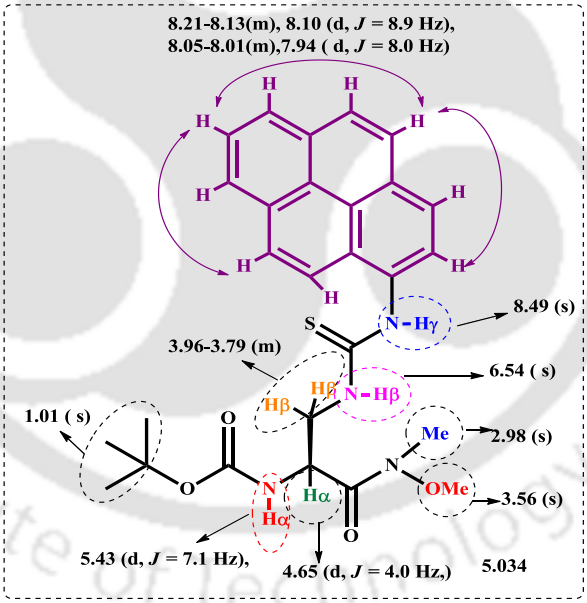


Figure 5.11. ^{1}H-NMR characterisation of pyrenylthioureayl alanine (**^{PyTU}Ala**).

5.7.2. Photophysical Study

After synthesising and characterising the probe **3.063** (**^{PyTU}Ala)**, we turned our attention to study the photophysical properties of the probe in presence of various metal cations in aqueous medium.

5.7.2.1. Photophysical Properties

We have previously reported (in **Chapter 3**) the fluorescence behaviour of the probe **^{PyTU}Ala** in different solvents of varying polarity. In brief, it showed slightly distorted shapes of absorption bands compared to pyrene in all solvents (**Figure 5.12a)** whereas, in fluorescence an irregular trend indicating a modulated pyrene emission, had appeared at short wavelength range 390-398 nm and 410-425 nm in long wavelength range (**Figure 5.12b)**.[20] Following the absorption and emission response in ACN/water solvent mixture (**Figure 5.12c-d)** and the solubility of the sensor, we chosed 7:3 water: ACN solvent mixture for studying the interaction with metal ions. Hence, we used this solvent system throughout the titration of the metal ions.

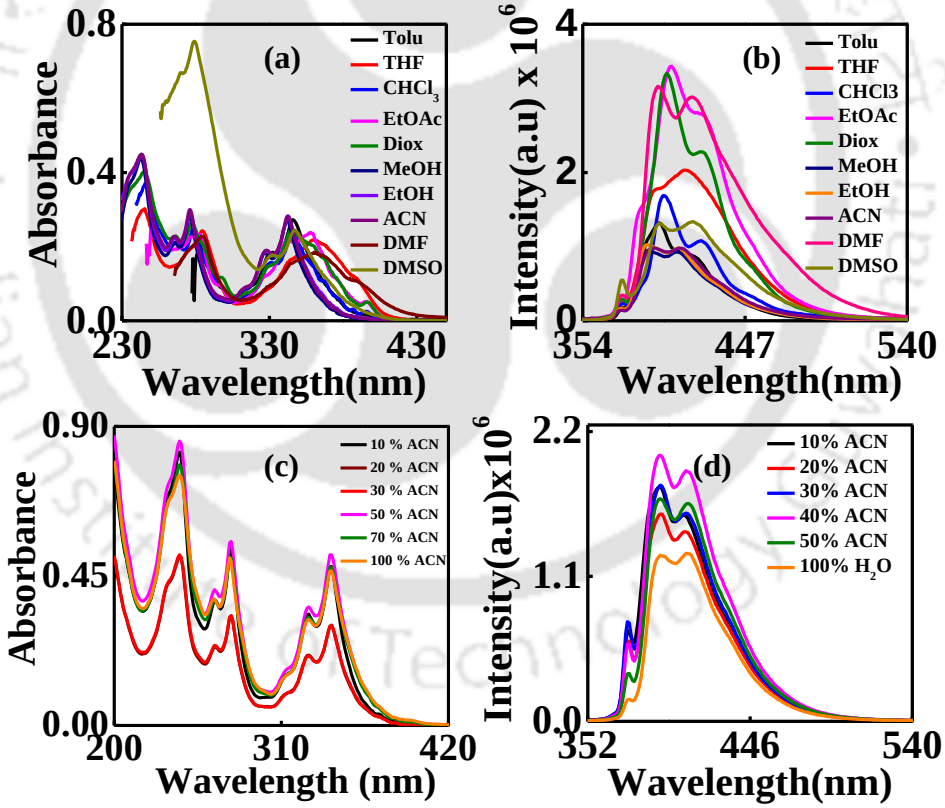


Figure 5.12. (a) UV-visible and (b) Emissions behaviour of the **^{PyTU}Ala** in different solvents (c) UV-visible and (d) Emissions behaviour of the **^{PyTU}Ala** at different percentage of water/acetonitrile (V/v).

5.7.2.2. UV-Visible Study

We next explored the sensing capability of **^{PyTU}Ala** towards Hg^{2+} & Cu^{2+} ion in presence of other metal ions like, Ag^{+} (**Figure 5.16a**), Cu^{+}(**Figure 5.17a**), Cd^{2+}(**Figure 5.18a**), Co^{2+}(**Figure 5.19a**), Fe^{3+}(**Figure 5.20a**), Mg^{2+}(**Figure 5.21a**), Mn^{2+}(**Figure 5.22a**), Ni^{2+}(**Figure 5.23a**), Pb^{2+} (**Figure 5.24a**), Zn^{2+}(**Figure 5.25a**). To do so, we added 1mM concentration of each metal ion into the probe's solution at fixed concentration and then photophysical properties were studied. Thus, upon addition of 1 equivalent of various metal ions (10 μM), as their chloride salts, to the solution of **3.063** the absorption spectra were found to be considerably modulated (**Figure 5.13a**) only in the presence of Hg^{2+} ion wherein we observed slight blue shift of pyrenyl absorption bands (2-3 nm) along with an appearance of an extra strong band at 359 nm (**Figure 5.13b**). The other metal ions did not show any significant change (**Figure 5.13a**) in the absorption spectra of **3.063** except for Cu^{2+} ion. The presence of Cu^{2+} ion induced appearance of new weak absorbance peak at 374 nm (**Figure 5.13c**). Thus, the UV-visible absorption supported the binding event of the probe with Hg^{2+} and Cu^{2+} ions.

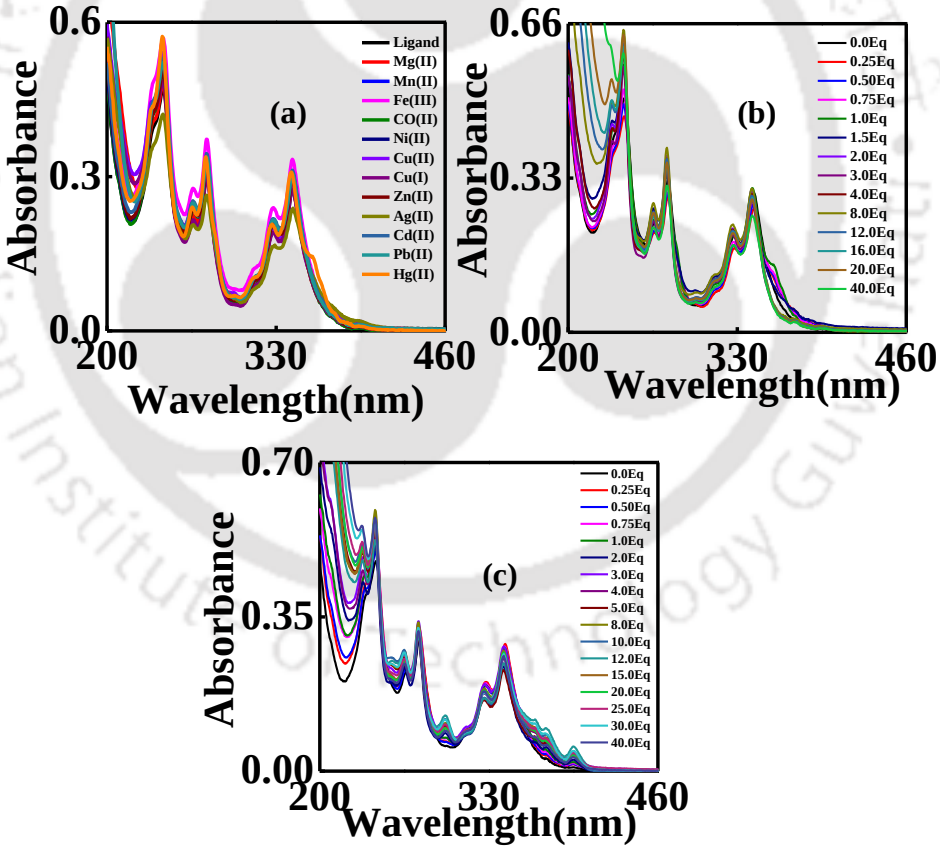


Figure 5.13. (a) UV-vis spectra of probe (**^{PyTU}Ala**) (10 μM) in presence of various metal ions (1.0 equiv) (b) UV-vis spectra of probe (**^{PyTU}Ala**) (10 μM) in presence of increasing Hg^{2+} concentration (c) UV-vis spectra of probe (**^{PyTU}Ala**) (10 μM) in presence of increasing Cu^{2+} concentration.

5.7.2.3. Fluorescence Study

The fluorescence study of the probe (**^{PyTU}Ala)** was carried out in presence and absence of various metal ions at different concentrations. We observed an abrupt change in the emission intensity of the probe in presence of Hg^{2+} ion compared to Cu^{2+} ion whereas, the other metal ions like, Ag^{+} (**Figure 5.16b**), Cu^{+}(**Figure 5.17b**), Cd^{2+}(**Figure 5.18a**), Co^{2+}(**Figure 5.19b**), Fe^{3+}(**Figure 5.20b**), Mg^{2+}(**Figure 5.21b**), Mn^{2+}(**Figure 5.22b**), Ni^{2+}(**Figure 5.23b**), Pb^{2+} (**Figure 5.24b**), Zn^{2+}(**Figure 5.25b**), and showed irregular and negligible change of the emission spectra at 407nm. Upon addition of Hg^{2+}, fluorescence intensity of the band drastically increased (**Figure 5.14b**) compared to Cu^{2+} ion (**Figure 5.14c**). The saturation for Hg^{2+} reached at 1 equivalent, while that of Cu^{2+} required 10 equivalents. These facts implied that both the Hg^{2+} and Cu^{2+} ions interacted significantly with the probe with a higher binding of Hg^{2+} ion. The fluorescence enhancement upon addition of Hg^{2+} was also supported from a time resolve fluorescence experiment showing enhancement of probe's life time from 3.1 ns to 3.6 ns (0.5 equivalent Hg^{2+}) to 4.4 ns (1.0 equivalent Hg^{2+} (**Figure 5.14d**). A two times enhancement of quantum yield was observed for the detection of Hg^{2+} (1equiv.) while for Cu^{2+} it was 1.5 times (**Table 5.1**). All these phenomena indicated a well-defined binding of Hg^{2+}/Cu^{2+} ion to the probe.

.

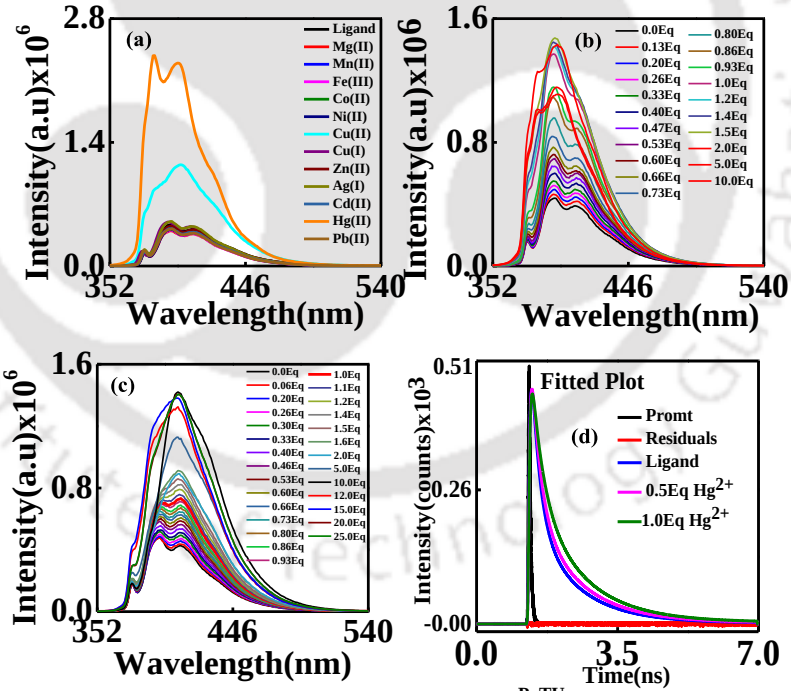


Figure 5.14. (a) Fluorescence spectra (λ_{ex} = 342 nm) of probe (**^{PyTU}Ala**) (10 μM) in presence of various metal ions (1.0 equiv.) (b) Fluorescence spectra (λ_{ex} = 342 nm) of probe (**^{PyTU}Ala**) (10 μM) at increasing Hg^{2+} concentration. (c) Fluorescence spectra (λ_{ex} = 342 nm) of probe (**^{PyTU}Ala**) (10 μM) at increasing Cu^{2+} concentration (d) Fluorescence life time of the probe in presence and absence of Hg^{2+} ion.

Table 5.1. Photophysical summary of ^{PyTU}**Ala** with [Hg^{2+}] and [Cu^{2+}] ion in Water/CH_3CN (7:3, V/v)

Hg^{2+} (Equiv.)	λ^{abs}_{max} (nm)	Abs	λ^{fl}_{mono} / λ^{fl}_{exci} (nm)	% Φ
0.00	342	0.115	394	4.61
0.25	342	0.083	394	5.60
0.50	342	0.084	393	6.83
0.75	342	0.109	394	7.23
1.0	342	0.124	393	**9.01**
Cu^{2+} (Equiv.)	λ^{abs}_{max} (nm)	Abs	λ^{fl}_{mono} / λ^{fl}_{exci} (nm)	% Φ
0.00	342	0.114	394	4.21
0.25	342	0.113	395	4.24
0.50	342	0.112	394	4.38
0.75	342	0.110	408	5.32
1.0	342	0.108	406	**5.96**

The fluorescence change of the probe (^{PyTU}**Ala)** in presence and/or absence of other metal ions were also investigated to test the sensitivity of the probe for Hg^{2+} ion. Then, we plotted the emission intensity of the probe (10 μM, 1 eqiv.) *Vs* various metal ions at fixed concentration (10 μM, 1 eqiv.) (**Figure 5.15**) and compared the change in fluorescence intensity {$(A - A_0)/A_0$ x 100%} of the probe upon addition of various metal ions. Thus, from the bar diagram it is clear that about 5.2 times enhancement of fluorescence intensity for Hg^{2+} ion and 2.6 times for Cu^{2+} ion was observed with respect to probe.

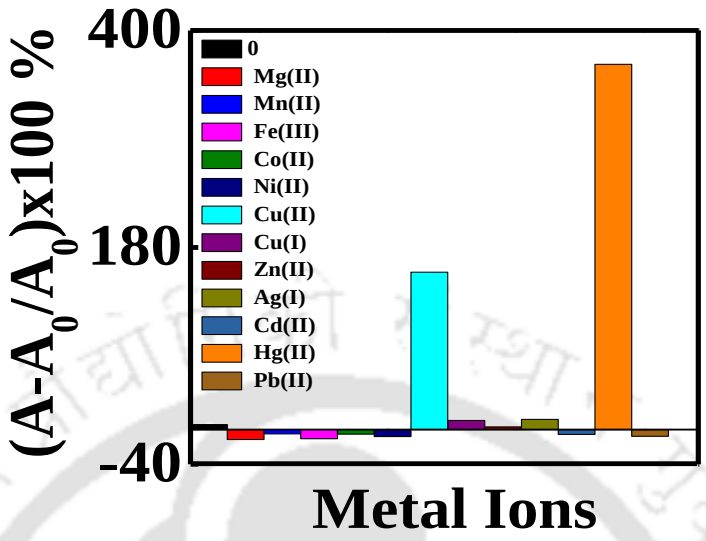


Figure 5.15. Fluorescence (λ_{ex} = 342 nm) intensity area changes ((A – A_0)/A_0 x 100%) of probe (**^{PyTU}Ala**) (10 μM) at 298 K upon addition of various metal ions (1.0 equiv.) [A_0 is fluorescence emission intensity area of free probe and A is the fluorescence intensity area of probe after adding metal ions].

5.7.2.4. UV-Visible and Fluorescence Spectra of ^{PyTU}Ala in Presence of other Metal ions

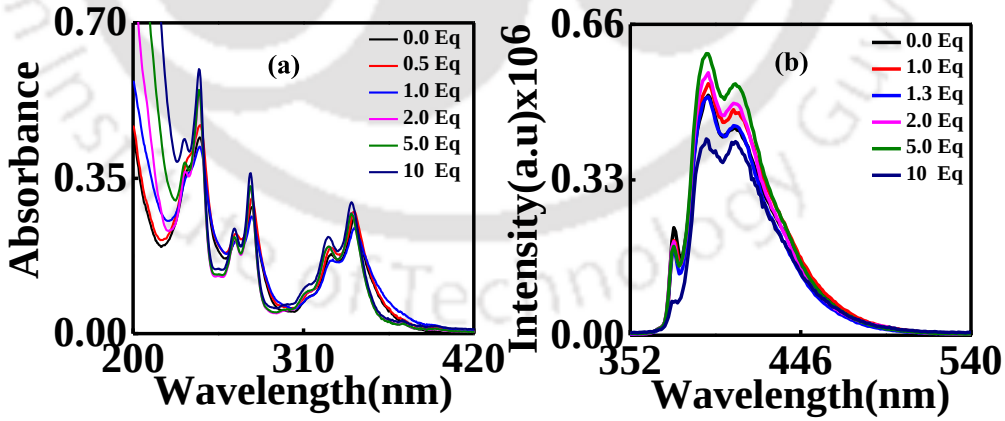


Figure 5.16. (a) UV-visible & (b) emission spectra (λ_{ex} = 342 nm) of **^{PyTU}Ala** (10μM) in presence of Ag^{1+} ion.

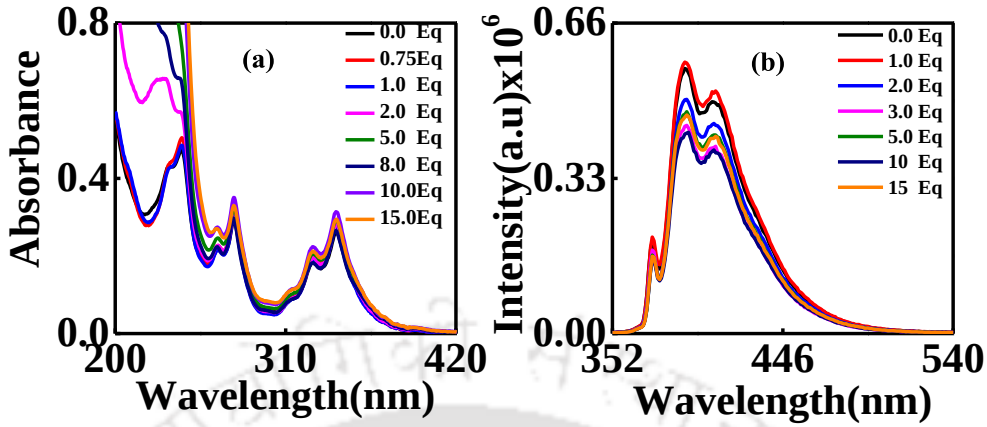


Figure 5.17. (a) UV-visible & (b) emission spectra (λ_{ex} = 342 nm) of ^{PyTU}**Ala** (10µM) in presence of Cu^{1+} ion.

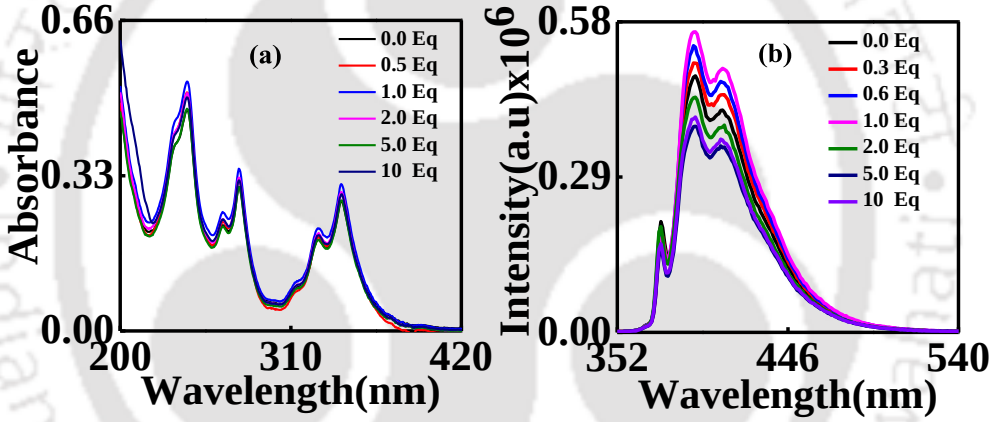


Figure 5.18. (a) UV-visible & (b) emission spectra (λ_{ex} = 342 nm) of ^{PyTU}**Ala** (10µM) in presence of Cd^{2+} ion.

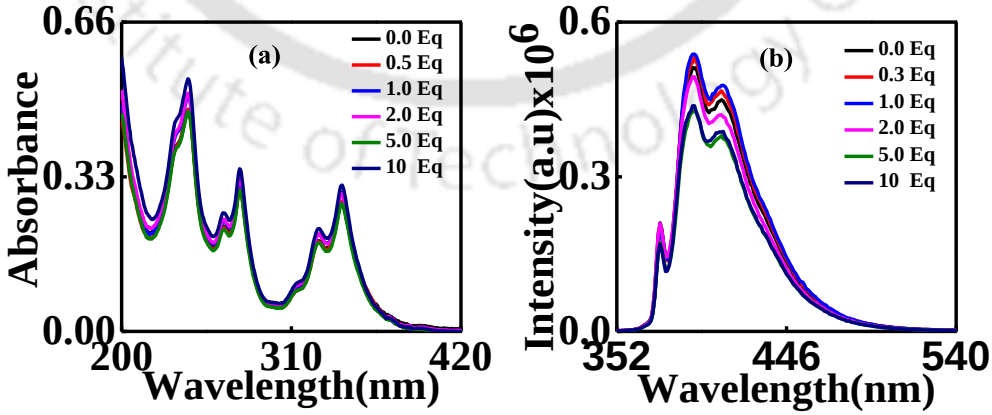


Figure 5.19. (a) UV-visible & (b) emission spectra (λ_{ex} = 342 nm) of ^{PyTU}**Ala** (10µM) in presence of Co^{2+} ion.

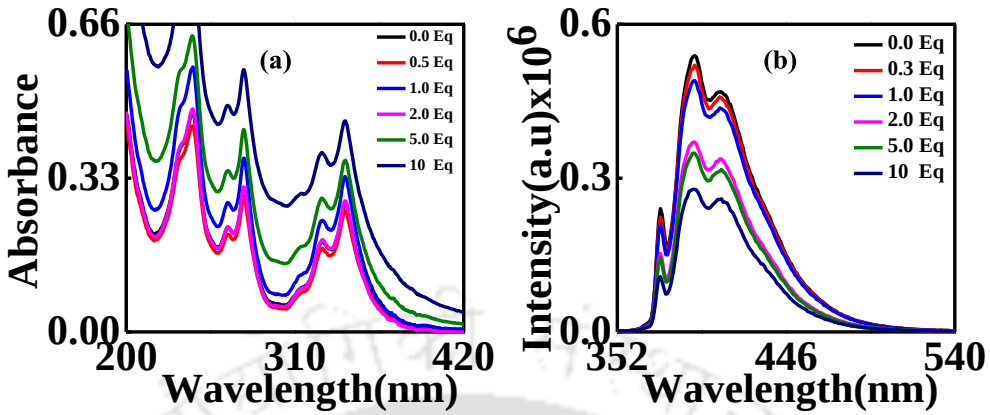


Figure 5.20. (a) UV-visible & (b) emission spectra (λ_{ex} = 342 nm) of **^{PyTU}Ala** (10µM) in presence of Fe^{3+} ion.

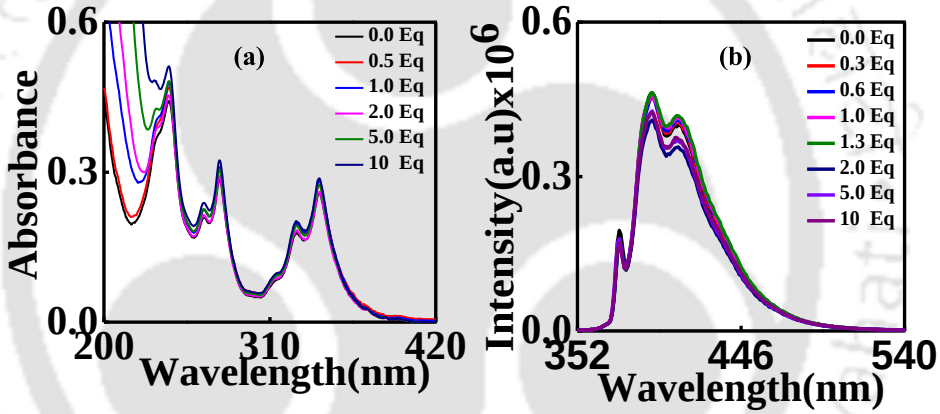


Figure 5.21. (a) UV-visible & (b) emission spectra (λ_{ex} = 342 nm) of **^{PyTU}Ala** (10µM) in presence of Mg^{2+} ion.

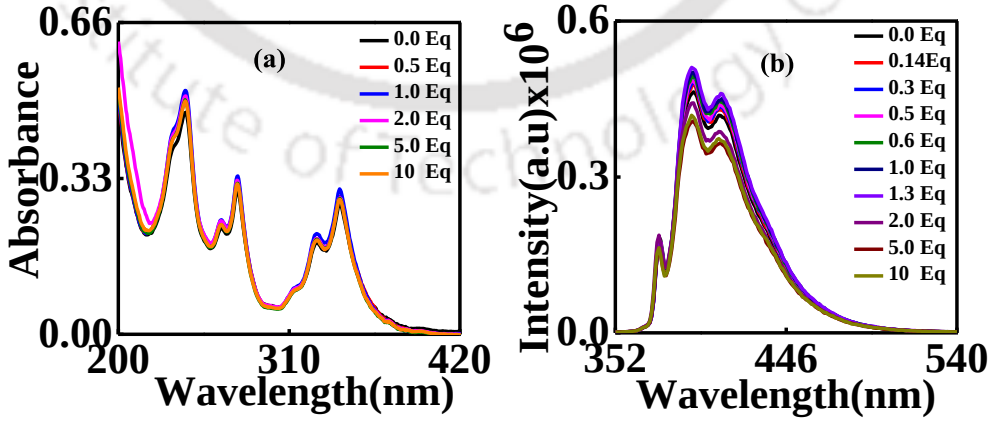


Figure 5.22. (a) UV-visible & (b) emission spectra (λ_{ex} = 342 nm) of **^{PyTU}Ala** (10µM) in presence of Mn^{2+} ion.

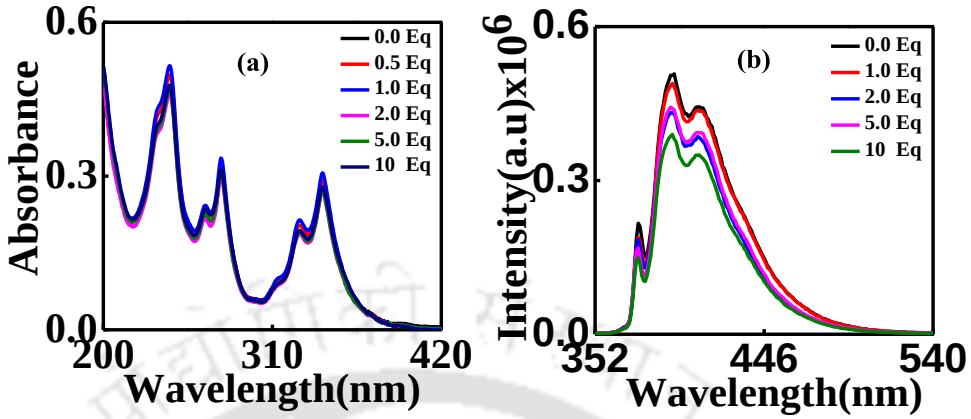


Figure 5.23. (a) UV-visible & (b) emission spectra (λ_{ex} = 342 nm) of **^{PyTU}Ala** (10µM) in presence of Ni^{2+} ion.

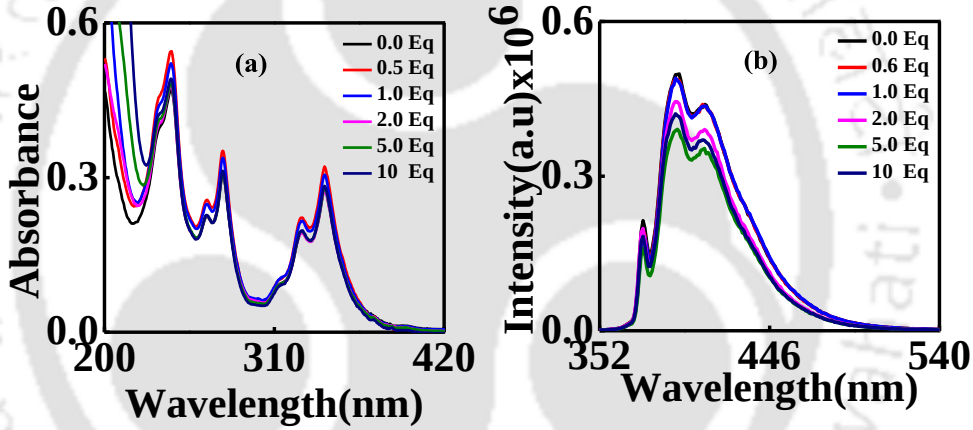


Figure 5.24. (a) UV-visible & (b) emission spectra (λ_{ex} = 342 nm) of **^{PyTU}Ala** (10µM) in presence of Pb^{2+} ion.

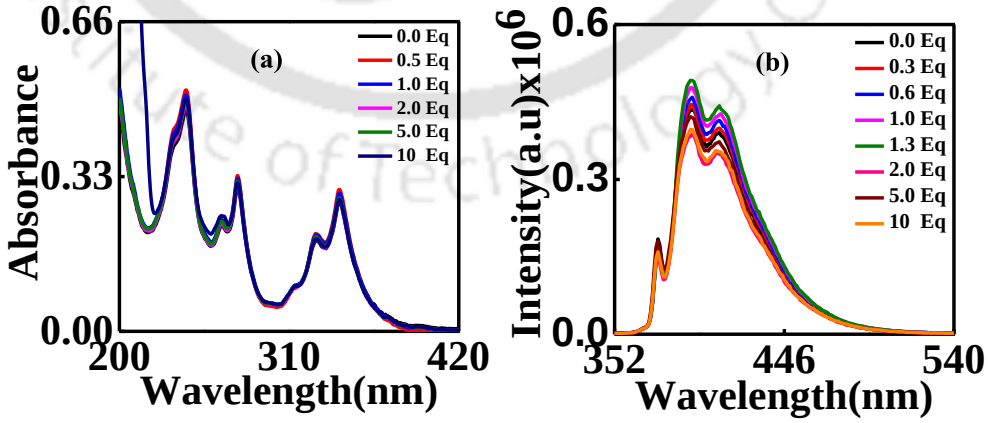


Figure 5.25. (a) UV-vis, (b) emission spectra (λ_{ex} = 342 nm) of **^{PyTU}Ala** (10µM) in presence of Zn^{2+} ion.

5.7.2.5. Binding Stoichiometry from Fluorescence Job's Plot for Hg^{2+} Ion

To quantify the stoichiometry between probe (^{PyTU}**Ala**) and the Hg^{2+} ion, the fluorescence Job's plot measurement was carried out. For this purpose, an equal concentration of probe (^{PyTU}**Ala**) and Hg^{2+} ion solutions were prepared separately in 70 % water/acetonitrile (V/v) medium. The equal concentration of the probe and the Hg^{2+} ion solutions were mixed in variable fractions of volume to maintain different stoichiometry at fixed total concentration. Then fluorescence emission spectra for different compositions of the probe were recorded (**Figure 5.26a**) and the maximum fluorescence intensity was found at 6:4 (probe: Hg^{2+}). To calculate the probe-Hg^{2+} complexation ratio, a graph between [probe-Hg^{2+}] vs. X_{probe} was plotted, where [probe-Hg^{2+}] = $\Delta A/A_0$ x [probe], {$\Delta A = (A–A_0)$}; where A and A_0 are the fluorescence intensity area of the probe before and after addition of Hg^{2+} ion to the probe respectively and X_{probe} is the mole fraction of probe, {X_{probe} = [probe(L)]/([probe(L)] + [Hg^{2+}(M)])}. The corresponding X_{probe} value where the intensity of probe-Hg^{2+} was maximum was taken as probe-to-metal stoichiometry. The Job's plot (**Figure 5.16b**) showed the point of maxima at the mole fraction of 0.64 of the probe, which clearly indicates that, the stoichiometry of the probe to Hg^{2+} ion in the complex is 2:1.

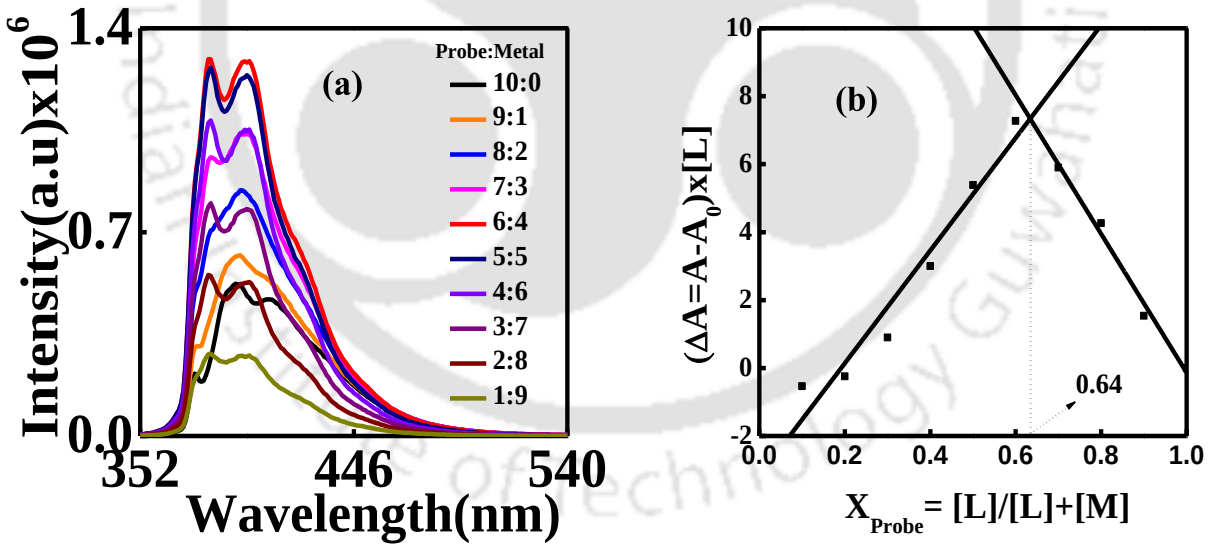


Figure 5.26. (a) Fluorescence job spectra (b) Fluorescence Job's plot of the probe, ^{PyTU}**Ala** in presence of Hg^{2+} ion in H_2O/CH_3CN.([L]= probe Conc.^{n} and [M]= Metal Conc.^{n})

5.7.2.6. Determination of Complexation Constant for Hg^{2+} Ion

After determining and confirming the 2:1 probe-Hg^{2+} complexation we, next, turned our attention to evaluate the binding constant by fluorescence spectroscopy. Thus, to determine binding constant we have constructed the well-known Benesi-Hildebrand plot using the following equation. In order to maintain Benesi-Hildebrand condition, we took probe concentration much greater than Hg^{2+} ion,

$$\frac{1}{(A-A_0)} = \frac{1}{(A-A_0)} + \frac{1}{(A_\infty-A_0)\,K\,[Probe]} \quad \text{---------- (1)}$$

where A_0, A and A_∞ are the emission intensity area of **^{PyTU}Ala** in the absence, at intermediate and at infinite concentration (where maximum intensity area was observed) of Hg^{2+} ion respectively. To determine the binding constant, $1/(A-A_0)$ vs. $1/[Probe]^2$ and $1/(A-A_0)$ vs. 1/[Probe] was plotted (**Figure 5.27a-b**). From the slope and the intercept of the linear fit plot of $1/(A-A_0)$ vs $1/[Probe]^2$ or 1/[Probe], the complexation constant (K_{com}) can be obtained as K_{com} = Intercept/Slope. The complexation constant was calculated, which was found to be $3.93x10^{10}$ (M^{-1}) and $9.15x10^4$ (M^{-1}) indicating a strong complexation between Hg^{2+} with probe **^{PyTU}Ala**.

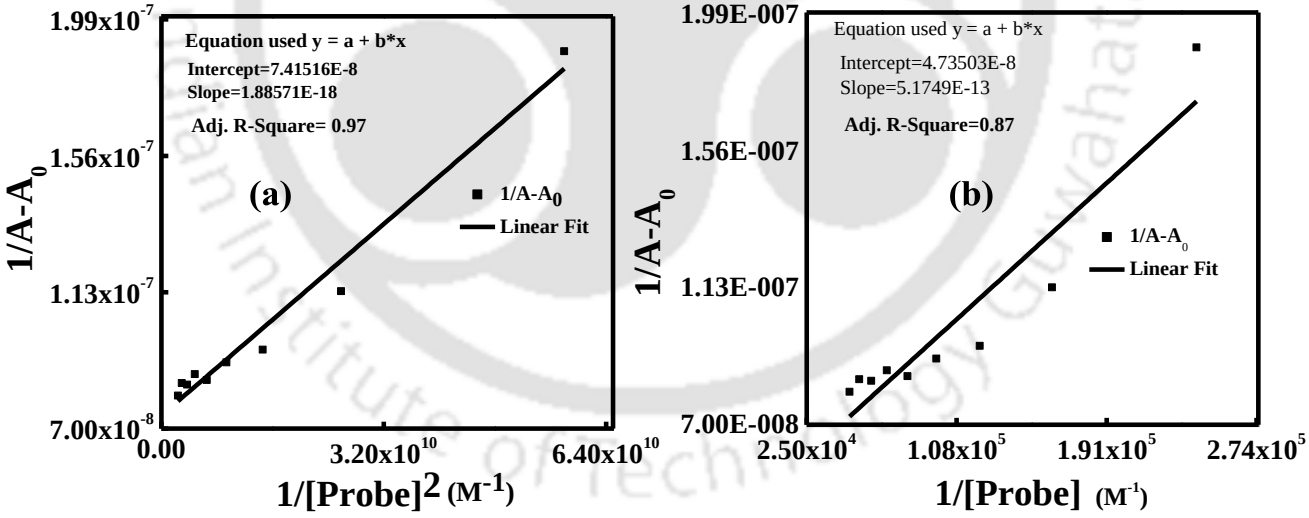


Figure 5.27. Benesi-Hildebrand plot (a) 1:2 and (b)1:1 of **^{PyTU}Ala** with Hg^{2+}.

5.7.2.7. Test for Selectivity in Sensing of Hg^{2+} and Cu^{2+} Ion in Presence of Other Interfering Metal Ions

Competition experiments were carried out to explore the utility of the probe as an ion-selective switch-on fluorescence chemosensor for Hg^{2+} and Cu^{2+} ions. Thus, to do so, the fluorescence spectra (**Figure 5.28a-b**) of the probe (10 µM, 1 equiv.) were individually taken in presence of 1 equivalent Hg^{2+} ion (10 µM, 1equiv.) along with excess concentration of other background metal ions (20 equivalent each). Similarly, the fluorescence spectra (**Figure 5.28c-d**) of the probe (10 µM) were taken individually in presence of 10 equivalent Cu^{2+} ion (100µM) along with other background metal ions in 50 equivalent concentration each. The fluorescence property of the probe with Hg^{2+} ion in presence of other background metal ions is negligibly changed. From the bar diagram it was clear that no metal ions had interference (**Figure 5.28**) in the detection of Hg^{2+}. However, in case of Cu^{2+} ion the probe fluorescence is quenched (**Figure 5.28c-d**) in presence of 50 equivalents Fe^{3+} ion which reflects that Fe^{3+} showed interference in the detection of Cu^{2+} ion (**Figure 5.28d**). These results suggested that probe **3.063** showed an excellent selectivity for detecting the Hg^{2+}.[18e-f] Thus, these results indicate that the probe could be used as a potential Hg^{2+} ion selective turn-on fluorescence sensor.

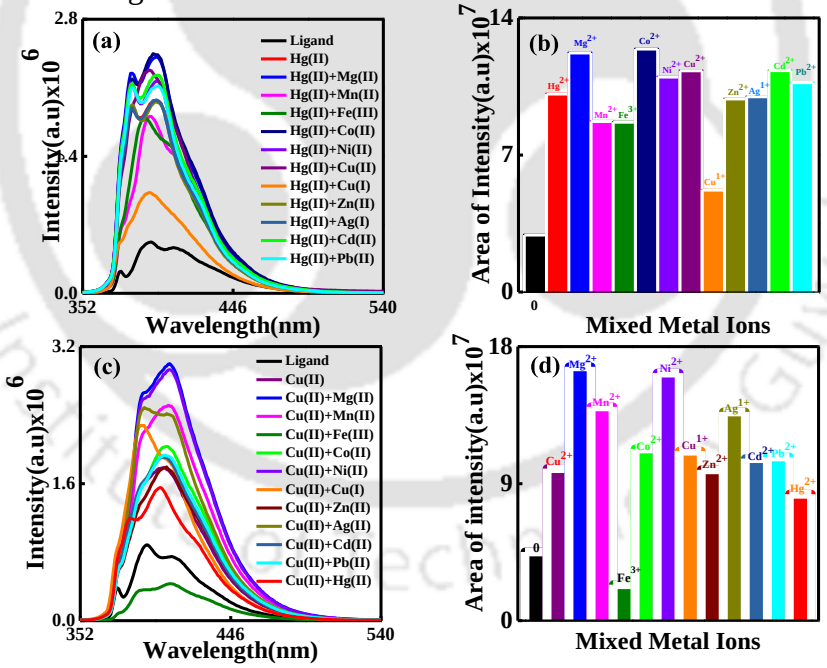


Figure 5.28. (a) Emission spectra (λ_{ex} = 342 nm) (b) fluorescence intensity of ^{PyTU}**Ala** (10µM) with Hg^{2+} ion (1 equiv.) in absence and presence of other metal ions (20 equiv.) (c) Emission spectra (λ_{ex} = 342 nm) (b) fluorescence intensity of ^{PyTU}**Ala** (10µM) with Cu^{2+} ion (10 equiv.) in absence and presence of other metal ions (50 equiv.)

5.7.2.8. Determination of the Detection Limit

We have calculated the detection limit of the probe ^{PyTU}**Ala** towards Hg^{2+} ion and Cu^{2+} ion by plotting fluorescence intensity area of probe in presence of metal ion vs. concentration of Hg^{2+} ion (**Figure 5.29a**) and Cu^{2+} ion (**Figure 5.29b**). The linear fit was drawn by taking the 11 points (0.66 µM to 7.27 µM) of the linear region and the slope was determined. The limit of detection (LOD) has been calculated using the equation 3σ/K; where σ denotes the standard deviation of fluorescence intensity area of the probe before adding metal ion and K represents the slope of the plot (**Table 5.2**).[18e-f,19] A good linear relationship indicated the probe's efficiency for the detection of the submilimole concentration of Hg^{2+} ions found in many chemical and biological systems.[18e-g] The detection limits were found to be 93 nM for Hg^{2+} and 318 nM for Cu^{2+} ion. Thus, our probe can be utilised for the detection of the submillimolar concentration range of Hg^{2+} compared to Cu^{2+} ion. [18e-f,19]

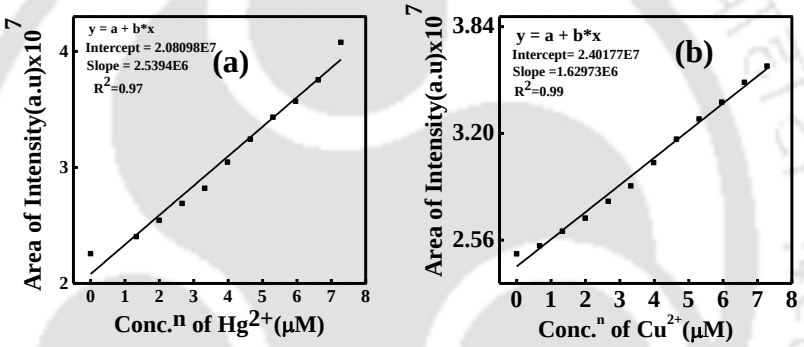


Figure 5.29. Plot of area of intensity of probe vs (a) conc.^{n} of Hg^{2+} ion (b) conc.^{n} Cu^{2+} ion.

Table 5.2. Limit of detection calculation

Repetition	Area of Fluorescence Intensity of ^{PyTU}**Ala**	
	Probe for Hg^{2+}	Probe for Cu^{2+}
1	23187454	24565622
2	23136312	24366915
3	23026781	24240279
4	23022650	24193494
5	23009187	24137542
6	23007423	24136398
7	23036752	24061474
8	22987616	24055790
9	23180717	24032943
Average(Mean)	23066099.122	24198939.493
Standard Deviation (σ)	79001.901	172856.635
Slope	2.54E+06	1.63E+06
LOD(3*σ/k)	0.0933uM =93.3nM (in 3ml)	0.318uM (In 3ml) =318nM (in 3ml)

5.7.3. Binding Stoichiometry for Cu^{2+} Ion from NMR

As the UV and fluorescence sensitivity of the probe was insignificant to quantify the stoichiometry between probe (^{PyTU}**Ala**) and the Cu^{2+} ion, we used NMR titration for the further investigation. For this purpose, we titrated known concentration of probe (1 eqiv.) with variable concentration of Cu^{2+} ion in d_6-DMSO. Then chemical shift of β-NH of the ^{PyTU}**Ala** at different equivalent Cu^{2+} ion was recorded (**Figure 5.30a**). Thus, $\Delta\delta$ vs. X_{Cu}^{2+} was plotted for various equivalent of Cu^{2+} ions, where $\Delta\delta = \delta - \delta_0$ [δ_0 = Chemical shift of β-NH and δ = Chemical shift of β–NH of the probe in absence and presence of metal respectively] and X_{Cu}^{2+} is the equivalent Cu^{2+} ion. The X_{Cu}^{2+} equivalent corresponding to constant $\Delta\delta$ value was taken as probe-to-metal stoichiometry, which was found to be 0.50 (**Figure 5.30b)**. This observation clearly indicated a 2:1 stoichiometry of the probe to Cu^{2+} ion complexation.

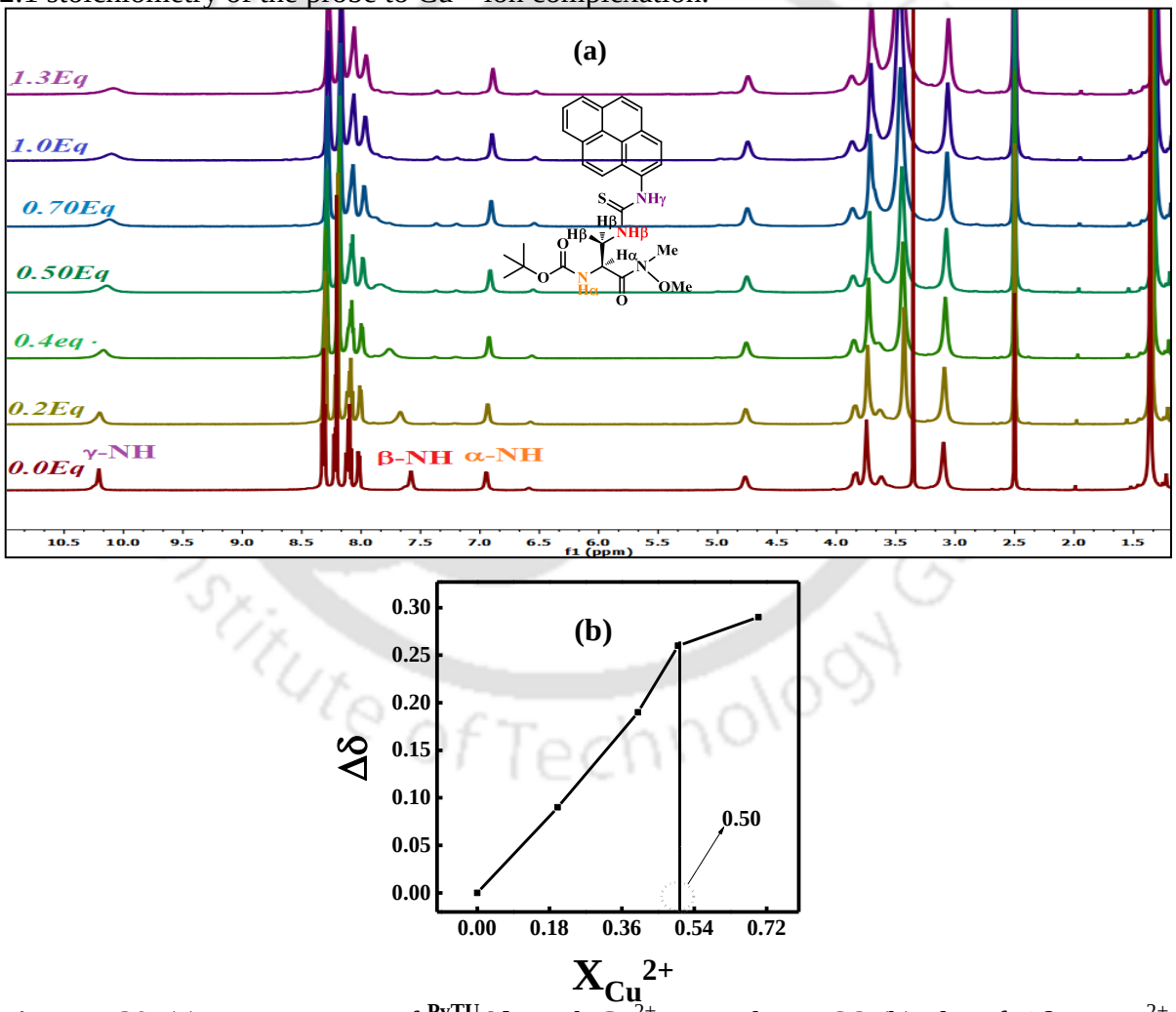


Figure 5.30. (a) NMR titration of ^{PyTU}**Ala** with Cu^{2+} ion in d_6-DMSO (b) Plot of $\Delta\delta$ Vs X_{Cu}^{2+} for binding stoichiometry.

5.7.4. Determination of complexation Constant for Cu^{2+} Ion from NMR

After determining and confirming the 2:1 probe-Cu^{2+} complexation, we next turn our attention to evaluate the binding constant by NMR spectroscopy. Thus, to determine binding constant, we have constructed Benesi-Hildebrand plot for NMR titration using the following equation. In order to maintain Besnesi-Hildbrand condition we took probe concentration much greater than Cu^{2+} ion concentration.

$$\frac{1}{(\delta-\delta_0)} = \frac{1}{(\delta-\delta_0)} + \frac{1}{(\delta_\infty-\delta_0)\,K\,[Cu^{2+}]} \quad \text{----------} \quad (2)$$

Where δ_0= Chemical shift (β-NH) of ligand in absence of metal; δ = Chemical shift (β-NH) of ligand in presence of Cu^{2+} at particular concentration and δ_∞ = Chemical shift (β-NH) of ligand when all of the Cu^{2+} are in the complex. To determine the complexation constant, $1/(\delta-\delta_0)$ vs. $1/[Cu^{2+}]$ was plotted (**Figure 5.31**). From the slope and the intercept of the linear fit plot of $1/(\delta-\delta_0)$ vs $1/[Cu^{2+}]$; complexation constant (K_{com}) can be obtained as, K_{com}= Intercept/Slope. The complexation constant was calculated, which was found to be 4.11×10^{-6} (M^{-1}) indicating a weak complexation between Cu^{2+} with probe **^{PyTU}Ala**.

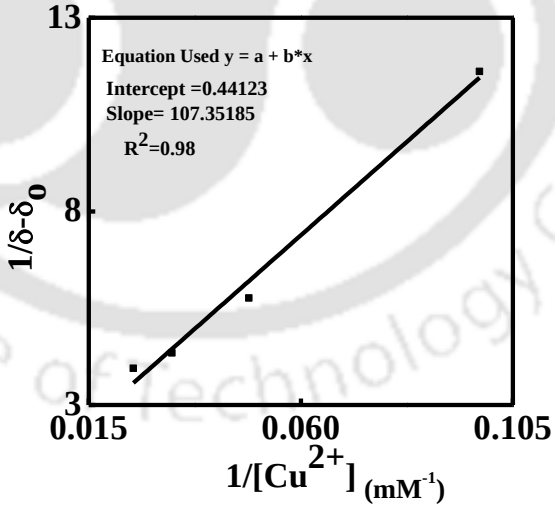


Figure 5.31. Benesi-Hildebrand plot of **^{PyTU}Ala** with Cu^{2+} from NMR chemical shift.

5.7.5. Support of 2:1 Complexation by Mass Spectral Analysis

To support the 2:1 probe-Hg^{2+} ion complexation, we have carried out a mass spectral analysis of the solution of the probe containing Hg^{2+} ion. Thus, the +ESI-Mass spectrum analysis revealed the formation of 2:1 probe-to-metal complex (**Figure 5.32**) with a clear isotopic distribution. Thus, we observed a peak at m/z = 1213.3582 which has matched with the calculated mass for $[(^{\mathbf{PyTU}}\mathbf{Ala})_2+Hg+H]^+$ (1213.3604; $C_{54}H_{59}N_8O_8S_2Hg$) for Hg^{2+} which corresponds to 2:1 imine complex. These results again support that one Hg^{2+} is coordinated by two *β*–NH of thoiurea of two probes ($^{\mathbf{PyTU}}\mathbf{Ala}$).

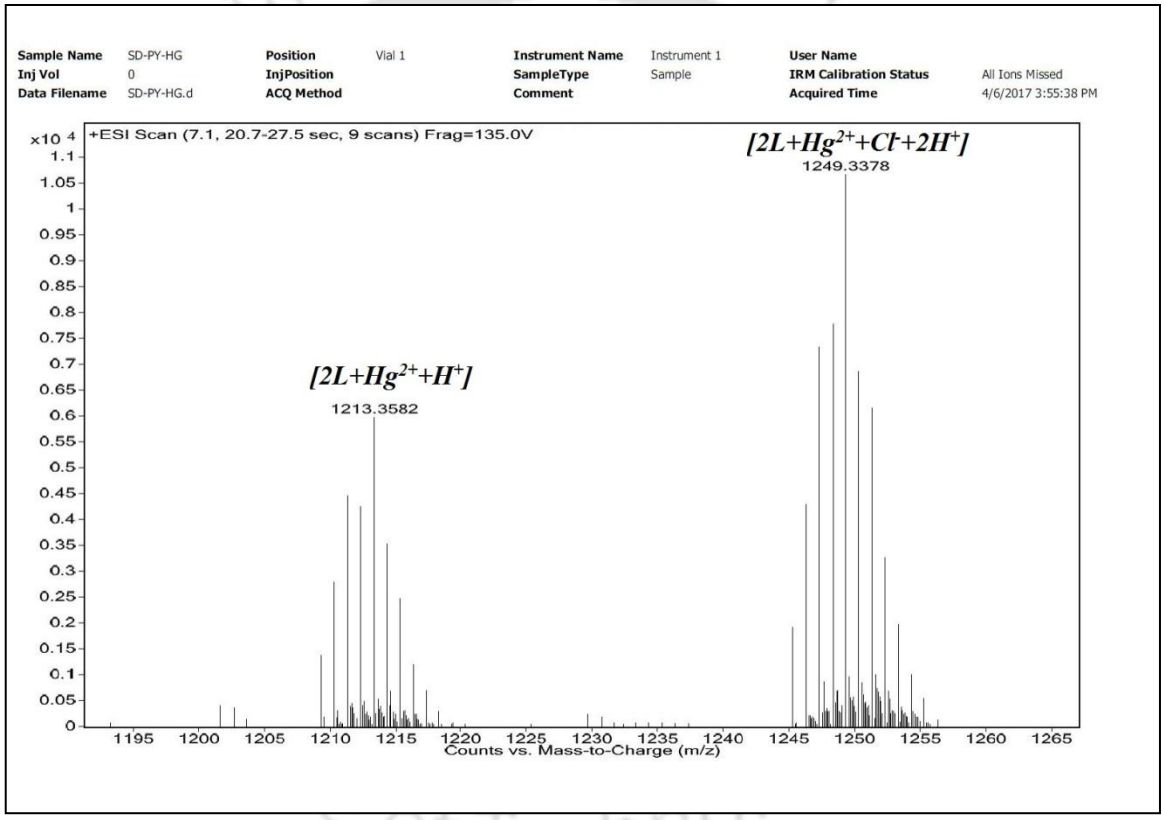


Figure 5.32. +ESI-Mass spectrum of the Probe in presence of Hg^{2+} ion

5.7.6. Explanation for Fluorescence Enhancement of ^{PyTU}Ala

The switch-on fluorescence response of the probe upon binding with the metal through -S-(Hg^{2+} or Cu^{2+})-S-linkage can be explained by considering the role of lone pairs on sulfur and nitrogen atoms in the thiourea moieties. The N-lone pair of directly linked γ-NH is involved in PET with the fluorophoric pyrene leading to quenching of fluorescence of parent pyrenyl in absence of metal ions (**II, Figure 5.33**).[1,21a] Further the lone of sulpher is also involved in PET through space with the pyrene and thus drastically quenching the fluorescence. However, upon complexation with the metal through S *via* the formation of imino thiourea (**III-IV, Figure 5.33**) the PET is no longer operative as the lone pairs are not available. Therefore the fluorescence enhancement is the result of metal binding.

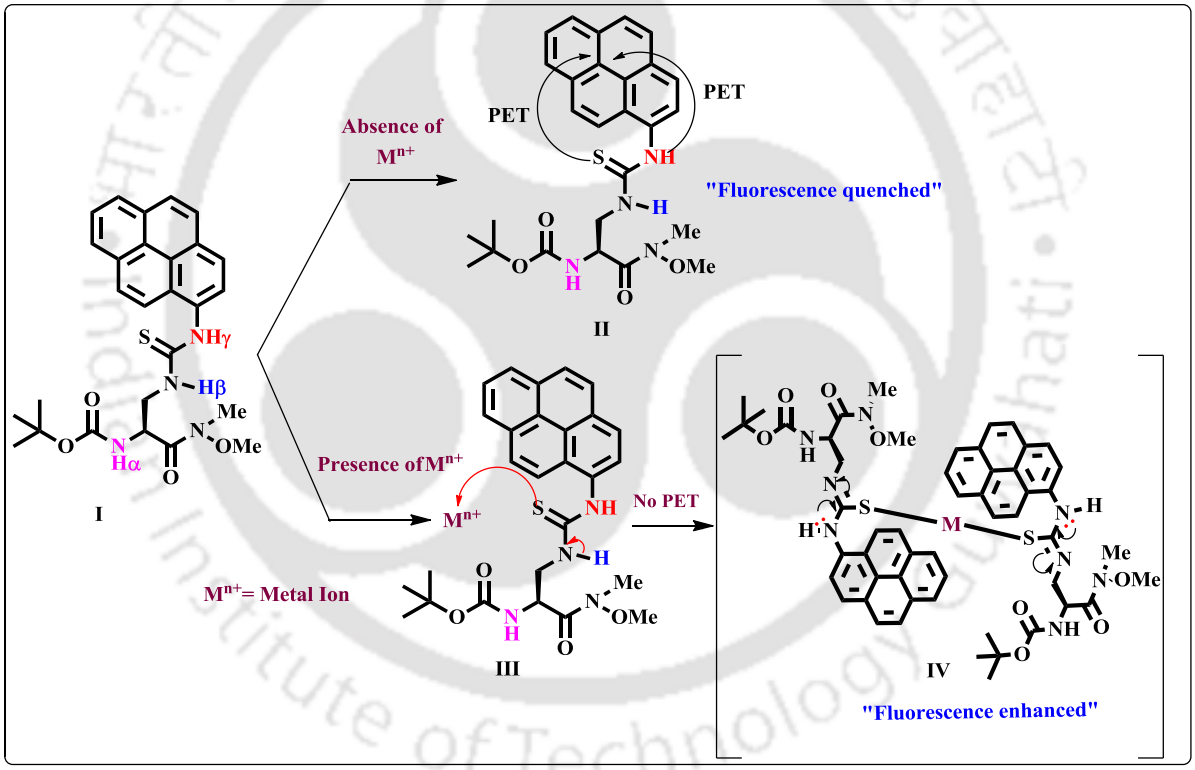


Figure 5.33. Probable mechanism for fluorescence quench and enhancement in (a) absence and (b) presence of metal ion.

5.7.7. Study of Possible Binding Mode via NMR, UV and IR

In order to support the mode of binding interaction between the probe and the metal or probable complex structure, we titrated the known concentration of probe (1 eqiv.) with variable concentrations of Hg^{2+} i.e 0 to 1 equivalent of ion in d_6-DMSO (**Figure 5.34**). We found that after the sequential addition of Hg^{2+} ion to the probe, the *β*-NH was shifted from *δ* 7.59 to 8.05 and *γ*–NH was shifted from *δ* 10.18 to 11.14 ppm in ^{1}H-NMR whereas α-NH remained at almost same position. We also found that when metal equivalent is 0.5, β-NH has vanished but *γ*-NH was still clearly visible in the NMR spectrum. So, this result is suggesting the probable formation of stable iminium complex with Hg^{2+} ion in 2:1 ratio which has been earlier supported by fluorescence and mass spectroscopy. For further confirmation of stable imine complex, we did variable temperature (VT) NMR (20 ^{o} C to 75 ^{o} C) experiment with 1 equivalent addition of Hg^{2+} ion and we found that with increasing temperature the *γ*-NH and α-NH protons shifted to up field (**Figure 5.35)** due to loss of intermolecular H-bonding. There is no reappearance of *β*-NH peak in the NMR spectrum which infers the formation of stable imine complex with Hg^{2+} ion in 2:1 ratio.

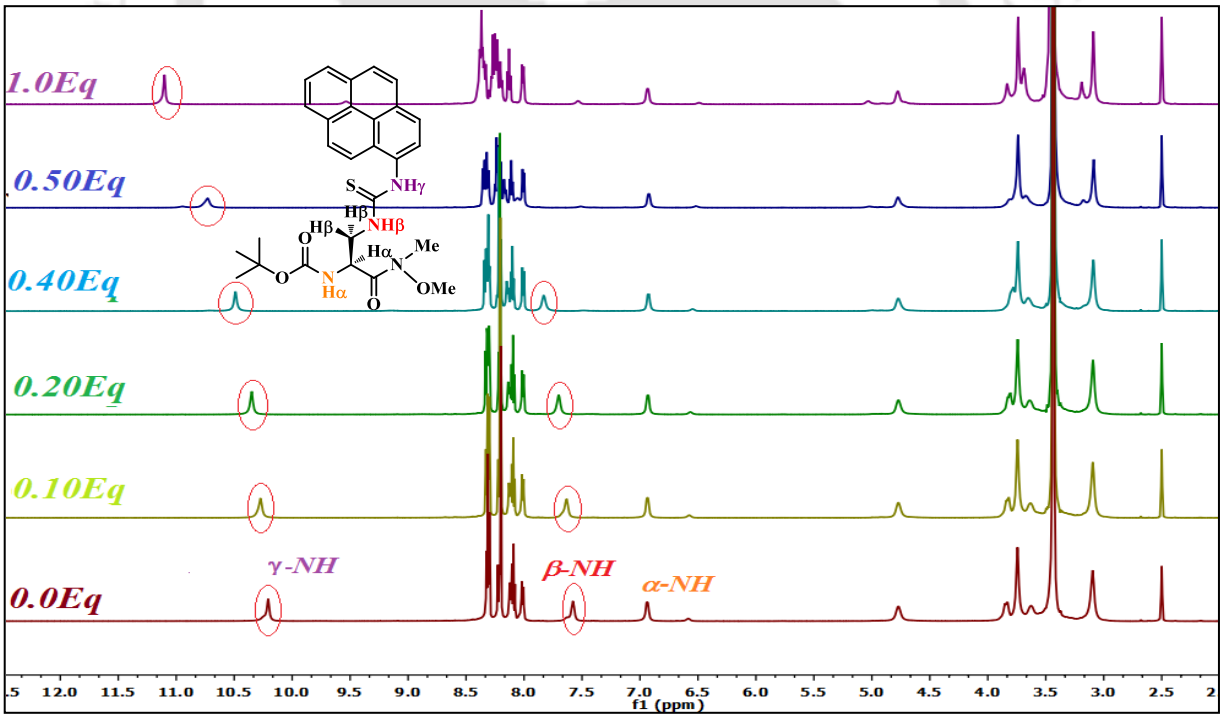


Figure 5.34. ^{1}H-NMR titration of probe ^{PyTU}**Ala** with Hg^{2+} ion up to 1 equivalent in d_6-DMSO.

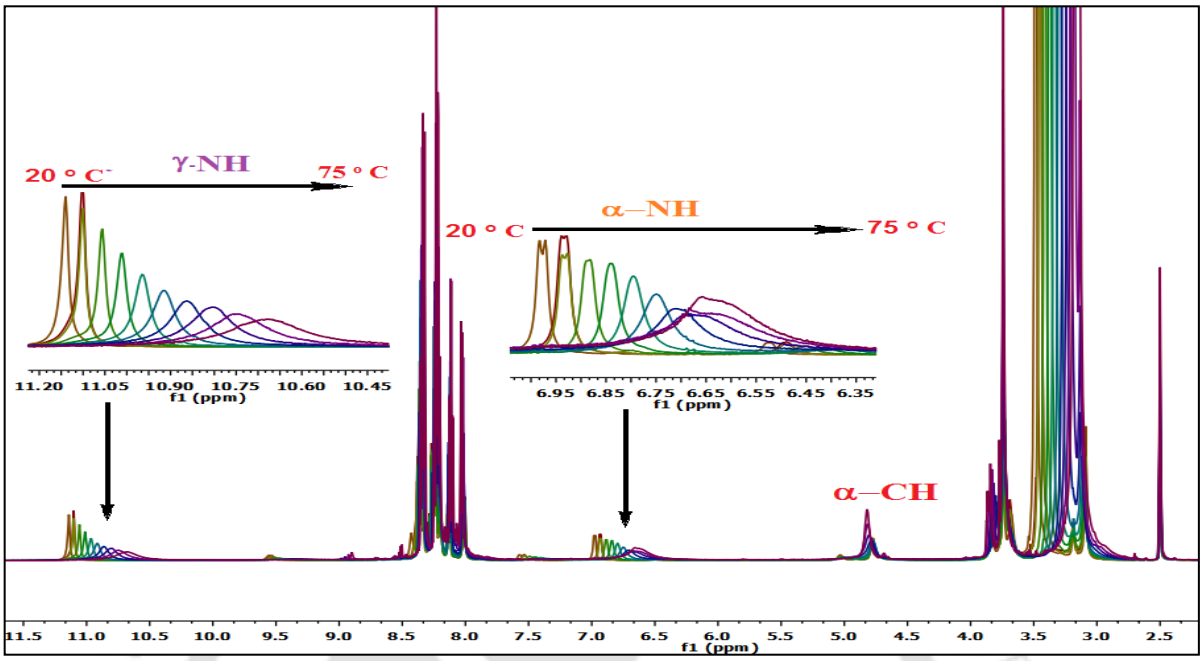


Figure 5.35. VT-NMR (20 ^{o} C to 75 ^{o} C) of probe ^{PyTU}**Ala** after 1 equivalent addition of Hg^{2+} ion in d_6-DMSO.

For further confirmation, we carried out real time UV-visible titration to check the imine formation. Interestingly, we found that during the addition of Hg^{2+} ion there is a generation of probable imine band at 358 nm (**Figure 5.36a**) at which the intensity increases up to 1 equivalent. Whereas, for the Cu^{2+} ion the imine band is appeared as broad (**Figure 5.36b)** at 10 equivalent concentrations.

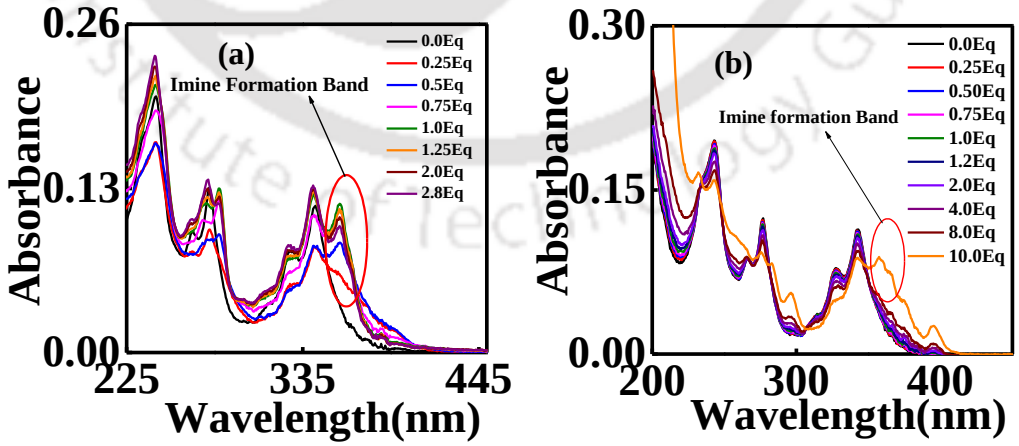


Figure 5.36. Real time UV-visible titration of the probe (a) in presence of Hg^{2+} ion and (b) in presence of Cu^{2+}ion.

Afterward, we used IR as a supportive tool for confirming the imine linked complex formation. Generally, the Schiff base compounds have characteristic stretching frequency at around 1615-1650 cm^{-1} which is characteristic to imine part vibration [22]. Similarly, when we took IR of the probe in presence of 1 equivalent of Hg^{2+} ion and 10 equivalent of Cu^{2+} ion, we found, there is clear appearance of strong band at 1613 cm^{-1} for Hg^{2+} ion and very weak broadband at 1621 cm^{-1} for Cu^{2+} ion, which denotes the formation of the iminium complex. However, the probe did not have any such characteristic peak at IR spectra (**Figure 5.37**) in the similar range. Thus, these results also support the formation of new imine linkage complex in presence of metal ions like Hg^{2+} and Cu^{2+}.

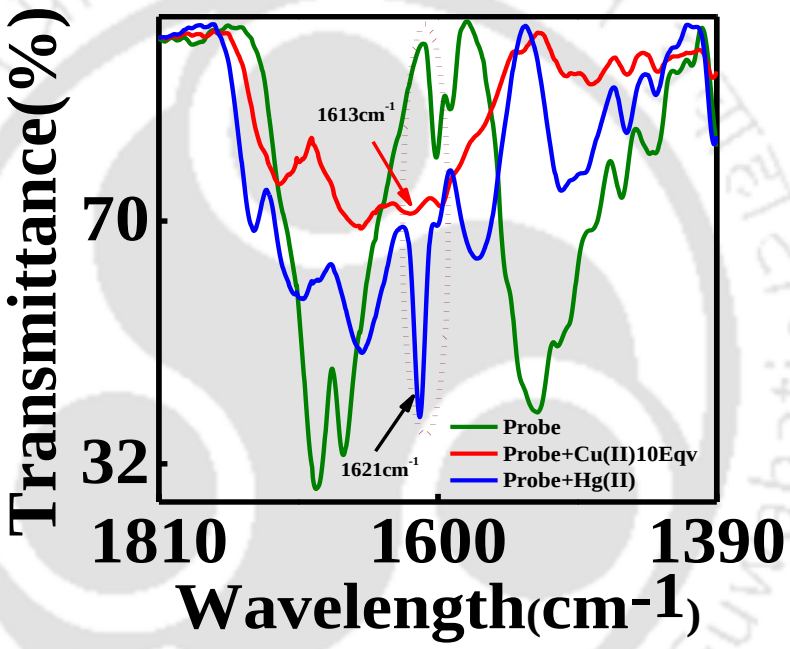
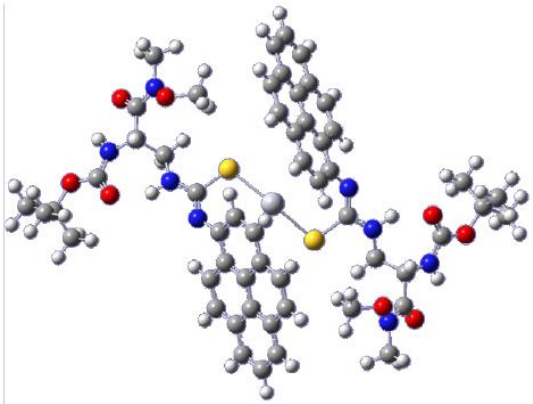


Figure 5.37. IR experiments of the probe in absence and presence of Hg^{2+} and Cu^{2+} ion.

Further support of this coordination came from an optimised geometry obtained from a density functional calculation using Gaussian 09 program package.[23] Thus, from a TD-DFT calculation at B3LYP/LANL2DZ level of theory, we observed that the two S atoms from two thiourea units had a significant amount of LUMO coefficient with maximum coefficient remaining with Hg^{2+} indicating a significant overlap of the Hg^{2+} and the probes' S atoms. Thus, theoretical study also supported the mode of binding (**Section 5.7.8**).

5.7.8. DFT Calculation

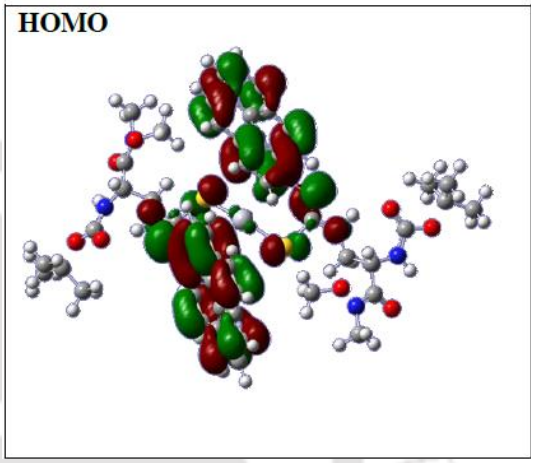


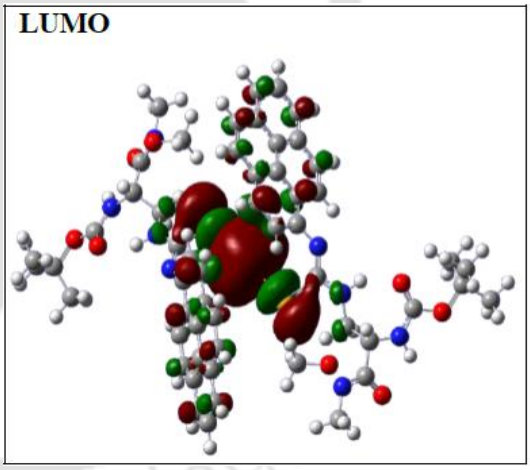
Calculation Method **B3LYP/LANL2DZ; E(RB3LYP) = -3195.06776612;**

Standard orientation: **Imaginary Freq = 0;**

Center Number	Atomic Number	Atomic Type	Coordinates (Angstroms) X	Y	Z
1	6	0	6.787199	-1.001986	-0.884282
2	7	0	7.746403	-0.163549	-1.612190
3	6	0	7.902605	1.169414	-1.373528
4	6	0	7.179731	-2.455237	-1.213073
5	6	0	5.300786	-0.740926	-1.289947
6	8	0	8.955925	1.670282	-2.105158
7	8	0	7.173068	1.840829	-0.592517
8	6	0	9.325458	3.141207	-2.075771
9	6	0	10.512656	3.187029	-3.049483
10	6	0	8.138452	3.976858	-2.590040
11	6	0	9.748788	3.529637	-0.647146
12	8	0	7.900172	-2.727056	-2.214113
13	7	0	6.699047	-3.448937	-0.390000
14	6	0	3.444188	0.565326	-0.145693
15	16	0	2.348913	-0.945422	-0.383308
16	6	0	7.173880	-4.837035	-0.458927
17	8	0	6.177931	-3.034531	0.890868
18	6	0	4.822113	-3.575246	1.140005
19	1	0	6.872033	-0.833183	0.194308
20	1	0	8.341711	-0.633619	-2.290829
21	1	0	4.702234	-1.631244	-1.090377
22	1	0	5.270404	-0.559885	-2.374794
23	1	0	10.897733	4.211708	-3.121284
24	1	0	11.319719	2.532419	-2.702228
25	1	0	10.205430	2.856744	-4.048031
26	1	0	8.446293	5.026373	-2.681951
27	1	0	7.290074	3.919903	-1.903494
28	1	0	7.823151	3.625416	-3.580110
29	1	0	10.553187	2.872382	-0.295078
30	1	0	8.905317	3.462439	0.044460
31	1	0	10.124644	4.561000	-0.646654
32	1	0	7.822470	-5.051038	0.399675
33	1	0	6.330921	-5.536312	-0.467804
34	1	0	7.742134	-4.945529	-1.385218
35	1	0	4.532427	-3.130097	2.093869
36	1	0	4.849397	-4.667255	1.228581
37	1	0	4.121826	-3.268713	0.357927
38	7	0	3.080379	1.729046	0.317407
39	7	0	4.743269	0.389253	-0.540847
40	1	0	5.380615	1.179589	-0.387435
41	6	0	1.978745	2.110363	1.063320
42	6	0	1.290561	3.330112	0.732572
43	6	0	1.574308	1.404433	2.239141
44	6	0	0.240077	3.811629	1.582324
45	6	0	1.647803	4.094874	-0.436841
46	6	0	0.558513	1.893572	3.073703
47	1	0	2.122616	0.509386	2.518658
48	6	0	-0.445848	5.029985	1.252393
49	6	0	-0.127269	3.090492	2.770500
50	6	0	0.986416	5.257480	-0.756418
51	1	0	2.463792	3.724468	-1.048735
52	1	0	0.299986	1.346569	3.978133
53	6	0	-0.083312	5.764049	0.072219
54	6	0	-1.493513	5.523090	2.105064
55	6	0	-1.176465	3.611253	3.613858
56	1	0	1.267056	5.823236	-1.642233
57	6	0	-0.772812	6.961974	-0.236758
58	6	0	-2.154999	6.725394	1.761266
59	6	0	-1.833653	4.775858	3.296951
60	1	0	-1.435822	3.060126	4.515653
61	6	0	-1.796449	7.434486	0.600675
62	1	0	-0.498014	7.518254	-1.130100
63	1	0	-2.939479	7.106534	2.412069
64	1	0	-2.616846	5.163425	3.945513
65	1	0	-2.311715	8.359514	0.353071
66	80	0	-0.000026	-0.000020	0.043394
67	16	0	-2.349014	0.945412	-0.383051
68	6	0	-3.444313	-0.565290	-0.145357
69	7	0	-3.080472	-1.728984	0.317793
70	7	0	-4.743391	-0.389253	-0.540509
71	6	0	-1.978778	-2.110248	1.063638
72	6	0	-5.300915	0.740862	-1.289701
73	1	0	-5.380700	-1.179640	-0.387197
74	6	0	-1.290681	-3.330062	0.732946
75	6	0	-1.574159	-1.404195	2.239323
76	6	0	-6.787326	1.001961	-0.884031
77	1	0	-4.702342	1.631193	-1.090241
78	1	0	-5.270556	0.559705	-2.374530
79	6	0	-0.240094	-3.811511	1.582608
80	6	0	-1.648113	-4.094951	-0.436325
81	6	0	-0.558280	-1.893282	3.073814
82	1	0	-2.122377	-0.509081	2.518804

Center Number	Atomic Number	Atomic Type	Coordinates (Angstroms) X	Y	Z
83	7	0	-7.746537	0.163564	-1.611977
84	6	0	-7.179849	2.455228	-1.212764
85	1	0	-6.872161	0.833122	0.194553
86	6	0	0.445763	-5.029915	1.252711
87	6	0	0.127430	-3.090256	2.770657
88	6	0	-0.986796	-5.257606	-0.755867
89	1	0	-2.464186	-3.724599	-1.048140
90	1	0	-0.299625	-1.346188	3.978152
91	6	0	-7.902538	-1.169468	-1.373581
92	1	0	-8.341923	0.633684	-2.290514
93	8	0	-7.900335	2.727096	-2.213758
94	7	0	-6.699113	3.448887	-0.389675
95	6	0	0.083045	-5.764100	0.072669
96	6	0	1.493544	-5.522942	2.105284
97	6	0	1.176726	-3.610949	3.613928
98	1	0	-1.267580	-5.823459	-1.641574
99	8	0	-8.955809	-1.670343	-2.105276
100	8	0	-7.172874	-1.840937	-0.592735
101	6	0	-7.173868	4.837014	-0.458544
102	8	0	-6.177924	3.034408	0.891133
103	6	0	0.772482	-6.962069	-0.236278
104	6	0	2.154967	-6.725290	1.761519
105	6	0	1.833857	-4.775595	3.297050
106	1	0	1.436207	-3.059740	4.515638
107	6	0	-9.325118	-3.141333	-2.076174
108	1	0	-7.822186	5.051095	0.400242
109	1	0	-6.330871	5.536242	-0.467720
110	1	0	-7.742393	4.945487	-1.384672
111	6	0	-4.822102	3.575124	1.140224
112	6	0	1.796238	-7.434502	0.601056
113	1	0	0.497539	-7.518447	-1.129514
114	1	0	2.939546	-7.106364	2.412240
115	1	0	2.617128	-5.163112	3.945548
116	6	0	-10.512349	-3.187132	-3.049847
117	6	0	-8.138001	-3.976686	-2.590673
118	6	0	-9.748329	-3.530127	-0.647616
119	1	0	-4.532272	3.129773	2.093948
120	1	0	-4.849414	4.667113	1.229034
121	1	0	-4.121903	3.268795	0.357986
122	1	0	2.311459	-8.359561	0.353477
123	1	0	-10.897281	-4.211852	-3.121839
124	1	0	-11.319492	-2.532710	-2.702425
125	1	0	-10.205215	-2.856601	-4.048343
126	1	0	-8.445672	-5.026233	-2.682786
127	1	0	-7.289599	-3.919732	-1.904155
128	1	0	-7.822803	-3.624986	-3.580684
129	1	0	-10.552821	-2.873074	-0.295383
130	1	0	-8.904844	-3.462940	0.043972
131	1	0	-10.124018	-4.561550	-0.647325





5.8. Conclusion

The newly developed pyrenylthioureayl alanine act as a novel and smart fluorescence switch-on sensor for Hg^{2+} and Cu^{2+} ions with more selectivity towards Hg^{2+} ion. The probe is also able to sense Hg^{2+} metal ion without any interference of other metal ions. The binding affinity of Hg^{2+} ion with probe is very high, about $3.93x10^{10}$ (M^{-1}) whereas for Cu^{2+} ion it is $4.11x10^{-6}$ (M^{-1}) which has also been confirmed by the optical and NMR titration. Binding stoichiometry of the probe is 2:1 (Probe: Metal) for both metal ions which is evident by fluorescence, NMR, ESI-MS and DFT optimisation. Furthermore, we have shown that quantum yield of ligand has increased much more in presence of metal ions. On the other hand, from the life time measurement it is clear that life time of probe has increased from 3.11 ns to 4.44 ns upon addition of Hg^{2+} (0 to 1 Eqiv.) ion which is supporting the previous results. The designed chemosensor is able to detect Hg^{2+} ion in 93nM range whereas for Cu^{2+} ion detection limit is 318nM in a semi aqueous medium. The low and promising detection limit of our present probe for selective sensing of Hg^{2+} ion, might give idea for the design of new, efficient, and simple fluorescence light-up probe for metal ion detection. So, our study proved that these types of unnatural amino acids have the highest sensitivity for the detection of transition metal ions.

5.9. Experimental Section

5.9.1. Synthesis and characterisation: *(S)-Pyrenylthioureayl alanine (***3.063***):* Synthesised and described in **Chapter 3.**

5.9.2. Photophysical Studies of the Amino Acids

UV-visible measurements: All the UV–visible spectra of **^{PyTU}Ala** (10 μM) were measured in different solvents using a UV-visible spectrophotometer with a cell of 1 cm path length. The measurements were carried out in absorbance mode. The absorbance values of the sample solutions were measured in the wavelength regime of 200–600 nm. All the sample solutions were prepared before an hour of the measurement.

Fluorescence experiments: All the sample solutions were prepared as described in UV measurement experiments. Fluorescence spectra were obtained using a fluorescence spectrophotometer at 25 °C using 1 cm path length cell. The excitation wavelengths for all the cases were set at the excitation maxima of each sample in each solvents and emission spectra were measured in the wavelength regime of 352–580nm. All the sample solutions were prepared before an hour of the measurement. Fluorescence emission spectra were recorded in a fluorescence spectrophotometer using a cell of 1 cm path length and 3 nm excitation/emission slit width at 25 ^{o} C. All the sample solutions were prepared just before doing the experiment. Total volume of 1.0 ml from a stock solution of 2 ml of 10 μM concentration for each case was used for fluorescence experiment in 1 ml cell. Fluorescence emissions were collected exciting the samples at the wave length corresponding to their absorption maxima. The fluorescence quantum yield (*Φf*) was determined using quinine sulphate as a reference with the known Φ_R = 0.54 in 0.1 molar solutions in sulphuric acid where we used 2nm excitation/emission slit width. The equation used for quantum yield (QY) calculation

$$\Phi_s = \Phi_R \times \frac{I_S}{I_R} \times \frac{A_R}{A_S} \times \frac{{\eta_S}^2}{{\eta_R}^2}$$

Where, Φ_S = sample's QY ; Φ_R = standard's QY (0.54 in 0.1 M H_2SO_4); I_S = area under the emission curve (range-352-640nm) of sample; I_R = area the emission curve (range-352-640nm) of standard; A_R = standard's absorbance; A_S = sample's absorbance at 342nm; η_S = refractive index of solvent, which was used for dispersion of sample (i.e. Water/Acetonitrile 7:3); η_R = refractive index of solvent, which was used for dispersion of standard (i.e. 1.33 for water) . Here we consider the ratio (η_S^2 /η_R^2) of refractive index is 1.

5.9.3. ^{1}H, ^{13}C and HSQC (^{1}H- ^{13}C) Characterisation of ^{PyTU}Ala

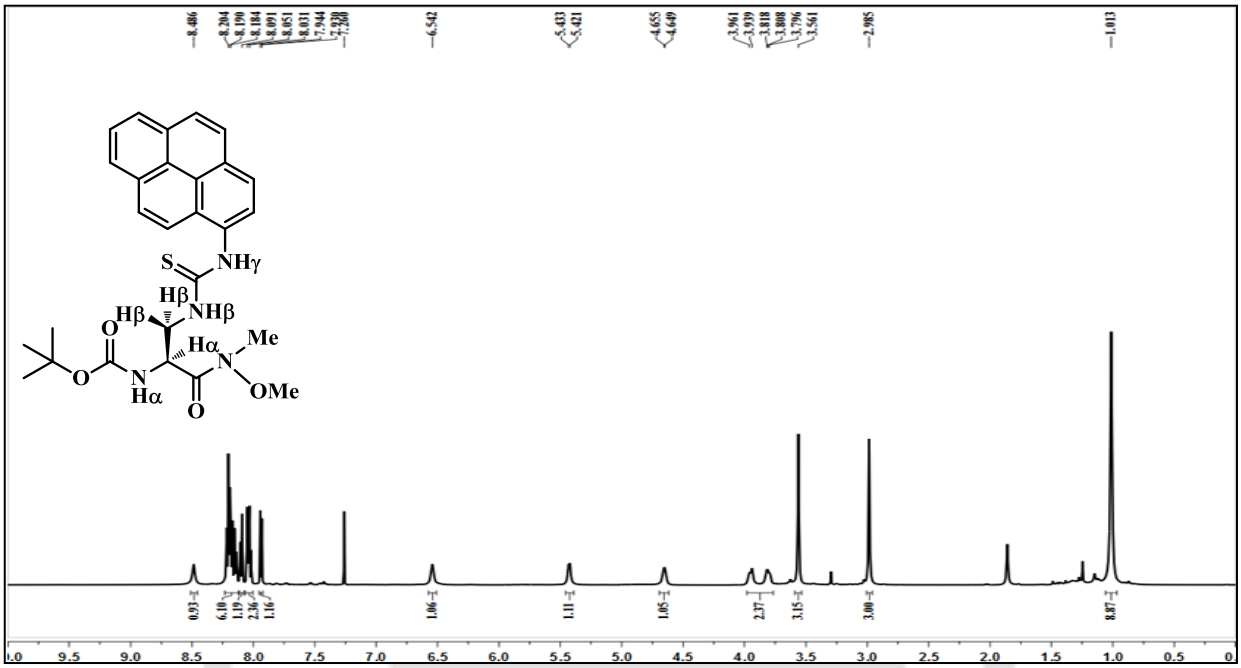


Figure 5.38. ^{1}H-NMR spectra of Probe ^{PyTu}**Ala.**

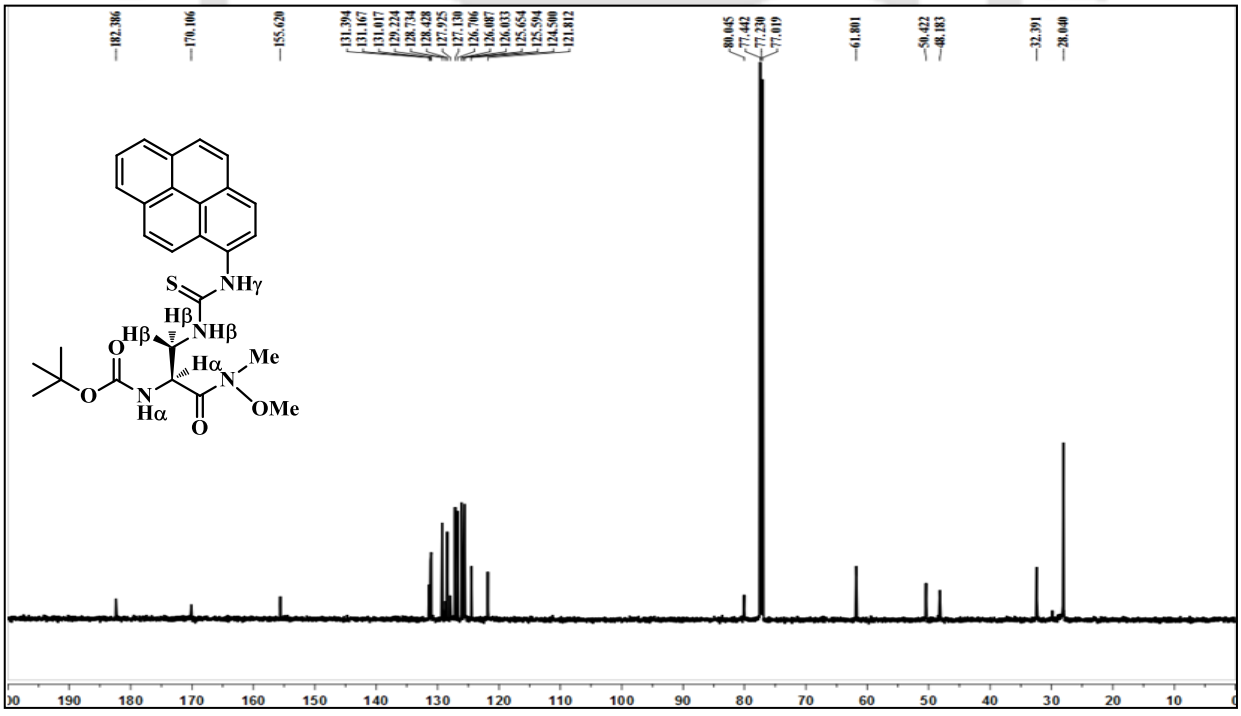


Figure 5.39. ^{13}C-NMR spectra of Probe ^{PyTu}**Ala.**

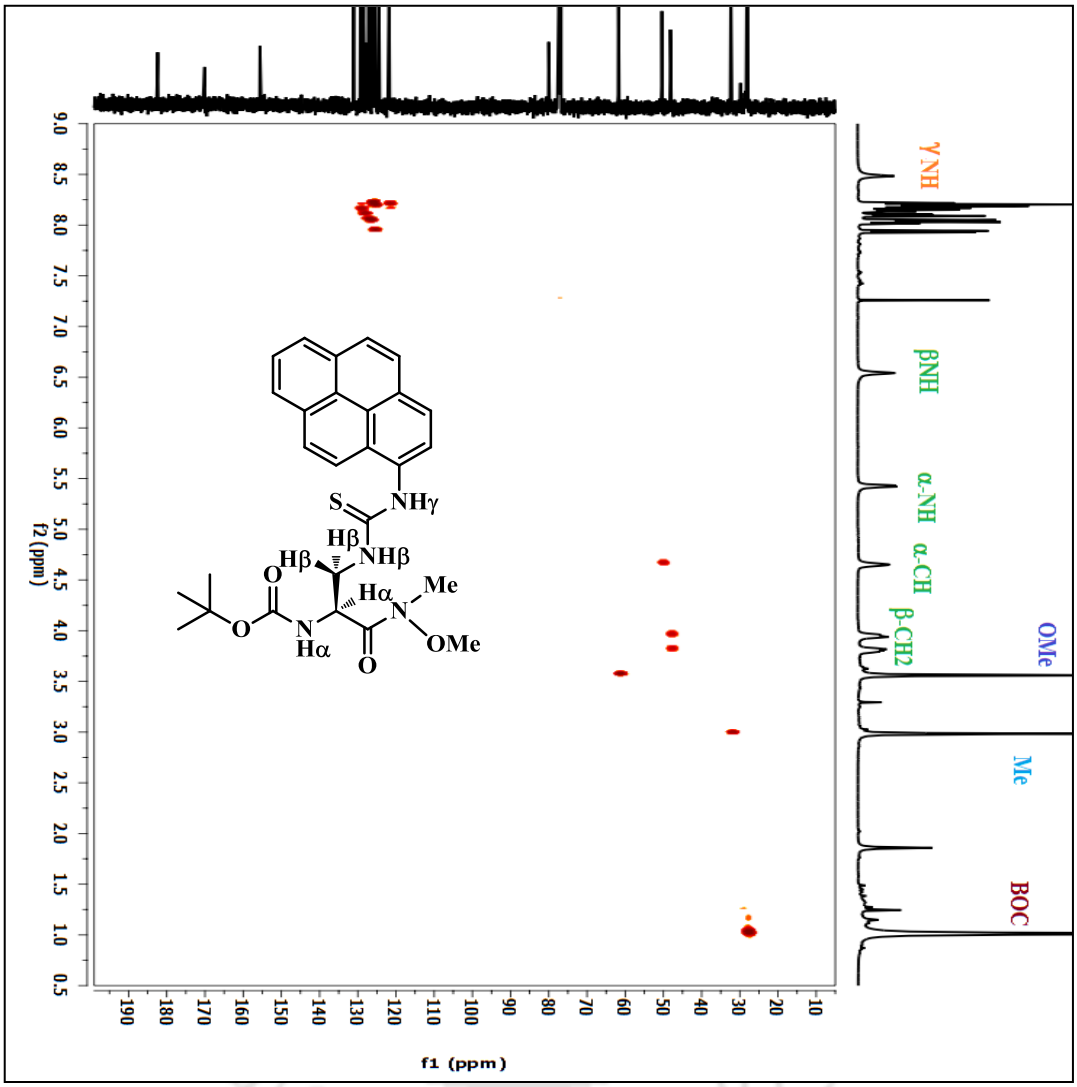


Figure 5.40. HSQC (^{1}H-^{13}C) NMR spectra of Probe **^{PyTu}Ala.**

5.10. References

1. (a) Lee, H.; Lee, H.-S.; Reibenspies, J.H.; Hancock, R.D. *Inorg. Chem.* **2012**, *51*, 10904. (b) Bencini, A.; Bernardo, M. A.; Bianchi, A.; Garcia-Espana, E.; Giorgi, C.; Luis, S.; Pina, F.; Valtancoli, B. *Adv. Supramol. Chem*. **2002**, *8*, 79. (c) Valeur, B.; Leray, I. *Coord. Chem. Rev*. **2000**, *205*, 3.(d) Czarnik, A. W. Trends *Org. Chem*. **1993**, *4*, 123.

2. (a) de Silva, A.P.; Fox, D.B.; Huxley, A.J.M.; Moody,T.S. *Coord. Chem. Rev*. **2000**, *41*, 205. (b) Nguyen, B.T.; Ansyln, E.V. *Coord. Chem. Rev*. **2006**, *250*, 3118.

3. Boening, D.W. *Chemosphere*. **2000**, *40*, 1335.

4. (a) Linder M. C.; Hazegh-Azam, M. *Am. J. Clin. Nutr*. **1996**, *63*, 797S. (b) Uauy, R.; Olivares, M.; Gonzalez, M. *Am. J. Clin. Nutr*. **1998**, *67*, 952S. (c) Harris Z. L.; Gitlin, J. D. *Am. J. Clin. Nutr*. **1996**, *63*, 836S. (d) Scheinberg I. H.; Sternlieb, I. *Am. J. Clin. Nutr*. **1996**, *63*, 842S.

5. (a) Li, G.-K.; Xu, Z.-X.; Chen, C.-F.; Huang, Z.-T. *Chem. Commun*. **2008**, 1774. (b) Yu, M.; Shi, M.; Chen, Z.; Li, F.; Li, X.; Gao, Y.; Xu, J.; Yang, H.; Zhou, Z.; Yi, T.; Huang, C. *Chem. Eur. J*. **2008**, *14*, 6892. (c) Zhou, Y.; Wang, F.; Kim, Y.; Kim, S.-J.; Yoon, *J. Org. Lett*. **2009**, *11*, 4442. (d) Mu, H.; Gong, R.; Ma, Q.; Sun, Y.; Fu, E. *Tetrahedron Lett*. **2007**, *48*, 5525 (e) Saleem, M.; Rafiq, M.; Hanif, M. *J. Fluoresc*. **2017**, *27*, 31.

6. (a) Zhang, G.; Zhang, D.; Yin, S.; Yang, X.; Shuai,Z.; Zhu, D. *Chem. Commun*. **2005**, 2161. (b) Misra, A.; Shahid, M.; Srivastava, P. *J Incl Phenom Macrocycl Chem.* **2011**, *69*, 119. (c) Guo, X.; Qian, X.; Jia, L. *J. Am. Chem. Soc*. **2004**, *126*, 2272. (b) Nolan,E.M.; Lippard, S.J.; *J. Am. Chem. Soc*. **2003**,*125*, 14270. (c) Hennrich, G.; Walther, W.; Resch-Genger, U.; Sonnenschein, H. *Inorg. Chem*. **2001**, *40*, 641. (d)Descalzo, A.B.; Martinez-manez,R.; Radeglia, R.; Rurack, K.; Soto, J. *J. Am. Chem. Soc*. **2003**, *125*, 3418. (e) Ono, A.; Togashi, H. *Angew. Chem., Int. Ed*. **2004**, *43*, 4300. (g) Wu, Q.; Anslyn, E.V. *J. Am. Chem. Soc*. **2004**, *126*, 14682.

7. (a) Martinez, R.; Zapata, F.; Caballero, A.; Espinosa, A.; Tarraga, A.; Molina, P. *Org. Lett*. **2006**, *8*, 3235. (b) Caballero, A.; Martinez,R.; Lioveras,V.; Ratera, I.; Vidal-Gancedo,J.; Wurst, K.; Tarraga, A.; Molina,P.; Veciana, J. *J. Am. Chem. Soc*. **2005**, *127*, 15666. (c) Martinez, R.; Espinosa, A.; Tarraga,A.; Molina, P. *Org. Lett*. **2005**, *7*, 5869.

8. (a) Li, G.-K.; Xu, Z.-X.; Chen, C.-F.; Huang, Z.-T. *Chem. Commun*. **2008**, 1774-1776. (b) McClure, D.S. *J. Chem. Phys*. **1952**, ***20***, 682. (c) Nolan, E. M.; Lippard, S. J. ; *J. Am. Chem. Soc.* **2003**, *125*, 14270. (d) Prodi, L.; Montali, M. *J. Am. Chem. Soc*. **2000**, *122*, 6769. (e) Martinez,R.; Espinosa,A. Tarraga,A.; Molina,P. *Org. Lett.* **2005,** *7*, 5869.

9. (a) Wang, J.; Qian, X.; *J. Org. Chem*. **2006**, *71*, 4308. (b) Zhou,Y.; Zhu,C-Y.; Gao, X.-S.; You, X-Y.; Yao, C. *Org. Lett*. **2010**, *12*, 2566.

10. (a) Qi, X.; Jun, E. J.; Xu, L.; Kim, S-J.; Hong, J. S. J.; Yoon,Y. J.; Yoon, J. *J. Org. Chem.* **2006**, *71*, 2881. (b) Xu, Z.; Xiao, Y.; Qian, X.; Cui, J.; Cui, D. *Org. Lett*. **2005**, *7*, 889. (c) Xu, Z.; Qian, X.; Cui, J. *Org. Lett*. **2005**, *7*, 3029. (d) Martı´nez, R.; Espinosa, A.; Ta´rraga, A.; Molina, P. *Org. Lett*. **2005**, *7*, 5869. (e) Martı´nez, R.; Zapata, F.; Caballero, A.; Espinosa, A.; Ta´rraga, A.; Molina, P. *Org. Lett*. **2006**, *8*, 3235. (f) Zhang, X.; Shiraishi, Y.; Hirai, T. *Org.*

Lett. **2007**, *9*, 5039. (g) Sheng, R.; Wang, P.; Gao, Y.; Wu, Y.; Liu, W.; Ma, J.; Li, H.; Wu, S.*Org. Lett.* **2008**, *10*, 5015. (h) Yu, M-M.; Li, Z-X.; Wei, L-H.; Wei, D-H.; Tang, M-S. Org. Lett. **2008**, *10*, 5115. (i) Zhou, Y.; Wang, F.; Kim, Y.; Kim, S-J.; Yoon. J. J. *Org. Lett.* **2009**, *11*, 4442.(j) Kim, H. J.; Hong, J.; Hong, A.; Ham, S.; Lee, J.; Kim, J. S. *Org. Lett.* **2008**, 10, 1963. (K) Jung, H. S.; Park, M.; Han, D. Y.; Kim, E.; Lee, C.; Ham, S.; Kim, J. S. *Org. Lett.* **2009**, 11, 3378.

11. (a) Du, J. J.; Fan, J. L.; Peng, X. J.; Sun, P. P.; Wang, J. Y.; Li, H. L.; Sun, S. G. *Org. Lett.* **2010**, *12*, 476.(b) Saha, S.; Mahato, P.; Reddy, G. U.; Suresh, E.; Chakrabarty, A.; Baidya, M.; Ghosh, S. K.; Das, A. *Inorg. Chem.* **2012**, *51*, 336. (c) Sivaraman,G.; Anand, T.; Chellappa, D.; *RSC Advances*. **2012**, *2*, 10605. (d) Wanichacheva,N.; Praikaew,P.; Suwanich,T.; Sukrat,K. *Spectrochimica Acta* A. **2014**, *118* , 908. (e) Yea,H.; Gea, F.; Chenb, X.-C.; Li,Y.; Zhanga, H.; Zhaoa,B.-X.; Miaob, J.-Y. *Sensors Actuators, B.* **2013**, *182*, 273. (f) Wan, C.-F.; Lin, H.-Y.; Chienb, C.; Wub, A.-T. *Luminescence,* **2013.** (g)Li, M.; Lu, H-Y.; Liu, R-L.; Chen, J-D.; Chen, C.-F. *J. Org. Chem.* **2012**, *77*, 3670. (h) De la Cruz-Guzman, M.; Aguilar-Aguilar, A.; Hernandez-Adame,L.; Bañuelos-Frias, A.; Medellín-Rodríguez, F J.; Palestino,G. *Nanoscale Res. Lett.* **2014**, *9*, 431

12. (a) Misra, A.; Shahid, M.; Srivastava,P.; Dwivedi,P.; *J Incl Phenom Macrocycl Chem.* **2011**, *69*, 119. (b) Udhayakumaria, D.; Velmathi, S.; Venkatesanb,P.; Wu, S.-P. *Anal. Methods.* **2015**, *7*, 1161. (c) Jun, M.E.; Roy,B.; Ahn, K.H.; *Chem. Commun*. **2011**, **47**, 7583. (d)Kwon, J. Y.; Jang, Y. J.; Lee, Y. J.; Kim, K.-M. ; Seo, M.-S.; Nam, W.; Yoon, J. *J. Am. Chem. Soc.* **2005**, *127*, 10107. (e) Kim, H.-N.; Lee, M.-H.; Kim, H.-J.; Kim, J.-S.; Yoon, J. *Chem. Soc. Rev*. **2008**, *37*, 1465. (f) Liu, B.; Tian, H. *Chem. Commun*. **2005**, 3156. (g) Vonlanthen, M.; Connelly, C.M.; Deiters, A.; Linden, A.; Finney, N.S. *J. Org. Chem.* **2014**, *79*, 6054. (h) Lu, X. Y.; Guo, Z. Q.; Feng, M. X.; Zhu, W. H. *ACS Appl. Mater. Interfaces.* **2012**, *4*, 3657. (i) Lee, M. H.; Lee, S. W.; Kim, S. H.; Kang, C.; Kim, J. S. *Org. Lett.* **2009**, *11*, 2101. (j)Srivastava, P.; Ali, R.; Razi, S. S.; Shahid, M.; Misra, A. *Sens. Actuators, B.* **2013**, *181*, 584.

13. (a) Zhang, X. L.; Xiao, Y.; Qian, X. H. *Angew. Chem. Int., Ed.* **2008**, *47*, 8025.(b) Shang, G. Q.; Gao, X.; Chen, M. X.; Zheng, H.; Xu, J. G. *J. Fluoresc.* **2008** *18*, 1187. (c) Liu, Y. L.; Lv, X.; Zhao, Y.; Chen, M. L.; Liu, J.; Wang, P.; Guo, W. *Dyes Pigm.* **2012**, *92*, 909.

14. (a) Kim, J.S.; Quang, D.T. Chem. Rev. **2007**, *107*, 3780–3799. (b) de Silva, A. P.; Gunaratne, H. Q. N.; Gunnlaugsson, T.; Huxley, A. J. M.; McCoy, C. P.; Rademacher, J. T.; Rice, T. E. *Chem. Rev*. **1997**, *97*, 1515. (c) Fabbrizzi, L. *Coord. Chem. Rev*. **2000**, *205*, 1-232. (d) Ellis, A. B.; Walt, D. R. *Chem. Rev*. **2000**, *100*, 2477-2738

15. (a) Yoon, S.; Miller, E. W.; He, Q.; Do, P. H.; Chang, C. J. *Angew. Chem., Int. Ed*. **2007**, *46*, 6658. (b) Blake, A. J.; Bencini, A.; Caltagirone, C.; De Filippo, G.; Dolci, L. S.; Garau, A.; Isaia, F.; Lippolis, V.; Mariani, P.; Prodi, L.; Montalti, M.; Zaccheroni, N.; Wilson, W. *Dalton Trans*. **2004**, *17*, 2771. (c) Aragoni, M. C.; Arca, M.; Bencini, A.; Blake, A. J.; Caltagirone, C.; Decortes, A.; Demartin, F.; Devillanova, F. A.; Faggi, E.; Dolci, L. S.; Garau, A.; Isaia, F.; Lippolis, V.; Prodi, L.; Wilson, W.; Valtancoli, B.; Zaccheroni, N. *Dalton Trans*. **2005**, *18*, 2994. (d) Tamayo, A.; Pedras, B.; Lodeiro, C.; Escriche, L.; Casab_o, J.; Capelo, J. L.; Covelo, B.; Kivekas, R.; Sillanpaa, R. *Inorg. Chem.* **2007**, *46*, 7818.

16. (c) Descazo, A. B.; Martínez-Máñez, R.; Redeglia, R.; Rurack, K.; Soto, J. *J. Am. Chem. Soc.* **2003**, *125*, 3418. (f) Li, M.-J.; Chu, B. W.-K.; Zhu, N.; Yam, V. W.-W. *Inorg. Chem*. **2007**, *46*, 720. (g) Li, M.-J.; Ko, C.-C.; Duan, G.-P.; Zhu, N.; Yam, V. W.-W. *Organometallics.* **2007**, *26*, 6091. (j) Yuan, M.; Li, Y.; Li, J.; Li, C.; Liu, X.; Lv, J.; Xu, J.; Liu, H.; Wang, S.; Zhu, D. *Org. Lett.* **2007**, *9*, 2313. (k) Ho, M.-L.; Chen, K.-Y.; Lee, G.-H.; Chen, Y.-C.; Wang, C.-C.; Lee, J.-F.; Chung, W.-C.; Chou, P.-T. *Inorg. Chem*. **2009**, *48*, 10304.

17. (a) Meng, X.M.; Zhu, M.Z.; Guo, Q.X.; *Chin. Chem. Lett*. **2007**, *18*, 1209. (b) Pearson, R.G. *J. Am. Chem. Soc*. **1963**, ***85***, 3533. (c)Rurack, K.; Bricks, J.L.; Schulz, B.; Maus, M.; Reck, G.; Resch-Genger, U. *J. Phys. Chem. A* **2000**, ***104***, 6171. (d) Yang, H.; Liu, Z.Q.; Zhou, Z.G.; Shi, E.X.; Li, F.Y.; Du, Y.K.; Yi, T.; Huang, C.H. *Tetrahedron* ,**2006**, ***47,*** 2911.

18. (a) Bag, S. S.; Jana, S.; Pradhan, M. K. *Bioorg. Med. Chem.* **2016**, *24*, 3579. (b) Bag, S. S.; Jana, S.; Pradhan, M. K. Pal, S. *RSC Adv*. **2016**, *6*, 72654. (c) Bag, S. S.; Jana, S.; Yashmeen, A.; De, S. *Chem. Commun.* **2015**, *51*, 5242. (d) Bag, S. S.; Jana, S.; Yashmeen, Senthilkumar, K.; Bag, R. *Chem. Commun.* **2014**, *50*, 433. (e) Ghosh, U.; Bag, S.S.; Mukherjee, C. *Sens. Actuators, B*. **2017**, *238*, 903. (f) Bag, S.S.; Kundu, R.; Talukdar S. *Tetrahedron Lett*. **2012**, *53*, 5875. (g) Shortreed, M.; Kopelman, R.; Kuhn, M.; Hoyland, B. *Anal.Chem.*, **1996**, *68*, 1414.

19. a) Ha-Thi, M.-H.; Penhoat, M.; Michelet,V.; Leray, I. *Org. Lett.* **2007**, *9*, 1133. (b) Shortreed, M.; Kopelman, R. ; Kuhn, M.; Hoyland, B.; *Anal. Chem.* **1996**, *68*, 1414. (c) Sulistiawaty, L.; Sugiarti, S.; Darmawan, N.; *Indones. J. Chem.* **2015**, *15*, 1. (d) *Guidelines for drinking-water quality*, 3rd ed.; World Health Organization: Geneva, **2004**; p 188. (e) Roy, B. C. Chandra, B.; Hromas, D.; Mallik, S. *Org. Lett.* **2003,** *5*, 11.

20. (a) Klonis, N; Sawyer, W. H. *Photochem. Photobiol.* **2003**, *77*, 502. (b) Shah, N. M.; Mehta, D. H. J. *Ind. Chem. Soc.* **1954**, *31*, 784. (c) Grabowksi, Z. R.; Rotkiewicz, K.; Rettig, W. *Chem. Rev*. **2003**, *103*, 3899. (d) Badger, G. M.; Walker, I. S. *J. Chem. Soc*. **1956**, 122.

21. (a) Al-Kady, A. S.; Gaber, M.; Hussein, M. M.; Ebeid, E.-Z. M *J. Phys. Chem. A* **2009**, *113*, 9474 and references therein. (b) J. R. Lakowicz,Principles of Fluorescence Spectroscopy; Springer: Singapore, **2006**; p 642.

22. Cinarli, A.; Gürbüz, D.; Tavman, A.; Birteksöz.A.S *Bull. Chem. Soc. Ethiop*. **2011**, *25*, 407.

23. Frisch, M. J., Gaussian 09, Revision D.01; Gaussian, Inc.: Wallingford, CT, **2013**.

This dissertation has a total of **5 Chapters. Chapter 2,3** explains the detailed synthesis of unnatural amino acids such as triazolyl, thiocyanyl, isothiocyanyl, thioureayl and aminotetrazolyl amino acids. The as synthesized amino acids were involved in synthesis of various side chain modified peptides as described in **Chapter 4**. Additionally, this chapter also includes, site specific ligation/ labeling study of short/long peptides under physiological conditions. In the final chapter; **Chapter 5,** application of thioureayl alanine unnatural amino acids as a potential metal sensing agent was shown.

Chapter 1 is a critical review of applications of UNAAs in genetic incorporation, site specific protein modification and fluorophoric unnatural amino acids as potential sensors. This chapter is basically a surmise of all the chapters of the present dissertation.

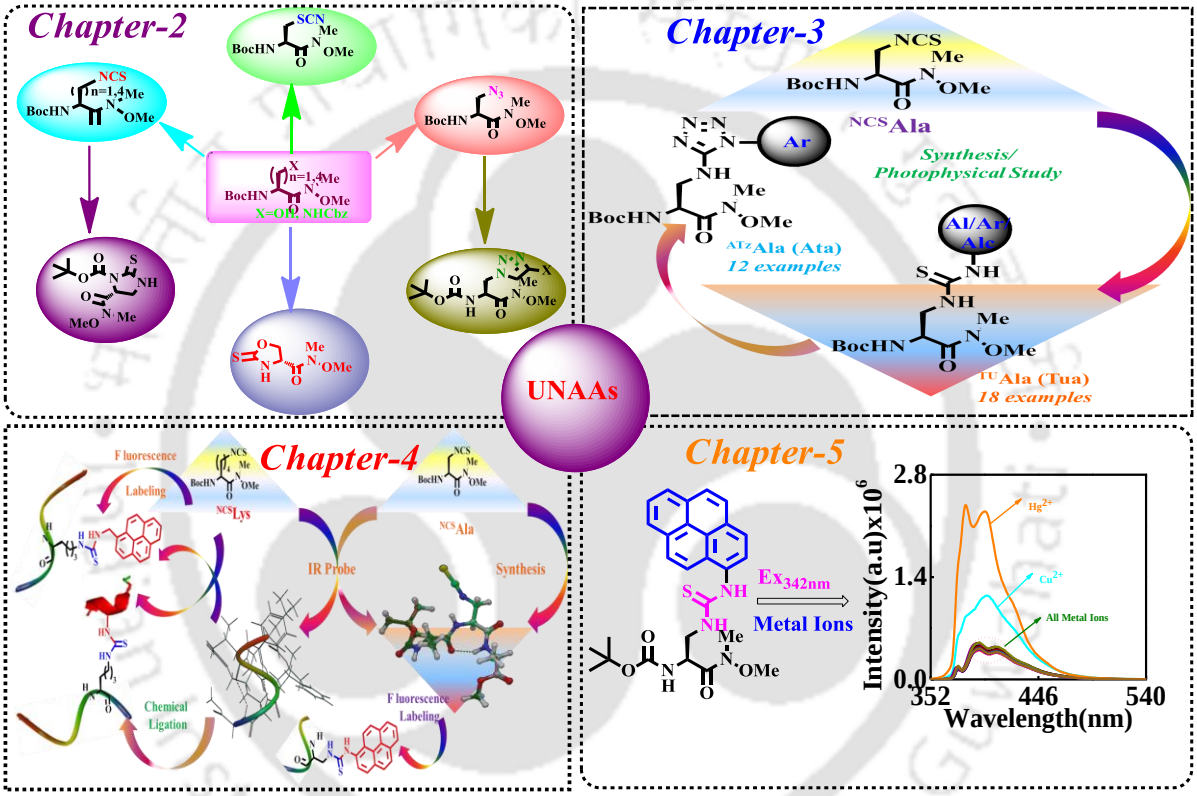
Chapter 2 elaborates the synthesis of triazolyl, thiocyanyl, and isothiocyanyl unnatural amino acids *via* 1. ‘‘click’’ chemistry, 2. nucleophilic substitution reaction and 3. decomposition of dithiocarbamic acids salts respectively. Our design concept is focused on possible site specific incorporation of triazolyl amino acids, isothiocyanyl and thiocyanyl amino acids (as a new class of side chain modified sulfur-containing amino acids).

Chapter-3 describes the application of isothiocyanyl alanine (**^{NCS}Ala)** in the synthesis of thioureayl alanines and then to aminotetrazolyl alanine as other new classes of unnatural amino acids. Our design concept exploited the electrophilicity of –NCS functionality. The UV-visible and fluorescence study of the **^{TU}Ala, ^{MeTU}Aa** and **^{TzA}Ala** derivatives are explored further to observe their photophysical properties.

Chapter 4 deals with development of isothiocyanyl alanine/lysine as solvatochromic IR responsive probes for possible application in modifying the local structures and electrostatic microenvironment of a short peptide containing **^{NCS}Ala/ ^{NCS}Lys** by exploring the infrared absorption property of –NCS functionality. The electrophilic property of **^{NCS}Ala** and **^{NCS}Lys** in short tri-/hexa-peptides was exploited to covalently label the chromophoric/fluorophoric amine and a detailed conformational study was undertaken.

Chapter 5 The synthesized fluorescent pyrenylthioureayl alanine amino acid was successfully applied as a switch on sensor for detection of Hg^{2+} and Cu^{2+} ions. Interestingly, we also found out that the specific sensing efficiency of the probe remained unaltered in presence of other metal ions.

Diagrammatic Representation of Thesis Work



ChemComm

COMMUNICATION

View Article Online
View Journal | View Issue

Triazolo-β-aza-ε-amino acid and its aromatic analogue as novel scaffolds for β-turn peptidomimetics†

Subhendu Sekhar Bag,* Subhashis Jana, Afsana Yashmeen and Suranjan De

Cite this: *Chem. Commun.*, 2015, **51**, 5242

Received 24th October 2014,
Accepted 10th December 2014

DOI: 10.1039/c4cc08414d

www.rsc.org/chemcomm

Triazolo-β-aza-ε-amino acid and its aromatic analogue (^{Al}TAA/^{Ar}TAA) in the peptide backbone mark a novel class of conformationally constrained molecular scaffolds to induce β-turn conformations. This was demonstrated for ^{Al}TAA in a Leu-enkephalin analogue and in a designed pentapeptide wherein the FRET process was established. Restricted rotation induced chirality and turn conformation into the achiral aromatic amino acid scaffold, ^{Ar}TAA, which in a short tripeptide backbone acted as a β-turn mimic as a β-sheet folding nucleator.

Introduction of a conformationally constrained nonpeptide isostere into peptide backbones in order to achieve desirable secondary structures along with pharmacologically viable peptide-based drug candidates is of great interest in recent times.[1] Among the various secondary structures,[2,3] β-turns[3] are important targets for mimicry, because they serve as recognition sites in peptides and proteins as well as they allow a protein chain to fold back upon itself to form a compact structure.[1c] Considerable efforts have thus been invested in delineating the impact of appended molecular scaffolds on the one hand and nucleating turn mimics on the other hand, on the conformational preferences of proteins and peptides in solution.[1c] Despite an exponential growth on the development of constrained non-peptidic molecular scaffolds, very few peptidomimetic drugs have been developed, necessitating an overhaul in the existing design principles.[1–3]

As part of our ongoing research efforts on the design of unnatural biomolecular building blocks[4] *via* click chemistry and β-turn peptidomimetics,[4c] we report herein on the simple synthesis and application of triazolo-β-aza-ε-amino acid (**1**, ^{Al}TAA) and its aromatic analogue (**2**, ^{Ar}TAA) as new and novel constrained molecular scaffolds (Fig. 1). Similar to the sugar amino acids,[3j,k] the rigid frameworks of pseudo-aromatic triazole units prompted us to use triazolo amino acids as novel molecular scaffolds in peptidomimetic studies.[5] These two molecules with constrained backbone angles, $\omega(i)$ and $\phi(i+1)$, are expected to induce folded conformations in linear peptides. The triazolo amino acids are advantageous with respect to their metabolic inertness, easy associability with biological targets and tolerance to various reaction conditions used in peptide synthesis. Moreover, the triazolo unit acts as a *trans*-amide mimetic, which makes the scaffolds more prone to nucleate β-turn structures while present in a short peptide backbone. Though click chemistry has been utilised in mimicking the protein's secondary structures, to the best of our knowledge, triazolo amino acids as scaffolds have not been explored.[6] We envisioned that incorporation of ^{Al}TAA/^{Ar}TAA into the backbone of a linear peptide, such as the Leu-enkephalin analogue, might adopt β-turn conformation.[3,7]

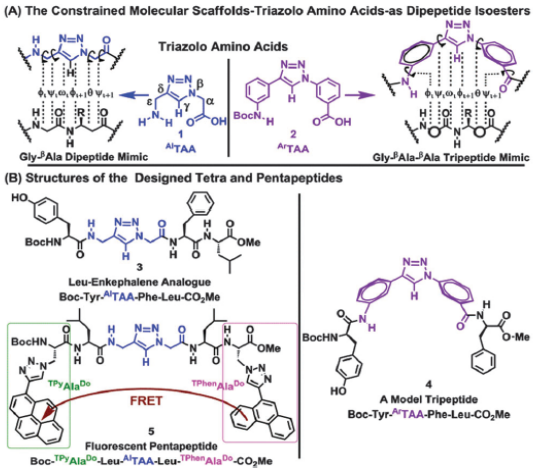


Fig. 1 (A) The constrained molecular scaffolds-triazolo amino acids (**1**, ^{Al}TAA and **2**, ^{Ar}TAA) and (B) structures of the designed peptidomimetics **3**–**5**.

Bioorganic Chemistry Laboratory, Department of Chemistry, Indian Institute of Technology Guwahati, North Guwhati-781039, Assam, India.
E-mail: ssbag75@iitg.ernet.in; Fax: +91-361-258-2349; Tel: +91-361-258-2324

† Electronic supplementary information (ESI) available: Synthesis, characterisation data, spectroscopic data, macromodel study and ^{1}H and ^{13}C NMR spectra. CCDC 997089. For ESI and crystallographic data in CIF or other electronic format see DOI: 10.1039/c4cc08414d

JOC The Journal of Organic Chemistry

Article

pubs.acs.org/joc

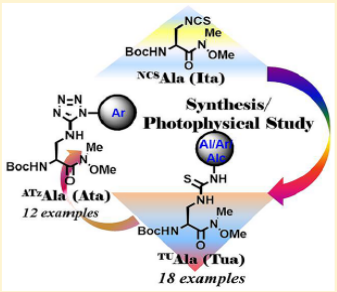
Isothiocyanyl Alanine as a Synthetic Intermediate for the Synthesis of Thioureayl Alanines and Subsequent Aminotetrazolyl Alanines

Subhendu Sekhar Bag* and Suranjan De

Bioorganic Chemistry Laboratory, Department of Chemistry, Indian Institute of Technology Guwahati 781039, India

Supporting Information

ABSTRACT: The synthesis of unnatural amino acids with small side-chain functionalities usable for further transformations is highly demanding for the expansion of the genetic code and other possible biotechnological applications. To this end, we wanted to report the utility of an unexplored unnatural amino acid, isothiocyanyl alanine (^{NCS}Ala = Ita), for the synthesis of another class of unnatural amino acids, thioureayl alanines (^{TU}Ala = Tua). The synthesis of a third class of unnatural amino acids, amino tetrazolyl alanines (^{ATz}Ala = Ata), in a very good yield was subsequently achieved utilizing thioureayl alanines. Thus, a variety of aliphatic- and aromatic-substituted thioureayl alanines and aromatic-substituted amino tetrazolyl alanines were successfully synthesized in good to excellent yields. The photophysical properties of three of the fluorescent unnatural amino acids from two classes were also studied and presented herein.



INTRODUCTION

In the journey of the expansion of the genetic code and other possible biotechnological applications, the synthesis of unnatural amino acids with various novel functional roles is highly demanding. In these developments, a large number of unnatural amino acids (UNAAs) have been introduced.[1] Many of the reported unnatural amino acids have also been incorporated into proteins.[1,2] Furthermore, several of them have been utilized as probes for investigating protein structures, functions, and dynamics and have even been applied to study interbiomolecular interactions in both in vitro and in vivo.[2−6] We also have reported fluorescent triazolyl UNAAs as labels for investigating the conformation of a short peptide and addressing fundamental photophysical aspects.[7] However, it has often been observed that the large sizes of labels cause problems of structural perturbation, leading these labels to be unsuitable as protein probes. Therefore, nonperturbing small functional groups as side-chain labels are highly desirable as probes over typical fluorescent probes.[8] Hence, the synthesis of small side-chain-modified unnatural amino acids is highly important.

As a part of our continuous research efforts in the design of unnatural amino acids via click chemistry,[7] we thought that it would be worthwhile to generate an amino acid with a small and reactive functional group that can further be exploited as a precursor for other novel families of UNAAs. Thus, we became interested in exploiting the unexplored isothiocyanyl (−NCS) functional group for the generation of isothiocyanyl alanine (^{NCS}Ala = Ita) as a new family of UNAAs. The logics behind our choice are the following: (a) isothiocyanates are a very important class of chromophores long known for their analytical use, such as in the determination of the primary structures of peptides and proteins and in amino acid analysis;[9] (b) recently, these chromophores have attracted much research attention because of their stability, synthetic easiness, and wide applications in organic synthesis, such as for the synthesis of thioureas and thioamides and as precursors for various sulfur-containing biologically active heterocycles.[10] For example, isothiocyanates derived from the hydrolysis of plant defense compounds, i.e., glucosinolates, have frequently been shown to act as insecticides.[11] Moreover, the unique Cotton effect of chiral isothiocyanates could be utilized as an alternative for X-ray diffraction studies for stereochemical assignments.[12]

In peptidomimetic chemistry, the bioisosteric replacement of amide carbonyls by thiocarbonyls has been profoundly reported for the synthesis of modified peptides with new properties compared to their parent counterparts.[13] Thus, several substituted peptidyl ureas and a few peptidyl thioureas have been reported.[14,15] Although there are only a few reports of isothiocyanates derived as α-isothiocyanato alkyl esters from amino acids/peptides, side-chain isothiocyanyl amino acids have not been reported in the literature.[15,16] Therefore, we envisaged that the hitherto unexplored isothiocyanyl alanine (^{NCS}Ala) would be worthwhile for exploiting its synthetic applications for the generation of new classes of amino acid analogues. Furthermore, side-chain-modified ^{NCS}Ala with a reactive −NCS functionality could be useful in designing side-chain fluorophoric thiouereayl amino acids/peptides and for labeling these with a fluorophore. Therefore, we wanted to report the synthesis of ^{NCS}Ala and its various synthetic applications.

RESULTS AND DISCUSSION

Synthesis of Isothiocyanyl Alanine (^{NCS}Ala = Ita). To explore our aim of the present investigation, we first synthesized

Received: August 19, 2017
Published: October 24, 2017

DOI: 10.1021/acs.joc.7b02103
J. Org. Chem. 2017, 82, 12276−12285

Bioorganic & Medicinal Chemistry Letters 28 (2018) 1404–1409

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry Letters

journal homepage: www.elsevier.com/locate/bmcl

Multipurpose isothiocyanyl alanine/lysine: Use as solvatochromic IR probes and in site specific labeling/ligation of short peptides

Subhendu Sekhar Bag *, Suranjan De

Bioorganic Chemistry Laboratory, Department of Chemistry, Indian Institute of Technology Guwahati, Guwahati 781039, Assam, India

ARTICLE INFO

Article history:
Received 6 December 2017
Revised 27 January 2018
Accepted 10 February 2018
Available online 12 February 2018

Keywords:
Isothiocyanyl alanine
Isothiocyanyl lysine
IR sensitive probe
Site specific labeling
Peptide ligation

ABSTRACT

The solvatochromic IR responsivity of small side chain –NCS in two unexplored unnatural amino acids, isothiocyanyl alanine (^{NCS}**Ala = Ita**) and lysine (^{NCS}**Lys = Itl**), without perturbing the conformation is demonstrated in two designed short tripeptide (BocAla-^{NCS}**Ala**-Ala-OMe) and hexapeptide (BocLeu-Val-Phe-Phe-^{NCS}**Lys**-Gly-OMe). Demonstration of site specific fluorescent labeling in both the peptides and ligation type reaction in ^{NCS}**Lys** indicates the novelty of these two amino acids as alternative to the available canonical amino acids.

Side chain modified unnatural amino acids (UNAAs) have attracted current research activity as novel spectral probes of peptides/proteins. For example, fluorescent UNAAs,[1] EPR active UNAAs[2a–c] and photoaffinity labels[2c–e] have been reported in recent years. As a part of our ongoing research we also have contributed to the generation of fluorescent triazolyl UNAAs as peptide labels.[3] However, large sizes of the labels often caused problems of structural perturbation. Therefore, simple and small functional groups as side chain labels would be more beneficial leading to minimal or no structural perturbation of the peptide secondary structure. Intrinsic stretching vibration of diatomic functional groups as chromophores, thus, have attracted much recent interest as Infrared (IR) spectral probes of peptides/proteins.[4] However, choosing an appropriate IR probe for a specific application is a challenging task. In spite of great success of –CN as IR probe,[5a–c] it often suffers from shortcomings such as small dipole strength and interference by Fermi resonance.[5d,e] On the other hand, because of large dipole strength $-N_3$ has got popularity as IR probe.[5a,c,f] However, broad band width, short vibrational lifetime and overlapped stretching with accidental Fermi resonance are its drawbacks.[5g] Furthermore, many of the IR-chromophores, such as, –SCN, has been introduced via derivatization of cysteine which could be a problem when present in multiples in a protein.[6] Therefore, novel side chain possessing stronger IR intensity and does not rely on post synthetic modification would be more attractive.

As a part of our continuous research efforts in the design of unnatural amino acids,[3] the limitations of the currently used IR probes and related literature report[4a] have led us to think –NCS as a new family of possible high sensitive and intensive IR stretch probe. Although, there exists only one example wherein phenylisothiocyanate has been genetically introduced in a protein,[7] there is no report of IR spectroscopic studies in peptide or protein. We thought that the stronger IR transition dipole strength of –NCS compared to reported >C=O and comparable to $-CN/N_3$ group, would show promise as useful IR probes for studying peptide's/protein's microenvironment. Recently, we reported the synthesis of ^{NCS}**Ala** (**Ita**) and its versatile application as a synthetic intermediate for the generation of another class of unnatural amino acid, thioureayl alanines (^{TU}**Ala = Tua**).[8a] In this particular, we want to report the applications of ^{NCS}**Ala** (**1**) and another unexplored unnatural amino acid, ^{NCS}**Lys** (**2**) (Fig. 1) with threefold aims in mind: (a) study of impact of small –NCS group on the peptide's secondary structures, (b) showcasing –NCS as solvatochromic IR-responsive [2100–2270 (med) cm^{-1}] probe and (c) exploiting the electrophilicity of –NCS of ^{NCS}**Ala**/^{NCS}**Lys** in labeling of peptide and peptide ligation. This would be advantageous over the generation of IR sensitive –NCS via derivatisation of a native amino acid in a protein/peptide.[6]

To explore our aims, we first synthesised ^{NCS}**Ala** (**Ita**) following our recently published protocol.[8a] In short, the N, C-diprotected serine **3** was converted to diprotected amino-alanine **6**

* Corresponding author.
E-mail address: ssbag75@iitg.ernet.in (S.S. Bag).

https://doi.org/10.1016/j.bmcl.2018.02.021
0960-894X/

Pyrenylthioureayl Alanine as a Selective Switch-on Fluorescent Sensor for Hg^{2+} and Cu^{2+}

Subhendu Sekhar Bag*[,1,2] and Suranjan De[1]

A new and novel pyrenylthioureayl alanine amino acid (1, ^{PyTU}Ala) is explored for sensing of Hg^{2+} ion with a switch-on fluorescence response. The probe also shows binding interaction with Cu^{2+} ion in absence of other interfering metals with an enhancement of emission. In both the cases 1:2 metal-ligand binding stoichiometry is evident from Job's plot analyses. The detection limits are found to be 93 nM for Hg^{2+} and 318 nM for Cu^{2+} ion. The association constant for Hg^{2+} ion is very high in the order of 10^{10} M^{-1}.

Research in the field of fluorescence sensing of heavy metal as well as transition metal ions is rapidly increasing because of their adverse effects on biological systems and on environment.[1] Among the various heavy metal ions-based pollutants, the mercury ion (Hg^{2+}) is in the frontline to adversely impact the environment and human health.[2] On the other hand, the transition metal ion copper (Cu^{2+}) is important for some physiological processes.[3a-b] However, an excess amount of Cu^{2+} is highly toxic to organisms.[3c-d] Therefore design of fluorescent chemosensors for the sensing of both Hg^{2+} and Cu^{2+} is highly important for the protection of both environment and biological systems. A tremendous number of literatures exists which report the design of various types of chemosensors for the detection of either Cu^{2+} or Hg^{2+} ion.[4] Many efforts have been put forth in recent times to design the fluorescence sensors specific for the sensing of both Hg^{2+} and Cu^{2+} ion detection.[5] Few reports of chemosensors also exist which relies the sensing via change in both color and fluorescence.[6] However, many of them suffer from shortcomings such as fluorescence quenching by paramagnetic Cu^{2+} and Hg^{2+} ions mostly due to electron transfer mechanism[7a-b] and spin–orbit coupling[7c] respectively. Moreover, many organic chemosensors for Hg^{2+} have poor water solubility and suffers from surrounding effects and by the interference of other metal ions with low limit of detection.[8] Therefore, the development of fuorescence switch-on chemosensors for the detection of Cu^{2+} and/or Hg^{2+} is highly desirable.

Figure 1. (a) UV-visible, (b) chemical structure and (c) fluorescence emission spectra of pyrenyl thioureayl amino acid (**1**, ^{PyTU}**Ala**) in various organic solvents (conc. 10 μM).

A fruitfull design of a fluorescence chemosensor involves attachment of a metal binding unit with a fluorophore through a spacer. Binding of a target metal display noticeable change in the photophysical response of the fluorophore [9] Further, the selectivity of a metal sensor is solely dependent on the nature of the coordinating ligands in the receptor unit and the interaction thereof. Therefore, many of the metal sensors containing N/O atoms[9] and few containing S-donor[10] in the metal binding site have been reported. Recently, the potentiality of mixed-donor-S containing ligands are considered for the design of metal sensors. [11] In particular, thiophilic nature of Hg^{2+} and Cu^{2+} attracted very much for the synthesis of sensors containing S as ligating atom. In this respect, in recent time, few thiourea based sensor have been reported as fruitful as ensors for Hg^{2+} and Cu^{2+} ions and other heavy metals.[12] However, newer design of more effective thiourea-based sensor would be interesting and is still in high demand.

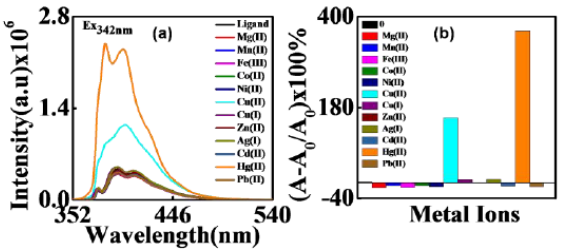


Fig. 2. (a) Fluorescence emission spectra of the sensor **1** (10 μM) in absence and in presence of **1** equivalent (10 uM) concentration of various metal ions. (b) Bar diagram.

Keeping the above idea in mind and our continuing research efforts toward the design of unnatural fluorescent amino acids[13a-d] and sensors for biologically and environmentally important heavy and transition metal ions,[13e-f] we thought that it would be worthwhile to explore fluorescent pyrenylthioureayl alanine amino acid (**1**, ^{PyTU}**Ala**) as a possible sensor for heavy and transition metal ions (Fig. 1). We envisaged that thiourea-moiety would allow strong coordination with Hg^{2+} and/or Cu^{2+} due to their thiophilic nature and hence would generate a detectable fluorescence signal from pyrenyl unit. Therefore, in this particular, we want to report the pyrenylthioureayl alanine amino acid (**1**, ^{PyTU}**Ala**) as a potential sensor for the detection of heavy and transition metal–ions. Thus, we investigated the binding interaction of our newly synthesized thioureayl amino acid for various heavy and transition metal–ions by spectroscopic methods. Among the various metal ions, the probe showed a selective affinity for Hg^{2+} ion in the partial aqueous medium. However, the sensor also showed affinity for Cu^{2+} ion in the absence of interfering metal ions. In both the cases the sensing events