

**Synthesis and Structural Studies of
4(*R/S*)-Aminoproline Polypeptide
and Collagen Mimetics**

THESIS SUBMITTED TO
THE UNIVERSITY OF PUNE

FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN
CHEMISTRY

BY
MAHESH V. SONAR

RESEARCH SUPERVISOR
DR. KRISHNA N. GANESH

**DIVISION OF ORGANIC CHEMISTRY
NATIONAL CHEMICAL LABORATORY
PUNE 411008**

NOVEMBER 2011

CERTIFICATE

This is to certify that the work presented in the thesis entitled "**Synthesis and Structural Studies of 4(*R/S*)-Aminoproline Polypeptide and Collagen Mimetics**" submitted by **Mahesh V. Sonar** was carried out by the candidate at the National Chemical Laboratory, Pune, under my supervision. Such materials as obtained from other sources have been duly acknowledged in the thesis.

Prof. Krishna N. Ganesh, FNA, FNASc

(Research Supervisor)

Director, IISER, Pune

J. C. Bose Fellow

National Chemical Laboratory

Pune-411008

November 2011

Candidate's Declaration

I hereby declare that the thesis entitled "**Synthesis and Structural Studies of 4(*R/S*)-Aminoproline Polypeptide and Collagen Mimetics**" submitted for the degree of Doctor of Philosophy in Chemistry to the University of Pune has not been submitted by me to any other University or Institution. This work was carried out at the National Chemical Laboratory, Pune, India.

Mahesh V. Sonar

November 2011

Research Fellow

Organic Chemistry Division

National Chemical Laboratory

Pune - 411008

*Dedicated to my beloved father
(Aannasaheb) who inspired me*



Acknowledgments

Any human accomplishment is the culmination of numerous contributions and endeavour. The present thesis is no exception. I take this opportunity to thank the special people whose kind support was the reason I could complete the task with confidence. It gives me immense pleasure to express my deep sense of gratitude to my research supervisor Prof. Krishna N. Ganesh for his advice, guidance, support and encouragement during every stage of this work. I do sincerely acknowledge freedom rendered to me by him for independent thinking, planning and executing the research. His endless enthusiasm and receptive attitude will always remain a source of inspiration for me.

I am grateful to Dr. V. A. Kumar, for her valuable support and allowing the laboratory facilities during important stages of this work. I am also grateful to Dr. Moneesha Fernandes, Dr. H. N. Gopi, Dr. Srinivas Hotha, Dr. H. V. Thulasiram, Dr. Arab Mukherjee and Mrs. Anita Gunjal for their help and encouragement during the course of this work. I am also thankful to Mrs. M. V. Mane for HPLC analysis, Mrs. Shantakumari and Dr. Mahesh Kulkarni for mass analysis. My sincere thanks go to Dr. Chandrabhas Narayana (JNCASR, Bangalore) and his research group for helping me in Raman Spectroscopic studies and interpretations.

I would also like to acknowledge Dr. C. V. Rode, Dr. S. P. Chavan and Dr. U. R. Kalkote with their research groups for endless support during my stay at NCL as a project assistant.

Research becomes most enjoyable with wonderful lab mates around. I was fortunate to have a vivacious blend of talented and unique lab mates who can rise to any occasion in their own different way, in addition to being co-operative & helpful. The association of educative seniors like Dr. Umashankara, Dr. Khirud, Dr. Raman (Anna), Dr. Amit Patwa (Sr. bheja fry), Dr. Gitali Devi, Dr. Madhuri, Dr. Gokhale and Dr. Roopa Mitra (rukhsana) helped me soak up the essentials and trivia. I have been fortunate to have friendship with Dr. Ashwani Sharma (indore bhaisab), who stood the test of time and I am grateful to him for always encouraging me in whatever I choose to do. I never forgot the days spend him and his wife Deepti and their little angle Sia. There were times being funny and playful with labmates like Shreedhar (gunturgade), Manaswini (MNS) and Pradnya.

I enjoyed working with juniors like Tanpreet, Deepak (bheja fry 2), Nitin (hard eater), Vijay (VDC), Satish (bodyguard) and Madan. Lab # 209 will be an unforgettable place in my research life. I experienced the positive “neighboring group participation” from Seema, Namrata, Kiran, Venu, Manoj, Anjan and Tanaya.

I was always looking forward to lunch at *katta*, thanks to the cheerful company of the *katta* members over the period of time. Friends make the think go easy and life beautiful. My special thanks go to friends like Sachin (wani), Dhananjay (salman), Umesh, Prasad, Lalit and Pramod for their cheerful atmosphere in friend circle. I never forgot the days spend with room partners Panjab and Ravi. I would like extend my thanks to Aaba, Sharad, Kishor, Roshan, Alok, Nagesh, Nilesh, Prakash, Kailash, Rajendra, Swaroop, Kapil and Mudit.

It is impossible to express my sense of gratitude for my parents Anna and Aai for their wonderful gifts of teaching, understanding and a perfect upbringing to launch my flight of thoughts to fulfill my aspiration. It is my parent’s prayer, constant struggle and relentless hard work to overcome the odds of life, which has inspired me to pursue life with a greater optimism. No word would suffice to express my gratitude and love to my sisters, Akka, Mai, Lalita tai, Shobha tai, Bharati tai. I would like to extend my gratitude to my uncle, aunty, cousins, brother-in-laws, father and mother-in-laws and my all nephews.

Words fail me to express my appreciation to my wife Yamini whose dedication, love and persistent confidence in me, has taken the load off my shoulder. I am also grateful to my wife for the unconditional support and encouragement in all my effort and being always there for me. My success now and always will be dedicated to my family. I would like to express my love for my daughter (nanhi pari) Kanishka (saii). I never forgot her childhood activities which make me relax after hectic schedule. I would like to express my apologize towards my daughter, that I am unable to gave sufficient time during her childhood.

Mahesh V. Sonar

Contents

Abbreviations	I
Abstract	IV

Chapter 1

Synthesis and Conformational Studies of 4(*R/S*)-Aminoproline Polypeptides

1	Introduction	2
1.1	Conformational Analysis of Polypeptides	2
1.1.1	Ramachandran plot	3
1.2	Protein Secondary Structures: the α Helix and the β Sheet	4
1.2.1	The α helix	4
1.2.2	The β sheet	5
1.3	Other Secondary Structures: 3_{10} helix, β hairpin and turns	6
1.3.1	The 3_{10} helix	6
1.3.2	The β -hairpin	7
1.3.3	Turns	7
1.4	Polyproline Conformation	8
1.4.1	Peptidyl prolyl <i>cis-trans</i> isomerization	8
1.5	Types of Polyproline Conformation	10
1.5.1	A survey of left-handed polyproline II helices	11
1.5.2	Side-chain rotamer preference	12
1.5.3	Conformational preference and <i>cis-trans</i> isomerization of 4(<i>R/S</i>)-substituted proline	13
1.5.4	Conformational preferences of β - and γ -aminated proline analogues	14
1.6	Factors Affecting Polyproline Conformation	15
1.6.1	Stereoelectronic effect on polyproline conformation	15
1.6.2	Azidoproline as conformational directing element and functionalizable site	16
1.6.3	Effect of <i>O</i> -Galactosylation of 4 <i>R</i> -Hydroxyproline on PPII conformation	17
1.6.4	Effects of terminal functional groups on the stability of the PPII structure	17
1.6.5	Temperature induced transition between polyproline II and I helices	18

1.6.6	Effect of H ₂ O and D ₂ O on polyproline II conformation	19
1.6.7	Effect of urea on polyproline II conformation	19
1.6.8	Effect of salt on polyproline structure	20
1.6.9	Effect of aliphatic alcohol on polyproline conformation	21
1.7	Biological Significance of Polyproline Conformation	21
1.7.1	Oligoproline as cell penetrating agents	22
1.7.2	Polyproline conformation in elastic function	23
1.7.3	Self-assembly processes and PPII conformation	25
1.7.4	Biophysical reasons for why proline is a common binding motif	25
1.7.5	Polyproline peptidomimetics	26
2	Methodology	27
2.1	Peptide Bond Formation: Methods and Strategies	27
2.1.1	Racemisation suppressants	30
2.1.2	Solution phase peptides synthesis	30
2.1.3	Solid phase peptide synthesis (SPPS)	31
2.2	The Choice of Solid Support (Resin)	34
2.3	Characterization of Polyproline Structures	36
2.3.1	Polyproline II conformation by ¹⁵ N NMR spectroscopy	36
2.3.2	Circular dichroism	37
2.3.3	Polyproline II conformation by VCD	39
2.3.4	Polyproline II conformation by ROA	39
3	Present work: Rationale	40
3.1	Results	41
3.1.1	Synthesis of fully protected (2 <i>S</i> ,4 <i>R</i>)-and (2 <i>S</i> ,4 <i>S</i>)-aminoproline monomers	41
3.1.1a	Synthesis of (2 <i>S</i> ,4 <i>R</i>)-N ¹ -(Fmoc)-4NH-(<i>t</i> -Boc)-aminoproline (8)	41
3.1.1b	Synthesis of (2 <i>S</i> ,4 <i>S</i>)-N ¹ -(Fmoc)-4NH-(<i>t</i> -Boc)-aminoproline (12)	43
3.1.1c	Synthesis of homo oligopeptides 1-3 by using solid phase protocol	44
3.1.2	Determination of the peptide concentration in stock solution	46
3.1.3	p <i>K</i> _a of 4-amino group of 4 <i>R/S</i> -Aminoproline monomers and oligomers	47
3.2	CD Spectroscopic Studies	48
3.2.1	Identification of PPII by UV-CD spectroscopy	49
3.2.2	Concentration dependent CD spectroscopy for peptides 1-3	49
3.2.3	Effect of protonation of the 4-amino group on	51

	PPII helical content	
3.2.4	CD dependent thermal denaturation of peptides 1-3 at different pHs	54
3.2.5	Effect of urea on peptides 1-3	58
3.2.6	Effect of salt (NaCl) on peptides 1-3	61
3.2.7	Effect of trifluoroethanol on peptides 1-3	64
3.2.8	Identification of β -structure by Raman Spectroscopy	66
3.2.9	Water-induced switching of β -structure to PPII conformation in the 4S- <i>amp</i> ₉	68
3.2.10	Effect of <i>n</i> -propanol on peptides 1-3	69
3.3	Discussion	72
3.4	Conclusion	78
3.5	Experimental	80
3.6	Reference	94
3.7	Appendix 1	100

Chapter 2

Collagen Host-Guest Peptides: Position Dependent Substitution of 4(*R/S*)-Aminoproline Triplets

1	Introduction	120
1.1	Collagen Superfamily	122
1.1.1	Three dimensional structure and crystal packing of collagen	124
1.2	Terminology Used in Higher Structure of Collagen	125
1.2.1	Nucleation and modulation of collagen fibrillogenesis	126
1.2.2	Biosynthesis of collagen	127
1.3	Factors Affecting Collagen Triple Helix Structure	129
1.3.1	Pyrrolidine ring conformation in proline and substituted proline	129
1.3.2	<i>Gauche</i> effect on the ring-pucker preferences	130
1.3.3	Positional preferences of amino acids in collagen	132
1.3.4	Interstrand hydrogen bonds	135
1.3.5	$n \rightarrow \pi^*$ Interaction	136
1.3.6	<i>Cis-trans</i> isomerization	137
1.3.7	Glycine substitution in model peptides and its correlation with disease	138
1.3.8	Electrostatic interactions in the triple-helix	138
1.3.9	Effect of charged termini on the stability of triple-helices	139

1.3.10	Effect of ionizable side-chains	140
1.3.11	Effect of ethylene glycol on the stability of triple-helices	140
1.4	Collagen Mimetics	141
1.4.1	Collagen mimetics with unnatural amino acids	143
1.5	Collagen in Aging and Disease	144
1.6	Biomedical Applications of Collagen	146
1.6.1	Collagen-based drug delivery systems	148
1.6.2	Collagen-based systems for tissue engineering	150
1.6.3	Collagen cross-links	152
1.7	Characterization of Triple Helical Structures	153
1.7.1	X-Ray crystallography of triple helical peptides	153
1.7.2	NMR spectroscopy	154
1.7.3	Mass spectroscopy	158
1.7.4	SDS-PAGE	158
1.7.5	Characterization of triple-helical structures by circular dichroism	158
1.7.6	Differential scanning calorimetry	160
1.7.7	Surface morphology	161
2	Present work: Rationale	163
3	Results	165
3.1	Solution <i>versus</i> solid phase peptide synthesis – block trimer approach	165
3.1.1	Synthesis of collagen model [Pro-Hyp-Gly] tripeptide	166
3.1.2	Synthesis of glycine benzyl ester	167
3.1.3	Synthesis of [Pro-4 <i>R</i> -Amp-Gly] tripeptide	167
3.1.4	Synthesis of [Pro-4 <i>S</i> -amp-Gly] tripeptide	171
3.1.5	Synthesis of [4 <i>S</i> -amp-4 <i>R</i> -Amp-Gly] Tripeptide	173
3.1.6	Solid phase peptide synthesis	175
3.2	Characterization of Peptides by CD-Spectroscopy	179
3.3	CD Spectroscopic Studies of N- and C-Capped Peptides 4-12	180
3.3.1	CD spectroscopic studies of C-terminus modified peptides 4-6	180
3.3.2	CD spectroscopic studies of middle modified peptides 7-9	186
3.3.3	CD spectroscopic studies of N-terminal modified peptides 10-12	191
3.4	CD Spectroscopic Studies for Non-N-Acetylated	197

	Peptides 13-21	
3.4.1	CD spectroscopic studies of C-terminus modified peptides 12-15	198
3.4.2	CD spectroscopic studies of middle modified peptides 16-18	202
3.4.3	CD spectroscopic studies of N-terminal modified peptides 19-21	201
4	Discussion	213
5	Conclusion	218
6	Experimental	219
7	References	243
8	Appendix	252

Chapter 3

Collagen peptide (Pro-Hyp-Gly)_n analogues with α -aminoisobutyric acid substituted glycine: (Pro-Hyp-Aib)_n

1	Introduction	295
1.1	Substitution of Glycine	295
1.2	Importance of α -Aminoisobutyric Acid (Aib)	296
1.3	Present work: Rationale	298
1.3.1	Synthesis and characterization of peptides 30-33	298
1.3.2	Solid phase peptide synthesis	299
1.4	Conformational Study by CD Spectroscopy	301
1.4.1	CD spectra of N-capped peptides 30-31	301
1.4.2	CD spectra of uncapped peptides 30-31	301
1.4.3	Concentration dependent CD spectroscopy for peptides 30-33	302
1.4.4	CD thermal denaturation study of peptides 30-33 in buffer at pH 7.2	305
1.4.5	Effect of ethylene glycol on the stability of triple-helices	310
1.4.6	CD spectroscopy study of peptides 30-33 in buffer: ethylene glycol (1:1)	310
1.4.7	CD spectroscopy study of peptides 30-33 in ethylene glycol	315
1.5	Discussion	320
1.6	Conclusion	322
1.7	Experimental	323
1.8	References	328
1.9	Appendix	331

Abbreviations & Symbols

Ac	Acetyl
Ac ₂ O	Acetic anhydride
Aib	α-Aminoisobutyric acid
Ala	Alanine
Amp	(2 <i>S</i> ,4 <i>R</i>)-Aminoproline
amp	(2 <i>S</i> ,4 <i>S</i>)-aminoproline
Arg	Arginine
Asn	Asparagine
Azp	(2 <i>S</i> ,4 <i>R</i>)-Azidoproline
azp	(2 <i>S</i> ,4 <i>S</i>)-azidoproline
(Boc) ₂ O	Boc anhydride
°C	Degree Celsius
Cbz	Benzyloxycarbonyl
CBMIT	1,1'-Carbonylbis(3-Methylimidazolium), Triflate
CD	Circular Dichroism
CDI	1,1'-Carbonyldiimidazole
CMPs	Collagen mimetic peptides
CMBI	2-chloro-1,3-dimethyl-1 <i>H</i> -benzimidazolium hexafluorophosphate
COSY	Correlation Spectroscopy
CRPs	Collagen Related Peptides
Cys	Cystine
D-	Dextro-
DCC	<i>N,N'</i> -Dicyclohexylcarbodiimide
DCM	Dichloromethane
DSC	Differential scanning calorimetry
Δ <i>G</i>	Change in Gibb's free energy
Δ <i>H</i>	Change in enthalpy
Δ <i>T</i> _m	Difference in melting temperature
DIAD	Diisopropyl azodicarboxylate
DIPEA	<i>N,N</i> - Diisopropylethylamine
DMF	<i>N,N</i> -Dimethylformamide
DMS	Dimethyl sulfide
DNA	Deoxyribonucleic acid
<i>E</i>	<i>entgegen</i>
ECM	Extracellular matrix
EDC	1-ethyl-3-(3-dimethylaminopropyl) carbodiimide
EG	Ethylene glycol
eq.	equivalents
ESI-MS	Electro Spray Ionization Mass Spectrometry
Et	Ethyl
FACIT	Fibril associated with interrupted triple-helices

Flp	(2 <i>S</i> ,4 <i>R</i>)-Fluoroproline
flp	(2 <i>S</i> ,4 <i>S</i>)-fluoroproline
Fmoc	9-Fluorenylmethoxycarbonyl
g	Gram
Glu	Glutamine
Gly	Glycine
h	Hours
HATU	O-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HBTU	(2-(1 <i>H</i> -Benzotriazole-1-yl)-1,1,3,3-tetramethyl-uronium-hexafluorophosphate)
HOAt	1-Hydroxy-7-azabenzotriazole
HOBt	N- Hydroxybenzotriazole
HPLC	High Performance Liquid Chromatography
Hyp	(2 <i>S</i> ,4 <i>R</i>)-Hydroxyproline (<i>trans</i> -4-hydroxy-L-proline)
hyp	(2 <i>S</i> ,4 <i>S</i>)-hydroxyproline (<i>cis</i> -4-hydroxy-L-proline)
Hz	Hertz
<i>in situ</i>	In the reaction mixture
<i>in vivo</i>	Within the living organism
<i>in vitro</i>	outside the living organism
IR	Infra Red
K	Kelvin/Kilo/Binding Constant
kJ	kilojoules
$K_{trans/cis}$	Ratio of <i>trans/cis</i>
L-	Levo-
LCMS	Liquid Chromatography-Mass Spectrometry
Leu	Leucine
Lys	Lysine
M	Molar
MACIT	Membrane associated with interrupted triple-helices
MALDI-TOF	Matrix Assisted Laser Desorption Ionization-Time of Flight
MBHA	p-methoxybenzhydrylamine
MECN	Acetonitrile
MeOH	Methanol
Mep	(2 <i>S</i> ,4 <i>R</i>)-Methylproline
mep	(2 <i>S</i> ,4 <i>S</i>)-methylproline
Mg	milligram
mL	millilitre
MHz	megahertz
min	minutes
μ	Micron
μl	Microliter
μM	Micromolar
mL	Microliter
mmol	millimolar
MS	Mass spectrometry

MsCl	methanesulfonylchloride
Nd:YAG	Neodymium-doped yttrium aluminium garnet; Nd:Y ₃ Al ₅ O ₁₂
ν	nu (frequency)
NaN ₃	Sodium azide
Nm	Nanometer
NMP	N-Methyl-2-pyrrolidone
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Magnetic Overhauser Effect
Pd-C	Palladium on carbon
Phe	Phenylalanine
Pro	L-Proline
PNA	Peptide Nucleic Acid
ppm	parts per million
PPII	Polyproline II
PPI	Polyproline I
psi	pound per square inch
PyBOP	Benzotriazol-1-yloxytri(pyrrolidino)-phosphonium-hexafluorophosphate
<i>R</i>	Rectus
ROA	Raman optical activity
$R_{p/n}$	Ratio of positive to negative band intensities
RP	Reverse phase
RNA	Ribonucleic acid
<i>S</i>	Sinister
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
Ser	Serine
SPPS	Solid Phase Peptide Synthesis
τ	Tau
TBTU	O-(benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
<i>t</i> -Boc	<i>tert</i> -butyloxycarbonyl
TFA	Trifluoroacetic acid
TFMSA	Trifluoromethane sulfonic acid
TFE	2,2,2-Trifluoroethanol
THF	Tetrahydrofuran
TIS	Triisopropylsilane
TMS	Tetramethylsilane
T_m	Melting temperature
Trp	Tryptophan
TSTU	2-succinimido-1,1,3,3-tetramethyluronium tetrafluoroborate
Tyr	Tyrosine
Val	Valine
VCD	Vibrational Circular Dichroism
v/v	volume to volume ratio
Z	zusammen

Abstract

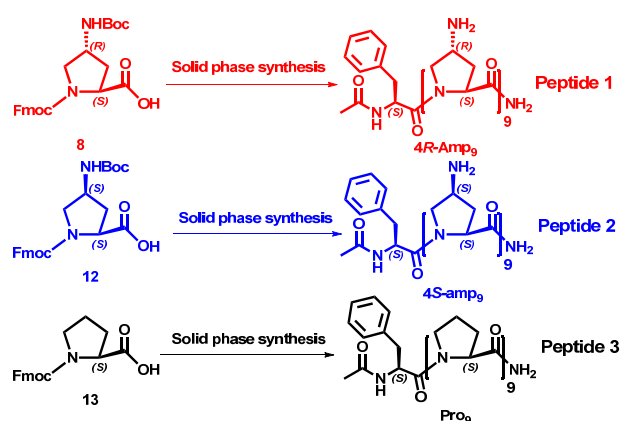
The thesis entitled “**Synthesis and Structural Studies of 4(*R/S*)-Aminoproline Polypeptide and Collagen Mimetics**” comprises of studies towards the design and synthesis of peptides based on 4(*R/S*)-amino substituted prolines in polyproline and collagen peptide sequences and study of their behavior under variable physical environments. This may have importance in understanding their biological functions. The work describes different behavior of 4-amino polyproline peptides under different pH, solute and solvent conditions. The investigation of β -structure in 4*S*-amp₉ may gives new directions of future research work in protein engineering. The substitution of 4-*R/S* aminoproline in collagen hybrid and collagen host guest peptides show stereo and position dependent effects on triple helix forming ability at different pHs and solvent conditions. The replacement of α -aminoisobutyric acid for glycine in collagen peptides decreases the triple helix stability of the derived peptides. The thesis is divided in three chapters.

Chapter 1: Synthesis and conformational studies of 4(*R/S*)-aminoproline polypeptides

The polyproline type conformations PPI and PPII have been recognized for their presence in both folded and unfolded protein structures.¹ Recent studies of polyproline model peptides having non-hydrogen bonding substituents such as 4(*R/S*)-F/ 4(*R/S*)-N₃ have shown that these peptides form polyproline single helices which are more stable than helix from their corresponding unsubstituted proline oligomer. In view of these findings, it was proposed to study the effect of 4*R/S*-amino group, which unlike fluorine or azide can participate in intramolecular and intermolecular hydrogen bonding. The 4-NH₂ group may also provide a pH-dependent switch for polyproline stability.

The 4*R*-aminoproline (**8**) and *S*-aminoproline (**12**) monomers were synthesized by reported literature procedure.² These were incorporated into the homo oligopeptides **1** (4*R*-Amp₉) and **2** (4*S*-amp₉) by solid phase synthesis on Rink amide resin, using Fmoc chemistry with *t*-Boc protection for the 4-amino function on proline (Scheme 1). The readily available N-Fmoc proline (**13**) was used for the synthesis of control proline oligomer **3** (Pro₉).

Scheme 1



The conformational studies of oligopeptides were carried out by CD spectral analysis as a function of pH, temperature, different solutes (urea and sodium chloride) and in solvents (trifluoroethanol (TFE) and n-propanol) as aliphatic alcohol.

The pH dependent studies show that the peptide **1** (4R-Amp₉) has highest polyproline content in all pHs. The peptide **1** (4R-Amp₉) shows higher PPII helical content at acidic pH (4.0-5.0) which decreases as the pH increases upto neutral pH and it does not change in alkaline pH (9.0-10.0). In case of peptide **2** (4S-amp₉), the PPII helicity at acidic pH (4.0-5.0) was low but enhanced at alkaline pH (9.0-10.0). The PPII helicity of peptide **3** (Pro₉) remained constant with pH.

The CD thermal denaturation study was carried out for peptides (**1-3**) at different pHs and the T_m of these peptides are tabulated in Table 1.

Table 1

pH	Peptide 1 (4R-Amp ₉)	Peptide 2 (4S-amp ₉)	Peptide 3 (Pro ₉)
4.0	52	27	37
5.0	51	30	38
7.2	50	41	38
9.0	50.5	47	38
10.0	50.5	48	38

T_m values are (± 0.5 °C)

T_m values indicate that peptides **1** (4R-Amp₉) and **3** (Pro₉) were almost invariant in their thermal stability as a function of pHs. The peptide **2** (4S-amp₉) showed increase in thermal stability with increase in pH. In general as compared to control proline peptide **3** (Pro₉), the peptide **1** (4R-Amp₉) stabilizes PPII helix in both acidic as well as basic pH, while the peptide **2** (4S-amp₉) destabilized PPII helix at acidic pH and stabilizes PPII helix at basic pH.

It is well known that urea promotes PPII helicity while salt (NaCl) disrupts the PPII helix. The effects of these solutes on the conformation of peptides (**1-3**) were examined by CD spectroscopy. It was found that the PPII helicity of peptides **1** (*4R-Amp₉*) and **3** (*Pro₉*) increases smoothly with increase in concentration of urea, while it enhanced rapidly in case of peptide **2** (*4S-amp₉*). The addition of salt slightly decreases the PPII helicity for peptides **1** (*4R-Amp₉*) and **3** (*Pro₉*), while drastic decrease was seen in case of peptide **2** (*4S-amp₉*).

The hydrophobic solvents like trifluoroethanol and n-propanol affect the conformational properties of peptide and hence the effect of these solvents on peptides (**1-3**) was investigated by CD spectroscopic study. Peptides **1** (*4R-Amp₉*) and **3** (*Pro₉*) retained PPII conformation in trifluoroethanol and n-propanol. In these solvents, very interestingly β -structure formed in case of peptide **2** (*4S-amp₉*) which transformed to PPII by addition of tiny amount of water (Figure 1).³ The presence of β -structure in TFE for peptide **2** (*4S-amp₉*) was supported by Raman spectroscopy. The β -structure in polyproline peptides is hitherto unknown and the present results will add a new design principle to a growing repertoire of strategies for engineering peptide secondary structural motifs for new biomaterials. The structural conversion illustrates a fine balance between stereoelectronic and H-bonding effects in novel tuning of the secondary structure of *4R/S*-aminoproline polypeptides is discussed in this chapter.

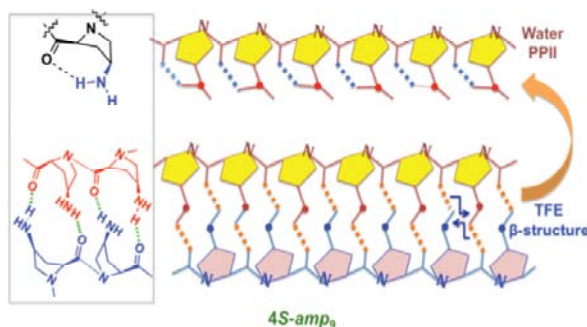


Figure 1: β -structure formed in case of peptide **2** (*4S-amp₉*) which is transformed to PPII by addition of tiny amount of water

Chapter 2: Collagen host-guest peptides: position dependent substitution of 4(*R/S*)-aminoproline triplets

Collagen is an ancient structural protein found on earth till now,⁴ present in connective tissues of higher organisms. Each strand of collagen consists of repeating tripeptide motif X-Y-Gly where X and Y are proline (Pro) and *4R*-hydroxyproline (Hyp) respectively (Figure 2). The collagen triple-helix consists of

three parallel left handed polyproline II-like helices, supercoiled around each other in a right-handed manner.

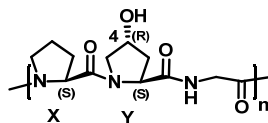


Figure 2: Collagen peptide sequence

It was previously reported from literature² that 4*R*-amino substitution on proline at Y position in collagen peptide stabilizes the triple helix at both acidic and basic pHs. In order to understand the molecular origin, hybrid collagen and host-guest model peptides (Figure 3) were synthesized to probe the position dependent effects.

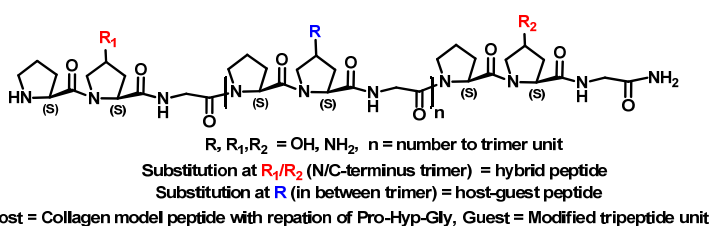
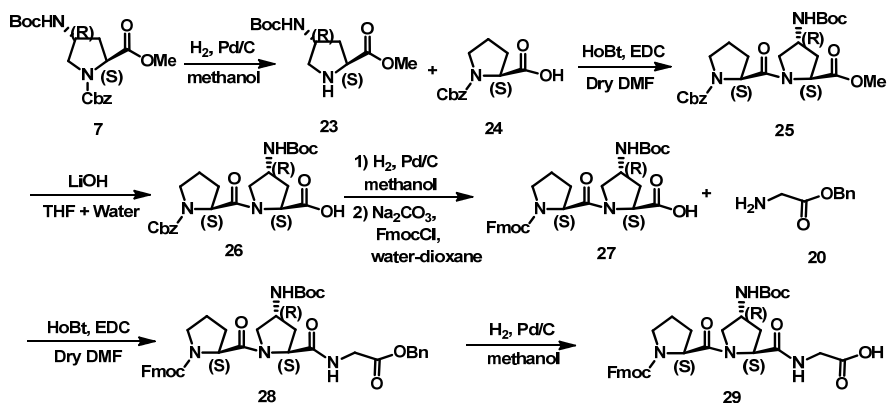


Figure 3: Collagen hybrid and host-guest peptide

The tripeptides were synthesized by solution phase approach. The synthesis of tripeptide **29** is shown in Scheme 2.

Scheme 2



Similar type of synthetic approach was used for synthesis of other tripeptides (**22**, **35** and **41**).

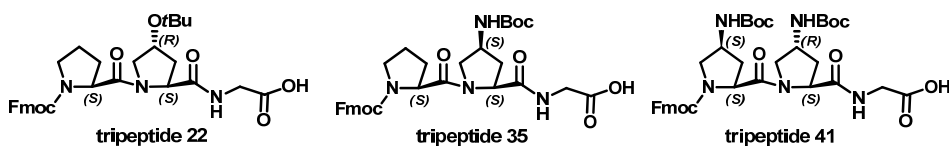


Figure 4: Collagen model tripeptides

The trimers were used for synthesis of oligomers on Rink amide resin by using

standard solid phase peptide synthesis protocol. The N-capped acetylated and non-acetylated oligomers were synthesized with various combinations of the tripeptide. The CD thermal denaturation study were carried out for peptides (4-21) at different pHs, in ethylene glycol and the T_m of these peptides are tabulated for N-capped acetylated in Table 2 and for non-acetylated in Table 3.

Table 2: CD T_m of N-capped acetylated peptides (4-12)

Peptide	pH 4.0	pH 7.2	pH 10.0	EG
Ac-Phe-[Pro-Hyp-Gly] ₅ -Pro-4R-Amp-Gly-NH ₂ (Peptide 4)	46	41	40	41
Ac-Phe-[Pro-Hyp-Gly] ₅ -Pro-4S-amp-Gly-NH ₂ (Peptide 5)	19	23	25	43
Ac-Phe-[Pro-Hyp-Gly] ₅ -4S-amp-4R-Amp-Gly-NH ₂ (Peptide 6)	41	37	39	40
Ac-Phe-[Pro-Hyp-Gly] ₂ -Pro-4R-Amp-Gly-[Pro-Hyp-Gly] ₃ -NH ₂ (Peptide 7)	45	40	39	42
Ac-Phe-[Pro-Hyp-Gly] ₂ -Pro-4S-amp-Gly-[Pro-Hyp-Gly] ₃ -NH ₂ (Peptide 8)	19	24	26	41
Ac-Phe-[Pro-Hyp-Gly] ₂ -4S-amp-4R-Amp-Gly-[Pro-Hyp-Gly] ₃ -NH ₂ (Peptide 9)	42	39	40	39
Ac-Phe-Pro-4R-Amp-Gly-[Pro-Hyp-Gly] ₅ -NH ₂ (Peptide 10)	44	39	38	43
Ac-Phe-Pro-4S-amp-Gly-[Pro-Hyp-Gly] ₅ -NH ₂ (Peptide 11)	20	25	26.5	40
Ac-Phe-4S-amp-4R-Amp-Gly-[Pro-Hyp-Gly] ₅ -NH ₂ (Peptide 12)	43	40	41	38

T_m values are (\pm 0.5 °C)

Table 3: CD T_m of non acetylated peptides (13-21)

Peptide	pH 4.0	pH 7.2	pH 10.0	EG
H ₂ N-Phe-[Pro-Hyp-Gly] ₅ -Pro-4R-Amp-Gly-NH ₂ (Peptide 13)	42	38	39	37
H ₂ N-Phe-[Pro-Hyp-Gly] ₅ -Pro-4S-amp-Gly-NH ₂ (Peptide 14)	31	34	38	37
H ₂ N-Phe-[Pro-Hyp-Gly] ₅ -4S-amp-4R-Amp-Gly-NH ₂ (Peptide 15)	41	37	34	38
H ₂ N-Phe-[Pro-Hyp-Gly] ₂ -Pro-4R-Amp-Gly-[Pro-Hyp-Gly] ₃ -NH ₂ (Peptide 16)	41	37	38	38
H ₂ N-Phe-[Pro-Hyp-Gly] ₂ -Pro-4S-amp-Gly-[Pro-Hyp-Gly] ₃ -NH ₂ (Peptide 17)	32	35	38	36
H ₂ N-Phe-[Pro-Hyp-Gly] ₂ -4S-amp-4R-Amp-Gly-[Pro-Hyp-Gly] ₃ -NH ₂ (Peptide 18)	42	38	35	37
H ₂ N-Phe-Pro-4R-Amp-Gly-[Pro-Hyp-Gly] ₅ -NH ₂ (Peptide 19)	40	36	37	39
H ₂ N-Phe-Pro-4S-amp-Gly-[Pro-Hyp-Gly] ₅ -NH ₂ (Peptide 20)	33	36	39	35
H ₂ N-Phe-4S-amp-4R-Amp-Gly-[Pro-Hyp-Gly] ₅ -NH ₂ (Peptide 21)	43	39	36	36

T_m values are (\pm 0.5 °C)

From the results, it is seen that the replacement of even one 4-OH prolyl residue by 4R/S-NH₂ has significant in stabilizing or destabilizing the collagen triple helix structure.

Chapter 3: Collagen peptide (Pro-Hyp-Gly)_n analogues with α -aminoisobutyric acid substituted glycine: (Pro-Hyp-Aib)_n

It is well known that α,α -dimethyl glycine (Aib) induces helicity⁵ in peptides. Hence it was proposed to study the effect of rigidification of α -carbon of glycine on collagen triple helix. The peptides (30-33) were synthesized by using linear coupling approach for studying the effect of 2-aminoisobutyric acid on collagen triple helix.

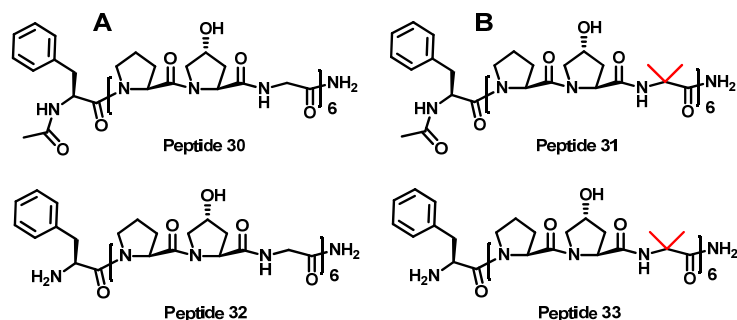


Figure 5: (A) Collagen model peptide with glycine substitution (B) Collagen model peptide with α -aminoisobutyric acid (Aib) substitution.

The CD thermal denaturation study was carried out for peptides (**33-33**) in buffer (pH 7.2), in ethylene glycol and the T_m of these peptides are tabulated in Table 4.

Table 4: T_m values of peptides **30-33**.

Peptide	Buffer (pH 7.2)	EG:water (1:1)	Ethylene glycol
AcPhe(Pro-Hyp-Gly) ₆ (Peptide 30)	27	32.5	44
AcPhe(Pro-Hyp-Aib) ₆ (Peptide 31)	nd	27	31
H ₂ N-Phe(Pro-Hyp-Gly) ₆ (Peptide 32)	31.5	26	31.5
H ₂ N-Phe(Pro-Hyp-Aib) ₆ (Peptide 33)	nd	19	22

nd = triplex not detected, T_m values are (± 0.5 °C)

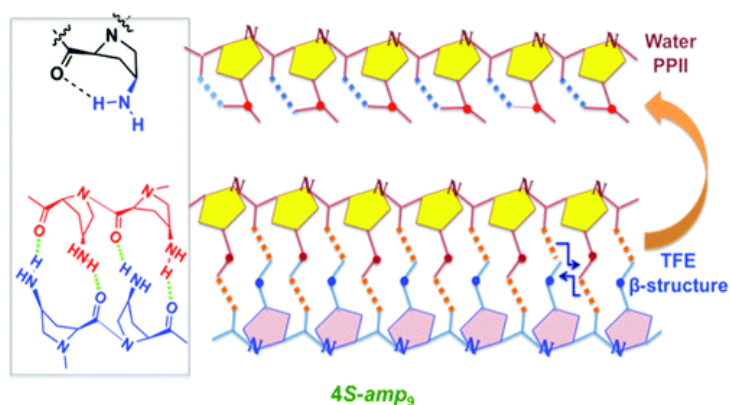
It was found that 2-aminoisobutyric acid (Aib) destabilizes the collagen triple helix in aqueous buffer. Interestingly Aib substituted peptides (**31**, **33**) show weak triple helix formation in ethylene glycol and ethylene glycol-water system.

References

- Cowan, P. M.; McGavin, S. *Nature* **1955**, *176*, 501–503. (b) Traub, W.; Shmueli, U. *Nature* **1963**, *198*, 1165–1166.
- (a) Babu, I. R.; Ganesh, K. N. *J. Am. Chem. Soc.* **2001**, *123*, 2079–2080. (b) Umashankara, M.; Babu, I. R., Ganesh, K. N. *Chem. Comm.* **2003**, 2606–2607.
- Sonar, M. V.; Ganesh, K. N. *Org. Lett.* **2010**, *12*, 5390–5393.
- Schweitzer, M. H. et. al *Science* **2007**, *316*, 277–280.
- Balaram, P., et. al. *J. Am. Chem. Soc.* **2006**, *128*, 7916–7928.

Chapter 1

Synthesis and Conformational Studies of 4(*R/S*)-Aminoproline Polypeptides



This chapter gives an introduction to polyproline conformation in important biopolymers, factors affecting PPII conformation and their importance in biological processes. The focus is on the role of C4-substitution on proline and its effect on prolyl-peptide bond isomerization. The effect of stereochemistry, pH, solute and solvent effects on polyproline conformation is discussed. The application of circular dichroism and Raman spectroscopy was used for identification of β -structure. The overview on the background literature is followed by rationale for undertaking the present research work, which indicate that 4(*S*)-aminoproline polypeptides exhibit unseen β -structure in hydrophobic medium (TFE) in contrast to PPII helical form in aqueous medium.

1 Introduction

Nature is an excellent source of inspiration in the design of stimulus responsive materials. Different type of mechanisms found in proteins, in which a response to environmental changes leads to conformational transitions, have been extensively studied. Proteins themselves may not be the most likely structures to prepare materials due to their fragile nature and high cost of production. However, peptides are more synthetically accessible and have been evaluated as building blocks for stimuli responsive materials.¹ Often a change in the peptide properties is caused by a conformational transition, leading to e.g. a spatial redistribution of hydrophobic and hydrophilic residues. Such induced changes can be explored to control the self-assembly of peptides. These peptides are potentially useful in the field of tissue engineering, drug delivery, biosensing and so-called smart biomaterials, i.e. materials with defined and controlled properties. The stimuli that can induce the desired changes are e.g. pH, temperature, metal ions, light, enzymes and ion concentrations. Due to the diversity in properties of their constituent amino acids, peptides are easily designed to behave as stimulus dependent materials. Moreover, by the incorporation of non-natural amino acids or other small organic moieties the scope can be extended even further.

1.1 Conformational analysis of polypeptides

Proteins are long chain polypeptides that adopt many different conformations; yet, the sequence of their amino acid residues directs the folding to a particular native state conformation.² Polypeptide conformations can be described in terms of 3 main chain torsion angles; (a) the torsion angle about C_{α} -N σ -bond ϕ (*phi*), (b) the angle about the σ -bond between carbonyl group and C_{α} ψ (*psi*), and (c) the angle about the amide bond ω (Figure 1). The σ -bonds (except in the imino acid proline) are relatively flexible and the preferred values for ϕ and ψ angles depend primarily on the nature of the α -substituent. Allowed values for ϕ and ψ are graphically revealed when ϕ is plotted *versus* ψ in a Ramachandran plot, introduced by G. N. Ramachandran.³ The torsional angles of each residue in a peptide define the geometry of its attachment to its two adjacent residues by positioning its planar peptide bond relative to the two adjacent planar peptide bonds. Thus the torsional angles determine the conformation of the residues and the peptide as a whole.

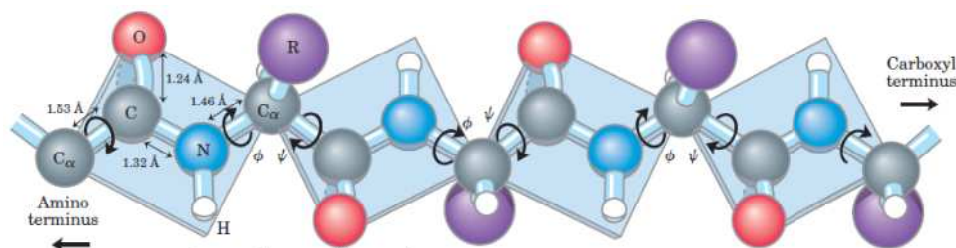


Figure 1: Illustration of peptide plane and the torsion angles⁴

1.1.1 Ramachandran plot

The sterically allowed values of ϕ and ψ can be determined by calculating the distances between the atoms of a tripeptide at all values of ϕ and ψ for the central peptide unit. Sterically forbidden conformations, such as the one shown in Figure 2A, are those in which any nonbonding interatomic distance is less than its corresponding Van der Waals distance. Such information is summarized in a Ramachandran plot (Figure 2B), named after its inventor, G. N. Ramachandran.³

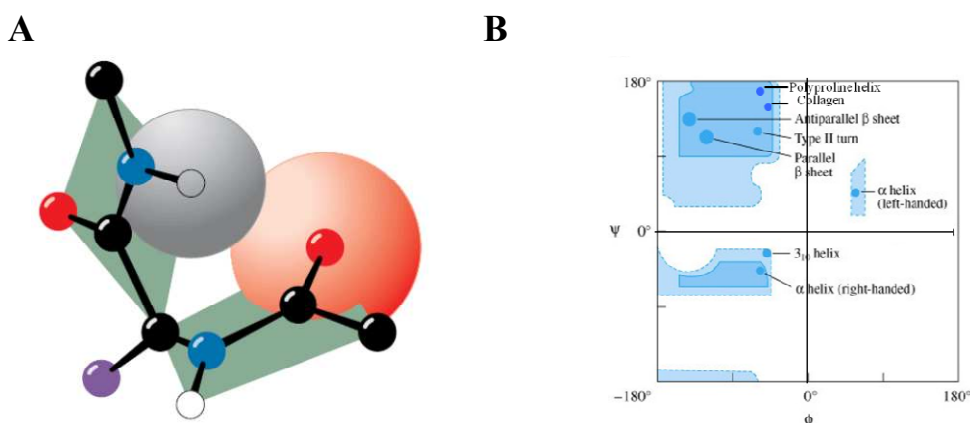


Figure 2: (A) Steric interference between adjacent residues, the collision between carbonyl oxygen and the following amide hydrogen prevents the conformation $\phi = -60^\circ$, $\psi = -30^\circ$. (B) Ramachandran plot for variety of peptide structures.⁵

Most areas of the Ramachandran diagram (most combinations of ϕ and ψ) represent forbidden conformations of a polypeptide chain. Only three small regions of the diagram are physically accessible to most residues. The observed ϕ and ψ values of accurately determined structures nearly always fall within these allowed regions of the Ramachandran plot. There are, however, some notable exceptions:

- The cyclic side chain of proline limits its range of ϕ values to angles of around -60° , making it the most conformationally restricted amino acid residue.

- Gly, the only residue without a C_{β} atom, is much less sterically hindered than the other amino acid residues. Hence, its permissible range of ϕ and β covers a larger area of the Ramachandran diagram. At glycine residues, polypeptide chains often assume conformations that are forbidden to other residues.

The three dimensional structures of proteins are of great help to understand the precise details of its biological function. The basic properties of peptides that ultimately also define their structure are determined by the primary amino acid sequence. The residues provide diversity *via* non-covalent and covalent interactions including hydrophobic interactions, aromatic stacking, hydrogen bonding, disulfide bridges and electrostatic interactions. The interactions of the individual amino acids are weak, but as an ensemble they can give rise to a stable secondary structure. The most common secondary structures are α -helix and β -sheets, though other structures such as the 3_{10} helix, polyproline helix and the π -helix also exist. The type of secondary structure formed depends on the primary structure since different amino acids have different secondary structure propensities.

1.2 Protein Secondary Structures: the α Helix and the β Sheet

A few elements of protein secondary structure are so widespread that they are immediately recognizable in proteins with widely differing amino acid sequences. Both α helix and the β sheet are such elements which are called the secondary structures as they are composed of sequences of residues with repeating ϕ and ψ values.

1.2.1 The α helix

The α helix (Figure 3) is right-handed and has 3.6 residues per turn with a pitch of 5.4 Å. The α helices of proteins have an average length of ≈ 12 residues, which corresponds to over three helical turns, and a length of ≈ 18 Å.

In an α helix, intrastrand hydrogen bonds between backbone NH and C=O in residues i and $i + 4$ can be formed resulting in a stable secondary structure. This results in a strong hydrogen bond which has the optimum N...O distance of 2.8 Å. Amino acid side chains project outward and downward from the helix, thereby avoiding steric interference with the polypeptide backbone and with each other. The core of the helix is tightly packed with its atoms in Van der Waals contact.

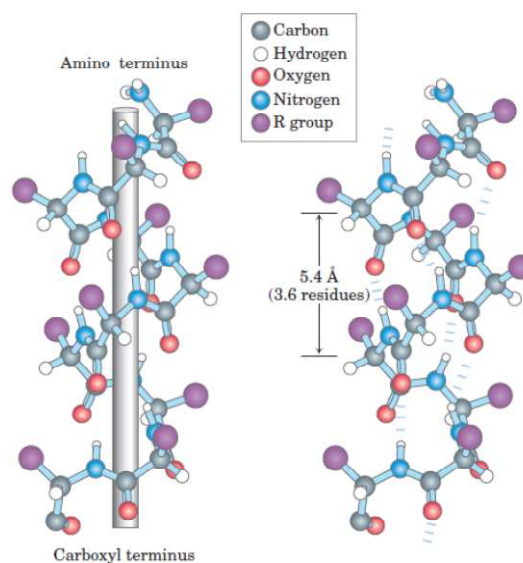


Figure 3: The α helix, dashed lines indicate hydrogen bonds between polypeptide strands⁴

1.2.2 The β sheet

Like α helix, β sheet uses the full hydrogen-bonding capacity of the polypeptide backbone. However in β sheets, hydrogen bonding occurs between the neighbouring polypeptide chains (interstrand) rather than in the same chain as in an α helix (intrastrand).

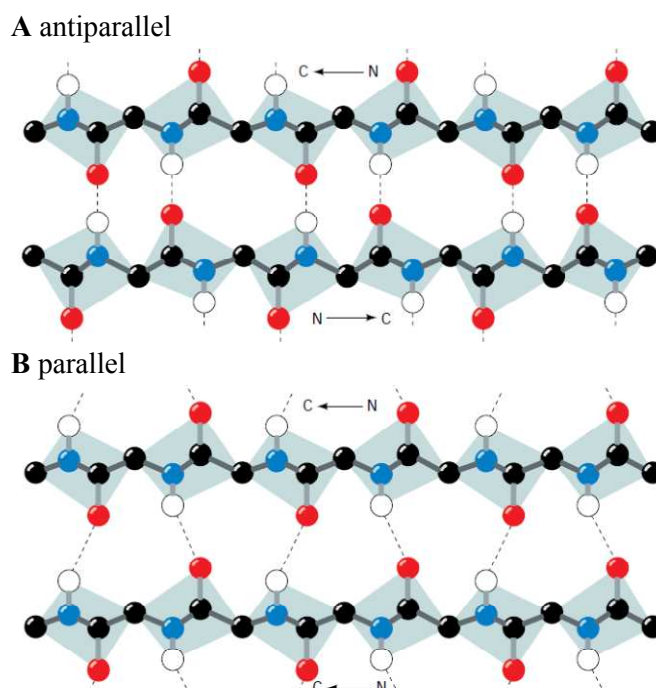


Figure 4: The β Sheets with dashed lines indicating hydrogen bonds between polypeptide strands. (A) An antiparallel β sheet. (B) A parallel β sheet.⁵

β sheets are divided into two categories.

- The antiparallel β sheet, in which the neighbouring hydrogen-bonded polypeptide chains run in opposite directions (Figure 4A).
- The parallel β sheet, in which both the hydrogen-bonded chains extend in the same direction (Figure 4B).

The conformations in which the folded β structures are optimally hydrogen bonded differ from that in the fully extended form. The β sheets have a rippled or pleated edge-on appearance known as “pleated sheets.” Successive side chains of a polypeptide chain in a β sheet extend in opposite sides of the sheet with a two-residue repeat distance of 7.0 Å.

1.3 Other Secondary Structures: 3_{10} helix, β hairpin and turns

1.3.1 The 3_{10} helix

The 3_{10} helix (Figure 5A) secondary structure is found rarely in proteins. The amino acids in a 3_{10} helix are arranged in a right-handed form with each amino acid taking a 120° turn on the helix axis corresponding to three residues per turn, and a translation of 2.0 Å along the helical axis. It has 10 atoms in the ring formed by making the hydrogen bond with the N-H group of i amino acid and the C=O group of the $i + 3$ amino acid. The repeated $i + 3 \rightarrow i$ (Figure 5B) hydrogen bonding defines a 3_{10} -helix. The related structures include the α helix ($i + 4 \rightarrow i$ hydrogen bonding) and the π -helix ($i + 5 \rightarrow i$ hydrogen bonding).

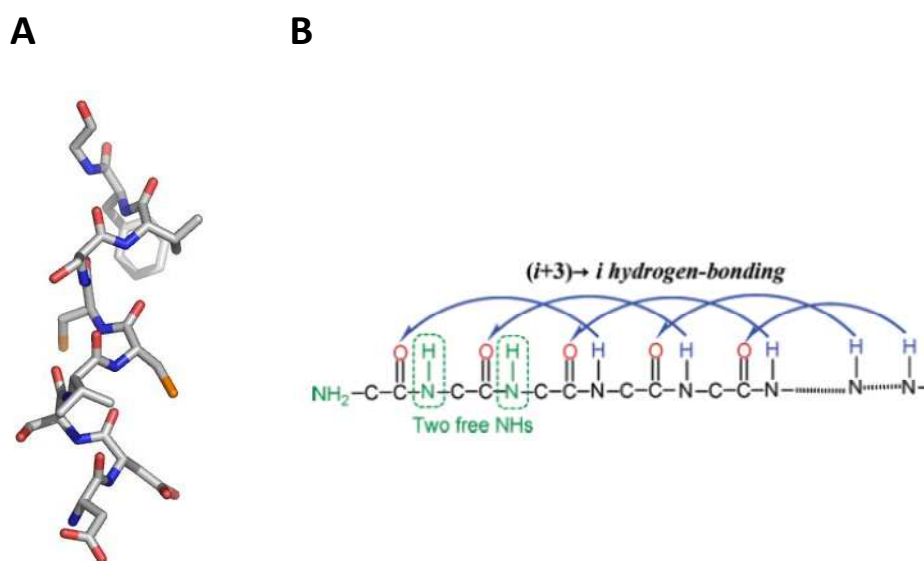


Figure 5: (A) A representative model of 3_{10} helix. (B) Intramolecular hydrogen-bonding pattern of a 3_{10} -helix.

1.3.2 The β -hairpin

The β -hairpin (Figure 6) structure is the simplest protein motif involving two β -strands arranged as a hairpin. The motif consists of two adjacent strands oriented in an antiparallel arrangement (where the N-terminus of one sheet is adjacent to the C-terminus of the next) and linked by a short loop of two to five amino acids. β hairpins can occur in isolation or as part of a series of hydrogen bonded strands that collectively comprise a β sheet/ β barrel.⁶

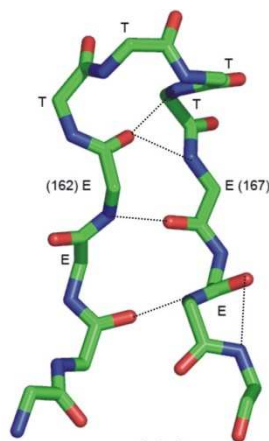


Figure 6: A representative model of β hairpin.⁵

1.3.3 Turns

A turn is a structural motif in which the C^α atoms of two residues are separated by a few (1 to 5) peptide bonds (Figure 7) and come in close proximity ($< 7 \text{ \AA}$), without the corresponding residues forming a regular secondary structure such as an α helix or β sheet. Contrary to helices, the backbone dihedral angles are not (roughly) constant for all the residues in the turn. Although the close approach of the two terminal C^α atoms is usually correlated with the forming of one or two hydrogen bonds between the corresponding residues, such hydrogen bond is not strictly required in the definition of the turn.

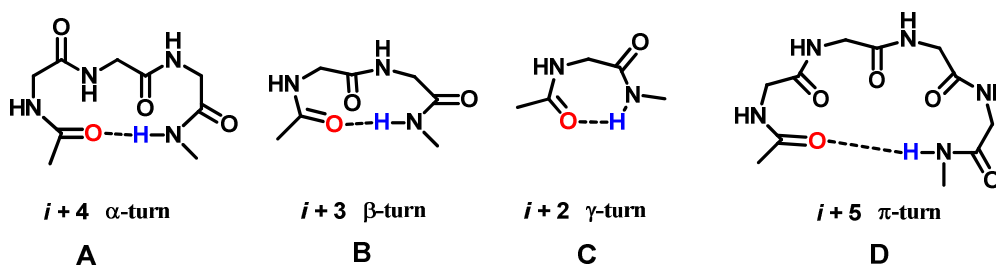


Figure 7: (A) Representation of α turn. (B) Representation of β turn. (C) Representation of γ turn. (D) Representation of π turn.

The turns are classified according to the separation between the two end residues.

- In an α -turn, the end residues are separated by *four* peptide bonds ($i \rightarrow i \pm 4$).
- In a β -turn (the most common form), by three bonds ($i \rightarrow i \pm 3$).
- In a γ -turn, by two bonds ($i \rightarrow i \pm 2$).
- In a δ -turn, by one bond ($i \rightarrow i \pm 1$).
- In a π -turn, by five bonds ($i \rightarrow i \pm 5$).

1.4 Polyproline Conformation

Apart from the regular secondary structures and turns, the other characterized important secondary structure is the Polyproline II (PPII) helix which is present in both folded and unfolded proteins.⁷

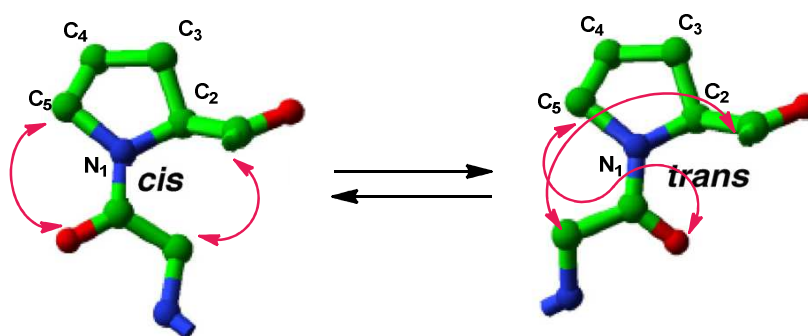


Figure 8: Prolyl *cis-trans* isomerization

The peptide bond of proline residues in proteins exist in *cis* and *trans* conformations (Figure 8). This conversion can be catalyzed by ubiquitous enzymes called peptidyl prolyl *cis-trans* isomerases (PPIases) or rotamases.⁸

1.4.1 Peptidyl prolyl *cis-trans* isomerization

The fundamental repeating unit of a protein backbone is the six atom peptide plane, which contains three atoms from each of two adjacent amino acid residues. This planar unit can be classically viewed as 40% double-bond character of the N-C' bond due to delocalization of the nitrogen lone electron pair. Rotation about the central bond of the peptide plane is highly restricted. Although the conformation of a protein backbone is defined by three torsion angles per amino acid residue (ϕ , ψ and ω), for all the amino acid residues except proline, the *trans* conformation ($\omega = 180^\circ$) is far more energetically favorable than the *cis* ($\omega = 0^\circ$), and protein structure is typically expressed in terms of the ϕ and ψ torsion angles with $\omega = 180^\circ$. However, because of

the unique backbone linked five-membered ring of proline and the corresponding imide peptide bond, the *cis* and *trans* isomers are closer in free energy for the X-Pro peptide bond (where X is any residue), and the *cis* conformation appears with a frequency of 5–6% in protein structures.⁹

Linus Pauling¹⁰ proposed a simple approximate relationship for the strain energy associated with rotation about the C-N bond of a peptide plane.

$$E = 30 \sin^2 \delta \text{ (kcal mol}^{-1}\text{)}$$

where δ is the dihedral angle measuring deviations from the *trans* conformation.

The analysis of 3,938 high-resolution protein structures demonstrate the remarkable accuracy of this early approximation.¹¹ In agreement with this prediction, small deviations from planar geometry ($\pm 5^\circ$) are frequently observed in actual protein structures with deviations of 15–20° being rare. Based on Pauling's prediction, the 90° (*syn*) transition state has an energy that is 30 kcal mol⁻¹ higher than that of the 180° (*trans*) conformation, which indicates that conversion from the *trans* to *cis* state would be expected to occur only rarely. Indeed, NMR measurements for *cis*-to-*trans* conversion in peptides show activation barriers ΔG (Figure 9) ranging from 14–24 kcal mol⁻¹, while the rate measurements by NMR and by proteolytic or protease-free assays confirm slow *cis* to *trans* conversion rates on the order of 0.002 s⁻¹ at 25 °C.¹² The lower free energy of the *trans* state relative to the *cis* state would correspond to a higher activation barrier for the *trans*-to-*cis* conversion, thereby bringing the activation barrier ΔG closer to Pauling's predicted 30 kcal mol⁻¹.

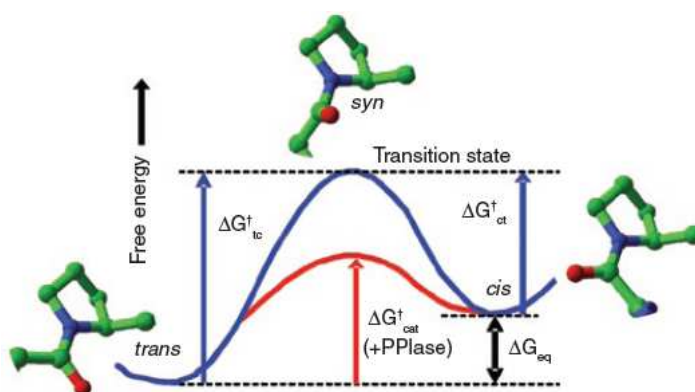


Figure 9: Energy diagram for prolyl *cis*-*trans* isomerization. Conformational exchange proceeds through the twisted 90-*syn* high-energy transition state in the intrinsic (blue curve) or PPIase-catalyzed (red curve) reactions.

1.5 Types of Polyproline Conformation

The polyproline helix is a type of protein secondary structure, which occurs in proteins comprising of repeating proline residues. The polyproline type II (PPII) helix is prevalent in both folded and unfolded proteins⁷ and plays an important role in a wide variety of biological processes, such as signal transduction, transcription, immune response, and cell motility.¹³ Each strand of collagen triplex with the Pro-Hyp-Gly tripeptide repeat unit adopts a left-handed PPII like conformation.¹⁴ Oligoprolines and their derivatives have found utility as cell penetrating agents¹⁵ and as molecular spacers in biomimetic systems for energy/electron transport.¹⁶ The PPII helix has also been proposed as an important local conformation in disordered states of proteins that are relevant for protein folding.¹⁷

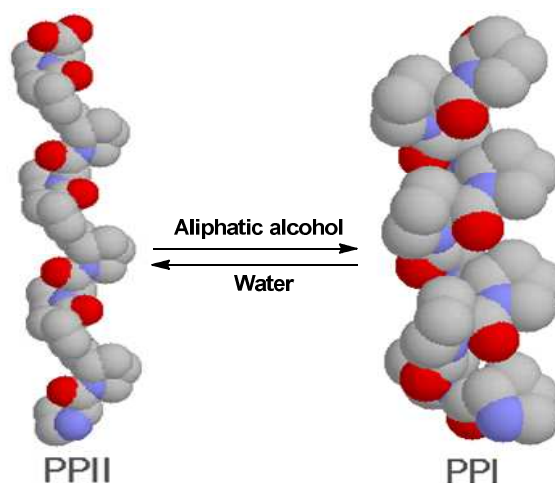


Figure 10: Solvent induced switching between PPII \leftrightarrow PPI conformations

The PPII helix is a fully extended left-handed structure with all amide bonds in the *trans* conformation, while the right handed PPI helix is compact with all amide bonds in the *cis* conformation.⁷ It is well-known that polyprolines adopt PPII conformation in water and PPI conformation in hydrophobic solvents (Figure 10).¹⁸

In contrast to the α -helix and β -sheet, neither the PPII nor the PPI helix is stabilized by intramolecular or intermolecular hydrogen bonds due to lack of amine (donor atom) functionality.

Table 1 shows a comparison between PPII and PPI conformations in peptides.¹⁹

Table 1

Parameter	PPII conformation	PPI conformation
Direction of helix	Left handed	Right handed
Amide bond conformation	<i>trans</i>	<i>cis</i>
Nature of helix	Fully extended	More compact
Dihedral angles	$\omega=180^\circ$, $\phi=-75^\circ$, $\psi=+145^\circ$	$\omega=0^\circ$, $\phi=-75^\circ$, $\psi=+160^\circ$
Helical pitch	9.4 Å per turn, 3.3 proline residues per turn	5.6 Å per turn, 3.3 proline residues per turn
Orientation of amide bond in peptide backbone	Nearly perpendicular to the helix axis	Nearly parallel to the helix axis
Preferred solvent	Water	Aliphatic alcohols

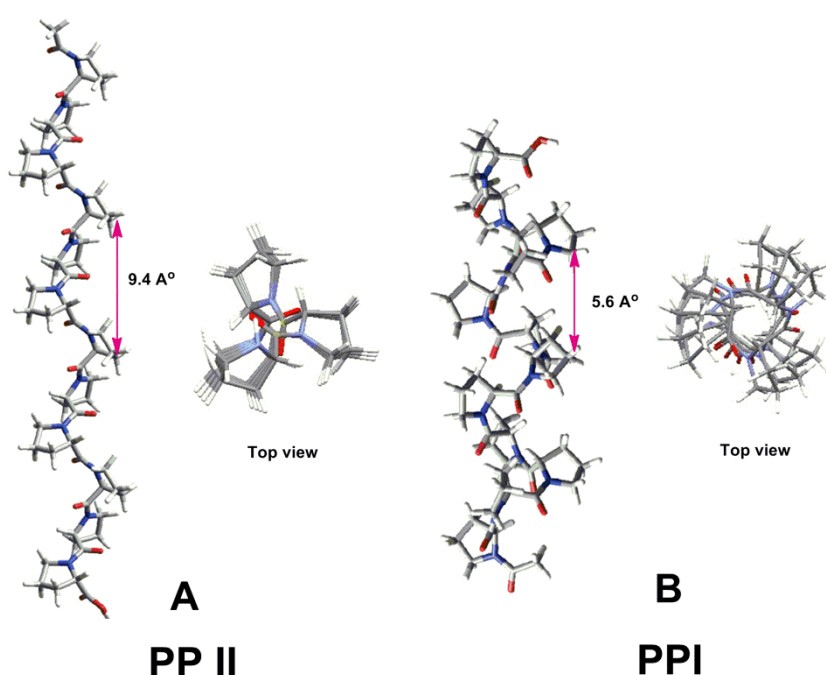


Figure 11: (A) Model structure of polyproline II (PPII) conformation along with top view. (B) Model structure of polyproline I (PPI) conformation along with top view.^{19f}

1.5.1 A survey of left-handed polyproline II helices

Adzhubei and Sternberg²⁰ in their first systematic search, found 96 PPII helices in a protein databank of 80 proteins. They are not only composed of proline successions but some PPII helices have no proline at all, e.g. short stretches of polyglutamines form PPII conformation. Stapley and Creamer²¹ presented a survey of 274 non-homologous polypeptide chains from proteins of known structure for regions that form PPII structures. They have shown that such regions are rare, but the majority of proteins contain at least one PPII helix. Most PPII helices are shorter than five residues, although the longest one contained 12 amino acids. Proline predominates in

PPII, but glutamine and positively charged residues are also favoured. The basis of glutamine prevalence is its ability to form an $i, i+1$ side-chain to main-chain hydrogen bond with the backbone carbonyl oxygen of the preceding residue; this helps to fix the ψ angle of the Glutamine and the ϕ and ψ of the preceding residue in PPII conformations and explains why Glutamine is favoured at the first position in a PPII helix.*

1.5.2 Side-chain rotamer preference

Proline is a cyclic imino acid and the bridging of the α -carbon atom to the main chain amide nitrogen by 3 methylene bridge imposes additional constraint on the main chain torsion angles ϕ and ψ . Further, pyrrolidine ring conformation of Pro is defined in terms of four endocyclic torsion angles χ_1, χ_2, χ_3 and χ_4 (Figure 12A).

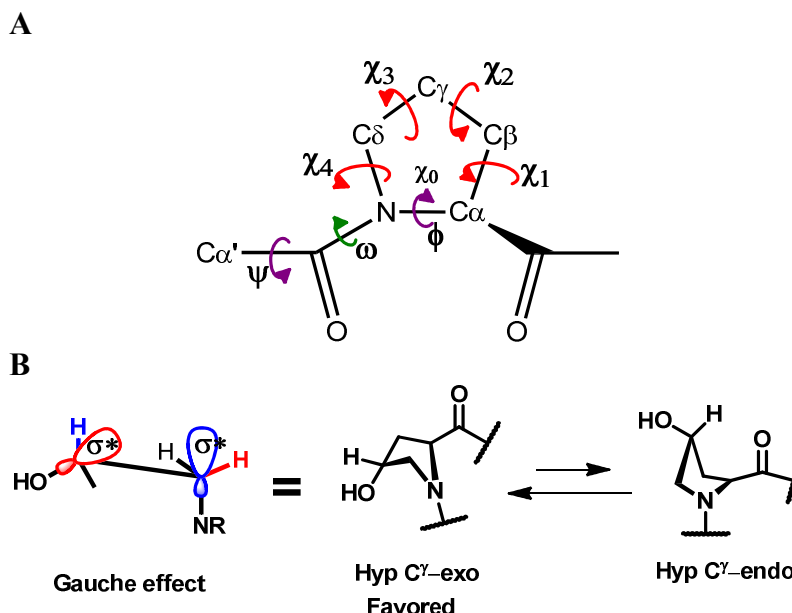


Figure 12: (A) Main chain torsion angles and endocyclic torsion angles in imino acid proline (Pro). (B) Orbital overlaps contributing to stabilization are indicated in blue and red. The favoured C γ -exo pucker of Hyp residue allows for the efficient recognition of hydroxylated product.

It is well known that backbone dihedral angles influence the preferred side-chain χ_1 and χ_2 rotamers.²² These rotamers are classified as *trans* ($-120^\circ > \chi_1 > 120^\circ$), *gauche*⁺

* Berkholz and Karplus recently proposed that its common name could be changed to a more general form, i.e., “polypeptide-II”. This would maintain the familiar acronym, avoid the misleading association with only Proline, and be consistent with the observation that it is a prominent conformation in unfolded polypeptide chains.

($0^\circ > \chi_1 > -120^\circ$) and *gauche*⁻ ($-120^\circ > \chi_1 > 0^\circ$).²³ The PPII region of ϕ space generally precludes the *gauche*⁻ rotamer through clashes with the backbone amide hydrogen and perhaps its hydrogen bond acceptor. The extended nature of the PPII helix means that contacts between the side chains are uncommon and that there is very little effect on rotamer preference from neighbouring side chains. However, aromatic side chains favour the *trans* rotamer while for others *gauche*⁺ is preferred. The side-chain rotamer preferences are similar to those obtained from a survey of all random coil residues including non- structures from a equivalent region of ϕ/ψ space.

1.5.3 Conformational preference and *cis-trans* isomerization of 4(*R/S*)-substituted proline

Song and Kang²⁴ reported the conformational preference and prolyl *cis-trans* isomerization of 4(*R/S*)-substituted proline dipeptides, *N*-acetyl-*N'*-methylamides of 4(*R/S*)-hydroxy-L-proline and 4(*R/S*)-fluoro-L-proline (Ac-Hyp-NHMe and Ac-Flp-NHMe, respectively), using quantum calculations. It was found that the 4*R*-substitution by electron-withdrawing groups did not result in significant changes in backbone torsion angles as well as endocyclic torsion angles of the prolyl ring. However, small changes in backbone torsion angles ϕ and ψ and the decrease of bond lengths $r(C^\beta-C^\gamma)$ or $r(C^\beta-C^\delta)$ appear to enhance the relative stability of the *trans* *exo*-puckered conformation and alter the relative stabilities of transition states for prolyl *cis-trans* isomerization. The population of *trans* *exo*-puckered conformation increased in the order Ac-Pro-NHMe < Ac-Hyp-NHMe < Ac-Flp-NHMe in chloroform and water. This increase in population for *trans* *exo*-puckered (Figure 13) conformations in solution is attributed to an increase in population for the polyproline-II like conformations with *exo* puckering.

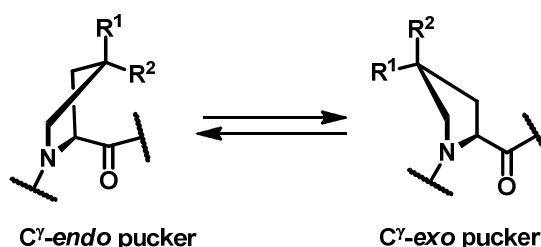


Figure 13: Ring conformations of 4-substituted proline residues. The *C*^γ-*endo* conformation is favoured strongly by steric effects when $R^1 = \text{Me}$, $R^2 = \text{H}$ (mep) or stereoelectronic effects when $R^1 = \text{H}$ and $R^2 = \text{F}$ (flp). Similarly, the *C*^γ-*exo* conformation is favored strongly by steric effects when $R^1 = \text{H}$, $R^2 = \text{Me}$ (Mep) or stereoelectronic effects when $R^1 = \text{OH}$, $R^2 = \text{H}$ (Hyp) or $R^1 = \text{F}$, $R^2 = \text{H}$ (Flp)

For 4*S*-substitution by electron-withdrawing groups, opposite effect is obtained. The population of *cis endo*-puckered conformations increased in the order Ac-Pro-NHMe < Ac-Hyp-NHMe < Ac-Flp-NHMe in aliphatic alcohols. This increase in population for *cis endo*-puckered conformations in solution is attributed to an increase in population for the polyproline-I like conformations with *endo* puckering.

1.5.4 Conformational preferences of β - and γ -aminated proline analogues

Aleman and co-workers²⁵ investigated theoretically, the effect of incorporation of an amino group at the $C^\beta(3)$ - or $C^\gamma(4)$ -positions of the pyrrolidine ring. This affects the intrinsic conformational properties of the proline. Specifically, a conformational study of the *N*-acetyl-*N'*-methylamide derivatives of four isomers of aminoproline, which differ not only in the β - or γ -position of the substituent but also in its *cis* or *trans* relative disposition, has been studied. A seven membered intramolecular hydrogen bond is possible between 4*S* amino group and carbonyl oxygen. The formation of this intra-residue hydrogen bond (Figure 14) explains the higher stability of the Ac- γ -*cis*Amp-NHMe and Ac- β -*cis*Amp-NHMe dipeptides, compared to that of the corresponding analogues with *trans* amino group.

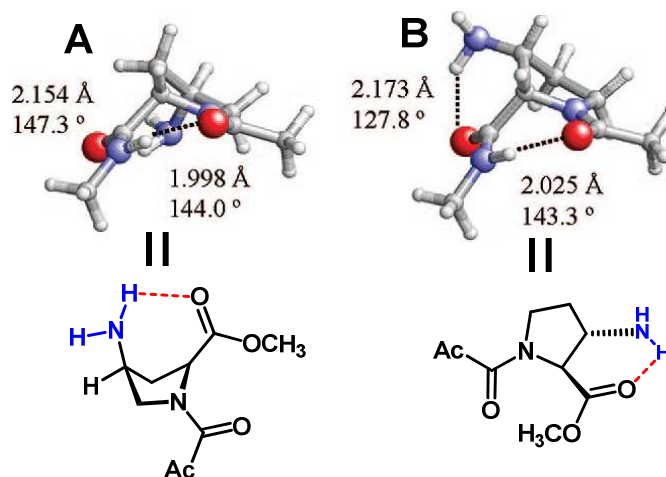


Figure 14: Representation of minimum energy conformations characterized in the gas phase for the Amp-containing dipeptides studied (A) Ac- γ -*cis*Amp-NHMe (B) Ac- β -*cis*Amp-NHMe

The conformational properties may be altered by protonation of amino group to a charged ammonium group. This equilibrium in aqueous solution can be used to control the conformation of the 4*R/S*-aminoproline derivatives by altering pH. The conformational preferences of the protonated 4*R/S*-aminoproline derivatives indicated that²⁶ (a) protonation reduces the backbone conformational flexibility and destabilizes

the *cis* form of the amide bond involving the pyrrolidine nitrogen, (b) the strong interactions of side chain amino group with carbonyl backbone induced significant pyramidalization at the amide nitrogen atom and (c) in neutral analogues, intra-residue interactions of side chain amino group and carbonyl backbone favour *cis* disposition of the substituent.

1.6 Factors Affecting Polyproline Conformation

1.6.1 Stereoelectronic effect on polyproline conformation

Stereoelectronic effect (Figure 15) is defined as the relationship between structure, conformation and reactivity resulting from the alignment of filled or unfilled electronic orbitals. Raines and co-workers²⁷ suggested that the backbone $n \rightarrow \pi^*$ interaction could play a key factor in stabilizing the PPII helices by enhancing stereoelectronic effects that favour the *trans* prolyl peptide bond and increase the folding stability of the triple-helical collagen structure comprised of PPII helices.

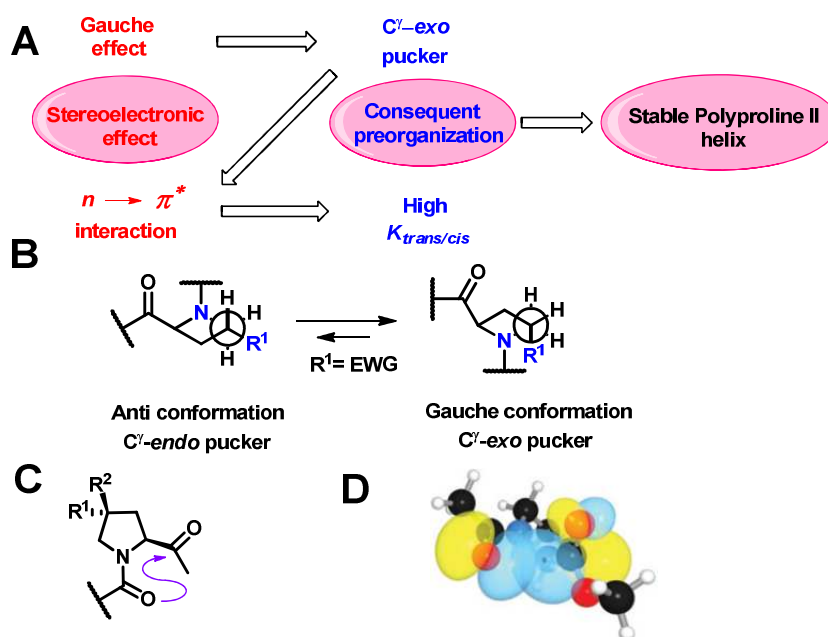


Figure 15: Stereoelectronic effects that stabilize the polyproline helix. (A) A gauche effect and an $n \rightarrow \pi^*$ interaction preorganize main chain torsion angles and enhance the stability. (B) A gauche effect, elicited by an electron withdrawing group (EWG) in the 4*R* position, stabilizes the C γ -exo ring pucker. (C) An $n \rightarrow \pi^*$ interaction stabilizes the *trans* isomer of the peptide bond but is substantial only when Pro derivatives are in the C γ -exo ring pucker (e.g., R¹ = OH or F, R² = H). (D) Depiction of overlap between n and π^* natural bond orbitals in a Pro residue with C γ -exo pucker.

In studies using short polyproline peptides, it has been shown that the stereoelectronic effects have a significant impact on polyproline conformation.²⁸ In

this study, substitution of proline by 4*R*-hydroxyproline (Hyp) or 4*R*-Fluoroproline (Flp) was found to enhance the backbone $n \rightarrow \pi^*$ interaction and stabilize PPII conformation by inhibiting the conversion of PPII to PPI. On the other hand, the 4*S*-fluoroproline (flp) substitution favoured conversion of PPII to PPI and destabilized PPII conformation due to steric effects. An electron-withdrawing substituent in the 4*R* position constrains the ϕ and ψ dihedral angles to be close to that in a PPII helix and favors its requisite *trans* peptide bond by enhancing the $n \rightarrow \pi^*$ interaction between O_{i-1} and C'_i .²⁹ An electron-withdrawing substituent in the 4*S* position obviates the $n \rightarrow \pi^*$ interaction and thereby alters the relative free energy in favour of the PPI helix. The absence of an electron-withdrawing substituent in proline itself leads to less preferences for the ϕ , ψ , and ω dihedral angles and consequently, an intermediate preference for a PPI or PPII helix.

Recently Horng and co-workers³⁰ used host–guest peptides to demonstrate that an electron-withdrawing substituent in the 4*R* or 4*S* position has a substantial impact on the kinetics of PPII→PPI conversion. The kinetic consequences from stereoelectronic effects lead to difference in the stability of PPII helices. Accordingly, stereoelectronic effects provide a rational means to modulate polyproline conformation.

1.6.2 Azidoproline as conformational directing element and functionalizable site

Wennemers and co-workers^{19b} demonstrated that 4(*R/S*)-azidoproline can be used both as conformation directing element of the PPII structure and as a functionalizable site for the development of proline-based molecular scaffolds. The conformational analyses in Ac[(4*R*)-Azp]₉OH and Ac[(4*S*)-azp]₉OH demonstrate that the PPII helix is stabilized by (4*R*)-Azp and destabilized by (4*S*)-azp, whereas the PPI helix has opposite behavior and is stabilized by (4*S*)-azp and destabilized by (4*R*)-Azp. The stabilizing effect of (4*R*)-Azp is due to an enhancement of the $n \rightarrow \pi^*$ interactions which have been proposed to stabilize the PPII conformation.

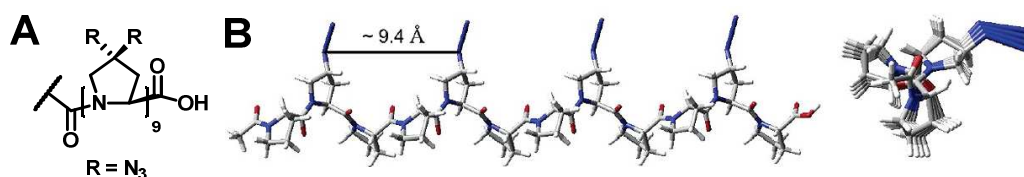


Figure 16: (A) 4(*R/S*)-azidoproline oligomer. (B) Model of an oligoproline PPII-helix with Azp residues in every third position.

These 4*R/S*-azido groups provide site for further functionalization which can be done by “click chemistry” leading to Azp containing polyprolines as attractive molecular scaffolds.

1.6.3 Effect of *O*-Galactosylation of 4*R*-Hydroxyproline on PPII conformation

The hydroxyproline-rich glycoproteins (HRGPs) are the major structural proteins of the extracellular matrix of algae and land plants. They are characterized by a rigid PPII conformation and extensive *O*-glycosylation of 4*R*-hydroxy-L-proline (Hyp) residues, which is a unique post-translational modification of proteins. Schweizer and co-workers³¹ investigated the effects of naturally occurring β -*O*-galactosylation of Hyp in a series of well-defined model peptides Ac-(Pro)₉-NH₂, Ac-(Hyp)₉-NH₂ and Ac-[Hyp(β -D-galactose)]₉-NH₂ (Figure 17A). They demonstrated that contiguous *O*-glycosylation of Hyp residues causes a dramatic increase in the thermal stability of the PPII helix according to analysis of thermal melting curves. Molecular modeling indicates that the increase in conformational stability may be due to a regular network of inter-glycan and glycan-peptide hydrogen bonds (Figure 17B), in which the carbohydrate residues form a hydrophilic “overcoat” of the PPII helix.

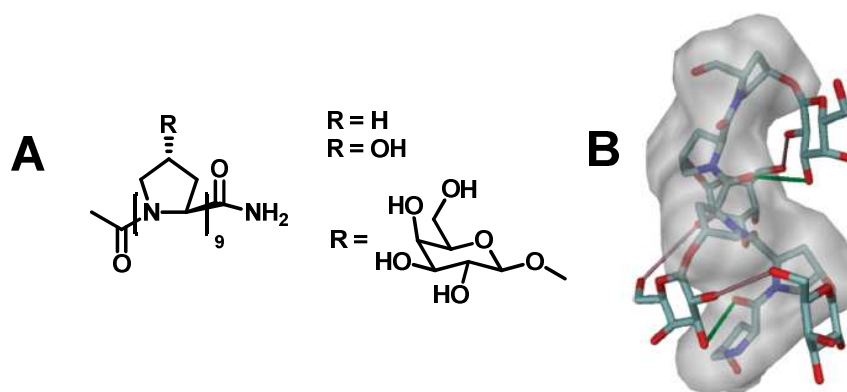


Figure 17: (A) Model polyproline peptides. (B) Molecular modeling of truncated of *O*-Galactosylation peptides shows the surface area of the PPII backbone (shaded) with the D-galactose residues lying in the grooves of the PPII helix. Also shown are inter-glycan H-bonds (C6(OH)_i---(OH)C_{2i+1}) (pink) and glycan-peptide carbonyl backbone H-bonds (C3-(OH)_i---OC_{i-2}) (green).³¹

1.6.4 Effects of terminal functional groups on the stability of the PPII structure

The conformational stability of the PPII helix with respect to the functional groups at the C- and N-terminus was examined both experimentally and theoretically.^{19f} The oligoproline (a) AcN-[Pro]₁₂-CONH₂ (b) H-[Pro]₁₂-CONH₂ (c) AcN-[Pro]₁₂-CO₂H

and (d) H-[Pro]₁₂-CO₂H with charged and capped termini served as model compounds shown (Figure 18).

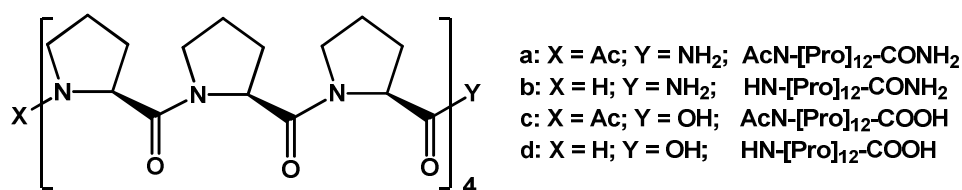


Figure 18: Oligoprolines a-d with different functional groups at the C- and N-terminal.

The conversion of PPII to the PPI was used as a measure of the conformational stabilities. The CD spectroscopic studies show that a positively charged N-terminus and a negatively charged C-terminus destabilize the PPII helix and favour the PPI helix, while the capped termini favour the PPII over the PPI helix. These experimental findings are supported by the energy differences computed by *ab initio* methods between the PPII and PPI helices of oligoprolines (Figure 18, **a-d**). The electrostatic interactions between the terminal charges and the amide dipoles stabilize both helices. They are significantly stronger in the PPI helix where the amide bonds are oriented linear to the helix axis as compared to the PPII helix in which the amides are perpendicular to the axis. The negative charge at the C-terminus has a more pronounced effect on the relative stability as compared to a positive charge at the N-terminus. This is due to destabilization of the PPII helix by repulsive interaction between the C-terminal carboxylate with the neighboring amide bond. Studies at different pHs demonstrated how changes in the protonation state can be used to deliberately stabilize the PPII helix over the PPI helix or *vice versa*.

1.6.5 Temperature induced transition between polyproline II and I helices

The conformational properties of the polyproline I (PPI) helix in oligoprolines as a function of temperature has been examined.³² Oligoproline H-(Pro)₁₂-NH₂ adopts a pronounced PPI conformation with all *cis* amide bonds in *n*-propanol. CD spectroscopic studies revealed that a conformational change from the PPI to the PPII helix (Figure 19) takes place upon heating from 10 to 90 °C and revert to the PPI helix upon cooling. This conformational transition cycle is characterized by a strong hysteresis which suggests that the *cis-trans* isomerization of the amide bonds involved in the transition from the PPI to the PPII helix and *vice versa* is slow compared to the rate of change in the temperature. The transition PPI → PPII occurs with a positive

standard enthalpy of $\approx 33.5 \text{ kJ mol}^{-1}$ and increasing standard entropy of $\approx 102 \text{ J mol}^{-1} \text{ K}^{-1}$. The oligoprolines are thus good as thermally switchable synthetic scaffolds.

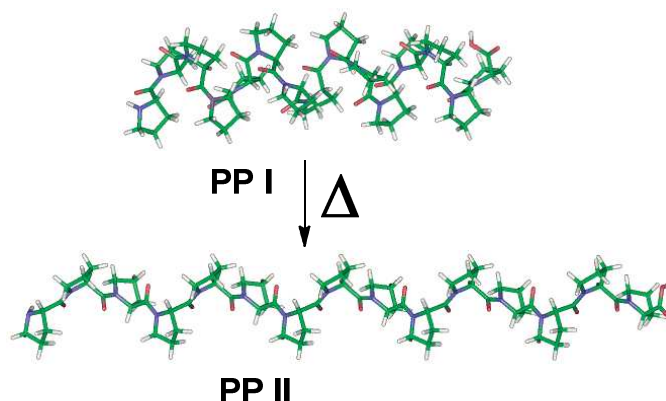


Figure 19: Temperature induced switching between polyproline II and I helices.³²

1.6.6 Effect of H₂O and D₂O on polyproline II conformation

The interaction of solvent with a polypeptide chain is one of the primary factors controlling the protein folding and stability. Estimates of the effects of water on peptide folding can be obtained from solvent perturbation experiments. Chellgren and Creamer³³ showed that D₂O stabilizes the left-handed polyproline II (PPII) helical conformation of peptides relative to H₂O. In host-guest peptides of sequence acetyl (Pro)₃-X-(Pro)₃-Gly-Tyr-NH₂ with one or more alanine or valine in X-position, D₂O stabilizes the PPII helices relative to H₂O. This arises from helices disrupting the bulk water less than other secondary structures and since D₂O is more ordered than H₂O, disruption of bulk water structure would have a greater effect in D₂O.

1.6.7 Effect of urea on polyproline II conformation

It is commonly assumed that urea denatures proteins by promoting backbone disorder, resulting in a random-coil behavior. Creamer *et.al.*³⁴ demonstrated that even denatured states possess significant amounts of locally ordered backbone structure and that urea promotes polyproline II helix formation in oligomers of proline, alanine, valine and some proteins.³⁵

The mechanism by which urea promotes polyproline II helical structure is not clear. Nozaki and Tanford³⁶ and Robinson and Jencks³⁷ (Figure 20) proposed that urea interacts favourably with the polypeptide backbone influencing the conformation to favour the PPII helical form. This effect is modulated by the nature of the side chains, leading to the observed sequence dependence. Tiffany and Krimm³⁸ suggested that

urea and guanidine hydrochloride interact with proteins through hydrogen bonding to the backbone carbonyl group.

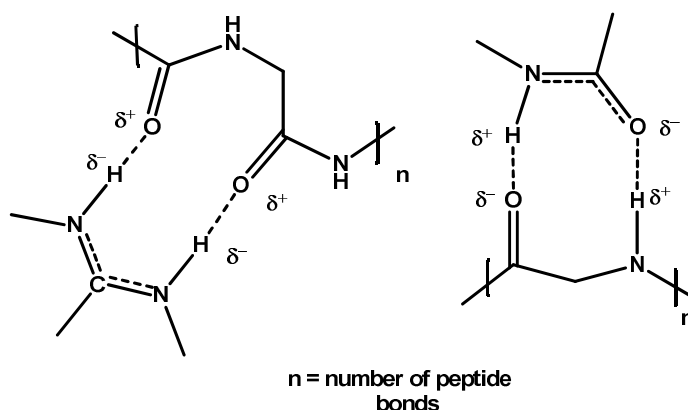


Figure 20: Mechanism of rigidification by urea.

1.6.8 Effect of salt on polyproline structure

Mattice and Mandelkern³⁹ found that the PPII helical content of proline oligomers or peptides decreased with increasing salt concentration. PPII helical structures are disrupted by addition of sodium chloride in case of proline and polylysine oligomers.⁴⁰ It has also been demonstrated that neutralization of side chain charges in polylysine at higher pH induces α -helix formation, suggesting the important role of electrostatic interaction in stabilizing the PPII helices.⁴¹ This may occur by increase in the rotational freedom about the -C=O and C^α in poly-L-proline,⁴² an increase in rotational freedom about the C=O-NH bond,⁴³ reducing the energy difference between the *cis* and *trans* isomers of planar peptide bond.⁴⁴ It has been suggested that there may be direct binding of the salt to the peptide group.⁴⁵ The barrier to *cis-trans* isomerization in polyproline is due in part to electrostatic interactions.⁴⁶ The decrease in PPII structure was recently proposed as a result of chaotropic action of concentrated salt.⁴⁰ The existence of metal ion effect on amide groups in the presence of water was proposed by Jencks and co-workers⁴⁷ (Figure 21).

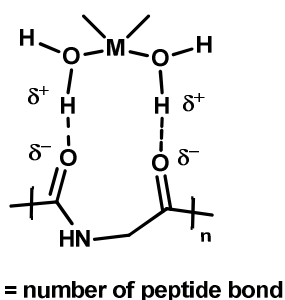


Figure 21: Mechanism of salt binding to peptide backbone

1.6.9 Effect of aliphatic alcohol on polyproline conformation

Solvent can play a role in inducing a conformational change through the *cis-trans* isomerism about X-Pro peptide bonds, where X is any amino acid residue. All amide bonds in proteins and peptides prefer to exist predominantly in the *trans* conformation in water, while about 25% of X-Pro groups adopt *cis* conformation in water.⁴⁸ The relative amount of the *cis* conformation varies with the nature of the solvent⁴⁹ and dramatic manifestation of the solvent effect on proline peptides is the polyproline I \leftrightarrow II interconversion. It is well-known that polyprolines adopt PPII conformation in water and PPI conformation in hydrophobic solvents (Figure 22). While the PPII conformation is stabilized by solvation of the backbone,⁵⁰ the way water influences the tendency of peptides to adopt a PPII conformation is still not well understood.

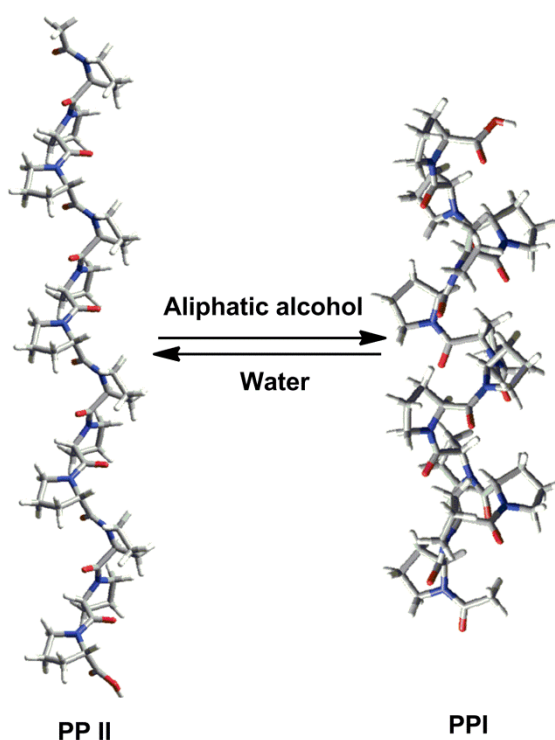


Figure 22: Solvent induced switching between PPII \leftrightarrow PPI conformations.^{19f}

1.7 Biological Significance of Polyproline Conformation

The PPII helix is widespread in nature and plays an important role in many biological processes such as signal transduction, transcription, immune response and cell motility. The single strands of collagen with the typical [Pro-Hyp-Gly]_n repeat unit adopt a PPII-like conformation. The PPII helix has also been proposed as an important local conformation in disordered states of proteins that are relevant for protein folding.

The oligoprolines have been used as molecular spacers for studying, for example, energy or electron transport.

The importance of proline *cis/trans* isomerization as rate-limiting step in *in vitro* protein folding has been well characterized.⁵¹ In interleukin-2 tyrosine kinase (*itk*), *cis-trans* isomerism about the Asn286-Pro287 amide bond is responsible for the coexistence of two different *itk* conformations that display distinct biological activities.⁵²

The isomerisation process of Xaa-Pro peptide bonds can be catalyzed by peptidyl prolyl *cis/trans* isomerase (PPIase) enzymes,⁵³ found to be involved in cell signalling and cell replication. It is also implicated in the induction of severe diseases such as cancer, AIDS, Alzheimer's disease and other neurodegenerative disorders.⁵⁴ Proline isomerization also functions as molecular switch due to its potential ability to control protein activity within the confines of the intrinsic conformational exchange.

Many proline-rich proteins participate in delivering actin monomers to specific cellular locations such as ruffles, filopodia and microspikes, necessary for cell motility. The proline-rich proteins involved in regulating actin polymerization are classified according to their role in recruiting profilin to the membrane. The proline-rich motifs that bind to profilin and/or EVH1 domains are designated as PRM1 and PRM2.

1.7.1 Oligoproline as cell penetrating agents

It has been reported that proline-rich peptides⁵⁵ and proline dendrimers⁵⁶ can be internalized by eukaryotic cells. The most important advantage of proline-rich peptides in biological systems is their solubility in water. In this context, Royo and co-workers⁵⁷ synthesized the *cis-γ*-amino-L-proline oligomers functionalized at the proline α -amine with several groups that mimic the side chains of natural amino acids, including alanine, leucine, and phenylalanine. These γ -peptides enter into different cell lines (COS-1 and HeLa) *via* an endocytic mechanism. In addition to their capacity for cellular uptake, these unnatural short length oligomers offer advantages over the well-known cell penetrating TAT peptide, being less toxic and protease resistance.

Chmielewski and co-workers⁵⁸ have synthesized cell penetrating agents that introduce cationic and hydrophobic moieties along the backbone of a polyproline helix in an amphiphilic manner. This was done by O-alkylation of hydroxyproline monomer

to yield a scaffold that displays both hydrophobic and cationic moieties. Introduction of an isobutyl group onto hydroxyproline led to a proline-based mimic of leucine, whereas functionalization with amino or guanidinium groups led to proline-based mimics of lysine or arginine respectively (Figure 23). The addition of groups had little effect on the backbone structure. Dramatic increases in uptake were found with MCF-7 cells when upto six guanidinium groups were positioned on the polyproline helix, whereas only modest increases in cellular uptake were observed with the amine-containing polyproline compounds as compared to their flexible counterparts. Amphiphilicity played a key role in the enhanced cell translocation, as scrambled versions of the designed agents, with hydrophobic and cationic groups on all faces of the helix, were only as effective as their flexible peptide counterparts. Interestingly, the most potent agent, P11LRR, demonstrated efficiently by an order of magnitude in cellular uptake as compared to that of the well-studied Tat peptide, with lower cytotoxicity.

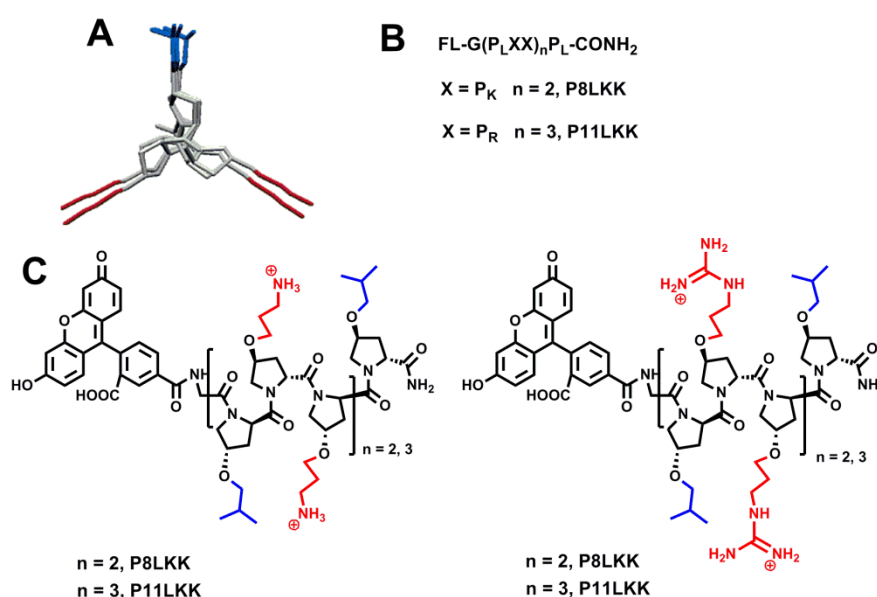


Figure 23: (A) Top view of an amphiphilic polyproline helix containing modified side chains (red, cationic; blue, hydrophobic), (B) Sequences, and (C) Structures of modified polyproline oligomers containing amino (PK) or guanidinium (PR) functionality.

1.7.2 Polyproline conformation in elastic function

PPII conformation is widely present in titin, abductin, and elastin, which are elastomeric proteins and therefore it could play a pivotal role in the genesis of elasticity. The giant elastic protein titin, whose complete domain organization has been revealed,⁵⁹ gives rise to an elastic sarcomere matrix in striated muscle. Precisely, the I-

band region of the sarcomere is involved in the elastic response upon stretch and is composed of two main domains. The first domain is made of tandem repeats of about 100 residue immunoglobulin and the second domain is a motif consisting of tandem repeat of four amino acid residues, PEVK. The extension of the PEVK segment is important for the elastic response of striated muscle to passive stretch and behaves mechanically as an entropic spring.⁶⁰ Structural studies of PEVK motif indicated the likely presence of PPII structure with flexible joints. In aqueous solution PEVK contains multiple PPII helices in equilibrium which are flexible and make significant entropic contribution to elasticity.

Table 2 represents list of proteins comprising significant stretches of PPII conformation and technique used for conformation studies.

Table 2: Proteins comprising significant stretches of PPII conformation⁶¹

Protein	Technique
Human tropoelastin	CD
Exon 30 of human tropoelastin	CD
Abductin	CD
Titin	CD
Collagen	CD, X-ray
C hordein	CD
Antigenic peptide of foot and mouth disease virus (FMDV)	CD, X-ray
Wheat glutenin	CD
Bowman-Birk inhibitor	ROA
Amyloidogenic prefibrillar intermediate of human lysozyme	ROA
Protein kinase inhibitor	X-ray
p85 subunit of P13-kinase	CD
Profilin and class II major histocompatibility complexes	X-ray
Mucin	CD
γ -zein protein of maize	CD
Casein milk proteins	ROA
α -synuclein proteins	ROA
Tau proteins	ROA
Syp (Tyr phosphatase)	X-ray
Ligand-acceptor complex of SH3 of Src	X-ray
Ligand-acceptor complex of SH3-5(Tyr kinase)	X-ray
Antigens(mixture)-MHC class II	X-ray

Another elastomeric protein abductin, is a major constituent of the abductor muscle in the bivalve shellfish "*Pecten jacobaeus*." The ligament lets the mollusc swim by opposing the opening movement of the bivalve of the shellfish. The sequence of abductin⁶² shows the motif FGGMGGGNAG contained protein in tandem repeat. This

decapeptide studied by CD at different temperatures suggests the presence of a mixture of PPII conformation and random coil in equilibrium which may be important in relation to the mechanism of elasticity of the protein.

Polyproline conformation is also present in elastin, the protein responsible for the elasticity of many vertebrate tissues, including skin, lungs, ligaments, and large arteries.⁶³ The elasticity of this protein is entropy driven, in restoring the force responsible for the elastic behavior originating from an increase of disorder accompanying the transition of the stretched to the relaxed state. The octapeptide ALGGGALG of elastin containing GG adopts mainly the PPII conformation which seems to be important for the elastic behavior of the protein

1.7.3 Self-assembly processes and PPII conformation

PPII conformation is prevalent in elastomer proteins such as elastin, abductin and lamprin. The common characteristic of all three proteins is both the presence in the extracellular matrix and the propensity to the formation of fibrils. Lamprin is the most important matrix protein of lamprey annular cartilage. The sequence of the protein has revealed the presence of repetitive sequences such as GGLGY with occasional replacements of Y.⁶⁴ The CD spectra of GGLGA shows quasi-extended helix of polyproline II-type helices which are able to interact inter molecularly, giving rise to the formation of fibrillar supramolecular structures.

1.7.4 Biophysical reasons for why proline is a common binding motif

Proline is unique among the 20 common amino acids in having the side-chain cyclized onto the backbone nitrogen atom. This means that the conformation of proline itself is limited, with backbone ϕ angles of $\approx -65^\circ$. It also restricts the conformation of the residue preceding the proline because of the bulk of the *N*-substituent and results in a strong preference for a β -sheet conformation.⁶⁵ As a consequence, polyproline sequences tend to adopt the PPII helix, which is an extended structure with three residues per turn. This implies that the two prolines in the core, PXP, are on the same face of the helix and are thus well placed to interact with the protein. The PPII helix is an unusual structure: the prolines form a continuous hydrophobic strip round the surface of the helix, while the backbone carbonyls present ideal hydrogen bonding sites, being both conformationally restricted (and therefore poorly hydrated) and

electron-rich. Therefore, PPII helices present an easily accessible hydrophobic surface, as well as a good hydrogen-bonding site. The accessibility of PPII helices is greatly enhanced by the fact that they are frequently found either at the NH₂ or carboxyl termini of proteins where they form extended structures that have been described as ‘sticky arms’.⁶⁶

Because proline-rich regions are exposed, their on-rates and off-rates for binding targets can be very fast. However, due to these fast rates, the complexes are not structurally very well defined on a nanometer scale. Therefore, proline-rich sequences are commonly found in situations requiring the rapid recruitment or interchange of several proteins, such as during initiation of transcription, signaling cascades, and cytoskeletal rearrangements. The role of proline-rich regions is not to provide a structurally defined complex but rather to bring proteins together in such a way that subsequent interactions are more probable.

1.7.5 Polyproline peptidomimetics

With the elucidation of peptide sequences that bind to specific domains, it should be possible to use peptide ligands to interfere with specific SH3 domains in the cells. Ligands for the SH3 domains of Src and Lyn have been injected into oocytes⁶⁷ or electroporated into mast cells⁶⁸ and led to a biological response. Although most peptides fail to cross the membrane on their own, in the future it should be possible to link them to peptide segments of *Antennapedia*,⁶⁹ Kaposi fibroblast growth factor, or HIV Tat,⁷⁰ which do have the capacity to cross the plasma membrane. A number of proline-rich peptides which are antimicrobial and adopt a PPII conformation, such as indolicidin and bactenecin,⁷¹ act by crossing cell membranes on their own. When neutrophils are incubated with the PR-39 which is a proline- and arginine-rich antimicrobial peptide, NADPH oxidase production of superoxide anion O₂⁻ is blocked. The peptide is presumed to act by binding to the SH3 domains of the NADPH oxidase subunits, thereby blocking assembly of a functional enzyme.⁷² Several research groups attempted to create peptidomimetics of the scaffolding prolines of SH3 ligands. This could have major consequences for drug design, because it allows a search for novel functionalities to replace the proline that can interact with new regions of the SH3 and hence enhance binding. Novel oligomers containing proline analogs have been observed to adopt a PPII conformation.

2 Methodology

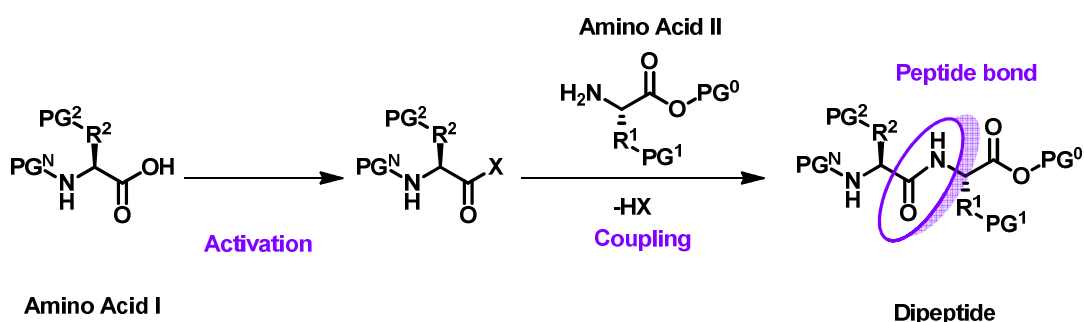
2.1 Peptide Bond Formation: Methods and Strategies

The key step in the peptide synthesis is the formation of the amide bond between two amino acid segments. Carboxy acid function is activated as acyl halides, acyl azides, acylimidazoles, anhydrides, esters etc. There are different ways of coupling reactive carboxy derivatives with an amine:

- ❖ An intermediate acylating agent (R-CO-) formed is isolated then subjected to aminolysis.
- ❖ A reactive acylating agent is formed from the acid in a separate step, followed by immediate treatment with the amine.
- ❖ The acylating agent is generated *in situ* from the acid in the presence of the amine, by the addition of an activating or coupling agent.

The activation of the carboxyl group of the first amino acid is followed by a nucleophilic attack of the amino group of the second amino acid to form the peptide bond. In order to prevent undesirable peptide bond formation, the amino group of the first amino acid and functional side chain groups need to be blocked. Repeated de-blocking, activation and coupling leads to the desired final peptide sequence (Scheme 1).

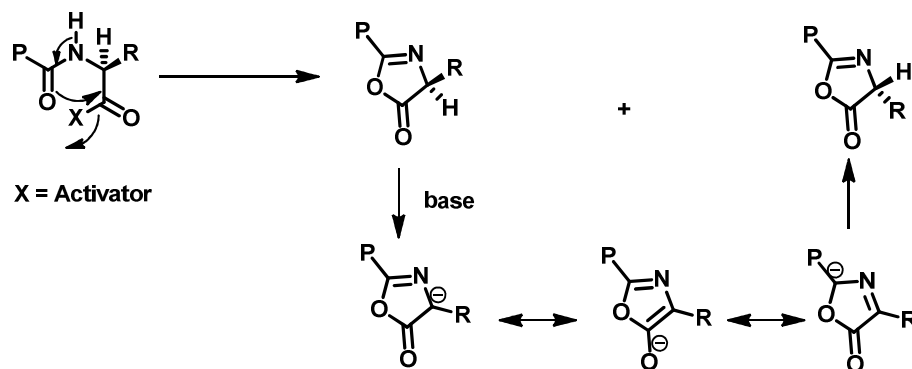
Scheme 1: Simplified general mechanism of peptide bond formation.



However, due to the presence of various functional groups in natural and unnatural amino acids and particularly the requirement for full retention of chiral integrity, the coupling of amino acids and peptides under mild conditions can be challenging. A plethora of coupling reagents has been developed superseding each other in efficiency

and suitability for specific applications. Racemisation can occur at the C-terminal amino acid residue in the course of a coupling reaction due to the ionization of the α -hydrogen and the formation of an oxazolone intermediate (Scheme 2).

Scheme 2: Racemization mechanism in the course of peptide bond formation



In recent years, peptide coupling reactions have significantly advanced with the development of new peptide coupling agents. DCC as a peptide-coupling reagent is replaced by onium-type coupling reagents to smoothly incorporate sterically hindered amino acids including N-methylated and α,α -dialkylated amino acids into the corresponding peptides. Racemisation suppressants are also used as additives to the peptide coupling reagents which also act as a rate enhancer.⁷³

Table 3 evaluates the advantages, disadvantages, and effectiveness of various peptide coupling reagents, based on phosphonium, uronium, immonium, carbodiimide, imidazolium, organophosphorous, acid halogenating reagents.

Table 3a: Imidazolium based coupling reagents

Sr. No.	Coupling Reagent	Structure of Coupling Reagent	Advantages	Disadvantages	Effectiveness
A	CDI		CBMIT is particularly useful in peptide coupling reactions with sterically hindered amino acids, no sign of racemisation in the presence of $\text{CuCl}_2/\text{Cu}(\text{OTf})_2$.	CBMIT is moisture sensitive and air sensitive, choice of solvent is restricted to nitro methane	New esterification reagents to avoid the toxic HMPA by-product of the BOP reagent, best result in the formation of a dipeptide
B	CBMIT				
C	CMBI				

Table 3b: Uronium based coupling reagents

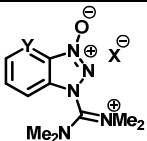
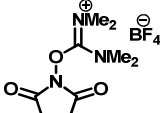
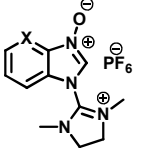
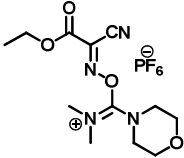
Sr. No.	Coupling Reagent	Structure of Coupling Reagent	Advantages	Disadvantages	Effectiveness
A	HBTU/ TBTU/ HATU	 <p>X = PF₆, Y = H HBTU X = BF₄, Y = H TBTU X = PF₆, Y = N HATU</p>	TSTU and TNTU were recognized as useful peptide coupling reagents in aqueous reactions, successful optimization for synthesis of PNA analogues, modification of HBTU provided several new peptide coupling reagents, excellent choice for solution-phase synthesis	COMU shows a less hazardous safety profile than benzotriazole-based reagents, such as HATU and HBTU, which in addition exhibit unpredictable autocatalytic decompositions and therefore a higher risk of explosion	Useful for Kg scale synthesis, use for solid as well as solution phase synthesis, difference in activities of these compounds could be explained by the hydrogen bond from the additional nitrogen atom of HOAt, stabilising the activated ester intermediate <i>via</i> the anchimeric assistance effect
B	TSTU				
C	HAMDU/ HBMDU	 <p>X = N HAMDU X = CH HBMDU</p>			
D	COMU				

Table 3c: Phosphonium-based coupling reagents

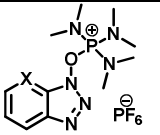
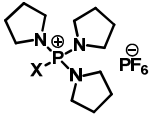
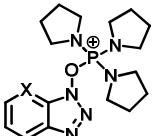
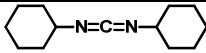
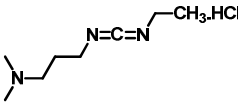
Sr. No.	Coupling Reagent	Structure of Coupling Reagent	Advantages	Disadvantages	Effectiveness
A	BOP/AOP	 <p>X = CH BOP X = N AOP</p>	Racemization is minimal, useful for preparing esters under mild conditions, reactions are rapid, useful in difficult coupling, such as coupling N-methyl amino acids or α,α -dialkylglycines	BOP must be handled with caution as highly carcinogenic Hexamethylphosphoramide is formed as a by-product in coupling reactions.	To avoid the racemization and side reactions that can occur with carbodiimide reagents, many alternative reagents were developed to generate OBt esters <i>in situ</i> .
B	PyCloP/PyBroP	 <p>X = Cl PyCloP X = Br PyBroP</p>			
C	PyBOP/PyAOP	 <p>X = CH PyBOP X = N PyAOP</p>			

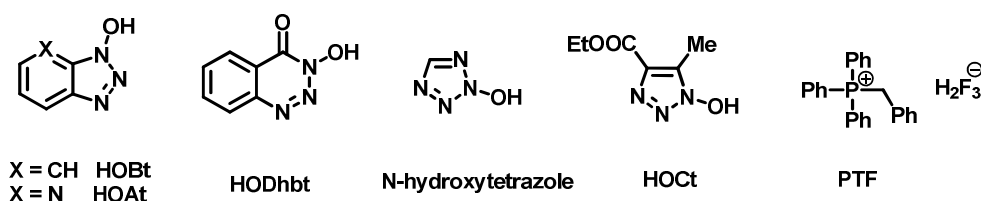
Table 3d: Carbodiimide based coupling reagents

Sr. No.	Coupling Reagent	Structure of Coupling Reagent	Advantages	Disadvantages	Effectiveness
A	DCC		Very useful in solution phase reactions, convert primary amides to nitriles,	DCC form its urea side product which is not soluble in any organic solvent,	Commonly used to prepare, amides, esters and acid anhydrides from carboxylic acids.
B	EDC		moderate activity and they are reasonably cheap, EDC its urea by-product are water soluble	some time causes a partial racemization of the amino acid	

2.1.1 Racemisation suppressants

Geiger⁷⁴ first reported the use of HOBt as a racemisation suppressant in peptide coupling reactions with carbodiimide coupling reagents. Several additives such as HOBt, HOAt, HODhbt, N-hydroxytetrazole, HOt, and PTF have been developed which not only suppress racemisation, but also enhance the reactivity.

HOAt has been reported to be more efficient than HOBt because of an anchimeric assistance caused by the pyridine ring.⁷⁵ Later, N-hydroxytriazoles and N-hydroxytetrazole were examined for their coupling efficiency.⁷⁶ Ramage and co-workers⁷⁷ reported the coupling reaction of dipeptide with DIC and the newly designed HOt for a racemisation study. The racemisation with EDC/HOAt activation was negligible for all amino acids except histidine.

**Figure 24:** Some examples of racemisation suppressant in peptide coupling reactions

2.1.2 Solution phase peptides synthesis

This technique has been used for the synthesis of small peptides composed by only a few amino acid residues. Its main advantage is that the intermediate products can be isolated and purified after each step of synthesis, deprotected and recombined to obtain larger peptides of the desired sequence. This technique is highly flexible with respect to the chemistry of coupling and the combination of the peptidic blocks. New strategies for synthesis in solution have been developed, going from the design of

functional groups for the side chains and condensation of fragments for the synthesis of large molecules⁷⁸ to the use of new coupling reagents.⁷⁹

2.1.3 Solid phase peptide synthesis (SPPS)

In solid phase synthesis chemical transformations are carried out on solid support by sequential addition of α -amino and side-chain protected amino acid residues to an insoluble polymeric support. Acid labile Boc group or base labile Fmoc group is used for N- α -protection. After removal of the protecting group, the next protected amino acid is added by using either a coupling reagent or pre-activated protected amino acid derivative. The resulting peptide is attached to the resin *via* a linker or directly through its C-terminus and may be cleaved to yield a peptide acid or amide, depending on the linker. Side-chain protecting groups are often chosen so as to cleave them simultaneously with detachment of the peptide from the resin. Cleavage of the Boc protecting group is achieved by trifluoroacetic acid (TFA) and the Fmoc protecting group by piperidine or diethyl amine. Final cleavage of the peptidyl resin and the side-chain deprotection requires strong acid, such as hydrogen fluoride (HF) or trifluoromethanesulphonic acid (TFMSA) in case of Boc chemistry and 20% TFA in DCM in case of Fmoc chemistry.

Table 4 shows the merits and demerits for solid phase peptide synthesis *versus* solution phase peptide synthesis.

Table 4: Solid phase *versus* solution phase peptide synthesis

Solid phase peptide synthesis	Solution phase peptide synthesis
Good for short chain as well as long chain peptides	Good for only short chain peptides
Limitation for large-scale synthesis of peptides	The large-scale synthesis of peptides can be carried out
Isolation and purification is not possible	Need to isolate and purify intermediates
Problems of handling large amounts of expensive resins or large excesses of reagents	Easy to handle the reagents
Generally racemisation not observed	In some cases racemisation observed
This technique is highly flexible with respect to the chemistry of coupling and the combination of the peptidic blocks	This technique has limited flexibility with respect to the chemistry of coupling and the combination of the peptidic blocks

2.1.3a Linkers for peptide synthesis

Polypeptide synthesis by solid phase technique is more effective when a specific combination of the linker/handle-resin was used. These linkers can help the cleavage of the peptide from the support under specific selected conditions that enable the peptide free from side reactions. The covalent linkage between the growing peptide chain and the polymer support is one of the factors that determine the purity of the peptide. Some of the linkers are shown in Figure 25.

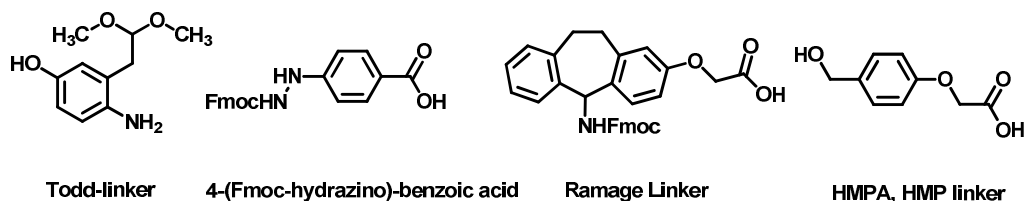


Figure 25: Linkers used for peptide synthesis

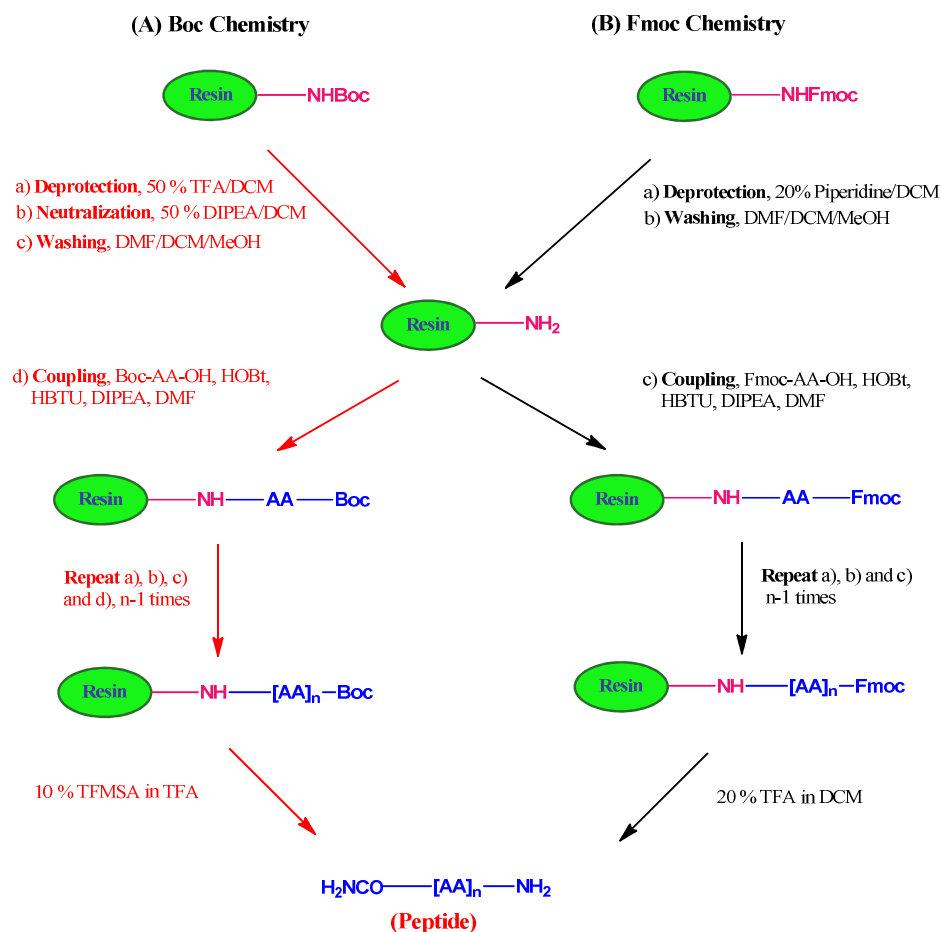
2.1.3b Comparison of Fmoc and Boc chemistry

There are two major protocols for the routine synthesis of peptides by solid phase peptide synthesis Fmoc strategy (base-labile α -amino protecting group) and Boc strategy (acid-labile protecting group). Each method involves different resins and amino acid side-chain protection and consequent cleavage/deprotection steps. In both protocols, the linker group that joins the peptide to the resin is chosen such that the side chain protecting groups and the linker are cleaved in one step at the end of peptide synthesis. Fmoc chemistry is known for generating peptides of higher quality and in greater yield than *t*-Boc chemistry.

The Fmoc SPPS overcomes the limitation of repetitive TFA acidolysis of Boc-group deprotection that could lead to alternation of sensitive peptide bonds as well as acid catalyzed side reactions. For Fmoc deprotection, the growing peptide is subjected to mild base treatment using piperidine/diethylamine during Fmoc deprotection and 20% TFA in DCM is required only for the final cleavage and deprotection of peptidyl resin. By contrast, the cleavage and deprotection in Boc strategy requires the use of dangerous HF or TFA/TFMSA. The removal of a Boc group with TFA yields a positively-charged α -amino group, whereas the removal of an Fmoc group yields a neutral α -amino group. The steric hindrance of the positively charged α -amino group limits the formation of secondary structure on the resin. Finally, the Fmoc method is considered orthogonal, since α -amino group deprotection is with base, while the final cleavage from the resin is with acid. The Boc method utilizes acid for both

deprotection and cleavage from the resin. Scheme 3 shows the general protocol for solid phase peptide synthesis of Fmoc strategy (base-labile alpha-amino protecting group) and Boc strategy (acid-labile protecting group).

Scheme 3: General protocol for SPPS *via* (A) Boc chemistry (B) Fmoc chemistry



Both the Fmoc and Boc methods offer merits and demerits as shown in Table 5.

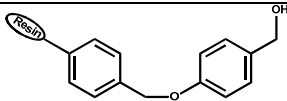
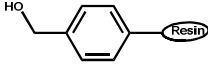
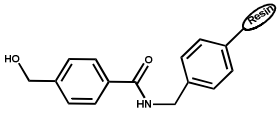
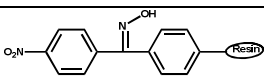
Table 5: Boc *versus* Fmoc chemistry for peptide synthesis

Conditions	Boc	Fmoc
Requires special equipment	Yes	No
Cost of reagents	Low	Higher
Solubility of peptides	Higher	Lower
Purity of hydrophobic peptides	High	May be lower
Problems with aggregation	Less frequently	More frequently
Synthesis time	~20 min/amino acid	~20-60 min/amino acid
Final deprotection	HF/ TFA/TFMSA	20% TFA in DCM
Safety	Potentially dangerous	Relatively safe
Orthogonal	No	Yes

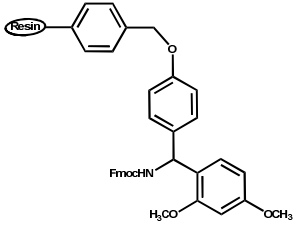
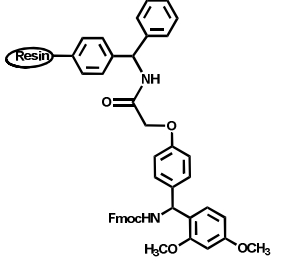
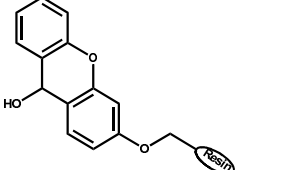
2.2 The Choice of Solid Support (Resin)

Several polymeric supports are available which can be derivatized with functional groups to produce a highly stable linkage to the peptide being synthesized⁸⁰ and to obtain peptides with different functionalities in the terminal carboxyl group (*i.e.*: amide, acid, thioester).⁸¹ Some examples are the *p*-methoxybenzhydrylamine (MBHA), 4-hydroxymethylphenylacetamidomethyl (PAM) and hydroxymethyl functionalized resins used for *t*-Boc/Bzl and the 4-(2',4'-dimethoxyphenyl-aminomethyl)-phenoxyethyl polystyrene (Rink), 2-chlorotrityl chloride and diphenyldiazomethane functionalized resins used for Fmoc/*t*Bu. Table 6 represents different type of functionalized resins.

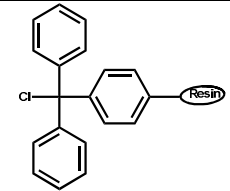
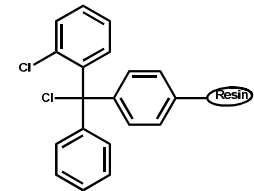
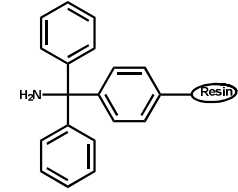
Table 6: OH Functionalized resin

Functionalized Resins	Structure of Resin	Chemistry use	Cleavage conditions	Advantages	Final products
Wang Resin		Fmoc	100% TFA	It can be use for immobilization of acids and phenols, good stability to a variety of reaction conditions	Peptide acids
Hydroxymethyl Resin		Fmoc	70 °C in acetic acid with zinc chloride and 10% TFA in DCM	Immobilization of carboxylic acid, Less cross-linking carbamate linker of hydroxymethyl polystyrene can be used as an equivalent of Cbz, is stable to a wide range of acidic, basic & reductive conditions	Peptide acids
PAM Resin		Boc	TFA/TFMSA	Greater acid stability, suitable for synthesizing medium and large peptides	Peptide acids
Oxime Resin					
Oxime Resin		Boc	NaOH, N ₂ H ₄ , NH ₃ , RNH ₂ , ROH	Oxime linkage not completely stable to TFA, the use of 25% TFA in DCM is recommended for removal of Boc groups	Protected peptide acids, protected peptide hydrazides, peptide amides

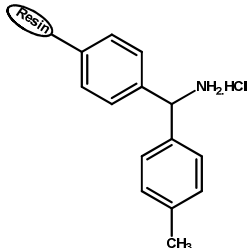
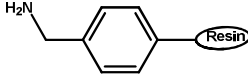
Fmoc-NH Functionalized Resins

Rink amide Resin		Fmoc	20% TFA in DCM	Allow the attachment of various nucleophiles with a wide range of functional groups to the resin, such as amines, alcohols & thiols	Peptide carboxamides
Rink Amide-MBHA Resin		Boc/Fmoc	95% TFA in DCM + scavenger	Use in both Boc as well as Fmoc chemistry, less acid sensitive	Peptide carboxamides
Sieber Resin		Fmoc	TFA/DCM/ EDT (2:98:0.1)	Excellent support for the synthesis of protected peptide amides & carboxamides	Protected peptide carboxamides

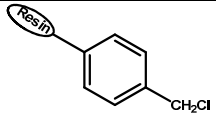
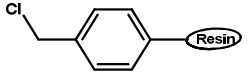
Trityl Resin

Trityl Chloride Resin		Fmoc	1-5% TFA in DCM containing 5% TIS	Highly acid-labile support for the immobilization of alcohols, sugars, amines, hydroxylamines, imidazoles, carboxylic acids & thiols	Protected peptide acid
2-Chlorotrityl Chloride Resin		Fmoc	Acetic acid /TFE/DCM (1,8,8,v/v) HFIP/DCM (1,4, v,v) 0.5 TFA in DCM	Racemization during the loading of the first amino acid (using DIPEA) is avoided with trityl linkers	Protected peptide acid
Amino Trityl Resin		Fmoc	50% TFA in DCM	Synthesis of amides	Peptide carboxamides

NH₂ Functionalized resin

MBHA Resin		Boc/Fmoc	TFA/ TFMSA	The excellent swelling properties, used in both Boc as well as Fmoc chemistry	Peptide carboxamides
Amino methyl Resin		Boc/Fmoc	TFA/ TFMSA	Use in both Boc as well as Fmoc chemistry, scavenger resin in solution phase synthesis to remove excess acids, alkylating agents & other electrophiles	Peptide carboxamides

Cl Functionalized resin

Merrifield Resin		Boc	HF/ TFMSA HBr/TFA	For anchoring protected amino acids by nucleophilic displacement of chlorine give resin-bound benzyl esters	Peptide acids
Macroporous Chloromethyl Resin		Boc	HF/ TFMSA	Matrix for solid phase scavengers is superior to the Merrifield resin when non swelling solvent like acetonitrile used	Peptide acids

2.3 Characterization of Polyproline Structures

Various methods for determination of polyproline conformation are available in literature which includes NMR,⁸² UV, Resonance Raman spectroscopy⁸³ and Chiroptical spectroscopy⁶¹ etc.

2.3.1 Polyproline II conformation by ¹⁵N NMR spectroscopy

Lam and Hsu⁸⁴ reported the NMR of a polyproline II helix. With temperature and sequence corrections on the predicted random coil ¹⁵N chemical shifts, a significant ¹⁵N chemical shift deviation was observed for the model peptide. The ¹⁵N chemical shift differences also correlate well with the molar ellipticities (at 220 nm) of the CD spectra at different temperatures, indicating that the ¹⁵N chemical shift is a sensitive

probe for helices. The average $^3J_{\text{HN}\alpha}$ and $^1J_{\text{C}\alpha\text{H}\alpha}$ values of the model peptide are determined to be 6.5 and 142.6 Hz, respectively, which are consistent with the values calculated from the geometry of helices. They measure the amide ^{15}N chemical shifts and incorporating temperature and sequence effect corrections, which can be used together with $^3J_{\text{HN}\alpha}$ and $^1J_{\text{C}\alpha\text{H}\alpha}$ to differentiate helices from random coils with high confidence. Based on these observations, a strategy was developed for probing left-handed PPII helical structures from other secondary structures.

2.3.2 Circular dichroism

Circular dichroism arises from the fact that chiral molecules interact differently with right and left circularly polarized light. The majority of biological molecules are chiral and CD is useful in determining the PPII helical content of proteins and conformational analysis of nucleic acids and their interaction with other ligands. CD spectra are particularly valuable in characterization of the self assembly of strands to duplex/triplex/tetraplex for higher order structure and distinguish conformation of hybrid structures such as DNA-peptide complexes. CD spectra can be readily used to estimate the fraction of a molecule that is in the alpha-helix, the beta-sheet, the beta-turn structure, the PPII or PPI conformations and random coil (Figure 26).⁸⁵

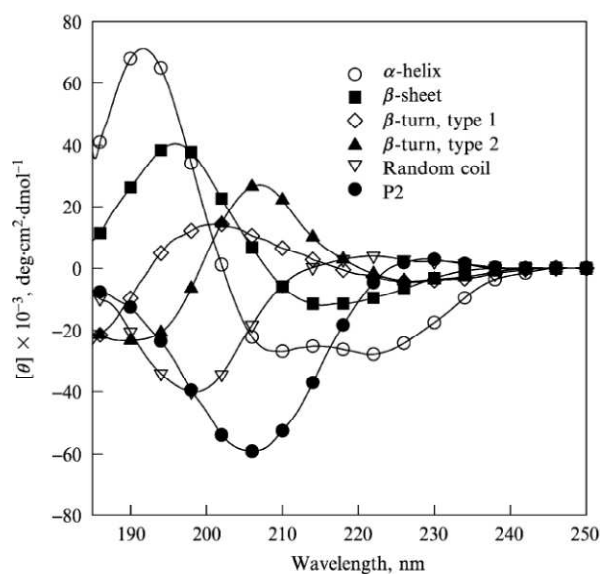


Figure 26: Representative circular dichroism curves corresponding to common secondary structural protein elements.⁸⁵

In general, this phenomenon is seen in absorption bands of any optically active molecule. As a consequence, structure of biological molecules imparts a distinct CD

profile specific to its composite secondary structures. The alpha helix of proteins and the double helix of nucleic acids have CD spectral signatures characteristic of their structures. Thus CD can be used to probe structural changes arising from varying solvent conditions, temperature, pH, ionic strength and interaction with ligands. Circular dichroism is a useful technique to study the kinetics of protein folding and protein–ligand interactions. It is an indirect method useful for measurement of thermodynamics and binding. The far-UV CD spectrum of proteins can reveal important characteristics of their secondary structure.

In addition to measuring in aqueous systems, far-UV CD can be measured in organic solvents e.g. ethanol, methanol, trifluoroethanol (TFE) which induce structural changes of proteins (beta-sheets and alpha helices).

2.3.2a Identification of polyproline II by UV-CD spectroscopy

PPII exhibits typical CD spectra characterized by a negative band around 200-210 nm and a weaker positive band at about 217-227 nm (Figure 27).⁸⁶ This positive band is shifted toward 225 nm for peptides rich in proline/hydroxyproline.⁸⁷ Polypeptides such as poly-Lys and poly-Glu⁸⁸ with ionized side chains considered to adopt random (unordered) conformations at neutral pH show CD spectra with strong negative CD band below 200 nm and a positive one at 218 nm. An additional negative band near 170 nm is also observed in vacuum UV.⁸⁹

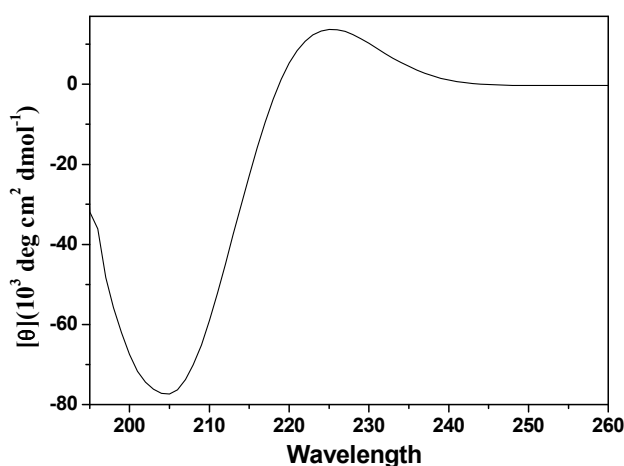


Figure 27: CD spectra of polyproline II helix structure.

The CD spectra with a positive band between 220 to 230 nm and a negative band between 200 to 210 nm are hallmarks of the PPII conformation in which the intensity of the positive band is proportional to the PPII helical content.⁹⁰ The thermal denaturation study of polyproline peptides can be monitored by observing the changes in ellipticity at 222-225 nm with temperature. The thermal denaturation curves uniformly show a decrease in molar ellipticity with increase in temperature for the peptides.

2.3.3 Polyproline II conformation by VCD

VCD (Vibrational Circular Dichroism) spectroscopy mainly measures the difference in absorption of left and right circularly polarized IR radiation by molecular vibrational transitions and therefore it has the advantage of combining the spectroscopic detail of IR with the stereochemical information by CD. Both the spectrum of random coil and the spectrum of PPII conformation exhibit a negative VCD couplet (+, - with decreasing frequency), the difference between them being the shift of the carbonyl stretch of the amide group of PPII, which is 23 cm^{-1} lower than that of random coil conformation.⁹¹

2.3.4 Polyproline II conformation by ROA

Another useful technique in addition to CD has been used for studying polypeptide structures in aqueous solution. The Raman optical activity (ROA) measures the vibrational optical activity by means of a small difference in the intensity of Raman scattering from chiral molecules for right- and left-circularly polarized incident laser light. ROA is complementary to VCD, because it is able to clearly discriminate PPII bands in the presence of superimposed bands from random coil conformation. The Raman spectrum of a polypeptide is subdivided into three main regions of interest:

- 1) $870\text{--}1150\text{ cm}^{-1}$ associated with the vibrations of the backbone $\text{C}\alpha\text{-C}$ and $\text{C}\alpha\text{-N}$;
- 2) $1230\text{--}1350\text{ cm}^{-1}$ containing the amide III region vibrations associated with normal modes of various combinations of the $\text{C}\alpha\text{-H}$ and N-H deformations together with the $\text{C}\alpha\text{-C}$ and $\text{C}\alpha\text{-N}$ stretches;
- 3) $1630\text{--}1700\text{ cm}^{-1}$ associated with C=O stretching modes.

3 Present work: Rationale

The work in this chapter is mainly devoted to the synthesis and structural characterization of a new class of synthetic polyproline helix composed with 4*R/S*-aminoprolines to evaluate the effect of 4-NH₂ substitution on polyproline conformation.

It was previously reported from this laboratory that 4-amino substitution on proline in collagen peptide stabilizes the triple helix⁹² at both acidic and basic pHs. In collagen, the tripeptide repeat unit [Pro-Hyp-Gly]_n has glycine whose amide linkage is involved in an interchain H-bond, leading to the triple helix structure. In contrast, polyproline peptides which lack amide NH are unable to form a triplex *via* interchain H-bonds and end up as a single helix of PPI or PPII type. Unlike other 4-substituents on proline studied so far (OH, F, N₃),^{27,19b} the ionizable 4-NH₂ group is a good probe to examine the pH effects on polyproline conformation. The 4-aminoproline at acidic pH in protonated form (NH₃⁺) may influence the proline puckering *via* its electron withdrawing effects.⁹³ At basic pH, the NH₂ group has the ability to form hydrogen bonding. In view of the importance of polyproline in biology and the influence of 4-substitution, this chapter is intended to study 4-aminoproline polypeptides. Such study would lead to an understanding of the importance of ionizable substituent such as 4-NH₂ group in dictating the PPII conformation and its influence of ring puckering. The effect of 4(*R/S*)-aminoproline in influencing the PPI and PPII conformation in proline polypeptides, thereby modulating the stability of PPII conformation is of importance. A recent computational study^{25, 26} suggested that in 4*S*-aminoproline, intramolecular hydrogen bond is feasible between 4*S*-NH₂ and the carbonyl carbon of same proline moiety. This leads to increase the *trans/cis* ratio of peptide bond, thereby enhancing stability of polyproline conformation.

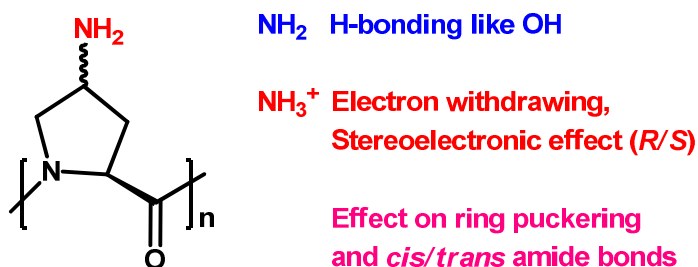


Figure 28: Ionizable amino group influence the proline puckering *via* H-bonding or electronic effect

The specific objectives of this chapter are

- Synthesis of $(2S,4R)$ - N^1 -(Fmoc)-4NH-(*t*-Boc)-aminoproline (**8**) and $(2S,4S)$ - N^1 -(Fmoc)-4NH-(*t*-Boc)-aminoproline monomers.
- Synthesis of 4(*R/S*)-aminoproline homopeptides, 4*R* *Amp*₉ (peptide **1**) and 4*S* *amp*₉ (peptide **2**) by solid phase protocol.
- Synthesis of control proline homopeptides *Pro*₉ (peptide **3**) by solid phase synthesis protocol.
- Cleavage of these peptides from the solid support, purification and characterization.
- Investigation of conformation of these peptides using concentration dependent and temperature dependent CD spectroscopy.
- Evaluate the effect of different pH, solutes and solvent on these peptides by using CD spectroscopy.

3.1 Results

3.1.1 Synthesis of fully protected $(2S,4R)$ and $(2S,4S)$ aminoproline monomers

To achieve the synthesis of peptides (**1**, **2**), the monomers $(2S,4R)$ - N^1 -(Fmoc)-4NH-(*t*-Boc)-aminoproline **8** and $(2S,4S)$ - N^1 -(Fmoc)-4NH-(*t*-Boc)-aminoproline **12** with a combination of N^1 -Fmoc and N^4 -Boc protection were chosen. The peptide synthesis can be achieved by choosing a combination of alternative N-Boc and N-Fmoc protection. The peptide **3** synthesized from Fmoc proline **17**.

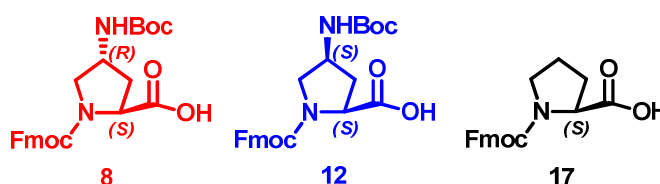


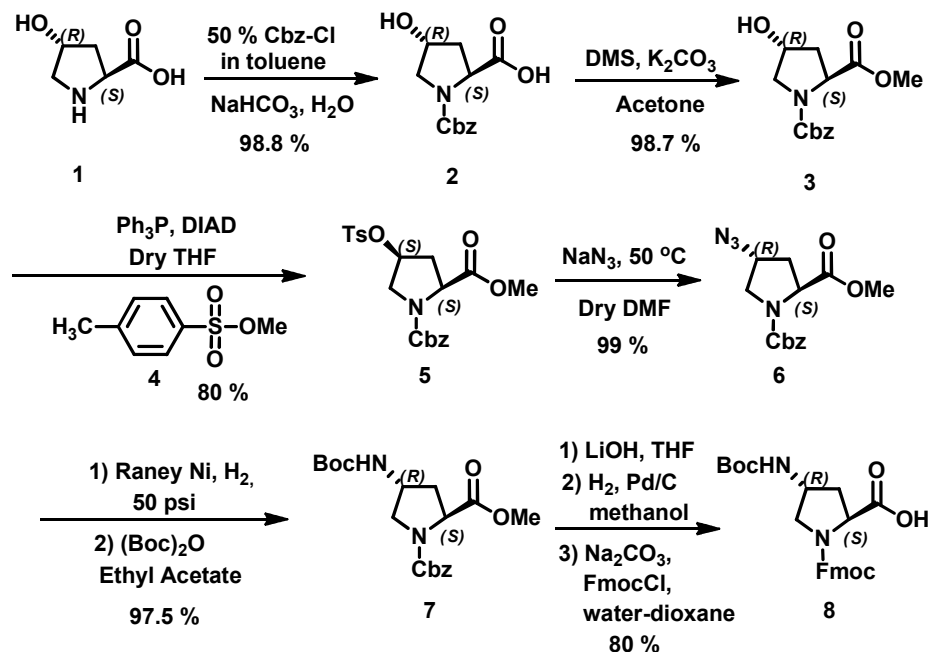
Figure 29: Structures of monomers

3.1.1a Synthesis of $(2S,4R)$ - N^1 -(Fmoc)-4NH-(*t*-Boc)-aminoproline (**8**)

The synthesis of orthogonally protected $(2S,4R)$ - N^1 -(Fmoc)-4NH-(*t*-Boc)-aminoproline (**8**) monomer was achieved in eight steps from the naturally occurring *trans*-4-hydroxyproline **5** (Scheme 4). The reaction of *trans* 4*R*-hydroxyproline with benzyloxycarbonyl chloride in water/dioxane, in the presence of Na_2CO_3 , yielded N-

benzyloxycarbonyl-4-hydroxyproline **2**. The appearance of aromatic ($\delta = 7.2-7.3$ and benzylic $\delta = 5.1$) signals in ^1H NMR spectrum of compound **2** confirmed its structure.

Scheme 4



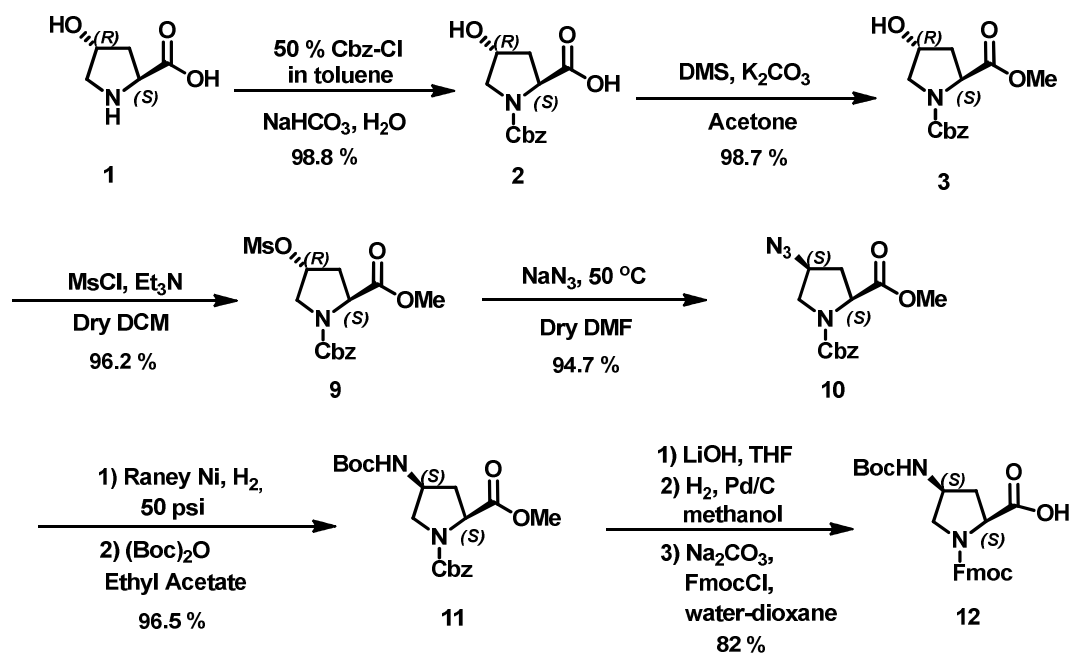
The compound **2** upon treatment with DMS in acetone/K₂CO₃ yielded the corresponding methyl ester **3**, which showed two signals at $\delta = 3.75$ and 3.54 in the ^1H NMR spectrum corresponding to the methyl (-CH₃) group of the ester (minor and major rotamers respectively). The 4*R*-OH group of compound **3** was converted into 4*S*-O-tosylate **5** by using Mitsunobu reaction with DIAD/PPh₃ and methyl-*p*-toluenesulfonate (**4**), the reaction accompanied by an inversion of configuration at C-4. The reaction of 4*S*-O-tosylate **5** with NaN₃ in DMF resulted in a second inversion of stereochemistry at C-4 to yield (2*S*, 4*R*)-N¹-Cbz-azidoproline methyl ester **6** supported by appearance of ^1H NMR signals at $\delta 7.23-7.29$ for aromatic protons, at $\delta 4.93-5.19$ for benzylic protons and at $\delta 4.35-4.43$ for C ^{γ} proton. A peak at 2108 cm^{-1} characteristic of azide was seen in IR spectrum of compound **6** which was selectively reduced to the corresponding amine using Raney-Ni without affecting the N¹-Cbz group. This was followed by *in situ* Boc protection, with Boc anhydride to yield N⁴-*t*-Boc compound **7**. The structure was supported by the appearance of ^1H signal at $\delta 1.40$ for *t*-Boc protons and at $\delta 4.70$ for NH proton. The methyl ester of compound **7** was hydrolysed by using LiOH in THF and water. The deprotection of N¹-Cbz group was achieved by hydrogenation using 10% Pd/C in methanol. The disappearance of

aromatic signals of Cbz group in ^1H NMR and the appearance of ninhydrin positive spot on TLC indicated complete deprotection of Cbz group. The N^1 -amine was protected as Fmoc, by reaction with Fmoc-Cl to yield the fully protected 4*R*-aminoproline **8**. The appearance of ^1H signals in the aromatic region typical of Fmoc group (δ 7.28-7.37 and 7.53-7.74) and mass peaks at 453 (M + H) in the LCMS confirmed the structural integrity of the monomer **8**.

3.1.1b Synthesis of (2*S*,4*S*)- N^1 -(Fmoc)-4NH-(*t*-Boc)-aminoproline (**12**)

The synthesis of orthogonally protected (2*S*,4*S*)- N^1 -(Fmoc)-4NH-(*t*-Boc)-aminoproline (**12**) monomer was achieved in a similar way in eight steps starting from the naturally occurring *trans*-4-hydroxyproline **1** (Scheme 5). The reaction of *trans*-4*R*-hydroxyproline with benzyloxycarbonylchloroformate in water/dioxane in the presence of Na_2CO_3 , yielded N-benzyloxycarbonyl-4-hydroxyproline **2**. This was confirmed by the appearance of aromatic (δ = 7.2-7.3 and benzylic δ = 5.1) signals in ^1H NMR spectrum. Compound **2** upon treatment with DMS in acetone/ K_2CO_3 yielded the corresponding methyl ester **3**. The 4-OH group in **3** was converted to the corresponding mesyl derivative **9** by treatment with methanesulfonylchloride. The treatment of mesylate **9** with NaN_3 in DMF yielded the 4*S* azide compound **10**, the reaction accompanied by inversion of stereochemistry at C-4. A peak at 2108 cm^{-1} corresponding to the azide function seen in the IR spectrum confirmed the conversion of mesyl to azide. The azide was selectively reduced to the corresponding amine using Raney-Ni as a catalyst without affecting the N^1 -Cbz group. The 4-aminogroup was *in situ* protected using Boc anhydride to yield N^4 -*t*-Boc compound **11** as shown by the appearance ^1H signal at δ 1.44 for *t*-Boc protons and at δ 5.2 for NH proton. The methyl ester of compound **11** was hydrolysed by using LiOH in THF and water. The N^1 -benzyloxycarbonyl group of compound **11** was removed by hydrogenation with 10% Pd-C as a catalyst to yield free amine product. The N^1 amine protection was done by using Fmoc-Cl to yield the fully protected 4*S*-aminoproline monomer **12**. The appearance of ^1H signals in the aromatic region indicative of Fmoc group (δ 7.25-7.38 and 7.52-7.73) and mass peaks at 453 (M + H) in the LCMS confirmed the structural integrity of the monomers **12**.

Scheme 5

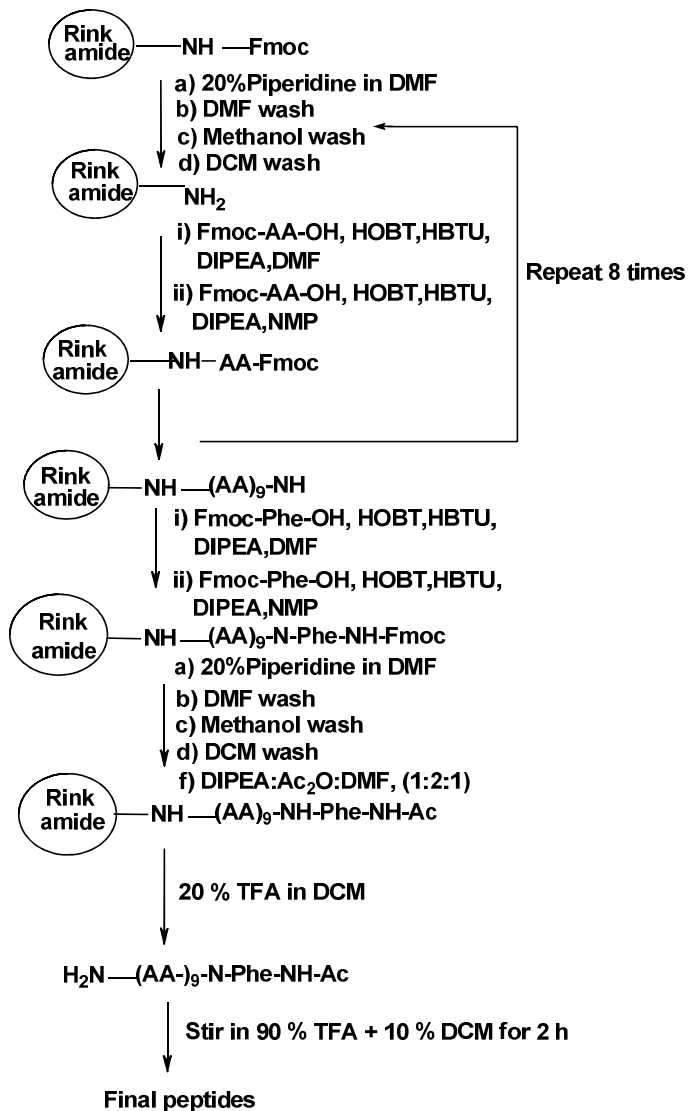


3.1.1c Synthesis of homo oligopeptides 1-3 by using solid phase protocol

The target peptides were synthesized by manual solid phase synthesis on the readily available rink amide resin using standard Fmoc chemistry (Scheme 6), in which the cleavage from resin directly leads to the peptide with C-terminal amide. The commercially available Kieselghur supported N,N-dimethylacrylamide resin (with Rink-amide linker) was used. The resin bound Fmoc group was cleaved with 20% piperidine in DMF and the monomers were coupled as free acids using *in situ* activation procedure with amino acid (3 eq.), HBTU as a coupling reagent and HOBt and DIPEA as catalyst and racemisation-suppressant. The coupling reaction was repeated by using *N*-Methyl-2-pyrrolidone (NMP) as solvent. The deprotection reactions were done by using 20% piperidine in DCM and monitored using qualitative Chloranil test⁹⁴ for iminoacids. The terminal amino group of the final peptide was capped with Ac₂O and the peptide was cleaved from the resin using 20% TFA in DCM. As the *t*-Boc removal requires stronger acidic conditions, the deprotection of sidechain *t*-Boc on Amp residues of the peptide was carried out with 90% TFA in DCM. Since *t*-Butyl cation formed during the deprotection of *t*-Boc from the final peptide can lead to N-alkylation of the amines, (0.1% TIS (triisopropylsilane) was used as a scavenger) to prevent side reactions during peptide-cleavage and *t*-Boc deprotection. The N-terminal acetylated and C-terminal amidated peptides were

purified on semi-preparative RP-18 HPLC column using water-acetonitrile gradient containing 0.1% TFA.

Scheme 6: Solid phase synthesis for **1-3** peptide synthesis



The oligopeptides **1-3** were synthesized from appropriate N-Fmoc-protected monomers (2*S*,4*R*)-N¹-(Fmoc)-4NH-(*t*-Boc)-aminoproline (**8**), (2*S*,4*S*)-N¹-(Fmoc)-4NH-(*t*-Boc)-aminoproline (**12**) and the readily available N-Fmoc proline monomers assembled on the solid phase is shown in Scheme 7. All oligomers were purified by HPLC and characterized by MALDI-TOF. The purity of the peptides as determined using analytical RP-18 HPLC was found to be greater than 98%. The structural integrity of the peptides was further confirmed by MALDI-TOF mass spectrometry data which agreed closely with the calculated values (Table 7)

Scheme 7: Synthesis of homooligopeptides 1-3 from proline and 4*R/S*-aminoproline monomer

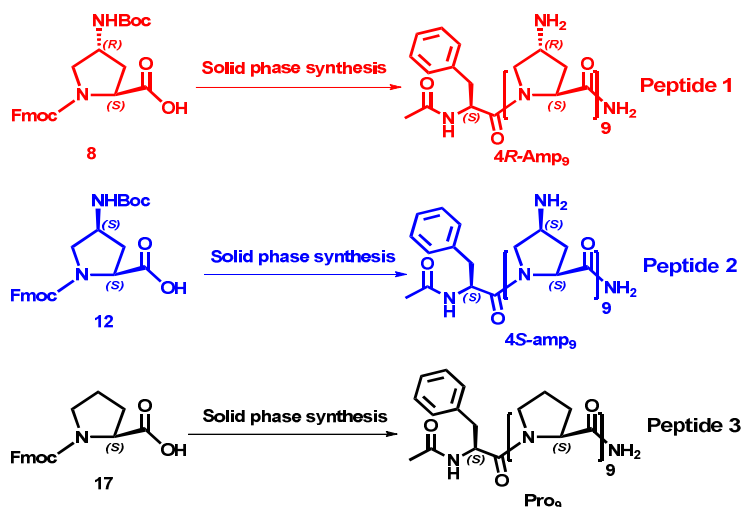


Table 7: Calculated and observed masses for peptides 1-3

Peptide	Mol. Formula	Mass (calculated)	Mass (observed)
1 (4 <i>R</i> -Amp ₉)	C ₅₆ H ₈₆ N ₂₀ O ₁₁	1214	1214.56 [M] ⁺
			1236.64 [M + Na] ⁺
			1252.74 [M + K] ⁺
2 (4 <i>S</i> -amp ₉)	C ₅₆ H ₈₆ N ₂₀ O ₁₁	1214	1214.53 [M] ⁺
			1236.78 [M + Na] ⁺
			1252.74 [M + K] ⁺
3 (Pro ₉)	C ₅₆ H ₇₇ N ₁₁ O ₁₁	1078	1080.48 [M + 1] ⁺
			1101.52 [M + Na] ⁺
			1117.45 [M + K] ⁺

All peptides synthesized were end-capped as amides to eliminate the possible interchain N- and C-terminal electrostatic repulsions.^{19f} They were used to evaluate the actual effect of 4*R/S* amino group on polyproline conformation.

3.1.2 Determination of the peptide concentration in stock solution

Determination of the exact concentration of the peptide solutions usually poses a problem. Even after several hours of drying under vacuum, peptides retain significant amounts of water, the amount of which varies with drying conditions and times.⁹⁵ Basic amino acids retain counter ions such as acetyl and trifluoroacetyl, arising from additives used during cleavage and purification procedures. Methods such as

quantitative amino acid analysis has been employed to determine the exact concentration of the peptide solutions. However these methods are tedious and the concentration of peptide stock solution may change with time. To discriminate these problems in the present study, the amino acid phenylalanine (Phe) was included at the N-terminal side for all peptides. Since it has aromatic side chain, the concentration of peptide stock-solutions can be determined directly by using UV-absorbance ($\epsilon_{259} = 2 \times 10^2 \text{ m}^{-1} \text{ cm}^{-1}$). All peptides contained phenylalanine at N-terminus and hence any effect on polyproline helix stability from this residue is same for all peptides.

3.1.3 pK_a of 4-amino group of 4*R/S*-Aminoproline monomers and oligomers

Both (2*S*,4*R*) and (2*S*,4*S*)-aminoproline monomers (**8**, **12**) and their homooligomers (**1**, **2**) contain free exocyclic amino group, which in free form get easily protonated under physiological conditions. In order to evaluate the pH dependence of PP II helical stability, it is necessary to ascertain the protonation state through the pK_a of exocyclic amino group of both 4*R*-Amp and 4*S*-amp monomers and their homooligomers. For this purpose, pH changes were determined by titrating⁹⁶ α -N-protected 4*R*-Amp and 4*S*-amp monomers and their homooligomers with 0.01N aqueous NaOH.

The direct reading of pH meter upon titration was plotted against the volume of added NaOH (Figures 29 and 30). This plot shows two sharp pH transitions, one corresponding to the deprotonation of α -carboxylic acid group and a second corresponding to deprotonation of exocyclic NH_3^+ group. The maxima of each transition correspond to the pK_a of the corresponding ionizations. Table 8 shows the pK_a of 4*R/S*-Aminoproline monomers and their homopeptides.

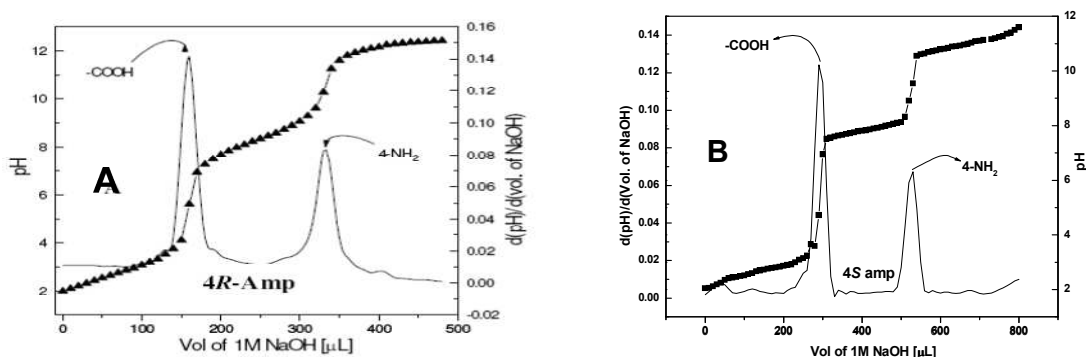


Figure 29: pK_a titration with 0.01 M solution of NaOH for monomers (A) 4*R*-Amp, (B) 4*S*-amp

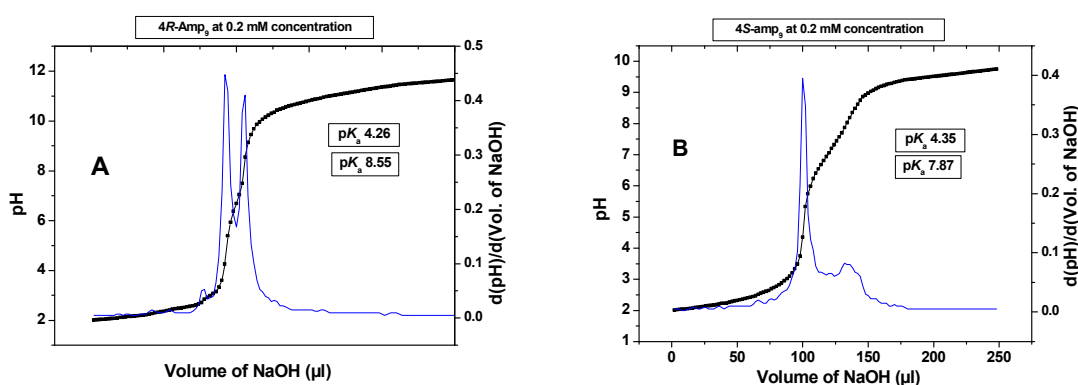


Figure 30: pK_a titration with 0.02 M solution of NaOH for peptides (A) 4*R*-Amp₉ (**1**), (B) 4*S*-amp₉ (**2**)

Table 8: The pK_a of 4*R/S*-aminoproline monomers and their homopeptides

Monomers/peptides	pK_a	ΔpK_a (<i>R-S</i>)
4 <i>R</i> -Amp (8)	10.2	0.9
4 <i>S</i> -amp (12)	9.3	
4 <i>R</i> -Amp ₉ (peptide 1)	8.55	0.7
4 <i>S</i> -Amp ₉ (peptide 2)	7.83	

Both 4*R*-Amp and 4*S*-amp monomers display two sharp pH transitions corresponding to ionization of carboxylic acid and the protonation of 4-NH₂ group. Amongst these, the exocyclic 4-amino group has significance difference in the pK_a values. The pK_a value of amino group in 4*R*-Amp monomer is 10.2 while the pK_a value of 4*S*-amino group monomer is 9.3, thus 4*R*-amino group being more basic than 4*S*-amino group. Due to this difference in pK_a , the 4-NH₂ groups in these monomers undergo protonation to different extents at physiological pH. The peptides 4*R*-Amp₉ (**1**) and 4*S*-amp₉ (**2**) also show two sharp pH transitions and their exocyclic 4-amino group has significant differences in the pK_a values. The pK_a value of 4-amino group in 4*R*-Amp₉ is 8.55 while the pK_a in 4*S*-amp₉ is 7.83. The 4-amino groups in monomer are thus more basic (higher pK_a) than the corresponding oligomers (lower pK_a).

3.2 CD Spectroscopic Studies

There are several methods for determination of polyproline conformation known in literature which include NMR, UV, Resonance Raman spectroscopy and chiroptical spectroscopies. In present study the circular dichroism spectroscopy was used to investigate the conformational behavior of peptides **1-3**.

3.2.1 Identification of PPII by UV-CD spectroscopy

In order to investigate the conformational nature of peptides 4*R*-Amp₉ (peptide 1), 4*S*-amp₉ (peptide 2) and *Pro*₉ (peptide 3), the CD spectral analyses were carried out as a function of temperature, pH, urea, salt and solvents (buffer, trifluoroethanol and *n*-propanol). The CD spectra of all three peptides 1-3 were recorded with 100 μM peptide concentration at pH 7.2 in Na-phosphate buffer and shown in Figure 31.

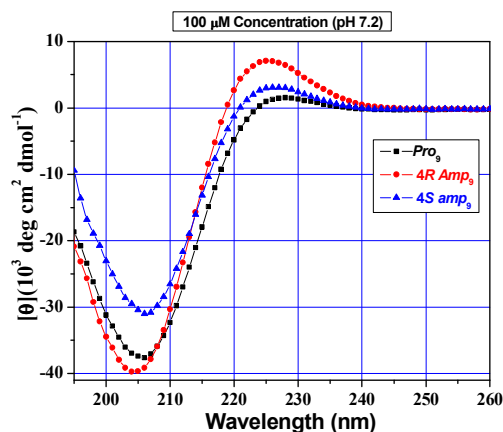


Figure 31: CD profiles of peptides 1, 4*R*-Amp₉ (●), peptide 2, 4*S*-amp₉ (▲) and peptide 3, *Pro*₉ (■), all at 100 μM (pH 7.2).

The CD spectra of peptides 1-3 (Figure 31) has a positive band between 220 to 230 nm and a negative band between 200 to 210 nm which are the hallmarks of the PPII conformation wherein the intensity of the positive band is proportional to the PPII helical content.^{90,35} The conformational averaging resulting from the presence of other local structure can lead to negative ellipticity in this region. It is seen from Figure 31 that both peptide 1 (4*R*-Amp₉) and peptide 2 (4*S*-amp₉) oligomers have more positive intensity in the region (225-227 nm) than the unsubstituted proline oligomer *Pro*₉ (3), with 4*R*-Amp₉ having maximum PPII helix content. The intensity of the positive band at 225-227 nm is seen to decrease in the order 4*R*-Amp₉ 1 > 4*S*-amp₉ 2 > *Pro*₉ 3. The peptide 1 (4*R*-Amp₉) also shows a more negative ellipticity (205 nm) than peptide 2 (4*S*-amp₉) and peptide 3 (*Pro*₉).

3.2.2 Concentration dependent CD spectroscopy for peptides 1-3

In collagen peptides, glycine NH involved in interchain hydrogen bonding, which leads to formation of the triple helix structure. In contrast, polyproline peptides lack amide NH and hence are unable to form a triplex *via* interchain H-bonds, ending up as

a single helix of PPI or PPII type. Formation of triple-helical structure is a concentration dependent phenomenon and as shown by Goodman *et al.*⁹⁷ a single stranded chain shifts to the triple-helical conformation with increase in concentration in solution. The percentage of triple-helical structure has maximum when the concentration is greater than its critical triple-helical concentration. The magnitude of the ratio of positive to negative band intensity in the CD spectra of collagen peptides ($R_{p/n}$) has been proposed to quantitate the extent of triple-helical strength. In peptides **1** and **2**, the presence of amino group may influence the formation of secondary structures. Figure 32 (A-C) shows the CD spectra of peptides **1-3** recorded at 25 °C at pH 7.2 (Na-phosphate buffer) in the concentration range of 100–500 μ M.

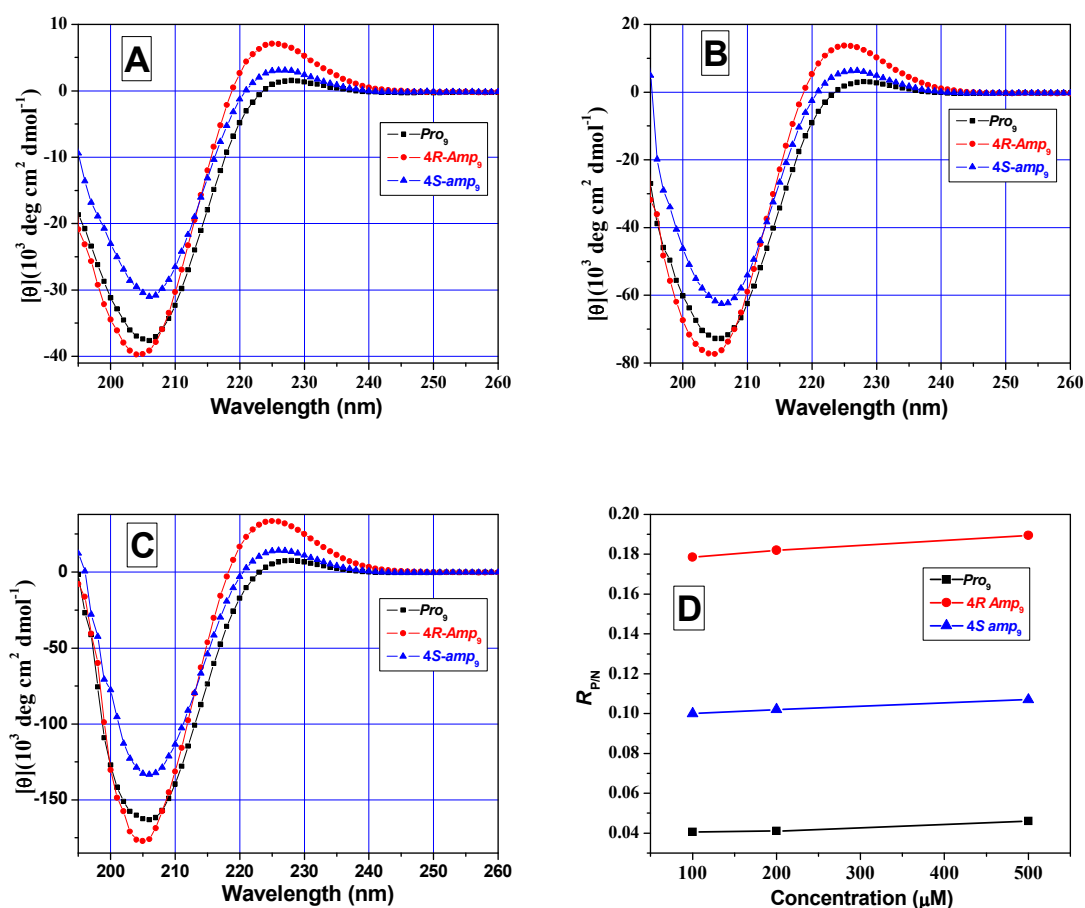


Figure 32: (A) CD profiles of peptides **1-3** at 100 μ M concentration (B) CD profiles of peptides **1-3** at 200 μ M concentration (C) CD profiles of peptides **1-3** at 500 μ M concentration (D) Ratio of intensity of positive to negative band intensity in the CD spectra.

In the entire concentration range, peptides **1-3** show positive maxima at 224-225 nm and negative band at 203-205 nm with increasing concentration. Importantly all

the spectral traces pass through an isoelliptic point around 213-215 nm. The ratio of positive to negative intensity of CD spectra increases with higher concentration in a linear manner. Figure 32D shows a plot of $R_{p/n}$ values derived from these spectra against concentration of peptides **1-3** at pH 7.2. The $R_{p/n}$ value increase only slightly with increase in concentration linearly without any saturation.

3.2.3 Effect of protonation of the 4-amino group on PPII helical content

The effect of protonation of the 4-amino group on PPII helical content was examined by the CD spectra of peptides **1-3** recorded at different pH values (4.0-10.0). All peptides were used at the same concentration (100 μ M) for CD experiments. Buffers used for CD measurement at different pH's were Na-acetate buffer for pH (4.0-5.0) Na-phosphate buffer for pH (7.2) and Borate buffer for pH (9.0-10.0).

3.2.3a Effect of pH on peptide 4R-Amp₉ (I): The polyproline II helix is the only secondary structure known to have a positive band in the wavelength range of 218-228 nm. The intensity of the positive band is taken to be proportional to the PPII helical content.⁹⁰ The CD spectra for peptide **1** (4R-Amp₉) as a function of pH are shown in Figure 33A. The intensity of positive band in CD spectra at 225 nm is plotted as a function of pH and shown in Figure 33B.

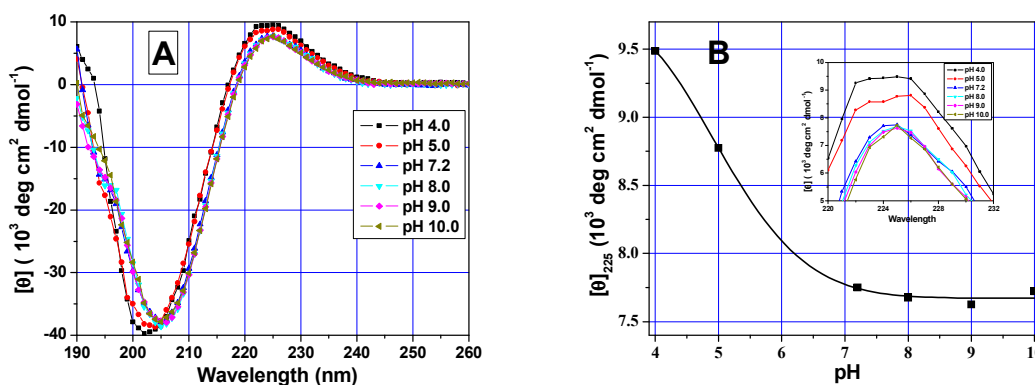


Figure 33: (A) CD profiles of peptide **1** (4R-Amp₉) as a function of pH (B) Intensity of the positive band of CD spectra of peptide **1** (4R-Amp₉) as a function of pH (Inset shows zoom positive region of CD spectra at different pHs).

The CD spectra of peptide **1** (4R-Amp₉) at different pHs show that in acidic pH (4.0-5.0), it has more PPII helix content than at neutral pH (7.2). As the pH increases, the PPII helix content decreases upto pH 7.2 and remains constant in basic pH 8.0-10.0

range. The positive ellipticity at 225 nm for peptide **1** (4*R*-Amp₉) decreased by >15% with increasing pH upto 7.2 and did not change further until pH 10.0. The data suggests that in acidic pH (4.0-5.0) at which the amino groups are protonated (NH₃⁺), peptide **1** (4*R*-Amp₉) stabilizes the PPII form compared to that in unionized form (pH 8.0-10.0).

3.2.3b Effect of pH on peptide 2 (4*S*-amp₉): The CD spectra for peptide **2** (4*S*-amp₉) as a function of pH are shown in Figure 34A. The intensity of positive band in CD spectra at 225 nm plotted as a function of pH is shown in Figure 34B. The data show an opposite behavior to that of peptide **1** (4*R*-Amp₉). In acidic pH (4.0-5.0) it has less PPII helix content than at neutral pH (7.2) and as pH increases, the PPII helix content also increases upto pH 9.0-10.0 and thereafter remains constant. The positive intensity in CD spectra for peptide **2** (4*S* amp₉) at 225 nm is enhanced in the pH range 4.0-10.0 in a sigmoidal fashion. At acidic pH, the PPII helicity of peptide **2** (4*S*-amp₉) was low (20% of peptide **1** (4*R*-Amp₉)) but increased by 2-fold at pH 10.0. The data suggests that at acidic pH (4.0-5.0), the protonated NH₃⁺ form of peptide **2** (4*S*-amp₉) does not stabilize the PPII helix and in the unionized form (pH 8.0-10.0), the stability of 4*S*-amp₉ **2** increases with increase in pH.

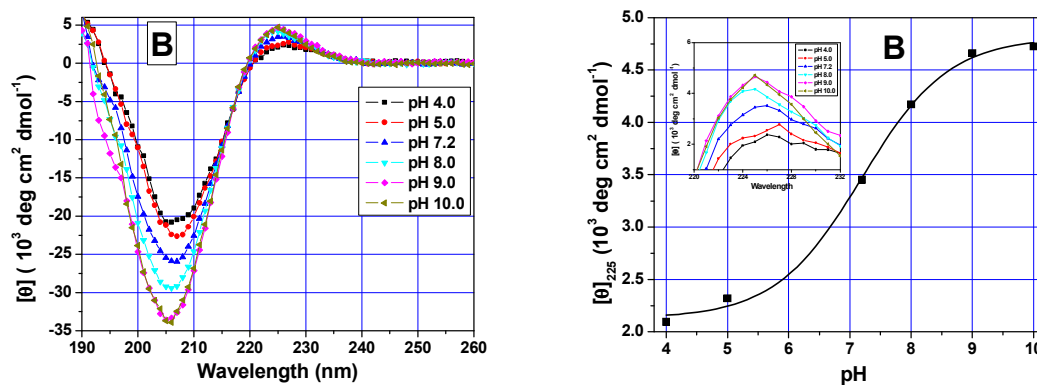


Figure 34: (A) CD profiles of peptide **2** (4*S*-amp₉) as a function of pH (B) Intensity of the positive band of CD spectra of peptide **2** (4*S*-amp₉) as a function of pH (Inset shows zoom positive region of CD spectra at different pHs).

3.2.3c Effect of pH on peptide 3 (Pro₉): The CD spectra for peptide **3** (Pro₉) as a function of pH are shown in Figure 35A. The intensity of positive band in CD spectra at 227 nm plotted as a function of pH shown in Figure 35B. The positive ellipticity of peptide **3** (Pro₉) remained constant with pH. The data suggests that PPII helicity is not

affected in case of peptide **3** (*Pro*₉) with pH change and the altered behaviour of 4*R*-*Amp*₉ and 4*S*-*amp*₉ as a function of pH is entirely due to protonation of ionizable substituent at C4.

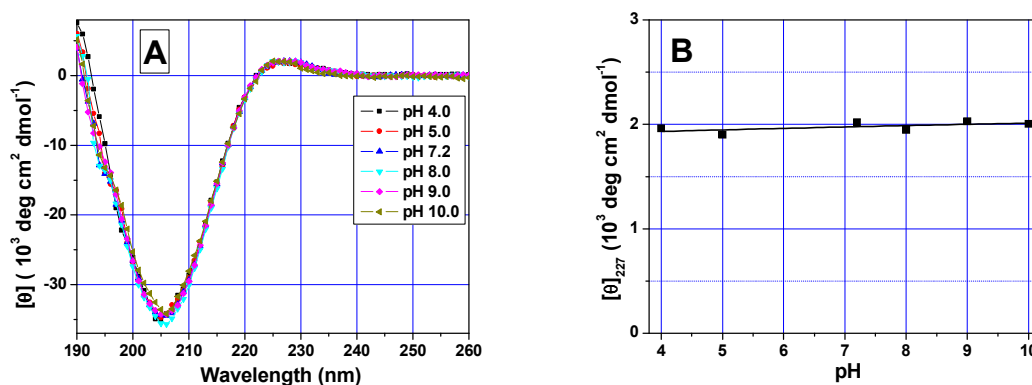


Figure 35: (A) CD profiles of peptide **3** (*Pro*₉) as a function of pH (B) Intensity of the positive band of CD spectra of peptide **3** (*Pro*₉) as a function of pH.

3.2.3d Comparison of peptides 1-3 as a function of pH

Figure 36 shows the comparative behavior of peptides **1-3** as a function of pH. The PPII content of peptide **1** (4*R*-*Amp*₉) in acidic pH (4.0-5.0) is higher than that at neutral and basic pH. The positive ellipticity for peptide **1** (4*R*-*Amp*₉) decreases with increasing pH and does not change in basic pH. The peptide **1** (4*R*-*Amp*₉) shows highest PPII helical content among the peptides **1-3** at acidic pH range. In the case of peptide **2** (4*S*-*amp*₉), the PPII helicity at acidic pH was low (20% of peptide **1** (4*R*-*Amp*₉)) but doubled at basic pH (9.0-10.0). The positive ellipticity of peptide **3** (*Pro*₉) remained constant with pH.

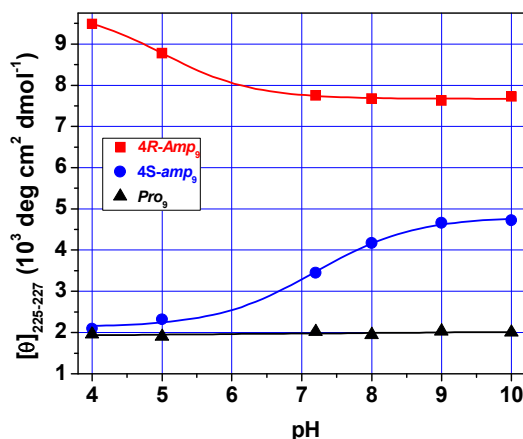


Figure 36: Intensity of the positive band of CD spectra of peptides **1-3** as a function of pH.

These results indicate that at acidic pH (4.0-5.0) peptide **1** (*4R-Amp*₉) is stabilized in PPII form, whereas peptide **2** (*4S-amp*₉) destabilized the PPII conformation. At basic pH, the stability of peptide **1** (*4R-Amp*₉) remained constant while that of peptide **2** (*4S-amp*₉) increased at higher pH. This suggests a combined role of both stereochemistry and protonation of the 4-amino group in modulating the PPII helicity of 4(*R/S*)-aminoproline polypeptides **1** and **2**.

3.2.4 CD dependent thermal denaturation of peptides 1-3 at different pHs

Intramolecular hydrogen bonding between the side chain amine and backbone carbonyl oxygen of amide is known to stabilize the PPII conformation in charged peptides.²¹ To examine the factors that stabilize the PPII helix in ionizable 4(*R/S*)-aminoproline polypeptides, pH-dependent thermal stability (T_m) were carried out for peptides **1-3** (Figure 37-39).

3.2.4a Peptide 1 (*4R-Amp*₉): Figure 37A shows the CD spectra recorded for peptide **1** (*4R-Amp*₉) in different pHs at temperature 20 °C. Figure 37A shows temperature dependent CD spectra for peptide **1** (*4R-Amp*₉) recorded from 5 to 90 °C at pH 7.2. An isoelliptic point is observed around 212-214 nm in CD spectra at different pHs. The CD-thermal denaturation plot of molar ellipticity at 223-225 nm *versus* temperature at different pHs for peptide **1** is shown in Figure 37C. Figure 37D shows the first derivative curve obtained from the sigmoidal fit of data in Figure 37C and the T_m values for peptide **1** obtained from minima in plots of Figure 37D is shown in Table 9. The T_m values indicate that at acidic pH, peptide **1** (*4R-Amp*₉) forms a slightly more stable polyproline II helix structure and as the pH increases, the stability of the peptide **1** (*4R-Amp*₉) slightly decreases till pH 7.2 and remains constant thereafter at higher pH. The change in stability with experimental errors is only marginal.

Table 9: The T_m values for peptide 1 (*4R-Amp*₉) at different pHs

pH	T_m (°C)
4.0	52
5.0	51
7.2	50.5
9.0	50
10.0	50

T_m values are (± 0.5 °C)

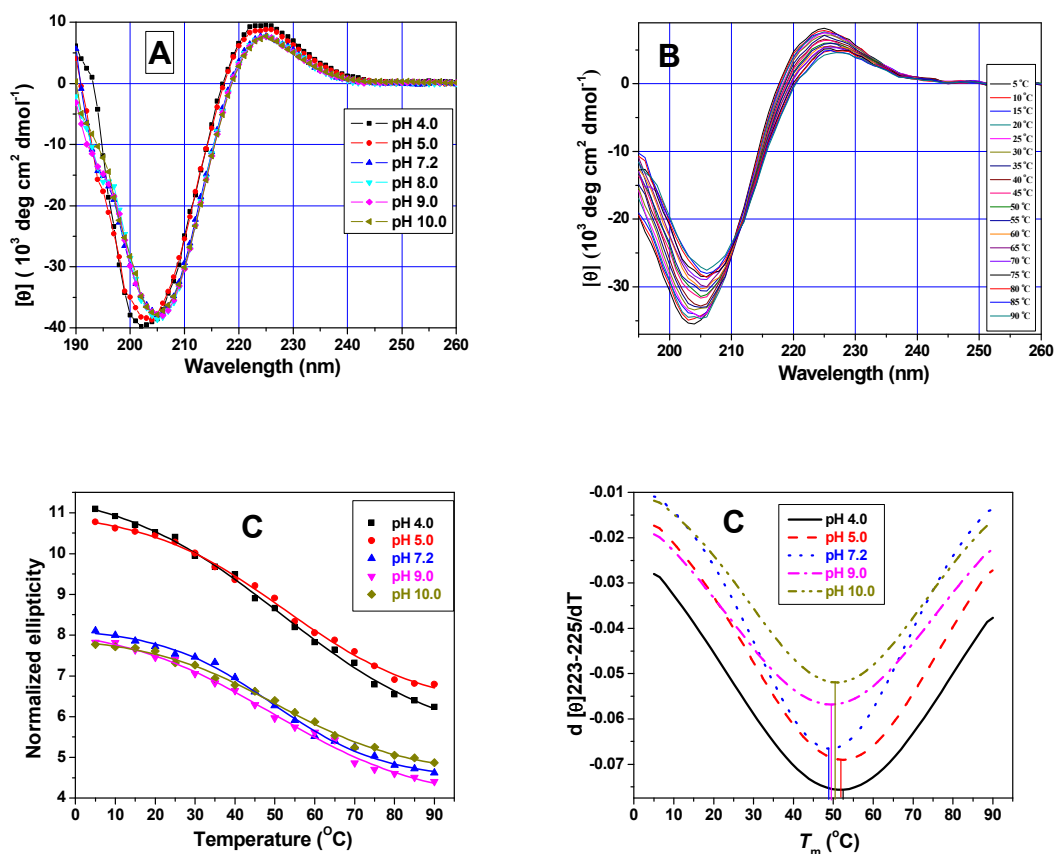


Figure 37: (A) CD spectra of peptide **1** (*4R-Amp*₉) recorded at different pHs (B) CD spectra at 5-90 °C for peptide **1** (*4R-Amp*₉) at pH 7.2. (C) CD-thermal denaturation plot of molar ellipticity at 223-225 nm vs. temperature at different pHs for peptide **1** (*4R-Amp*₉). (D) First derivative curve of peptide **1** (*4R-Amp*₉)

3.2.4b Peptide 2 (*4S-amp*₉): Figure 38A shows the CD spectra recorded for peptide **2** in different pHs at 20 °C. Figure 38B shows the CD spectral changes of peptide **2** (*4S-amp*₉) upon heating in solution from 5 to 90 °C at pH 10.0. An isoelliptic point is observed around 212-214 nm in CD spectra at different pHs. The CD-thermal denaturation plot of molar ellipticity at 225-227 nm *versus* temperature at different pHs for peptide **2** is shown in Figure 38C. The first derivative curve obtained from the sigmoidal fit of data is shown in Figure 38D. The T_m values for peptide **2** obtained from minima in plot of Figure 37D at different pHs is shown in Table 10. The values indicate that at acidic pH peptide **2** (*4S-amp*₉) forms a less stable polyproline II helix structure and as the pH increases the stability of PPII helix also increases.

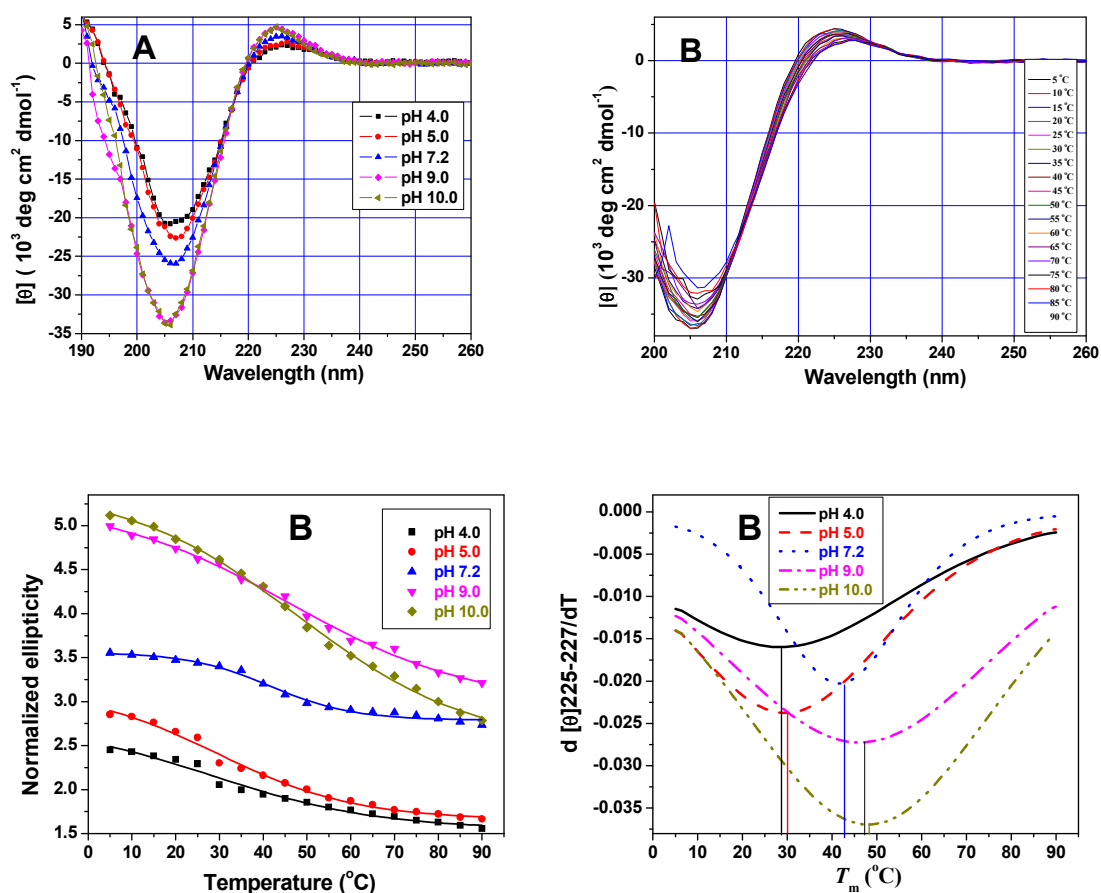


Figure 38: (A) CD spectra of peptide 2 (*4S-amp₉*) recorded at different pHs (B) CD spectra at 5-90 °C for peptide 2 (*4S-amp₉*) at pH 10.0. (C) CD-thermal denaturation plot of molar ellipticity at 225-227 nm vs. temperature at different pH for peptide 2 (*4S-amp₉*) (D) First derivative curve of peptide 2 (*4S-amp₉*).

Table 10: The T_m values for peptide 2 (*4S-amp₉*) at different pHs

pH	T_m (°C)
4.0	28
5.0	30
7.2	43
9.0	47
10.0	48

T_m values are (± 0.5 °C)

3.2.4c Peptide 3 (*Pro₉*): The CD-thermal denaturation plot of molar ellipticity at 227 nm versus temperature at different pHs for peptide 3 is shown in Figure 39A. Figure 39B shows the first derivative curve obtained from the sigmoidal fit of data of Figure 39A and the T_m values for peptide 3 obtained from minima in plot of Figure 39B. The data indicate that the T_m of peptide 3 (*Pro₉*) remain constant with change in pH as

expected. At all pHs, peptide **3** forms a moderately stable polyproline II helix structure with T_m of 37 °C.

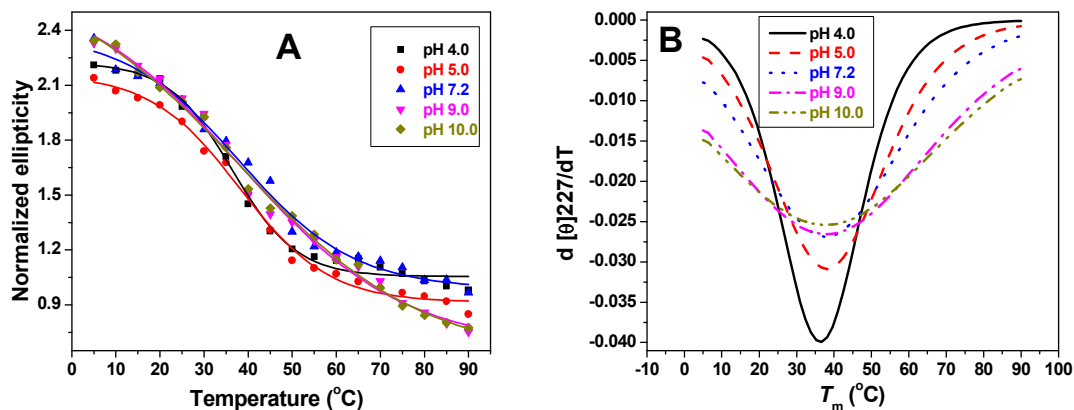


Figure 39: (A) CD-thermal denaturation plot of molar ellipticity at 227 nm vs. temperature at different pH for peptide **3** (*Pro*₉). (B) First derivative curve of peptide **3** (*Pro*₉)

3.2.4d Comparison of peptides T_m as a function of pH

The comparative T_m of peptides **1-3** obtained at different pHs is shown in Table 11. A comparison of T_m data (Table 11 and Figure 40) reveals the following: (i) *4R-Amp*₉ **1** has maximum T_m at all pHs and is almost invariant; (ii) *4S-amp*₉ **2** has the lowest T_m among the peptides at pH 4.0, but gradually increases with higher pH, to values slightly lower than the T_m of *4R-Amp*₉ **1** at pH 10.0; (iii) *Pro*₉ **3** has T_m intermediate to that of peptides **1** and **2** at pH 4.0 and remained constant in the pH range; and (iv) *4(R/S)*-aminoproline peptides have T_m values higher than that of peptide **3** at pH 10.0.

Table 11: T_m values (± 0.5) of peptides **1-3** at different pHs

pH	Peptide 1 (<i>4R-Amp</i> ₉)	Peptide 2 (<i>4S-amp</i> ₉)	Peptide 3 (<i>Pro</i> ₉)
4.0	52	27	37
5.0	51	30	38
7.2	50	41	38
9.0	50.5	47	38
10.0	50.5	48	38

Buffers used: pH 4.0-5.0, Na-acetate buffer; pH 7.2, Na-phosphate buffer; pH 9.0-10.0, Borate buffer

Thus at pH 4.0, the 4-NH_3^+ group stabilized the PPII helix most in *4R*-form and least in *4S*-form, while at pH 10.0, the stability is similar for both *4R* and *4S*-NH₂. The *4S-amp*₉ **2** exhibited significant pH-dependent PPII stability, which is maximum in the

unionized amino form at alkaline pH. This suggests a crucial combined role for both stereochemistry and the protonation of the 4-amino group in eliciting the PPII helicity of 4(*R/S*)-aminoproline polypeptides **1** and **2**.

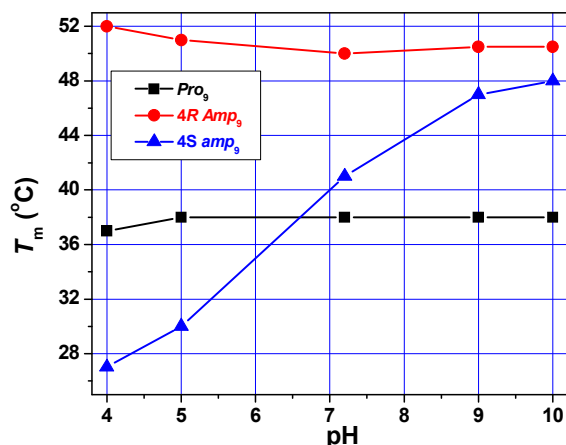


Figure 40: Melting temperature (T_m) of peptides **1-3** as a function of pH.

The thermal denaturation (self melting) studies demonstrate that at acidic pH (4.0-5.0) protonated NH_3^+ in 4*R*-*Amp*₉ **1** stabilizes the PPII conformation, whereas in 4*S*-*amp*₉ **2** it destabilizes the PPII conformation. Similarly at basic pH (9.0-10.0), the unionized NH_2 , the PPII helical stability of 4*R*-*Amp*₉ **1** remains constant while it drastically increases for 4*S*-*amp*₉ **2** with increase in pH. The stability of *Pro*₉ **3** remains unchanged as a function of pH which is consistent with previous results.

3.2.5 Effect of urea on peptides 1-3

PPII helix content is known to increase in the presence of urea, which interacts favourably with the polypeptide backbone. Hence the effect of varying concentrations of urea on peptides **1-3** was investigated through CD spectral analysis at pH 7.2. All peptides had same peptide concentration (100 μ M) for CD experiments. The buffer used for CD experiment was Na-phosphate buffer at pH 7.2 with different concentration of urea. The absorbance properties of high concentrations of urea prevented collection of spectra below 215 nm wavelength. The intensity of the positive band is taken to be proportional to the PPII helical content.

3.2.5a Peptide 1 (4*R*-*Amp*₉): The polyproline II helix is the only secondary structure known to have a positive band in the wavelength range of 218-228 nm.⁹⁰ The CD

spectra (positive band) for peptide **1** (*4R-Amp₉*) as a function of urea concentration are shown in Figure 41A. The positive band intensity in CD spectra at 224 nm plotted as a function of urea concentration is shown in Figure 41B. It is seen from the data that as the concentration of urea increases, the PPII helix content of peptide **1** (*4R-Amp₉*) also increases linearly. The PPII helicity of peptide **1** (*4R-Amp₉*) was enhanced by 15-20% by addition of 1M urea. The intensities of the positive band increases linearly by 90-100% upto 6 M urea concentration, demonstrating that urea stabilized the PPII helix in case of peptide **1** (*4R-Amp₉*).

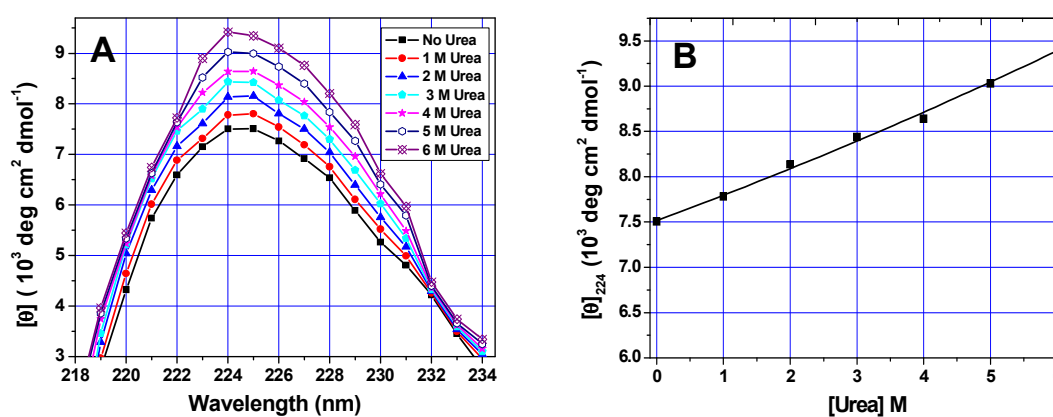


Figure 41: (A) CD profile of peptide **1** (*4R-Amp₉*) as a function of concentration of urea at pH 7.2 (B) Intensity of positive band of CD spectra as a function of concentration of urea.

3.2.5b Peptide 2 (*4S-amp₉*): The CD spectra (positive band) for peptide **2** (*4S-amp₉*) as a function of urea concentration are shown in Figure 42A. The positive band intensity in CD spectra at 224 nm plotted as a function of urea concentration is shown in Figure 42B. It is seen from data that the peptide **2** (*4S-amp₉*) has less PPII helical content in the absence of urea. As the concentration of urea is increased upto 0.25 M, the PPII helical content for peptide **2** (*4S-amp₉*) rapidly increases. At 1 M urea, the PPII helical content is comparable with that of peptide **1** (*4R-Amp₉*) and changes little afterwards upto 6 M urea. The low PPII helical content of peptide **2** (*4S-amp₉*) was enhanced enormously (>200%) by addition of 1M urea. The intensities of the positive band increases linearly by 300% upto 6 M urea concentration and indicate that urea greatly stabilizes the PPII helix in case of peptide **2** (*4S-amp₉*).

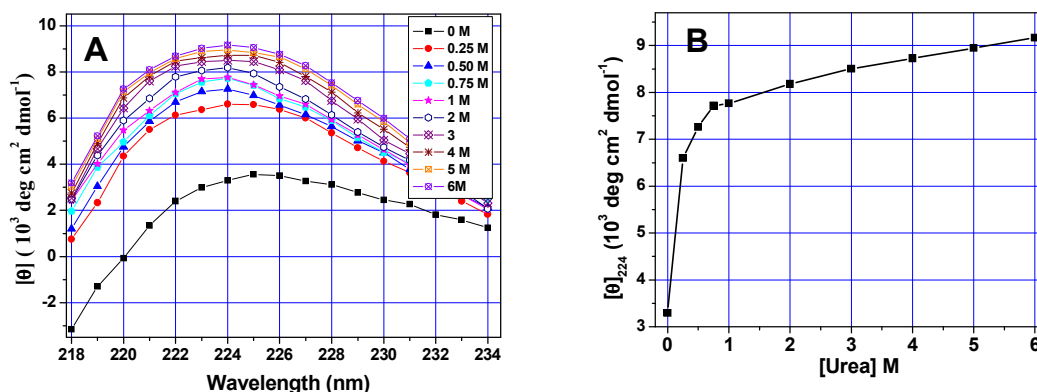


Figure 42: (A) CD profile of peptide 2 (4S-amp₉) as a function of concentration of urea at pH 7.2 (B) Intensity of positive band of CD spectra as a function of concentration of urea.

3.2.5c Peptide 3 (Pro₉): The CD spectra (positive band) for peptide 3 (Pro₉) as a function of urea concentration are shown in Figure 43A. The positive band intensity in CD spectra at 226 nm plotted as a function of urea concentration is shown in Figure 43B. It is seen that as the concentration of urea increased, the PPII helical content of peptide 3 (Pro₉) also increased linearly. The PPII helical content of peptide 3 (Pro₉) was enhanced 100% by addition of 6M urea. These results demonstrate that urea stabilized the PPII helix in case of peptide 3.

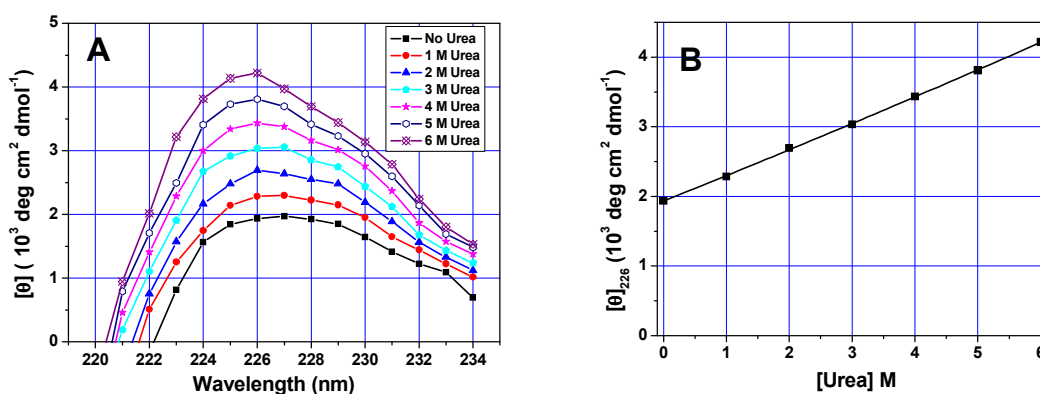


Figure 43: (A) CD profile of peptide 3 (Pro₉) as a function of concentration of urea at pH 7.2 (B) Intensity of positive band of CD spectra as a function of concentration of urea.

3.2.5d Comparative behavior of peptides 1-3 as a function of urea concentration

The positive band intensity in CD spectra of peptides 1-3 between 224-226 nm plotted as a function of urea concentration is shown in Figure 44. In all peptides, urea increased the PPII helix content linearly in the range 1M to 6M concentrations.

However, in peptide **2** (*4S-amp*₉) which initially had less PPII helix content in the absence of urea, the increase of PPII helical content was very rapid upto 1M urea at which the PPII content became comparable with that of peptide **1** (*4R-Amp*₉) and aligned with it afterwards. The low PPII helicity of peptide **2** (*4S-amp*₉) was enhanced enormously by three folds upon addition of 6 M urea (pH 7.2), while that of peptide **1** (*4R-Amp*₉) and peptide **3** (*Pro*₉) increased it by one fold. These results demonstrate that urea greatly stabilized the PPII helix in case of both 4R-and 4S-aminopropyl peptides to a greater extent than with the control prolyl polypeptide.

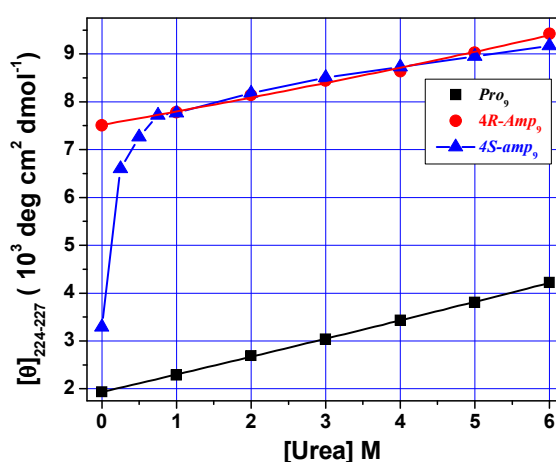


Figure 44: Intensity of CD spectra positive band of peptides **1-3** as a function of urea concentration

3.2.6 Effect of salt (NaCl) on peptides 1-3

The PPII structure is known to be disrupted by the addition of sodium chloride.^{39,40} Further neutralization of side chain charges in polylysine at higher pH induces α -helix formation, suggesting the important role of electrostatic interaction in stabilizing the helices. Hence the effect of varying concentrations of salt on peptides **1-3** was investigated through CD spectral analysis at pH 7.2. All CD experiments were done at the same peptide concentration (100 μ M) in Na-phosphate buffer (5 mM) at pH 7.2 containing different concentration of salt (NaCl). The spectra could not be recorded below wavelength of 210 nm due to high absorbance properties of the salt solution.

3.2.6a Peptide 1 (*4R-Amp*₉): The CD spectra (positive band) for peptide **1** (*4R-Amp*₉) as a function of salt (NaCl) concentration are shown in Figure 45A. The intensity of positive band in CD spectra at 225 nm plotted as a function of salt concentration shown in Figure 45B. It is observed that as the concentration of salt increases, the PPII

helical content of peptide **1** (*4R-Amp₉*) slightly decreases upto 4-5% upon addition of 5M salt. Thus salt slightly disrupts the PPII helix in case of peptide **1** (*4R-Amp₉*).

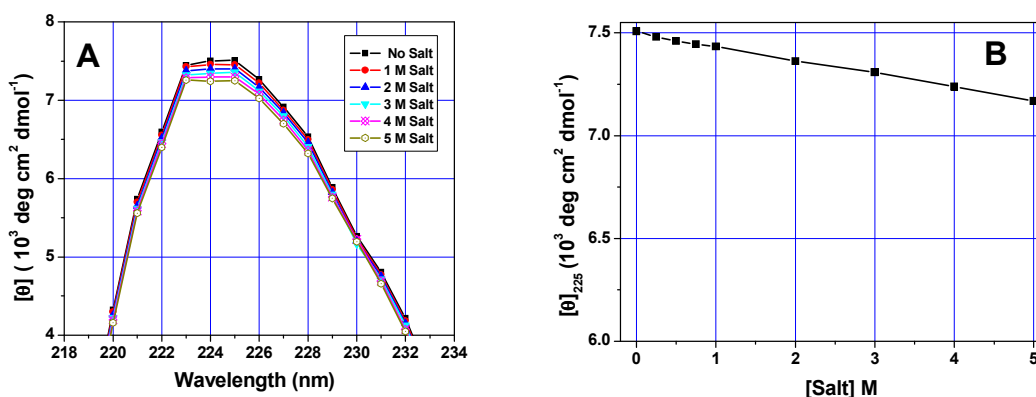


Figure 45: (A) CD profile of peptide **1**(*4R-Amp₉*) as a function of concentration of salt at pH 7.2 (B) Intensity of positive band of CD spectra's as a function of concentration of salt.

3.2.6b Peptide 2 (*4S-amp₉*): The CD spectra (positive band) for peptide **2** (*4S-amp₉*) as a function of salt concentration are shown in Figure 46A. The intensity of positive band at 225 nm in CD spectra plotted as a function of salt concentration is shown in Figure 46B. It is observed that as the concentration of salt increases upto 0.25M, the PPII helical content for peptide **2** (*4S-amp₉*) drastically decreased by 50% upto 1M salt. Thus salt greatly disrupts the PPII helix in case of peptide **2** (*4S-amp₉*).

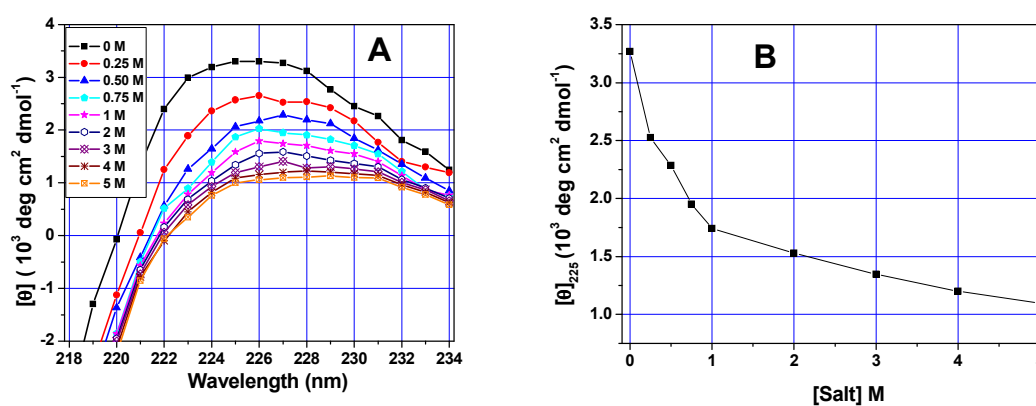


Figure 46: (A) CD profile of peptide **2** (*4S-amp₉*) as a function of concentration of salt at pH 7.2 (B) Intensity of positive band of CD spectra's as a function of concentration of salt.

3.2.6c Peptide 3 (*Pro₉*): The CD spectra (positive band) for peptide **3** (*Pro₉*) as a function of salt concentration is shown in Figure 47A. The positive band intensity at 227 nm in CD spectra plotted as a function of salt (NaCl) concentration is shown in

Figure 47B. It is seen that as the concentration of salt is increased, the PPII helical content of peptide **3** (*Pro*₉) decreases linearly by 4-5% by addition of 5M salt. Thus salt slightly disrupts the PPII helix in case of peptide **3** (*Pro*₉).

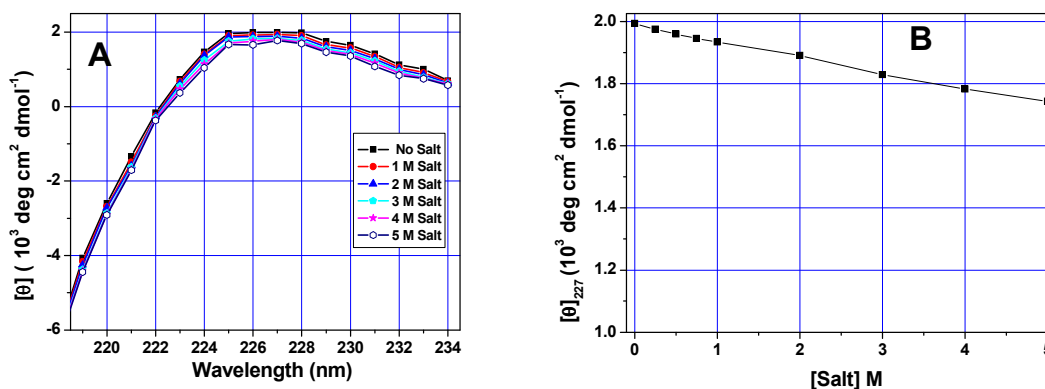


Figure 47: (A) CD profile of peptide **3** (*Pro*₉) as a function of concentration of salt at pH 7.2 (B) Intensity of positive band of CD spectra's as a function of concentration of salt

3.2.6d Comparative effect of salt on PPII stability of peptides 1-3

PPII helix content is known to decrease in presence of salt. A comparison of the intensity of positive band as a function of salt concentration for peptides **1-3** is shown in Figure 48. In all peptides, as the salt concentration increases, the PPII helical content decreased in the range of 1M to 5M concentration. However, in peptide **2** (*4S-amp*₉) PPII helical content decreased very rapidly upto 1M, while the decrease of PPII helical content of peptide **1** (*4R-Amp*₉) and peptide **3** (*Pro*₉) was gradual and a mere 4-5%. These results suggest that salt specifically destabilized the PPII helical content in peptide **2** (*4S-amp*₉).

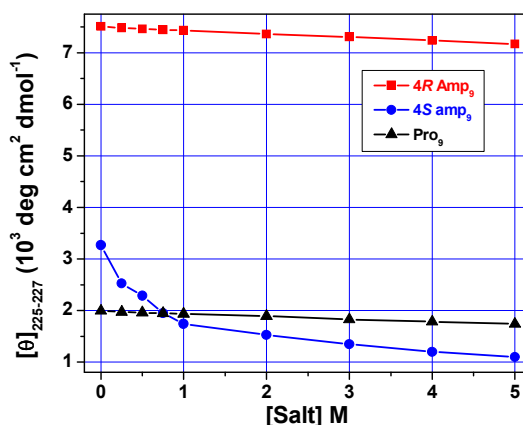


Figure 48: Intensity of CD spectra positive band as a function of salt concentration.

3.2.7 Effect of trifluoroethanol on peptides 1-3

Solvent plays a key role in modulating the H-bonding effects and hence the CD spectra of peptides **1-3** were recorded in a non-aqueous fluorinated solvent trifluoroethanol (TFE). Because the PPII \leftrightarrow PPI conversion is slow, samples were incubated in trifluoroethanol (TFE) for 6 days before measurements.³⁰ The difference in the observed intensity of CD spectra between the immediately recorded and that after 6 days of incubation was negligible.

The CD spectra recorded for peptides **1-3** in trifluoroethanol (TFE) are shown in Figure 49. The peptides **1** (*4R-Amp₉*) and **3** (*Pro₉*) show CD spectra in TFE typical of PPII conformation with a positive band around 223-226 and a negative band at 204-205 nm. Very interestingly, the CD profile of peptide **2** (*4S-amp₉*) in TFE is unlike the PPII pattern and has a bisignate shape with a negative maximum around 214 nm, a broad shoulder at 228 nm and a positive band around 197 nm. This characteristic of CD spectra was seen only in 100% TFE.

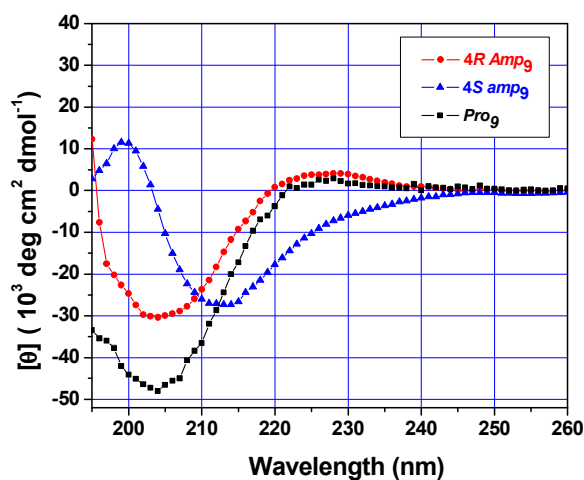


Figure 49: CD spectra profile of peptides 1-3 in trifluoroethanol (TFE).

Very interestingly this pattern corresponds to CD signature of β -structure⁹⁸ and is an unexpected but significant result requiring detailed studies. This was sought from concentration dependent CD spectroscopy study.

3.2.7a Effect of peptide concentration: Since the β -structure arises from the inter-strand interactions, it should be favored at a higher peptide concentration.⁹⁹ Hence the effect of increasing the peptide concentration for peptides **1-3** from 50 μ M to 250 μ M in TFE as solvent were recorded. Figure 50 shows the CD profiles for

peptide **2** (*4S-amp*₉) recorded from 50 to 250 μM concentration in TFE. The spectra exhibited a nice growth in the negative band intensity at 210 nm at 50 μM , accompanied by its shift to 216 nm at 250 μM and a large increase of the positive band at 200 nm. The spectral profile was also affected by the peptide concentration. At higher peptide concentration (250 μM), the peptide **2** (*4S-amp*₉) has the bisignate shape typical of a β -structure (minima at 216 nm, maxima at 198 nm). The distinctly different shape of the CD-spectra at 50 μM and 250 μM indicates different secondary structure for peptide **2** at these concentrations. An isodichroic point[†] at 208 nm for the peptide **2** (*4S-amp*₉) was seen at 208 nm, indicative of transformation between two states under these conditions. The ellipticity-amplitude at 210-216 nm of the peptide at high concentration is about 4 times lower than that of the peptide at low concentration.

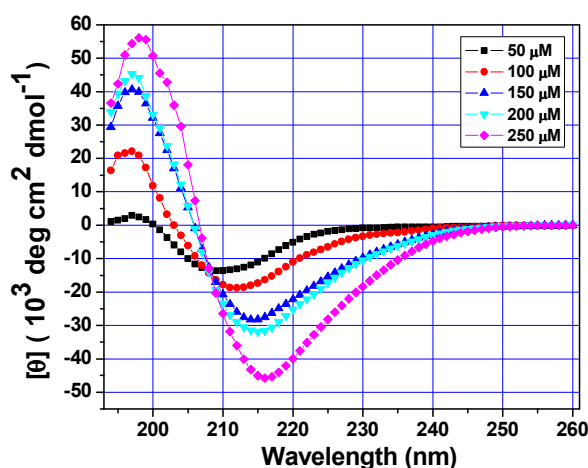


Figure 50: Increasing peptide concentration of peptide **2** (*4S-amp*₉) from 50 to 250 μM .

The effect of increasing the peptide concentration for peptides **1** (*4R-Amp*₉) and **3** (*Pro*₉) from 50 μM to 250 μM in TFE as solvent was similarly studied, and shown in Figure 51A and 51B respectively. In case of peptide **1** (*4R-Amp*₉) and peptide **3** (*Pro*₉), increasing the peptide concentration lead to a mere enhancement of PPII form without any other spectral changes. As expected, the CD amplitude increases with increasing peptide concentration. The ellipticity-amplitude at 223-227 nm of the CD spectra at high peptide concentration rises as the peptide concentration is increased. No drastic change in CD pattern was seen in peptides **1** (*4R-Amp*₉) and **3** (*Pro*₉). An isoelliptic

[†] An isodichroic/isobestic suggest that there are two different states or species present.

point[‡] at zero $[\theta]$ was seen for the peptide **1** (4*R*-*Amp*₉) at 220 nm and for peptide **3** (*Pro*₉) at 223 nm when peptide concentration varied from 50 μM to 250 μM . The results imply that peptide **1** (4*R*-*Amp*₉) and peptide **3** (*Pro*₉) retain PPII form in trifluoroethanol.

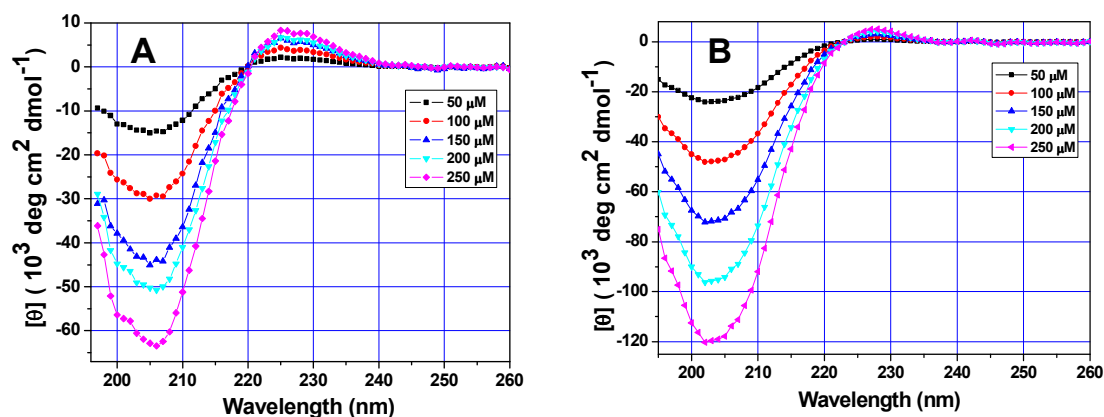


Figure 51: Increasing peptide concentration for (A) Peptide **1** (4*R*-*amp*₉) from 50 to 250 μM . (B) Peptide **3** (*Pro*₉) from 50 to 250 μM .

The overall results assert that peptide **2** (4*S*-*amp*₉) specifically assumes a β -structure, while peptides **1** (4*R*-*Amp*₉) and **3** (*Pro*₉) retain PPII conformation in TFE. Peptide **2** (4*S*-*amp*₉) is perhaps the first case of a polyproline peptide exhibiting a β -structure. Since formation of a β -structure by any polyproline peptide derivative is unprecedented in literature, additional confirmation from another technique is needed.

3.2.8 Identification of β -structure by Raman Spectroscopy

In order to support the formation of β -structure by an independent technique, Raman spectroscopic studies were carried out for peptide **1** (4*R*-*Amp*₉) and peptide **2** (4*S*-*amp*₉) in TFE. Raman spectroscopy has been abundantly used in analysis of protein conformations wherein one-to-one correspondence between the amide bands and the type of secondary structure has been established.¹⁰⁰ The peptides **1-2** were dissolved in TFE and Raman spectra were recorded using 532 nm frequency-doubled Nd:YAG laser. The blank TFE Raman spectra were subtracted from original spectra shown in Figure 52. The position of amide I band around 1630–1640 cm^{-1} corresponds to α -helix, 1640–1660 cm^{-1} corresponds to random coil and 1660–1675 cm^{-1}

[‡] An isoeliptic point arises due to change in concentration of peptide. Only the intensity of CD band increases or decreases with concentration.

corresponds to β -structure. It is seen from Figure 52 that for the peptide **2** (*4S-amp*₉), the amide I and amide III bands were observed at 1670 cm^{-1} and 1628 cm^{-1} respectively which clearly suggests the formation of β -structure. Significantly the characteristic band seen for peptide **2** at 1670 cm^{-1} is absent in the Raman spectra of peptide **1** (*4R-Amp*₉). A comprehensive Raman band assignment was performed and the results are tabulated in Table 12.

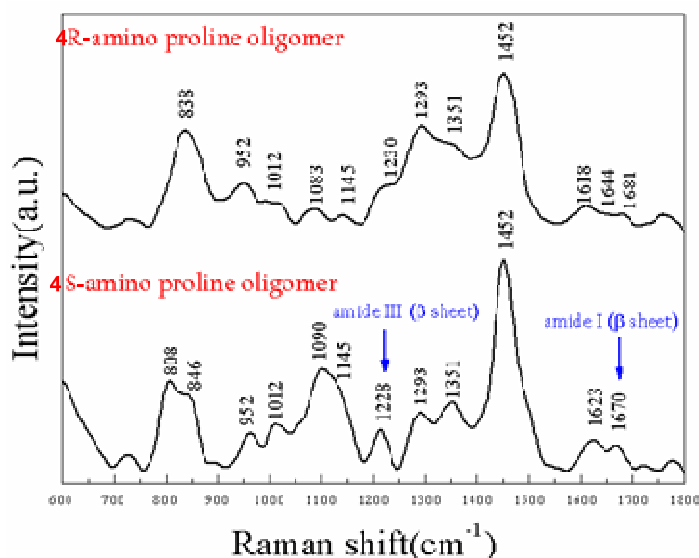


Figure 52: Raman spectra of peptide **1** (*4R-Amp*₉) and peptide **2** (*4S-amp*₉) from 600 cm^{-1} to 1800 cm^{-1} .

Table 12: Raman band assignments for peptides **1** and **2**

Peptide 1 (<i>4R-Amp</i> ₉)		Peptide 2 (<i>4S-amp</i> ₉)	
Raman shift (cm^{-1})	Band assignment	Raman shift (cm^{-1})	Band assignment
838	δ (proline ring)	846	δ (proline ring)
952	δ (N-H)	952	δ (N-H)
1012	ν (N-C) + δ (N-C-C)	1012	ν (N-C) + δ (N-C-C)
1083	ν (C-C)	1090	ν (C-C)
1145	τ (CH_2)	1145	τ (CH_2)
1230	amide III	1228	amide III - β structure
1351	ν (COO^-)	1351	1351
1452	δ (CH_2)	1452	1452
1618	ring (C-C)	1623	1623
1644	amide I – random coil	1670	amide I – β structure
1681 (very weak mode)	amide I – random coil/ β structure		

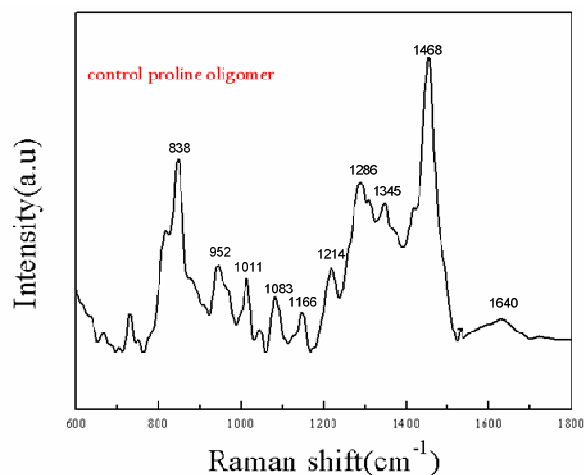


Figure 53: Raman spectra peptide 3 *Pro*₉ from 600 cm^{-1} to 1800 cm^{-1} .

Figure 53 is the Raman spectra obtained from the control proline oligomer recorded with the same experimental parameters as above. The amide I band for the control peptide 3 (*Pro*₉) is around 1640 cm^{-1} corresponding to a random coil/polyproline conformation.

3.2.9 Water-induced switching of β -structure to PPII conformation in the 4S-amp₉

In order to examine the origin of β -structure seen exclusively for peptide 2 (4S-amp₉) in trifluoroethanol, aliquots of aqueous phosphate buffer solution (pH 7.2) was titrated into TFE solution and the changes monitored by CD spectra is shown in Figure 54.¹⁰¹

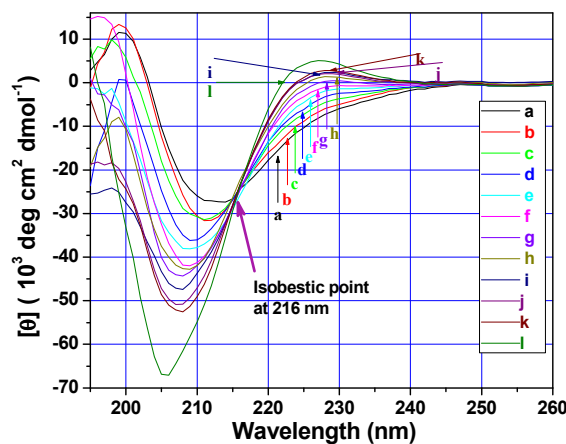


Figure 54: CD spectra of peptide 2 (4S-amp₉) in TFE with incremental addition of phosphate buffer (pH 7.2) from (a) 0.1% to (k) 1.0% in 0.1% steps and (l) 2.0%.

The bisignate CD spectrum characteristic of a β -structure gradually shifted to spectral line shape that corresponds to polyproline II structure. The negative band at 214 nm slowly shifted to 205 nm, accompanied by growing broad negative shoulder at 228 nm into a positive band at about 224 nm, typical of PPII helical form. The isobestic point seen at 216 nm is indicative of the conversion of peptide **2** (4*S*-amp₉) from β -structure in 100% TFE to full PPII form with 0.8% buffer in TFE.

The CD spectra were also recorded in different percentage of mixtures of aqueous phosphate buffer (5 mM, pH 7.2) and trifluoroethanol (TFE) and shown in Figure 55. It is seen that polyproline PPII helical content increases by addition of even 10% TFE and a very small difference was observed upto 95% TFE in buffer. The ellipticity-amplitude of the CD spectra slightly increased as the TFE content increased upto 90-95%. No PPI conformation was observed in different percentage of TFE in water. This suggests that solvent plays an important role in determining the conformation and stability of peptide secondary structure.

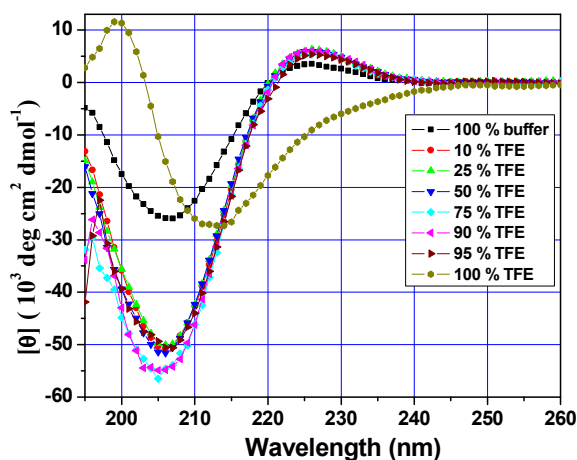


Figure 55: Different percentage of TFE in Na-Phosphate Buffer

The results clearly confirm that peptide **2** (4*S*-amp₉) assumes a β -structure in TFE that is transformed to the PPII form in aqueous medium, unlike peptide **1** (4*R*-Amp₉) which retains the PPII form in both conditions.

3.2.10 Effect of *n*-propanol on peptides 1-3

The PPII helix is the predominant conformation of oligoproline in water, whereas the PPI helix is favoured in solvents such as *n*-propanol.¹⁹ The amount of *n*-propanol needed to switch the PPII helix to the PPI helix therefore reflects the ease of the

conformational change between the PPII and the PPI form and can serve as a measure of their relative stabilities. Hence the CD spectra of peptides **1-3** were recorded in an aliphatic alcohol such as *n*-propanol. All peptides had same peptide concentration (100 μ M) for CD experiment. Because the PPII \leftrightarrow PPI conversion is slow, samples were incubated in different mixtures of aqueous phosphate buffer and *n*-propanol for 6 days before measurements.

3.2.10a Peptide 1 (4R-Amp₉): The CD spectra of peptide **1** (4R-Amp₉) recorded in different percentage mixtures of aqueous phosphate buffer (5 mM, pH 7.2) and *n*-propanol is shown in Figure 56. It is seen that PPII helical content increases by addition of 10% *n*-propanol and a very small difference was observed upto 50% *n*-propanol in buffer. The ellipticity-amplitude of the CD spectra slightly increases as the *n*-propanol content increases in mixture upto 50%. When *n*-propanol percentage increases from 50% to 100%, PPII helical content decreased rapidly. Interestingly no PPI conformation was seen in different percentages of *n*-propanol in buffer. This suggests that peptide **1** (4R-Amp₉) retains its PPII conformation even in *n*-propanol.

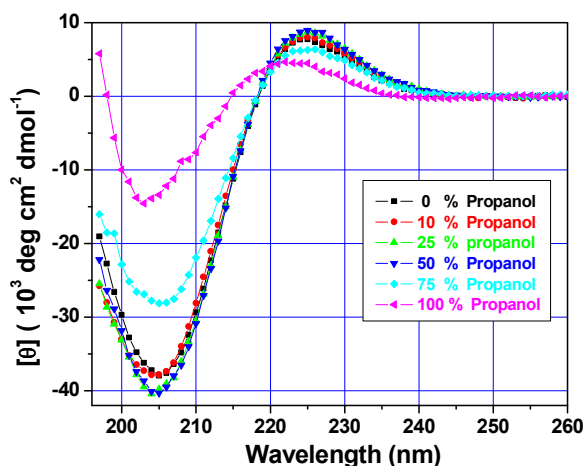


Figure 56: CD profile of peptide **1** (4R-Amp₉) in different percentage of *n*-propanol in buffer.

3.2.10b Peptide 2 (4S-amp₉): The CD spectra of peptide **2** (4S-amp₉) recorded in different percentage of mixtures of aqueous phosphate buffer (5 mM, pH 7.2) and *n*-propanol is shown in Figure 57. It is seen that PPII helical content slightly increases by addition of even 10% *n*-propanol and very small difference was observed upto 50%. The ellipticity-amplitude of the CD spectra slightly increases as the *n*-propanol content is increased upto 50%. When percentage of *n*-propanol increases upto 75%, PPII

helical content decreases. In 100% *n*-propanol it has bisignate shape showing a negative maximum around 214 nm and a broad shoulder at 228 nm that is typical of β -structure with a positive band around 200 nm. No PPI conformation was seen in different percentages of *n*-propanol in buffer. This suggests that peptide 2 *4S-amp*₉ retains its PPII conformation in *n*-propanol upto 75%, but in 100 % *n*-propanol it forms a β -structure.

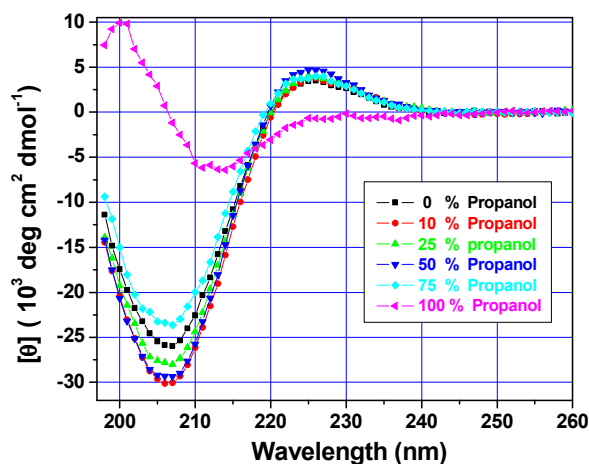


Figure 57: CD profile of peptide 2 (*4S-amp*₉) in different percentage of *n*-propanol in buffer.

3.2.10c Peptide 3 (*Pro*₉): The CD spectra of peptide 3 (*Pro*₉) was recorded in different percentage of mixtures of aqueous phosphate buffer (5 mM, pH 7.2) and *n*-propanol is shown in Figure 58.

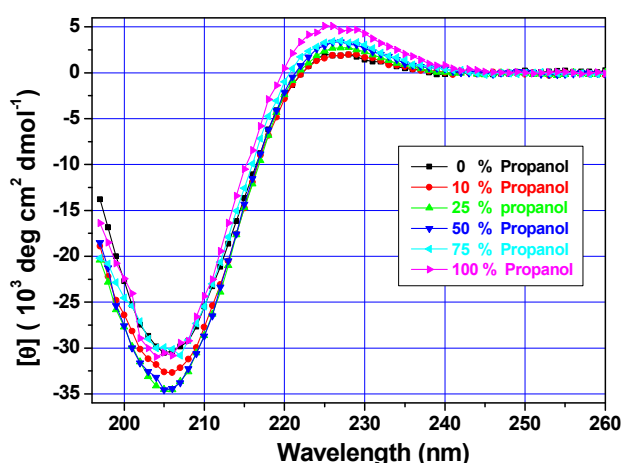


Figure 58: CD profile of peptide 3 (*Pro*₉) in different percentage of *n*-propanol in buffer.

It is seen that PPII helical content increased by addition of 10% *n*-propanol and comparable difference was observed upto 100% *n*-propanol in buffer. The ellipticity-amplitude of the CD spectra slightly increased as the *n*-propanol content increases in the mixture upto 100%. No PPI conformation was observed in different percentage of *n*-propanol in buffer. This suggests that peptide **3** (*Pro*₉) retains its PPII conformation in *n*-propanol.

3.3 Discussion

The pK_a values of amino group in 4*R*-Amp monomer and 4*S*-amino monomer are 10.2 and 9.3 respectively, while the pK_a of peptides **1** (4*R*-Amp₉) and **2** (4*S*-amp₉) are 8.55 and 7.83 (Table 8). The 4-amino group has higher pK_a in monomers (more basic, less acidic) than the homopeptides. The pK_a of 4-amino group in 4*R* and 4*S* monomers is also different, indicating that both monomers are present in different amount of protonated form at physiological pH. This may arise from stereoelectronic effects within the molecules leading to stabilization from delocalization of the charge and reducing the energy of the system. The increased acidity of 4-NH₂ group may arise either through H-bonding,¹⁰² space field effect, bond inductive effect, solvation or stereoelectronic anomeric and gauche effects.¹⁰³ However it is interesting to note that the pK_a of exocyclic 4-amino group changes when present in a peptide chain, due to other local micro environmental effects.

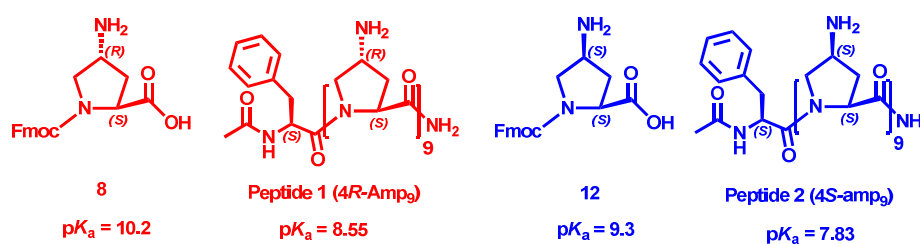


Figure 59: pK_a of monomers **8** and **12** and peptides **1-2**.

Peptides **1-3** form PPII helix at physiological pH as indicated by positive band at 222 to 225 nm and a negative band at 200 to 210 nm (Figure 31) which is the established pattern of the PPII conformation. Since the intensity of positive band in CD spectrum is proportional to the PPII helical content, it emerges from Figure 31 that both peptide **1** (4*R*-Amp₉) and peptide **2** (4*S*-amp₉) have more positive intensity in the region (225-227 nm) than the unsubstituted proline peptide **3** (*Pro*₉). Peptide **1** (4*R*-Amp₉) has maximum PPII helix content and the intensity of the positive band at 225-

227 nm decreases in the order peptide **1** (*4R-Amp₉*) > peptide **2** (*4S-amp₉*) > peptide **3** (*Pro₉*). The peptide **1** with 4-amino group in *4R* stereochemistry seems to favour PPII form more than the peptide **2** (*4S-amp₉*) with 4-amino group in *4S* stereochemistry which is less favourable for PPII conformation.

The peptide **1** (*4R-Amp₉*) also shows more negative ellipticity (205 nm) than peptide **2** (*4S-amp₉*) and peptide **3** (*Pro₉*) suggesting the presence of relatively more diverse non-PPII conformers. The peptide **2** (*4S-amp₉*) shows less negative ellipticity than peptides **1** (*4R-Amp₉*) and peptide **3** (*Pro₉*).

In ratio of intensity of positive to negative bands in CD spectra (Figure 32) for all peptides (**1-3**) show a linear increase with concentration without saturation. Since this indicates that the peptide strands do not associate into higher order and are present in single chain helix form.

The relative effects of protonation (Figure 36) on peptides **1** and **2** (*4R/S-Amp₉*) is seen at acidic pH. Peptide **1** with *4R* stereochemistry shows higher PPII helical content and positive ellipticity decreasing upto physiological pH and remains constant in basic pH range. The peptide **2** (*4S-amp₉*) shows an opposite behavior in acidic pH with a lower PPII helical content which increases as pH is increased to basic range. The positive ellipticity of peptide **3** (*Pro₉*) remained constant with pH. This suggests a combined role of both stereochemistry and protonation status of the 4-amino group in modulating the PPII helicity of 4(*R/S*)-aminoproline polypeptides **1** and **2**.

The thermal denaturation (self melting) studies (Table 11) demonstrate that at acidic pH (4.0-5.0) protonated NH_3^+ form stabilizes the PPII conformation in peptide **1** (*4R-Amp₉*) more than that in peptide **2** (*4S-amp₉*). Similarly at basic pH (9.0-10.0), the unionized NH_2 form does not alter the stability of peptide **1** (*4R-Amp₉*) but the stability of peptide **2** (*4S-amp₉*) increases with increase in pH. The stability of peptide **3** (*Pro₉*) remains unchanged as a function of pH, as expected since it does not have any ionizable function.

At acidic pH, 4- NH_2 gets protonated as NH_3^+ and acts like electron withdrawing group. The stereoelectronic effect (Figure 60) of *4R-X* substituents (*R* stereochemistry + electron withdrawing effect) on proline is known to strongly favour the PPII form by enhancing the *trans*-amide content.²⁸ This may be the plausible reason for stabilization of peptide **1** (*4R-Amp₉*), while the peptide **2** (*4S-amp₉*) having opposite stereocenter has low stability at acidic pH.

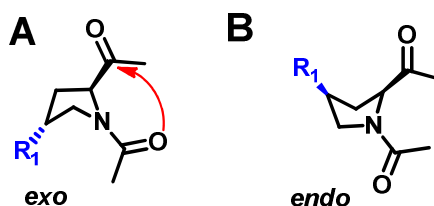


Figure 60: An $n \rightarrow \pi^*$ interaction stabilizes the *trans* isomer of the peptide bond (A) Substantial only when Pro derivatives are in the C^γ -*exo* ring pucker (*R* stereochemistry) (B) when Pro derivatives are in C^γ -*endo* ring pucker (*S* stereochemistry) $n \rightarrow \pi^*$ interaction are not operative.

At basic pH, NH_2 acts like hydroxyl group which has ability to form hydrogen bonding (Figure 61). The data suggest that the stability of peptide **1** (*4R-Amp*₉) remains constant and the stability of peptide **2** (*4S-amp*₉) increases with increase in pH. This indicates that the stabilizing effect of **2** (*4S-amp*₉) may be due to intramolecular hydrogen bonding interaction between side chain amino group and the backbone carbonyl oxygen which has been proposed to stabilize PPII conformation.²⁵

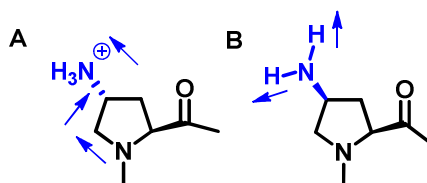


Figure 61: (A) At protonated state amino group exhibits electron withdrawing effect (B) At basic pH amino group has ability to form hydrogen bonding like hydroxy group.

PPII helical content is known to increase in presence of urea, which interacts favorably with the polypeptide backbone. Hence the effect of varying concentrations of urea on 4-aminoproline polypeptides was investigated through CD spectral analysis at pH 7.2 (Figure 44). The change in CD profile indicates that the ensemble of backbone conformations is changing for peptide as the urea concentration is increased. Urea greatly stabilized the PPII helix in case of both *4R*- and *4S*-aminoprolyl peptides to a greater extent than with the control prolyl polypeptide. The initial large increase with peptide **2** (*4S-amp*₉) may be due to reorientation of the amino group to form urea mediated hydrogen bonding with the amide carbonyl carbon (Figure 62) and thereby favoring the PPII helix. Such type of complementary H-bonding is found in literature.^{35,37} The steady increase of PPII helical content with urea for peptide **1** (*4R-Amp*₉) and peptide **3** (*Pro*₉) suggests a rigidification of peptide backbone by bidentate H-bonding by urea. Larger changes seen specifically for peptide **2** (*4S-amp*₉) in the

presence of urea and upon increasing the pH suggest the combined role of H-bonding and stereoelectronic effects in formally dictating the PPII conformation.

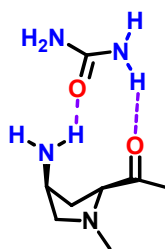


Figure 62: Urea mediated hydrogen bonding in peptide **2** (*4S-amp*₉)

PPII helical content is known to decrease in presence of salt. Hence the effect of varying concentrations of salt on 4-aminoproline polypeptides was investigated through CD spectral analysis at pH 7.2 (Figure 48). The results suggest that salt destabilized the PPII helix in case of *4R/S*-aminoprolyl peptides and control prolyl polypeptide. These observations are consistent with the findings that the PPII helical content of proline homopolymers or peptides decreased with increasing salt concentration.¹⁰⁴ The initial larger decrease with peptide **2** (*4S-amp*₉) may be due to salt immediately disrupting the hydrogen bonding network which results in a decrease of PPII helical content. The sudden significant decrease of PPII helical content in peptide **2** (*4S-amp*₉), strengthens the importance of weak intramolecular H-bonds in retention of PPII form. While the linear decrease of PPII helical content with salt for peptide **1** (*4R-Amp*₉) and peptide **3** (*Pro*₉) suggest that the decrease may be due to chaotropic action of concentrated NaCl which decreases the secondary structure content. In case of peptide **3** (*Pro*₉), there is no case of screening of charges because this peptide does not possess any formal charges or hydrogen bonding. This result is due to the chaotropic action of concentrated NaCl.

Larger changes seen specifically for peptide **2** (*4S-amp*₉) in the presence of urea, salt and upon increasing the pH suggest the combined role of H-bonding and stereoelectronic effects in dictating the PPII conformation. Recently, *4S*-(OH/NH₃⁺) groups on proline were shown to form intramolecular H-bonds with the amide carbonyl (Figure 63), increasing the *trans/cis* amide ratio and thereby promoting PPII conformation in the derived polypeptides.¹⁰⁵

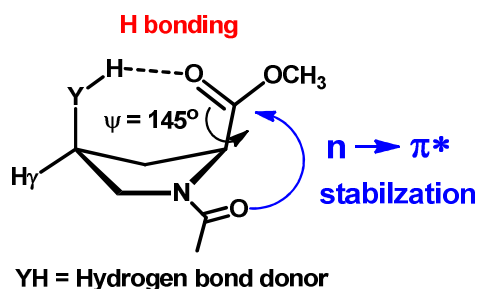


Figure 63: C4-*endo* conformations of 4*S*-configured derivatives at Y angles of approximately 145° with H-bond donors at C4.

In order to evaluate the effect of non-aqueous solvent on conformations of poly-4(*R/S*)-aminoprolines, the CD spectra of peptides **1-3** were recorded in trifluoroethanol (TFE) (Figure 49). The CD profiles indicate that peptide **1** (4*R*-*Amp*₉) and peptide **3** (*Pro*₉) retain their PPII conformation. Very interestingly, the CD profile of peptide **2** (4*S*-*amp*₉) in TFE exhibited a pattern, characteristic of a β -structure. Such formation of β -structure in any polyproline peptides under any conditions is unprecedented in the literature since they lack H-bond donor sites. Hypothetically formation of the β -structure is possible in case of poly-4*S*-OH proline, but not known in literature. Hence additional confirmation was done by concentration dependent CD spectroscopy. In this experiment the isodichroic point seen at 208 nm for the peptide **2** (4*S*-*amp*₉) suggests a two-state conversion from a low concentration PPII secondary structure into a β -structure at higher concentration, arising from *intermolecular* hydrogen bonding increasing with concentration. Further confirmation was seen from Raman spectroscopy data, in which the bands at 1228 and 1670 cm⁻¹ specifically seen for peptide **2** (4*S*-*amp*₉) support the formation of β -structure.

All these results can be rationalised by following description. In peptide **2** (4*S*-*amp*₉), the NH₂ group can form an *intramolecular* H-bond with the amide carbonyl of the same proline moiety which promotes the PPII conformation.¹⁰⁵ The possibility of the 4*S*-NH₂ group engaging the amide carbonyl of another chain of 4*S*-*amp*₉ through an *intermolecular* H-bond would lead to antiparallel β -structure. Such an *interchain* H-bonded structure would be facilitated at higher peptide concentration. Even at high concentration in aqueous solution, no β -structure is seen due to competition of H₂O molecules which disrupt intermolecular H-bonds seen for β -structure. The proposed synthon for β -structure is shown in Figure 64.

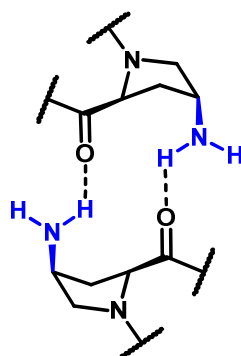


Figure 64: Proposed synthon for structure of antiparallel β -structure

The formation of β -structure for peptide **2** (*4S-amp*₉) was also observed in *n*-propanol, while peptide **1** (*4R-Amp*₉) and peptide **3** (*Pro*₉) retain PPII conformation in *n*-propanol. No PPI form was seen for any of the peptides **1-3**, under different conditions of pH, urea, salt, *n*-propanol and TFE. This is due to fact that all the peptides (**1-3**) are capped as N-terminus acetyl and C-terminal amide, which favour the PPII over the PPI helix. The electrostatic interactions between the terminal charges and the amide dipoles stabilize both helices. They are significantly stronger in the PPI helix where the amide bonds are oriented linear to the helix axis as compared to the PPII helix in which the amides are perpendicular to the axis.

3.3.1 Water-induced switching of β -structure to PPII conformation in the *4S-amp*₉

When aqueous phosphate buffer (pH 7.2) was titrated into TFE solution of peptide **2** (*4S-amp*₉) in tiny incremental steps of 0.1%, the 214 nm negative band slowly shifted to 205 nm, accompanied by a growing of the broad negative shoulder at 228 nm into a positive band at about 224 nm (Figure 55) typical of PPII form. The isobestic point seen at 215 nm is indicative of the conversion of peptide **2** (*4S-amp*₉) from β -structure in 100% TFE to full PPII form with 0.8% buffer in TFE. The overall results imply that peptide **2** (*4S-amp*₉) assumes a β -structure in TFE that is transformed to the PPII form in aqueous medium.

The β -structure arises from *interchain* hydrogen bonds involving *4S-NH*₂ and amide carbonyl, which are broken in water and rearranged to *intramolecular* H-bonding that favors the PPII form *via* enriching the *trans* amide geometry.

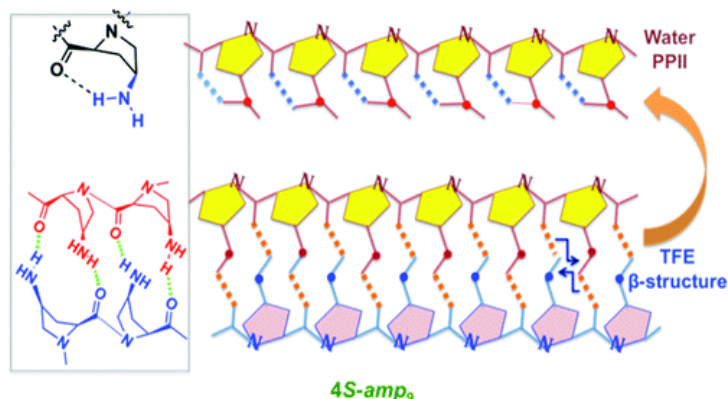


Figure 65: Water induced switching β -structure to PPII conformation in peptide **2** ($4S$ - amp_9)

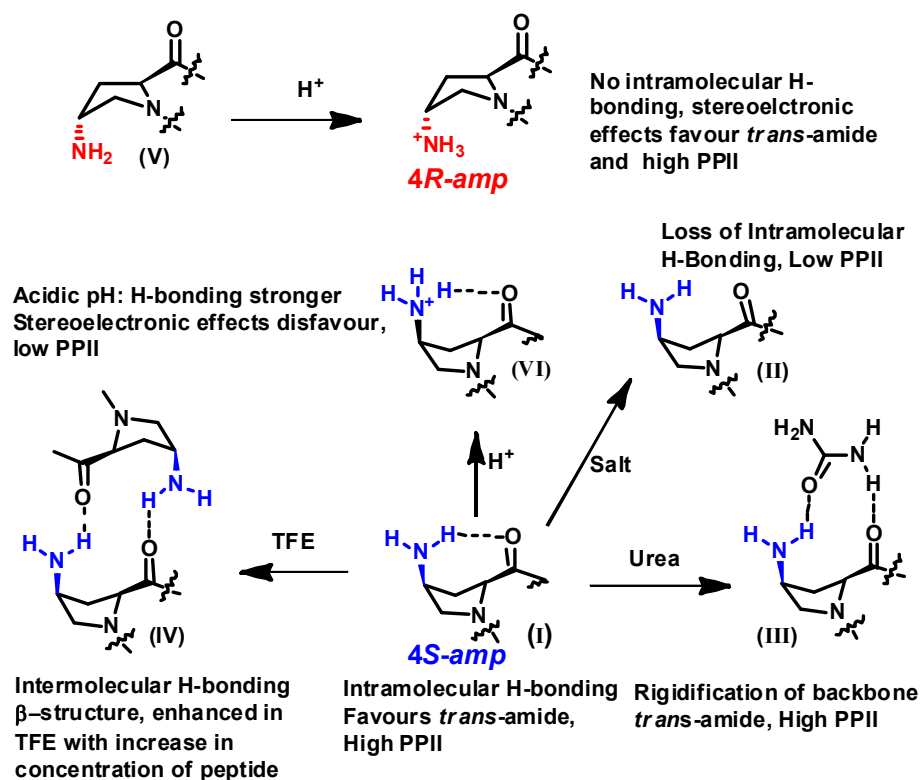
3.3.2 Stereoelectronic versus H-bonding effects in peptide **2** ($4S$ - amp_9)

The stereoelectronic effect of $4R$ -X substituents on proline is known to strongly favour the PPII form, over that of $4S$ -X substituents by enhancing the *trans*-amide content.²⁸ Although the intramolecular H-bonding in $4S$ - NH_3^+ favors the *trans*-amide form,^{105b} the unfavourable stereoelectronic effect of $4S$ - NH_3^+ (unlike $4R$ - NH_3^+) strongly negates the benefit of H-bonding, leading to a low PPII form for peptide **2** ($4S$ - amp_9) at acidic pH. At higher pH, the effect of intramolecular H-bonding dominates a weaker stereoelectronic effect of $4S$ - NH_2 , promoting higher PPII content in peptide **2** ($4S$ - amp_9). While the conversion of PPII to PPI takes many hours/days due to a slow conversion of amide from *trans* to *cis* form,³⁰ the switching of β -structure to PPII form is fast, within minutes, suggesting that the amide bond of peptide **2** ($4S$ - amp_9) is also in *trans* form in the β -structure.

3.4 Conclusion

A plausible molecular picture for various structural transitions under different conditions for peptide **1-2** is depicted in Scheme 8. The *intramolecular* H-bonding of $4S$ - NH_2 with amide carbonyl possible only in peptide **2** ($4S$ - amp_9) (I) promotes PPII form in buffer, while the presence of salt weakens the PPII by breaking the H-bond (II). Urea rigidifies the backbone by complementary H-bonding (III) to *cis*-disposed $4S$ - NH_2 group and amide carbonyl to enhance the PPII form. In a fluorinated solvent TFE, the intramolecular H-bonding between $4S$ - NH_2 and the amide carbonyl switches to interchain H-bonding (IV) giving rise to antiparallel β -structure. The *trans* disposition of $4R$ - NH_2 and amide carbonyl group in $4R$ - Amp_9 (V) is not conducive to form either intramolecular or interchain H-bonding in the derived peptide.

Scheme 8



In summary, it is demonstrated here that peptide 2 (4S-amp₉) adapts an unusual β -structure in TFE unlike most polyproline peptides which prefer the PPI form in hydrophobic/ fluorinated media. The β -structure arises from *interchain* hydrogen bonds involving 4S-NH₂ and amide carbonyl, which are broken in water and rearranged to *intramolecular* H-bonding that favors the PPII form *via* enriching the *trans* amide geometry. This structural conversion illustrates a fine balance between stereoelectronic and H-bonding effects in novel tuning of the secondary structure of 4R/S-aminoproline polypeptides. β -structure in polyproline peptides is hitherto unknown, and the present results will add a new design principle to a growing repertoire of strategies for engineering peptide secondary structural motifs for new biomaterials and nanoassemblies.¹⁰⁶

In future, more research needs to be done to unambiguously establish these results in terms of 2D-NMR studies, theoretical studies, host-guest peptide system, peptide length and use of these peptide for biological study *etc.* to obtain better insight into this aspect. The results have a direct bearing on the current interest in polyproline peptides and stimulus responsive materials.

3.5 Experimental

Materials and reagents were of the highest commercially available grade and used without further purification. Reactions monitored by thin layer chromatography (TLC) were carried out on precoated silica gel 60 F254 plates (E. Merck). TLCs were visualized under UV light, iodine and/or ninhydrin spray followed by heating upto 110 °C with heat gun. Silica gel 60-120 and 100-200 mesh (Merck) was used for routine column chromatography. Elution was done with ethyl acetate/petroleum ether or dichloromethane/methanol mixture depending upon the compound polarity and chemical nature. Flash chromatography was done using 230-400 mesh silica. All solvents were distilled under an inert atmosphere with appropriate desiccant. Reactions in aqueous medium and workup processes were done using double distilled water. Unless otherwise noted, all reactions were carried out at room temperature. IR spectra were recorded on an Infrared Fourier Transform Spectrophotometer using chloroform, nujol or neat sample. The optical rotation values were measured on JASCO P2000 polarimeter. ¹H and ¹³C NMR spectra were recorded on Bruker AV 500 and AV 400, AV 200 spectrometers. Chemical shifts are reported in ppm using TMS and CDCl₃ as a reference. Spectra were analyzed using ACD spectviewer software from ACD labs. NMR spectra of compounds show two sets of peaks arising from *cis-trans* isomerization of N-CO bond. Analytical HPLC was performed using a LiChrospher 100 RP-18e 5 μm (250 mm x 4 mm) column from Merck. Preparative HPLC was carried out on a LiChrospher RP-18e 5 μm (250 mm x 10 mm) column from Merck. Mass spectra were obtained by ESI-MS technique on AP-QSTAR spectrometer. Jasco J-715 (Japan) instrument was used for CD measurements. All graphs presented for CD spectra's are drawn by Origin 8 software.

Reagents for the buffer preparation such as NaCl, Na₂BO₄, NaH₂PO₄, Na₂HPO₄, Citric acid and Na-Borate were obtained from Sigma and were of Molecular biology reagent grade. Double distilled water was demineralized using Millipore MilliQ deionizer and was used for the preparation of buffers. pH was adjusted using NaOH and HCl in case of Borate buffer and using acetic acid for acetate buffer.

Resins for solid phase peptide synthesis and Fmoc-protected amino acids were bought from Novabiochem and were used without further purification.

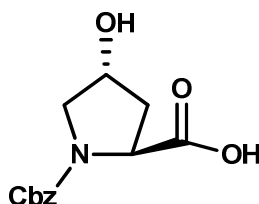
The term “concentrated under reduced pressure” refers to the removal of solvents and other volatile materials using a rotary evaporator at water aspirator pressure (<20

Torr) while maintaining the water-bath temperature below 40 °C. Residual solvent was removed from samples at high vacuum (<0.1 Torr). The term “high vacuum” refers to vacuum achieved by a mechanical belt-drive oil pump.

3.5.1 Synthesis of compounds 2-12

Compounds 2-12 was synthesized by reported literature procedure.⁹²

(2*S*, 4*R*)-N1-(benzyloxycarbonyl)-4-hydroxyproline (2)

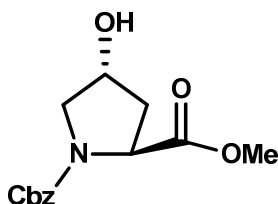


A solution of *trans*-4-hydroxy-L-proline **1** (7 g, 53.4 mmol) in 10% aqueous NaHCO₃ (80 mL) was cooled in an ice bath to which solution of Cbz-Cl in toluene (50%) was added. The reaction mixture was stirred for 12 h at room temperature, and solvent toluene was removed under vacuum. The aqueous solution was cooled to 0 °C and washed with diethyl ether for 3-4 times to remove unreacted Cbz-Cl. The aqueous layer was acidified to pH 2 with aqueous HCl (1N) and then extracted with ethyl acetate (3 x 50 mL). The combined organic solution was washed with water followed by saturated brine solution. The organic layer was dried over anhydrous Na₂SO₄ and concentrated under reduced pressure to yield white sticky solid of compound **2**. This compound was used for the next reaction without purification. Yield: 13.56 g, (98.8%). Molecular Formula, C₁₃H₁₅NO₅. [Mass: calculated 265, observed 266.05(M+1)]. [α]: -86.70 (c 1.0 CH₂Cl₂).

¹H NMR (CDCl₃, 200 MHz) δ: 2.06-2.13 (1H, m), 2.62-2.32 (1H, m), 3.53-3.61 (2H, m), 4.14-4.51 (1H, m), 7.26-7.36 (5H, m).

¹³C NMR δ: 37.8, 38.7, 49.0, 54.3, 57.6, 60.4, 76.7, 127.0, 127.7, 128.4, 136, 155.1, 175.1.

(2*S*, 4*R*)-N1-(benzyloxycarbonyl)-4-hydroxyproline methyl ester (3)

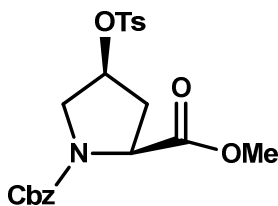


A solution of compound **2** (13 g, 49 mmol) in anhydrous acetone (75 mL), anhydrous K_2CO_3 (16.9 g, 122 mmol) and dimethylsulphate 5.6 mL (58.8 mmol) was stirred in a flask. The mixture was then refluxed under nitrogen for 4 h. The acetone was removed under vacuum and the resulting residue was dissolved in water followed by extraction with ethyl acetate (3 x 50 mL). The combined organic layer was washed with water (3 x 50 mL) followed by saturated brine solution (3 x 30 mL) then dried over Na_2SO_4 and concentrated under vacuum. The crude material was purified by silica gel chromatography (50% ethyl acetate/hexane) afford compound **3** as colourless thick oil. Yield: 13.52 g, (98.78%). Molecular Formula, $C_{14}H_{17}NO_5$. [Mass: calculated 279, observed 280.04 (M+1)]. $[\alpha]$: -61.15 (c 1.0 CH_2Cl_2).

1H NMR ($CDCl_3$, 200 MHz) δ : 2.04-2.28 (2H, m), 3.32 (1H, m), 3.51-3.73 (5H, m), 4.41(2H, m), 5.11 (2H, m), 7.32 (5H, m).

^{13}C NMR δ : 37.8, 38.6, 51.7, 54.7, 57.6, 66.8, 69.2, 127.3, 128.0, 135.7, 135.9, 154.3, 172.8.

(2S, 4S)-N1 (t-butoxy-carbonyl)-4-(p-toluenesulfonyloxy) proline methyl ester (5)



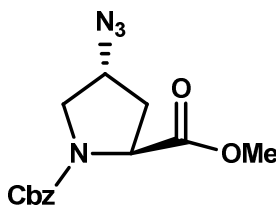
The compound **3** (7.8 g, 28.7 mmol), PPh_3 (8.27 g, 31.5 mmol) and methyl-*p*-toluenesulfonate (4.76 mL, 31.54 mmol) was dissolved in dry THF (100 mL) and cooled to 0 °C on ice bath under nitrogen. The mixture was stirred for 30 min at 0 °C. Diethyl azodicarboxylate (6.37 mL, 31.5 mmol) was added slowly with syringe. The reaction mixture was stirred at 0 °C for first 4 h then stirred at room temperature for 8 h after which THF was removed under vacuum. The resulting orange colored thick oil was dissolved in ethyl acetate (20 mL) to which petroleum ether (500 mL) was added. After triturating with spatula, the solution was kept at room temperature for 2 h. The white powder settled was filtered and the residue was washed with diethyl ether (6 x

60 mL) to obtained compound **5**. Yield: 10.50 g, (80%). Molecular Formula, C₂₁H₂₃NO₇S. [Mass: calculated 433, observed 434.12 M+1]. [α]: -24.44 (c 1.0 CH₂Cl₂).

¹H NMR (CDCl₃ 200 MHz) δ: 1.80 (1H, s), 2.34-2.40 (1H, m), 2.45 (3H, s), 2.48 (1H, s), 3.58 (2H, s), 3.65-3.73 (3H, m), 4.42-4.55 (1H, m), 5.02-5.23 (3H, m), 7.28-7.36 (7H, m), 7.73-7.77 (2H, d, *J* = 8.46)

¹³C NMR δ: 21.66, 36.8, 37.15, 52.07, 52.36, 52.48, 57.31, 57.56, 67.29, 67.37, 78.76, 127.71, 127.91, 127.97, 128.12, 128.51, 130.01, 133.52, 136.27, 145.25, 153.99, 154.36, 171.28, 171.53.

(2*S*, 4*R*)-N1-(benzyloxycarbonyl)-4-azido proline methyl ester (6**)**

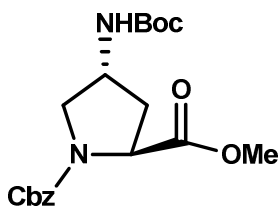


The compound **5** (10 g, 23 mmol) and NaN₃ (12 g, 284.8 mmol) was dissolved in dry DMF (80 mL). The solution was stirred for 8 h at 55-60 °C under nitrogen atmosphere. Water (620 mL) was added to reaction mixture and the aqueous layer was extracted with ethyl acetate (3 x 150 mL). The combined organic layer was washed with water (3 x 200 mL) followed by brine. The organic layer dried over anhydrous Na₂SO₄ and concentrated under vacuum. The crude product obtained was purified by silica gel chromatography (40% ethyl acetate/hexane) to afford **6** as a colourless thick oil. Yield: 7 g, (99%). Molecular Formula, C₁₄H₁₆N₄O₄. [Mass: calculated 304, observed 304 M⁺]. [α]: -45.75 (c 1.0 CH₂Cl₂).

¹H NMR (CDCl₃ 200 MHz) δ: 2.12-2.26 (1H, d, *J* = 13.3), 2.31-2.41 (1H, m), 3.46 (1H, m), 3.55 (1H, s), 3.66 (3H, s), 4.10 (1H, d, *J* = 2.76), 4.35-4.41 (1H, ddd, *J* = 8.78, 8.53), 4.90-5.18 (2H, m), 7.23-7.29 (5H, m).

¹³C NMR δ: 35.09, 36.11, 51.19, 51.49, 52.56, 56.71, 57.54, 57.81, 59.22, 67.29, 67.39, 78.00, 127.91, 128.02, 128.53, 136.31, 154.11, 154.53, 170.60, 171.82.

(2*S*,4*R*)-N1-(benzyloxycarbonyl)-4-(*t*-butoxycarbonylamino) proline methyl ester (7)

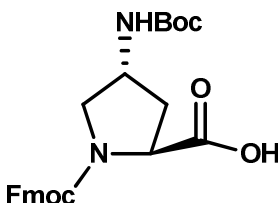


Compound **6** (7 g, 23 mmol) was dissolved in dry ethyl acetate (10 mL) to which Raney-Nickel (4 mL suspension in ethyl acetate) and di-*tert*-butyl-dicarbonate (6.1 g, 27 mmol) were added. The mixture was subjected to hydrogenation in Parr shaker (H_2 gas, 60-psi). After completion the reaction mixture was filtered through Whatman paper and the filtrate was concentrated under reduced pressure followed purification by silica gel chromatography to obtain compound **7**. Yield: 8.5 g, (97.6%). Molecular Formula, $C_{19}H_{26}N_2O_6$. [Mass: calculated 378, observed 378 M^+]. $[\alpha]$: -29.41 (c 1.0 CH_2Cl_2).

1H NMR ($CDCl_3$ 400 MHz) δ : 1.40 (9H,s), 2.20-2.20 (2H m), 3.31, 3.39 (1H, ddd, $J = 11.04, 10.79$), 3.70 (3H, s), 3.77-3.83 (1H, m), 4.27 (1H, m), 4.35-4.44 (1H, s), 4.97-5.26 (2H, m), 7.24-7.32 (5H, m).

^{13}C NMR δ : 28.33, 34.35, 35.30, 35.83, 36.93, 48.90, 49.54, 51.87, 52.26, 52.46, 57.60, 57.81, 59.23, 67.31, 67.42, 80.02, 127.89, 128.10, 128.45, 136.23, 136.34, 154.16, 154.80, 155.15, 172.40, 172.60.

(2*S*,4*R*)-N1-(fluorenylmethyloxycarbonyl)-4-(*t*-butoxycarbonylamino) proline (8)



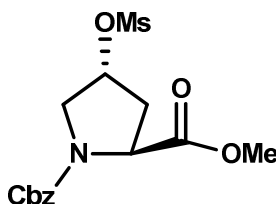
The ester compound **7** (7.27 g, 19 mmol) was subjected to hydrolysis using LiOH (3 eq in THF:H₂O, 1:1) for 1 h. THF was removed under vacuum and the aqueous layer was washed with ethyl acetate (3 x 50 mL) to remove THF and unreacted organic compound. The residual aqueous layer was acidified by KHSO₄ to pH 2 and extracted with ethyl acetate (3 x 60 mL) followed by concentration. The crude product obtained was purified by silica gel chromatography (65% ethyl acetate/hexane) to afford a white solid powder. This (7 g, 20 mmol) was dissolved in dry methanol (15 mL) to

which of 10% Pd/C (0.7 g) was added. The mixture was subjected to hydrogenation under H₂ gas (60-psi pressure) for 6 h in Parr reactor. Water (200 mL) was added to the reaction mixture which was filtered through Whatman filter paper and the filtrate was concentrated under reduced pressure. The product obtained as a white solid powder was dissolved in water: dioxane, 1:1 (60 mL). The pH was maintained at 10 by addition of 10% Na₂CO₃. The reaction mixture was stirred at 0 °C for 15 minutes. Fmoc-Cl (7.45 g, 29 mmol) was added in portion wise during 45 minutes maintained the temperature at 0 °C for first 4 h and then allowed to come at room temperature and was stirred for 18 h. The dioxane was removed under vacuum and the aqueous layer was washed with ethyl acetate (3 x 80 mL). The aqueous layer was acidified with KHSO₄ to pH 2 followed by extraction with ethyl acetate (3 x 70 mL). Concentration of solvent gave crude product which was purified by silica gel chromatography (80% ethyl acetate/hexane) to afford compound **8** as white solid. Yield: 6.95 g, (80%). Molecular Formula, C₂₅H₂₈N₂O₆. [Mass: calculated 452, observed 475 (M + Na⁺)]. [α]: -43.75 (c 1.0 CH₂Cl₂).

¹H NMR (CDCl₃ 400 MHz) δ: 1.44 (9H, s), 2.16 (1H, s), 2.32 (1H, s), 3.30-3.34 (1H, q), 3.77-3.89 (1H, q), 4.12-4.22 (1H, t), 4.29-4.36 (2H, m), 4.43 (1H, d), 4.66-4.75 (1H, t), 5.28 (1H, s), 7.23-7.39 (4H, m), 7.51-7.57 (2H, m), 7.68-7.74 (2H, dd, *J* = 7.28).

¹³C NMR δ: 28.36, 29.71, 34.47, 35.47, 36.99, 47.12, 48.61, 49.55, 51.89, 53.46, 57.36, 57.78, 67.73, 67.96, 80.33, 119.99, 125.10, 127.14, 127.77, 141.32, 143.58, 143.93, 154.45, 155.34, 155.46, 174.89, 175.89.

(2*S*,4*R*)-N1-(benzyloxycarbonyl)-4-(methanesulfonyloxy) proline methyl ester (9**)**



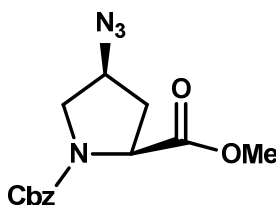
A solution of methyl ester **3** (13 g, 46.5 mmol) and triethylamine (9.74 mL, 69.9 mmol) in dry dichloromethane (100 mL) was cooled to 0 °C in ice bath under nitrogen. While stirring, methanesulfonyl chloride (4.32 mL, 55.9 mmol) was added drop-wise over a period of 3 h at 0 °C. After completion of reaction, the mixture was washed with

water followed by saturated aqueous brine solution. The organic layer was dried over anhydrous Na_2SO_4 and concentrated under vacuum. The crude material was purified by silica gel chromatography (30% ethyl acetate/hexane) to afford mesylated compound **9** as colorless thick oil. Yield: 16 g, (96.18%). Molecular Formula, $\text{C}_{15}\text{H}_{19}\text{NSO}_7$. [Mass: calculated 357, observed 357.05 (M^+)]. $[\alpha]$: -43.66 (c 1.0 CH_2Cl_2).

^1H NMR (CDCl_3 200 MHz) δ : 219-2.33 (1 H, m), 2.63-2.67 (1 H, m), 3.01 (3H, d), 3.76 (3H, s), 3.83 (1H, d), 3.98-4.15 (1H, m), 4.44-4.56 (1H, dd, $J = 7.96$), 4.98-5.26 (2H, m), 7.30-7.34 (5H, s).

^{13}C NMR δ : 36.1, 37.2, 38.5, 52.2, 57.1, 57.3, 67.3, 77.4, 127.7, 128.3, 136.0, 153.8, 154.3, 172.1.

(2S, 4S)-N1- (benzyloxycarbonyl)-4-azidoprolinemethyl ester (10)

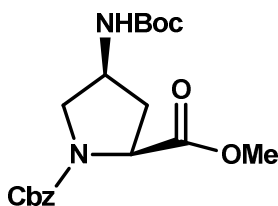


A solution of mesyl compound **9** (15.5 g, 43.4 mmol) and NaN_3 (22.57 g, 347.3 mmol) in dry DMF (120 mL) was stirred at 70 °C for 8 h under Nitrogen. Water (620 mL) was added to reaction mixture and the aqueous layer was extracted with ethyl acetate (3 x 150 mL). The combined organic layer was washed with water (6 x 50 mL) followed by brine and dried over anhydrous Na_2SO_4 . It was concentrated under vacuum and the crude material was purified by silica gel chromatography (40% ethyl acetate/hexane) to afford the azide compound **10** as colorless thick oil. Yield: 12.5 g, (94.7%). Molecular Formula, $\text{C}_{14}\text{H}_{16}\text{N}_4\text{O}_4$. [Mass: calculated 304, observed 305.05 ($\text{M}+1$)]. $[\alpha]$: -23.32 (c 1.0 CH_2Cl_2).

^1H NMR (CDCl_3 200 MHz) δ : 2.15-2.27 (1H, m), 2.36-2.50 (1H, m), 3.50-3.63 (2H, m), 3.76 (3H, s), 4.13-4.22 (1H, m), 4.44-4.53 (1H, m), 5.03-5.23 (2H, m), 7.31-7.38 (5H, m).

^{13}C NMR: 34.9, 35.9, 51.3, 52.2, 57.4, 57.6, 59.0, 67.1, 125.1, 127.9, 128.3, 128.8, 136.1, 153.9, 154.3, 171.4, 171.6.

(2S,4S)-N1-(benzyloxycarbonyl)-4-(*t*-butoxycarbonylamino) proline methyl ester (11)

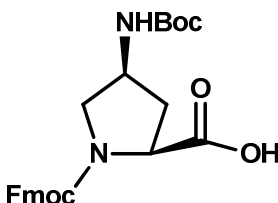


The azide **10** (12 g, 39.47 mmol) was dissolved in of dry ethyl acetate (14 mL) to which Raney-Nickel (6 mL of suspension in ethyl acetate) and di-*t*-butyl-dicarbonate (9.47 g, 43.43 mmol) were added. The reaction mixture was subjected to hydrogenation (55 psi pressure of H₂ gas) for 4 h in Parr-reactor. The reaction mixture was filtered and filtrate was concentrated under reduced pressure followed by purification by silica gel chromatography to obtained compound **11** as a colorless liquid. Yield: 14.4 g, (96.5%). Molecular Formula, C₁₉H₂₆N₂O₆. [Mass: calculated 378, observed 379.05 (M +1)]. [α]: -32.44 (c 1.0 CH₂Cl₂).

¹H NMR (CDCl₃ 400 MHz) δ: 1.42 (9H, s), 1.95-2.01 (1H, t), 2.46-2.49 (1H, t), 3.54-3.59 (2H, m), 3.78 (3H, s), 4.33-4.41 (2H, m), 5.01-5.20 (2H, m), 5.42 (1H, s), 7.28-7.34 (5H, m).

¹³C NMR δ: 29.36, 35.18, 35.89, 36.91, 49.12, 50.01, 51.52, 52.64, 53.45, 56.71, 57.63, 58.00, 67.37, 78.01, 79.72, 127.88, 128.00, 128.51, 136.25, 154.05, 154.74, 155.15, 173.06, 173.94.

(2S,4S)-N1-(fluorenylmethyloxycarbonyl)-4-(*t*-butoxycarbonylamino) proline (12)



The methyl ester **11**, 5.20 g (13 mmol) was subjected to hydrolysis using LiOH (3 eq. inTHF:H₂O, 1:1) for 1 h. THF was removed under vacuum and the aqueous layer was washed with ethyl acetate (3 x 50 mL) to remove THF and unreacted organic compound. The crude product obtained was purified by silica gel chromatography (70% ethyl acetate/hexane) to afford a white solid powder. This 5 g (13 mmol) was dissolved in dry methanol (15 mL) to which of 10% Pd/C (0.5 g) was added. The mixture was subjected to hydrogenation under H₂ gas (60-psi pressure) for 6 h in Parr

reactor. Water (200 mL) was added to the reaction mixture which was filtered through Whatman filter paper and the filtrate was concentrated under reduced pressure. The product obtained as a white solid powder was dissolved in water: dioxane, 1:1 (60 mL). The pH was maintained at 10 by addition of 10% Na₂CO₃. The reaction mixture was stirred at 0 °C for 15 minutes. Fmoc-Cl (5.326 g, 20 mmol) was added in portion wise during 45 minutes maintained the temperature at 0 °C for first 4 h and then allowed to come at room temperature and stirred for 18 h. The dioxane was removed under vacuum and the aqueous layer was washed with ethyl acetate (3 x 80 mL). The aqueous layer was acidified with KHSO₄ to pH 2 followed by extraction with ethyl acetate (3 x 70 mL). Concentration of solvent gave crude product which was purified by silica gel chromatography (80% ethyl acetate/hexane) to afford compound **12** as white solid. Yield: 5.09 g, (82%). Molecular Formula, C₂₅H₂₈N₂O₆. [Mass: calculated 452, observed 475.30 (M + Na⁺)]. [α]: -33.78 (c 1.0 CH₂Cl₂).

¹H NMR (CDCl₃, 400 MHz) δ: 1.44 (9H, s), 2.16-2.29 (1H, m), 2.35-2.47 (1H, m), 3.53 (1H, d, *J* = 11.04), 3.65 (1H, t), 4.21-4.25 (2H, t, *J* = 6.78), 4.31 (2H, m), 4.44-4.50 (2H, m), 5.34 (1H, s), 7.28-7.38 (4H, m), 7.52-7.57 (2H, dd, *J* = 7.53), 7.71-7.75 (2H, t, *J* = 8.28).

¹³C NMR δ: 28.2, 35.1, 36.9, 47.0, 49.9, 53.0, 57.4, 58.0, 67.8, 68.0, 119.0, 119.8, 124.8, 126.9, 127.6, 141.1, 143.4, 143.7, 143.9, 154.4, 155.2, 155.5, 174.6.

3.5.2 Peptide synthesis

All peptides were synthesized manually in a sintered vessel equipped with a stopcock. The readily available Rink amide resin with loading value 0.5-0.6 mmol/g was used and standard Fmoc chemistry was employed. The resin bound Fmoc group was first deprotected with 20% piperidine in DMF and the coupling reactions were carried out using *in situ* active ester method, using HBTU as a coupling reagent and HOBT as a recemization suppresser and DIPEA as a catalyst. All the materials used were of peptide synthesis grade (Sigma-Aldrich) and was used without further purification. Analytical grade DMF was purchased from Merck (India) and was distilled over P₂O₅ under vacuum at 45°C, stored over 4Å molecular sieves for 2 days before using for peptide synthesis.

3.5.2a Resin functionalization

The resin (2',4'-Dimethoxyphenyl-Fmoc-aminomethyl)-phenoxy, Catalog number 855047, 100-200 mesh) (100 mg) from Nova biochem was taken in sintered vessel (25 mL) and rinsed with 5 mL of dry DCM and filtered. The process was repeated 3 to 4 times and the resulting resin was kept for 2 h in DCM (10 mL) for swelling. The DCM was removed and rinsed 3 times with dry DMF and kept 2 h in dry DMF (10 mL) for swelling before the first coupling. The deprotection of 'Fmoc group' attached to the resin was done with 20% piperidine in DMF (3 x 5 mL) before proceeding for first amino acid coupling.

- The resin was washed and swollen in dry DCM for at least 2 h.
- Further washing and swelling with dry DMF for 2 h.
- Immediate coupling of 1st amino acid desired in C-terminus of peptide.

3.5.2b General method for solid phase peptide synthesis

All peptides were assembled on solid phase method by one amino acid per coupling. All amino acids were well dried over P₂O₅ in vacuum desiccator before coupling. Fmoc protecting group was used for main chain α -amino group. The *t*-Boc protection was used for side chain amine protection and was cleaved with 20% TFA in DCM for final cleavage of peptide from the resin. The peptide obtained after cleavage was stirred in 95% TFA in DCM for 2 h for complete deprotection *t* Boc.

3.5.2c Synthesis protocol for solid phase synthesis

The resin was pre-swollen overnight and the following steps were performed for each cycle.

- Wash with DMF 4 x 5 mL.
- 20% piperidine in DMF 2 x 5 mL (15 min for each) for deprotection of Fmoc group.
- Wash with DMF 3 x 5 mL, MeOH 3 x 5 mL and DCM with 3 x 5 mL.
- Test for complete deprotection (chloranil test).
- Coupling reaction with amino acid, DIPEA, HOBt and HBTU (3 eq.) in DMF (1 mL).
- Repeat of the coupling reaction in NMP for better yield.
- Test for completion of coupling reaction (chloranil test).

This cycle was repeated for every amino acid.

3.5.2d General procedure for peptide couplings on Rink Amide Resin

Fmoc-Xxx-OH (3 eq), HBTU (3 eq) and HOBT (3 eq) dissolved in DMF/NMP followed by *i*Pr₂NEt (7-8 eq) were added to the amino-functionalized resin in DMF. The mixture was kept for 2 h and last 5 min bubbled with N₂ and washed with DMF (3x), MeOH (3x) and DCM (3x). The loading value for peptide synthesis is taken as 0.5~0.6.

3.5.2e General procedure for Fmoc deprotection

20% piperidine in DMF was added to the resin and the reaction mixture was kept for 15 min, drained and the piperidine treatment was repeated 3 times. Finally the resin was washed with DMF (3x), MeOH (3x) and DCM (3x).

3.5.2f Chloranil test

This sensitive test has been developed for reliable detection of secondary amino groups but it will also detect primary amines. A few beads of resin were taken in a small test tube and were washed with methanol. To this of 2% acetaldehyde in DMF and 2% chloranil in DMF (5 drops each) were added. After a short mixing, the mixture was left at room temperature for 5 min and the beads inspected. A dark blue to green colour on by beads indicates a positive test (presence of NH group). A colourless to yellowish beads indicates then the test to be negative.

3.5.2g General procedure for acetylation

Triethylamine (20 eq) and acetic anhydride (20 eq) were added to the resin in DMF (\approx 100 mM). The mixture was kept for 1 h followed by bubbled with N₂ for 5 min and washed with DMF (3x), MeOH (3x) and DCM (3x).

3.5.2h Preparation of resin with peptide for cleavage

After the last coupling/acetylation, the resin was washed sequentially with DMF (5 x 10 mL), DCM (5 x 10 mL), toluene (5 x 10 mL) and finally with methanol (5 x 10 mL) and dried with nitrogen gas for 3 min. The resin sintered flask was dried in a vacuum desiccator over P₂O₅.

3.5.2i General procedure for cleavage of peptides from the solid support

The dry peptide-resin (20 mg) was taken in round-bottomed flask to which of 20% TFA in DCM (10 mL) and Triisopropylsilane (as scavengers) (2-3 drops) were added.

The resulting mixture was kept for 2 h by gentle shaking. The mixture was filtered through a sintered funnel and the resin was washed with 3 x 5 mL of above solution. The filtrate was collected in pear shape round-bottom flask and evaporated under reduced pressure. The resin was washed with MeOH (3 X 5 mL) and the washings were evaporated to dryness. The crude peptide obtained containing a N⁴-*t*Boc group, was deprotected by stirring the peptide solution with 95% TFA in DCM (10 mL) for 2 h. The TFA: DCM mixture was removed under reduced pressure. The residue obtained was dissolved in anhydrous methanol (0.4 mL) and to it anhydrous diethyl ether (4 x 1.5 mL) was added. The off-white precipitate obtained was centrifuged. The precipitation procedure was repeated twice to obtain peptide as a colourless powder.

3.5.3 High performance liquid chromatography (HPLC)

Peptides (**1-3**) were purified by reverse phase-HPLC on Waters 600 equipped with 2998-Photodiode array detector (PDA). Semi-preparative RP-C18 columns (250 x 10 mm, 10 μ m) of Altech-Altima make were used for peptides. The solvent system comprised of MeCN: Water (5:95) with 0.1% TFA for solution A and for solution B MeCN: Water (50:50), 0.1% TFA. A gradient of 0-100% at a flow rate of 3 mL/min was used to elute the peptide and the eluant was monitored at 220 nm. The peak corresponding to the peptide was collected and the fractions were frozen. Subsequently these peptides were concentrated by using speed vacuum. The purity of the final peptides were further analyzed on the Merck LiChrospher 100 RP-18 (250 x 4 mm, 5 μ M) column by using a gradient flow of 0 to 100% B in 30 min at a flow rate of 1.5 mL/min. The spectra acquisition, analysis and processing was done on Waters Empower-2154 software. The absorbance of the eluant was monitored at its corresponding wavelength and the purity was obtained from the integrator output. The purities of the hence purified peptides were found to be more than 95%.

3.5.4 MALDI-TOF characterization

MALDI-TOF mass spectra were obtained on either Voyager-Elite instrument (PerSeptive Biosystems Inc., Farmingham, MA) equipped with delayed extraction or on Voyager-De-STR (Applied Biosystems) instrument. Sinapinic acid and α -cyano-4-hydroxycinnamic acid (CHCA) both were used as matrix for peptides of which CHCA

was found to give satisfactory results. A saturated matrix solution was prepared with typical dilution solvent (50:50:0.1 Water: MeCN: TFA) and spotted on the metal plate along with the oligomers. The metal plate was loaded to the instrument and the analyte ions are then accelerated by an applied high voltage (15-25 kV) in reflector mode, separated in a field-free flight tube and detected as an electrical signal at the end of the flight tube. HPLC purified peptides were characterized through this method and were observed to give good signal to noise ratio, mostly producing higher molecular ion signals.

3.5.5 Circular dichroism (CD) spectroscopy

CD spectrometric study was carried out on JASCO J-715 spectropolarimeter using cylindrical, jacketed quartz cell (10 mm path length), which was connected to Julabo-UC-25 water circulator. To facilitate analysis of uncertainties, each set of spectra were measured using at least three individually prepared solutions. CD spectra were recorded using a spectral bandwidth of 1.0 nm at 25 °C with a time constant of 1 s and a step resolution of 1 nm. All the spectra were corrected for respective buffer condition and are typically averaged over 5-10 scans. CD data are given as mean residual molar ellipticities $[\theta] 10^3 \text{ deg cm}^2 \text{ dmol}^{-1}$. The spectra are the result of 5-10 accumulations. A quartz cell with a path length of 1 cm was used with solutions containing approximately 0.9 mL (100 μM) peptide solutions. For the spectra in buffer the blank spectrum of the solution was subtracted. All samples were equilibrated for at least 10 h before measurement except some TFE experiment.

Resolution:	1 nm
Band width:	1.0 nm
Sensitivity:	100 mdeg
Response:	1 sec
Speed:	100 nm/min
Accumulation:	5-10

All peptides had same peptide concentration (100 μM) for temperature dependent CD experiment. Buffers use for different pHs were Na-acetate buffer for pH 4.0-5.0; Na-phosphate buffer for pH 7.2 and Borate buffer pH 9.0-10.0. Prior to spectroscopic analysis samples annealed by immersing into a water bath at 90 °C and were cooled to

room temperature in the bath over a period of 4 h. These samples were incubated at room temperature for 12 h followed by 4 h at 2 °C. An additional ½ h incubation time was allowed in the instrument at the initial measurement temperature. The temperature was varied in steps of 5 °C and the spectra were recorded at each step. An equilibration period of 5 min was allowed at each temperature. Data processing and curve fitting was performed using MicroCal Origin 8.0 software. Ellipticity at specified wavelength for each temperature was plotted, normalized data was fitted to a sigmoid curve and the T_m (melting temperature) values are derived from the first derivative curve of the fit.

For the effect of urea the spectra were collected at 25 °C with a 1 nm resolution and a scan rate of 100 nm min⁻¹. Spectra are the averages of 10 scans. To facilitate analysis of uncertainties, each set of spectra were measured using at least three individually prepared solutions. The absorbance properties of high concentrations of urea prevent collection of spectra at wavelengths below about 215 nm.

For the effect of salt the buffer used for CD experiment was 5 mM Na-phosphate buffer at pH 7.2 in different concentration of salt (NaCl). Spectra were collected at 25 °C with a 1 nm resolution and a scan rate of 100 nm min⁻¹. Spectra are the averages of 10 scans. To facilitate analysis of uncertainties, each set of spectra were measured using at least three individually prepared solutions. The spectra could not be collected below a wavelength of 210 nm because of the absorbance properties of the salt solution.

For different solvent (TFE/*n*-propanol) experiment the CD spectra recorded in 100% TFE. Two set of spectra were collected a) Immediately after the sample preparation and b) Because the PPII↔PPI conversion is slow, samples were incubated in trifluoroethanol (TFE) for 6 d before measurements. Spectra were collected at 20 °C with a 1 nm resolution and a scan rate of 100 nm min⁻¹. Spectra are the averages of 10 scans and to facilitate analysis of uncertainties, each set of spectra were measured using at least three individually prepared solutions. Only very little difference was

observed in the intensity of CD spectra, with immediately recorded and after 6 day incubation collected spectra.

3.5.6 pK_a determination of monomers and oligomers

The pH of 0.1M solution (1mL) of compound in water was adjusted to 2.0 using conc. HCl. This solution was titrated with 10 µl aliquots of 1M NaOH. After each addition of NaOH solution, pH was recorded after the reading reached a stable. The pK_a values were derived from the first derivative curve of the data.

3.6 References

- 1 (a) Chockalingam, K.; Blenner M. and Banta, S. *Protein Eng. Des. Sel.*, **2007**, *20*, 155–161. (b) Mart, R. J.; Osborne R. D.; Stevens, M. M. and Ulijn, R. V. *Soft Matter*, **2006**, *2*, 822–835.
- 2 Anfinsen, C. B. *Science* **1973**, *181*, 223-230.
- 3 Creighton, T. E. *Proteins: Structures and Molecular properties*, 2nd ed.; W. H. Freeman & Co., **1993**, 173-174.
- 4 Figure taken from book Lehninger – *Biochemistry*, **2004**, 4th edition.
- 5 Figure taken from book Voet and Voet- *Biochemistry*, **2002**, 2nd edition.
- 6 Blanco, F. J.; Rivas, G.; Serrano, L. *Nat Struct Biol.*, **1994**, *1*, 584–590.
- 7 (a) Cowan, P. M.; Mcgavin, S. *Nature* **1955**, *176*, 501–503. (b) Traub, W.; Shmueli, U. *Nature* **1963**, *198*, 1165–1166.
- 8 Lu, K. P.; Finn, G.; Lee, T. H. and Nicholson L. K. *Nat Chem Biol.*, **2007**, *10*, 619-629.
- 9 (a) Stewart, D.E.; Sarkar, A. & Wampler, J.E. *J. Mol. Biol.* **1990**, *214*, 253–260. (b) Pal, D. and Chakrabarti, P. *J. Mol. Biol.* **1999**, *294*, 271–288.
- 10 Corey, R.B. and Pauling, L. *Proc. R. Soc. Lond. B*, **1953**, *141*, 10–20.
- 11 Edison, A.S. *Nat. Struct. Biol.*, **2001**, *8*, 201–202.
- 12 (a) Yaffe, M.B. *et al. Science*, **1997**, *278*, 1957–1960. (b) Thali, M. *et al. Nature*, **1994**, *372*, 363–365. (c) Christoph, C. and Wuthrich, K. *NMR Biopolymers*, **1981**, *20*, 2623–2633.
- 13 (a) Rath, A.; Davidson, R.; Deber, C. M. *Biopolymers (Pept. Sci.)* **2005**, *80*, 179–185. (b) Holt, M. R.; Koffer, A. A. *Trends Cell Biol.* **2001**, *11*, 38–46. (c) Kay, B. K.; Williamson, M. P.; Sudol, M. *FASEB J.* **2000**, *14*, 231–241.
- 14 (a) Brodsky, B.; Thiagarajan, G.; Madhan, B.; Kar, K. *Biopolymrs* **2008**, *89*, 345–353. (b) Shoulders, M. D.; Raines, R. T. *Annu. Rev. Biochem.* **2009**, *78*, 929–958.

- 15 (a) Farrera-Sinfreu, J.; Giralt, E.; Castel, S.; Albericio, F.; Royo, M. *J. Am. Chem. Soc.* **2005**, *127*, 9459–9468. (b) Fillon, Y. A.; Anderson, J. P.; Chmielewski, J. *J. Am. Chem. Soc.* **2005**, *127*, 11798–11803.
- 16 (a) Doose, S.; Neuweiler, H.; Barsch, H.; Saer, M. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 17400–17405. (b) Schuler, B.; Lipman, E. A.; Steinbach, P. J.; Klumke, M.; Eaton, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 2754–2759.
- 17 For reviews, see: (a) Shi, Z.; Chen, K.; Liu, Z.; Kallenbach, N. R. *Chem. Rev.* **2006**, *106*, 1877–1897. (b) Woody, R. W. *Adv. Biophys. Chem.* **1992**, *2*, 37–79. (c) Dukor, R. K.; Keiderling, T. A.; *Biopolymers*, **1991**, *31*, 1747–1761. (d) Krimm, S.; Tiffany, M. L. *Isr. J. Chem.* **1974**, *12*, 189–200.
- 18 (a) Knof, S.; Engel, J. *Isr. J. Chem.* **1974**, *12*, 165–177. (b) Mutter, M.; Wohr, T.; Gioria, S.; Keller, M. *Biopolymers* **1999**, *51*, 121–128. (c) Kakinoki, S.; Hirano, Y.; Oka, M. *Polym. Bull.* **2005**, *53*, 109–115.
- 19 (a) Erdmann, R. S.; Kuemin, M.; Wennemers, H. *Chimia* **2009**, *63*, 197–200. (b) Kuemin, M.; Sonntag, L. S.; Wennemers, H. *J. Am. Chem. Soc.* **2007**, *129*, 466–467. (c) Knof, S.; Engel, J. *Isr. J. Chem.* **1974**, *12*, 165–177. (d) Mutter, M.; Wohr, T.; Gioria, S.; Keller, M. *Biopolymers*, **1999**, *51*, 121–128. (e) Kakinoki, S.; Hirano, Y.; Oka, M. *Polym. Bull.* **2005**, *53*, 109–115. (f) Kuemin, M.; Schweizer, S.; Ochsenfeld, C. and Wennemers, H. *J. Am. Chem. Soc.*, **2009**, *131*, 15474–15482.
- 20 Adzhubei, A. A.; Sternberg, M. J.; *J Mol Biol*, **1993**, *229*, 472–493.
- 21 Stapley, B. J. and Creamer, T. P. *Protein Science*, **1999**, *8*, 587–595.
- 22 (a) McGregor, M. J.; Islam, S. A.; Sternberg, M. E.; *J Mol Biol*, **1987**, *198*, 295–310. (b) Muñoz, V.; Serrano, L.; *Proteins Struct Funct Genet*, **1994**, *20*, 301–311. (c) Dunbrack, R. L. (Jr); Karplus, M.; *Nature Struct Biol*, **1994**, *1*, 334–340. (d) Swindells, M. B.; MacArthur, M. W.; Thornton, J. M.; *Nature Struct. Biol.*, **1995**, *2*, 596–603.
- 23 IUPAC-IUB Commission on Biochemical Nomenclature. Abbreviations and symbols for the description of the conformation of polypeptide chains. *J Mol Biol*, **1970**, *52*, 1–17.
- 24 Song, K. and Kang, Y. K., *J. Phys. Chem. B*, **2006**, *110*, 1915–1927.
- 25 Ortega, A. F.; Casanovas, J.; Nussinov, R. and Alema'n, C., *J. Phys. Chem. B*, **2008**, *112*, 14045–14055.
- 26 Ortega, A. F.; Casanovas, J.; Assfeld, S. and Alema'n, C. *J. Org. Chem.* **2009**, *74*, 3101–3108.
- 27 Bretscher, L. E.; Jenkins, C. L.; Taylor, K. M.; DeRider, M. L. and Raines, R.T., *J Am Chem Soc*, **2001**, *123*, 777–778.
- 28 Cherng, J. H. and Raines, R. T., *Protein Sci.* **2006**, *15*, 74–83.
- 29 (a) DeRider, M.L.; Wilkens, S.J.; Waddell, M.J.; Bretscher, L.E.; Weinhold, F.; Raines, R.T. and Markley, J. L. *J. Am. Chem. Soc.* **2002**, *124*, 2497–2505. (b) Taylor, C.M.; Hardre', R. and Edwards, P.J.B. *J. Org. Chem.*, **2005**, *70*: 1306–1315.

- 30 Chiang, Y.-C.; Lin, Y.-J.; Horng, J. C. *Protein Sci.* **2009**, *18*, 1967–1977.
- 31 Owens, N. W.; Stetefeld, J.; Lattova, E. and Schweizer, F. *J. Am. Chem. Soc.* **2010**, *132*, 5036–5042.
- 32 Kuřmin, M.; Engelb, J. and Wennemers, H., *J. Pept. Sci.*, **2010**, *16*, 596–600.
- 33 Chellgren, B. W. and Creamer T. P. *J. Am. Chem. Soc.* **2004**, *126*, 14734-14735.
- 34 Aurora, R.; Creamer, T. P.; Srinivasan, R.; Rose, G. D. *J. Biol. Chem.* **1997**, *272*, 1413-1416.
- 35 Whittington, S. J.; Chellgren, B. W.; Hermann, V. M.; Creamer, T. P. *Biochemistry*, **2005**, *44*, 6269–6275.
- 36 Nozaki, Y. and Tanford, C. *J. Biol. Chem.*, **1963**, *238*, 4074-4081.
- 37 Robinson, D. R. and Jencks, W. P. *J. Am. Chem. Soc.* **1965**, *87*, 2462–2469.
- 38 Tiffany, M. L. and Krimm, S. *Biopolymers*, **1973**, *12*, 575-587.
- 39 Mattice, W.L. and Mandelkern, L., *Biochemistry*, **1970**, *9*, 1049–1058.
- 40 Rucker, A. L.; Creamer, T. P. *Proteins* **2002**, *11*, 980-985.
- 41 Arunkumar, A. I.; Kumar, T. K. *Biochim. Biophys. Acta.* **1997**, *1338*, 69-76.
- 42 (a) Steinberg, I. Z.; Harrington, W. F.; Berger, A.; Sela, M. and Katchalski, E. *J. Am. Chem. Soc.* **1960**, *82*, 5263. (b) Kurtz, J. and Harrington, W. F. *J. Mol. Biol.* **1966**, *17*, 440. (c) Schleich, T. and von Hippel, P. H. *Biopolymers*, **1969**, *7*, 861.
- 43 (a) Mandelkern, L.; Halpin, J. C.; Diorio, A. F. and Posner, A. S. *J. Am. Chem. Soc.* **1962**, *84*, 1383 (b) Kurtz, J. and Harrington, W. F. *J. Mol. Biol.*, **1966**, *17*, 440.
- 44 Harrington, W. F. and Sela, M., *Biochim. Biophys. Acta.*, **1958**, *27*,24.
- 45 Mandelkern, L. and Stewart, W. E. *Biochemistry*, **1964**, *3*, 1135.
- 46 Holzwarth, G. and Backman, K., *Biochemistry*, **1969**, *8*, 883.
- 47 Robinson, D. R. and Jencks, W. P. *J. Am. Chem. Soc.*, **1965**, *87*, 2462–2469.
- 48 Zimmerman, S. S. and Scheraga, H. A., *Macromolecules*, **1976**, *9*, 408- 416.
- 49 (a) Grathwohl, C.; Wiřthrich, K. *Biopolymers*, **1976**, *15*, 2025-2041 (b) Rich, D. H.; Jasensky, R. D. *J. Am. Chem. Soc.* **1980**, *102*, 1112-1119.
- 50 (a) Kelly, M. A.; Chellgren, B.W.; Rucker, A.L.; Troutman, J. M.; Fried, M.G.; Miller, A.F. and Creamer, T.P. *Biochemistry*, **2001**, *40*, 14376–14383. (b) Eker, F.; Griebenow, K. and Schweitzer-Stenner, R. *J. Am. Chem. Soc.* **2003**, *125*, 8178–8185. (C) Rucker, A.L. and Creamer, T.P., *Protein Sci.*, **2002**, *11*, 980–985. (d) Chen, K.; Liu, Z.G. and Kallenbach, N.R. *Proc. Natl. Acad. Sci.*, **2004**, *101*, 15352–15357.

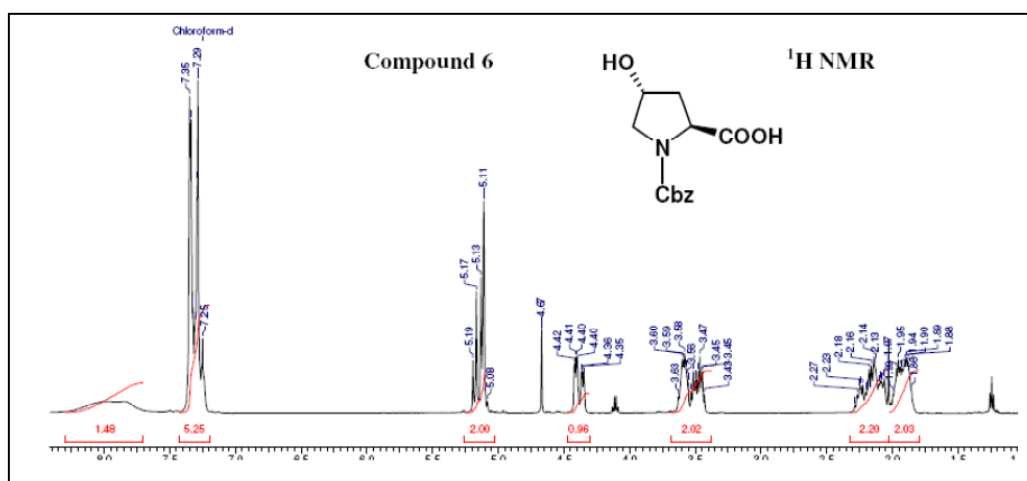
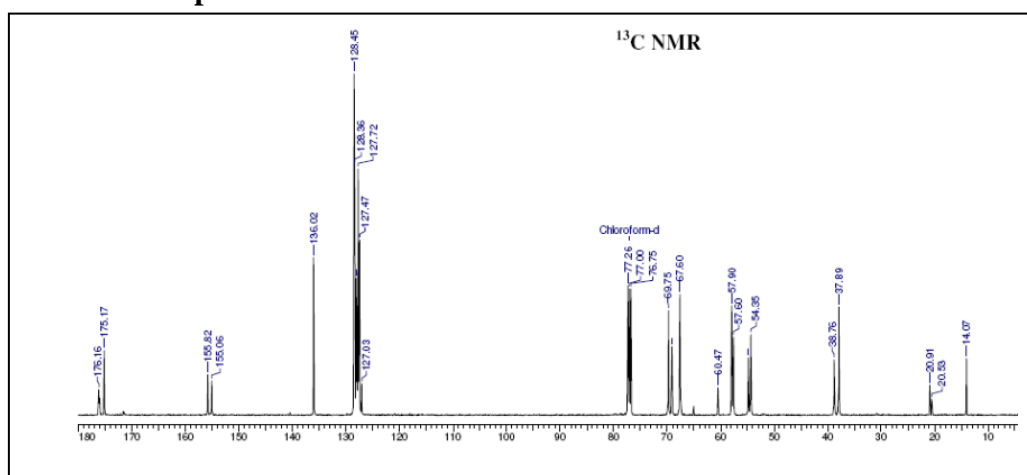
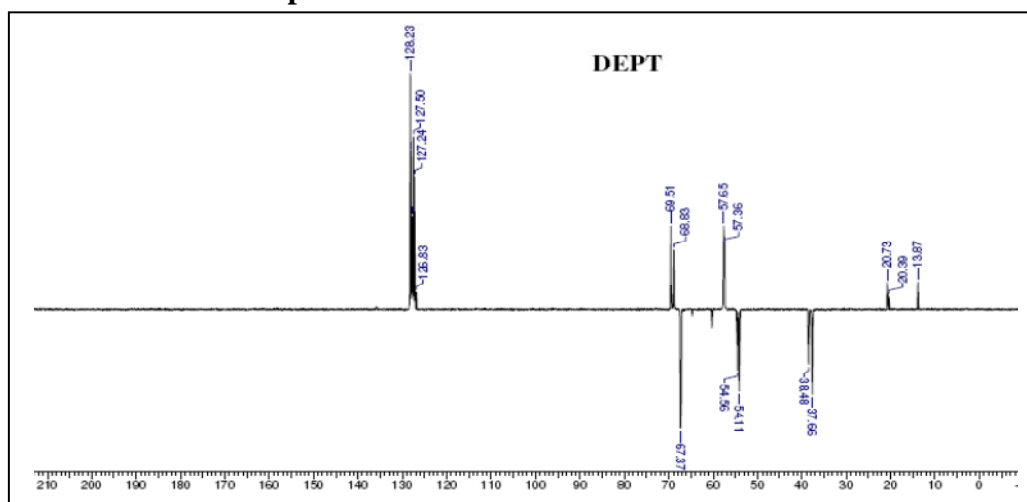
- 51 (a) Eckert, B.; Martin, A.; Balbach, J.; Schmid, F. X. *Nat. Struct. Mol. Biol.* **2005**, *12*, 619-623. (b) Wedemeyer, W.J.; Welker, E.; Scheraga, H. A. *Biochemistry* **2002**, *41*, 14637-14644. (c) Reimer, U.; Scherer, G.; Drewello, M.; Kruber, S.; Schutkowski, M.; Fischer, G. *J. Mol. Biol.* **1998**, *279*, 449-460.
- 52 Brazin, K. N.; Mallis, R. J.; Fulton, D. B.; Andreotti, A. H. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 1899.
- 53 Schmid, F. X.; Mayr, L. M.; Mücke, M. Schönbrunner, E. R. *Advan. Protein Chem.* **1993**, *44*, 25-66.
- 54 Dugave, C. and Demange, L. *Chem. Rev.* **2003**, *103*, 2475-2532.
- 55 (a) Sadler, K.; Eom, K. D.; Yang, J. L.; Dimitrova, Y.; Tam, J. P. *Biochemistry* **2002**, *41*, 14150-14157. (b) Fernández Carneado, J.; Kogan, M. J.; Castel, S.; Pujals, S.; Giralt, E. *Angew. Chem., Int. Ed.* **2004**, *43*, 1811-1814. (c) Fernández-Carneado, J.; Kogan, M. J.; Pujals, S.; Giralt, E. *Biopolymers* **2004**, *76*, 196-203.
- 56 Crespo, L.; Sanclimens, G.; Montaner, B.; Pe'rez-Toma's, R.; Royo, M.; Pons, M.; Albericio, F. and Giralt, E. *J. Am. Chem. Soc.* **2002**, *124*, 8876-8883.
- 57 Farrera-Sinfreu, J.; Giralt, E.; Castel, S.; Albericio, F.; Royo, M. *J. Am. Chem. Soc.* **2005**, *127*, 9459-9468.
- 58 Fillon, Y. A.; Anderson, J. P.; Chmielewski, J. *J. Am. Chem. Soc.* **2005**, *127*, 11798-11803.
- 59 Labeit, S and Kolmerer, B. *Science*, **1995**, *270*, 293-296.
- 60 Gutierrez-Cruz, G.; Van Heerden, A. H.; Wang, K. *J Biol Chem*, **2001**, *276*, 7442-7449.
- 61 Bochicchio, B.; Tamburro, A. M. *Chirality*, **2002**, *14*, 782-792.
- 62 Wang, C. Q. and Bayley, H. *Curr Biol.*, **1997**, *7*, 677-678.
- 63 Royce PM, Steinmann B (editors). *Connective tissue and its heritable disorders.* **1993**, New York:, Wiley.
- 64 Robson, P.; Wright, G. M.; Sitarrz, E.; Maiti, A.; Rawat, M.; Youson, J. H. and Keeley F. W. *J Biol Chem*, **1993**, *268*, 1440-1447.
- 65 MacArthur, M. W. and Thornton, J. M. *J. Mol. Biol.*, **1991**, *218*, 397-412.
- 66 Williamson, M. P. *Biochem. J.*, **1994**, *297*, 249-260.
- 67 Sparks, A. B.; Quilliam, L. A.; Thorn, J. M.; Der, C. J. and Kay, B. K. *Biol. Chem.*, **1994**, *269*, 23853-23856.
- 68 Stauffer, T. P.; Martenson, C. H.; Rider, J. E.; Kay, B. K. and Meyer, T. *Biochemistry*, **1997**, *36*, 9388-9394.

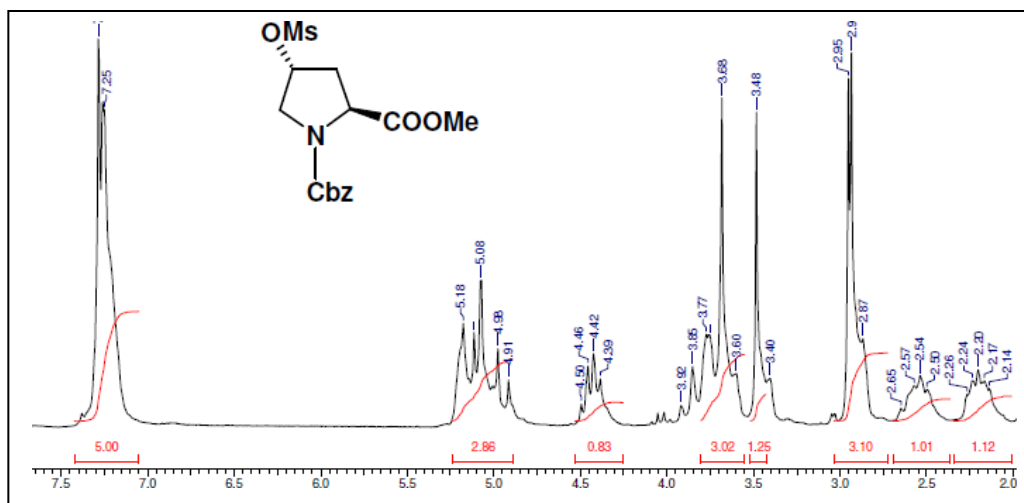
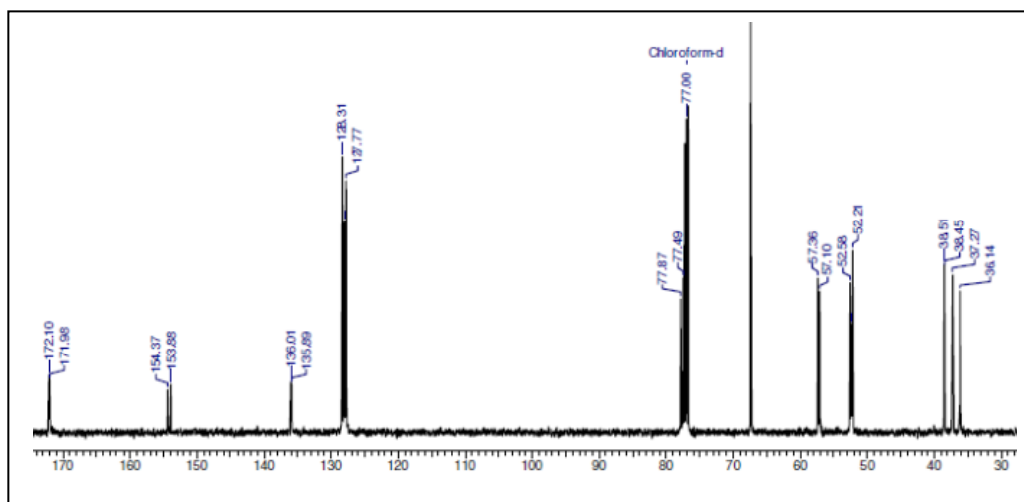
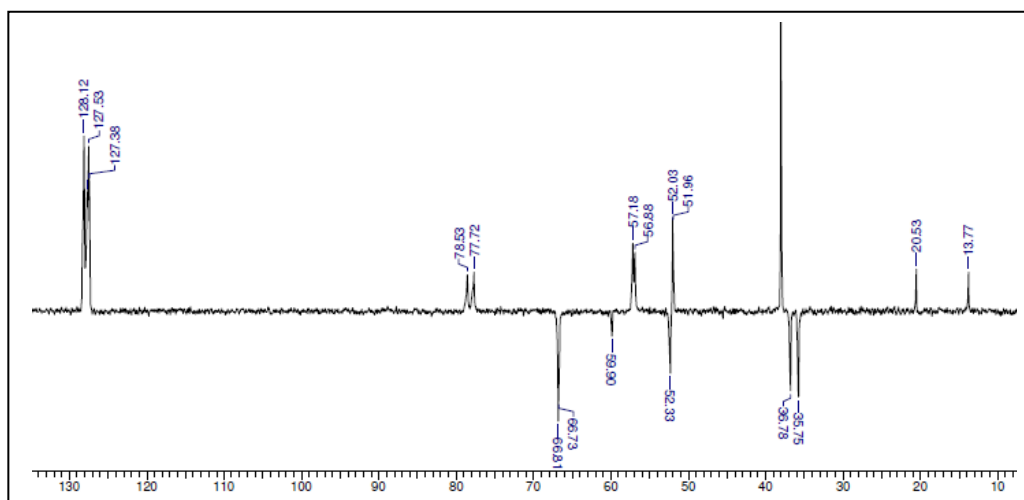
- 69 (a) Derossi, D.; Joliot, A. H.; Chassaing, G. and Prochiantz, A. *J. Biol. Chem.*, **1994**, *269*, 10444–10450. (b) Derossi, D.; Calvet, S.; Trembleau, A.; Brunissen, A.; Chassaing, G. and Prochiantz, A.; *J. Biol. Chem.*, **1996**, *271*, 18188–18193. (c) Fenton, M.; Bone, N. and Sinclair, A. *J. Immunol. Meth.*, **1998**, *212*, 41–48.
- 70 Vives, E.; Brodin, P. and Lebleu, B. *J. Biol. Chem.*, **1997**, *272*, 16010–16017.
- 71 Raj, P. A.; Marcus, E. and Edgerton, M. *Biochemistry*, **1996**, *35*, 4314–4325.
- 72 Shi, J.; Ross, C. R.; Leto, T. L. and Blecha, F. *Proc. Natl. Acad. Sci. USA*, **1996**, *93*, 6014–6018.
- 73 (a) Elmore, D. T., Peptide Synthesis. Amino Acids, Peptides, Proteins. 2002, *33*, 83–134. (b) Humphrey, J. M.; Chamberlin, A. R. *Chem. Rev.* **1997**; *97*, 2243–2266
- 74 (a) König, W. and Geiger, R. *Chem. Ber.*, **1970**, *103*:788–798 (b) Robertson, N.; Jiang, L.; Ramage R. *Tetrahedron*, **1999**, *55*, 2713–2720.
- 75 Jiang, L.; Davision, A.; Tennant, G.; Ramage, R.; *Tetrahedron*, **1998**, *54*, 14233–14254.
- 76 Carpino, L.A.; Ionescu, D.; El-Faham, A.; Beyermann, M.; Henklein, P.; Hanay, C. *Org. Lett.*, **2003**, *5*, 975–977.
- 77 (a) Nishiyama, Y.; Tanaka, M.; Saito, S.; Ishizuka, S.; Mori, T.; Kurita, K. *Chem. Pharm. Bull.* **1999**; *47*, 576–578 (b) Nishiyama, Y.; Ishizuka, S.; Kurita, K. *Tetrahedron Lett.*, **2001**, *42*, 8789–8791.
- 78 Nishuchi Y.; *et. al.*, *PNAS USA*, **1998**, *95*, 13549-13554.
- 79 Hiebl J., *et. al.*, *The Journal of Peptide Research*, **1999**, *54*, 54-65.
- 80 Barlos, K.; Chatzi, O.; Gatos, D. and Stavropoulos, G., *International Journal of Peptide and Protein Research*, **1991**, *37*, 513-520.
- 81 Chen, Y.; Edward, A. and Kent, S. *J. Am. Chem. Soc.*, **1999**, *121*, 8720-8727.
- 82 Sik, L.; Victor, L.; *Biopolymers*, **2003**, *69*, 270–28.
- 83 Aleksandr, V.; Nataliya, S.; Sergei, V. and Sanford, A., *J. Am. Chem. Soc.* **2005**, *127*, 7712-7720
- 84 Lam, S. L. and Hsu, V. L. *Biopolymers*, **2003**, *69*, 270–281.
- 85 (a) Whitmore, L.; Wallace, B. A. *Biopolymers*, **2008**, *89*, 392–400. (b) Greenfield, N. J. *Nature protocols*, **2006**, *1*, 2876–90. (c) Analysis of Circular Dichroism Data, Norma J. Greenfield, *METHODS IN ENZYMOLOGY*, *383*, 282-317.
- 86 Drake, A. F.; Siligardi, G.; Gibbons, W. A. *Biophys Chem*, **1988**, *31*, 143–146.
- 87 Venugopal, M. G.; Ramshaw, J.; Zhu, D.; Brodsky, B. *Biochemistry*, **1994**, *33*, 7948–7956.
- 88 Greenfield, N. and Fasman, G. D. *Biochemistry*, **1969**, *8*, 4108–4116.
- 89 Johnson, W. C.; Jr, Tinoco, I. Jr.; *J Am Chem Soc*, **1972**, *94*, 4389–4392.

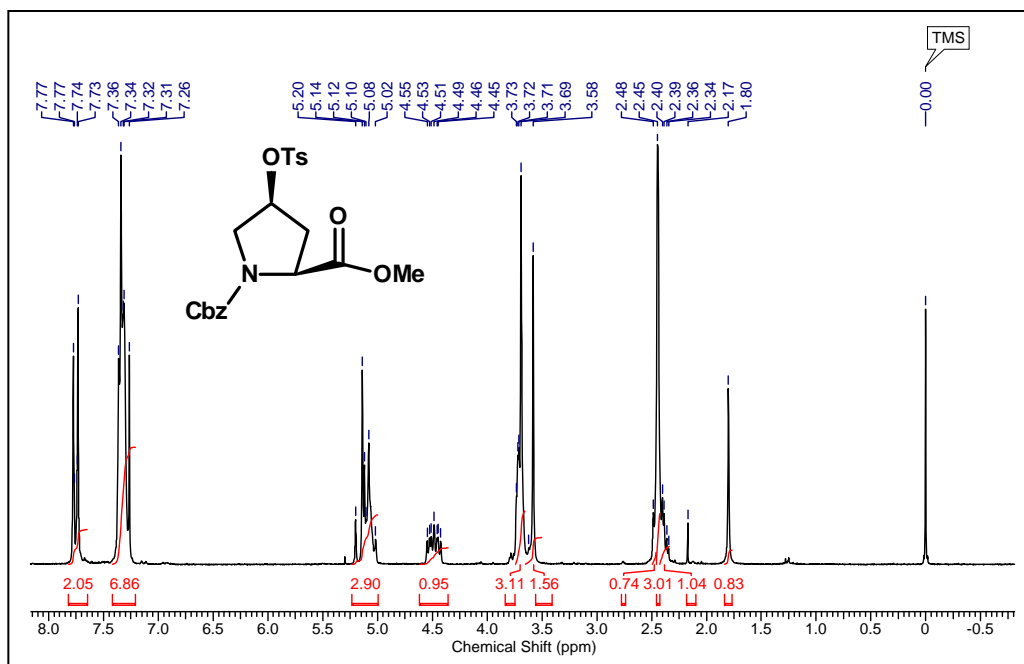
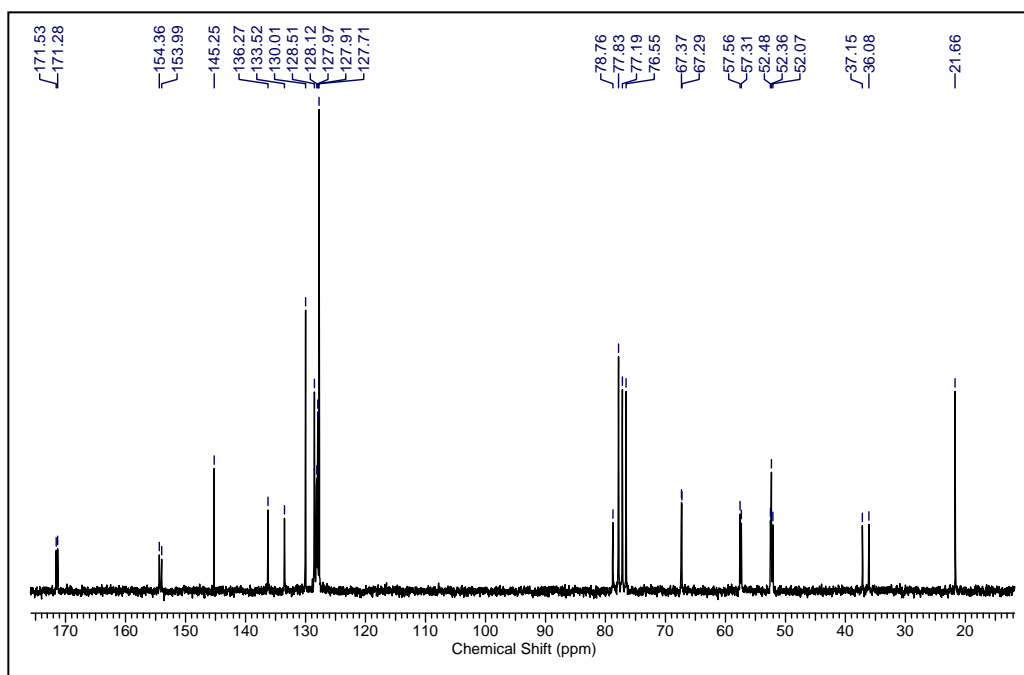
- 90 (a) Woody, R. W. *Adv. Biophys. Chem.* **1992**, *2*, 37–79. (b) Robert W. Woody, *J. Am. Chem. Soc.*, **2009**, *131*, pp 8234–8245
- 91 Dukor, R. K. and Keiderling, T. A. *Biopolymers*, **1991**, *31*, 1747–1761.
- 92 (a) Babu, I. R.; Ganesh, K. N. *J. Am. Chem. Soc.* **2001**, *123*, 2079–2080. (b) Umashankara, M.; Babu, I. R.; Ganesh, K. N. *Chem. Comm.* **2003**, 2606–2607.
- 93 (a) Holmgren, S. K.; Taylor, K. M.; Bretscher, L. E.; Raines, R. T. *Nature* **1998**, *392*, 666–667. (b) Hodges, J. A.; Raines, R. T. *J. Am. Chem. Soc.* **2003**, *125*, 9262–9263.
- 94 Vojtkovsky, T. *Pept. Res.* **1995**, *8*, 236–237.
- 95 Bodansky, M.; Bodansky, A. *The Practice of Peptide Synthesis* **1984**, Springer-Verlog, Berlin.
- 96 Glastone, S. A. *Text Book of Physical Chemistry 2nd ed.* Macmillan India Ltd., pp 1002–1005.
- 97 Feng, Y.; Melacini, G.; Taulane, J. P.; Goodmann, M. *J. Am. Chem. Soc.* **1996**, *118*, 10351–10358.
- 98 Seebach, D.; Overhand, M.; Kiihnle, F. N. M.; Martinoni, B. *Helv. Chim. Acta* **1996**, *79*, 913–941.
- 99 Luo, Z; Akerman, B; Zhang, S and Norden, B. *Soft Matter*, **2010**, *6*, 2260–2270.
- 100 (a) Podstawka, E.; Ozaki, Y.; Proniewicz, L. M. *Appl. Spect.* **2005**, *59*, 1516–1526. (b) Podstawka, E.; Ozaki, Y.; Proniewicz, L. M. *Appl. Spect.* **2004**, *58*, 570–580.
- 101 Sonar, M. V. and Ganesh, K. N. *Org. Lett.*, **2010**, *12*, 5390–5393.
- 102 Izatt, R. M.; Hansen, L. D.; Rytting, J. H.; Chirstensen, J. J. *J. Am. Chem. Soc.* **1965**, *87*, 2760–2761.
- 103 (a) Reynolds, W. F.; *Prog. Phys. Org. Chem.* **1983**, *14*, 165–203 (b) Thibaudeau, C.; Plavec, J.; Chattopadhyaya, J. *J. Org. Chem.* **1996**, *61*, 266–286.
- 104 Rucker, A. L.; Creamer, T. P. *Proteins* **2002**, *11*, 980–985.
- 105 (a) Shoulders, M. D.; Kotch, F. K.; Choudhary, A.; Guzei, I. A.; Raines, R. T. *J. Am. Chem. Soc.* **2010**, *132*, 10857–10865. (b) Kuemin, M.; Nagel, Y. A.; Schweizer, S.; Monnard, F. W.; Ochsenfeld, C.; Wennemers, H. *Angew. Chem., Int. Ed.* **2010**, *49*, 6324–6327.
- 106 (a) Tsai, C.J.; Zheng, J.; Zanuy, D.; Haspel, N.; Wolfson, H.; Alema'n, C.; Nussinov, R. *Proteins* **2007**, *68*, 1–12. (b) Zhao, X.; Pan, F.; Xu, H.; Yaseen, M.; Shan, H.; Hauser, C. A. E.; Zhang, S.; Lu, J. R. *Chem. Soc. Rev.* **2010**, *39*, 3480–3498. (c) Apostolovic, B.; Maarten, D. M.; Klok, H.-M. *Chem. Soc. Rev.* **2010**, *39*, 3541–3550.

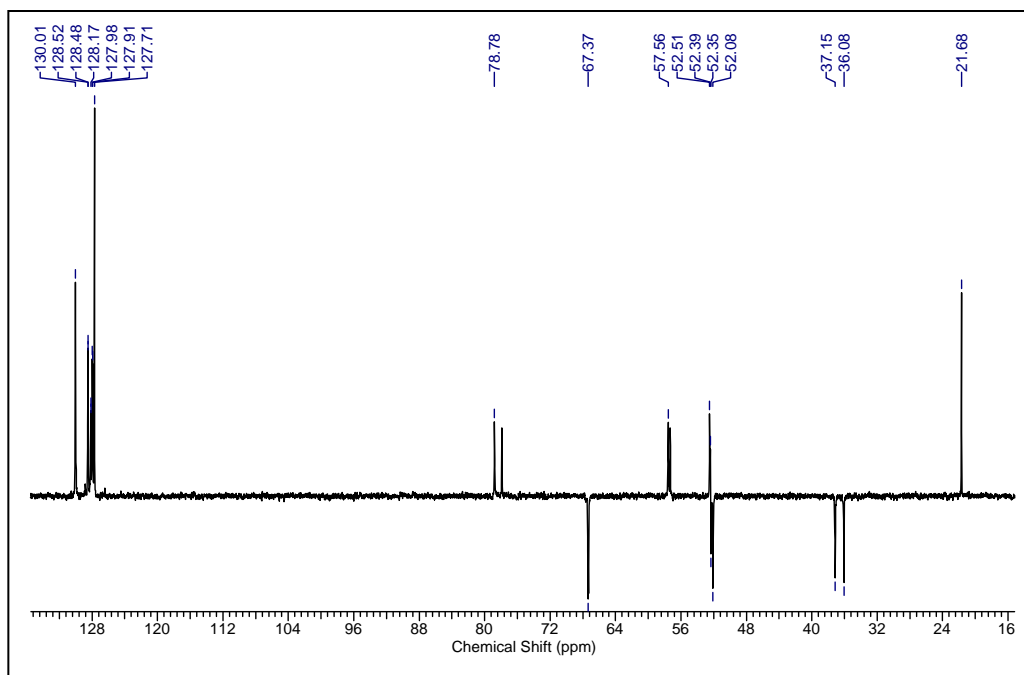
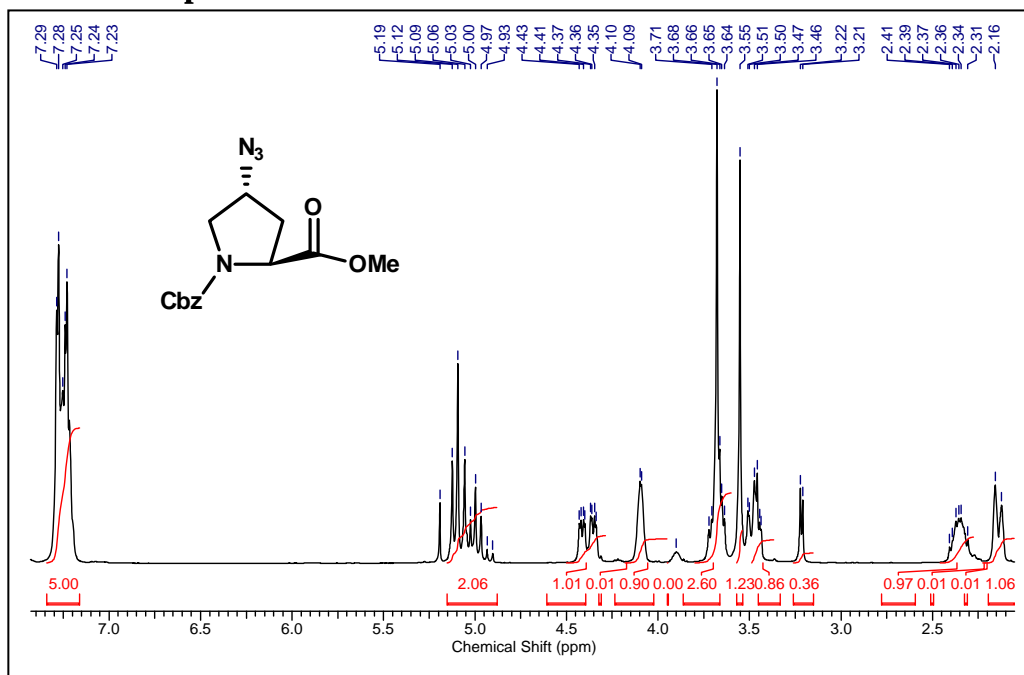
3.7 Appendix 1: Characterization data of synthesized compounds and peptides

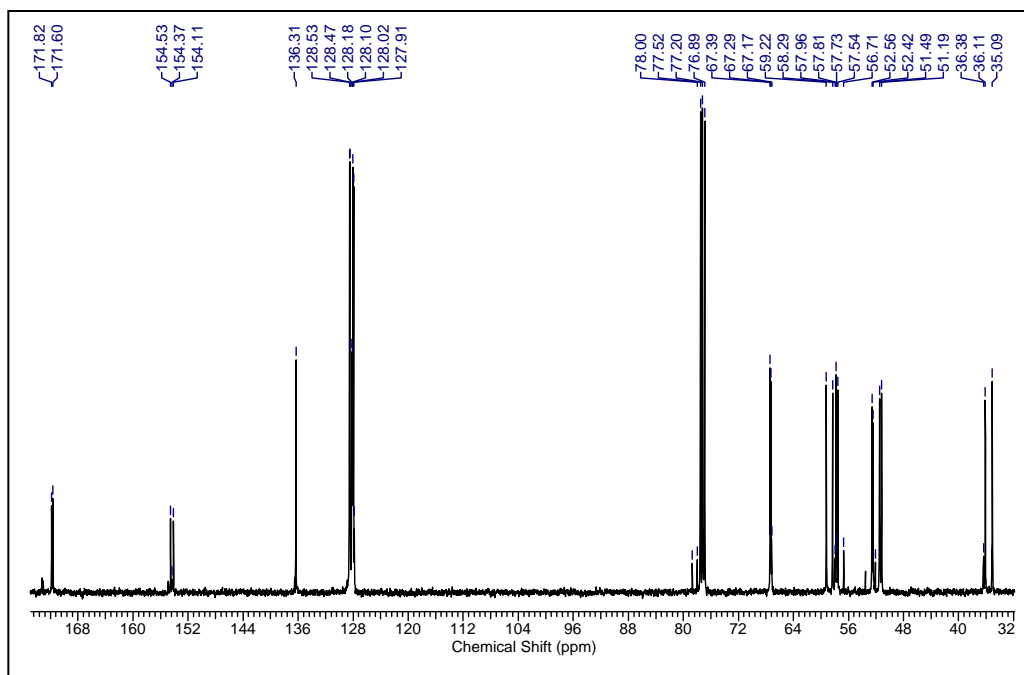
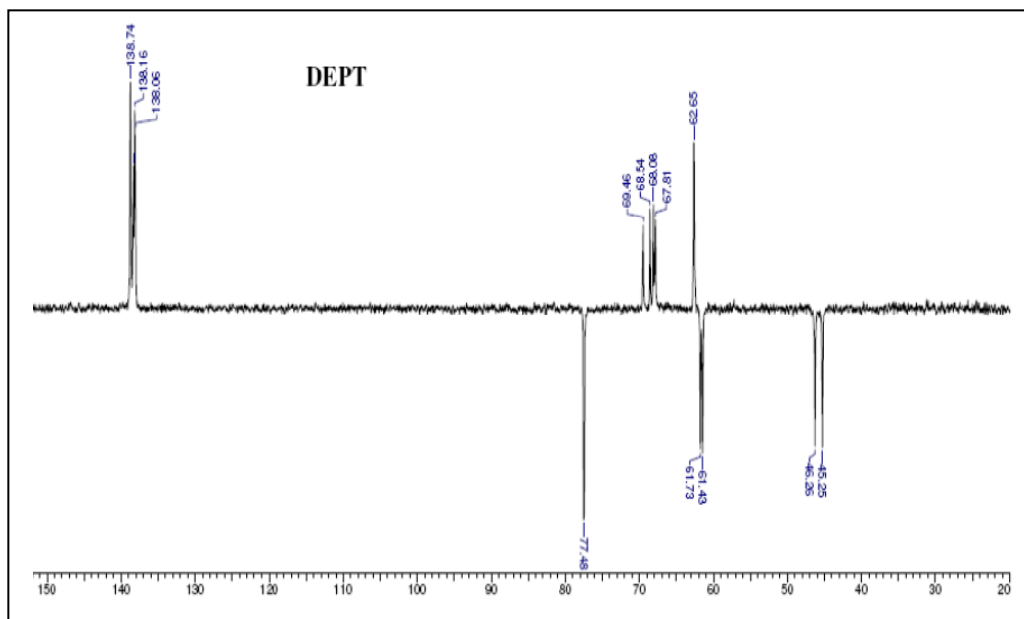
Entry	Page No.
^1H , ^{13}C and DEPT NMR spectra of compounds (2-12)	101-114
HPLC of Peptides (1-3)	114-115
MALDI-TOF of peptides (1-3)	116-118

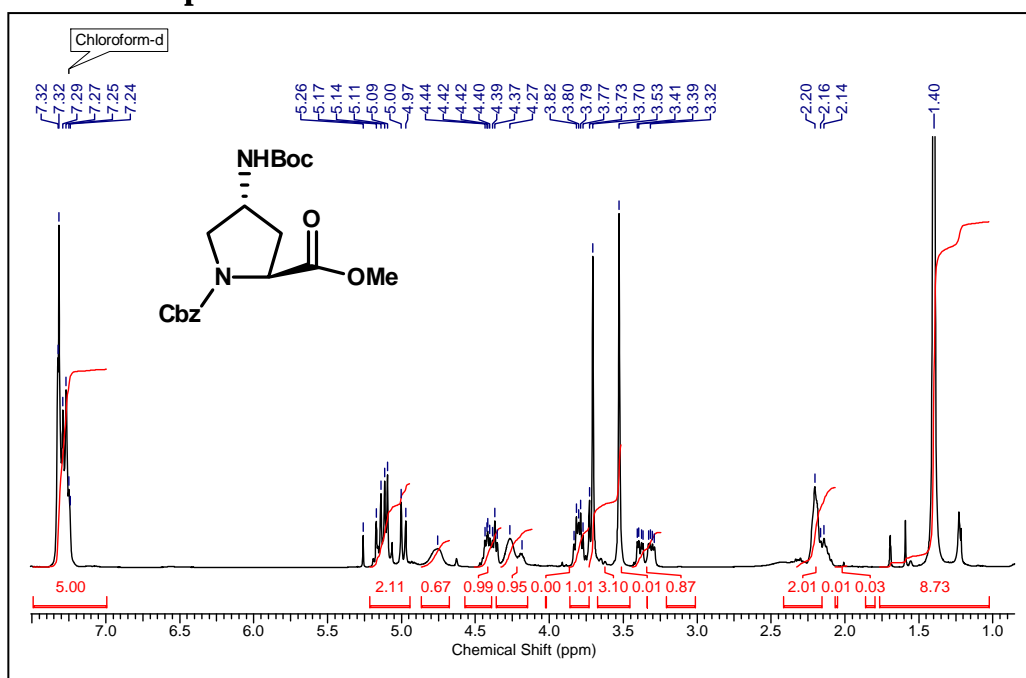
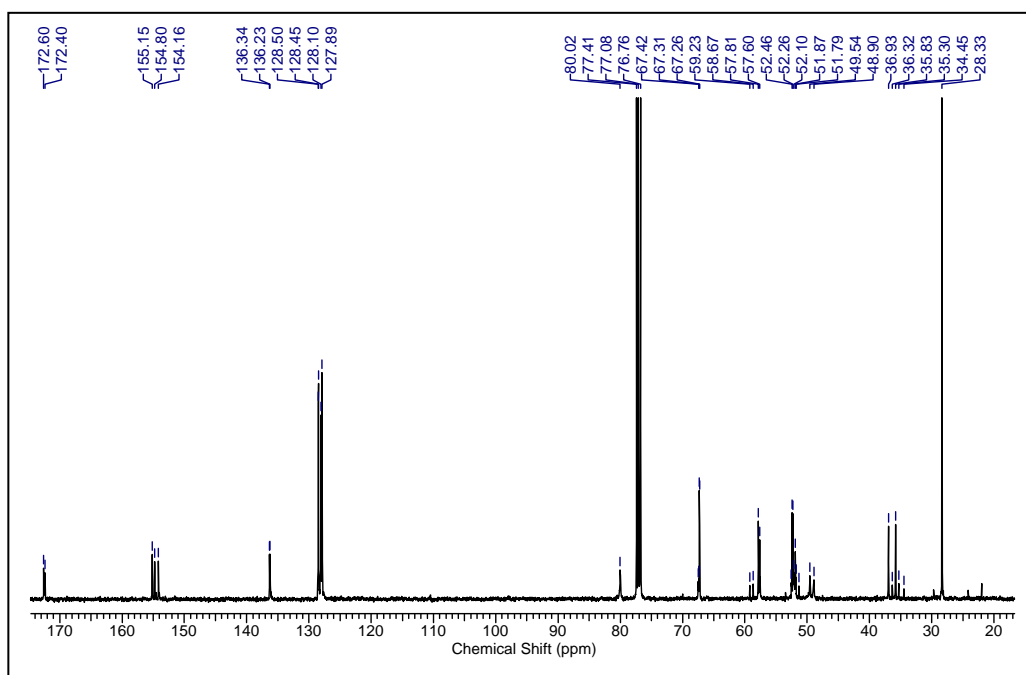
(A) ^1H , ^{13}C and DEPT NMR spectra of compounds (2-12) ^1H NMR of compound 2 ^{13}C NMR of compound 2DEPT ^{13}C NMR of compound 2

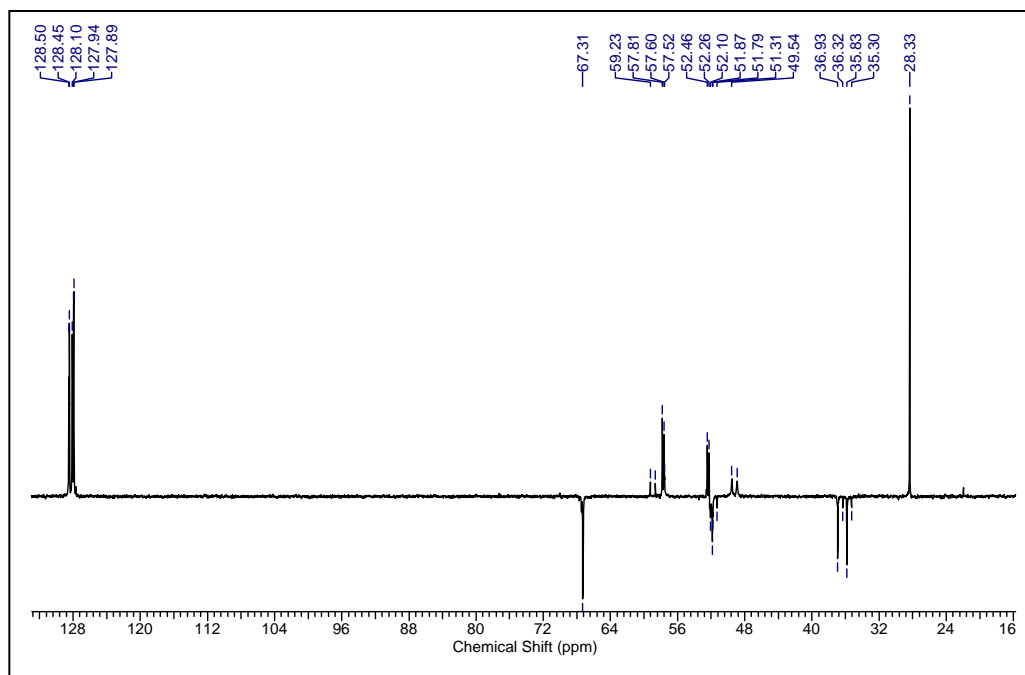
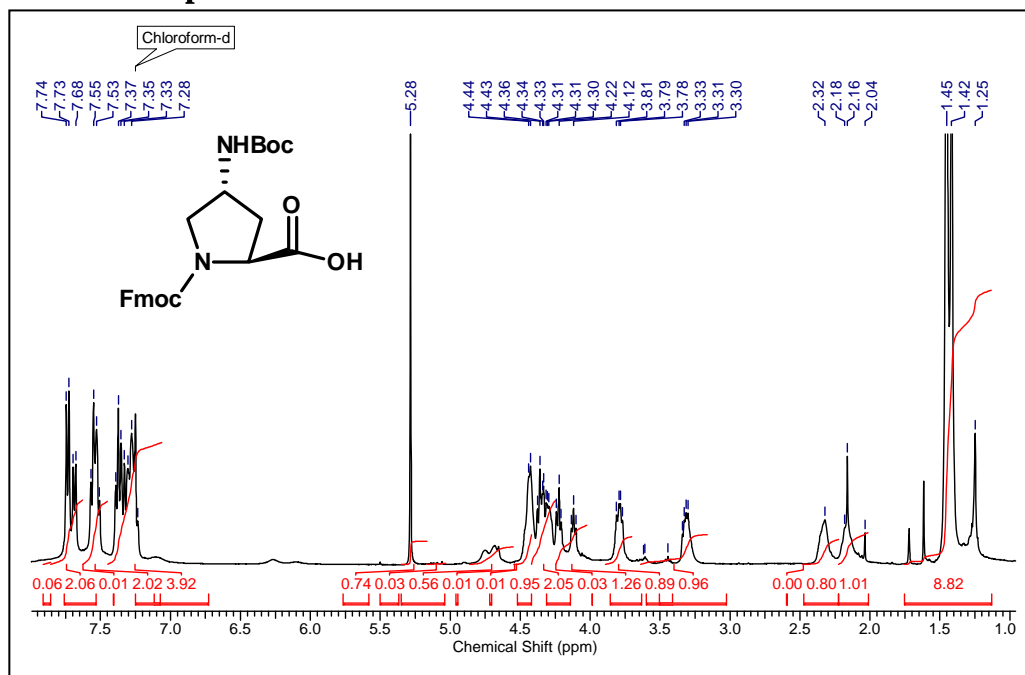
^1H NMR of compound 3 **^{13}C NMR of compound 3****DEPT ^{13}C NMR of compound 3**

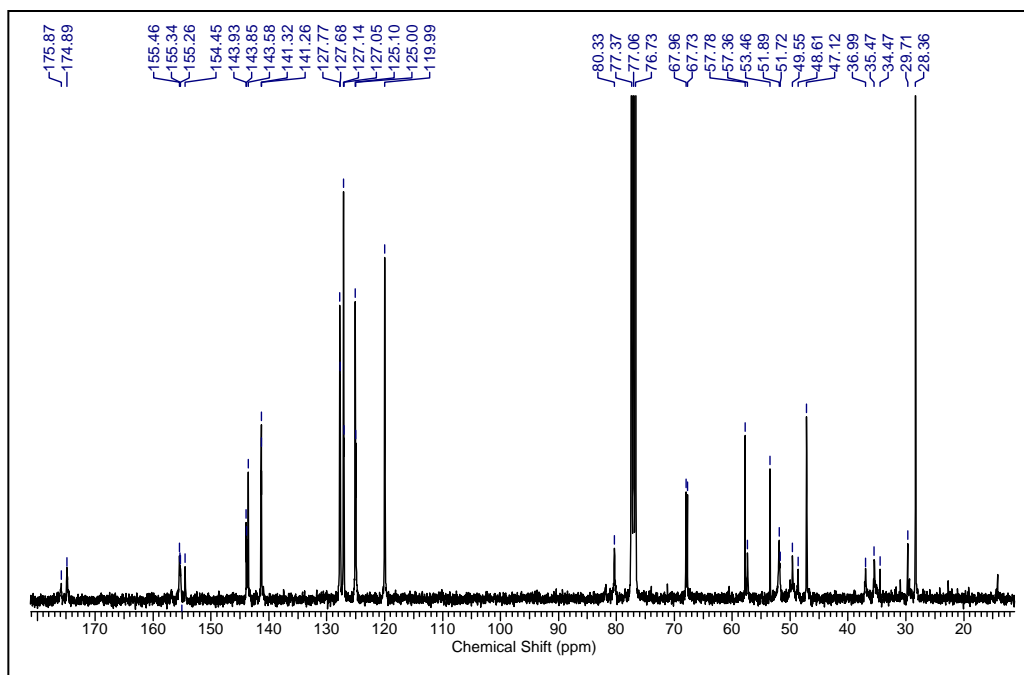
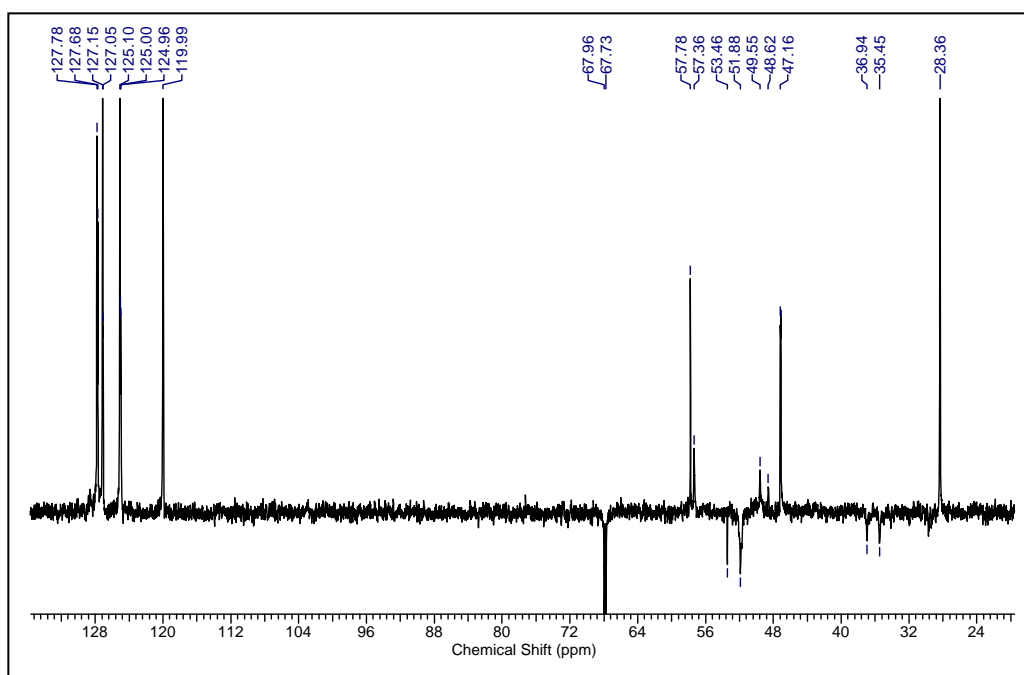
^1H NMR of compound 5 **^{13}C NMR of compound 5**

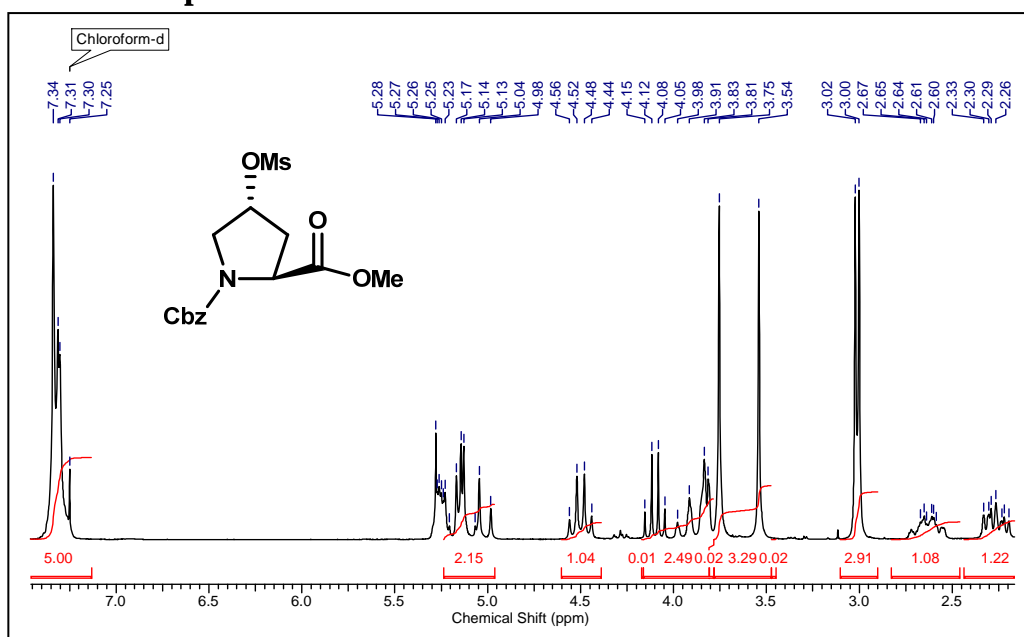
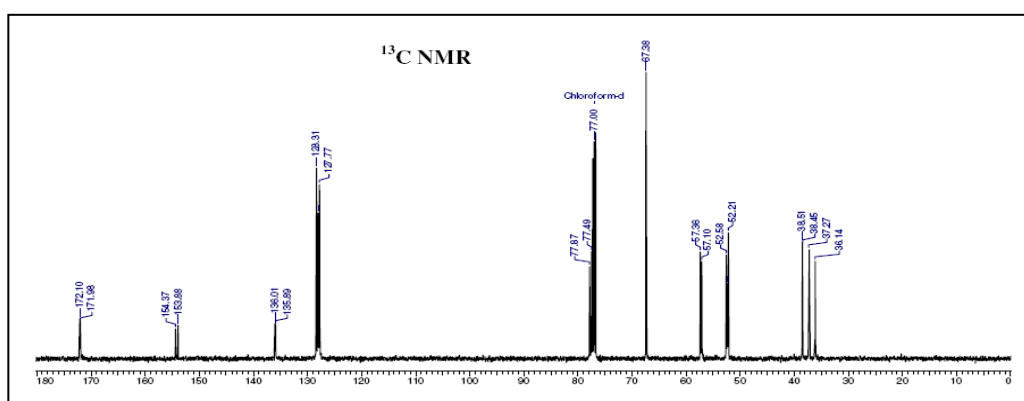
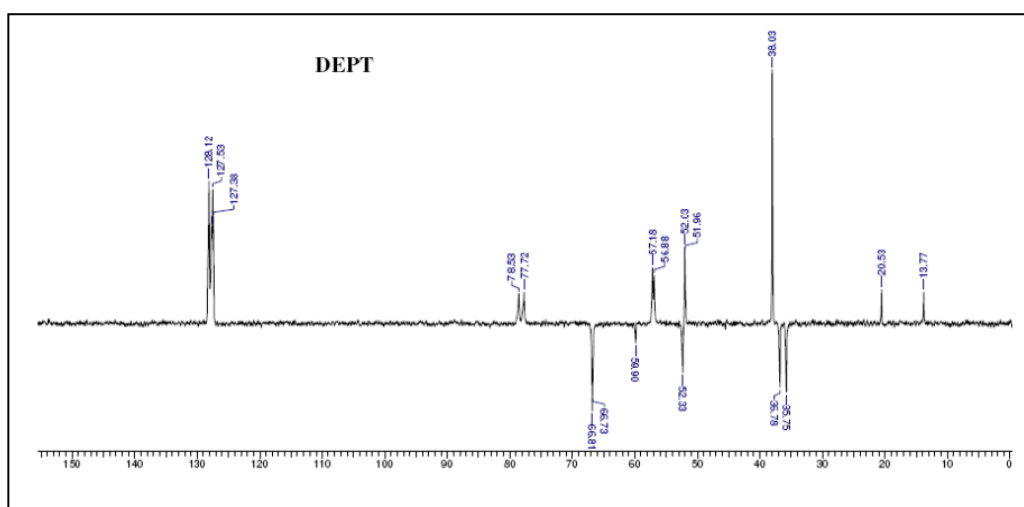
DEPT ^{13}C NMR of compound 5 ^1H NMR of compound 6

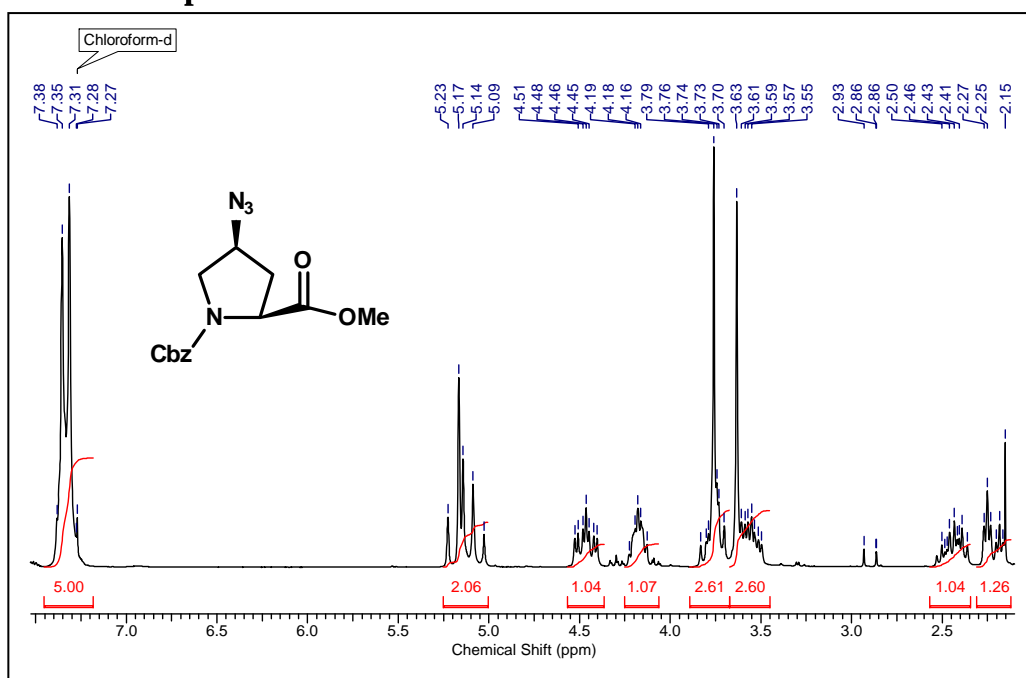
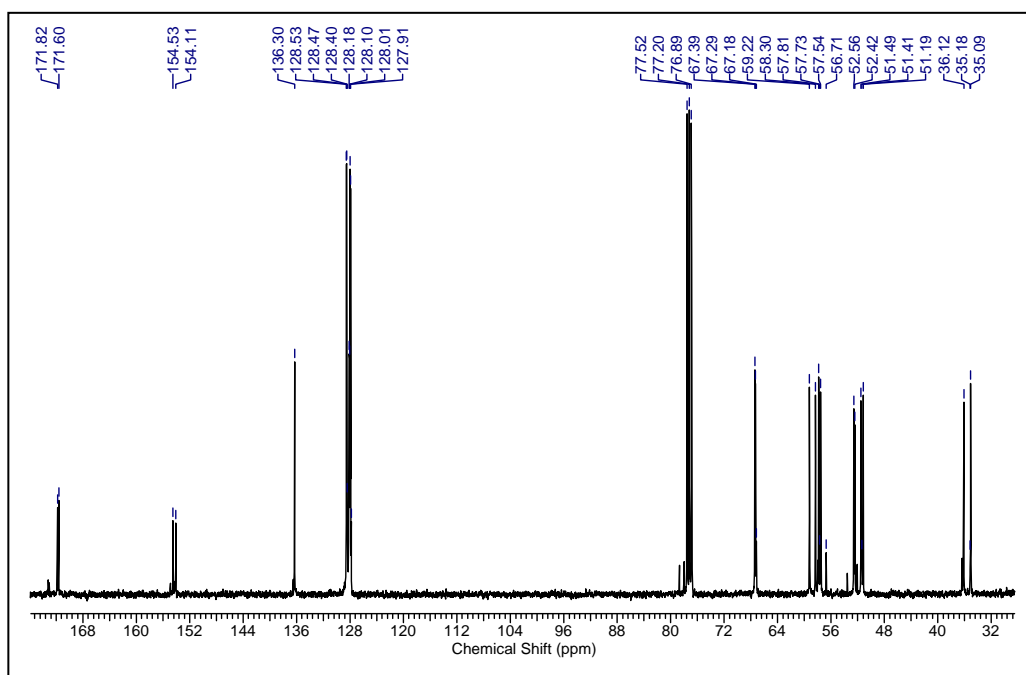
^{13}C NMR of compound 6**DEPT ^{13}C NMR of compound 6**

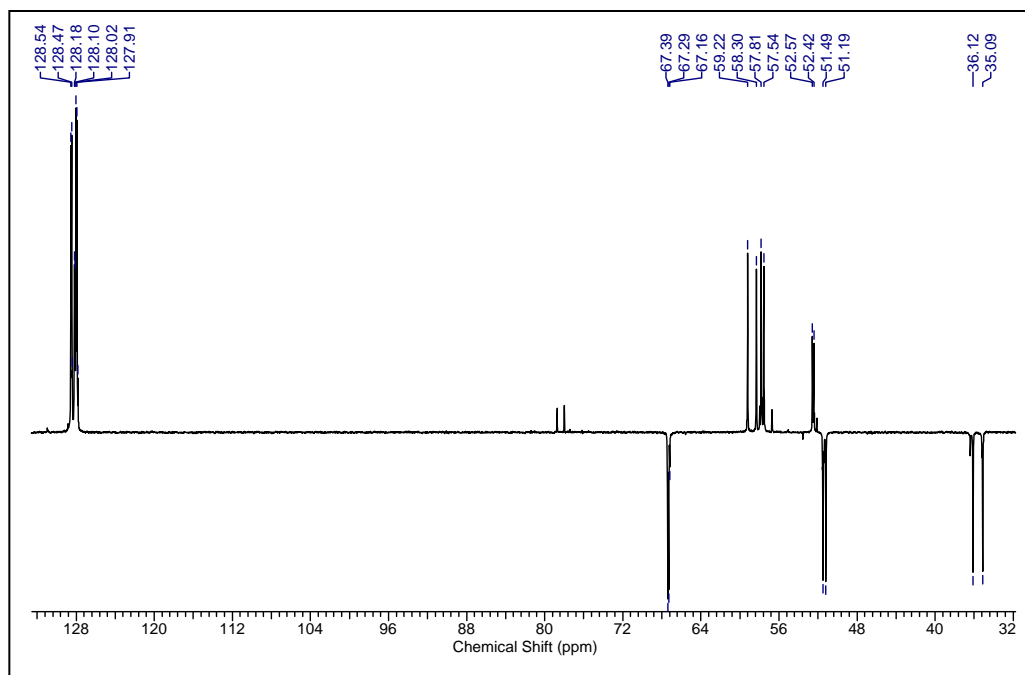
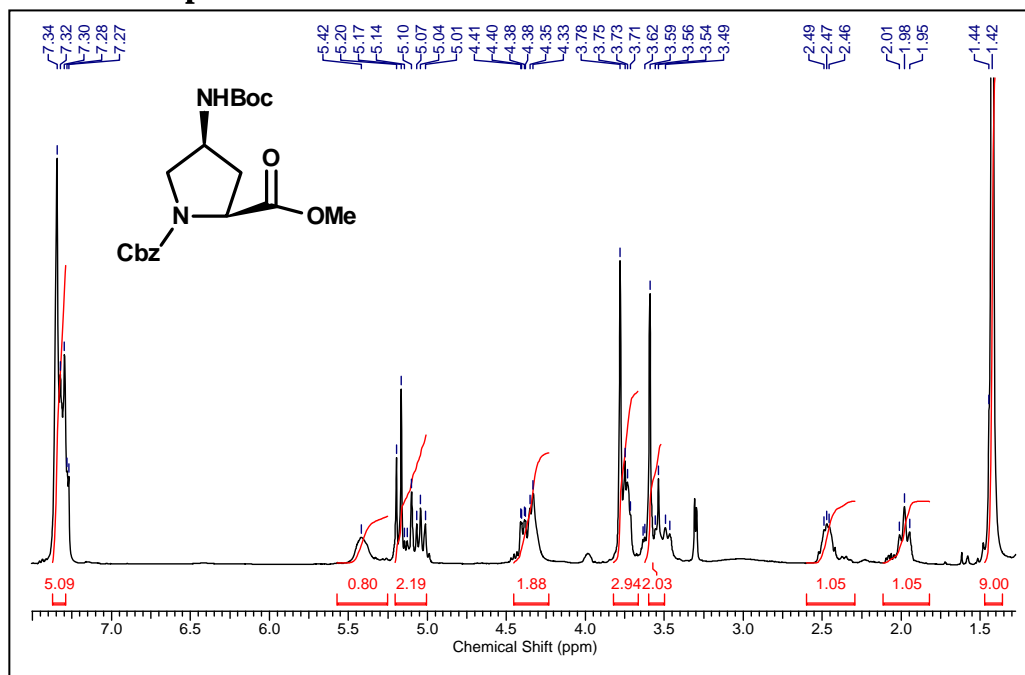
^1H NMR of compound 7 **^{13}C NMR of compound 7**

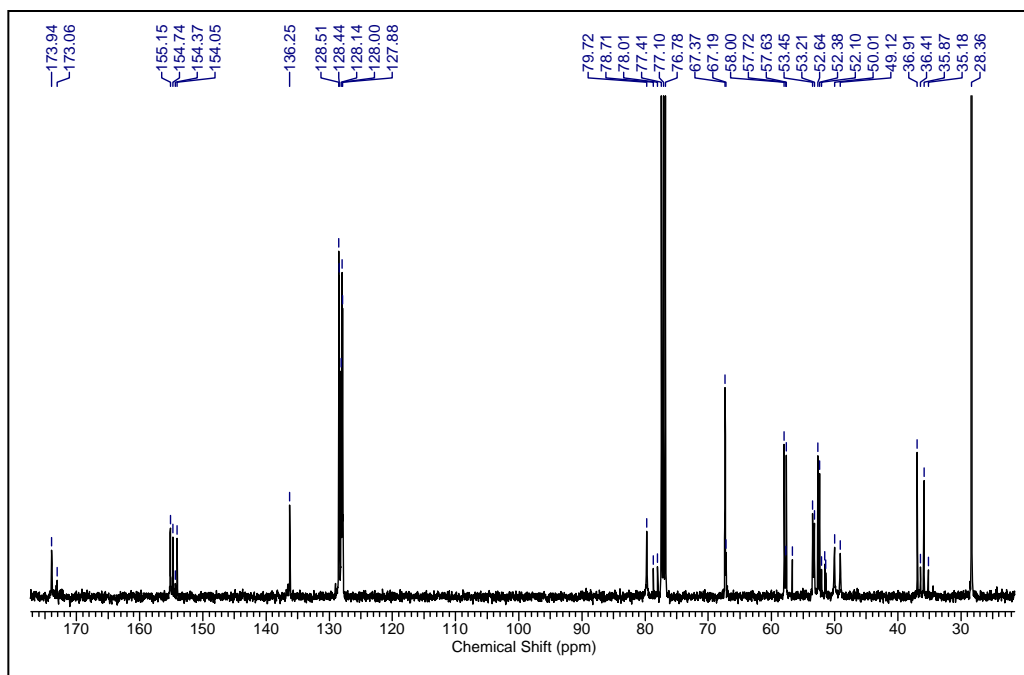
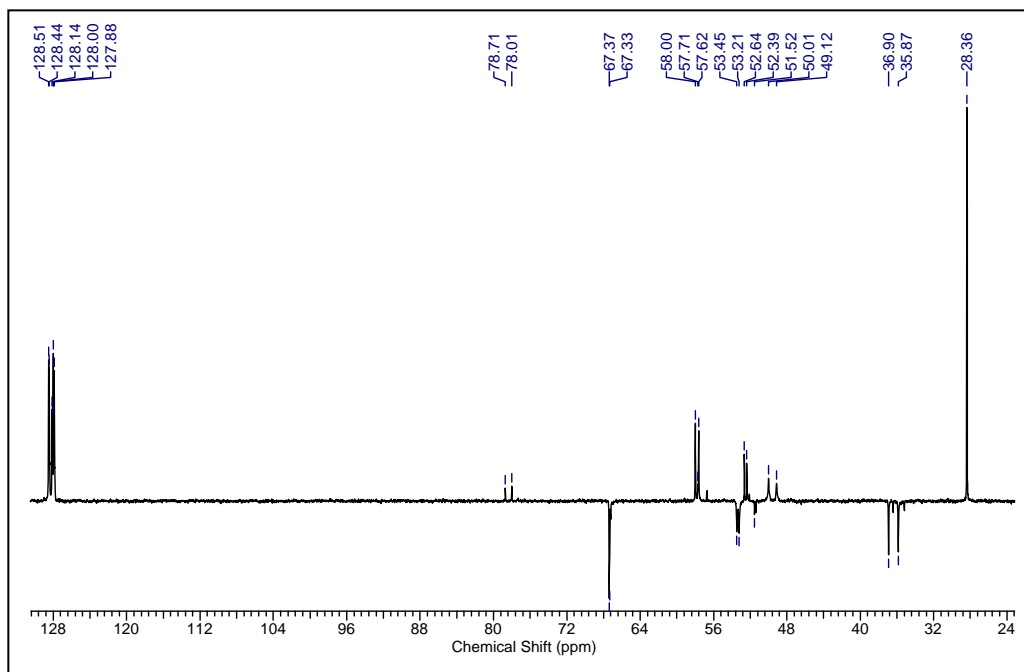
DEPT ^{13}C NMR of compound 7 ^1H NMR of compound 8

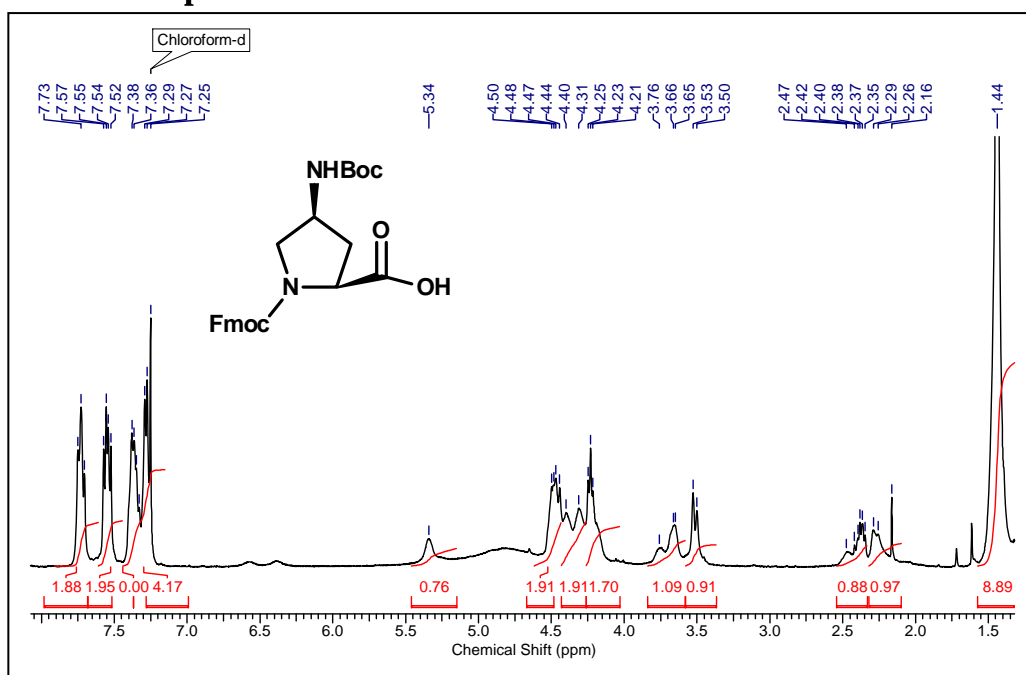
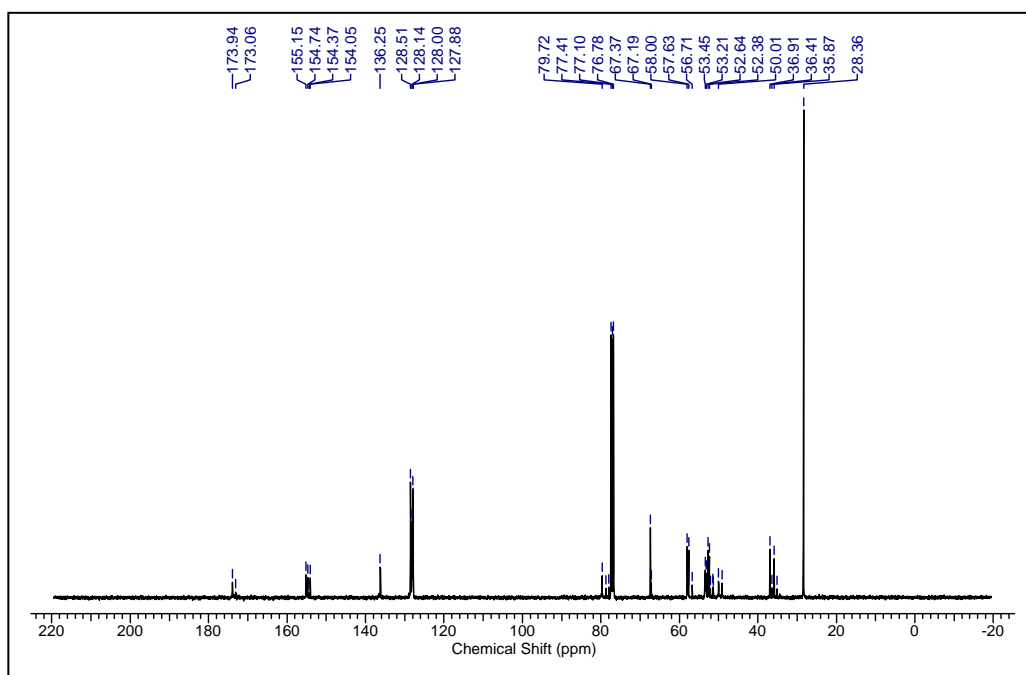
^{13}C NMR of compound 8**DEPT ^{13}C NMR of compound 8**

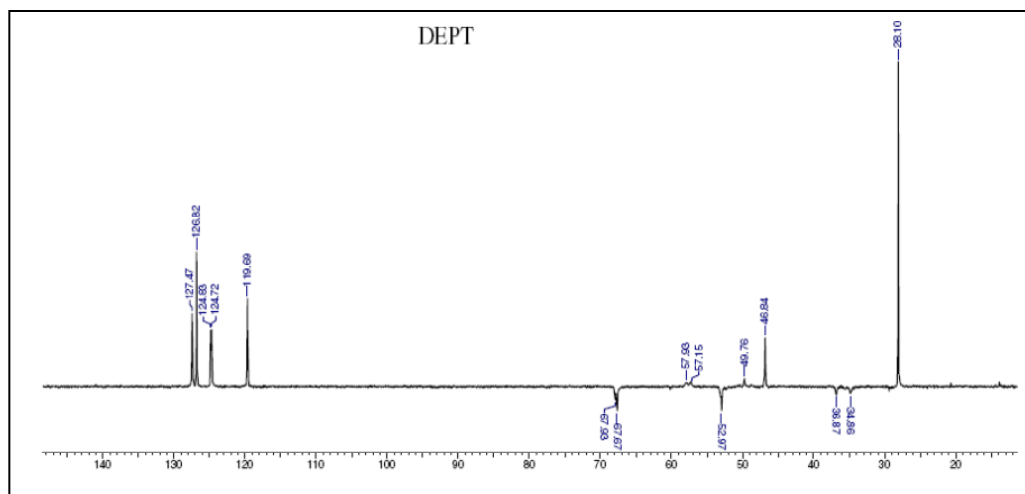
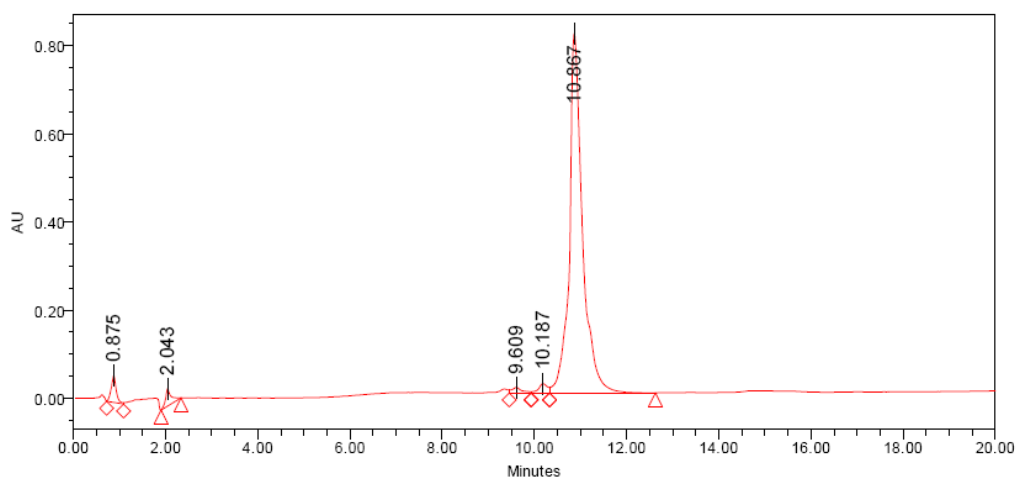
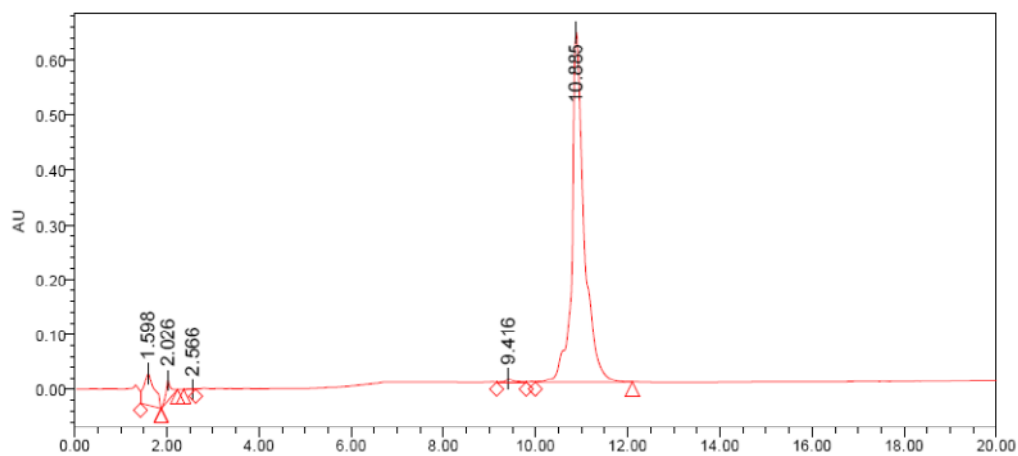
^1H NMR of compound 9 **^{13}C NMR of compound 9****DEPT ^{13}C NMR of compound 9**

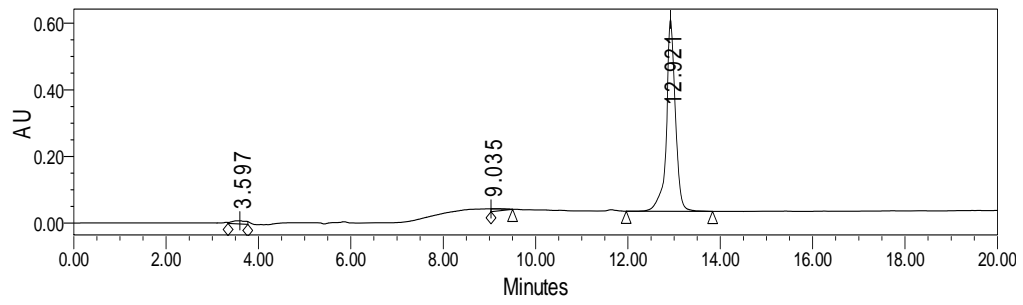
^1H NMR of compound 10 **^{13}C NMR of compound 10**

DEPT ^{13}C NMR of compound 10 ^1H NMR of compound 11

^{13}C NMR of compound 11**DEPT ^{13}C NMR of compound 11**

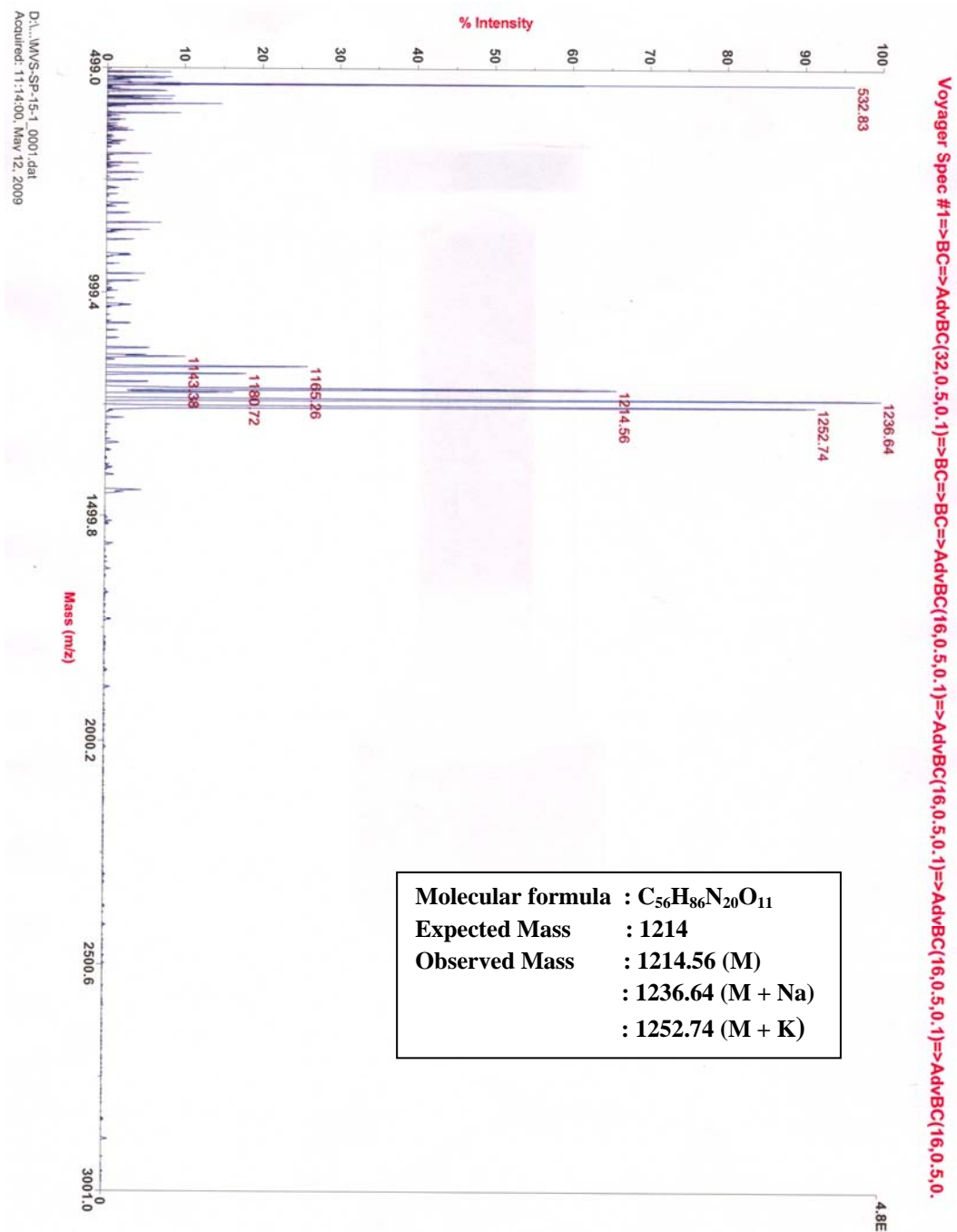
^1H NMR of compound 12 **^{13}C NMR of compound 12**

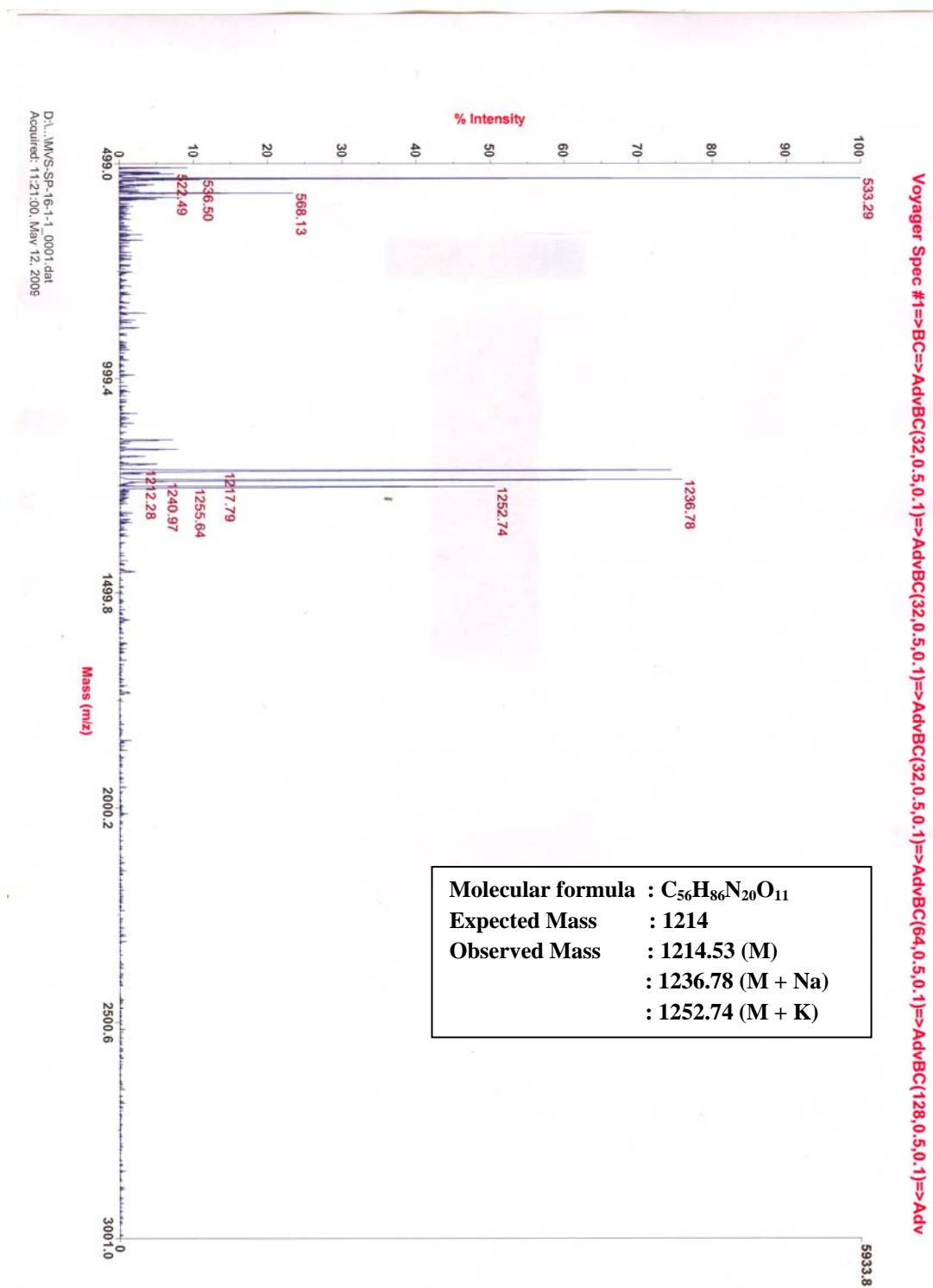
DEPT ^{13}C NMR of compound 12**(B) HPLC of peptides 1-3****HPLC of peptide 1 (4*R*-Amp₉)****HPLC of peptide 2 (4*S*-amp₉)**

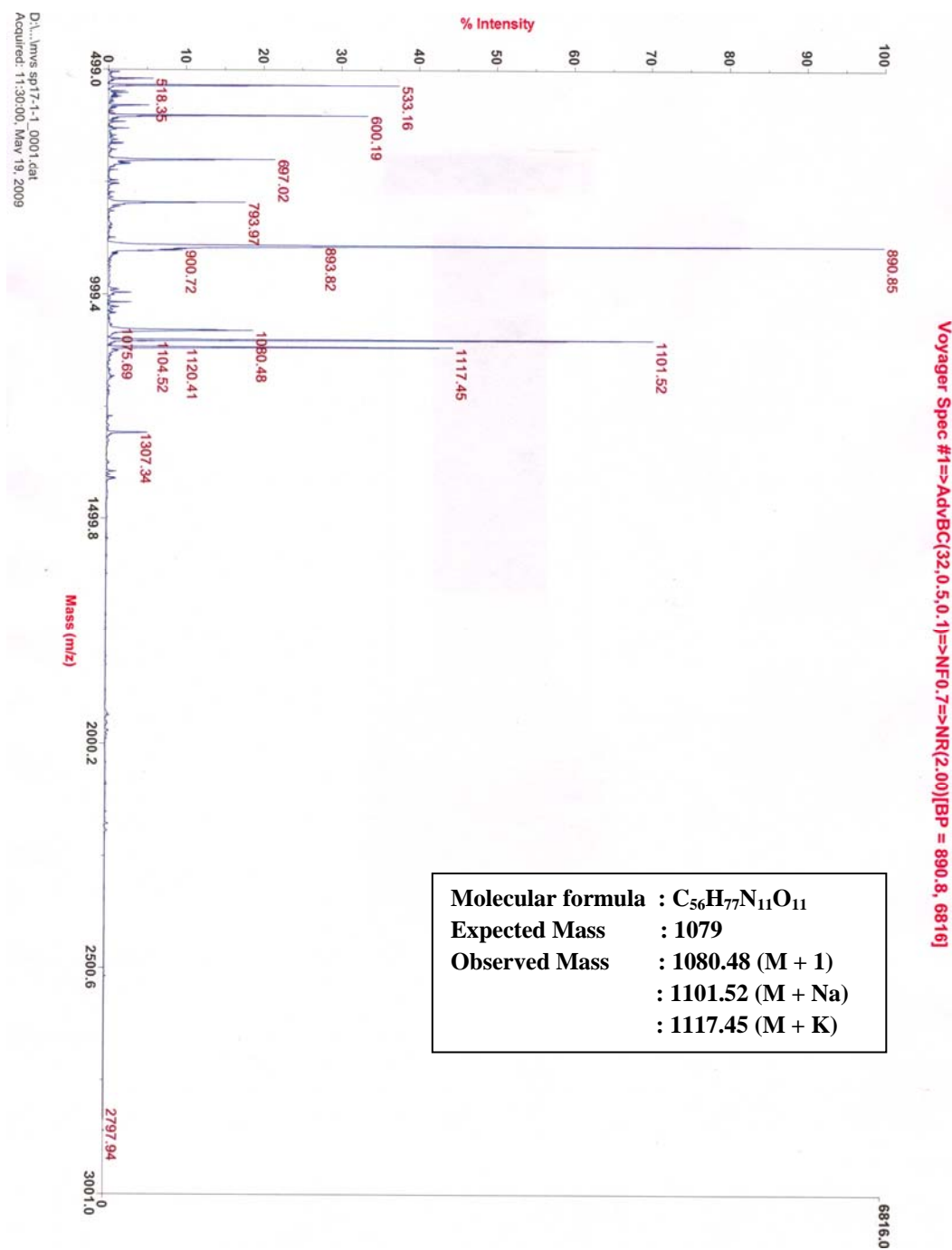
HPLC of peptide 3 (Pro₉)

(C) MALDI-TOF of peptides 1-3

MALDI-TOF of peptide 1 (*4R-Amp₉*)

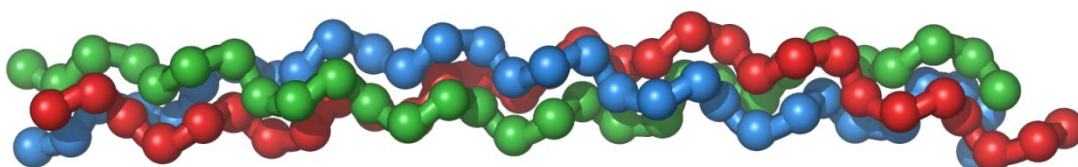


MALDI-TOF. of peptide 2 (4S-amp₉)

MALDI-TOF. of peptide 3 (*Pro*₉)

Chapter 2

Collagen Host-Guest Peptides: Position Dependent Substitution of 4(*R/S*)-Aminoproline Triplets



This chapter describes the role of C4-substituted aminoprolines on conformational properties of collagen and collagen-like peptides. The C4 aminoproline substitution at various positions of collagen peptide shows different biophysical effects as a function of position. This section demonstrates that the pH and solvent effects play an important role in determining the stability of C4-aminoproline collagen peptides.

1 Introduction

Collagen is an ancient structural protein found on earth till now,¹ present in connective tissues of higher organisms. The term ‘Collagen’ is derived from the Greek word for glue and was initially described as that constituent of connective tissue, which yields Gelatine on boiling. Collagen is a fascinating system of proteins that undergo a multi-step, hierarchical self-assembly which starts from individual peptide chains that assemble into a canonical triple helix. These triple helices then assemble into higher order structures which are often, but not always, fibrous in nature. Collagens are very abundant proteins in animal kingdom and are mainly located in the extracellular matrix (ECM).² The collagen is the most abundant protein in the human body, this fibrous protein is present in all connective tissues such as skin, bone, cartilage, tendons, basement membrane, blood vessels and imparts stability, integrity to these tissues to performing somewhat the same function as cellulose molecules in plants.

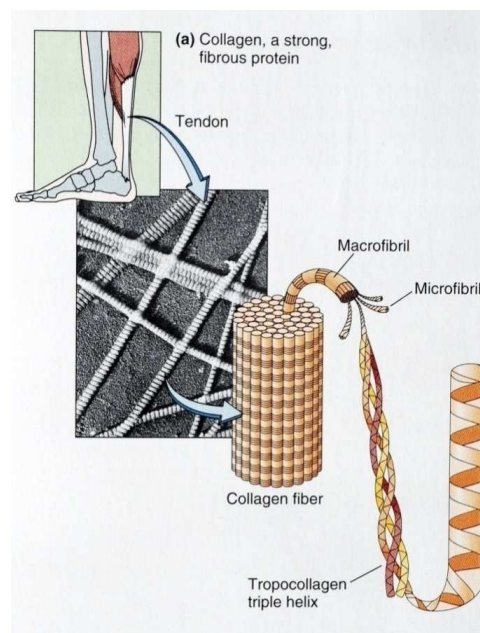


Figure 1: Classification of collagen as a structure.

In humans, collagen comprises one third of the total protein, accounts for three quarters of the dry weight of skin and is the most prevalent component of the extracellular matrix. Most collagens assist in anchoring cells to the extracellular matrix and some function in cellular regulation. Twenty-eight different types of collagen composed of at least 46 distinct polypeptide chains have been identified in vertebrates,

and many other proteins contain collagenous domains.³ These fibrils are organized in many different ways: they form molecular cables, that strengthen the tendons, large resilient sheets, which support the skin and internal organs, as mineralized aggregates in bone and teeth. Among these, collagen types I-III are the most abundant and form fibrils of similar structure. Figure 1 shows a typical collagen fiber originating from human tendon (a tough band of fibrous connective tissue that usually connecting muscle to bone) split into microfibril to triple-helix to the polypeptide chains.

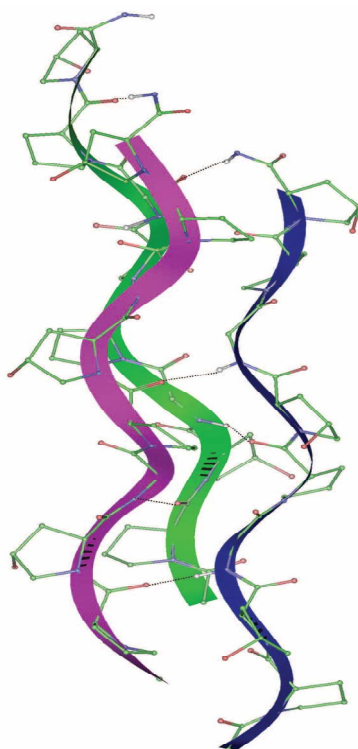


Figure 2: Ball and stick diagrams showing two projections of the currently accepted triple helical structure for collagen with one inter-chain hydrogen bond per tripeptide.⁴

The defining feature of collagen is an elegant structural motif in which three parallel polypeptide strands in a left-handed, polyproline II type (PPII) helical conformation coil about each other with a one-residue stagger to form a right-handed triple helix (Figure 2). The tight packing of PPII helices within the triple helix mandates that every third residue be Gly, resulting in a repeating X-Y-Gly sequence, where X and Y can be any amino acid. This repeat occurs in all types of collagen, although it is disrupted at certain locations within the triple-helical domain of nonfibrillar collagens.⁵ The amino acids in the X and Y positions of collagen are often

(2*S*)-proline (Pro, 28%) and (2*S*,4*R*)-4-hydroxyproline (Hyp, 38%), respectively. The other commonly found amino acids are Ala, Lys, Arg, Leu, Val, Ser and Thr.⁶ The repeat unit of Pro-Hyp-Gly is most common triplet (10.5%) in collagen.⁷ A large variation in the length of the triple helix can be found in nature, the shortest collagen described is 14 nm long (minicollagen of the nematocyst wall in hydra) and the longest ones 2400 nm (cuticle collagen of annelids). In animals, individual collagen triple helices, known as tropocollagen (TC), assemble in a complex, hierarchical manner that ultimately leads to the macroscopic fibers and networks observed in tissue, bone, cartilage, tendon, dentin and basement membranes (Figure 3).

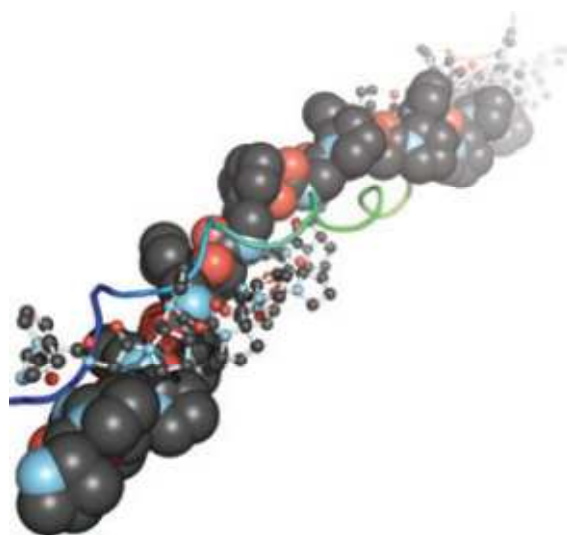


Figure 3: View down the axis of a (Pro-Pro-Gly)₁₀ triple helix [PDB entry 1k6f (7)] with the three strands depicted in space-filling, ball-and-stick, and ribbon representation.

1.1 Collagen superfamily

Collagens have been classified according to their α chains.⁸ Homotrimeric collagens (*i.e.*, types II and III) have three α chains of identical sequence. Heterotrimeric collagens have either two α chains of identical sequence (designated $\alpha 1$) and one α chain of differing sequence (designated $\alpha 2$), such as type I, or three α chains of differing sequence (designated $\alpha 1$, $\alpha 2$, and $\alpha 3$), such as type VI. Collagens are further classified into sub-families, based on their quaternary structure. These sub-families include fibrillar, fibril associated with interrupted triple-helices (FACIT), short chain, basement membrane, multiplexins, and membrane associated with interrupted triple-helices (MACIT).⁹ The most common collagens (types I, II, III, V, and XI) have fibrillar structures. The types of collagen, their distribution, composition,

and pathology are listed in Table 1. It is noteworthy that, although the three polypeptide chains in the triple helix of each collagen type can be identical, heterotrimeric triple helices are more prevalent than are homotrimeric triple helices.

Table 1: Classification of Collagen¹⁰

Type	Class	Composition	Distribution	Pathology
I	Fibrillar	$\alpha 1[I]2\alpha 2[I]$	Abundant and widespread: dermis, bone, tendon, ligament	OI, Ehlers–Danlos syndrome, Osteoporosis
II	Fibrillar	$\alpha 1[II]3$	Cartilage, vitreous	Osteoarthritis, Chondrodysplasias
III	Fibrillar	$\alpha 1[III]3$	Skin, blood vessels, intestine	Ehlers-Danlos syndrome, arterial Aneurysms
IV	Network	$\alpha 1[IV]2\alpha 2[IV]$ $\alpha 3[IV]\alpha 4[IV]\alpha 5[IV]$ $\alpha 5[IV]2\alpha 6[IV]$	Basement membranes	Alport syndrome
V	Fibrillar	$\alpha 1[V]3$ $\alpha 1[V]2\alpha 2[V]$ $\alpha 1[V]\alpha 2[V]\alpha 3[V]$	Widespread: bone dermis, cornea, placenta	Ehlers-Danlos syndrome
VI	Network	$\alpha 1[VI]\alpha 2[VI]\alpha 3[VI]$ $\alpha 1[VI]\alpha 2[VI]\alpha 4[VI]$	Widespread: bone, cartilage, cornea, dermis	Bethlem myopathy
VII	Anchoring fibrils	$\alpha 1[VII]2\alpha 2[VII]$	Dermis, bladder	Epidermolysis bullosa acquisita
VIII	Network	$\alpha 1[VIII]3\alpha 2[VIII]3$ $\alpha 1[VIII]2\alpha 2[VIII]$	Widespread: dermis, brain, heart, kidney	Fuchs endothelia corneal dystrophy
IX	FACIT	$\alpha 1[IX]\alpha 2[IX]\alpha 3[IX]$	Cartilage, cornea, vitreous	Osteoarthritis, multiple epiphyseal dysplasia
X	Network	$\alpha 1[X]3$	Cartilage	Chondrodysplasia
XI	Fibrillar	$\alpha 1[XI]\alpha 2[XI]\alpha 3[XI]$	Cartilage, intervertebral disc	Chondrodysplasia, osteoarthritis
XII	FACIT	$\alpha 1[XII]3$	Dermis, tendon	—
XIII	MACIT	—	Endothelial cells, dermis, eye, heart	—
XIV	FACIT	$\alpha 1[XIV]3$	Widespread: bone dermis, cartilage	—
XV	Multiplexin	—	Capillaries, testis kidney, heart	—
XVI	FACIT	—	Dermis, kidney	—
XVII	MACIT	$\alpha 1[XVII]3$	Hemidesmosomes in epithelia	Generalized atrophic epidermolysis bullosa
XXVIII	Multiplexin	—	Basement membrane, liver	Knobloch syndrome
XIX	FACIT	—	Basement membrane	—
XX	FACIT	—	Cornea (chick)	—
XXI	FACIT	—	Stomach, kidney	—
XXII	FACIT	—	Tissue junctions	—
XXIII	MACIT	—	Heart, retina	—
XXIV	Fibrillar	—	Bone, cornea	—
XXV	MACIT	—	Brain, heart, testis	Amyloid formation
XXVI	FACIT	—	Testis, ovary	—
XXVII	Fibrillar	—	Cartilage	—
XXVIII	—	—	Dermis, sciatic nerve	Neurodegenerative disease

1.1.1 Three dimensional structure and crystal packing of collagen

X-Ray diffraction analyses of biological samples provided the first glimpse into the unique structure of collagen.¹¹ In the 1950's, a model for the collagen triple helix was proposed by Kartha and Ramachandran¹² and refined by Rich and Crick¹³ to one that is essentially correct. In 1994, Berman, Brodsky *et. al* used X-ray diffraction analysis to determine the first high-resolution structure of a triple-helical collagen mimic (Figure 4).¹⁴ This structure confirmed the existence of not only interstrand N–H_(Gly)⋯O=C_(X) hydrogen bonds but also that C^α–H_(Gly/Yaa)⋯O=C_(X/Gly) hydrogen bonds stabilizing the triple helix.¹⁵

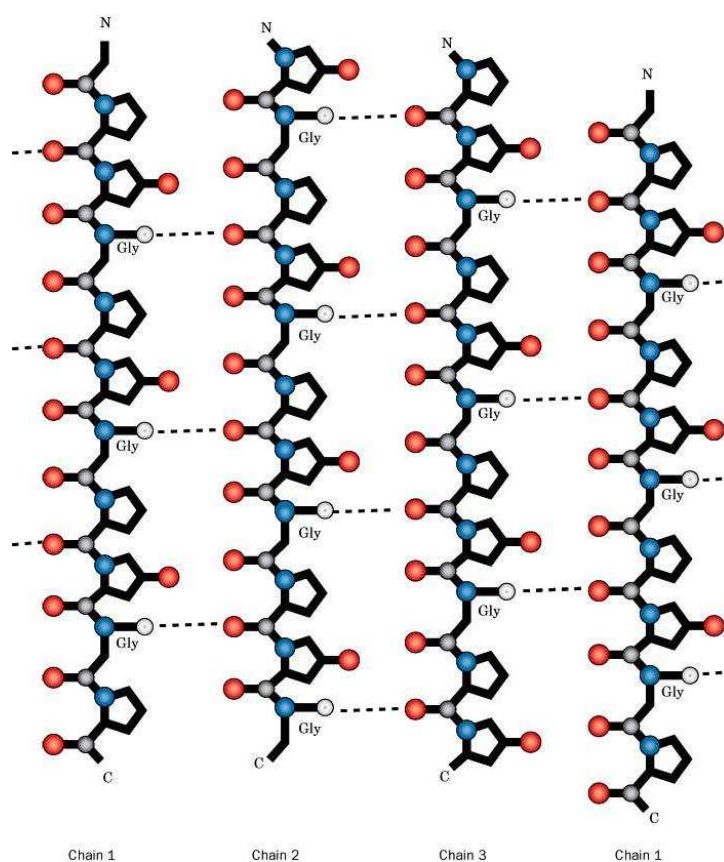


Figure 4: Schematic view of Rich & Crick model with one hydrogen bond per trimer repeat (Gly-Pro-Hyp) repeat.¹⁶

Most X-ray crystallographic studies on collagen-related peptides have been performed on proline-rich collagenous sequences. All the resulting structures have a 7/2 helical pitch (20.0-Å axial repeat), in contrast to the 10/3 helical pitch (28.6- Å axial repeat) predicted for natural collagen by fiber diffraction.¹⁷ On the basis of X-ray crystal structures of proline-rich collagen-related peptides, and in accordance with an

early proposal regarding the helical pitch of natural triple helices,¹⁸ Okuyama *et. al.*¹⁶ postulated that the correct average helical pitch for natural collagen is 7/2 (Figure 5). The generality of this hypothesis is unclear, as few regions of natural collagen are proline rich. The actual helical pitch of collagen is likely to vary across the domains and among different types of natural collagen.

Specifically, the helical pitch could be 10/3 in proline-poor regions and 7/2 in proline-rich regions. This proposal is supported by the observation that proline-poor regions within crystalline collagen-related peptides occasionally display a 10/3 helical pitch.¹⁹ Variability in the triple-helical pitch of native collagen could play a role in the interaction of collagenous domains with other biomolecules.²⁰

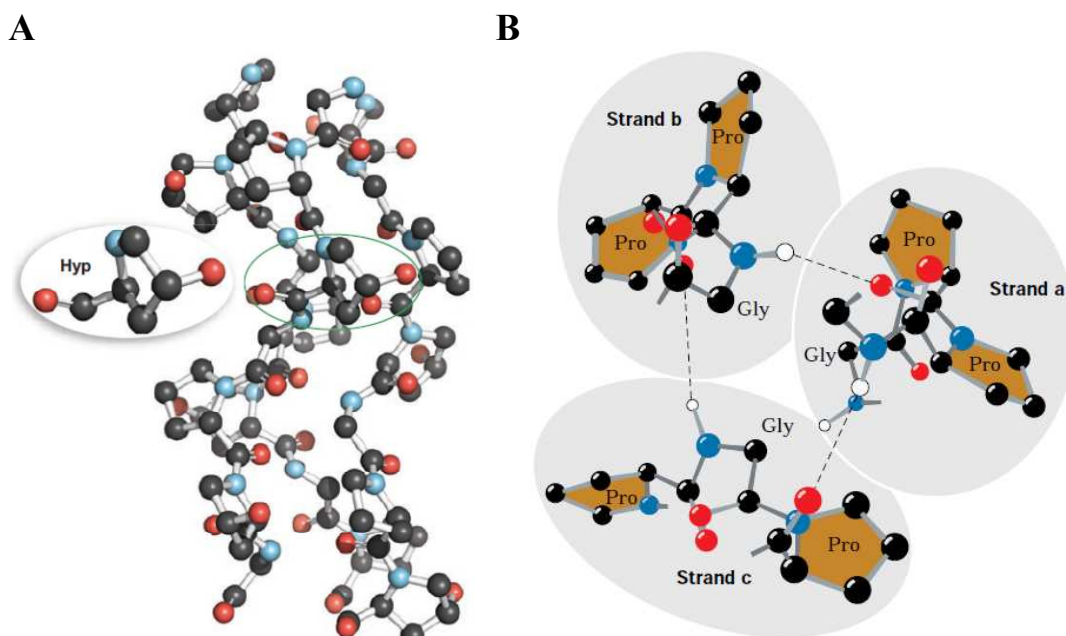


Figure 5: (A) Three-dimensional structure of a fragment of a collagen triple helix composed of (Pro–Hyp–Gly)_n strands (PDB 1v4f; Okuyama *et. al.*, 2004). Inset: Close-up of a Hyp residue showing the characteristic C γ -*exo* ring pucker. (B) Hydrogen bonding in the collagen triple helix. This view down the helix axis shows one Gly and two Pro residues (X and Y) in each chain.

1.2 Terminology Used in Higher Structure of Collagen

The tertiary structure of collagen refers to the fundamental unit originally known as tropocollagen (the monomeric collagen triple helix after proteolysis of collagen propeptides/the region where triplex form). Three polypeptide chains are intertwined to form a right-handed triple-helix with a pitch of approximately 8.6 nm. The rod-shaped triple helix has an average molecular weight of approximately 300 kDa, a length of 300 nm with a diameter of 1.5 nm.²¹ This extreme ratio of the dimensions

gives rise to high viscosity in solutions and high mobility in electrical fields. In addition, there are regions of 9–26 amino acids at the amino and carboxyl terminal chain ends of the molecules that are not incorporated into the helical structure. These non-helical regions are denoted as telopeptides. Procollagen is the hydroxylated form of collagen prior to collagen propeptides cleavage. Atelocollagen (a water-soluble form of collagen), which is produced by elimination of the telopeptides moieties using pepsin, has demonstrated its potential as a drug carrier, especially for gene delivery.²²

1.2.1 Nucleation and modulation of collagen fibrillogenesis

Collagen fibrillogenesis requires completion of two stages of self-assembly: nucleation and fiber growth. Collagen fibrillogenesis starts only after procollagen N- and C-proteinases cleave the collagen propeptides at each triple helix terminus to generate tropocollagen single chain. The C-terminal propeptides are essential for proper triple-helix formation but prevent fibrillogenesis.²³ After cleavage of the propeptides, tropocollagen monomers are composed of a lengthy triple-helical domain consisting of a repeating X-Y-Gly sequence flanked by short, non triple-helical telopeptides (Figure 6).

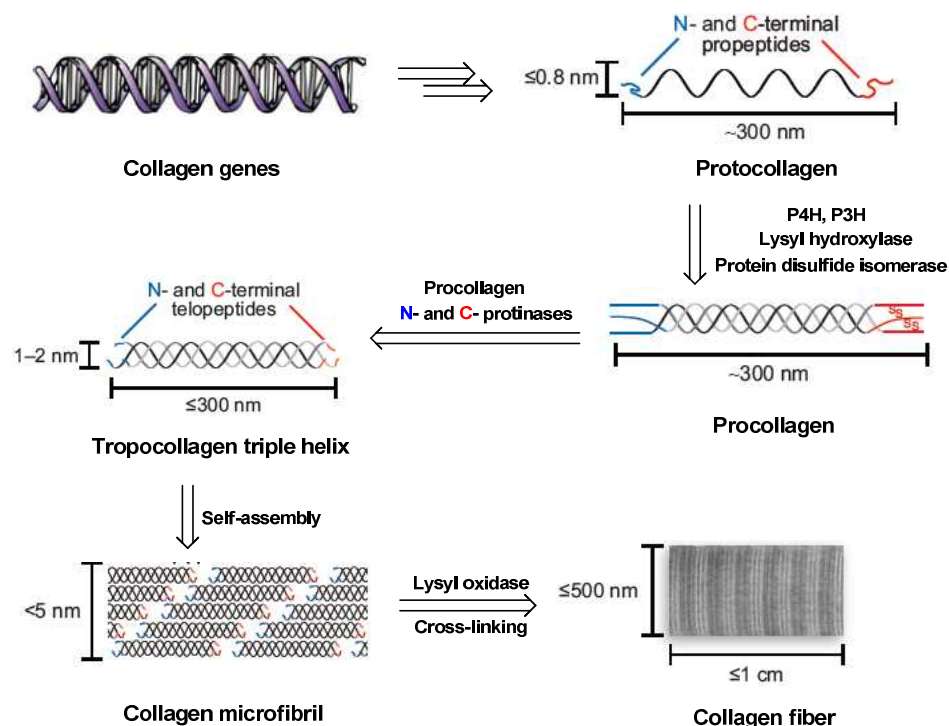


Figure 6: Synthetic route to collagen fibers, which are the major component of skin. Size and complexity is increased by post-translational modifications and self-assembly. Oxidation of lysine side chains leads to the spontaneous formation of desmosine and isodesmosine cross-links.

The C-terminal telopeptides of tropocollagen are important for triggering the proper fibrillogenesis. Prockop and Fertala²⁴ suggested that collagen self-assembly into fibrils is driven by the interaction of C-terminal telopeptides with specific binding sites on triple-helical single chain. The addition of synthetic telopeptide mimics can inhibit collagen fibrillogenesis, presumably by preventing the interaction between collagen telopeptides and tropocollagen monomers. Triple helices lacking the telopeptides can, however, assemble into fibrils with proper morphology.²⁵ Thus, collagen telopeptides could accelerate fibril assembly and establish the proper register within microfibrils and fibrils but might not be essential for fibrillogenesis.

Collagen telopeptides have another role in stabilizing mature collagen fibrils. Lys side chains in the telopeptides are cross-linked subsequent to fibril assembly, forming desmosine and isodesmosine cross-links between Lys and hydroxylysine residues with the aid of lysyl oxidase. The cross-linking process endows mature collagen fibrils with strength and stability, but is not involved in fibrillogenesis. Thus, although collagen telopeptides might not be essential for nucleating collagen fibrillogenesis, their absence greatly weakens the mature fibril owing to the lack of cross-links within and between triple helices.²⁶

1.2.2 Biosynthesis of collagen

The biosynthesis of collagen is well studied.²⁷ Collagen is synthesized in the endoplasmic reticulum (ER) as procollagen, which is the precursor protein that bears propeptide domains at either end of the triple helical domain. The processes by which procollagen is synthesized in the lumen of the ER include unique steps that are not found in the biosynthesis of globular proteins. The biosynthetic pathway is shown in Figure 7.

First, each polypeptide chain of procollagen (pro α -chains) finds its correct partners, which enables the formation of the distinct types of procollagen. Second, triple helix-formation of long X-Y-Gly repeats starts at a defined region, which results in the formation of a correctly aligned triple helix and thereby prevents mis-staggering. The most characteristic step is the formation of the triple helix. This step involves specific post-translational modifications, in particular, the prolyl 4-hydroxylation of the Y-position amino acids that stabilizes the triple helical conformation. The

formation of the triple helix is a slow process compared to the folding of globular proteins, including *cis-trans* isomerization of the many prolyl and hydroxyprolyl peptide bonds. Recent advances have indicated that these processes are assisted by a set of the ER-resident molecular chaperones, such as protein disulfide isomerase (PDI), peptidyl prolyl *cis-trans* isomerases (PPIases), heat-shock protein (Hsp)47, and prolyl 4-hydroxylase (P4H). The intracellular trafficking of procollagen molecules has also been shown to involve a pathway distinct from that utilized by small secretory proteins. The polypeptide chains are then secreted into extracellular matrix, where they fold into a triple-helix.²⁸

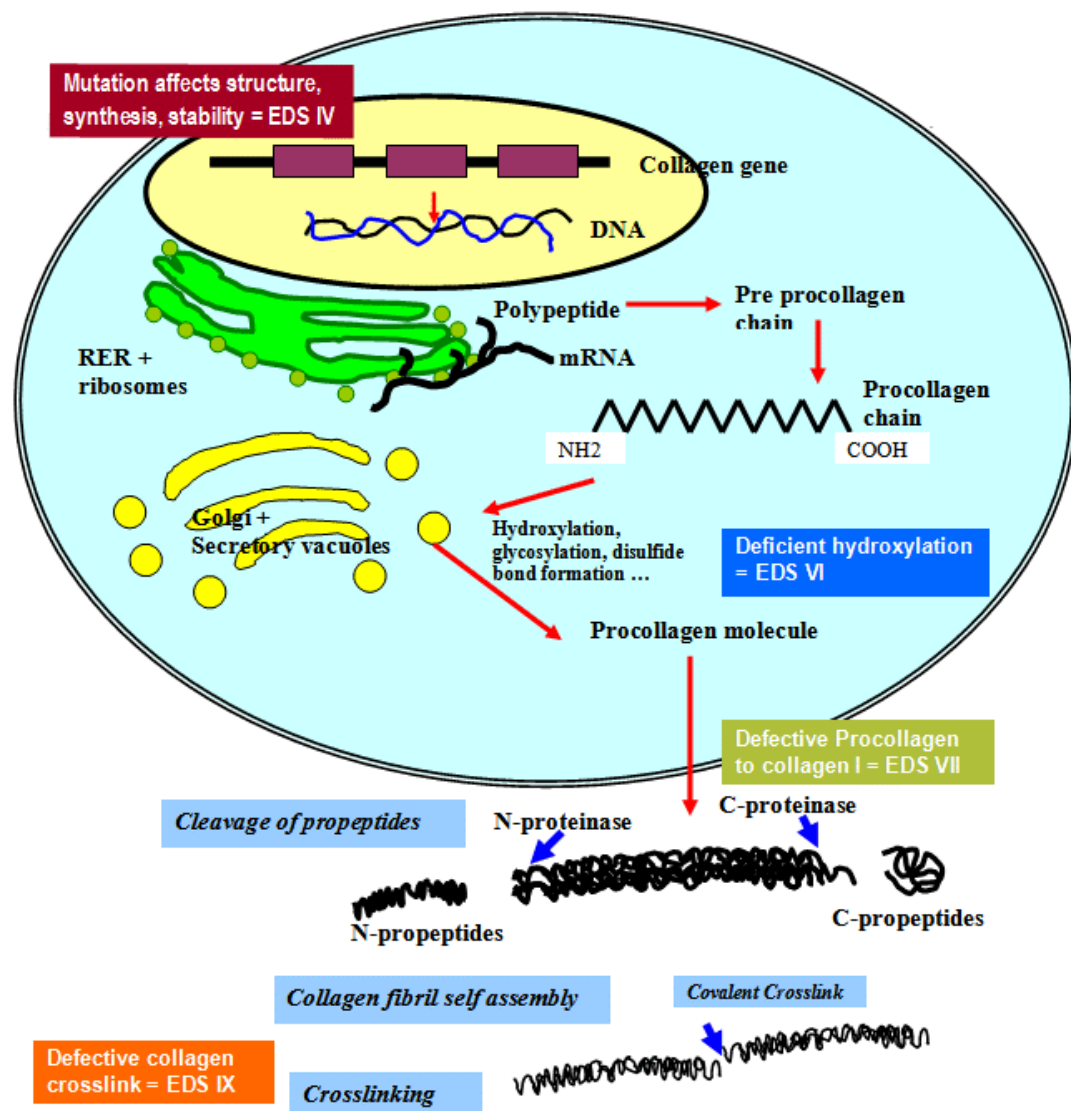


Figure 7: The biosynthetic pathway of collagen.

1.3 Factors Affecting Collagen Triple Helix Structure

1.3.1 Pyrrolidine ring conformation in proline and substituted proline

Proline is a cyclic imino acid and the bridging of the α -carbon atom to the main chain amide nitrogen by 3 methylene bridge imposes further constraint on the main chain torsion angles φ and ψ . Proline and substituted-proline rings exhibit two types of ring-pucker, and in analogy to *ribo* and *deoxyribo* sugars, they are named as *N* (γ -*exo*) and *S* (γ -*endo*) puckers (Figure 8). These ring-puckers are also expressed in terms of *endocyclic* torsion angles. In proline the two puckers are almost equally preferred and the energy barrier to inter-conversion is very low.²⁹

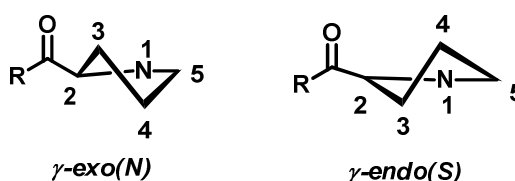


Figure 8: Ring-puckers of the pyrrolidine ring in proline and substituted prolines.

The pH dependence of *exo-endo* interconversion of the pyrrolidine ring of L-proline has also been studied in detail.³⁰ The *endo-exo* equilibrium of proline shifts towards more *endo* in basic solution, because of the steric effect of the nitrogen lone pair. At increased pH the preferred rotamer has the lone pair in the *cis* position to the carboxylic group and its smaller volume requirements, favors more *endo* population.

Extensive study of both the thermodynamics and kinetics of *E-Z* isomerization of 4-substituted proline model compounds has shown that the nature and stereochemistry of electron withdrawing substituent at 4-position has a marked effect on the proline ring conformation and peptidyl-prolyl bond isomerization.³¹ For example, the electronegative substituents in the 4*R*-stereochemistry of the proline increase the stability of *trans* isomer, whereas electronegative substituents in the 4*S*-stereochemistry decrease its stability. In Ac-X-OCH₃ model compounds, it was observed that an electron-withdrawing 4-substituent exerts a significant inductive (-I) effect on pyrrolidine ring nitrogen. This -I effect, transmitted through σ -bonds, increases pyramidalization of the ring nitrogen, decreases the bond order of N-CO bond, and thereby altering the *E-Z* isomerization of the peptidyl-prolyl bond. Thus, in 4-substituted prolines, depending on the steric and electronic effects exerted by the 4-

substituent, pyrrolidine ring may prefer any one of the ring-puckers. For example *trans*-4-hydroxy-L-proline (Hyp) shows almost exclusively γ -*exo* pucker, as found in both crystal structure and in solution.²⁹ This pucker preference has been attributed to the phenomenon of *gauche* effect between hydroxyl group and amide-N-atom.³²

It has also been demonstrated that conformational stability arises from stereoelectronic effects that fix the pucker of the pyrrolidine ring and thereby preorganize the peptide backbone as the steric effect of a 4-methyl group can confer stability similar to that from a 4-fluoro group in the opposite configuration (Figure 9).³³ The C^{γ} -*endo* conformation is favoured strongly by steric effects when $R^1 = \text{Me}$, $R^2 = \text{H}$ (mep) or stereoelectronic effects when $R^1 = \text{H}$ and $R^2 = \text{F}$ (flp). Similarly, the C^{γ} -*exo* conformation is favoured strongly by steric effects when $R^1 = \text{H}$, $R^2 = \text{Me}$ (Mep) or stereoelectronic effects when $R^1 = \text{OH}$, $R^2 = \text{H}$ (Hyp) or $R^1 = \text{F}$, $R^2 = \text{H}$ (Flp)

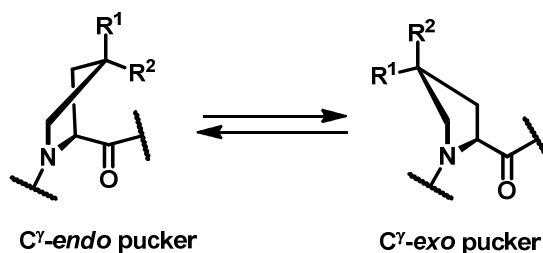


Figure 9: Ring conformations of 4-substituted proline residues.

1.3.2 *Gauche* effect on the ring-pucker preferences

Gauche effect may be described as “the preference of two electronegative atoms X and Y in vicinally substituted ethanes to remain *gauche* with respect to each other rather than *anti*”.³⁴ This effect is surprising since, due to a combination of dipole repulsion and steric effect between the electronegative atoms the vicinal substituents are expected to remain *anti* to each other. Many molecules containing N, O, P, S, F or Cl show a preference for *gauche* conformation. Though the origin of the *gauche* effect is not very clear and *ab initio* quantum chemical calculations underestimate the *gauche* effect still σ -Hyper-conjugation³⁵ and bent-bonds³⁶ have been proposed to explain the phenomenon of the *gauche* effect.

The structure and reactivity of an organic molecule also relies on the stereochemistry of its bonded and non-bonded electron pairs.³⁷ Stereoelectronic (*gauche*) effects, arise from the mixing of an electron pair with the antibonding σ^* of an adjacent polar bond (C-X, where X = N or O) and are important in stabilizing the

conformation of nucleic acids and carbohydrates.³⁸ For example, the multiple *gauche* effects (X-C-C-X) arising from 2' oxygen distinguishes stable conformations of RNA.RNA and RNA.DNA duplexes where anomeric effect (X-C-X) is shown to enhance the stability of the α -isomer of the glycosides.³⁹

In 4-substituted prolines, the steric repulsion between the amide-ring nitrogen and the 4-substituent should result in a pseudo-equatorial positioning of the 4-substituent i.e., *anti* with respect to the ring amide-nitrogen. For example, 4*R*-hydroxyproline and 4*R*-fluoroproline may be expected to exhibit γ -*endo* ring-pucker (Figure 10A). However, X-ray crystal structure, ^1H - ^1H and ^{19}F - ^1H coupling constant analyses confirm the pseudo-axial positioning of fluorine and the resulting γ -*exo* ring-pucker. This suggests that the *gauche* effect may be a dominating factor in determining the ring-pucker preference for proline with 4-electronegative substituent. Similarly, in 4*S*-fluoroproline with an opposite stereochemistry at C-4 the *gauche* effect leads to γ -*endo* ring-pucker.⁴⁰

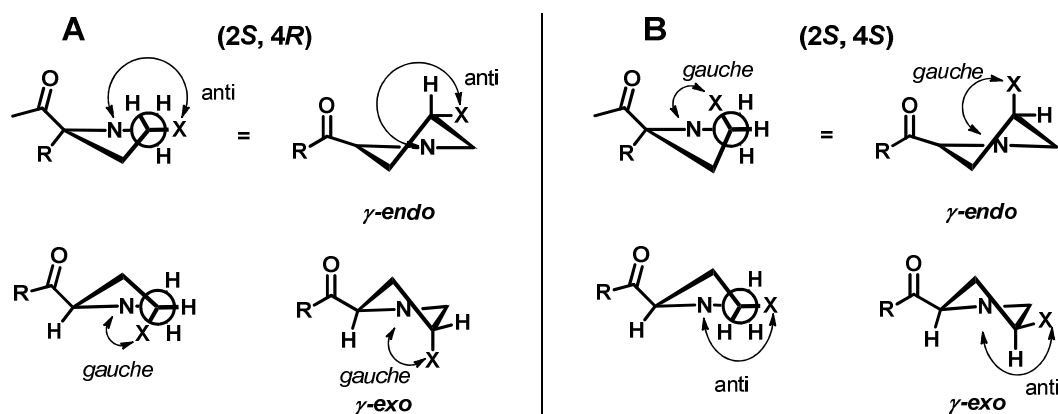


Figure 10: Newman and saw-horse projections depicting the *gauche* effect and the resulting pucker preferences in the prolines with electronegative (A) 4*R*-substituent, (B) 4*S*-substituent

It has also been proposed that restricted rotation about the N-C $^{\alpha}$ bond leads to a favorable C=O \cdots C=O electrostatic interaction,⁴¹ which leads to stability almost as strong (80%) as backbone hydrogen bond in a protein⁴². This stabilization involves n \rightarrow π^* interaction *via* electron donation from the oxygen lone pair of the (*i* - 1) amide C=O to the antibonding orbital of the C=O belonging to the proline (*i*) residue. Moreover in evidence for the significance of this interaction, quantum mechanical rather than electrostatic effect was seen to dominate with an estimated contribution of 0.7 kcal mol $^{-1}$ to the stability of the *trans* conformation at 298K.⁴³

1.3.3 Positional preferences of amino acids in collagen

In principle, there are 400 different possible X-Y-Gly triplets for composing peptide with 20 amino acids, but only a limited number are actually found in collagen sequences.⁴⁴ Several attempts have been made to measure the stability of collagen model peptides with various amino acids and the triple helical propensities.⁴⁵ In one approach, a set of host triple-helical peptides with the sequence (Pro-Hyp-Gly)₃-X-Y-Gly-(Pro-Hyp-Gly)₄ (Figure 11), where XYG is the guest triplet have been used to elucidate positional preference of various amino acids commonly found in natural collagen. All 20 amino acids were tested individually at both X or Y position with the help of T_m measurements and thermodynamic analyses of the host sequence. The results have revealed a striking relationship between the nature of the amino acid and the triple helix stability.⁴⁶

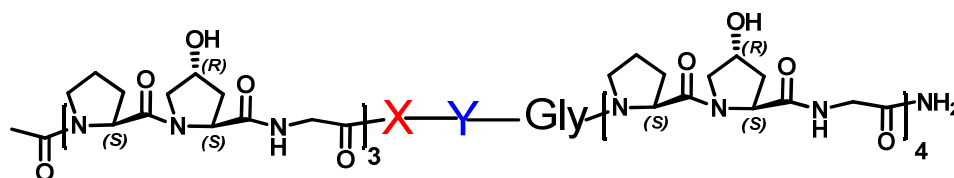


Figure 11: Amino acids substituted as guest in X and Y position

Pro was found to be the most stabilizing residue ($T_m = 47.3$ °C) in X position and also offered the lowest enthalpic contribution ($\Delta H^\circ = 435$ kJ mol⁻¹) in the entire series. The host peptides containing charged amino acids (Glu, Lys, Arg and Asp) were among the most stable peptides but these residues possessed higher enthalpic contribution than Pro, which reflected side-chain interactions with available backbone carbonyl or solvent. Trp was found to be the most destabilizing amino acid in X position ($T_m = 31.9$ °C). Among the Y position replacements Hyp/Arg were most stabilizing amino acids with nearly equal T_m values of 47.3 °C and 47.2 °C respectively. But, Arg peptide showed a higher ΔH° of denaturation than Hyp (610 kJ mol⁻¹ and 435 kJ mol⁻¹ respectively), suggesting a different mechanism for stabilization. The aromatic amino acids Tyr, Phe and Trp as usual were most destabilizing (T_m values 30.2 °C, 28.3 °C and 26.1 °C respectively).

Thermodynamic analysis has shown that both Hyp in Y position and Pro in X position have the same enthalpy contribution ($\Delta H^\circ = 435$ kJ mol⁻¹). It is also interesting to see that the most destabilizing amino acid Trp has the highest enthalpy

value (593 kJ mol^{-1} in X position and 670 kJ mol^{-1} in Y position). Analysis of amino acid sequences of fibril-forming collagens in SWISS-PROT database showed that the frequency of occurrence of each of the amino acids in fibril-forming collagens correlate well with triple-helix stability.

1.3.3a Proline in collagen peptides: In each strand of human collagen, $\approx 22\%$ of all residues are either Pro or Hyp. The abundance of these residues reorganizes the individual strands in a PPII conformation, thereby decreasing the entropic cost for collagen folding.⁴⁷ Despite their stabilizing properties, Pro derivatives also have certain damaging consequences for triple-helix folding and stability that partially offset their favourable effects. For example, Pro has a secondary amino group and forms tertiary amides within a peptide or protein. Tertiary amides have a significant population of both the *trans* and the *cis* isomers (Figure 12), whereas all peptide bonds in collagen are *trans*. Thus, before a $(\text{Pro-Hyp-Gly})_n$ strand can fold into a triple helix, all the *cis* peptide bonds must isomerize to *trans*. In collagen structure the factors dictating triple helix structure and stability are intertwined in a complex manner.

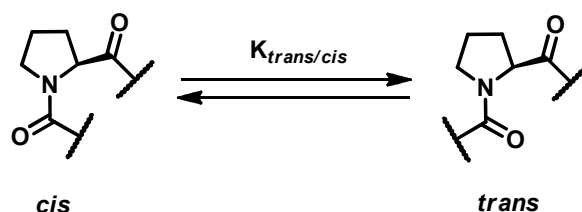


Figure 12: Pro *cis-trans* isomerization.

The ring pucker of Pro in the X position is important for triple-helix stability. Typically, Pro residues in the X position of biological collagen are not hydroxylated and usually display the C^γ -*endo* ring pucker.⁴⁸ By employing C^γ -substituents, both the *gauche* effect and steric effects can be availed to preorganize the C^γ -*endo* ring pucker. Installation of flp, (2*S*,4*S*)-4-aminoproline (amp), (2*S*,4*S*)-4-chloroproline (clp) or (2*S*,4*R*)-4-methylproline (mep) residues (all of which prefer the C^γ -*endo* ring pucker) in the X position of collagen is stabilizing relative to Pro, but installation of Flp, Amp, Clp or Hyp (which prefer the C^γ -*exo* ring pucker) is destabilizing.^{10,49} These results suggest that preorganizing the C^γ -*endo* ring pucker in the X position of collagen related peptides stabilizes triple helices. This results from the fact that the Pro

derivatives with a C^γ -endo ring-pucker have ϕ and ψ main chain torsion angles similar to those observed in the X position of triple helices.

Pro residues in the Y position of procollagen triplets are modified by prolyl 4-hydroxylase (P4H), a nonheme iron enzyme that catalyzes the post-translational and stereoselective hydroxylation of the unactivated γ -carbon of Pro residues in the Y position of collagen sequences to form Hyp (Figure 13). The Hyp is essential for the formation of collagen *in vivo*.

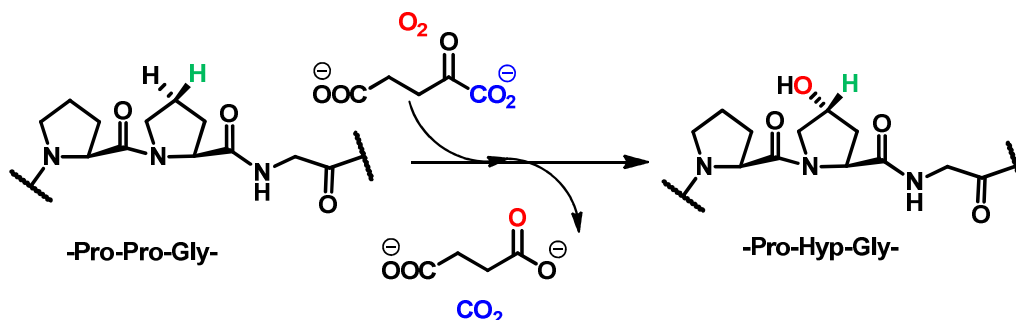


Figure 13: Reaction catalyzed by prolyl 4-hydroxylase (P4H). Pro residues in the Y position of collagen strands are converted into Hyp prior to triple-helix formation.

1.3.3b Hydroxyproline in collagen peptides: The hydroxylation of Pro residues in the Y position of collagen increases the thermal stability of triple helices. This stabilization occurs when the resultant Hyp is in the Y position⁵⁰ but not in the X position, or when the hydroxyl group is in the 4S configuration (2S,4S)-4-hydroxyproline (hyp).⁵¹ These findings led to the proposal that the 4R configuration of a prolyl hydroxyl group is alone privileged in enabling the formation of water-mediated hydrogen bonds that stitch together the folded triple helix. Such water bridges between Hyp and main chain heteroatoms were observed by Berman *et.al.*⁵² in their X-ray crystallographic studies of collagen related peptides. The frequency of Hyp in most natural collagen is, however, too low to support an extensive network of water bridges. For example, four or more repeating triads of X–Hyp–Gly occur only twice in the amino acid sequence of human type I collagen.

In the X position, a Pro residue with a C^γ -endo pucker generally stabilizes a triple helix, whereas one with a C^γ -exo pucker destabilizes a triple helix. For example, (Hyp-Pro-Gly)_n triple helices are far less stable than (Pro-Pro-Gly)_n triple helices⁵³ because Hyp prefers the C^γ -exo ring pucker and thus preorganizes the ϕ and ψ torsion angles improperly for the X position of a collagen triple helix. Surprisingly (Hyp-Hyp-Gly)₁₀

triple helices are actually slightly more stable than (Pro-Hyp-Gly)₁₀ triple helices⁵⁴ despite the Hyp residues in the X position of (Hyp-Hyp-Gly)₁₀ displaying the *C*^γ-*exo* ring pucker in the triple helix.

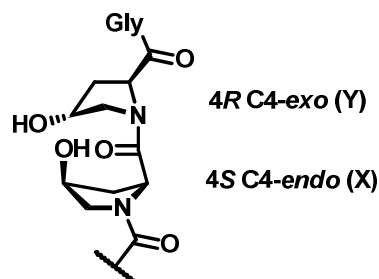


Figure 14: Position dependent preferred proline puckers in collagen.

The crystal structures of (Hyp-Hyp-Gly)₁₀ show that the main chain torsion angles in the X position of a (Hyp-Hyp-Gly)_n triple helix adjust to accommodate a *C*^γ-*exo* ring pucker in that position. Recently, it has been shown that (hyp-Pro-Gly)₁₀ forms less stable triple helices than the 4*S* substituted flp, clp at X position (which prefer the *C*^γ-*endo* ring pucker).⁵⁵ This is due to the fact that 4*S*-hydroxylation of Pro to form hyp enforces a *C*^γ-*endo* ring pucker and a transannular hydrogen bond between the hydroxyl moiety and the carbonyl of hyp distorts the main-chain torsion angle that accompanies a *C*^γ-*endo* ring pucker, thereby affecting the interstrand H-bonding result in destabilization of triple helices.

1.3.4 Interstrand hydrogen bonds

The ubiquity of collagen makes the ladder of recurrent N–H(Gly)⋯O=C(X) hydrogen bonds (amide carbonyl between two proline ring) that form within the triple helix the most abundant amide–amide hydrogen bond in proteins. Replacing the Y–Gly amide bond with an ester in a host-guest collagen related peptides (Figure 15) enabled estimation of the strength of each amide–amide hydrogen bond as $\Delta G^{\circ} = -2.0$ kcal/mol.⁵⁶ Boryskina and co-workers⁵⁷ used a variety of other experimental techniques to assess this same parameter, estimating the strength of each amide–amide hydrogen bond within a poly(Pro-Pro-Gly) as $\Delta G^{\circ} = -1.8$ kcal/mol and within native collagen as $\Delta G^{\circ} = -1.4$ kcal/mol. Very recently it was demonstrated that the interstrand hydrogen bonds are significantly more important for the stability of the collagen triple helix than the ring puckering,⁵⁸ by introducing 4*S*-acetamido proline (acp) and 4*S*-formamido proline (fmp), in which intramolecular hydrogen bonding

leads to a preference for the C^4 -endo ring pucker. The interference of the intramolecular hydrogen bond within 4*S*-acp or 4*S*-amp with the interstrand hydrogen bond results in destabilization of triple helix.

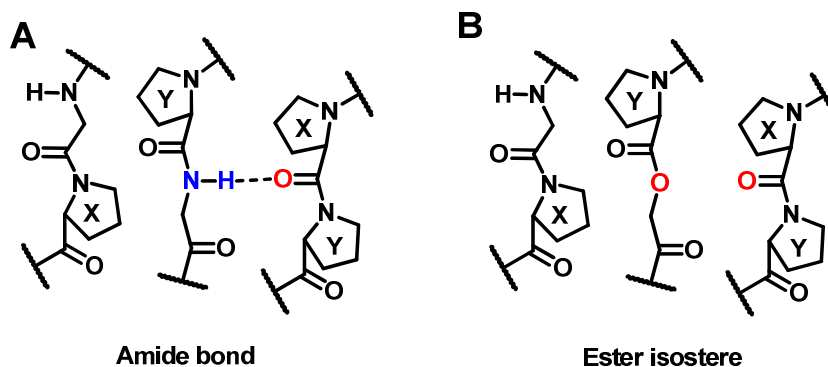


Figure 15: Importance of interstrand hydrogen bonds for collagen triple-helix stability.

1.3.5 $n \rightarrow \pi^*$ Interaction

A general aspect of collagen related peptides is that Pro residues with either a C^{γ} -endo or C^{γ} -exo ring pucker stabilizes triple helices in the X and Y positions, respectively. The appropriate ring pucker, enforced by a stereoelectronic or steric effect, preorganizes the φ and ψ torsion angles to those required for triple-helix formation. The importance of the hydrogen bond in protein structure has competition with $n \rightarrow \pi^*$ interaction that arises from electron delocalization analogous to that of the hydrogen bond.^{10,59}

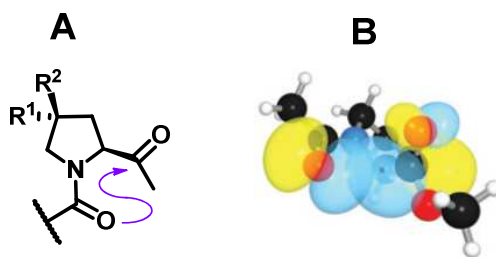


Figure 16: (a) An $n \rightarrow \pi^*$ interaction stabilizes the *trans* isomer of the peptide bond but is substantial only when Pro derivatives are in the C^{γ} -exo ring. (b) Depiction of overlap between n and π^* natural bond orbitals in a Pro residue with C^{γ} -exo pucker.¹⁰

The effect of a 4-X substituent on Pro ring pucker and the peptide bond isomerization ($K_{trans/cis}$) was explained from stereoelectronic effect and $n \rightarrow \pi^*$ interaction.⁶⁰ In $n \rightarrow \pi^*$ interaction, the amide oxygen of a peptide bond (O_{i-1}) donates electron density from its lone pairs into the antibonding orbital of the carbonyl in the subsequent peptide bond ($C_i = O_i$) (Figure 16). The C^{γ} -exo ring pucker of a Pro residue

provides a more favourable $O_{i-1} \dots C_i = O_i$ distance and angle for $n \rightarrow \pi^*$ interaction than does the C^γ -endo pucker. Importantly, $K_{trans/cis}$ for peptidyl prolyl amide bond is determined by the pyrrolidine ring pucker and not generally affected by the identity of substituents in the 4-position of the pyrrolidine ring.⁶¹ Because $n \rightarrow \pi^*$ interaction can occur only if the peptide bond containing O_{i-1} is *trans* it has an impact on the value of $K_{trans/cis}$ ratio for main chains with appropriate torsion angles. Thus, imposing a C^γ -exo pucker on a pyrrolidine ring in the Y position of a collagen like peptide preorganizes not only the φ and ψ angles for triple-helix formation, but also the ω angle. Indeed, single $n \rightarrow \pi^*$ interaction can stabilize the *trans* conformation by $\Delta G^\circ = -0.7$ kcal/mol.⁶²

1.3.6 Cis-trans isomerization

The peptide bond has partial double bond character and is planar.⁶³ The amide bond can adopt either in the *cis* ($\omega = 0^\circ$) or the *trans* ($\omega = 180^\circ$) conformation. With the exception of proline and its derivatives, the *cis*-content of a specific amide bond in unfolded proteins and peptides is very low. Secondary amides, *N*-methylformamide and *N*-methylacetamide, were used as models to study the equilibrium constants of *cis/trans* isomerization ($K_{cis/trans}$) in various solvents. The percentage of isomers with the *cis*-conformation is in the range of 1.4% to 10.5% with the lower percentage being seen in aqueous solutions.⁶⁴ Scherer *et al.*⁶⁵ studied peptides containing Gly, Ala, Phe, and Tyr in aqueous solution and found that only 0.11% to 0.48% of the amide bonds are in the *cis*-conformation.

The amide bond preceding the Pro residue has a higher, the *cis*-fraction content. It was reported that the *cis*-content can be as high as 10% to 30% in short unfolded Pro containing peptides, which means that the equilibrium constant between the *cis* and *trans*-conformation ($K_{cis/trans}$) is in the range of 0.11 to 0.43.⁶⁶ Schoetz *et al.*⁶⁷ reported that the *cis*-content of the amide bond can be as high as 46% in the dipeptide Phe-Pro as determined by dynamic capillary electrophoresis (DCE) study. The equilibrium constant was measured for model compounds containing 4-substituted prolyl derivatives in D_2O by NMR and the values of $K_{cis/trans}$ are in the reported range.⁶⁸ The $K_{cis/trans}$ value is 0.19 when a residue is attached to Pro, corresponding to 16% *cis*, and is 0.087 when a residue is attached to Hyp in unfolded type I collagen.⁶⁹

The stability of the collagen triple helix is determined by the free energy difference (ΔG°) between the unfolded state and the triple helix, which means that changes in the unfolded state contribute to the stability of the triple helix. The collagen triple helix requires all peptide bonds to be in the *trans*-conformation, so a change in the *cis* to *trans* ratio in the unfolded state will directly affect the triple helix stability.

1.3.7 Glycine substitution in model peptides and its correlation with disease

The Gly residue in the X-Y-Gly repeat is invariant in natural collagen and favorable substitutions are unknown in collagen related peptides. The host-guest triple-helical peptides with sequence Ac-(Pro-Hyp-Gly)₃-Pro-Hyp-Z-(Pro-Hyp-Gly)₄-Gly-Gly-amide were used to examine the influence of the residues replacing Gly on triple-helix stability with Z = Gly, Ala, Arg, Asn, Glu, Cys, Ser, or Val. Any substitution for Z = Gly (melting temperature, $T_m = 45^\circ\text{C}$) results in a dramatic destabilization of the triple helix. For Ala and Ser, T_m decreases to 10°C and for the Arg-, Val-, Glu- and Asp-containing peptides, $T_m < 0^\circ\text{C}$. The order of disruption of different Gly replacements in these peptides can be represented as Ala < Ser < Arg < Val < Glu < Asp. The rank of destabilization of substitutions for Gly in these Gly-Pro-Hyp-rich homotrimeric peptides shows a significant correlation with the severity of natural *osteogenesis imperfecta* (OI) mutations in the $\alpha 1$ chain of type I collagen.

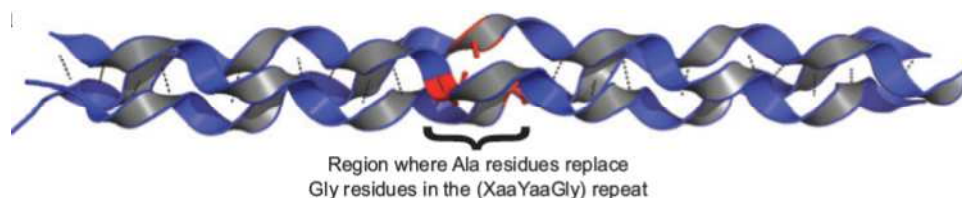


Figure 17: Impact of a Gly→Ala substitution on the structure of a collagen triple helix formed from the collagen-related peptide (Pro-Hyp-Gly)₄-(Pro-Hyp-Ala)-(Pro-Hyp-Gly)₅ [Protein Data Bank (PDB) entry 1cag (19)]. The Ala residues (*red*) disturb the structure. Mutations leading to such structural irregularities are common in osteogenesis imperfecta and can be lethal.¹⁰

1.3.8 Electrostatic interactions in the triple-helix

Charged residues are thought to play an important role in the triple-helix interactions as evidenced by their large proportion, high degree of conservation and asymmetric distribution.⁷⁰ The rod-like structure of the triple-helix allows large ratio of charged to hydrophobic residues than seen in globular proteins.⁷¹ Ionic residues Lys, Arg, Glu and Asp constitute 15-20% of the amino acid composition in fibrillar

collagens and about 40% of all X-Y-Gly triplets contain at least one charged residue. An asymmetric distribution of charged residues is found along the $(X-Y-Gly)_n$ collagen chain. Negatively charged residues are predominantly in the X position, while positively charged residues are predominantly in the Y position. It has been suggested that the preference for X or Y positions reflects the hydrogen bonding and electrostatic interactions between the residues in the triple-helix, as well as the steric factors. Charged residues are frequently clustered along the chain with the majority of basic residues located within 1 or 2 residues of acidic residue.⁷² The distribution of charged residues is highly conserved among the members of the fibril forming family.⁷³

The unique conformation of the triple-helix gives charged residues the potential to be involved in inter and intra-chain ion pairs. The thermal stability of collagens shows pH dependences, the stability is greater when all side chains are ionized, suggesting a supporting role for ion pairs in molecular stability.⁷⁴ All the non glycine residues in the triple-helix are substantially exposed to solvent, with X position being more exposed than the Y position. Theoretical calculations, chemical modification studies, pH dependence of fibril formation and ion binding studies indicate that electrostatic interactions are important in the interaction between neighbouring collagen molecules associated in periodic fibrils.⁷⁵ Charged residues have also been implicated in the binding of collagen to other extracellular matrix components such as cell-surface integrins.

1.3.9 Effect of charged termini on the stability of triple-helices

Orientation of individual chains in the triple-helix was first studied by Berg *et. al.*⁷⁶ who examined the changes in the T_m of the H-(Pro-Pro-Gly)₁₀-OH triple-helix as a function of pH. The T_m of the triple-helix was greatest at either low pH (<3) or high pH (>9), i.e., in fully protonated or fully non-protonated form. Brodsky *et. al.*⁷⁷ have studied the pH dependence of triple-helical strength of (Pro-Hyp-Gly)₁₀ which was greatest at acidic and basic pH, with $T_m = 61$ °C at pH 1.0 and $T_m = 62$ °C at pH 11.0. The triple-helical stability decreased around neutral pH with a minimum T_m of 56 °C at pH 6.0. This indicated that (Pro-Pro-Gly)_n model peptides form parallel triple helix similar to that of natural collagen. Because at extreme pH values only one of the termini is exclusively ionized, at any intermediate pH both the termini are ionized to

varying degrees. It was argued that in a parallel triple-helix, the presence of similar charged residues at the termini results in increased electrostatic repulsion of the chains, thus decreasing the triple-helical stability at intermediate pH values.

1.3.10 Effect of ionizable side-chains

Electrostatic interaction in triple-helix due to ionizable side chains has been studied by Brodsky *et. al.*⁷⁷ by introducing a zwitter-ionic dipeptide in the (Pro-Hyp-Gly)₅-(Glu-Lys-Gly)-(Pro-Hyp-Gly)₅ (EK containing peptide) and also by using the T3-487-peptide -(Pro-Hyp-Gly)-Ile-Thr-Gly-Ala-Arg-Gly-Leu-Ala-Gly-Pro-Hyp-Gly (Pro-Hyp-Gly)₃ corresponding to the 18 residues sequence of type III collagen with C-terminal Gly-Pro-Hyp tail. Variation of pH in the range 1.0-13.0 lead to 8-9 °C changes in the T_m of EK containing peptide and the T3-487 peptide. The greatest triple-helical stability was seen at pH where both acidic and basic residues are ionized. The NMR study on CB2 peptide 94-39 amino acid fragment of $\alpha 1$ chain of calf skin collagen has shown that the Arg residue in Y position is involved in hydrogen bonding with the adjacent chain in triple-helix.⁷⁸ The guanidinium hydrogens of Arg side chain were seen to be involved in slow exchange with water and hence thought to be the structural elements of triple-helix.

1.3.11 Effect of ethylene glycol on the stability of triple-helices

Intrachain peptide (amide) hydrogen bonding makes a major contribution to protein structure/stability and is effective only in the absence of accessible competing water. Water molecules bridge the carbonyl oxygen atoms and amide protons of different peptide links to catalyze the formation and reversal of peptide hydrogen bonds as well as formation of long-lived linkages stabilizing protein-ligand and protein-protein interfaces.⁷⁹ The internal molecular motions in proteins, necessary for biological activity, are very dependent on the degree of plasticizing, which is determined by the level of hydration.⁸⁰ Thus internal water enables the folding of proteins and is only expelled from the hydrophobic central core when finally squeezed out by cooperative protein chain interactions.⁸¹ It is possible to strengthen the intramolecular peptide hydrogen bonds using polar solvent not competing with hydrogen bonding of amide nitrogen. Polyols and sugars are known to offer protection to most proteins, including collagen triple-helix against thermal denaturation.⁸² Polyols contain hydrophilic

carbon chains that prevent the water from entering into the main chain domain of proteins, thereby enhancing the hydrogen bond strength between amide hydrogen and main chain carbonyl oxygen atoms. This study is very useful in concluding quantitatively the stability of protein structure by hydrogen bonding. If the stability of the protein structure is largely dependent on hydrogen bonding, then such proteins form a more stable structure in polyols compared to aqueous buffer condition. Ethylene glycol stabilizes helical structure and therefore can be very useful in amplifying and detect very weak triple-helical propensities.⁸³

1.4 Collagen Mimetics

Traditionally, collagen mimetic peptides (CMPs) have been used for elucidating the structure of the collagen triple helix and the factors responsible for its stabilization. The wealth of fundamental knowledge on collagen structure and cell–extracellular matrix (ECM) interactions accumulated over the past has led to exploration of the potential of CMPs to induce higher order assembly and modulate biological function of natural collagens for biomedical applications. The collagen triple helix has become a promising structural motif for engineering self-assembled, hierarchical constructs similar to natural tissue scaffolds which are expected to exhibit unique or enhanced biological activities. For this purpose the primary peptide sequence can be manipulated by incorporating different residues and sequences and the effect of such alternation can be investigated. By implementing a *de novo* approach, it is possible to design structures to induce triple helicity and enhance collagen stability.

Synthetic collagen model systems, also known as collagen mimetic peptides (CMPs), have been very useful in elucidating collagen structure and factors responsible for the stabilization of the triple helix. Among them, CMPs based on (Pro-Pro-Gly)_n and (Pro-Hyp-Gly)_n trimers have been widely studied, and their collagen-like triple helical structure and melting behaviours are documented in the literature. Unlike collagen molecules, CMPs exhibit reversible melting behavior due to their small size when heat denatured (melted) collagen is cooled whereas the collagen protein regains only a fraction of its original triple helical content, turning into gelatin. In contrast, CMPs regain 100% of its original triple helical structure despite a relatively slow folding rate allowing full thermodynamic characterization of folding

and melting processes. The small size of CMPs is also conducive to various structural analyses such as X-ray crystallography and NMR spectroscopy.

The thermal stability of CMPs is modulated by both the number of trimeric repeats and the amino acid composition. Longer peptides exhibit higher thermal stability with T_m of (Pro-Hyp-Gly) $_n$ rising from 37 °C for $n = 7$ to 43 °C and 69 °C for $n = 8$ and 10, respectively.⁸⁴ Raines *et.al.*⁸⁵ reported that the thermal stability of a CMP can be enhanced by incorporation of 4*R*-fluoroproline (Flp) in place of 4*R*-hydroxyproline (Hyp). In addition, 4*R*-azidoproline was found to have conformational properties and stabilizing effect similar to Flp and Hyp.⁸⁶ From these work, two distinct mechanisms by which amino acid side chains stabilize the collagen triple helix have been identified: (i) the hydration layer around the triple helix promoted by Hyp and charged amino acids, and (ii) the stereoelectronic effect of Hyp that stabilizes the *trans* peptide bond conformation of collagen strands. The glycosylating (Pro-Thr-Gly) $_n$ sequence has been reported to dramatically increase the triple helical stability in model collagen peptides and it is speculated that the (Pro-Thr-Gly) $_n$ (O-glycosylate) is the main component responsible for triple helix formation rather than the un-glycosylated precursor protein.⁸⁷

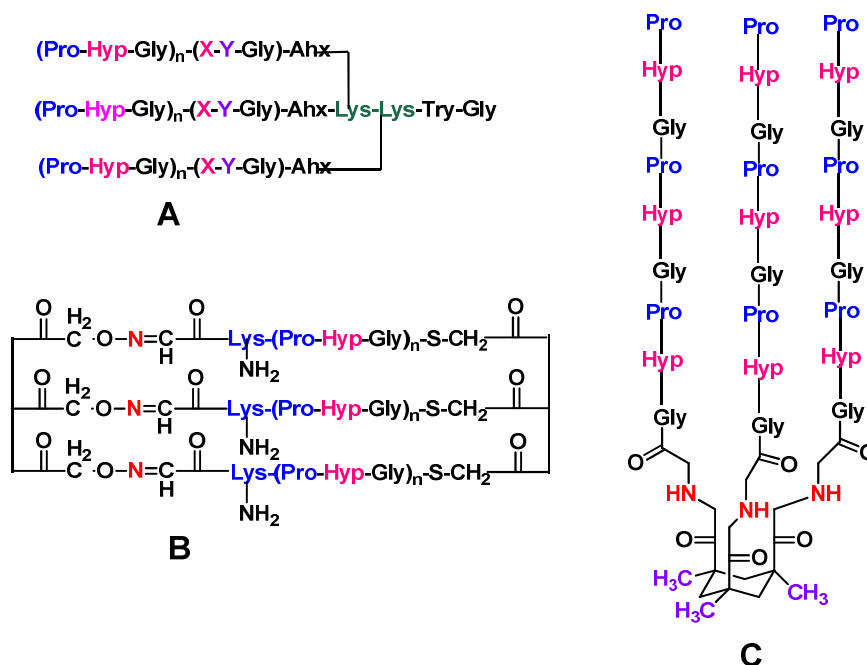


Figure 18. Template assembled collagen peptides. (A) Lysine template. (B) Lysyl-lysine template. (C) KTA template.

One of the well-known methods to stabilize a collagen triple helix is to link individual CMPs on a single template. An ideal template should have three functional

groups that can covalently link three peptide chains, and be flexible enough to allow proper packing of the three chains with correct amino acid register. Templates that have been used in the past are peptide templates with reactive amino acids (e.g. Lys or Glu), Kemp's tri-acid, and tris (2-aminoethyl) amine (TREN) succinic acid derivatives.⁸⁸ When various CMPs were conjugated to these templates, their T_m were elevated mainly because of reduction in entropy differences between folded and unfolded states of CMPs. Lately, Goodman *et.al.*⁸⁸ had developed a TRIS-based template (Figure 18) with an additional amino group that can be further utilized for attaching fluorescent tags or therapeutic agents for biomedical applications.⁸⁹

The use of CMPs for various biomedical applications requires ability to produce thermally stable CMPs in large quantity. Therefore understanding the origin of the structural stability within the collagen triple helix as well as the stabilization effect of various amino acids is critical for engineering synthetic collagens.

1.4.1 Collagen mimetics with unnatural amino acids

In attempts to understand and modify the triple-helical strength of collagen, several unnatural amino acids have been incorporated into the X and Y positions of collagen sequence (Figure 19).

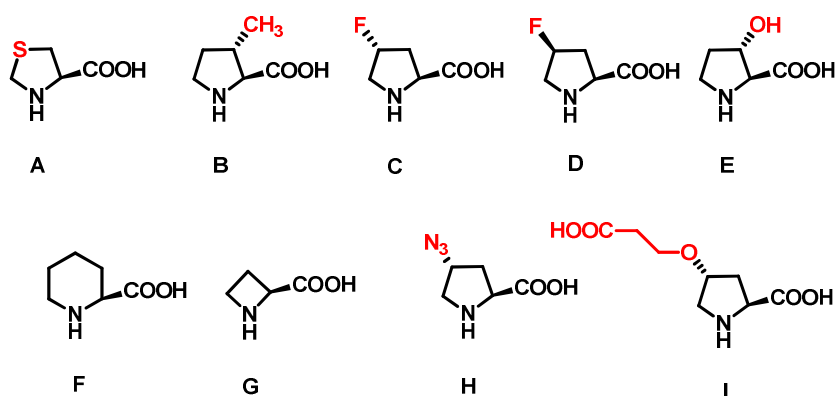


Figure 19: Various proline surrogates that have been incorporated in to collagen sequences. (A) Thiozolidine (Thz), (B) 3S-Methylproline (MePro), (C) 4R-Fluoroproline (Flp), (D) 4S-Fluoroproline (flp), (E) 3S-Hydroxyproline, (F) L-piperidine-2-carboxylic acid (Pipec), (G) Azitidine (Aze). (H) 4R-Azidoproline (Azp), (I) 4R-2-carboxyethoxy proline.

To extract the importance of stereochemistry of 4-hydroxyl group 4S-hydroxyproline (hyp) containing peptides (Pro-hyp-Gly)₁₀, (hyp-Pro-Gly)₁₀ have been

synthesized and studied for their triple-helix forming abilities. Variable temperature polarimetry and CD spectroscopy have shown that these peptides do not form triple-helices.⁹⁰ Synthesis of sequential polypeptides poly(Ala-Gly-Thz), Poly(Gly-Pipec-Ala) and poly-(Gly-Aze-Ala) have been reported.⁹¹ Peptide (Pro-MePro-Gly)_n was synthesized and tested for the *Proline-hydroxylase* inhibition activity.⁹² However, the conformational properties of these peptides are not known. Later work on poly(Pro-Aze-Gly), poly(Aze-Pro-Gly) and poly(Aze-Aze-Gly) has shown that poly(Pro-Aze-Gly) exhibits triple-helical character as per CD, vibrational spectroscopy and molecular modelling studies.

Several studies have evaluated the propensity of different amino acids to adopt the triple-helical conformation. A high proportion of charged residues is found in the sequence of the triple-helical domain and is presumed to participate in folding interactions.⁹³ Since residues in the X and Y are partially exposed to the solvent, charged residue has a potential for intra and intermolecular interactions. Modelling studies indicate that oppositely charged residues within a single chain form salt bridges when separated by one or two residues.⁹⁴ Subsequently, electrostatic interactions between opposite residues are sterically possible between the collagen polypeptide chains when separated by one residue along the chain.⁹⁵ Recently Chmielewski *et.al.*⁹⁶ had engineered a pH-responsive collagen triple helix peptide by introducing carboxylate groups at 4R position which protonated under acidic conditions. This design is complementary to the use of electrostatic interactions to promote the stabilization of triple helical peptides. Wennemers *et.al.*⁹⁷ demonstrated that 4R-Azidoproline has a similar stabilizing effect on the collagen triple helix as 4R-hydroxyproline and that functionalized collagen model peptides are readily accessible by “click” chemistry.

1.5 Collagen in Aging and Disease

The overall shape and function, in terms of flexibility and locomotion of the skeletal system depends on a basic framework of collagen fibers.⁹⁸ Collagen fibers provide mechanical strength that confers form, while allowing flexibility between various organs of the body. The biological diversity in the function of collagenous

tissue is primarily due to several genetically distinct classes of collagens that are tissue specific.

During aging, several chemical changes occur in the collagenous framework and these changes reflect in the physical properties of the fibers. The changes in the physico-chemical properties of collagen fibers that occur in the old age cause increase in stiffness of skin, tendon, bone and joints. The major change is an increase in the rigidity of the tissue, with the fibers ultimately becoming brittle.⁹⁹ Collagen synthesis decreases steadily with maturation and with subsequent aging drops 10-fold in majority of tissues. Collagen has low turnover in tissues, and the chemical changes in the matrix can be significant in old age. The changes in the mechanical properties of collagenous tissue in old age is ascribed to the cross-linking of the fibers to form a large polymeric network, resulting in decreased elasticity and increased brittleness of the tissue.¹⁰⁰ Cross-linking is essential for the strength of the collagenous tissue and is carried out in a controlled fashion by the enzyme *lysyl oxidase* during maturation of the tissue (growth phase of the organism). The *lysyl oxidase* oxidatively determines the lysine or hydroxyproline residue in the N-terminal regions of one peptide to an aldehyde. This aldehyde in turn reacts with the ϵ -amino group of lysine residues in the in the C-terminal region of the peptide in the adjacent fiber to form a reduced Schiff base, thus resulting in cross-linking. However, the decreased metabolic turnover of collagen with aging allows a second indirect cross-linking to occur through the reaction with glucose and its oxidation product, a process referred to as *glycation*. The open chain aldehyde form of glucose reacts with the free ϵ -amino acid side chain of lysine in the collagen peptides to form a *glycosyl-lysine*. This glycosylamine is then stabilized by spontaneous Amadori rearrangement to form a keto-imine, aminodeoxyketoase (Figure 20).¹⁰¹

The reaction of a pentose, for example ribose, with the side chain of Arg and Lys residues on different peptide chains results in the formation of *pentosidine* crosslinks.¹⁰² Several other mechanisms operate leading to chemical modifications that reduce the flexibility of collagenous tissue. Recently, type-I collagen has been shown to undergo β -isomerization of Asp-Gly bond within the C-terminal region and the extent of isomerization was shown to increase with age.¹⁰³

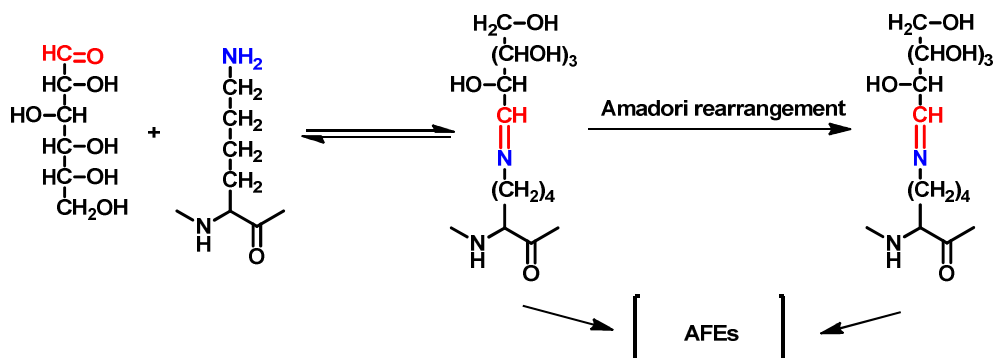


Figure 20: Reaction of glucose and lysine in collagen peptide to form a Schiff's base, which spontaneously undergoes an Amadori rearrangement to give aminodeoxyketose.

The critical role of collagen is clearly illustrated by the wide spectrum of diseases caused by more than 1000 mutations that have thus far been identified in 22 genes for 12 out of 20 collagen types.¹⁰⁴ Mutations that alter the expression or primary structure of collagen are the predominant causes of severe skeletal defects such as *Osteogenesis imperfecta* or brittle-bone diseases, and *Chondrodysplasias*.¹⁰⁵ Mutations that have milder effects on the synthesis or structure of protein are found in more common diseases such as *Osteoporosis* and *Osteoarthritis*. Deficiencies in the post-translational modification of the collagen are also known to cause several heritable disorders. For examples, in Ehlers-Danlos syndrome type VI, the collagen fiber fragility arises from the hydroxylation deficiency. Conversely, over-hydroxylation of collagen, as seen in several forms of *Osteogenesis imperfecta*, results in retarded triple-helix formation, thus altering collagen fibril formation and consequently decreasing the bone strength.¹⁰⁶ Among the acquired diseases, transient increases in hydroxylation occur in collagen during bone fracture-repair and *Osteoporosis*.¹⁰⁷ In all these cases, there is an increase in the number of hydroxylated cross-links and evidence for a decrease in fiber size. Several subtypes of Alport syndrome, Bethlem myopathy, certain subtypes of Epidermolysis bullosa, Knobloch syndrome, Arterial aneurysms, Osteoarthritis and Intervertebral disk disease have also been ascribed to the imperfections in the collagenous tissue.

1.6 Biomedical Applications of Collagen

Collagen is regarded as one of the most useful biomaterials. The excellent biocompatibility and safety due to its biological characteristics, such as

biodegradability and weak antigenicity made collagen the primary resource in medical applications. The main applications of collagen as drug delivery systems are collagen shields in ophthalmology, sponges for burns/wounds, mini-pellets and tablets for protein delivery, gel formulation in combination with liposomes for sustained drug delivery, as controlling material for transdermal delivery, and nanoparticles for gene delivery and basic matrices for cell culture systems. It was also used for tissue engineering including skin replacement, bone substitutes, and artificial blood vessels and valves. Table 2 summarizes the major characteristics of collagen, which are suitable for medical applications.

Table 2: Biomedical applications of collagen¹⁰⁸

Advantages	Disadvantages
Available in abundance and easily purified from living organisms (constitutes more than 30% of vertebrate tissues)	High cost of pure type I collagen
Non-antigenic	Variability of isolated collagen (e.g. crosslink density, fiber size, trace impurities, etc.)
Biodegradable and bio reabsorbable	Hydrophilicity which leads to swelling and more rapid release
Non-toxic and biocompatible	Variability in enzymatic degradation rate as compared with hydrolytic degradation
Synergic with bioactive components	Complex handling properties
Biological plastic due to high tensile strength and minimal expressibility	Side effects, such as bovine spongiform encephalopathy (BSF) and mineralization
Hemostatic - promotes blood coagulation	
Formulated in a number of different forms	
Biodegradability can be regulated by cross-linking	
Easily modifiable to produce materials as desired by utilizing its functional groups	
Compatible with synthetic polymers	

Relatively few collagen related peptides (CRPs) have been tested as biomaterials. Goodman *et.al.*¹⁰⁹ showed that peptoid-containing collagen related peptides have a notable ability to bind to epithelial cells and fibroblasts, particularly when displayed on a surface. CRPs are also useful for inducing platelet aggregation, which can aid the wound-healing process.¹¹⁰

A key step toward utilizing collagenous biomaterials for therapeutic purposes is the development of collagen related peptides that can either adhere to or bury themselves within biological collagen. Most efforts toward these objectives have relied on immobilization of CRPs on an unrelated substance. The collagen related peptides functionalized with gold nanoparticles was prepared and demonstrated binding of the gold nanoparticles to the gap region of natural collagen.¹¹¹ The collagen related peptides displayed on latex nanoparticles can stimulate human platelet aggregation with potency similar to that of type I collagen. The triple-helical fibrils obtained *via* aromatic interactions had a similar level of thrombogenic activity to the CRPs immobilized on latex nanoparticles.¹¹² The single strands of CRPs and polyethylene glycol conjugated CRPs bind to collagen films even without immobilization on nanoparticles and are of potential use in collagen imaging and wound-healing applications.¹¹³

1.6.1 Collagen-based drug delivery systems

1.6.1a Film/sheet/disc: The main application of collagen films is as barrier membrane. Collagen film/sheet/disc has been used for the treatment of tissue infection, such as infected corneal tissue or liver cancer. Soluble ophthalmic insert in the form of a wafer or a film was introduced as a drug delivery system for the treatment of infected corneal tissue using a high dose of antibiotic agents, such as gentamicin and tetracycline. The wafer route of administration gave the highest tissue concentration of incorporated drugs. The microfibrinous collagen sheets as a local delivery carrier for the treatment of cancer was evaluated. The local application of collagen sheets loaded with anticancer agent, ectopocide, resulted in a relatively long maintenance of drug concentrations at the target site, which, in this case, is the liver. Collagen film has been used for the treatment of tissue infection, such as infected corneal tissue or liver cancer. The local application of collagen sheets loaded with anticancer agent, ectopocide (VP-16), resulted in a relatively long maintenance of drug concentrations at the target site. Collagen film and matrix were used as gene delivery carriers for promoting bone formation. A composite of recombinant human bone morphogenetic protein 2 (rhBMP 2) and collagen was developed to monitor bone development and absorbent change of carrier collagen.¹⁰⁸

1.6.1b Collagen shields: The collagen shield was originally designed for bandage contact lenses, which are gradually dissolved in cornea. The idea of using a shield or hydrogel lens delivery device has led to the development of various drug delivery systems for ophthalmic applications. The collagen corneal shield is fabricated from porcine sclera tissue that closely resembles collagen molecule of human eye. The cross-linked collagen using glutaraldehyde or chromium tanning can serve as a drug reservoir and provide more desirable drug delivery than non cross-linked collagen shields by increasing the contact time between the drug and the cornea. Collasomes, in the form of collagen pieces, were developed by adding long hydrocarbon side chains to the collagen. This modification increases not only the hydrophobicity of the collagen but also the total surface area, which effectively decrease the diffusion rate of hydrophilic drug molecules from the collagen matrix. Collasomes can be formulated with various constituents and chemically alternated with the addition of lipid (Lacrisomes) for the treatment of dry eyes.¹⁰⁸

1.6.1c Collagen sponges: Collagen sponges have been very useful in the treatment of severe burns and as a dressing for many types of wounds, such as pressure sores, leg ulcers and decubitus ulcer as well as for in *vitro* test systems. Human collagen membrane has been a major resource of collagen sponge used as a biological dressing. Collagen sponges have ability to easily absorb large quantities of tissue exudates, smooth adherence to the wet wound bed with preservation of low moist climate. Experiments using sponge implantation demonstrated a rapid recovery of skin from burn wounds by an intense infiltration of neutrophils into the sponge. Coating of a collagen sponge with growth factor further facilitates dermal and epidermal wound healing. The main drawbacks of sponges appear to be difficulty of assuring adequate supplies and their preservation. Other problems arose from their poor mechanical strength, ineffectiveness in the management of infected wounds and burns.

Other methods for drug delivery systems include collagen films as a calcifiable matrix system, gel, hydrogel, liposomes-collagen, nanoparticles/ nanospheres, aqueous indictable and collagen conjugates.¹⁰⁸

1.6.2 Collagen-based systems for tissue engineering

1.6.2a Collagen as skin replacement: Collagen based implants have been widely used as vehicles for transportation of cultured skin cells or drug carriers for skin replacement and burn wounds. Since sponge implant was originally developed for recovery of skin and was very efficient in that purpose, various types of artificial skin were developed as a form of sponge. Cultured skin substitutes developed on collagen lattice were also used for skin replacement and skin wounds. Reconstituted type I collagen is suitable for skin replacement and burn wounds due to their mechanical strength and biocompatibility. Chronic wounds resulting from diabetes have been successfully cured with allogenic cultured skin substitutes prepared from cryo-preserved skin cells. The modified sponge for artificial skin was developed by combining fibrillar collagen with gelatin. Dehydrothermal crosslinks were used to stabilize collagen-based sponge physically and metabolically. Some limitations inherent to cultured skin substitutes, such as deficient barrier function *in vitro* and delayed keratinisation after grafting in comparison to native skin autografts.¹⁰⁸

1.6.2b Collagen as bone substitutes: Among the many tissues in the human body, bone has been considered as a powerful marker for regeneration and its formation serves as a prototype model for tissue engineering based on morphogenesis.

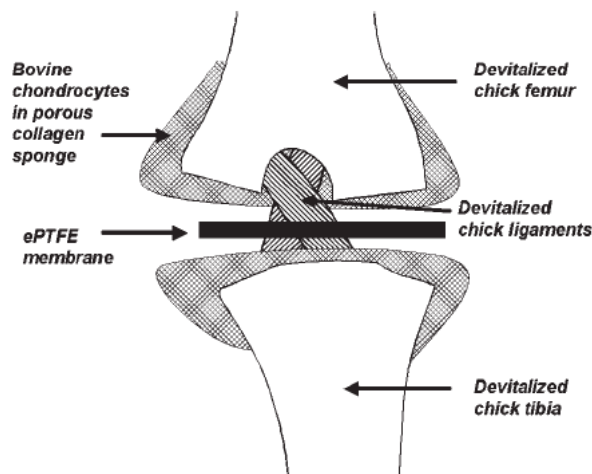


Figure 21: Composite chimeric joint. Two porous collagen sponges containing bovine chondrocytes are placed against shaved articular surfaces of femur and tibia of a devitalized embryonic chick knee. An ePTFE (expanded polytetrafluoroethylene) membrane is positioned to maintain the joint space.

Collagen has been used as implantable carriers for bone inducing proteins, such as bone morphogenetic protein 2 (rhBMP-2). Collagen itself was used as bone substitutes due to its osteoinductive activity. Type I collagen cross linked *N*-telopeptide was used as a marker of bone resorption and clinically used as a marker of bone metastasis of prostate cancer and breast cancer. Collagen in combination with other polymers or chemicals was also used for orthopaedic defects. Demineralized bone collagen was used as a bone graft material for the treatment of acquired and congenital orthopaedic defects either by itself or in combination with hydroxyapatite.¹⁰⁸

1.6.2c Collagen as bioengineered tissues: Collagen gel as human skin substitutes have demonstrated its usefulness in tissue engineering and led to the development of bioengineered tissues, such as blood vessels, heart valves and ligaments.¹¹⁴ Collagen shows hemostatic properties that promote blood coagulation and play an important role in tissue repair process. Collagen sponge or gel initiates adhesion and aggregation of platelets that lead to a thrombus formation.¹¹⁵ Monomeric collagen does not activate platelet aggregation, while polymeric collagen having a regular arrangement of the molecules with a length of around 1 μm does activate it. Arginine side chains of collagen seemed to be responsible for its interaction with platelets.¹¹⁶

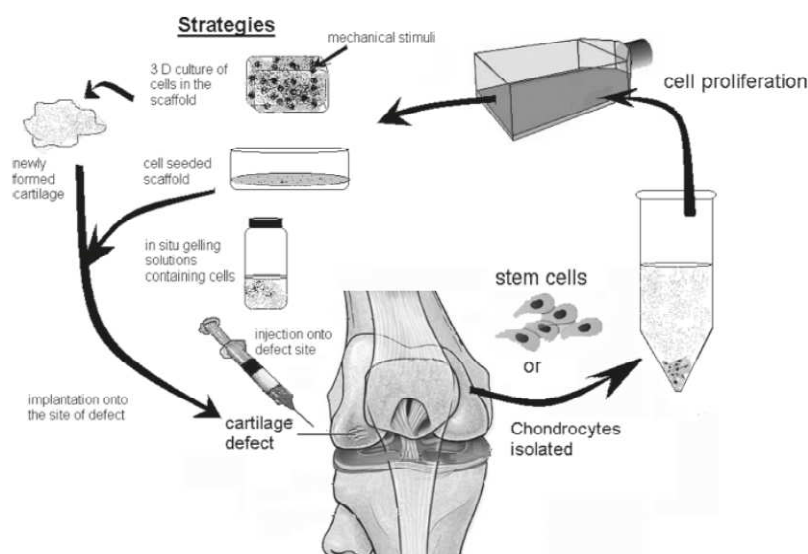


Figure 22: Schematic representation of methods of cartilage tissue engineering.

Natural collagenous materials were used for surgical repair and abdominal wall repair by taking advantage of their inherent low antigenicity and their ability to integrate with surrounding tissues.¹¹⁷ Moreover, new generations of collagen-based

biological tissue are practical and remodelable due to its simple membranous configuration, relative uniformity and abundant availability. These characteristics are employed in a new type of surgical adhesive made from porcine collagen and polyglutamic acid, developed for sealing air leakage from the lung, which takes a relatively long period for recovery.¹¹⁸ Recent progress in tissue engineering may lead to well-characterized and reproducible biomaterials from natural collagenous materials.

1.6.3 Collagen cross-links

The mechanical strength of collagen fibrils depends on a highly regulated mechanism of intermolecular cross-linking. The basis of this cross-linking from the most primitive to the most advanced multi-cellular animals and across a diversity of vertebrate tissue types, is the formation of covalent bonds from aldehydes produced from lysyl and hydroxylysyl side chains by lysyl oxidase. In past it has become clear that such bonds form not only between collagen molecules of the same type in homopolymeric fibrils but also between different types of collagen molecule that have evolved to interact and form heteromeric structures. Cross-linking amino acids and peptides containing them from collagen degradation, have received attention as bone resorption biomarkers in clinical studies and drug trials in the osteoporosis field.

Variations in cross-linking chemistry appear to be more tissue-specific than collagen type-specific. This is understandable if a single cell type synthesizing multiple collagens passes them through the same processing enzymes and endoplasmic reticular pathway. The basic pathway of cross-linking is regulated primarily by the hydroxylation pattern of telopeptide and triple-helix domain lysine residues. The flanking sequences around the cross-linking lysine residues, however, can affect the ensuing chemistry. An example is the participation of a histidine residue in the formation of the mature trivalent cross-links found in skin type I collagen. A histidine in the $\alpha_2(I)$ chain reacts with a vicinal aldimine cross-link formed between a lysine aldehyde and a hydroxylysine residue in two 4D-staggered collagen I molecules.¹¹⁹

1.7 Characterization of Triple Helical Structures

1.7.1 X-Ray crystallography of triple helical peptides

In 1954, Ramachandran and Kartha provided the first¹² correct model for the structure of collagen as three parallel left-handed polypeptide chains, held together by interchain hydrogen bonds, which was later modified by themselves to be a coiled-coil structure. Soon after Rich & Crick¹³ tailored this model to propose another model where each of the strands in triple-helix has left-handed coil (PPII) which intertwined one hydrogen bond per every trimer repeat to form a right-handed triple-helix conformation (Figure 23)

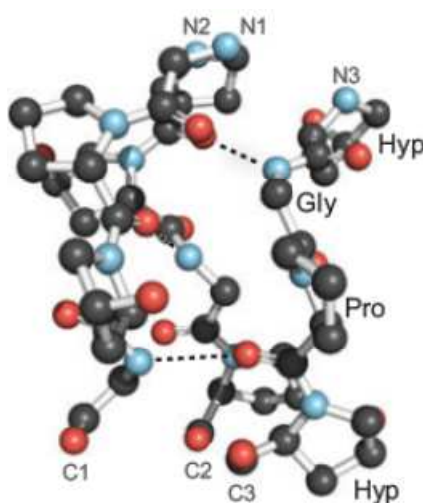


Figure 23: Ball-and-stick image of a segment of collagen triple helix (PDB entry 1cag 19), highlighting the ladder of interstrand hydrogen bonds.

But high resolution structures of triple-helix peptides obtained by X-ray crystallography provided confirmation and new insights into the molecular conformation. The first crystal structure was obtained by Okuyama *et al.*¹²⁰ for (Pro-Pro-Gly)₁₀, followed by structures for peptides with varying sequences from a number of laboratories (Table 3).¹²¹ These high-resolution structures of collagen-like peptides confirm the basic triple-helical model and show details of hydration, hydrogen bonding, and helical parameters.

Most X-ray crystallographic studies on collagen-related peptides have been performed on proline-rich collagenous sequences. All the resulting structures have a 7/2 helical pitch (20.0-Å axial repeat), in contrast to the 10/3 helical pitch (28.6-Å axial repeat) predicted for natural collagen by fiber diffraction.

Table 3: List of examples of collagen peptides with high resolution structures solved by X-Ray crystallography

	Peptide sequences	PDB ID	Properties
<i>Imino acid polytripeptides</i>	(Pro-Pro-Gly) _n where n= 9, 10 (Pro-Hyp-Gly) ₁₀ (Hyp-Hyp-Gly) ₉	1IA31, 1K6F, 2CUO, 1V7H, 1YMB8	7/2 symmetry, NH(Gly)...CO(X)
<i>Peptides with collagen sequences:</i> LOG2	(Pro-Hyp-Gly) ₄ -(Lue-Hyp-Gly) ₂ -(Pro-Hyp-Gly) ₄ (Gly-Pro-Hyp) ₂ -Gly-Phe-Hyp-Gly- Glu-Arg-(Gly-Pro-Hyp) ₃		NH(Gly)...CO(X) NH(X)..W..CO(Gly)
Integrin binding peptide	(Pro-Hyp-Gly) ₃ -Ile-Thr-Gly-Ala-Arg- Gly-Leu-Ala-Gly-(Pro-Hyp-Gly) ₄	1DZ1, 1Q7D	7/2 symmetry in ends, more relaxed in central imino acid poor zone
T3-785 peptide		1BKV	
<i>Peptides with Imperfect Gly-X-Y repeat:</i>			
Peptide with natural interruption	(Pro-Hyp-Gly) ₃ -Pro-Hyp-Gly-Pro- Gly-(Pro-Hyp-Gly) ₃ (Pro-Hyp-Gly) ₄ -Pro-Hyp-Ala-(Pro- Hyp-Gly) ₅	1E18	Altered H bonding at interruption site, good helix at both ends, but out of register
Peptides with Gly→Ala mutation		1CAG, 1CGD	

In addition to X-ray diffraction analysis, nuclear magnetic resonance (NMR) spectroscopy has been used to probe the structure of the collagen triple helix.

1.7.2 NMR spectroscopy

The knowledge of the structure and dynamics of the collagen triple helix is important for understanding its many interactions with receptors and other matrix molecules as well as the perturbations that are caused by collagen disease mutations.¹²² Collagen-like peptides are well suited for NMR studies because of the high solubility of both the single helix and the triple helix forms, relatively low molecular weight and slow folding. However, the triple helix form of these peptides has intrinsic features that complicate the structure determination. Such features include rod-like shape, repetitive sequence, multichain composition and symmetry. The triple helix is an extended uniform conformation, giving rise to a rod-like molecule, so that even a 30-residue peptide has an axial ratio of greater than 8:1. This leads to anisotropic tumbling which results in the broadening of NMR peaks expected for a much higher molecular weight molecule and the dependence of NOE intensities on the internuclear ¹H-¹H vector and the long axis of the molecule.¹²³

Multidimensional NMR studies have been reported on simple repeating poly tripeptides, such as (Pro-Hyp-Gly)₁₀ and shorter cross-linked peptides containing repeating Gly-Pro-Hyp or Gly-Pro-Nleu sequences.¹²⁴ Proton NMR including TOCSY, COSY and NOESY experiments have been used to obtain sequence specific

assignments for the repeating triplet unit in these peptides, but overlapping resonances prevent assignments to individual chains.¹²⁴ Because of the inability to distinguish different chains and the identical repetitive sequence within a given chain, a NOE cross peak may arise from interactions between atoms in the same chain and/or between atoms from different chains. For example, an NOE between Gly-NH and Hyp-C α H (Figure 24) can arise from either an interchain or an intrachain interaction. Direct structural determination based on NOE derived distance constraints is therefore not possible.

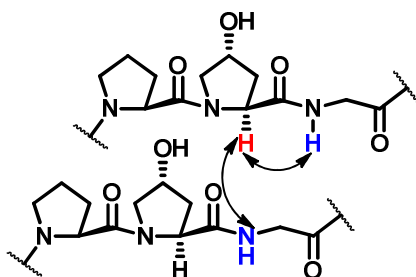


Figure 24: NOE between Gly-NH and Hyp-C α H can arise from either an interchain or an intrachain interaction.

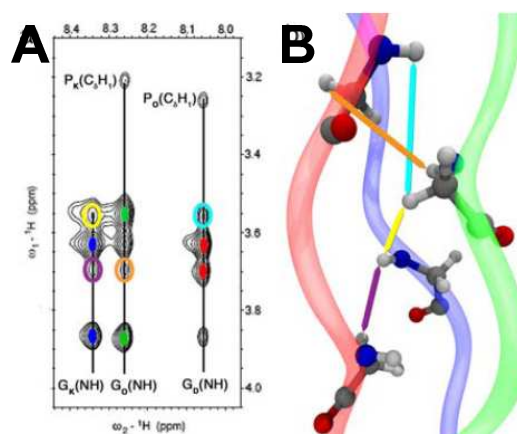


Figure 25: Two-dimensional NOESY spectrum and homology model. (A) NOESY spectrum shows the resonances from the glycine amide hydrogen to the α -protons. The vertical lines mark the NH chemical shifts; the solid colored ellipses indicate the position of the intra-strand cross-peaks (red stands for D, green for O, and blue for K), and the open ellipses indicate the unambiguous inter-strand interactions. The P(C δ H1)-G(NH) cross-peaks are also shown. (B) Homology model highlighting the atoms that give rise to the interstrand cross-peaks in the spectrum using colored arrows.¹²⁵

An initial structural approach is taken by using the X-ray derived triple-helix model to determine which NOEs arise uniquely from interchain interactions. The observation of these NOEs confirms that the basic triple-helix model is present in solution in (Pro-Hyp-Gly)₁₀ and in the cross-linked peptides. The repetitive sequence and the

stabilizing nature of Gly-Pro-Hyp triplets makes the peptide (Pro-Hyp-Gly)₁₀ a useful standard to compare with peptides with more varied sequences that model collagen.

Nuclear magnetic resonance (NMR) spectroscopy has been used to probe the structure of the collagen triple helix in solution, its dynamics and folding kinetics.¹²⁶ In general, the structural data obtained by NMR spectroscopy and X-ray diffraction analysis are in good agreement. Recently Baum *et. al.*¹²⁷ reported ¹⁵N relaxation NMR measurements on a collagen-like model peptide which revealed that the orientations of the Gly N-H bonds relative to the protein backbone have an unanticipated geometry and suggest that the H bonding may be responsible for this effect.

The structure and remodeling of collagen in *vivo* is critical to the pathology and healing of many human diseases as well as to normal tissue development and regeneration. In recent studies, it has been shown that collagen matrix structure has a major impact on cell and tissue outcomes related to cellular aging and differentiation potential. Collagen structure is complex because of both diversity of source materials, chemistry and structural hierarchy. The analysis should include as a starting point:

- ❖ Structural detail—mainly focused on molecular mass, purity, helical content, and bulk thermal properties.
- ❖ Chemical features—mainly focused on surface elemental analysis and hydrophobicity.
- ❖ Morphological features at different length scales.

The application of these analytical techniques to the characterization of collagen biomaterial matrices is critical in order to appropriately correlate biological responses from different studies with experimental outcomes in *vitro* or in *vivo*.¹²⁸ Here some methods are discussed for characterization of collagen triple helical structure.

Table 4 includes a technique summary listing for each method, the collagen format needed, solution concentrations, information gathered, and major limitations of the method.

Table 4: Technique summary listing for each method¹²⁸

Collagen description	Mass spectroscopy	SDS-PAGE	CD	DSC	XPS	Contact angle	AFM	SEM	ESEM
Material format	Single strand	Solution	Solution in cuvette or film on plate	Dehydrated film in DSC pan	Films, fibers, gels	Films or fibers	Films or fibers	Films, fibers, gels, sponges	Films, fibers, gels, sponges
Concentration, solvent	10 μM , water	10-20 μg , water	0.125 $\mu\text{g } \mu\text{L}^{-1}$ solution or 100 μg dry weight film	5 mg dry weight	Films: 15-780 $\mu\text{g cm}^{-1}$	Films: 15-780 $\mu\text{g cm}^{-1}$	Films: 15-780 $\mu\text{g cm}^{-1}$	Films: 15-780 $\mu\text{g cm}^{-1}$	Films: 15-780 $\mu\text{g cm}^{-1}$
Information	Mass	Size	Secondary and tertiary structure	Denaturation temperature, heat capacity	Atomic composition	Hydrophobicity	Molecular topography	Topography	Topography
Major limitation	Mass range, specialized equipment	Denaturation	Deconvolution software limited	Large mass needed	Specialized equipment	Hydration, surface smoothness	Specialized equipment	Artifacts due to dehydration	Resolution of features

1.7.3 Mass spectroscopy

Collagen proteins have been reported with molecular masses from 283 to 300 kDa.¹²⁹ Mass spectroscopy can be employed to provide the molecular mass data on collagen as well as the identification of telopeptides and other potential contaminants in collagen preparations that can directly impact cell functions. Matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectroscopy is commonly used.

Collagen chains \approx 94 kDa to 98 kDa exceed the working range of many mass spectrometers. Therefore digestion of the \approx 300 kDa collagen to \approx 94 to 98 kDa fragments is required. The lack of data reported for digested collagen and the complication due to its high molecular mass provide significant challenges in mass spectroscopy assessments of collagens, thus gel electrophoresis is more commonly utilized. Complications also arise from low-molecular-mass collagen telopeptides (\approx 6–14 kDa⁴²) that usually require a separate MALDI-TOF matrix for appropriate analysis. Mass spectroscopy is also useful to track bone remodeling and the formation of new bone collagen, where the presence of telopeptides is common. The purity of the collagen source material can also be assessed using mass spectroscopy.

1.7.4 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) is the most commonly used to assess collagen source material purity and breakdown. Coomassie blue or silver stains are commonly used to visualize the protein bands. Small sample amounts from nanogram to microgram are needed and molecular mass banding patterns are obtained. Subsequent western blots can be used to assess the specificity of collagen type using monoclonal antibodies.¹³⁰

1.7.5 Characterization of triple-helical structures by circular dichroism

Circular dichroism (CD) spectroscopy is one of the most widely employed tools in the characterization of the triple-helix structure and its strength. In solution, collagen like triple-helix and polyproline II like helices exhibit fingerprint CD spectra.¹³¹ These spectra are characteristic by the presence of a large negative band around 200 nm, a crossover at around 213 nm, and a small positive band around 215–227 nm (Figure 26). The presence of triplex or single helix can be determined by concentration dependent CD spectroscopy. The percentage of triple-helical structure is close to a

maximum when the concentration is greater than its critical triple-helical concentration. The magnitude of the ratio of positive to negative band intensity in the CD spectra of triple helical peptides ($R_{p/n}$) has been proposed to quantitate the triple-helical strength. The sharp co-operative transition state was observed for temperature dependent CD spectroscopy.

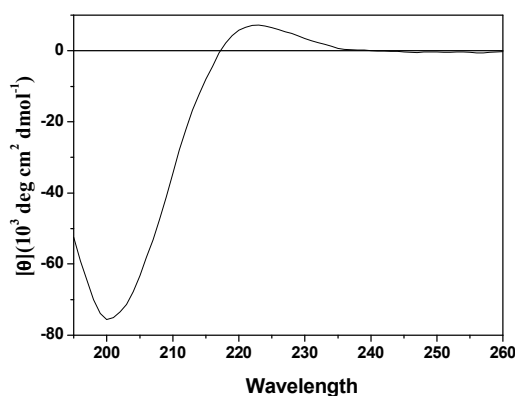


Figure 26: CD spectra of collagen triple-helix structure.

1.7.5a Thermal stability of collagen triple helix: The challenges of characterizing the thermal stability of the collagen triple helix are the lack of the reversibility of the thermal transition and the presence of multiple folding–unfolding steps during the thermal transition. The irreversibility exhibited by only proteins, while the short peptides exhibit reversibility during the thermal denaturation. The subdomain of collagen generally refers to a segment of the long, rope-like triple helical molecule that can unfold cooperatively as an independent unit whose properties. These are considered essential for the molecular recognition during the self-assembly of collagen and during the interactions of collagen with other macromolecules.¹³²

The thermal studies of short, synthetic peptides have established the molecular basis of the sub domains of collagen protein by demonstrating the close dependence of the thermal stability of a triple helix on the constituent amino acid residues at the X and the Y positions in the repeating sequence of X-Y-Gly. The typical thermal denaturation pattern for collagen like peptides is monitored CD spectroscopy is shown in Figure 27. The quality of CD data depends on sample concentration and temperature. Sufficient temperature control is also required to avoid denaturation under experimental conditions. Ellipticity data noting the angle of polarization of light, reported in milli degrees, can be converted to mean residue ellipticity [degree cm^2

dmol^{-1}]; however, the molecular mass of the sample is required for further conversion to molar ellipticity [$\text{dL mol}^{-1} \text{dm}^{-1}$]. Despite the complications associated with collecting and comparing data on collagen structure by CD, this remains a powerful tool to assess helicity and thus degree of naturation/denaturation of a sample used in biological studies. The ellipticity determined by CD provides a measure of this structural feature and thus an assessment of the native/denatured state of the collagen preparation.

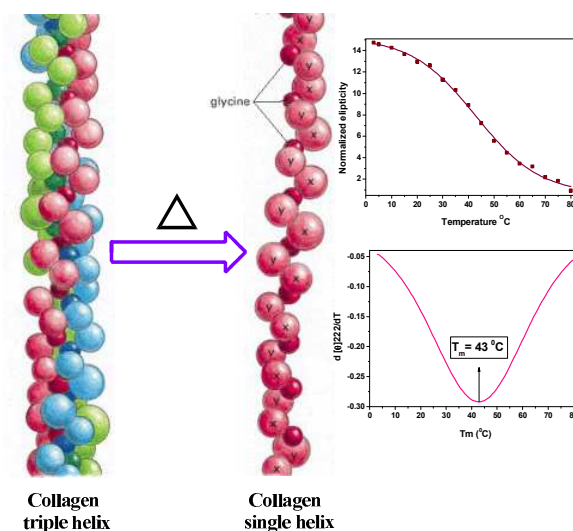


Figure 27: Collagen triple helix melts to single helix

In each case, the accuracy of the quantitative assessment of helicity depends strongly on the accuracy of the solution concentration, molecular mass, and amino acid content. These data can often be problematic for collagens, as they are often not well defined. A major limitation to the interpretation of CD data for collagens and other fibrous proteins is that the current algorithms used for conversion of signals to structural information are based on globular proteins.

1.7.6 Differential scanning calorimetry

Differential scanning calorimetry (DSC) provides direct determination of enthalpy (ΔH) by measuring the temperature dependence of partial heat capacity. With known masses and temperature changes, the sample heat capacity and melting temperature can be calculated.¹³³ DSC is frequently used to characterize the bulk thermal characteristics of a biomaterial,¹³⁴ including cross linked collagen.¹³⁵ Thermal properties of collagen-based scaffolds provide information on transitions in the structural state, reflecting initial primary (chemistry) sequence, structural state, degree

of cross linking and also purity of samples. Most often the researcher determines the apparent T_m , as the thermal unfolding of large proteins like collagen is usually irreversible.

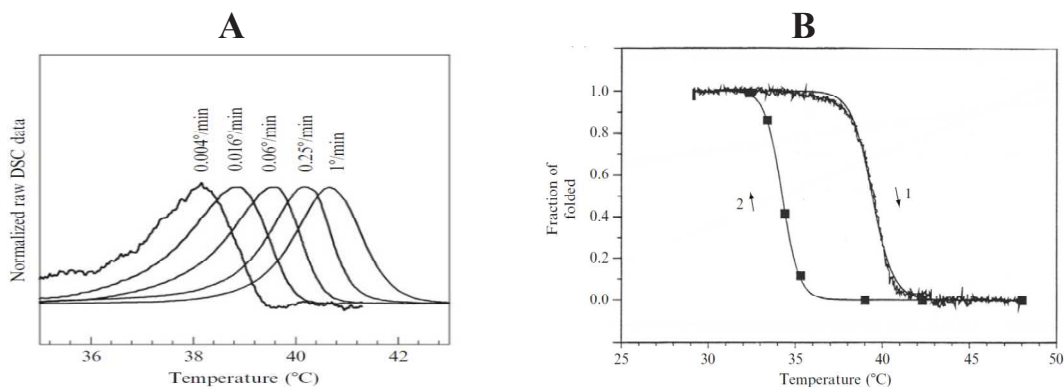


Figure 28: (A) The thermal unfolding of collagen measured by DSC at different heating rate in 0.2 M phosphate/0.5 M glycerol (pH 7.4). (B) Thermal unfolding and refolding profiles of collagen III¹³²

1.7.7 Surface morphology

Cells respond to surface morphology or the roughness. In order to characterize the surface topology of a collagen-based biomaterial, a variety of surface imaging tools can be used including atomic force microscopy (AFM), scanning electron microscopy (SEM), environmental scanning electron microscopy (ESEM), and light microscopy. Each technique offers advantages and disadvantages associated with sample preparation and resolution, thus often combinations of several microscopy techniques are considered. The collection of SEM data requires appropriate sample preparation such as sputter coating with gold, which can dampen surface resolution. The soft nature of collagen can result in cracking of the coating during imaging. Therefore, ESEM is often a more suitable choice, since the problems with SEM are avoided, and resolution is usually sufficient although not as good as SEM.

1.7.7a Atomic force microscopy

Nanometer-scale resolution is achieved with AFM, providing input on scales related to surface receptors and protein interactions. Contact imaging *via* tapping mode using the microscopic probe on the surface of a sample is accompanied by measurements of force deflection of the cantilever on which the probe is mounted, to generate the readout of surface morphology. AFM has been used to characterize the surface roughness of collagen coated with polystyrene and oxidized polystyrene,¹³⁶ and to characterize surface roughness with addition of collagen films on poly (ϵ -

caprolactone).¹³⁷ AFM has also been used to pattern surfaces with collagens and collagen peptides in the dip pen lithography mode with line resolution to 30- to 50-nm line widths.¹³⁸

1.7.7b Scanning electron microscopy

Depending on the specific model, SEM magnification of 3000 to 30000 can be achieved and can be used to image the collagen substrate and cells grown on these surfaces. The shape of human lung fibroblasts, IMR-90 cells, growing on collagen has been related to cell age using SEM, with older cells exhibiting a larger more irregular morphology.¹³⁹ Methods have been developed relating ratios of maximum to minimum cell length to characterize cell morphology in the study of stromal cell spreading.¹⁴⁰ In collagen scaffolds engineered for artificial dermis applications, SEM has been used to quantitate the pore sizes of the scaffolds as well as the extent of collagen cross linking.

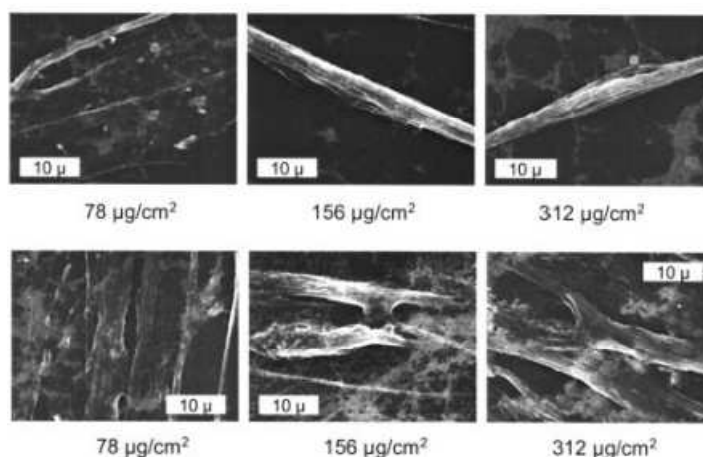


Figure 29: Scanning electron microscopy on nondenatured collagen films, increasing surface roughness is observed as a function of increasing collagen concentration.¹²⁸

1.7.7c Environmental Scanning Electron Microscopy

ESEM uses a high vacuum, high humidity chamber to image samples without sputter-coating. Both native and denatured collagen-cell samples have been imaged. For example, to characterize self-assembled fibrils of collagen composites for bone-tissue engineering, ESEM was used. In tissue engineering arterial constructs with collagen coatings, ESEM was used to image cells.¹⁴¹

2 Present work: Rationale

The work in this part of thesis is mainly devoted to the synthesis and structural characterization of new class of synthetic collagen model peptides with 4*R/S*-aminoproline substitution at different positions and evaluation of the position dependent effects of 4-NH₂ group on proline on stability of collagen model peptides.

The structure of collagen was proposed almost 50 years ago, but a complete understanding of various factors contributing to the stability of collagen triplex is still lacking. A number of experimental evidences pointed out the importance of Hyp in the Y position which allows water mediated hydrogen bonds between the hydroxyl group of Hyp and the peptide carbonyl group of another chain. However, this reasoning does not explain the observed greater thermal stability of 4*R*-Fluoroproline substituted collagen, since C-F has a poor tendency to form hydrogen bonds. Alternative proposal invoked the ring N-C5-C4-X *gauche* effect from (OH/F)- substituent that imposes a C4-*exo* pucker on the pyrrolidine ring to attain the required *trans* conformation for prolyl-peptide bond. This does not explain the incompatibility of 4*R*-substituted proline in X position. The crystal structure data and molecular modelling study on triplex ([Pro-Pro-Gly]₁₀) suggested that proline in X-position prefers to adopt a C4-*endo* (down) pucker, whereas in the Y position prefers C4-*exo* (up) pucker. This indicates that such differential proline puckering in X and Y positions are sterically necessary for a favorable triplex packing arrangement. The pyrrolidine ring is influenced by the nature and stereochemistry of the 4-substituents.

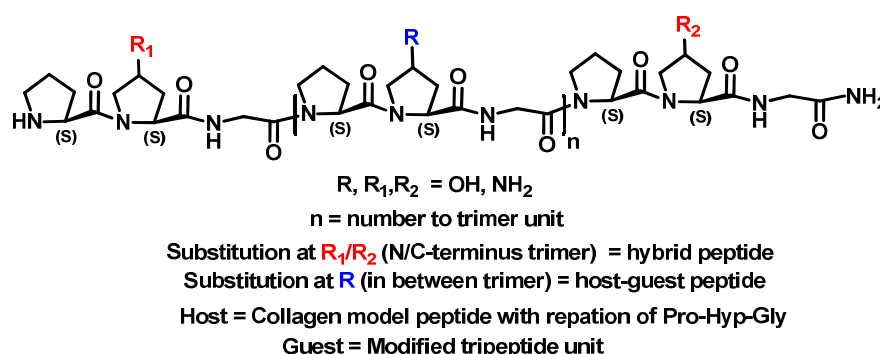


Figure 30: Collagen hybrid and host-guest peptide

In this context it was thought worthwhile to study the effect of replacing the one of the 4-OH by more electronegative group NH₂, which unlike fluorine is also a potent hydrogen bond. The NH₂ group is similar in size to OH and its higher basicity

causes its protonation at physiological pH. The ionizable 4-NH₂ group is a good probe to examine the pH effect on collagen triple helix. The ionizable 4-aminoproline at acidic pH in protonated form (NH₃⁺) may influence the proline puckering *via* its electron withdrawing effects.¹⁴² At basic pH, the NH₂ group has the ability to form hydrogen bonding. It was previously reported from this laboratory that 4*R*-amino substitution on proline in collagen peptide stabilizes the triple helix at both acidic and basic pHs.¹⁴³ In this work the hydroxy group was replaced by amino group for whole sequence with 4*R/S* stereochemistry. In order to understand the molecular origin of these effects, hybrid collagen and host-guest models (Figure 30) are needed to probe the position dependent effects. This would be also focus on the controlled collagen cross linking derived from amino group.

The specific objectives of this chapter are

- Synthesis of N¹-Fmoc-Pro-4*R*-Amp-Gly (**29**), N¹-Fmoc-Pro-4*S*-amp-Gly (**35**) and N¹-Fmoc-4*S*-amp-4*R*-Amp-Gly (**41**) trimers (Figure 31).
- Synthesis of collagen model (control) N¹-Fmoc-Pro-Hyp-Gly (**22**) trimer.
- Synthesis of collagen model peptides with 4*R/S*-amino substitution at different positions (peptides **4-21**) by solid phase synthesis protocol.
- Cleavage of these peptides from the solid support, purification and characterization.
- Investigation of conformation of these peptides using, concentration dependent and temperature dependent CD spectroscopy.
- Evaluate the effect of different pH and solvent on these peptides by using CD spectroscopy.

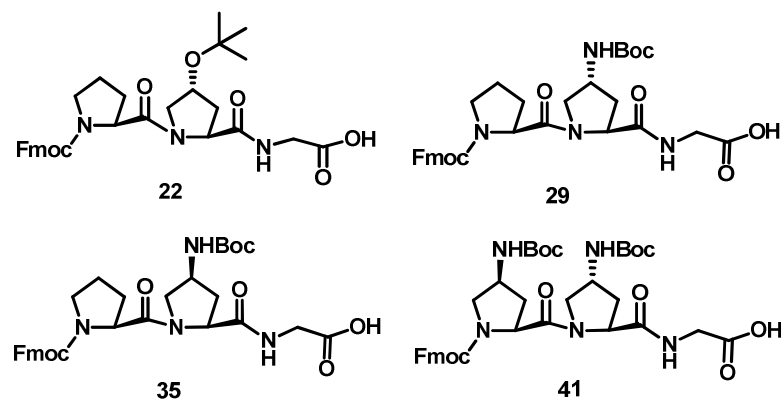


Figure 31: Structures of tripeptides **22**, **29**, **35** and **41**.

3 Results

3.1 Solution *versus* solid phase peptide synthesis – block trimer approach

Solution-phase peptide synthesis is a classical approach to peptide synthesis, as this approach is convenient for synthesis of oligomers in higher scale. Its main advantage is that the intermediate products can be isolated and purified after each step of synthesis, deprotected and recombined to obtain larger peptides of the desired sequence. This procedure is highly flexible with respect to the chemistry of coupling and the combination of the peptidic blocks. New strategies for synthesis in solution have been developed, going from the design of functional groups for the side chains and condensation of fragments for the synthesis of large molecules¹⁴⁴ to the use of new coupling reagents.¹⁴⁵

Various collagen model peptides (CMPs) comprising of typically 18–30 residues have been prepared in order to investigate their structural properties.¹⁴⁶ Since the Pro-Hyp-Gly motif is the most common natural repeat unit within collagen, it is also the most common motif used within CMPs. Thus, the effective synthesis of CMPs with at least six repeat units of Pro-Hyp-Gly is an important goal.

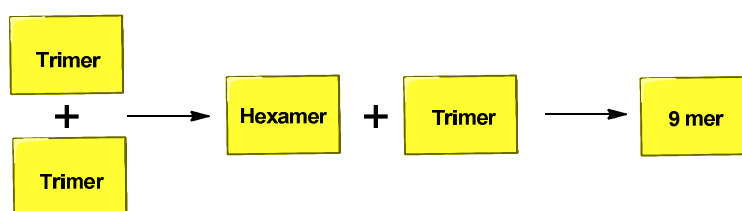


Figure 32: Block coupling approach

Two different strategies for the synthesis of CMPs are used: single Fmoc-protected amino acids are coupled using solid-phase peptide synthesis (SPPS) or trimeric building blocks (Figure 32) such as Fmoc-Pro-Hyp-Gly-OH are prepared in solution phase and then utilized in SPPS.¹⁴⁷ The first method relies on large excesses of the Fmoc-protected amino acid building blocks and the coupling reagents (5–10 equiv) in every coupling step. In this method large amounts of reagent and solvent are used for the synthesis of final oligomer. This approach can lead to the formation of significant amounts of truncated fragments in which the trimeric repeat unit is not preserved. These impurities can disrupt triple helix formation and are often difficult to separate from the desired CMP. The second method requires the preparation of the trimeric building block (e.g., Fmoc-Pro-Hyp-Gly-OH) in solution phase.¹⁴⁸ This approach is

advantageous since fewer coupling steps with less equivalents of the building block are necessary. As a result, fewer side products form, thereby facilitating the purification of the desired CMP.

Solid phase synthesis is a process by which chemical transformations can be carried out on solid support in order to prepare a wide range of synthetic compounds. It offers many advantages over conventional solution phase synthesis in terms of efficiency as well as convenient work-up and purification procedures. In solution phase peptide synthesis, particularly for longer sequences, the repetition of coupling and deprotection cycles can become very labour intensive and require the isolation of all peptide intermediates. Hence to achieve the synthesis of collagen model peptide, combined approach of solution and solid phases are used.

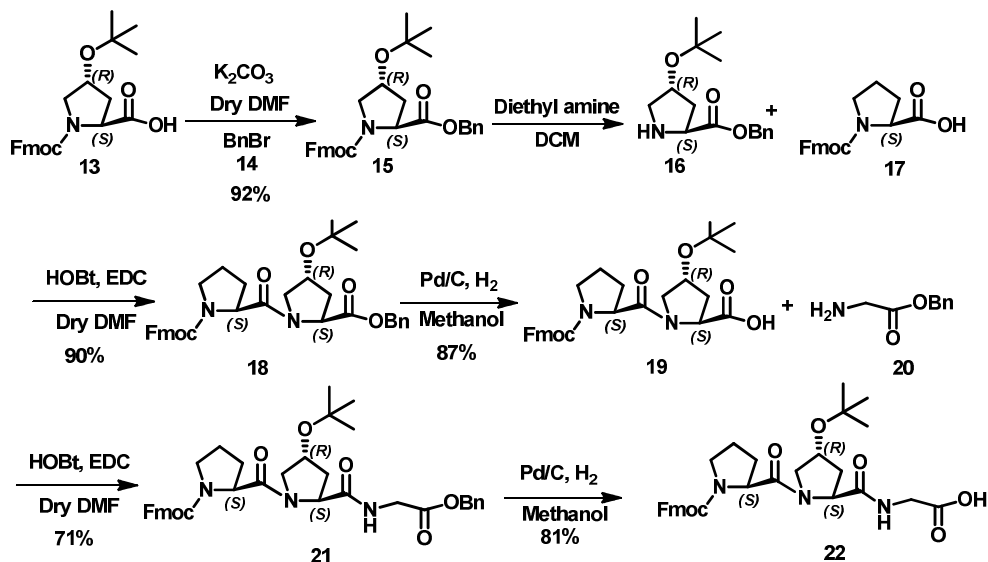
Here the synthesis of all target peptides was achieved by synthesis of tripeptide in solution phase and then utilized it in solid phase synthesis.

3.1.1 Synthesis of collagen model [Pro-Hyp-Gly] tripeptide

The synthesis of collagen model tripeptide (Fmoc-Pro-Hyp-Gly-OH) was achieved in five steps from the commercially available N¹-Fmoc O⁴-*tert* butyl hydroxyproline **13** (Scheme 1). Upon treatment of N¹-Fmoc O⁴-*tert* butyl hydroxyproline with benzyl bromide in K₂CO₃/DMF yielded N¹-Fmoc O⁴-*tert* butyl hydroxyproline benzyl ester (**15**). The deprotection of N¹-Fmoc group in compound **15** was achieved by using diethylamine in presence of DCM to yield compound **16**. The free amine **16** was coupled with N-Fmoc proline **17** in presence of HOBt, EDC in DMF to give the dipeptide **18**. The ¹H NMR of this compound showed aromatic protons ($\delta = 7.30$ to 7.77) for Fmoc and benzyl groups and increase in the number of proline CH₂ protons. The benzyl group in compound **18** was deprotected under mild hydrogenation condition using 5% Pd/C in methanol for 3-4 hours to give compound **19**. The dipeptide free acid **19** was coupled with glycine benzyl ester **20** (Scheme 2) in presence of HOBt, EDC in DMF to yield the tripeptide **21**. This compound showed in ¹H NMR signals for aromatic protons ($\delta = 7.21$ to 7.78) for Fmoc and benzyl groups. The increase in number of glycine CH₂ proton ($\delta = 4.04$) supported the formation of product. The structural integrity of the tripeptide **21** was confirmed by the mass peaks at 654 (M⁺ + 1) in the LCMS. The deprotection of benzyl group in tripeptide **21** was

achieved by hydrogenation using 5% Pd-C in methanol. The mass peak at 563 (M^+) and 586 ($M + Na^+$) in LCMS confirmed the structural integrity of the compound **22**.

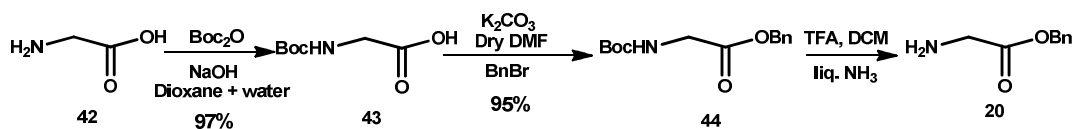
Scheme 1



3.1.2 Synthesis of glycine benzyl ester

The synthesis of glycine benzyl ester **20** was achieved¹⁴⁹ in 3 steps as shown in Scheme 2. Upon treatment of glycine with Boc anhydride in presence of NaOH, dioxane and water yielded Boc glycine **43** that was treated with benzyl bromide in presence of K_2CO_3 /DMF to yield Boc glycine benzyl ester **44**. The deprotection of Boc group was achieved by using trifluoroacetic acid followed by aqueous ammonia workup to yield the compound **20**.

Scheme 2

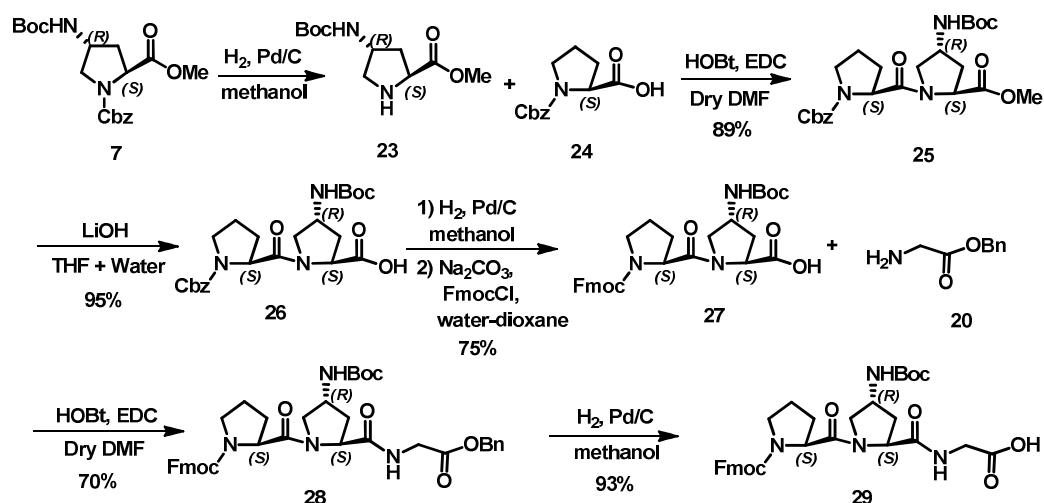


3.1.3 Synthesis of [Pro-4R-Amp-Gly] tripeptide

The synthesis of tripeptide (Fmoc-Pro-4R-Amp-Gly-OH) was achieved in six steps from N-Cbz-4R-Aminoproline methyl ester **7** (Chapter 1, Scheme 1) as shown in Scheme 3. The deprotection of N-Cbz group in compound **7** was done by hydrogenation using 10% Pd-C to yield the compound **23**. This was coupled with N-Cbz proline **24** in presence of HOBT, EDC/DMF to give the dipeptide **25**. The 1H - 1H

COSY (Figure 33) assignment confirmed the structural integrity of the dipeptide **25**. The methyl ester of dipeptide **25** was hydrolysed by using LiOH in THF:water to give the corresponding free acid dipeptide **26**. The deprotection of N-Cbz group in compound **26** was achieved by hydrogenation using 10% Pd/C and converted to N-Fmoc derivative **27** by treatment with Fmoc-Cl/Na₂CO₃. Compound **27** coupled with glycine benzyl ester (**20**) in presence of HOBt, EDC to obtained the tripeptide **28**. The mass peaks at 697 (M⁺ + 1) in the LCMS characterized the tripeptide **28**, supported by ¹H-¹H COSY NMR data. The deprotection of benzyl group in compound **28** was achieved by hydrogenation using 5% Pd-C. The mass peak at 607 (M⁺) and 630 (M + Na⁺) in LCMS and ¹H NMR data confirmed the structure of the compound **29**.

Scheme 3



3.1.3a ¹H-¹H COSY assignment of N-Cbz-Pro-4R-Amp dipeptide 25: The proline- α H (P₂ ring) is easily assignable as the most downfield shifted signal (δ 4.63) and its identification triggers the assignment of all *J*-coupled protons in ring by ¹H-¹H COSY spectral analysis. Figure 33A illustrates the typical 2D NMR assignment of vicinal connectivity by the use of ¹H-¹H COSY has been given in Figure 33B. The peak for Pro- α H (P₂ ring) at δ 4.63 shows two cross peaks (1-2) corresponding to its coupling with neighbours Pro- β H₂, having protons (β H^a δ 2.39 and β H^b 2.09). Both Pro- β H₂ shows cross peaks (3-4) with pro- γ H (γ H δ 4.30). This further connects to two the other cross peaks (5-6) of δ protons of proline (ProH^a δ 3.97 and ProH^b δ 3.79). The pro- γ H (γ H δ 4.30) also shows connectivity (7) with NH proton of *t*-Boc (δ 5.70). This completes the assignment of prolyl ring protons for proline ring 2 (P₂).

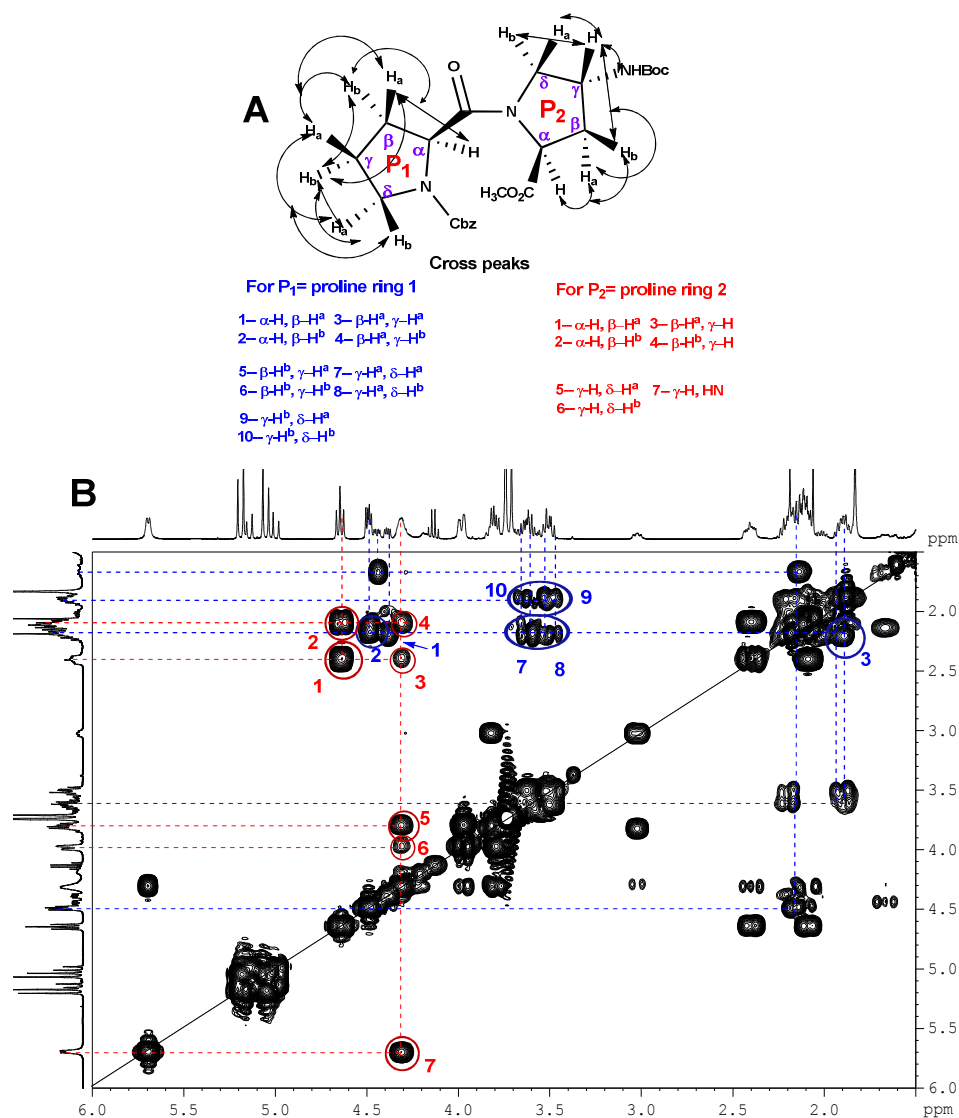


Figure 33: (A) Correlations expected with respect to 2D ^1H - ^1H COSY NMR spectra of dipeptide **25**. (B) 2D ^1H - ^1H COSY NMR spectra of pyrrolidine ring protons in dipeptide **25** (400 MHz)

The assignment of Pro- α H (P₁ ring) which starts from downfield shifted signal (δ 4.44 (*minor*) and δ 4.48 (*major*), arises due to *trans/cis* (75:25) isomerisation of peptide bond. The peak for Pro- α H (P₁ ring) at δ 4.48 (*major*) shows cross peak (2) corresponding to its coupling with neighbour Pro- β H₂, having proton (β H^a δ 2.17). This shows vicinal coupling (3) with γ H^a at δ 1.89 and further connects to two of other cross peaks (7,8) corresponds to δ protons of proline (ProH^a δ 3.52 and ProH^b δ 3.62). Similarly for Pro- α H (P₁ ring) at δ 4.48 (*minor*) shows cross peak (1) with neighbouring Pro- β H₂, having proton (β H^b δ 1.67) which further shows cross peak with its neighbour γ H^b at δ 2.1. This connects to two other cross peaks (9, 10) with δ

protons of proline (ProH^a δ 3.52 and ProH^b δ 3.62), thus completing the assignment of all prolyl ring protons.

3.1.3b ¹H-¹H COSY assignment of N-Cbz-Pro-4R-Amp-Gly-Bn-ester tripeptide 29: The correlations expected with respect to 2D ¹H-¹H COSY NMR spectra of tripeptide **29** are shown in Figure 34A. The ¹H-¹H assignment of glycine NH and α protons is shown in Figure 34B. The NH proton which is easily assignable as the most downfield shifted signal (δ 8.40) shows two cross peaks (1-2) corresponding to its coupling with neighbours α carbon protons (α H^a δ 4.12 and α H^b 3.92). The ¹H-¹H COSY assignment of all prolyl ring protons show similar pattern as in dipeptide **25** is shown in Figure 34C.

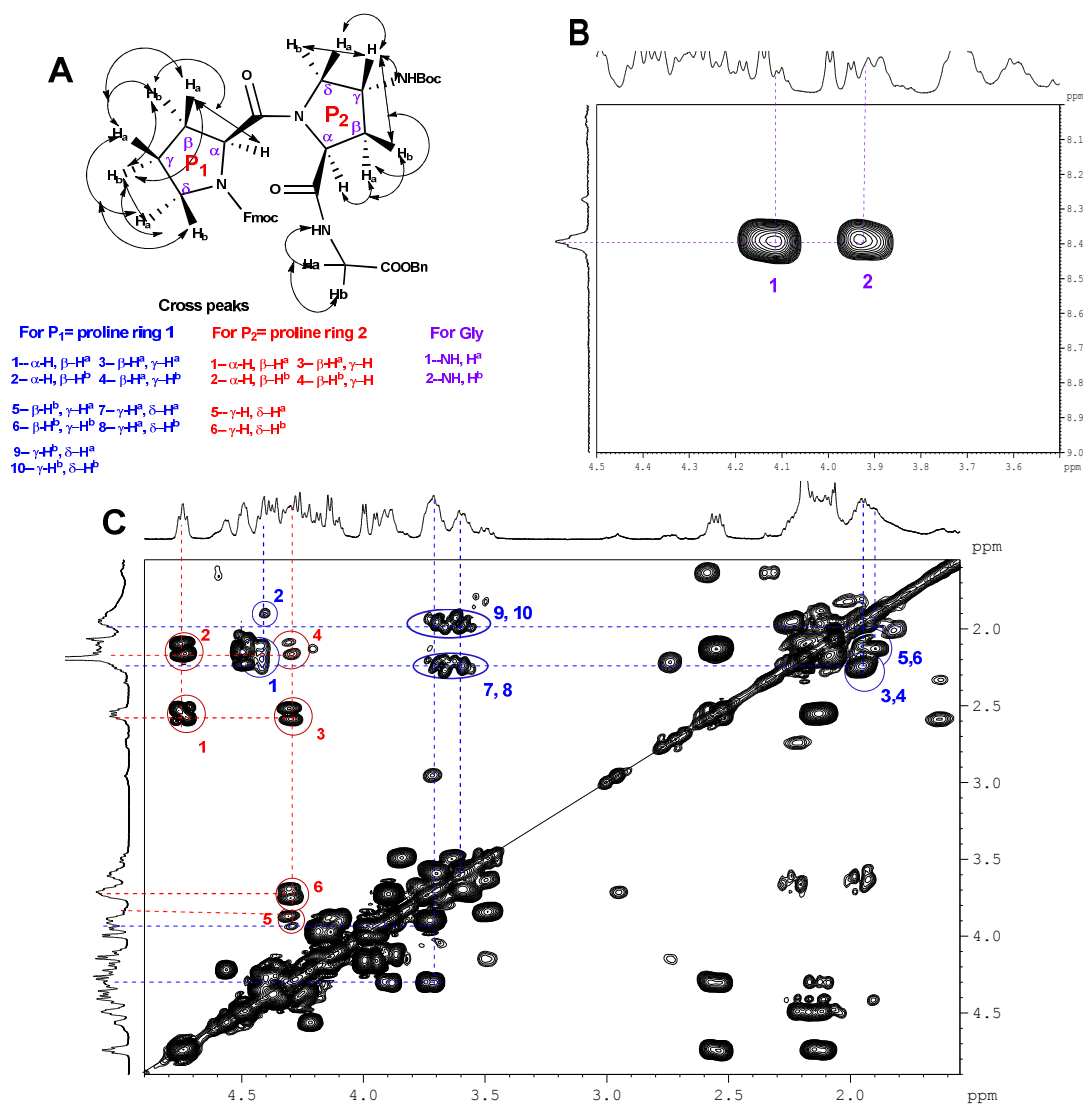
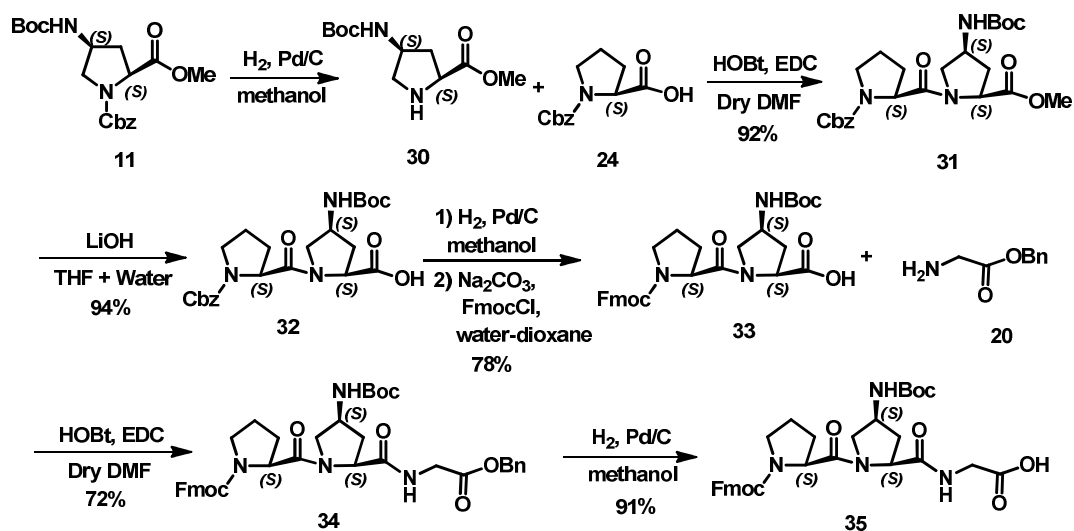


Figure 34: (A) Correlations expected with respect to 2D ¹H-¹H COSY NMR spectra of tripeptide **29**. (B) Gly NH correlation with α -carbon protons (C) 2D ¹H-¹H COSY NMR spectra of pyrrolidine ring protons in tripeptide **29** (400 MHz)

3.1.4 Synthesis of [Pro-4*S*-amp-Gly] tripeptide

The synthesis of tripeptide (Fmoc-Pro-4*S*-amp-Gly-OH) was achieved in six steps from *N*¹-Cbz-4*S*-aminoproline methyl ester **11** (Chapter 1, Scheme 2) as shown in Scheme 4. The deprotection of *N*-Cbz group of compound **11** was achieved by hydrogenation using 10% Pd-C in methanol to yield compound **30**. This was coupled with *N*-Cbz proline **24** in presence of HOBt, EDC to obtain the dipeptide **31**. The ¹H-¹H COSY (Figure 35) assignment characterized structure of the dipeptide **31**. The methyl ester in dipeptide **31** on hydrolysis using LiOH gave the corresponding free acid **32**. The deprotection of *N*-Cbz group in compound **32** was done by hydrogenation using 10% Pd/C and the amine was protected as Fmoc by reaction with Fmoc-Cl to yield the compound **33**. This was coupled with glycine benzyl ester (**20**) in presence of HOBt, EDC to obtain the tripeptide **34**. The mass peaks at 697 (*M*⁺ + 1) in the LCMS and NMR data confirmed the structure of the tripeptide **34**. The deprotection of benzyl ester group was achieved by hydrogenation using 5% Pd-C to obtain the free acid **35**. The mass peak at 607 (*M*⁺), 630 (*M* + Na⁺) in LCMS and NMR data confirmed the assigned structure of the compound **35**.

Scheme 4



3.1.4a ¹H-¹H COSY assignment of N-Cbz-Pro-4*S*-amp dipeptide **31:** The Proline- α H (P₂ ring) which is easily assignable as the most downfield shifted signal (δ 4.58) and its identification triggers the assignment of all *J*-coupled protons in ring by ¹H-¹H COSY spectral analysis. Figure 35A illustrates the typical 2D NMR assignment of vicinal connectivity by the use of ¹H-¹H COSY (Figure 35B).

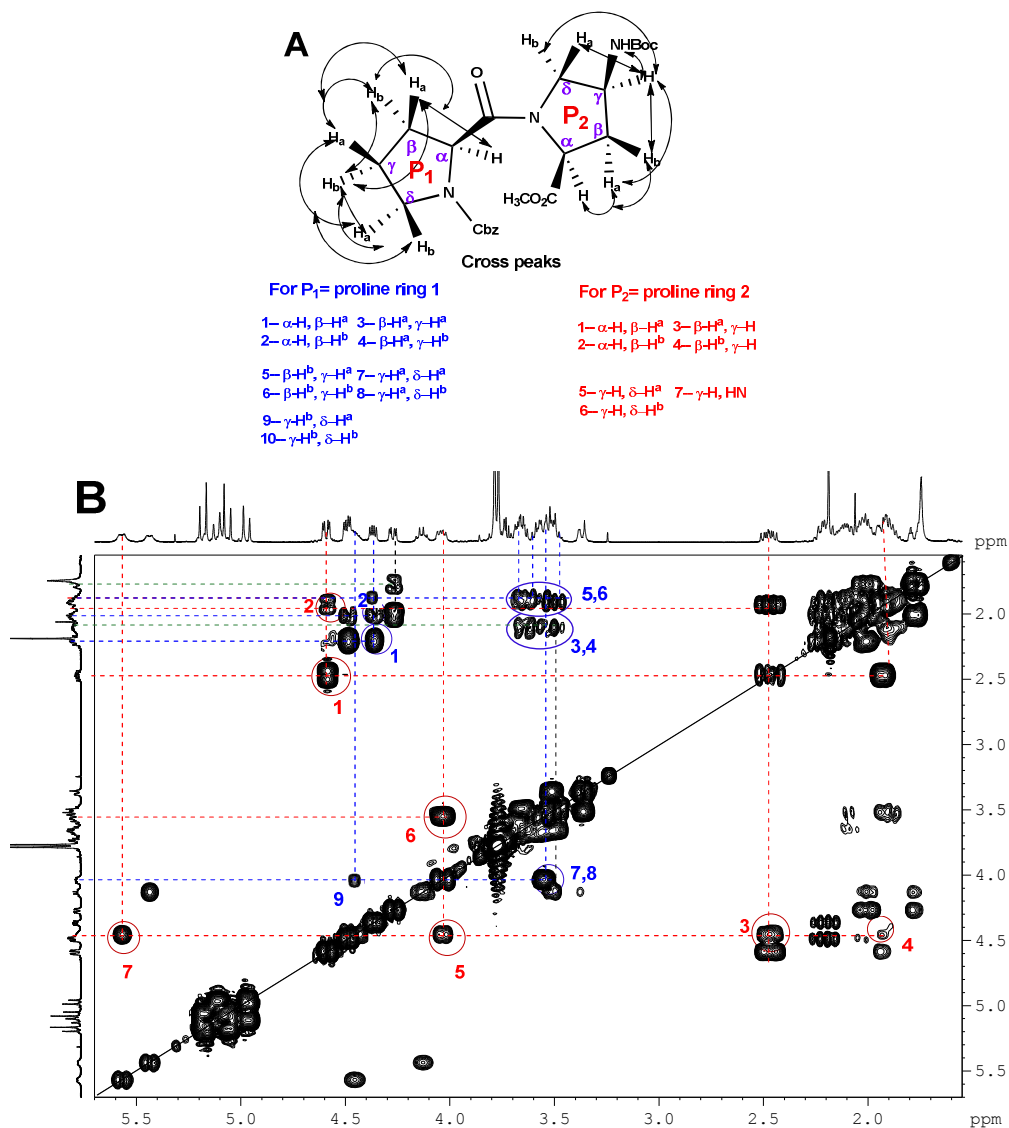


Figure 35: (A) Correlations expected with respect to 2D ^1H - ^1H COSY NMR spectra of dipeptide **31**. (B) 2D ^1H - ^1H COSY NMR spectra of pyrrolidine ring protons in dipeptide **31** (400 MHz)

The Pro- αH (P₂ ring) at δ 4.58 shows two cross peaks (1-2) corresponding to its coupling with neighbours Pro- βH_2 , having protons (βH^a δ 1.92 and βH^b 2.47). Among these, the relative chemical shift values assisted in identifying the upfield components (cross peak 1). The Pro- βH^b shows cross peaks (4) with pro- γH (γH δ 4.44). The pro- γH (δ 4.44) shows cross peak (7) connectivity with NH proton of *t*-Boc (δ 5.58). The pro- γH (δ 4.44) connects to two other cross peaks (5, 6) of δ protons of proline (ProH^a δ 4.04 and ProH^b δ 3.55) that show geminal coupling with each other. This completes the assignment of prolyl ring protons for proline ring 2.

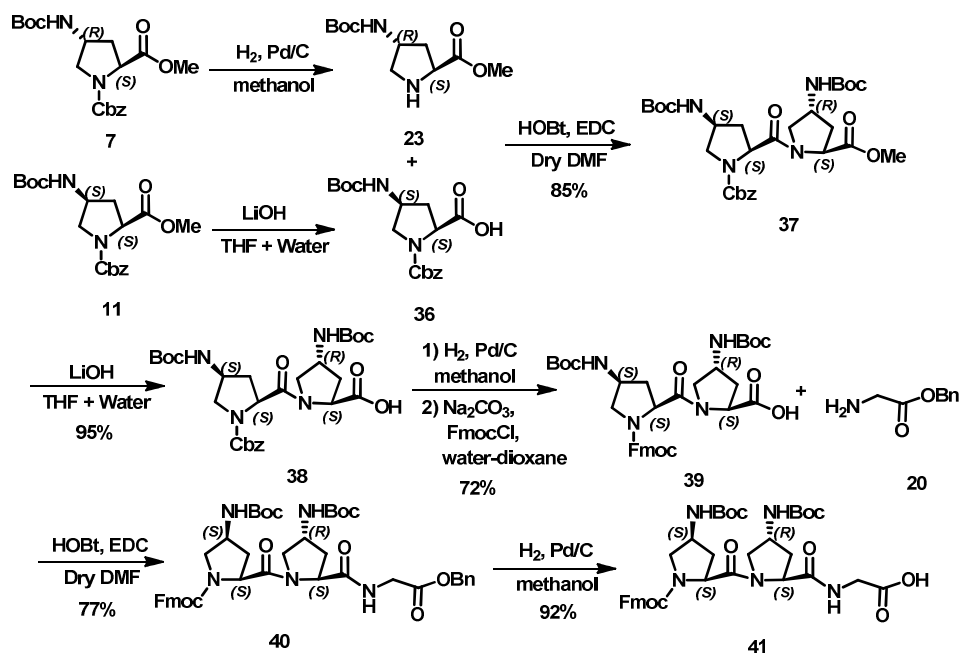
The assignment of Pro- αH (P₁ ring) which starts from downfield shifted signal (δ 4.38 and δ 4.26 arising from *trans/cis* (51:49) isomerisation of peptide bond. The Pro-

α H (P_1 ring) at δ 4.38 shows cross peaks (1, 2) corresponding to its coupling with neighbours Pro- β H₂, having proton (β H^a δ 2.00 and β H^b δ 1.88). Both β protons show geminal coupling with each other. The proton β H^a δ 1.88 show connectivity with neighbouring protons γ H^a at δ 3.44-3.55 and γ H^b at δ 3.60-3.65. The γ proton show geminal coupling, further connect to two peaks of δ protons of proline (ProH^a δ 4.04 and ProH^b δ 4.46). Similarly, Pro- α H (P_1 ring) at δ 4.26 shows cross peak with neighbours Pro- β H₂, having protons (β H^a δ 1.78 and β H^b δ 2.00) that are geminally coupled. The β H^b at δ 2.00 shows connectivity (5,6) with γ H^a at δ 3.54 and γ H^b at δ 3.65 and further connects to two cross peaks of δ protons of proline (ProH^a δ 4.04 and ProH^b δ 4.46). This completes the assignment of all prolyl ring protons.

3.1.5 Synthesis of [4*S*-amp-4*R*-Amp-Gly] Tripeptide

The synthesis of tripeptide (Fmoc-4*S*-amp-4*R*-Amp-Gly-OH) was achieved in seven steps from N¹-Cbz-N⁴ *R*-Aminoproline methyl ester **7** (Chapter 1, Scheme 1) and N¹-Cbz-N⁴ *S*-Aminoproline methyl ester **11** (Chapter 1, Scheme 2) and is shown in Scheme 5.

Scheme 5



The deprotection of N-Cbz group of compound **7** by hydrogenation using 10% Pd-C in methanol gave **23**. The methyl ester in compound **11** was hydrolysed by using LiOH to the free acid **36**. The amine compound **23** was coupled with acid **36** in presence of HOBt, EDC to obtain dipeptide **37**. The appearance of 18 protons for *t*-

Boc and the ^1H - ^1H COSY (Figure 36) assignment confirmed the structure of the dipeptide **37** methyl ester which was hydrolysed with LiOH to the free acid **38** that was hydrogenated using 10% Pd/C to the amine. This was reacted with Fmoc-Cl to yield N-Fmoc dipeptide free acid **39**. This was coupled with glycine benzyl ester **20** in presence of HOBt, EDC to obtain the tripeptide **40**. The mass peaks at 812 ($M^+ + 1$) in the LCMS and NMR data confirmed the assigned structure of the tripeptide **40**. The deprotection of benzyl group in **40** was achieved by hydrogenation using 5% Pd/C to result in compound **41**. The mass peaks at 722 (M^+), 745 ($M + \text{Na}^+$) in LCMS and NMR data confirmed the structure of the compound **41**.

3.1.5a ^1H - ^1H cosy assignment of N-Cbz-4S-amp-4R-Amp dipeptide 37: The proline- αH (P_2 ring) is easily assignable as the most downfield shifted signal (δ 4.65) and its identification generates the assignment of all J -coupled protons in ring by ^1H - ^1H COSY spectral analysis. Figure 36A illustrates the typical 2D NMR assignment of vicinal connectivity by the use of ^1H - ^1H COSY and given in Figure 36B. The peak for Pro- αH (P_2 ring) at δ 4.65 shows two cross peaks (1-2) corresponding to its coupling with neighbours Pro- βH_2 , having protons (βH^a δ 2.48 and βH^b 2.08). Both the Pro- βH_2 show cross peaks (3,4) with pro- γH (γH δ 4.29) which further connects to two other cross peaks (5,6) of δ protons of proline (ProH^a δ 4.14 and ProH^b δ 3.78). The pro- γH (γH δ 4.30) also shows connectivity (7) with NH proton of *t*-Boc (δ 5.76). This completes the assignment of prolyl ring protons for proline ring 2.

Similarly proline αH (P_1 ring) is assignable at δ 4.55 and helps further assignment. The Pro- αH (P_1 ring) at δ 4.55 shows two cross peaks (1-2) corresponds to Pro- βH_2 , having protons (βH^a δ 2.42 and βH^b 2.02). Both the Pro- βH_2 show cross peaks (3,4) with pro- γH (γH δ 4.42). This further connects to two other cross peaks (5,6) of δ protons of proline (ProH^a δ 3.72 and ProH^b δ 3.59). The pro- γH (γH δ 4.42) shows connectivity (7) with NH proton of *t*-Boc (δ 6.21, 6.40). This may arise due to *cis-trans* isomerization of peptide bond. This completes the assignment of prolyl ring protons for proline ring 1.

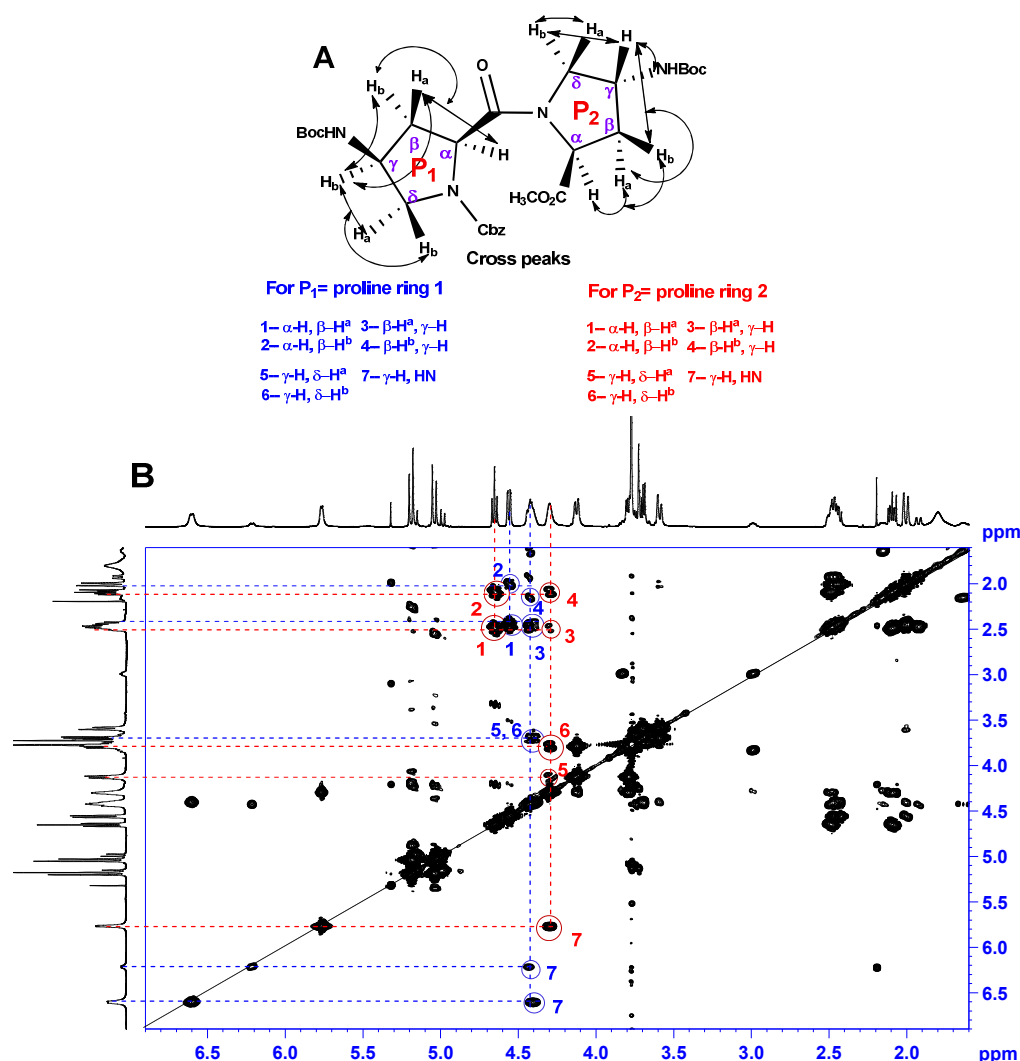


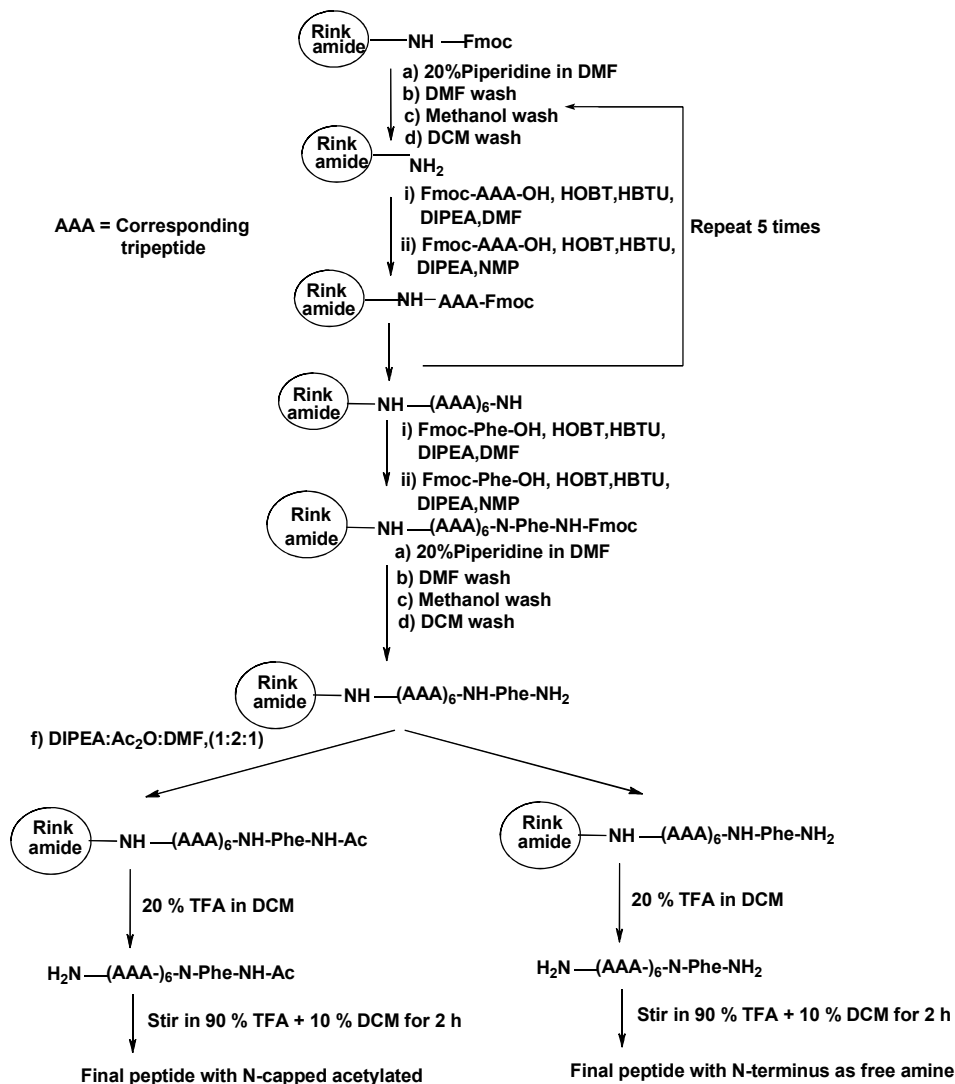
Figure 36: (A) Correlations expected with respect to 2D ¹H-¹H COSY NMR spectra of dipeptide 37. (B) 2D ¹H-¹H COSY NMR spectra of pyrrolidine ring protons in dipeptide 37 (400 MHz)

3.1.6 Solid phase peptide synthesis

The target peptides (4-21) were synthesized by manual solid phase synthesis on readily available rink amide resin using standard Fmoc chemistry (Scheme 6), followed by cleavage to yield the peptides as C-terminal amides. The commercially available Kieselghur supported N,N-dimethylacrylamide resin (with Rink-amide linker) with Fmoc group was cleaved by 20% piperidine in DMF. The trimer peptides as free acids were coupled using *in situ* activation procedure using 3 eq. of tripeptide, HBTU and HOBt/DIPEA in DMF. The coupling reaction was repeated second time using *N*-Methyl-2-pyrrolidone (NMP) as solvent. The N-Fmoc deprotection reactions were done by using 20% piperidine in DCM. The deprotection and coupling reactions

were monitored by using qualitative Chloranil test¹⁵⁰ for iminoacids and Ninhydrin (Kaiser)¹⁵¹ test for glycine.

Scheme 6: Schematic representation of SPPS for 4-21 peptide synthesis

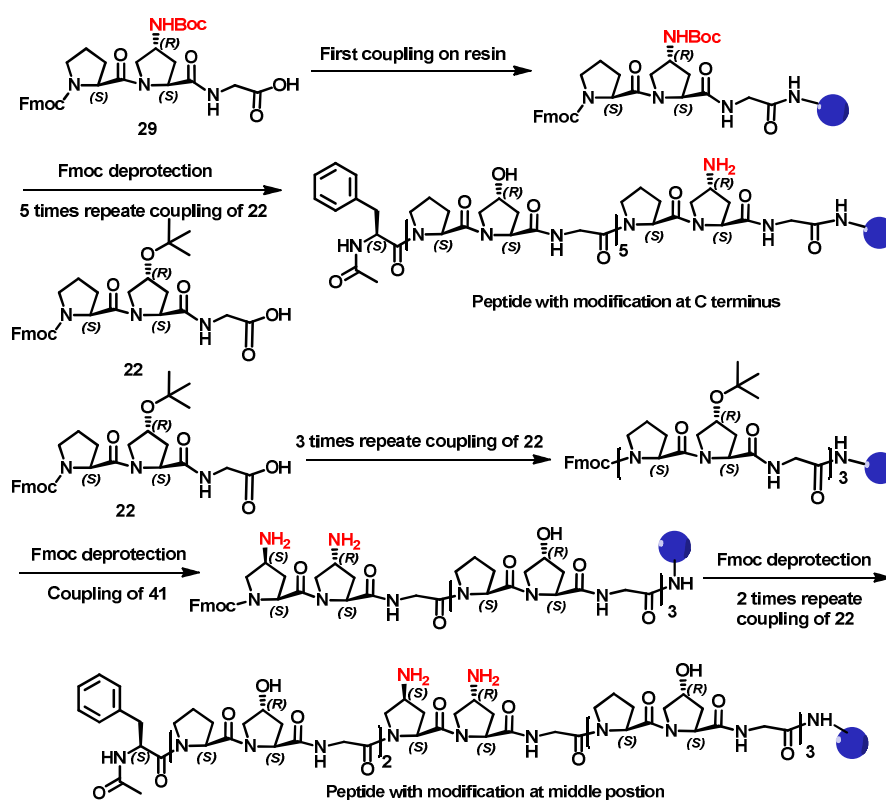


After completion of synthesis, the resin was divided into two equal parts. The terminal amino group in one portion of the resin was capped with Ac₂O. The peptide was cleaved from the resin using 20% TFA in DCM. As the *t*-Boc and *t*-Butyl groups removal requires stronger acidic conditions, the deprotection of side chain *t*-Boc on Amp and *t*-Butyl on hydroxyproline residues of the peptide was carried out with 90% TFA in DCM. As *t*-Butyl cation formed during the deprotection of *t*-Boc from the final peptide can lead to N-alkylation of the amines. To prevent such side reactions during peptide-cleavage and *t*-Boc deprotection, 0.1% TIS (triisopropylsilane) was used as a scavenger. The peptides were purified on semi-preparative RP-18 HPLC column using

water-acetonitrile gradient to which 0.1% TFA was added. Peptides **4-12** were blocked as N-terminal acetylated and C-terminal aminated to remove the electrostatic interactions. Peptide **13-21** kept as N-terminal free amine to evaluate the effect of charge on triplex forming abilities.

The hybrid and host-guest peptides were synthesized to probe the position dependent effects. The trimers were used for synthesis of oligomers on Rink amide resin. The peptides were synthesized with various combinations of the tripeptide. Scheme 7 shows representative synthesis of oligomer by using tripeptide.

Scheme 7: Representative synthesis of peptides 4-21



The peptides **4-21** were synthesized from the appropriate N-Fmoc-protected Pro-Hyp-Gly (**22**), Pro-4*R*-Amp-Gly (**29**), Pro-4*S*-amp-Gly (**35**) and 4*S*-amp-4*R*-Amp-Gly (**41**) assembled in the solid phase. The amino acid phenylalanine was included at the N-terminus to enable accurate determination of peptide concentrations by UV absorbance at 259 nm ($\epsilon = 200 \text{ M}^{-1} \text{ cm}^{-1}$). All the peptides (**4-21**) were purified by HPLC and characterized by MALDI-TOF. The purity of the peptides was determined using analytical RP-18 HPLC and the purity of peptides was found more than 98%.

The structural integrity of the peptides was further confirmed by MALDI-TOF mass spectrometry which agreed closely with the calculated values (Table 5-6).

Table 5: N-capped acetylated peptides (4-12)

Peptides	Mass (cal)	Mass (obs)
Ac-Phe-[Pro-Hyp-Gly] ₅ -Pro-4R-Amp-Gly-NH ₂ (4)	1808	1831.33 [M + Na] ⁺ 1847.32 [M + K] ⁺
Ac-Phe-[Pro-Hyp-Gly] ₅ -Pro-4S-amp-Gly-NH ₂ (5)	1808	1832.27 [M + Na] ⁺ 1848.24 [M + K] ⁺
Ac-Phe-[Pro-Hyp-Gly] ₅ -4S-amp-4R-Amp-Gly-NH ₂ (6)	1823	1846.29 [M + Na] ⁺ 1862.26 [M + K] ⁺
Ac-Phe-[Pro-Hyp-Gly] ₂ -Pro-4R-Amp-Gly-[Pro-Hyp-Gly] ₃ -NH ₂ (7)	1808	1832.27 [M + Na] ⁺ 1848.24 [M + K] ⁺
Ac-Phe-[Pro-Hyp-Gly] ₂ -Pro-4S-amp-Gly-[Pro-Hyp-Gly] ₃ -NH ₂ (8)	1808	1832.25 [M + Na] ⁺ 1848.21 [M + K] ⁺
Ac-Phe-[Pro-Hyp-Gly] ₂ -4S-amp-4R-Amp-Gly-[Pro-Hyp-Gly] ₃ -NH ₂ (9)	1823	1824.43 [M+1] ⁺ 1846.32 [M + Na] ⁺ 1862.30 [M + K] ⁺
Ac-Phe-Pro-4R-Amp-Gly-[Pro-Hyp-Gly] ₅ -NH ₂ (10)	1808	1832.33 [M + Na] ⁺ 1848.38 [M + K] ⁺
Ac-Phe-Pro-4S-amp-Gly-[Pro-Hyp-Gly] ₅ -NH ₂ (11)	1808	1830.41 [M + Na] ⁺ 1848.35 [M + K] ⁺
Ac-Phe-4S-amp-4R-Amp-Gly-[Pro-Hyp-Gly] ₅ -NH ₂ (12)	1823	1824.35 [M] ⁺ 1847.34 [M + Na] ⁺ 1863.32 [M + K] ⁺

Table 6: Peptides (13-21) free amine at N terminus

Peptides	Mass (cal)	Mass (obs)
H ₂ N-Phe-[Pro-Hyp-Gly] ₅ -Pro-4 <i>R</i> -Amp-Gly-NH ₂ (13)	1767	1767.35 [M] ⁺ 1789.33 [M + Na] ⁺ 1805.30 [M + K] ⁺
H ₂ N-Phe-[Pro-Hyp-Gly] ₅ -Pro-4 <i>S</i> -amp-Gly-NH ₂ (14)	1767	1767.31 [M] ⁺ 1789.31 [M + Na] ⁺ 1805.28 [M + K] ⁺
H ₂ N-Phe-[Pro-Hyp-Gly] ₅ -4 <i>S</i> -amp-4 <i>R</i> -Amp-Gly-NH ₂ (15)	1781	1782.39 [M] ⁺ 1804.38 [M + Na] ⁺ 1820.38 [M + K] ⁺
H ₂ N-Phe-[Pro-Hyp-Gly] ₂ -Pro-4 <i>R</i> -Amp-Gly-[Pro-Hyp-Gly] ₃ -NH ₂ (16)	1767	1768.39 [M] ⁺ 1789.42 [M + Na] ⁺ 1806.40 [M + K] ⁺
H ₂ N-Phe-[Pro-Hyp-Gly] ₂ -Pro-4 <i>S</i> -amp-Gly-[Pro-Hyp-Gly] ₃ -NH ₂ (17)	1767	1768.31 [M] ⁺ 1790.28 [M + Na] ⁺ 1806.26 [M + K] ⁺
H ₂ N-Phe-[Pro-Hyp-Gly] ₂ -4 <i>S</i> -amp-4 <i>R</i> -Amp-Gly-[Pro-Hyp-Gly] ₃ -NH ₂ (18)	1781	1783.34 [M] ⁺ 1805.37 [M + Na] ⁺ 1820.26 [M + K] ⁺
H ₂ N-Phe-Pro-4 <i>R</i> -Amp-Gly-[Pro-Hyp-Gly] ₅ -NH ₂ (19)	1767	1768.30 [M] ⁺ 1790.29 [M + Na] ⁺ 1806.26 [M + K] ⁺
H ₂ N-Phe-Pro-4 <i>S</i> -amp-Gly-[Pro-Hyp-Gly] ₅ -NH ₂ (20)	1767	1768.32 [M] ⁺ 1790.29 [M + Na] ⁺ 1806.23 [M + K] ⁺
H ₂ N-Phe-4 <i>S</i> -amp-4 <i>R</i> -Amp-Gly-[Pro-Hyp-Gly] ₅ -NH ₂ (21)	1781	1783.32 [M] ⁺ 1805.39 [M + Na] ⁺ 1820.21 [M + K] ⁺

3.2 Characterization of Collagen Triple-Helical Structure

In solution, collagen like triple-helix and polyproline II like helices exhibit fingerprint CD spectra. The spectra are characterised by the presence of a large negative band around 200 nm, a crossover at around 213 nm and a small positive band around 215-227 nm. The presence of triple α or single helix can be determined by concentration dependent CD spectroscopy. The percentage of triple-helical structure is close to a maximum when the concentration is greater than its critical triple-helical concentration. The magnitude of the ratio of positive to negative band intensity in the CD spectra of triple helical peptides ($R_{p/n}$) has been proposed to quantitate the triple

helical strength. The sharp co-operative transition state has been observed for temperature dependent CD spectroscopy.

In order to investigate the conformational behavior of peptides 4-21, CD spectra were measured under various pH values and solvent conditions.

3.3 CD Spectroscopic Studies of N- and C-Capped Peptides 4-12

Peptides 4-12 were blocked as N-terminal acetylated and C-terminal aminated to remove the electrostatic interactions.

3.3.1 CD Spectroscopic Studies of C-Terminus Modified Peptides 4-6

Ac-Phe-[Pro-Hyp-Gly]₅-Pro-4*R*-Amp-Gly-NH₂ (Peptide 4)

Ac-Phe-[Pro-Hyp-Gly]₅-Pro-4*S*-amp-Gly-NH₂ (Peptide 5)

Ac-Phe-[Pro-Hyp-Gly]₅-4*S*-amp-4*R*-Amp-Gly-NH₂ (Peptide 6)

3.3.1a Concentration dependent CD spectroscopy of peptides 4-6: Figure 37 shows the CD spectra of peptides 4-6 taken at 25 °C in the concentration range of 50 μM – 500 μM at pH 7.2.

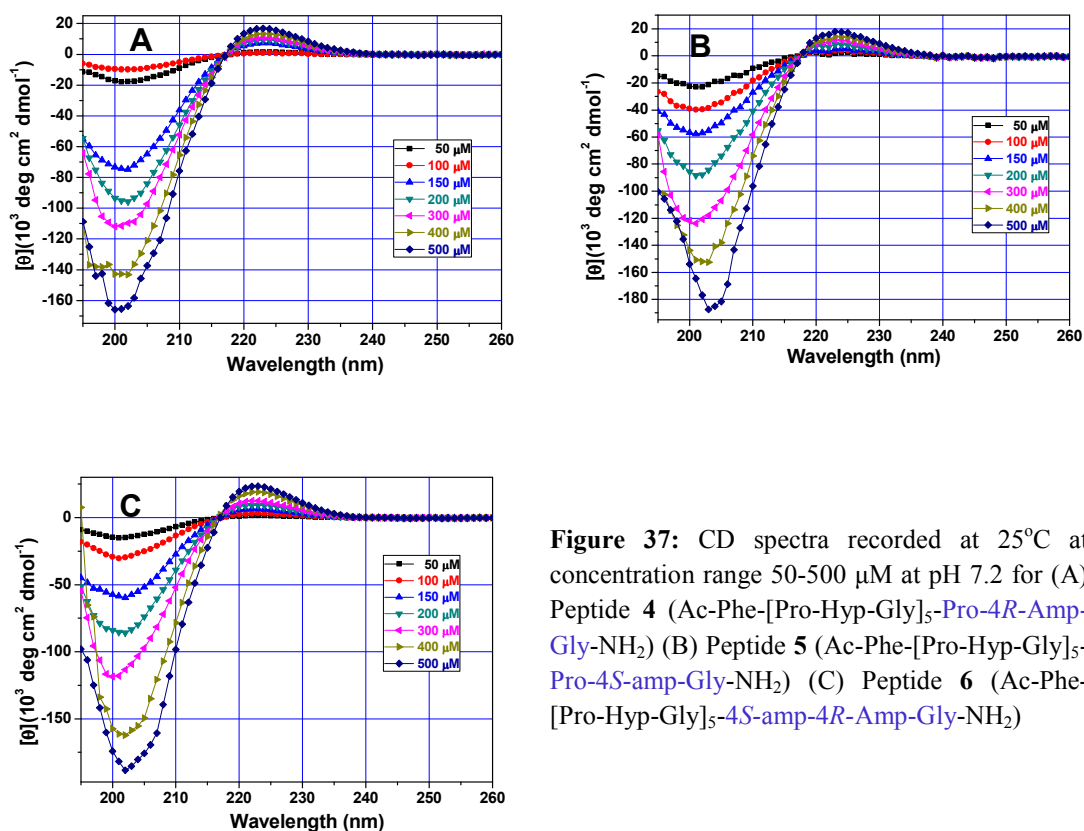


Figure 37: CD spectra recorded at 25°C at concentration range 50-500 μM at pH 7.2 for (A) Peptide 4 (Ac-Phe-[Pro-Hyp-Gly]₅-Pro-4*R*-Amp-Gly-NH₂) (B) Peptide 5 (Ac-Phe-[Pro-Hyp-Gly]₅-Pro-4*S*-amp-Gly-NH₂) (C) Peptide 6 (Ac-Phe-[Pro-Hyp-Gly]₅-4*S*-amp-4*R*-Amp-Gly-NH₂)

The peptides **4-6** show similar positive and negative maxima between 222-224 nm and 201-203 nm respectively. Importantly, all the spectral traces pass through an isobestic point at 217-218 nm.

Figure 38 shows plots of $R_{p/n}$ values derived from CD spectra against concentration of peptides **4-6** at pH 7.2. The $R_{p/n}$ values increase rapidly and reach saturation at 200 μM and remain constant thereafter. A critical triple-helical concentration of $\sim 200 \mu\text{M}$ was derived from these plots. Hence, all further studies were performed at a concentration of 0.2 mM where peptides **4-6** are associated in triple helical form.

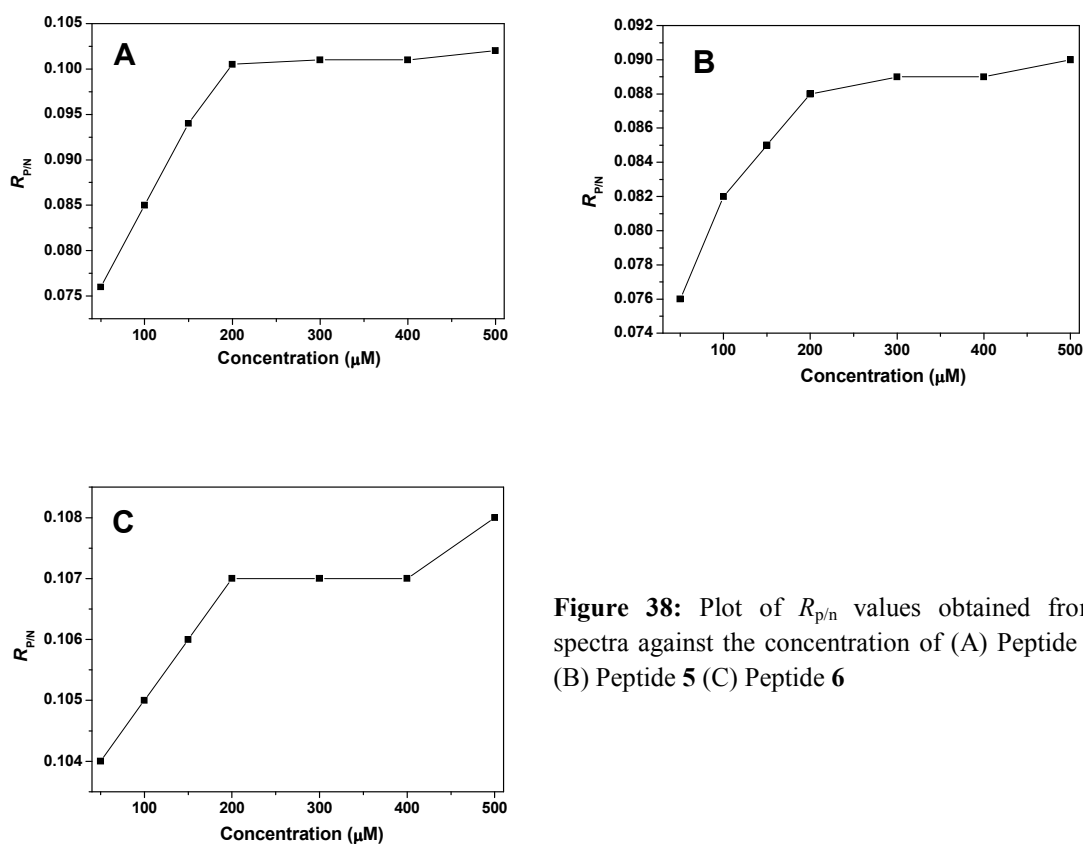


Figure 38: Plot of $R_{p/n}$ values obtained from spectra against the concentration of (A) Peptide **4** (B) Peptide **5** (C) Peptide **6**

2.3.1b CD thermal denaturation study of peptides 4-6 at different pHs: It was reported earlier from this laboratory that T_m of collagen oligomer AcPhe[Pro-Hyp-Gly]₆ **22** to be 27 °C and invariant with change in pH. The T_m 4*R*-aminoproline peptide AcPhe[Pro-4*R*-Amp-Gly]₆ **23** was higher and it changed with pH. In contrast, the 4*S*-aminoproline peptide AcPhe[Pro-4*S*-amp-Gly]₆ **24** showed no such behaviour indicating that it does not form a triplex. The T_m of the peptide AcPhe[4*S*-amp-4*R*-

Amp-Gly]₆ **25** was enhanced and shows pH dependent stability effect. The T_m values for these peptides are shown in Table 7.¹⁴³

Table 7: T_m of peptide **22-25** in °C at different pHs.

Peptide	pH/solvent				
	4.0	7.2	10.0	12.0	EG
AcPhe[Pro-Hyp-Gly] ₆ (peptide 22)	27	28	27	27	44
AcPhe[Pro-4 <i>R</i> -Amp-Gly] ₆ (peptide 23)	60	56.5	26	49	23
AcPhe[Pro-4 <i>S</i> -amp-Gly] ₆ (peptide 24)	nd	nd	nd	nd	nd
AcPhe[4 <i>S</i> -amp-4 <i>R</i> -Amp-Gly] ₆ (peptide 25)	61	46.6	40.5	34	37

nd = triplex not detected, T_m values are (± 0.5 °C)

In order to get information about the relative triple-helical strength of peptides **4-6**, CD-spectroscopic studies as a function of temperature was carried out by observing the changes in ellipticity at 222-225 nm with temperature. Upon heating a decrease in molar ellipticity with increase in temperature was observed for all the peptides.

Peptide 4 (Ac-Phe-[Pro-Hyp-Gly]₅-Pro-4*R*-Amp-Gly-NH₂): Figure 39A shows the CD-thermal denaturation plot of molar ellipticity at 223-224 nm *versus* temperature at different pHs for peptide **4**. Figure 39B shows the first derivative curve obtained from the sigmoidal fit of data of Figure 39A and the T_m values for peptide **4** obtained from minima of Figure 39B.

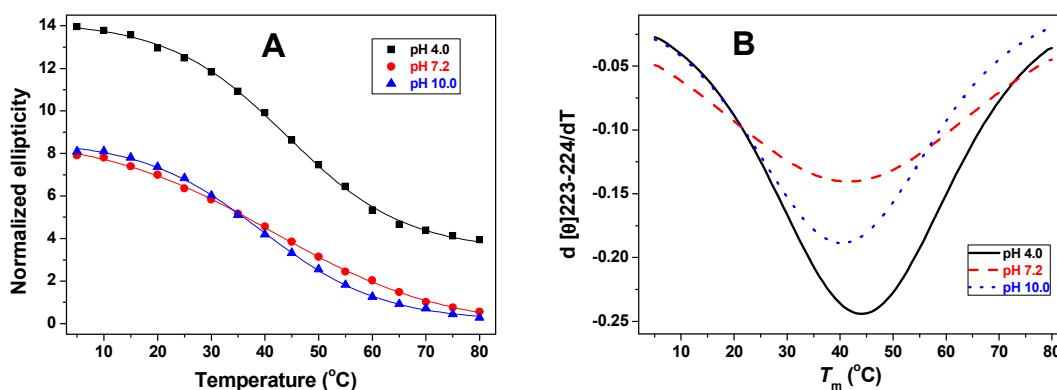


Figure 39: (A) CD-thermal denaturation plot of molar ellipticity at 223-224 nm vs. temperature at different pHs for peptide **4** (Ac-Phe-[Pro-Hyp-Gly]₅-Pro-4*R*-Amp-Gly-NH₂). (B) First derivative curve of peptide **4** as a function of pH.

At all pHs, the peptide **4** shows cooperative melting transition with increasing temperature. This indicates that peptide **4** is associated in triple-helical structure at different pHs. Table 8 shows T_m obtained for peptide **4** at different pH and ΔT_m compared with control collagen peptide **22** and fully modified peptide **23**. The T_m

values indicate that at pH 4.0, peptide **4** forms a most stable triple-helical structure with T_m of 46 °C. The T_m of peptide **4** decreases to 41 °C as the pH increases to 7.2 and 40 °C at pH 10.0.

Table 8: T_m of peptide **4** (Ac-Phe-[Pro-Hyp-Gly]₅-Pro-4*R*-Amp-Gly-NH₂)

pH	T_m	ΔT_m (4-22)	ΔT_m (4-23)
4.0	46	+17	-14
7.2	41	+13	-15.5
10.0	40	+13	+14

T_m values are (± 0.5 °C)

Peptide 5 (Ac-Phe-[Pro-Hyp-Gly]₅-Pro-4*S*-amp-Gly-NH₂): Figure 40A shows the CD-thermal denaturation plot of molar ellipticity at 222-224 nm *versus* temperature at different pHs for peptide **5**. Figure 40B shows the first derivative curve obtained from the sigmoidal fit of data of Figure 40A and the T_m values for peptide **5** obtained from minima of Figure 40B at different pHs is shown in Table 9. The T_m of peptide **5** increases with increase in pH.

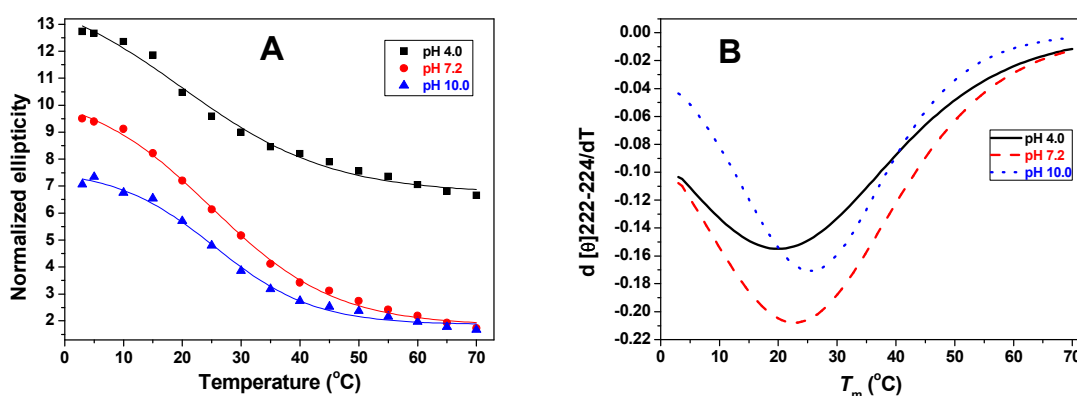


Figure 40: (A) CD-thermal denaturation plot of molar ellipticity at 222-224 nm vs. temperature at different pHs for peptide **5** (Ac-Phe-[Pro-Hyp-Gly]₅-Pro-4*S*-amp-Gly-NH₂). (B) First derivative curve of peptide **5** as a function of pH.

Table 9: T_m of peptide **5** (Ac-Phe-[Pro-Hyp-Gly]₅-Pro-4*S*-amp-Gly-NH₂)

pH	T_m	ΔT_m (4-22)	ΔT_m (4-24)
4.0	19	-8	+19
7.2	23	-4	+23
10.0	25	-2	+25

T_m values are (± 0.5 °C)

Peptide 6 (Ac-Phe-[Pro-Hyp-Gly]₅-4*S*-amp-4*R*-Amp-Gly-NH₂): Figure 41A shows the CD-thermal denaturation plot of molar ellipticity at 222-224 nm *versus* temperature at different pHs for peptide 6. Figure 41B shows the first derivative curve obtained from the sigmoidal fit of data of Figure 41A and the T_m values for peptide 6 obtained from minima of Figure 41B at different pHs is shown in Table 10.

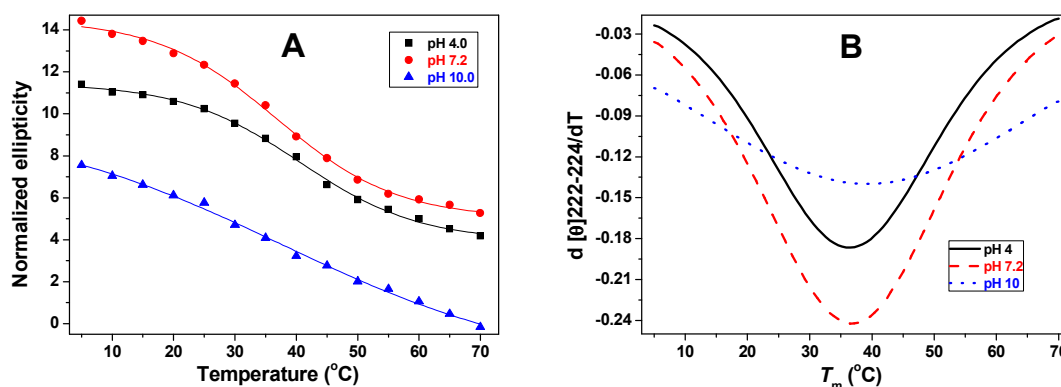


Figure 41: (A) CD-thermal denaturation plot of molar ellipticity at 222-224 nm *vs.* temperature at different pHs for peptide 6 (Ac-Phe-[Pro-Hyp-Gly]₅-4*S*-amp-4*R*-Amp-Gly-NH₂). (B) First derivative curve of peptide 6 as a function of pH.

Table 10: T_m of peptide 6 (Ac-Phe-[Pro-Hyp-Gly]₅-4*S*-amp-4*R*-Amp-Gly-NH₂)

pH	T_m	ΔT_m (4-22)	ΔT_m (4-25)
4.0	41	+14	-20
7.2	37	+9	-10.5
10.0	39	+12	-1

T_m values are (± 0.5 °C)

3.3.1c CD thermal denaturation study of peptides 4-6 in ethylene glycol: Ethylene glycol stabilizes collagen triple helices and is useful for detecting weak triple helical propensities¹⁵² and hence CD thermal denaturation studies for peptides 4-6 were carried out in ethylene glycol. Figure 42A shows the CD-thermal denaturation plot of molar ellipticity at 222-224 nm *versus* temperature in ethylene glycol for peptides 4-6. Figure 42B shows the first derivative curve obtained from the sigmoidal fit of data of Figure 42A and the T_m values obtained from minima of Figure 42B of peptides 4-6 is shown in Table 11.

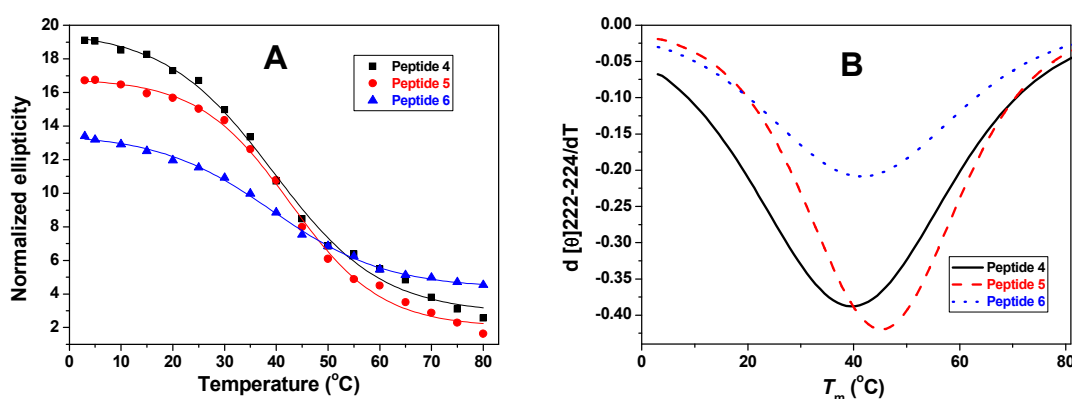


Figure 42: (A) CD-thermal denaturation plot of molar ellipticity at 222-224 nm vs. temperature in ethylene glycol for peptides 4-6. (B) First derivative curve of peptides 4-6 in ethylene glycol.

Table 11: T_m of peptides 4-6 in ethylene glycol

Peptide	T_m
4	41
5	43
6	40

T_m values are (± 0.5 °C)

3.3.1d Comparison of stability of peptides 4-6 with peptides 22-25: The comparative ΔT_m values for peptides 4-6 with corresponding control peptides 22-25 is represented as a bar diagram in Figure 43.

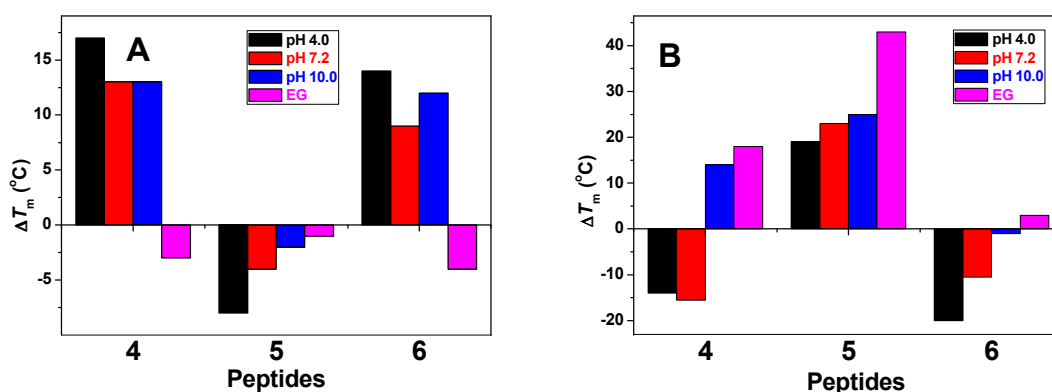


Figure 43: (A) ΔT_m of peptides 4-6 compared with control collagen peptide 22 (AcPhe[Pro-Hyp-Gly]₆) (B) ΔT_m of peptides 4-6 compared with corresponding fully modified peptides 23-25.

The following features are noticed from the data.

- The increase in T_m for single 4R-amino prolyl modification (peptide 4) compared with control collagen peptide 22 at all pHs is about 50% of that with fully modified peptide 23.

(ii) The decrease in T_m for single 4*S*-amino prolyl modification (peptide **5**) is 3 °C to 8 °C as compared with control collagen peptide **22**, while triplex was not detected for fully 4*S*-amino prolyl modified peptide **24** at all pHs.

(iii) For the chimeric peptide **6**, the increase in T_m at acidic and neutral pHs compared with control collagen peptide **22** is about half as compared with that of fully modified peptide **25**. At basic pH, the T_m for peptide **6** is nearly equal to T_m of peptide **25**.

3.3.2 CD Spectroscopic Studies of Middle Modified Peptides 7-9

Ac-Phe-[Pro-Hyp-Gly]₂-Pro-4*R*-Amp-Gly-[Pro-Hyp-Gly]₃-NH₂ (Peptide **7**)

Ac-Phe-[Pro-Hyp-Gly]₂-Pro-4*S*-amp-Gly-[Pro-Hyp-Gly]₃-NH₂ (Peptide **8**)

Ac-Phe-[Pro-Hyp-Gly]₂-4*S*-amp-4*R*-Amp-Gly-[Pro-Hyp-Gly]₃-NH₂ (Peptide **9**)

3.3.2a Concentration dependent CD spectra of peptides 7-9: Figure 44 show the CD spectra with varying concentrations (50 μ M –500 μ M) of peptides **7-9**.

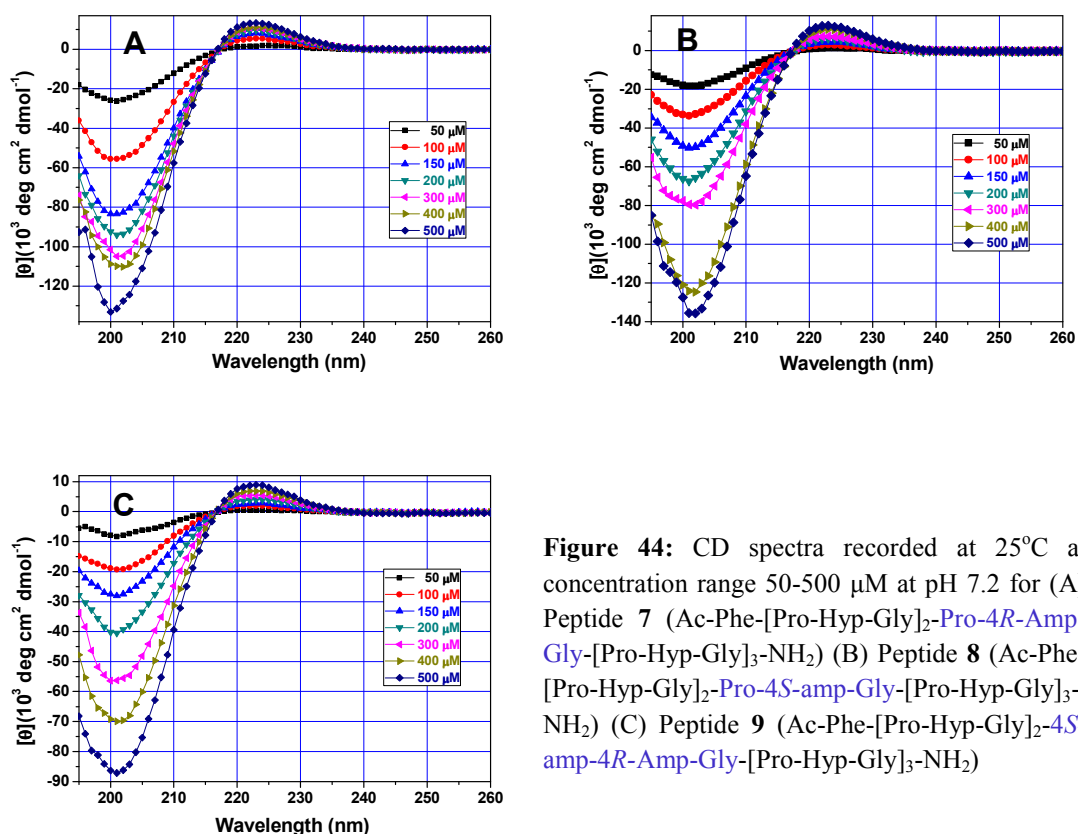


Figure 44: CD spectra recorded at 25°C at concentration range 50-500 μ M at pH 7.2 for (A) Peptide **7** (Ac-Phe-[Pro-Hyp-Gly]₂-Pro-4*R*-Amp-Gly-[Pro-Hyp-Gly]₃-NH₂) (B) Peptide **8** (Ac-Phe-[Pro-Hyp-Gly]₂-Pro-4*S*-amp-Gly-[Pro-Hyp-Gly]₃-NH₂) (C) Peptide **9** (Ac-Phe-[Pro-Hyp-Gly]₂-4*S*-amp-4*R*-Amp-Gly-[Pro-Hyp-Gly]₃-NH₂)

These peptides show isobestic point at 216-218 nm. Figure 45 show a plot of $R_{p/n}$ values derived from CD spectra against concentration of peptides **7-9**. A critical triple-

helical concentration of between 200-300 μM was derived from these plots for peptides 7-9.

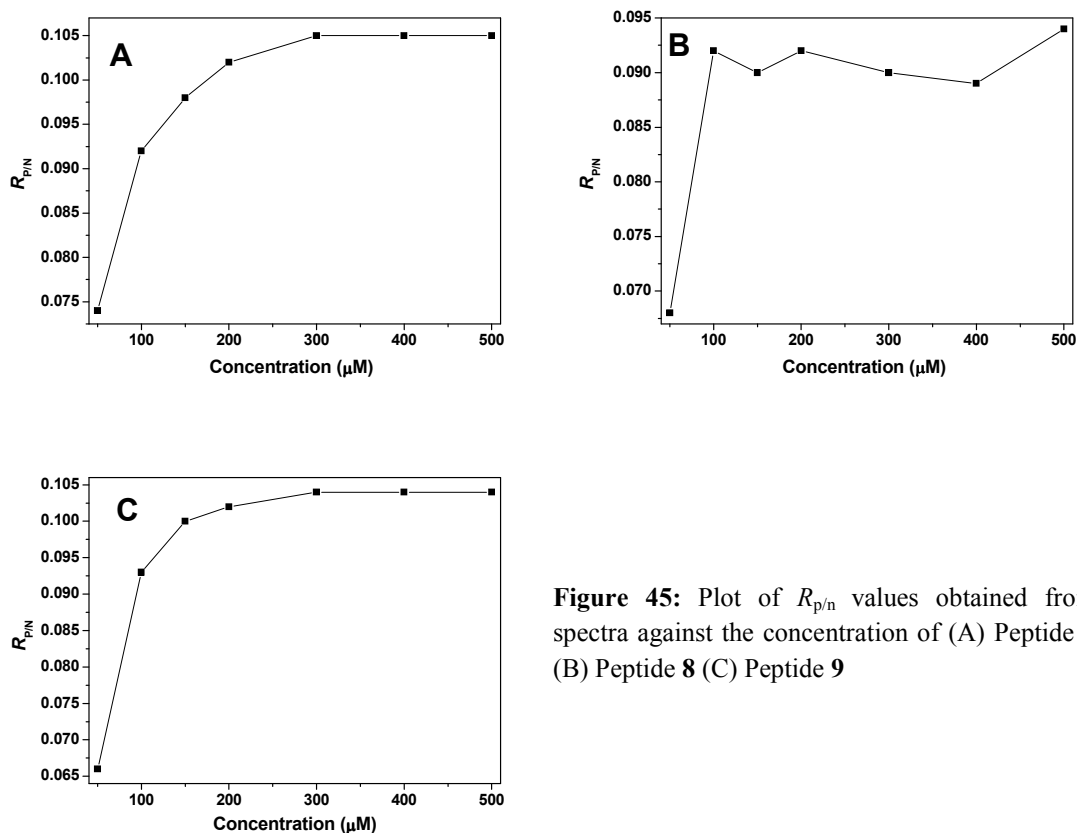


Figure 45: Plot of $R_{p/n}$ values obtained from spectra against the concentration of (A) Peptide 7 (B) Peptide 8 (C) Peptide 9

3.3.2b CD thermal denaturation study of peptides 7-9 at different pHs: In order to get information about the relative triple-helical strength of peptides 7-9, CD-spectroscopic studies as a function of temperature was carried out.

Peptide 7 (Ac-Phe-[Pro-Hyp-Gly]₂-Pro-4R-Amp-Gly-[Pro-Hyp-Gly]₃-NH₂): The CD-thermal denaturation plot of molar ellipticity at 222-224 nm *versus* temperature at different pHs for peptide 7 is shown in Figure 46A. Figure 46B shows the first derivative curve obtained from the sigmoidal fit of data of Figure 46A and the T_m values for peptide 7 obtained from minima of Figure 46B at different pHs is shown in Table 12. The T_m of peptide 7 decreases with increase in pH.

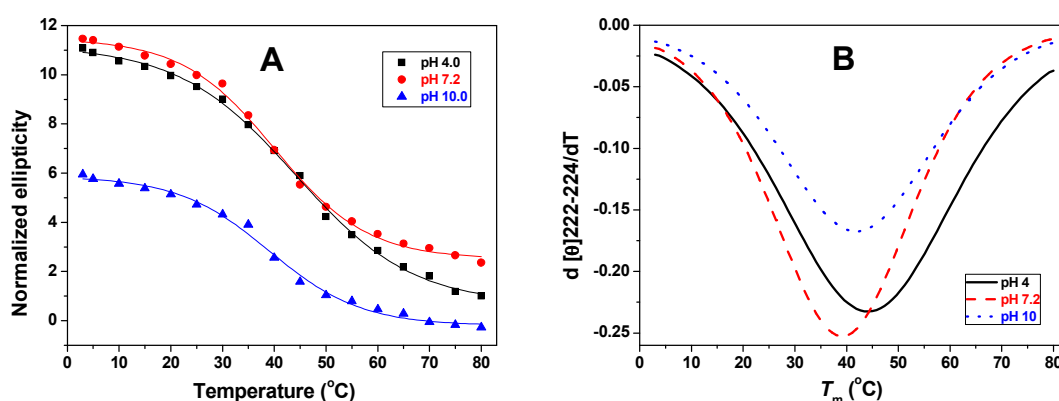


Figure 46: (A) CD-thermal denaturation plot of molar ellipticity at 222-224 nm vs. temperature at different pHs for peptide **7** (Ac-Phe-[Pro-Hyp-Gly]₂-Pro-4*R*-Amp-Gly-[Pro-Hyp-Gly]₃-NH₂). (B) First derivative curve of peptide **7** as a function of pH.

Table 12: T_m of peptide **7** (Ac-Phe-[Ac-Phe-[Pro-Hyp-Gly]₂-Pro-4*R*-Amp-Gly-[Pro-Hyp-Gly]₃-NH₂)

pH	T_m	ΔT_m (Peptide 7-4)
4.0	45	-1
7.2	40	-1
10.0	39	-1

T_m values are (± 0.5 °C)

Peptide 8 (Ac-Phe-[Pro-Hyp-Gly]₂-Pro-4*S*-amp-Gly-[Pro-Hyp-Gly]₃-NH₂): Figure 47A shows the CD-thermal denaturation plot of molar ellipticity at 222-224 nm *versus* temperature at different pHs for peptide **8** (upto 45 °C, beyond this temperature triplex was not detected).

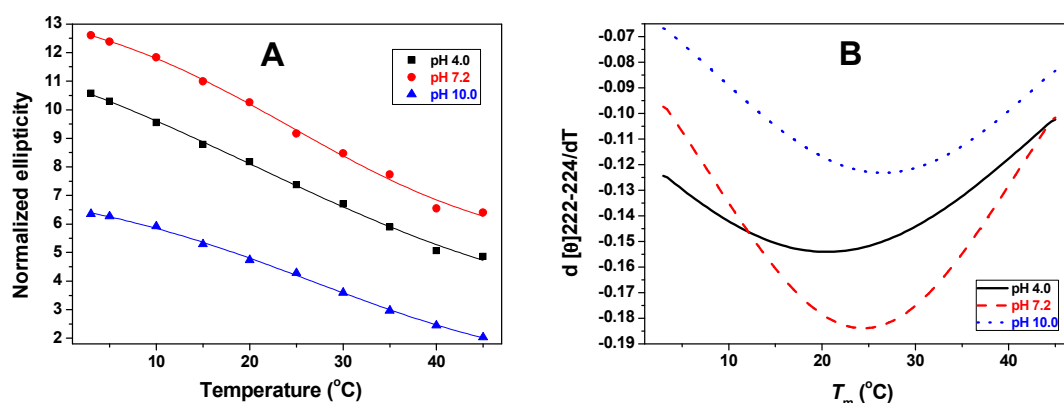


Figure 47: (A) CD-thermal denaturation plot of molar ellipticity at 222-224 nm vs. temperature at different pHs for peptide **8** (Ac-Phe-[Pro-Hyp-Gly]₂-Pro-4*S*-amp-Gly-[Pro-Hyp-Gly]₃-NH₂). (B) First derivative curve of peptide **8** as a function of pH.

Figure 47B shows the first derivative curve obtained from the sigmoidal fit of data of Figure 47A and the T_m values for peptide **8** obtained from minima of Figure 47B at different pHs is shown in Table 13. The T_m of peptide **8** increases with increase in pH.

Table 13: T_m of peptide **8** (Ac-Phe-[Ac-Phe-[Pro-Hyp-Gly]₂-Pro-4*S*-amp-Gly-[Pro-Hyp-Gly]₃-NH₂)

pH	T_m	ΔT_m (Peptide 8-5)
4.0	20	+1
7.2	24	+1
10.0	26	+1

T_m values are (± 0.5 °C)

Peptide 9 (Ac-Phe-[Pro-Hyp-Gly]₂-4*S*-amp-4*R*-Amp-Gly-[Pro-Hyp-Gly]₃-NH₂): The CD-thermal denaturation plot of molar ellipticity at 222-223 nm *versus* temperature at different pHs for peptide **9** is shown in Figure 48A. Figure 48B shows the first derivative curve obtained from the sigmoidal fit of data of Figure 48A and the T_m values for peptide **9** obtained from minima of Figure 48B at different pHs is shown in Table 14.

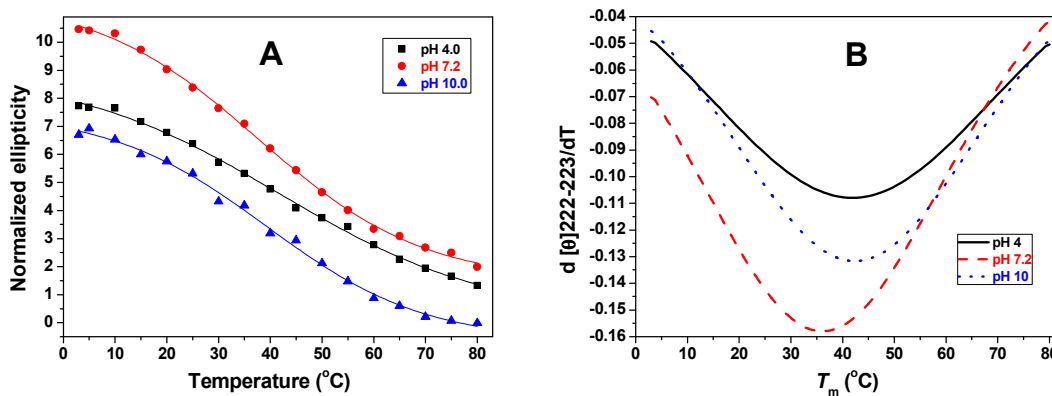


Figure 48: (A) CD-thermal denaturation plot of molar ellipticity at 222-223 nm vs. temperature at different pHs for peptide **9** (Ac-Phe-[Pro-Hyp-Gly]₂-4*S*-amp-4*R*-Amp-Gly-[Pro-Hyp-Gly]₃-NH₂). (B) First derivative curve of peptide **9** as a function of pH.

Table 14: T_m of peptide **9** Ac-Phe-[Pro-Hyp-Gly]₂-4*S*-amp-4*R*-Amp-Gly-[Pro-Hyp-Gly]₃-NH₂

pH	T_m	ΔT_m (Peptide 9-6)
4.0	42	+1
7.2	38	+1
10.0	40	+1

T_m values are (± 0.5 °C)

3.3.2c CD thermal denaturation study of peptides 7-9 in ethylene glycol: Figure 49A shows the CD-thermal denaturation plot of molar ellipticity at 222-223 nm *versus* temperature in ethylene glycol for peptides 7-9. Figure 49B shows the first derivative curve obtained from the sigmoidal fit of data of Figure 49A and the T_m values obtained from minima of Figure 49B of peptides 7-9 is shown in Table 15.

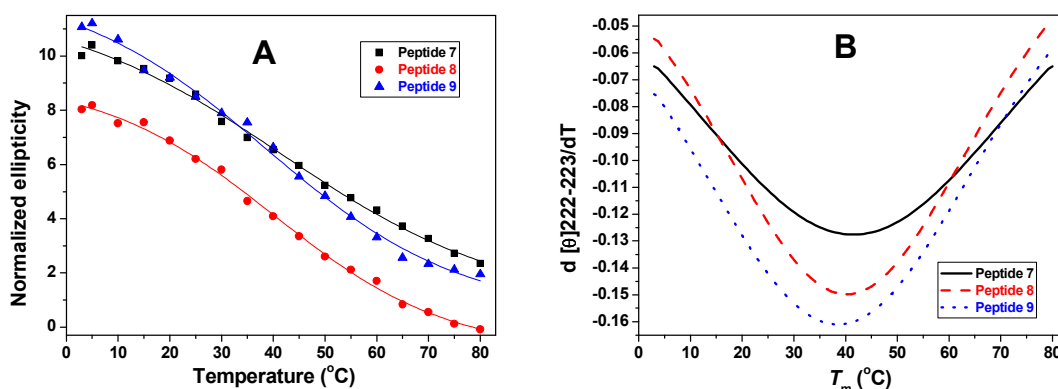


Figure 49: (A) CD-thermal denaturation plot of molar ellipticity at 222-223 nm vs. temperature in ethylene glycol for peptides 7-9. (B) First derivative curve of peptides 7-9 in ethylene glycol.

Table 15: T_m of peptides 7-9 in ethylene glycol

Peptide	T_m	ΔT_m compared with C-terminal modified peptides
7	42	+1 (Peptide 7-4)
8	42	-1 (Peptide 8-5)
9	39	-1 (Peptide 9-6)

T_m values are (± 0.5 °C)

3.3.2d Comparison of stability of peptides 7-9 as a function of pH: A comparative study of T_m for peptides 7-9 as a function of pH is shown in Figure 50.

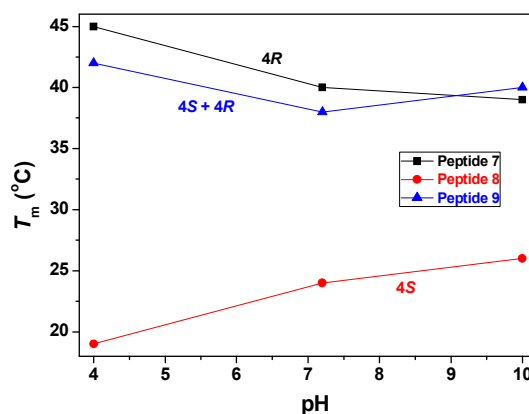


Figure 50: T_m of peptides 7-9 as a function of pH.

A comparison of T_m data (Figure 50) reveals following: (i) Peptide 7 at pH 4.0 shows a T_m of 45 °C which was maximum and decreased to 40 °C at pH 7.2 and to 39 °C at pH 10.0 (ii) An opposite behavior was observed for peptide 8 which shows minimum T_m 19 °C at pH 4.0, which increases to 24 °C at pH 7.2 and to 26 °C at pH 10.0. (iii) Peptide 9 at pH 4.0 shows a T_m of 42 °C which is maximum and decreases to 38 °C at pH 7.2 and a slight increase to 40 °C at pH 10.0. At pH 10.0, peptide 9 shows higher triple helix stability than peptide 7.

3.3.3 CD Spectroscopic Studies of N-Terminal Modified Peptides 10-12

Ac-Phe-Pro-4*R*-Amp-Gly-[Pro-Hyp-Gly]₅-NH₂ (Peptide 10)

Ac-Phe-Pro-4*S*-amp-Gly-[Pro-Hyp-Gly]₅-NH₂ (Peptide 11)

Ac-Phe-4*S*-amp-4*R*-Amp-Gly-[Pro-Hyp-Gly]₅-NH₂ (Peptide 12)

3.3.3a Concentration dependent CD spectroscopy for peptides 10-12: Figure 51 show the CD spectra with varying concentrations (50 μM –500 μM) of peptides 10-12.

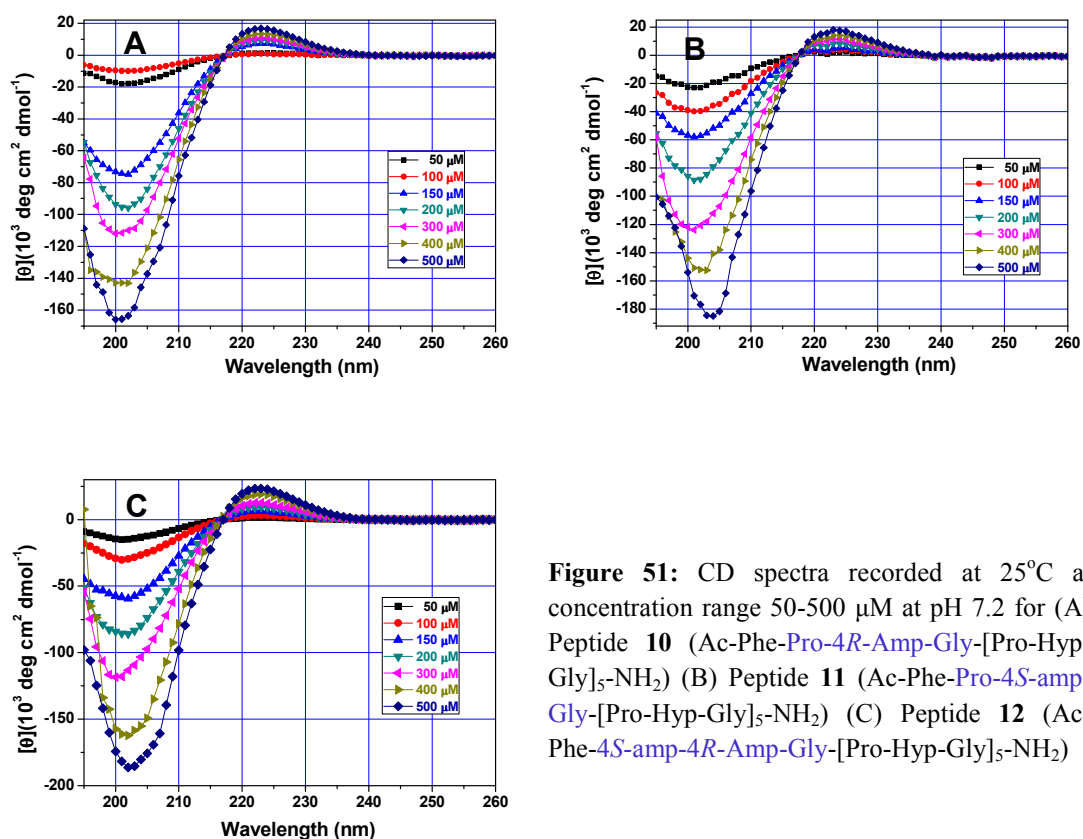


Figure 51: CD spectra recorded at 25°C at concentration range 50-500 μM at pH 7.2 for (A) Peptide 10 (Ac-Phe-Pro-4*R*-Amp-Gly-[Pro-Hyp-Gly]₅-NH₂) (B) Peptide 11 (Ac-Phe-Pro-4*S*-amp-Gly-[Pro-Hyp-Gly]₅-NH₂) (C) Peptide 12 (Ac-Phe-4*S*-amp-4*R*-Amp-Gly-[Pro-Hyp-Gly]₅-NH₂)

These peptides show isobestic point at 216-218 nm. Figure 52 show a plot of $R_{p/n}$ values derived from CD spectra against concentration of peptides 10-12.

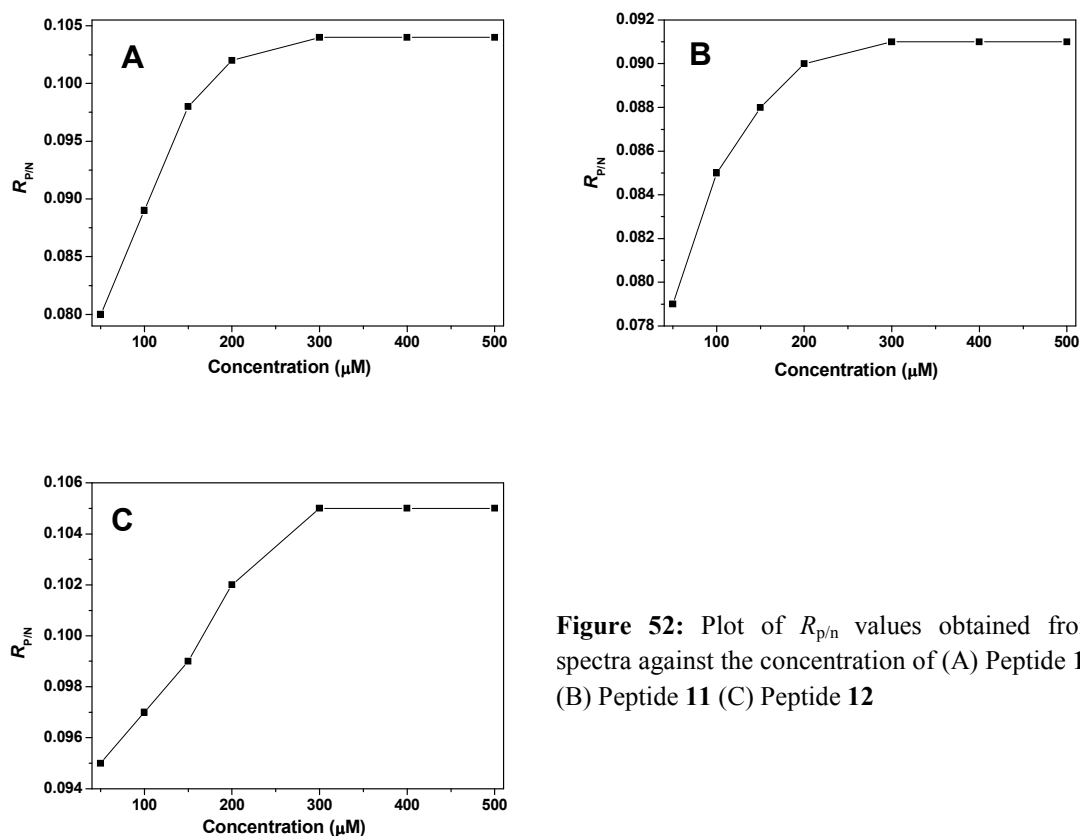


Figure 52: Plot of $R_{p/n}$ values obtained from spectra against the concentration of (A) Peptide 10 (B) Peptide 11 (C) Peptide 12

3.3.3b CD thermal denaturation study of peptides 10-12 at different pHs: In order to get information about the relative triple-helical strength of peptides 10-12, CD-spectroscopic studies as a function of temperature was carried out.

Peptide 10 (Ac-Phe-Pro-4*R*-Amp-Gly-[Pro-Hyp-Gly]₅-NH₂): The CD-thermal denaturation plot of molar ellipticity at 222-223 nm *versus* temperature at different pHs for peptide 10 is shown in Figure 53A. Figure 53B shows the first derivative curve obtained from the sigmoidal fit of data of Figure 53A and the T_m values for peptide 10 obtained from minima of Figure 53B at different pHs is shown in Table 16. The T_m of peptide 10 decreases with increase in pH.

Table 16: T_m of peptide 10 Ac-Phe-Pro-4*R*-Amp-Gly-[Pro-Hyp-Gly]₅-NH₂

pH	T_m	ΔT_m (Peptide 10-4)
4.0	44	-2
7.2	39	-2
10.0	38	-2

T_m values are (± 0.5 °C)

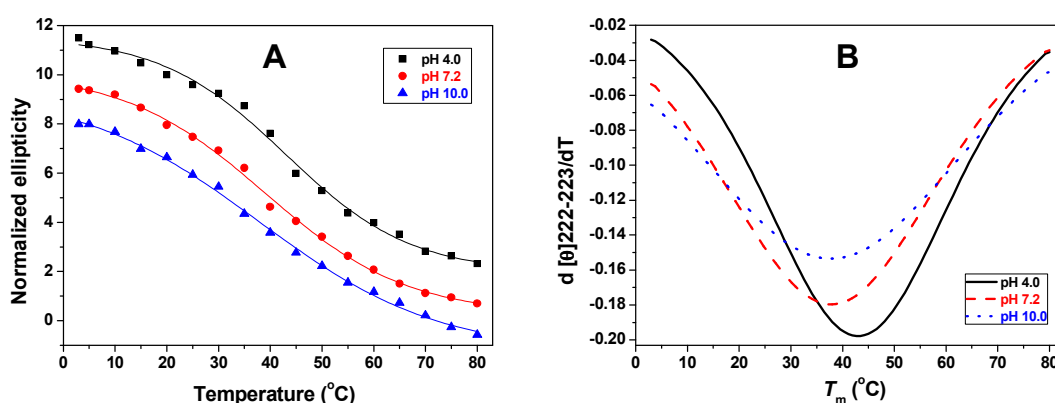


Figure 53: (A) CD-thermal denaturation plot of molar ellipticity at 222-223 nm vs. temperature at different pHs for peptide **10** (Ac-Phe-Pro-4R-Amp-Gly-[Pro-Hyp-Gly]₅-NH₂). (B) First derivative curve of peptide **10** as a function of pH.

Peptide 11 (Ac-Phe-Pro-4S-amp-Gly-[Pro-Hyp-Gly]₅-NH₂): Figure 54A shows the CD-thermal denaturation plot of molar ellipticity at 222-223 nm *versus* temperature at different pHs for peptide **11**. Figure 54B shows the first derivative curve obtained from the sigmoidal fit of data of Figure 54A and the T_m values for peptide **11** obtained from minima of Figure 54B at different pHs is shown in Table 17. The T_m of peptide **11** increases with increase in pH.

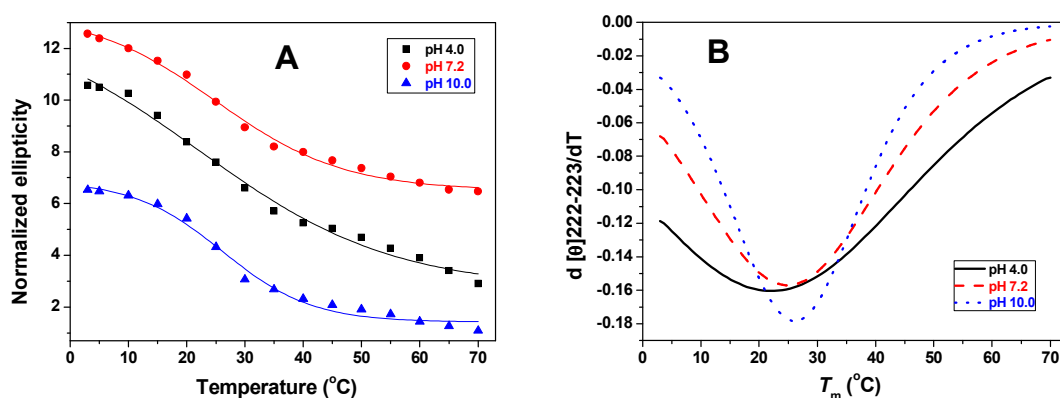


Figure 54: (A) CD-thermal denaturation plot of molar ellipticity at 222-223 nm vs. temperature at different pHs for peptide **11** (Ac-Phe-Pro-4S-amp-Gly-[Pro-Hyp-Gly]₅-NH₂). (B) First derivative curve of peptide **11** as a function of pH.

Table 17: T_m of peptide **11** (Ac-Phe-Pro-4S-amp-Gly-[Pro-Hyp-Gly]₅-NH)

pH	T_m	ΔT_m (Peptide 11-5)
4.0	20	+2
7.2	25	+2
10.0	26.5	+1.5

T_m values are (± 0.5 °C)

Peptide 12 (Ac-Phe-4*S*-amp-4*R*-Amp-Gly-[Pro-Hyp-Gly]₅-NH₂): The CD-thermal denaturation plot of molar ellipticity at 222-223 nm *versus* temperature at different pHs for peptide **12** is shown in Figure 55A. Figure 55B shows the first derivative curve obtained from the sigmoidal fit of data of Figure 55A and the T_m values for peptide **12** obtained from minima of Figure 55B at different pHs is shown in Table 18.

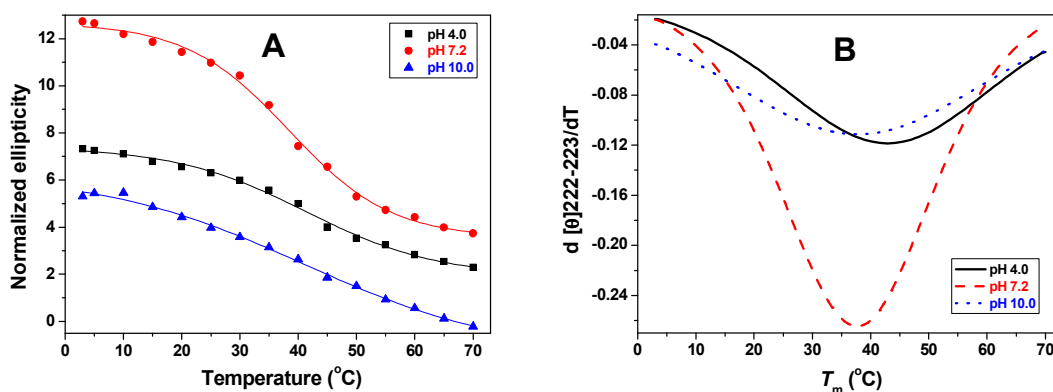


Figure 55: (A) CD-thermal denaturation plot of molar ellipticity at 222-223 nm vs. temperature at different pHs for peptide **11** (Ac-Phe-4*S*-amp-4*R*-Amp-Gly-[Pro-Hyp-Gly]₅-NH₂). (B) First derivative curve of peptide **11** as a function of pH.

Table 18: T_m of peptide **12** Ac-Phe-4*S*-amp-4*R*-Amp-Gly-[Pro-Hyp-Gly]₅-NH₂

pH	T_m	ΔT_m (Peptide 12-6)
4.0	43	+2
7.2	40	+2
10.0	41	+2

T_m values are (± 0.5 °C)

3.3.3c Comparison of stability of peptides 10-12 as a function of pH: A comparative study of T_m for peptides **10-12** as a function of pH is shown in Figure 56.

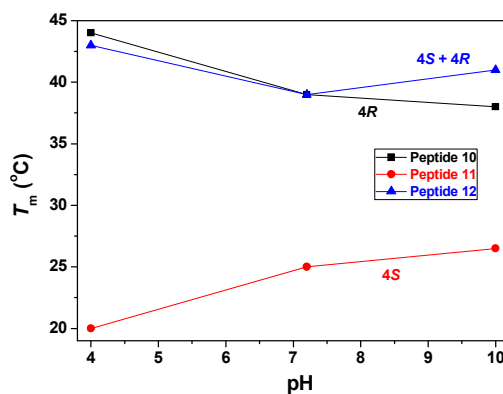


Figure 56: T_m of peptides **10-12** as a function of pH.

A comparison of T_m data (Figure 56) reveals following:

(i) Peptide **10** at pH 4.0 shows a T_m of 44 °C which has maximum and decreases to 39 °C at pH 7.2 and to 38 °C at pH 10.0. (ii) An opposite behavior was observed for peptide **11** which shows minimum T_m 20 °C at pH 4.0, which increases to 25 °C at pH 7.2 and to 26.5 °C at pH 10.0 (iii) Peptide **12** at pH 4.0 shows a T_m of 43 °C which is maximum and decreases to 39 °C at pH 7.2 and slightly increase to 41 °C at pH 10.0. At pH 10.0, peptide **12** shows higher triple helix stability than peptide **10**.

3.3.3d CD thermal denaturation study of peptides 10-12 in ethylene glycol: Figure 57A shows the CD-thermal denaturation plot of molar ellipticity at 222-223 nm *versus* temperature in ethylene glycol for peptides **10-12**. Figure 57B shows the first derivative curve obtained from the sigmoidal fit of data of Figure 57A and the T_m values obtained from minima of Figure 57B of peptides **10-12** is shown in Table 19.

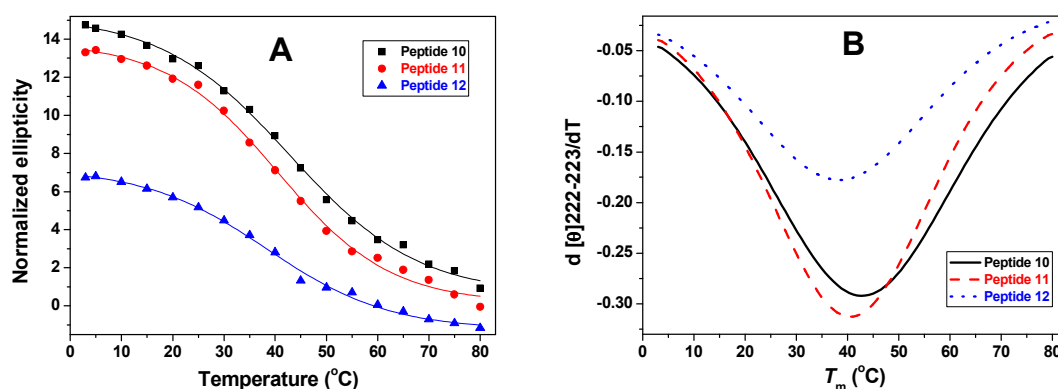


Figure 57: (A) CD-thermal denaturation plot of molar ellipticity at 222-223 nm vs. temperature in ethylene glycol for peptides **10-12**. (B) First derivative curve of peptides **10-12** in ethylene glycol

Table 19: T_m of peptides **10-12** in ethylene glycol

Peptide	T_m	ΔT_m compared with C-terminal modified peptide
7	43	+2 (Peptide 10-4)
8	40	-3 (Peptide 11-5)
9	38	-2 (Peptide 12-6)

T_m values are (± 0.5 °C)

3.3.4a Comparative study of C-terminus, middle and N-terminus modifications at different pHs: The comparative T_m values of peptides **4-12** at pH 4.0, pH 7.2 and pH 10.0 are shown in Figure 58. From the figure it is seen that when a triplet modification is done at Y position by 4*R*-aminopropyl (peptides **4**, **7** and **10**), the T_m

slightly decreases from C-terminus to N-terminus modified peptides. When a triplet modification is done at Y position by 4S-aminoprolyl (peptides 5, 8 and 11), the T_m slightly increases from C-terminus to N-terminus modified peptides. When a triplet modification is done at X position by 4S-aminoprolyl and 4R-aminoprolyl at Y position (peptides 6, 9 and 12), the T_m slightly increases from C-terminus to N-terminus modified peptides.

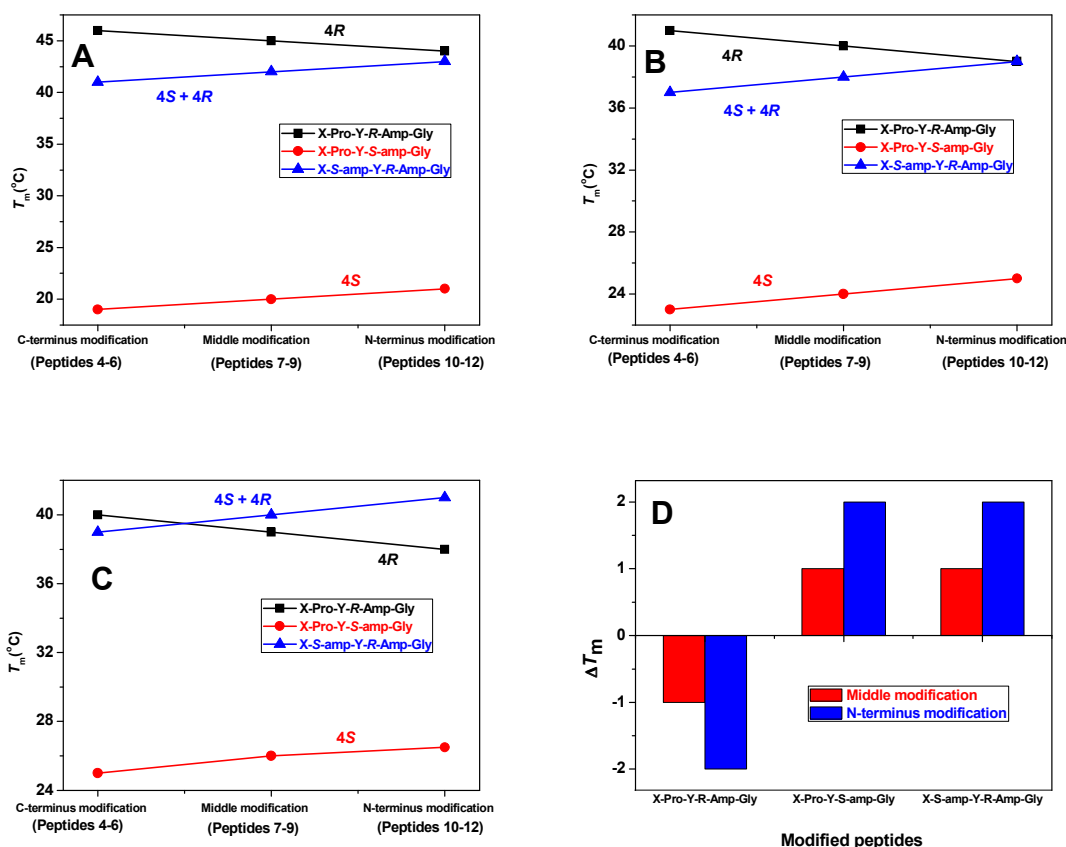


Figure 58: Comparative study of stability of peptides 4-12 with respect to C-terminus, middle and N-terminus modification. (A) At pH 4.0 (B) At pH 7.2 (C) At pH 10.0 (D) ΔT_m of middle modified and N-terminus modified peptides compared with C-terminus modified peptides in all pHs.

3.3.4b Comparative study of C-terminus, middle and N-terminus modifications in ethylene glycol:

The comparative T_m values of peptides 4-12 in ethylene glycol are shown in Figure 59. It is seen that when a triplet modification is done at Y position by 4R-aminoprolyl (peptides 4, 7 and 10), the T_m slightly increases from C-terminus to N-terminus modified peptides. When a triplet modification is done at Y position by 4S-aminoprolyl (peptides 5, 8 and 11), the T_m slightly decreases from C-terminus to N-terminus modified peptides. When a triplet modification is done at X position by 4S-

aminopropyl and 4*R*-aminopropyl at Y position (peptides **6**, **9** and **12**), the T_m slightly decreases from C-terminus to N-terminus modified peptides. This behavior observed in ethylene glycol is opposite to that seen in aqueous system.

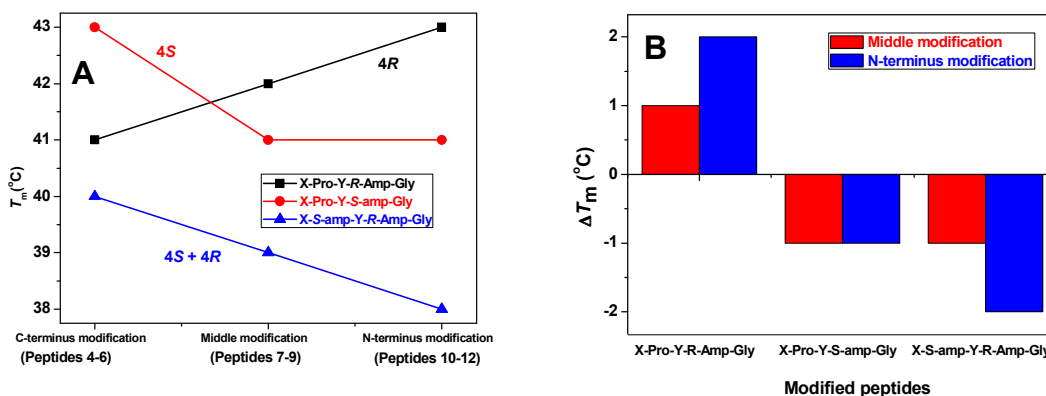


Figure 59: Comparative study of stability of peptides **4-12** with respect to C-terminus, middle and N-terminus modification. (A) T_m in ethylene glycol (B) ΔT_m of middle modified and N-terminus modified peptides compared with C-terminus modified peptides in ethylene glycol.

From the CD spectroscopic studies of N-capped peptides, it is clear that stereochemistry and solvent play an important role for elucidating the stability of collagen triple helix having only one modified triplet. The modified sites (C-terminal, middle, N-terminal) have very little effect on stability of collagen triple helix.

3.4 CD Spectroscopic Studies for Non-N-Acetylated Peptides 13-21

The contributions from terminal effects to triplex stability may relatively differ in non-acetylated peptides **13-21** due to the presence of additional ionizable NH_2 groups. Hence CD-spectroscopic studies were performed for non-acetylated peptides.

It was reported earlier from this laboratory that, the T_m of uncapped collagen oligomer $\text{H}_2\text{N-Phe[Pro-Hyp-Gly]}_6$ **26** changes with pH and the T_m of 4*R*-aminoproline peptide $\text{H}_2\text{N-Phe[Pro-4R-Amp-Gly]}_6$ **27** was higher and changes with pH. In contrast, the 4*S*-aminoproline peptide $\text{H}_2\text{N-Phe[Pro-4S-amp-Gly]}_6$ **28** showed no such behaviour indicating that it does not form a triplex. The T_m of $\text{H}_2\text{N-Phe[4S-amp-4R-Amp-Gly]}_6$ peptide **29** was enhanced and shows pH dependent stability effect. The T_m values for these peptides are shown in Table 20.¹⁴³

Table 20: T_m of peptides **26-29** in °C at different pHs.

Peptide	pH/solvent				
	4.0	7.2	10.0	12.0	EG
H ₂ N-Phe[Pro-Hyp-Gly] ₆ (peptide 26)	23	21.6	15.5	39.6	31
H ₂ N-Phe[Pro-4 <i>R</i> -Amp-Gly] ₆ (peptide 27)	42	31.3	18.5	56.6	23
H ₂ N-Phe[Pro-4 <i>S</i> -amp-Gly] ₆ (peptide 28)	nd	nd	nd	nd	nd
H ₂ N-Phe[4 <i>S</i> -amp-4 <i>R</i> -Amp-Gly] ₆ (peptide 28)	45	40	34	42	36

Nd = triplex not detected, T_m values are (± 0.5 °C)

3.4.1 CD Spectroscopic Studies of C-Terminus Modified Peptides 12-15

H₂N-Phe-[Pro-Hyp-Gly]₅-Pro-4*R*-Amp-Gly-NH₂ (Peptide 13)

H₂N-Phe-[Pro-Hyp-Gly]₅-Pro-4*S*-amp-Gly-NH₂ (Peptide 14)

H₂N-Phe-[Pro-Hyp-Gly]₅-4*S*-amp-4*R*-Amp-Gly-NH₂ (Peptide 15)

3.4.1a Concentration dependent CD spectroscopy of peptides 13-15: Figure 60 shows the CD spectra recorded at different pHs and in ethylene glycol for peptides 13-15.

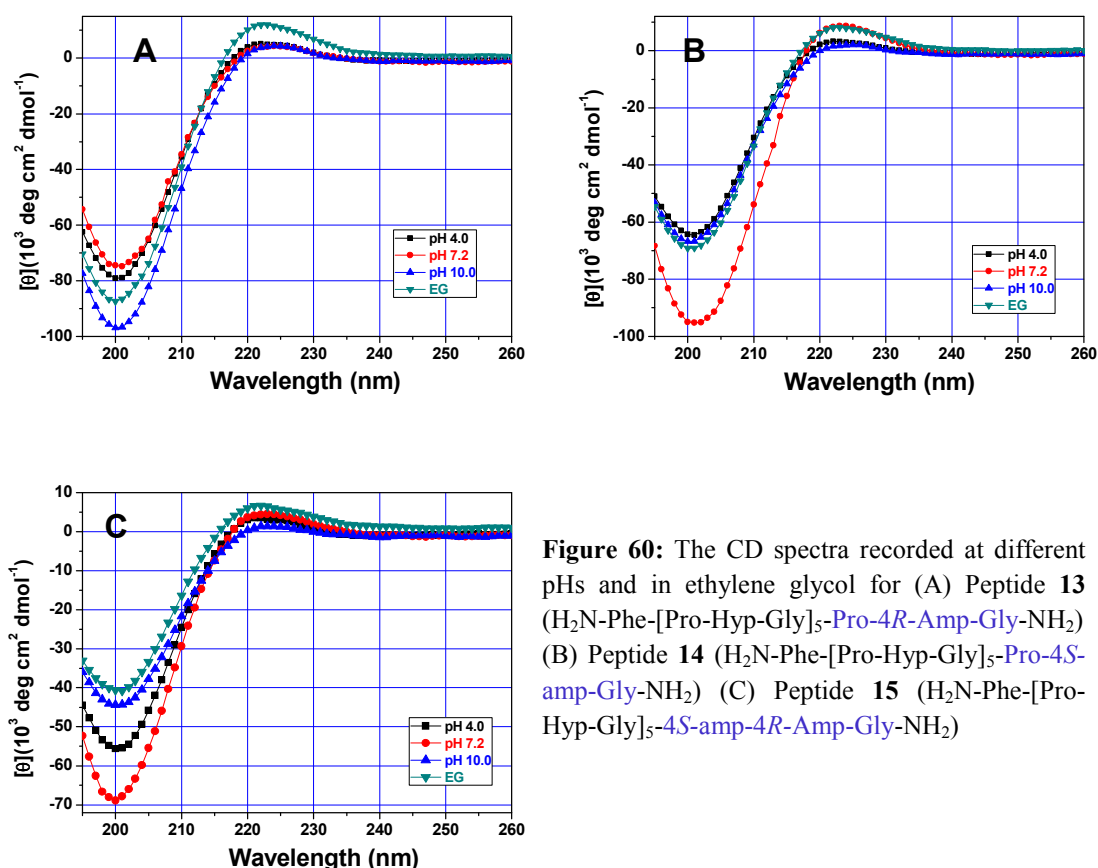


Figure 60: The CD spectra recorded at different pHs and in ethylene glycol for (A) Peptide 13 (H₂N-Phe-[Pro-Hyp-Gly]₅-Pro-4*R*-Amp-Gly-NH₂) (B) Peptide 14 (H₂N-Phe-[Pro-Hyp-Gly]₅-Pro-4*S*-amp-Gly-NH₂) (C) Peptide 15 (H₂N-Phe-[Pro-Hyp-Gly]₅-4*S*-amp-4*R*-Amp-Gly-NH₂)

The peptides 13-15 show positive maxima at 222-224 nm and negative minima at 201-204 nm respectively. Importantly all the spectral traces pass through an isobestic

point at 217-218 nm. A critical triple-helical concentration is derived at 200-300 μM from these plots.

3.4.1b CD thermal denaturation study of peptides 13-15 at different pHs: In order to get information about the relative triple-helical strength of peptides **13-15**, CD-spectroscopic studies as a function of temperature was carried out.

Peptide 13 ($\text{H}_2\text{N-Phe-[Pro-Hyp-Gly]}_5\text{-Pro-4R-Amp-Gly-NH}_2$): Figure 61A shows the CD-thermal denaturation plot of molar ellipticity at 222-224 nm *versus* temperature at different pHs for peptide **13**. Figure 61B shows the first derivative curve obtained from the sigmoidal fit of data of Figure 61A and the T_m values for peptide **13** obtained from minima of Figure 61B at different pHs is shown in Table 21.

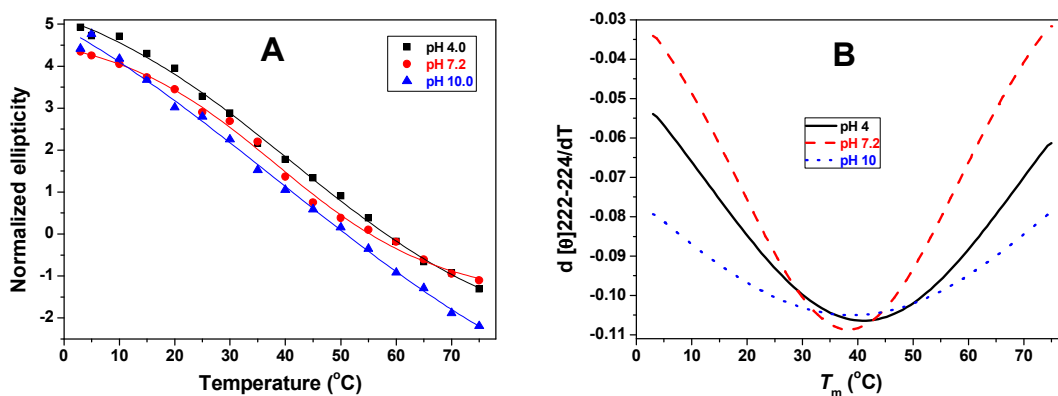


Figure 61: (A) CD-thermal denaturation plot of molar ellipticity at 222-224 nm vs. temperature at different pHs for peptide **13** ($\text{H}_2\text{N-Phe-[Pro-Hyp-Gly]}_5\text{-Pro-4R-Amp-Gly-NH}_2$). (B) First derivative curve of peptide **13** as a function of pH.

Table 21: T_m of peptide **13** ($\text{H}_2\text{N-Phe-[Pro-Hyp-Gly]}_5\text{-Pro-4R-Amp-Gly-NH}_2$)

pH	T_m	ΔT_m (13-26)	ΔT_m (13-27)
4.0	42	+19	0
7.2	38	+16.5	-6.7
10.0	39	+23	+20.5

T_m values are (± 0.5 °C)

Peptide 14 ($\text{H}_2\text{N-Phe-[Pro-Hyp-Gly]}_5\text{-Pro-4S-amp-Gly-NH}_2$): The CD-thermal denaturation plot of molar ellipticity at 222-224 nm *versus* temperature at different pHs for peptide **14** is shown in Figure 62A. Figure 62B shows the first derivative curve obtained from the sigmoidal fit of data of Figure 62A and the T_m values for

peptide **14** obtained from minima of Figure 62B at different pHs is shown in Table 22. The T_m of peptide **14** increases with increase in pH.

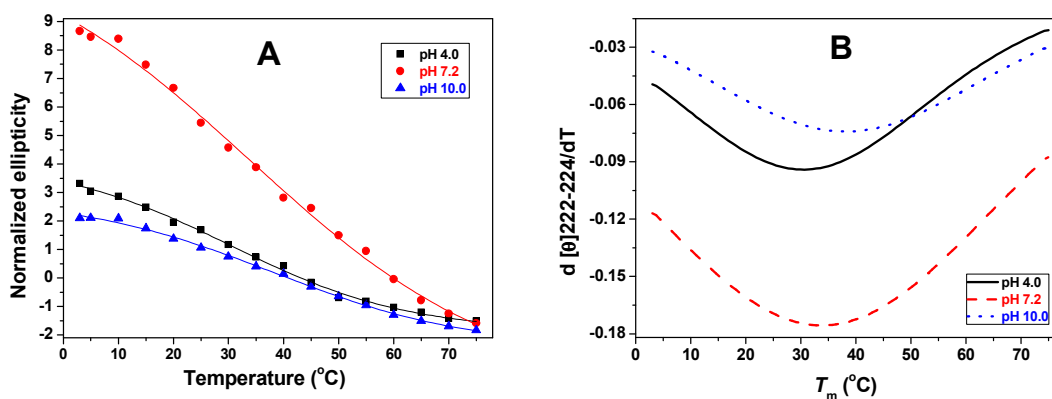


Figure 62: (A) CD-thermal denaturation plot of molar ellipticity at 222-224 nm vs. temperature at different pHs for peptide **14** ($\text{H}_2\text{N-Phe-[Pro-Hyp-Gly]}_5\text{-Pro-4S-amp-Gly-NH}_2$). (B) First derivative curve of peptide **14** as a function of pH.

Table 22: T_m of peptide **14** ($\text{H}_2\text{N-Phe-[Pro-Hyp-Gly]}_5\text{-Pro-4S-amp-Gly-NH}_2$)

pH	T_m	ΔT_m (14-26)	ΔT_m (14-28)
4.0	31	+8	+31
7.2	34	+12.4	+34
10.0	38	+22.5	+38

T_m values are (± 0.5 °C)

Peptide 15 ($\text{H}_2\text{N-Phe-[Pro-Hyp-Gly]}_5\text{-4S-amp-4R-Amp-Gly-NH}_2$): Figure 63A shows the CD-thermal denaturation plot of molar ellipticity at 222-224 nm *versus* temperature at different pHs for peptide **15**.

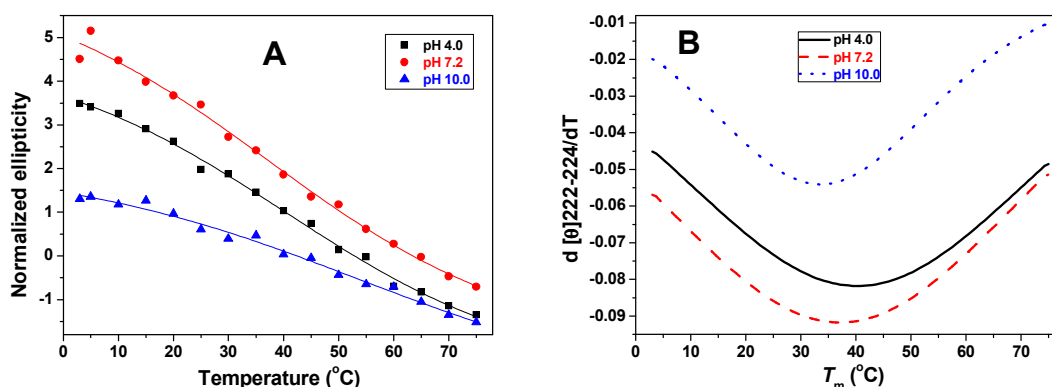


Figure 63: (A) CD-thermal denaturation plot of molar ellipticity at 222-224 nm vs. temperature at different pHs for peptide **15** ($\text{H}_2\text{N-Phe-[Pro-Hyp-Gly]}_5\text{-4S-amp-4R-Amp-Gly-NH}_2$). (B) First derivative curve of peptide **15** as a function of pH.

Figure 63B shows the first derivative curve obtained from the data of Figure 63A and the T_m values for peptide **15** obtained from minima of Figure 63B at different pHs is shown in Table 23. The T_m of peptide **15** decreases with increase in pH.

Table 23: T_m of peptide **15** (H₂N-Phe-[Pro-Hyp-Gly]₅-4*S*-amp-4*R*-Amp-Gly-NH₂)

pH	T_m	ΔT_m (15-26)	ΔT_m (15-29)
4.0	41	+18	-4
7.2	37	+15.5	-15.5
10.0	34	+18.5	-18.5

T_m values are (± 0.5 °C)

3.4.1c CD thermal denaturation study of peptides 13-15 in ethylene glycol: Figure 64A shows the CD-thermal denaturation plot of molar ellipticity at 222-223 nm *versus* temperature in ethylene glycol for peptides **13-15**. Figure 64B shows the first derivative curve obtained from the sigmoidal fit of data of Figure 64A and the T_m values obtained from minima of Figure 64B of peptides **13-15** is shown in Table 24.

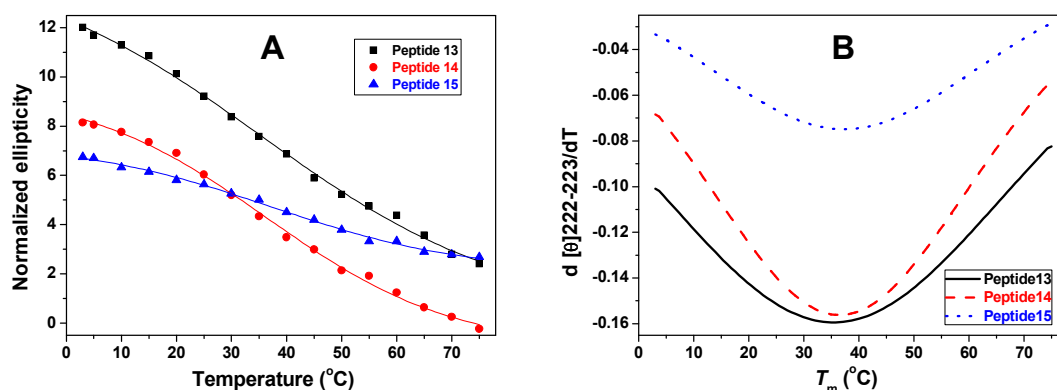


Figure 64: (A) CD-thermal denaturation plot of molar ellipticity at 222-223 nm vs. temperature in ethylene glycol for peptides **13-15**. (B) First derivative curve of peptides **13-15** in ethylene glycol.

Table 24: T_m of peptides **13-15** in ethylene glycol

Peptide	T_m
13	37
14	37
15	38

T_m values are (± 0.5 °C)

3.4.1d Comparison of stability of peptides 13-15 with peptides 26-29: The comparative ΔT_m values for peptides **13-15** with corresponding control peptides **26-29** is represented as a bar diagram in Figure 65.

A comparison of T_m data reveals the following:

(i) The increase in T_m for single 4*R*-aminoprolyl modification (peptide **13**) at all pHs is more than that of fully 4*R*-amino modified peptide **27** and control collagen peptide **26**.

(ii) The T_m of single 4*S*-aminoprolyl modification (peptide **14**) at all pHs is greater than control collagen peptide **26**, while triplex was not detected in fully modified 4*S*-aminoprolyl peptide **28**.

(iii) The chimeric peptide **15** with single triplet modification (4*S*-aminoprolyl at X position and 4*R*-aminoprolyl at Y position) shows higher T_m as compared with control collagen peptide **26** and slightly less stability than that of fully modified peptide **29** at all pHs.

(iv) Peptides **13-15** show higher T_m compared to fully modified peptides **27-29** in EG.

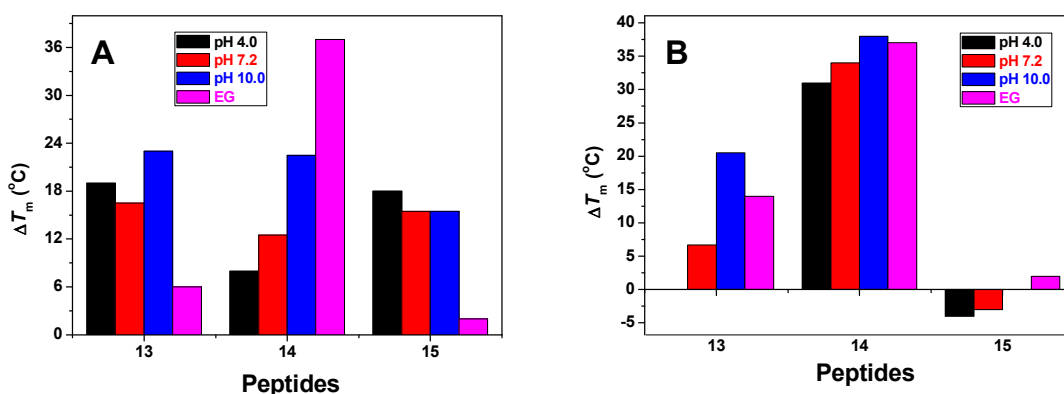


Figure 65: (a) ΔT_m of peptides **13-15** compared with control collagen peptide **26** ($\text{H}_2\text{N-Phe[Pro-Hyp-Gly]}_6$) (b) ΔT_m of peptides **13-15** compared with corresponding fully modified peptides **27-29**.

3.4.2 CD Spectroscopic Studies of Middle Modified Peptides 16-18

$\text{H}_2\text{N-Phe-[Pro-Hyp-Gly]}_2\text{-Pro-4R-Amp-Gly-[Pro-Hyp-Gly]}_3\text{-NH}_2$ (Peptide **16**)

$\text{H}_2\text{N-Phe-[Pro-Hyp-Gly]}_2\text{-Pro-4S-amp-Gly-[Pro-Hyp-Gly]}_3\text{-NH}_2$ (Peptide **17**)

$\text{H}_2\text{N-Phe-[Pro-Hyp-Gly]}_2\text{-4S-amp-4R-Amp-Gly-[Pro-Hyp-Gly]}_3\text{-NH}_2$ (Peptide **18**)

3.4.2a Concentration dependent CD spectroscopy of peptides 16-18: Figure 66 shows the CD spectra recorded at different pHs and in ethylene glycol for peptides **16-18**. The peptides **16-18** show positive maxima at 221-225 nm and negative minima at 200-205 nm respectively. Importantly, all the spectral traces pass through an isobestic point at 217-218 nm. A critical triple-helical concentration is derived at 150-300 μM from these plots.

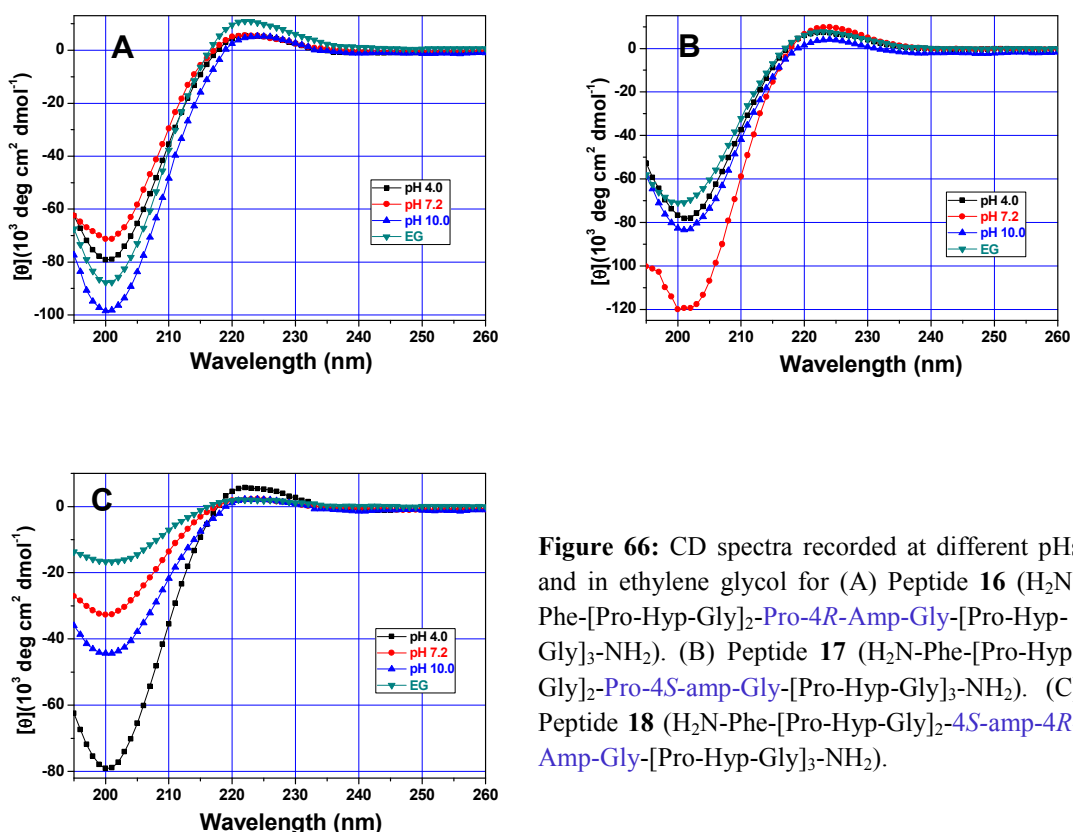


Figure 66: CD spectra recorded at different pHs and in ethylene glycol for (A) Peptide **16** ($\text{H}_2\text{N-Phe-}[\text{Pro-Hyp-Gly}]_2\text{-Pro-4R-Amp-Gly-}[\text{Pro-Hyp-Gly}]_3\text{-NH}_2$). (B) Peptide **17** ($\text{H}_2\text{N-Phe-}[\text{Pro-Hyp-Gly}]_2\text{-Pro-4S-amp-Gly-}[\text{Pro-Hyp-Gly}]_3\text{-NH}_2$). (C) Peptide **18** ($\text{H}_2\text{N-Phe-}[\text{Pro-Hyp-Gly}]_2\text{-4S-amp-4R-Amp-Gly-}[\text{Pro-Hyp-Gly}]_3\text{-NH}_2$).

3.4.2b CD thermal denaturation study of peptides 16-18 at different pHs: In order to get information about the relative triple-helical strength of peptides **16-18**, CD-spectroscopic studies as a function of temperature was carried out.

Peptide 16 ($\text{H}_2\text{N-Phe-}[\text{Pro-Hyp-Gly}]_2\text{-Pro-4R-Amp-Gly-}[\text{Pro-Hyp-Gly}]_3\text{-NH}_2$): The CD-thermal denaturation plot of molar ellipticity at 222-224 nm *versus* temperature at different pHs for peptide **16** is shown in Figure 67A. Figure 67B shows the first derivative curve obtained from the sigmoidal fit of data of Figure 67A and the T_m values for peptide **16** obtained from minima of Figure 67B at different pHs is shown in Table 25.

Table 25: T_m of peptide **16** ($\text{H}_2\text{N-Phe-}[\text{Ac-Phe-}[\text{Pro-Hyp-Gly}]_2\text{-Pro-4R-Amp-Gly-}[\text{Pro-Hyp-Gly}]_3\text{-NH}_2$)

pH	T_m	ΔT_m (Peptide 16-13)
4.0	41	-1
7.2	37	-1
10.0	38	-1

T_m values are (± 0.5 °C)

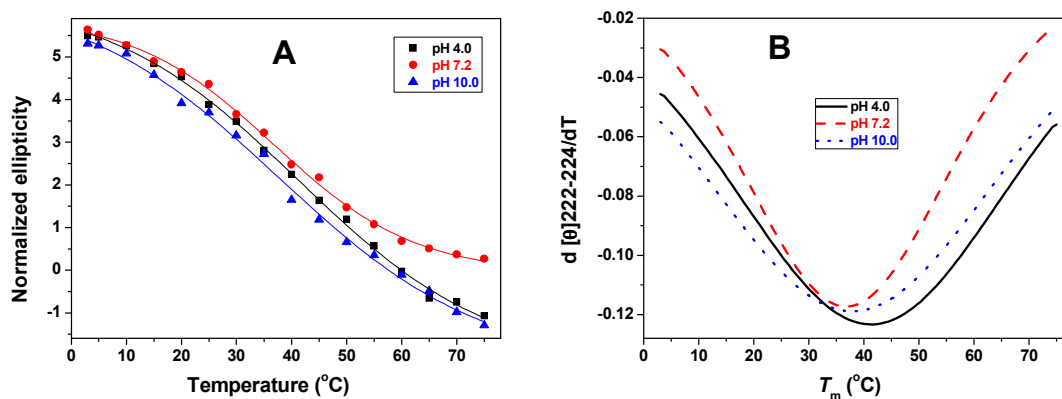


Figure 67: (A) CD-thermal denaturation plot of molar ellipticity at 222-224 nm vs. temperature at different pHs for peptide **16** ($\text{H}_2\text{N-Phe-}[\text{Pro-Hyp-Gly}]_2\text{-Pro-4R-Amp-Gly-}[\text{Pro-Hyp-Gly}]_3\text{-NH}_2$). (B) First derivative curve of peptide **16** as a function of pH.

Peptide 17 ($\text{H}_2\text{N-Phe-}[\text{Pro-Hyp-Gly}]_2\text{-Pro-4S-amp-Gly-}[\text{Pro-Hyp-Gly}]_3\text{-NH}_2$): The CD-thermal denaturation plot of molar ellipticity at 222-224 nm *versus* temperature at different pHs for peptide **17** is shown in Figure 68A. Figure 68B shows the first derivative curve obtained from the sigmoidal fit of data of Figure 68A and the T_m values for peptide **17** obtained from minima of Figure 68B at different pHs is shown in Table 26. The T_m of peptide **19** increases with increase in pH.

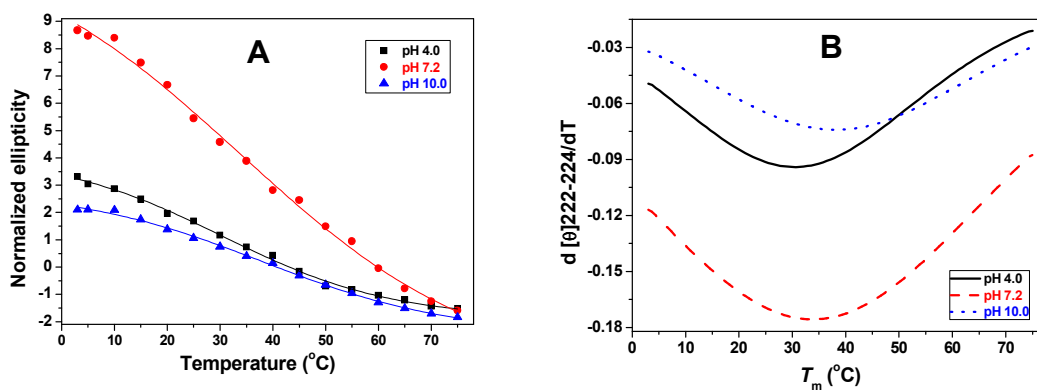


Figure 68: (A) CD-thermal denaturation plot of molar ellipticity at 222-224 nm vs. temperature at different pHs for peptide **17** ($\text{H}_2\text{N-Phe-}[\text{Pro-Hyp-Gly}]_2\text{-Pro-4S-amp-Gly-}[\text{Pro-Hyp-Gly}]_3\text{-NH}_2$). (B) First derivative curve of peptide **17** as a function of pH.

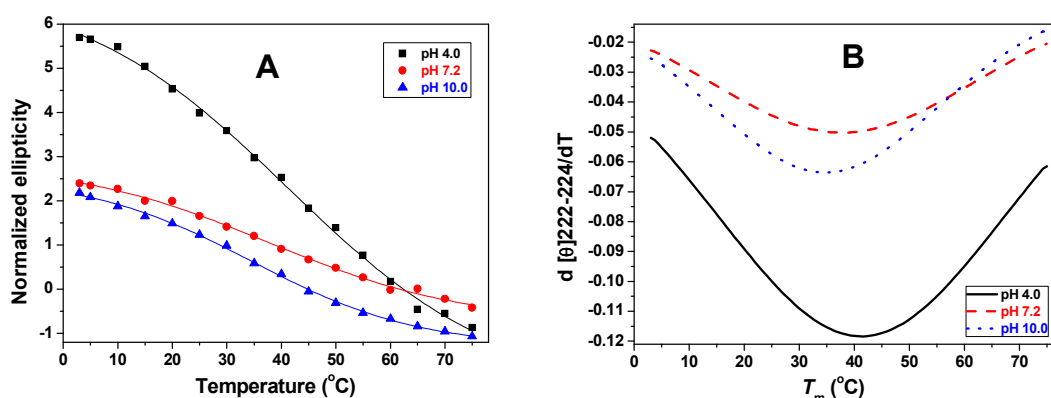
Table 26: T_m of peptide **17** (H₂N-Phe-[Ac-Phe-[Pro-Hyp-Gly]₂-Pro-4*S*-amp-Gly-[Pro-Hyp-Gly]₃-NH₂])

pH	T_m	ΔT_m (Peptide 17-14)
4.0	32	+1
7.2	35	+1
10.0	38	+1

T_m values are (± 0.5 °C)

Peptide 18 (H₂N-Phe-[Pro-Hyp-Gly]₂-4*S*-amp-4*R*-Amp-Gly-[Pro-Hyp-Gly]₃-NH₂):

Figure 69A shows the CD-thermal denaturation plot of molar ellipticity at 222-224 nm *versus* temperature at different pHs for peptide **18**. Figure 69B shows the first derivative curve obtained from the sigmoidal fit of data of Figure 69A and the T_m values for peptide **18** obtained from minima of Figure 69B at different pHs is shown in Table 27. The T_m of peptide **18** decreases with increase in pH.

**Figure 69:** (A) CD-thermal denaturation plot of molar ellipticity at 222-224 nm vs. temperature at different pHs for peptide **18** (H₂N-Phe-[Pro-Hyp-Gly]₂-4*S*-amp-4*R*-Amp-Gly-[Pro-Hyp-Gly]₃-NH₂). (B) First derivative curve of peptide **18** as a function of pH.**Table 27:** T_m of peptide **18** H₂N-Phe-[Pro-Hyp-Gly]₂-4*S*-amp-4*R*-Amp-Gly-[Pro-Hyp-Gly]₃-NH₂

pH	T_m	ΔT_m (Peptide 18-15)
4.0	42	+1
7.2	38	+1
10.0	35	+1

T_m values are (± 0.5 °C)

3.4.2c CD thermal denaturation study of peptides 16-18 in ethylene glycol: Figure 70A shows the CD-thermal denaturation plot of molar ellipticity at 222-223 nm *versus* temperature in ethylene glycol for peptides **16-18**. Figure 70B shows the first

derivative curve obtained from the sigmoidal fit of data of Figure 70A and the T_m values obtained from minima of Figure 70B of peptides **16-18** is shown in Table 28.

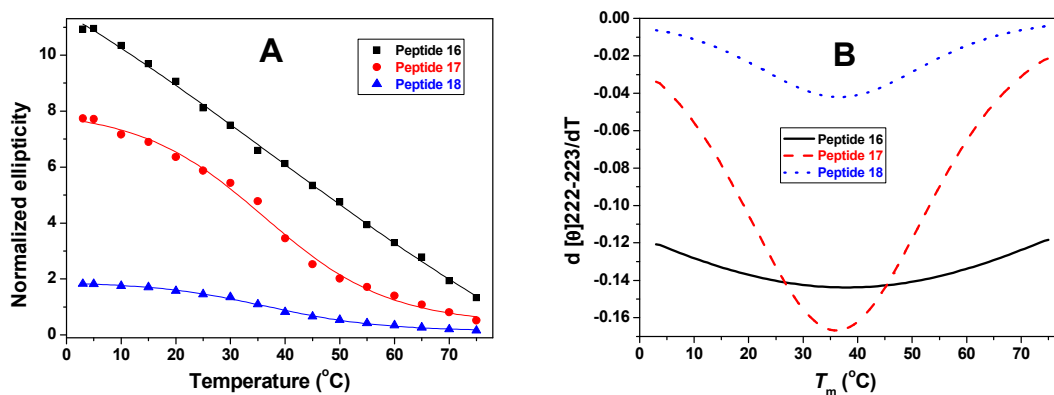


Figure 70: (A) CD-thermal denaturation plot of molar ellipticity at 222-223 nm vs. temperature in ethylene glycol for peptides **16-18**. (B) First derivative curve of peptides **16-18** in ethylene glycol.

Table 28: T_m of peptides **16-18** in ethylene glycol

Peptide	T_m	ΔT_m compared with C-terminal modified peptides
7	38	+1 (Peptide 16-13)
8	36	-1 (Peptide 17-14)
9	37	-1 (Peptide 18-15)

T_m values are (± 0.5 °C)

3.4.2c Comparison of stability of peptides 16-18 as a function of pH: A comparative study of T_m for peptides **16-18** as a function of pH is shown in Figure 71.

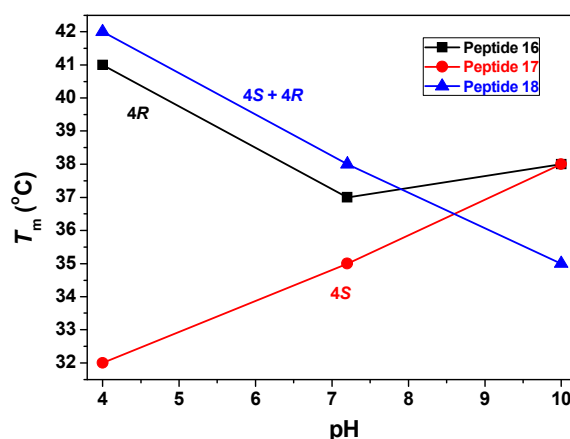


Figure 71: T_m of peptides **16-18** as a function of pH.

A comparison of T_m data (Figure 71) reveals the following: (i) Peptide **16** at pH 4.0 shows a T_m of 41 °C which was maximum and decreases to 37 °C at pH 7.2 and slightly increase to 38 °C at pH 10.0. (ii) An opposite behavior was observed for

peptide **17** which shows minimum T_m 32 °C at pH 4.0, which increases to 35 °C at pH 7.2 and 38 °C at pH 10.0. (iii) Peptide **18** at pH 4.0 shows a T_m of 42 °C which is maximum and decreases to 38 °C at and 35 °C at pH 10.0. At pH 4.0, peptide **18** shows higher triple helix stability than peptide **16**. It is surprising that at pH 10.0, peptide **16** shows less triple helix stability than peptide **17** and **18**.

3.4.3 CD Spectroscopic Studies of N-Terminal Modified Peptides 19-21

H₂N-Phe-Pro-4*R*-Amp-Gly-[Pro-Hyp-Gly]₅-NH₂ (Peptide **19**)

H₂N-Phe-Pro-4*S*-amp-Gly-[Pro-Hyp-Gly]₅-NH₂ (Peptide **20**)

H₂N-Phe-4*S*-amp-4*R*-Amp-Gly-[Pro-Hyp-Gly]₅-NH₂ (Peptide **21**)

3.4.3a Concentration dependent CD spectroscopy for peptides 19-21: Figure 72 shows the CD spectra recorded at different pHs and in ethylene glycol for peptides **19-21**.

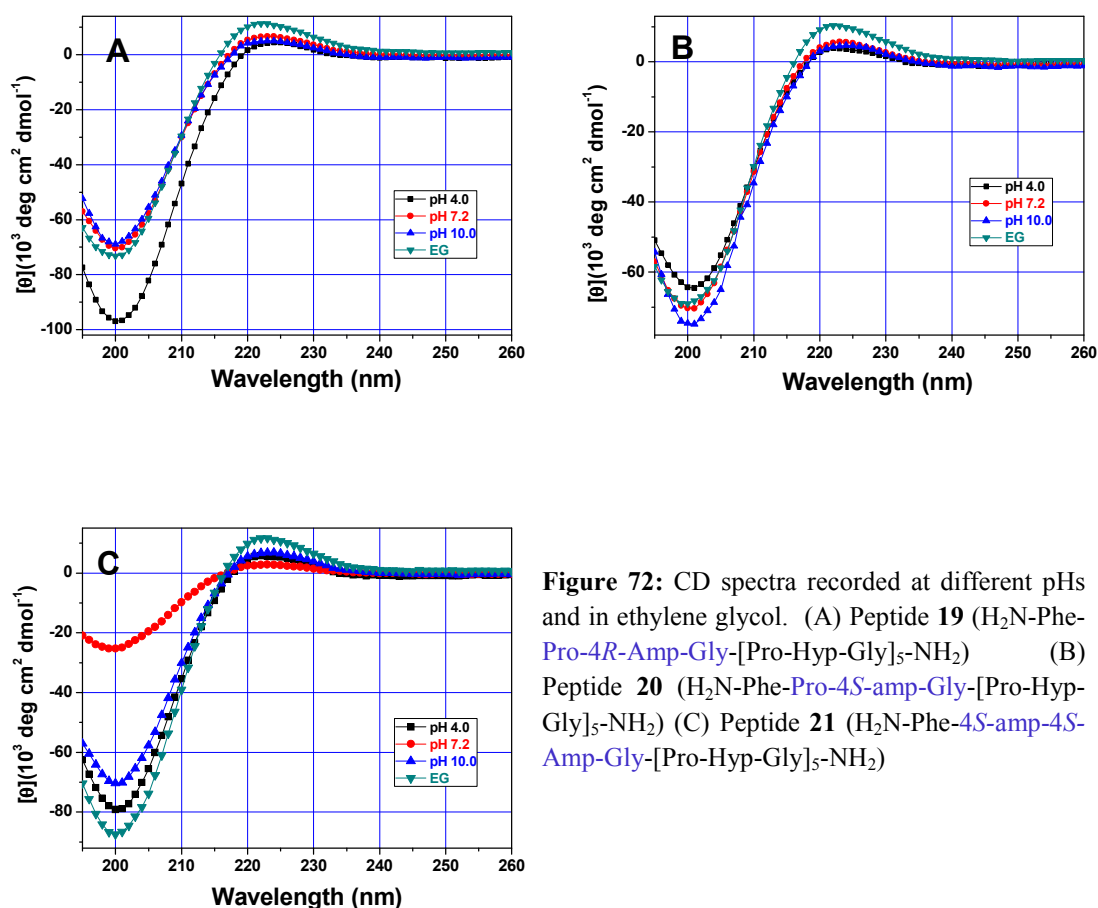


Figure 72: CD spectra recorded at different pHs and in ethylene glycol. (A) Peptide **19** (H₂N-Phe-Pro-4*R*-Amp-Gly-[Pro-Hyp-Gly]₅-NH₂) (B) Peptide **20** (H₂N-Phe-Pro-4*S*-amp-Gly-[Pro-Hyp-Gly]₅-NH₂) (C) Peptide **21** (H₂N-Phe-4*S*-amp-4*R*-Amp-Gly-[Pro-Hyp-Gly]₅-NH₂)

The peptides **19-21** show positive maxima at 222-225 nm and negative minima at 200-204 nm respectively. Importantly all the spectral traces pass through an isobestic

point at 215-218 nm. A critical triple-helical concentration is derived at 200-300 μM from these plots.

3.4.3b CD thermal denaturation study of peptides 19-21 at different pHs: In order to get information about the relative triple-helical strength of peptides **19-21**, CD-spectroscopic studies as a function of temperature was carried out.

Peptide 19 ($\text{H}_2\text{N-Phe-Pro-4R-Amp-Gly-[Pro-Hyp-Gly]}_5\text{-NH}_2$): The CD-thermal denaturation plot of molar ellipticity at 222-224 nm *versus* temperature at different pHs for peptide **19** is shown in Figure 73A. Figure 73B shows the first derivative curve obtained from the sigmoidal fit of data of Figure 73A and the T_m values for peptide **10** obtained from minima of Figure 73B at different pHs is shown in Table 29.

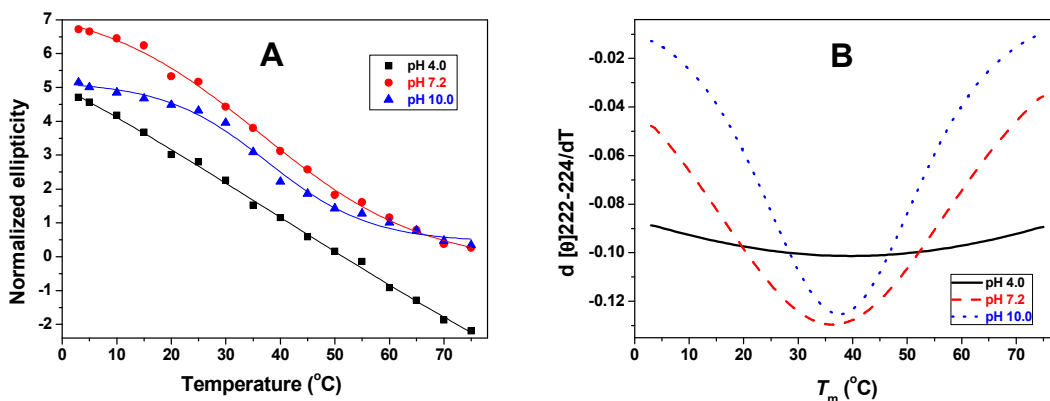


Figure 73: (A) CD-thermal denaturation plot of molar ellipticity at 222-223 nm vs. temperature at different pHs for peptide **19** ($\text{H}_2\text{N-Phe-Pro-4R-Amp-Gly-[Pro-Hyp-Gly]}_5\text{-NH}_2$). (B) First derivative curve of peptide **19** as a function of pH.

Table 29: T_m of peptide **19** $\text{H}_2\text{N-Phe-Pro-4R-Amp-Gly-[Pro-Hyp-Gly]}_5\text{-NH}_2$

pH	T_m	ΔT_m (Peptide 19-13)
4.0	40	-2
7.2	36	-2
10.0	37	-2

T_m values are (± 0.5 °C)

Peptide 20 ($\text{H}_2\text{N-Phe-Pro-4S-amp-Gly-[Pro-Hyp-Gly]}_5\text{-NH}_2$): Figure 74A shows the CD-thermal denaturation plot of molar ellipticity at 222-224 nm *versus* temperature at different pHs for peptide **20**. Figure 74B shows the first derivative curve obtained from the sigmoidal fit of data of Figure 74A and the T_m values for peptide **20** obtained

from minima of Figure 74B at different pHs is shown in Table 30. The T_m of peptide **20** increases with increase in pH.

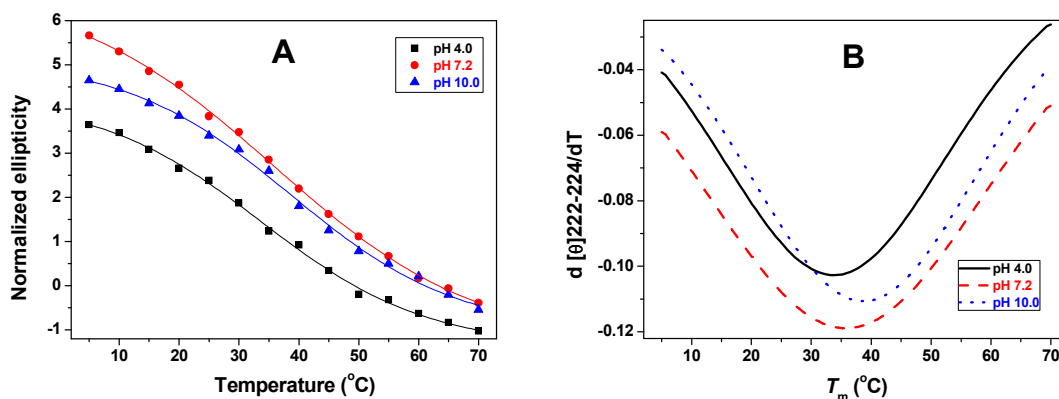


Figure 74: (A) CD-thermal denaturation plot of molar ellipticity at 222-224 nm vs. temperature at different pHs for peptide **20** ($\text{H}_2\text{N-Phe-Pro-4S-amp-Gly-[Pro-Hyp-Gly]}_5\text{-NH}_2$). (B) First derivative curve of peptide **20** as a function of pH.

Table 30: T_m of peptide **20** ($\text{H}_2\text{N-Phe-Pro-4S-amp-Gly-[Pro-Hyp-Gly]}_5\text{-NH}$)

pH	T_m	ΔT_m (Peptide 20-14)
4.0	33	+2
7.2	36	+2
10.0	39	+2

T_m values are (± 0.5 °C)

Peptide 21 ($\text{H}_2\text{N-Phe-4S-amp-4R-Amp-Gly-[Pro-Hyp-Gly]}_5\text{-NH}_2$): Figure 75A shows the CD-thermal denaturation plot of molar ellipticity at 222-224 nm *versus* temperature at different pHs for peptide **21**

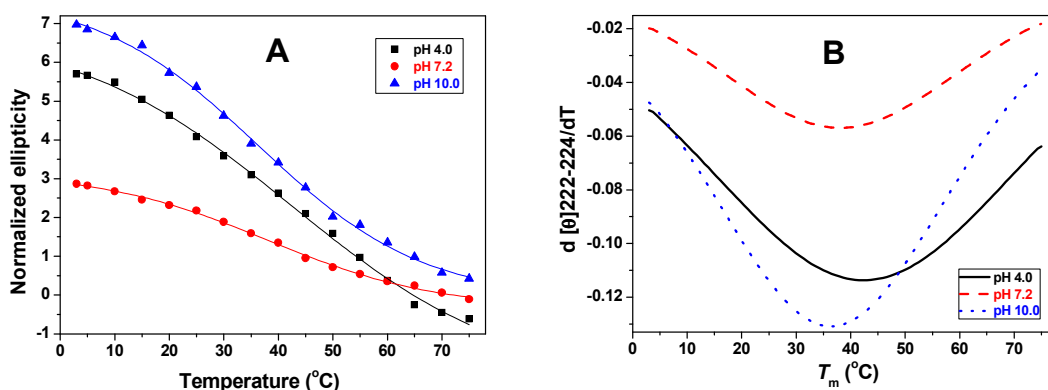


Figure 75: (A) CD-thermal denaturation plot of molar ellipticity at 222-224 nm vs. temperature at different pHs for peptide **21** ($\text{H}_2\text{N-Phe-4S-amp-4R-Amp-Gly-[Pro-Hyp-Gly]}_5\text{-NH}_2$). (B) First derivative curve of peptide **21** as a function of pH.

Figure 75B shows the first derivative curve obtained from the sigmoidal fit of data of Figure 75A and the T_m values for peptide **21** obtained from minima of Figure 7B at different pHs is shown in Table 31. The T_m of peptide **21** decreases with increase in pH.

Table 31: T_m of peptide **21** H₂N-Phe-4*S*-amp-4*R*-Amp-Gly-[Pro-Hyp-Gly]₅-NH₂

pH	T_m	ΔT_m (Peptide 21-15)
4.0	43	+2
7.2	39	+2
10.0	36	+2

T_m values are (± 0.5 °C)

3.4.3c Comparison of stability of peptides 19-21 as a function of pH: A comparison of T_m data (Figure 76) reveals the following: (i) Peptide **19** at pH 4.0 shows a T_m of 40 °C which decreases to 36 °C at pH 7.2 and slightly increases to 38 °C at pH 10.0. (ii) An opposite behavior was observed for peptide **20** which shows minimum T_m 33 °C at pH 4.0, which increases to 36 °C at pH 7.2 and 39 °C at pH 10.0. (iii) Peptide **21** at pH 4.0 shows a T_m of 43 °C which was maximum and decreases to 39 °C at pH 7.2 and 36 °C at pH 10.0. At pH 4.0, peptide **21** shows higher stability of triplex than peptide **19**.

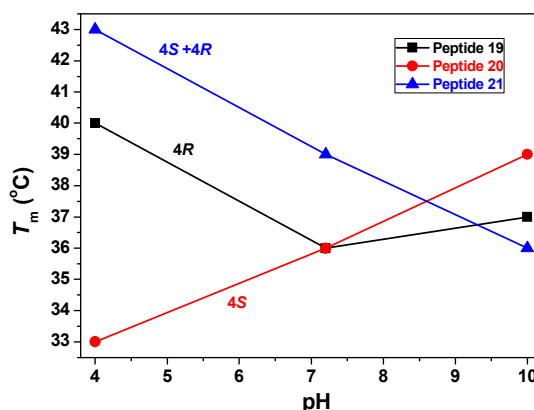


Figure 76: T_m of peptides **19-21** as a function of pH.

3.4.3d CD thermal denaturation study of peptides 19-21 in ethylene glycol: Figure 77A shows the CD-thermal denaturation plot of molar ellipticity at 222-223 nm *versus* temperature in ethylene glycol for peptides **19-21**. Figure 77B shows the first derivative curve obtained from the sigmoidal fit of data of Figure 77A and the T_m values obtained from minima of Figure 77B of peptides **19-21** is shown in Table 32.

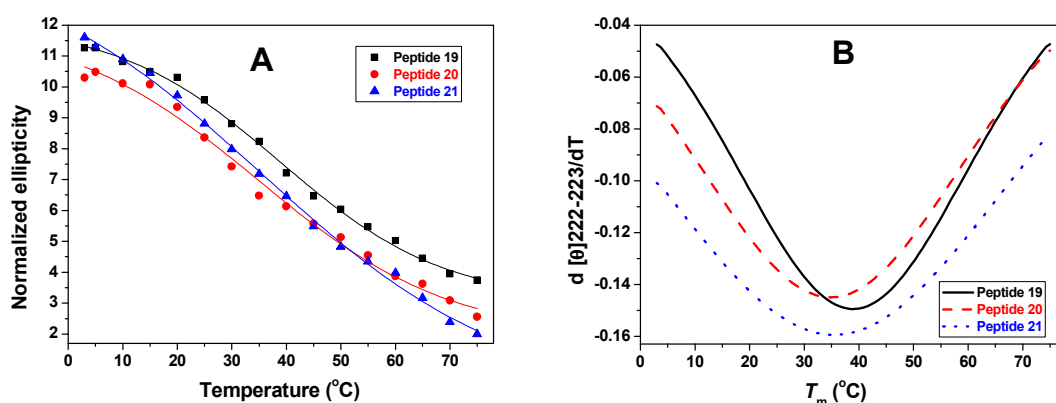


Figure 77: (A) CD-thermal denaturation plot of molar ellipticity at 222-223 nm vs. temperature in ethylene glycol for peptides **19-21**. (B) First derivative curve of peptides **19-21** in ethylene glycol.

Table 32: T_m of peptides **19-21** in ethylene glycol

Peptide	T_m	ΔT_m compared with C-terminal modified peptide
7	39	+2 (Peptide 19-13)
8	35	-2 (Peptide 20-14)
9	36	-2 (Peptide 21-15)

T_m values are (± 0.5 °C)

3.4.4a Comparative study of C-terminus, middle and N-terminus modifications at different pH: The comparative T_m values of peptides **13-21** at pH 4.0, pH 7.2 and pH 10.0 are shown in Figure 78. From the figure it is seen that when a triplet modification is done at Y position by 4*R*-aminopropyl (peptides **13**, **16** and **19**), the T_m slightly decreases from C-terminus to N-terminus modified peptides. When a triplet modification is done at Y position by 4*S*-aminopropyl (peptides **14**, **17** and **20**), the T_m slightly increases from C-terminus to N-terminus modified peptides. When a triplet modification is done at X position by 4*S*-aminopropyl and 4*R*-aminopropyl at Y position (peptides **15**, **18** and **21**), the T_m slightly increases from C-terminus to N-terminus modified peptides.

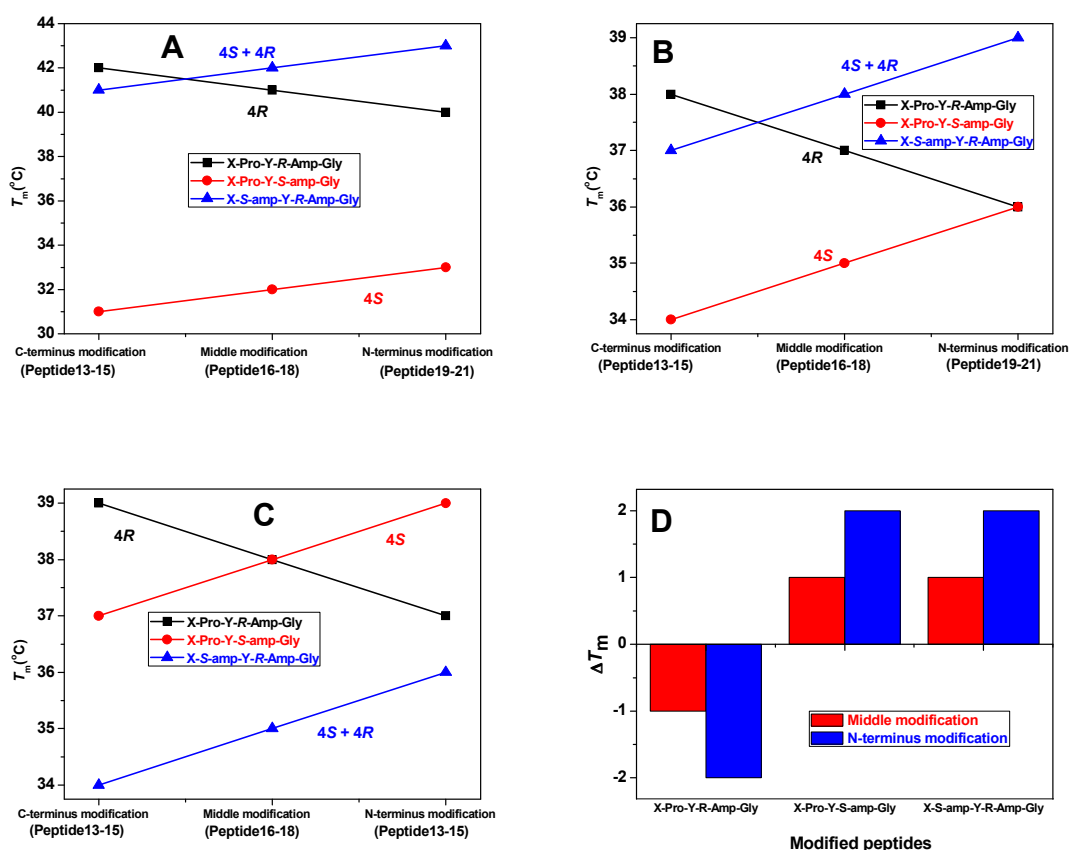


Figure 78: Comparative study of stability of non acetylated peptides 13-21 with respect to C-terminus, middle and N-terminus modification. (A) T_m at pH 4.0 (B) T_m at pH 7.2 (C) T_m at pH 10.0 (D) ΔT_m of middle modified and N-terminus modified peptides compared with C-terminus modified peptides in all pH.

3.4.4b Comparative study of C-terminus, middle and N-terminus modifications in ethylene glycol:

The comparative T_m values of peptides 13-21 in ethylene glycol are shown in Figure 79. From the figure it is seen that when a triplet modification is done at Y position by 4R-aminoprolyl (peptides 13, 16 and 19), the T_m slightly increases from C-terminus to N-terminus modified peptides. When a triplet modification is done at Y position by 4S-aminoprolyl (peptides 14, 17 and 20), the T_m slightly decreases from C-terminus to N-terminus modified peptides. When a triplet modification is done at X position by 4S-aminoprolyl and 4R-aminoprolyl at Y position (peptides 15, 18 and 21), the T_m slightly decreases from C-terminus to N-terminus modified peptides. This behavior observed in ethylene glycol is opposite to that seen in aqueous system.

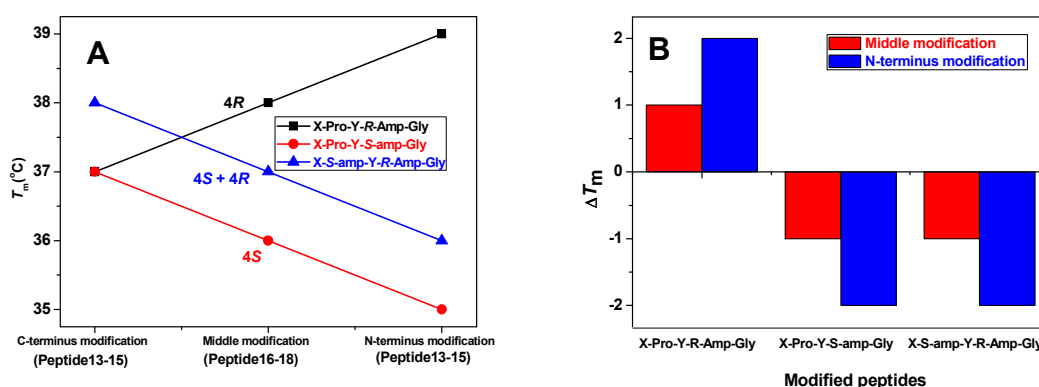


Figure 79: (A) Comparative study of stability of peptides **13-21** with respect to C-terminus, middle and N-terminus modification in EG. (B) ΔT_m of middle modified and N-terminus modified peptides compared with C-terminus modified peptides in EG.

From the CD spectroscopic studies of non-acetylated peptides (**13-21**), it is clear that the charge on terminal ends plays important role on the stability of collagen triple helix. The stereochemistry and solvent are also important for elucidating the stability of collagen triple helix having only one modified triplet. The modified sites (C-terminal, middle, N-terminal) have very little effect on stability of collagen triple helix.

4 Discussion

The *4R/S*-aminoproline containing peptides **4-21** form a polyproline II like structure at all pHs as indicated by a weak positive band between 222-225 nm and a strong negative band between 200-205 nm in the CD spectra.

In the concentration dependent triple-helix formation study, the peak positions are nearly independent of the concentration and all the spectral traces pass through an isobestic point between 212-218 nm for peptides **4-21**. This CD spectral feature indicates that within the concentration range measured *i.e.* 50-500 μM , all the peptides **4-21** exist in the triple-helical conformation. The smooth and progressive increase in the $R_{p/n}$ values with concentration also suggests a two-state equilibrium, which rapidly moves towards saturation to a single-state with increase in the concentration. The observed sigmoidal transitions in the CD thermal denaturation studies provide additional evidence for the two state (triple-helix to coil) transition.

The T_m value of non-ionizable peptide **22** Ac-Phe-(Pro-Hyp-Gly)₆ is irreverent to changes in pH while the T_m of Y site modified ionizable peptide **23** Ac-Phe-(Pro-4R-

Amp-Gly)₆ changes with pH. The T_m values of peptide **23** Ac-Phe-(Pro-4*R*-Amp-Gly)₆ decreases with an increase in pH from 3.0-10.0 but increase at pH 12 (Table 7). Since the pK_a of 4-NH₂ group of 4*R*-aminoproline is 10.2 (Chapter 1), the 4*R*-Amp exists in fully protonated (NH₃⁺) form at pH 3.0, pH 7.0. At pH 12.0 the 4*R*-Amp exists in the free amine (NH₂) form. At pH 10.0 which is very close to pK_a of 4*R*-Amp, only partial protonation of 4-NH₂ group may occur leading to decrease in stability of triple-helix.

The Y position modified peptide **24** Ac-Phe-(Pro-4*S*-amp-Gly)₆ does not form triple helix since it adopts a C4-*endo* puckering of pyrrolidene ring, while 4*R*-Amp adopts *exo* puckering that favoured triplex formation in peptide **23**. This stereoselectivity favouring the *R* stereoisomer over *S* is proposed to arise from the overlap of an electron pair with the antibonding σ^* of an adjacent polar bond (C-X, where X=N or O).¹⁵³ In literature it has been well demonstrated that the stereoelectronic effect of the 4-substituent plays a major role in determining the pucker of the pyrrolidine ring of proline and hence the conformational stability of proteins.¹⁵⁴

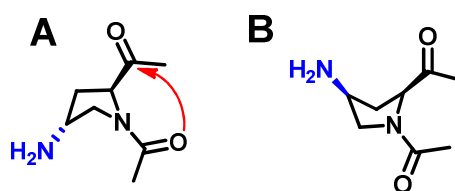


Figure 80: An $n \rightarrow \pi^*$ interaction stabilizes the *trans* isomer of the peptide bond (A) Substantial only when Pro derivatives are in the C ^{γ} -*exo* ring pucker (*R* stereochemistry) (B) When Pro derivatives are in C ^{γ} -*endo* ring pucker (*S* stereochemistry) $n \rightarrow \pi^*$ interaction are not operative.

The T_m values of X and Y site modified chimeric peptide **25** (Ac-Phe-(4*S*-amp-4*R*-Amp-Gly)₆) decrease with increase in pH and the cause of the stabilization is the C4-*endo* pyrrolidene conformation adopted by 4*S*-Amp that is inherently favoured at the X-position. The pH dependent stabilities in both 4*R* and 4*S*-aminoprolines also suggest that protonation of NH₂ is a prerequisite for triplex formation in the X-position, while it is not so in the Y-position.

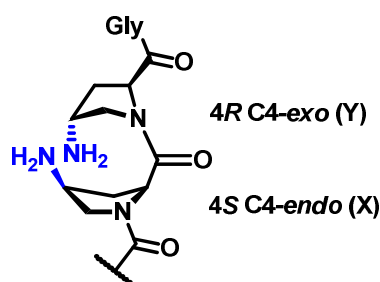


Figure 81: Position dependent preferred proline puckers in [4*S*-amp-4*R*-Amp-Gly] peptide.

It is possible that the conformation of pyrrolidine ring may also depend on the protonation status of the 4-amino group and influence the *cis*–*trans* amide rotameric equilibrium.¹⁵³

The 4*R*-aminoproline at Y position (peptides **4**, **7** and **10**) having single modified triplet shows enhanced stability at all pHs compared to the analogous collagen model peptide **22** Ac-Phe-(Pro-Hyp-Gly)₆ (Figure 43). This suggests that the 4*R*-NH₂ group in both protonated and unprotonated forms have a positive role in stabilizing the structure of collagen peptide. The increase in T_m for single 4*R*-amino modified triplet (peptide **4**, **7** and **10**) at all pHs is nearby half as compared with full modified peptide **23** Ac-Phe-(Pro-4*R*-Amp-Gly)₆ at all pHs (Figure 43). This suggests that even one 4*R*-NH₂ modification has a significant positive role in stabilizing the structure of collagen peptide. The QM calculation also confirms that *endo*-*exo* pyrrolidine ring arrangement respectively at X and Y positions is favorable for tight packing of triplex structure.¹⁵⁵ In addition, the ability of *exo* puckered 4*R*-Amp residue to prefer *trans* conformation of peptide bond¹⁵⁶ favors the stabilization of triplex structure. The T_m of peptides **4**, **7** and **10**, decreases with increase in pH. The higher stability of these peptides at lower pH could arise from an increased electronegativity and the hydrogen bonding potential of the protonated amine moiety. At higher pH, the stabilizing effect may be a consequence of hydrogen bonding and the absence of electrostatic repulsion in nonprotonated 4-NH₂ group.

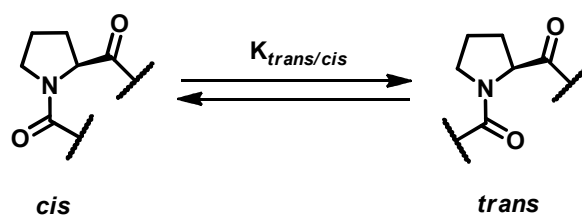


Figure 82: *trans* isomerization favoured the collagen triple helix.

The 4*S*-aminoproline at Y position (peptides **5**, **8** and **11**) having single modified triplet show slight decrease in stability at all pHs compared to the analogous collagen model peptide **22** Ac-Phe-(Pro-Hyp-Gly)₆ (Figure 43), while the 4*S*-aminoproline peptide **24** (Ac-Phe-(Pro-4*S*-amp-Gly)₆) does not form triple helix. This is due to fact that the all 4*R*-hydroxyl groups adopt C4-*exo* puckering, while the 4*S*-amp adopts C4-*endo* puckering leading to small decrease in collagen triple helical structure. The T_m of peptides **5**, **8** and **11** increases with increase in pH, suggests that the protonation of

NH₂ is important for triple helix structure. In literature, it was demonstrated that the ionizable 4-aminoproline at acidic pH in protonated form (NH₃⁺) influences the proline puckering *via* its electron withdrawing effects.¹⁴² At acidic pH 4*S*-amino group gets protonated (4*S*-NH₃⁺) and the accompanying unfavourable stereoelectronic effect (unlike 4*R*-NH₃⁺) leads to low stability of collagen triplex. At basic pH, the weaker stereoelectronic effect of 4*S*-NH₂ does not influence much on the stability of collagen triplex.

When a single triplet modification is done at X position by 4*S*-aminoprolyl and 4*R*-aminoprolyl at Y position (peptides **6**, **9** and **12**), these peptides show slight decrease in stability at all pHs compared to the analogous collagen model peptide **22** Ac-Phe-(Pro-Hyp-Gly)₆ (Figure 43). This is due to the preferred puckering simultaneously adopted by 4*S*-amp and 4*R*-Amp at X and Y positions in accordance with the two prolines in natural collagen (Pro-Hyp-Gly)_n sequence. This helps in close packing of individual chains to form a stable triplex-helical structure resulting in a higher thermal stability of chimeric peptides. The *T_m* values of peptides **6**, **9** and **12** decreases with increase in pH and the cause of the stabilization is the *C4-endo* pyrrolidene conformation adopted by 4*S*-amp that is inherently favoured at the X-position. The pH dependent stabilities in both 4*R*/4*S*-aminoprolines also suggest that protonation of NH₂ is a prerequisite for triplex formation in the X-position, while it is not so in the Y-position. It is possible that the conformation of pyrrolidine ring may depend on the protonation status of the 4-amino group and influence the *cis-trans* amide rotameric equilibrium. The change of ring pucker in going from acidic to basic pH for aminoproline may not be mutually compatible for peptides **6**, **9** and **12** which leads to a decrease in thermal stability of triple-helix at higher pH.

In ethylene glycol, the peptide **22** Ac-Phe-(Pro-Hyp-Gly)₆ forms more stable triple helix (*T_m* = 44 °C) than in buffer system. The most likely mechanism for stabilization of a collagen-like triple-helical structure involves the correlation of hydrogen-bonding strengths of the solvents with the increase in ordered structure. The weaker hydrogen bonding solvents provide a more favorable environment in which the chains associate through interchain hydrogen bonds required for stabilization of the triple-helical collagen-like conformation.

However peptides **23** Ac-Phe-(Pro-4*R*-Amp-Gly)₆, *T_m* = 23 °C and **25** Ac-Phe-(4*S*-amp-4*R*-Amp-Gly)₆, *T_m* = 37 °C show lower stability than the peptide **22**. This reversal

of stability in ethylene glycol arises from the poor solvation of the EG for the charged and polar amino residue present at Y position in peptides **23** and at X and Y positions in peptide **25**. The additional interstrand repulsion arising from the charge on 4-NH₃⁺ groups in peptide **23** and **25** remains unscreened in less polar ethylene glycol. This repulsive interaction aggravated by the absence of salt may result in decreased thermal stability of triple-helix.

Peptides **4-12** show slightly low stability compared to that peptide **22** Ac-Phe-(Pro-Hyp-Gly)₆ in ethylene glycol. This may due to the presence of only single amino modified triplet in the peptides **4-12**, result in only slight change, while the fully modified peptides show the opposite effect in ethylene glycol.

The non acetylated peptides **13-21** behave similarly as peptides **4-12** with some low T_m (Figure 65). This observed decrease in stability of uncapped peptides is due to presence of similar charged residues at termini in a parallel triplex that leads to repulsions and results in destabilization. Collagen being a parallel triple-helix with one residue shift, these NH₃⁺ groups lie in close proximity to each other which results in large electrostatic repulsive interactions between the chains. The terminal effects may be nulled by end capping, which significantly stabilizes the collagen triplex helix.

The single triplet 4*R*-aminoprolyl modified peptides (**13**, **16** and **19**) shows higher triplex stabilities at either extremes of pHs (4.0 and 10.0) compared to that at intermediate pH (7.2), which arise from the N- and C-terminal effects.¹⁵⁷ At extremes pH, only one of the terminal groups is completely ionized (NH₃⁺/COOH at pH 4.0 and COO⁻/NH₂ at pH 10.0), while at intermediate pH values both groups are ionized to varying degrees. The presence of similar charged residues at termini in a parallel triplex leads to its destabilization (Figure 83).

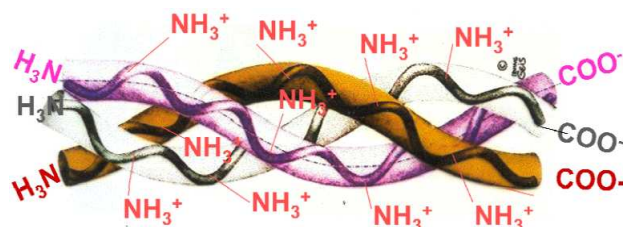


Figure 83: Effect of terminal groups on collagen triple helix.

The single triplet 4*S*-aminoprolyl modified peptides (**14**, **17** and **20**) show increase in collagen stability with increase in pH. At acidic pH, 4*S*-amino group gets protonated to 4*S*-NH₃⁺ and the predominant unfavourable stereoelectronic effect leads to lower stability of collagen triplex. While at basic pH, the weaker stereoelectronic effect of 4*S*-NH₂ does not influence much on the stability of collagen triplex.

The chimeric peptides (**15**, **18** and **21**) with single triplet modification (4*S*-aminoprolyl at X position and 4*R*-aminoprolyl at Y position) show decrease in collagen stability with increase in pH. At acidic pH, 4*S* + 4*R*-amino group get protonated and the predominant favourable stereoelectronic effects with C-4 *endo* pucker at X position and C-4 *exo* pucker at Y position leads to higher stability of collagen triple helix.

The triplex stabilities of single triplet 4*R*-aminoprolyl modified peptides (**4**, **7** and **10**) slightly decrease ($\Delta T_m = -1-2$ °C) from C-terminus to N- terminus modification at different pHs (Figure 58). The triplex stabilities of 4*S*-aminoproline (**5**, **8** and **11**) and the 4*S* + 4*R*-aminoproline containing chimeric peptides (**6**, **9** and **12**) slightly increase ($\Delta T_m = +1-2$ °C) from C-terminus to N- terminus modification at different pHs. In ethylene glycol, an opposite behavior was observed for peptides **4-21**. The difference observed $\Delta T_m \approx 1-2$ is insufficient for any conclusion since may arise from solvation of peptide backbone, depending on the stereochemistry and puckering of proline ring.

5 Conclusion

In conclusion, it is shown that replacement of even a single 4-OH prolyl residue by 4*R*-NH₂ has significant positive role in stabilizing the triplex structure of collagen to nearby half the magnitude of stability seen in full modification at all pHs. The higher stability at lower pH could arise from increased electronegativity and the hydrogen bonding potential of the protonated amine moiety. At higher pH, the stabilizing effect may be a consequence of hydrogen bonding and the absence of electrostatic repulsion in nonprotonated 4-NH₂ group. The single 4*S*-NH₂ modification shows a small decrease in stability compared to collagen model peptide at all pHs. The stability of 4*S*-NH₂ peptides increase with increase in pH and this is perhaps due to unfavourable stereoelectronic effect of 4*S*-NH₃⁺ group at acidic pH. The single 4*S* + 4*R*-amino chimeric peptides show enhanced collagen triple helix stability, this is due to the preferred puckering adopted by 4*S*-amp/4*R*-Amp at X and Y positions in accordance

with the two prolines. In ethylene glycol, the modified peptides show comparable stability with control collagen model peptide due to presence of only one amino modified triplet. The non acetylated peptides behave similarly as acetylated peptides with some lower stability. The observed decrease in stability of uncapped peptides is due to presence of similar charged residues at termini in a parallel triplex that leads to repulsions and results in destabilization. The substitution at different position (N/C termini or middle modification) does not have any effect on collagen triple helix structure. The observed little difference in stability due to solvation of peptide backbone also depends on stereochemistry and puckering of proline ring.

The results reinforce the important role of amino group for tuning the required stability of modified collagen peptides. Future potential of this work lies in rationally combining the stabilities offered to design collagens analogues or chimeric collagens of unusual stability for applications in collagen based biomaterials.¹⁵⁸ It is also possible to prepare controlled collagen crosslink peptides by using only single amino group present in these modified peptides for biological application. The results in this work also have implications for peptidomimetics design and the control of peptide conformations through 4-substituent effects on proline.

6 Experimental

Materials and reagents were of the highest commercially available grade and used without further purification. Reactions monitored by thin layer chromatography (TLC) were carried out on precoated silica gel 60 F254 plates (E. Merck). TLCs were visualized under UV light, iodine and/or ninhydrin spray followed by heating up to 110 °C with heat gun. Silica gel 60-120 and 100-200 mesh (Merck) was used for routine column chromatography. Elution was done with ethyl acetate/petroleum ether or dichloromethane/methanol mixture depending upon the compound polarity and chemical nature. Flash chromatography was done using 230-400 mesh silica gel. All solvents were distilled under an inert atmosphere with appropriate desiccant. Reactions in aqueous medium and workup processes were done using double distilled water. Unless otherwise noted, all reactions were carried out at room temperature. IR spectra were recorded on an Infrared Fourier Transform Spectrophotometer using chloroform, nujol or neat sample. ¹H and ¹³C NMR spectra were recorded on Bruker AV 500 and AV 400, AV 200 spectrometers. Chemical shifts are reported in ppm using TMS and

CDCl₃ as a reference. Spectra were analyzed using ACD spectviewer software from ACD labs. NMR spectra of compounds show two sets of peaks arising from *cis-trans* isomerization of N-CO bond. Analytical HPLC was performed using a LiChrospher 100 RP-18e 5 μM (250 mm x 4 mm) column from Merck. Preparative HPLC was carried out on a LiChrospher RP-18e 5 μM (250 mm x 10 mm) column from Merck. Mass spectra were obtained by ESI-MS technique on AP-QSTAR spectrometer. Jasco J-715 (Japan) instrument was used for CD measurements. All graphs presented for CD spectra are drawn by Origin 8 software.

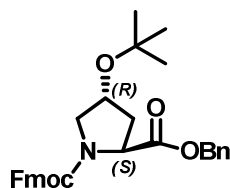
Reagents for the buffer preparation such as NaCl, Na₂BO₄, NaH₂PO₄, Na₂HPO₄, Citric acid and Na-Borate were obtained from Sigma and were of Molecular biology reagent grade. Double distilled water was demineralised using Millipore MilliQ deionizer and was used for the preparation of buffers. pH was adjusted using NaOH and HCl in case of borate buffer and using acetic acid for acetate buffer.

Resins for solid phase peptide synthesis and Fmoc-protected amino acids were bought from Novabiochem and were used without further purification.

The term “concentrated under reduced pressure” refers to the removal of solvents and other volatile materials using a rotary evaporator at water aspirator pressure (<20 Torr) while maintaining the water-bath temperature below 40 °C. Residual solvent was removed from samples at high vacuum (<0.1 Torr). The term “high vacuum” refers to vacuum achieved by a mechanical belt-drive oil pump.

6.1 Synthesis of compounds 15-41

(2*S*, 4*R*)-N1-Fmoc-4*R*-O-*t*-Bu-hydroxyproline benzyl ester (15)

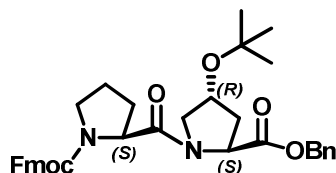


This compound was prepared by known literature procedure.¹⁴⁹ A mixture of N1-Fmoc-*trans*-4-O-*tert*-butyl-hydroxy-L-proline **1** (5 g, 12.2 mmol) and K₂CO₃ (5.6 g, 36.6 mmol) was stirred in anhydrous DMF (30 mL) under nitrogen atmosphere for 10 min at 0 °C. Addition of benzyl bromide (1.6 mL, 13.4 mmol) was done drop-wise in time period of 10 min. The reaction mixture was stirred for 2 h at 0 °C temperature. After completion of reaction water (250 mL) was added to remove DMF and the

aqueous layer was extracted with ethyl acetate (3 x 70 mL). The combined organic layer was washed with water (3 x 50 mL) followed by brine, dried over anhydrous Na_2SO_4 and concentrated under vacuum. The crude product obtained was purified by silica gel chromatography (20% ethyl acetate/hexane) to afford compound **15** as colourless thick oil. Yield 5.61 g, (92%). Molecular Formula, $\text{C}_{31}\text{H}_{33}\text{NO}_5$. [Mass: calculated 499.6; observed 522.30 ($\text{M} + \text{Na}^+$)].

^1H NMR (CDCl_3 , 200 MHz) δ_{H} : 1.16-1.18 (9H, d), 2.04-2.21 (3H, s), 3.32-3.40 (1H, m), 3.77-3.82 (1H, m), 4.10-4.40 (4H, m), 4.51 (1H, m), 5.10-5.20 (2H, m), 7.24-7.37 (9H, m), 7.5-7.62 (2H, m), 7.75-7.78 (2H, m).

N1-Fmoc-proline-4*R*-O-*t*Bu-proline benzyl ester dipeptide (**18**)



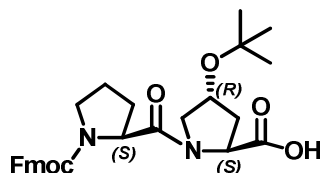
The compound **15** (5.5 g, 11 mmol) was dissolved in anhydrous DCM (6 mL) and stirred at 0 °C for 5 min. The addition of diethylamine (7 mL, 66 mmol) was done over 10 min and the reaction mixture was stirred for 1.5 h. After completion of reaction, the resulting reaction mixture was concentrated under vacuum. Silica gel column was run to elute the corresponding free amine compound **16** which was dissolved in anhydrous DMF at 0 °C. Simultaneously a mixture of N1-Fmoc-Proline **17** (4.1 g, 12 mmol), HOBt (1.63 g, 12 mmol) and EDC (2.3 g, 12 mmol) in anhydrous DMF (30 mL) was stirred at 0 °C for 15 min under nitrogen atmosphere. The free amine compound **16** in anhydrous DMF was slowly added to the reaction mixture over duration of 15 min. The reaction mixture was stirred for 2 h at 0 °C and then at ambient temperature for 4 h. After completion of reaction (monitored by TLC) water (100 mL) was added and the aqueous layer was extracted with ethyl acetate (3 x 80 mL). The combined organic layer was washed with NaHCO_3 (50 mL) to remove the unreacted free acid followed by washing with KHSO_4 (50 mL) to remove the unreacted free amine. The combined organic layers was washed with NaHCO_3 (50 mL) followed by water (3 x 50 mL) and brine. The organic layer was dried over anhydrous Na_2SO_4 and concentrated under vacuum. The crude product obtained was purified by silica gel chromatography (35% ethyl acetate/hexane) to afford dipeptide **18** as a white solid. Yield 7.2 g, (90%).

Molecular Formula, C₃₆H₄₀N₂O₆. [Mass: calculated 596.71; observed 619.38 (M + Na⁺)].

¹H NMR (CDCl₃, 400 MHz) δ_H: 1.09-1.16 (9H, d), 1.95-1.97 (2H, m), 1.98-2.02 (1H, m), 2.07-2.12 (2H, m), 2.13-2.17 (2H, m), 3.51-3.60 (2H, m), 3.73-3.79 (1H, m), 4.20-4.27 (1H, m), 4.29-4.35 (2H, m), 4.37-4.44 (1H, m), 4.51-4.59 (2H, m), 4.73-4.76 (1H, m), 5.08-5.25 (2H, m), 7.30-7.42 (9H, m), 7.57-7.67 (2H, m), 7.76 (2H, d, *J* = 7.53).

¹³C NMR: 23.1, 24.2, 28.2, 29.0, 30.0, 36.3, 46.8, 47.3, 47.4, 52.8, 53.0, 57.6, 57.6, 58.1, 66.9, 66.9, 67.0, 67.4, 69.4, 69.5, 119.9, 125.2, 127.0, 127.6, 127.6, 128.1, 128.6, 135.7, 141.3, 143.9, 144.3, 144.5, 154.9, 170.8, 170.9, 172.0, 172.1.

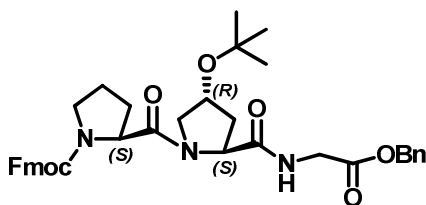
N1-Fmoc-proline-4*R*-O-*t*Bu-proline acid dipeptide (**19**)



The dipeptide **18** (6 g, 10 mmol) was carefully soluble in methanol (15 mL) and subjected to hydrogenation to remove benzyl ester by using 10% Pd/C (0.6 g). The mixture was shaken in Parr hydrogenation apparatus under H₂ gas (55 psi pressure) for 2-2.5 h. Careful monitoring by TLC was necessary to prevent hydrogenolysis of the Fmoc group. The suspension was filtered through a pad of Celite and concentrated under reduced pressure. The crude product obtained was purified by silica gel chromatography (80% ethyl acetate/hexane) to afford dipeptide acid **19** as a white solid powder. Yield 4.44 g, (87%). Molecular Formula, C₂₉H₃₄N₂O₆. [Mass: calculated 506.69; observed 529.58 (M + Na⁺)].

¹H NMR (CDCl₃, 400 MHz) δ_H: 1.12-1.19 (9H, t), 1.83 (1H, m), 1.94-2.05 (2H, m), 2.07-2.13 (1H, m), 2.15-2.25 (2H, m), 2.29-2.34 (1H, m), 3.49-3.62 (2H, m), 3.68-3.74 (2H, m), 4.20 (1H, m), 4.28 (1H, m), 4.30 (1H, d), 4.33-4.38 (1H, m), 4.40-4.42 (1H, m), 4.44-4.47 (1H, m), 4.50-4.54 (1H, m), 6.39 (1H, bs), 7.29-7.42 (4H, m), 7.52-7.64 (2H, m), 7.75 (2H, s, *J* = 7.79).

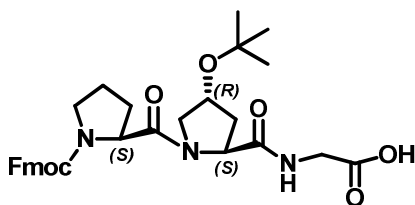
¹³C NMR: 23.2, 24.3, 27.3, 29.1, 29.8, 30.1, 35.6, 46.9, 47.2, 47.4, 53.2, 57.7, 58.2, 58.4, 67.1, 67.6, 69.5, 69.6, 74.3, 74.4, 119.9, 120.0, 124.9, 125.2, 125.4, 127.1, 127.7, 141.2, 141.4, 143.9, 144.4, 155.1, 155.4, 172.2, 172.7, 174.0, 174.2.

N1-Fmoc-proline-4R-O-*t*Bu-proline-glycine benzyl ester tripeptide (21)

A mixture of dipeptide acid **19** (4 g, 7 mmol), HOBt (1.17 g, 8 mmol) and EDC (1.66 g, 8 mmol) in anhydrous DMF (20 mL) was stirred at 0 °C for 15 min under nitrogen atmosphere. The free amine glycine benzyl ester **20** (2.31 g, 14 mmol) soluble in anhydrous DMF was slowly added to the reaction mixture during 15 min. The reaction mixture was stirred for 2 h at 0 °C and then at ambient temperature for 4 h. After completion of reaction, water (100 mL) was added and the aqueous layer was extracted with ethyl acetate (3 x 70 mL). The combined organic layer was washed with NaHCO₃ (40 mL) to remove unreacted free acid followed by washing with KHSO₄ (40 mL) to remove unreacted free amine. The combined organic layer again was washed with NaHCO₃ (40 mL) followed by water (3 x 40 mL) and brine. The organic layer dried over anhydrous Na₂SO₄ and concentrated under vacuum. The crude product obtained was purified by silica gel chromatography (65% ethyl acetate/hexane) to afford tripeptide the **21** as a white solid. Yield 3.66 g, (71%). Molecular Formula, C₃₈H₄₃N₃O₇. [Mass: calculated 653.76; observed 676.58 (M + Na⁺)].

¹H NMR (CDCl₃, 400 MHz) δ_H: 1.18-1.21 (9H, t), 1.75-1.90 (2H, m), 1.93-2.05 (2H, m), 2.12-2.18 (2H, m), 2.44 (1H, m), 2.47-2.62 (1H, m), 3.48-3.60 (2H, m), 3.64-3.72 (2H, m), 3.84-3.92 (1H, m), 4.01-4.05 (1H, dd, *J* = 5.77, 6.27), 4.08-4.16 (1H, m), 4.22 (1H, t), 4.37-4.43 (2H, m), 4.48-4.59 (2H, m), 4.61 & 4.80 (1H, dd, *J* = 8.28, 8.03), 4.98-5.19 (2H, m), 7.22 (1H, m), 7.30-7.43 (8H, m), 7.54-7.66 (2H, m), 7.77 (2H, d, *J* = 7.28), 8.30 (1H, t, *J* = 5.27, 5.23).

¹³C NMR: 24.4, 28.2, 28.3, 29.2, 29.7, 30.3, 34.5, 41.2, 46.8, 47.1, 47.3, 52.8, 57.8, 58.4, 66.9, 67.0, 67.4, 67.5, 70.2, 74.1, 120.0, 120.0, 125.1, 125.3, 127.1, 127.7, 127.7, 128.1, 128.3, 128.5, 128.6, 141.3, 143.9, 144.2, 154.9, 155.3, 169.2, 169.5, 171.5, 171.7, 172.5.

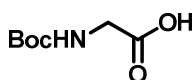
N1-Fmoc-proline-4*R*-O-*t*Bu-proline-glycine acid tripeptide (22)

The tripeptide benzyl ester **21** (3.5 g, 5 mmol) was dissolved in methanol (7-8 mL) and subjected to hydrogenation to remove benzyl ester by using 0.35 g of 10% Pd/C. The mixture was shaken in Parr hydrogenation apparatus under H₂ gas (55 psi pressure) for 2-2.5 h. Careful monitoring by TLC was necessary to prevent hydrogenolysis of the Fmoc group. The suspension was filtered through a pad of Celite and concentrated under reduced pressure. The crude product obtained was purified by silica gel chromatography (85-90% ethyl acetate/hexane) to afford dipeptide free acid **19** as white solid powder. Yield 2.44 g, (81%). Molecular Formula, C₃₁H₃₇N₃O₇. [Mass: calculated 563.64; observed 586.51 (M + Na⁺)].

¹H NMR (CDCl₃, 400 MHz) δ_H: 1.19-1.22 (9H, t), 1.73-1.87 (2H, m), 1.95-2.07 (2H, m), 2.10-2.16 (2H, m), 2.50-2.64 (1H, m), 3.52-3.64 (2H, m), 3.67-3.75 (2H, m), 3.81-3.87 (1H, m), 4.03-4.05 (1H, m), 4.11-4.19 (1H, m), 4.25 (1H, t), 4.39-4.45 (2H, m), 4.5-4.59 (2H, m), 4.61 & 4.80 (1H, dd, *J* = 8.28, 8.03), 7.20-7.43 (4H, m), 7.51-7.62 (2H, m), 7.85 (2H, d).

¹³C NMR: 24.0, 28.6, 28.9, 29.4, 29.8, 30.4, 34.4, 41.3, 46.7, 47.3, 47.3, 52.7, 57.8, 58.4, 67.4, 67.5, 70.2, 74.0, 119.9, 120.0, 125.2, 127.1, 127.8, 128.6, 141.3, 143.9, 144.1, 155.0, 155.4, 169.2, 169.4, 171.5, 171.7, 172.5.

Compound **20**, **43** and **44** were synthesized by known literature procedure.¹⁴⁹

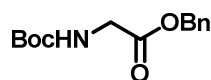
N-Boc glycine (43)

The mixture of glycine (50 g, 66 mmol) soluble in water (250 mL), dioxane (200 mL) and NaOH solution (40 g in 300 mL of water) was stirred at 0 °C. Boc anhydride (175 g, 0.799 mmol) was added drop wise during 30 min at 0 °C. The reaction was stirred for 2.5 h, the reaction was monitored by TLC 20% MeOH in 80% CHCl₃. After completion of reaction, the mixture was concentrated under reduced pressure at 80 °C

to 250 mL. To the reaction mixture ethyl acetate (200 mL) was added. The mixture was acidified with 20% KHSO₄ to pH 2. The organic layer was separated and the aqueous layer was extracted with ethyl acetate (2 x 100 mL). The combined organic layer was dried over anhydrous Na₂SO₄ and concentrated under vacuum. The product obtained was Boc glycine **43** as white solid. Yield 113 g, (97%). Molecular Formula, C₇H₁₃NO₄. [Mass: calculated 175.18; observed 198.23 (M + Na⁺)].

¹H NMR (CDCl₃, 200 MHz) δ_H: 1.44-1.45 (9H, s), 3.96 (2H, d), 5.10 (1H, s), 7.99 (1H, s).

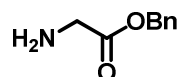
N-Boc glycine benzyl ester (**44**)



A mixture of N-Boc glycine **43** (10 g, 57.1 mmol) and 23.65 g (171 mmol) of K₂CO₃ was stirred in anhydrous DMF (80 mL) under nitrogen atmosphere for 10 min at 0 °C. Addition of benzyl bromide (7.5 mL, 62.8 mmol) was done drop-wise in time period of 10 min. The reaction mixture was stirred for 2 h at 0 °C temperature. After completion of reaction water (250 mL) was added to remove DMF and the aqueous layer was extracted with ethyl acetate (3 x 100 mL). The combined organic layer was washed with water (3 x 100 mL) followed by brine, dried over anhydrous Na₂SO₄ and concentrated under vacuum. The crude product obtained was purified by silica gel chromatography (15% ethyl acetate/hexane) to afford **44** as white crystals. Yield 14.38 g, (95%). Molecular Formula, C₁₄H₁₉NO₄. [Mass: calculated 265.30; observed 265.41 (M⁺)].

¹H NMR (CDCl₃, 200 MHz) δ_H: 1.42-1.52 (9H, d), 3.92 (2H, d), 4.51 (1H, s) 5.17 (2H, s), 7.30-7.39 (5H, s).

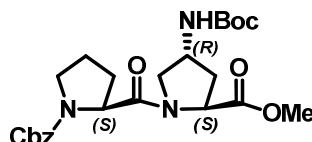
Glycine benzyl ester (**20**)



A mixture of N-Boc glycine benzyl ester **44** (5 g, 18 mmol) soluble in 7-8 mL DCM and trifluoroacetic acid (8.7 mL, 113 mmol) was stirred at 0 °C for 2.5 h. After completion of reaction DCM (25 mL) was added and the reaction mixture was stirred at ice cold condition. The diluted liquid ammonia was added very slowly till pH

became 11.0-11.5. The mixture was extracted with DCM (3 x 25 mL), the combine organic layer dried over anhydrous Na₂SO₄ and concentrated under vacuum. The crude product obtained **20** was directly used for reaction in very short time span.

N1-Cbz-proline-4R-NH-*t*-boc-proline methyl ester dipeptide (**25**)



Methanol was carefully added to a mixture of N1Cbz-4R-Aminoproline methyl ester **7**^{*} (5 g, 13 mmol) and 10% Pd/C (0.5 g). The resulting black suspension was shaken in Parr hydrogenation apparatus under H₂ gas (65 psi pressure) for 6-7 h. After completion of the reaction, the suspension was filtered through Whatman paper. The resulting filtrate was concentrated under vacuum. A filter silica gel column was run to elude the corresponding free amine compound which was dissolved in anhydrous DMF at 0 °C. Simultaneously a mixture of N1-Cbz-Proline **24** (3.94 g, 15 mmol), HOBt (2.13 g, 15 mmol) and EDC (3.03 g, 15 mmol) in anhydrous DMF (30 mL) was stirred at 0 °C for 15 min under nitrogen atmosphere. The free amine compound in anhydrous DMF was slowly added to the reaction mixture during 15 min. The reaction mixture was stirred for 2 h at 0 °C and then at ambient temperature for 4 h. After completion of reaction water (150 mL) was added and the aqueous layer was extracted with ethyl acetate (3 x 70 mL). The combined organic layer was washed with NaHCO₃ (3 x 60 mL) to remove unreacted free acid followed by washing with KHSO₄ (60 mL) to remove unreacted free amine. The combined organic layer once again was washed with NaHCO₃ (60 mL) followed by water (3 x 70 mL) and brine. The organic layer dried over anhydrous Na₂SO₄ and concentrated under vacuum. The crude product obtained was purified by silica gel chromatography (40% ethyl acetate/hexane) to afford dipeptide **25** as white solid. Yield 5.58 g, (89%). Molecular Formula, C₂₄H₃₃N₃O₇. [Mass: calculated 475.53; observed 498.35 (M + Na⁺)].

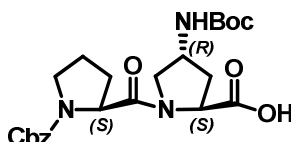
¹H NMR (CDCl₃, 400 MHz) δ_H: 1.44 (9H, d), 1.82 (1H, s), 1.85-1.91 (1H, m), 1.99-2.02 (1H, m), 2.05-2.23 (4H, m), 2.36-2.40 (1H, m), 3.46-3.57 (1H, m), 3.58-3.68 (1H, m), 3.70-3.73 (3H, d), 3.77-3.82 (1H, m), 3.97 (1H, d, *J* = 10.07), 4.10-4.15 (1H, dd, *J*

^{*} Synthesis of compound **7** is shown in Chapter 1

= 7.03), 4.30-4.39 (1H, m), 4.42-4.49 (1H, m), 4.63 (1H, t, $J = 7.78, 8.03$), 4.97-5.19 (2H, m), 5.68 (1H, d, $J = 6.02$), 7.31-7.39 (5H, m).

^{13}C NMR: 23.6, 24.5, 28.4, 29.4, 30.0, 34.2, 35.2, 46.8, 47.2, 50.5, 52.1, 52.3, 57.2, 57.6, 67.1, 79.6, 80.2, 127.9, 128.4, 128.5, 128.8, 130.9, 136.2, 136.9, 154.0, 155.1, 155.5, 171.5, 172.3.

N1-Cbz-proline-4R-NH-*t*-boc-proline acid dipeptide (26)

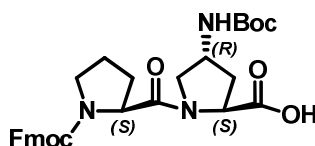


The dipeptide ester compound **25** (5 g, 10 mmol) was subjected to hydrolysis by using LiOH (0.75 g, 30 mmol) in THF:H₂O (70 mL) for 1 h. After completion of reaction, THF was removed under vacuum and the aqueous layer was washed with ethyl acetate to remove organic impurities. The residual aqueous layer was acidified by KHSO₄ to pH 2 and extracted with ethyl acetate (3 x 60 mL). The organic layer dried over anhydrous Na₂SO₄ and concentrated under vacuum. The crude product obtained was purified by silica gel chromatography (65% ethyl acetate/hexane) to afford dipeptide free acid **26** as white solid. Yield 4.60 g, (95%). Molecular Formula, C₂₃H₃₁N₃O₇. [Mass: calculated 461.51; observed 484.42 (M + Na⁺)].

^1H NMR (CDCl₃, 400 MHz) δ_{H} : 1.38 (9H, s), 1.79-1.85 (1H, m), 1.97-2.01 (4H, m), 2.05-2.13 (2H, m), 2.17 (1H, t, $J = 5.77$), 3.42-3.46 (1H, m), 3.53-3.59 (1H, m), 3.70-3.74 (1H, t), 3.79 (1H, d), 4.29-4.33 (1H, m), 4.44-4.47 (1H, m), 4.59-4.62 (1H, t, $J = 7.03$), 4.95-5.13 (2H, m), 5.95 (1H, d, $J = 6.53$), 7.23-7.30 (5H, m), 10.59 (1H, s).

^{13}C NMR: 20.8, 23.7, 24.4, 28.3, 28.4, 29.3, 33.7, 34.6, 46.9, 47.3, 50.3, 51.8, 52.2, 53.5, 57.4, 57.6, 57.5, 67.3, 79.7, 127.8, 127.9, 128.1, 128.2, 128.4, 128.6, 136.3, 154.1, 155.3, 155.7, 172.0, 172.3, 174.0, 174.7, 176.4.

N1-Fmoc-proline-4R-NH-*t*-boc-proline acid dipeptide (27)



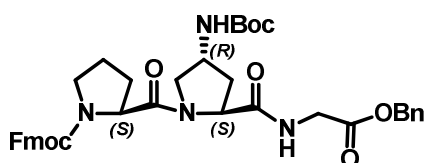
Methanol was carefully added to a mixture of dipeptide free acid **26** (4 g, 8 mmol) and 10% Pd/C (0.4 g). The resulting black suspension was shaken in Parr

hydrogenation apparatus under H₂ gas (65 psi pressure) for 6-7 h. After completion of reaction the suspension was filtered through Whatman paper. The resulting filtrate was concentrated under vacuum. The product obtained as a white solid was dissolved in water: dioxane (50 mL) by addition of 10% Na₂CO₃. The pH was maintained at 10 and the reaction mixture was stirred at 0 °C. Fmoc-Cl (3.36 g, 13 mmol) was added in portion wise during 45 min. The temperature of reaction mixture was maintained at 0 °C for first 4 h and then at ambient temperature for 12 h. The dioxane was removed under vacuum and the aqueous layer was washed with ethyl acetate (3 x 40 mL) to remove organic impurities. The aqueous layer was acidified with KHSO₄ to pH 2 followed by extraction with ethyl acetate (3 x 50mL). The organic layer dried over anhydrous Na₂SO₄ and concentrated under vacuum. The crude product obtained was purified by silica gel chromatography (60% ethyl acetate/hexane) to afford Fmoc protected dipeptide acid **27** as a white solid flappy powder. Yield 3.57 g, (75%). Molecular Formula, C₃₀H₃₅N₃O₇. [Mass: calculated 549.61; observed 472.52 (M + Na⁺)].

¹H NMR (CDCl₃, 400 MHz) δ_H: 1.40 (9H, d), 1.92-1.97 (2H, m), 2.08 (2H, m), 2.15-2.26 (2H, m), 2.32 (2H, s), 3.56 (1H, m), 3.69 (1H, t), 3.77 (1H, t), 3.89 (1H, d, *J* = 7.78), 4.25 (1H, d, *J* = 7.03), 4.29-4.43 (2H, t), 4.38-4.40 (1H, d, *J* = 7.03), 4.48-4.51 (1H, m), 4.66-4.70 (1H, t, *J* = 7.28), 5.72 (1H, s), 7.29-7.40 (4H, m), 7.55-7.65 (2H, dd, *J* = 7.53), 7.76-7.78 (2H, d, *J* = 7.28).

¹³C NMR: 28.3, 28.3, 29.5, 29.7, 34.3, 46.9, 47.1, 47.3, 47.6, 50.2, 52.5, 57.4, 57.7, 57.8, 67.8, 79.7, 119.8, 120.0, 124.9, 125.1, 125.2, 127.1, 127.7, 141.3, 143.7, 144.0, 155.2, 155.6, 172.8, 173.1.

N1-Fmoc-proline-4*R*-NH-*t*-boc-proline-glycine benzyl ester tripeptide (**28**)



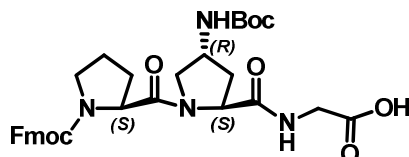
A mixture of dipeptide free acid **27** (3 g, 5 mmol), HOBt (0.92 g, 7 mmol) and EDC (1.31 g, 7 mmol) in anhydrous DMF (20 mL) was stirred at 0 °C for 15 min under nitrogen atmosphere. The free amine glycine benzyl ester **20** (1.80 g, 10 mmol) soluble in anhydrous DMF was slowly added to the reaction mixture during 15 min. The reaction mixture was stirred for 2 h at 0 °C and then at ambient temperature for 4

h. After completion of reaction water (100 mL) was added and the aqueous layer was extracted with ethyl acetate (3 x 60 mL). The combined organic layer was washed with NaHCO₃ (50 mL) to remove unreacted free acid followed by washing with KHSO₄ (50 mL) to remove unreacted free amine. The combined organic layer once again was washed with NaHCO₃ (50 mL) followed by water (3 x 50 mL) and brine. The organic layer dried over anhydrous Na₂SO₄ and concentrated under vacuum. The crude product obtained was purified by silica gel chromatography (70% ethyl acetate/hexane) to afford tripeptide **28** as white solid. Yield 2.66 g, (70%). Molecular Formula, C₃₉H₄₄N₄O₈. [Mass: calculated 696.79; observed 719.54 (M + Na⁺)].

¹H NMR (CDCl₃, 400 MHz) δ_H: 1.40-1.44 (9H, d), 1.89-1.97 (2H, m), 2.02-2.06 (1H, m), 2.08-2.14 (2H, m), 2.18 (4H, s), 2.23-2.33 (1H, m), 3.50-3.59 (1H, m), 3.87-3.90 (1H, m), 3.93-3.99 (1H, dd, *J* = 4.77), 4.09-4.13 (1H, dd, *J* = 4.52, 5.77), 4.16-4.18 (1H, d, *J* = 6.02), 4.16-4.21 (1H, m), 4.25 (1H, m), 4.30 (1H, m), 4.34-4.42 (2H, m), 4.48 (1H, t, *J* = 7.28), 4.71-4.71 (1H, t, *J* = 7.03, 6.78), 5.13-5.20 (2H, m), 5.56 (1H, s), 7.29-7.42 (9H, m), 7.55-7.63 (2H, m), 7.77 (2H, d, *J* = 7.53).

¹³C NMR: 24.7, 28.3, 29.7, 30.9, 33.3, 41.4, 46.9, 47.2, 47.7, 50.5, 52.7, 57.9, 57.9, 67.1, 67.1, 97.5, 119.9, 120.0, 125.1, 125.2, 127.0, 127.7, 128.2, 128.4, 128.5, 128.6, 135.2, 141.3, 143.8, 144.9, 155.1, 155.4, 169.5, 170.8, 173.0.

N1-Fmoc-proline-4*R*-NH-*t*-boc-proline-glycine acid tripeptide (**29**)

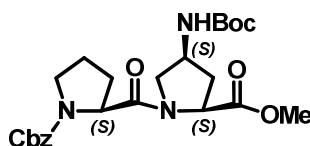


Methanol was carefully added to a mixture of tripeptide benzyl ester compound **28** (2.5 g, 3 mmol) and 10 % Pd/C (0.25 g). The resulting black suspension was shaken in Parr hydrogenation apparatus under 55 psi pressure of H₂ gas for 3-4 h. Careful monitoring by TLC was necessary to prevent hydrogenolysis of the Fmoc group. The suspension was filtered through a Whatman paper and concentrated under reduced pressure. The crude product obtained was purified by silica gel chromatography (85-90% ethyl acetate/hexane) to afford tripeptide acid **29** as a white solid powder. Yield 2.02 g, (93%). Molecular Formula, C₃₂H₃₈N₄O₈. [Mass: calculated 606.67; observed 629.52 (M + Na⁺)].

$^1\text{H NMR}$ (CDCl_3 , 400 MHz) δ_{H} : 1.38 (9H, s), 1.86 (2H, m), 2.05 (2H, s), 2.15 (2H, m), 2.20 (1H, m), 3.52 (2H, m), 3.61 (1H, m), 3.76 (2H, s), 3.90 (1H, m), 3.99 (1H, m), 4.22 (2H, m), 4.29 (2H, m), 4.35 (1H, d), 4.69 (1H, s), 5.30 (1H, s), 5.83 (1H, s), 7.29-7.40 (4H, m), 7.51-7.60 (2H, m), 7.75 (2H, d, $J = 7.28$).

$^{13}\text{C NMR}$: 24.6, 28.2, 28.5, 29.4, 34.1, 41.6, 47.1, 50.4, 52.8, 58.2, 58.6, 67.7, 79.7, 119.97, 125.10, 125.21, 127.06, 127.73, 141.26, 143.70, 144.06, 155.17, 155.77, 171.64, 172.44, 175.04.

N1-Cbz-proline-4S-NH-*t*-boc-proline methyl ester dipeptide (31)



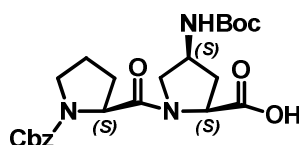
Methanol was carefully added to a mixture of N1-Cbz-4S-aminoproline methyl ester **11**[†] (6 g, 15 mmol) and 10% Pd/C (0.5 g). The resulting black suspension was shaken in Parr hydrogenation apparatus under H_2 gas (65 psi pressure) for 6-7 h. After completion of reaction the suspension was filtered through Whatman paper. The resulting filtrate was concentrated under vacuum. The silica gel column was run to elude the corresponding free amine compound which was dissolve in anhydrous DMF at 0 °C. Simultaneously a mixture of N1-Cbz-Proline **24** (4.74 g, 19 mmol), HOBT (2.56 g, 19 mmol) and EDC (3.64 g, 19 mmol) in anhydrous DMF (35 mL) was stirred at 0 °C for 15 min under nitrogen atmosphere. The free amine compound in anhydrous DMF was slowly added to the reaction mixture during 15 min. The reaction mixture was stirred for 2 h at 0 °C and then at ambient temperature for 4 h. After completion of reaction water (200 mL) was added and the aqueous layer was extracted with ethyl acetate (3 x 100 mL). The combined organic layer was washed with NaHCO_3 (3 x 100 mL) to remove unreacted free acid followed by washing with KHSO_4 (100 mL) to remove unreacted free amine. The combined organic layer once again was washed with NaHCO_3 (100 mL) followed by water (3 x 100 mL) and brine. The organic layer dried over anhydrous Na_2SO_4 and concentrated under vacuum. The crude product obtained was purified by silica gel chromatography (40% ethyl acetate/hexane) to afford dipeptide **31** as white solid. Yield 7 g, (92%). Molecular Formula, $\text{C}_{24}\text{H}_{33}\text{N}_3\text{O}_7$. [Mass: calculated 475.53; observed 498.32 ($\text{M} + \text{Na}^+$)].

[†] Synthesis of compound **11** is shown in Chapter 1.

¹H NMR (CDCl₃, 400 MHz) δ_H: 1.44 (9H, d), 1.73-1.79 (2H, m), 1.84-1.94 (2H, m), 1.96-2.05 (2H, m), 2.07-2.14 (1H, m), 2.15-2.24 (2H, m), 2.42-2.50 (1H, m), 3.34-3.37 (1H, d), 3.45-3.59 (2H, m), 3.62-3.67 (1H, m), 3.69-3.74 (1H, m), 3.76 (3H, d), 4.01-4.15 (1H, m), 4.24-4.37 (1H, m), 4.46-4.49 (1H, m), 4.56-4.59 (1H, dd, *J* = 9.79), 4.94-5.18 (2H, m), 5.41-5.57 (1H, dd, *J* = 8.03, 7.78), 7.29-7.35 (5H, m).

¹³C NMR: 22.6, 23.6, 24.2, 28.3, 29.2, 30.2, 34.3, 34.6, 46.7, 47.2, 50.4, 50.7, 52.5, 53.3, 53.6, 57.4, 57.5, 57.7, 58.1, 66.9, 67.2, 79.8, 127.7, 127.9, 128.0, 128.2, 128.3, 136.5, 136.7, 153.9, 154.9, 155.1, 171.2, 171.6, 173.6, 173.9.

N1-Cbz-proline-4*S*-NH-*t*-boc-proline acid dipeptide (**32**)

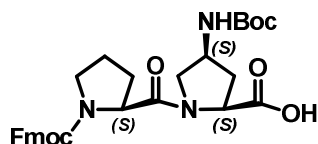


The dipeptide ester compound **31** (6.5 g, 13 mmol) was subjected to hydrolysis by using LiOH (0.95 g, 39 mmol) in THF:H₂O (70 mL) for 1 h. After completion of reaction, THF was removed under vacuum and the aqueous layer was washed with ethyl acetate to remove organic impurities. The residual aqueous layer was acidified by KHSO₄ to pH 2 and extracted with ethyl acetate (3 x 70 mL). The organic layer dried over anhydrous Na₂SO₄ and concentrated under vacuum. The crude product obtained was purified by silica gel chromatography (65% ethyl acetate/hexane) to afford dipeptide free acid **32** as white solid. Yield 5.92 g, (94%). Molecular Formula, C₂₃H₃₁N₃O₇. [Mass: calculated 461.51; observed 484.45 (M + Na⁺)].

¹H NMR (CDCl₃, 400 MHz) δ_H: 1.38-1.40 (9H, d), 1.82-2.00 (2H, m), 2.02-2.15 (2H, m), 2.17-2.20 (1H, m), 2.30-2.33 (1H, t), 3.39 (1H, m), 3.43-3.53 (1H, m), 3.58-3.66 (2H, m), 3.97 (1H, m), 4.30-4.33 (1H, m), 4.42-4.45 (1H, m), 4.70-4.73 (1H, m), 5.03-5.16 (2H, m), 5.41 & 5.55 (1H, dd, *J* = 6.27 & 6.07), 7.24-7.33 (5H, m).

¹³C NMR: 23.7, 24.3, 28.3, 29.3, 30.2, 33.0, 33.3, 46.8, 47.2, 50.4, 50.7, 54.1, 57.4, 58.1, 58.5, 58.8, 67.2, 67.5, 79.9, 127.8, 128.0, 128.5, 128.5, 136.3, 136.5, 154.0, 155.1, 155.5, 173.0, 173.3, 173.6, 173.9.

N1-Fmoc-proline-4*S*-NH-*t*-boc-proline acid dipeptide (**33**)

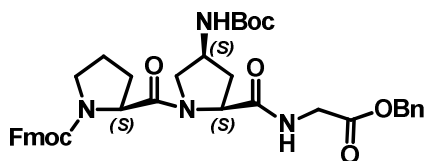


Methanol was carefully added to a mixture of dipeptide free acid **32** (5 g, 10 mmol) and 10% Pd/C (0.5 g). The resulting black suspension was shaken in Parr hydrogenation apparatus under H₂ gas (65 psi pressure) for 6-7 h. After completion of reaction the suspension was filtered through Whatman paper. The resulting filtrate was concentrated under vacuum. The product obtained as a white solid was dissolved in water: dioxane (60 mL) by addition of 10% Na₂CO₃. The pH was maintained at 10 and the reaction mixture was stirred at 0 °C. Fmoc-Cl (4.20 g, 16 mmol) was added in portion wise during 45 min. The temperature of reaction mixture was maintained at 0 °C for first 4 h and then at ambient temperature for 12 h. The dioxane was removed under vacuum and the aqueous layer was washed with ethyl acetate (3 x 60 mL) to remove organic impurities. The aqueous layer was acidified with KHSO₄ to pH 2 followed by extraction with ethyl acetate (3 x 60 mL). The organic layer dried over anhydrous Na₂SO₄ and concentrated under vacuum. The crude product obtained was purified by silica gel chromatography (60% ethyl acetate/hexane) to afford Fmoc protected dipeptide free acid **33** as a white solid flappy powder. Yield 4.64 g, (78%). Molecular Formula, C₃₀H₃₅N₃O₇. [Mass: calculated 549.61; observed 472.48 (M + Na⁺)].

¹H NMR (CDCl₃, 400 MHz) δ_H: 1.43 (9H, s), 1.70-1.90 (2H, m), 1.94-2.08 (2H, m), 2.17-2.25 (2H, m), 2.34-2.42 (1H, m), 3.04-3.33 (1H, m), 3.42-3.59 (2H, m), 3.71 (1H, d), 3.90-4.01 (1H, dd, *J* = 4.77, 6.26), 4.21-4.31 (2H, m), 4.39-4.43 (1H, m), 4.48-4.60 (1H, m), 4.68-4.71 (1H, d, *J* = 9.03), 5.49-5.60 (1H, dd, *J* = 6.02, 5.02), 7.29-7.41 (4H, m), 7.53-7.63 (2H, m), 7.76 (2H, d, *J* = 7.03), 8.13 (1H, s).

¹³C NMR: 23.1, 24.5, 28.3, 29.4, 29.7, 30.1, 32.6, 46.9, 47.1, 47.5, 50.6, 50.8, 53.6, 54.5, 57.4, 58.1, 58.5, 59.3, 65.6, 67.7, 80.1, 119.7, 119.8, 120.0, 125.1, 125.2, 127.1, 127.1, 127.2, 127.6, 127.7, 141.2, 141.3, 143.6, 143.7, 144.0, 154.2, 155.1, 156.5, 172.9, 173.4, 173.6, 174.0.

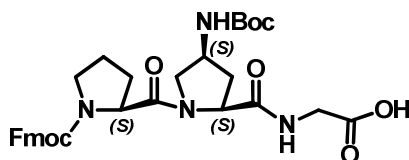
N1-Fmoc-proline-4S-NH-*t*-boc-proline-glycine benzyl ester tripeptide (**34**)



A mixture of Fmoc protected dipeptide free acid **23** (4 g, 7 mmol), HOBt (1.23 g, 9 mmol) and EDC (1.74 g, 9 mmol) in anhydrous DMF (25 mL) was stirred at 0 °C for 15 min under nitrogen atmosphere. The free amine glycine benzyl ester **20** (2.40 g, 14 mmol) soluble in anhydrous DMF was slowly added to the reaction mixture during 15 min. The reaction mixture was stirred for 2 h at 0 °C and then at ambient temperature for 4 h. After completion of reaction water (120 mL) was added and the aqueous layer was extracted with ethyl acetate (3 x 70 mL). The combined organic layer was washed with NaHCO₃ (150 mL) to remove unreacted free acid followed by washing with KHSO₄ (150 mL) to remove unreacted free amine. The combined organic layer once again was washed with NaHCO₃ (150 mL) followed by water (3 x 150 mL) and brine. The organic layer was dried over anhydrous Na₂SO₄ and concentrated under vacuum. The crude product obtained was purified by silica gel chromatography (70% ethyl acetate/hexane) to afford tripeptide **34** as white solid. Yield 3.65 g, (72%). Molecular Formula, C₃₉H₄₄N₄O₈. [Mass: calculated 696.79; observed 719.64 (M+ Na⁺)].

¹H NMR (CDCl₃, 400 MHz) δ_H: 1.43-1.45 (9H, s), 1.91-1.98 (2H, m), 2.04 (1H, s), 2.13-2.18 (2H, m), 2.30 (1H, m), 3.49-3.53 (2H, m), 3.71 (1H, m), 3.90-4.04 (2H, m), 4.10-4.13 (1H, m), 4.19-4.28 (2H, m), 4.33-4.43 (3H, m), 4.58-4.84 (1H, m), 5.14-5.19 (2H, m), 6.39-6.53 (1H, dd, *J* = 7.33 & 7.33), 7.30-7.44 (9H, m), 7.51-7.64 (2H, m), 7.75-7.78 (2H, d, *J* = 7.07), 7.89-7.92 (1H, t).

N1-Fmoc-proline-4S-NH-*t*-boc-proline-glycine acid tripeptide (**35**)



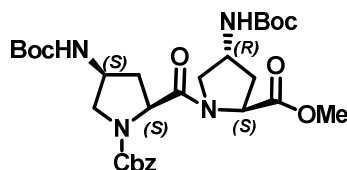
Methanol was carefully added to a mixture of tripeptide benzyl ester compound **34** (3.2 g, 4 mmol) and 10% Pd/C (0.32 g). The resulting black suspension was shaken in Parr hydrogenation apparatus under H₂ gas (55 psi pressure) for 3-4 h. Careful monitoring by TLC was necessary to prevent hydrogenolysis of the Fmoc group. The

suspension was filtered through a Whatman paper and concentrated under reduced pressure. The crude product obtained was purified by silica gel chromatography (85-90% ethyl acetate/hexane) to afford tripeptide free acid **35** as a white solid powder. Yield 2.53 g, (91%). Molecular Formula, C₃₂H₃₈N₄O₈. [Mass: calculated 606.67; observed 629.48 (M + Na⁺)].

¹H NMR (CDCl₃, 400 MHz) δ_H: 1.40-1.45 (9H, d), 1.90-1.99 (2H, m), 2.06 (1H, s), 2.12-2.19 (2H, m), 2.28 & 2.57 (1H, d), 2.90-3.18 (1H, m), 3.54 (2H, m), 3.68 (2H, m), 3.97 (1H, m), 4.11 (1H, s), 4.23 (2H, m), 4.30-4.34 (2H, m), 4.44 (1H, m), 4.51 (1H, s), 4.76 (1H, m), 5.32 (1H, s), 6.64 (1H, s), 7.31-7.41 (4H, m), 7.51-7.63 (2H, m), 7.78 (2H, d, *J* = 7.28).

¹³C NMR: 28.2, 28.4, 29.4, 29.7, 30.1, 46.9, 47.1, 47.5, 50.6, 53.4, 54.5, 57.6, 58.3, 59.4, 59.7, 65.7, 67.6, 79.5, 119.8, 119.9, 124.5, 125.1, 125.2, 127.1, 127.7, 141.3, 143.6, 143.8, 144.1, 154.6, 155.0, 155.3, 155.6, 172.0, 172.7, 175.2.

N1-Cbz-4*S*-NH-*t*-boc-proline-4*R*-NH-*t*-boc-proline methyl ester dipeptide (**37**)



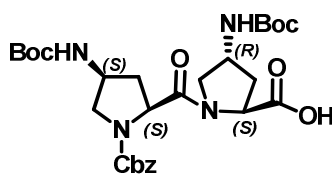
Methanol was carefully added to a mixture of N1-Cbz-4*R*-Aminoproline methyl ester **7** (5 g, 13 mmol) and 10% Pd/C (0.5 g). The resulting black suspension was shaken in Parr hydrogenation apparatus under H₂ gas (65 psi pressure) for 6-7 h. After completion of reaction the suspension was filtered through Whatman paper. The resulting filtrate was concentrated under vacuum. A filter silica gel column was run to elude the corresponding free amine compound **23** which was dissolve in anhydrous DMF at 0 °C. Simultaneously the N1-Cbz-4*S*-aminoproline methyl ester **11** (5.2 g, 13 mmol) was subjected to hydrolysis by using LiOH (0.95 g, 39 mmol) in THF:H₂O (80 mL) for 1 h. After completion of reaction, THF was removed under vacuum and the aqueous layer was washed with ethyl acetate to remove organic impurities. The residual aqueous layer was acidified by KHSO₄ to pH 2 and extracted with ethyl acetate (3 x 70 mL). The organic layer was dried over anhydrous Na₂SO₄ and concentrated under vacuum. The crude product obtained was purified by silica gel chromatography (55% ethyl acetate/hexane) to afford free acid **36** as a white solid.

The free acid compound **36** (4.73 g, 13 mmol), HOBt (2.13 g, 15 mmol) and EDC (3.03 g, 15 mmol) in anhydrous DMF (45 mL) was stirred at 0 °C for 15 min under nitrogen atmosphere. The free amine compound **23** (3.17 g, 13 mmol) soluble in anhydrous DMF was slowly added to the reaction mixture in 15 min time duration. The reaction mixture was stirred for 2 h at 0 °C and then at ambient temperature for 4 h. After completion of reaction water (150 mL) was added and the aqueous layer was extracted with ethyl acetate (3 x 80 mL). The combined organic layer was washed with NaHCO₃ (3 x 150 mL) to remove unreacted free acid followed by washing with KHSO₄ (150 mL) to remove unreacted free amine. The combined organic layer once again was washed with NaHCO₃ (150 mL) followed by water (3 x 100 mL) and brine. The organic layer was dried over anhydrous Na₂SO₄ and concentrated under vacuum. The crude product obtained was purified by silica gel chromatography (45% ethyl acetate/hexane) to afford dipeptide **37** as a white solid. Yield 6.51 g, (85%). Molecular Formula, C₂₉H₄₂N₄O₉. [Mass: calculated 590.67; observed 613.45 (M + Na⁺)].

¹H NMR (CDCl₃, 400 MHz) δ_H: 1.43 (9H, s), 1.44-1.45 (9H, d), 1.78-1.92 (2H, m), 1.97-2.18 (2H, m), 2.40-2.49 (2H, m), 3.67 (1H, d, *J* = 11.28), 3.69-3.71 (1H, t), 3.75 (3H, s), 3.77-3.80 (1H, m), 4.10 (1H, d, *J* = 9.08), 4.28 (1H, s), 4.40-4.43 (2H, t), 4.53-4.55 (1H, d, *J* = 9.35), 4.62-4.65 (1H, t, *J* = 7.98, 8.25), 4.96-5.18 (2H, m), 5.74 (1H, d, *J* = 5.23), 6.21 & 6.58 (1H, dd, *J* = 8.80, 7.15), 7.31-7.40 (5H, m).

¹³C NMR: 28.4, 28.4, 29.7, 34.1, 35.0, 35.4, 36.6, 49.2, 50.1, 50.6, 51.7, 52.5, 54.9, 56.5, 56.9, 57.3, 57.5, 67.5, 79.3, 79.6, 91.1, 94.0, 98.8, 113.0, 115.3, 123.6, 128.1, 128.2, 128.5, 128.7, 136.0, 136.5, 153.7, 155.1, 155.4, 161.2, 171.3, 171.7, 172.1.

N1-Cbz-4*S*-NH-*t*-boc-proline-4*R*-NH-*t*-boc-proline acid dipeptide (**38**)



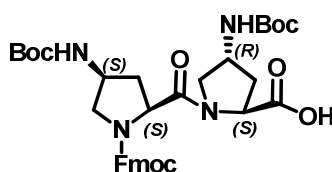
The dipeptide ester compound **37** (6 g, 10 mmol) was subjected to hydrolysis by using LiOH (0.72 g, 30 mmol) in THF:H₂O (60 mL) for 1 h. After completion of reaction, THF was removed under vacuum and the aqueous layer was washed with ethyl acetate to remove organic impurities. The residual aqueous layer was acidified by KHSO₄ to pH 2 and extracted with ethyl acetate (3 x 60 mL). The organic layer

dried over anhydrous Na_2SO_4 and concentrated under vacuum. The crude product obtained was purified by silica gel chromatography (60% ethyl acetate/hexane) to afford dipeptide free acid **38** as a white solid. Yield 5.56 g, (95%). Molecular Formula, $\text{C}_{28}\text{H}_{40}\text{N}_4\text{O}_9$. [Mass: calculated 576.64; observed 599.55 ($\text{M} + \text{Na}^+$)].

^1H NMR (CDCl_3 , 400 MHz) δ_{H} : 1.42 (9H, s), 1.44 (9H, s), 1.97-2.05 (1H, m), 2.22 (1H, m), 2.43-2.46 (2H, m), 2.74 (1H, m), 3.48-3.54 (1H, t), 3.71-3.79 (2H, m), 4.08-4.18 (1H, m), 4.29-4.44 (2H, m), 4.53 (1H, d, $J = 6.02$), 4.65-4.69 (1H, t, $J = 7.78$, 7.53), 4.96-5.18 (2H, m), 6.00 (1H, s), 6.55 (2H, d, $J = 8.53$), 7.34-7.38 (5H, d).

^{13}C NMR: 28.4, 29.6, 34.7, 35.0, 49.1, 49.9, 50.4, 52.5, 53.4, 54.1, 56.3, 56.8, 57.5, 67.6, 79.5, 79.7, 128.1, 128.2, 128.5, 128.7, 136.0, 153.9, 155.2, 155.7, 172.1, 173.6.

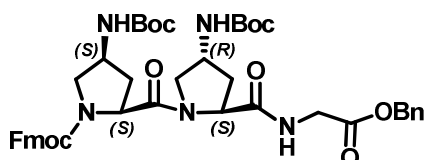
N1-Cbz-4S-NH-*t*-boc-proline-4R-NH-*t*-boc-proline acid dipeptide (**39**)



Methanol was carefully added to a mixture of dipeptide free acid **38** (5 g, 8 mmol) and 10% Pd/C (0.5 g). The resulting black suspension was shaken in Parr hydrogenation apparatus under H_2 gas (65 psi pressure) for 6-7 h. After completion of reaction the suspension was filtered through Whatman paper. The resulting filtrate was concentrated under vacuum. The product obtained as a white solid was dissolved in water: dioxane (80 mL) by addition of 10% Na_2CO_3 . The pH was maintained at 10 and the reaction mixture was stirred at 0 °C. Fmoc-Cl (4.20 g, 16 mmol) was added in portion wise during 45 min. The temperature of reaction mixture was maintained at 0 °C for first 4 h and then at ambient temperature for 12 h. The dioxane was removed under vacuum and the aqueous layer was washed with ethyl acetate (3 x 80 mL) to remove organic impurities. The aqueous layer was acidified with KHSO_4 to pH 2 followed by extraction with ethyl acetate (3 x 80 mL). The organic layer dried over anhydrous Na_2SO_4 and concentrated under vacuum. The crude product obtained was purified by silica gel chromatography (55% ethyl acetate/hexane) to afford Fmoc protected dipeptide free acid **39** as white solid flappy powder. Yield 4.14 g, (72%). Molecular Formula, $\text{C}_{35}\text{H}_{44}\text{N}_4\text{O}_9$. [Mass: calculated 664.75; observed 687.68 ($\text{M} + \text{Na}^+$)].

¹H NMR (CDCl₃, 400 MHz) δ_H: 1.38-1.45 (18H, d), 1.99-2.05 (2H, m), 2.17 (2H, s), 2.33 (2H, m), 3.67 (2H, m), 3.76 (2H, d), 4.07 (1H, s), 4.21-4.23 (2H, d, *J* = 6.27), 4.31 (2H, d, *J* = 6.53), 4.38 (1H, s), 5.63 (2H, s), 6.06 & 6.36 (1H, d), 6.76 (1H, s), 7.29-7.41 (4H, m), 7.53-7.59 (2H, dd, *J* = 7.28, 7.03), 7.74 (2H, d, *J* = 7.53).

N1-Cbz-4*S*-NH-*t*-boc-proline-4*R*-NH-*t*-boc-proline-glycine-benzyl ester tripeptide (40)



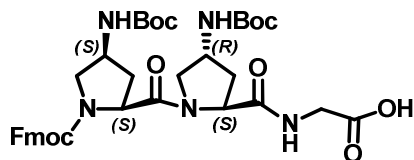
A mixture of Fmoc protected dipeptide free acid **39** (3.5 g, 5 mmol), HOBT (0.85 g, 7 mmol) and EDC (1.21 g, 7 mmol) in anhydrous DMF (25 mL) was stirred at 0 °C for 15 min under nitrogen atmosphere. The free amine glycine benzyl ester **20** (1.73 g, 10 mmol) soluble in anhydrous DMF was slowly added to the reaction mixture during 15 min. The reaction mixture was stirred for 2 h at 0 °C and then at ambient temperature for 4 h. After completion of reaction water (100 mL) was added and the aqueous layer was extracted with ethyl acetate (3 x 60 mL). The combined organic layer was washed with NaHCO₃ (150 mL) to remove unreacted free acid followed by washing with KHSO₄ (150 mL) to remove unreacted free amine. The combined organic layer once again was washed with NaHCO₃ (150 mL) followed by water (3 x 150 mL) and brine. The organic layer dried over anhydrous Na₂SO₄ and concentrated under vacuum. The crude product obtained was purified by silica gel chromatography (65% ethyl acetate/hexane) to afford tripeptide **40** as a white solid. Yield 3.30 g, (77%). Molecular Formula, C₄₄H₅₃N₅O₁₀. [Mass: calculated 811.91; observed 834.78 (M + Na⁺)].

¹H NMR (CDCl₃, 400 MHz) δ_H: 1.40 (9H, s), 1.45 (9H, s), 1.73 (1H, m), 1.97-2.05 (2H, m), 2.20 (1H, m), 2.43-2.48 (2H, m), 3.64 (1H, m), 3.71-3.74 (1H, m), 3.76-3.78 (1H, m), 3.99 (1H, d, *J* = 4.77), 4.02 (1H, d, *J* = 5.27), 4.12 (1H, t), 4.16-4.24 (2H, m), 4.29 (1H, m), 4.35 (2H, m), 4.43 (1H, m), 4.53 (1H, d, *J* = 8.28), 4.70 (1H, m), 4.93-5.16 (2H, m), 5.71 (1H, s), 6.21 & 6.52 (1H, d), 7.29-7.42 (9H, m), 7.53-7.60 (2H, m), 7.77 (2H, d, *J* = 7.53).

¹³C NMR: 22.7, 24.7, 28.3, 28.5, 29.4, 29.7, 31.9, 33.0, 34.5, 35.0, 41.4, 46.9, 47.0, 50.1, 50.6, 52.8, 54.6, 57.1, 58.1, 67.2, 68.0, 79.4, 79.5, 120.0, 125.0, 125.2, 127.1,

127.8, 128.2, 128.4, 128.6, 128.7, 135.1, 141.2, 141.3, 143.6, 143.8, 155.2, 155.5, 169.5, 170.7, 172.7.

N1-Cbz-4S-NH-*t*-boc-proline-4R-NH-*t*-boc-proline-glycine acid tripeptide (**41**)



Methanol was carefully added to a mixture of tripeptide benzyl ester compound **40** (3 g, 3 mmol) and 10% Pd/C (0.30 g). The resulting black suspension was shaken in Parr hydrogenation apparatus under H₂ gas (55 psi pressure) for 3-4 h. Careful monitoring by TLC was necessary to prevent hydrogenolysis of the Fmoc group. The suspension was filtered through a Whatman paper and concentrated under reduced pressure. The crude product obtained was purified by silica gel chromatography (85-90% ethyl acetate/hexane) to afford tripeptide free acid **41** as a white solid powder. Yield 2.45 g, (92%). Molecular Formula, C₃₇H₄₇N₅O₁₀. [Mass: calculated 721.79; observed 744.68 (M + Na⁺)].

¹H NMR (CDCl₃, 400 MHz) δ_H: 1.44 (9H, d), 1.48 (9H, s), 1.89 (1H, m), 1.97 (2H, s), 2.20 (1H, m), 2.25 (1H, m), 2.61 (1H, m), 3.53 (1H, m), 3.72 (1H, s), 3.75 (2H, m), 3.83 (2H, t), 4.12-4.21 (2H, m), 4.25 (2H, t), 4.36 (1H, t), 4.48 (1H, m), 4.56 (2H, t), 5.48 (1H, s), 7.32-7.42 (4H, m), 7.58-7.64 (2H, m), 7.80 (2H, d, *J* = 7.32).

¹³C NMR: 22.0, 27.4, 27.5, 29.4, 34.1, 34.9, 43.1, 48.1, 48.3, 50.1, 52.2, 53.4, 59.0, 67.1, 67.7, 79.1, 79.2, 119.7, 124.5, 124.7, 124.8, 126.9, 126.9, 127.0, 127.5, 129.0, 135.0, 141.2, 143.6, 144.1, 154.6, 155.1, 156.3, 171.6, 171.8, 171.9, 175.2.

6.2 Peptide Synthesis

All peptides were synthesized manually in a sintered vessel equipped with a stopcock. The readily available Rink amide resin with loading value 0.5-0.6 mmol/g was used and standard Fmoc chemistry was employed. The resin bound Fmoc group was first deprotected with 20% piperidine in DMF and the coupling reactions were carried out using *in situ* active ester method, using HBTU as a coupling reagent and HOBt as a racemization suppresser and DIPEA as a catalyst. All the materials used were of peptide synthesis grade (Sigma-Aldrich) and was used without further purification. Analytical grade DMF was purchased from Merck (India) and was

distilled over P₂O₅ under vacuum at 45°C, stored over 4Å molecular sieves for 2 days before using for peptide synthesis.

6.2a Resin functionalization

The Rink amide resin (2',4'-Dimethoxyphenyl-Fmoc-aminomethyl)-phenoxy) (Nova biochem, 100-200 mesh) was taken in sintered vessel (20 mL) and rinsed with anhydrous DCM (5 mL) and filtered. The process was repeated 3 to 4 times and the resulting resin was kept for 2 h in DCM (10 mL) for swelling. The DCM was removed by filtration and rinsed 3 times with anhydrous DMF and kept 2 h in anhydrous DMF (10 mL) for swelling before the first coupling. The deprotection of 'Fmoc group' attached to the resin was done with 20% piperidine in DMF (3 x 5 mL) before proceeding for first tripeptide block coupling.

- The resin was washed and swollen in dry DCM for at least 2 h.
- Further washing and swelling with dry DMF for 2 h.
- Immediate coupling of 1st trimeric block desired in C-terminus of peptide.

6.2b General method for solid phase peptide synthesis

All peptides were assembled on solid phase method by tripeptide block coupling. All tripeptide free acids were dried over P₂O₅ in vacuum desiccator before coupling. Fmoc protecting group was used for main chain α -amino group. The *t*-Boc protection was used for side chain amine protection and was cleaved with 20% TFA in DCM for final cleavage of peptide from the resin. The peptide obtained after cleavage was stirred in 95% TFA in DCM for 2 h for complete deprotection of Boc group.

6.2c Synthesis protocol for solid phase synthesis

The resin was pre-swollen overnight and the following steps were performed for each cycle.

- Wash with DMF 4 x 5 mL.
- Wash with 20% piperidine in DMF 3 x 5 mL (15 minutes for each) for deprotection of Fmoc group.
- Wash with DMF 3 x 5 mL, wash with MeOH 3 x 5 mL and wash with DCM with 3 x 5 mL.
- Test for complete deprotection (chloranil test).

- Coupling reaction with each tripeptide (3 eq), DIPEA (7 eq), HOBt (3 eq) and HBTU (3 eq) in minimum volume of DMF.
- Same coupling reaction repeated in NMP for better yield.
- Test for completion of coupling reaction (chloranil test).

This cycle was repeated for every tripeptide block.

6.2d General procedure for peptide couplings on Rink Amide Resin

Fmoc-xxx-OH (3 eq), HBTU (3 eq) and HOBT (3 eq) dissolved in DMF/NMP followed by *i*Pr₂NEt (7-8 eq) were added to the amino-functionalized resin in DMF (1 mL) (\approx 100 mM concentration). The mixture was stand for 2 h and last 5 min bubbled with N₂ and washed with DMF (3x), MeOH (3x) and CH₂Cl₂ (3x) was omitted. The loading value for peptide synthesis is taken as 0.5~0.6.

6.2e General procedure for Fmoc deprotection

20% piperidine in DMF was added to the resin and the reaction mixture was kept for 15 min, drained and the piperidine treatment repeated for 3 times. Finally the resin was washed with DMF (3x), MeOH (3x) and CH₂Cl₂ (3x).

6.2f General procedure for acetylation

Triethylamine (20 eq) and acetic anhydride (20 eq) were added to the resin in DMF (\approx 100 mM). The mixture was kept for 1 h followed by bubbled with N₂ for 5 min and washed with DMF (3x), MeOH (3x) and CH₂Cl₂ (3x).

6.2g Preparation of resin with peptide for cleavage

After the last coupling/acetylation was over the resin was washed with DMF (5 x 10 mL), DCM (5 x 10 mL) and finally with MeOH (5 x 10 mL) and dried with nitrogen gas for 3 minutes. The resin containing sintered flask was dried in a vacuum desiccator over P₂O₅.

6.2h General procedure for cleavage of peptides from the solid support

The dry peptide-resin (20 mg) was taken in round-bottomed flask to which of 20% TFA in DCM (10 mL) and triisopropylsilane (as scavengers) (2-3 drops) were added.

The resulting mixture was kept for 2 h by gentle shaking. The mixture was filtered through a sintered funnel and the resin was washed with 3 x 5 mL of above solution. The filtrate was collected in pear shape round-bottom flask and evaporated under reduced pressure. The resin was washed with MeOH (3 X 5 mL) and the washings were evaporated to dryness. The crude peptide obtained containing a N⁴-*t*Boc group, was deprotected by stirring the peptide solution with 95% TFA in DCM (10 mL) for 2 h. The TFA:DCM mixture was removed under reduced pressure. The residue obtained was dissolved in anhydrous methanol (0.4 mL) and to it anhydrous diethyl ether (4 x 1.5 mL) was added. The off-white precipitate obtained was centrifuged. The precipitation procedure was repeated twice to obtain peptide as a colourless powder.

6.3 High performance liquid chromatography (HPLC)

Peptides (**4-21**) were purified by reverse phase-HPLC on Waters 600 equipped with 2998-Photodiode array detector (PDA). Semi-preparative RP-C18 columns (250 x 10 mm, 10 μ M) of Allteck-Alltima make were used for peptides. The solvent system comprised of MeCN:Water (5:95) with 0.1% TFA for solution A and for solution B MeCN:Water (50:50), 0.1% TFA. A gradient of 0-100% at a flow rate of 3 mL/min was used for semi preparative HPLC to elute the peptide and the eluant was monitored at 220 nm. The peak corresponding to the peptide was collected and the fractions were frozen. Subsequently these peptides were concentrated by using speed vacuum. The purity of the final peptides were further analyzed on the Merck LiChrospher 100 RP-18 (250 x 4 mm, 5 μ M) column by using a gradient flow of 0 to 100% B in 20 min at a flow rate of 1.5 mL/min. The oven was heated to 50 °C to prevent triple helix formation for non acetylated peptides **13-21**. The spectra acquisition, analysis and processing was done on Waters Empower-2154 software. The absorbance of the eluant was monitored at its corresponding wavelength and the purity was obtained from the integrator output. The purities of the hence purified peptides were found to be more than 95%.

6.4 MALDI-TOF characterization

MALDI-TOF mass spectra were obtained on either Voyager-Elite instrument (PerSeptive Biosystems Inc., Farmingham, MA) equipped with delayed extraction or

on Voyager-De-STR (Applied Biosystems) instrument. Sinapinic acid and α -cyano-4-hydroxycinnamic acid (CHCA) both were used as matrix for peptides of which CHCA was found to give satisfactory results. A saturated matrix solution was prepared with typical dilution solvent (50:50:0.1 Water:MeCN:TFA) and spotted on the metal plate along with the oligomers. The metal plate was loaded to the instrument and the analyte ions are then accelerated by an applied high voltage (15-25 kV) in reflector mode, separated in a field-free flight tube and detected as an electrical signal at the end of the flight tube. HPLC purified peptides were characterized through this method and were observed to give good signal to noise ratio, mostly producing higher molecular ion signals.

6.5 Circular dichroism (CD) spectroscopy

CD spectrometric studies were carried out on JASCO J-715 spectropolarimeter using cylindrical, jacketed quartz cell (1 mm path length), which was connected to Julabo-UC-25 water circulator. CD spectra were recorded using a spectral bandwidth of 1.0 nm at 25 °C with a time constant of 1 s and a step resolution of 1 nm. All the spectra were corrected for respective buffer condition and are typically averaged over 5-10 scans. CD data are given as mean residual molar ellipticities $[\theta] 10^3 \text{ deg cm}^2 \text{ dmol}^{-1}$. The spectra are the result of 5-10 accumulations. A quartz cell with a path length of 1 mm was used with solutions containing approximately 0.2 mL (50-500 μM) peptide solutions. For the spectra in buffer the blank spectrum of the solution was subtracted. All samples were equilibrated for at least 10 h before measurement.

Resolution:	1 nm
Band width:	1.0 nm
Sensitivity:	100 mdeg
Response:	1 sec
Speed:	100 nm/min
Accumulation:	5-10

For CD thermal denaturation studies, samples were annealed in water bath at 95 °C and slowly cooled to room temperature over a period of 6 h prior to spectroscopic analysis. These samples were then incubated at 4 °C for 12 h followed by additional half hour incubation in the instrument at the initial measurement temperature. The temperature was varied in steps of 5 °C and the spectra were recorded at each step. An equilibration period of 5 min was allowed at each temperature. Data processing and curve fitting was performed using MicroCal Origin 8.0 software. Ellipticity at specified wavelength for each temperature was plotted, normalized data was fitted to a sigmoid curve and the T_m (melting temperature) values are derived from the first derivative curve of the fit.

7 References

- 1 Schweitzer, M. H.; *et. al Science* **2007**, *316*, 277-280.
- 2 Engel, J. *Science* **1997**, *277*, 1787-1786.
- 3 (a) Brinckmann, J. *Top. Curr. Chem.* **2005**, *247*, 1–6. (b) Veit, G.; Kobbe, B.; Keene, D. R.; Paulsson, M.; Koch, M. and Wagener R. *J. Biol. Chem.*, **2006**, *281*, 3494–504.
- 4 Bhattacharjee, A. and Bansal, M. *Life*, **2005**, *57*, 161–172.
- 5 Brazel, D.; Oberb"auamer, I.; Dieringer, H.; Babel, W.; Glanville, R. W. *Eur. J. Biochem.* **1987**, *168*, 529–536.
- 6 Woodhead-Galloway, J. in *Collagen: The Anatomy of a Protein* Edward Arnold (Publishers) Ltd. London, 10-19.
- 7 Ramshaw, J.; Shah, N.K.; Brodsky, B. *J. Struct. Biol.* **1998**, *122*, 86–91.
- 8 Gordon, M. K. and Hahn R. A., *Cell Tissue Res.*, **2010**, *339*, 247–257.
- 9 Cole W. G. *Prog. Nucl. Acid Res. Mol. Biol.*, **1994**, *47*, 29–80.
- 10 Shoulders, M. D.; Raines, R. T. *Annu. Rev. Biochem.* **2009**, *78*, 929–958.
- 11 (a) Rougvie, M. A. and Bear, R. S. *J. Am. Leather Chem. Ass.*, **1953**, *48*, 735. (b) Ramachandran, G. N. *Chemistry of Collagen*, Academic Press, London, 1967. (c) Fraser, R. D. and Suzuki, E. *J. Mol. Biol.*, **1979**, *129*, 463.
- 12 (a) Ramachandran, G. N. and Kartha, G. *Nature*, **1954**, *174*, 269. (b) Ramachandran, G. N. and Kartha, G. *Nature*, **1955**, *176*, 593. (c) Ramakrishnan, C. *Protein Sci.*, **2001**, *10*, 1689.
- 13 (a) Rich, A. and Crick, F. H. C. *Nature*, **1955**, *176*, 915. (b) Rich, A. and Crick, F. H. C. *J. Mol. Biol.*, 1961, **3**, 483. (c) Riddihough, G. *Nat. Struct. Biol.*, **1998**, *5*, 858.

- 14 Bella, J.; Eaton, M.; Brodsky, B. and Berman, H. M. *Science*, **1994**, 266, 75.
- 15 Bella, J. and Berman, H. M. *J. Mol. Biol.*, **1996**, 264, 734–742.
- 16 Okuyama, K.; Xu, X.; Iguchi, M. and Noguchi, K. *Biopolymers*, **2006**, 84, 181–91.
- 17 Cowan, P.M.; McGavin, S. and North, A.C. *Nature*, **1955**, 176, 1062–1064.
- 18 Cohen, C. and Bear, R. S. *J. Am. Chem. Soc.* **1953**, 75, 2783–2784.
- 19 (a) Kramer, R. Z.; Bella, J.; Mayville, P.; Brodsky, B. and Berman, H. M. *Nat. Struct. Biol.* **1999**, 6, 454–457. (b) Boudko, S.; Engel, J.; Okuyama, K.; Mizuno, K.; Bächinger, H.P. and Schumacher, M. A. *J. Biol. Chem.* **2008**, 283, 32580–32589.
- 20 (a) Sweeney, S. M.; Guy, C. A.; Fields, G. B.; San Antonio, J. D. *Proc. Natl. Acad. Sci. USA*, **1998**, 95, 7275–7280 (b) Di Lullo, G. A.; Sweeney, S. M.; Korkkko, J.; Ala-Kokko, L. and San Antonio, J. D. *J. Biol. Chem.* **2002**, 277, 4223–4231. (c) Sweeney, S. M.; Orgel, J. P.; Fertala, A.; McAuliffe, J. D. and Turner, K. R. *J. Biol. Chem.* **2008**, 283, 21187–21197.
- 21 Nimni, M.E.; Harkness, R.D.; Collagen Vol. I – Biochemistry, CRC Press, Boca Raton, FL, **1988**, pp. 1–79.
- 22 Kohmura, E. *et.al. Brain Res.* **1999**, 849, 235–238.
- 23 Kadler, K. E.; Hojima, Y. and Prockop, D. J. *J. Biol. Chem.*, **1987**, 262, 15696–15701.
- 24 Prockop, D. J and Fertala, A. *J. Biol. Chem.*, **1998**, 273, 15598–15604.
- 25 Kuznetsova, N. and Leikin S. *J. Biol. Chem.* **1999**, 274, 36083–36088.
- 26 Eyre D. R, Paz M. A, Gallop P. M., *Annu. Rev. Biochem.*, **1984**, 53, 717–748.
- 27 (a) Prockop, D. J.; Kivrikko, K. I.; Tuderman, L.; Guzman, N. A. *New Eng. J. Med.* **1979**, 301, 13-23. (b) Prockop, D. J. *New Eng. J. Med.* **1992**, 326, 540-546. (c) Brodsky, B., Shah, N. K. *FASEB J.* **1995**, 9, 153-1546.
- 28 For review see, Koide, T. and Nagata, K. *Top Curr Chem*, **2005**, 247, 85-114.
- 29 (a) Haasnoot, C. A. G.; De Leeuw, A. A. M.; De Leeuw, H. P. M.; Altona, C. *Biopolymers* **1981**, 20, 1211-1245. (b) Flores-Ortega, A., Casanovas, J.; Assfeld, X., Alemán, C. *J. Org. Chem.* **2009**, 74, 3101–3108.
- 30 Jankowski, K.; Soler, F.; Ellenberger, M. *J. Mol. Struct.*, **1978**, 48, 63-68.
- 31 Eberhardt, E. S.; Loh, S. N. and Raines, R. T. *Tetrahedron Lett.* **1993**, 33, 3055-3056.
- 32 Bretscher, L. E.; Jenkins, C. L.; Taylor, K. M.; DeRider, M. L.; Raines, R. T. *J. Am. Chem. Soc.* **2001**, 123, 777-778.
- 33 Shoulders, M. D.; Hodges, J. A.; Raines, R. T., *J. Am. Chem. Soc.* **2006**, 128, 8112-8113.

- 34 (a) Wolfe, S. *Acc. Chem. Res.* **1972**, *5*, 102. (b) Brunck, T. K.; Weinhold, F. *J. Am. Chem. Soc.* **1979**, *101*, 1700-1790. (c) Senderowitz, H.; Aped, P.; Fuchs, B. *J. Comput. Chem.* **1993**, *14*, 944. (d) Houk, K. N.; Eksterwicz, J. E.; Wu, Y. D.; Fuglesang, C. D.; Mitchell, D. B. *J. Am. Chem. Soc.* **1993**, *115*, 4170.
- 35 Parker, D.; Senanayake, K.; Vepsäläinen, J.; Williams, S.; Batsanov A. S.; Howard, J. A. K. *J. Chem. Soc., Perkin Trans. 2*, **1997**, *8*, 1445-1452.
- 36 Wiberg, K. B. *Acc. Chem. Res.* **1996**, *29*, 229-234.
- 37 Deslongchamps, P. *Stereoelectronic Effects in Organic Chemistry*; Pergamon Press: New York, **1983**.
- 38 Gorenstein, D. G. *Chem. Rev.* **1987**, *87*, 1047-1077.
- 39 Plavec, J.; Thibaudeau, C.; Chattopadhyaya, J. *J. Am. Chem. Soc.* **1994**, *116*, 6558-6560.
- 40 Gerig, J. T.; McLeod, R. S. *J. Am. Chem. Soc.* **1973**, *95*, 5725-5729.
- 41 Zimmerman, S. S.; Scheraga, H. A. *Macromolecules* **1976**, *9*, 408-416.
- 42 (a) Maccallum, P. H.; Poet, R.; Milner-White, E. J. *J. Mol. Biol.* **1995**, *248*, 361-373. (b) Maccallum, P. H.; Poet, R.; Milner-White, E. J. *J. Mol. Biol.* **1995**, *248*, 374-384.
- 43 (a) DeRider, M. L.; Wilkens, S. J.; Waddell, M. J.; Bretscher, L. E.; Weinhold, F.; Raines, R. T.; Markely, J. L. *J. Am. Chem. Soc.* **2002**, *124*, 2497-2505. (b) Hinderaker, M. P.; Raines, R. T. *Protein Sci.* **2003**, *12*, 1188-1194.
- 44 Dolz, R.; Heidemann, E. *Biopolymers*, **1986**, *25*, 1069-1080.
- 45 Heidemann, E.; Roth, W. *Adv. Polym. Sci.* **1982**, *43*, 143-203.
- 46 Persikov, A. V.; Ramshaw, J. A.; Brodsky, B. *J. Biol. Chem.* **2005**, *280*, 19343-19349.
- 47 Cram D. J. *Science*, **1988**, *240*, 760-767.
- 48 Vitagliano, L.; Berisio, R.; Mazzarella, L. and Zagari, A. *Biopolymers*, **2001**, *58*, 459-464.
- 49 Umashankara, M.; Babu, I. R.; Ganesh, K. N. *Chem. Commun.* **2003**, 2606-2607.
- 50 (a) Berg, R. A.; Prockop, D. J. *Biochem. Biophys. Res. Comm.* **1973**, *52*, 115-120 (b) Sakakibara, S.; Inouye, K.; Shudo, K.; Kishida, Y.; Kobayashi, Y.; Prockop, D. J. *Biochim. Biophys. Acta*, **1973**, *303*, 198-202.
- 51 (a) Inouye, K.; Sakakibara, S.; Prockop, D. J. *Biochim. Biophys. Acta*, **1976**, *420*, 133-141 (b) Jiravanichanun, N.; Nishino, N.; Okuyama, K. *Biopolymers*, **2006**, *81*, 225-233.
- 52 Bella, J.; Brodsky, B.; Berman, H. M. *Structure*, **1995**, *3*, 893-906.
- 53 Inouye, K.; Kobayashi, Y.; Kyogoku, Y.; Kishida, Y.; Sakakibara, S.; Prockop, D. J. *Arch. Biochem. Biophys.* **1982**, *219*, 198-203

- 54 (a) Berisio, R.; Granata, V.; Vitagliano, L.; Zagari, A. *J. Am. Chem. Soc.*, **2004**, *126*, 11402–11403. (b) Mizuno, K.; Hayashi, T.; Peyton, D. H.; Bächinger, H. P. *J. Biol. Chem.*, **2004**, *279*, 38072–38078.
- 55 Shoulders, M. D.; Kotch, F. K.; Choudhary, A.; Guzei, I. A.; Raines, R. T. *J. Am. Chem. Soc.* **2010**, *132*, 10857–10865.
- 56 Jenkins, C. L.; Vasbinder, M. M.; Miller, S. J.; Raines, R. T. *Org. Lett.* **2005**, *7*, 2619–2622.
- 57 Boryskina, O. P.; Bolbukh, T. V.; Semenov, M. A.; Gasan, A. I.; Maleev, V. Y. *J. Mol. Struct.* **2007**, *827*, 1–10.
- 58 Erdmann, R. S. and Wennemers, H. *Angew. Chem. Int. Ed.* **2011**, *123*, 6967–6970.
- 59 Bartlett, G. J.; Choudhary, A.; Raines, R. T.; Woolfson, D. N. *Nat. Chem. Bio.* **2010**, *8*, 615–620.
- 60 (a) DeRider, M. L. *et. al. J. Am. Chem. Soc.* **2002**, *124*, 2497–2505. (b) Hinderaker, M. P.; Raines, R. T. *Protein Sci.* **2003**, *12*, 1188–1194.
- 61 Jenkins, C. L.; Lin, G.; Duo, J.; Rapolu, D.; Guzei, I. A. *J. Org. Chem.*, **2004**, *69*, 8565–8573.
- 62 Hodges, J. A. and Raines, R. T. *Org. Lett.*, **2006**, *8*, 4695–4697.
- 63 Pauling, L.; Corey, R. B.; Branson, H. R., *Proc. Natl. Acad. Sci. USA* **1951**, *37*, (4), 205–211.
- 64 (a) Radzicka, A.; Pedersen, L.; Wolfenden, R., *Biochemistry* **1988**, *27*, 4538–4541. (b) Li, P.; Chen, X. G.; Shulin, E.; Asher, S. A., *J. Am. Chem. Soc.* **1997**, *119*, 1116–1120.
- 65 Scherer, G.; Kramer, M. L.; Schutkowski, M.; Reimer, U.; Fischer, G., *J. Am. Chem. Soc.* **1998**, *120*, 5568–5574.
- 66 (a) Gratwohl, C.; Wuthrich, K., *Biopolymers* **1976**, *15*, 2025–2041. (b) Reimer, U.; Scherer, G.; Drewello, M.; Kruber, S.; Schutkowski, M.; *J. Mol. Biol.* **1998**, *279*, 449–460.
- 67 Schoetz, G.; Trapp, O.; Schurig, V., *Electrophoresis* **2001**, *22*, 2409–2415.
- 68 Renner, C.; Alefelder, S.; Bae, J. H.; Budisa, N.; Huber, R.; Moroder, L., *Angew. Chem. Int. Ed.* **2001**, *40*, 923–925.
- 69 Sarkar, S. K.; Young, P. E.; Sullivan, C. E.; Torchia, D. A., *Proc. Natl. Acad. Sci. USA* **1984**, *81*, 4800–4803.
- 70 (a) Li, S. T.; Ggloub, E.; Kartz, E. P. *J. Mol. Biol.* **1975**, *98*, 835–839. (b) Traub, B. L.; Piez, K. A. *J. Mol. Biol.* **1976**, *108*, 705–732. (c) Kartz, E. P.; David, C. W. *Biopolymers* **1990**, *29*, 791–798. (d) Kartz, E. P.; David, C. W. *J. Mol. Biol.* **1992**, *228*, 963–969.
- 71 Cohen, C.; Parry, D. A. D. *Trends Biochem.Sci.* **1988**, *11*, 245–248.
- 72 (a) Traub, W.; Fietzek, P.P. *FEBS Letters* **1976**, *68*, 245–249. (b) Jones, E. Y.; Miller, A. *J. Mol. Biol.* **1991**, *218*, 209–219.
- 73 Bruns, R. R.; Gross. *J. Biochemistry* **1973**, *12*, 808–815.

- 74 Barber, H. E.; Mackay, D. *Arch. Biochem. Biophys.* **1996**, *177*, 466-468.
- 75 (a) Halmes, D. J. S.; Miller, A.; Parry, D. A. D.; Piez, D. A.; Woodhead- Galloway, J. *J. Mol. Biol.* **1973**, *79*, 137-148. (b) Trus, B. L.; Piez, K. A. *J. Mol. Biol.* **1976**, *108*, 705-732. (c) Li, M.; Fan, P.; Brodsky, B.; Baum, J. *Biochemistry* **1993**, *32*, 7377-7387.
- 76 Berg, R. A.; Olsen, B. R.; Prockop, D. J.; Klump, H. *J. Biol. Chem.* **1970**, *245*, 5759-5763.
- 77 Venugopal, M. G.; Ramshaw, J. A. M.; Braswell, E.; Zhu, D.; Brodsky, B. *Biochemistry* **1994**, *33*, 7948-7956.
- 78 Cansonni, R.; Santomo, L.; Tenni, R.; Longhi, R.; Zetta, L. *FEBS Letters* **1998**, *436*, 243-246.
- 79 Ikura, T.; Urakubo, Y.; Nobutoshi, I. *Chem. Phys.* **2004**, *307*, 111-119.
- 80 Pacaroni, A.; Cinelli, S.; Cornicchi, E.; de Francesco, A.; Onori, G. *Chem. Phys. Lett.* **2005**, *410*, 400-403.
- 81 Cheung, M. S.; Garcia, A. E.; Onuchic, J. N. *Proc. Nat. Acad. Sci. USA* **2002**, *99* 685-690.
- 82 (a) Gerisma, S. Y.; Stuur, E. R. *Int. J. Pept. Protein Res.* **1972**, *4*, 377-383. (b) Back, J. F.; Oakenfull, D.; Smith, M. B. *Biochemistry* **1979**, *18*, 5191-5196. (c) Arakawa, T.; Timasheff, S. N. *Biochemistry* **1982**, *21*, 6536-6544.
- 83 (a) Harrap, B. S. *Int. J. Pept. Protein Res.* **1969**, *1*, 2527-2532. (b) Gekko, K.; Koga, S. *J. Biochem.* **1983**, *94*, 199-205. (c) Feng, Y.; Melacini, G.; Taulane, J. P.; Goodman, M. *J. Am. Chem. Soc.* **1996**, *118*, 10351-10358.
- 84 Brodsky, B.; Thiagarajan, G.; Madhan, B. and Kar K., *Biopolymers*, **2008**, *89*, 345-353.
- 85 (a) Holmgren, S. K.; Taylor, K. M.; Bretscher, L. E. and Raines R. T. *Nature*, **1998**, *392*, 666-667. (b) Holmgren, S. K.; Taylor, K. M.; Bretscher, L. E. and Raines R. T. *Chem. Biol.*, **1999**, *6*, 63-70.
- 86 (a) Sonntag, L.S.; Schweizer, S.; Ochsenfeld, C. and Wennemers, H. *J. Am. Chem. Soc.*, **2006**, *128*, 14697-14703.
- 87 Bann, J. G.; Peyton, D. H. and Bachinger H. P. *FEBS Lett.*, **2000**, *473*, 237-240.
- 88 (a) Fields, C. G.; Mickelson, D. J.; Drake, S. L.; McCarthy, J. B. and Fields G. B. *J. Biol. Chem.*, **1993**, *268*, 14153-14160. (b) Feng, Y.; Melacini, G.; Taulane, J. P. and Goodman M. *J. Am. Chem. Soc.*, **1996**, *118*, 10351-10358. (c) Kwak, J.; DeCapua, A.; Locardi, E. and Goodman M. *J. Am. Chem. Soc.*, **2002**, *124*, 14085-14091.
- 89 Cai, W.; Wong, D.; Kinberger, G. A.; Kwok, S. W.; Taulane, J. P. and Goodman, M., *Bioorg. Chem.*, **2007**, *35*, 327-337.
- 90 Inouye, K.; Sasakibara, S.; Prockop, D. J. *Biochim. Biophys. Acta*, **1976**, *420*, 133-141.
- 91 Fairwather, R.; Jones, J. H. *J. Chem. Soc. Perkin I* **1972**, 2475-2481.

- 92 Hutton, J. J.; Marglin, A.; Witkop, B.; Kurtz, J.; Berger, A.; Udenfreind, S. *Arch. Biochim. Biophys.* **1968**, *125*, 779-785.
- 93 Zagari, A.; Nemethy, G.; Scheraga, H. A. *Biopolymers*, **1990**, *30*, 967.
- 94 Li S-T.; Golub, E.; Katz, E. P. *J. Mol. Biol.* **1975**, *98*, 791.
- 95 Katz, E. P.; David, C. W. *Biopolymers* **1990**, *29*, 791.
- 96 Lee S. G., Lee J. Y., and Chmielewski J., *Angew. Chem. Int. Ed.* **2008**, *47*, 8429–8432.
- 97 Erdmann R. S. and Wennemers H., *J. Am. Chem. Soc.* **2010**, *132*, 13957–13959.
- 98 Alexander, R. M. *Sci. Prog.* **1981**, *67*, 109-130.
- 99 (a) Trop, S.; Arridge, R. S. C.; Armenides, C. D.; Baer, E. In *Structure of Fibrous Biopolymers* Atkins, E. D. T.; Keller, A. Eds. **1975**, 26, pp. 197-221, Butterworth, London. (b) Uitto, J. *Dermatol. Clin.* **1986**, *4*, 433-436.
- 100 Bailey, A. J.; Paul, R. G.; Knott, L. *Mechanisms of Ageing and Devel.* **1998**, *106*, 1-56.
- 101 (a) Hodge, H. J. *Agric. Food Sci.* **1953**, *1*, 928-943. (b) Robins, S. P.; Bailey, A. *J. Biochem. Biophys. Res. Commun.* **1972**, *48*, 76-84. (c) Siegel, R. C.; *Int. Rev. Connect. Tissue Res.* **1979**, *8*, 73-188
- 102 Kato, H.; Hayase, F.; Shin, D. B.; Oimomi, M.; Baba, S. In *The Maillard Reaction in Aging Diabetes and Nutrition* Baynes, J. W.; Monnier, V. M. Eds. **1989**, pp. 69-84.
- 103 (a) Sell, D. R.; Monnier, V. M. *J. Biol. Chem.* **1989**, *264*, 21597-21602. (b) Dyer, D. G.; Blackledge, J. A.; Thorpe, S. R.; Baynes, J. W. *J. Biol. Chem.* **1991**, *266*, 11654-11660. (c) Grandhee, S. K.; Monnier, V. M. *J. Biol. Chem.* **1991**, *266*, 11649-11653.
- 104 Fledelius, C.; Johnsen, A. H.; Cloos, P.A.C; Bonde, M; Qvist, P. *J. Biol. Chem.* **1997**, *272*, 9755-9763.
- 105 Myllyharju, J.; Kivirikko, K. *I. Ann. Med.* **2001**, *33*, 7-21.
- 106 Prockop, D. *J. Matrix Biol.* **1998**, *16*, 519-528.
- 107 Krisch, E.; Krieg, T.; Remberger, K.; Fendel, H.; Bruckner, P.; Muller, P. K. *Eur. J. Clin. Investig.* **1981**, *11*, 39-47.
- 108 Lee, C. H.; Singla, A.; Lee, Y. *Inte. J. Phar.*, **2001**, *221*, 1–22.
- 109 Johnson, G.; Jenkins, M.; McLean, K. M.; Griesser, H.J.; Kwak, J. *et al. J. Biomed. Mater. Res.*, **2000**, *51*, 612–624.
- 110 (a) Cejas, M. A.; Chen, C.; Kinney, A.; Maryanoff, B. E. *Bioconjug. Chem.*, **2007**, *18*, 1025–1027. (b) Smethurst, P. A. *et al. J. Biol. Chem.*, **2007**, *282*, 1296–1304.

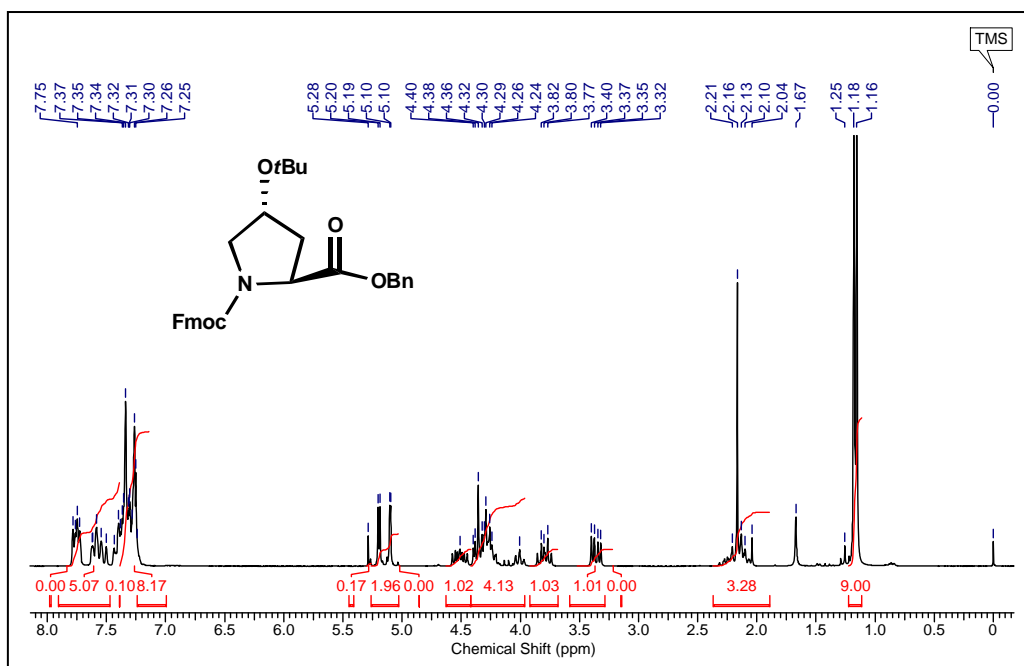
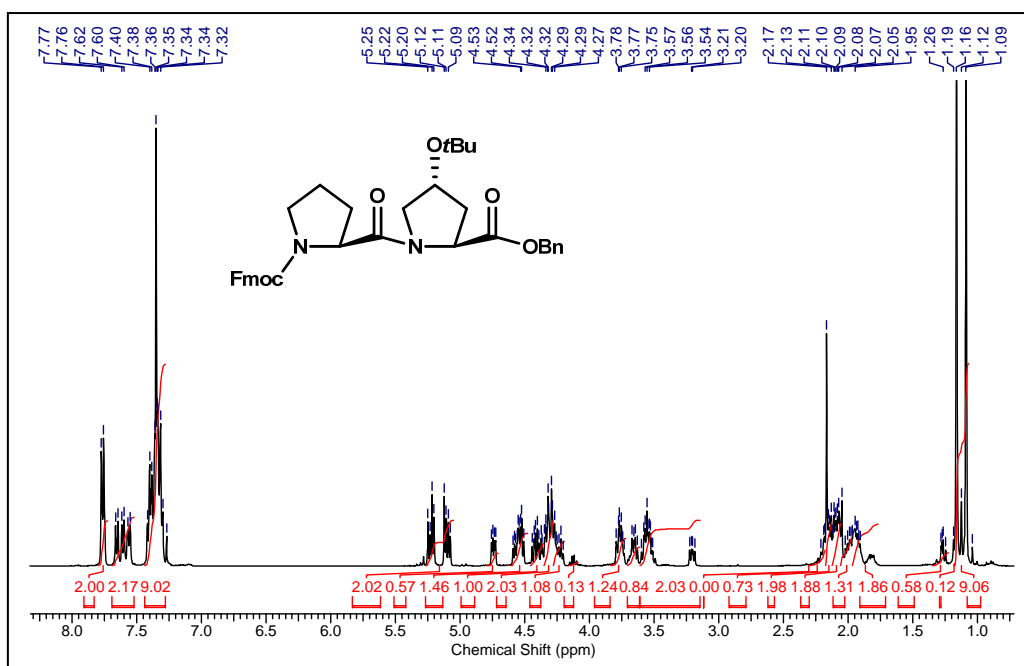
- 111 Mo, X.; An, Y.; Yun, C-S.; Yu, S. M. *Angew. Chem. Int. Ed.*, **2006**, *45*, 2267–2270.
- 112 Cejas, M. A. *et al. Proc. Natl. Acad. Sci. USA*, **2008**, *105*, 8513–8518.
- 113 (a) Wang, A.Y.; Mo, X.; Chen, C. S.; Yu, S. M.; *J. Am. Chem. Soc.* **2005**, *127*, 4130–4131. (b) Wang, A.; Foss C.; Leong S, Pomper, M. G.; Yu, S. M. *Biomacromolecules*, **2008**, *9*, 1755–1763.
- 114 Auger, F.A.; Rouabhia, M.; Goulet, F.; Berthod, F.; Moulin, V.; Germain, L. *Med. Biol. Eng. Comput.* **1998**, *36*, 801–812.
- 115 Miyata, T.; Taira, T.; Noishiki Y. *Clin. Mater.*, **1992**, *9*, 139–148.
- 116 Wang, C.L.; Miyata, T.; Weksler, B.; Rubin, A.L.; Stenzel, K.H. *Biochim. Biophys. Acta.* **1978**, *544*, 555–567.
- 117 Van der Laan, J.S.; Lopez, G.P.; VanWachem, P.B.; Nieuwenhuis, P.; Ratner, B.D.; Bleichrodt, R.P.; Schakenraad, J.M. *Int. J. Artif. Organs*, **1991**, *14*, 661–666.
- 118 Sekine, T.; Nakamura, T.; Shimizu, Y.; Ueda, H.; Matsumoto, K. *J. Biomed. Mater. Res.* **2001**, *54*, 305–310.
- 119 Eyre, D. R.; Wu, J. *Top Curr Chem*, **2005**, *247*, 207–229.
- 120 Okuyama, K.; Arnott, S.; Takayanagi, M.; Kakudo, M. *J. Mol. Biol.* **1981**, *152*, 427–443.
- 121 (a) Brodsky, B.; Persikov, A. V. *Adv Protein Chem* **2005**, *70*, 301–339. (b) Okuyama, K.; Wu, G.; Jiravanichanun, N.; Hongo, C.; Noguchi, K. *Biopolymers* **2006**, *84*, 421–432. (c) Okuyama, K.; Narita, H.; Kawaguchi, T.; Noguchi, K.; Tanaka, Y.; Nishino, N. *Biopolymers* **2007**, *86*, 212–221.
- 122 (a) Myllyharju, J.; Kivirikko, K. I. *Trends Genet.* **2004**, *20*, 33. (b) Lauer-Fields, J. L.; Juska, D.; Fields, G. B. *Biopolymers* **2002**, *66*, 19. (c) Di Lullo, G. A.; Sweeney, S. M.; Korkko, J.; Ala-Kokko, L.; San Antonio, J. D. *J. Biol. Chem.* **2002**, *277*, 4223. (d) Li, Y.; Brodsky, B.; Baum, J. *J. Biol. Chem.* **2009**, *284*, 20660.
- 123 (a) Withka, J.M.; Swaminathan, S. & Bolton, P.H. *J. Magn. Res.* **1990**, *89*, 386–390. (b) Fan, P.; Li, M.-H.; Brodsky, B. & Baum, J. *Biochemistry*, **1993**, *32*, 13299–13309.
- 124 (a) Li, M.-H.; Fan, P.; Brodsky, B. & Baum, J. *Biochemistry*, **1993**, *32*, 7377–7387. (b) Brodsky, B.; Li, M.-H.; Long, C.G.; Apigo, J. & Baum, J. *Biopolymers*, **1992**, *32*, 447–451. (c) Melacini, G.; Feng, Y. & Goodman, M. *J. Am. Chem. Soc.* **1996**, *118*, 10359–10364. (d) Melacini, G.; Feng, Y. & Goodman, M. *J. Am. Chem. Soc.* **1996**, *118*, 10725–10732.
- 125 Fallas, J. A.; Gauba, V. and Hartgerink J. D. *JBC*, **2009**, *84*, 26851–26859.
- 126 Melacini, G.; Bonvin, J.; M. Goodman, R.; Kaptein, R.; *J. Mol. Biol.*, **2000**, *300*, 1041.
- 127 Xiao, J. and Baum, J., *J. Am. Chem. Soc.* **2009**, *131*, 18194–18195.
- 128 Leah, C.; Erin Z.; Bernardo P.; Kaplan, D., *Journal of Biomedical Materials Research Part B: Applied Biomaterials*, **2008**, 264–285.

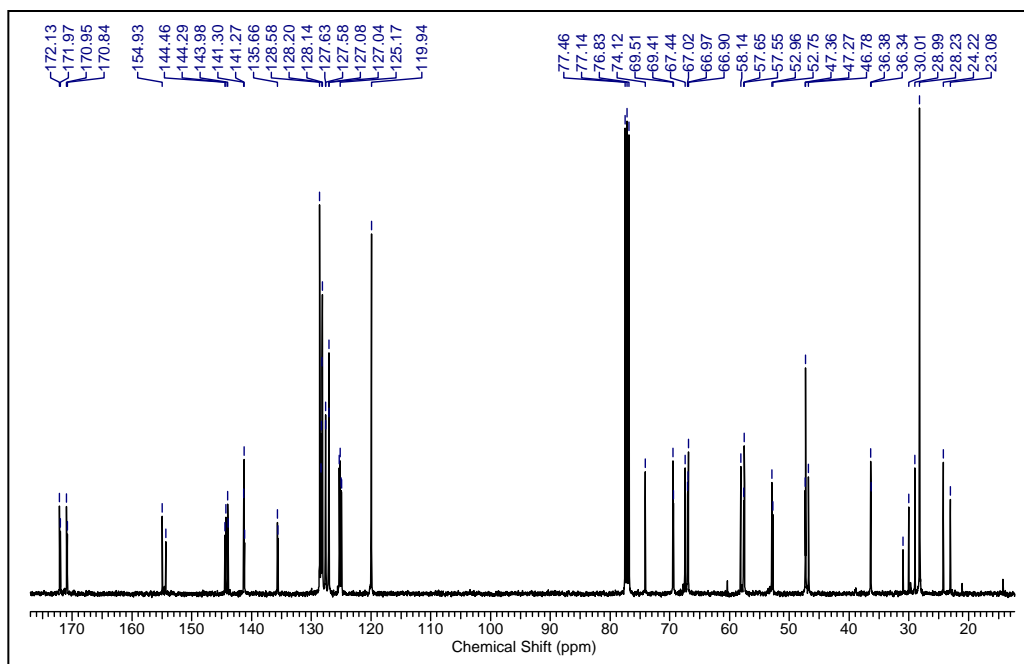
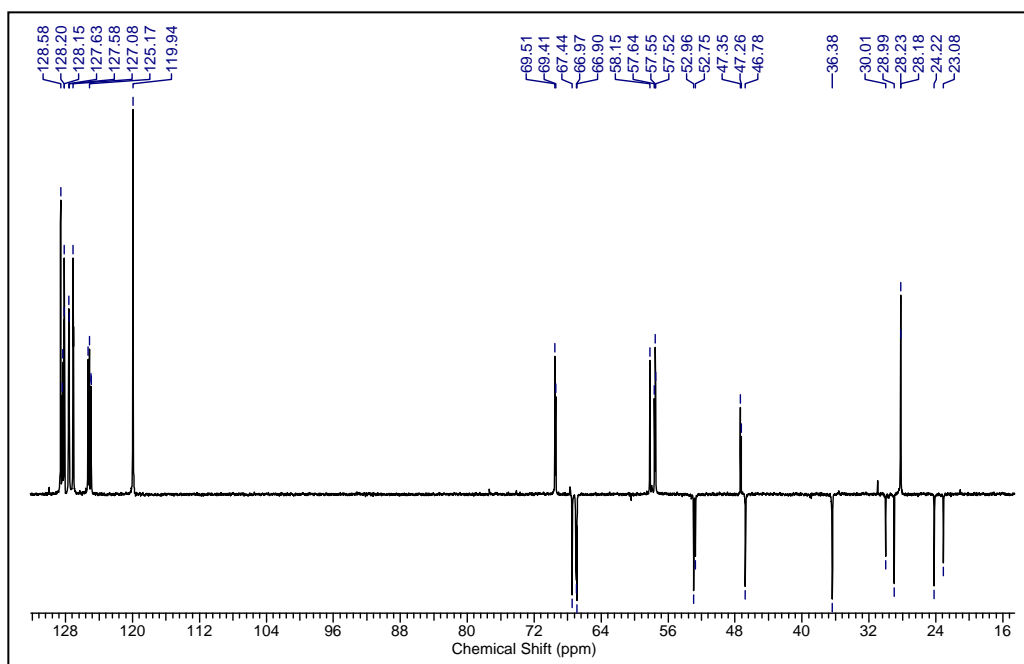
- 129 (A) Dreisewerd, K.; Rohlfing, A.; Spottke, B.; Urbanke, C.; Henkel, W. *Anal Chem.* **2004**, *76*, 3482–3491. (B) Schonherr, E.; Witsch-Prehm, P.; Harrach, B.; Robenek, H.; Rauterberg, J.; Kresse, H. *J Biol Chem*, **1995**, *270*, 2776–2783.
- 130 Wei, L.; Sun, X, J.; Wang, Z.; Chen, Q. *Arthr Res Ther*, **2006**, *8*, 37.
- 131 (a) Kobayashi, Y.; Isemura, T. *Progress in Polymer science* **1992** Vol. 3, Okamura, S.; Takayanagi, M.; Eds., Kodnasha Limited, Tokyo & John Wiley & Sons, NY, pp 315-380. (b) Inuoye, K.; Sakakibara, S.; Prockop, D. J. *Biochem. Biophys. Acta* **1976**, *420*, 133-141. (c) Weber, R. W.; Nitschmann, H. *Helv. Chem. Acta* **1978**, *61*, 701-708. (d) Inuoye, K.; Kobayashi, Y.; Kyogoku, Y.; Kisida, Y.; Sakakibara, S.; Prockop, D. J. *Arch. Biochem. Biophys.* **1982**, *219*, 198-203. (e) Engel, J.; Chen, H.-T.; Prockop, D. J.; Klump, H. *Biopolymers* **1977**, *16*, 601-622.
- 132 Yujia, X. *Methods in Enzymology*, **2009**, *466*, 211-232.
- 133 Jelesarov, I.; Bosshard, H. R. *J Mol Recogn*, **1999**, *12*, 3–18.
- 134 Flandin, F.; Buffevant, C.; Herbage, D. *Biochim Biophys Acta*, **1984**, *791*, 205–211.
- 135 Kumar, T. R.; Shanmugasundaram. N.; Babu, M. *J Biomater Sci Polym Ed*, **2003**, *14*, 689–706.
- 136 Hickman, D. *et.al. J Biotechnol*, **2000**, *79*, 245–257.
- 137 Cheng, Z.; Teoh, S. H. *Biomaterials*, **2004**, *25*, 1991–2001.
- 138 Wilson, D.L.; Martin, R.; Hong, S.; Cronin-Golomb, M.; Mirkin, C. A.; Kaplan, D. L. *Proc Natl Acad Sci USA*, **2001**, *98*, 13660–13664.
- 139 Cristofalo, V. J.; Pignolo, R. J.; *Physiol Rev*, **1993**, *73*, 617–638.
- 140 Qiu, Q.; Sayer, M.; Kawaja, M.; Shen, X.; Davies. J. *J Biomed Mater Res*, **1998**, *42*, 117–127.
- 141 Baguneid, M. *et. al. Biotechnol Appl Biochem*, **2004**, *39*, 151–157.
- 142 (a) Holmgren, S. K.; Taylor, K. M.; Bretscher, L. E.; Raines, R. T. *Nature*, **1998**, *392*, 666-667. (b) Hodges, J.; Ronald T. Raines, *J. Am. Chem. Soc.* **2003**, *125*, 9262-9263.
- 143 (a) Babu, I. R.; Ganesh, K. N. *J. Am. Chem. Soc.* **2001**, *123*, 2079-2080.
- 144 Nishiuchi, Y. *et.al. PNAS*, **1998**, *95*, 13549-13554.
- 145 Hievl, J. *et. al. The Journal of Peptide Research*, **1999**, *54*, 54-65.
- 146 (a) Lee, S.-G.; Lee, J. Y.; Chmielewski, J. *Angew. Chem. Int. Ed.* **2008**, *47*, 8429. (b) Dai, N.; Wang, X. J.; Etkorn, F. A. *J. Am. Chem. Soc.* **2008**, *130*, 5396. (c) Gauba, V.; Hartgerink, J. D. *J. Am. Chem. Soc.* **2008**, *130*, 7509. (d) Gauba, V.; Hartgerink, J. D. *J. Am. Chem. Soc.* **2007**, *129*, 2683. (e) Gauba, V.; Hartgerink, J. D. *J. Am. Chem. Soc.* **2007**, *129*, 15034. (f) Bachmann, A.; Kiefhaber, T.; Boudko, S.; Engel, J.; Bachinger, H. P. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13897. (g) Beck, K.; Chan, V. C.; Shenoy, N.; Kirkpatrick, A.; Ramshaw, J. A. M.; Brodsky, B. *Proc. Natl. Acad. Sci. U.S.A.*

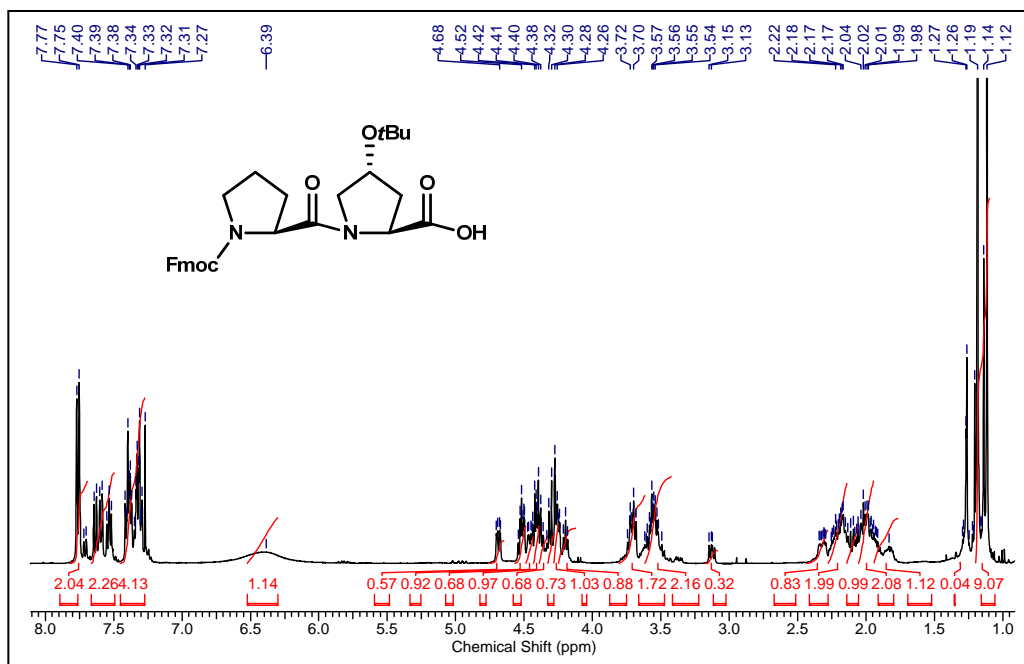
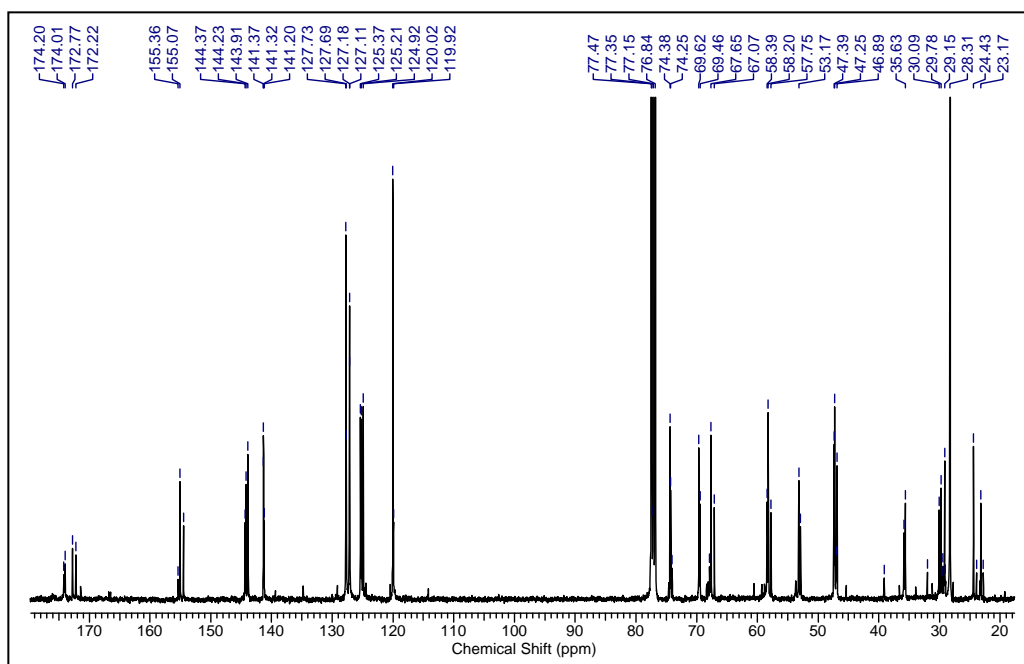
- 2000, 97, 4273. (h) Persikov, A. V.; Ramshaw, J. A. M.; Brodsky, B. *Biopolymers* **2000**, 55, 436. (i) Roman S. Erdmann, Helma Wennemers, *Synthesis* **2009**, 143-147.
- 147 (a) Cejas, M. A.; Kinney, W. A.; Chen, C.; Vinter, J. G.; Almond, H. R. Jr.; Balss, K. M.; Maryanoff, C. A.; Schmidt, U.; Breslav, M.; Mahan, A.; Lacy, E.; Maryanoff, B. E. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, 105, 8513. (b) Przybyla, D. E.; Chmielewski, J. *J. Am. Chem. Soc.* **2008**, 130, 12610. (c) Rele, S.; Song, Y.; Apkarian, R. P.; Qu, Z.; Conticello, V. P.; Chaikof, E. L. *J. Am. Chem. Soc.* **2007**, 129, 14780. (d) Cejas, M. A.; Kinney, W. A.; Chen, C.; Leo, G. C.; Tounge, B. A.; Vinter, J. G.; Joshi, P. P.; Maryanoff, B. E. *J. Am. Chem. Soc.* **2007**, 129, 2202. (e) Kar, K.; Amin, P.; Bryan, M. A.; Persikov, A. V.; Mohs, A.; Wang, Y.-H.; Brodsky, B. *J. Biol. Chem.* **2006**, 281, 33283. (f) Kotch, F. W.; Raines, R. T. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, 103, 3028.
- 148 (a) Ottl, J.; Musiol, H. J.; Moroder, L. *J. Pept. Sci.* **1999**, 5, 103. (b) Inouye, K.; Sakakibara, S.; Prockop, D. J. *Biochim. Biophys. Acta* **1976**, 420, 133. (c) Sakakibara, S.; Inouye, K.; Shudo, K.; Kishida, Y.; Kobayashi, Y.; Prockop, D. J. *Biochim. Biophys. Acta* **1973**, 303, 198.
- 149 For a monograph on solution-phase methods see Bodansky, M.; Bodansky, A. *The practice of peptide synthesis* **1984**, Springer-Verlog, Berlin.
- 150 Vojtkovsky, T. *Pept. Res.* **1995**, 8, 236-237.
- 151 Kaiser, E.; Colescott, R.L.; Bossinger, C.D.; Cook, P.I. *Anal. Biochem.* **1970**, 34, 595-598.
- 152 Brown, F. R.; III; Di Corato, A.; Lorenzi, G. P.; Blout, E. R. *J. Mol. Biol.* **1972**, 63, 85-99.
- 153 Mooney, S. D.; Kollman, P. A. and Klein, T. E. *Biopolymers*, **2001**, 64, 63.
- 154 Bretscher, L. E.; Jenkins, C. L.; Taylor, K. L.; DeRider, M. L.; Raines, R. T. *J. Am. Chem. Soc.* **2001**, 123, 777-778.
- 155 Improta, R.; Mele, F.; Crescenzi, O.; Benzi, C. and Barone, V. *J Am Chem Soc*, **2002**, 124, 7857-7865.
- 156 Kuemin, M.; Nagel, Y. A.; Schweizer, S.; Monnard, F. W.; Ochsenfeld, C. and Wennemers, H. *Angew Chem Int Ed*, **2010**, 49, 6324-6327.
- 157 Venugopal, M. G.; Ramshaw, J. A. M.; Braswell, E.; Zhu D.; Brodsky, B. *Biochemistry* **1994**, 33, 7948-7956.
- 158 *Collagen Biomaterials*, J. A. Werkmeister and J. A. M. Ramshaw, eds., Elsevier Science, Barking, Essex, 1992.

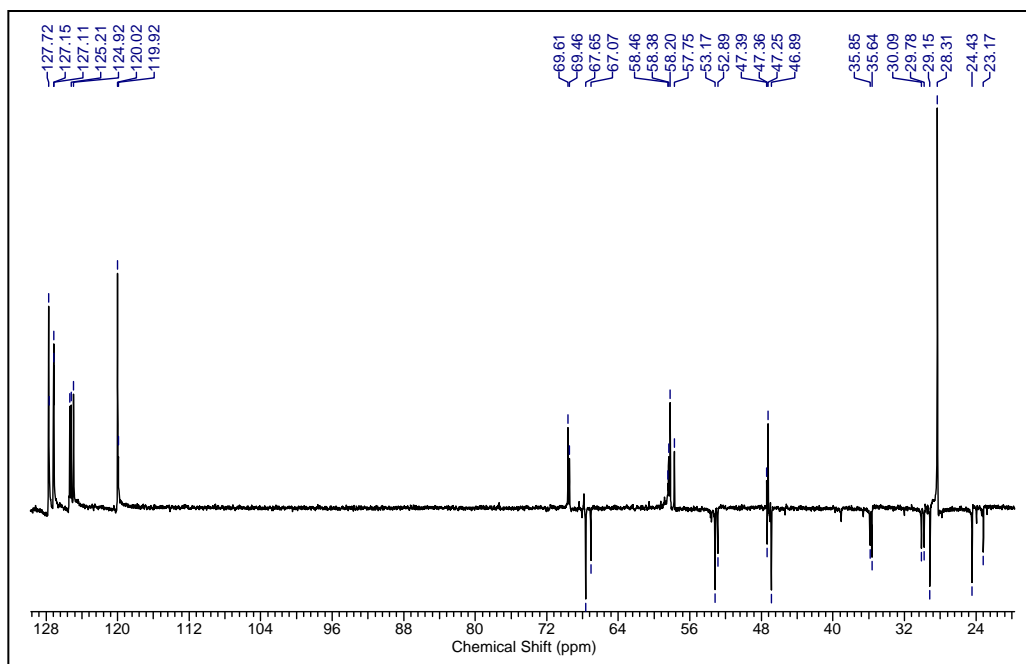
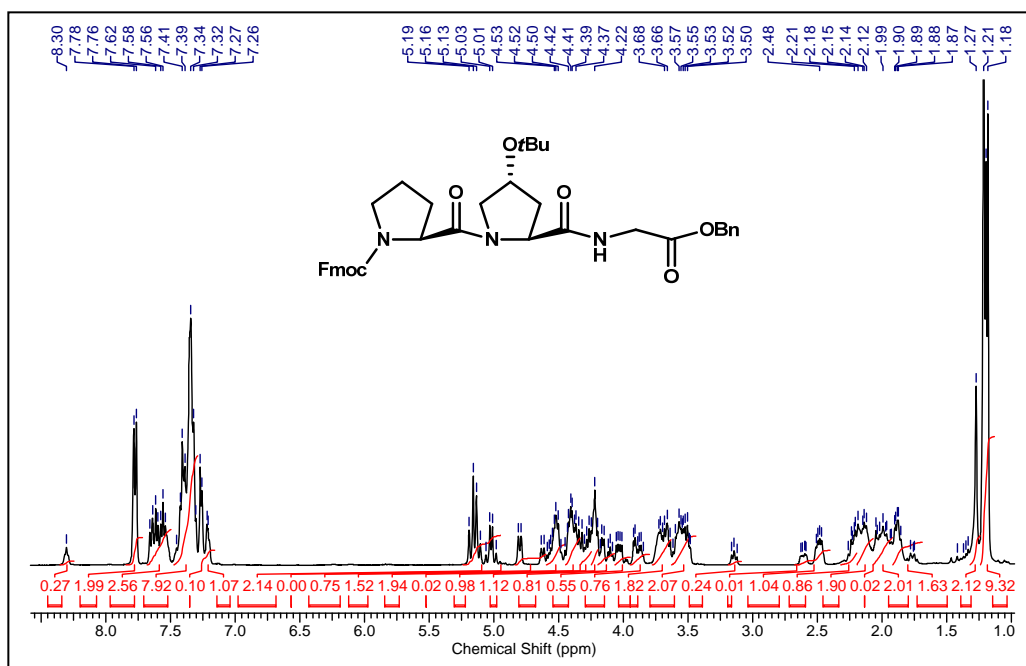
8 Appendix 2: Characterization data of synthesized compounds and peptides

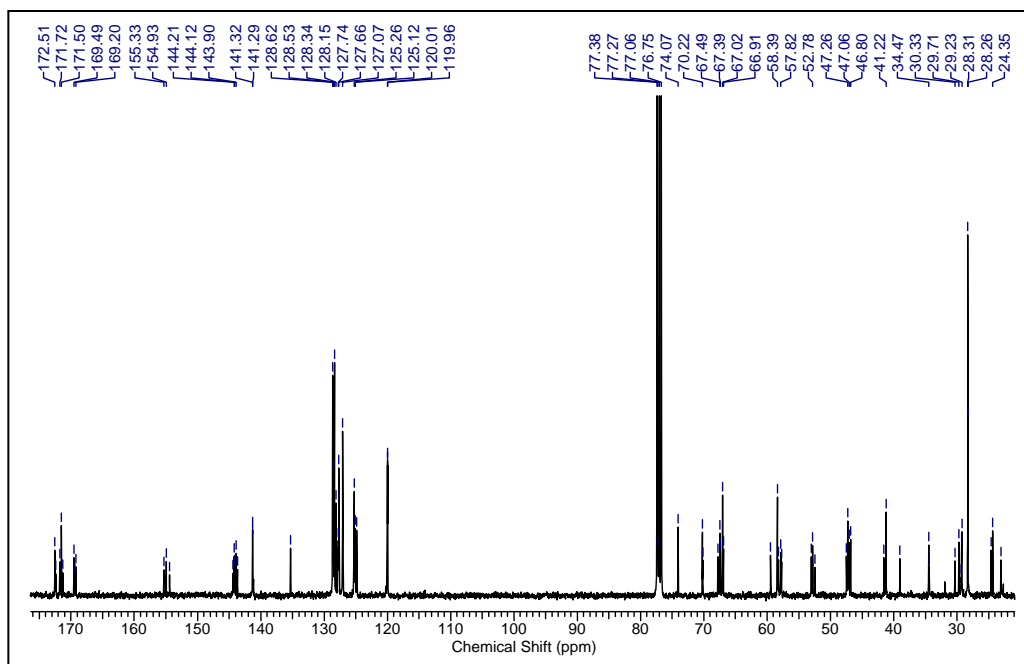
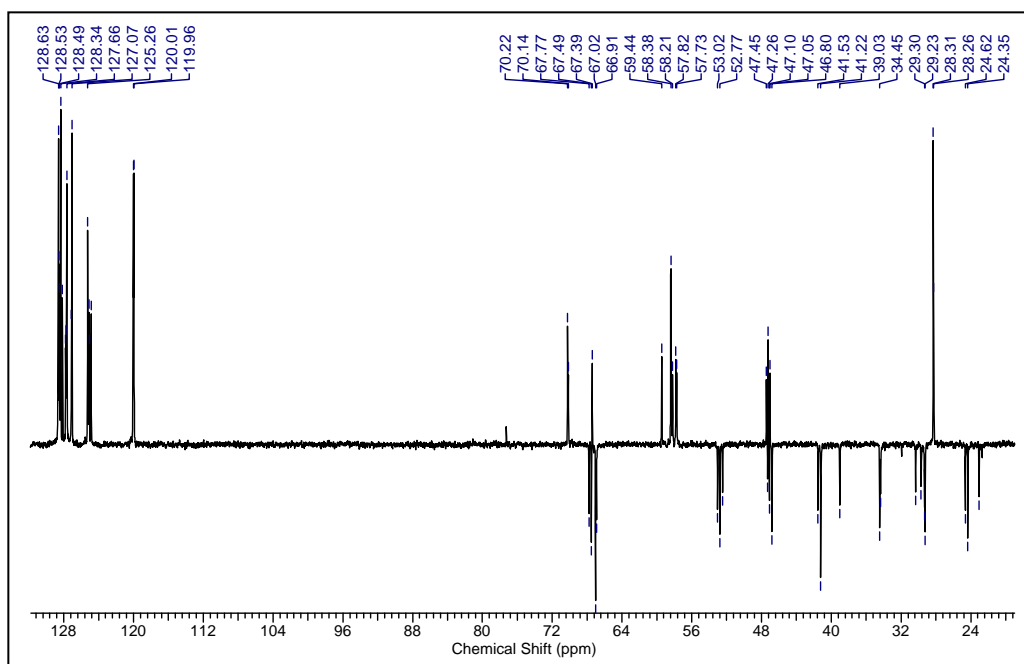
Entry	Page No.
^1H , ^{13}C and DEPT NMR spectra of compounds (15-41)	253-278
HPLC of Peptides (4-21)	278-284
MALDI-TOF of peptides (4-21)	285-293

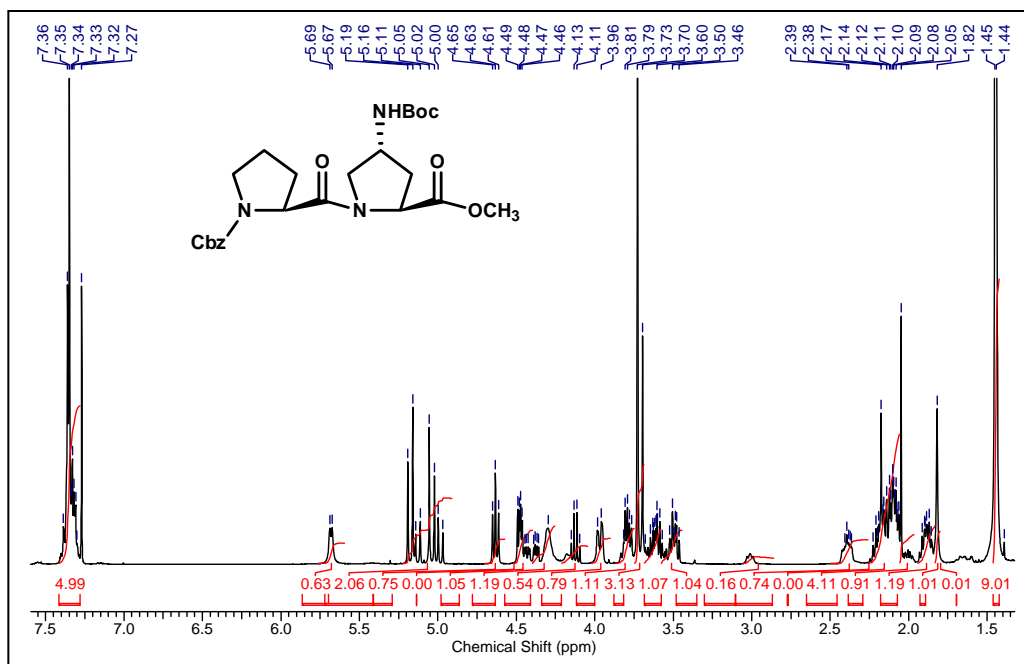
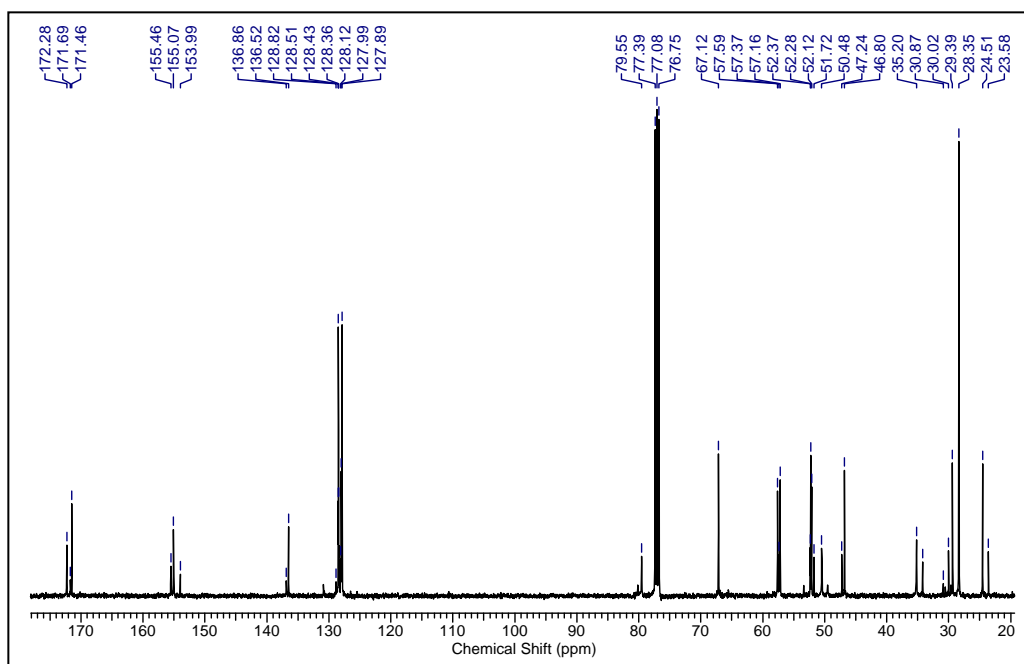
(A) Characterization data of synthesized compounds (15-41)**¹H NMR of compound 15****¹H NMR of compound 18**

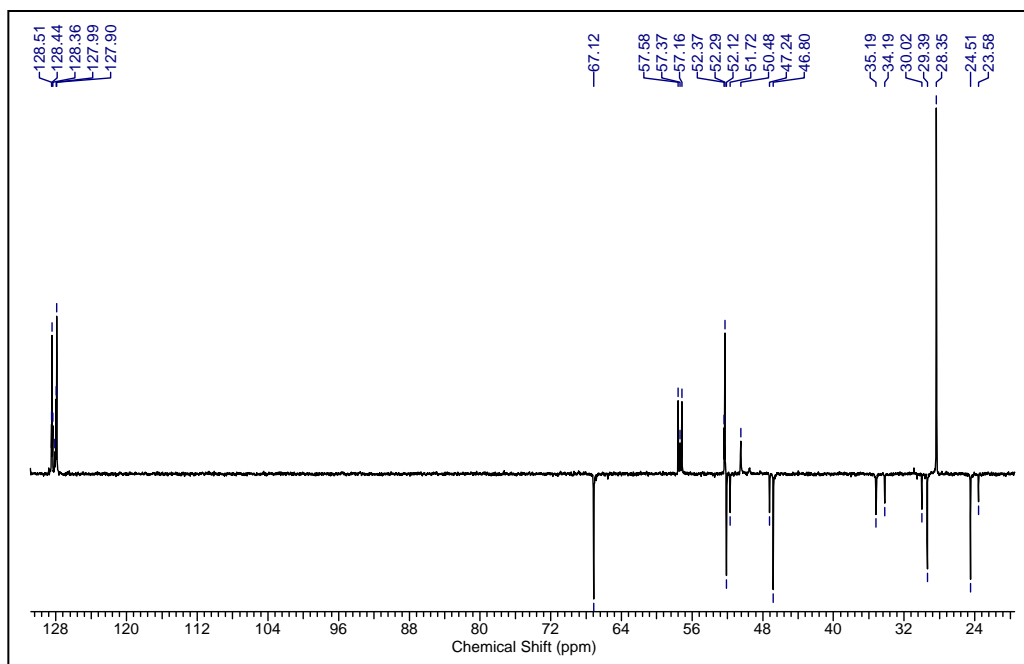
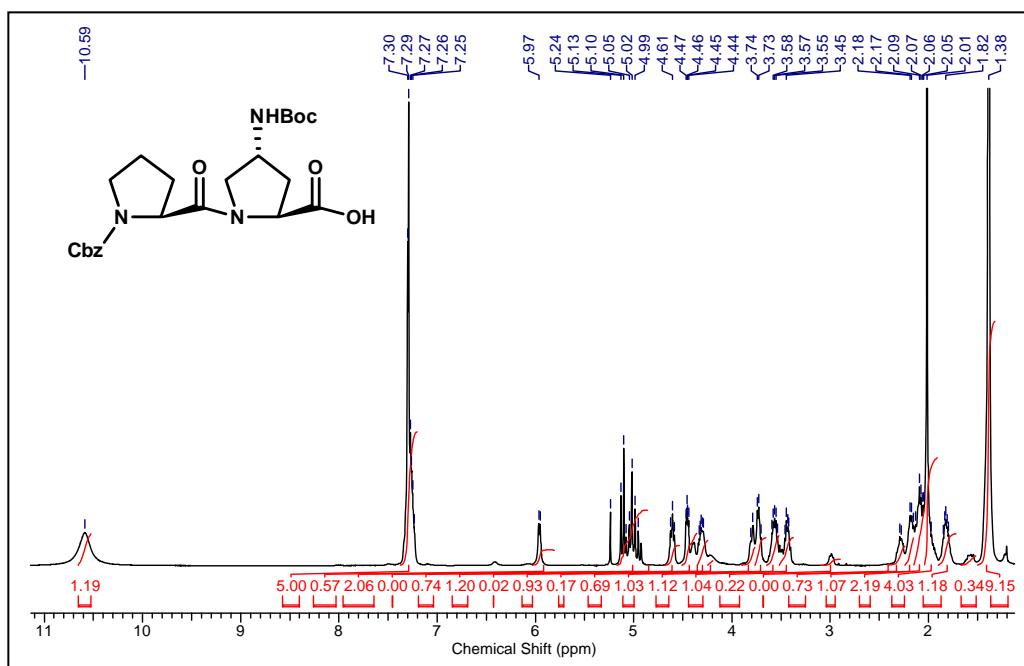
^{13}C NMR of compound 18**DEPT ^{13}C NMR of compound 18**

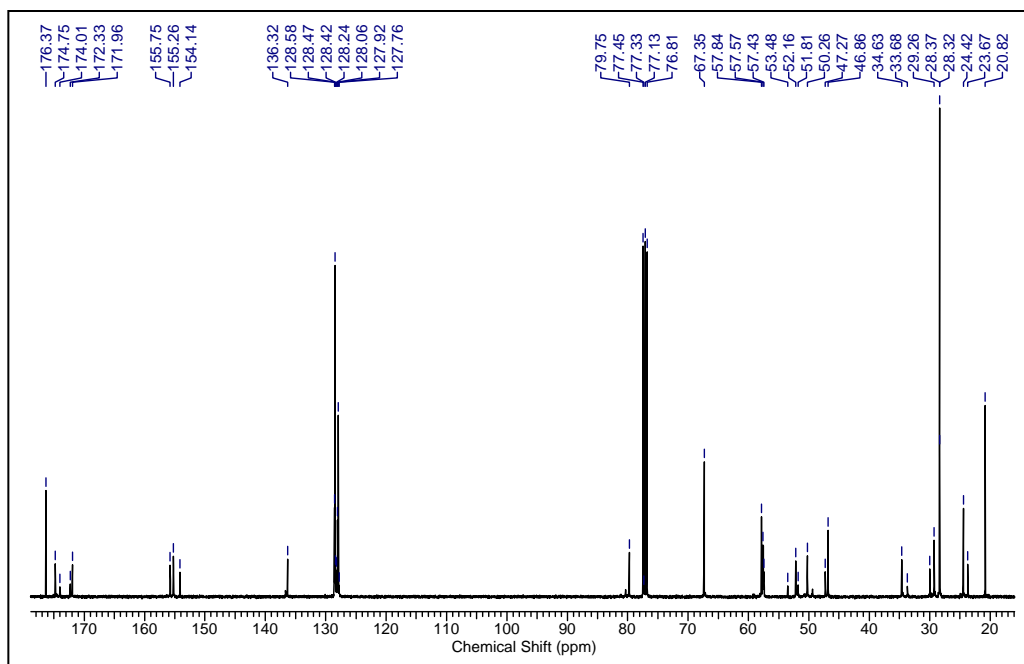
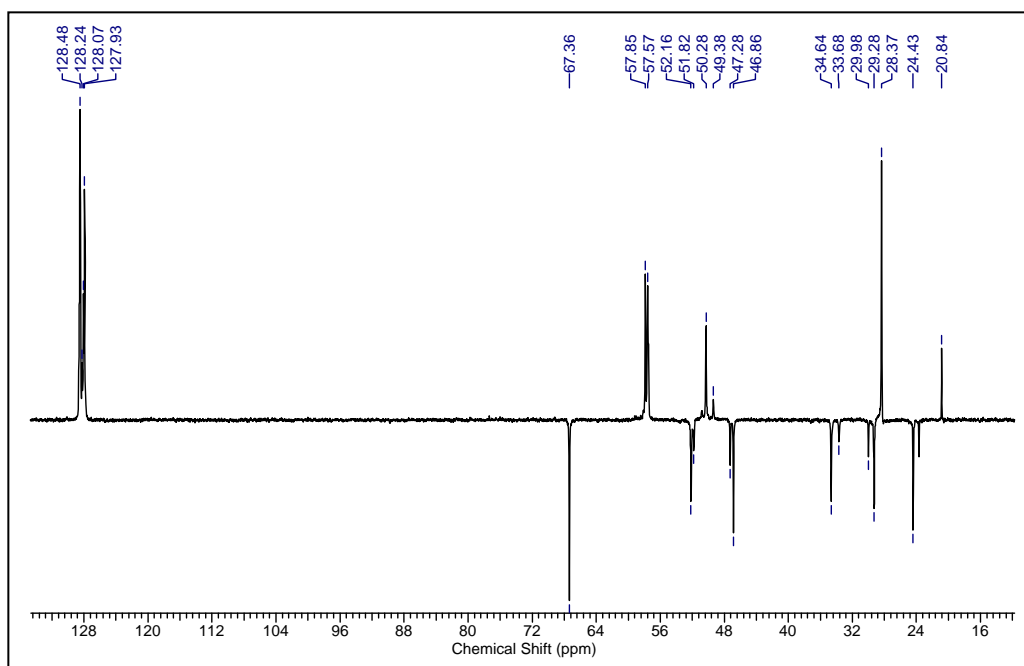
^1H NMR of compound 19 **^{13}C NMR of compound 19**

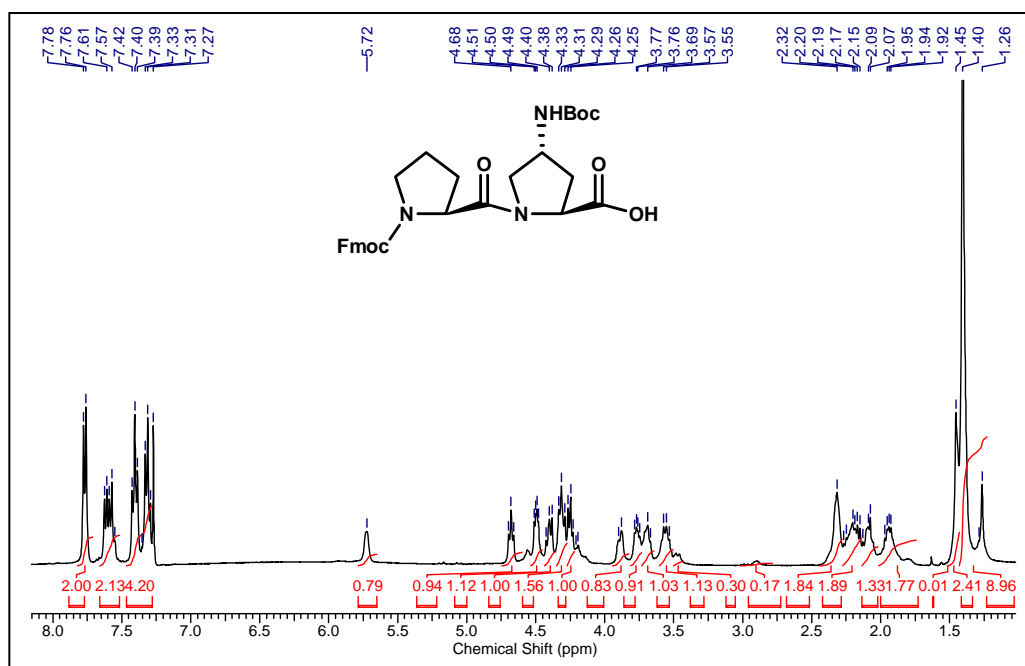
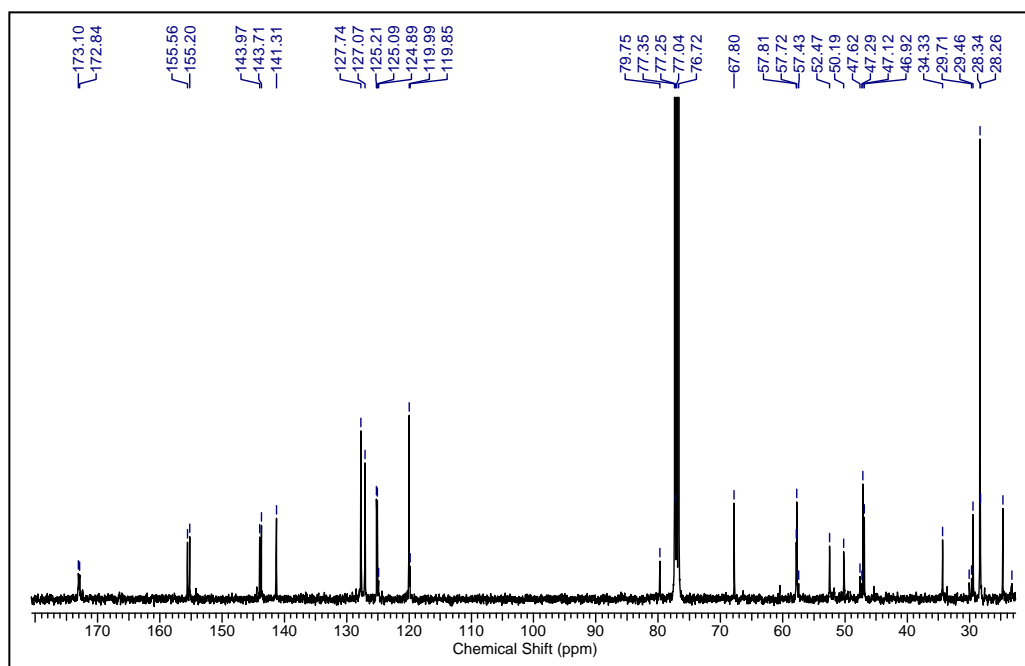
DEPT ^{13}C NMR of compound 19 ^1H NMR of compound 21

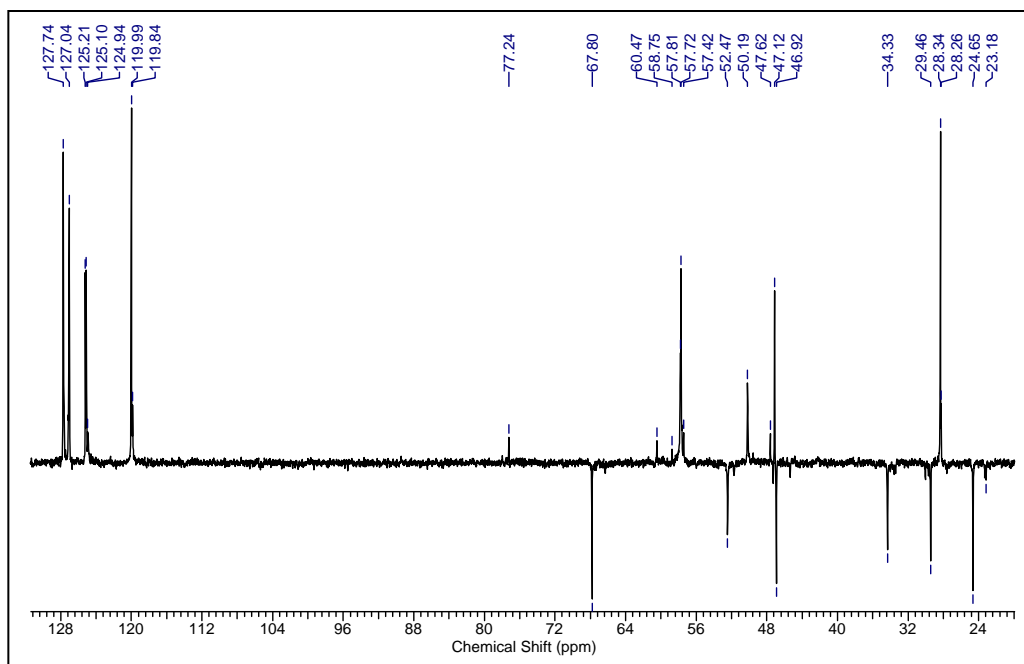
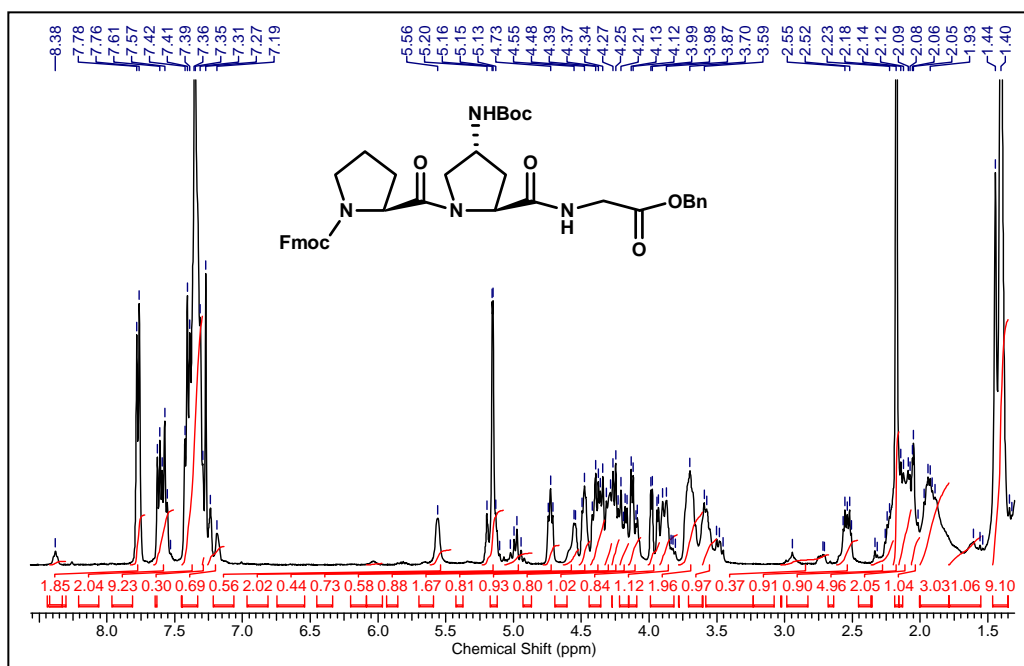
^{13}C NMR of compound 21**DEPT ^{13}C NMR of compound 21**

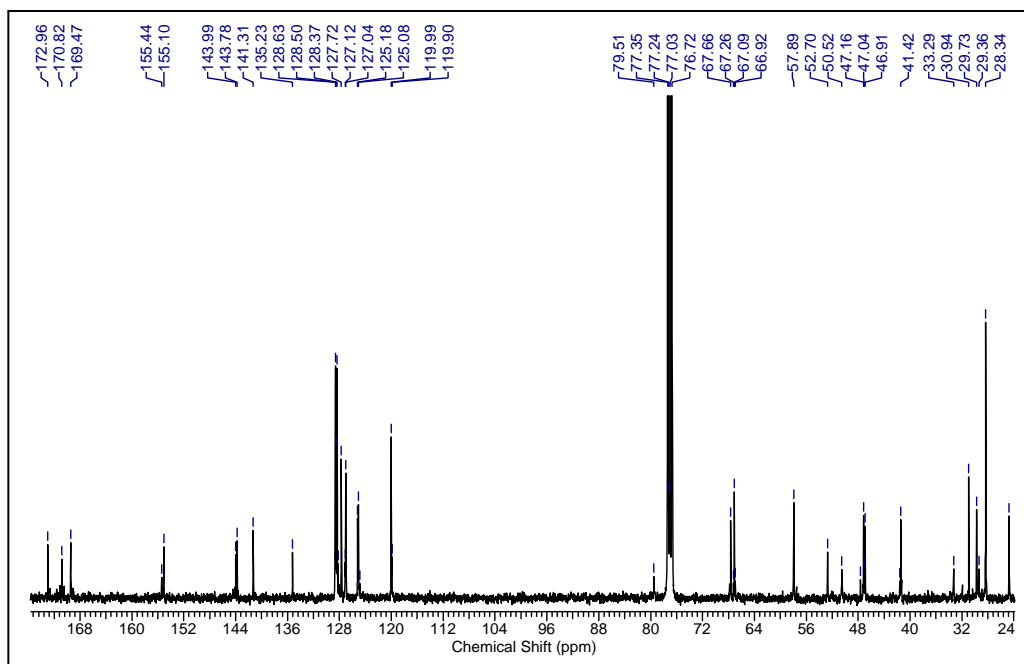
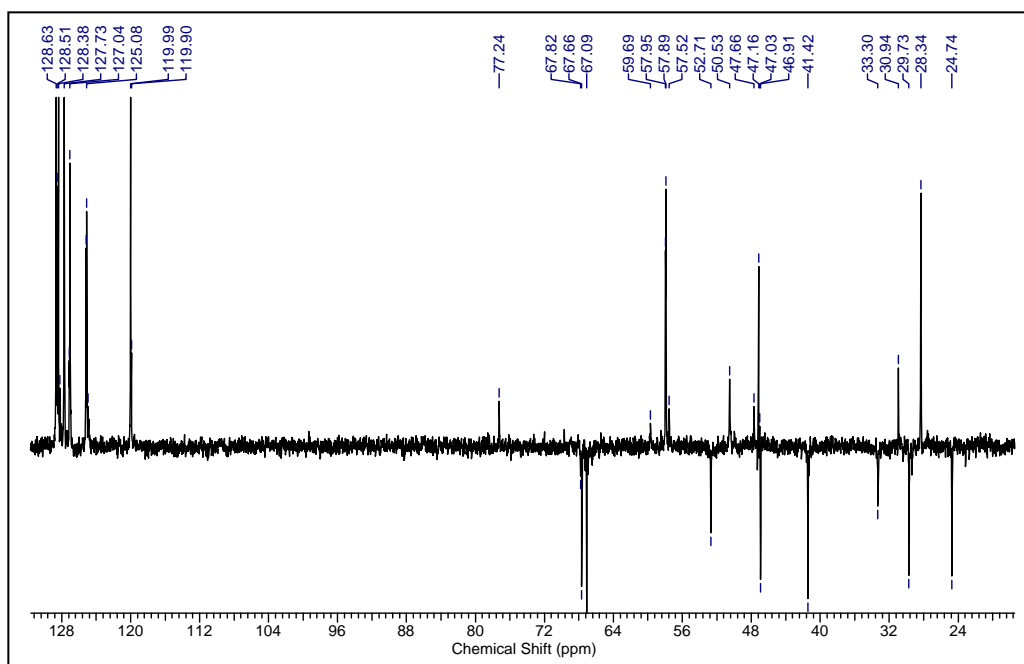
¹H NMR of compound 25**¹³C NMR of compound 25**

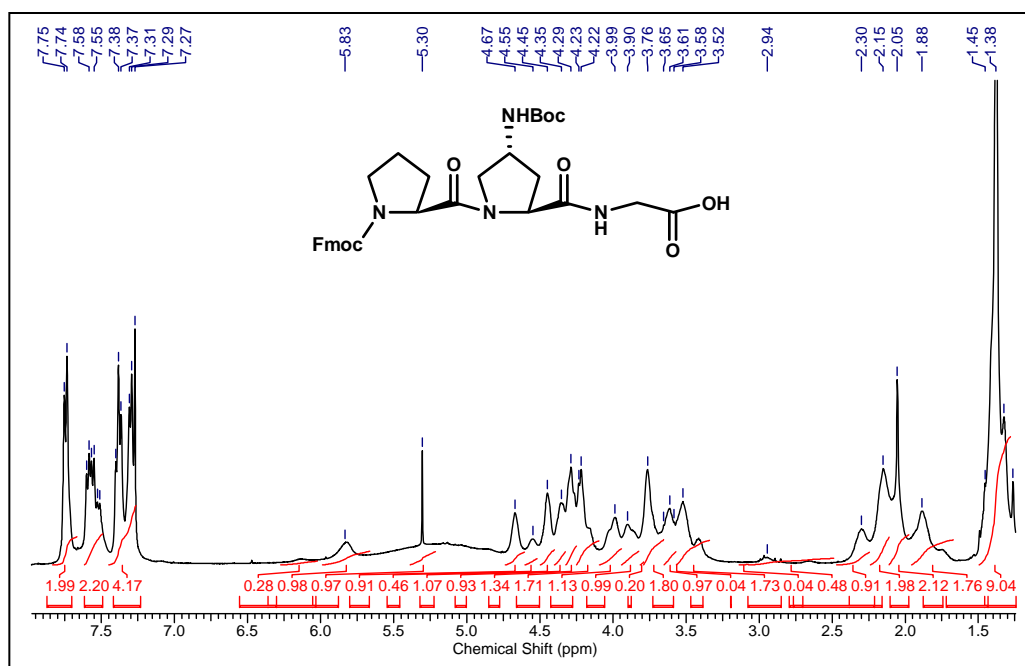
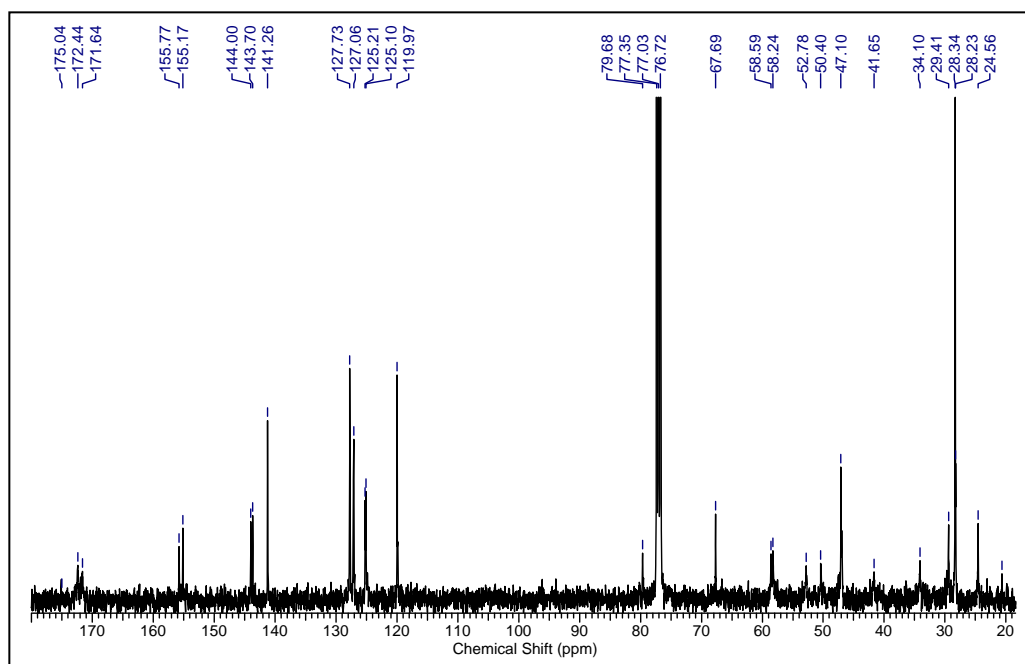
DEPT ^{13}C NMR of compound 25 ^1H NMR of compound 26

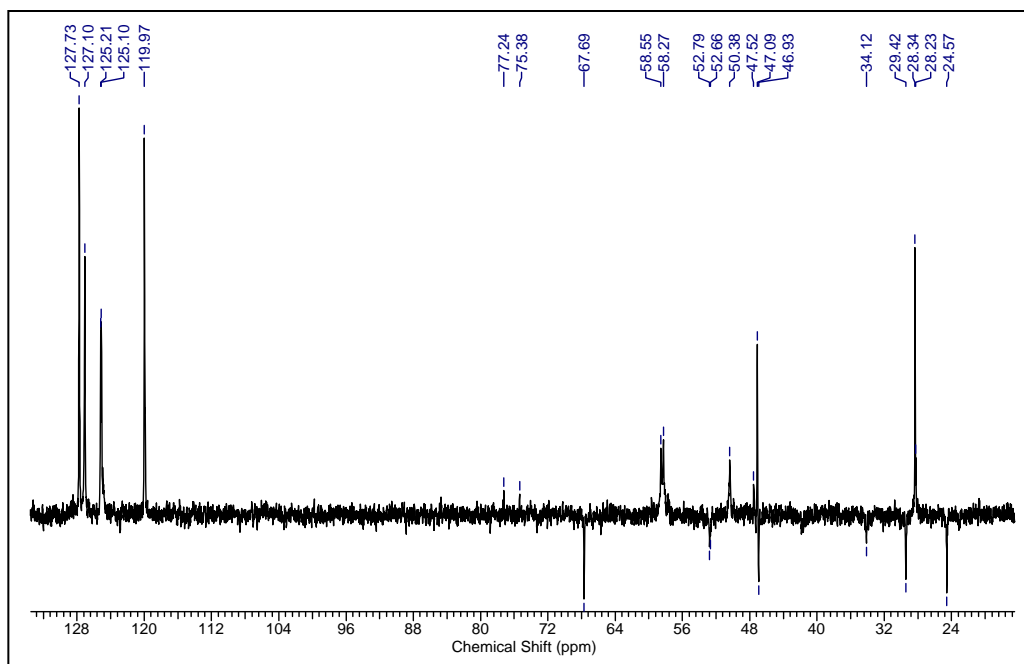
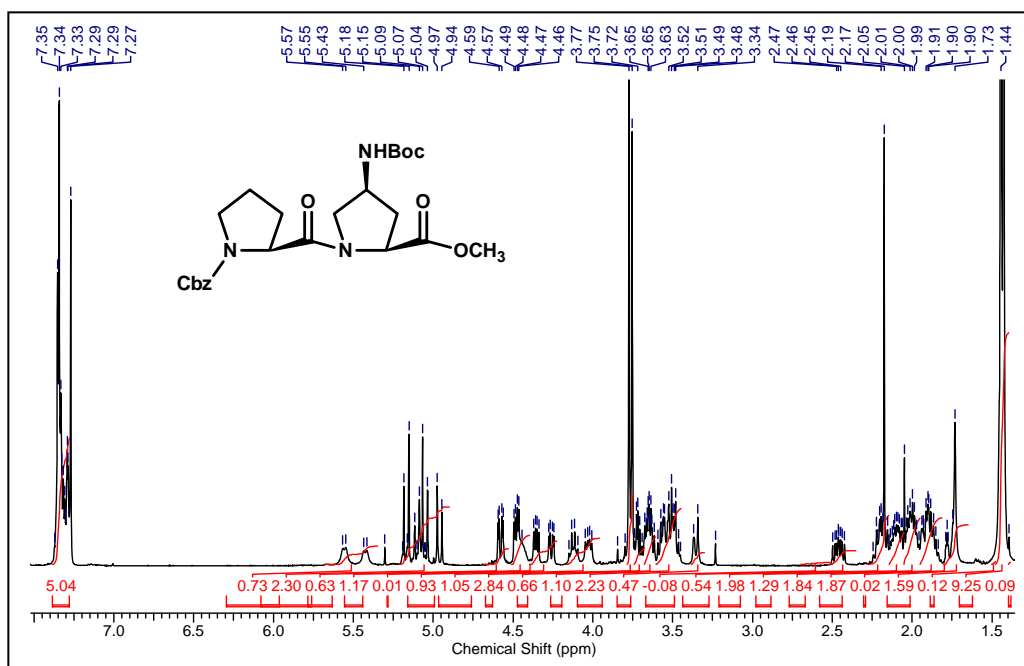
^{13}C NMR of compound 26**DEPT ^{13}C NMR of compound 26**

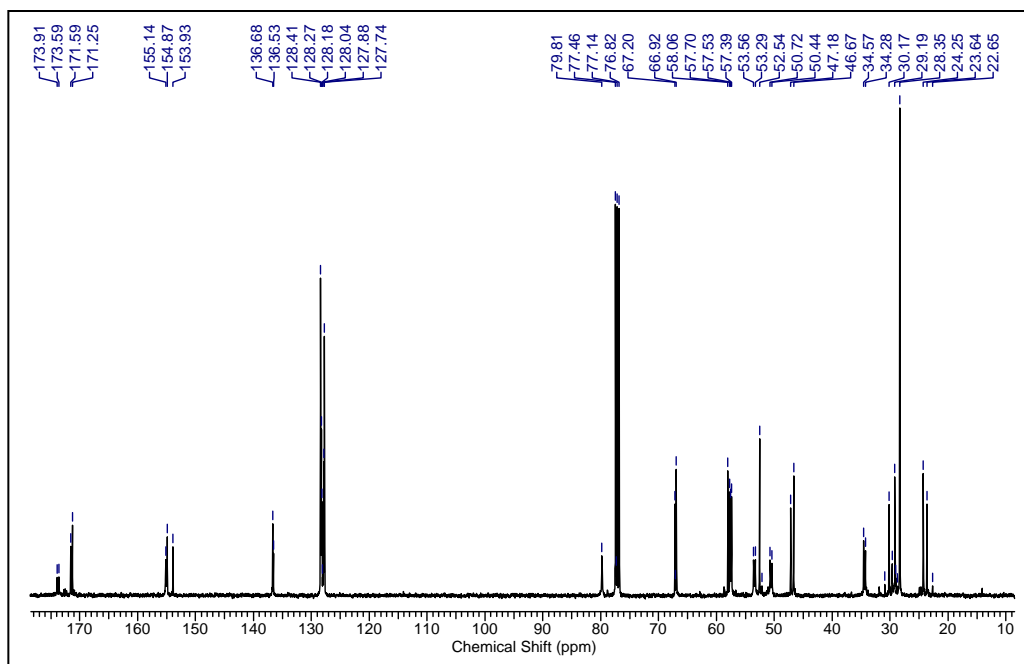
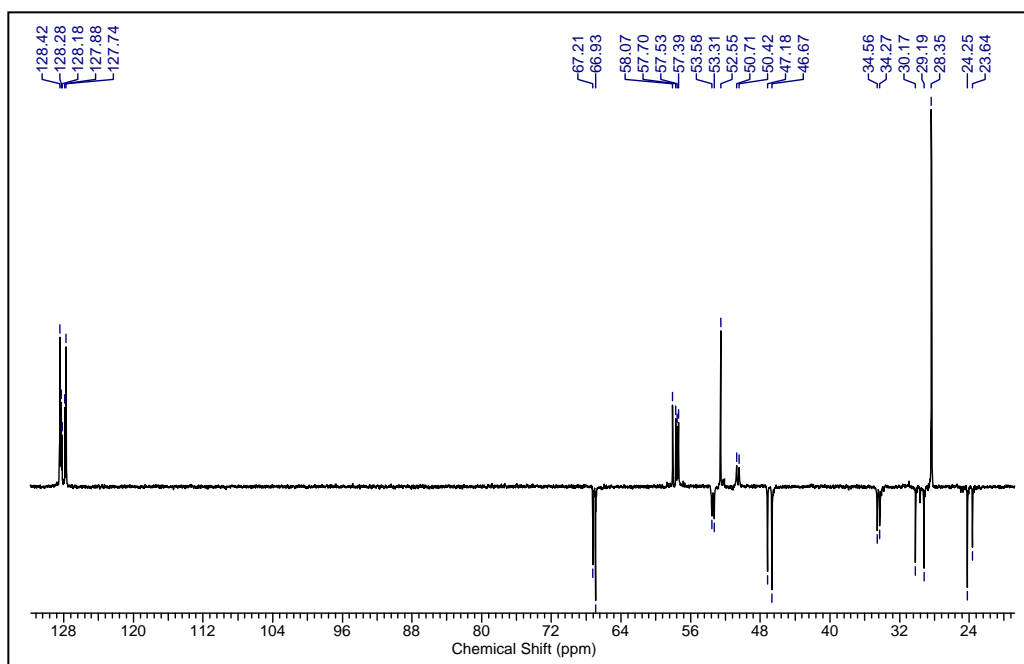
¹H NMR of compound 27**¹³C NMR of compound 27**

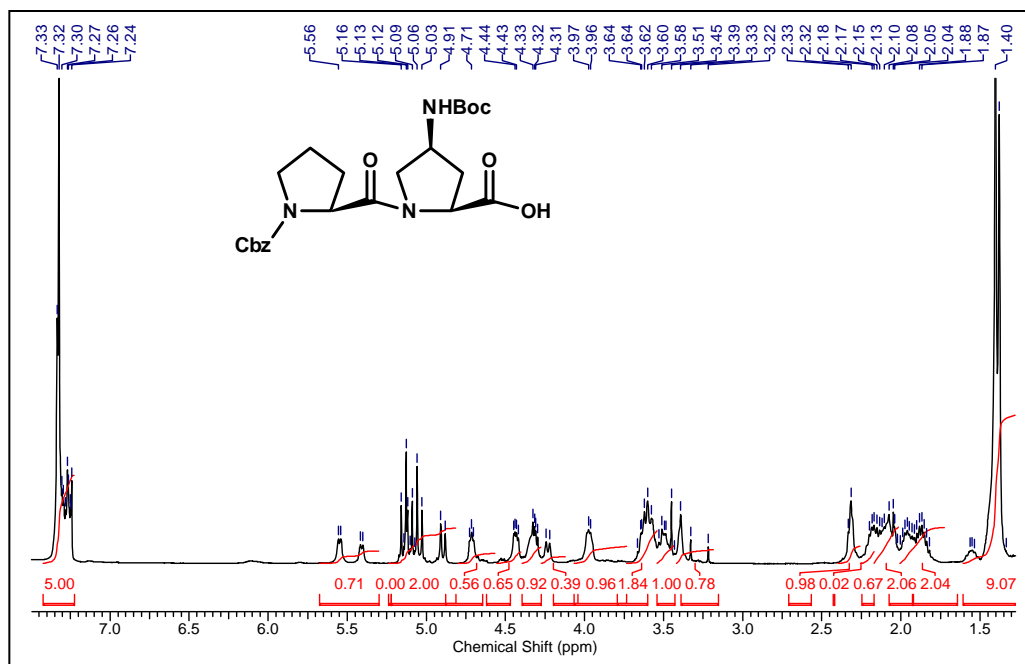
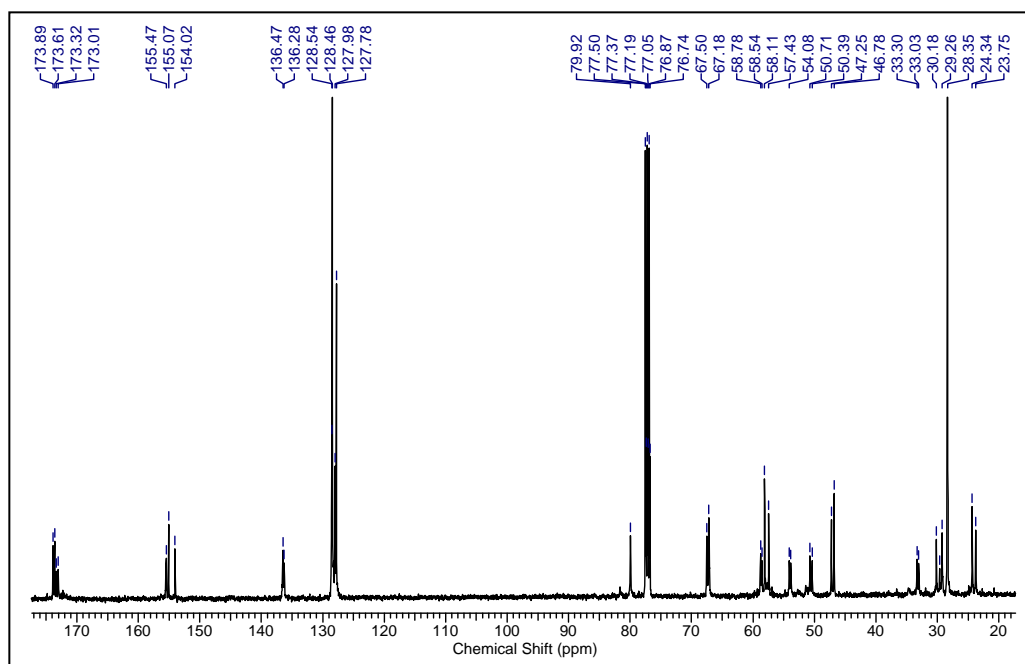
DEPT ^{13}C NMR of compound 27 ^1H NMR of compound 28

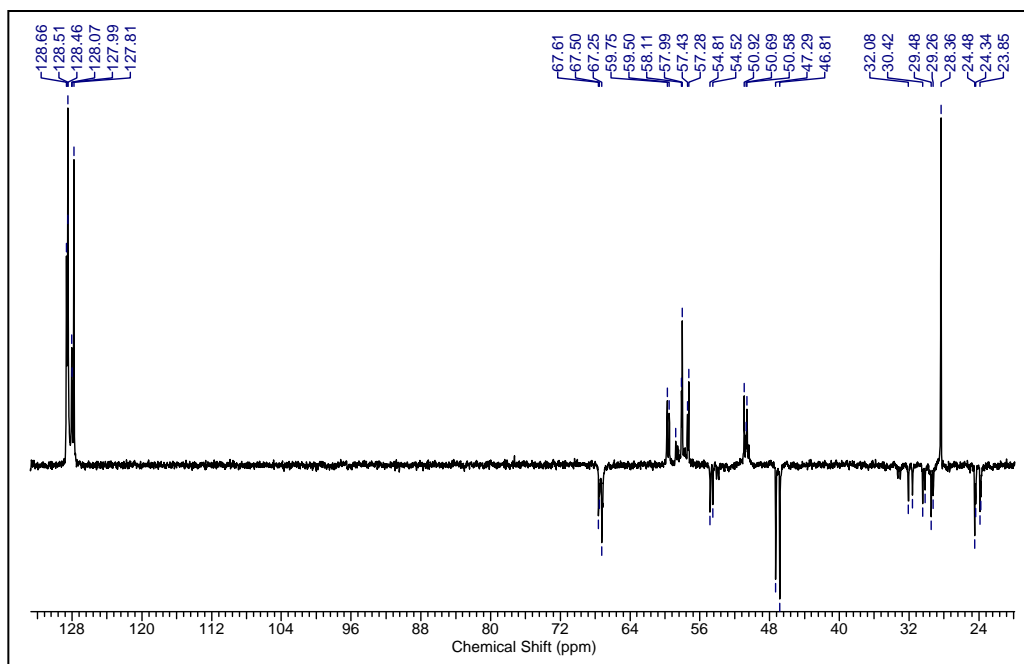
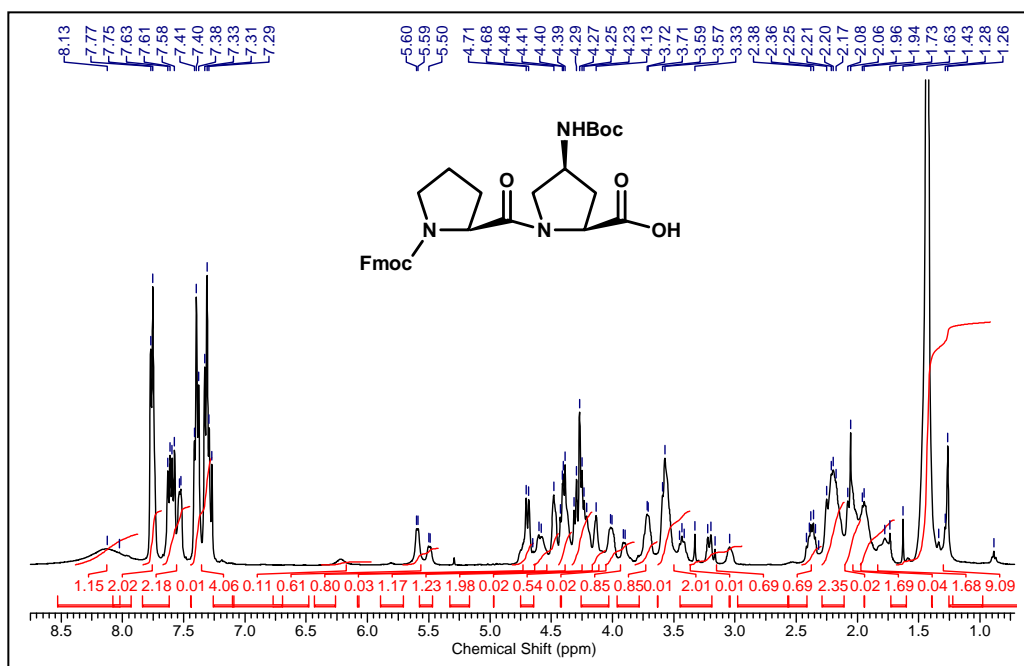
^{13}C NMR of compound 28**DEPT ^{13}C NMR of compound 28**

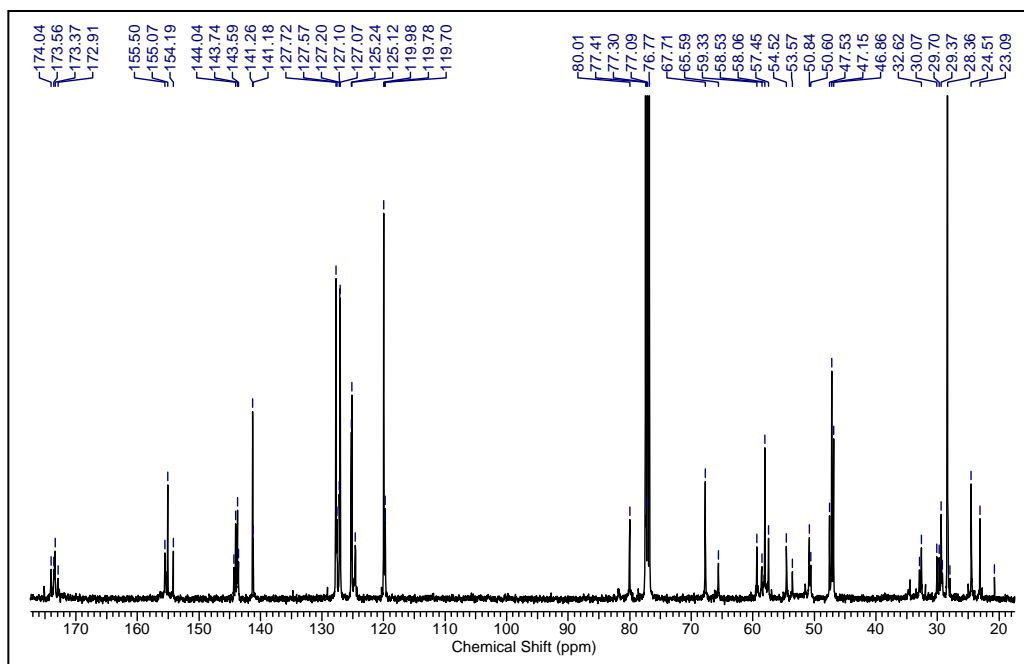
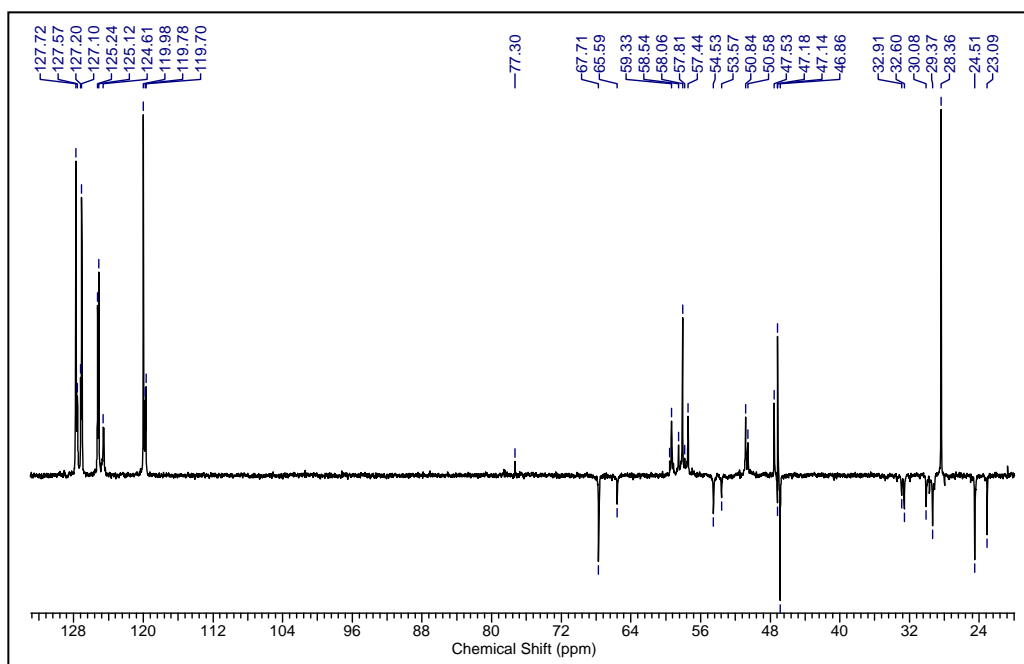
^1H NMR of compound 29 **^{13}C NMR of compound 29**

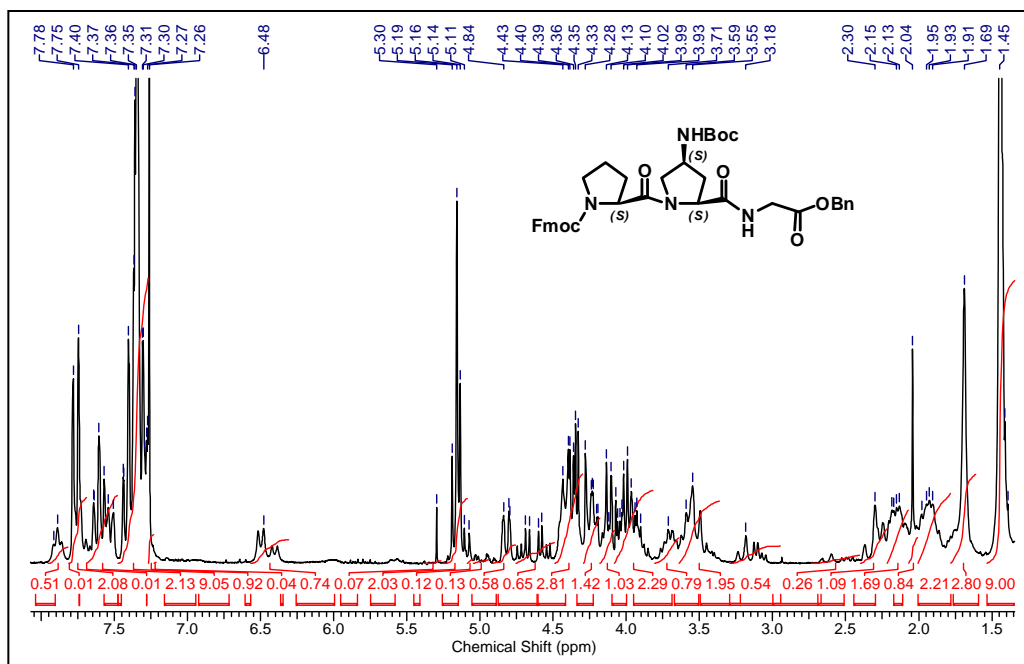
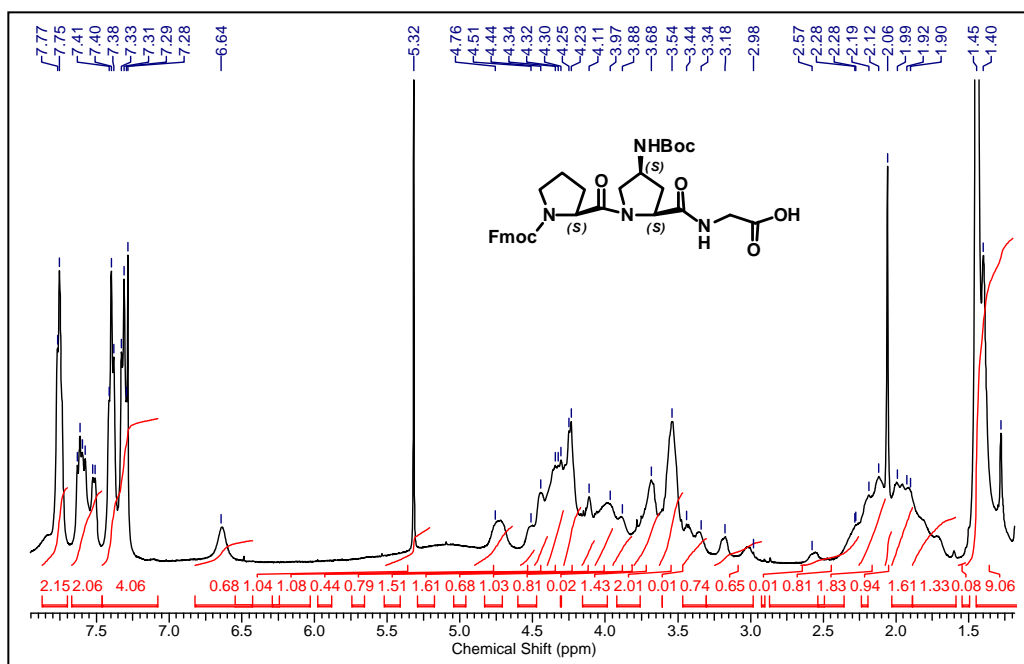
DEPT ^{13}C NMR of compound 29 ^1H NMR of compound 31

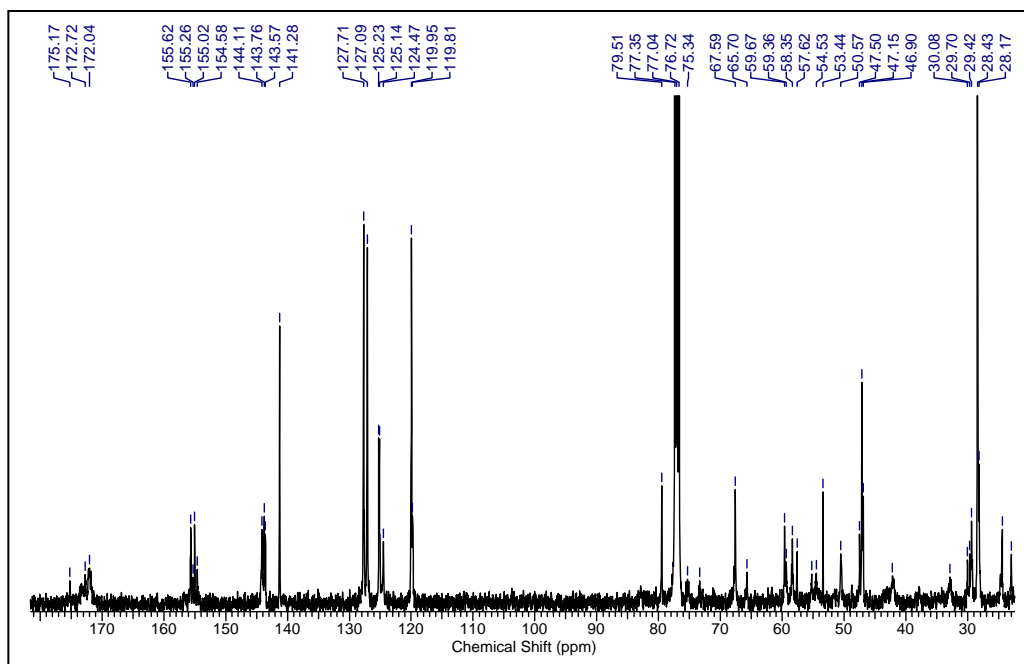
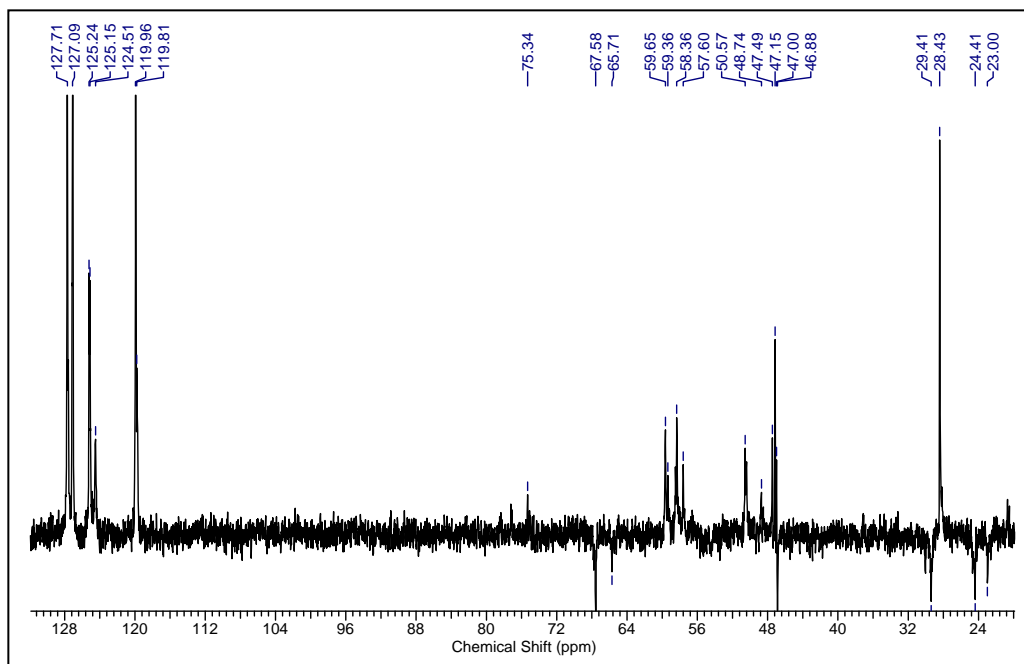
^{13}C NMR of compound 31**DEPT ^{13}C NMR of compound 31**

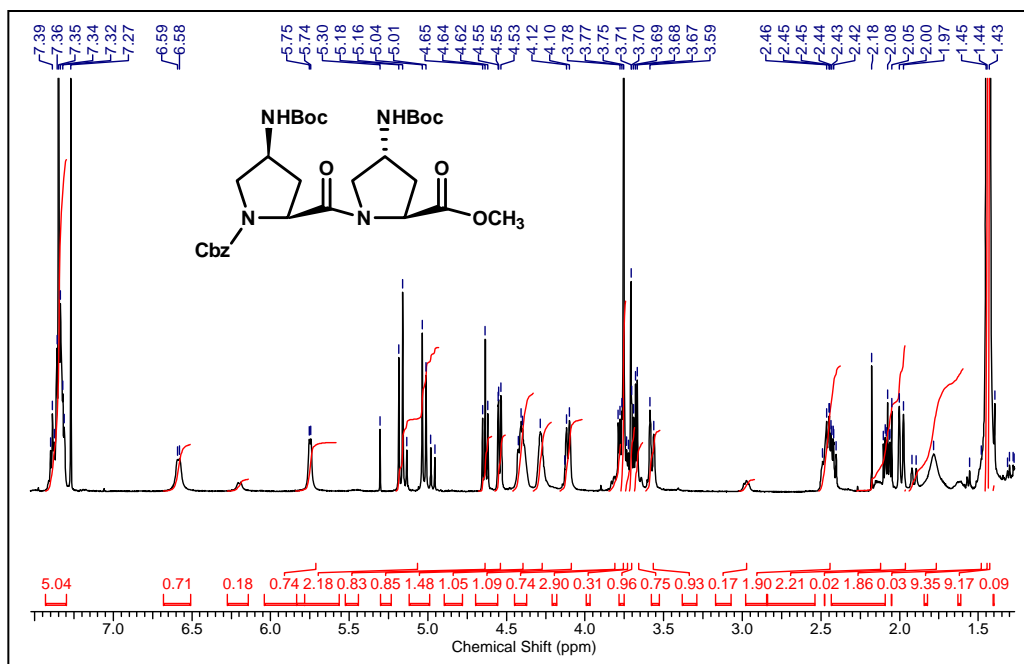
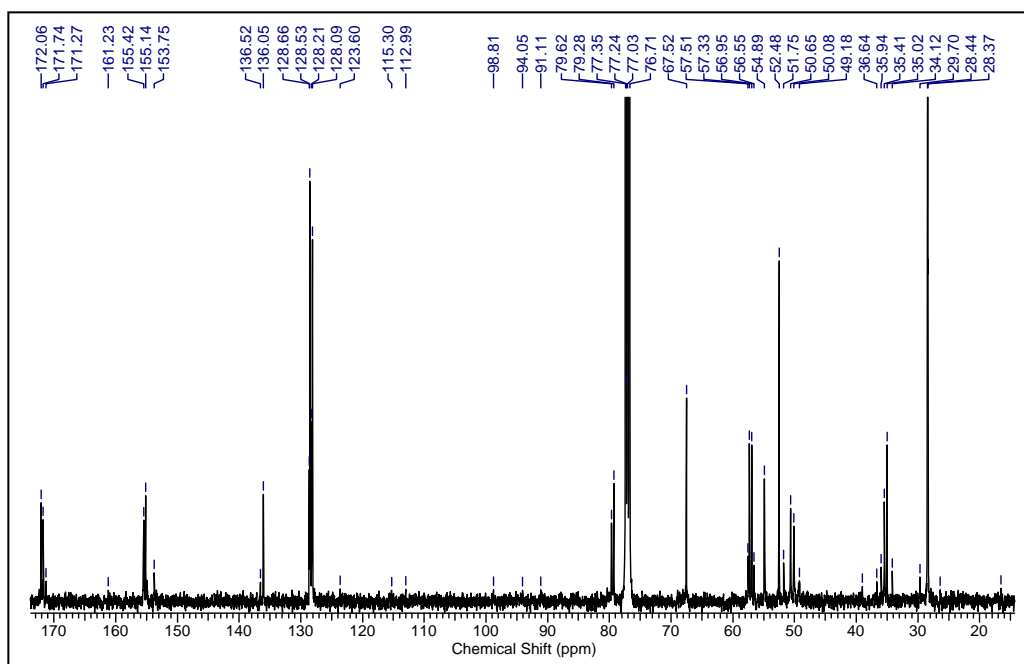
^1H NMR of compound 32 **^{13}C NMR of compound 32**

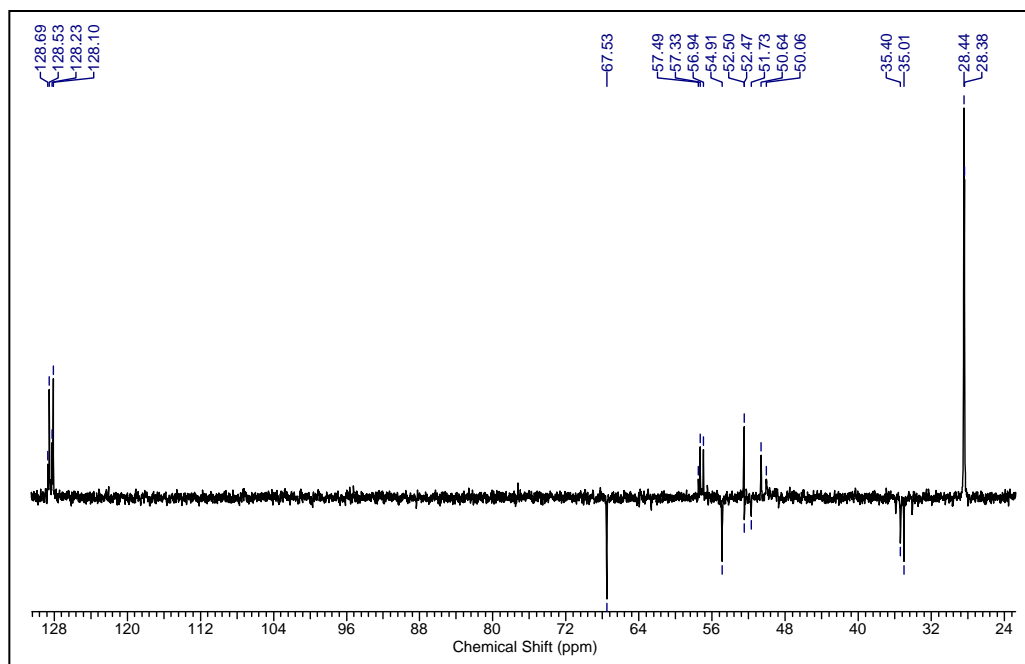
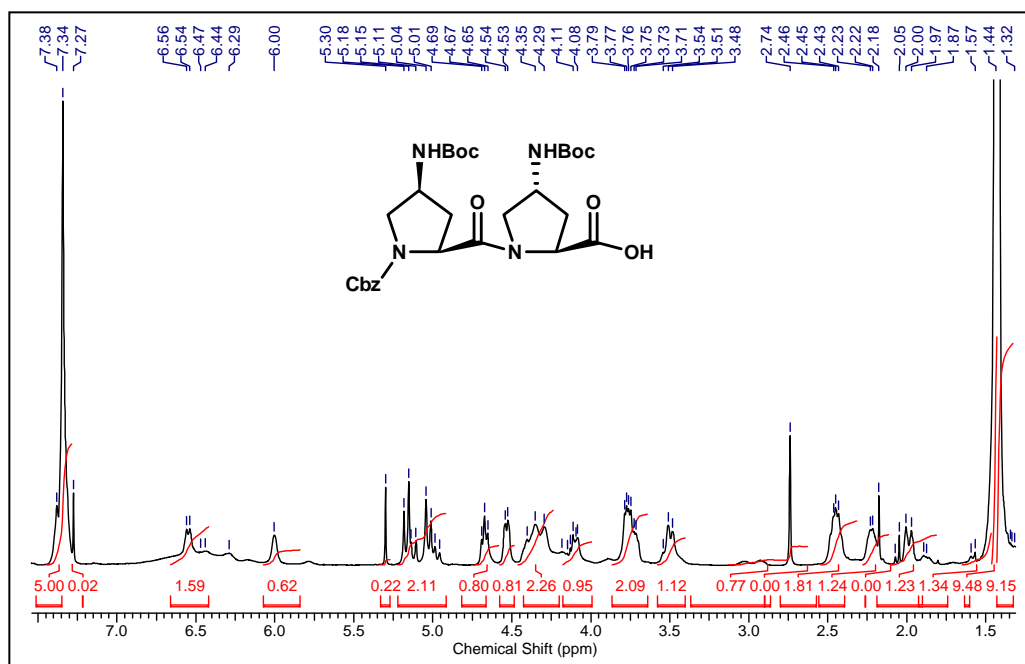
DEPT ^{13}C NMR of compound 32 ^1H NMR of compound 33

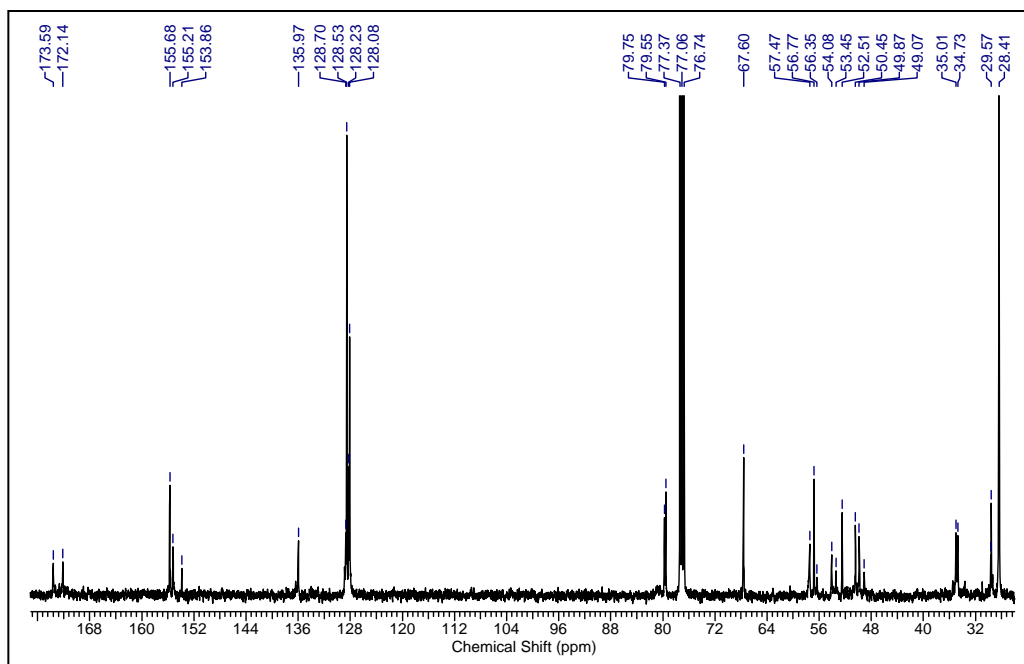
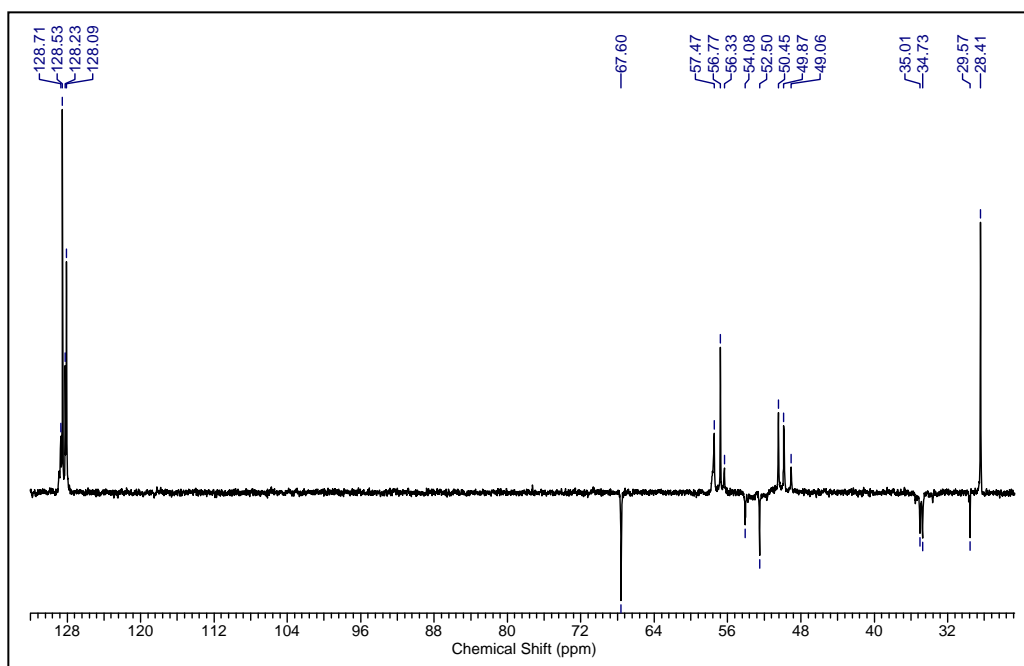
^{13}C NMR of compound 33**DEPT ^{13}C NMR of compound 33**

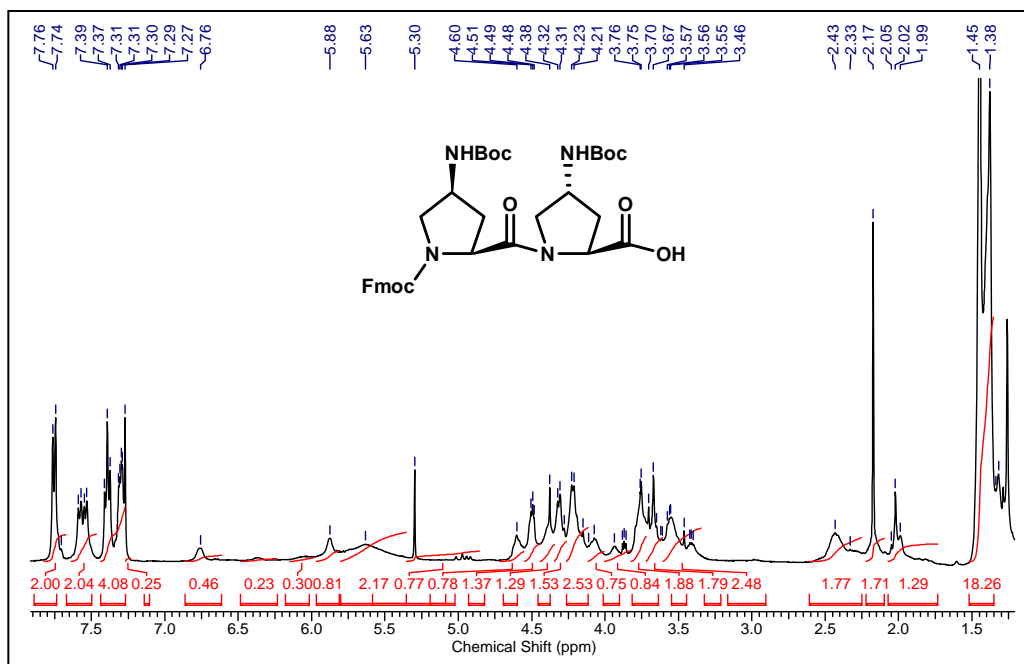
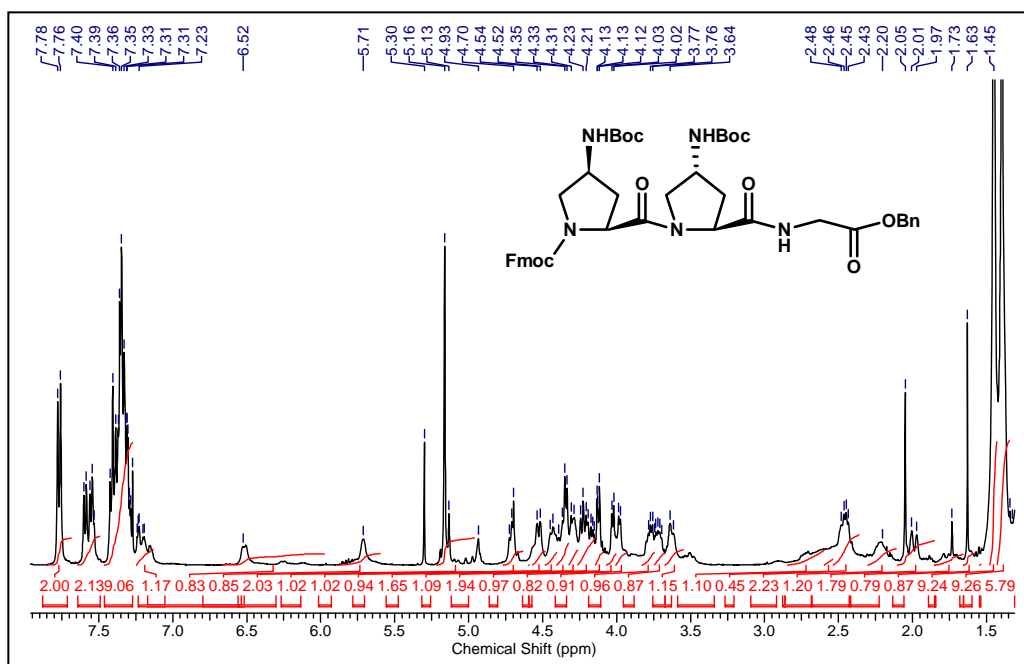
¹H NMR of compound 34**¹H NMR of compound 35**

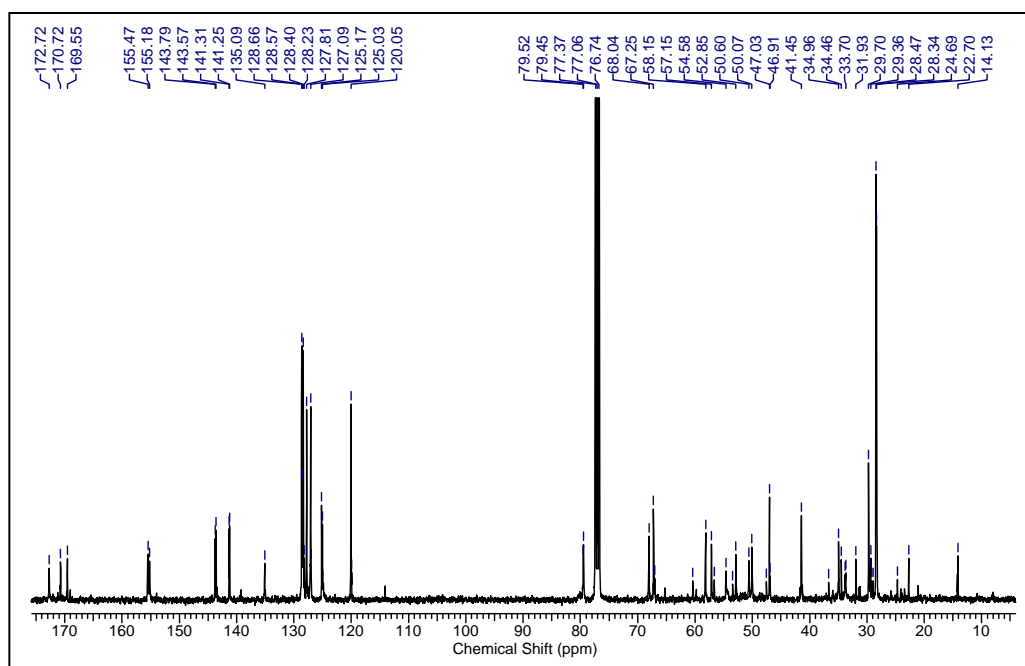
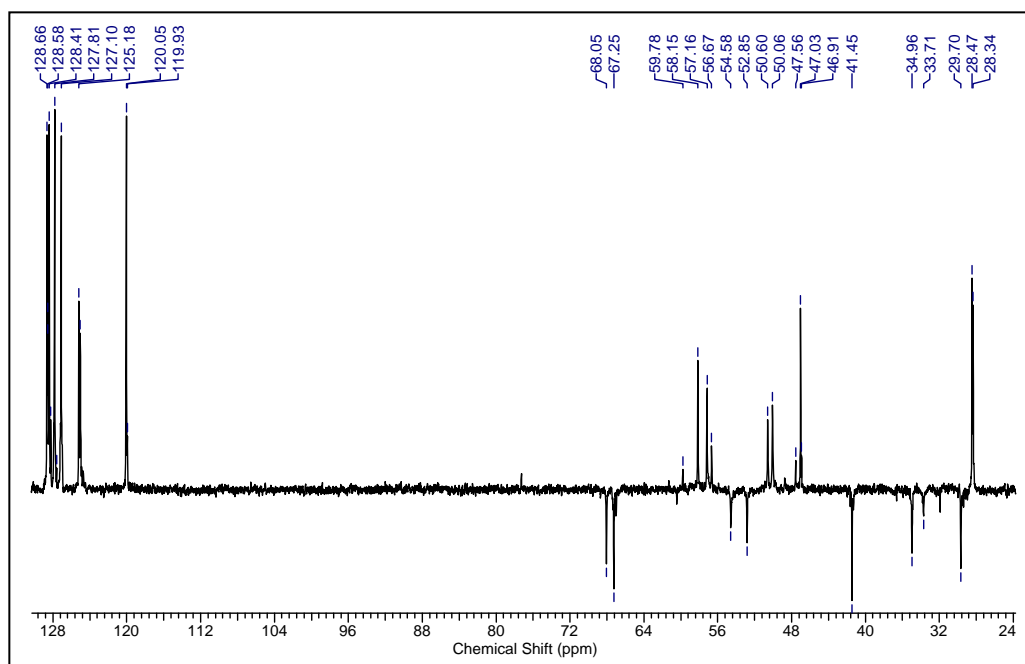
^{13}C NMR of compound 35**DEPT ^{13}C NMR of compound 35**

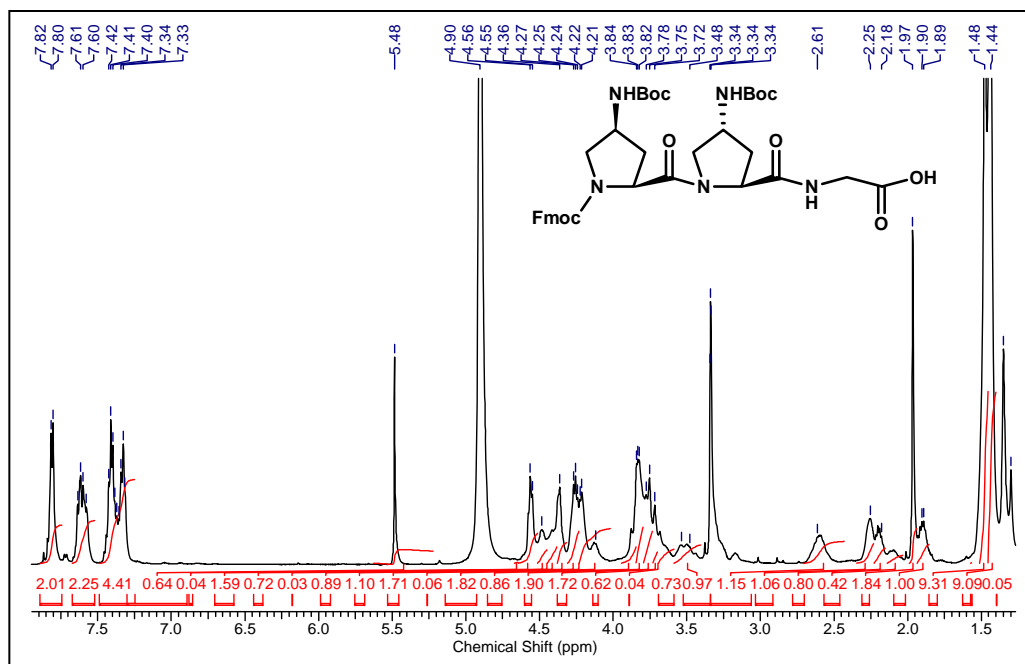
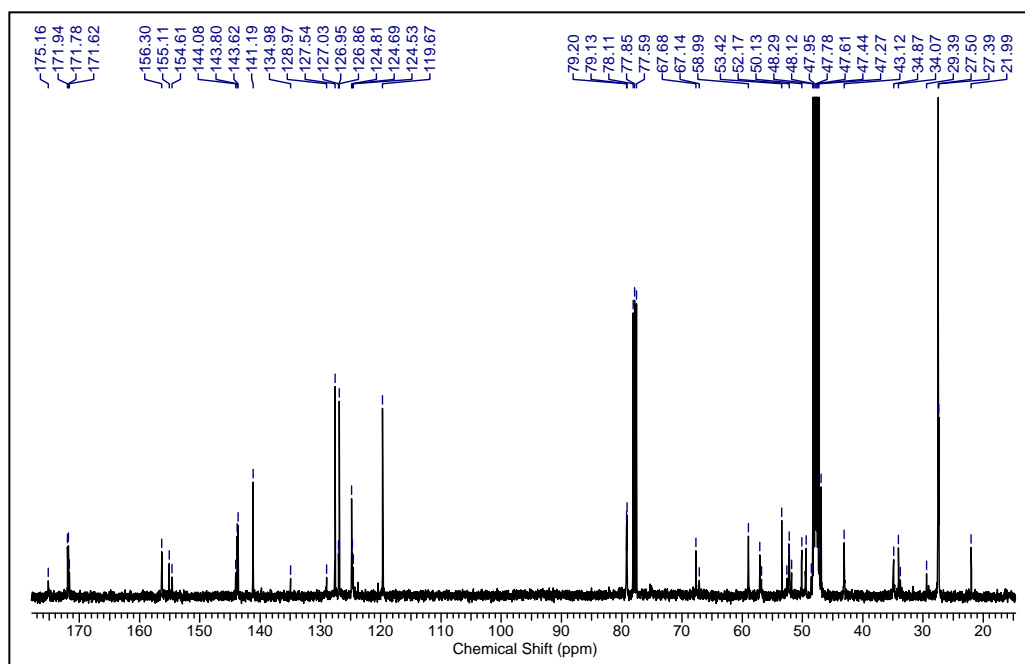
¹H NMR of compound 37**¹³C NMR of compound 37**

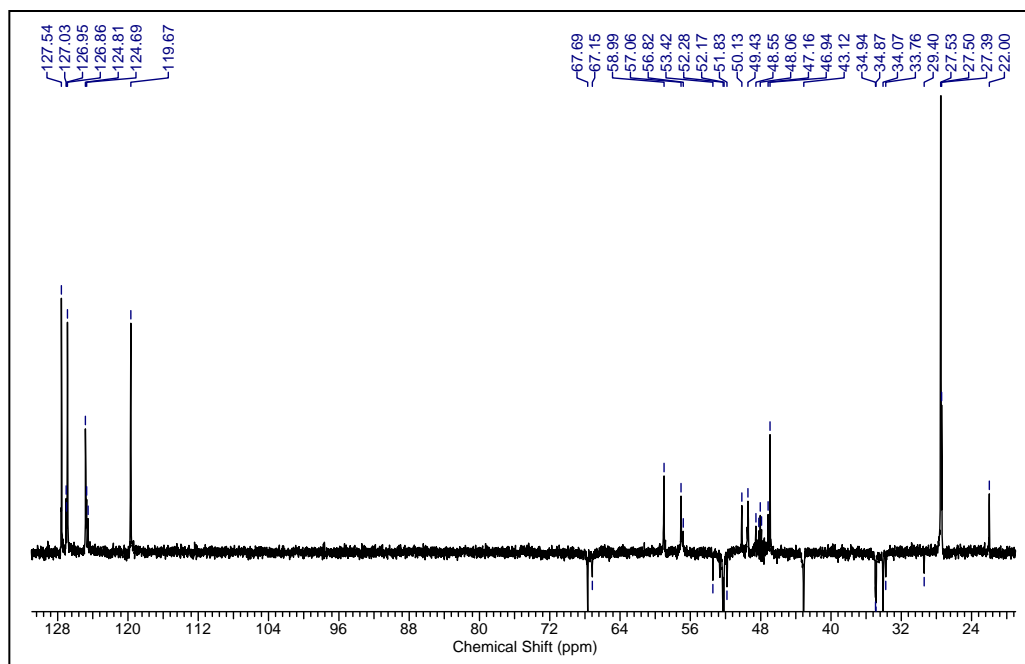
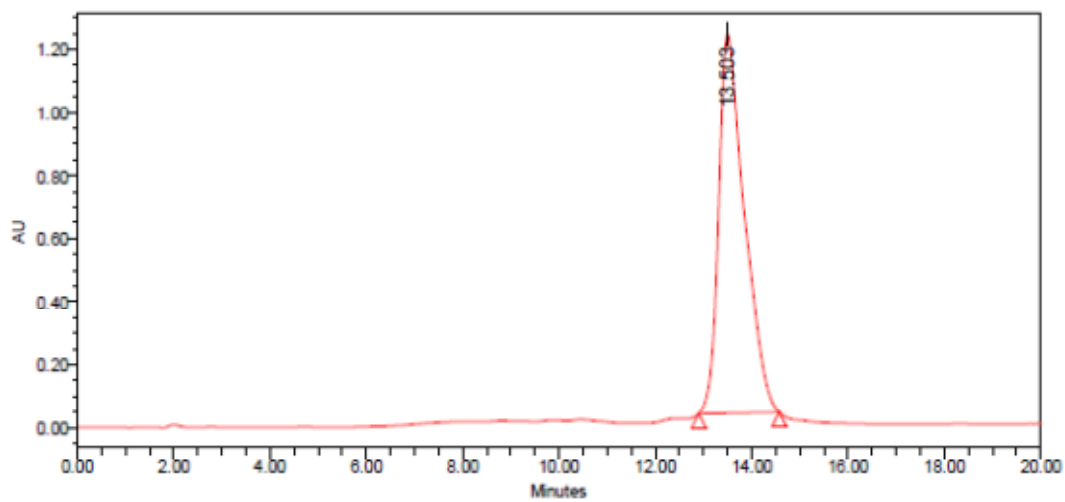
DEPT ^{13}C NMR of compound 37 ^1H NMR of compound 38

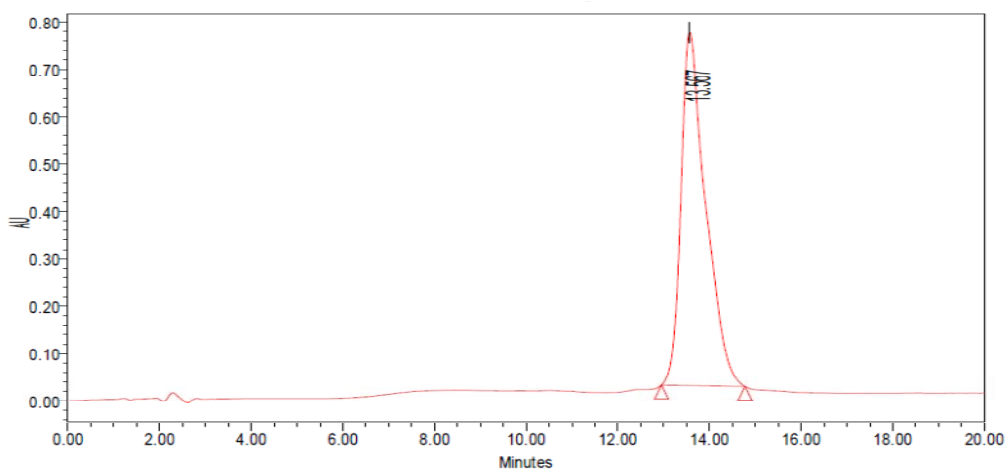
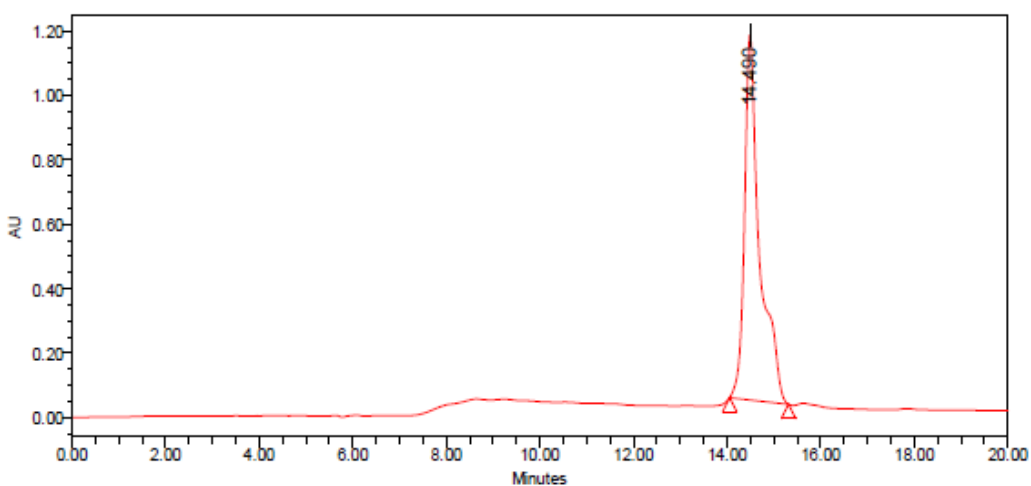
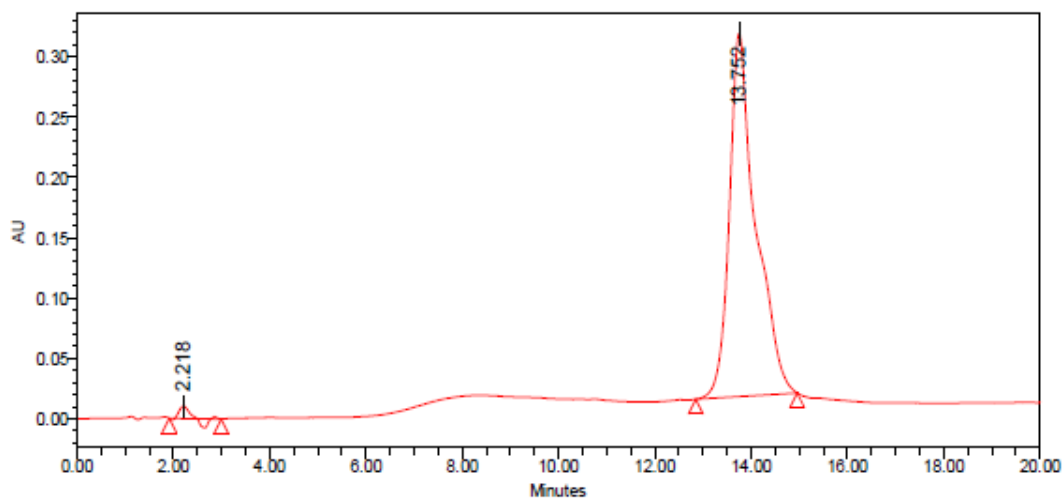
^{13}C NMR of compound 38**DEPT ^{13}C NMR of compound 38**

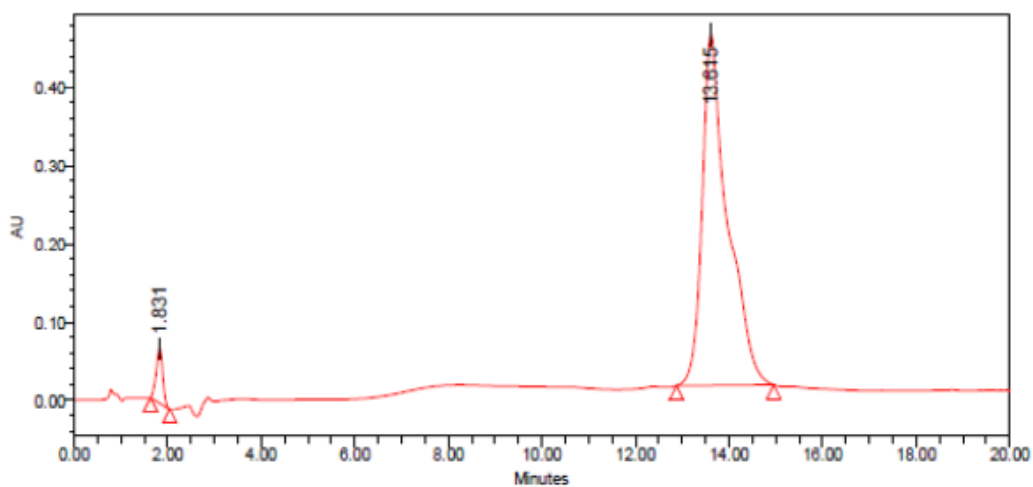
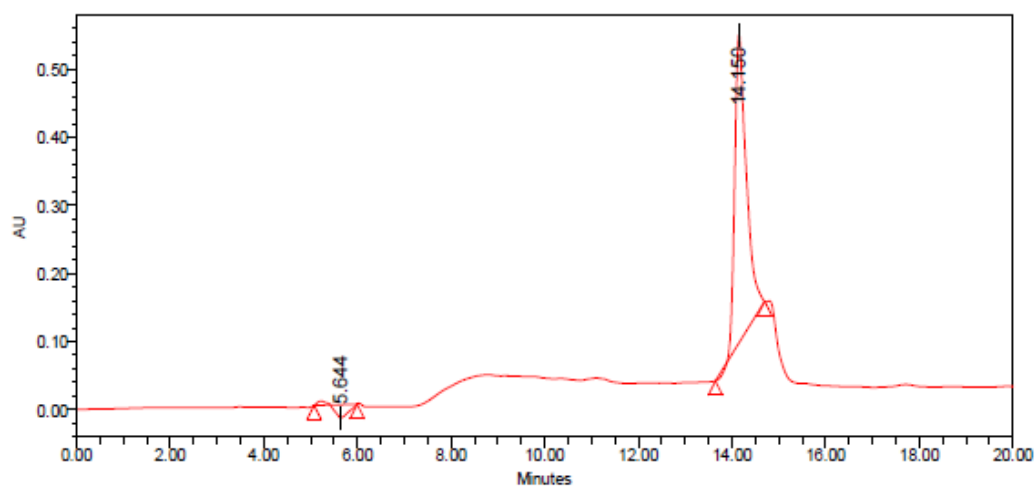
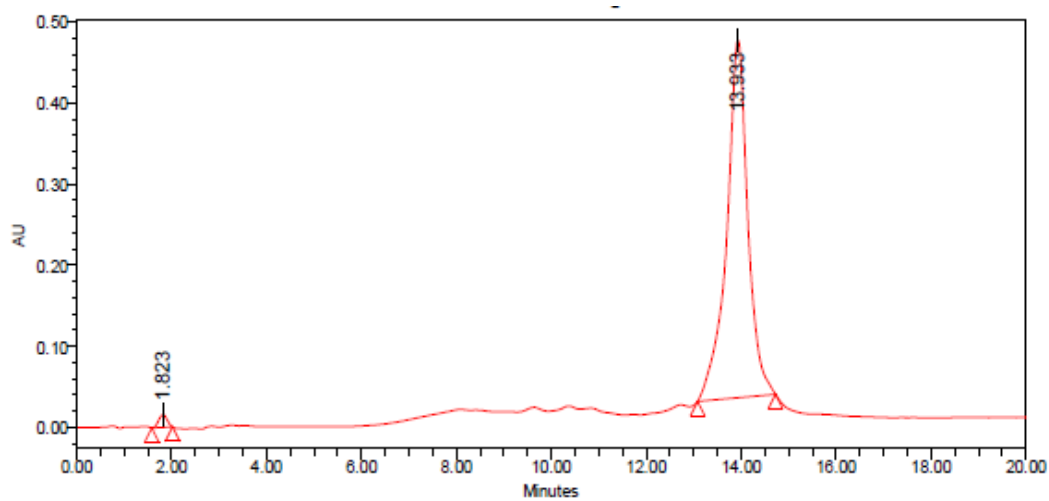
¹H NMR of compound 39**¹H NMR of compound 40**

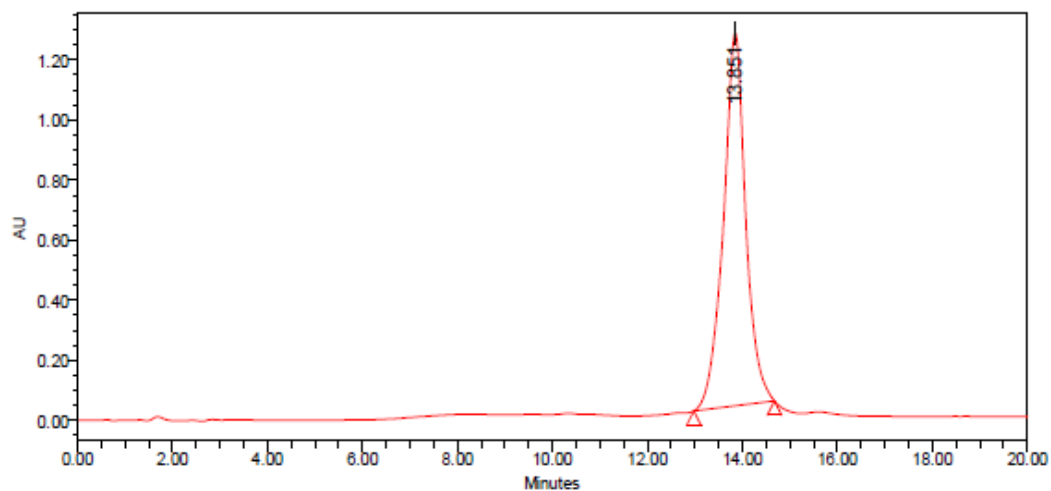
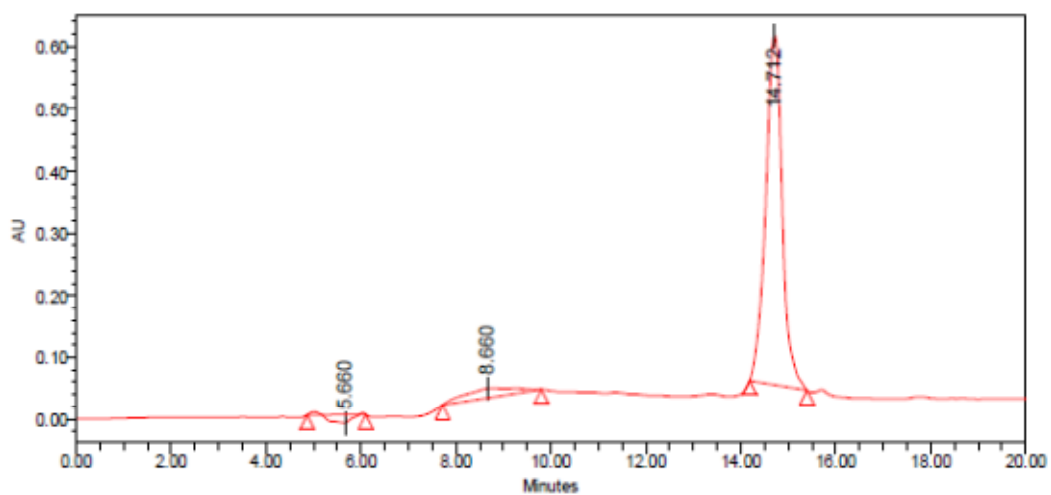
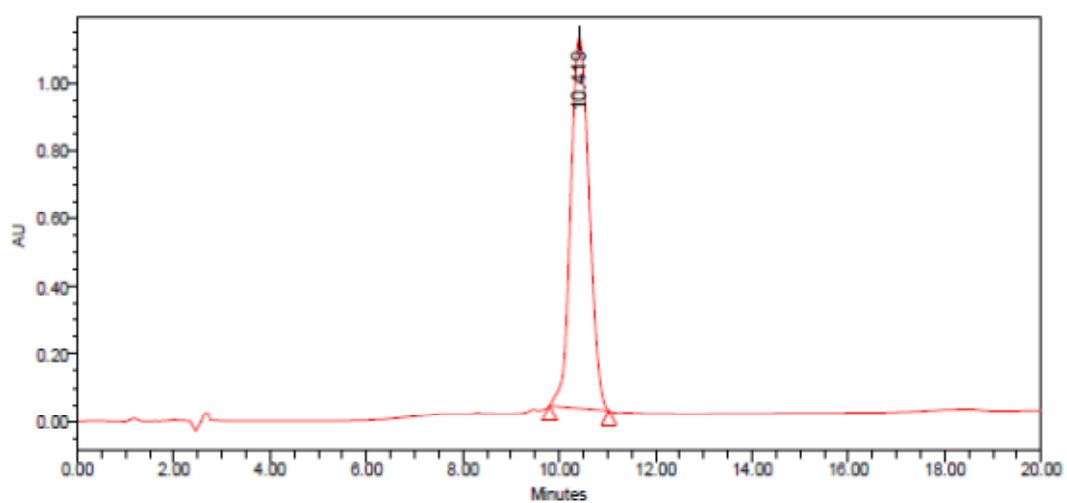
^{13}C NMR of compound 40**DEPT ^{13}C NMR of compound 40**

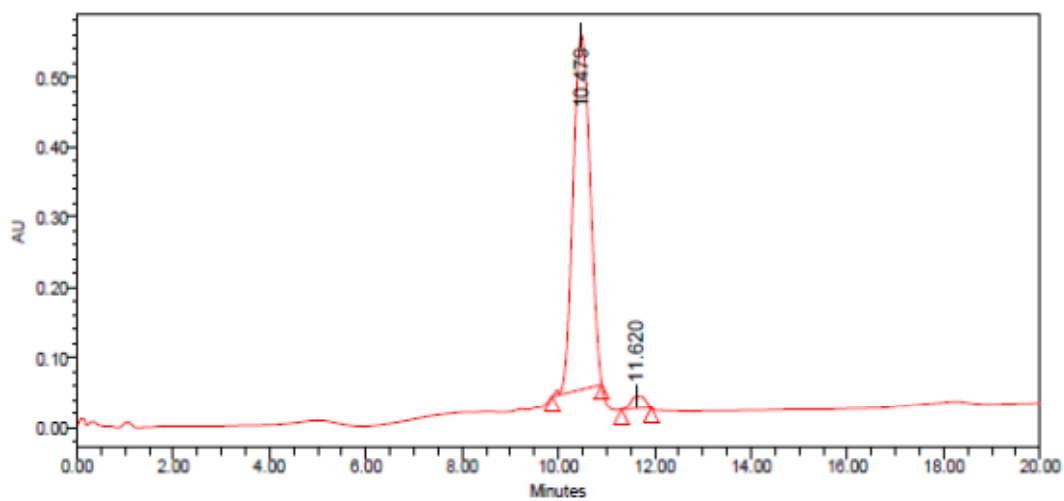
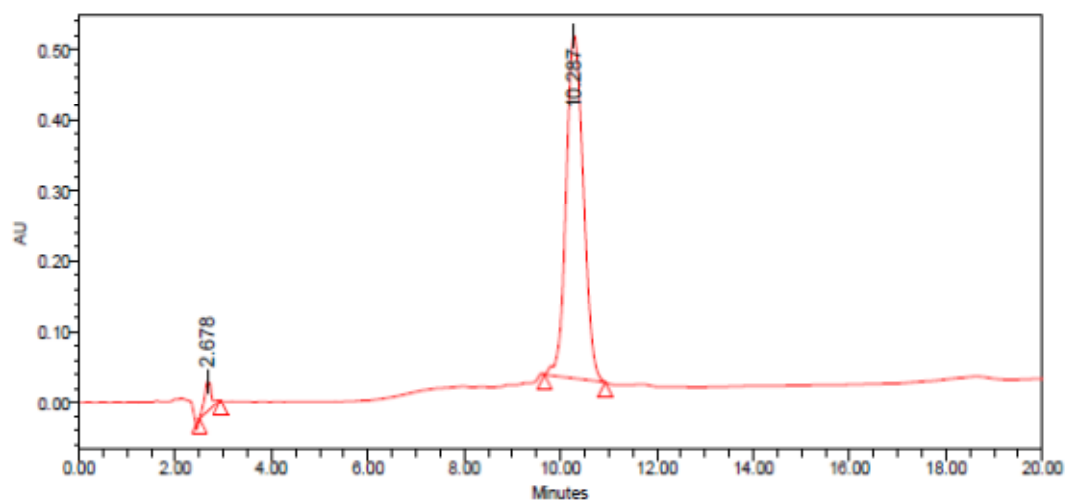
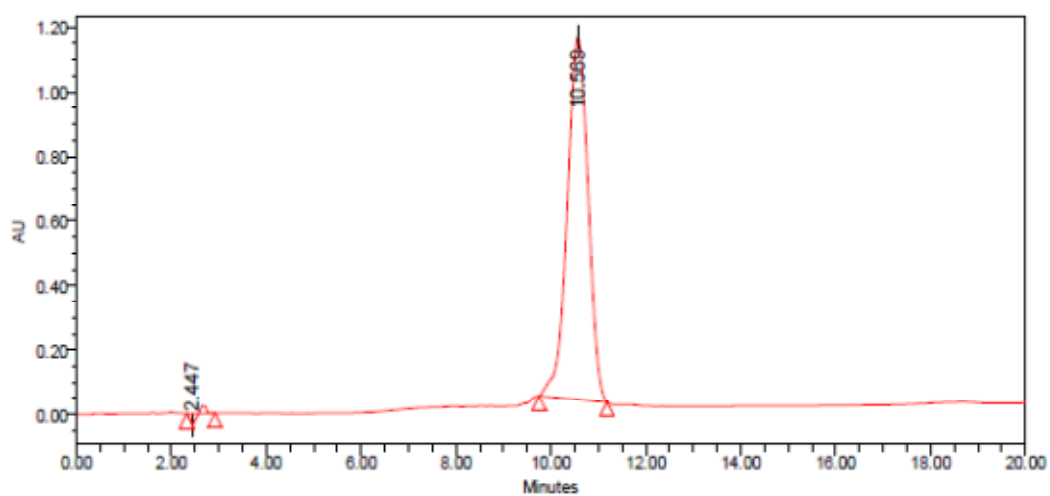
^1H NMR of compound 41 **^{13}C NMR of compound 41**

DEPT ^{13}C NMR of compound 41**(B) HPLC of peptides 4-21****HPLC of peptide 4 (Ac-Phe-[Pro-Hyp-Gly]₅-Pro-4R-Amp-Gly-NH₂)**

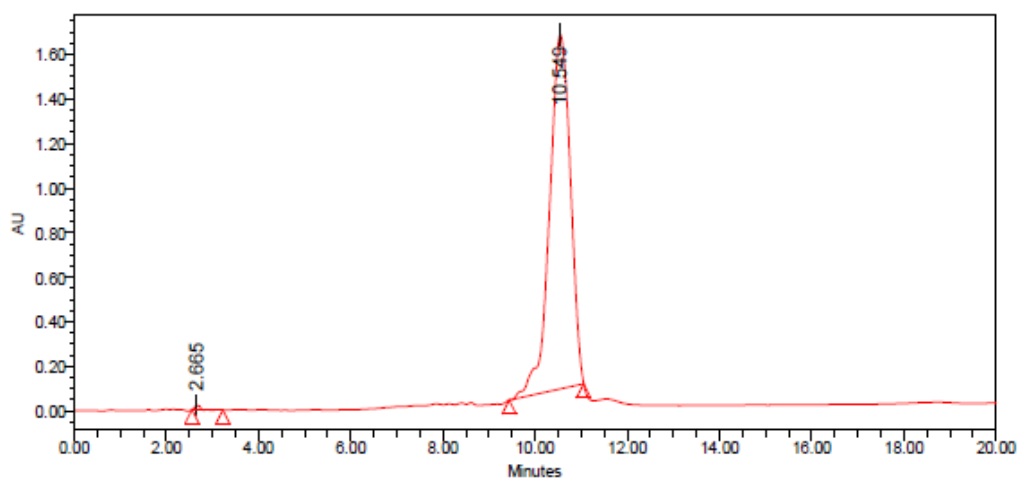
HPLC of peptide 5 (Ac-Phe-[Pro-Hyp-Gly]₅-Pro-4S-amp-Gly-NH₂)**HPLC of peptide 6 (Ac-Phe-[Pro-Hyp-Gly]₅-4S-amp-4R-Amp-Gly-NH₂)****HPLC of peptide 7 (Ac-Phe-[Pro-Hyp-Gly]₂-Pro-4R-Amp-Gly-[Pro-Hyp-Gly]₃-NH₂)**

HPLC of peptide 8 (Ac-Phe-[Pro-Hyp-Gly]₂-Pro-4S-amp-Gly-[Pro-Hyp-Gly]₃-NH₂)**HPLC of peptide 9 (Ac-Phe-[Pro-Hyp-Gly]₂-4S-amp-4R-Amp-Gly-[Pro-Hyp-Gly]₃-NH₂)****HPLC of peptide 10 (Ac-Phe-Pro-4R-Amp-Gly-[Pro-Hyp-Gly]₅-NH₂)**

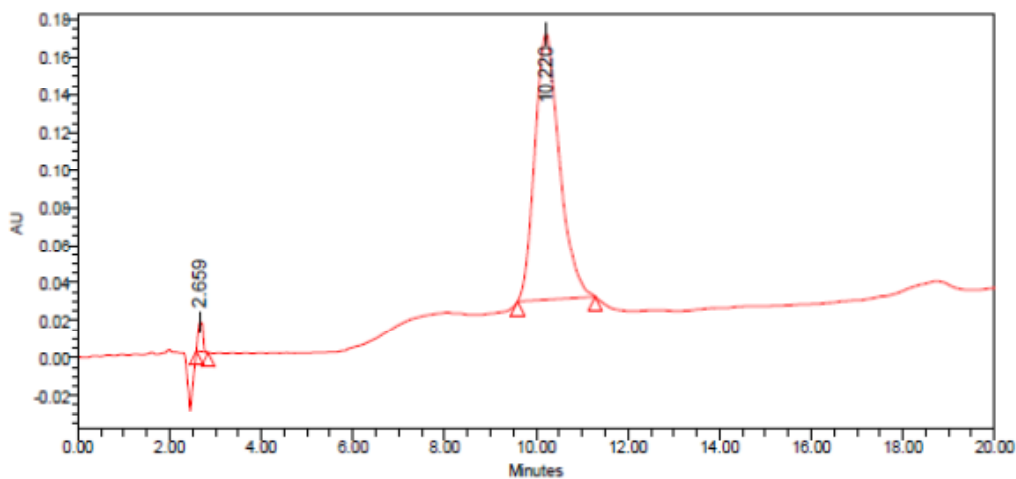
HPLC of peptide 11 (Ac-Phe-Pro-4S-amp-Gly-[Pro-Hyp-Gly]₅-NH₂)**HPLC of peptide 12 (Ac-Phe-4S-amp-4R-Amp-Gly-[Pro-Hyp-Gly]₅-NH₂)****HPLC of peptide 13 (H₂N-Phe-[Pro-Hyp-Gly]₅-Pro-4R-Amp-Gly-NH₂)**

HPLC of peptide 14 (H₂N-Phe-[Pro-Hyp-Gly]₅-Pro-4S-amp-Gly-NH₂)**HPLC of peptide 15 (H₂N-Phe-[Pro-Hyp-Gly]₅-4S-amp-4R-Amp-Gly-NH₂)****HPLC of peptide 16 (H₂N-Phe-[Pro-Hyp-Gly]₂-Pro-4R-Amp-Gly-[Pro-Hyp-Gly]₃-NH₂)**

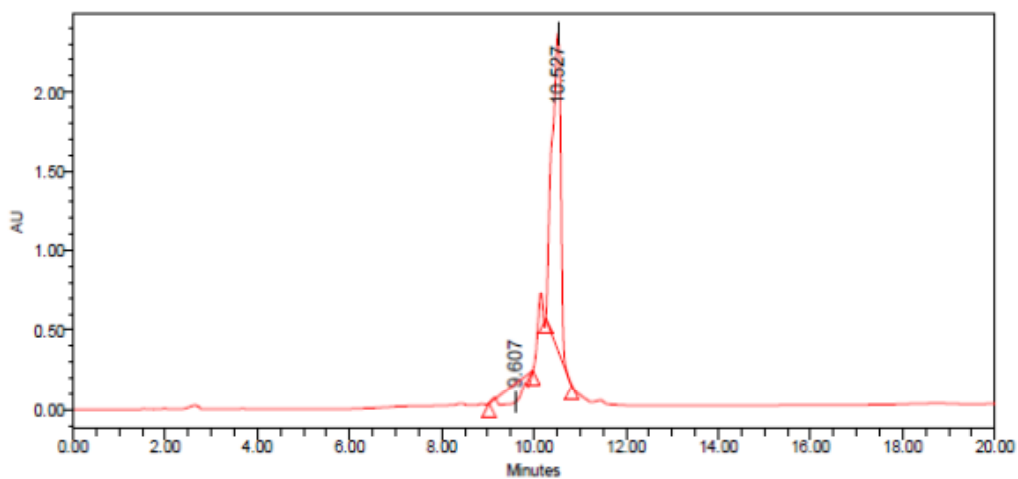
HPLC of peptide 17 (H₂N-Phe-[Pro-Hyp-Gly]₂-Pro-4S-amp-Gly-[Pro-Hyp-Gly]₃-NH₂)

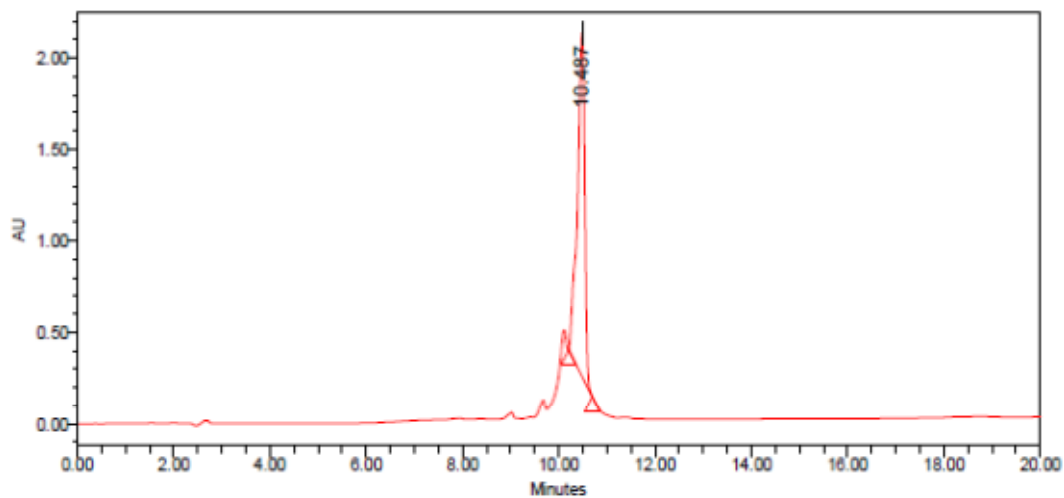
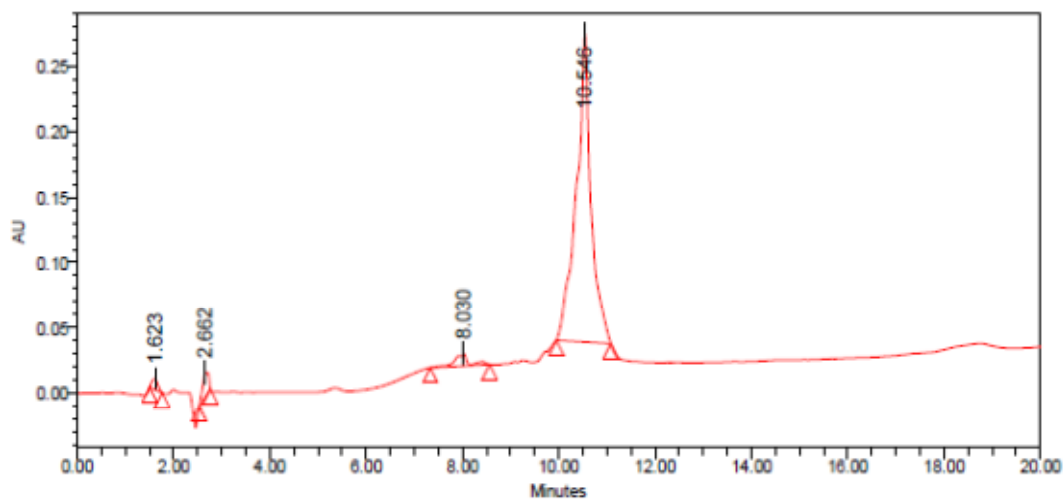


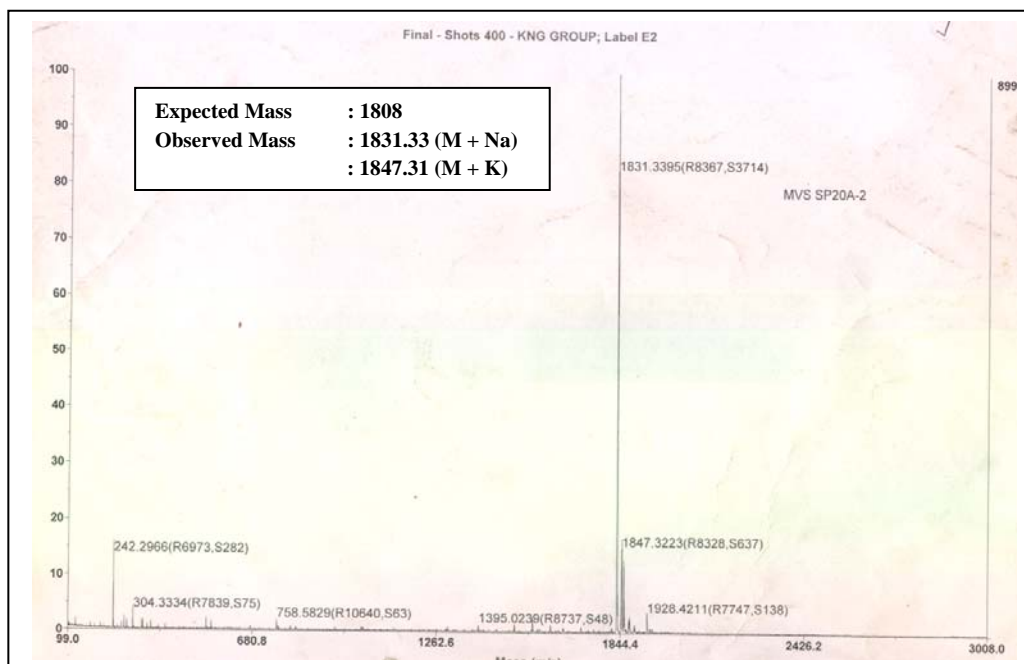
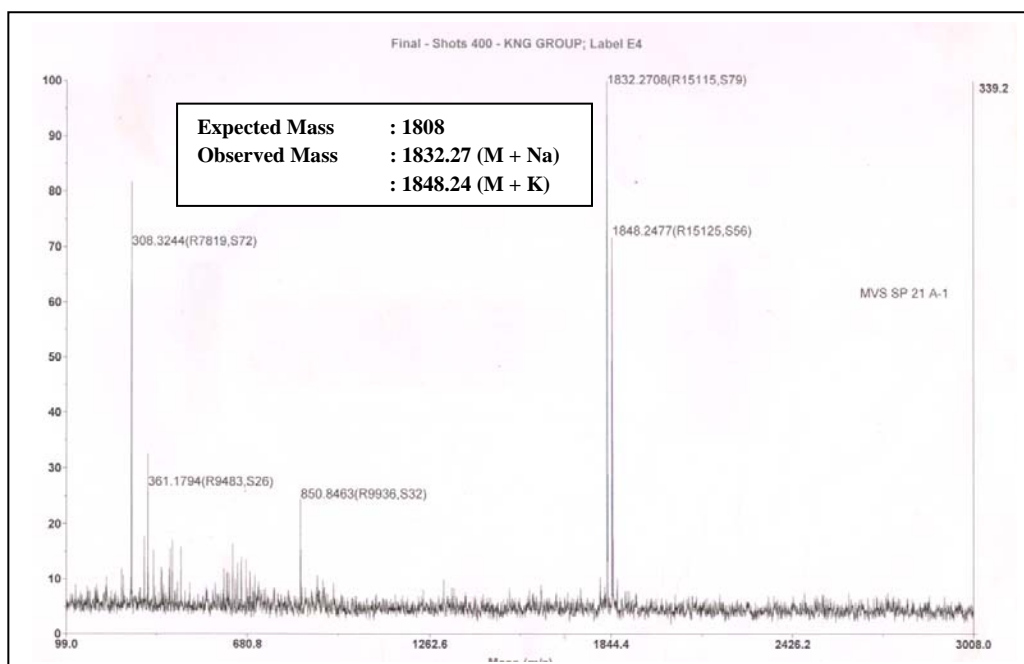
HPLC of peptide 18 (H₂N-Phe-[Pro-Hyp-Gly]₂-4S-amp-4R-Amp-Gly-[Pro-Hyp-Gly]₃-NH₂)



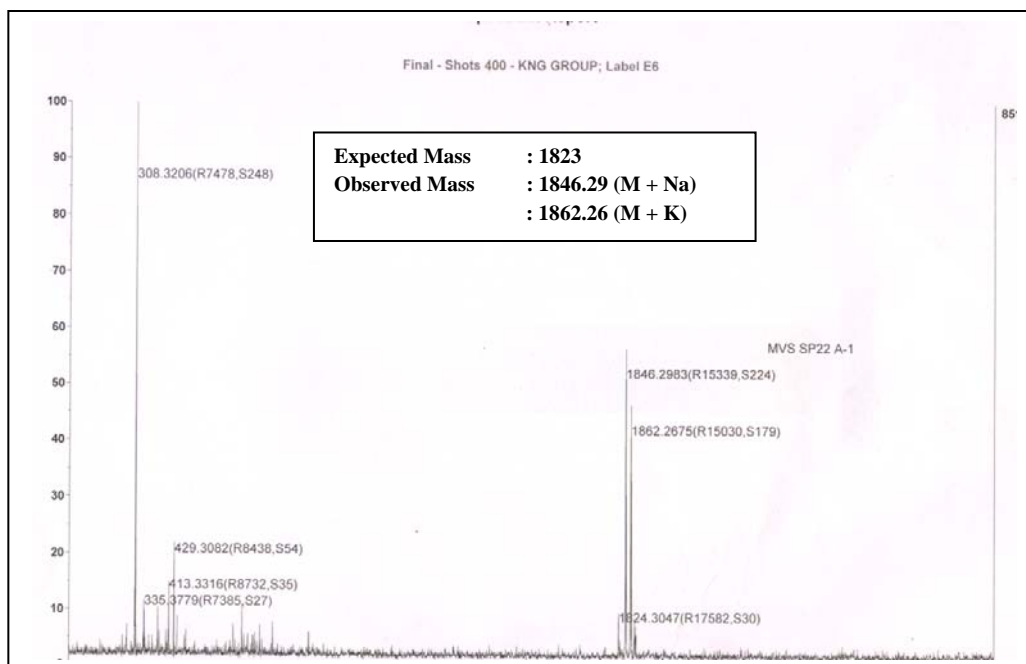
HPLC of peptide 19 (H₂N-Phe-Pro-4R-Amp-Gly-[Pro-Hyp-Gly]₅-NH₂)



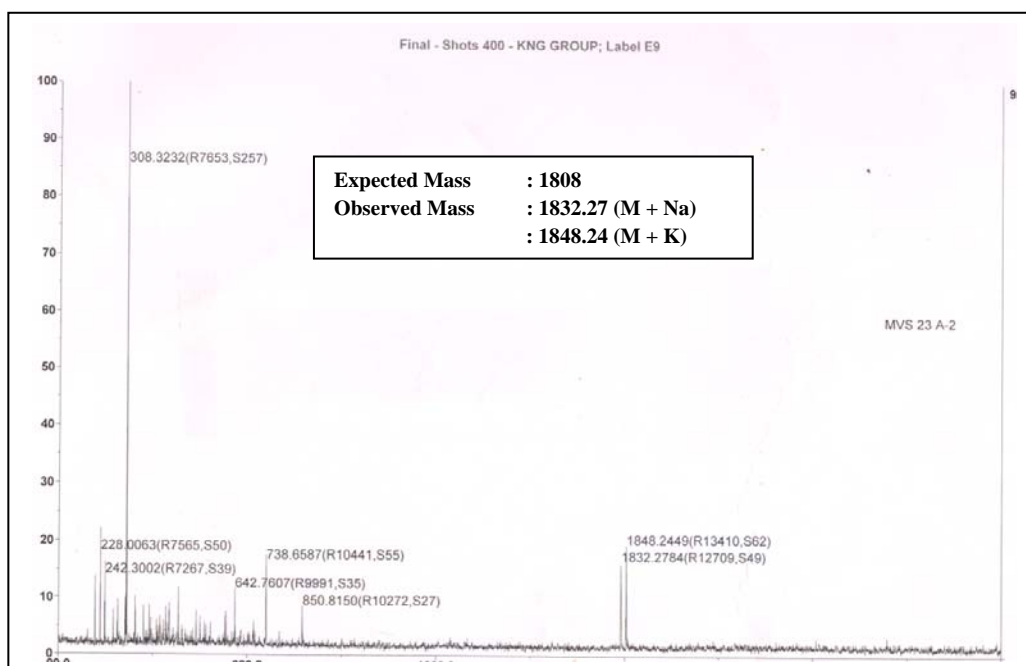
HPLC of peptide 20 (H₂N-Phe-Pro-4S-amp-Gly-[Pro-Hyp-Gly]₅-NH₂)**HPLC of peptide 21 (H₂N-Phe-4S-amp-4R-Amp-Gly-[Pro-Hyp-Gly]₅-NH₂)**

(C) MALDI-TOF of peptides 4-21**MALDI-TOF of peptide 4 (Ac-Phe-[Pro-Hyp-Gly]₅-Pro-4R-Amp-Gly-NH₂)****MALDI-TOF of peptide 5 (Ac-Phe-[Pro-Hyp-Gly]₅-Pro-4S-amp-Gly-NH₂)**

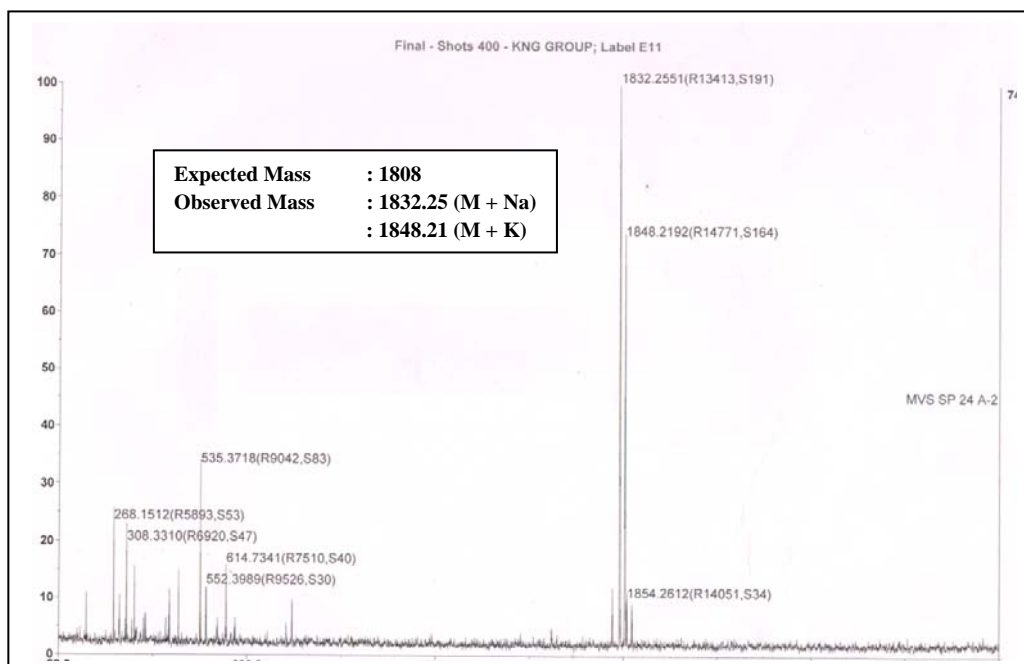
MALDI-TOF of peptide 6 (Ac-Phe-[Pro-Hyp-Gly]₅-4S-amp-4R-Amp-Gly-NH₂)



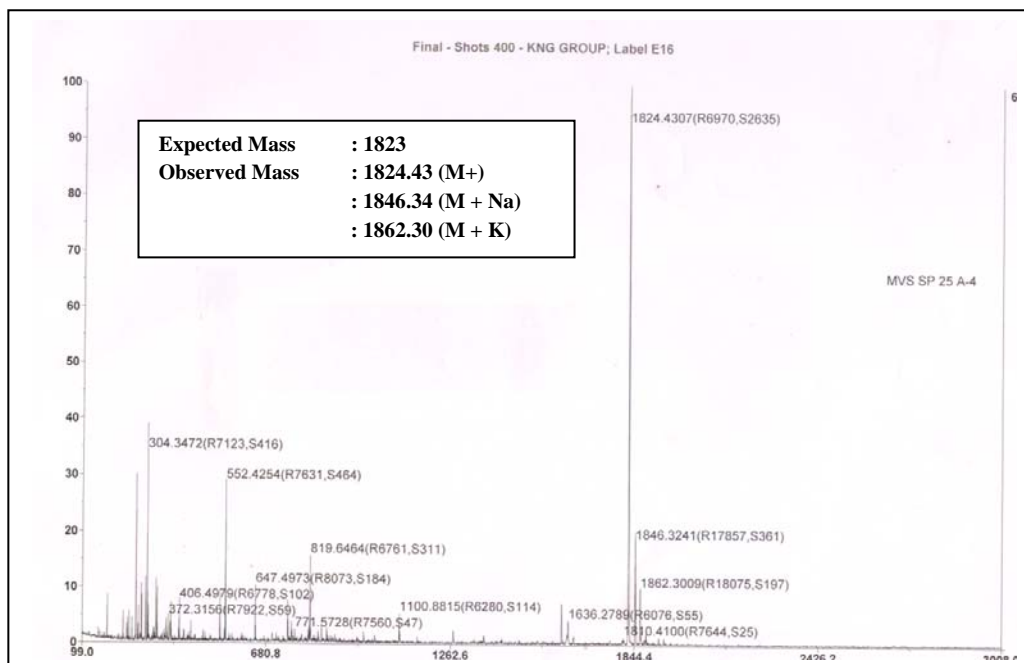
MALDI-TOF of peptide 7 (Ac-Phe-[Pro-Hyp-Gly]₂-Pro-4R-Amp-Gly-[Pro-Hyp-Gly]₃-NH₂)



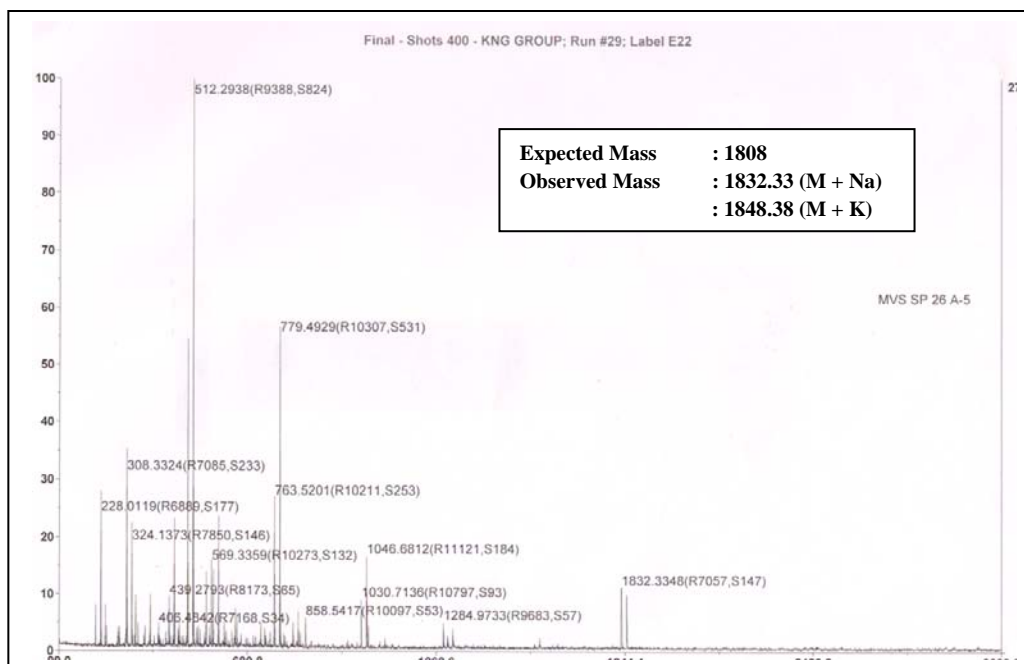
MALDI-TOF of peptide 8 (Ac-Phe-[Pro-Hyp-Gly]₂-Pro-4S-amp-Gly-[Pro-Hyp-Gly]₃-NH₂)



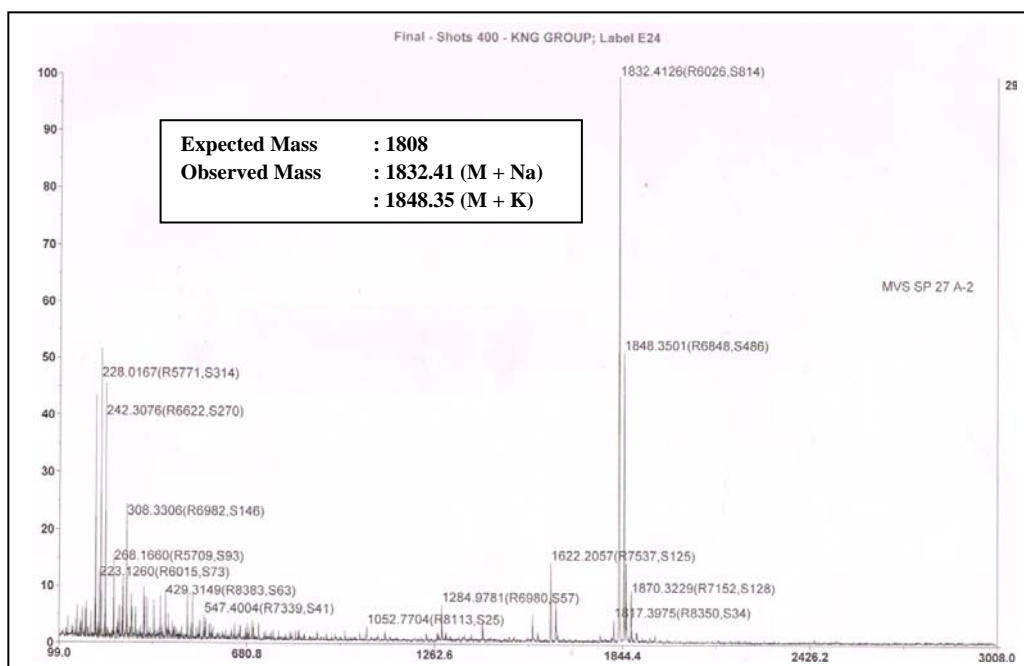
MALDI-TOF of peptide 9 (Ac-Phe-[Pro-Hyp-Gly]₂-4S-amp-4R-Amp-Gly-[Pro-Hyp-Gly]₃-NH₂)



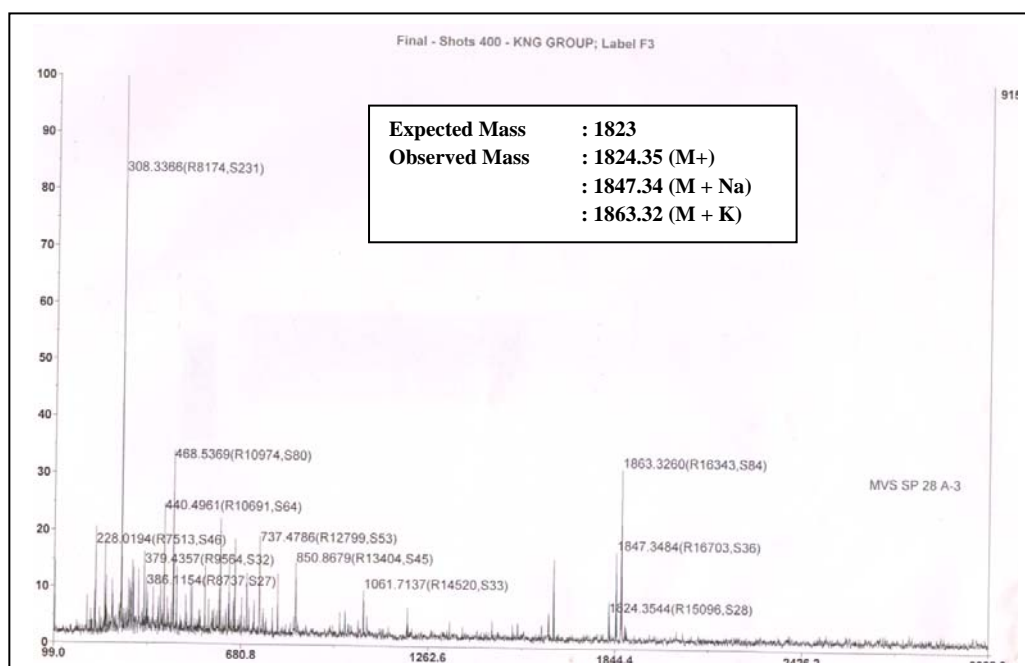
MALDI-TOF of peptide 10 (Ac-Phe-Pro-4R-Amp-Gly-[Pro-Hyp-Gly]₅-NH₂)



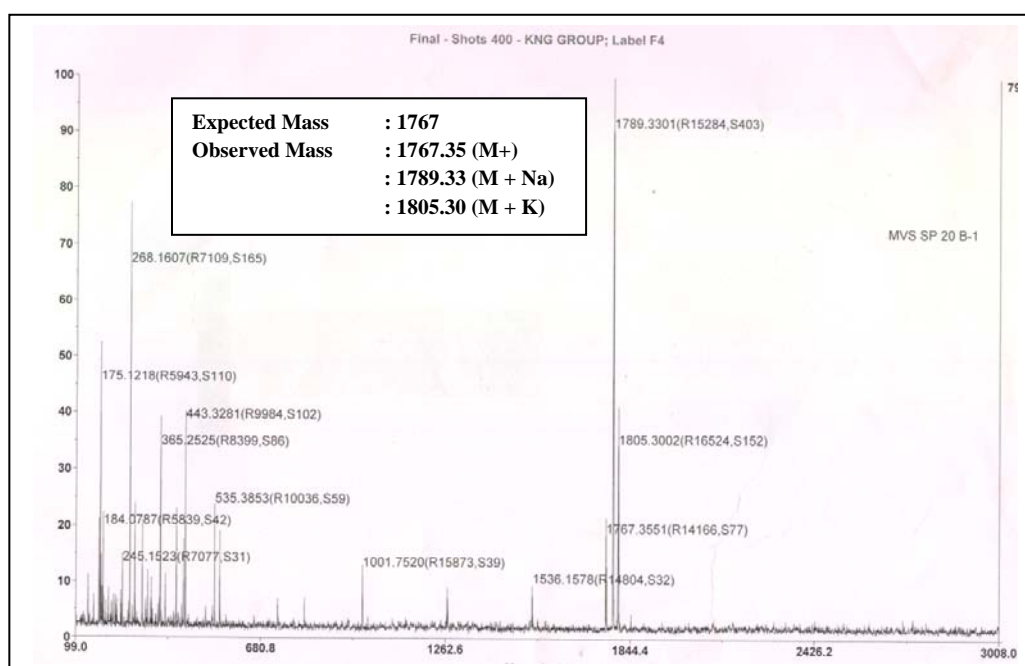
MALDI-TOF of peptide 11 (Ac-Phe-Pro-4S-amp-Gly-[Pro-Hyp-Gly]₅-NH₂)

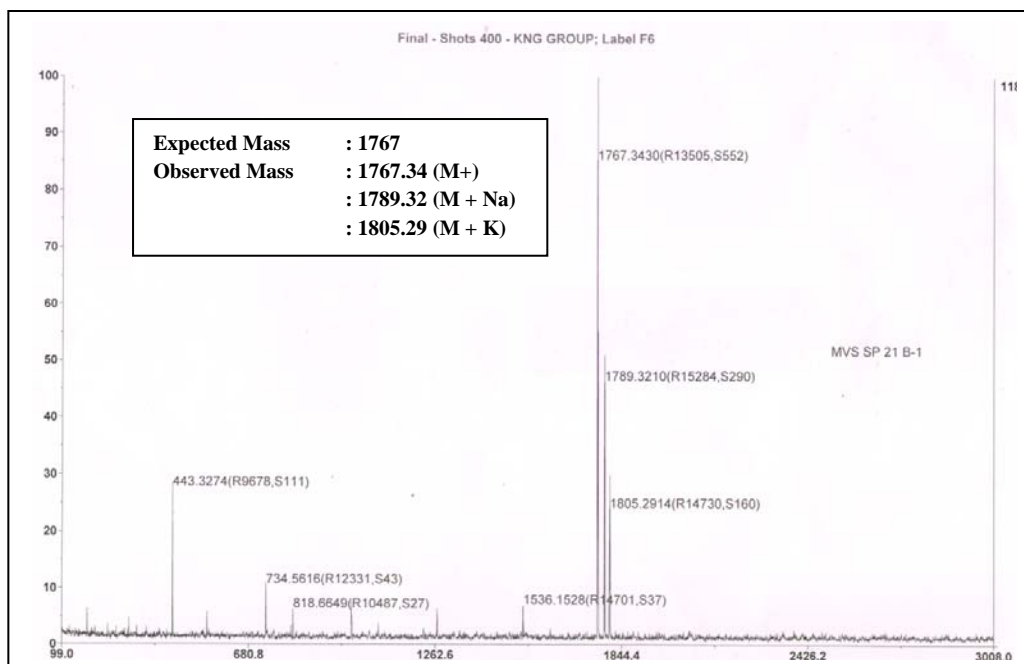
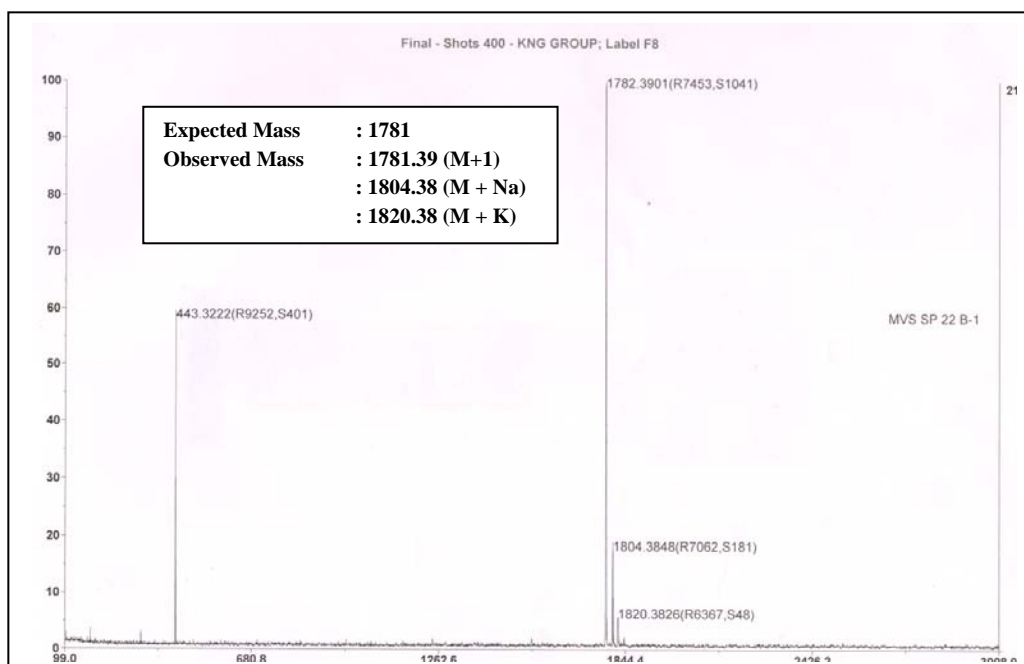


MALDI-TOF of peptide 12 (Ac-Phe-4S-amp-4R-Amp-Gly-[Pro-Hyp-Gly]₅-NH₂)

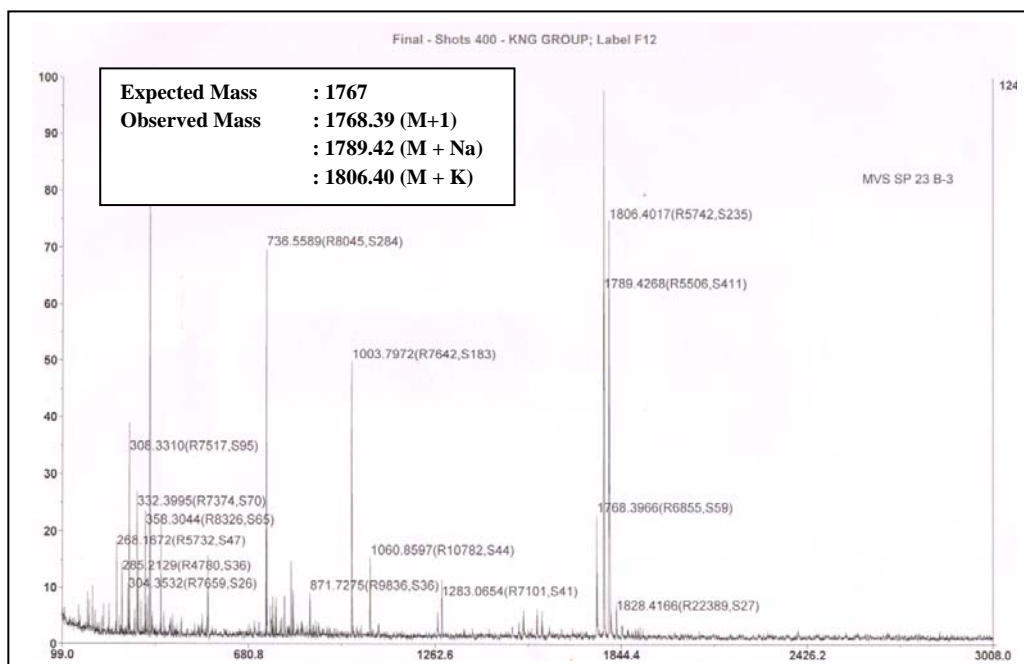


MALDI-TOF of peptide 13 (H₂N-Phe-[Pro-Hyp-Gly]₅-Pro-4R-Amp-Gly-NH₂)

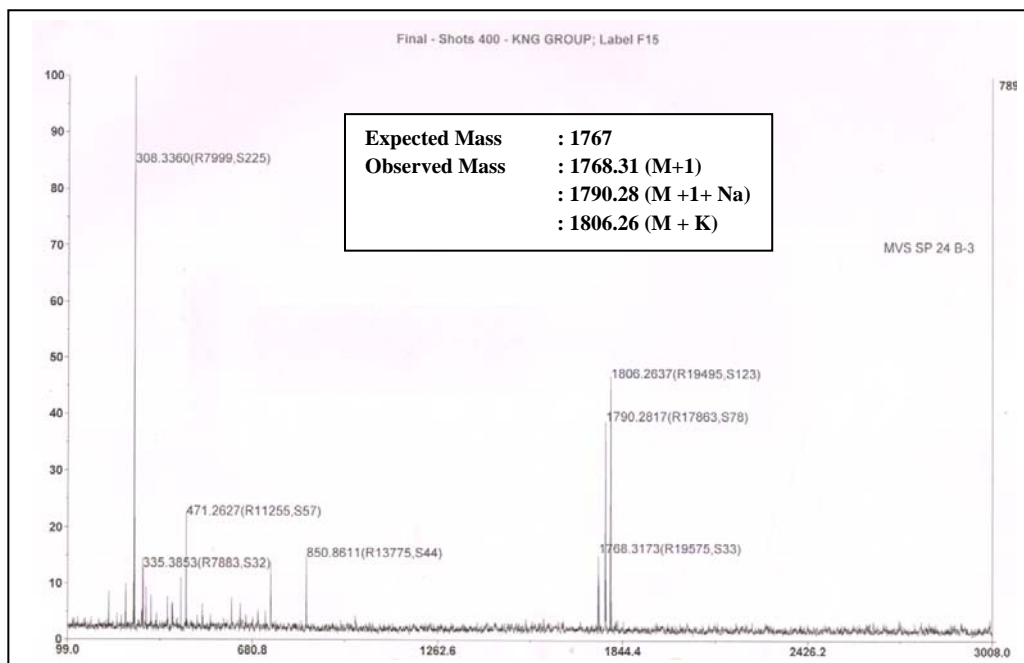


MALDI-TOF of peptide (14 H₂N-Phe-[Pro-Hyp-Gly]₅-Pro-4S-amp-Gly-NH₂)**MALDI-TOF of peptide 15 (H₂N-Phe-[Pro-Hyp-Gly]₅-4S-amp-4R-Amp-Gly-NH₂)**

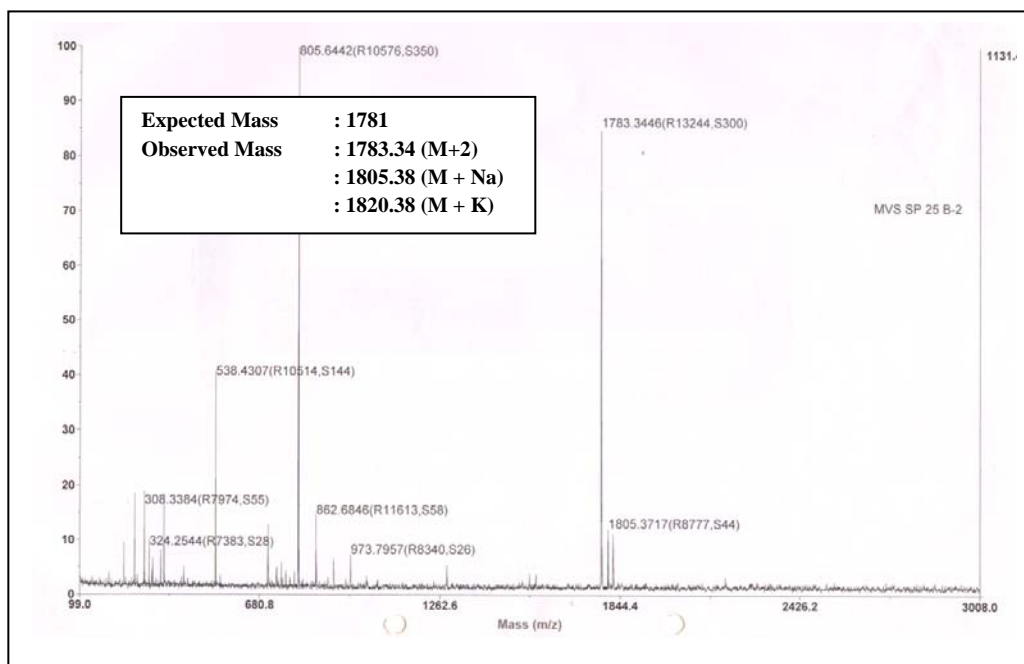
MALDI-TOF of peptide 16 (H₂N-Phe-[Pro-Hyp-Gly]₂-Pro-4R-Amp-Gly-[Pro-Hyp-Gly]₃-NH₂)



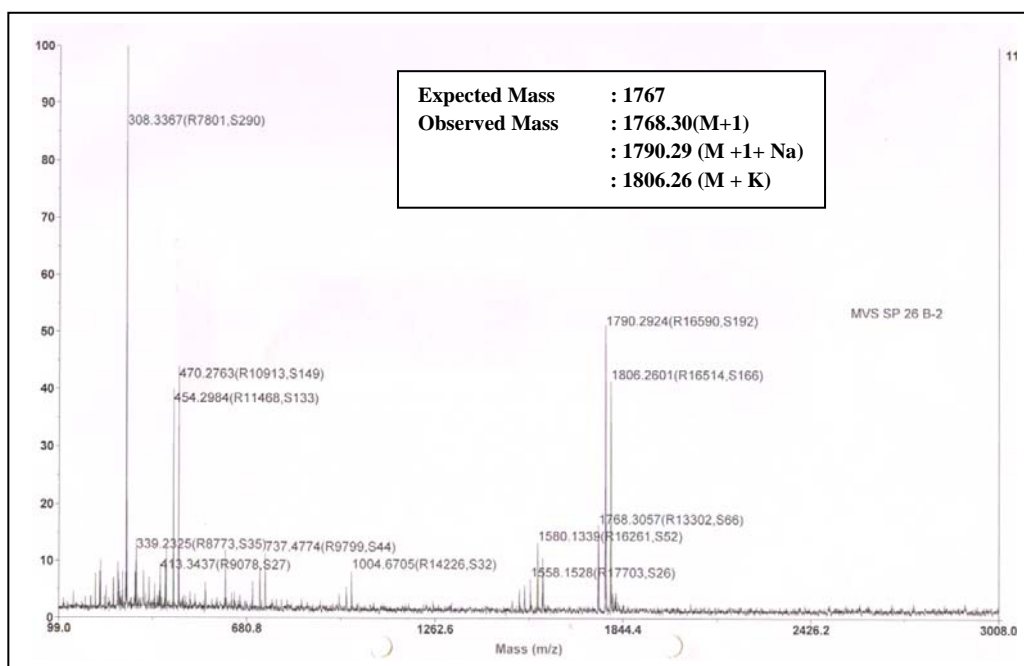
MALDI-TOF of peptide 17 (H₂N-Phe-[Pro-Hyp-Gly]₂-Pro-4S-amp-Gly-[Pro-Hyp-Gly]₃-NH₂)



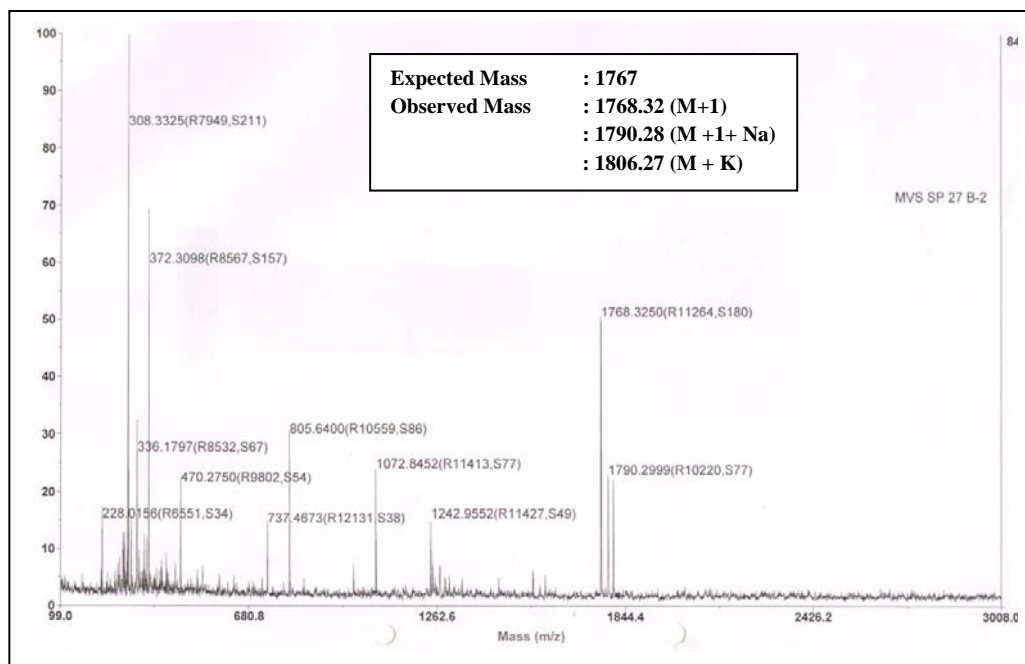
MALDI-TOF of peptide 18 (H₂N-Phe-[Pro-Hyp-Gly]₂-4S-amp-4R-Amp-Gly-[Pro-Hyp-Gly]₃-NH₂)



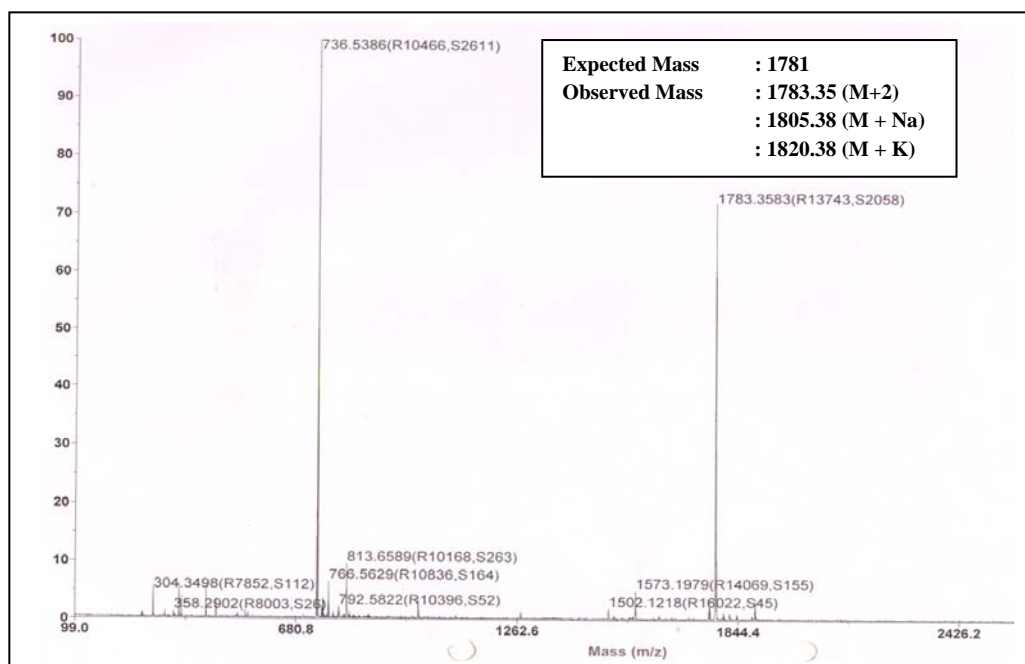
MALDI-TOF of peptide 19 (H₂N-Phe-Pro-4R-Amp-Gly-[Pro-Hyp-Gly]₅-NH₂)



MALDI-TOF of peptide 20 (H₂N-Phe-Pro-4S-amp-Gly-[Pro-Hyp-Gly]₅-NH₂)

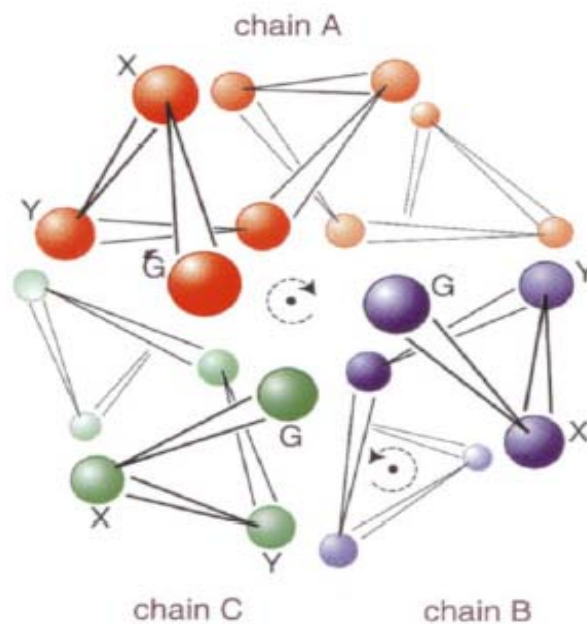


MALDI-TOF of peptide 21 (H₂N-Phe-4S-amp-4R-Amp-Gly-[Pro-Hyp-Gly]₅-NH₂)



Chapter 3

Collagen peptide $(\text{Pro-Hyp-Gly})_n$ analogues with α - aminoisobutyric acid substituted glycine: $(\text{Pro-Hyp-Aib})_n$



This chapter describes the role of substitution of 2-aminoisobutyric acid in place of glycine in collagen tripeptide. It demonstrates how the rigidification of α -carbon in glycine destabilizes the collagen triple helix.

1 Introduction

1.1 Substitution of Glycine

The absence of a stereocenter in glycine endows its Ramachandran plot with internal C_2 symmetry, and the absence of a side chain provides access to most ϕ ($C_{i-1}-N_i-C_i^\alpha-C_i$) and ψ ($N_i-C_i^\alpha-C_i-N_{i+1}$) torsion angles. Many of the angles accessible to glycine residues are accessible to D-amino acids but not L-amino acids. Accordingly, glycine can be a suitable target for substitution with a D-amino acid. Raleigh and co-workers¹ replaced glycine residues with D-amino acids in several globular proteins and found that the incorporation of D-amino acids can greatly enhance conformational stability. In addition, D-amino acids can be better than L-amino acids as glycine surrogates, retaining protein function without perturbing protein structure. The advantage of D-amino acid substitutions have a common feature—the target glycine residue always has $\phi > 0^\circ$.

A recent computational study by Dannenberg and co-workers² suggested that replacing the glycine residues in collagen strands with D-alanine or D-serine residues would stabilize the triple helix. Moreover, this work suggested that D-serine would have a larger stabilizing effect than D-alanine because of the formation of a hydrogen bond between its side-chain hydroxyl group and a carbonyl group in another strand of the triple helix (Figure 1).

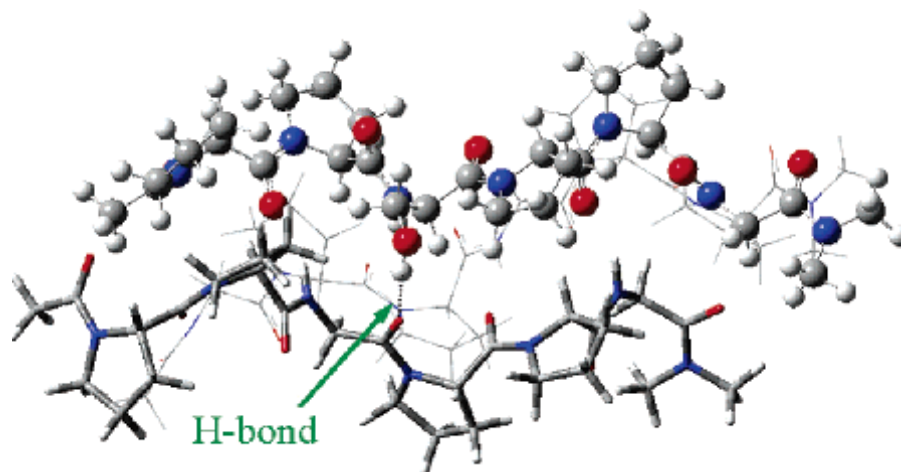


Figure 1: Structure of the triple helix containing one D-Serine in place of Gly. The additional H-bond is indicated by green arrow²

1.2 Importance of α -Aminoisobutyric Acid (Aib)

α -Aminoisobutyric acid (Aib) has two methyl groups on α carbon of glycine and due to this it rigidifies the α -carbon of glycine. It is well established that the natural, nonprotein amino acid, α -aminoisobutyric acid (Aib, α -methylalanine) has a very high helix propensity (helix inducing property) and a high potential as a β -sheet (strand) breaker.³ It is a common residue in peptides produced by microbes.⁴ Examples are linear peptides of 15-20 residues, such as alamethicin, antiamoebin, and emerimicin, which produce voltage-gated channels in lipid membranes. This suggests the biological importance of α -aminoisobutyric acid.⁵ Aib can be considered as being produced by the replacement of the $C\alpha$ proton of an alanine residue by a methyl group which significantly limits the rotational degree of freedom with respect to the dihedral backbone coordinates ϕ and ψ (rigidification of α carbon of glycine).

The Ramachandran plot reported by Marshall and Bosshard⁶ and by Burgess and Leach⁷ indicate two troughs of the potential energy landscape centered at ($\phi = -57^\circ$, $\psi = -47^\circ$) and ($\phi = 57^\circ$, $\psi = 47^\circ$), corresponding to right- and left-handed α -helices. Aib is *per se* achiral (the achiral Aib residue can adopt both left (α_L)- and right (α_R)-handed helical conformations), but differential sampling of the two enantiomeric conformations corresponding to the above clusters in the Ramachandran plot are induced by neighboring chiral amino acid residues. The two clusters encompass both α -helices and 3_{10} -helices with a more frequent sampling of the latter for shorter peptides.

The presence of geminal dialkyl substituents at the tetrahedral $C(\alpha)$ -atom imposes major steric restrictions on the energetically accessible conformational space, mostly limiting the Aib residue to the helical region of the Ramachandran map ($\phi = 60^\circ \pm 20^\circ$; $\psi = 30^\circ \pm 20^\circ$).⁸ Figure 2 illustrates a conformational energy map for Aib, computed using classical potential-energy functions.⁹ The deepest minima lie in the right- and left-handed $3_{10/\alpha}$ -helical regions of the ϕ , ψ space. These minima correspond to the region of overlap obtained by superposition of the Ramachandran maps for L-Ala and D-Ala residues. In addition to the helical regions, minima are also observed in the fully extended and semi extended regions of the ϕ , ψ space. Conformations corresponding to β -strand structures ($\phi \approx 120^\circ$, $\psi \approx 120^\circ$) are strongly disallowed

because of short contacts between one of the methyl groups at C(β) and the C=O group of the preceding residue, and the NH group of succeeding residue.

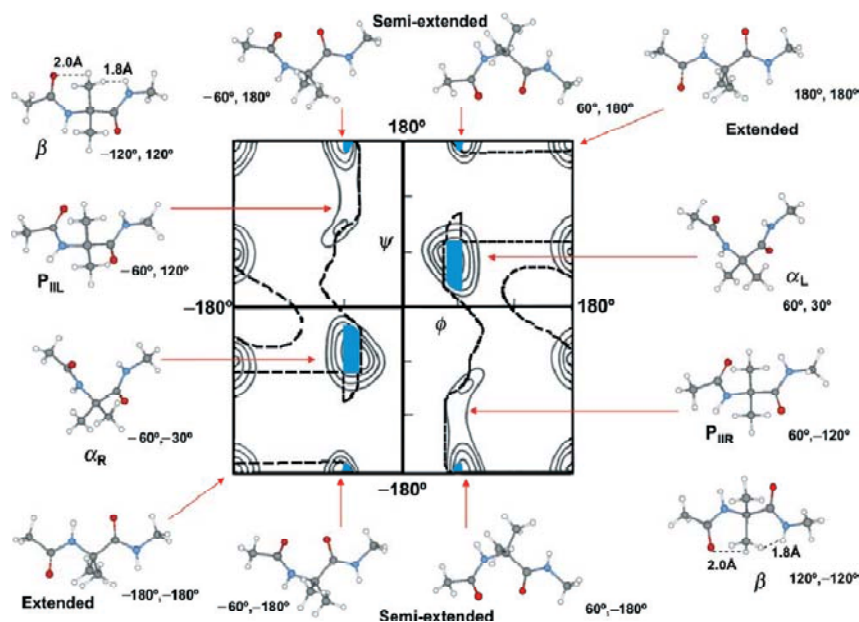


Figure 2: Superposition of the allowed region of the ϕ , ψ space determined by hard sphere criteria Ac-Aib-NHMe and conformational energy map for an Aib residue.⁹

The homooligomers of α -aminoisobutyric acid (Aib) generally adopt 3_{10} -helix which accounts for approximately 10% of helical residues in crystal structures (PDB).¹⁰ The 3_{10} -helical segments in proteins are generally short with an average length of four residues and differ from the canonical α -helix not only with respect to its dihedral coordinates but also by a $(i, i + 3)$ hydrogen-bonding arrangement.¹¹ Augspurger *et.al.*¹² investigated several Aib-based, aliphatic peptides of intermediate size in organic solvents and obtained a very thermostable 3_{10} structure. De Filippis *et.al.*¹³ confirmed and explained the thermostability of Aib with the reduction of entropy of this amino acid residue in the unfolded state. Numerous studies have been conducted to identify the conditions at which Aib stabilizes a 3_{10} -helix in peptides. Basically, a peptide gains an extra internal hydrogen bond in a 3_{10} -helix while steric interactions between residues favour α -helical torsional angles. At a critical length depending on the dielectric constant of the solvent, the two interactions approximately balance, and populations of both helices coexist. For this reason, short Aib-based peptides generally prefer 3_{10} -helices, whereas α -helices are more likely for longer peptides.¹⁴

1.3 Present work: Rationale

The work in this chapter of thesis is mainly devoted to the synthesis and structural characterization of a new class of synthetic collagen peptides with α -aminoisobutyric acid (Aib) as a glycine surrogate.

In view of the interesting properties of Aib peptides, the stability of collagen triple helix having glycine replaced by sterically unfavoured α -aminoisobutyric acid (Aib) is studied. The presence of geminal dialkyl substituents on a carbon atom results in a restriction of the flanking torsional angles. The collagen oligomer mimic containing α -aminoisobutyric acid in place of the canonical glycine has been synthesized. For comparison, control collagen peptide was synthesized with normal glycine substitution.

The specific objectives of this chapter are

- Synthesis of Aib collagen and control collagen peptides AcPhe(Pro-Hyp-Gly)₆ (Peptide **30**), AcPhe(Pro-Hyp-Aib)₆ (Peptide **31**), H₂NPhe(Pro-Hyp-Gly)₆ (Peptide **32**) and H₂NPhe(Pro-Hyp-Aib)₆ (Peptide **33**) by solid phase synthesis protocol
- Cleavage of peptides from the solid support, purification and characterization.
- Investigation of triple helix forming abilities of peptides **30-33** by using concentration and temperature dependent CD spectroscopy.

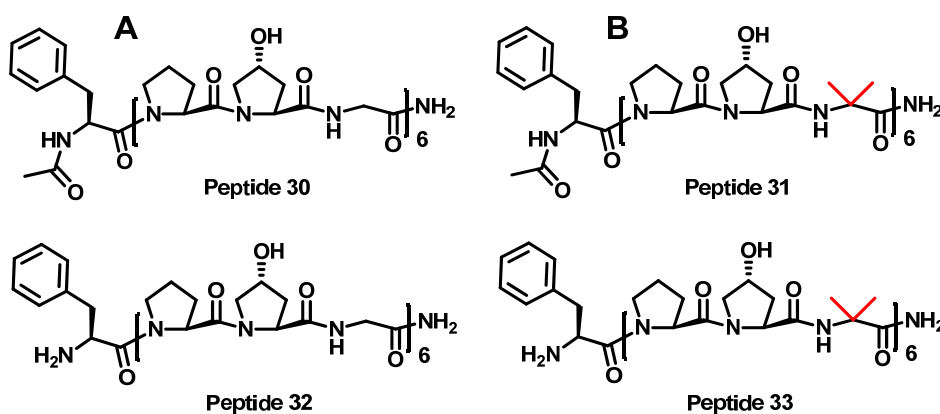


Figure 3: (A) Collagen model peptide with glycine substitution (B) Collagen model peptide with α -aminoisobutyric acid (Aib) substitution.

1.3.1 Synthesis and characterization of peptides 30-33

The synthesis of peptides **30** to **33** was achieved from the commercially available N-Fmoc proline (**17**), N-Fmoc *O-tert* butyl hydroxyproline (**13**), N-Fmoc glycine (**45**)

and N-Fmoc α -aminoisobutyric acid (**46**). The synthesis was done by solid phase synthesis of the commercially available Rink amide resin by a linear coupling approach (Scheme 1).

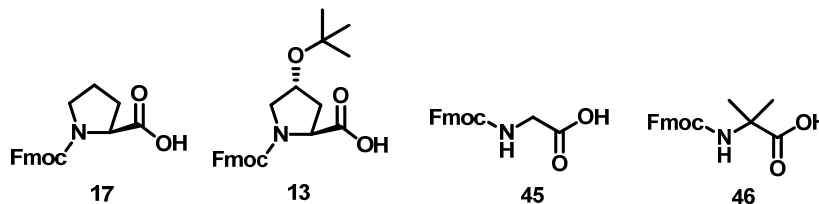
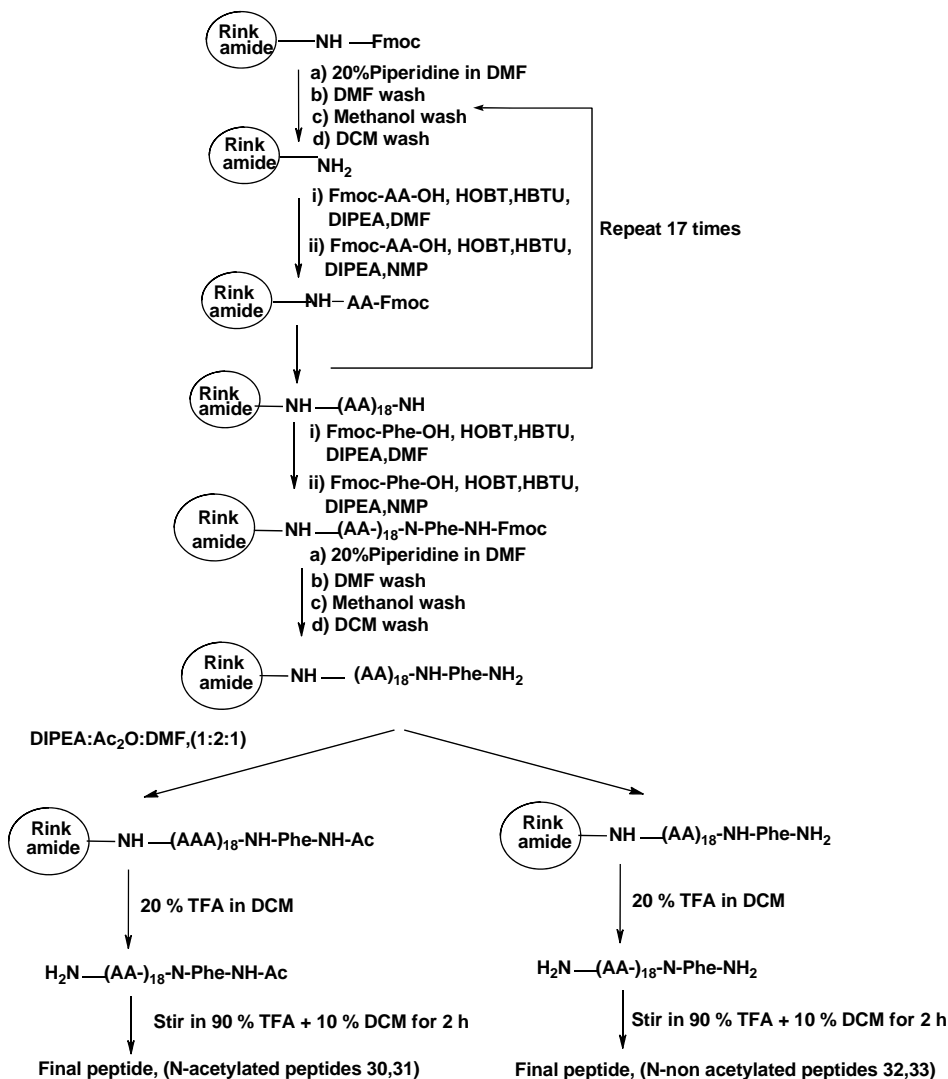


Figure 4: Monomers used for peptide synthesis

1.3.2 Solid phase peptide synthesis

The target peptides (**30** - **33**) were synthesized by manual solid phase synthesis (Scheme 1) on readily available rink amide resin using standard Fmoc chemistry, followed by cleavage to yield the peptides as C-terminal amides. The commercially available Kieselghur supported N,N-dimethylacrylamide resin (with Rink-amide linker) with Fmoc group was cleaved by 20% piperidine in DMF. The monomers were coupled as free acids using *in situ* activation procedure with 3 eq. of amino acid, HBTU, HOBT and 8 eq. of DIPEA. The coupling reaction was repeated second time using *N*-Methyl-2-pyrrolidone (NMP) as solvent. The N-Fmoc deprotection reaction was done by using 20% piperidine in DCM. The deprotection and coupling reactions were monitored by using qualitative Ninhydrin (Kaiser)¹⁵ test for glycine and Chloranil test¹⁶ for iminoacids. After completion of synthesis the resin was divided into two equal parts. One portion of the resin was capped with Ac₂O and second portion of the resin remain kept as amine function. The peptide was cleaved from the resin using 20% TFA in DCM. As the *t*-butyl group removal requires stronger acidic conditions, the deprotection of side chain *t*-butyl group on hydroxyproline residues of the peptide was carried out with 90% TFA in DCM. The *t*-butyl cation formed during the deprotection can lead to N-alkylation of the amines and to prevent such side reactions, 0.1% TIS (triisopropylsilane) was used as a scavenger. The peptides were purified on semi-preparative RP-C-18 HPLC column using water-acetonitrile gradient to which 0.1% TFA added as ionizer. Peptides **30** and **31** are blocked as N-terminal acetylated and C-terminal amidated to remove terminal electrostatic repulsion. Peptide **32** and **33** kept as N-terminal free amine to evaluate the effect of charge on triplex forming abilities.

Scheme 1: Schematic representation of SPPS of peptides **30-33**

The peptides (**30-33**) were purified on semi-preparative RP-18 HPLC column using water-acetonitrile gradient to which 0.1% TFA added. The purity of the peptides were found to be greater than 97% and the structural integrity of the peptides were further confirmed by MALDI-TOF mass spectrometry which agreed closely with the calculated values (Table 1).

Table 1: Calculated and observed masses for peptides **30-33**

Peptide	Mol. Formula	Mass (cal)	Mass (obs)
Peptide 30 AcPhe(Pro-Hyp-Gly) ₆	C ₈₃ H ₁₁₆ N ₂₀ O ₂₆	1809.97	1831.5874 [M + Na] ⁺ 1847.5459 [M + K] ⁺
Peptide 31 AcPhe(Pro-Hyp-Aib) ₆	C ₉₅ H ₁₄₀ N ₂₀ O ₂₆	1978.29	2000.2661 [M + Na] ⁺ 2016.2404 [M + K] ⁺
Peptide 32 H ₂ N-Phe(Pro-Hyp-Gly) ₆	C ₈₁ H ₁₁₄ N ₂₀ O ₂₅	1767.89	1767.8311 [M] ⁺ 1789.7845 [M + Na] ⁺ 1805.811 [M + K] ⁺
Peptide 33 H ₂ N-Phe(Pro-Hyp-Aib) ₆	C ₉₃ H ₁₃₈ N ₂₀ O ₂₅	1936.20	1959.9584 [M + Na] ⁺ 1973.8965 [M + K] ⁺

1.4 Conformational Study by CD Spectroscopy

1.4.1 CD spectra of N-capped peptides 30-31: The CD spectra of peptides **30-31** were recorded at 200 μ M concentration at pH 7.2 at 20 °C and are shown in Figure 5. The CD data reveal that at 200 μ M, both the peptides show positive band at 220-225 nm and negative band at 200-205 nm which is a characteristic CD hallmark of collagen peptide. The peptide AcPhe(Pro-Hyp-Aib)₆ **31** has a CD spectrum similar to that of peptide AcPhe(Pro-Hyp-Gly)₆ **30**, but having higher ellipticity of maxima (10.3 $\times 10^3$ deg cm² dmol⁻¹) and lower minima (-25.8 $\times 10^3$ deg cm² dmol⁻¹). This indicates more extended structure adopted by peptide **31** than the control collagen peptide **30**.¹⁷

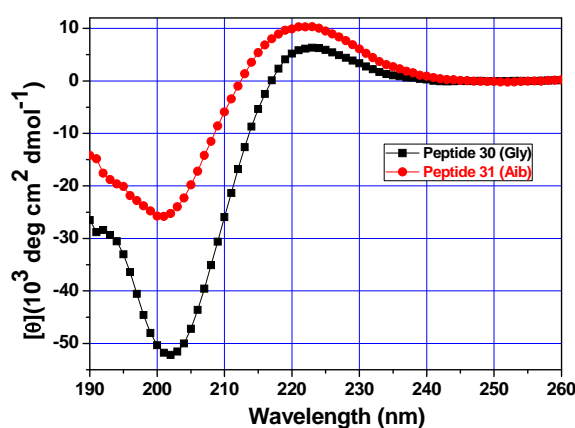


Figure 5: CD profiles of peptide **30** (■), peptide **31** (●), all at 200 μ M (pH 7.2, Na-Phosphate Buffer)

1.4.2 CD spectra of uncapped peptides 32-33: The CD spectra of peptides **32-33** were recorded at 200 μ M concentration in pH 7.2 at 20 °C and are shown in Figure 6.

The CD data reveal that at 200 μM , both the peptides show positive band at 220-225 nm and a negative band at 200-205 nm which is a characteristic of collagen peptide. The peptide $\text{H}_2\text{N-Phe(Pro-Hyp-Aib)}_6$ **33** has a CD spectrum similar to that of peptide $\text{H}_2\text{N-Phe(Pro-Hyp-Gly)}_6$ **32**, but with a higher ellipticity of maxima and lower minima. This indicates a more extended structure adopted by Aib peptide **33** than control collagen peptide **32**.

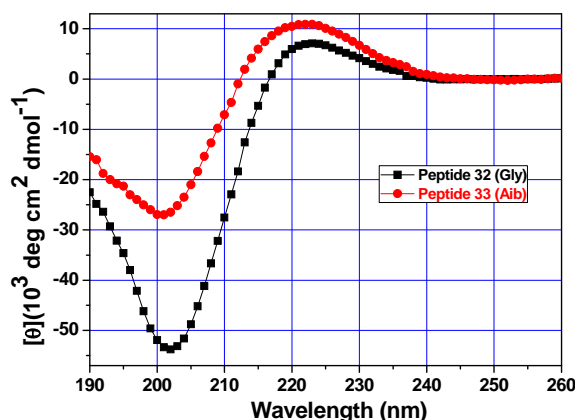


Figure 6: CD profiles of peptide **32** (■), peptide **33** (●), all at 200 μM (pH 7.2, Na-Phosphate Buffer)

1.4.3 Concentration dependent CD spectroscopy for peptides 30-33

Formation of triple-helical structure is a concentration dependent phenomenon which was shown by Goodman *et al.*¹⁸ Single stranded chains assemble to the triple-helical conformation with increase in peptide concentration in solution. The percentage of triple-helical structure is close to a maximum when the concentration is greater than its critical triple-helical concentration. The magnitude of the ratio of positive to negative band intensity in the CD spectra of collagen peptides ($R_{p/n}$) is used to quantitate the triple-helical strength. In the present study, this parameter has been used to determine the formation of concentration dependent triple-helix at pH 7.2 for all peptides.

1.4.3a Peptide 30 (AcPhe(Pro-Hyp-Gly)₆): Figure 7A shows the CD spectra of peptide **30** recorded at 20 °C in the concentration range of 50 μM –500 μM at pH 7.2. In the entire concentration range, peptide **30** shows similar positive and negative maxima at 223 nm and 202 nm respectively which are characteristics of collagen like triple-helical structure, the magnitude of positive and negative bands varying as a

function of concentration. Importantly, all the spectral traces pass through an isobestic point at 217 nm. Figure 7B shows a plot of $R_{p/n}$ values derived from these spectra against concentration of peptide **30** at pH 7.2. The $R_{p/n}$ value increases rapidly from 0.073 upto 0.104 to reach a near saturation at 300 μM and remains nearly constant thereafter. A critical triple-helical concentration of $\sim 300 \mu\text{M}$ is derived from these plots. Hence, all further studies are performed at a concentration of 0.3 mM peptides where peptides are in fully associated form.

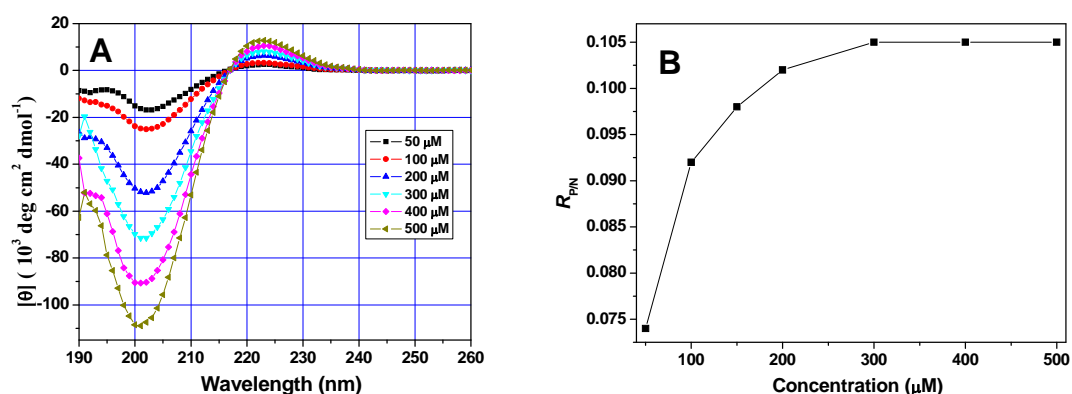


Figure 7: (A) CD spectra of peptide **30** at 20 °C at concentration from 50-500 μM at pH 7.2. (B) Plot of $R_{p/n}$ values obtained from these spectra against the concentration of peptide **30**.

1.4.3b Peptide 31 AcPhe(Pro-Hyp-Aib)₆: Figure 8A shows the CD spectra of Aib peptide **31** recorded at 20 °C in the concentration range of 50 μM – 500 μM at pH 7.2.

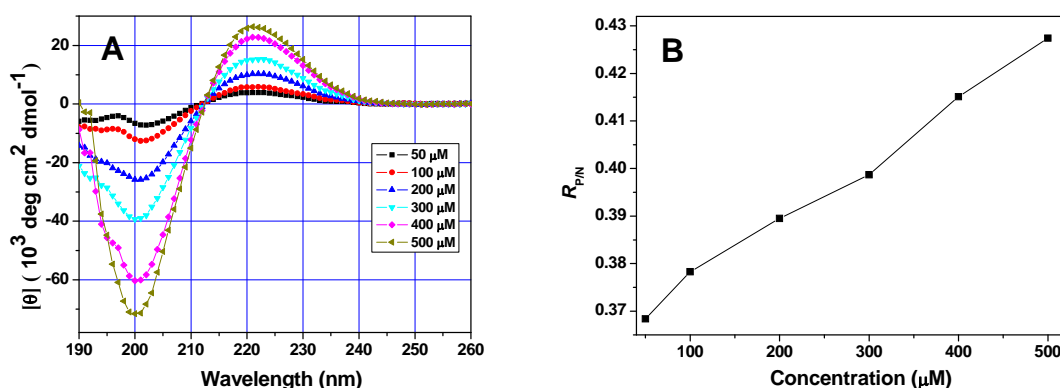


Figure 8: (A) CD spectra of peptide **31** at 20 °C at concentration from 50-500 μM at pH 7.2. (B) Plot of $R_{p/n}$ values obtained from these spectra against the concentration of peptide **31**.

In the entire concentration range, peptide **31** shows positive maxima shifted from 223 nm to 221 nm and negative band shifted from 203 nm to 201 nm with increasing

concentration. All the spectral traces pass through an isoelliptic point at 213 nm which is also shifted to lower wavelength compared to peptide **30**. Figure 8B shows a plot of $R_{p/n}$ values derived from these spectra against concentration of peptide **31** at pH 7.2. The ratio of intensity of positive to negative band of CD spectra increase with higher concentration in a linear manner. The $R_{p/n}$ value increases rapidly from 0.37 to 0.43 linearly with increase in concentration without reaches saturation. This is characteristic of collagen like (PPII) single-helix rather than triplex.

1.4.3c Peptide 32 H₂N-Phe(Pro-Hyp-Gly)₆: Figure 9A shows the CD spectra of peptide **32** recorded in the concentration range of 50 μ M – 500 μ M at pH 7.2. In the entire concentration range, peptide **32** shows similar positive and negative maxima at 224 nm and 203 nm respectively. Like peptide **30**, all the spectral traces pass through an isobestic point at 217 nm. Figure 9B shows a plot of $R_{p/n}$ values derived from these spectra against concentration of peptide **32** at pH 7.2. The $R_{p/n}$ value increases rapidly from 0.075 upto 0.104 to reach a near saturation at 300 μ M and remains nearly constant thereafter. Hence, all further studies are performed at a concentration of 0.3 mM peptides where peptides are in fully associated form.

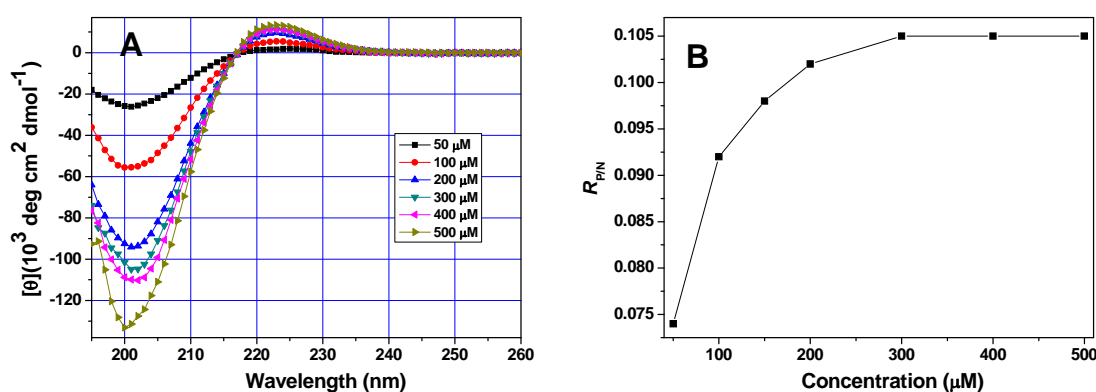


Figure 9: (A) CD spectra of peptide **32** at 20 °C at concentration from 50-500 μ M at pH 7.2. (B) Plot of $R_{p/n}$ values obtained from these spectra against the concentration of peptide **32**.

1.4.3d Peptide 33 H₂N-Phe(Pro-Hyp-Aib)₆: Figure 10A shows the CD spectra of Aib peptide **33** recorded in the concentration range of 50 μ M – 500 μ M at pH 7.2. In the entire concentration range, Aib peptide **33** shows positive maxima shifted from 224 nm to 222 nm and negative band shifted from 204 nm to 202 nm with increasing concentration. All the spectral traces pass through an isoelliptic point at 214 nm which

is shifted to lower wavelength compared to peptide **32**. Figure 10B shows a plot of $R_{p/n}$ values derived from these spectra against concentration of peptide **33** at pH 7.2. The $R_{p/n}$ value increases rapidly from 0.26 upto 0.326 linearly with increase in concentration without reaching saturation.

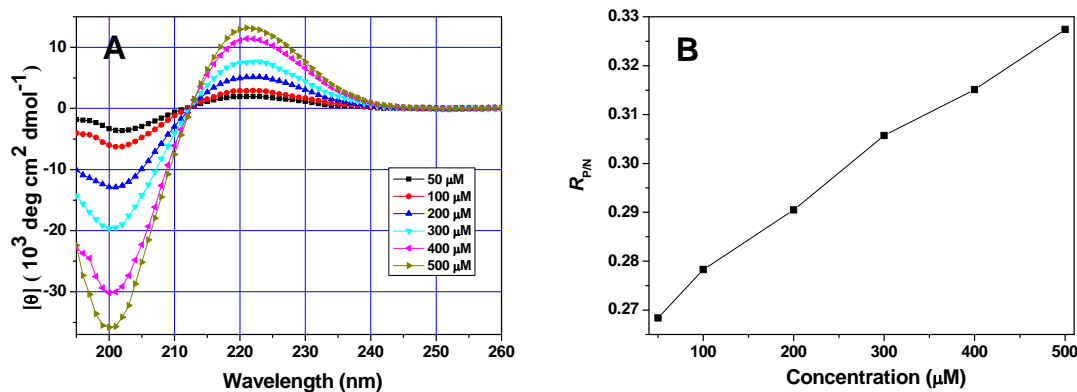


Figure 10: (A) CD spectra of peptide **33** at 20 °C at concentration from 50-500 μM at pH 7.2. (B) Plot of $R_{p/n}$ values obtained from these spectra against the concentration of peptide **33**.

The $R_{p/n}$ values observed for peptides **30** - **33** at pH 7.2 are shown in Table 2.

Table 2

Peptide	+ve band (nm)	-ve band (nm)	$R_{p/n}$
AcPhe(Pro-Hyp-Gly) ₆ 30	223	202	0.105
AcPhe(Pro-Hyp-Aib) ₆ 31	221-223	201-203	0.37-0.43
H ₂ N-Phe(Pro-Hyp-Gly) ₆ 32	224	203	0.104
H ₂ N-Phe(Pro-Hyp-Aib) ₆ 33	222-224	202-204	0.265-326

1.4.4 CD thermal denaturation study of peptides 30-33 in buffer at pH 7.2

In order to get information about the relative strengths of peptides **30-33** CD-spectroscopic study as a function of temperature was carried out by observing the changes in ellipticity at 222 nm - 225 nm with temperature. Upon heating, a decrease in molar ellipticity with increase in temperature was observed for all the peptides.

1.4.4a Peptide 30 AcPhe(Pro-Hyp-Gly)₆: Figure 11A shows the temperature dependent CD spectra for peptide **30** recorded from 5 to 90 °C. As the temperature increases from 5 to 90 °C, the positive maxima and negative minima of molar ellipticity decreases. The isobestic point seen at 213 nm is indicative of the

transformation of peptide from one conformation to another conformation. The CD-thermal denaturation plot of molar ellipticity at 223-224 nm *versus* temperature at pH 7.2 for peptide **30** is shown in Figure 11B. Figure 11C shows the first derivative curve obtained from the sigmoidal fit of data of Figure 11B and the T_m values for peptide **30** obtained from minima in plot of Figure 11C. At pH 7.2, peptide **30** shows cooperative melting transition with increasing temperature. This indicates that peptide **30** AcPhe(Pro-Hyp-Gly)₆ is associated in a triple helical structure at physiological pH with $T_m = 27^\circ\text{C}$.

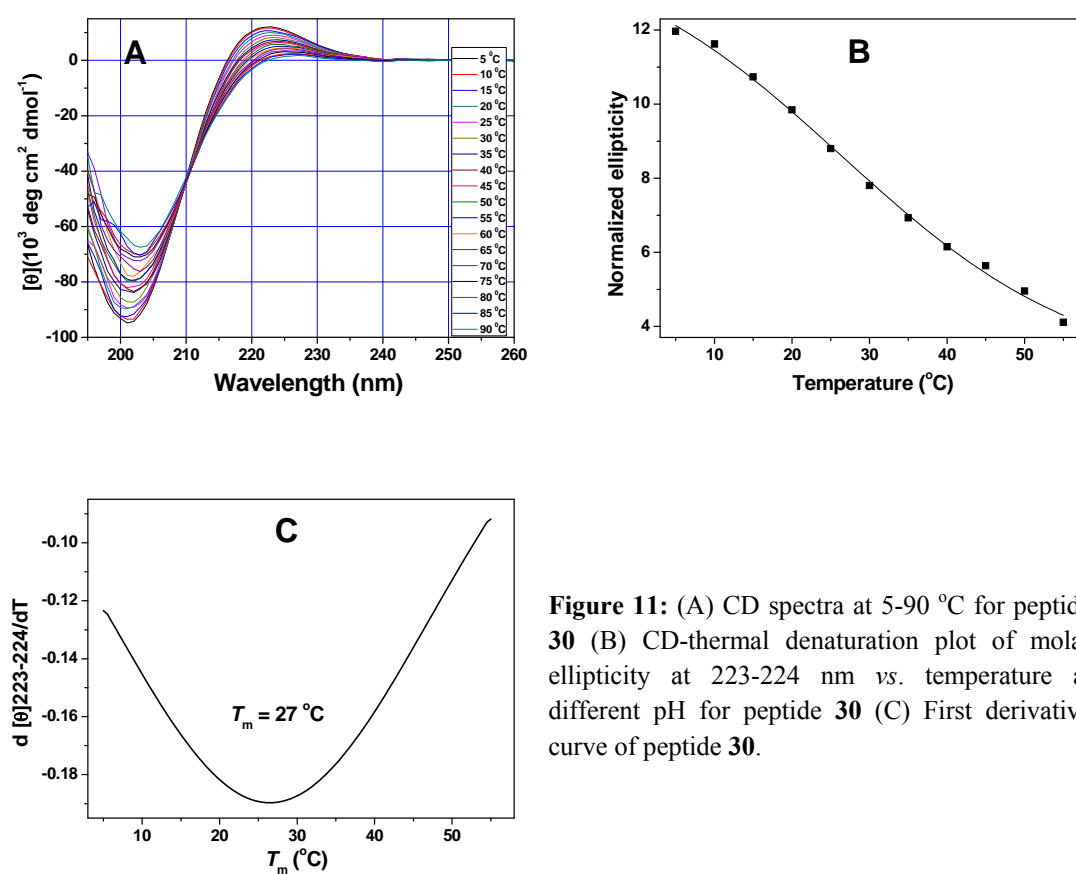


Figure 11: (A) CD spectra at 5-90 °C for peptide **30** (B) CD-thermal denaturation plot of molar ellipticity at 223-224 nm *vs.* temperature at different pH for peptide **30** (C) First derivative curve of peptide **30**.

1.4.4b Peptide 31 AcPhe(Pro-Hyp-Aib)₆: Figure 12A shows the CD spectra obtained by heating the solution from 5 to 90 °C, which indicates that the positive maxima and negative minima of molar ellipticity decreases with temperature. No isobestic point was observed at 212-218 nm which is indicative of the triple helix formation. Figure 12B shows the CD-thermal denaturation plot of molar ellipticity at 223 nm *versus* temperature at pH 7.2 for peptide **31**. At pH 7.2, the peptide **31** upon heating a solution

results in only a linear decrease in ellipticity rather than the cooperative transition which is characteristic of triple-helix denaturation. This indicates that peptide **31** AcPhe(Pro-Hyp-Aib)₆ is present in a single helical structure rather than triple helix structure at physiological pH. Figure 12C shows the first derivative curve obtained from the sigmoidal fit of data of Figure 12B for peptide **31** which is a straight line. This in combination with spectral behaviour in concentration dependent CD experiment, confirmed that peptide **31** is in single helix chain rather than triplex in water at pH 7.2

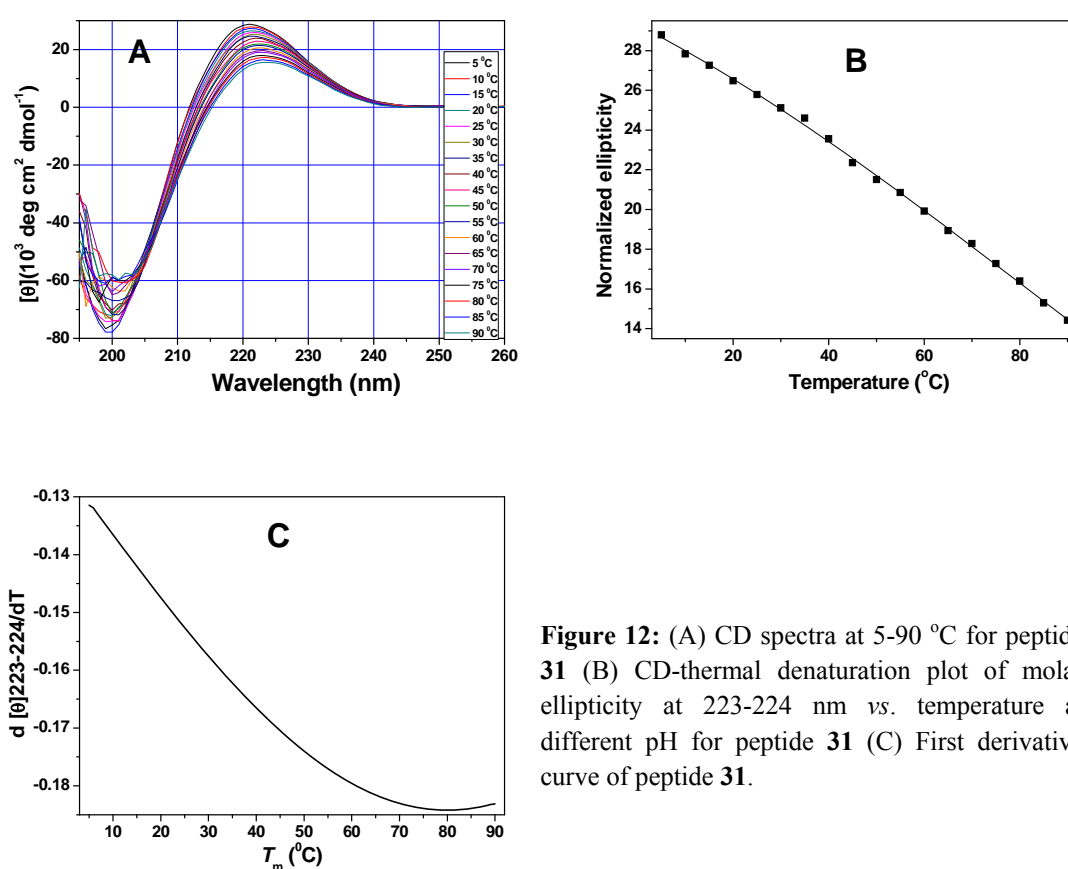


Figure 12: (A) CD spectra at 5-90 °C for peptide **31** (B) CD-thermal denaturation plot of molar ellipticity at 223-224 nm vs. temperature at different pH for peptide **31** (C) First derivative curve of peptide **31**.

1.4.4c Peptide 32 H₂N-Phe(Pro-Hyp-Gly)₆: Figure 13A shows the temperature dependent CD spectra for peptide **32** recorded from 5 to 70 °C. As the temperature increases from 5 to 90 °C, the positive maxima and negative minima of molar ellipticity decreases. The isobestic point seen at 208 nm is indicative of the transformation of peptide form one conformation to another conformation. Figure 13B shows the CD-thermal denaturation plot of molar ellipticity at 222 nm *versus*

temperature at pH 7.2 for peptide **32**. Figure 13C shows the first derivative curve obtained from the sigmoidal fit of data of Figure 13B and the T_m values for peptide **32** obtained from minima in plot of Figure 13C. At pH 7.2, peptide **32** shows cooperative melting transition with increasing temperature. This indicates that peptide **32** H₂N-Phe(Pro-Hyp-Gly)₆ is associated in a triple helical structure at physiological pH with $T_m = 23$ °C.

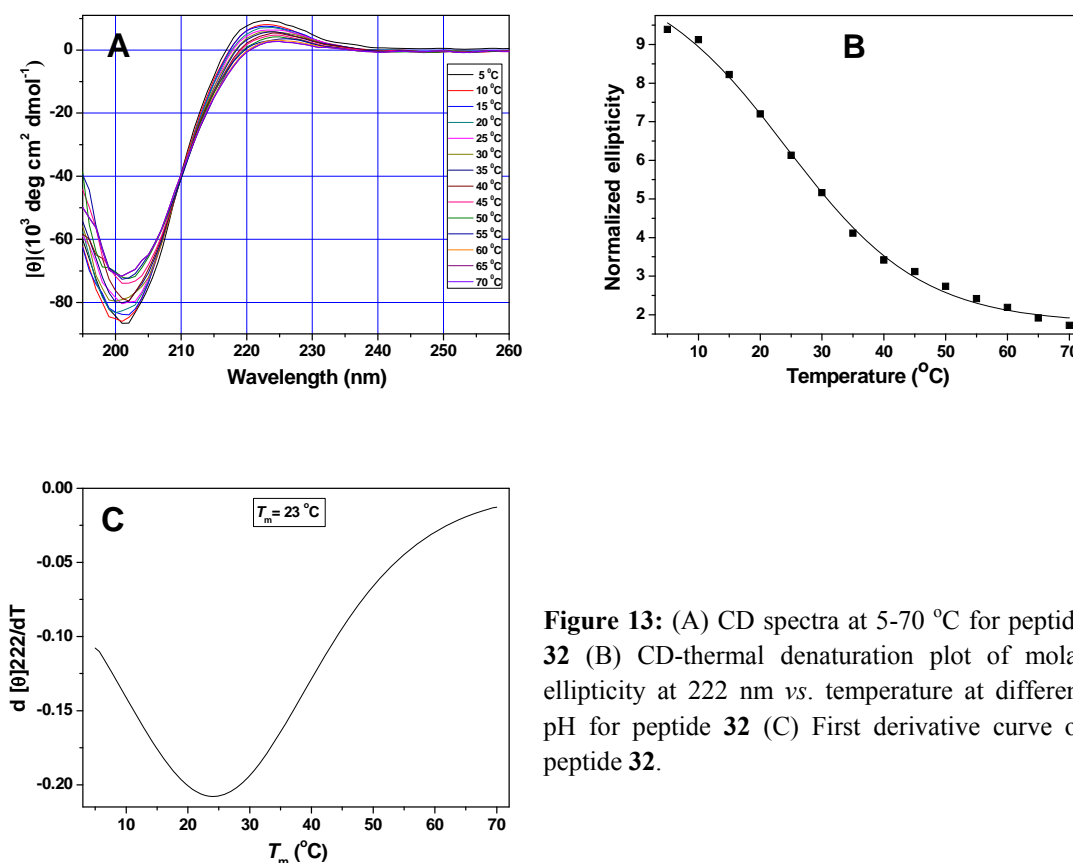


Figure 13: (A) CD spectra at 5-70 °C for peptide **32** (B) CD-thermal denaturation plot of molar ellipticity at 222 nm vs. temperature at different pH for peptide **32** (C) First derivative curve of peptide **32**.

1.4.4d Peptide 33 H₂N-Phe(Pro-Hyp-Aib)₆: Figure 14A shows the CD spectra obtained by heating the solution from 5 to 80 °C, which indicates that the positive maxima and negative minima of molar ellipticity decreases with temperature. No isobestic point was observed at 208-218 nm which is indicative of the triple helix formation. Figure 14B shows the CD-thermal denaturation plot of molar ellipticity at 222 nm versus temperature at pH 7.2 for peptide **33**. At pH 7.2, the peptide **33** upon heating a solution results in only a linear decrease in ellipticity rather than the cooperative transition which is characteristic of triple-helix denaturation. This

indicates that peptide **33** H₂N-Phe(Pro-Hyp-Aib)₆ is present in a single helical structure rather than triple helix structure at physiological pH. Figure 14C shows the first derivative curve obtained from the sigmoidal fit of data of Figure 14B for peptide **33** which is a straight line, this confirmed that peptide **33** is in single helix chain rather than triplex in water at pH 7.2

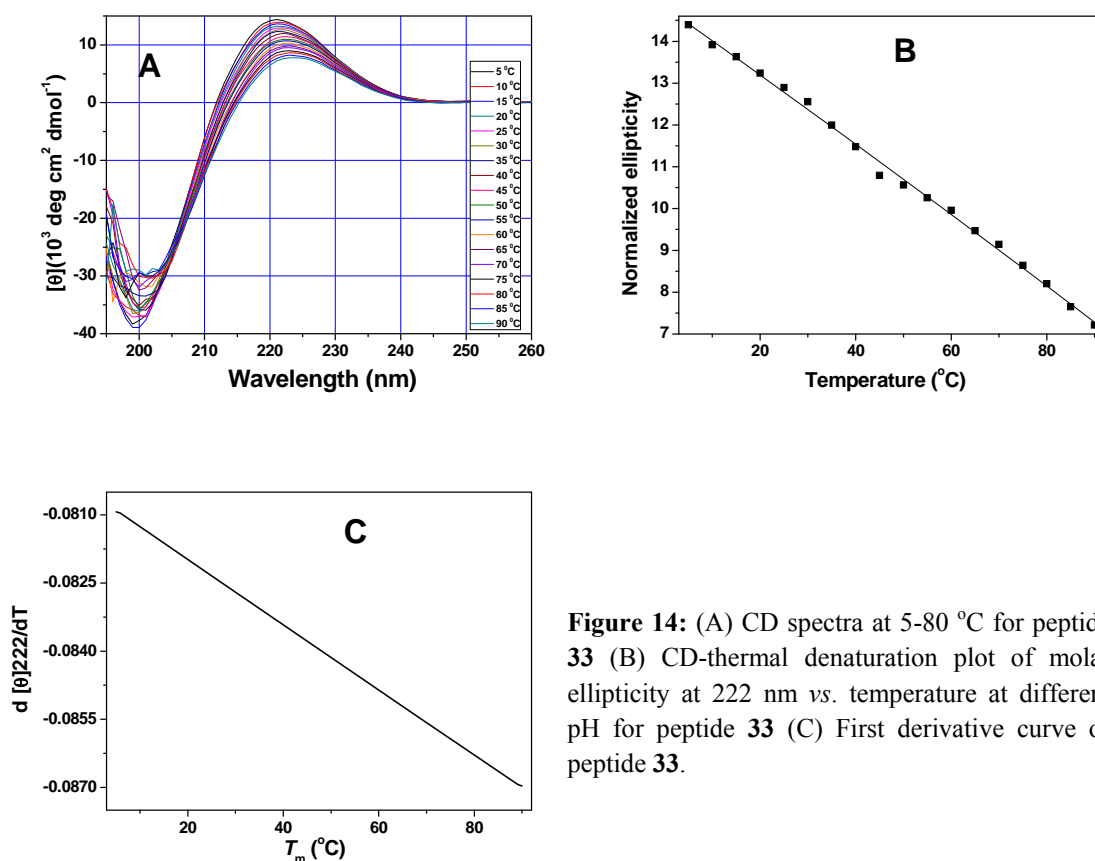


Figure 14: (A) CD spectra at 5-80 °C for peptide **33** (B) CD-thermal denaturation plot of molar ellipticity at 222 nm vs. temperature at different pH for peptide **33** (C) First derivative curve of peptide **33**.

Table 3 shows T_m values of peptides **30** to **33** at physiological pH conditions. From the data it is clear that, replacement of glycine by Aib does not form triple helix structure and remains in a single chain conformation in water.

Table 3: T_m values for peptide **30-33** in buffer at physiological pH

Peptides	T_m °C at pH 7.2
AcPhe(Pro-Hyp-Gly) ₆ 30	27
AcPhe(Pro-Hyp-Aib) ₆ 31	nd
H ₂ N-Phe(Pro-Hyp-Gly) ₆ 32	23
H ₂ N-Phe(Pro-Hyp-Aib) ₆ 33	nd

nd = triplex not detected, T_m values are (± 0.5 °C)

1.4.5 Effect of ethylene glycol on the stability of triple-helices

Intrachain peptide (amide) hydrogen bonding makes a major contribution to protein structure and stability and is more effective in the absence of accessible competing water. The presence of water molecules close-by causes the peptide hydrogen bonds to lengthen.¹⁹ Water molecules bridge the carbonyl oxygen atoms and amide protons of different peptide bonds to catalyze the formation and reversal of peptide hydrogen bonds as well as formation long-lived linkages stabilizing protein-ligand and protein-protein interfaces.²⁰ The internal molecular motions in proteins, necessary for biological activity, are very dependent on the degree of plasticizing, which is determined by the level of hydration.²¹ Thus internal water enables the folding of proteins when expelled from the hydrophobic central core, and squeezed out by cooperative protein chain interactions.²² If the stability of the protein structure is largely dependent on hydrogen bonding, then such proteins form a more stable structure in polyols compared to aqueous buffer condition. Ethylene glycol stabilizes helical structure and therefore can be very useful in amplifying and detecting very weak triple-helical propensities.²³ In order to evaluate the stability effect of ethylene glycol, CD spectroscopic studies were carried out for peptides **30-33**.

1.4.6 CD spectroscopy study of peptides 30-33 in buffer: ethylene glycol (1:1)

1.4.6a Concentration dependent CD spectroscopy of peptides 30-33: Figure 15 shows the CD spectra of peptides **30-33** recorded in water: EG (50:50 v/v).

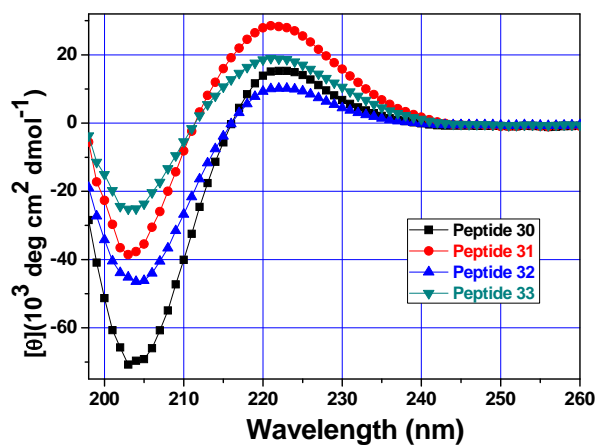


Figure 15: CD spectra of peptide **30-33** recorded in Water: EG (50:50, v/v).

The CD spectra of peptides **30-33** resemble that of native collagen, the positive maxima appearing at 221-223 nm and negative minima at 203 nm respectively. The $R_{p/n}$ values are in the range of triple-helical conformation and shown in Table 4.

Table 4: $R_{p/n}$ values of peptides **30-33** in 50% water and ethylene glycol.

Peptides	+ve band (nm)	-ve band (nm)	$R_{p/n}$
AcPhe(Pro-Hyp-Gly) ₆ 30	222	203	0.216
AcPhe(Pro-Hyp-Aib) ₆ 31	221	203	0.739
H ₂ NPhe(Pro-Hyp-Gly) ₆ 32	223	202	0.176
H ₂ NPhe(Pro-Hyp-Aib) ₆ 33	222	203	0.423

In order to determine the relative triple-helical strength of peptides **30-33** CD-thermal denaturation study was carried out in water: EG (50:50, v/v).

1.4.6b CD thermal denaturation study of peptide 30 AcPhe(Pro-Hyp-Gly)₆: Figure 16A shows the CD spectra obtained by heating the solution from 5 to 80 °C in water: EG (50:50, v/v).

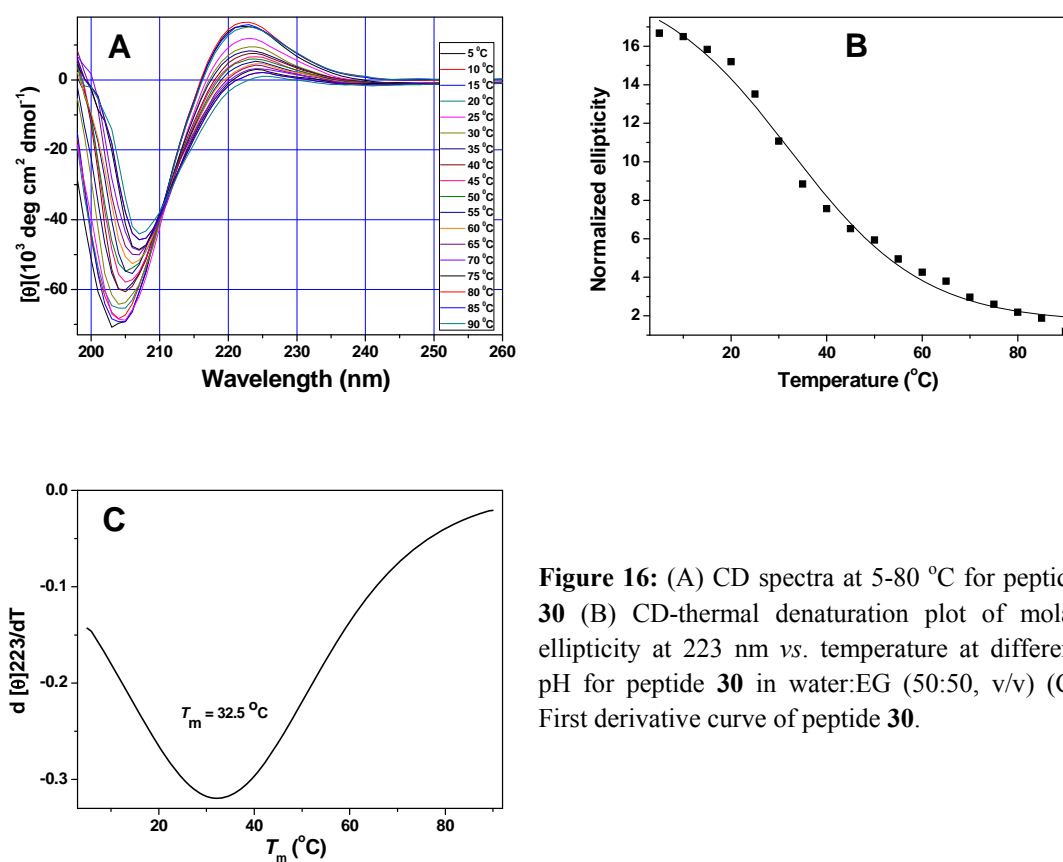


Figure 16: (A) CD spectra at 5-80 °C for peptide **30** (B) CD-thermal denaturation plot of molar ellipticity at 223 nm vs. temperature at different pH for peptide **30** in water:EG (50:50, v/v) (C) First derivative curve of peptide **30**.

The isobestic point seen at 212 nm is indicative of the conversion from one conformation to another conformation. Figure 16B shows the CD-thermal denaturation plot of molar ellipticity at 222 nm *versus* temperature in water: EG (50:50, v/v) for peptide **30**. Figure 16C shows the first derivative curve obtained from the sigmoidal fit of data of Figure 16B and the T_m values for peptide **30** obtained from minima in plot of Figure 16C. In water: EG (50:50, v/v), peptide **30** shows cooperative melting transition with increasing temperature. This indicates that peptide **30** is associated in a triple helical structure in water: EG (50:50, v/v) with $T_m = 32.5$ °C ($\Delta T_m = + 5.5$ °C compared with buffer system).

1.4.6c CD thermal denaturation study of peptide 31 AcPhe(Pro-Hyp-Aib)₆: CD spectra obtained by heating the solution from 5 to 80 °C in water:EG (50:50,v/v) is shown in Figure 17A.

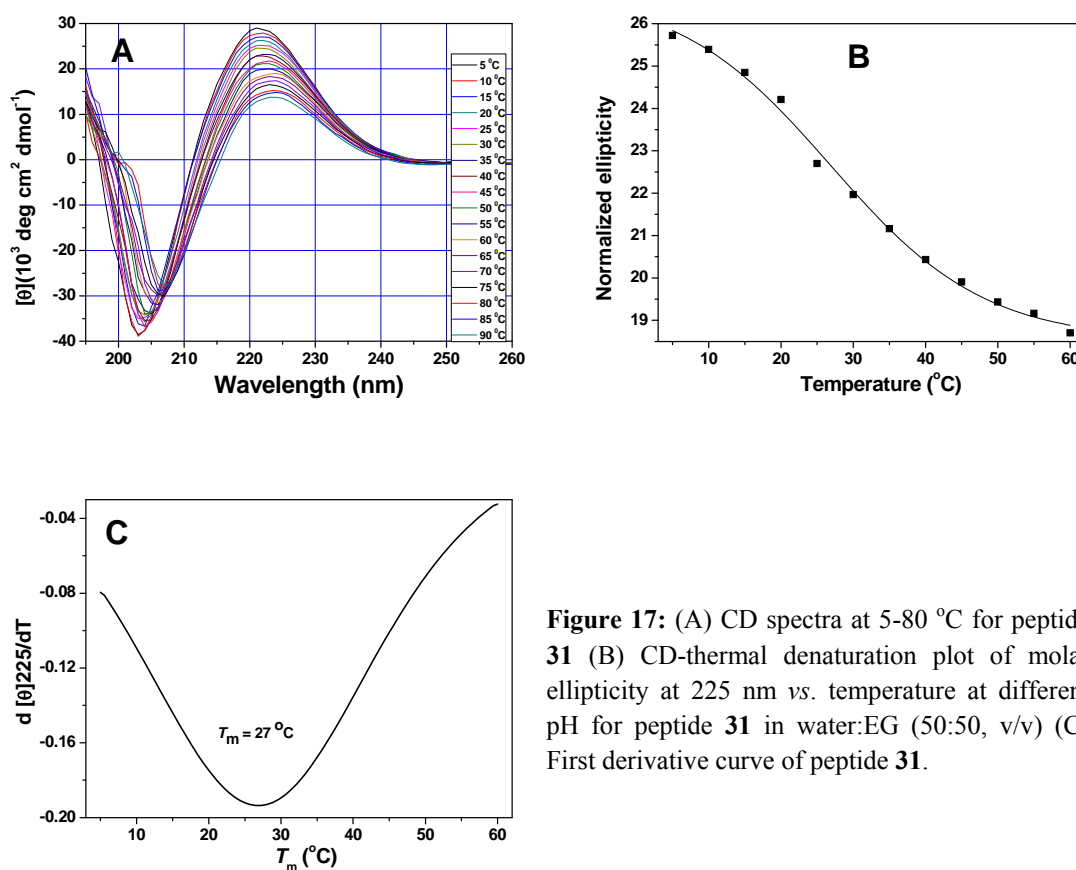


Figure 17: (A) CD spectra at 5-80 °C for peptide **31** (B) CD-thermal denaturation plot of molar ellipticity at 225 nm vs. temperature at different pH for peptide **31** in water:EG (50:50, v/v) (C) First derivative curve of peptide **31**.

The isobestic point seen at 207 nm is indicative of the conversion from one conformation to another conformation. Figure 17B shows the CD-thermal denaturation

plot of molar ellipticity at 225 nm *versus* temperature in Water: EG (50:50, v/v) for peptide **31**. Figure 17C shows the first derivative curve obtained from the sigmoidal fit of data of Figure 17B and the T_m values for peptide **31** obtained from minima in plot of Figure 17C. In water: EG (50:50, v/v), peptide **31** shows cooperative melting transition with increasing temperature. This indicates that peptide **31** is associated in a triple helical structure in water: EG (50:50, v/v) with $T_m = 27$ °C, while peptide **31** does not form triplex in 100% buffer.

1.4.6d CD thermal denaturation study of peptide **32** $\text{H}_2\text{N-Phe(Pro-Hyp-Gly)}_6$:

Figure 18A shows the CD spectra obtained by heating the solution from 5 to 70 °C in water: EG (50:50, v/v). The isobestic point seen at 209 nm is indicative of the conversion from one conformation to another conformation.

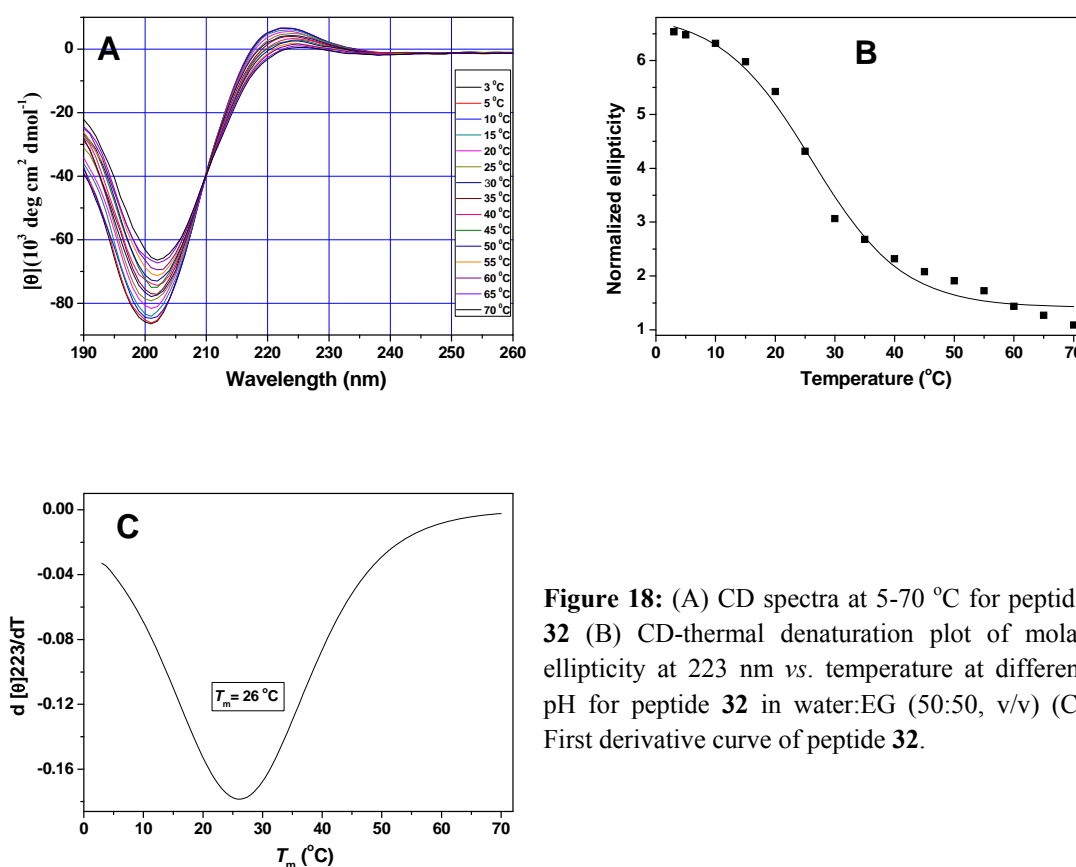


Figure 18: (A) CD spectra at 5-70 °C for peptide **32** (B) CD-thermal denaturation plot of molar ellipticity at 223 nm vs. temperature at different pH for peptide **32** in water:EG (50:50, v/v) (C) First derivative curve of peptide **32**.

Figure 18B shows the CD-thermal denaturation plot of molar ellipticity at 223 nm *versus* temperature in water: EG (50:50, v/v) for peptide **32**. Figure 18C shows the first derivative curve obtained from the sigmoidal fit of data of Figure 18B and the T_m

values for peptide **32** obtained from minima in plot of Figure 18C. In water:EG (50:50,v/v), peptide **32** shows cooperative melting transition with increasing temperature. This indicates that peptide **32** is associated in a triple helical structure in water:EG (50:50,v/v) with $T_m = 27\text{ }^\circ\text{C}$ ($\Delta T_m = 4\text{ }^\circ\text{C}$ compared with buffer system).

1.4.6e CD thermal denaturation study of peptide **33** $\text{H}_2\text{N-Phe(Pro-Hyp-Aib)}_6$:

The CD spectra obtained by heating the solution from 5 to 70 $^\circ\text{C}$ in water:EG (50:50,v/v) is shown in Figure 19A. The isobestic point seen at 209 nm is indicative of the conversion of one species to another species.

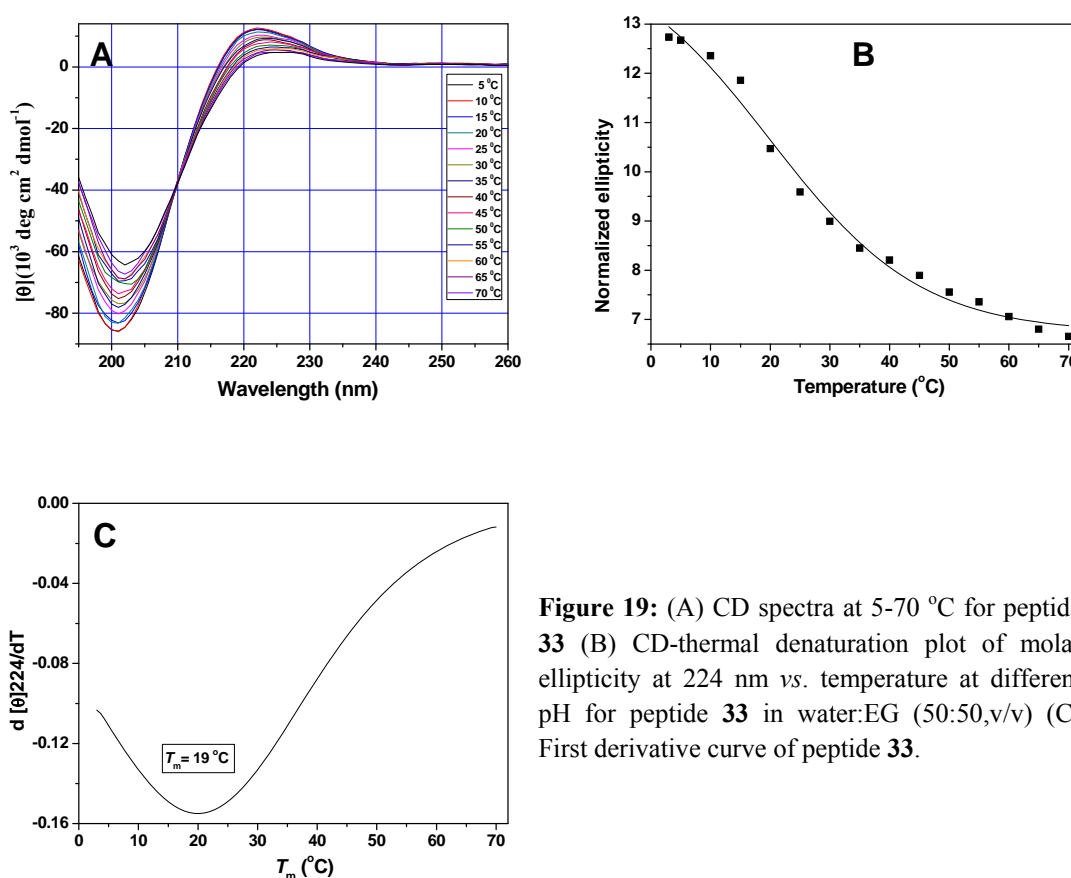


Figure 19: (A) CD spectra at 5-70 $^\circ\text{C}$ for peptide **33** (B) CD-thermal denaturation plot of molar ellipticity at 224 nm vs. temperature at different pH for peptide **33** in water:EG (50:50,v/v) (C) First derivative curve of peptide **33**.

Figure 19B shows the CD-thermal denaturation plot of molar ellipticity at 224 nm *versus* temperature in water: EG (50:50, v/v) for peptide **33**. Figure 19C shows the first derivative curve obtained from the sigmoidal fit of data of Figure 19B and the T_m values for peptide **33** obtained from minima in plot of Figure 19C. In water: EG (50:50, v/v), peptide **33** shows cooperative melting transition with increasing temperature. This indicates that peptide **33** is associated in a triple helical structure in

water: EG (50:50, v/v) with $T_m = 19$ °C, while peptide **33** does not form triplex in 100% buffer.

Table 5 shows T_m values of peptides **30-33** in water: EG (50:50, v/v). From CD thermal denaturation studies it is clear that, replacement of glycine by Aib forms weak triple helix in water: EG (50:50, v/v).

Table 5: T_m values of peptides **30-33** in water: EG (50:50, v/v)

Peptide	T_m °C
AcPhe(Pro-Hyp-Gly) ₆ 30	32.5
AcPhe(Pro-Hyp-Aib) ₆ 31	27
H ₂ N-Phe(Pro-Hyp-Gly) ₆ 32	26
H ₂ N-Phe(Pro-Hyp-Aib) ₆ 33	19

T_m values are (± 0.5 °C)

1.4.7 CD spectroscopy study of peptides 30-33 in 100% ethylene glycol

1.4.7a Concentration dependent CD spectroscopy of peptides 30-33: The CD spectra of peptides **30-33** recorded in ethylene glycol is shown in Figure 20 and resemble that of native collagen. The positive maxima appearing at 221-223 nm and negative minima at 203 nm. The $R_{p/n}$ values are in the range of triple-helical conformation and shown in Table 6.

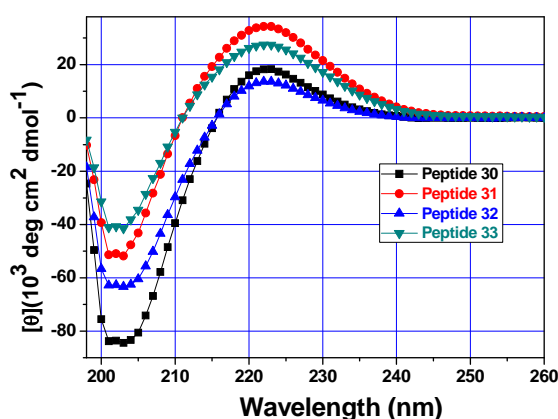


Figure 20: CD spectra of peptide **30-33** recorded in 100% ethylene glycol.

Table 6: $R_{p/n}$ values of peptides **30-33** in 100% ethylene glycol.

Peptides	+ve band (nm)	-ve band (nm)	$R_{p/n}$
AcPhe(Pro-Hyp-Gly) ₆ 30	223	203	0.236
AcPhe(Pro-Hyp-Aib) ₆ 31	222	203	0.668
H ₂ NPhe(Pro-Hyp-Gly) ₆ 32	222	202	0.156
H ₂ NPhe(Pro-Hyp-Aib) ₆ 33	223	202	0.359

In order to determine the relative triple-helical strength of peptides **30-33** CD-thermal denaturation study was carried out in 100% ethylene glycol.

1.4.7b CD thermal denaturation study of peptide 30 AcPhe(Pro-Hyp-Gly)₆: To evaluate the effect of ethylene glycol, the thermal denaturation studies were carried in 100% ethylene glycol.

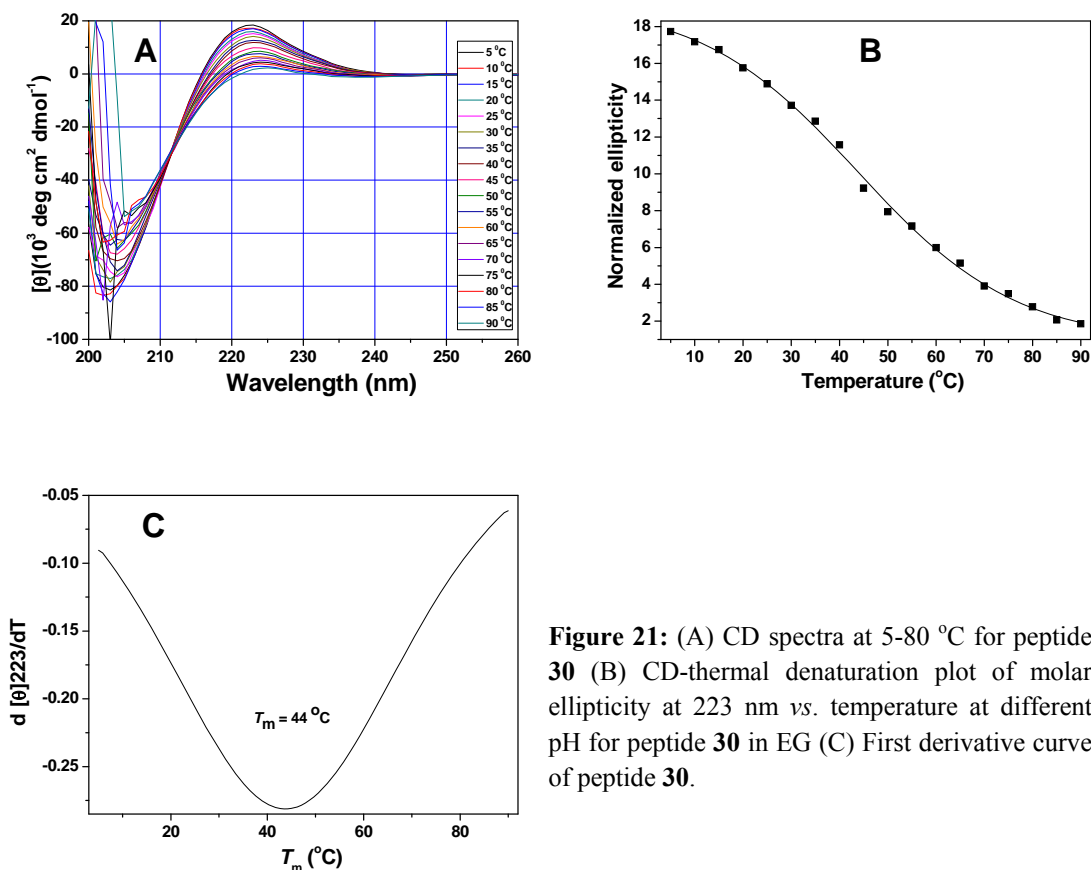


Figure 21: (A) CD spectra at 5-80 °C for peptide **30** (B) CD-thermal denaturation plot of molar ellipticity at 223 nm *versus* temperature at different pH for peptide **30** in EG (C) First derivative curve of peptide **30**.

Figure 21A shows the CD spectra obtained by heating the solution from 5 to 80 °C in ethylene glycol with isobestic point at 213 nm. Figure 21B shows the CD-thermal denaturation plot of molar ellipticity at 223 nm *versus* temperature in ethylene glycol for peptide **30**. Figure 21C shows the first derivative curve obtained from the

sigmoidal fit of data of Figure 21B and the T_m values for peptide **30** obtained from minima in plot of Figure 21C. In ethylene glycol, peptide **30** shows cooperative melting transition with increasing temperature. This indicates that peptide **30** is associated in a triple helical structure with $T_m = 44$ °C ($\Delta T_m = +17$ °C compared with buffer system).

1.4.7c CD thermal denaturation study of peptide 31 AcPhe(Pro-Hyp-Aib)₆: Figure 22A shows CD spectra obtained by heating the solution from 5 to 80 °C in ethylene glycol with isobestic point at 213 nm. Figure 22B shows the CD-thermal denaturation plot of molar ellipticity at 223 nm *versus* temperature in ethylene glycol for peptide **31**. Figure 22C shows the first derivative curve obtained from the sigmoidal fit of data of Figure 22B and the T_m values for peptide **31** obtained from minima in plot of Figure 22C. In ethylene glycol, peptide **31** shows cooperative melting transition with increasing temperature, which indicates that peptide **31** is associated in a triple helical structure with $T_m = 31$ °C.

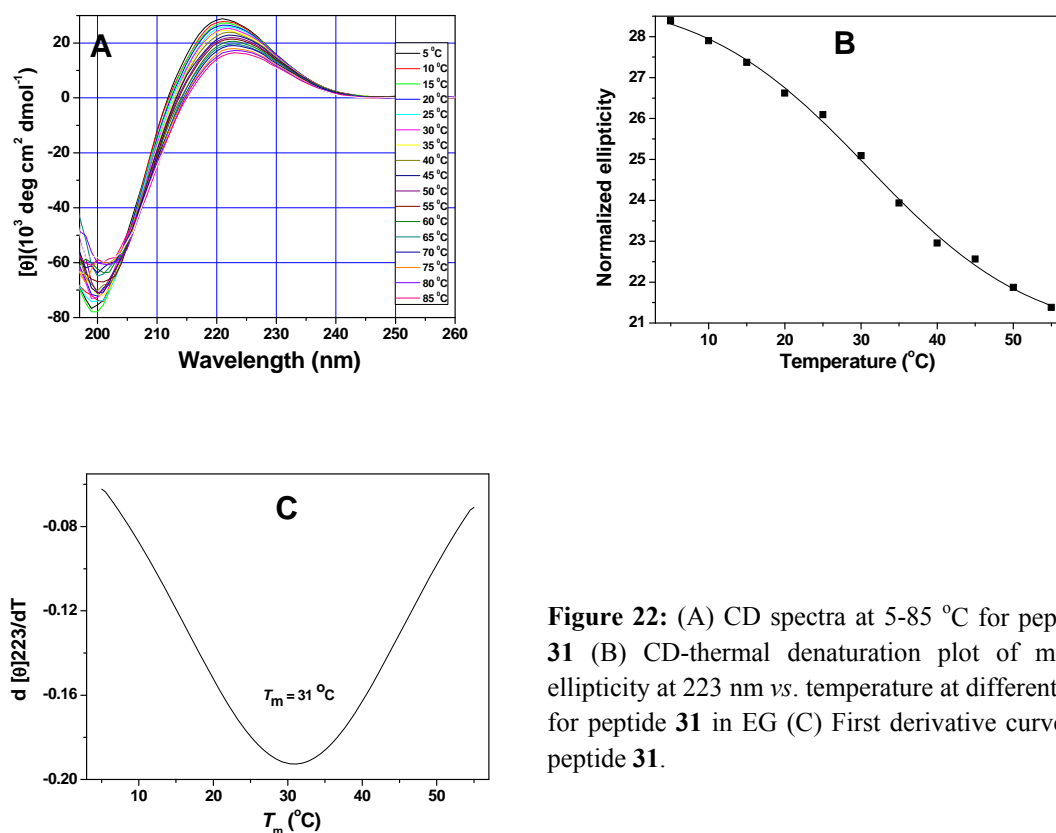


Figure 22: (A) CD spectra at 5-85 °C for peptide **31** (B) CD-thermal denaturation plot of molar ellipticity at 223 nm vs. temperature at different pH for peptide **31** in EG (C) First derivative curve of peptide **31**.

1.4.7c CD thermal denaturation study of peptide 32 H₂N-Phe(Pro-Hyp-Gly)₆: The CD spectra obtained by heating the solution from 5 to 70 °C in ethylene glycol with isobestic point at 212 nm is shown in Figure 23A. Figure 23B shows the CD-thermal denaturation plot of molar ellipticity at 223 nm *versus* temperature in ethylene glycol for peptide 32. Figure 23C shows the first derivative curve obtained from the sigmoidal fit of data of Figure 23B and the T_m values for peptide 32 obtained from minima in plot of Figure 23C. In ethylene glycol, peptide 32 shows cooperative melting transition with increasing temperature, which indicates that peptide 32 is associated in a triple helical structure with $T_m = 31.5$ °C.

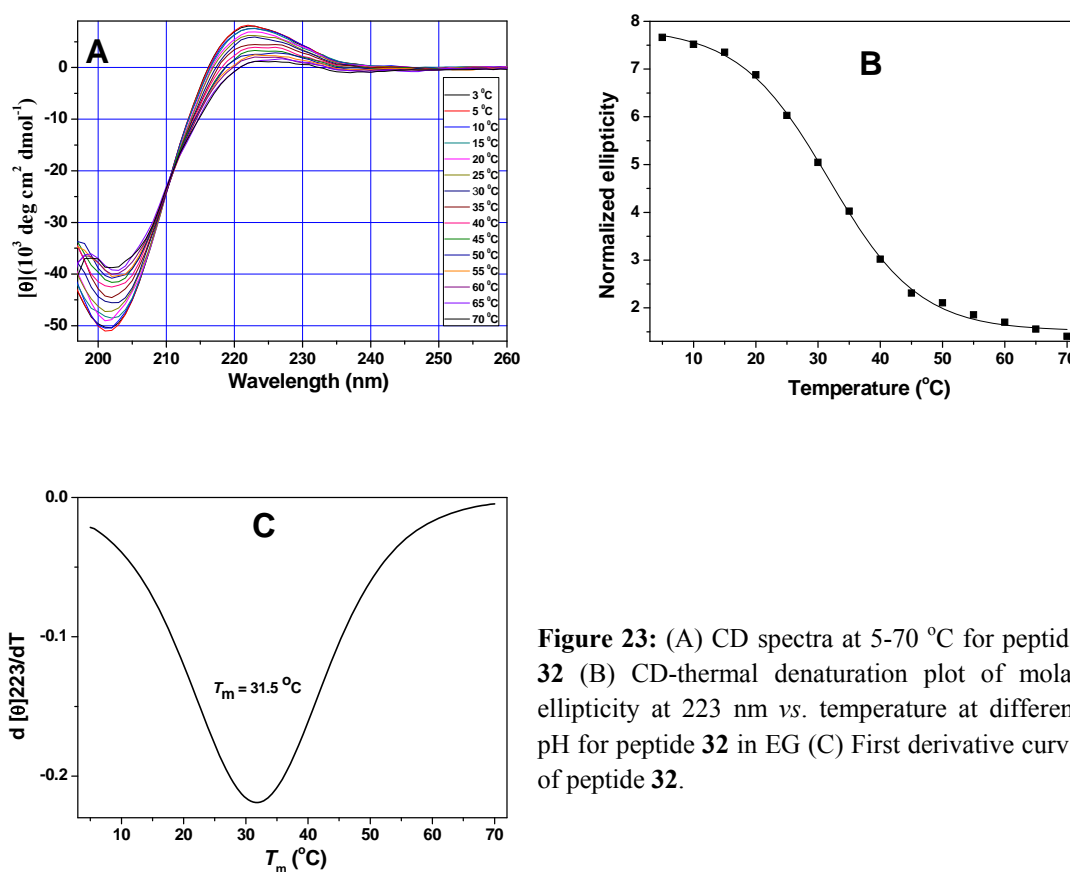


Figure 23: (A) CD spectra at 5-70 °C for peptide 32 (B) CD-thermal denaturation plot of molar ellipticity at 223 nm vs. temperature at different pH for peptide 32 in EG (C) First derivative curve of peptide 32.

1.4.7d CD thermal denaturation study of peptide 33 H₂N-Phe(Pro-Hyp-Aib)₆: The CD spectra obtained by heating the solution from 5 to 80 °C in ethylene glycol with isobestic point at 212 nm is shown in Figure 24A. The CD-thermal denaturation plot of molar ellipticity at 223 nm *versus* temperature in ethylene glycol for peptide 33 is shown in Figure 24B. Figure 24C shows the first derivative curve obtained from the

sigmoidal fit of data of Figure 24B and the T_m values for peptide **33** obtained from minima in plot of Figure 24C. In ethylene glycol, peptide **33** shows cooperative melting transition with increasing temperature, which indicates that peptide **33** is associated in a triple helical structure with $T_m = 22$ °C.

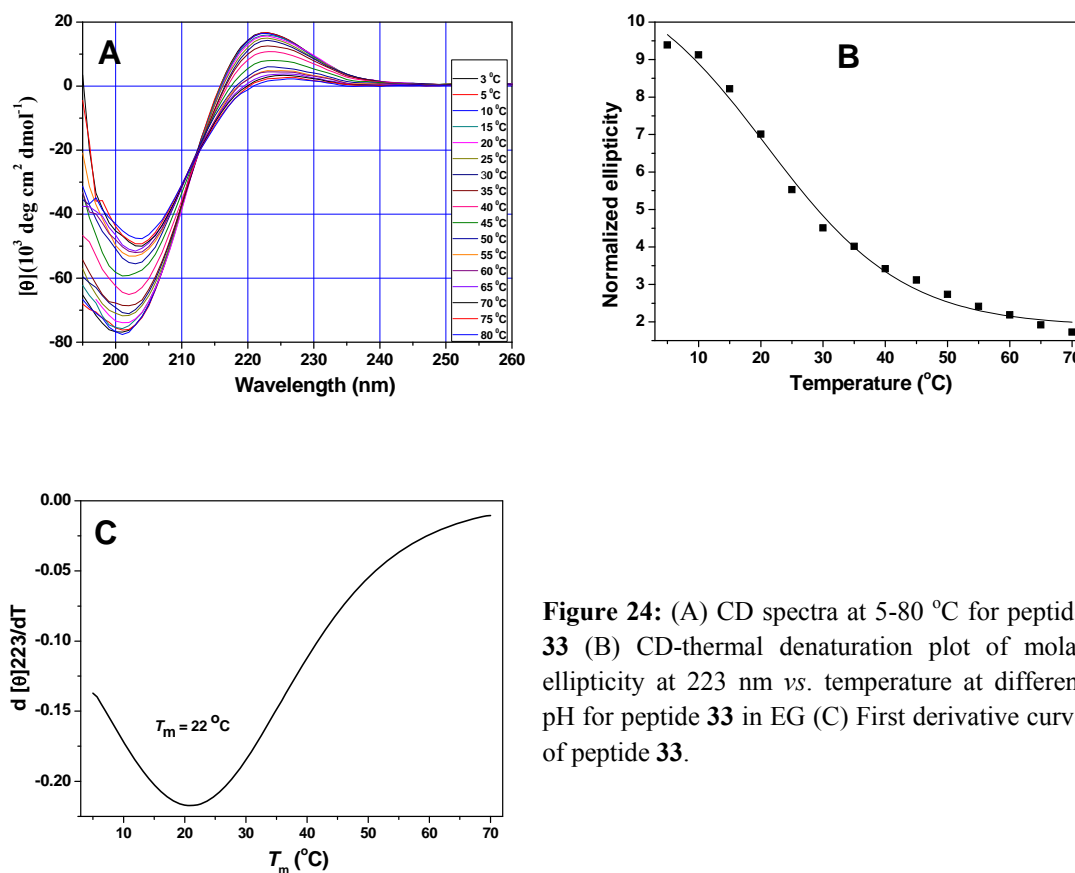


Figure 24: (A) CD spectra at 5-80 °C for peptide **33** (B) CD-thermal denaturation plot of molar ellipticity at 223 nm vs. temperature at different pH for peptide **33** in EG (C) First derivative curve of peptide **33**.

Table 7 shows comparative CD- T_m values of peptides **30-33** in buffer at physiological pH, water: EG (1:1) and 100% ethylene glycol

Table 7: T_m values of peptides **30-33** in different system.

Peptide	Buffer (pH 7.2)	EG:water (ΔT_m)	Ethylene glycol (ΔT_m)
AcPhe(Pro-Hyp-Gly) ₆ 30	27	32.5	44
AcPhe(Pro-Hyp-Aib) ₆ 31	nd	27	31
H ₂ N-Phe(Pro-Hyp-Gly) ₆ 32	23	26 (-6.5)	31.5 (-12.5)
H ₂ N-Phe(Pro-Hyp-Aib) ₆ 33	nd	19 (-8)	22 (-9)

nd = triplex not detected, T_m values are (± 0.5 °C), bracket values are $\Delta T_m = T_m$ of (N-capped peptides – N-nonacetylated peptide)

From above thermal denaturation data it is clear that the replacement of glycine by Aib in collagen model peptides **31** and **33** leads to formation of weak triple helix structure in ethylene glycol or EG:H₂O systems and such structure is absent in water.

1.5 Discussion

The control collagen glycine substituted peptides (**30**, **32**) at physiological pH form polyproline-II like structure as indicated by a weak positive maxima at 222-224 nm and strong negative band at 201-204 nm in CD spectra respectively. The peak positions are nearly independent of the concentration and most importantly, all the spectral traces pass through an isoelliptic point at 213-217 nm for peptides **30** and **32** (Figure 7 and 9). Smooth and progressive increase in the $R_{p/n}$ values upto 300 μ M concentration and then rapidly moves towards saturation for peptides **30** and **32** with concentration also suggests a two state equilibrium. This suggests that the peptides **30** and **32** form triple helix structure as seen by the minimum acceptable criterion.²⁴ The observed sigmoidal transition in variable temperature CD measurements for peptides **30** and **32** provides additional evidence for the two state triple helix \leftrightarrow single helix transitions. The T_m values obtained at physiological pH for peptides **30** and **32** are 27 °C and 23 °C respectively.

The collagen Aib substituted peptides (**31**, **33**) at physiological pH form PPII like structure as indicated by weak positive maxima at 222-223 nm and strong negative band at 201-203 nm in CD spectra respectively. The peak positions are nearly independent of the concentration and all the spectral traces pass through the isoelliptic point at 213-214 nm for peptides **31** and **33** (Figure 8 and 10). These CD spectral properties indicate that within the concentration range measured i.e., 50-500 μ M peptides exists in the same conformational state. The presence of higher CD ellipticity of maxima and lower ellipticity of minima for peptide **31** and **33** suggest that these peptides adopt more extended structure than peptide **30** and **32**.¹⁷ The $R_{p/n}$ value increases linearly with increase in peptide concentration suggesting that Aib peptides **31** and **33** exist in a single helical structure at physiological pH. The observed linear decrease of molar ellipticity in the temperature dependent CD measurements for peptides **31** and **33** provides additional evidence that these peptides exist in single helical structure rather than triple helix structure at physiological pH. Thus peptides **31**

and **33** are present in only one kind of conformation in water at pH 7.2. Comparable observation was made by Raines *et.al.*²⁵ for collagen peptides with D-alanine/D-serine substitution for glycine.

In water: EG (1:1), the peptides **30** and **32** form more stable triple helix ($T_m = 32.5$ °C and 26 °C respectively) than in buffer system. The most likely mechanism for stabilization of a collagen-like triple-helical structure involves the correlation of hydrogen-bonding strengths of the solvents with the increase in ordered structure. The weaker hydrogen bonding solvents may provide a more favorable environment in which the chains associate through interchain hydrogen bonds required for stabilization of a triple-helical conformation in collagen-like peptides.²⁶

The progressive increase in the $R_{p/n}$ values with concentration upto 300 μ M and followed by saturation seen for peptides **31** and **33** suggests a two state equilibrium. The observed sigmoidal transition in H₂O: EG system in the temperature dependent CD measurements for peptides **31** and **33** provides additional evidence for the two state triple helix \leftrightarrow single helix transition. This suggests that the peptides **31** and **33** form weak triple helix structures with $T_m = 27$ °C and 19 °C respectively.

In ethylene glycol, the peptides **30** and **32** form more stable triple helix ($T_m = 44$ and 31.5 °C respectively) than in buffer system and EG: water (1:1).

The progressive increase in the $R_{p/n}$ values with concentration upto 300 μ M and beyond towards saturation suggests a two state equilibrium for peptides **31** and **33**. The observed sharp sigmoid transition in the temperature dependent CD experiments for peptides **31** and **33** gave further evidence for the two state triple helix \leftrightarrow single helix transitions. This suggests that peptides **31** and **33** form weak triple helix structures with T_m of 31 °C and 22 °C respectively. From literature²⁷ it is well known that ethylene glycol stabilizes triple helical structures through interchain hydrogen bonding and therefore can be very useful to amplify and detect very weak triple-helical propensities.

In this study the uncapped peptides (**30** and **31**) show lower triplex stability (T_m) than capped peptides (**32** and **33**). This observed decrease in stability of uncapped peptides is due to the presence of similar charged residues at termini in a parallel triplex that leads to repulsion and hence result in destabilization. The terminal effects may be nulled by end capping, which significantly stabilizes the triplex helix.²⁸

The data in this study indicates that α -aminoisobutyric acid (Aib) substitution in place of glycine greatly destabilizes the collagen triple helix in aqueous medium. These results suggest that the α -aminoisobutyric acid residues in collagen is not a surrogate for a glycine and support the notion that the main-chain torsion angles of a glycine residue in the native structure (especially, $\phi > 0^\circ$) are critical determinants for its beneficial substitution with a α -aminoisobutyric acid in a protein. This is supported by glycine to alanine substitution known to destabilize the collagen triple helix due to deleterious steric effects imposed by its side chain.²⁹ The lower stability of Aib substituted peptides may be due to a more extended structure adopted by peptides **31** and **33** in which the steric bulk of gemdimethyl groups increases the distance between carbonyl oxygen and NH of Aib residue affecting the interstrand hydrogen bonding between carbonyl oxygen and NH of Aib. In ethylene glycol, competitive hydrogen bonding with solvent is less, leading somewhat favourable interchain hydrogen bonds, resulting in weak triple helix structure of peptide **31** and **33**.

1.6 Conclusion

In conclusion, it is shown that the α -aminoisobutyric acid (Aib) substitution in place of glycine greatly destabilises the collagen triple helix in water. The Aib substituted peptides **31** and **33** remain in polyproline II conformation while the glycine substituted peptides form triple helix structure in buffer. The weak triple helix structure is detected in Aib substituted peptides in water: EG (1:1) and in 100% ethylene glycol. The results reinforce the significant role of glycine in collagen triplex structure. The terminal effect plays an important role in determining the stability of collagen triple helix and may be nulled by end capping, which leads to stabilization of the triplex helix significantly. More research needs to be done to verify these results in terms of theoretical studies and host-guest peptide systems to obtain better insight into this aspect. The results have a direct bearing on the current interest in collagen structure and mimetics. The properties of this analogue may have significant applications in the design of new collagen based biomaterials.³⁰

1.7 Experimental

1.7.1 Peptide synthesis

All peptides were synthesized manually in a sintered vessel equipped with a stopcock. The readily available Rink amide resin with loading value 0.5-0.6 mmol/g was used and standard Fmoc chemistry was employed. The resin bound Fmoc group was first deprotected with 20% piperidine in DMF and the coupling reactions were carried out using *in situ* active ester method, using HBTU as a coupling reagent and HOBT as a racemization suppresser and DIPEA as a catalyst. All the materials used were of peptide synthesis grade (Sigma-Aldrich) and was used without further purification. Analytical grade DMF was purchased from Merck (India) and was distilled over P₂O₅ under vacuum at 45°C, stored over 4Å molecular sieves for 2 days before using for peptide synthesis.

1.7.1a Resin functionalization

The resin (2',4'-Dimethoxyphenyl-Fmoc-aminomethyl)-phenoxy, catalog number 855047, 100-200 mesh) (100 mg) from Nova biochem was taken in sintered vessel (25 mL) and rinsed with 5 mL of dry DCM and filtered. The process was repeated 3 to 4 times and the resulting resin was kept for 2 h in DCM (10 mL) for swelling. The DCM was removed and rinsed 3 times with dry DMF and kept 2 h in dry DMF (10 mL) for swelling before the first coupling. The deprotection of 'Fmoc group' attached to the resin was done with 20% piperidine in DMF (3 x 5 mL) before proceeding for first amino acid coupling.

- The resin was washed and swollen in dry DCM for at least 2 h.
- Further washing and swelling with dry DMF for 2 h.
- Immediate coupling of 1st amino acid desired in C-terminus of peptide.

1.7.1b General method for solid phase peptide synthesis

All peptides were assembled on solid phase method by one amino acid per coupling. All amino acids were well dried over P₂O₅ in vacuum desiccator before coupling. Fmoc protecting group was used for main chain α -amino group. The *t*-Boc protection was used for side chain amine protection and was cleaved with 20% TFA in

DCM for final cleavage of peptide from the resin. The peptide obtained after cleavage was stirred in 95% TFA in DCM for 2 h for complete deprotection *t* Boc.

1.7.1c *Synthesis protocol for solid phase synthesis*

The resin was pre-swollen overnight and the following steps were performed for each cycle.

- Wash with DMF 4 x 5 mL.
- 20% piperidine in DMF 2 x 5 mL (15 min for each) for deprotection of Fmoc group.
- Wash with DMF 3 x 5 mL, MeOH 3 x 5 mL and DCM with 3 x 5 mL.
- Test for complete deprotection (chloronil test).
- Coupling reaction with amino acid, DIPEA, HOBT and HBTU (3 eq.) in DMF (1 mL).
- Repeat of the coupling reaction in NMP for better yield.
- Test for completion of coupling reaction (chloronil test).

This cycle was repeated for every amino acid.

1.7.1d *General procedure for peptide couplings on Rink Amide Resin*

Fmoc-Xxx-OH (3 eq), HBTU (3 eq) and HOBT (3 eq) dissolved in DMF/NMP followed by *i*Pr₂NEt (7-8 eq) were added to the amino-functionalized resin in DMF. The mixture was kept for 2 h and last 5 min bubbled with N₂ and washed with DMF (3x), MeOH (3x) and DCM (3x). The loading value for peptide synthesis is taken as 0.5~0.6.

1.7.1e *General procedure for Fmoc deprotection*

20% piperidine in DMF was added to the resin and the reaction mixture was kept for 15 min, drained and the piperidine treatment was repeated 3 times. Finally the resin was washed with DMF (3x), MeOH (3x) and DCM (3x).

1.7.1f *General procedure for acetylation*

Triethylamine (20 eq) and acetic anhydride (20 eq) were added to the resin in DMF (\approx 100 mM). The mixture was kept for 1 h followed by bubbled with N₂ for 5 min and washed with DMF (3x), MeOH (3x) and DCM (3x).

1.7.1g Preparation of resin with peptide for cleavage

After the last coupling/acetylation, the resin was washed sequentially with DMF (5 x 10 mL), DCM (5 x 10 mL), toluene (5 x 10 mL) and finally with methanol (5 x 10 mL) and dried with nitrogen gas for 3 min. The resin sintered flask was dried in a vacuum desiccator over P₂O₅.

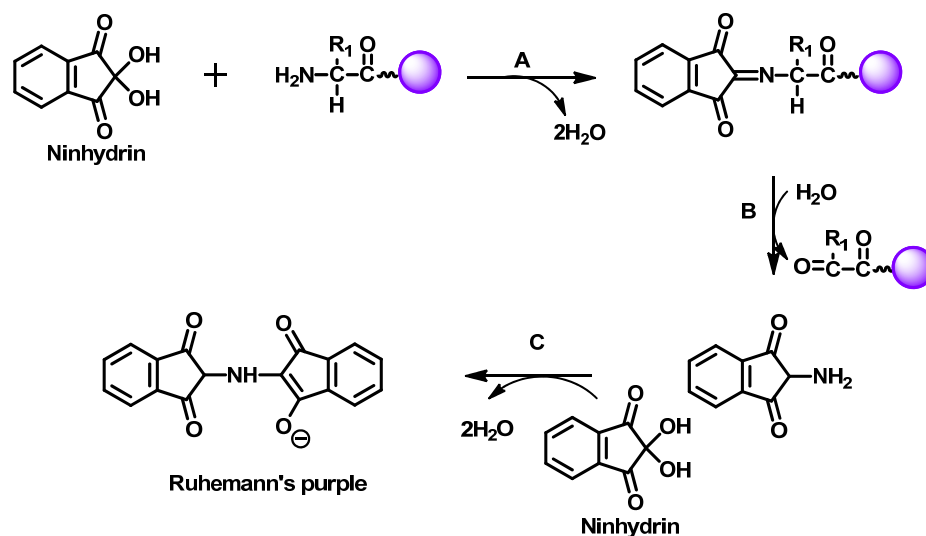
1.7.1h General procedure for cleavage of peptides from the solid support

The dry peptide-resin (20 mg) was taken in round-bottomed flask to which of 20% TFA in DCM (10 mL) and Triisopropylsilane (as scavengers) (2-3 drops) were added. The resulting mixture was kept for 2 h by gentle shaking. The mixture was filtered through a sintered funnel and the resin was washed with 3 x 5 mL of above solution. The filtrate was collected in pear shape round-bottom flask and evaporated under reduced pressure. The resin was washed with MeOH (3 X 5 mL) and the washings were evaporated to dryness. The crude peptide obtained was stirring the peptide solution with 95% TFA in DCM (10 mL) for 2 h. The TFA: DCM mixture was removed under reduced pressure. The residue obtained was dissolved in anhydrous methanol (0.4 mL) and to it anhydrous diethyl ether (4 x 1.5 mL) was added. The off-white precipitate obtained was centrifuged. The precipitation procedure was repeated twice to obtain peptide as a colourless powder.

The important tests for primary and secondary amine detection are Kaiser's and Chloranil tests are discussed below.

1.7.2 Kaiser's test

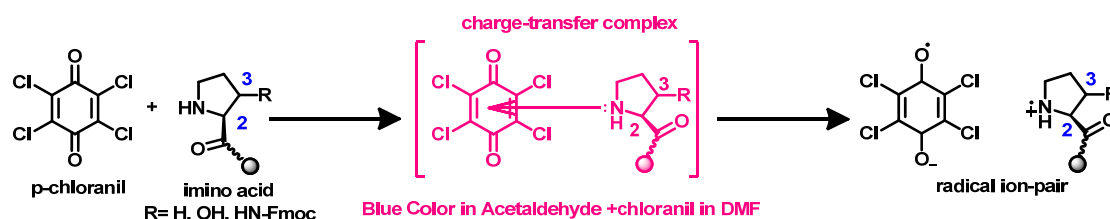
Kaiser's test is the most widely used qualitative test for the presence or absence of free amino groups (deprotection/coupling). It is used in monitoring the *t*-Boc deprotection and amide bond (peptide bond) formation steps in solid phase peptide synthesis. The *t*-Boc-deprotection step leads to a positive Kaiser's test, where in the resin beads as well as the solution are blue in color (Rheumann's purple). On the other hand, upon completion of the coupling reaction, the Kaiser's test is negative, the resin beads remaining colourless (Scheme 2).



Scheme 2: (A) Reaction of ninhydrin (trioxohydrindene hydrate) with the amino group of a bound residue generates the Schiff's base. (B) Hydrolysis generates the aldehyde and another amine. (C) Amine reacts with a second molecule of ninhydrin to give an equilibrium mixture of the anion depicted and its tetraoxo form with a maximum of absorbance at 570 nm.

1.7.3 Chloranil test

This sensitive test has been developed for reliable detection of secondary amino groups³¹ but it will also detect primary amines. Only the beads will be colored in case of a positive test. A few beads of resin were taken in a small test tube and were washed with methanol. To this 5 drop of 2% acetaldehyde in DMF and 2% chloranil in DMF were added. After a short mixing, the mixture was left at room temperature for 5 min and the beads inspected. If dark blue to green colour obtained by beads then the test was positive, if colourless to yellowish beads colour obtained by beads then the test was negative. The probable mechanism of the process is through charge-transfer complex between chloranil molecule and free amine of the imino acids.



Scheme 3: Proposed mechanism for the reaction between free amine group of peptide on solid phase and chloranil to produce charge-transfer complex.

1.7.4 High performance liquid chromatography (HPLC)

Peptides (**30-33**) were purified by reverse phase-HPLC on Waters 600 equipped with 2998-Photodiode array detector (PDA). Semi-preparative RP-C18 columns (250 x 10 mm, 10 μ m) of Alteck-Altima make were used for peptides. The solvent system comprised of MeCN:Water (5:95) with 0.1% TFA for solution A and for solution B MeCN:Water (50:50), 0.1% TFA. A gradient of 0-100% at a flow rate of 3 ml/min was used for semi preparative HPLC to elute the peptide and the eluant was monitored at 220 nm. The peak corresponding to the peptide was collected and the fractions were freeze-dried. Subsequently these peptides were concentrated by using speed vacuum. The purity of the final peptides were further analyzed on the Merck LiChrospher 100 RP-18 (250 x 4 mm, 5 μ m) column by using a gradient flow of 0 to 100% B in 20 min at a flow rate of 1.5 ml/min. The oven was heated to 50 °C to prevent triple helix formation for non acetylated peptides **13-21**. The spectra acquisition, analysis and processing was done on Waters Empower-2154 software. The absorbance of the eluant was monitored at its corresponding wavelength and the purity was obtained from the integrator output. The purities of the hence purified peptides were found to be more than 95%.

1.7.5 MALDI-TOF characterization

MALDI-TOF mass spectra were obtained on either Voyager-Elite instrument (PerSeptive Biosystems Inc., Farmingham, MA) equipped with delayed extraction or on Voyager-De-STR (Applied Biosystems) instrument. Sinapinic acid and α -cyano-4-hydroxycinnamic acid (CHCA) both were used as matrix for peptides of which CHCA was found to give satisfactory results. A saturated matrix solution was prepared with typical dilution solvent (50:50:0.1 Water:MeCN:TFA) and spotted on the metal plate along with the oligomers. The metal plate was loaded to the instrument and the analyte ions are then accelerated by an applied high voltage (15-25 kV) in reflector mode, separated in a field-free flight tube and detected as an electrical signal at the end of the flight tube. HPLC purified peptides were characterized through this method and were observed to give good signal to noise ratio, mostly producing higher molecular ion signals.

1.7.6 Circular dichroism (CD) spectroscopy

CD spectrometric studies were carried out on JASCO J-715 spectropolarimeter using cylindrical, jacketed quartz cell (1 mm path length), which was connected to Julabo-UC-25 water circulator. CD spectra were recorded using a spectral bandwidth of 1.0 nm at 25 °C with a time constant of 1 s and a step resolution of 1 nm. All the spectra were corrected for respective buffer condition and are typically averaged over 5-10 scans. CD data are given as mean residual molar ellipticities $[\theta] 10^3 \text{ deg cm}^2 \text{ dmol}^{-1}$. The spectra are the result of 5-10 accumulations. A quartz cell with a path length of 1 mm was used with solutions containing approximately 0.2 ml (50-500 μM) peptide solutions. For the spectra in buffer the blank spectrum of the solution was subtracted. All samples were equilibrated for at least 10 h before measurement except some TFE experiment.

For CD thermal denaturation studies samples were annealed in water bath at 95 °C and slowly cooled to room temperature over a period of 6 h prior to spectroscopic analysis. These samples were then incubated at 4 °C for 12 h followed by additional half hour incubation in the instrument at the initial measurement temperature. The temperature was varied in steps of 5 °C and the spectra were recorded at each step. An equilibration period of 5 min was allowed at each temperature. Data processing and curve fitting was performed using MicroCal Origin 8.0 software. Ellipticity at specified wavelength for each temperature was plotted, normalized data was fitted to a sigmoid curve and the T_m (melting temperature) values are derived from the first derivative curve of the fit.

1.8 References

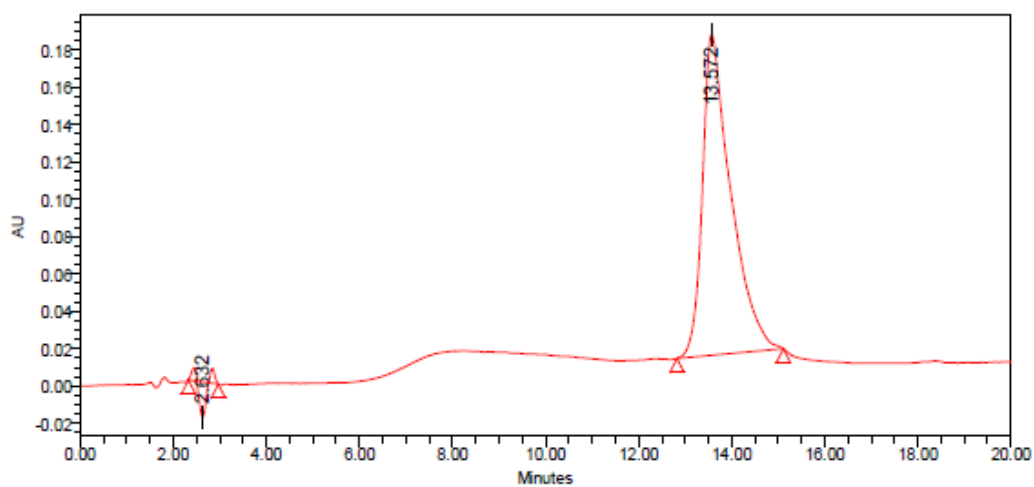
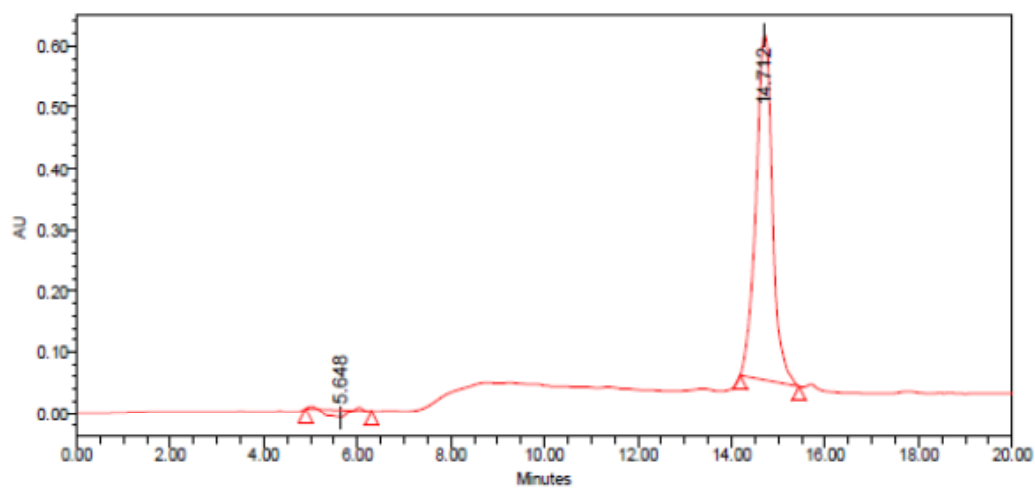
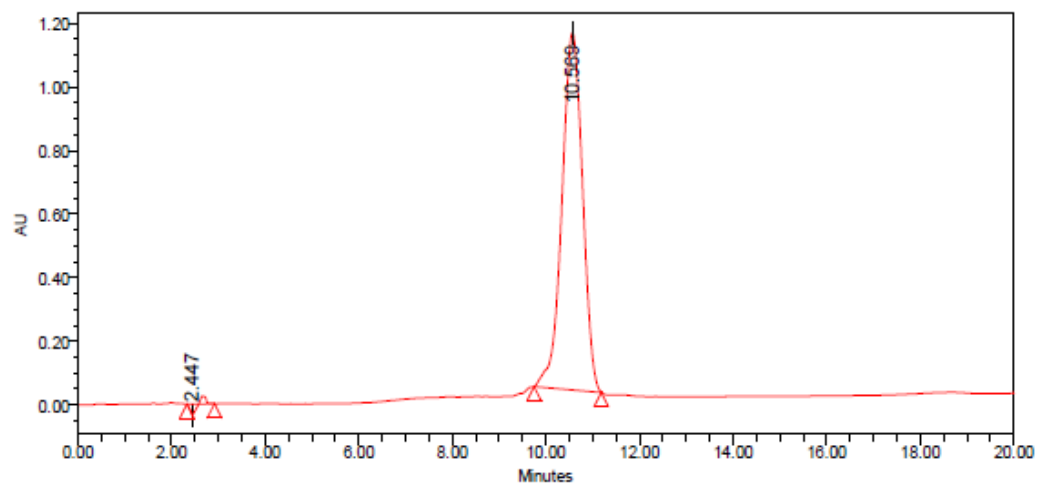
- 1 (a) Anil, B.; Song, B.; Tang, Y. and Raleigh, D.P. *J. Am. Chem. Soc.*, **2004**, *126*, 13194–13195. (b) Anil, B.; Craig-Schapiro, R. and Raleigh, D.P. *J. Am. Chem. Soc.*, **2006**, *128*, 3144–3145.
- 2 Tsai, M.; Xu, Y. and Dannenberg, J.J., *J. Am. Chem. Soc.*, **2005**, *127*, 14130–14131.
- 3 (a) Karle, I. L.; Balaram, P. *Biochemistry* **1990**, *29*, 6747. (b) Moretto, V.; Crisma, M.; Bonora, G. M.; Toniolo, C.; Balaram, H.; Balaram, P. *Macromolecules* **1989**, *22*, 2939. (c) Kumita, J. R.; Weston, C. J.; Choo-Smith, L.-P.; Wooley, G. A.; Smart, O. S. *Biochemistry* **2003**, *42*, 4492.
- 4 (a) Closse, A.; Huguenin, R. *Helv. Chim. Acta* **1974**, *57*, 533. (b) Flippen, J. L.; Karle, I. L. *Biopolymers* **1976**, *15*, 1081.

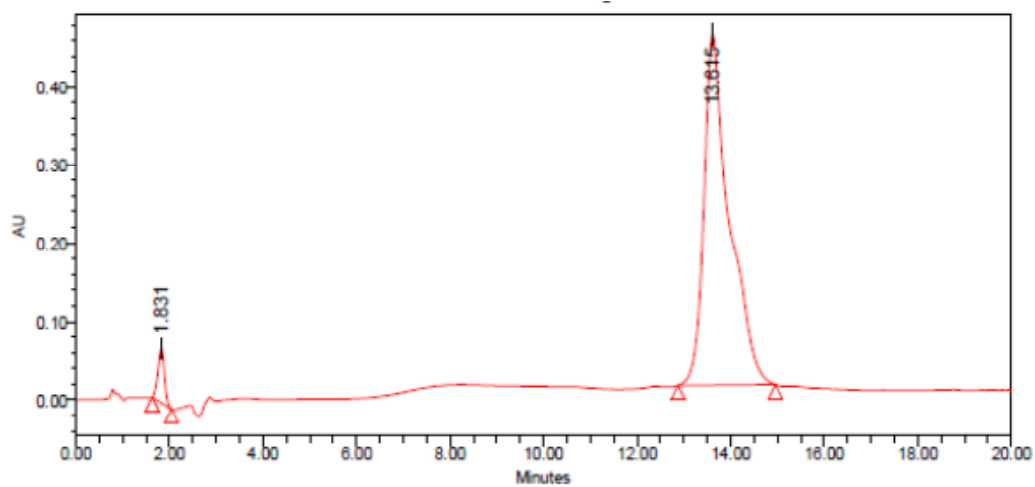
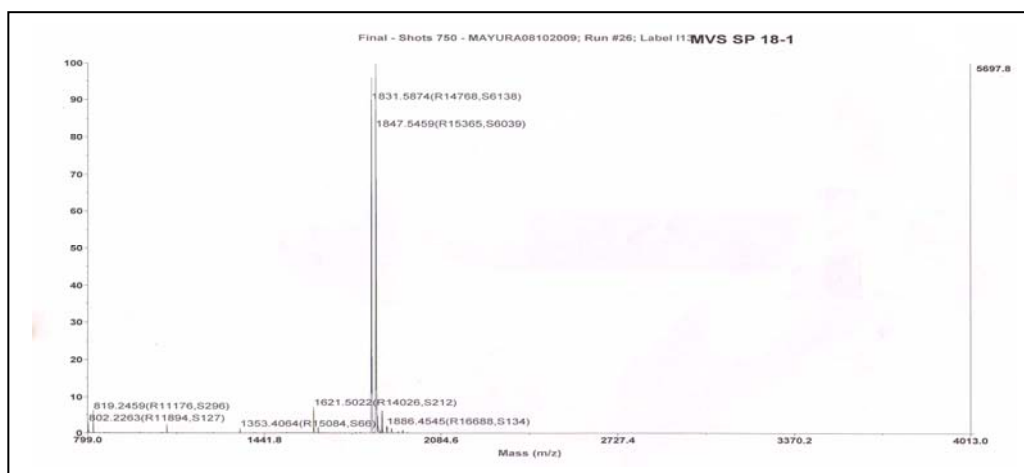
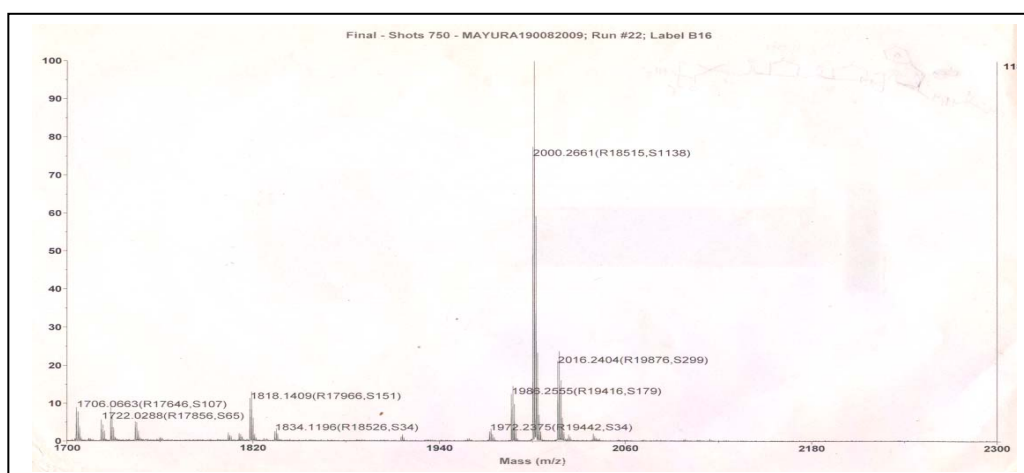
- 5 (a) Marshall, G. D.; Beunsen, D. D. In *Biomembrane Electrochemistry*; Blank, M., Vodanoy, I., Eds; Advances in Chemistry Series 235; American Chemical Society: Washington, DC, **1994**; p 259. (b) Pandey, R. C.; Meng, H.; Cook, J. C.; Rinehart, K. L. *J. Am. Chem. Soc.* **1977**, *99*, 5203.
- 6 Marshall, G. R.; Bosshard, H. E. *Circ. Res.* **1972**, *30/31* (Suppl. II.), 143.
- 7 Burgess, A. W.; Leach, S. J. *Biopolymers* **1973**, *12*, 2599.
- 8 Ramachandran, G. N. and Sasisekharan, V. *Adv. Protein Chem.* **1968**, *23*, 283.
- 9 Aravinda, S.; Shamala N. and Balaram P. *Chem Biodivers.*, **2008**, *5*, 1238-1262.
- 10 Barlow, D. J.; Thornston, J. M. *J. Mol. Biol.* **1988**, *201*, 601.
- 11 Prasad, B. V. V.; Balaram, P. *CRC Crit. Rev. Biochem.* **1984**, *16*, 307.
- 12 Augspurger, J. D.; Bindra, V. A.; Scheraga, H. A.; Kuki, A. *Biochemistry*, **1995**, *34*, 2566.
- 13 De Fillipis, V.; De Antoni, F.; Frigo, M.; de Laureto, P. P.; Fontana, A. *Biochemistry*, **1998**, *37*, 1686.
- 14 (a) Bosch, R.; Jung, G.; Schmitt, H.; Winter, W. *Biopolymers* **1995**, *24*, 979. (b) Tomicolo, C.; Bonora, G. M.; Bovaso, A.; Bendetti, E.; DiBlasio, B.; Pavone, V.; Pedone, C. *J. Biomol. Struct.* **1985**, *3*, 585.
- 15 Kaiser, E.; Colescott, R.L.; Bossinger, C.D.; Cook, P.I. *Anal. Biochem.* **1970**, *34*, 595-598.
- 16 Kaiser, E.; Bossinger, D. D.; Colescott, R. L.; Olsen, D. B.; *Anal. Chim. Acta*, **1980**, *188*, 149-151.
- 17 Whittington, S. J.; Chellgren, B. W.; Hermann, V. M.; Creamer, T. P. *Biochemistry*, **2005**, *44*, 6269–6275.
- 18 Feng, Y.; Melacini, G.; Taulane, J. P.; Goodman, M. *J. Am. Chem.Soc.* **1996**, *118*, 10351-10358.
- 19 Fernandez, F.; Scherega, H. A. *Proc. Natl. Acad. Sci.*, **2003**, *100*, 113-118.
- 20 Ikura, T.; Urakubo, Y.; Nobutoshi, I. *Chem. Phys.* **2004**, *307*, 111-119.
- 21 Pacaroni, A.; Cinelli, S.; Cornicchi, E.; de Francesco, A.; Onori, G. *Chem. Phys. Lett.* **2005**, *410*, 400-403.
- 22 Cheung, M. S.; Garcia, A. E.; Onuchic, J. N. *Proc. Nat. Acad. Sci. USA* **2002**, *99* 685-690.
- 23 (a) Harrap, B. S. *Int. J. Pept. Protein Res.* **1969**, *1*, 2527-2532. (b) Gekko, K.; Koga, S. *J. Biochem.* **1983**, *94*, 199-205. (c) Feng, Y.; Melacini, G.; Taulane, J. P.; Goodman, M. *J. Am. Chem.Soc.* **1996**, *118*, 10351-10358.
- 24 Feng, Y.; Melacini, G.; Taulane, J. P.; Goodman, M. *J. Am. Chem.Soc.* **1996**, *118*, 10351-10358.
- 25 Horng, J. C.; Kotch, F. W. and Raines R. T. *Protein Sci.*, **2007**, *16*, 208–215.

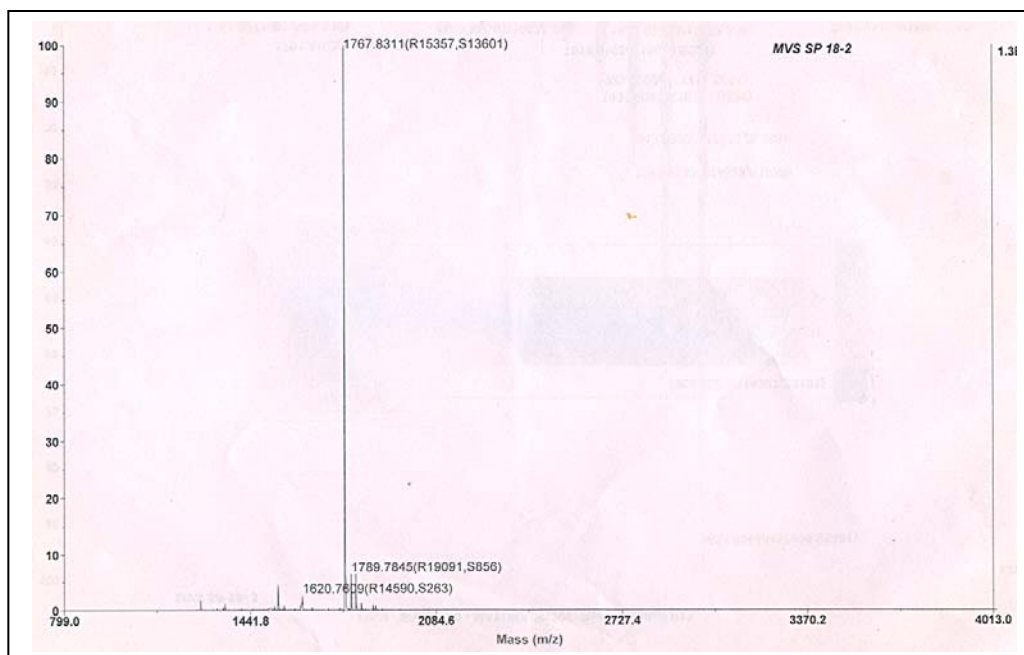
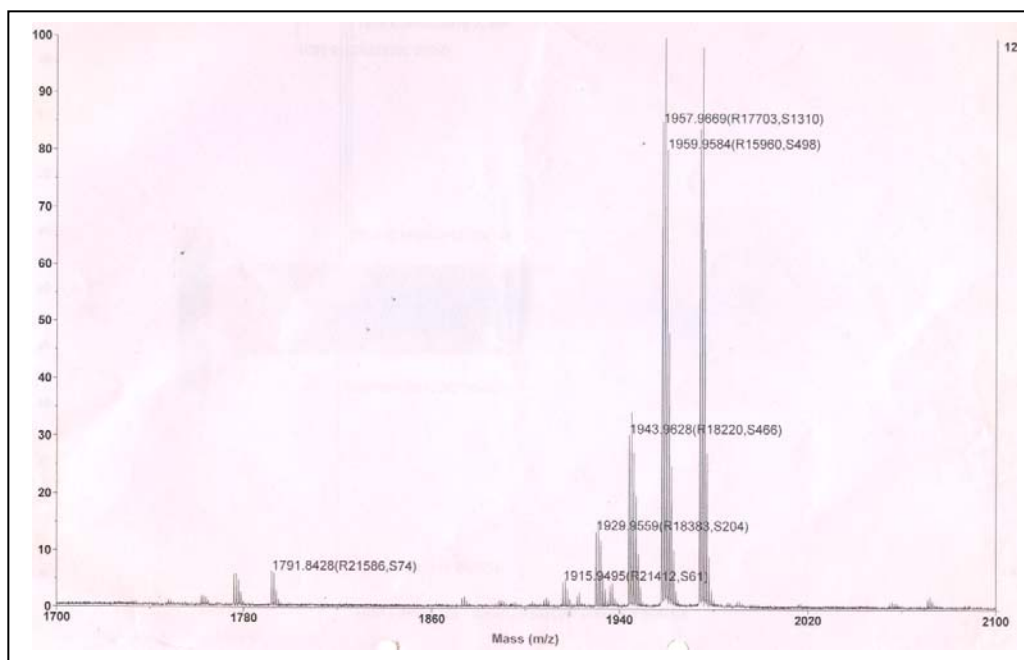
- 26 Brown, F. R.; Di Corato, A.; Lorenzi, G. P.; Blout, E. R. *J. Mol. Biol.* **1972**, *63*, 85-99.
- 27 (a) Brown, F. R., III; Carver, J. P.; Blout, E. R. *J. Mol. Biol.* **1969**, *39*, 307-313. (b) Vuilleumier, S.; Mutter, M. *Biopolymers* **1993**, *33*, 389-400.
- 28 (a) Babu, I. R.; Ganesh, K. N. *J. Am. Chem. Soc.* **2001**, *123*, 2079-2080. (b) Umashankara, M.; Babu, I. R.; Ganesh, K. N. *Chem. Comm.* **2003**, 2606-2607.
- 29 Beck, K.; Chan, V. C.; Shenoy, N.; Kirkpatrick, A.; Ramshaw, J.; Brodsky, B. *Proc. Natl. Acad. Sci.*, **2000**, *97*, 4273-4278.
- 30 Werkmeister, J. A.; Ramshaw, J. A. M. (Ed.) *Collagen Biomaterials*, Elsevier Science, Barking, Essex, **1992**.
- 31 Vojkovsky, T. *Pept. Res.* **1995**, *8*, 236-237.

1.9 Appendix 3: Characterization data of synthesized peptides

Entry	Page No.
HPLC of Peptides (30-33)	332-333
MALDI-TOF of peptides (30-33)	333-334

(A) HPLC of peptides 30-33**HPLC of peptide 30 (AcPhe(Pro-Hyp-Gly)₆)****HPLC of peptide 31 (AcPhe(Pro-Hyp-Aib)₆)****HPLC of peptide 32 (H₂N-Phe(Pro-Hyp-Gly)₆)**

HPLC of peptide 33 (H₂N-Phe(Pro-Hyp-Aib)₆)**(B) MALDI-TOF of peptides 30-33****MALDI-TOF of peptide 30 (AcPhe(Pro-Hyp-Gly)₆)****MALDI-TOF of peptide 31 (AcPhe(Pro-Hyp-Aib)₆)**

MALDI-TOF of peptide 32 (H₂N-Phe(Pro-Hyp-Gly)₆)**MALDI-TOF of peptide 33 (H₂N-Phe(Pro-Hyp-Aib)₆)**

Water-Induced Switching of β -Structure to Polyproline II Conformation in the 4*S*-Aminoproline Polypeptide via H-Bond Rearrangement

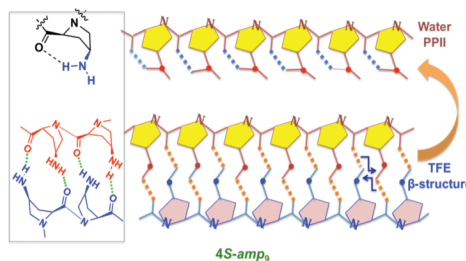
Mahesh V. Sonar[†] and Krishna N. Ganesh^{*†‡}

Division of Organic Chemistry, National Chemical Laboratory, Dr Homi Bhabha Road, Pune 411008, India, and Indian Institute of Science Education and Research, 900, NCL Innovation Park, Dr Homi Bhabha Road, Pune 411008, India

kn.ganesh@iiserpune.ac.in

Received September 14, 2010

ABSTRACT



4*S*-Aminoproline polypeptide 2 forms unusual β -structure in trifluoroethanol that switches to the polyproline II (PPII) form in aqueous medium, while 4*R*-aminoproline peptide 1 retains PPII form in both solvents. This first instance of a polyproline derivative showing a β -structure is attributed to competitive pH-dependent (4-NH₃⁺/NH₂) stereoelectronic effect (4*R* vs 4*S*) and the overriding importance of stereospecific intra/intermolecular H-bonding in (2,4)-*cis*-4*S*-aminoproline in contrast to (2,4)-*trans*-4*R*-aminoproline oligomers.

The polyproline type II (PPII) helix is a prevalent conformation in both folded and unfolded proteins¹ and plays an important role in a wide variety of biological processes, such as signal transduction, transcription, immune response, and cell motility.² Each strand of collagen triplex with the Pro-Hyp-Gly tripeptide repeat unit adopts a left-handed PPII-like conformation.³ Oligoprolines and their derivatives have

found utility as cell penetrating agents⁴ and as molecular spacers in biomimetic systems for energy/electron transport.⁵ The PPII helix is a fully extended left-handed structure with all amide bonds in the *trans* conformation, while the right-handed PPI helix is compact, with all amide bonds in the *cis* conformation.¹ It is well-known that polyprolines adopt PPII conformation in water⁶ and PPI conformation in hydrophobic solvents of short chain aliphatic alcohols.⁶ It has been well demonstrated that the stereoelectronic effect

[†] National Chemical Laboratory.

[‡] Indian Institute of Science Education and Research.

(1) (a) Cowan, P. M.; McGavin, S. *Nature* **1955**, *176*, 501–503. (b) Traub, W.; Shmueli, U. *Nature* **1963**, *198*, 1165–1166.

(2) (a) Rath, A.; Davidson, R.; Deber, C. M. *Biopolymers (Pept. Sci.)* **2005**, *80*, 179–185. (b) Holt, M. R.; Koffer, A. A. *Trends Cell Biol.* **2001**, *11*, 38–46. (c) Kay, B. K.; Williamson, M. P.; Sudol, M. *FASEB J.* **2000**, *14*, 231–241.

(3) (a) Brodsky, B.; Thiagarajan, G.; Madhan, B.; Kar, K. *Biopolymers* **2008**, *89*, 345–353. (b) Shoulders, M. D.; Raines, R. T. *Annu. Rev. Biochem.* **2009**, *78*, 929–958.

(4) (a) Farrera-Sinfreu, J.; Giral, E.; Castel, S.; Albericio, F.; Royo, M. *J. Am. Chem. Soc.* **2005**, *127*, 9459–9468. (b) Fillon, Y. A.; Anderson, J. P.; Chmielewski, J. *J. Am. Chem. Soc.* **2005**, *127*, 11798–11803.

(5) (a) Doose, S.; Neuweiler, H.; Barsch, H.; Saer, M. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 17400–17405. (b) Schuler, B.; Lipman, E. A.; Steinbach, P. J.; Klumke, M.; Eaton, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 2754–2759.

(6) (a) Knof, S.; Engel, J. *Isr. J. Chem.* **1974**, *12*, 165–177. (b) Mutter, M.; Wöhr, T.; Gioria, S.; Keller, M. *Biopolymers* **1999**, *51*, 121–128. (c) Kakinoki, S.; Hirano, Y.; Oka, M. *Polym. Bull.* **2005**, *53*, 109–115.

of the 4-substituent plays a major role in determining the pucker of the pyrrolidine ring of proline and hence the conformational stability of proteins.⁷

In this context, we reported earlier that 4-amino substitution on proline in collagen peptide stabilizes the triple helix⁸ at both acidic and basic pHs. In collagen, the tripeptide repeat unit [Pro-Hyp-Gly] has glycine whose amide linkage is involved in an interchain H-bond, leading to the triple helix structure. In contrast, polyproline peptides lack amide NH and hence are unable to form a triplex via interchain H-bonds, ending up as a single helix of PPI or PPII type. Recently, 4S(OH/NH₃⁺) groups on proline were shown to form intramolecular H-bonds with the amide carbonyl, increasing the *trans/cis* amide ratio and thereby promoting PPII conformation in the derived polypeptides.⁹ Unlike other 4-substituents on proline studied so far (OH, SH, CH₃, F), the ionizable 4-NH₂ group is a good probe to examine the pH effects on polyproline conformation. Herein, we report the novel behavior of 4S-aminoproline polypeptide to form a novel β -structure in trifluoroethanol (TFE) that switches to PPII form in aqueous medium. This property exclusive to 4S-aminoproline polypeptide arises from a stereospecific *intramolecular* H-bonding that stabilizes the PPII form, while the unusual β -structure results from *interchain* H-bonding. To our knowledge, this is perhaps the first report of formation of β -structure in any polyproline derivatives and its switch over to PPII form induced by water.

Synthesis and Conformational Studies of 4-Aminoproline Oligomers. The oligopeptides **1–3** were synthesized from appropriate N-Fmoc-protected monomers assembled in the solid phase, purified by HPLC, and characterized by MALDI-TOF (for details, see Supporting Information). The CD spectral analyses were carried out as a function of temperature, pH, urea, and solvents (buffer and trifluoroethanol, TFE). All three peptides (100 μ M, pH 7.2) show CD spectra (Figure 1) with a positive band between 220 and 230 nm and a negative band between 200 and 210 nm that are the established patterns of the PPII conformation.¹⁰ The intensity of the positive band at 225–227 nm is proportional to the PPII helical content which is seen to decrease in the order 4*R*-Amp₉ **1** > 4*S*-amp₉ **2** > Pro₉ **3**.

The effect of protonation of the 4-amino group on PPII helical content was examined by the CD spectra of peptides **1–3** (Figure 2A) recorded at different pH (4.0–10.0). The positive ellipticity at 225 nm for 4*R*-Amp₉ **1** decreased by 10% with increasing pH up to 7.2 and did not change further until pH 10.0. In the case of 4*S*-amp₉ **2**, positive intensity was enhanced in the pH range 4.0–10.0 in a sigmoidal

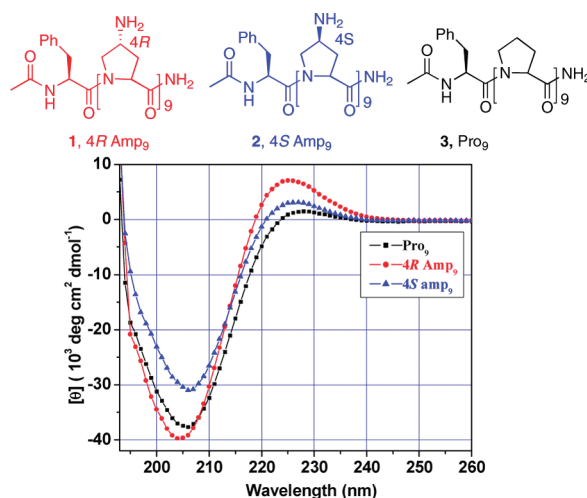


Figure 1. CD profiles of polypeptides **1**, 4*R*-Amp₉ (●); **2**, 4*S*-amp₉ (▲); and **3**, Pro₉ (■), all at 100 μ M (pH 7.2).

fashion. At acidic pH (4.0–5.0), the PPII helicity of 4*S*-amp₉ **2** was low (20% of 4*R*-Amp₉ **1**) but increased by 2-fold at pH 10.0. The ellipticity of peptide Pro₉ **3** remained

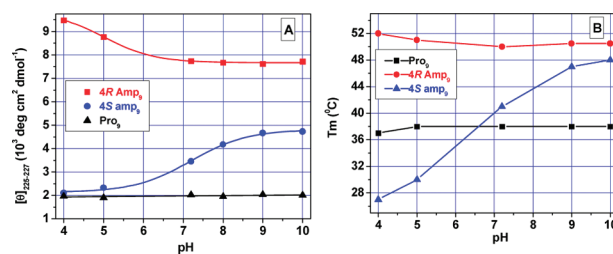


Figure 2. (A) Intensity of the positive band of CD spectra of peptides **1–3** as a function of pH. (B) Thermal stability of PPII helices in peptides **1–3** as a function of pH, followed at 225 nm.

constant with pH. This suggests a pivotal role for both stereochemistry and the protonation status of the 4-amino group in eliciting the PPII helicity of 4(*R/S*)-aminoproline polypeptides **1** and **2**.

The pH-dependent thermal stability (T_m) of PPII helices in peptides **1–3** was measured from temperature-dependent CD spectral data (Figure 2B and Supporting Information). It is seen that (i) 4*R*-Amp₉ **1** has maximum T_m at all pHs; almost invariant (ii) 4*S*-amp₉ **2** has the lowest T_m among the peptides at pH 4.0 but increased gradually with raise in pH to 10.0 to a value closer to the T_m of 4*R*-Amp₉ **1**; and (iii) Pro₉ **3** with intermediate T_m at pH 4.0 remained constant over the pH range. The 4-NH₃⁺ group at pH 4.0 stabilized the PPII helix most in the 4*R*-form and least in the 4*S*-form, while 4-NH₂ at pH 10.0 stabilized both 4*R*- and 4*S*-peptides to a similar extent. The 4*S*-amp₉ **2** thus exhibited significant pH-dependent PPII stability that is maximum in the unionized amino form.

(7) Bretscher, L. E.; Jenkins, C. L.; Taylor, K. L.; DeRider, M. L.; Raines, R. T. *J. Am. Chem. Soc.* **2001**, *123*, 777–778.

(8) Babu, I. R.; Ganesh, K. N. *J. Am. Chem. Soc.* **2001**, *123*, 2079–2080. (b) Umashankara, M.; Babu, I. R.; Ganesh, K. N. *Chem. Commun.* **2003**, 2606–2607.

(9) (a) Shoulders, M. D.; Kotch, F. K.; Choudhary, A.; Guzei, I. A.; Raines, R. T. *J. Am. Chem. Soc.* **2010**, *132*, 10857–10865. (b) Kuemin, M.; Nagel, Y. A.; Schweizer, S.; Monnard, F. W.; Ochsenfeld, C.; Wennemers, H. *Angew. Chem., Int. Ed.* **2010**, *49*, 6324–6327.

(10) Woody, R. W. *Adv. Biophys. Chem.* **1992**, *2*, 37–79.

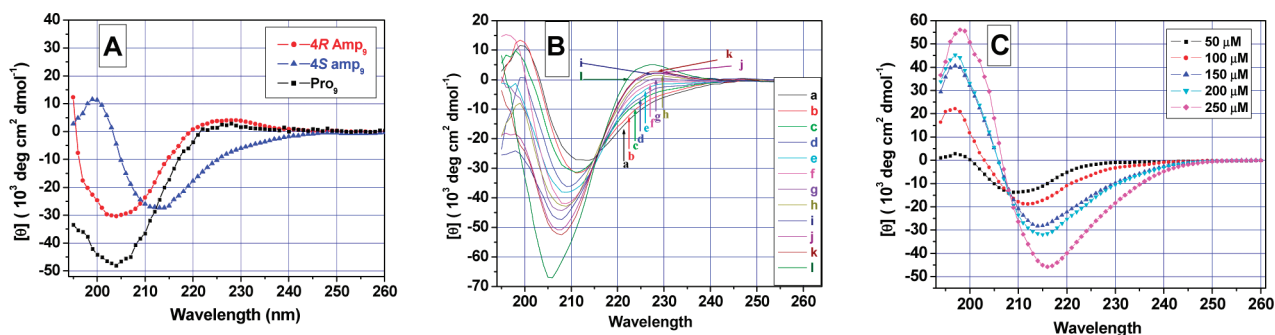


Figure 3. (A) CD spectra of peptides **1–3** in trifluoroethanol (TFE). (B) CD spectra of **4S-amp₉** **2** in TFE with incremental addition of phosphate buffer (pH 7.2) from (a) 0.1% to (k) 1.0% in 0.1% steps and (l) 2.0%. (C) Increasing concentration of **2** from 50 to 250 μM .

Urea is known to enhance the PPII helical content through rigidification of the polypeptide backbone.¹¹ The low PPII helicity of **4S-amp₉** **2** was enhanced enormously (>300%) by addition of 1 M urea (pH 7.2), while that of **4R-Amp₉** **1** and (**Pro₉**) **3** increased by a mere 15–20% (Supporting Information). The larger changes seen specifically for **4S-amp₉** **2** in the presence of urea and upon increasing the pH suggest the combined role of H-bonding and stereoelectronic¹² effects in dictating the PPII conformation.

Solvent plays a key role in modulating the H-bonding effects, and hence the CD spectra of peptides **1–3** were recorded in a fluorinated solvent trifluoroethanol (TFE) (Figure 3A). The **4R-Amp₉** **1** and **Pro₉** **3** show CD spectra typical of PPII form. Very interestingly, the CD spectra of **4S-amp₉** **2** in TFE are unlike the PPII pattern, showing a negative maximum around 214 nm and a broad shoulder at 228 nm that is typical of β -structure.¹³

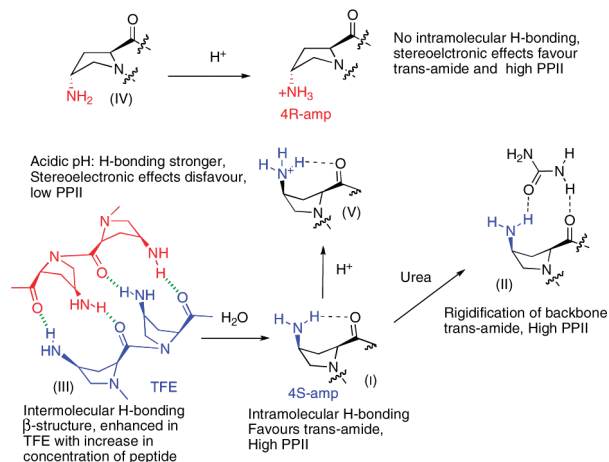
When aqueous phosphate buffer (pH 7.2) was titrated into TFE solution of **4S-peptide 2** in tiny incremental steps of 0.1%, the 214 nm negative band slowly shifted to 205 nm, accompanied by a growing of the broad negative shoulder at 228 nm into a positive band at about 224 nm (Figure 3B) typical of PPII form. The isosbestic point seen at 215 nm is indicative of the conversion of **4S-peptide 2** from β -structure in 100% TFE to full PPII form with 0.8% buffer in TFE.

Upon increasing the concentration of **4S-amp₉** **2** from 50 to 250 μM in TFE, the CD spectra exhibited a nice growth in the negative band intensity at 210 nm, accompanied by its shift to 216 nm and a large increase of the positive band at 200 nm (Figure 3C). As expected for intermolecular hydrogen bonding increasing with concentration, this strongly points to a consolidation of β -structure in the peptide **4S-amp₉** **2**. In the case of **4R-Amp₉** **1**, increasing the peptide concentration leads to enhancement of the PPII form without any other changes. The overall results imply that **4S-amp₉** **2** assumes a β -structure in TFE that is transformed to the PPII form in aqueous medium, unlike **4R-amp₉** **1** which retains the PPII form in both conditions.

The formation of β -structure in polyproline peptides under any conditions is unprecedented in the literature since they lack H-bond donor sites. In **4S-amp₉** **2**, the NH_2 group can form an intramolecular H-bond with the amide carbonyl of the same proline moiety, promoting a PPII conformation.^{9b} The possibility of the **4S-NH₂** group engaging the amide carbonyl of another chain of **4S-amp₉** through an intermolecular H-bond would lead to β -structure. Such an interchain H-bonded structure should be facilitated at higher peptide concentration.

A plausible molecular picture for this conversion is depicted in Scheme 1. The intramolecular H-bonding of **4S-**

Scheme 1. Solvent-Derived Rearrangement of H-Bonds in **4S-amp₆** Leading to Conformational Switch



NH_2 with amide carbonyl possible only in **4S-amp₉** **2** (I) promotes the PPII form in buffer. Urea rigidifies the backbone¹¹ by complementary H-bonding (II) among the *cis*-disposed **4S-NH₂** group and the amide carbonyl to strengthen the PPII form. In a fluorinated solvent TFE, the intramolecular H-bonding between **4S-NH₂** and the amide carbonyl switches to interchain H-bonding (III) giving rise to anti-parallel β -structure. The *trans* disposition of **4R-NH₂** and

(11) (a) Robinson, D. R.; Jencks, W. P. *J. Am. Chem. Soc.* **1965**, *87*, 2462–2469. (b) Whittington, S. J.; Chellgren, B. W.; Hermann, V. M.; Creamer, T. P. *Biochemistry* **2005**, *44*, 6269–6275.

(12) Jia-Cherng, H.; Raines, R. T. *Protein Sci.* **2006**, *15*, 74–83.

(13) Seebach, D.; Overhand, M.; Kiihnl, F. N. M.; Martinoni, B. *Helv. Chim. Acta* **1996**, *79*, 913–941.

the amide carbonyl group in 4*R*-*Amp*₉ (IV) is not conducive to formation of either intramolecular or strong interchain H-bonding in the derived peptide.

Stereoelectronic vs H-Bonding Effects in 4*S*-*amp*₉ 2. The stereoelectronic effect of 4*R*-X substituents on proline is known to strongly favor the PPII form, over that of 4*S*-X substituents by enhancing the *trans*-amide content.¹² Although the intramolecular H-bonding in 4*S*-NH₃⁺ favors the *trans*-amide form,^{9b} the unfavorable stereoelectronic effect of 4*S*-NH₃⁺ (unlike 4*R*-NH₃⁺) strongly negates the benefit of H-bonding (V), leading to a low PPII form for **2** at acidic pH. At higher pH, the effect of intramolecular H-bonding dominates a weaker stereoelectronic effect of 4*S*-NH₂, promoting higher PPII content in peptide **2**. While the conversion of PPII to PPI takes many hours/days due to a slow conversion of amide from *trans* to *cis* form,¹⁴ the switching of β -structure to PPII form is fast, within minutes, suggesting that the amide bond of **2** is also in *trans* form in the β -structure. No PPI form was seen for any of the peptides **1–3**, under different conditions of pH, *n*-propanol and TFE.

In conclusion, it is demonstrated here that 4*S*-*amp*₉ **2** adapts an unusual β -structure in TFE unlike most polyproline peptides which prefer the PPI form in hydrophobic/fluorinated media. The β -structure arises from interchain

hydrogen bonds involving 4*S*-NH₂ and amide carbonyl, which are broken in water and rearranged to *intramolecular* H-bonding that favors the PPII form via enriching the *trans*-amide geometry. This structural conversion illustrates a fine balance between stereoelectronic and H-bonding effects in novel tuning of the secondary structure of 4*R/S*-aminoproline polypeptides. β -Structure in polyproline peptides is hitherto unknown, and the present results will add a new design principle to a growing repertoire of strategies for engineering peptide secondary structural motifs for new biomaterials and nanoassemblies.¹⁵

Acknowledgment. MVS acknowledges an award of fellowship by Lady Tata Memorial Trust and CSIR (New Delhi). KNG holds JC Bose Fellowship of DST (New Delhi).

Supporting Information Available: Experimental synthesis procedures, HPLC, mass spectral data, and CD spectra of peptides **1–3** under different conditions of pH, urea, salt, and concentration. This material is available free of charge via the Internet at <http://pubs.acs.org>.

OL1021993

(14) Chiang, Y.-C.; Lin, Y.-J.; Horng, J.-C. *Protein Sci.* **2009**, *18*, 1967–1977.

(15) (a) Tsai, C.-J.; Zheng, J.; Zanuy, D.; Haspel, N.; Wolfson, H.; Alema'n, C.; Nussinov, R. *Proteins* **2007**, *68*, 1–12. (b) Zhao, X.; Pan, F.; Xu, H.; Yaseen, M.; Shan, H.; Hauser, C. A. E.; Zhang, S.; Lu, J. R. *Chem. Soc. Rev.* **2010**, *39*, 3480–3498. (c) Apostolovic, B.; Maarten, D. M.; Klok, H.-M. *Chem. Soc. Rev.* **2010**, *39*, 3541–3550.