"SYNTHESIS AND BIOPHYSICAL STUDIES OF 4-SUBSTITUTED PROLINE CONTAINING COLLAGEN PEPTIDES: EFFECT OF SUBSTITUENTS ON TRIPLEX STRUCTURE"

THESIS SUBMITTED TO

THE UNIVERSITY OF PUNE

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

CHEMISTRY

BY

M. UMASHANKARA

NATIONAL CHEMICAL LABORATORY

PUNE 411008

OCTOBER 2006

In memories of

BHARATA RATNA **SIR M. VISVESVARAYA**

1861-1962

and

RASHTRA KAVI **KUVEMPU**

1904-1994

ACKNOWLEDGEMENTS

With deep felt sense of gratitude, I thank my supervisor, Dr. Krishna N. Ganesh for his whole hearted support, constant encouragement, timely help and personal freedom rendered to me during my research period. His endless enthusiasm and receptive attitude will always remain a source of inspiration. He has been very patient and tolerant towards my erratic behavior and set me back in the right course without ever hurting me.

I am indebted to my parents who have been a source of encouragement and I am grateful for all their sacrifice.

I am grateful to Dr. Rajmohan, Mrs. Usha Phalgune and Dr. Swapna Ravindranatan without whose expertise in NMR spectroscopy the fifth chapter would not have been possible. Their scientific temper and pleasing personalities make them great people to work with.

I thank my seniors Dr. Mrs. Meena Sharma and Dr. Ramesh Babu they tought me solidphase peptide syntheis, running HPLC, CD spectrometer etc… during my starting days of PhD course.

My special thanks for Dr. (Mrs) Vijayanti Anil Kumar, Dr. A. A. Natu, Dr. Moneesha D' Costa and Mrs. Anita Gunjal. They were always there to help me.

I take this opportunity to place on record my thanks to Dr. Vairamani, IICT, Hyderabad and Prof. Balram, IISc, Bangalore, Dr. Kumbar, Dr, Mahesh Kulkarni and Dr, Shanta Kumari for their help in recording the MALDI-TOF and FAB mass spectra.

The kind help extended by the staff of the NMR facility at NCL, Mr.Samuel, Mr. Sathe also must be acknowledged here. I am especially grateful to Mrs. M. V. Mane and Mrs. S. S. Kunte for carrying out HPLC analysis,

I express my sincere thanks to Prof. Rangappa and Prof. D. Channegowda ,Univrcity of Mysore, Mysore, Karnataka, for their concern and guiding in my M Sc course.

This page would be incomplete without the mention of my seniors Dr.Pradeeppallan, Dr. NagendraSharma, Dr. PallaviLonlar, Dr. DineshDandekar, Dr. T .Govindaraju and RammanVyasaBattar, who have extended their friendship and helped me with various capacities.

I express my heartful thanks to my lab mates GouriShankar, Praveen, Kirhud, Amit, Snil.G, KosgiSridhar, AshwiniSharma, MaheshSonar, Sachin, Smita, madhuri, Gitali, Roopa, Manaswini, Pradnya, Nasrin, Satish, and Dr. Susmita for their help and co-operation

I also thank Pawar, Bhumkar and Sunil, More for their assistance.

My friends in G. J. hostel and in NCL have made my stay at NCL a lively and memorable one and I am indebted to them for their support.

I am grateful to the Director, NCL, for giving me the opportunity to work in this premier research institute

The financial support in the form of fellowship provided by CSIR, New Delhi, is also acknowledged.

Finally I thank NCL staff and all the people who helped and supported me in my research period

M. Umashankara

CANDIDATE'S DECLARATION

I hereby declare that the thesis entitled **"SYNTHESIS AND BIOPHYSICAL STUDIES OF 4-SUBSTITUTED PROLINE CONTAINING COLLAGEN PEPTIDES: EFFECT OF SUBSTITUENTS ON TRIPLEX STRUCTURE"** submitted for the award of 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 by me at the National Chemical Laboratory, Pune, India.

(**M. Umashankara**)
National Chemical Laboratory
Pune 411 008
India

October 2006

CONTENTS

List of publication		**i**
Abstract		**ii-xiv**
Abbreviations and symbols		**xv-xviii**

CHAPTER 1

Collagen Structure And Mimetics: An Introduction

1.1	**Introduction**	**1**
1.2	**Collagen super family**	**2**
1.3	**Biosynthesis of collagen**	**3**
1.4	**Positional preferences of amino acids in collagen**	**7**
1.5	**Three-dimensional structure of collagen**	**9**
1.6	**Use of synthetic polypeptide models**	**11**
1.7	**Crystal packing of collagen polypeptide models**	**14**
1.8	**Role of trans-4-hydroxyproline in collagen structure**	**16**
1.9	**Conformation of polypeptide chains**	**19**
1.10	**Pyrrolidine ring conformation in proline and substituted proline**	**22**
1.11	**Gauche effect on the ring-pucker preferences**	**23**
1.12	**Effect of 4-substituents on the pyrrolidine ring pucker**	**25**
1.13	**Effect of 4-substituents on E-Z isomerization of prolyl-peptide bond**	**26**
1.14	**Electrostatic interactions in the triple-helix**	**30**
1.15	**Effect of charged termini on the stability of triple-helices**	**31**
1.16	**Effect of ionizable side chains**	**32**

1.17	**Characterization of triple-helical structure**	**32**
1.18	**Collagen mimetics**	**34**
1.19	**Collagen mimetics with unnatural aminoacids**	**35**
1.20	**Template-assembled collagen structure**	**37**
1.21	**Collagen in aging and disease**	**40**
1.22	**Collagen biomaterials**	**44**
1.23	**Present work**	**45**
1.24	**Objective in under taking work**	**47**
1.25	**References**	**50**

CHAPTER 2

Investigation Of Compatibility Of 4*S*- And 4*R*-Aminoprolines In X And Y Position Of The Collagen Peptides $(X\text{-}Y\text{-}Gly)_n$

2.1	**Introduction**	**60**
2.1.1	Trans-cis isomerization	62
2.2	**Objective of present work**	**64**
2.3	**Results**	**65**
2.3.1	Synthesis of fully protected (2*S*,4*R*) and (2*S*,4*S*)-aminoproline monomers	65
2.3.2	Solidphase peptide synthesis (SPPS)	68
2.3.3	Comparison of Fmoc and Boc chemistry	69
2.4	**Present work**	**70**
2.4.1	End group ionizing effect	73
2.4.2	Determination of the peptide concentration in stock solution	74
2.4.3	Determination of pKa of 4-aminogroup of Amp and amp monomers	74

	Characterization of triple-helical structure	**76**
2.5.1	Concentration dependence on triple-helix formation of peptides 20-22	76
2.5.2	Salt effects on triple-helical stability	88
2.5.3	Effect of ethylene glycol on the stability of triple-helices	91
2.5.4	Chain length dependent triple-helix formation	93
2.5.5	Concentration effect on triple-helix formation of Ac-Phe(amp-Pro-Gly)$_6$-NH$_2$	94
2.6	**Discussion**	**96**
2.7	**Conclusions**	**100**
2.8	**Experimental section**	**101**
2.8.1	Peptide synthesis	102
2.8.2	Resin functionalization	102
2.8.3	General method for solid phase peptide synthesis	102
2.8.4	Synthesis protocol for solid phase peptide synthesis	103
2.8.5	Kaiser's test	103
2.8.6	Chloronil test	104
2.8.7	Preparation of resin with peptide for cleavage	104
2.8.8	Peptide cleavage from the resin	104
2.8.9	HPLC purification and analysis	105
2.8.10	CD spectroscopy and thermal denaturation studies	106
2.8.11	CD thermal denaturation studies	106
2.8.12	pKa determination of Amp and amp	107
2.8.13	Synthesis of compound 6-18	107
2.9	**References**	**116**
	Appendix - I	**120**

CHAPTER 3

TWO AMINOPROLINES WITH A DIFFERENCE

Part A

Stereo Electronic Effects Of 4-*R*/*S*-Aminoproline On Stability In Collagen Peptides $[Pro(X)\text{-}Pro(Y)\text{-}Gly]_n$.

3.A.1	**Introduction**	**138**
3.A.1.1	Pyrrolidine ring stability by stereochemistry of 4-substituents	142
3.A.2	**Aim and rationale of present work**	**144**
3.A.3	**Results**	**146**
3.A.3.1	Solid phase peptide synthesis of chimeric collagen peptides	147
3.A.4	**Biophysical studies**	**150**
3.A.4.1	Concentration dependent triple-helix formation of peptides 23-26	150
3.A.4.2	Characterization of triple-helical structure by CD spectroscopy	152
3.A.4.3	Effect of ethylene glycol on the stability of peptides 23-26 triple-helices	160
3.A.4.4	Salt effects on triple-helix stability of chimeric peptides 23-24	161
3.A.5	**Discussion**	**163**

DUAL COMPATIBILITY OF 4*R*-AMINOPROLINE

Part B

Hyper Stable Collagen Peptides $(X\text{-}Y\text{-}Gly)_n$ With Modified 4R-Iminoacids Both At X And Y Positions

3.B.1	**Introduction**	**166**
3.B.1.1	Preorganization of peptide back-bone angles	167

3.B.1.2	Reciprocity of steric and steroelectronic effects	167
3.B.1.3	Duality of 4*R*-hydroxyproline	170
3.B.1.4	Crystal structure of Ac(4*R*-Hyp-4*R*-Hyp-Gly)$_9$-NH_2	171
3.B.1.5	Puckering alternation and triple-helix stability	173
3.B.2	**Aim and rationale of the present work**	**175**
3.B.3	**Synthesis of chimeric collagen peptides**	**176**
3.B.4	**Biophysical studies**	**178**
3.B.4.1	Concentration dependent triple-helix formation of peptides 27-30	178
3.B.4.2	Characterization of triple-helical structure by CD spectroscopy	180
3.B.4.3	Effect of ethylene glycol on the stability of peptides 27-30 triple-helices	189
	Salt effects on triple-helix stability of chimeric peptides	190
3.B.5	**Discussion**	**191**
3.B.6	**Conclusion**	**195**
3.B.7	**Experimental section**	**196**
3.B.7.1	Synthesis of compound 7-11	197
3.B.8	**References**	**201**
	Appendix - II	**204**

CHAPTER 4

4-(N-Formyl)-Aminoproline Collagen Peptides: Probes For Studying The Role Of Protonated 4-Amino Group In Triplex Stability

4.1	**Introduction**	**217**
4.1.1	Proline ring pucker and amide bond geometry	217

4.1.2	Conformational preferences of Y position of the pyrrolidine ring pucker	218
4.1.3	Solvation of proline residue in triple-helix	220
4.2	**Rationale and objective of the present work**	**221**
4.3	**Results**	**222**
4.3.1	Synthesis of (2*S*,4*R*) and (2*S*,4*S*) -4-Nformyl-aminoprolines	222
4.3.2	Solid phase synthesis of peptides 31-34	223
4.4	Biophysical studies	226
4.4.1	Characterization of triple-helical structure by CD spectroscopy	226
4.4.2	Sample preparation	226
4.4.3	Effect of ethylene glycol on the stability of peptides 31-34 triple-helices	233
4.5	**Discussion**	**235**
4.6	**Experimental section**	**238**
4.6.1	General procedures	238
4.7	**References**	**247**
	Appendix – III	248

CHAPTER 5

Prolyl- Peptide- Bond Isomerization In 4*R*/*S* Aminoproline: Effect Of Ph On The Prolyl-Peptide Bond Isomerization In Aminoproline Containing Collagen Peptides

5.1	**Introduction**	**263**
5.1.1	Factors affecting conformation of prolyl amide bond in proline-containing peptides	264
5.1.2	Conformational control of pyrrolidine ring	267

5.2	**Rationale and objectives of the present work**	**269**
5.3	**Results**	**270**
5.4	**Study of conformational changes of model compounds 12-15**	**272**
5.4.1	Inductive effect of 4-substituents on the prolyl-peptide bond-order in IR-spectroscopic studies	272
5.4.2	Circular dichroisum (CD) spectroscopic studies	276
5.4.3	NMR spectroscopy study	277
5.4.4	Thermodynamics of prolylpeptide bond isomerization: Variable temperature NMR studies	281
5.4.5	Pyrrolidine ring conformation derived from J_{H-H} coupling of compounds 12-15	288
5.5	**Determination of pKa value of 4-aminoproline containing collagen peptides**	**298**
5.5.1	pKa calculation methods	298
5.5.2	Molecular dynamics (MD)-based method	299
5.5.3	Using pH titration curve	299
5.6	**Discussion**	**304**
5.6.1	Infrared spectroscopy	305
5.6.2	^{1}H-NMR spectroscopy	306
5.6.3	Thermodynamic analysis	307
5.6.4	pKa of 4-amino group in 4-aminoproline collagen peptide	308
5.7	**Implication for collagen stability**	**309**
5.8	**Conclusion**	**317**
5.9	**Experimental section**	**318**
5.9.1	Synthesis of compounds 12-15	318
5.9.2	NMR experiments	320
5.9.3	Conformational analysis	321
5.10	**Reference**	**322**
	Appendix - IV	**326**

PUBLICATION

Two prolines with a difference: constraining stereoelectronic effects of 4*R*/*S*-aminoprolines on triplex stability in collagen peptides [Pro(X)-Pro(Y)-Gly]$_n$.

M.Umashankara, I. Ramesh Babu, Krishna N. Gasesh*. *Chem. Commun*. **2003**, 2606-2607.

ABSTRACT

The thesis comprises of studies towards the synthesis of 4*R*/*S*-aminoproline containing cationic collagen peptides and evaluation of their triplex forming abilities at different pHs. The thesis is divided into five chapters.

Chapter 1: **Collagen structure and mimetics: An introduction**

Collagen is a major structural protein found 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 4*R*-hydroxyproline (Hyp) respectively (Fig 1). The collagen triple-helix consists of three parallel left handed polyproline II-like helices, supercoiled about each other in a right-handed manner. Although its structure was proposed almost 50 years back, a comprehensive understanding of various factors contributing to the stability of collagen triple-helix is still lacking. The initial hypothesis of triple-helix stabilizing influence of 4-OH group of proline through water mediated hydrogen bonds was challenged by the observation that 4-F proline stabilizes triple-helix.

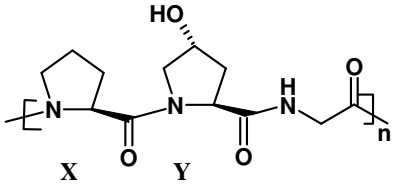


Figure 1: Collagen peptide sequence

Several chemical modifications have been attempted to understand and improve the stability of collagen triple-helices, which may render their use as collagen biomaterials. Some of these include pre-orienting three collagen peptide chains by linking them to a scaffold such as Kemp's triacid, a metal co-ordinating ligand or by replacing Hyp residues by 4*R*-fluoroproline (Flp).

This chapter reviews the current literature on collagen structure and mimetics. The current understanding of structural aspects of collagen in terms of aminoacid composition and hydrogen bonding, relating to its three dimensional structure, is reviewed. Collagen and collagen mimetics as biomaterials, and the role of collagen in diseases are described briefly. The chapter ends with rationale for the present work, based on the literature findings and earlier work from this laboratory.

Chapter 2: Investigation of compatibility of 4*S* and 4*R* aminoproline in X and Y position of the collagen peptides $(X\text{-}Y\text{-}Gly)_n$

Recent studies on $(Pro.Flp.Gly)_n$ collagen model peptides with non-hydrogen bonding 4R-fluoroproline (Flp) in Y position of the X-Y-Gly repeat have shown that these peptides form much stronger triple helices than their corresponding Hyp peptides. This weakens the initial idea that solvation effect may stabilize the triple-helix via water-mediated hydrogen bonds between OH and amide carbonyl groups. Molecular modeling of triplex of $(Pro.Pro.Gly)_{10}$ and its crystal structure has suggested that Pro in Y position prefers to adopt C_4-exo (*up*) pucker, whereas Pro in X position adopt C_4-endo (*down*) pucker (Fig 2). The nature and stereochemistry of the substituents on C_4 influence the pyrrolidine ring pucker and also affect the triple-helix stability. It has been shown that electronegative substituent in 4*R*-position of proline favors the C_4-exo pucker and 4*S*-position of Pro favors the C_4-endo pucker.

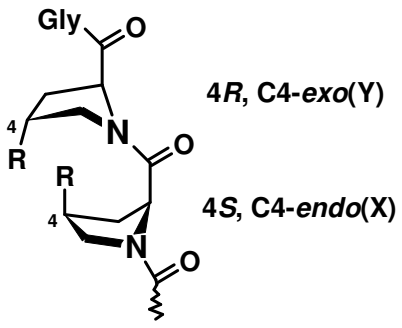
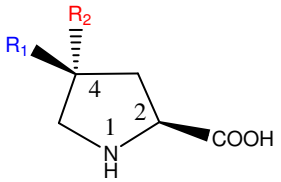


Figure 2a: Position dependent preferred proline puckers in collagen.

In a previous study from this lab[6] it was reasoned that the replacement of 4-OH group by 4-amino group in proline might enable one to understand the true effects of 4-proline substitution on collagen structure. NH_2 group is OH-like in terms of hydrogen bonding and at acidic pH; the NH_3^+ would be more like F in terms of electronegativity effect. Thus the NH_2 group in 4-aminoproline may provide a pH dependent switch for the triple-helix stability. Analysis of the dependence of pyrolidine ring conformation of the 4*R*-aminoproline (Amp) and 4*S*-aminoproline (amp) by vicinal ^{1}H-^{1}H coupling constant in ^{1}H-NMR, indicated that 4*R*-Amp prefers C_4-exo (*up*) and 4*S*-amp prefers C_4-endo (*down*) pucker at neutral pH.



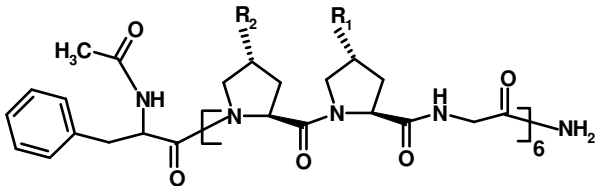
Amp → R_1 = H; R_2 = NH_2. 4*R*-aminoproline
Hyp → R_1 = H; R_2 = OH. 4*R*-hydroxyproline
amp → R_1 = NH_2; R_2 = H. 4*S*-aminoproline
Pro → R_1 = R_2 = H. Proline

Figure 2b: Different 4-Subtituted prolines

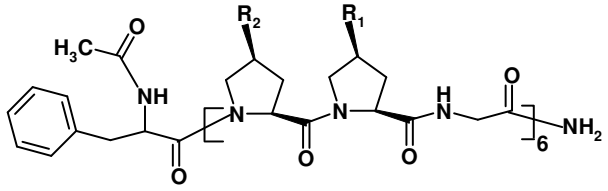
In view of these findings, it was thought worthwhile to study the effect of 4*R* and 4*S*- aminoproline (Fig 2b) in X and Y position of the collagen peptide $(X\text{-}Y\text{-}Gly)_n$. This chapter describes the synthesis, and study of triple-helical propensity of collagen with 4*R* and 4*S* aminoprolines in the Y and X positions of the X-Y-Gly repeat.

The (2*S*,4*R*) and (2*S*,4*S*) N^{α} –Fmoc-N^{γ}-Boc- 4-aminoprolines were synthesized from the naturally occurring *trans*-4-hydroxy-L-proline and used to construct the aminoproline containing collagen peptides, Ac-Phe$(Pro\text{-}Amp\text{-}Gly)_6$-$NH_2$ **19**, Ac-Phe$(Pro\text{-}amp\text{-}Gly)_6$-$NH_2$ **21**, Ac-Phe$(Amp\text{-}Pro\text{-}Gly)_6$-$NH_2$ **20** and Ac-Phe$(amp\text{-}Pro\text{-}Gly)_6$-$NH_2$ **22**, (Fig 3) . These peptides were used to study the triple-helical stability using temperature dependent CD spectroscopy and the results are described. The stability of 4-aminoproline containing collagen triple-helix at different pH and salt

conditions has been studied (Table 1) and compared with their corresponding controls Ac-Phe(Pro-Hyp-Gly)$_6$-NH_2 (**15**) and Ac-Phe(Pro-Pro-Gly)$_6$-NH_2 (**16**) analogues.



19, R_1= NH_2, R_2= H; (**Pro-Amp**) **15**, R_1= OH, R_2= H; (**Pro-Hyp**)
20, R_1= H, R_2= NH_2; (**Amp-Pro**) **16**, R_1= R_2= H; (**Pro-pro**)



21, R_1= NH_2, R_2 = H; (**Pro-amp**) **22**, R_1= H, R_2= NH_2; (**amp-Pro**)

Figure 3: Collagen peptides with Amp and amp at either Y and X positions

Table 1. Triple-helical thermal stabilities (T_m) of collagen peptides [in ^oC]

X-Y→	(Pro-Amp) **19**	(Pro-amp) **21**	(Amp-Pro) **20**	(amp-Pro) **22**	(Pro-Hyp) **15**	(Pro-pro) **16**
pH 3	60	-nt-♀	36	44	27	-nt-♀
pH 7	54.7	-nt-	33	37	28	-nt-
pH 9	26	-nt-	-nt-	34	27	-nt-
pH 12	46	-nt-	-nt-	-nt-	27	-nt-

♀ no transition observed

Aminoproline has positive role in stabilizing the collagen triple-helices in the protonated form (pH 3) resulting in higher triple-helical strength compared to the non-protonated form (pH 12). The results presented here demonstrate that the 4*R*-aminoproline stabilizes the collagen triplex when present in the X-position, and 4S-amp is better than 4*R*-amp.

The 4-aminoproline (X/Y, *S*/*R*) peptides **19-22** always showed better stability than the 4-hydroxyproline analogue 1**5**. The peptide 1**6** without any 4-substituted

proline did not show any triple-helix formation. The cause of the stabilization is the C4-endo pucker conformation adopted by 4*S*-amp that is inherently favored at X position (Fig-1). The properties of these analogues may have significance in the design of new collagen biomaterials.

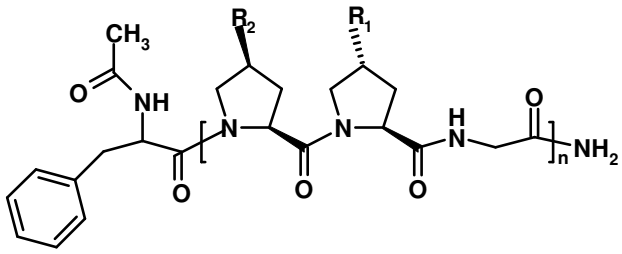
***Chapter 3*: TWO AMINOPROLINES WITH A DIFFERENCE:**

Part-A: Stereoelectronic effects of 4R/S-aminoproline on stability in collagen peptides [Pro(X)-Pro(Y)-Gly]$_n$.

The results of chapter 2 demonstrated that the 4*R/S*-aminoprolines stabilize the collagen triple helix when present in the X-position and 4*S*-amp is better than 4*R*-Amp. The pH dependent stabilities of both 4*R* and 4*S* aminoproline suggested that protonation of NH_2 is prerequisite for triplex formation in the X-position, while it is not so in Y-position. It is possible that the conformation of pyrrolidine ring is also dependent on the protonation status of 4-amino group and the cis-*trans* amide rotomeric equilibrium. The peptides in Chapter 2 have 4-substituted prolines in either X or Y position, it was therefore thought to study the chimeric collagen peptides (**7-10**) wherein both X and Y positions have 4-substituted Prolines to simultaneously exploit the stabilities offered by the two diastereomers to form high stability triplex.

This chapter describes the synthesis, and study of triple-helical propensity of chimeric collagen peptides with 4*R*-NH_2/OH proline at Y position and 4*S*- NH_2/OH proline at X positions of the Gly-X-Y repeat. The protected amino acids, (2*S*,4*R*) and (2*S*,4*S*) N^{α} –Fmoc-N^{γ}-Boc- 4-aminoprolines and (2*S*,4*R*) and (2*S*,4*S*)- N^{α} –Fmoc-hydroxyprolines were used for synthesis of the chimeric peptides, Ac-Phe(amp-Amp-Gly)$_6$-NH_2 **23**, Ac-Phe(amp-Hyp-Gly)$_6$-NH_2 **24**, Ac-Phe(hyp-Amp-Gly)$_6$-NH_2 **25** and Ac-Phe(hyp-Hyp-Gly)$_6$-NH_2 **26**, (**Fig 4**). All peptides were identified by their MALDI-

TOF spectra. The study of the triple-helical stability of these peptides was done using temperature dependent CD spectroscopy at different pH.



23, R_1= NH_2, R_2= NH_2; (**amp-Amp**)
24, R_1= OH, R_2= NH_2; (**amp- Hyp**)
25, R_1= NH_2, R_2= OH; (**hyp-Amp**)
26, R_1= OH, R_2= OH; (**hyp-Hyp**)

Figure 4: Chimeric collagen peptides with Amp/Hyp at Y and anp/hyp at X positions

The chimeric peptides **23** (amp-Amp-Gly) and **24** (amp-Hyp-Gly) containing 4S-aminoproline at X position and 4R (OH/NH_2) at Y-position were found to form more stable triple-helix (ΔT_m 17^0C and 5.8^0C respectively at pH 3) compared to corresponding reference peptide with only one 4*S*-substituted proline at X-position model peptide Ac-Phe(amp-Pro-Gly)$_6$-NH_2 **22** (Table **2**).

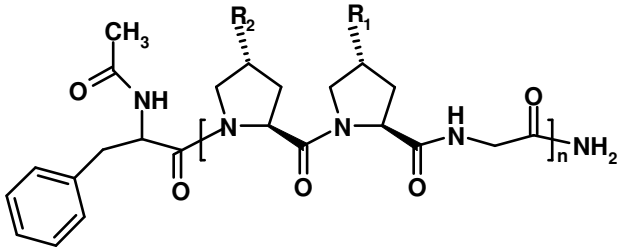
Table 2.Triple-helical thermal stabilities (T_m) of collagen peptides [in ^{o}C]

X-Y→	23 (amp-Amp)	24(amp- Hyp)	25 (hyp-Amp)	26 (hyp-Hyp)
pH 3	61	49.8	-nt-	-nt-
pH 7	46.6	39.3	-nt-	-nt-
pH 9	40.5	37	-nt-	-nt-
pH 12	34	-nt-	-nt-	-nt-

Peptides **25** (hyp-Amp-Gly) and **26** (hyp-Hyp-Gly) containing 4*S*-hydroxyproline at X position failed to show any thermal transition indicating the absence of triple-helix formation. This study reveals that 4*S*-aminoproline is well compatible at X position with either 4*R*-amino or 4*R*-hydroxyproline at Y-position, while 4*S*-hyroxyproline is not compatible.

***Part-B*: Dual compatibility of 4*R*-aminoproline: Hyperstable collagen peptides (X-Y-Gly) with modified iminoacids at both X and Y positions.**

First Part of this chapter demonstrated that 4*R* and 4*S* aminoprolines are compatible when present simultaneously at Y and X positions of collagen model peptide X-Y-Gly. The cause of the stabilization presumably arises from the C_4-endo and C_4-exo pyrrolidine conformation adopted by 4*S* and 4*R* aminoprolines, which are inherently favored at X and Y positions respectively. The formation of triple-helix by protonated 4*R*-aminoproline (Amp) when present in X position is surprising since its preferred C_4-exo pucker conformation is not favored at X position. Zagari *et al*. in their study of peptide (Hyp-Hyp-Gly)$_{10}$ showed that 4*R*-hydroxyproline is compatible at both X and Y positions to form a more stable triple helix. In view of this result, this chapter examines the role of different 4*R*-subtitute prolines in X and Y positions. (Fig 5).



27 R_1= R_2= NH_2. (**Amp-Amp**)
29 R_1 =OH, R_2=NH_2. (**Amp-Hyp**)
30 R_1= R_2= OH. (**Hyp-Hyp**)
28 R_1= NH_2, R_2= OH. (**Hyp-Amp**)

Figure 5: Chimeric collagen peptide with Amp/Hyp at both X and Y positions

Table: 3 Triple-helical thermal stabilities (T_m) of collagen peptides [in ^oC]

X-Y→	27 (Amp-Amp)	29 (Amp-Hyp)	30 (Hyp-Hyp)	28 (Hyp- Amp)
pH 3	51.5	43.4	31.8	55.8
pH 7	49.8	36	32	51.6
pH 9	25	29	30.9	31
pH 12	-nt-	-nt-	30	53

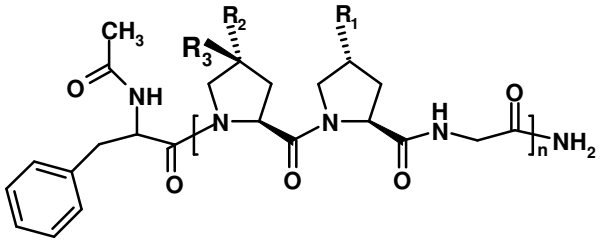
Triple-helix stability of the synthesized peptides (**27**-**30**) at different pH conditions were measured using temperature dependent CD spectroscopy (Table 3). The chimeric peptide Ac-Phe(Hyp-Hyp-Gly)$_6$-NH$_2$ **30** and Ac-Phe(Amp-Hyp-Gly)$_6$-NH$_2$ **29** were found to form stable triple-helix compared to their native model peptide Ac-Phe(Pro-Hyp-Gly)$_6$-NH$_2$. This result suggests that 4*R*-Amp has dual compatibility in stabilizing the collagen triple-helix when present in X or Y-positions, while 4*S*-amp tolerated only in X-site with Pro, Amp and Hyp in Y-site. The origin of triplex stabilization of Amp and amp seems to arise from a complex combination of stereoelectronic and electrostatic effects, with the latter dominantly contributing to hyperstability of 4-aminocollagen triple-helix.

Chapter 4: 4-(N-Formyl)-aminoproline collagen peptides: probes for studying the role of protonated 4-amino group in triplex stability

Though 4*R*-aminoproline is compatible at both X and Y position, the stability of the triple-helix of collagen peptide containing 4*R*-aminoproline is highly pH dependent. When present at Y position, 4*R*-Amp forms a stable triple helix in entire pH range (3-12), while at X position triplex is formed at only acidic pH. In comparison, 4*S*-amp also forms a stable triple helix only at acidic pH but no triplex at pH 12. These results suggest that the protonation of the NH$_2$ group 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 pyrolidine ring is dependent on the protonation status of the 4-amino group and also influences the *cis-trans* amide rotameric equilibrium. This chapter examines the possible effect of protonation of 4-NH$_2$ group on the formation and stability of collagen triple helix. The 4-NH$_2$ group of 4-aminoproline was protected as amide via formyl derivative, which cannot be protonated easily. The 4-formyl protection group is stable at

all pH conditions. Formyl group being small, any sterric interference of this group is only minimal during triplex formation with other strands. The 4-formyl protected 4-aminoproline, was synthesized from naturally occurring trans-4-hydroxy proline, and was utilized to synthesize the following peptides (Fig 6).

Triple-helix stability of above peptides was studied at varying pH condition using temperature dependent CD spectroscopy (Table 4). The triple-helical stability of 4N-formylproline peptides are pH independent, but less stable compared to 4-aminoproline peptide **19** (pro-Amp-Gly) (ΔT_m = 18, 13.7 and 5.5oC at pH 3, 7and 12) (Table-1). Further fAmp was not tolerated in X position with Pro in Y. This shows that protonation of amino group of 4-aminoproline plays an important role for stabilizing collagen triple heli˘



31 $R_2 = R_3 = H$, $R_1 = NHCHO$. (**Pro-fAmp**)
32 $R_3 = R_1 = H$, $R_2 = NHCHO$. (**fAmp-Pro**)
33 $R_1 = R_2 = H$, $R_3 = NHCHO$. (**famp-Pro**)
34 $R_2 = H$, $R_1 = NH_2$, $R_3 = NHCHO$. (**famp-Amp**)

Figure 6: Collagen peptides with 4N-formyl protected aminoproline

Table: 4 Triple-helical thermal stabilities (T_m) of collagen peptides [in ^oC]

X-Y→	31 (Pro-fAmp)	32 (fAmp-Pro)	33 (famp-Pro)	34 (famp-Amp)
pH 3	42	-nt-	23.2	58
pH 7	41	-nt-	21	56
pH 9	40	-nt-	21	24
pH 12	40.5	-nt-	20	44

Chapter 5: Prolyl-peptide bond isomerization in 4(*R/S*)- aminoproline: effect of pH on the peptidyl-prolyl bond isomerization in collagen peptides

Cis-trans (*E-Z*) isomerization of the amide bond (Fig 7) plays an important role in protein folding mechanisms. In most amino acids, *Z*-conformation is greatly favored over the *E* conformation. In proline- an imino acid, *Z*- conformation is only slightly favored over *E*- isomer. It was proposed that triple-helix stabilization arises form electron withdrawing inductive effect of 4-substitution on proline ring.

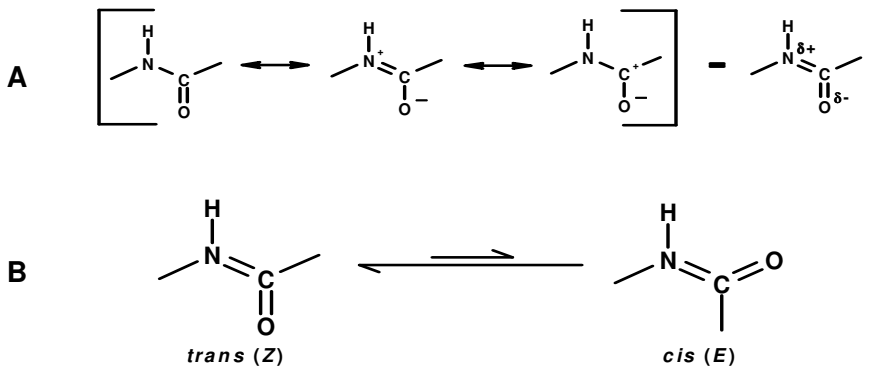


Figure 7: Resonance phenomenon in amide bon **A**; cis-trans isomerization of the peptide bond B; Longer backward arrow indicates that the trans isomer is greatly favored over cis isomer

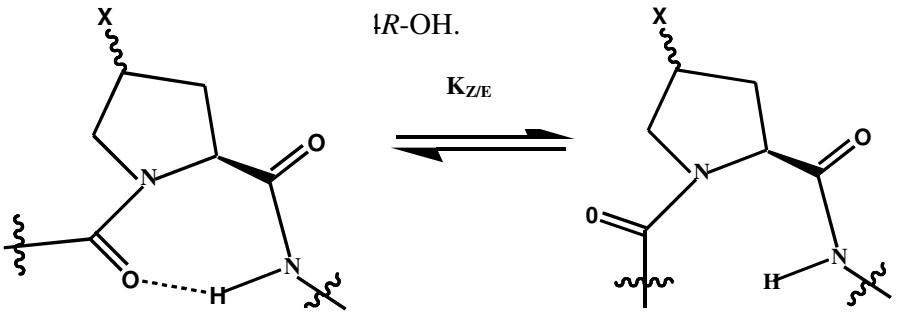
It was also argued that a **-I** (electronegative) group in the 4-position of proline increases the *Z*/*E* ratio (increases the equilibrium constant $K_{E/Z}$) of the peptidyl-prolyl bond, which pre-organizes the individual collagen chains in to an extended polyproline-II conformation, as required for the triple-helix formation. Studies on Ac-Xaa-OMe model compounds and collagen model peptides have shown that both $K_{E/Z}$ values and the triple-helix stability increases in the order Pro> Hyp> Flp. This is the same as the order of increasing electronegativity of the 4-substituents (H<OH<F).

The protonated amine group is more electronegative like F compared to the non protonated free amine. Hence it is possible to obtain the two conformations (*E* and *Z*)

from single compound aminoproline by just changing the pH condition. Since the stability of the aminoproline containing collagen peptides is highly pH dependent, it is also possible that the conformation of the pyrrolidine ring is also dependent on the protonation status of 4-amino group and influence the cis-trans amide rotomeric equilibrium. In the view of the results shown in the previous chapters, a study was undertaken to estimate the electronic, steric and stereochemical influence of the Amp and to decipher the contribution in the triple-helical stabilization in terms of intra-residue effects viz. *Z-E* isomerization and pyrrrolidine ring-pucker. This chapter describes the synthesis and study of the peptidyl-prolyl bond isomerization of 4*S* and 4*R*-aminoprolines in acidic and basic conditions. IR and variable temperature NMR spectroscopy have been used to estimate the effect of 4-aminogroup on the *Z-E* isomerization at protonated and nonprotonated forms. In addition, conformational analysis of the pyrrolidine ring puckers

in these compounds has been carried out using ^{1}H-^{1}H coupling constant data. The results have shown that the presence of 4*R*-NH_3^+ group on proline ring results in de

R-OH.



***trans* isomer** ***cis* isomer**

X = 4*R*-NH_2 **12**; 4*S*-NH_2 **13**; A*R*-N_3 **14**; 4*S*-N_3 **15**.

Table 5: $K_{Z/E}$ values and pyrroldine ring puckering of model compounds **12-15** at pH 2.0 and 12.0

	$K_{Z/E}$ 30 °C		**Ring puckering**	
compound	**pH 2.0**	**pH 12.0**	**pH 2.0**	**pH 12.0**
Ac-Amp-OMe **12**	6.89	6.06	C^{γ}-*endo*	C^{γ}-*exo*
Ac-amp-OMe **13**	2.63	1.33	C^{γ}-*endo*	-
Ac-Azp-OMe **14**	6.53	6.41	C^{γ}-*exo*	C^{γ}-*exo*
Ac-azp-OMe **15**	3.43	2.56	C^{γ}-*endo*	C^{γ}-*endo*

Abbreviations and Symbols

A	Adenine
aq.	aqueous
Ac	Acetyl
Ala	Alanine
AMBER	Assisted Model Building with Refinement
Amp	(2*S*,4*R*)-aminoproline
amp	(2*S*,4*S*)-aminoproline
Amp^+	(2*S*,4*R*)-aminoproline-protonated
amp^+	(2*S*,4*S*)-aminoproline-protonated
anhy.	Anhydrous
Arg	Arginine
Asp	Aspartic acid
Azp	(2*S*,4*R*)-azidoproline
azp	(2*S*,4*S*)-azidoproline
t-Boc	*tert*-butyloxycarbonyl
cm	Centimeter
Cbz	Benzyloxycarbonyl
CD	Circular Dichroism/dichroic
COSY	Correlation Spectroscopy
CT-DNA	Calf Thymus DNA
Cys	Cysteine
^0C	Celsius
DCM	Dichloromethane
DEAD	Dethylazodicarboxylate
DIPCDI	Diisopropylcarbodiimide
DIPEA	Diisopropylethylamine
DMF	N,N-dimethylformamide
DMSO	Dimethylsulfoxide
DCC	Dicyclohexyl carbodimide
E	*entgegen*

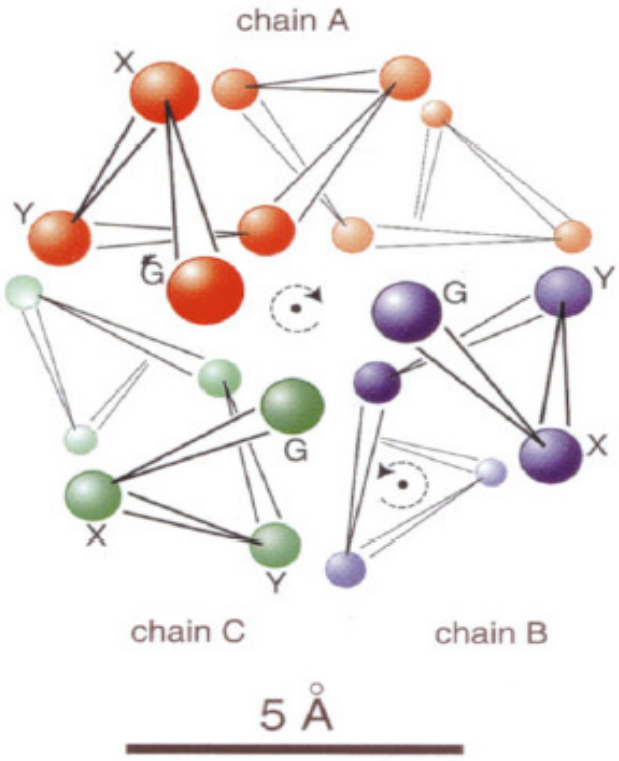
EB	Ethidium bromide
EG	ethyleneglycol
eq.	equivalents
FAB	Fast Atom Bombardment
Flp	(2*S*,4*R*)-fluoroproline
flp	(2*S*,4*S*)-fluoroproline
Fmoc	9-fluorenylmethyloxycarbonyl
FPLC	Fast Protein Liquid Chromatography
Gln	Glutamine
G	guanine
Glu	Glutamic acid
Gly	Glycine
HBTU	*O*-benzotriazol-1-yl-*N,N,N',N'*-tetramethyluronium hexafluorophosphate
His	Histidine
HOBt	1-hydroxybenzotriazole
HPLC	High Performance Liquid Chromatography
hrs	hours
Hyp	(2*S*,4*R*)-hydroxyporline (*trans*-4-hydroxy-L-proline)
hyp	(2*S*,4*S*)-hydroxyproline
Hz	Hertz
Ile	isoleucine
IR	Infra Red
K	Kelvin
$K_{Z/E}$	equilibrium constant for *Z*/*E* isomerization
Leu	Leucine
Lys	Lysine
M	molar
ma	major isomer
MALDI-TOF	Matrix Assisted Laser Desorption Ionization – Time of Flight
MD	Molecular Dynamics
MF	Merrifield Resin
mg	milligram
MHz	Megahertz

mi	minor isomer
µ	Micron
µM	micromolar
mL	milliliter
mM	millimolar
mm	millimeter
MMFF	Merck Molecular Force Field
mmol	millimoles
MS	Mass Spectrometry / Mass Spectrum
N	Normal
nm	nanometer
NMR	Nuclear Magnetic Resonance
NOESY	Nuclear Overhauser Effect Spectroscopy
Pro	Proline
ps	picosecond
p-TSA	*p*-toluenesulfonic acid
R	Rectus
RNA	Ribonucleic acid
RT	Room temperature
S	Sinister
Ser	Serine
SPPS	Solid Phase Peptide Synthesis
T	Thymine
TBTU	*O*-Benzotriazol-1-yl-N,N,N',N'-tetramethyluronium tetrafluoroborate
TEA	Triethylamine
TFA	Trifluoroacetic acid
TFMSA	Trifluoromethanesulfonic acid
THF	Tetrahydrofuran
Thr	Threonine
TMS	Tetramethylsilane
Trp	Tryptophan
UV	Ultraviolet

UV-Vis	Ultraviolet-Visible
Val	Valine
Z	*zusammen*

Chapter 1

Collagen structure and mimetics: An introduction



1.1: Introduction

Collagens are very abundant proteins in animal kingdom and are mainly located in the extracellular matrix.[1] The term 'Collagen' is derived from the Greek word for glue and was initially described as "that constituent of connective tissue, which yields Gelatin on boiling". About one quarter of all the protein in most animals is Collagen. It is the major constituent of all connective tissues in vertebrate as well as invertebrate animals,[2] performing in the connective tissues of animals some-what the same function as cellulose molecules in plants. Skin, tendon, bone cartilage, cornea, and teeth all contain collagen fibrils. 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.

The molecular hallmarks of collagen proteins are the multiple repetitions of Gly-X-Y amino acid sequences and unique triple-helical structure built by three polypeptide chains. Each collagen polypeptide chain contains more than 1000 amino acid residues. Depending up on the amino acid composition and the supramolecular assembly, collagen families are classified into different subfamilies. Upto now 42 different polypeptide chains have been identified, which are encoded by 41 specific genes and compose 27 unique collagen types.[3] Among these, Collagen I-III are the most abundant and form fibrils of similar structure.

The primary structure of collagen is composed of approximately 300 repeats of the trimer Gly-X-Y where X and Y can be any amino acid. The most commonly found amino acids in X and Y positions are Proline (Pro) and ***4R***- hydroxyproline (Hyp). They account for nearly 20% and glycine accounts for 30% of total amino acid content of

natural collagen. The other commonly found amino acids being Ala, Arg, Leu, Lys, Ser, Thr, and Val.[4]

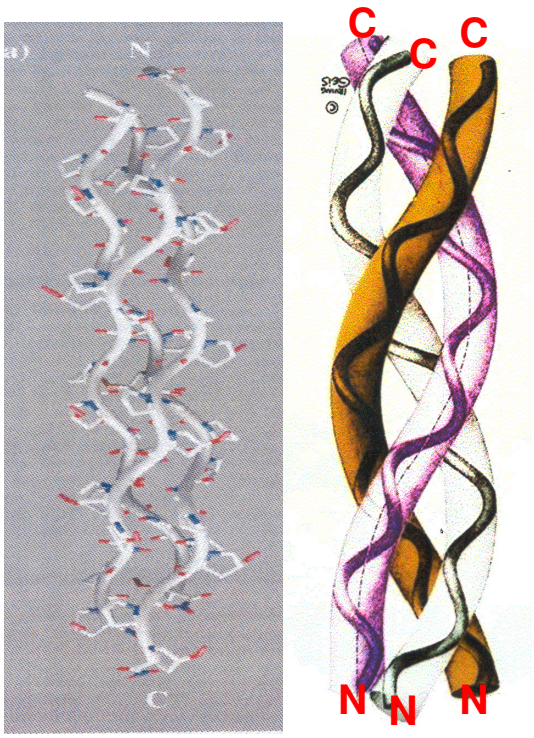


Figure 1. Structure of collagen

The secondary structure of collagen is an extended left-handed helix polypeptide chain called polyproline-II conformation, with all peptide bonds in *trans* conformation.[5] In the tertiary structure of collagen, three polyproline-II helices are intertwined in a parallel fashion with one residue shift to form a right-handed super helix called "collagen triple-helix"[6](Fig 1).

1.2: Collagen super family

The collagen superfamily is highly complex and shows a remarkable diversity in molecular and supra molecular organization, tissue distribution and function. The number of triplet repeats ranges from a few dozen to 510 depending on collagen type.[7] Some

types of collagens are homotrimeric proteins that consist of three identical α-chains, such as type-II and III, while others are heterotrimers of different α-chains. For examples, type-I collagen consist of two α1(I) chains and one α2 (I) chain. The collagen proteins are classified into the following subfamilies.[8]

1) Fibril-forming collagens. (Type, II, III, V, XI), and I

2) Network-forming collagens. (Type, IV, VIII, and X)

3) Fibril-associated collagens with interrupted triple-helices. (FACITs type, IX, XII, XIV, XVI, and XIX,)

4) Transmembrane collagens. (Type XIII, and XVII).

1.3: Biosynthesis of collagen

The biosynthesis of collagen is well studied.[9] Collagen is synthesized in endoplasmic reticulum (ER) as procollagen (Fig 2), which is the precursor protein that bears propeptide domains at either end of the triple helical domain. The process by which procollagen is synthesized in the lumen of the ER includes unique steps. First each polypeptide chain of procollagen (Pro-α-chains) finds its correct patterns, which enables the formation of the distinct types of procollagen. Second, triple-helix formation of long X-Y-Gly, repeats begins with C-propeptide trimerization, [the C-propeptides of Pro- α -chains are non-Collagenous domains containing approximately 250 amino acid residues] (Fig 3)

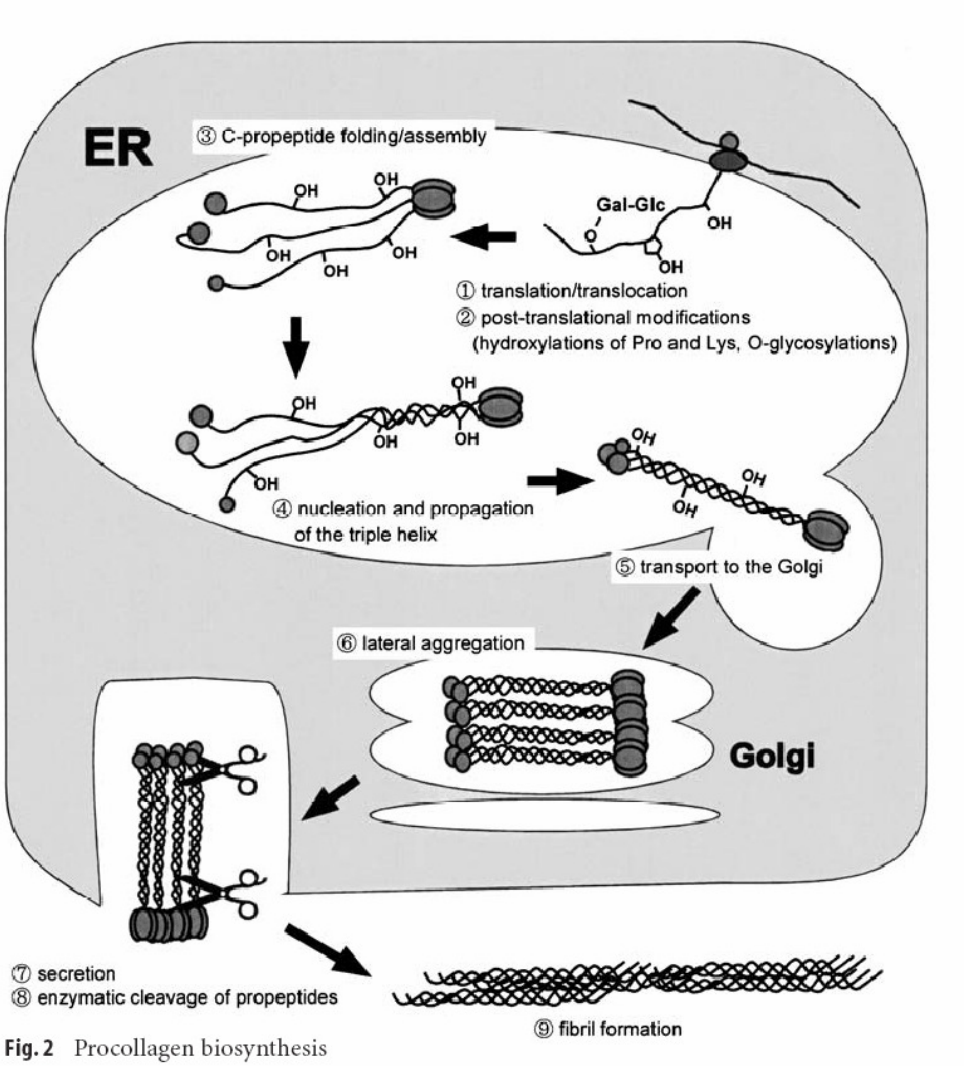


Fig. 2 Procollagen biosynthesis

which result in the formation of a correctly aligned triple-helix and thereby prevents mis- staggering. The most characteristic step is the formation of triple-helix. This step involves specific post-transnational modifications, in particular the prolyl-4 hydroxylation of Y-position amino acids through the action of enzymes *Prolyl 4-hydroxyase* and *lysyl hydoxyase* (Fig 4 & Fig 5). The polypeptide chains are then secreted into extra cellular matrix, where they fold into a triple-helix.[10]

Figure 3

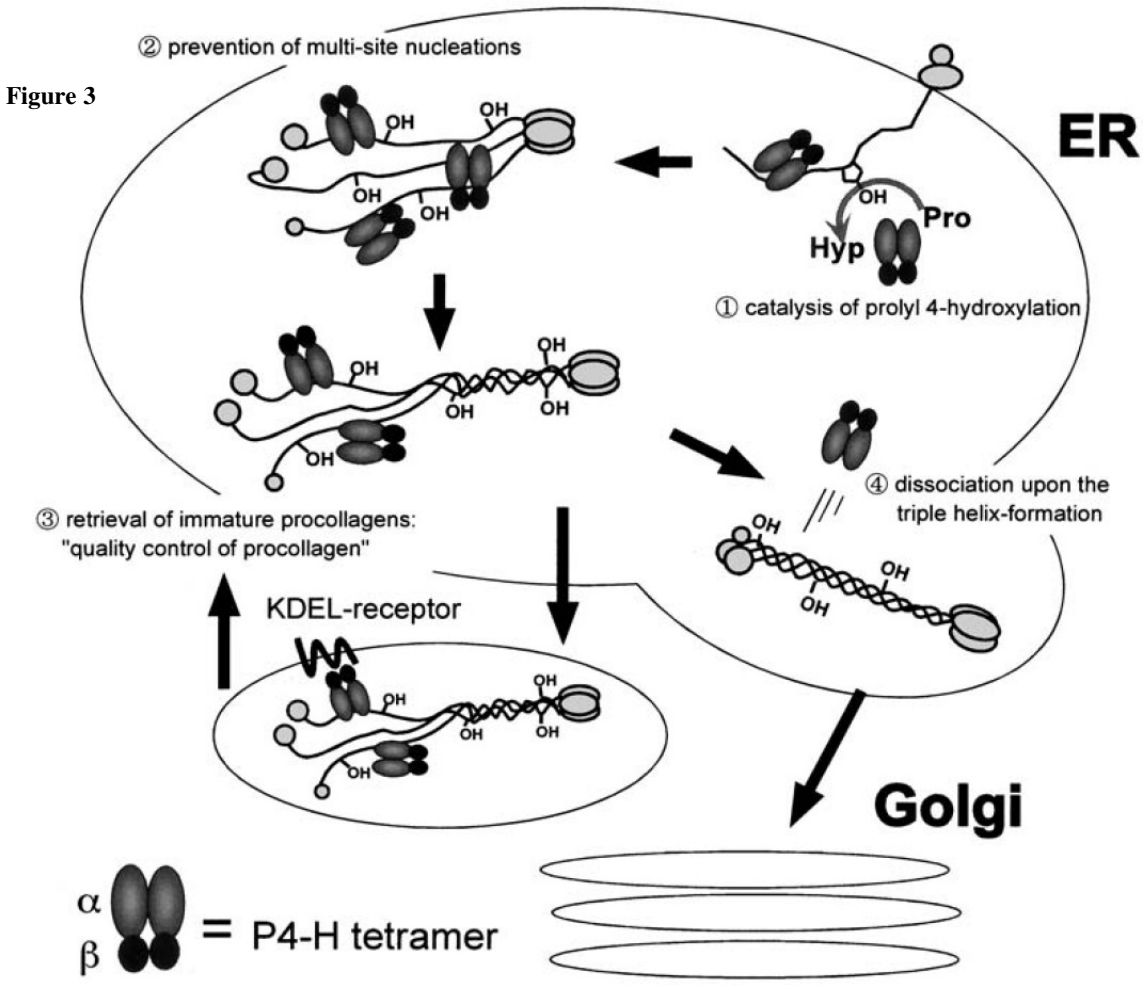


Figure 4

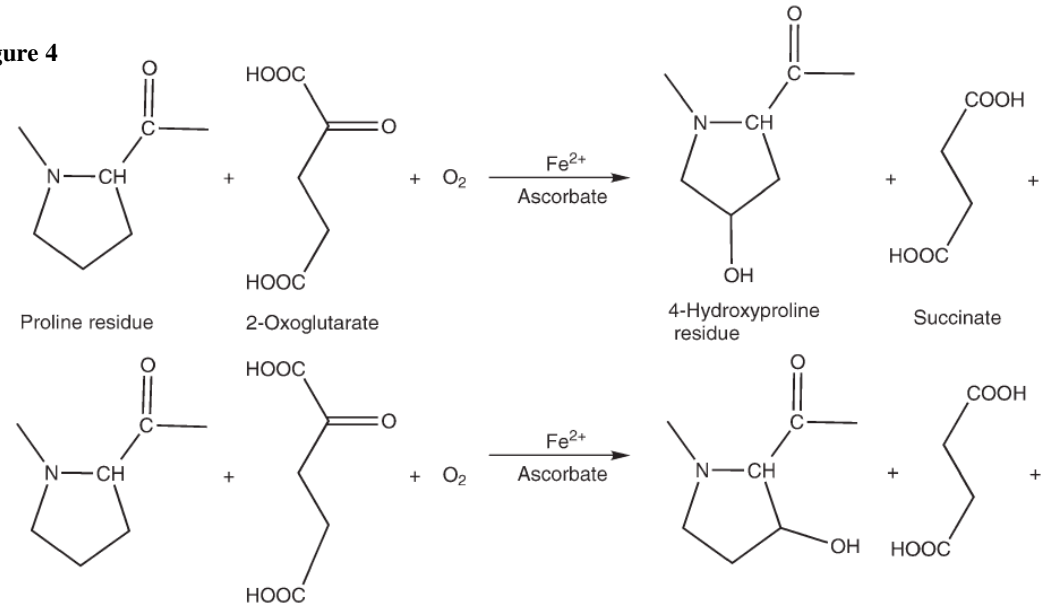


Figure 3. Possible role of P4-H as an procollagen-specific molecular chaperone.[10]
Figure 4. Reaction catalyzed by collagen prolyl 4-hydroxylase, and 3-hydroxylase.[10]

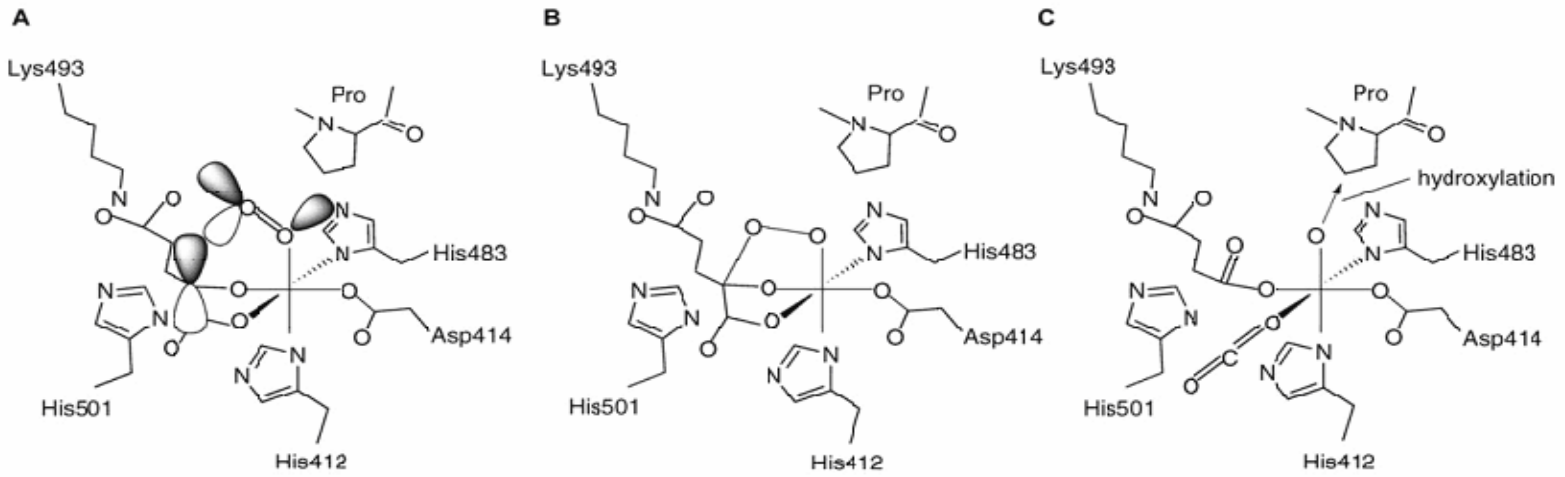


Fig.5 Schematic representation of the first half-reaction of collagen prolyl 4-hydroxylase and the critical residues at the catalytic site of the human α(I) subunit. Fe^{2+} is coordinated with the enzyme by His412, Asp414 and His483. Subsite I of the 2-oxoglutarate binding site consists of Lys493, which ionically binds the C5 carboxyl group of 2-oxoglutarate, while subsite II consists of two *cis* -positioned equatorial coordination sites of the enzyme-bound Fe^{2+} and is chelated by the C1 carboxyl and C2 oxo functions of the 2-oxoglutarate. Molecular oxygen is bound end-on in an axial position, producing a dioxygen unit. (A) One of the electron-rich orbitals of the dioxygen is directed to the electron-depleted orbital at C2 of the 2-oxoglutarate bound to the iron. (B) A nucleophilic attack on C2 generates a tetrahedral intermediate, with loss of the double bond in the dioxygen unit and of the double-bond characteristics of the oxo-acid moiety. (C) Elimination of CO_2 coincides with the formation of succinate and a ferryl ion, which hydroxylates a proline residue in the peptide substrate in the second half-reaction. His501 is an additional important residue, probably being involved in both coordination of the C1 carboxyl group of 2-oxoglutarate with Fe^{2+} and cleavage of the tetrahedral ferryl intermediate.

Taken from ref 9

1.4. Positional preferences of amino acids in collagen

Though 400 different X-Y-Gly triplets are possible for 20 naturally occurring amino acids in X and Y position, only a limited number are actually found in collagen sequences.[11] For instant, in the $(X\text{-}Y\text{-}Gly)_{338}$ of α1 chain of type-I collagen, only 22 triplets occurs four or more times, with the commonly found triplets being Pro-Hyp-Gly (12%), Gly-Pro-Ala (9%) and Gly-Ala-Hyp (6%), indicating a positional preference for Hyp in Y position of the triplet, and Leu, Phe being found almost exclusively at X position.[12] Modeling studies suggest potential interactions and steric factors as an explanation for these positional preferences.[13,14] Several attempts have been made to measure the stability of collagen model peptides with various amino acids and the triple-helical propensities of amino acid.[15] These studies were complicated by small difference in thermal stability among the peptides due to the presence of charges at free *N* and *C*-termini. In a more recent approach, a set of host triple-helical peptides with the sequence $Ac\text{-}(Pro\text{-}Hyp\text{-}Gly)_3\text{-}XYG\text{-}(Pro\text{-}Hyp\text{-}Gly)_4\text{-}Gly\text{-}Gly\text{-}NH_2$ where XYG is the guest triplet, have been used to elucidate positional preference of various amino acid commonly found in natural collagen. A triple-helix propensity scale has been proposed for all the most commonly occurring amino acids in fibril-forming collagens. The host peptide $Ac\text{-}(Pro\text{-}Hyp\text{-}Gly)_8\text{-}Gly\text{-}Gly\text{-}CONH_2$ was used to characterize the triple-helix forming ability of all the amino acids.[16] T_m measurements and thermodynamic analyses of the host sequence $Ac\text{-}(Pro\text{-}Hyp\text{-}Gly)_3\text{-}X\text{-}Hyp\text{-}Gly\text{-}(Pro\text{-}Hyp\text{-}Gly)_4\text{-}Gly\text{-}Gly\text{-}CONH_2$ have revealed a striking relationship between the nature of the amino acid and the triple-helix stability. Proline (Pro) was found to be the most stabilizing residue (T_m =47.3 ^{0}C) in X position and also offered the lowest enthalpic

Table-1:♣ Melting temperature and enthalpies of host-guest peptides Ac-(Gly-Pro-Hyp)$_3$-GXY-(Gly-Pro-Hyp)$_4$-Gly-Gly-NH_2 with all common amino acids in the X position, together with their frequency of occurrence in fibril forming collagen.

Gly-X-Hyp	T_m (^{0}C)	ΔH (KJ/mol)	Occurrence (%)
Pro	47.3	435	32.9
Glu	42.9	590	13.0
Ala	41.7	480	11.1
Lys	41.5	540	3.6
Arg	40.6	520	2.8
Gln	40.4	565	2.9
Asp	40.1	520	4.9
Leu	39.0	437	7.8
Val	38.9	518	2.6
Met	38.6	452	0.9
Ile	38.4	624	2.0
Asn	38.3	502	2.1
Ser	38.0	506	4.9
His	36.5	580	1.6
Thr	36.2	506	1.8
Cys	36.1	423	0.0
Tyr	34.3	629	0.5
Phe	33.5	514	3.0
Gly	33.2	575	1.6
Trp	31.9	593	0.0

Table-2:♣ Melting temperature and enthalpies of host-guest peptides Ac-(Gly-Pro-Hyp)$_3$-GXY-(Gly-Pro-Hyp)$_4$-Gly-Gly-NH_2 with all common amino acids in the Y position, together with their frequency in fibril forming collagen.

Gly-X-Hyp	T_m (^{0}C)	ΔH (KJ/mol)	Occurrence (%)
Hyp	47.3	435	34.0
Arg	47.2	610	11.4
Met	42.6	436	9.0
Ile	41.5	559	2.1
Gln	41.3	559	6.9
Ala	40.9	502	10.6
Val	40.0	481	4.3
Glu	39.7	630	2.0
Thr	39.7	647	4.2
Cys	37.7	471	0.0
Lys	36.8	400	9.0
His	35.7	497	0.5
Ser	35.0	435	4.0
Asp	34.0	776	4.8
Gly	32.7	665	0.7
Leu	31.7	514	1.7
Asn	30.3	640	2.1
Tyr	30.2	657	0.0
Phe	28.3	557	0.2
Trp	26.1	670	0.0

♣ Taken from reference 16

contribution (ΔH^0 = 435 KJ mol^{-1}) in the entire series. The host peptides with charged amino acids (Glu, Lys, Arg and Asp) were among the most stable peptides. All these residues possessed higher enthalpic contribution than Pro, which reflects side chain interactions with available backbone carbonyls or solvent. Most destabilizing residues were found to be aromatic amino acids together with Glycine (Gly). Trp was found to be the most destabilizing amino acid in X position (T_m =31.9 ^{0}C). A similar set of 20 peptides containing host triplet Pro-Y-Gly was used to determine the triple helical propensities of 20 amino acids and Hyp. Hyp/Arg was among the most stabilizing amino acids with nearly equal T_m values of 47.3 ^{0}C and 47.2 ^{0}C respectively. Further, Arg peptide showed a higher ΔH^0 of denaturation than Hyp (610 KJ mol^{-1} and 435KJ mol^{-1} respectively), suggesting a different mechanism for stabilization by Arg. The aromatic amino acids Try, Phe and Trp were the most destabilizing (T_m values 30.2 ^{0}C, 28.3 ^{0}C and 26.1 ^{0}C respectively). Thermodynamic analysis has shown that both Hyp in Y and Pro in X position have the same enthalpy contribution (ΔH^0 = 435 KJ mol^{-1}). It is also interesting to see that the most destabilizing amino acid Trp has the highest enthalpy value (ΔH^0 = 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 correlates well with triple-helix stability.

1.5: Three-dimensional structure of collagen

Due to the more complex nature of the fibrous proteins, more guesses were made on the structural models of collagen than any other protein. In 1954, Ramachandran and Kartha[17] provided first correct model for the structure of collagen.

According to their model (Fig 6A), collagen consists of three parallel left-handed polypeptide chains, which are related by 3-fold symmetry and held together by interchain hydrogen bonds. Each chain had 3_2 symmetry, staggered with respect to one another by about $3A^0$ and positioned in such a fashion that the residue in the interior of the triple-helix could only be glycine, facilitating juxtaposition of three polypeptide chains in close proximity. To explain the X-ray diffraction pattern of stretched collagen fibers, Ramachandran and Kartha[17a] modified their original model by introducing a coiling of the three chains around a common axis visualizing the structure of collagen to be a coiled-coil structure.[18] This retained all the essential features of the original model of the location of amino acid residue, the orientation of N-H and C=O bonds in a sequence (X-Y-Gly), with (i) Gly N-H, hydrogen bonded to C=O of the Y-residue and (ii) N-H group of X-residue when it was not imino acid hydrogen bonded with C=O of the neighboring chains.

Rich and Crick[19] disagreed with this proposal, and found through calculation that only one interchain hydrogen bond Gly-NH….O=C (X), could be formed and the second interchain hydrogen bond disallowed short contact distances. Later, Ramachandran and Chandrasekharan suggested that a second interchain hydrogen bond involving NH group of X position could be mediated by a water molecule and the possibility of interchain CH(Gly)….O=C bonds as in polyglycine-II also in collagen. Finally, Rich and Crick modified this model to include only one hydrogen bond per every trimer repeat. According to this model (Fig 6B) each of the strands in triple-helix has left-handed coiled-coil conformation derived form polyproline-II and polyglycine-II hybrid to form a right-handed triple-helix.

This may be designated as a single left-handed helix with 10/3 screw symmetry (3 1/3 residue per turn) having a fiber-repeating period of $28.6A^0$.

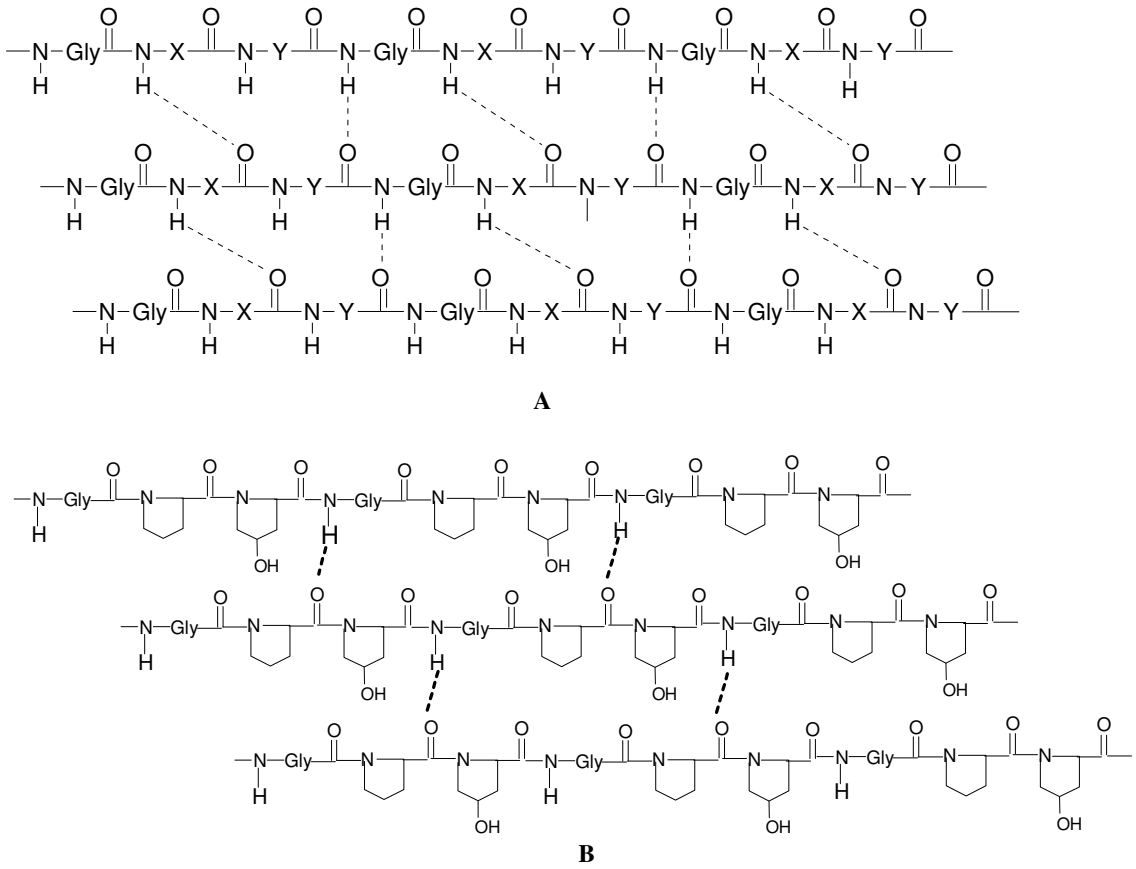


Figure 6: Schematic view of **A)** Rmachandran and Karuna model involving two hydrogen bonds per trimer repeat when the X residue is not an iminoacid. **B)** Rich and Crick model with one hydrogen bond per trimer repeat (Pro- Hyp-Gly). This model is valid both for an amino and imino acid in the X position.

1.6: Use of Synthetic polypeptide models

The fibrous nature of collagen has prevented the growth of single crystals for high-resolution studies. After 1960, this limit has been overcome by using collagen-like synthetic polypeptides among them, the X-ray diffraction pattern analysis of poly-(Pro-Gly-Pro) by Yonath and Traub[20] and single crystal X-ray diffraction analysis of (Pro-Pro-Gly)$_{10}$ peptide[21] may be considered (Fig 8). These studies lead Okuyama[22] *et al*. to propose a slightly different structure of collagen triple-helix.

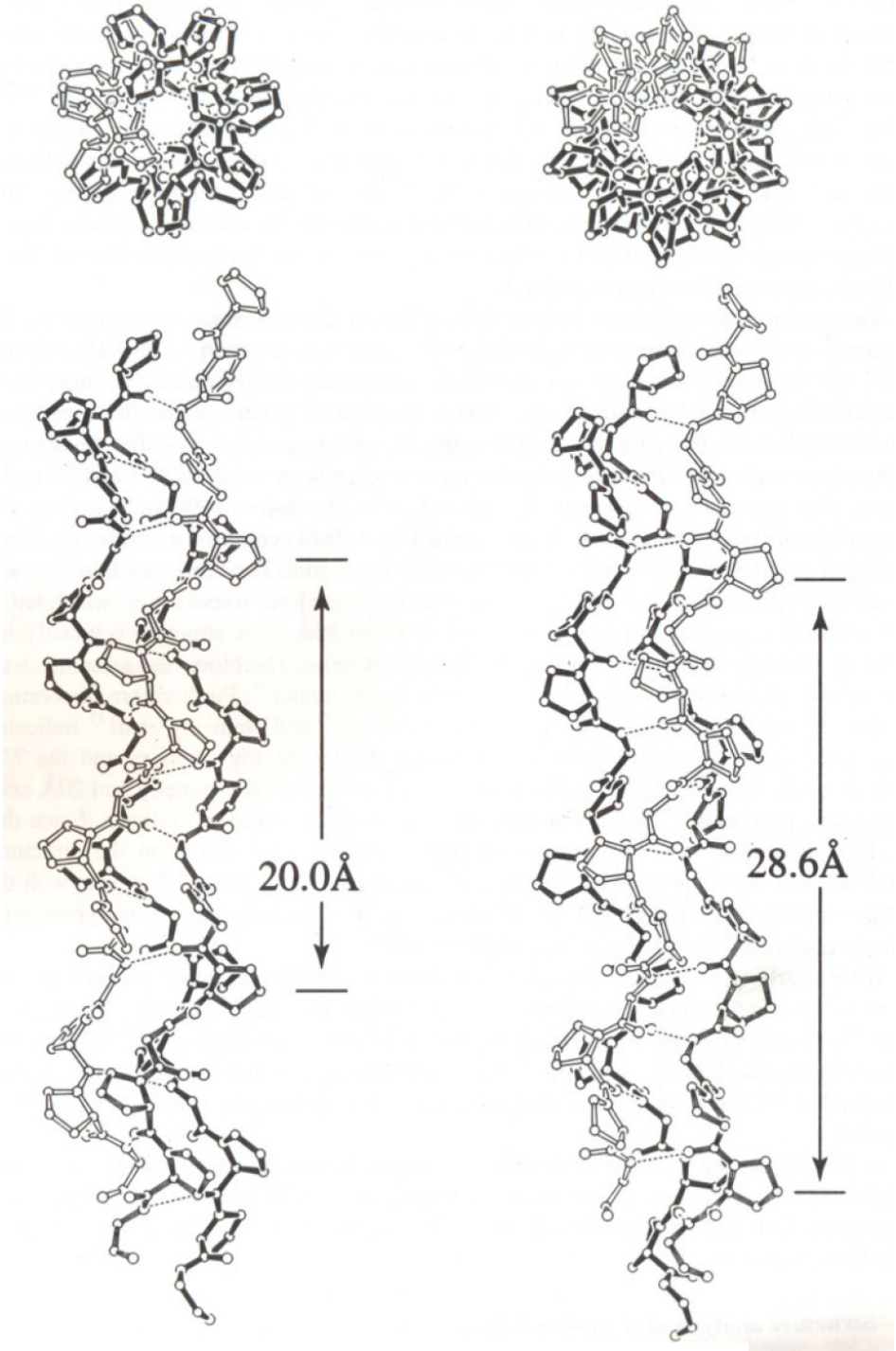


Figure 7: Two proposed helical structures for collagen.[19]

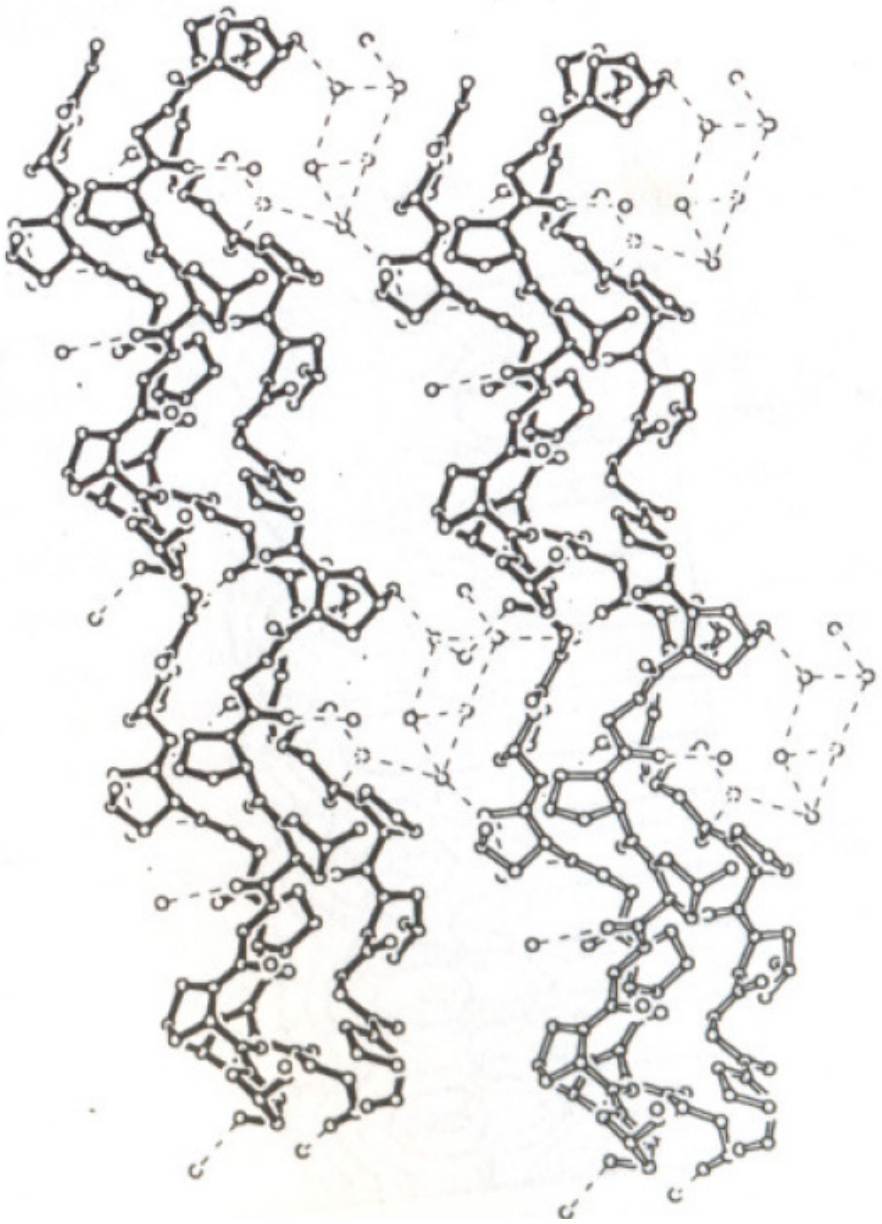


Figure 8. The molecular structure of (Pro-Pro-Gly)$_{10}$ together with interaction between triple-helices *Via* Water.[17a]

According to the Okuyama model, the molecular structure of collagen consist three peptide strands in which each strand has 7/1 helical symmetry and pitch length of $60A^0$ (Fig 7B). This may be designated as a single left-handed 7/2 helix with a fiber-repeating period of $20A^0$ leading to a more compact structure compare to Rich and Crick model Later several single X-ray crystal structures of synthetic collagen model peptides of type $(Pro\text{-}Pro\text{-}Gly)_n$ and $(Pro\text{-}Hyp\text{-}Gly)_n$ have been reported.[23] These structures have conformed the Okuyama model for the collagen.

It may however be noted that the Rich and Crick, and Okuyama model differ only in terms of helical parameter and they are not qualitatively different from each other with respect to the hydrogen bonding pattern. It appears that various structural variations in the triple-helix are within the main framework of triple-helical structure proposed for collagen. Recently, Karmer[24] *et al*. have determined the crystal structure of collagen like peptide $(Pro\text{-}Hyp\text{-}Gly)_3$-Ile-Thr-Gly-Ala-Arg-Gly-Leu-Ala-Gly-Pro-Hyp-Gly-$(Pro\text{-}Hyp\text{-}Gly)_3$ in which a 12 amino acid sequence 785-796 of human type-III collagen is flanked by two $(Pro\text{-}Hyp\text{-}Gly)_3$ peptides, has revealed subtle variations in the triple helical parameters. For example imino acid rich terminal regions showed 7-fold symmetry (Okuyama model) and the central region showed a 10-fold symmetry (Rich and Crick model) indicating the sequence dependence of triple-helical parameter. Collagen triple-helix seems to accommodate such minor structural variations of symmetry imposed by different amino acid sequence.

1.7: Crystal packing of collagen polypeptide models.

In the analysis of a X-ray structures, crystal packing is often considered as an undesired complication, since it can alter the molecular structure. By contrast, collagen

like polypeptides can provide interesting information on the interactions occurring among the triple-helices. Indeed there are indications that, in the crystal state, triple-helices pack in a similar manner as they do in collagen fibrils. The crystal packing of the polypeptide models thus far studied presents a number of analogies. The lateral packing of these models is either *pseudo-tetragonal*, as in the case of polypeptide (Pro-Pro-Gly)$_{10}$, or *Quasi-exagonal*, as for Gly→Ala, EKG and POG (Table 2). In consistent with the most recent model of the collagen assembly,[25] collagen-like polypeptides containing Hyp mainly pack *Quas*i-exagonally. Furthermore, all of the known structures of Hyp containing polypeptides exhibit direct hydrogen bonding interactions between Hyp hydroxyl groups of adjacent triple-helices,[26,27] except for the structure of Gly→Ala.[28] In this case, the lack of direct intermolecular hydrogen bonding interactions may be attributed to the loss of axial triple-helix repetition, which is due to the existence of a bulge in the center of the molecule.[28] The above results have suggested that Hyp-Hyp interactions may play a role in lateral assembly of triple-helices in collagen fibrils.[29, 26,27]

The difficulties to obtain ordered crystal of collagen-like polypeptides along with the limited number of conformational variables, prompted a number of theoretical calculations on simple sequences first[30,31] and subsequently on real collagen sequences.[32,33] The determination of single crystal X-ray structure of the collagen-like model with sequence (Pro-Hyp-Gly)$_4$-Pro-Hyp-Ala-(Pro-Hyp-Gly)$_5$ has been a real breakthrough.[34b] Indeed, not only has this provided the first high-resolution structure of a collagen triple-helix, but also shown the structural effects of a disease causative mutation. In particular, this structure has confirmed the main features of the Okuyama model.

Table2. Available crystal structure of collagen-like polypeptides [34a-f]

Polypeptide model	Sequence	Resolution (A^0)	PDP code
PPG (average model)	$(PPG)_{10}$	2.2	-
Gly→Ala	$(POG)_4POA(POG)_5$	1.9	1cgd
PPG (average model)	$(PPG)_{10}$	1.6	1a3j
PPG (average model)	$(PPG)_{10}$	1.97	1a3i
PPG (average model)	$(PPG)_{10}$	1.9	-
POG (average model)	$(POG)_{10}$	1.9	-
T3-785	$(POG)_3$ ITGARGLAGPOG $(POG)_3$	2.0	1kbv
EKG	$(POG)_4EKG(POG)_5$	1.75	1qsu
POG (average model)	$(POG)_{10}$	1.4	-
PPG (average model)	$(PPG)_{10}$	1.3	1g9w
$[(PPG)_{10}]_3$ (full length)	$(PPG)_{10}$	1.3	1k6f

Furthermore, it has evidenced that local untwisting of the triple-helix occurs in the region of the Gly→Ala mutation, where interstitial water molecules mediate the characteristic hydrogen bonds of collagen triple-helices. The structure of Gly→Ala also

showed that the hydroxyl group of Hyp residues is extensively involved in hydrogen bonding network with water molecules. This finding supported the idea[35] that water molecules play a major role in the stabilization of the triple-helix.

1.8: Role of trans-4-hydroxyproline in collagen structure.

Collagen triple-helices exhibit high thermal stability, which is important to the great tensile strength of its fibers. The thermal stability of triple-helical collagen is enhanced by the hydroxylation of the proline residues in the Y position of the X-Y-Gly triplets. Proline hydroxylation occurs as a major post–transnational event, as more than 100 prolines are hydroxylated in each collagen polypeptide chain, which consists of about 1000 residues. In this context, it is worth noting that content of hydroxyprolines in the collagen sequence clearly correlates with the upper limit of the environmental temperature in which a particular organism lives. During biosynthesis, proline residues are incorporated both in the X and the Y positions. However, only those in the Y positions are hydroxylated in vertebrate collagen to produce the specific diastereoisomer 4*R*-hydroxyproline (Hyp). By contrast, prolines both in X and Y positions may be hydroxylated to 3-hydroxyproline in some basement membrane collagens. A large amount of experiments have evidenced that the proline hydroxylation occurring in Y position has a solid impact on triple-helix stability.[36,37] On the other hand, polypeptides, having repeating sequence Hyp-Pro-Gly, with 4*R*-Hyp in X position, do not associate in triple-helix.[38]

Table 3. Melting temperatures of the triple-helices

Polypeptide model	T_m (K)	Reference
[(Pro-Pro-Gly) $_{10}$] $_3$	314	42
[(Pro-Flp-Gly) $_{10}$] $_3$	364	42
[(Pro-4R-Hyp-Gly) $_{10}$] $_3$	342	42
[(4R-Hyp-Pro-Gly) $_{10}$] $_3$	n.d	38
[(Pro-4S-Hyp-Gly) $_{10}$] $_3$	n.d	39
[(4S-Hyp-Pro-Gly) $_{10}$] $_3$	n.d	39
[(4R-Hyp-Thr-Gly) $_{10}$] $_3$	291.2	43
[(4R-Hyp-Thr-Gly (βGal)) $_{10}$] $_3$	232.3	43
[(Pro-Thr-Gly (βGal)$_{)\ 10}$] $_3$	312.0	43

n.d indicates not determined T_m, since polypeptides do not fold in triple-helix

Furthermore it has been shown that triple-helix folding is also inhibited by the presence of the diastereoisomer 4*S*-hydroxyproline (4*S*-Hyp)[39] in either X or Y position. Therefore, as shown in Table 3, the impact of the proline hydroxylation depends on the sequence position and on the diastereoisomer produced.

To explain the stabilizing role of 4*R*-Hyp in Y position several hypotheses have been put forward since early 1950s, when the triple-helix structure was not known. At that time extra stabilization was intuitively attributed to direct interaction between hydroxylgroup of Hyp.[40] However when the triple-helix structure appeared, this hypothesis was discarded. More recently, such enhanced triple-helical stability has been

attributed to the extensive hydration induced by the presence of Hyp.[35] As said above, the existence of many water sites was demonstrated by the crystal structure of the collagen triple-helix model.[41] In particular the extra stability was attributed to the formation of water-mediated hydrogen bonds between the hydroxyl group of 4*R*-Hyp in Y position and the carbonyl group of adjacent chain. However, this stabilization model was weakened by the observation that the substitution of 4*R*-Hyp with 4*R*-Fluoro-L-proline (Flp) in collagen like polypeptides enhanced the thermal stability of the triple-helix, despite very low tendency of fluorine to form hydrogen bonds. Therefore an alternative explanation was provided, that the hyperstability was thought to arise from the strong inductive effects of the most electronegative atom, fluorine, which imposes C_4-*exo* pucker on the pyrrolidine ring to attain the required *trans* conformation for prolyl-peptide bond through ring NH-C_5-C_4-X gauche effect.

4*R*-Hyp residue probably in the same way as fluoroproline stabilizes the collagen triple-helix by means of a stereoelectronic effect which fixes the pyrrolidine ring pucker and thus preorganize all three torsion angles.[42] More recent data have shown that, when compared to Pro, 4*S*-Hyp favours the *cis* state of the peptide bond. This finding is in accordance with the observation that 4S-Hyp destabilizes the triple-helix both in X and Y position.[43] However, both the hydration hypothesis and the *cis-trans* equilibrium hypothesis failed to explain the triple-helix destabilization induced by 4*R*-Hyp in X position. Therefore both stereospecificity and position dependence of prolyl-hydroxylation are explained by another mechanism, based on the conformational preferences of iminoacids[30].

1.9: Conformation of polypeptide chains.

Pauling et al[44] analyzed the geometry and dimensions of peptide bonds in the crystal structures of molecule containing either one or few peptide bonds. It should be noted that the C-N bond length of the peptide bond is 10% shorter than that found in usual C-N amine bonds[45] (Fig 9).

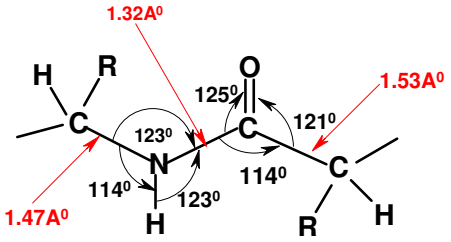


Figure 9: Bond angles and bond length of peptide chain.

This is because the peptide bond has some double bond character (40%) due to resonance which occurs with amides (Fig 10A). As a consequence of this resonance, all peptide bonds in protein structures are found to be almost planar.

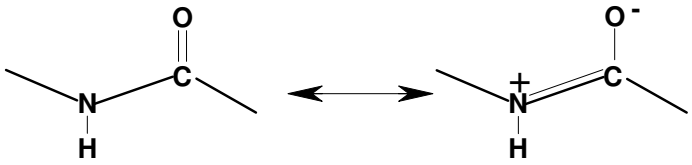


Figure 10A. Peptide bond isomerization

This rigidity of the peptide bond reduces the degrees of freedom of the polypeptide during folding. The peptide bond nearly always has the *trans* configuration since it is more favorable than *cis,* which is found to occur with proline residues.[46] Steric hindrance between the functional groups attached to the Cα atoms will be greater in the *cis* conformation (Fig 10B).

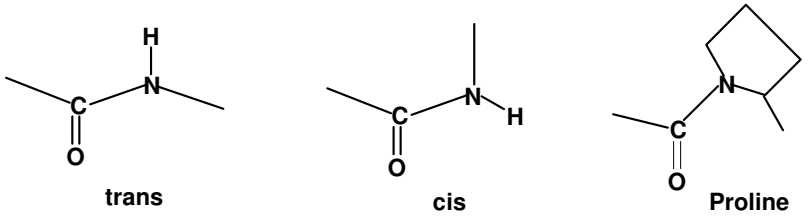


Figure10B: Two conformational existence of peptide bond

The entire conformation of the protein can be described by angles of internal rotation. Each set of four successive atoms in the main chain defines a dihedral angle. for each residue i (except for the N and C termini) the angle Φ_i is the angle defined by the atoms C_{i-1}-N- C_α-C, and the angle ψ_i is the angle defined by atoms N- C_α- C-N_{i+1}. ω_i is the angle around the peptide bond itself, defined by the atoms $C_{\alpha(i)}$-C-N_{i+1}- $C_{\alpha(i+1)}$. ω is restricted to 180^0(*trans*) or 0^0(*cis*) with slight deviation in some cases.

Table 4. Main-chain Torsion angles for various conformations in peptides of L-amino acids.[a,b]

Φ(deg)	**Rotation about N-C^α (to N-H)**	**Ψ(deg)**	**Rotation about C^α-C (to C-O)**
0	C^α-C *trans*	0	C^α-N *trans*
+60	C^α-H *cis*	+60	C^α-R *cis*
+120	C^α-R *trans*	+120	C^α-H *trans*
+180	C^α-C *cis*	+180	C^α-N *cis*
-120	C^α-H *trans*	-120	C^α-R *trans*
-60	C^α-R *cis*	-60	C^α-H *cis*

a. *trans* to N_i-H_i is the same as *cis* to N_i-C_{i-1}; *trans* to C_i-O_i is the same as *cis* to C_i-N_{i+1}.

b. For the description of D-amino acids, inter change C^α-H and C^α-R in the table.

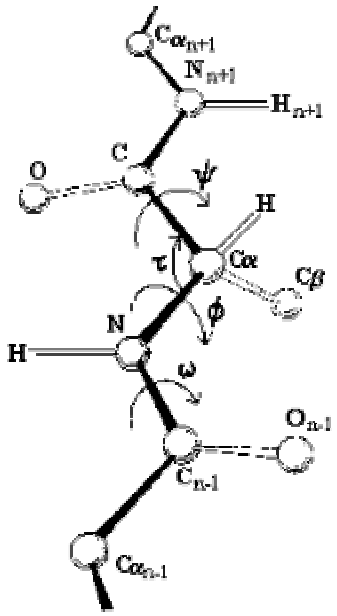


Figure 11: Perspective drawing of a segment of polypeptide chain comprising three peptide units. The recommended notations for atoms and bond angles are indicated. The polypeptide chain is shown in the fully extended conformation.

Only two torsional angles, Φ and ψ determine the backbone conformation of an amino acid residue in a peptide chain (Fig 11). By extension, these angles determine the overall backbone conformation of an entire protein. The conformation adopted by a particular residue in a protein structure reflects a combination of environmental factors, including secondary and tertiary interactions, but also intrinsic residual preferences for Φ, ψ angle conformations.[47,48] These intrinsic preferences are of particular interest as they may give insight into initial ordering of a peptide or protein backbone during the first steps along the protein folding pathway.

Statistical studies of high-resolution structures in the PDB have been widely used to examine the amino acid secondary structure propensities[49-52] with a view to improve protein structure prediction.[53-55] A novel approach to deconvolute the various competing factors which determine the intrinsic Φ, ψ, propensities of amino acids lies in

the analysis of residues in coil regions of protein structures, that is, residues that are not in alpha-helices and beta-strands.[48] From the analysis of residues located in regions of regular secondary structure, intrinsic conformational preferences can be identified which are free from the regular interactions associated with beta-sheets and alpha-helices. The striking conclusion from such an analysis is that the distribution of backbone Φ and ψ angles within the coil regions is far from random. The most populated regions of the Ramachandran plot corresponds with the range of angles normally associated with beta-sheet and alpha-helical secondary structure.

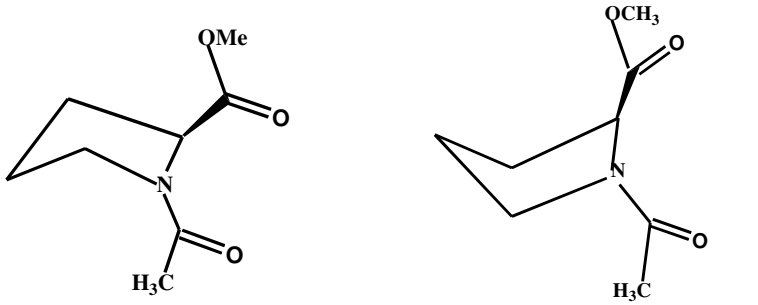
Table 5. Approximate Torsion angels for some regular structure.[49]

Regular structures	Φ (deg)	ψ (deg)	ω (deg)	Reference
Right handed α-helix (α-Poly(L-alanine))	-57	-47	+180	56
Left-Handed α-helix	+57	+47	+180	56
Parallel-chain pleated sheet	-119	+113	+180	57
Antiparallel-chain pleated sheet (β-poly(L-alanine))	-139	+135	-178	58
Polyglycine II	-80	+150	+180	59
Collagen	-51, -76, -45	+153, +127, +148	+180	60
Poly (L-proline) I	-83	+158	0	61
Poly (L-proline) II	-78	+149	+180	62

1.10: Pyrrolidine ring conformation in proline and substituted proline

Proline is a cyclic amino acid and the bridging of the α- carbon atom to the main chain nitrogen by 3 methylene bridge imposes further constraint on the main chain tortion 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 (Fig 12). These ring puckers are also expressed in terms of *endocyclic* torsional angles. In proline, both puckers are almost equally preferred and the energy barrier to inter-conversion is very low.[63]



γ- endo (down) pucker **γ- exo (up) pucker**

Figure 12: Pyrrolidine ring conformations

However in 4-substituted prolines, depending on the steric and electronic effects exerted by the 4- substitutent, pyrrolidine ring may prefer any one of the ring-puckers. For example *trans*-4-hydroxy-proline shows almost exclusively γ-*exo* pucker, as found in its crystal structure and in solution.[63] This pucker preference has been attributed to the phenomenon of hydroxyl-amide gauche effect.[64,65]

1.11 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 in *gauche* orientation with respect to each other rather than *anti*" orientation.[66] This effect is surprising since, a combination of dipole repulsion and steric effect between the electronegative atoms in vicinal substituents are expected to make them remain anti to each other (Fig 13). Many molecules containing N, O, P, S, F or Cl show a preference for *gauche* conformation. For example in difluoroethane the p orbital character of the both C-F bonds is increased due to the large electronegativity of fluorine, as a result electon density builds up above

and below to the left and right of the central C-C bond. The resulting reduced orbital overlap can be partially compensated when a gauche conformation is assumed, forming a bent bond. This was evidenced by increased -C-C-F- bond angles (by 3.2°) due to steric repulsion between the fluorine atoms and increased –F-C-C-F dihedral angles (from the default 60° to 71°).

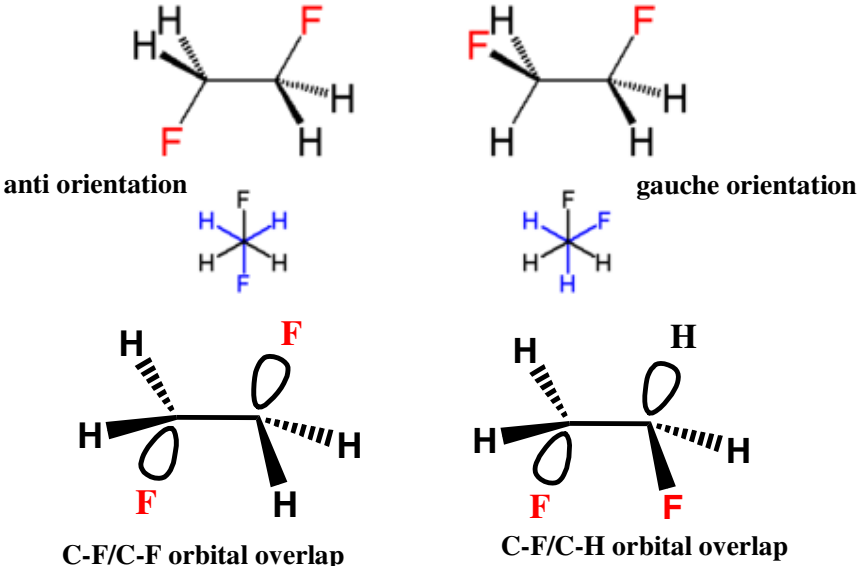


Figure 13: gauche effect in 1,2-difluoro ethane molecule.

The origin of the gauche effect is not very clear and *ab initio* quantum chemical calculations underestimate the *gauche* effect. σ –Hyper-conjugation[67] and bent-bonds[68] have been proposed to explain the phenomenon of the *gauche* effect.

In 4-substituted prolines, the steric repulsion between the amide-ring nitrogen and 4-substituent should result in a pseudo-equatorial position of the 4-sudstituents i.e., *anti* to the ring amide-nitrogen. For example, 4R-hydroxyproline may be expected to exhibit *anti* orientation, however X-ray crystal structure, and ^{1}H-^{1}H coupling constant analysis conform the axial orientation and resulting γ-*exo* ring pucker.[69] This suggests that gauche effect may be a dominating factor in determining the ring-pucker preference for proline with 4-electronegative substituents.

1.12 Effect of 4-substituents on the pyrrolidine ring pucker

The insertion of a substituent at the C_γ position has strong influence on the conformation of five-membered pyrrolidine ring. The preferred conformation is one in which there are maximum number of stabilizing interactions are occurred. The 4*R*-substitution (4*R*-X), on the C_γ position leads to a strong preference for a C_γ-exo (up) conformation of the pyrrolidine ring.[64] Because in the C_γ-*exo* conformation, the 4*R*-X group is in axial position (Fig. 14A), hence gauche orientation of ring nitrogen and 4*R*-X groups with respect to C_γ-C_δ bond axis is possible. Morevover the bonding $\sigma(C_\beta\text{-}H)_{ax}$ and $\sigma(C_\delta\text{-}H)_{ax}$ orbitals donate electrons to $\sigma^*(C\gamma\text{-}O)$ antibonding orbital through hypercongujation effect.[65]

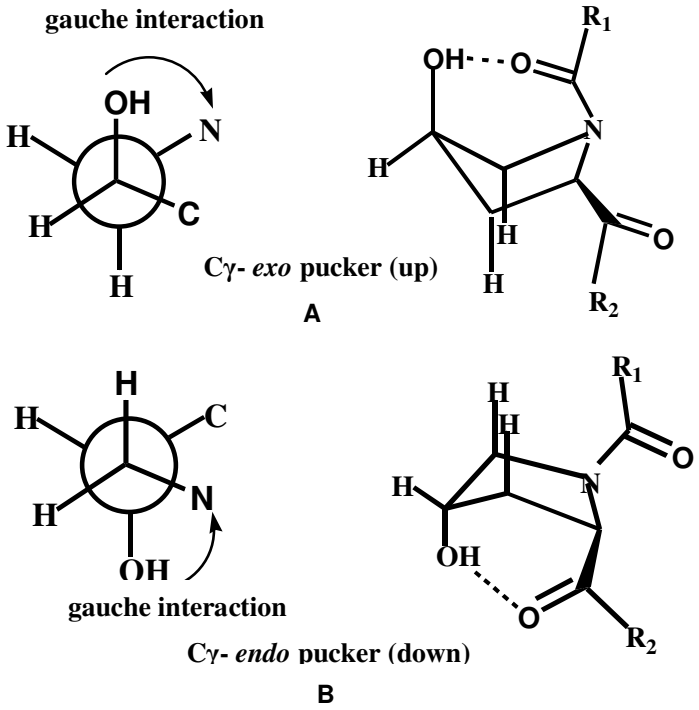


Figure 14: Effect of **A**; 4*R*- and **B**; 4*S*-substitutents on the pyrrolidine ring puckering preferences

This stereoelectronic effect was confirmed by the observed shorter bond distance of (C_β-C_γ) and (C_γ- C_δ) bonds (1.512 and 1.510 Å respectively) and the relatively long

C-O bond (1.425 Å) in Ac-Hyp-OMe. The C_γ-*endo* preferences of 4*R*-substituent does not allow for such stabilizing effects.

Similarly the 4*S*-substitution leads to a preference of C_γ-*endo* conformation for the pyrrolidine ring. In the C_γ-*endo* conformation, the 4*S*-substituent is in axial position which provides gauche interaction between ring nitrogen and 4S-substituent and hyperconjugative interactions between bonding $\sigma(C_\beta\text{-}H)_{ax}$ and $\sigma(C_\delta\text{-}H)_{ax}$ orbitals and $\sigma^*(C\gamma\text{-}O)$ antingbonding orbit (Fig 14B), like in case of 4*R*-substituent. Raines et.al. demonstrated that electron-withdrawing groups in the 4-position inductively withdrawn electron density from the peptide bond, as a result increasing N-pyramidalization, reducing the bond order of the C-N linkage, and thereby facilitating the pyrrolidine ring pucker which is lower in energy.

1.13: Effect of 4-substituents on E-Z isomerization of prolyl-peptide bond

As a secondary amine, proline has a much greater propensity than other amino acids to form *cis* (that is, *E*) peptide bonds (Fig. 15).[70-72a]

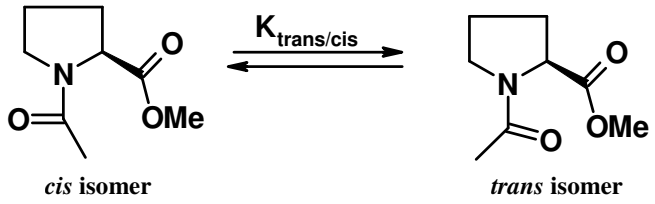


Figure 15: cis/tarns isomerization of peptide bond in proline

Proline and hydroxyproline residues are constrained by their pyrrolidine rings and this rigidity stabilizes triple-helical structure of collagen. Yet, the *trans* and *cis* conformations of the peptide bonds in proline and 4-substituted proline residues are nearly equal energy, which is expected to destabilize the triple-helix structure because all the peptide bonds in triple-helical collagen are in the *trans* conformation, even

though 4*R*-substituted prolines increase the conformational stability of collagen. This indicates that 4*R*-substituent not only alters the pyrrolidine ring preferences, also it influences the $K_{trans/cis}$ of peptide bond (Fig 16).

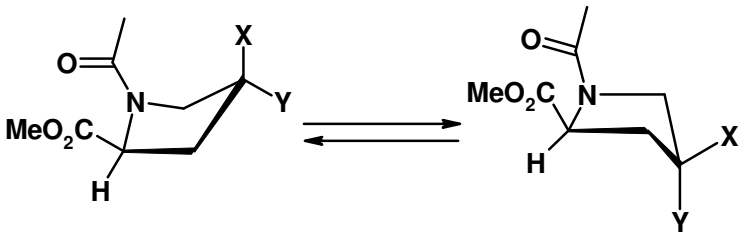


Figure16: Ring puckers in 4-substituted Ac-Pro-OMe. C^{γ}-endo pucker is favored when X= H, OH, or F and Y= H. C^{γ}-exo pucker is favored when X= H, and Y= OH or F.

The *cis* and *trans* peptide bonds are mainly affected by the angle φ. The *cis* peptide bond is characterized by more –ve value of φ angle than the *trans* peptide bond. The angle φ also correlates with the pyrrolidine ring pucker.[72b] The *exo* ring pucker has smaller ψ and a less negative φ values than the *endo* pucker. Minimization of steric strain between the two carbonyl groups and the C_{δ}-C_{γ} and C_{δ}-N bonds provides a rationale for these correlations.

In addition to affecting the ring pucker, the stereochemistry of a 4-substituent on the proline residue affects $K_{trans/cis}$ value of prolyl-peptide bond. The stabilization of trans isomer of 4*R*-substituent relative to simple proline arises from the effect of γ-exo ring pucker on the interaction between the two carbonyl groups (Fig 17A), as expected from a more optimal n→π* interaction (Bürgi-Dunitz trajectory).[72c-e] Because the exo ring pucker allows the lone pair of the amide oxygen (i-1) in to the π* orbital of the amide carbonyl group (i). In case of 4S-substituent, which prefers the γ-endo ring pucker, both the less optimal n→π* interaction and additional destabilizing interactions (steric strain between the two carbonyl groups and the C_{δ}-C_{γ} and C_{δ}-N bonds (Fig 17B)) are responsible for the decrease in the trans/cis ratio.

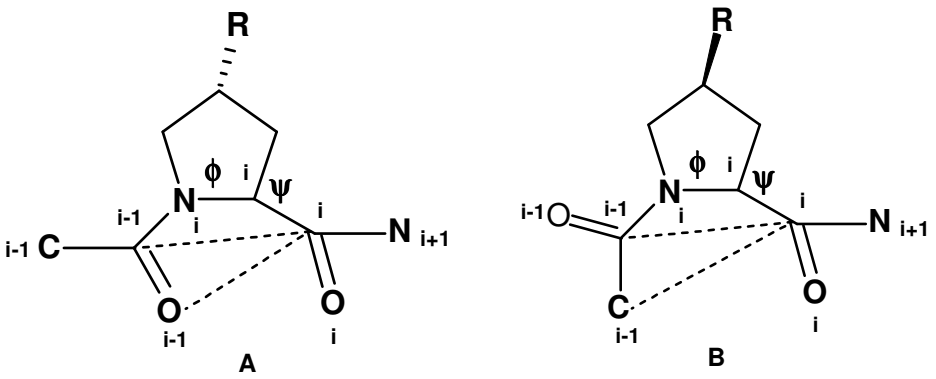


Figure 17: Schematic representation of nonbonding back bone interactions occurring in **A**; 4R and **B**; 4S- substituted proline.

Thus the pucker of the pyrolidine ring in a γ-substituted proline residue, its φ and ψ torsion angles, and its peptide bond trans/cis ratio are interdependent parameters. The interplay between these parameters is responsible for the contribution of a γ-substituted proline residue to the conformational stability of the collagen triple-helix.

A variety of methods have been developed to control the *trans/cis* ratio and pyrrolidine ring puckering preferences of proline, including buttressing the 2,[73] 3,[74] and 5-positions[75-78] with functional groups, replacing the prolyl peptide bond with an alkene isostere,[79-84] and including the amide in a ring system that is fused to the pyrrolidine ring.[77] These approaches enable torsional control of the amide bond but introduce steric bulk that could be undesirable. The pyrrolidine ring of proline exists in a variety of puckers, with $\mathbf{C}_\gamma$ being its most nonplanar constituents.[85,86] More recently, a study of peptidyl-prolyl bond isomerization by Raines *et al.*[87] has shown that the *Z* to *E* ratio ($K_{Z/E}$) increases with the increase in the electronegativity of the 4-substituents. The 4*R*-substituents increase stability of the *trans* isomer, whereas the 4*S*-substituents decrease that stability (Table-6). The negative inductive effect transmitted through the C^{γ}-C^{δ} bond to the amide bond, results in an increase of the Z isomer[87]. Such increase was observed in the model compound of the type Ac-Xaa-OMe[88] where Xaa is Proline,

Hydroxyproline, and Fluoroproline. In the crystal structure of these model compounds a significant decrease in the C^{γ}-C^{δ} bond length from proline to fluoroproline through hydroxyproline, was also observed. This decrease in the bond length and the concomitant decrease in steric repulsion between the C^{α} of the i+1 residue was suggested as the mechanism of Z isomer stabilization.[89]

Table 6. Values of $K_{trans/cis}$ for 4-substituted AcXaaOMe[a]

Xaa	X	Y	$K_{trans/cis}$
Flp	H	F	6.7
Hyp	H	OH	6.1
Pro	H	H	4.6
hyp	OH	H	2.4
flp	F	H	2.5

[a] values were measured in D_2O at 25°C by integration of ^1H NMR spectra (Taken from ref.87)

Based on this observation, Raines et. al. explained that the peptide $(Pro.Flp.Gly)_{10}$ formed more stable triple-helix than $(Pro.Hyp.Gly)_{10}$ which was in turn more stable than triple-helix of $(Pro.Pro.Gly)_{10}$. Fluorine in organofluorines C-F bond is non-hydrogen bonding and thus its stabilizing effect on the triple-helix possibly cannot be attributed to hydrogen bonding. It was argued that, due to unfavorable entropy loss involved in maintaining the hydrogen bonded networks, 4*R*-Hyp residue stabilizes the collagen triple-helix more by virtue of its inductive effect and not by hydrogen bonding. Incidentally (2*S*,4*S*)-hydroxyproline (hyp) containing peptides $(hyp.Pro.Gly)_{10}$ and $(Pro.hyp.Gly)_{10}$ do not form triple-helices. A similar destabilizing effect by (2*S*,4S)-fluoroproline (flp) in Y position was also observed. Thus the destabilizing of the triple-helix was attributed to low $K_{E/Z}$ values of amide bond.

1.14: 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.[90] The rod-like structure of the triple-helix allows large ratio of charged to hydrophobic residues than seen in globular proteins.[91] 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. There is an excess of basic residues over acidic ones, giving collagen an isoelectric point in the basic pH range. An asymmetric distribution of charged residues is found along the $(X\text{-}Y\text{-}Gly)_n$ 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.[92] The distribution of charged residues is highly conserved among the members of the fibril forming family.[93]

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.[94] 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 neighboring collagen molecules associated in periodic fibrils.[95] Charged residues have also been implicated in the binding of collagen to other extra cellular matrix components such as cell-surface integrins.

1.15: 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.*[96] 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 value (<3) or high pH value (>9), ie, fully protonated or fully non-protonated form. Venugopal et al.[97] have studied the pH dependence of triple-helical strength of (Pro-Hyp-Gly)$_{10}$ which was greatest at acidic and basic pH values, with T_m= 61^{0}C at pH 1.0 and T_m= 62^{0}C at pH 11.0. The triple-helical stability decreased around neutral pH with a minimum T_m of 56^{0}C at pH 6.0. This indicated that (Pro-Pro-Gly)$_n$ model peptides form parallel triple-helices 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.

1.16: Effect of ionizable side-chains

Electrostatic interaction in triple-helix due to ionizable side chains has been studied by Venugopal *et al.*[97] by introducing a zwitterionic 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-

$(Pro\text{-}Hyp\text{-}Gly)_3$ corresponding to the 18 residues sequence of type III collagen with C-terminal Gly-Pro-Hyp tail. Variation of pH in the range 1-13 lead to 8-9 ^{o}C changes in the T_m of EK containing peptide and the T3-487 peptide. The greatest triple-helical stability was seen at pH values where both acidic and basic residues are ionized.

Recent NMR study on CB2 peptide 94-39 amino acid fragment of α 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.[98] 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.

1.17: Characterization of triple-helical structure

An outstanding characteristic of the collagen fiber is the sharp contraction to about one quarter of its original length on heating.[99] This phenomenon was extensively investigated in the late 60's and the basis of the structural and kinetic relationships were well-established. Collagen triple-helix upon heating undergoes a helix↔coil transition resulting in change of various spectroscopic and physical properties. These include change in molar extinction coefficient, molar ellipticity, specific rotation and integrations of NMR signals of associated and dissociated forms. The thermal denaturation of associated triple-helical peptides obtained by monitoring one of these properties shows a sharp cooperative transition. The midpoint of such a transition (T_m) is used as a measure of triple-helical strengths of collagenous peptides.

A variety of experimental techniques have been used to detect and characterize triple-helical structures and these include infrared-vibrational spectroscopy,[100] Raman spectroscopy,[101] and two-dimensional NMR spectroscopy.[102] Techniques such as X-ray

diffraction and electron microscopy have been employed to study collagen fibrils and synthetic collagen-like polymers.[103] While X-ray crystallographic experiments have been used to illustrate triple helical packing of collagen structures in solid state, intrinsic viscosity,[106] equilibrium sedimentation and light scattering techniques have been employed to investigate formation and denaturation of triple helices in solution.[104,105] The susceptibility of the unfolded chains to enzymes compared to the helix or antibodies specific to the native configuration have also been used to follow denaturation. Kinetics of denaturation can also be studied using differential scanning calorimetry (DSC),[107] wherein both insoluble fibers and the corresponding molecules in solution can be studied. On heating in a calorimeter, the molecules undergo an abrupt helix to coil transition giving rise to a sharp peak at a defined temperature. The position, height, width, area and symmetry of the thermogram peak provide valuable information about the denaturation process. The area under peak gives the total heat involved in the transition and is known as the enthalpy of denaturation.

CD spectroscopy is frequently and extensively used to study the triple helicity of collagen mimetic structures.[108] Natural collagen has a unique CD spectrum where a small positive peak appears at 220nm, a crossover at 213nm and a large trough at 197nm. Polyproline II helices also posess similar absorbances in the CD spectrum as collagen triple-helices (Fig 18). However, collagen exhibits a cooperative melting transition with increasing temperature, whereas polyproline II helices do not[109].

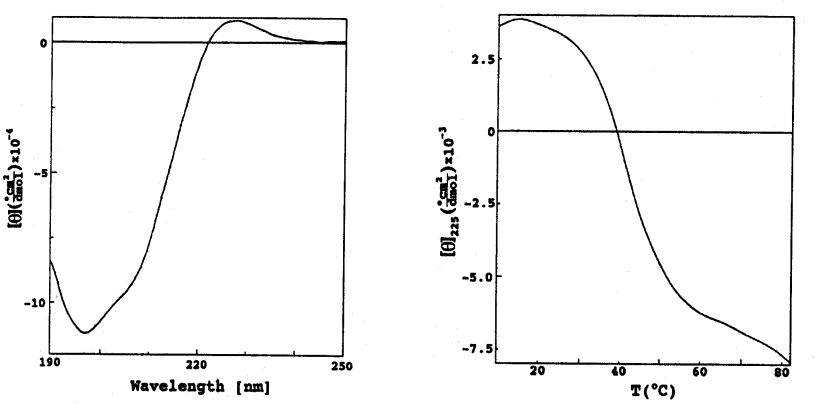


Figure 18: CD spectra at 10°C and triple-helix to coil transition curve for $(Pro.Hyp.Gly)_{10}$ in water

Recently, NMR spectroscopy has become an effective technique in detecting Triple-helical structures.[110,111] The NMR spectrum for a triple-helical structure will present a new set of resonances in the ID spectra, which is absent in the NMR spectra of non-triple helical structures. The triple-helical formation causes a slow proton exchange in D_2O to give observable amide proton NH signals in the NMR spectrum and gives rise to unique interchain nuclear Overhauser effects (NOEs).[109]

1.18: Collagen mimetics

The development of collagen research has expanded to elucidate the interaction involved in stabilization and folding of collagen-like structures and to design alternatives to natural collagen-based biomaterials. 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.

The development of solid phase chemistry has facilitated the synthesis of collagen-like peptides of varying sequences and chain lengths. Sakakibara et. al.

synthesized $(Pro\text{-}Pro\text{-}Gly)_{10}$ and $(Pro\text{-}Pro\text{-}Gly)_{20}$. According to optical rotation dispersion experiments these collagen mimetic peptides formed stable triple helices.[112,113] Other collagen peptide models have been synthesized and assessed for triple helicity, including $(Gly\text{-}Pro\text{-}Leu)_{10}$,[114] $(Gly\text{-}Pro\text{-}Pro)_5\text{-}(Gly\text{-}Pro\text{-}Ala)_n\text{-}(Gly\text{-}Pro\text{-}Pro)_5$,[115] $(Pro\text{-}Hyp\text{-}Gly)_{10}$,[116,117] and $(Hyp\text{-}Pro\text{-}Gly)_{10}$.[118] All these peptides except for $(Hyp\text{-}Pro\text{-}Gly)_{10}$ formed stable triple helical structures. It is interesting that $(Hyp\text{-}Pro\text{-}Gly)_{10}$ does not exhibit cooperative melting transition, whereas $(Pro\text{-}Hyp\text{-}Gly)_{10}$ forms a stable triple-helix in water. These studies show that the stabilizing effect of Hyp is observed when it is located in the Y position of X-Y-Gly trimer unit.

1.19: Collagen mimetics with unnatural aminoacids

In attempts to understand and modify the triple-helical strength of collagen, several unnatural aminoacids have been incorporated into the X and Y positions of collagen sequence (Fig. 19). To understand the importance of stereo-configuration of 4-hydroxylgroup 2*S*,4*S*-hydroxyproline (hyp) containing peptides $(Pro.hyp.Gly)_{10}$, $(hyp\text{-}Pro.Gly)_{10}$ 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.[119] Synthesis of sequential polypeptides poly-(Ala.Gly.Thz), Poly-(Gly.Pipec.Ala) and poly-(Gly.Aze.Ala) (Fig. 19) have been reported.[120] Peptide $(Gly.Pro.MePro)_n$, $(Pro.Gly.Flp)_nPro$ (n = 1- 4) were synthesized and tested for the *Proline-hydroxylase* inhibition activity.[121] However, the conformational properties of these peptides are not known. Later work on poly-(Gly.Pro.Aze), poly-(Gly.Aze.Pro) and poly-(Gly.Aze.Aze) has shown that poly-(Gly-Pro-Aze) exhibits triple-helical character according to CD and vibrational spectroscopy. Molecular modeling studies

have conformed the triple-helicity of poly(Gly-Pro-Aze). The result demonstrates that Aze, a homologue of Proline, can be made into triple-helix, although its stability is much lower than that of poly(Pro-Pro-Gly).

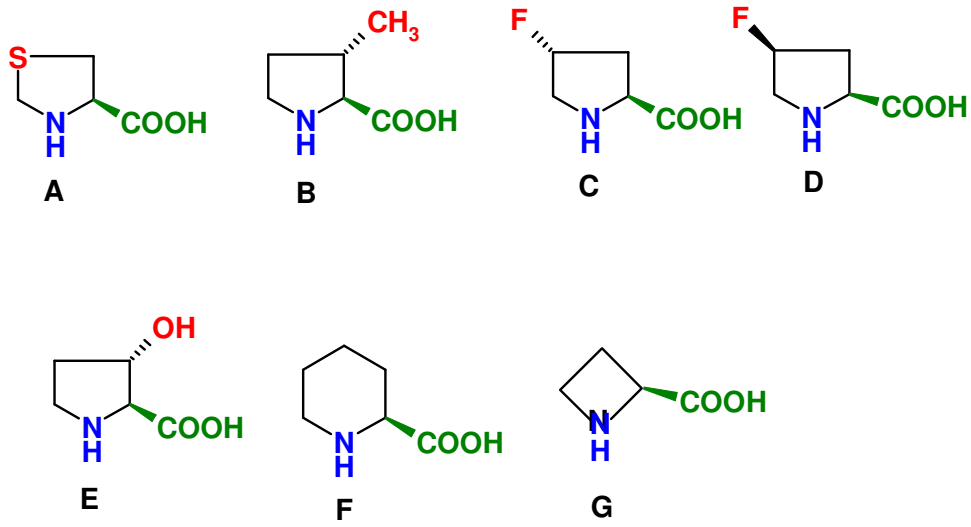


Figure 19: Various proline surrogates that have been incorporated in to collagen sequences. **A** Thiozilidine (Thz), **B** *trans*-3-Methylproline (MePro), **C** *trans*-4-Fluoroproline (Flp), **D** *cis*-4-Fluroproline (flp), **E** *trans*-3-Hydroxyproline, **F** L-piperidine-2-carboxylic acid (Pipec), **G** Azitidine (Aze).

Several recent studies have evaluated the propensity of different aminoacids 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.[122] Since residues in the X and Y are partially exposed to the solvent, charged residue has a potential for intra and intermolecular interactions. Modeling studies indicate that oppositely charged residues within a single chain form salt bridges when separated by one or two residues.[123] Subsequently, electrostatic interactions between opposite residues are sterically possible between the collagen polypeptide chains when separated by one residue along the chain.[124]

In addition to electrostatic interactions, Brodsky et al.[125] have also studied intermolecular hydrophobic interactions of non polar aminoacid residues for triple-helical stabilization. A series of Gly-X-Y trimer units are observed in a defined environment like Ac-(Gly-Pro-Hyp)$_3$-Gly-X-Y-(Gly-Pro-Hyp)$_3$-Gly-Gly-NH_2 (where X and Y are a combination of Pro, Hyp, Ala, Leu and Phe aminoacids, investigated). All collagen-like peptides formed triple-helices, with melting temperatures ranging between 21^0C and 44^0C. Thermodynamic calculations indicate these collagen mimetic peptides have a range of free energy values (ΔG = 9 Kcal/mol), which suggests entropy to be the dominant factor for increased triple-helical stability. Hydrophobic residues such as Leu and Phe are found preferably in the X position and relate to their increased potential for intermolecular hydrophobic interactions. Molecular modeling studies show that Phe in the X position favorably orients the side chain for inter helix hydrophobic interactions, with the ring face almost parallel to the axis of the helix. However, in the Y position, the Phe residue is partially buried in the helix, in an orientation unsuitable for intermolecular interaction[125].

1.20: Template-assembled collagen structures

Assembly of the collagen triple-helix can also be effected via *de novo* approach. Peptide folding into its secondary structures is the fundamental requirement to induce proper biological response or recognition and incorporation of template into the design of peptidomimetics can direct and reinforce the intramolecular folding of peptides. Mutter et al.[126,127] introduced the TASP, (template-assembled synthesis proteins) methodology in their synthesis of four helical bundles. They were able to demonstrate that template assembly of peptides enhanced secondary structure formation with high

conformational stability. DeGrado et al.[128,129] designed a four-helix bundle via metal ion-dependent modulation. This study showed the feasibility of peptides to bind to cofactors and assume native-like folding. Incorporating a template into the synthesis of collagen structure can favor intramolecular folding as opposed to single chain intermolecular folding, and stabilize the triple helicity by reducing any entropy loss in triple helical formation.[130]

Template-assembled analogs have proven to increase the thermal stability of collagen mimetic structure. Fields et al.[131,132] have synthesized a lysine-lysine dipeptide template to covalently anchor the C-terminal of the three polypeptide chains covalently. Biological sequence from type I and IV collagen were incorporated into a collagen mimetic structure composed of Gly-Pro-Hyp trimer units and stabilized by the lysine-lysine dipeptide template at C-termini. The unique feature of this amino acid based template is the incorporation of orthogonal protecting groups, which can allow selective deprotection/protection strategies for the synthesis of heterotrimeric collagen mimetic structures. Fields et al.[133] have developed a series of heterotrimeric 'mini-collagens' using their lysine-lysine template, whereby biological sequences of interest can be incorporated and studied in triple helical environments.

The use of templates to induce triple helical conformations has also included structures which are cross-linked collagen-like peptides joined at both the amino and carboxy termini. Tanaka *et al.*[134] synthesized lysine-lysine dipeptide template derivatives, using two chemoselective ligations to construct amino and carboxy cross-linked peptides. The thermal stabilities and standard free energies ΔG^0 for cross-linked sequences of Gly-Pro-Hyp repeats were higher than that of the corresponding single

cross-linked structures and single chains. The enthalpic and entropic values also increased with increasing number of repeats within the cross-linked peptides.[135] These novel structures have exhibited high thermal melting temperature and have opened up a new area of research for template assembly techniques (Fig. 20).

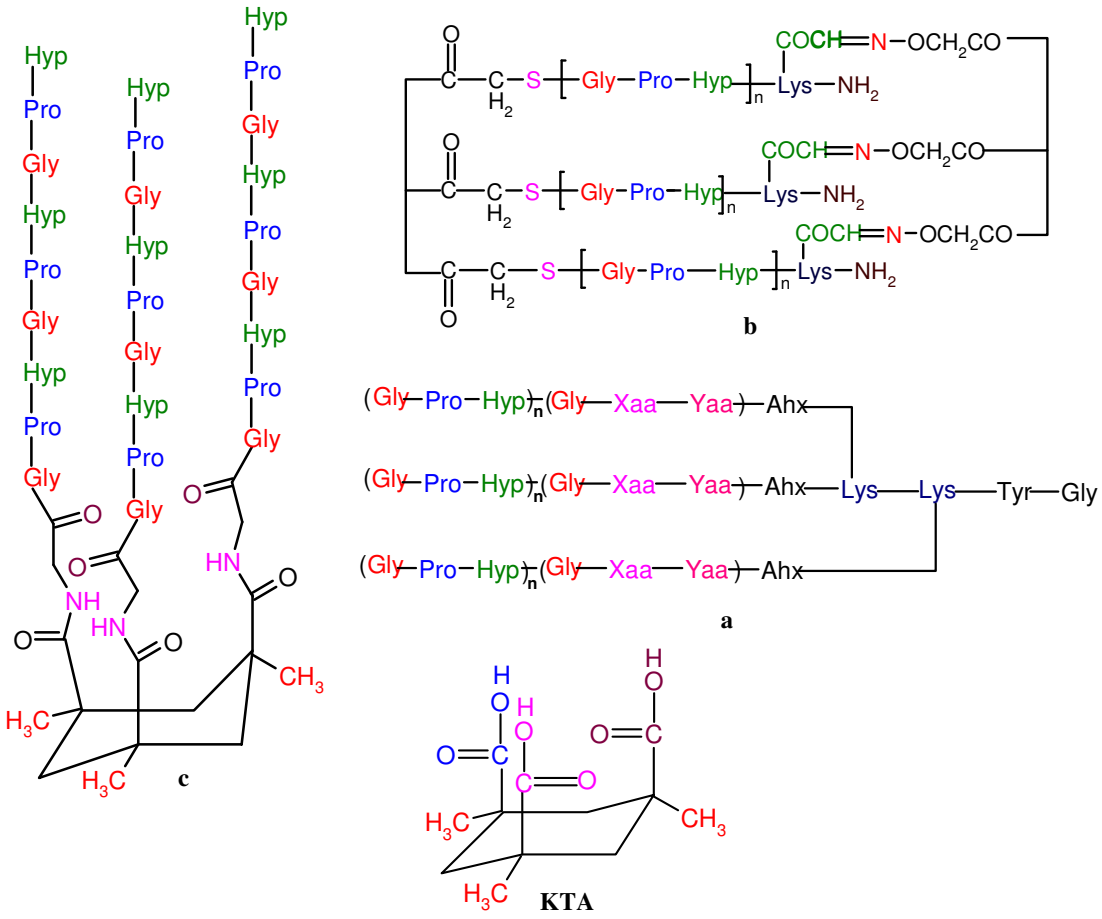


Figure 20.Template assembled collagen peptides. **a** lysine dimmer template, **b** lysyl-lysine template, **c** KTA template.[136]

In laboratories of Goodman, an integrated biophysical approach is utilized to characterize the triple helical collagen mimetic structures. The collagen mimetic structures investigated include single chain and template-assembled analogs and collagen mimetic structures composed of sequences containing non-natural residues. The template assembly motif has been applied to the collagen mimetic

program, where the Kemp triacid (KTA, 1,3,5,-methyl cyclohexane-1,3,5-tricarboxylicacid) was employed as the conformationally constrained template .

A glycine spacer was introduced between the template and the peptide chains to add flexibility and to compensate for the difference in space between the template and the collagen triple helix.[136,137] A series of structures Ac-(Gly-Pro-Hyp)$_n$-NH_2 and KTA-[Gly-(Gly-Pro-Hyp)$_n$-NH_2]$_3$ (n = 1, 3, 5, 6) were synthesized and assessed for triple helicity in aqueous solution by temperature-dependent optical rotation measurement and CD spectroscopy. The shorter single chain analogs exhibited no cooperative melting transition, whereas a transition was observed at 18^{o}C and 26^{o}C in water for Ac-(Gly-Pro-Hyp)$_n$-NH_2 where n = 5, 6, respectively. Similarly, the shorter template-assembled structure KTA-[Gly-(Gly-Pro-Hyp)-NH_2]$_3$ showed no melting transition, while KTA-[Gly-(Gly-Pro-Hyp)-NH_2]$_3$ where n = 3, 5, 6 exhibited transitions at 30^{o}C, 70^{o}C and 81^{o}C respectively.[137] A comparison between the single chain analogs and the template-assembled analogs defines the role of the template in inducing intramolecular folding and further stabilize the triple helical conformation.

1.21: 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.[138] Collagen fibers provide mechanical strength that confer 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.[139] Collagen synthesis decreases steadily with maturation and with subsequent aging drops 10-fold in majority of tissue. 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.[140] 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 determinate 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.[140] 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*.[141].This glycosylamine is then stabilized by spontaneous Amadori rearrangement to form a keto-imine, aminodeoxyketoase[141b,142].

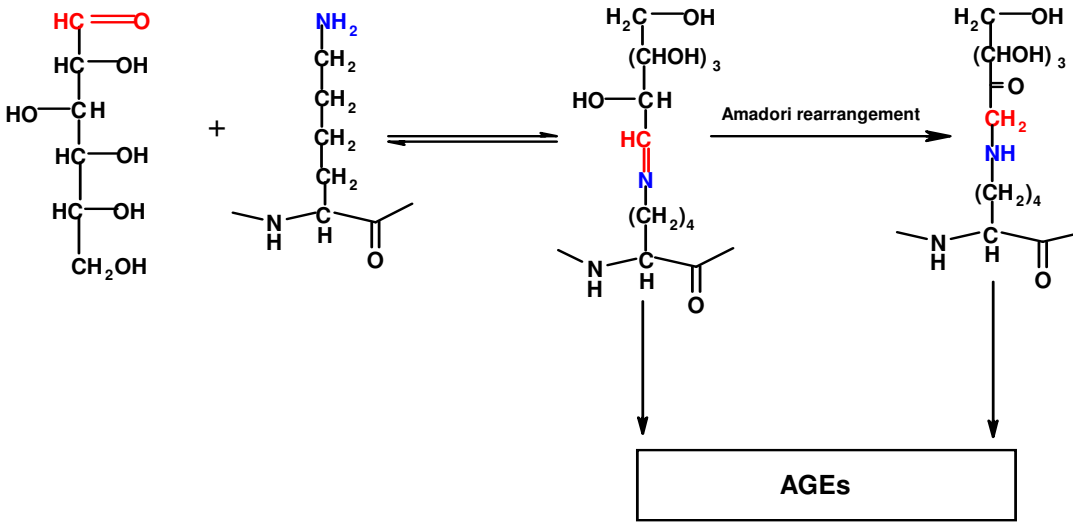


Figure 21: Reaction of glucose and lysine in collagen peptide to form a Schiff's base, which spontaneously undergoes an Amadori rearrangement to give aminodeoxyketose. Both the Schiff base and the ketose are belived to react further to form AGEs (advanced glycation end

the formation of more reactive sugars, such as 3-deoxyglycosone and the glyoxals, which then complex with lysine side-chain to form AGEs.[144]

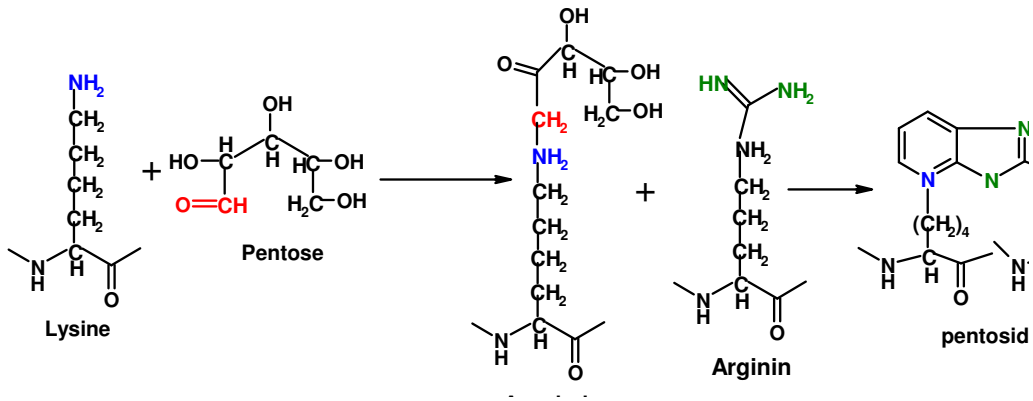


Figure22: Reaction of a pentose sugar with the side-chains of the Lys and Arg residues on different peptide chains to form the *pentosidine* cross-link

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* cross-links[145]. Several other mechanisms operate leading to chemical modifications that reduce the flexibility of collagenous tissue.[140] 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.[146]

The critical role of collagen is clearly illustrated by the wide spectrum of diseases caused by more than 1000 mutations that have thus for been identified in 22 genes for 12 out of 20 collagen types.[147] 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.*[148] Mutations that have milder effects on the synthesis or structure of protein are found in more common diseases such as *Osteporosis* 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 form of *Osteogenesis imperfecta*, results in retarded triple-helix formation, thus altering collagen fibril formation and consequently decreasing the bone strength.[149] Among the acquired diseases, transient increases in hydroxylation occur in collagen during bone fracture-repair and Osteoporosis.[150] 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 subtype of epidermolysis bullosa, knobloch syndrome, arterial aneurysms, Osteroarthrosis, and intervertebrals disk disease have also been ascribed to the imperfections in the collagenous tissue.[148]

1.22: Collagen biomaterials

Collagen and collagen based materials have extensive applications in biomedical devices and tissue engineering.[151] They are used, for example, in the fabrication of cardiac and aortal valves[151b] and vascular grafts,[152] and have been proposed for use in ophthalmic implants,[153] abdominal repair fabrics,[154] and ligaments.[155] Chemical stabilization of collagen is often used to reduce the potential of immune and inflammatory reactions, and to stabilize it against proteolytic degradation. Such collagen and collagen-based materials seem to have an inherent biocompatibility that is difficult to mimic in synthetic polymers. However, the biological origin of collagen, with the consequential risk of disease transmission, complicates their manufacture and can be an issue during the regulatory approval process. Furthermore, there can be ethical considerations for their use in which the collagen is derived from human tissue. The design of mimics of collagen may facilitate the development of wholly synthetic "collagen-based" biomaterials, which have the biocompatibility of native collagen while manufacturing and regulatory characteristics of synthetic polymer. Collagen-based peptides and intact collagen molecules also been used as surface coating in such applications as born repair matrices and corneal implants.[156] However, intact collagen neither has all the desirable properties as a biomaterial nor the necessary enzyme resistance. In order to improve the biomaterial properties, such as increased enzyme resistance and improved cell adhesion properties, several chemical modifications have been attempted. For example, chemical linking of PEG (polyethylene glycol) to natural collagen reduced the attachment of fibroblast and bacterial cells by 98% on the films containing collagen. Modifications that increased the net positive charges on the

collagen such as amidation of the –COOH group has increased the attachment of fibroblast while decreasing the bacterial cell adhesion.[157]

Recently, triple-helical collagen mimic peptides containing achiral amino acid N-isobutyl glycine (Nleu) (Fig. 23) have been tested for their cell binding activity. It was found that when immobilized on a surface the $(Gly\text{-}Pro\text{-}Nleu)_{10}\text{-}Gly\text{-}Pro\text{-}NH_2$ sequence stimulated attachment and growth of corneal epithelial cells and fibroblasts and migration of epithelial tissue. Interestingly peptide containing the sequence $(Gly\text{-}Nleu\text{-}Pro)_{10}\text{-}NH_2$ did not have cell binding activity.[158] The discovery that the triple-helical synthetic peptides containing the achiral amino acids can interact with cells opens up new opportunities in the design of collagen mimetic biomaterials.

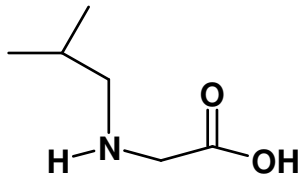


Figure 23: N-Isobutylglycine (Nleucine, Nleu)

1.23: Present work

Although the structure of collagen was proposed almost 50 years ago, a complete understanding of various factors contributing to the stability of collagen triplex is still lacking. A number of experimental evidences pointed out the necessicity of Hyp in the Y position for collagen triplex structure to allow 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 did not explain the incompatibility of 4*R*-substituted proline in X position. The recent crystal structure data and molecular modeling study on triplex $([Pro\text{-}pro\text{-}gly]_{10})_3$ suggested that proline in X-position prefers to adopt a C4-*endo* (down) pucker, whereas in the Y position prefers to 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.

In earlier study from this laboratory on collagen® it has showed that the replacement of Hyp by 4*R*-aminoproline enables one to understand the true effects of 4-proline substitution on collagen structure. NH_2 group is like OH in terms of hydrogen bonding, since it contains two hydrogen atoms it can form better hydrogen bond compared to -OH. In acidic pH protonated amine (NH_3^+) would be more like F in terms of electronegativity. 4*R*-Amp in Y-position remarkably stabilizes the collagen triple helix at both pH 3 (NH_3^+) and at pH 12 (NH_2), but forms a lower stability triplex at pH 9.

The analysis of the dependence of pyrrolidine ring conformation on the stereochemistry of 4-NH_3^+ by vicinal ^1H-^1H-coupling constant in ^1H-NMR of 4*R*-Amp and 4*S*-amp, shows that 4*R*-Amp prefers a C4-exo pucker for the pyrrolidine ring as 4*R*-

 C4-endo pucker, which is well suited for X position in the collagen model peptide X-Y-Gly.

1.24: Objective in undertaking work

The ionizable 4-aminoproline provides an extra handle for probing collagen structure through pH changes. This thesis comprises of a complete study of compatible

preferences of both 4*R* and 4*S* aminoprolines at X and Y positions of the collagen model peptide X-Y-Gly, individually as well as in combination with other 4-susbtituted proline (Hyp).

The present chapter has described the tremendous progress that has been made in understanding the collagen structure. Several chemical modifications that have been attempted to modulate and probe the role of various aminoacids and trans-4-hydroxyproline in particular, have been described. The importance of these studies lie in their application for the design of various collagen-biomaterials and understanding of various extracellular-matrix related diseases. The work in this thesis is mainly devoted to the synthesis and structural characterization of new a class of synthetic collagens with 4-aminoproline as a Hyp and Pro surrogate. 4-aminoproline is both α- and γ-amino acid. In addition, this proline surrogate possessing two stereo-centers can give rise to 4-stereoisomers. In the present work, (2*S*,4*R*) and (2*S*,4*S*) stereoisomers of this versatile building block have been utilized to construct the various collagen model peptides.

Chapter 2 describes the synthesis and conformational study of novel collagen peptides with 4*R* and 4*S*-amionprolines in the X and Y positions of the collagen repeat sequence X-Y-Gly. The ability of 4*R* and 4*S* aminoprolines to adopt triple-helical conformation has been compared with Hyp and Pro in the same positions. It is shown that 4*R*-aminoproline at Y and 4*S*-aminoproline at X position, form stable triple-helices, in addition, aminoproline containing collagen peptides act as a pH dependent conformational switch for triple helical stability.

Chapter 3 This chapter is divided into two parts. Part A presents the study of chimeric collagen peptide containing modified aminoacids both at X and Y positions. Individual

conformation preferences of 4*R* and 4*S* aminoprolines at X and Y positions are clubbed together to construct hyperstable collagen model peptide with 4*S*-substituted proline at X and 4*R*-substitute proline at Y positions. It is shown that 4*S*-amnioproline at X position is well compatible with 4*R*-amino (hydroxyl) proline at Y position.

Part B explains the dual compatible role of 4*R*-substituted prolines both at X and Y positions together. The preference of protonated 4*R*-aminoproline at X position to form a triple-helical structure in chapter-2 brings to test its compatibility at X position with other 4*R*-substituted prolines at Y position. It is observed that 4*R*-aminoproline is also well compatible at X position like 4*S*-aminoproline, but later one forms comparatively less stable triple helix.

Chapter 4 explains the role/effect of protonation of 4-aminogroup on the triplex formation. The 4-aminogroup was subjected to formyl protection in order to prevent getting protonated at acidic pH. It is shown that the 4-formyl protected aminoproline behaves like 4-hydroxyproline since its triplex stability is invariable with pH changes, as observed in case of 4-hydroxyproline. This shows that protonation of 4-aminogroup has a big role for formation and stability of triplex.

Chapter 5 explains the estimation of electronic, steric and stereochemical influence of the Amp and to decipher the contribution in the triple-helical stabilization in terms of intra-residue effects viz. *Z-E* isomerization and pyrrrolidine ring-pucker. This chapter describes the synthesis and study of the peptidyl-prolyl bond isomerization of 4*S* and 4*R*-aminoprolines in acidic and basic conditions. IR and variable temperature NMR spectroscopy have been used to estimate the effect of 4-amino group on the *Z-E* isomerization as protonated and nonprotonated forms.

1.25. References

1. Engel, J. *Science* **1997**, *277*, 1787-1786

2. Engel, J., Bachinger, H. P. in *Trends in Collagen*; eds. Chandrasekeren, G.; Yatihindra, N.; Ramasami, T. *Proceed.Indian Acad. Sciences* **1999**, *111*, 81-86.

3. (a) Linsenmayar, T. F. in Cell Biology of *Extracellular Matrix* 2nd ed. Hay, E.D., Ed., 1991, Pllenumpress, NY, PP. 7-44. (b) Prockop, D. J.: Kivirikko, K.I. *Ann. Rev. Biochem.* **1995**, *64*, 403-434.

4. Woodhead-Galloway, J. in *Collagen*: *The Anatomy of a Protein* Edward Arnold (Publishers) Ltd. London, PP. 10-19.

5. (a) Creighton, T. E. *Proteins* **199**3, W. H. Feeman & Co., NY, PP.45-60. (b) Ramachandran, G. N., Rmakrishnan, C. in *Biochemistry of Collagen* Ramchandran, G. N.; Reddi, A. H., Ed. Academic press, London, 1976, PP. 45-84. (c) Ramachandran, G. N., Rmakrishnan, C.; Venkatachalam, C. M. in *Conformations of Collagen* Ramachandran, G., Ed., Academic Press, London, **1967** PP. 429-438.

6. Ramachandran, G. N., in *Aspects of protein structure* Ramachandran, G. N., Ed Academic Press, NY, **1963** PP.39-55.

7. Bork, P.; Downing, A. K.; Kieffer, B.; Campbell, I. D. *Q Rev Biophys*, **1996**, *29*, 119.
8. Hohesester, E.; Engel, *J. Matrix Biol*, **2002**, 21, 115.

9. (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*, 1543-1546
10. For review see, Takaki Koide,; Kazuhiro Nagata. Collagen at glance, *Top Curr Chem,* **2005**, *247*, 85-114.

11. Dolz, R.; Heidemann, E. *Biopolymers* **1986**, *25*, 1069-1080.

12. (a) Fietzek, P.P.; Kuhn, K. *Mol. Cell. Biochem.* **1975**, *8*, 141-157. (b) Salem, G.; Traub, W. *FEBS Letters* **1975**, *51*, 94-99.

13. Bansal, M. *Int. J. Pep. Protein Res.* **1977**, *9*, 224-234.

14. Bansal, M.; Ramachandran, G. N. *Int. J. Pep. Protein Res.* **1978**, *11*, 75-81.

15. Heidmann, E.; Roth, W. *Adv. Polym. Sci.* **1982**, *43*, 143-203.

16. (a) Shah, N. K.; Ramshaw, J. A. M.; Kirkpatrik, A.; shah, C.; Brodsky, B. *Biochemistry* **1996**, *35*, 10262-10268. (b) Persikov, A. V.; Ramshaw, J. A. M.; Kirkpatrik, A.; Brodsky, B. *Biochemistry* **2000**, *39*, 14960-14967.

17. Ramachandran, G. N.; Kartha, G. *Nature* **1954**, *174*, 269-270. (a) Ramachandran, G. N.; Kartha, G. *Nature* **1955**, *176*, 593-595.

18. (a) Ramachandran, G. N.; Kartha, G. *Nature* **1955**, *176*, 593-595. (b) Ramachandran, G. N.; Kartha, G. *Proc. Indian Acad. Sci*. **1955**, *XLII*, 215-234.

19. Rich, A.; Crick, F. H. C. *Nature* **1995**, *176*, 915-916. (b) Rich, A.; Crick, F. H. C. *J. Mol. Biol*. **1961**, *3*, 483-506.

20. Yonath, A.; Traub, W. *J. Mol. Biol*. *1969, 43*, 461-477.

21. Sakakibara, S.; Kishida,Y.; Okuyama, K.; Tanaka, N.; Ashida, T.; Kakudo,M. *J.Mol.Biol*. **1972**, *65*, 371-373.

22. (a) Okuyama, K.; Tanaka, N.; Ashida, T.; Kakudo, M. *Bull. Chem. Soc. Jpn*.**1976**, *49*, 1805-1810. (b) Okuyama, K.; Tanaka, N.; Ashida, T.; Kakudo, M. *J. Mol. Biol*. **1981**, *152*, 427-443.

23. (a) Nagarajan, V.; Kamitori, S.; Okuyama, K. *J. Biochem*. **1998**, *124*, 1117-1123. (b) Kramer, R. Z.; Vitagliano, L.; Bella, J.; Berisio, R.; Mazzarella, L.; Brodsky, B.; Zagari, A.; Berman, H. M. *J. Mol. Biol*. **1998**, *280*, 623-628.

24. Karmer, R. Z.; Bella, J.; Mayville, P.; Brodsky, B.; Berman, H. M. *Nature Struct. Biol*. **1999**, *6*, 454-457.

25. Orgel, J. P.; Miller, A.; Irving, T. C.; Fischetti, R. F.; Hammersley, A. P.; Wess, T. J. *Structure* **2001**, *9*, 1061-1069.

26. Berisio, R.; Vitagoliano, L.; Mazzarella, L.; and Zagari, A. *Biopolymers* **2001**, *56*, 8-13.

27. Kramer, R.Z.; Venugopal, M. G.; Bella, J.; Mayvillle, P.; Brodsky, B.; Berman, H. M. *J. Mol. Biol*. **2000**, *301*,1191-1205.

28. Bella, J.; Eaton, M.; Brodsky, B.; Berman, H. M. *Science* **1994,** *266*, 75-81.

29. Karmer, R. Z.; Bella, J.; Brodsky, B.; Berman, H. M. *J. Mol. Biol*. **2001**, *311*, 131-147.

30. Miller, M. H.; Scheraga,H. A. *J. Polym. Sci*, **1976**, *54*, 171-200.

31. Miller, M. H.; Nemethyl, G.; Scheraga, H. Macromolecules, **1980**, *13*, 470-478.

32. Vitaglino, L.; Nemethyl, G.; Scheraga, H. A. *Biopolymer*, **1993**, *32*, 7354-7359.

33. Vitaglino, L.; Nemethyl, G.; Scheraga, H. A. *J. Mol. Biol*. **1993**, *247*, 69-80.

34. (a) Okuyama, K.; Arnott, S.; Takayanagari, M.; Kakudo, M. *J. Mol. Biol*. **1981**, *152*, 427-443. (b) Bella, J.; Eaton, M. Brodsky, B.; Berman. H. M. *Science*,

1994, *266*, 75-81. (c) Bella, J.; Brodsky, B.; Berman, H. M. *Structure* **1995**, *3*, 893-906. (d) Kramer, R. Z.; Vitagliano, L.; Bella, J.; Berisio, R.; Mazzarela, L.; Brodsky, B.; Zagari, A.; Berman, H. M. *J. Mol. Biol*. **1998**, *280*, 623-638. (e) Nagarajan, V.; Kamitori, S.; Okuyama, K. *J. Biochem*. **1998**, *124*, 1117-1123. (f) Nagarajan, V.; Kamitori, S.; Okuyama, K. *J. Biochem*. **1999**, *125*, 310-318.

35. Ramachandran, G. N.; Bansal, M.; Batnagar, R. S. *Biochim. Biophys.Acta*, **1973**, *322*, 166-171.

36. Privalov, P. L. *Adv. Protein. Chem*, **1982**, *35*, 1-104.

37. Brg, R. A.; Prockop, D. J. *Biopchem. Biophys. Res. Commun*. **1973**, *52*, 115-120.

38. Inouye, K.; Kobayashi, Y.; Kyogoku, Y.; Kishida, Y.; Sakakibara, S.; Prockop, D. J. *Arch. Biochem. Biophys*, **1982**, *219*, 198-203.

39. Inouye, K.; Sakakibara, K.; Prockop, D. J. *Biochim. Biophys.Acta*, **1976**, *420*, 133-141.

40. Gustavson, K. H.; *Nature*, **1955**, *175*, 70-74.

41. Bella, J.; Bordsky, B.; Berman. H. M. *Structure*, **1995**, *3*, 893-906.

42. Holmgren, S. K.; Taylor, K. M.; Bretscher, L. E.; Raines, R. T. *Nature*, **1998**, *392*, 666-667.

43. Bann, J. G.; Bachinger, H. P. *J. Biol. Chem*, **2000**, *275*, 24466-24469.

44. Pauling, L.; et al. *Proc. Natl.Acad. Sci. USA* **1951**, *37*, 205-211.

45. Flory, P. J. *Principles of polymer chemistry,* Cornell University Press, Ithaca, New York, 1953.
46. Ramachandran, G. N.; Sasisekharan, V. *Adv. Protein Chem.* **1968**, *23*, 283-487.

47. Thornton, J. M.; Jones, D. T.; MacArthur, M. W.; Orenbo, C. M.; Swindells, M. B. *Philos. T. Roy. Soc. B*, **1995**, *348*, 71-79.

48. Swindells, M. B.; Macarthur, M.W.; Thornton, J. M. *Nat. Struct. Biol*, **1952**, *2*, 596-603.

49. Chou, P. Y.; Fasman, G. D. *Adv. Enzymology*, **1978**, *47*, 45-148.

50. Munoz, V. Serrano, L. *Proteins: Struct.Funct. Genet*, **1994**, *20*, 301-311.

51. Gibart, J. F.; Robson, B. Garnier, J. *Biochemistry*, **1991**, *30*, 1578-1586.

52. Kang, H. S.; Kurochkina, N. A.; Lee, B. *J. Mol. Biol*, **1993**, *229*, 448-460.

53. Barton, G. J. *Curr. Opin. Struc. Biol*, **1995**, *5*, 372-376.

54. Pace, C. N.; Shirley, B. A.; Mcnutt, M.; Gajiwala, K. *FASEB J*, **1996**, *10*, 75-83.

55. Shortle, D. *FASEB J*, **1996**, *10*, 27-34.

56. Arnott, S.; Dover, S. D. *J. Mol. Biol*, **1967**, *30*, 209.

57. Schellman, J. A.; Schellman, C. *Proteins*, **1964**, *2*, 1.

58. Arnott, S.; Dover, S. D.; Elliott, A. *J. Mol. Biol*, **1967**, *30*, 201.

59. Ramachandran, G. N.; Sasisekharan, V.; Ramakrishnan, C. *Biochim. Biophys. Act,* **1996**, *122*, 168.

60. Yonath, A.; Trab, W. *J. Mol. Biol*, **1969**, *43*, 461.

61. Ramachandran, G. N.; Sasisekharan, V. *Advan. Protein Chem.* **1968**, *23*, 283.

62. Teaub, W.; Schmueli, U. *Aspects of protein structure*, **1963**, Ramachandran, G. N., Ed, *Academic press*, London, p 81.

63. Haasnoot, C. A. G.; De Leeuw, A. A. M.; De leeuw, H. P. M.; Altona, C. *Biopolymers*, **1981**, *20*, 1221-1245.

64. Ederhardt, E. S; Panasik, N. J. R.; Raines, R. T. *J. Am. Chem. Soc*. **1996**, *118*, 12261-12266.

65. Bretscher, L.E.; Jenkins, C. L.; Taylor, K. M.; De Rider, M. L.; Raines, R. T. J. *Am. Chem. Soc*. **2001**, *123*, 777-778.

66. (a) Wolfe, S. *Acc. Chem. Res*. **1972**, *5*, 102. (b) Brunck, T. K.; Weinhoid, F. *J. Am. Chem. Soc.* **1979**, *101*, 1700-1709. (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.

67. Parker, D.; Senanayake, K.; Vepsaillainen, J.; Williams, S.; Batsanov, A. S.; Howard, J. A. K. *J. Chem. Soc. Perkin Trans. 2,* **1997**, *8*, 1445-1452.

68. Wiberg, K. B. *Acc. Chem. Res*. **1996**, *26*, 229-234.

69. (a) Panasik, N. J. R.; Eberhardt, E. S.; Edison, A. S.; Powell, D. R.; Rains, R. T. *Int. J. Pept. Protein Res*. **1994**, *44*, 262-269. (b) Gerig, J. T.; McLeod, R,S. *J. Am. Chem. Soc*. **1973**, *95*, 5725-5729.

70. Stewart, D. E,; Sarkar, A.; Wampler, J. E. *J.Mol. Biol*. **1990**, *214*, 253-260.

71. Weiss, M. S.; Jabs, A.; Hilgenfeld, R. *Nat. Struct. Biol*. **1998**, *5*, 676.

72. (a) Jabs, A.; Weiss, M. S.; Hilgenfeld, R. *J. Mol. Biol*. **1999**, *286*, 291-304. (b) Impotra, R.; Benzi, C.; Barone, V. *J. Am. Chem. Soc.* **2001**, *123*, 12568-12577. (c) Bürgi, H. B.; Dunitz, J. D.; Shefter, E. *J. Am. Chem. Soc*. **1973**, *95*, 5065-5067. (d) Bürgi, H. B.; Dunitz, J. D.; Wipff, G. *J. Am. Chem. Soc*. **1974**, *96*, 1956-1957. (e) Bürgi, H. B.; Dunitz, J. D.; Lehn, J. M.; Wipff, G. *Tetrahedron* **1974**, *30*, 1563-1572.

73. Delaney, N. G.; Madisn, V. *J. Am. Chem. Soc*. **1982**, *104*, 6635-6641.

74. Beauoleil, E,; Sharma, R..; Michnick, S. W.; Lubell, W. D. *J. Org. Chem*. **1998**, *63*, 6572-6578.

75. Magaard, V. W.; Sanchez, R. M.; Bean, J. W.; Moore, M. L. *Tetrahedron Lett*. **1993**, *34*, 381-384.

76. An, S. S. A.; Lester, C. C.; Peng, J. L. Li, Y. J.; Rothwarf, D. M,; Welker, E.; Thannhauser, T. W.; Zhang, L. S.; Tam, J. P.; Scheraga, H. A. *J. Am. Chem. Soc*. **1999**, *121*, 11558-11566.

77. Halab, L.; Goselin, F.; Lubell, W. D. *Biopolymers (pep. Sci)* **2000**, *55*, 101-122.

78. Arold, U.; Hinderkar, M. P.; Koeditz, J; Golbik, R.; Ulbrich-Hoffmann, R,; Raines, R. T. *J. Am. Chem. Soc*. **2003**, *125*, 7500-7501.

79. Andes, C. J.; Macdonald, T. L.; Ocain, T. D.; Longhi, D. *J. Org. Chem*. **1993**, *58*, 6609-6613.

80. Welch, J. T.; Lin, J. *Tetrahedron* **1996**, *52*, 291-304.

81. Hart, S. A.; Sabat, M.; Etzkorn, F. A. *J. Org. Chem*. **1998**, *63*, 7580-7581.

82. Lin, J.; Toscano, P. J.; Welch, J. T. *Proc. Natl. Acad. Sci. U.S.A*. **1998**, *95*, 14020-14024.

83. Otaka, A.; Katagiri, F.; Kinoshita, T.; Odagaki, Y.; Oishi, S.; Tamamura, H.; Hamanaka, N.; Fujii, N. *J. Org. Chem*. **2002**, *67*, 6152-6161.

84. Wang, X. J.; Hart, S. A.; Xu, B.; Mason, M. D; Goodell, J. R.; Etzkorn, F. A. *J.Org. Chem*. **2003**, *68*, 2343-2349.

85. DeTar, D. F; Luthra, N. P. *J. Am. Chem. Soc*. **1977**, *99*, 1232-1244.

86. Vitagliano, L.; Berisio, R.; Mazzarella, L.; Zagari, A. *Biopolymers* **2001**, *58*, 459-464.

87. O'Hagan, D.; Bilton, C.; Howard,, J. A. K.; Knight, L.; Tozer, D. J. *J. Chem. Soc., Perkin Trans*. 2 **2000**, 605-607.

88. Berscher, L. E.; Jenkins, C. L.; Taylor, K. M.; DeRider, M. L.; Raines, R. T. *J. Am. Chem. Soc.* **2001**, *123*, 777-778.

89. Hinderaker, M. P.; Raines, R. T. *Protien Sci*. **2003**, *12*, 1188-1194.

90. (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.

91. Cohen, C.; Parry, D. A. D. *Trends Biochem.Sci*. **1988**, 11, 245-248.

92. (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.

93. Bruns, R. R,; Gross. *J. Biochemistry* **1973**, *12*, 808-815.

94. Barber, H. E.; Mackay, D. *Arch. Biochem. Biophys*. **1996**, *177*, 466-468.

95. (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.; Brosdky, B.; Baum, J. *Biochemistry* **1993**, *32*, 7377-7387.

96. Berg, R. A.; Olsen, B. R.; Prockop, D. J.; Klump, H. *J. Biol. Chem*. **1970**, *245*, 5759-5763.

97. Venugopal, M. G.; Ramshaw, J. A. M.; Braswell, E.; Zhu, D.; Brodsky, B. *Biochemistry* **1994**, *33*, 7948-7956.

98. Cansonni, R.; Santomo, L.; Tenni, R.; Longhi, R.; Zetta, L.*FEBS Letters* **1998**, *436*, 243-246.

99. (a) Gratzer, W. B.; Rhodes, W.; Fasman, G. D. *Biopolymers* **1963,** *1*, 319-330. (b) Wood, G. C. *Biochem. Biophys. Res. Commun*. **1963**, *13,* 95-98.

100. (a) Lazarev, Yu. A.; Lazareva, A. V.; Shidnev, V. A.; Esipova, N. G. *Biopolymers* **1998**, *17*, 1197-1214. (b) Khromova, T. B.; Lazarev, Yu. A. *Biophysics* **1986**, *31*, 227-232.

101. (a) Saito, H.; Tabeta, R.; Shoji, A.; Ozaki, T.; Ando, I.; Miyata, T. *Biopolymers* **1984**, *23*, 2279-2297. (b) Renugopalakrishnan, V.; Colette, T. W.; Carreira, L. A.; Bhatnagar, R. S. *Macromolecules* **1985**, *18*, 1786-1788.

102. Brodsky, B.; Li, M-H.; Long, C. G.; Apigo, J.; Baum, j. *Biopolymers* **1992**, *32*, 447-451.

103. (a) Miller, A.; Wray, J. S. *Nature* **1971**, *230*, 437. (b) Bruns, R. R.; Gross, J. *Biochemistry* **1973**, *12*, 808.

104. Fraser, R. D. B.; MacRae, T. P.; Suzuki, E. J. Mol. Biol. 1979, 129, 463.

105. Sakakibara, S.; Kisida, Y.; Okuyama, K.; Tanaka, N.; Asida, T.; Kakudo, M. *J. Mol. Biol.* **1972**, *65*, 371.

106. Burjanadze, T. V.; Bezhitadze, M. O. *Biopolymers* **1992**, *32*, 951.

107. Christopher, A. M.; Allen, J. B. *Proc. Indian Acad. Sci (Chem. Sci)* **1999**, *111*, 71-80.

108. Fields, C. G.; Grab, B.; Lauer, J. L.; Fields, G. B. *Anal. Biocem.* **1995**, *231*, 51.

109. Murry Goodman.; Juliann, K. Proc. *Indian Acad. Sci. (Chem. Sci.)* **1999**, *111*, 35-49.

110. Sutoh, K.; Noda, H. *Biopolymers* **1974**, *13*, 2461.

111. Katz, E. P.; David, C. W. *J. Mol. Biol*. **1992**, *228*, 963.

112. Hoppe, H-J.; Reid, K. B. M. *Protein Sci*. **1982**, *3*, 1143.

113, Sakakibara, S.; Kisida, Y.; Kikudihi, Y.; Sakai, R.; Kakiushi. K. *Bull. Chem. Soc. Jpn*. **1968**, *41*, 1273.

114. Kobayashi. Y.; Sakai, R.; Kakiuchi, K.; Isemura,T. *Biopolymers* **1970**, *9*, 415.

115. Satturin, A.; Tamburro, A. M.; Del Pra, A.; Bordingnon, E. *Int. J. Peptide Prot. Res*. **1975**, *7*, 425.

116. Suoth, K.; Noda, H. *Biopolymers* **1974**, *13*, 2461.

117. Li, M.; Fan, P.; Brodsky, B.; Baum, J. *Biochemistry* **1993**, *32*, 7377.

118. Brodsky, B.; Li, M.; Long, C. G.; Apigo, J.; Baum, J. *Biopolymers* **1992**, *32*, 447.

119. Inouye, K.; Sasakibara, S.; Prockop, D. J. *Biochim. Biophys. Acta* **1976**, *420*, 133-141.

120. Fairwather, R.; Jones, J. H. *J. Chem. Soc. Perkin I* **1972**, 2475-2481.

121. Hutton, J. J.; Marglin, A.; Witkop, B.; Kurtz, J.; Berger, A.; Undenfreind, S. *Arch. Biochim. Biophys*. **1968**, *125*, 779-785.

122. Zagari, A.; Nemethy, G.; Scheraga, H. A. *Biopolymers* **1990**, *30*, 967.

123. Li S-T.; Golub, E.; Katz, E. P. *J. Mol. Biol.* **1975**, *98*, 791.

124. Katz, E. P.; David, C. W. *Biopolymers* **1990**, *29*, 791.

125. Venugopal, M. G.; Ramshaw, J. W.; Brestscher, L. E.; Raines, R. T. *Nature (London)* **1998**, *392*, 666.

126. Mutter, M.; Tuchscherer, G. G.; Miller, C.; Altmann, K. H.; Carey, R. I.; Wyss, D. F.; Labhardt, A. M.; Rivier, J. E. *J.Am. Chem. Soc*. **1992**, *114*, 1463.

127. Grove, A.; Mutter, M.; Rivier, J. E. Montal, M. *J. Am. Chem. Soc*. **1993**, *115*, 5919.

128. Handel, T. M.; Williams, S. A.; DeGrado, W. F. *Science* **1993**, *261*, 879.

129. Choma, C. T.; Lear, J. D.; Nelson, M.J.; Dutton, P. L.; Robertson, D. E.; DeGrado, W. F.; *J. Am. Chem. Soc*. **1994,** *116*, 856.

130. Goodman, M.; Bhumralkar, M.; Jefferson, E. A.; Kwak, J.; Locardi, E. *Biopolymers* **1998**, *47*, 127.

132. Fields, C. G.; Grab, B.; Lauer, J. L.; Fields, G. B. *Anal. Biochem*. **1995**, *231*, 57.

133. Grab, B.; Miles, A. J.; Furch, L. T.; Fields G. B. *J.Biol. Chem*. **1965**, *271*, 12234.

134. Lauer, J. L.; Fields, G. B. *6th Naples workshop on Bioactive Peptides*. **1998**, Abstract no.*12*.

135. Tanaka, Y.; Suzuki, K.; Tanaka, T. *J. Peptide Res*. **1998**, *51*, 413.

136. Feng, Y.; Melacini, G.; Taulane, J. P.; Goodman, M. *J. Am. Chem. Soc*. **1996**, *118*, 10351.

137. Goodman, M.; Feng, Y.; Melacini, G.; Taulane, J. P. *J. Am. Chem. Soc.* **1996**, *118*, 5156.

138. Alexander, R. M. *Sci. Prog*. **1981**, *67*, 109-130.

139. (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.

140. Bailey, A. J.; Paul, R. G.; Knott, L. *Mechanisms of Ageing and Devel*. **1998**, *106*, 1-56.

141. Siegel, R. C.; *Int. Rev. Connect. Tissue Res*. **1979**, *8*, 73-188.

142. (a) Hodge, H. J. *Agric. Food Sci.* **1953**, *1*, 928-943. (b) Robins, S. P.; Bailey, A. *J. Biocem. Biophys. Res. Commun.* **1972**, *48*, 76-84.

143. (a) Baynes, J. W.; Monnier, V. M. *The Maillard Reaction in ageing, Diabetes, and Nutrition.* **1989**, A. R. Liss, NY. (b) Labuza, D. V.; Reineccius, G. A.; Monnier, V. M.; O'Brien, J,; Baynes, J. W. *Maillard Reactions in Chemistry Food and Health.* **1994**, RSC, Cambridge, UK. (c) Schlecher, E.; Wieland, O. H. *J. Chem. Clin. Biochem*. **1981**, *19*, 81-87.

144. (a) Cerami, A.; Viassara, H.; Brownlee, M. *Sci. Am*. **1987**, *256*, 82-88. (b) Ahmed, M. U,; Thorpe, S. R.; Baynes, J. W. *J. Biol. Chem.* **1986**, *261*, 4889-48894.

145. (a) 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.

146. (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. (c0 Grandhee, S. K.; Monnier, V. M. *J. Biol. Chem.* **1991**, *266*, 11649-11653.

147. Fledelius, C.; Johnesen, A. H.; Cloos, P.A.C; Bonde, M; Qvist, P. *J. Biol. Chem*. **1997**, *272*,9755-9763.

148. Myllyharju, J.; Kivirikko, K. *I. Ann. Med*. **2001**, *33*, 7-21.

149. prockop, D. *J. Matrix Biol*. **1998**,*16*, 519-28.

150. Krisch, E.; Krieg, T.; Remberger, K.; Fendel, H.; Bruckner, P.; Muller, P. K. Eur. J. Clin. Investig, 1981, 11, 39-47.

151. (a) Glimcher, M. J.; Shapiro, F.; Ellis, R. D.; Eyre, D. R. *J. Bone Joint Surg*. **1980**, *62,* 964-973. (b) Bailey, A. J.; Wotton, S. F.; Sims, T. J.; Thomopson, P. W. *Biochem. Biophys. Res. Commun.* **1992**, *185*, 801-805. (c) Batge, B.; Diebold, J.; Stein, H.; Bodo, M.; Muller, P. K. *Eur. J. Clin. Investig*. **1992**, **22**, 805-812.

152. (a) Ramshaw, J. A.; Wekmeister, J. A.; Glattauer, V. *Biotechnol. Genet. Eng. Rev*. **1996**, *13*, 335-382.

153. Ramashaw, J. A.; Peters, D. E.; Werkmeister, J. A.; Ketharenathan, V. *J. Biomed. Mater. Res.* **1989**, *23*, 649-660.

154. (a) Thompson, K. P.; Hanna, K.; Waring, G. O.III.; Gipson, I.; Liu, Y.; Gailitis, R. P.; Johnson-Wint, B.; Green, K. *Refract. Coneal. Surg*. **1991**, *7*, 240-248. (b) Kirkham, S. M.; Dangel, M. E. *Ophthalmic Sur.* **1991**, *22*, 455-461.

155. Cavallaro, J. F.; Kemp, P. D.; Kraus, K. H. *Biotechnol. Bioeng*. **1994**, *43*, 781-791.

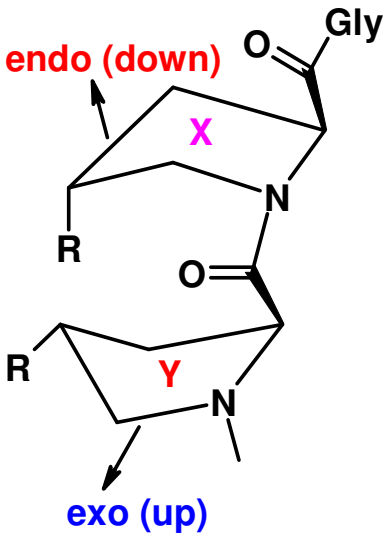
156. (a) Qian, J. J.; Bhatnagar, R. S. *J. Biomed. Mater. Res.* **1996**, *31*, 545-554. (b) Trinkaus-Randall, V.; Capecchi, J.; Newton, A.; Vadasz, A.; Leibowitz, H. M.; Franzblau, C. *Invest. Opthalmol. Vis. Sci.* **1998**, *29*, 393-400.

157. Tiller, J. C.; Bonner, G.; Pan, Li-Chun.; Klibanov, A. M. *Biotech. Bioeng*. **2001**, *3*, 246-252.

158. Johnson, G.; Jenkins, M.; McLean, K.; Griesser, H. J.; Kwak, J.; Goodman, M.; Steele, J. G. *J. Biomed. Mater. Res*. **2000**, *51*, 612-624.

Chapter 2

Investigation of compatibility of 4*S*- and 4*R*-aminoprolines in X and Y position of the collagen peptides $(X\text{-}Y\text{-}Gly)_n$



2. 1: Introduction

The triple-helix stability of collagen is greatly correlated with the Hyp (**1**) content.[1] Previous studies have shown that the peptide $(Pro\text{-}Hyp\text{-}Gly)_{10}$ forms a triplex with significantly higher transition temperature than the peptide $(Pro\text{-}Pro\text{-}Gly)_{10}$.[2] With this observation, it was thought that the thermal stability of the collagen triple-helix arises from the interchain hydrogen bonds between the amide group of glycine and the carboxyl group of X position residue through water mediation.[3] This thought was eviewed after the observation of formation of much stable triple-helix by replacing 4-hydroxyproline with 4-fluoroproline[4] (**2**), since fluorine linked to carbon does not make hydrogen bonds as strong as hydroxyl group.[5] Gauche effect was invoked to explain the higher strength of triplex derived from 4-fluoroproline.[6] The electronegative substituents on proline ring favor the *trans* peptide bond geometry (which is necessary for triple-helix formation) because the electron-withdrawing substituents on the 4-position of the proline affect the pyramidylization of the ring nitrogen, increasing its sp^3 character thereby influencing the rate of prolyl peptide bond isomerization. However, this reasoning devoid of stereochemical effects does not explain the incompatibility of 4*R*-hydroxyproline in X-position to form a triple-helix.[7] In earlier study from this laboratory the 4*R*-hydroxyproline was replaced with 4*R*-aminoproline (**3**), and found

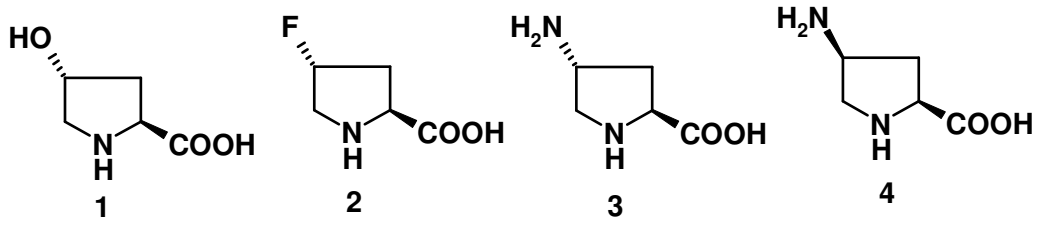


Figure 2.1: **1** 4*R*-hydroxyproline (Hyp); **2** 4*R*-fluoroproline (Flp); **3** 4*R*-aminoproline (Amp); **4** 4*S*-aminoproline (amp)

containing collagen peptide forms a triple-helix with greater stability than both fluoroproline and hydroxyproline.[8] The idea of replacement of aminoproline was that the easily ionizable amino group provides an extra handle for probing the collagen structure through pH changes. An acidic pH, the protonated amino group behaves like fluorine in terms of electronegativity, and at basic pH, the free amino group is more like –OH group in terms of hydrogen bonding property. The amino group can form better hydrogen bonds compared to –OH group, since it has two hydrogen atoms. The formation of stable triple-helix structure at both acidic (pH 3) and basic (pH 12) conditions leads to an ambiguity in concluding about the clear role of the electonegativity and/or hydrogen bonding in the triple-helix formation. The situation is unclear for the stability offered by 4-NH_2/NH_3^+ functions, where the conformational control may be exercised by a combination of hydrogen bonding, electronegativity and long range interstrand electrostatic contributions from the protonated NH_3^+ groups.

Recently, the crystal structure data[9] of triplex $[(Pro\text{-}Pro\text{-}Gly)_{10}]_3$ and its modeling study[10,11] showed that, proline in X- position prefers to adopt C4-*endo* (down) conformation while that in Y position prefers to adopt C4-*exo* (up) conformation (Fig.2.2A).

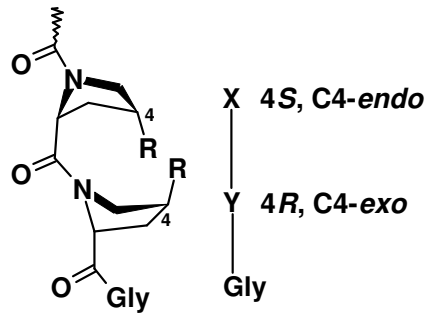


Figure 2.2A: position dependent preferred proline ring puckers in collagen

This was also supported by the theoretical calculations, which suggested that such differential proline puckering in X and Y positions are sterically necessary for a

favorable triple-helical packing arrangement. It is now emerging that the conformational preferences of the two prolines are strongly dictated by the nature of the C4-substituent.

For proline with no C4 substitution, the C4-*endo* and C4-*exo* puckers are isoenergetic. With 4*R*-OH (**1**) and 4*R*-F (**2**), the preferred equilibrium proline conformation is C4-*exo* which is also stabilized by a *guache* effect.[12] The situation is not that clear for the stability offered by 4-NH_2/NH_3^+ functions, where the conformational control may be exercised by a combination of electronegativity and electrostatic contributions from the protonated NH_3^+ groups. To decipher the influence of these factors, analysis of the dependence of pyrrolidine ring conformation on stereochemistry of 4-NH_2/NH_3^+ by vicinal ^{1}H-^{1}H-coupling constants in the ^{1}H-NMR of 4*R-trans* (**1**) and 4*S-cis* (**4**) aminoprolines. The results indicated that 4*R-trans*-aminoproline **1** prefers a C4-*exo* pucker for the pyrrolidine ring as in 4*R-trans*-hydroxyproline, whereas the preferred conformation for 4*S-cis*-aminoproline **4** is C4-*endo* (Fig. 2

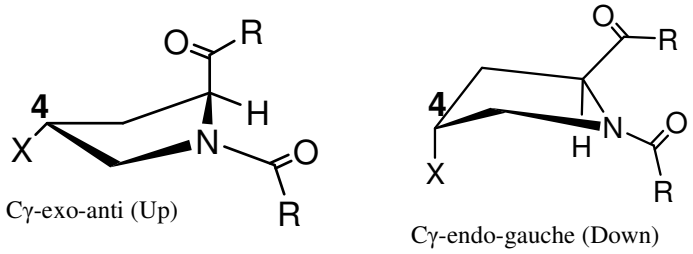


Figure 2.2B: Pyrroldine ring conformation.

2.1.1: *Trans-Cis* isomerization

Another factor correlated with triplex stability is the *trans-cis* isomerization rate constant of the tertiary amide bond which is also dependent on the nature of 4-substituent.[16]The thermodynamics of the *E*/*Z* isomerization of the peptidyl-prolyl bond can be effectively studied using variable temperature NMR spectroscopy which

prerequisites well separated signals for *E* and *Z* isomers in their ^{1}H and ^{13}C spectra. The analysis of $K_{trans/cis}$ for AcYaaOMe compounds where Yaa is Pro, Hyp, hyp, Flp and flp indicates that electronegativity and stereochemistry of 4-substitutent in the Yaa mimics has significant effect on $K_{trans/cis}$ (Table-1). Compared to flp residue, Pro residue is twice as likely and an Flp residue is three times as likely to have a *trans* peptide bond[17]. The correlation of $K_{trans/cis}$ values of AcYaaOMe with T_m values of $(ProYaaGly)_7$ triple-helices concluded that larger $K_{trans/cis}$ resulting with larger T_m. This indicates that Cγ substituents can enhance the conformational stability by favoring the *trans* isomer, thereby preorganizing individual strands to resemble more closely the strands in the triple-helix.

Table 1: Effect of 4*R* and 4*S* substituents of Residues on the conformational stability of Triple-helical collagen and on related parameters[16].

	AcYaaOMe		**Triple-helical(ProYaaGly)$_7$**
Yaa	**$K_{trans/cis}$**	**γ_{ester} (cm^{-1})**	**Tm (^{o}C)**
Flp	6.7	1748	45
Hyp	6.1	1746	36
Pro	4.6	1743	6-7
hyp	2.4	1725	< 5
flp	2.5	1754	< 2

In view of this, in this laboratory analysis of comparative kinetics of *trans-cis* isomerization in 4*R*/*S*-amino proline-N^1-acetamides (**3**,**4**) was done using variable temperature ^{1}H NMR. At 25^{o}C, $K_{Z/E}$ for 4*S*-aminoproline **4** (6.4) was higher than that of 4*R*-aminoproline **3** (3.7) suggesting a higher preponderance of *trans* amide in 4*S*-aminoproline **4** and hence a better propensity for triplex formation.

2.2: Objective of the present work

The C4-*exo* preferred conformation of 4*R*-aminoproline and its observed $K_{trans/cis}$ equilibrium constant are well suitable for Y position of the collagen peptide sequence (X-Y-Gly), which was evidenced by the formation of stable triple-helix by 4*R*-aminoproline at Y position. However, according to the observed pyrrolidine ring puckering in crystal structure of peptide $[(Pro\text{-}Pro\text{-}Gly)_{10}]_3$, the C4-*endo* preferred conformation of 4S-aminoproline is well suited for X position of the collagen peptide sequence(X-Y-Gly). The objective of the present work is to examine the compatibility of conformation of pyrrolidine ring adopted by 4*R* and 4*S* aminoprolines at both X and Y position of the collagen peptide $(X\text{-}Y\text{-}Gly)_n$**.** This chapter deals with the chemical synthesis and biophysical studies of the collagen peptides $(X\text{-}Y\text{-}Gly)_n$ having two modified amino acids, (2*S*,4*R*) and (2*S*,4*S*) 4-aminoprolines at X and Y position.

The specific objectives of this chapter are

1) Synthesis of N^1-Fmoc, N^4-Boc (2*S*, 4*R*) aminoproline and N^1-Fmoc, N^4-Boc (2*S*, 4*S*) aminoproline (Scheme 2.1 and 2.2)
2) Solid phase synthesis of collagen peptides incorporating above monomers using Fmoc chemistry.
3) Cleavage of the peptides from solid support followed by purification and characterization of the peptides.
4) Triple-helical stability study of peptides using CD spectrometry at different pHs, at different salt concentration conditions, at different concentrations of peptide, and at different solvent condition.

2.3: Results

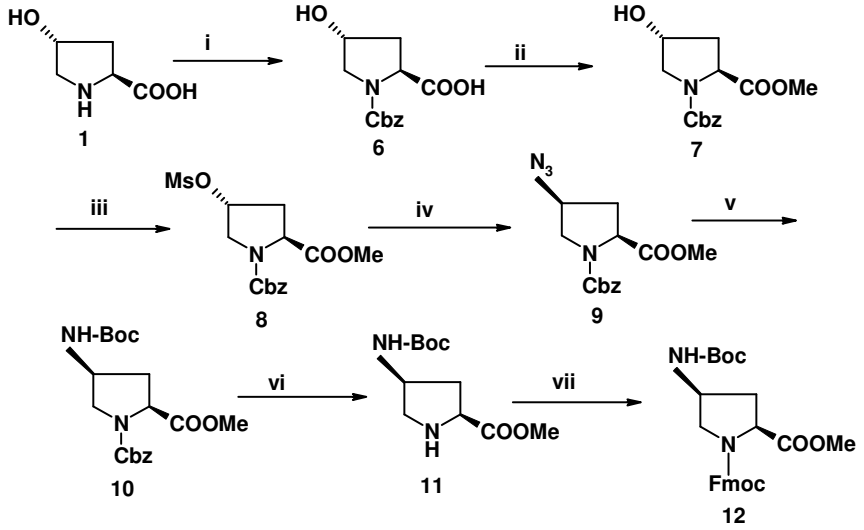
2.3.1: Synthesis of fully protected (2*S*,4*R*) and (2*S*,4*S*) aminoproline monomers

To achieve the synthesis of peptides (**19**-**22**), monomers **12** and **18** were chosen with a combination of N^{α}-Fmoc and N^{γ}-Boc protection. The peptide synthesis can also be achieved by choosing combination of alternative N^{α}-Boc and N^{γ}-Fmoc protection. The solid phase peptide synthesis using Fmoc strategy offers several advantages. The Fmoc group can be removed under mild basic condition using secondary base such as piperidine in DMF. This deprotection method excludes the change of solvent as both deprotection and coupling reaction are carried out in DMF, thus reducing the number of washing steps. The removal of t-Boc group requires harsher conditions such as 50% TFA/DCM and yields trifluoroacetate salt at the N-terminus, requiring an additional neutralizing step before the coupling reaction. In contrast, removal of Fmoc group results in free amine, which can be directly subjected to coupling reaction.

Synthesis of orthogonally protected (2*S*,4*R*) and (2*S*,4*S*) 4-aminoproline monomers (**12** and **18**) were achieved in nine steps from naturally occurring trans-4-hydroxyproline **5** (Scheme 2.1). Upon treatment of trans-4-hydroxyproline **1** with benzyloxychloroformate in water/dioxane, in the presence of Na_2CO_3, N^{α}-benzyloxycarbonyl-4-hydroxyproline **6** was obtained, as evidenced by the appearance of aromatic (δ 7.3-7.2) and benzylic (δ 5.1) signals in ^{1}H NMR spectrum. This upon treatment with DMS in anhydrous acetone with K_2CO_3 yielded corresponding methyl ester **7**, which showed two signals δ 3.75 and 3.54 (minor and major isomers respectively, because of rotamers) in the proton NMR spectrum corresponding to the ester -CH_3. The 4-OH group of **7** was converted to the corresponding mesyl derivative **8**

by treatment with methanesulfonylchloride, in DCM in presence of Et_3N (^{1}H NMR: δ 3.0, 3H SO_3CH_3). The treatment of mesylate compound **8** with NaN_3 in DMF at 55-60 ^{0}C resulted in an S_N2 displacement of 4-O-mesyl group to yield the azide **9**.

Scheme 2.1



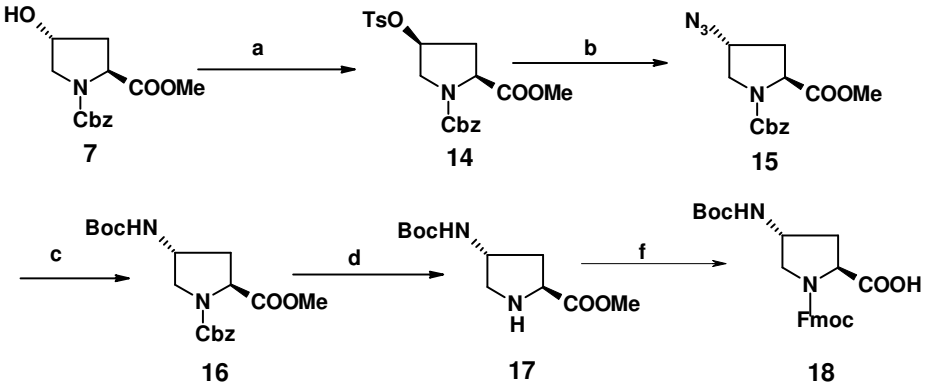
Reagents and conditions: i) 50% Cbz-Cl in toluene, $NaHCO_3$, H_2O; ii) K_2CO_3, DMS, acetone, reflux; iii) MsCl, Et_3N, DCM, 0^{o}C; iv) NaN_3, DMF, 50-55^{o}C, 8hrs; v) Raney-Ni/H_2, $(Boc)_2O$, ethylacetate; vi) 10%Pd-C/H_2 80psi, methanol,8hrs vii) a) 2N NaOH, MeOH, b) Fmoc-Cl, Na_2CO_3, water: dioxane (1:1).

A peak at 2108 cm^{-1} corresponding to azide functional group seen in 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 and the 4-aminogroup was in situ protected, using Boc anhydride to yield N^γ-t-Boc compound **10** (^{1}H NMR: δ 7.32-7.28, aromatic protons; δ 5.20-4.96, Benzylic protons; δ 1.40-1.38, t-Boc protons). Prolonged reaction time at high pressure of hydrogen gas (45 psi) resulted in partial deprotection of N^1- Cbz group and the resulting free amine reacted with $(t\text{-}Boc)_2O$ to give N,N-di-Boc product. The reaction was therefore continually monitored and was completed in 1.5-2 hrs. The N^1-benzyloxycarbonyl group of compound **10** was removed by hydrogenation with 10%

Pd-C as a catalyst to yield methyl ester **11**. This was hydrolyzed with 2N NaOH in methanol and N1 of the resulting product was protected with Fmoc using 9-fluorenylmethylchloroformate to yield the fully protected 4S-aminoproline (amp) monomer **12**.

The 4*R*-OH group of compound **7** was converted into 4*S*-O-tosylate **14** using Mitsunobu reaction with DIAD/PPh_3 and methyl-p-toluenesulfonate, resulting in an inversion of configuration at C-4. The reaction of tosylate **14** under S_N2 condition with NaN_3 in DMF resulted in a second inversion of stereochemistry at C-4, yielding (2*S*, 4*R*)-N^1-Cbz-azidoprolinemethylester **15**. (^{1}H NMR δ 7.4, aromatic protons; δ 5.4-5.35, benzylic protons; δ 4.25-4.15, C^γ, proton). A peak at 2108 cm^{-1} characteristic of azide was seen in IR spectrum of **15**. The azide **15** was selectively reduced to the corresponding amine using Raney-Ni without affecting the N^1-Cbz group followed by in situ Boc protection, with Boc anhydride, yield N^γ-t-Boc compound **16.**

Scheme 2.2



Reagents and conditions: a) PPh_3, DIAD, Ts-OMe, THF; b) NaN_3, DMF, 50-55^{o}C, 8hrs; c) Raney-Ni/H_2, $(Boc)_2O$, ethylacetate; d) 10%Pd-C/H_2 80psi, methanol,8hrs f) i) 2N NaOH, MeOH, ii) Fmoc-Cl, Na_2CO_3, water: dioxane (1:1).

The deprotection of N^1-Cbz group was achieved by hydrogenation using 10% Pd-C in methanol. The appearance of ninhydrin positive spot on TLC and disappearance of

aromatic signals of Cbz group in ^{1}H NMR indicated complete deprotection of Cbz group. The methyl ester group of resulting free amine compound was hydrolyzed with 2N NaOH in methanol, and the amine was protected with Fmoc, by reacting with 9-fluorenylmethylchloroformate to yield the fully protected 4*R*-aminoproline (Amp) **18** (Scheme 2.2). The appearance of signals in the aromatic region typical of Fmoc group (δ 7.56-745 (3m, 5H, ArH & CH of Fmoc), peaks at 453 (M + H) in the LCMS confirmed the structural integrity of the monomers **13** and **18**.

2.3.2: Solid phase peptide synthesis (SPPS)

Solid phase peptide synthesis[18] is based on 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 using either a coupling reagent or pre-activated protected amino acid derivative. The resulting peptide is attached to the resin, via a linker, 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 trifuoroacetic acid (TFA) and the Fmoc protecting group by piperidine. Final cleavage of the peptidyl resin and the side-chain deprotection requires strong acid, such as hydrogen fluoride (HF) or trifuoromethanesulphonic acid (TFMSA) in case of Boc chemistry, and TFA in Fmoc chemistry. Dichloromethane (DCM) and N,N-dimethylformamide (DMF) are the primary solvents used for resin deprotection, coupling and washing.

Peptide synthesis can be carried out in batchwise or continuous flow manner.[19] In the former technique, the peptidyl resin is contained in a filter reaction vessel and reagents are added and removed under manual or computer control. In the continuous flow method, the resin is contained in a column through which reagent and solvents are pumped continuously again under manual or automatic control.

2.3.3: Comparison of Fmoc and Boc chemistry

The development of Fmoc SPPS arose out of concern that repetitive TFA acidolysis in Boc-group deprotection could lead to alternation of sensitive peptide bonds as well as acid catalyzed side reactions. In Fmoc synthesis, the growing peptide is subjected to mild base treatment using piperidine during Fmoc deprotection and TFA is required only for the final cleavage and deprotection of peptidyl resin. By contrast, cleavage and deprotection in Boc strategy requires the use of dangerous HF.

Currently Fmoc chemistry is used for routine synthesis of peptides by solid phase method. 9-Fluorenylmethyloxycabonyl (Fmoc) group is chosen as N^{α} – protection, which can be cleaved efficiently by mild secondary base such as piperidine. Side chain protection is chosen as t-butoxycarbonyl (Boc) which is removed by 50% TFA in DCM and extremely stable to basic conditions. The linker group that connects or binds the peptides to the resin is chosen such that the side chains and the linker are cleaved in one step.

2.4: Present work

In the present work, peptides were synthesized by manual solid phase synthesis on readily available rink amide resin using standard Fmoc chemistry (Scheme 2.3), which upon cleavage directly yielded the peptide-C-terminal amide. Commercially available Kieselghur supported N,N-dimethylacrylamide resin (with Rink-amide linker) was used. The resin bound Fmoc group was cleaved with 20% piperidine/DMF and the monomers were coupled as free acids using *in situ* activation procedure with 3 eq. of aminoacid, HBTU as a coupling reagent and HOBt/DIPEA as catalyst and recemization-suppressant. The deprotection and coupling reactions were monitored using qualitative Ninhydrin (Kaiser)[20] test for Gly and Chloranil test[21] for iminoacids. In case of a positive color test after the first coupling which indicates incomplete reaction, the coupling was repeated. To avoid deletion of sequences, a capping step with Ac_2O/ Py in DCM was performed. The terminal amino group of the final peptide was capped with Ac_2O and the peptide was cleaved from the resin using 5% TFA in DCM. As the t-Boc removal requires stronger acidic conditions, the deprotection of side chain t-Boc on Amp residues of the peptide was carried out with 50% 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 and to prevent such side reactions during peptide-cleavage and t-Boc deprotection, 0.1% TIS (Triisopropylsilane) was used as a scavenger. The N-terminal acetaylated and C-terminal amidated peptides were purified on semi-preparative RP-4 HPLC column using water acetenitrile gradient to which 0.1% TFA added as modifier.

Scheme 2.3

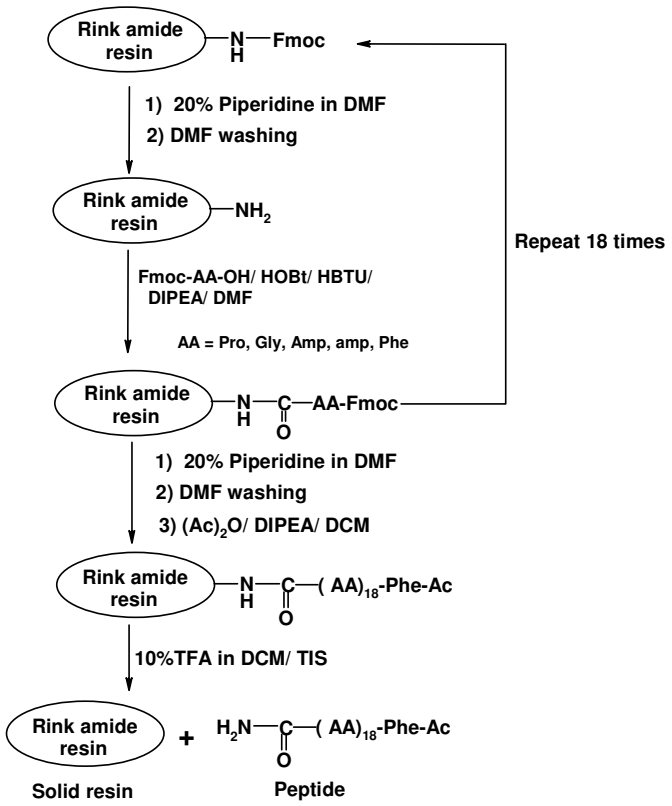


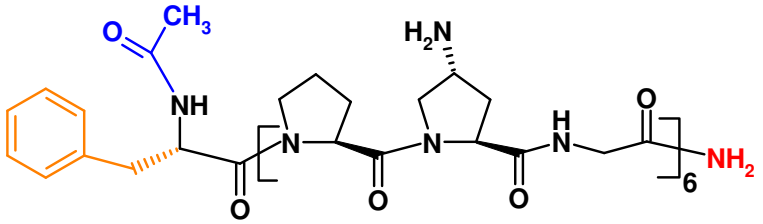
Figure 2.3: peptides used in the present study

The following peptides (Fig.2.4), were synthesized for the present study by incorporating the modified aminoacids (2*S*,4*R*) 4-aminoproline **18** and (2*S*,4*S*) 4-aminoproline **13**, at **X** and **Y** positions of the collagen model peptide (X-Y-Gly)$_n$. The peptides Ac-Phe(Pro-Amp-Gly)$_6$-NH$_2$ **19** and Ac-Phe(Pro-amp-Gly)$_6$-NH$_2$ **21** contain 4*R*-aminoproline (**Amp**) and 4*S*-aminoproline (**amp**) at **Y** position respectively, peptides Ac-Phe(Amp-Pro-Gly)$_6$-NH$_2$ **20** and Ac-Phe(amp-Pro-Gly)$_6$-NH$_2$ **22** contain 4*R*-aminoproline (**Amp**) and 4*S*-aminoproline(**amp**) at **X** position. The purity of the peptides was determined using analytical RP-18 HPLC and the peptides were found to be greater than 95% pure. The structural integrity of the peptides were further

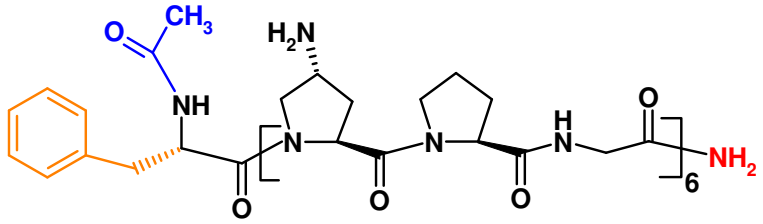
conformed by MALDI-TOF mass spectrometry which agreed closely with the calculated values (Table-2)

Table 2: Calculated and observed masses for peptides **19-22**

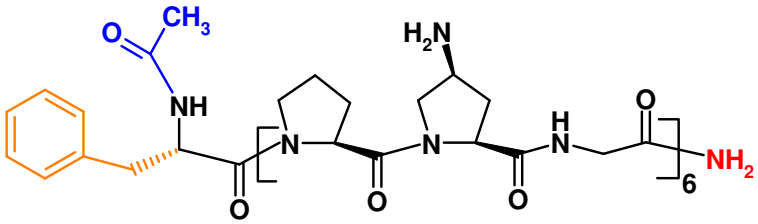
Peptide	Mol.formula	Mass (cal)	Mass (obs)
Ac-Phe(Pro-Amp-Gly)$_6$-NH$_2$ **19**	$C_{83}H_{124}N_{26}O_{20}$	1804.24	1808
Ac-Phe(Amp-Pro-Gly)$_6$-NH$_2$ **20**	$C_{83}H_{124}N_{26}O_{20}$	1804.24	1806
Ac-Phe(Pro-amp-Gly)$_6$-NH$_2$ **21**	$C_{83}H_{124}N_{26}O_{20}$	1804.24	1804
Ac-Phe(amp-Pro-Gly)$_6$-NH$_2$ **22**	$C_{83}H_{124}N_{26}O_{20}$	1804.24	1805



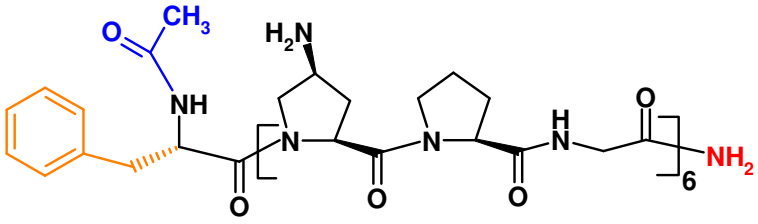
Peptide 19 (Pro-Amp-Gly)



Peptide 20 (Amp-Pro-Gly)



Peptide 21 (Pro-amp-Gly)



Peptide 22 (amp-Pro-Gly)

Figure 2.4: structure of peptide used in the present study; peptides **19** and **20** contain 4*R*-aminoproline in Y and X positions, peptide **21** and **22** contains 4*S*-aminoproline in Y and X position of the collagen peptide sequence (X-Y-Gly) respectively

2.4.1: End group ionization effect

Collagen is a parallel triple-helix with one residue shift register, all the N-termini of individual chains in the triple-helix point to one direction while the C-termini point to the other direction. At extremes of pH only one of the termini is completely ionized (-NH_3^+ at pH 3 and –COO^- at pH 12), while at any intermediate pH both groups are ionized to varying degree (Fig. 2.5). The presence of similar charged residues at the termini in a parallel triple-helix leads to destabilization.[22,23] To delineate and compare the relative contribution of 4-OH and 4-NH_2/NH_3^+ towards the triple-helix stability it was decided to cap the charges on the termini. For this purpose all peptides were capped at C-termini as amide and N-termini as acetyl group.

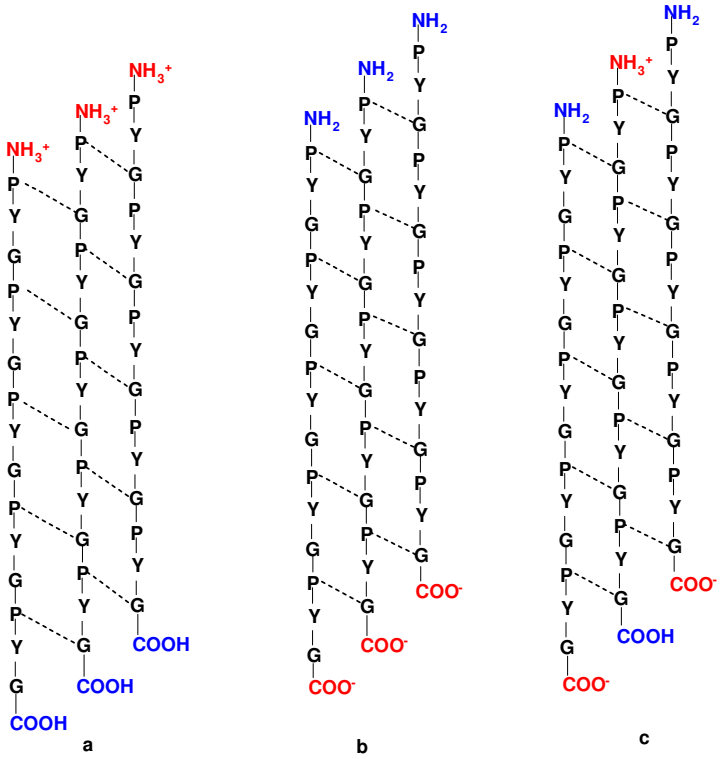


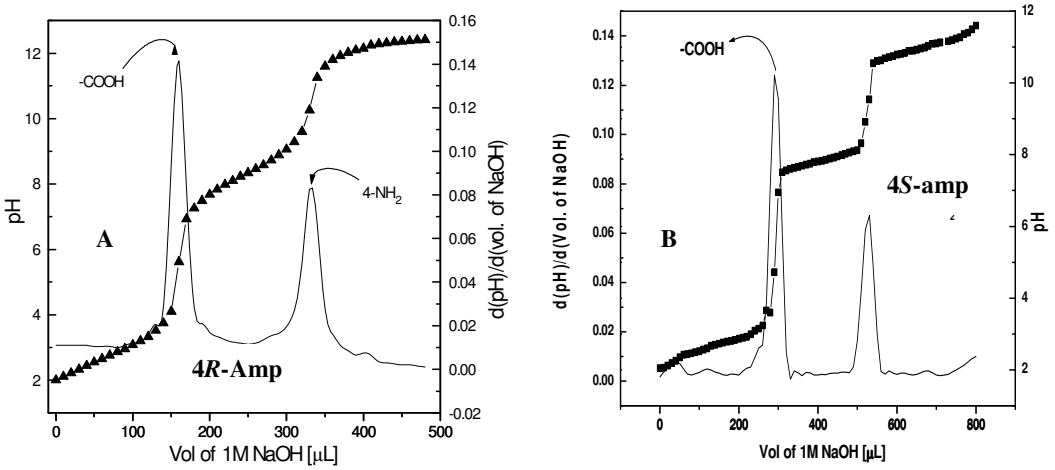
Figure 2.5: Schematic view of pH dependent repulsive forces at the termini in triple-helix, **a** protonated aminotermini at extreme acidic pH; **b** ionizied carboxy termini at extreme alkaline pH; **c** both groups are ionized to varying degree at intermediate pH.

2.4.2 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.[24] 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 have 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 this problem in the present study, the aminoacid 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 (ε_{259}=2 X 10^2 $mol^{-1}cm^{-1}$). All peptides containing phenylalanine and hence any effect on triple-helical stability from this residue is same for all peptides.

2.4.3: Determination of pKa of 4-amino group of Amp and amp monomers

Both (2*S*,4*R*) and (2*S*,4*S*)-aminoprolines contains free exocylic amino group, which is in free form in the peptide and can be easily protonated. In order to evaluate the pH dependence triple-helical strength, it is necessary to know the protonation state via the pKa of exocylic 4-amino group of both **Amp** and **amp** monomers. For this purpose pH changes was determined by titrating[25] the both α-N-protected **Amp** and **amp** monomers using 0.01N NaOH solution.



Direct pH meter reading up on titration was plotted against the volume of added NaOH (Fig 2.6). This plot shows two sharp pH transitions, one corresponding to the α-carboxylic acid group and a second corresponding to exocylic 4-aminogroup. Though both **Amp** and **amp** monomers display two sharp pH transitions, their exocylic 4-amino group has significance difference in the pKa value. The pKa value of **Amp** monomer is 10.2 while the pKa value of **amp** monomer is 9.3. This indicates that both monomers undergo protonation to different extents at physiological pH. This may be due to structural/electronic effects within the molecules that results in a stabilizing effect due to delocalization of the charge reducing the energy of the system. This results in increased acidity of the 4-NH_2 group either through H-bonding,[26] space field effect,[27a] bond inductive effect,[27a] solvation[27b] or stereoelectronic anomeric & gauche effect.[27c] However it should be noted that the pKa of exocylic 4-amino group may change when it is part of a peptide chain, due to other local environmental effects.

2.5: Characterization of triple-helical structures

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 structures exhibit fingerprint CD spectra[28]. These spectra are characteristic by the presence of a large negative band around 200nm, a crossover at around 213nm, and a small positive band around 215-227nm (Fig 2.7).

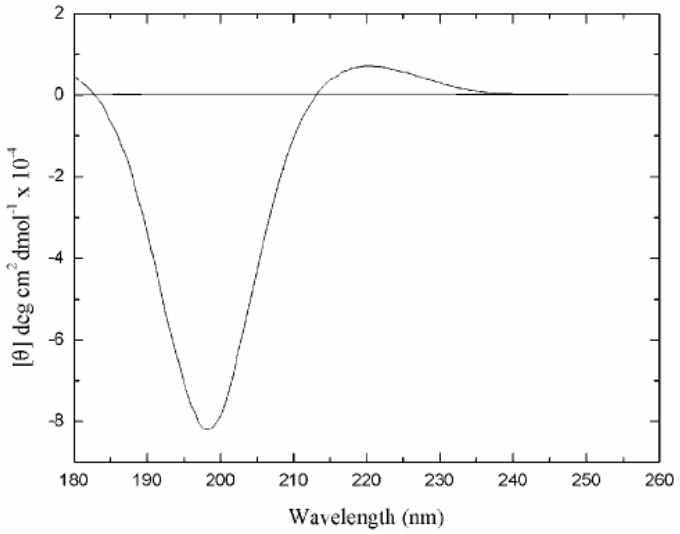


Figure 2.7: CD spectra of collagen triple-helix structure.

In order to investigate the conformational behavior of peptides **19**-**22**, CD spectra were measured under various pH values, salt concentrations and solvent conditions.

2.5.1: Concentration dependedence on triple-helix formation of peptides 20-22

Formation of triple-helical structure is a concentration dependent phenomenon. Goodman et al.[29] have shown that higher concentration for a single chain increases the triple-helical conformation 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_{pn}) has been proposed to quantitate the triple-helical strength[29]. In the present study, this parameter has been used to determine the formation of concentration dependent triple-helix at pH 3.0 for all peptides in the presence of 0.1M NaCl. Figure 2.8a shows the CD spectra of peptide **20** (Amp-Pro-Gly) taken at 25 ^{0}C in the concentration range of 0.05mM-0.30mM.

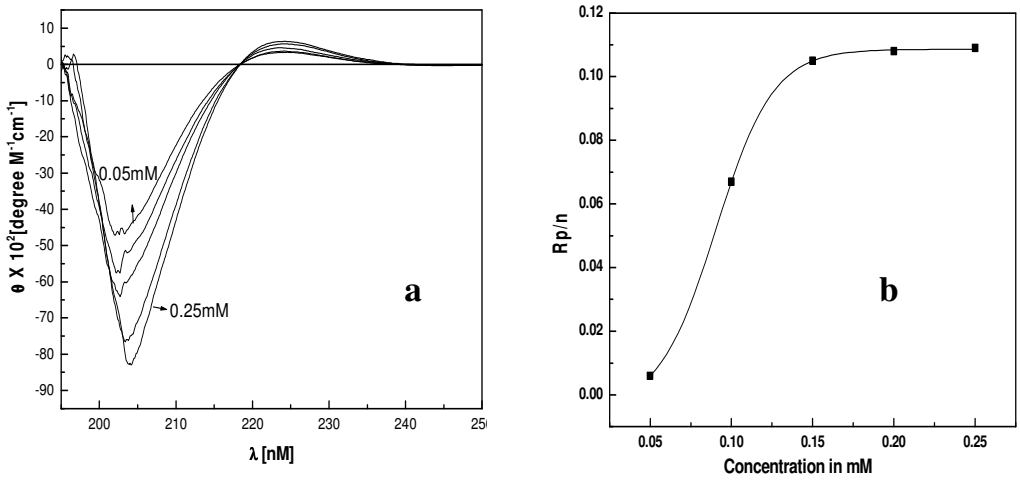


Figure 2.8: CD spectra of **a** peptide **20** (Amp-Pro-Gly) at 25^{o}C at concentration from 0.05mM-0.30mM in steps of 0.05mM (acetate buffer pH 3.0, 0.1M NaCl.); **b** plot of Rpn values deduced from this spectra against the concentration of peptide **20**.

In the entire concentration range peptide **20** (Amp-Pro-Gly) shows similar positive and negative maxima at 223nm and 204nm respectively, which are characteristics of collagen like triple-helical structure, with the magnitude of positive and negative bands varing as a function of concentration. Importantly, all the spectral traces pass through an isobestic point at 217nm. Figure 2.8b shows a plot of R_{pn} values derived from this spectra against concentration peptide **20** (Amp-Pro-Gly) at pH 3.0.

The R_{pn} value increases rapidly from 0.05 mM through 0.10 mM to reach a near saturation at 0.15 mM and remains nearly constant thereafter. A critical triple-helical concentration of ~0.15 mM is derived from these plots. Hence, all further studies are performed at a concentration of 0.2 mM peptides where peptides are in fully associated form.

Figure 2.9a shows the CD spectra with varying concentrations (0.05mM-0.25mM) of peptide **21** (Pro-amp-Gly). In the entire concentration range the peptide **21** (Pro-amp-Gly) displays the CD spectra with identical positive band ~225nm and negative band ~202nm. Like peptide 20 (Amp-Pro-Gly) this peptides also shows isobestic point at 218nm. 203nm.

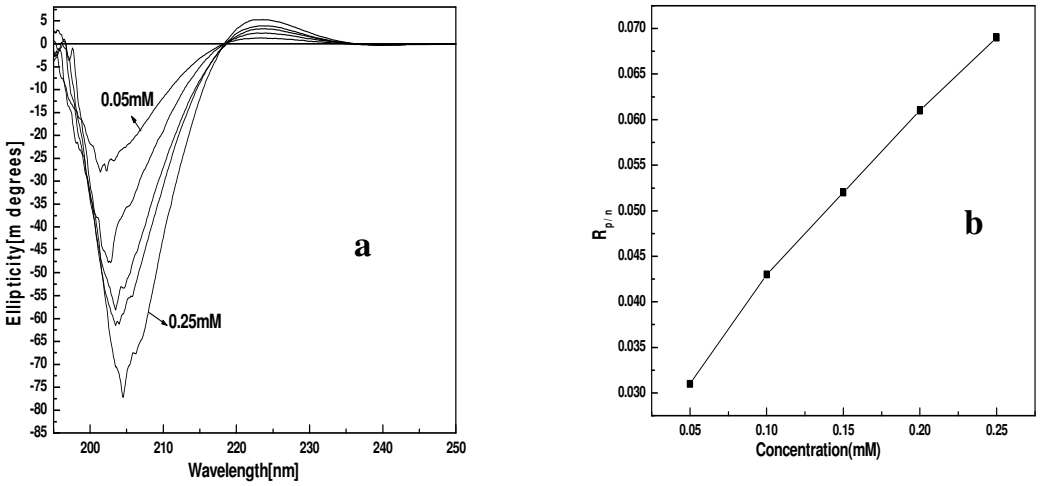


Figure 2.9: CD spectra of **a** peptide **21** (Pro-amp-Gly) at 25^{o}C at concentration from 0.05mM-0.30mM in steps of 0.05mM (acetate buffer pH 3.0, 0.1M NaCl.); **b** plot of Rpn values deduced from this spectra against the concentration of peptide **21.**

Despite very low R_{pn} valves, unlike peptide **20** (Amp-Pro-Gly) the plot of R_{pn} values against concentration of peptide **21** (Pro-amp-Gly) does not exhibit any saturation over the concentration range 0.05-0.25mM (Fig 2.9b). This may due to the single chain conformation of peptide **21** (Pro-amp-Gly).

The CD spectra of peptide **22** (amp-Pro-Gly) (Fig 2.10a) under the concentration range 0.05mM-0.25mM shows the identical positive maxima at 223nm and negative minima at 204nm. Interestingly, peptide **22** (amp-Pro-amp) shows two isobestic points at 216nm (for the concentration range 0.05-0.15mM) and at 217nm (for the concentration range 0.20-0.25) (see inset, Figure 2.15a).

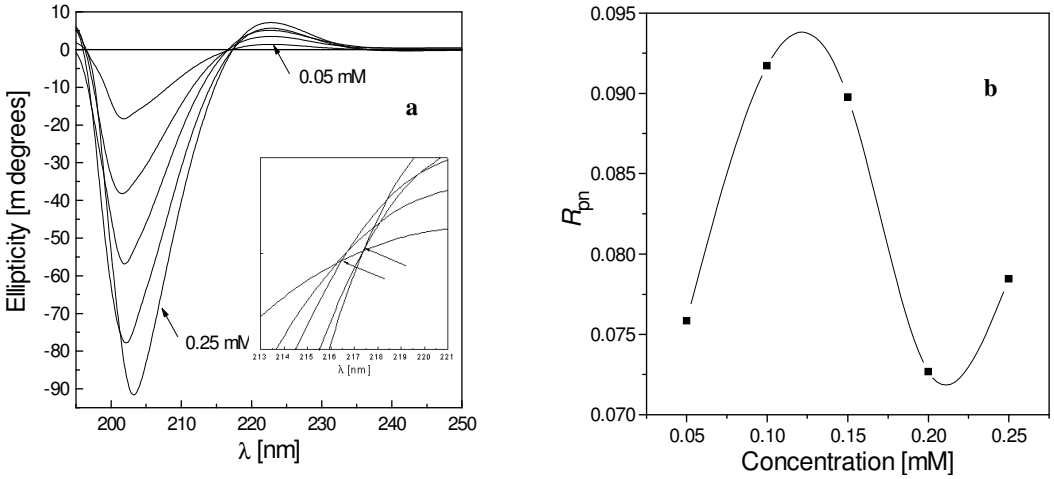


Figure 2.10: CD spectra of **a** peptide **22** (amp-Pro-Gly) at 25°C at concentration from 0.05mM-0. 25mM in steps of 0.05mM (acetate buffer pH 3.0, 0.1M NaCl.); **b** plot of Rpn values deduced from this spectra against the concentration of peptide **22**.

An attempt was made to determine the critical triple-helical concentration for the peptide **22** (amp-Pro-Gly) by plotting R_{pn} Vs the concentration of peptide **22** (Figure 2.10b). Unlike peptide **20** (Amp-Pro-Gly) the R_{pn} values does not increase progressively with concentration of peptide **22**.

The detection of the triple-helical conformations of collagen and collagen model peptides has been carried out using a variety of experimental techniques. For example, X-ray diffraction and electron microscopy have been used to study the size and shape of collagen fibrils as well as synthetic collagen like polymers.[30] X-ray crystallography has been employed to investigate the triple-helical packing of collagen segments and

synthetic collagen analogues in the solid state.[31] In solution, equilibrium sedimentation and light scattering have been utilized to study molecular weight-dependent properties, such as the formation and denaturation of triple-helices.[32]

CD spectroscopy is the most frequently used technique to study triple-helical conformations in solution. As described earlier, the natural collagen triple-helix has unique CD spectrum with small positive peak around 223nm, a crossover at 213nm, and a large trough around 197-205nm. It has been found that polyproline I and polyproline II like structures also exhibit CD spectra similar to that of triple-helices.

Figure 2.11 shows the CD spectra of 0.2mM solution of peptides **20** (Amp-Pro-Gly) in acidic (20mM acetate buffer pH 3.0) neutral (20mM phosphate buffer pH 7.0) and basic (20mM borate buffer pH 9.0 and pH 12) conditions in presences of 0.1M NaCl.

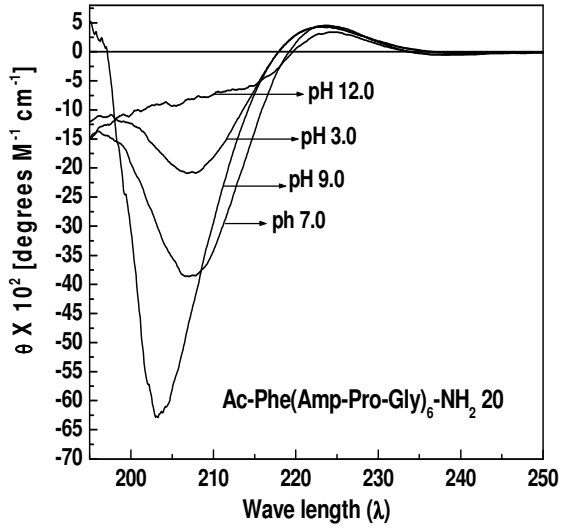


Figure 2.11: CD spectra of peptide **20** Ac-Phe(Amp-Pro-Gly)$_6$-NH$_2$ taken at different pH conditions.

The CD spectra of Peptide **20** (Amp-Pro-Gly) at acidic and neutral conditions resembles the CD spectra of native collagen. In both acidic and neutral conditions the

positive maxima appears at 223nm and negative minima appears at 204nm. But the cross over takes place at 213nm in acidic condition and at 215nm in neutral condition. The Rpn values in both acidic and basic conditions are in the range of triple-helical conformation, though higher Rpn value is observed in acidic compared to neutral condition (Table 3). The CD spectrum at pH 9.0 resembles the CD spectra of collagen peptide with respect to the appearance of positive maxima, negative minima and crossover point. But the Rpn value, unlike at acidic and neutral conditions is not in the range of triple-helical conformation. At pH 12.0 only positive maxima is appears at 213nm and does not show the negative minima.

Table 3: Rpn values of peptide **20** (Amp-Pro-Gly) at different pH.

Peptide 20 Ac-Phe(Amp-Pro-Gly)$_6$-NH$_2$			
pH	**+ve band (nm)**	**-ve band (nm)**	**Rpn**
3.0	223	206	0.31
7.0	223	206	0.24
9.0	223	203	0.097
12.0	223	-	-

The appearance of collagen like CD spectrum is neither a sufficient evidence for the presence of triple-helical structures, nor gives any information about the triple-helical strength. In order to determine the relative triple-helical strength of peptides **19-22** CD-thermal denaturation study was carried out.

Upon heating triple-helical structures undergo a cooperative triple-helix to coil transition. This change of conformation is reflected in several spectroscopic properties including degree of the molar ellipticity typically around 225nm in the CD spectra. Therefore, the triple-helix to coil transition was monitored by observing the changes in ellipticity at 225nm with temperature. The thermal denaturation curves so obtained show a decrease in molar ellipticity with increase in temperature for the peptides.

To confirm the triple-helix formation and its stability, thermal denaturation study was carried out for the peptide **20** (Amp-Pro-Gly) at both acidic and basic conditions. Figure 2.12 shows the CD-thermal denaturation curves of peptide **20** at different pH condition

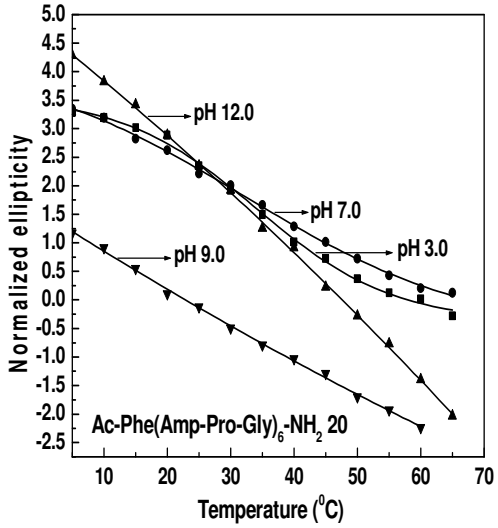


Figure 2.12: Thermal denaturation curves of peptide **20** Ac-Phe(Amp-Pro-Gly)$_6$-NH_2 at different pH conditions. pH 3.0 (20mM acetate buffer), pH 9.0 (20mM phosphate buffer, pH 9 and 12 (20mM borate buffer). All buffers contain 0.1M NaCl.

At pH 3.0 and 7.0 peptide **20** shows cooperative melting transition curves with increasing temperature. This indicates that peptide **20** (Amp-Pro-Gly) is associated in triple-helical structure at pH 3.0 and 7.0. The T_m values obtained from the integration of thermal denaturation curves indicate that at pH 3.0, the peptide **20** (Amp-Pro-Gly) forms a stable triple-helical structure with T_m of 36 ^{0}C compared to T_m of 33^{0}C at pH 7.0 (Table 4). At pH 9.0 and 12.0, peptide **20** shows a linear decrease in ellipticity with increasing temperature. This indicates that peptide **20** at basic condition does not associate into triple-helical conformation.

Table 4: CD-Tm values of peptide **20** (Amp-Pro-Gly) at different pH conditions.

Peptide 20 Ac-Phe(Amp-Pro-Gly)$_6$-NH$_2$	
pH	**T_m ^oC**
3.0	36
7.0	33
9.0	-nt-
12.0	-nt-

-nt- no transition

Figure 2.13 shows the CD spectra of 0.2mM solutions of peptides **21**Ac-Phe(Pro-amp-Gly)$_6$-NH$_2$ taken in acidic pH 3.0 (20mM acetate buffer), neutral pH 7.0 (20mM phosphate buffer) and basic pH 9.0 and 12.0 (20mM borate buffer) in presence of 0.1M NaCl. Under all pH conditions peptide **21** shows spectra characteristic of collagen like structure.

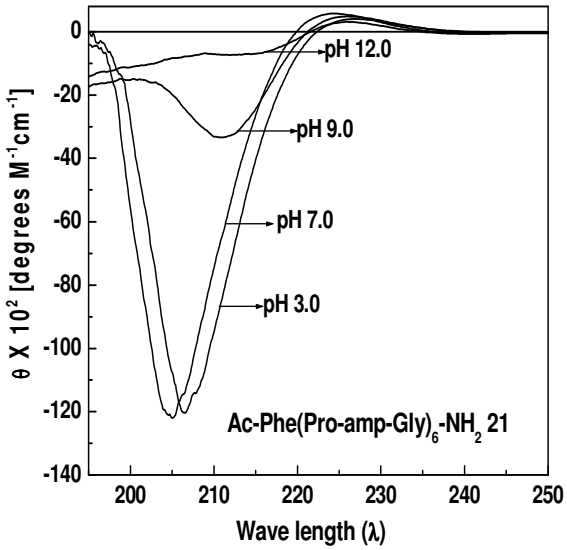


Figure 2.13: CD spectra of peptide **21** Ac-Phe(Pro-amp-Gly)$_6$-NH$_2$ taken at different pH conditions.

At pH 3.0 and 7.0 peptide **21** shows all positive bands at 223nm and very large negative bond around 205 and 206nm respectively. Similarly at pH 9.0 and 12.0 it shows small **Table 5:** Rpn values of peptide **21** (Pro-amp-Gly) at different pH. pn values at different pH conditions of peptide **21** (Pro-amp-Gly) are shown in Table 5.

Peptide 21 Ac-Phe(Pro-amp-Gly)$_6$-NH$_2$			
pH	**+ve band (nm)**	**-ve band (nm)**	**Rpn**
3.0	223	205	0.026
7.0	223	206	0.027
9.0	223	205	0.051
12.0	223	205	0.132

At acidic and neutral conditions peptide **21** (Pro-amp-Gly) shows very low Rpn values, at pH 9.0 it shows considerably high Rpn value but still it is not in the range of triple-helical structure. Interestingly at pH 12.0 peptide **21** shows a higher Rpn value in the range of triple-helical structure.

Figure 2.14 shows the CD-thermal denaturation curves of peptide **21** (Pro-amp-Gly) at different pH conditions.

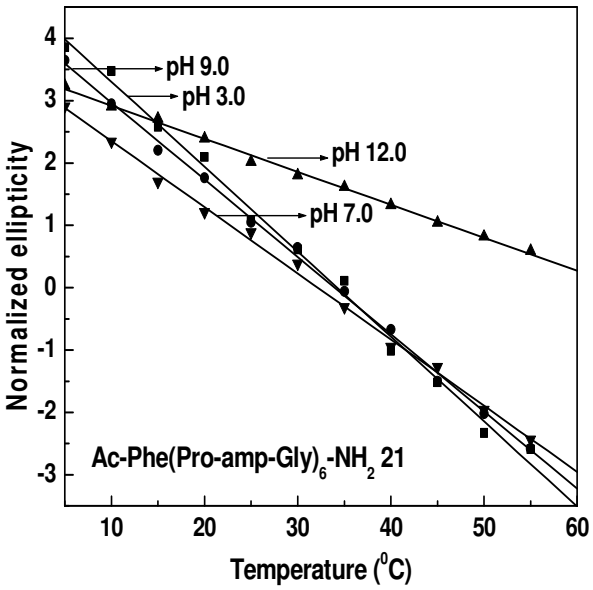


Figure 2.14: Thermal denaturation curves of peptide **21** Ac-Phe(Pro-amp-Gly)$_6$-NH$_2$ at different pH conditions. pH 3.0 (20mM acetate buffer), pH 9.0 (20mM phosphate buffer, pH 9 and 12 (20mM borate buffer). All buffers contain 0.1M NaCl.

Peptide **21** (Pro-amp-Gly) shows a linear decrease in ellipticity with increase in temperature for all pH conditions. This indicates that peptide **21** is not associated in triple-helical conformation under all pH conditions.

Figure 2.15 shows the CD spectra of 0.2mM solutions of peptides **22** Ac-Phe(amp-Pro-Gly)$_6$-NH_2 taken at acidic pH 3.0 (20mM acetate buffer), neutral pH 7.0 (20mM phosphate buffer) and basic pH 9.0 and 12.0 (20mM borate buffer) conditions in presence of 0.1M NaCl. At pH 3.0, 7.0 and 9.0 the peptide **22** (amp-Pro-Gly) shows CD spectra which is characteristic of collagen triple-helix.

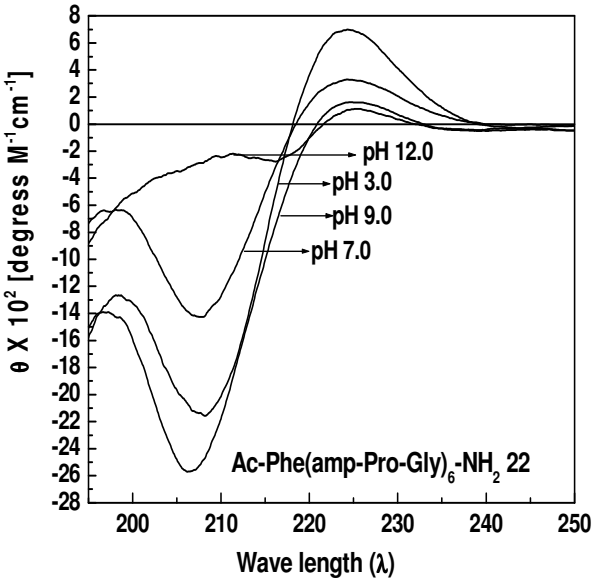


Figure 2.15: CD spectra of peptide **22** Ac-Phe(amp-Pro-Gly)$_6$-NH_2 taken at different pH conditions.

Under all pH conditions (except pH 12 in which negative band appears at 211nm) peptide **22** (amp-Pro-Gly) shows similar positive and negative band positions at 223nm and 206nm respectively. The Rpn values of peptide **22** (amp-Pro-Gly) at different pH conditions are shown in Table 6.

Table 6: Rpn values of peptide **22** (amp-Pro-Gly) at different pH conditions.

Peptide 22 Ac-Phe(amp-Pro-Gly)$_6$-NH$_2$			
pH	**+ve band (nm)**	**-ve band (nm)**	**Rpn**
3.0	223	206	0.39
7.0	223	206	0.26
9.0	223	206	0.18
12.0	223	211	0.19

Normalized ellipticity data at 225nm obtained from thermal denaturation study at different pH conditions for peptide **22** (amp-Pro-Gly) plotted against temperature is shown in Figure 2.16. Peptide **22** shows sigmoid transition at pH 3.0, 7.0 and 9.0 indicating the transition form triple-helix to coil structure at these pH conditions. It shows linear decrease in ellipticity with temperature at pH 12 indicating that at extreme basic condition, peptide **22** is not associated in triple-helical conformation.

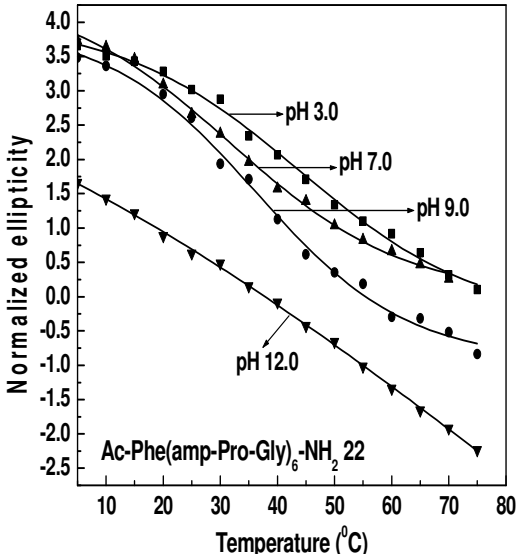


Figure 2.16: Thermal denaturation curves of peptide **22** Ac-Phe(amp-Pro-Gly)$_6$-NH$_2$ at different pH conditions. pH 3.0 (20mM acetate buffer), pH 9.0 (20mM phosphate buffer, pH 9 and 12 (20mM borate buffer). All buffers contain 0.1M NaCl.

The T_m data obtained for the transition from the first derivative curves of the sigmoid fit of ellipticity data are shown in Table 7. The T_m value indicate that at pH 3.0 peptide **22** (amp-Pro-Gly) forms a more stable triple-helical structure with T_m of 44^{o}C

followed by a T_m of 37^0C at pH 7.0 and T_m of 34^0C at pH 9.0' At pH 12.0, the peptide **22** is present in single chain polyproline II conformation.

Table 7: CD-Tm values of peptide **22** (amp-Pro-Gly) at different pH conditions.

Peptide 22 Ac-Phe(amp-Pro-Gly)$_6$-NH$_2$	
pH	T_m ^{0}C
3.0	44
7.0	37
9.0	34
12.0	-nt-

The T_m values of peptide **20** (Amp-Pro-Gly) and **22** (amp-Pro-Gly) decrease with increase in pH. But for peptide **19** (Pro-Amp-Gly) the T_m values decrease over the range 3.0-9.0, followed by an increase again at pH 12.0. Significantly, the triple-helical thermal stability of peptide **19** with 4*R*- Amp *at Y position* is always greater than that of peptide **20** with Amp at X position over the entire pH range. Moreover peptide **19** (Pro-Amp-Gly) forms a trip-helix structure in all pH conditions, but peptide **20** (Amp-Pro-Gly) with 4*R*-Amp *at X position* forms triple-helix structure only at acidic (pH 3.0) and neutral (pH 7.0) conditions. This suggests that 4*R*-NH_2 group, in both protonated and a non-protonated state has positive roles in stabilizing triple-helix at Y position. At X

Table 8: *CD-T_m values of peptides* ***19-22***

Peptide		**T_m values (^{0}C)**			
X - Y - Gly		**pH 3.0**	**pH 7.0**	**pH 9.0**	**pH 12.0**
Ac-Phe(Pro-Amp-Gly)$_6$-NH$_2$	**19**‖	60	54.7	26	46
Ac-Phe(Amp-Pro-Gly)$_6$-NH$_2$	**20**	36	33	-nt-	-nt-
Ac-Phe(Pro-amp-Gly)$_6$-NH$_2$	**21**	-nt-	-nt-	-nt-	-nt-
Ac-Phe(amp-Pro-Gly)$_6$-NH$_2$	**22**	44	37	34	-nt-

-nt- no transition, ‖ from reference 8.

position 4*R*-NH_2 group has positive role only in protonated form, and the non-protonated form of 4*R*-NH_2 group at X position destabilizes the triple helix. In contrast peptide **21** (Pro-amp-Gy) with 4*S*- amp *at Y position* doesnot form triple-helix over the

entire pH range, whereas peptide **22** (amp-Pro-Gly) with amp *at X position* forms a stable triple helix at acidic (pH 3.0), neutral (pH 7.0) and basic (pH 9.0) conditions, but doesn't form triple-helix at pH 12.0. This suggests that 4*S*-NH_2 group, in protonated has positive role in stabilizing triple-helix at X position, but at Y position 4*S*-amp destabilizes the triple-helix over the entire pH range That is in contrast with 4*R*-Amp peptide which stabilizes the triple-helical structure over entire pH range at Y position.

2.5.2: Salt effects on the triple-helical stability

Hydration is very important for the three-dimensional structure and activity of proteins.[33] In solution, they possess a conformational flexibility, which encompasses a wide range of hydration states, not seen in the crystal or in non-aqueous environments. Equilibrium between these states will depend on the activity of the water within its microenvironment *i.e.* the freedom that the water has to hydrate the protein.[34] Thus, protein conformations demanding greater hydration are favored by more active water (*e.g.* high density water containing many weak, bent and/or broken hydrogen bonds) and 'drier' conformations are relatively favored by lower activity water (*e.g.* low-density water containing many strong intra-molecular aqueous hydrogen bonds). Surface water molecules are held most strongly by the positively-charged basic amino acids. The exchange of surface water (and hence the perseverance of the local clustering and the overall system flexibility) is controlled by the exposure of the groups to the bulk solvent (*i.e.* greater exposure correlates to greater flexibility and freer protein chain movement).[35]

The interaction of protein molecule with solvent or with other molecules is determined primarily by its surface, with most favorable interactions provided by

charged or polar residues. The folding of proteins depends on the same factors as they control the junction zone formation in some polysaccharides; *i.e.* the incompatibility between the low-density water (LDW) and the hydrophobic surface that drives such groups to form the hydrophobic core. In addition, water acts as a lubricant,[36] so easing the necessary peptide amide-carbonyl hydrogen bonding changes. The biological activity of proteins appears to depend on the formation of a 2-D hydrogen-bonded network spanning most of the protein surface and connecting all the surface hydrogen-bonded water clusters.[37] Such a water network is able to transmit information around the protein and control the protein's dynamics, such as its domain motions.[38]

Additives such as salts, affect the solubility and secondary and tertiary structural stability of the proteins. The presence of salt and the resulting increase in ionic strength decreases the electrostatic free energy of the protein and enhances its structural stability. At high concentrations, salts decrease the solubility of proteins and denature the folded structure by disrupting the structure of water.[39]

Water is believed to be an integral component of collagen structure, as was confirmed by several spectroscopic and X-ray crystal structure studies. Solutes influencing the structure of water can also influence the stability of the triple-helix. The presence of salt is known to have significant effect on the stability of natural collagens and collagenous peptides. This effect was demonstrated almost 40 years ago for gelatin↔collagen transition in acidic solutions.[40] More recently several studies on collagen peptides using calorimetry have emphasized on such influences.[41]

In order to evaluate and elucidate the mechanism of triple-helix stabilization by (2*S*,4*R*) and (2*S*,4*S*) aminoprolines at Y and X position respectively, T_m measurements

of both peptide **19** (Pro-Amp-Gly) with 4*R*-Amp at Y position, and peptide **22** (amp-Pro-Gly) with 4*S*-amp at X position, were carried out at pH 7.0 with varying concentrations of NaCl. The T_m values are given in Table 9.

Conc of NaCl (mM)	**T_m values (^{0}C)**	
	Peptide 19	**peptide 22**
0	51	39
50	54	41
100	60	44
150	57	42

♀ 20 mM phosphate buffer

Figure 2.17 shows the plot of triple-helical stabilities against concentration of NaCl. The T_m values for both Amp peptide **19** (Pro-Amp-Gly) and amp peptide **22** (amp-Pro-Gly) show a steady increase upto 100mM and a slight decrease at higher salt concentrations. This stabilization by NaCl arises form the screening of interstrand electrostatic repulsion caused by 4-NH_3^+ groups. Beyond 150mM NaCl, breaking of the hydrogen bond network and water-structure may lead to a decrease in the triple-helical stability and hence a decrease in T_m.

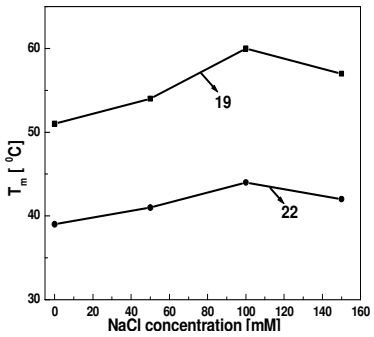


Figure 2.17: plot of T_m Vs concentration of NaCl of Peptide **19** and Peptide **22**

2.5.3: Effect of ethylene glycol on the stability of triple-helices

Intramolecular peptide (amide) hydrogen bonding makes a major contribution to protein structure and stability, but is effective only in the absence of accessible competing water. Even the presence of water molecules close-by causes the peptide hydrogen bonds to lengthen,[42] loosening the structure. Water molecules can 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 long-lived linkages stabilizing protein-ligand and protein-protein interfaces.[43] 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.[44] 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.[45] It is possible to strengthen the intramolecular peptide hydrogen bonds using polar solvent but not competing with hydrogen bonding of amide nitrogen. Polyols and sugars are known to offer protection to most proteins,[45] including collagen triple-helix against thermal denaturation.[46] 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.[47,29]

Figure 2.18 A shows the CD spectra of 0.2mM solution of peptides **19**-**22** taken in a 3:1 mixture of ethylene glycol and water. All the peptides show CD spectra which resembles the CD spectra of collagen like triple-helix in EG: Water system.

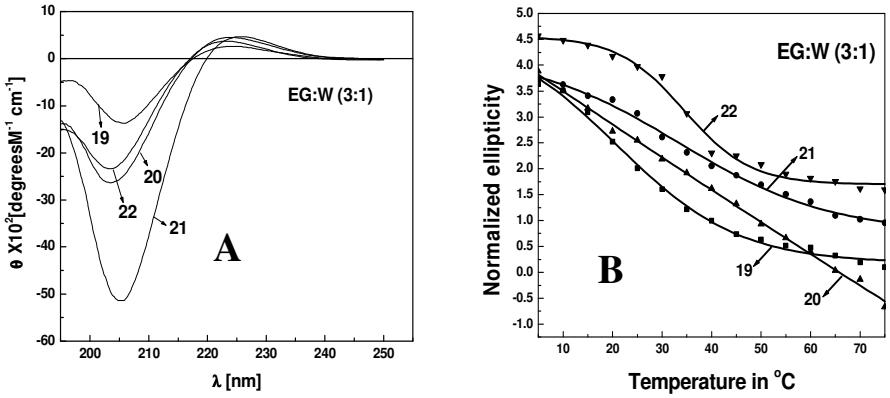


Figure 2.18 A. CD spectra at 10 ^{0}C and **B**. CD-Tm curves of 0.2mM solutions of peptides **19-22** in 3:1 (v/v) EG: H_2O mixture.

In comparison to aqueous solution, the spectrum of **21** (Pro-amp-Gly) with 4*S*-amp in Y position is characterized by the appearance of significant positive band at 225nm with a large negative band at 202nm in EG: H_2O (3:1) system.

Table 10: Rpn values and CD-T_m data of peptides **19-22** in 3:1(EG:H_20) system

Peptide	**Rpn**	**T_m ^{0}C**
Ac-Phe(Pro-Amp-Gly)$_6$-NH_2 **19**	0.14	23
Ac-Phe(Amp-Pro-Gly)$_6$-NH_2 **20**	0.19	35
Ac-Phe(Pro-amp-Gly)$_6$-NH_2 **21**	0.076	-nt-
Ac-Phe(amp-Pro-Gly)$_6$-NH_2 **22**	0.20	39

-nt- no transition

The CD-T_m (Fig 2.19B) for the peptide **19** (Pro-Amp-Gly) is lower than the peptide **22** (amp-Pro-Gly) (Table 10). This behavior of peptide **19** is opposite to the behavior in aqueous solutions where under all pH conditions, it forms more stable triple-helices than the other peptides. The peptide **22** with 4*S* amp at X position forms a more stable triple-helix in EG: H_2O (3:1), and also peptide **20** with 4*R*-Amp in X

position forms more stable triple-helix compared to the peptide **19** with same 4*R*-Amp at Y position.

2.5.4: Chain length dependent triple-helix formation

Goodman *et al*[29] have studied the propensity of formation of triple helix with increasing chain length of a peptide Ac-(Gly-Pro-Hyp)$_n$-NH$_2$ in different solvent conditions. When n =1 and 3 they did not observe triple-helical conformation either in H_2O or EG: H_2O (2:1), as indicated by low R_{pn} values and the lack of observable melting transition. On the other hand peptides with n = 6 and 9 form very stable triple-helical structures with melting temperatures of 36 ^{0}C (n = 6) and 67 ^{0}C (n = 9) in H_2O. At a concentration of 0.2mg/ml, peptide with n = 5 forms a stable triple-helical conformation in EG: H_2O (2:1) with a melting temperature of 32 ^{0}C. However in H_2O, a triple-helical structure is observed below temperature of 18 ^{0}C. This suggestsed that this compound represents a transition between the triple-helical and the non-triple-helical conformations in H_2O. These results indicate that at least five Gly-Pro-Hyp tripeptide repeats are required to form a stable triple-helical structure in H_2O, in the absence of a template. In order to find out the minimum trimer units required to form a collagen like triple helical structure in H_2O, we studied the peptides Ac-Phe (Pro-Amp-Gly)$_n$-NH$_2$ with n = 1-6, for their propensity to form a triple-helical structure at pH 3.0. Figure 2.19A shows the CD spectra of peptide Ac-Phe(Pro-Amp-Gly)$_n$-NH$_2$ [n = 1-5]. CD spectra of peptide with n = **2** does not show any negative band, which indicates that this peptide is not forming any triple-helix structure which is confirmed by the absence of a transition in CD-T_m study. Peptide with n = **3** displays a spectrum which resembles collagen triple helix, though the low R_{pn} (0.07) value and low T_m 19^{0}C indicate that this

peptide represents a transition point for switching the non triple-helical conformation to triple-helical form.

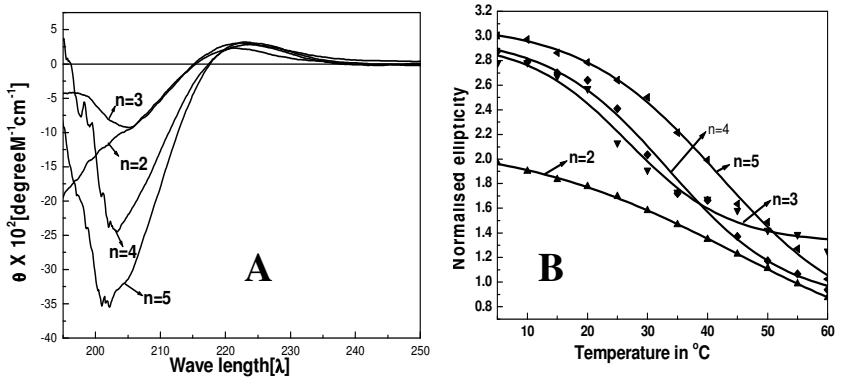


Figure 2.19 A. CD spectra **B**. thermal denaturation curves of peptide Ac-Phe(Pro-Amp-Gly)n-NH_2 (n = 2, 3, 4 and 5).

The peptides with n = 4 and 5 form a stable triple-helix with T_m 33.5 ^oC (n = 4) and 42.3 ^oC (n = 5) (Table 11). These results indicate that at least four Gly-Pro-Amp repeats are required to form a stable triple-helical structure in H_2O.

Table 11: Rpn and T_m values of peptide Ac-Phe(Pro-Amp-Gly)$_n$-NH_2 with variable length

Ac-Phe(Pro-Amp-Gly)$_n$-NH_2	**R_{pn}**	**T_m ^oC**
n= 2	-	-
n= 3	0.07	19
n= 4	0.109	33.5
n= 5	0.128	42.3

2.5.5: Concentration effects on triple-helix formation of Ac-Phe(amp-Pro-Gly)$_6$-NH_2

For single-chain acetyl-terminated collagen analogs, triple-helix formation involves the interaction of three molecules. Therefore, the concentration is expected to have a significant effect on triple-helical propensity, especially when the peptide chain

is short. As shown in section 2.35 peptide **20** (Amp-Pro-Gly) shows a gradual increase in their R_{pn} value thereby increasing in propensity of triple-helix with increasing concentration form 0.05mM to 0.15mM and thereafter reaches saturation. The peptide **22** (amp-Pro-Gly) displays a different kind of concentration dependent CD spectra, the R_{pn} value showing a maximum at 0.15mM and minimum at 0.20mM.

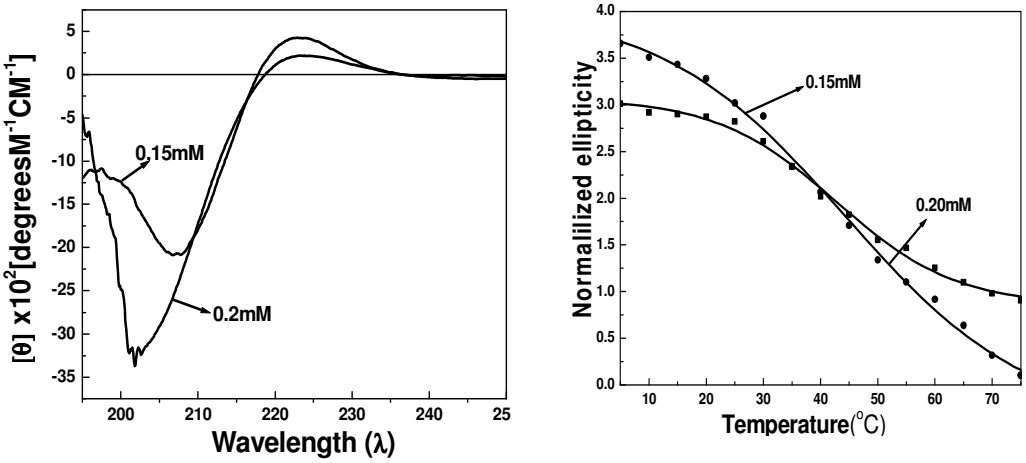


Figure 2.20: **A**; CD spectra, **B**; thermal denaturation curves of peptide-**22** Ac-Phe(amp-Pro-Gly)$_6$-NH_2 at concentration 0.15mM and 0.20mM respectively.

Table 12: CD-T_m values for peptide **22** at concentration of 0.15mM and 0.20mM.

Concentration (mM)	R_{pn} value	T_m (^oC)
0.15	0.19	45
0.20	0.097	44

To examine changes of concentration effects on triplex stability, CD-T_m study was carried out at two different concentrations.

Figer-2.20A shows the CD spectra of the peptide **22** (amp-Pro-Gly) at concentration 0.15mM and 0.20mM. A large difference in R_{pn} value and a slight shift in cross over point was seen at the two concentrations. The thermal denaturation experiment conducted at two different concentrations indicates a difference in T_m of 1^oC (Table 12).

2.6: Discussion

The 4*R*-amnioproline containing peptide **20** (Amp-Pro-Gly) with 4*R*-Amp at X position form a polyproline (PP) II like structure at pH 3.0 and pH 7.0 as indicated by a positive band at ~223nm and a strong negative band at ~200nm in the CD spectra. But peptide **20** (Amp-Pro-Gly) does not exhibit PP II like structure under pH 9.0 and 12.0. In the concentration dependent triple-helix formation study, the peak positions are nearly independent of the concentration and all the spectral traces passes through the same isobestic point at ~217nm for peptide **20**. This CD spectral property indicates that within the concentration range measured i.e. 0.05-0.30mM. The peptide **20** exists in the triple-helical conformation. Smooth and progressive increase in the R_{pn} values with concentration also suggests a two state equilibrium, which rapidly moves towards saturation to a single state with increase in the concentration. The peptide **21** (Pro-amp-Gly) with 4*S*-amp at Y position exhibits low R_{pn} and does not reach saturation within the concentration range of 0.05-0.30mM, suggesting that this peptide does not associate to triple-helix and exists in single chain (PP II) conformation. The non-progressive change of R_{pn} values with increase in concentration for peptide **22** (amp-Pro-Gly) with 4*S*-amp in X position may result from a higher order association at higher concentrations (> 0.20mM) of peptide. This view is also supported by the shift in CD-band position with concentration and the appearance of two distinct isobestic points, one at 216nm in the concentration range 0.05-0.15mM and a second one at 217nm in the concentration range 0.20-0.25mM.

The observed sigmoid transition in the variable temperature CD measurements provides additional evidence for the two state (triple-helix to coil) transition. The T_m

values of peptide **19** (Pro-Amp-Gly)[8] decreases with an increase in pH over the range 3.0-9.0, followed by an increase again at pH 12. Since the pKa of 4-NH_2 group of 4*R*-aminoproline is 10.2, at pH 3.0 and pH 7.0 the 4*R*-aminoproline exists in fully protonated (-NH_3^+) form, and at pH 12.0, it exists in the free amine (-NH_2) form. At pH 9.0, which is very close to its pKa, only partial protonation may occur, leading to a decrease in stability of triple-helix. The peptide **20** (Amp-Pro-Gly) with 4*R*-Amp at X position associated to a triplex only at pH 3.0 and 7.0, revealing that protonation of aminogroup is a prerequisite to form triple-helix for 4*R*-aminoproline when present at X position of the collagen peptide sequence (X-Y-Gly), but is not necessary when the same amino acid is present at Y position. The absence of sigmoid transition and very low R_{pn} value for peptide **21** (Pro-amp-Gly) clearly indicates that this peptide does not associate to a triple-helix under all pH conditions. Like 4*S*-hydroxyproline, 4*S*-aminoproline also destabilizes the triple-helical structure at Y position. This is due to the predominant *C4-endo* pyrrolidine ring puckering adopted by 4*S*-aminoproline, which is favorable at X position and not at Y position. This was further confirmed by the formation of stable triple-helical structure by this aminoacid at X position (peptide **22** amp-Pro-Gly) under pH 3.0, 7.0 and 9.0, but not at pH 12.0 where it destabilizes the triple-helix. The one important result is that, all aminoproline containing collagen peptides at both X and Y positions show pH dependent triple-helix stability, which is in contrast to the control peptide with 4-Hydroxyproline, for which the triple-helix stability is almost the same at all pH conditions.[8] Since all peptides were capped at both at N and C termini, the pH dependent triple-helix stability of 4-aminoproline containing collagen peptides are truly due to the protonation status of 4-amino group in aminoproline. Since NH_3^+ and NH_2 groups have different electronegativity, and the pyrrolidine ring

puckering is dependent on the electronegativity of the 4-substituent, it is possible that 4-aminoproline exhibits different pyrrolidine puckerings depending on whether it is protonated or nonprotonated. This may alter the difference in triple-helical stability at different pH conditions.

In EG:H_2O (3:1), the peptide **22** (amp-Pro-Gly) with 4*S*-aminoproline residue at X position form more stable triple-helices (T_m 39 ^oC) compared to the isomeric peptide **20** (Amp-Pro-Gly) with 4*R*-aminoproline residue (Tm 35 ^oC) at X position. All aminoproline containing peptides at both X and Y positions form less stable triplex structure in EG:H_2O compared to corresponding peptide in aqueous condition. This is opposite to the behavior of control peptide with 4*R*-hydroxyproline, which forms more stable structure in EG:H_2O system compared to aqueous buffer conditions. Also the peptide with 4-aminoproline in X position forms a more stable triple-helix compared to the peptide with 4-aminoproline at Y position. For example peptide **21** (Amp-Pro-Gly) forms more stable triple-helical structure in EG:H_2O system compared to peptide **19** (Pro-Amp-Gly) ($\Delta T_m = 12^0C$), but in aqueous system, peptide **19** forms a more stable structure under all pH conditions compared to peptide **20**. This may due to the fact that in collagen triple-helix, unlike the Y-residue which is buried inside, the residue in X position is more exposed and amenable to interactions with bulk solvent and other molecules. As the amp residue is present in the X position in peptide **22** (amp-Pro-Gly), such interactions are possible. Hence reversal of stability in EG:H_2O may arise from the difference in exposure of residues at X and Y positions in triple-helix to the solvent. Also the aminoproline containing peptides form less stable triple-helix in EG:H_2O system compared to that in aqueous buffer conditions. The additional interstrand electrostatic repulsions arising from the 4-NH_3^+ groups in aminoproline peptides remain

unscreened in EG:H_2O mixture. This repulsive interaction aggravated by the absence of salt may result in the decreased thermal stability for triple-helix of 4-aminoproline containing peptides. since EG:H_2O enhances the hydrogen bonding stability effects and decrease the electrostatic stability compared to H_2O, the formation of less stable triple-helical structure in EG:H_2O system indicates that electrostatic interactions play larger role for stabilizing triple-helical structure of 4-aminoproline containing collagen peptides. Interestingly despite a large increase in the positive band intensity in EG:H_2O, the peptide **21** (Pro-amp-Gly) still does not show a transition.

The presence of NaCl in the concentration range 0-150mM has a dramatic effect on the T_m values of both peptide **19** (Pro-Amp-Gly) and peptide **22** (amp-Pro-Gly) at pH 7.0. Higher concentration of salt seems to stabilize the triple-helical structure as evidenced by high T_m values for both peptides at 100mM NaCl. As pKa values of 4*R*-Amp and 4*S*-amp monomers are 10.5 and 10.3 respectively, the 4-NH_2 group would be largely protonated at pH 7.0. Collagens being a parallel triple-helix with one residue shift, the NH_3^+ groups lie in close proximity to each other which results in large electrostatic repulsive interactions between the chains. The presence of salt counter-ions offers a shielding effect resulting in decreased electrostatic repulsion between the chains. Hence stability of the triple-helix increases with increasing salt concentration from 0-100mM. Beyond 150mM the breaking of water-structure and poor solubility of peptide leads to a decrease in the triple-helical stability and a decrease in T_m for both peptides **19** and **22**. This result clearly indicates that the mechanism of triple-helix stabilization by 4-aminoproline containing collagen peptides at X and Y positions is qualitatively the same.

2.7: Conclusions

In conclusion, it is shown that the replacement of 4*R*-hydroxyproline residue in collagen peptide (Pro-Hyp-Gly) by 4*R*-aminoproline leads to a remarkable stabilization of the derived triple-helices. The results presented here demonstrate that the 4*R*/*S*-aminoprolines are one of the first proline derivatives that stabilize collagen triplex when present in the X-position and the (4*S*)amp is better than (4*R*)Amp. The cause of the stabilization is the C4-*endo* pyrrolidine conformation adopted by 4*S*-amp that is inherently favored at the X-position. The pH dependent stabilities in both 4*R* and 4*S* aminoprolines also suggest that protonation of NH_2 is a prerequisite for triplex formation in X position, while it is not so in Y-position. It is possible that conformation of pyrrolidine ring also depends on the protonation status of 4-amino group and influence the *cis-trans* amide rotameric equilibrium. The results reinforce the important role of stereoelectronic effects, well elucidated in 4-fluoroprolines, also to be determinants of stability of 4-aminoproline collagen. Future potential of this work lies in rationally combining the stabilities offered by the two diastereomers to design collagens (eg. amp.Hyp/Amp-Gly)$_n$ or chimeric collagens of unusual stability for applications in collagen based biomaterials. The results also have implications for peptidomimetic designs and control of peptide conformations through 4-substituent effects on proline. The observed pH and salt effects in these peptides suggest different mechanisms to be responsible for enhancing triple-helix stability at different conditions. The results have a direct bearing on the current interest in collagen structure and mimics. Since 4-OH group in natural collagen can be potentially transformed to 4-NH_2, the properties of this analogue may have significance in the design of new collagen based biomaterials.

2.8: Experimental section

All reagents were obtained from commercial sources and were used without further purification except in cases mentioned. Thin layer chromatography (TLC) was carried out on precoated silicagel 60 HF254 plates (E. Merk). TLCs were visualized under UV light, iodine and/or ninhydrine spray, followed by heating to 110 ^{o}C. ^{1}H NMR spectra were recorded at 200MHz on a Bruker AC-200 instrument controlled by an Aspect 2000 computer. ^{13}C NMR spectra were recorded on the same instrument. Spectra were analyzed using ACD spec-viewer software from ACD labs. All chemical shifts are with reference to TMS as internal standard and are expressed in δ-scale (ppm). NMR spectra of compounds show two sets of peaks arising from *cis-trans* isomerization of N-CO bond. FAB mass was recorded on VG Autospec (M series) instrument. MALDI-TOF mass spectra were obtained on a Voyager-Elite instrument (PerSeptive Biosystems Inc., Farmingham, MA) equipped with delayed extraction.

Reagents for the buffer preparation such as NaCl, Na_2BO_4, KH_2PO_4 and K_2HPO_4 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.

2.8.1: Peptide synthesis

All peptides were synthesized manually in a sintered vessel equipped with a stopcock. Readily available rink amide resin with loading value 0.7-1.0 mmol/g was used and standard Fmoc chemistry was employed. Resin bound Fmoc group was first deprotected with 20% piperidine in DMF and the coupling reaction 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 (Nova biochem) and was used without further purification. Analytical grade DMF was purchased from Merck (India) and was distilled over P_2O_5 under vacuum at 45^0C, stored over $4A^o$ molecular sieves for 2 days before using for peptide synthesis.

2.8.2: Resin functionalization

100mg resin was taken in sintered vessel equipped with a 14B joint stopper, and rinsed with 5 ml of dry DCM and filtered, the process was repeated 3 to 4 times and the resulting resin was kept overnight in 5ml DCM for swelling. DCM was filtered and rinsed 3 times with dry DMF, and kept 5hr in 3ml dry DMF for swelling before putting first coupling.

2.8.3: General method for solid phase peptide synthesis

All peptides were assembled using solid phase method by putting one amino acid per coupling. All amino acid were well dried over P_2O_5 in vacuum decicator before using for coupling. Fmoc protecting group was used for main chain α-amine group, and t-Boc protection was used for side chain amine group, and was cleaved with 50% TFA in DCM for final cleavage of peptide from the resin.

2.8.4: Synthesis protocol for solid phase synthesis

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

1) wash with DMF 4X 5ml;

2) deprotection of Fmoc group with 20% piperidine in DMF 2 X 5ml (20 minutes for each);

3) wash with DMF 3 X 5ml;

4) test for complete deprotection (ninhydrine or chloronil test);

5) coupling reaction with 3 eq. each aminoacid, DIPEA, HOBt and HBTU in minimum volume of DMF;

6) test for completion of coupling reaction (ninhydrin or chloronil test);

This cycle was repeated for every amino acid. The coupling and deprotection reactions were monitored by a combination of Kaiser's (ninhydrin) test and chloronil test. In case of positive test after coupling the second coupling was performed with same aminoacid before deprotection of Fmoc group followed by capping reaction with 1: 1: 1 mixture of acetic anhydride, DMAP and DCM.

2.8.5: Kaiser's test

Kaiser's test was used to monitor the Fmoc deptotection and coupling reactions of glycine (generally primary amines). The following three solutions

1) Solution **1:** 0.5g Ninhydrine dissolved in 100ml ethanol;
2) Solution **2:** 80g of phenol dissolved in 20ml of ethanol;
3) Solution **3:** 2ml of 0.001M aqueous KCN solution in 98ml of pyridine; were used.

Few beads of the resin were taken in a test tube, washed by agitating with methanol three times. To this 3-4 drops of each of above solutions were added. The mixture was heated at 100^0C for 1-2 minutes, and the color of the beads was noted. A blue color that start on the beads and diffuses in to the solution is a test for free amino

group (while deprotection reaction), and colorless beads even after heating for 4-5 minutes indicated a completion of the coupling reaction. The blank solution should remains straw yellow.

2.8.6: Chloronil test

A few beads of resin were taken in a test tube and were washed with methanol followed by toluene. To this three drops of saturated chloronil solution in toluene and 200µl of acetone were added. The mixture was shaken for 2-3 minutes. Blue or green color is observed on the resin beads if free amines are present.

2.8.7: Preparation of resin with peptide for cleavage

After the last coupling was over the resin was washed with DMF 5 X 10ml, DCM 5 X 10ml, then with toluene 5X 10ml and finally with methanol 5X10ml and flesh with nitrogen gas for 3minutes. The sintered flask was well stoperred and dried in a vacuum desiccator over P_2O_5.

2.8.8: Peptide cleavage from the resin

100mg of dry peptide-resin was taken in round-bottomed flask and to this 10ml 0f 5% solution of 95: 5 TFA:TIS in DCM was added. The resulting mixture was stirred gently using magnetic stirrer for 1hr; the mixture was filtered through a sintered funnel and the resin was further washed with 3 X 5ml of above solution. The filtrate was collected in round-bottomed flak and evaporated under reduced pressure. The resin was again washed with methanol 3X5ml and washings were collected in same RB and evaporated to dryness. The crude peptide obtained still contains a N^4-tBoc group, which was deprotected by stirring the peptide solution with 10ml of 50% of 95:5 mixture of

TFA:TIS in DCM for 2hrs. The TFA:DCM mixture was removed under reduced pressure; the resulting crude peptide was rinsed with anhydrous diethyl ether which was removed under vacuum. The residue obtained was dissolved in 0.5ml of anhydrous methanol and then the methanolic solution of the crude peptide was transferred to a centrifuge tube using micro pipette, and to it 20ml of anhydrous diethyl ether was added. The off-white precipitate obtained was centrifuged. The methanol-ether precipitation procedure was repeated twice resulting in almost colorless peptide.

2.8.9: HPLC purification and analysis

All peptides were purified by RP-HPLC on Waters Deltapak, RP-4 (15 μM) column connected to HP-1050 HPLC system equipped with JASCO UV-970 variable wavelength detector and HP-3380A integrator. A gradient of 0-30% buffer B at flow rate of 2ml/min was used to elute the peptide and the eluting was monitored at 220nm. The peak corresponding to peptide was collected and freeze-dried.

The purity of the final peptide was confirmed by analytical HPLC on a Merk Lichrospher, RP-18e (5μM) column by using a gradient of 0-100% B in 30 min at a flow rate of 1ml/min. The absorbance of the eluting was monitored at 220nm.

The gradient A = deionized water with 0.1% TFA;

The gradient B = 50:50 deionized water: acetonitrile with 0.1% TFA.

2.8.10: CD spectroscopy and thermal denaturation studies

CD spectrometry study was carried out on JASCO J-715 spectropolarimeter using cylindrical, jacketed quartz cell (1 mm path length), which was connected to Julabo-UC water circulater. Prior to spectroscopic analysis samples annealed by

immersing into a water bath at 90^0C and were cooled to room temperature in the bath over a period of 4hrs. These samples were incubated at room temperature for 12hr followed by 4hr at 2^0C. An additional ½ hr incubation time was allowed in the instrument at the initial measurement temperature. Spectra were recorded with a spectral resolution of 0.1nm, at a scan speed of 50nm/min and a response time 8sec. All the spectra were corrected for buffer condition and are typically averaged over 6-12 scans.

2.8.11: CD thermal denaturation studies

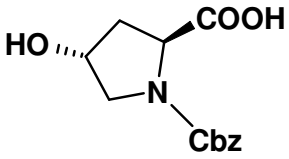
The samples were annealed according to the procedure described as above. The temperature was varied in steps of 2 or 5^0C and the spectra were recorded at each temperature. An equilibration period of 10min was allowed at each temperature. Data processing and curve fitting was performed using Microcal Origin software. Ellipticity at 225nm for each temperature was plotted, normalized data was fitted to a sigmoid curve and the T_m values are derived from the first derivative curve of the fit. Normalization of data was performed according to the equation $(\theta-\theta_{min}) / (\theta_{max} - \theta_{min})$, where θ_{max} and θ_{min} are maximum and minimum values in each data set.

2.8.12: pK_a determination of monomers Amp and amp

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

2.8.13: Synthesis of compounds 6-18

*(2S, 4R)-N^1 –(benzyloxycarbonyl)-4-hydroxyproline (**6**)*

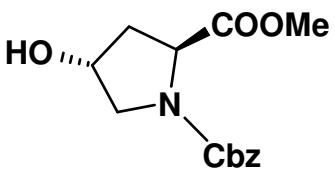


A solution of *trans*-4-hydroxy-L-proline (7g, 53.4 m mol) in 10% aq $NaHCO_3$ (80ml) was cooled in an ice bath. 50% solution of Cbz-Cl in toluene was added. The reaction mixture was stirred for 18hr, at room temperature. Toluene was removed under vacuum. The aqueous solution was cooled to 0^oC and was washed with diethylether, acidified to pH 3 with 10% HCl, and then extracted with ethyl acetate. The combined organic solution was washed with water followed by saturated brine solution, dried over anhydrous Na_2SO_4 and concentrated under reduced pressure to yield white solid of compound **6**. Yield 13.56g; 98.8%. Mo.F, $C_{13}H_{15}NO_5$, [Mass ;(calculated) 265; observed 266.05(M+1)]. $[\alpha]^D$ -86.70 (c 1.0 CH_2Cl_2). This compound was used for next reaction without purification.

^{1}H NMR ($CDCl_3$, 200 MHz): δ_H 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).

^{13}C NMR; 37.8, 38.7, 49. 54.3, 57.6, 60.4, 76.7, 127.0, 127.7, 128.4, 136., 155.1, 175.1.

*(2S, 4R)-N^1 – (benzyloxycarbonyl) –4- hydroxyproline methyl ester (**7**)*



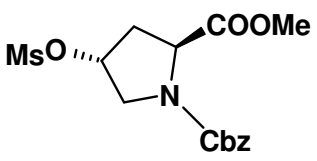
The stirred solution of compound **6 (**13g, 49 mmol) in anhydrous acetone (75ml) and 16.9g (122.mmol) of anhydrous K_2CO_3, dimethylsulphate 5.6ml (58.8mmol) was added in one shot. The mixture was then refluxed under nitrogen for 4hrs. The acetone was removed under vacuum; the resulting residue was dissolved in water then extracted with ethyl acetate.

The combined organic layer was washed with water, followed by saturated brine solution, dried over Na_2SO_4 and concentrated under vacuum. The crude material was purified by silica gel chromatography (50% ethyl acetate/hexane) afford compound **7** as a colorless thick oil. Yield 13.52g; 98.78%. Mo.F, $C_{14}H_{17}NO_5$, [Mass ;(calculated) 279; observed 280.04(M+1)]. $[\alpha]^D$ -61.15 (c 1.0 CH_2Cl_2).

^{1}H NMR ($CDCl_3$ 200 MHz) δ_H 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; 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, 4R)-N^1- (benzyloxycarbonyl)-4-(methanesulfonyloxy) proline methyl ester (**8**)*

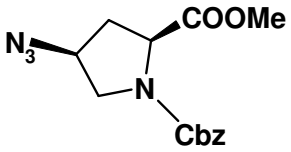


The solution of methylester **7** (13g, 46.5mmol) and triethylamine 9.74ml (69.89mmol) in dry dichloromethane (100ml) was cooled to 0^{o}C on ice bath under Argon. While stirring methanesulfonyl chloride 4.32ml, (55.91mmol) was added dropwise over a period of 3hrs at 0^{o}C. The reaction mixture was washed with water, followed by saturated 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) afford mesylated compound **8** as a colorless thick oil. Yield 16.0g; 96.18%. Mo.F, $C_{15}H_{19}NSO_7$, [Mass ;(calculated) 357; observed 357.05(M+)]. $[\alpha]^D$ -43.66 (c 1.0 CH_2Cl_2).

^{1}H NMR ($CDCl_3$ 200 MHz) δ_H 2.09-2.36 (1H, m) 2.42(1H, m), 2.79 (3H, s), 3.40 (1H, m) 3.60-3.84 (3H, s), 4.38 (1H, m), 4.84-5.10 (3H, m), 7.17-7.21 (5H, m).

^{13}C NMR; 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 (**9**)*

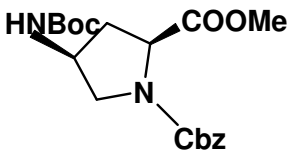


A solution of mesyl compound 8 (15.5 g, 43.4mmol) and NaN_3 22.57g, (347.3mmol) in dry DMF (120ml) was stirred at 70^{o}C for 8hr under Argon. DMF was removed under vacuum and the residue was dissolved in water. The aqueous layer was extracted with ethyl acetate. The combined organic layer was washed with water followed brine, dried over anhydrous Na_2SO_4 and concentrated under vacuum. The crude material was purified by silica gel chromatography (40% ethyl acetate/hexane elute) to afford azide compound **9** as a colorless thick oil. Yield 12.5g; 94.7%. Mo.F, $C_{14}H_{16}N_4O_4$, [Mass ;(calculated) 304; observed 305.05(M+1)]. $[\alpha]^D$ -23.32 (c 1.0 CH_2Cl_2).

^{1}H NMR ($CDCl_3$ 200 MHz) δ_H 2.21-2.32 (1H,m), 2.52-2.56 (1H,m), 3.61-3.63(2H, m), 3.71-3.84 (3H,s), 4.26-4.51(1H, m), 4.84-4.59 (1H, m), 5.13-5.28 (2H,m), 7.36-7.44 (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)-N^1- (benzyloxycarbonyl)-4-(t-butoxycarbonylamino) proline methyl ester (**10**)*

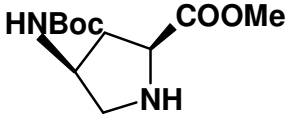


Azide Compound **9,** 12g (39.47mmol) was dissolved in of dry ethyl acetate (14 ml), 6ml of Raney-Nickel (suspension in ethylacetete) and di-t-butyldicarbonate (9.47g, 43.43mmol) were added. The reaction mixture was subjected to hydrogenation under 40 *psi* pressure of H_2 gas for 4 hrs in parr-hydrogen apparatus.

The reaction mixture was filtered through celite. Filtrate was concentrated under reduced pressure and purified by silica gel chromatography, obtained compound **10** as a colourless liquid. Yield 14.4g; 96.5%. Mo. F, $C_{19}H_{26}N_2O_6$, [Mass ;(calculated) 378; observed 379.05(M+)]. $[\alpha]^D$ -32.44 (c 1.0 CH_2Cl_2).

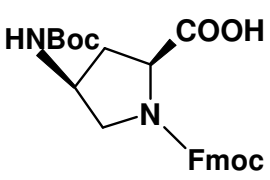
^{1}H NMR ($CDCl_3$ 200 MHz) δ_H 1.39 (9H, s), 2.21-2.22 (2H, m), 3.27-3.41 (3H, m), 3.71-3.84 (3H, s), 4.26-4.51(1H, m), 4.52-4.59 (1H, m), 5.13-5.28 (2H, m), 7.36-7.44 (5H, m). **^{13}C NMR**; 28.0, 35.4, 36.6, 48.7, 49.3, 51.5, 51.9, 57.3, 57.5, 66.9, 79.6, 127.5, 127.7, 128.1, 136.0, 154.9, 172.1, 172.3.

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



Compound **10** (14g, 37.2mmol) was dissolved in methanol (4ml) and to this 10% Pd-C (2g) was added. The mixture is subjected to hydrogenation in a parr-hydrogenation apparatus at a pressure of 55-60 *psi* for 8 hrs. The reaction mixture was filtered through celite. The filtrate was concentrated under reduced pressure offered pale yellow solid. Yield 8.5g, (96.59%),which was used for next reaction with out further purification.

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



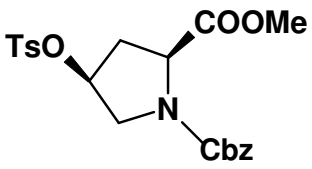
Above compound (6g 26.08mmol) was dissolved in methanol (10ml) and 2N NaOH (10ml) solution. The reaction mixture was stirred for 30min. then neutralized to pH 7.0 with $KHSO_4$ solution. Methanol was removed under

vacuum. The resulting aqueous solution was cooled to zero degree and 10% Na_2CO_3 (25ml) and dioxane (120ml) were added. 7.42g (28.69mmol.) of Fmoc-Cl, dissolved in dioxane (15 ml) was added drop wise about 45min. the resulting mixture was stirred for 10hrs. at room temperature, maintaining the solution's pH=8. Dioxane was removed under vacuum and extracted with ether. The aqueous layer was covered by ethyl acetate, and acidified with $KHSO_4$ solution to pH 3. The combined ethyl acetate layer was washed with water followed by brine, dried over anhydrous Na_2SO_4 and concentrated under vacuum. The crude material was purified by silica gel chromatography (80% ethyl acetate/hexane) afford **12** as white solid. Yield 4g, (33.92%). Mo. F, $C_{25}H_{28}N_2O_6$, [Mass ;(calculated) 452; observed 475.18 (M+Na)]. $[\alpha]^D$ -32.44 (c 1.0 CH_2Cl_2).

^{1}H NMR ($CDCl_3$ 200 MHz) δ_H 1.39 (9H, s), 2.18-2.26 (2H, m), 3.22-3.37(1H, m), 3.69-3.74 (1H, m), 4.13-4.40 (1H, m), 6.47 (1H, m), 7.33-7.77 (8H,m).

^{13}C NMR; 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.

*(2S, 4S)-N1 (t-butoxy carbonyl) –4-(P-toluene sulfonyloxy) proline methyl ester. (**14**)*



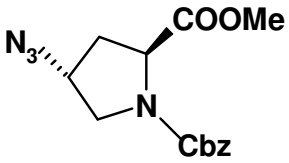
The mixture compound **7,** 8g (28.67mmol), PPh_3, 8.27g (31.54mmol) and methyl-*p*-toluene sulfonate,4.76ml (31.54mmol) dissolved in dry THF (100ml) was cooled to 0^{o}C on ice bath, under Argon. The mixture was stirred for 30min at ^{0}C; Diehtylazodicarboylate 6.37ml (31.54mmol) was added slowly with syringe. The reaction mixture was stirred at room temperature for 8hrs. Toluene was removed under vacuum. The resulting orange colored thick oil was dissolved in 150ml petroleum ether, by triturating with spatula; the resulting solution was kept overnight at room temperature. The white powder settled was filtered and the residue was washed with

petroleum ether followed by diethylether, offered compound **14** as white power. Yield 10.5g (80%). Mo. F, $C_{21}H_{23}NO_7S$, [Mass ;(calculated) 433; observed 434.12(M+)]. $[\alpha]^D$ -24.44 (c 1.0 CH_2Cl_2).).

^{1}H NMR ($CDCl_3$ 200 MHz) δ_H 1.80 (1H, s), 2.34-2.48 (2H, m), 2.45 (3H, s), 3.58 (3H, s), 3.69-3.73 (1H, m), 4.42-4.55 (1H, m), 5.02-5.23 (2H, m), 7-26-7.36 (7H, m), 7.73-7.77 (2H, d).

^{13}C NMR; 21.4, 35.8, 36.9, 51.8, 52.2, 52.2, 57.1, 57.3, 67.1, 67.1, 78.5, 127.5, 127.7, 127.9, 128.3, 129.8, 133.3, 136.0, 145.0, 153.8, 154.17, 171.0, 171.3.

*(2S, 4R)-N1- (benzyloxycarbonyl)-4-azido proline methyl ester (**15**)*

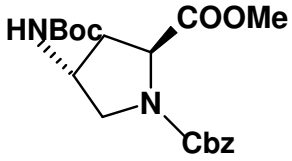


A solution of Compound **7,** 10g, (23mmol) and NaN_3, 12g, (284.8mmol) in dry DMF (80ml) was stirred at 55-60^{o}C for 8hr under Argon. DMF was removed under vacuum and residue was dissolved in water. The aqueous layer was extracted with ethyl acetate. The combined organic layer was washed with water-followed by brine, dried over anhydrous Na_2SO_4 and concentrated under vacuum. The crude product obtained was purified by silica gel chromatography (40% ethyl acetate/hexane elute) afford **16** as a colorless thick oil. Yield 7g, (99.7%). Mo.F, $C_{14}H_{16}N_4O_4$, [Mass ;(calculated) 304; observed 305.05(M+1)]. $[\alpha]^D$ -45.75 (c 1.0 CH_2Cl_2).

^{1}H NMR ($CDCl_3$ 200 MHz) δ_H 2.14-2.21 (1H, m), 2.31-2.35 (1H, m), 3.54-3.63 (2H, m), 3.57-3.75 (3H, s), 4.20-4.39 (1H,m), 4.41-4.49 (1H, m), 5.03-5.28 (2H, m), 7.29-7.34 (5H, m).

^{13}C NMR; 37.6, 38.6, 52.3, 52.5, 55.6, 57.6, 66.7, 127.4, 127.6, 128.0, 135.9, 153.8, 155.0, 172.1, 172.3.

*(2S, 4R)-N^1-(benzyloxycarbonyl)-4-(t-butoxycarbonylamino) proline methyl ester (**16**)*

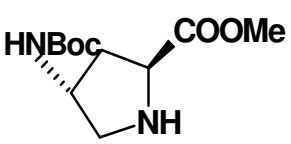


Compound **15,** 7g (23.mmol) was dissolved in dry ethyl acetate (10ml), 4ml of Raney-Nickel (suspension in ethylacetete) and di-t-butyldicarbonate 5.53g, (37mmol) were added. The mixture was subjected to hydrogenation under 40-psi pressure of H_2 gas for 4 hrs in par-hydrogen apparatus. The reaction mixture was filtered through celite, and filtrate was concentrated under reduced pressure, purified by silica gel chromatography. Obtain a compound **16.** Yield 8.5g (97.6%). Mo. F, $C_{19}H_{26}N_2O_6$, [Mass ;(calculated) 378; observed 379.05(M+)]. $[\alpha]^D$ -29.41 (c 1.0 CH_2Cl_2).

^{1}H NMR ($CDCl_3$ 200 MHz) δ_H 1.47 (9H,s), 1.93-2.14 (1H, m), 2.38-2.55 (1H,m), 3.51-3.63 (2H, m), 3.71-3.78 (3H, s), 4.31-4.40 (2H, m), 5.06-5.21 (2H, m), 5.37 (1H, m), 7.34 (5H, m).

^{13}C NMR; 27.9, 35.1, 36.2, 49.1, 51.7, 66.7, 127.4, 127.6, 128.0, 135.9, 153.8, 158.0, 172.1, 172.3.

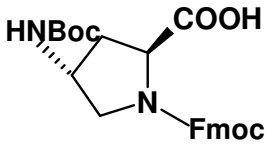
*(2S, 4S)-4-(t-butoxycarbonylamino) proline methyl ester (**17**)*



Compound **16,** 8g, (21.16mmol) was dissolved in methanol (10ml) and to this was added 10% Pd-C (1.5g). The mixture is subjected to hydrogenation in a parr-

hydrogenation apparatus at a pressure of 55-60 psi for 8 hrs. The reaction mixture was filtered through celite, filtrate was concentrated under reduced pressure, yield yellowish solid 4.8g (98.5%). This was used for next reaction without further purification.

*(2S, 4S)-N1-(fluorenyl methyloxycarbonyl)-4-(t-butoxycarbonylamino) proline (**18**)*



Above compound 4.5g (19.56mmol) was dissolved in methanol (10ml) and of 2N NaOH (10ml) solution. The reaction mixture was stirred for 30min. then neutralized to pH 7.0 with $KHSO_4$ solution. Methanol was removed under vacuum. The resulting aqueous solution was cooled to zero degree and 10% Na_2CO_3 (20ml) and dioxane (100ml) were added. 5.57g (21.52mmol.) of Fmoc-Cl, dissolved in dioxane (10ml) was added dropwise about 40min. the resulting mixture was stirred for 10hrs at room temperature, maintaining the solution's pH=8. Dioxane was removed under vacuum and extracted with ether. The aqueous layer was covered by ethyl acetate, and acidified with $KHSO_4$ solution to pH 3. The combined ethyl acetate layer was washed with water followed by brine, dried over anhydrous Na_2SO_4 and concentrated under vacuum. The crude material was purified by silica gel chromatography (80% ethyl acetate/hexane elute) afford compound **18** as white solid. Yield 3.3g, (37.5%). Mo. F, $C_{25}H_{28}N_2O_6$, [Mass ;(calculated) 452; observed 475.30(M+Na)]. $[\alpha]^D$ -43.75 (c 1.0 CH_2Cl_2).

^{1}H NMR ($CDCl_3$ 200 MHz) δ_H 1.43 (9H, s), 2.12-2.40 (1H, m), 2.42-2.44 (1H, m), 2.92-2.99 (1H, m), 3.44-3.36 (1H,m), 3.38-3.82 (1H, m), 3.84-4.28 (1H, m), 4.30-4.55 (5H, m), 7.31-7.80 (8H, m).

^{13}C NMR; 28.3, 35.4,36.7, 36.6, 47.0, 51.8, 57.4, 57.7, 67.8, 114.9, 125.0, 127.0, 127.6, 141.2, 143.5, 143.9, 154.4, 155.3, 174.5, 175.3, 178.

2.94: References

1) a) Kersteen, E.; Raines, R. *Biopoymers* **2001**, *59*, 24-28. b) Berg, R.; Prockop, D.; *Biochem Res Commun*. **1973**, *51*, 115-120. c) Holmgren, S.; Bretscher, L.; Taylor, K.; Raines, R. *Chem Biol*. **1999**, *6*, 63-70. d) sakakiba, S.; Inouye, K.; Kishida, Y.; Kobayash, Y.; Prockop, D. *Biochem. Biophys Acta* **1973**, *303*, 198-202.

2) a) Koboyashi, Y.; Sakai, R.; Kakiuchi, K.; Isemura, T. *Biopolymers* **1970**, *9*, 415-425. b) Sasakibara, S.; Kishida, Y.; Kakiuchi, Y.; Sakai, R.; Kakiuchi, K. *Bull. Chem. Soc. Jpn*. **1968**, *41*, 1273.

3) a) Ramachandran, G.; Bansal, M.; Bhatnagar, R. *Biochem Biophys Acta* **1973**, *322*, 166-171. b) Okuyama, K.; Arnott, S.; Takayangi, M.; Kakudo, M. *J. Mol Biol*. **1981**, *152*, 427-443. c) Migchels, C.; Berendse, H. *J. Chem. Phys*. **1973**, *59*, 296-305.

4) a) Eberhardt, E. S.; Panasik, N. Jr.; Rainer, R. T. *J. Am. Chem. Soc*. **1996**, *118*, 12261-12266. b) Holmgren, S. K.; Taylor, K. M.; Bretscher, L. E.; Raines, R. T. *Nature* **1998**, *392*, 666-667. c) Holmgren, S. K.; Bretscher, L.E.; Taylor, K. M.; Raines, R. T. *Chem. Biol*. **1998**, *6*, 63-70.

5) a) Howard, J. A. K.; Hoy, V. J.; O' Hagan, D.; Smith, G. T. *Tetrahedron* **1996**, *52*, 12613-12622. b) Dunitz, J. D.; Taylor, R. *Eur. J. Chem*. **1998**, *3*, 89-98.

6) Brestscher, L.; Jenkins, C.; Taylor, K.; DeRider, M.; Raines, R. T. *J. Am.Chem. Soc*. **2001**, *123*, 777-778.

7) a) Inouye, K.; Sakakibara, S.; Prockop, D. J. Biochem. Biophys. Acta 1976, 420, 133-141. b) Inouye, K.; Kobayashi, Y.; Kyogoku, Y.; Kishida, Y.; Sakakibara, S.; Prockop, D. *J. Arch. Biochem. Biophys*. **1982**, *219*, 198-203.

8) Babu, I. R.; Ganeh, K. N. *J. Am. Chem. Soc*. **2001**, *123*, 2079-2080.

9) a) Nagarajan, V.; Kamitori, S.; Okuyama, K. *J. Biochem*. **1999**, *125*, 310-318 b) Berisio, R.; Vitagliano, L.; Mazzarella, L.; Zagari, A. *Protein Science* **2002**, *11*, 262-270.

10) a) Import, R.; Mele, F.; Crescenzi, O.; Benzi, C.; Barone, V. *J. Am. Chem. Soc*. **2002**, *124*, 7857-7865. b) DeRider, M. L.; Weinhold, F.; Raines, R. T.; Markely, J. L. *J. Am. Chem. Soc*. **2002**, *124*, 2497.

11) Vitagliano, L.; Berisio, R.; Mastrangelo, A.; Mazzarella, L.; Zagari, A. *Protein Science* **2001**, *10*, 2627-2632.

12) Bretscher, L.; Jenkins, C.; Taylor, K. M.; DeRider, L. M.; Raines, R. *J. Am. Chem. Soc*. **2001**, *123*, 777-778.

13) Haasnoot, C. A.; De Leeuw, A. A. M.; De Leeuw, H. P. M.; Altona, C. *Biopolymers* **1981**, *20*, 1211-1245.

14) Karplus, M. *J. Chem. Phys.* **1959**, *30*, 11.

15) (a) Panasik, N., Jr.; Eberhardt, E. S.; Edison, A. S.; Powell, D. R.; Raines, R. T. *Int. J. Pept. Protein Res*. **1994**, *44*, 262-269. (b) Gerig, J. T.; McLeod, R. S. *J. Am. Chem. Soc*. **1973**, *95*, 5725-5729.

16) (a) Eberhardt, E. S.; Panasik, N., Jr.; Raines, R. T. *J,Am. Chem. Soc*. **1996**, *118*, 12261-12266. (b) Holmgren, S. K.; Taylor, K. M. Bretscher, L. E.; Raines, R. T. *Nature* **1998**, *392*, 666-667. (c) Holmgern, S. K.; Bretscher, l. e.; Taylor, K. M.; Raines, R. T. *Chem. Biol*. **1999**, *6*, 63-70.

17) Bretscher, L. E.; Jenkins, C. L.; Taylor, K. M. ; DeRider, M. L.; Raines, R. T. *J. Am. Chem. Soc.* **2001**, *123*, 777-778.

18) (a) For a monograph on solution-phase methods see Bodansky, M.; Bodansky, A. *The practice of peptide synthesis* **1984**, Springer-Verlog, Berlin. (b) For an account of solid-phase methods see Stewart, J. M.; Young, J. D. *Solid Phase Peptide Synthesis* **1969**, W. H. Freeman & Co, NY.

19) (a) Dryland, A.; Sheppard, R. C. *Tetrahedron* **1988**, *44*, 859. (b) Atherton, E. *J. Chem. Soc. Chem. Commun.* **1986**, 1763. (c) Cammeron, L. *J. Chem. Soc. Chem. Commun.* **1987**, *270*.

20) Kaiser, E.; Colescott, R.L.; Bossinger, C.D.; Cook, P.I. *Anal. Biochem*. **1970**, *34*, 595-598.

21) Kaiser, E.; Bossinger, D. D.; Colescott, R. L.; Olsen, D. B.; *Anal. Chim. Acta* **1980,** *188*, 149-151.

22) Berg, R. A.; Olsen, B. R.; Prockop, D. J. *J. Biol. Chem.* **1970**, *245*, 5759-5763.

23) Venugopal, M. G.; Ramshaw, J. A. M.; Braswell, E.; Zhu, D.; Brodsky, B. *Biochemistry* **1994**, *33,* 7948-7956.

24) Bodansky, M.; Bodansky, A. *The Practice of Peptide Synthesis* **1984**, pp 4 & 261, Springer-Verlog, Berlin.

25) Glastone, S. A. *Text Book of Physical Chemistry 2nd ed*. Macmillan India Ltd., pp 1002-1005.

26) Izatt, R. M.; Hansen, L. D.; Rytting, J. H.; Chirstensen, J. J. *J. Am. Chem. Soc.* **1965**, *87*, 2760-2761.

27) (a) Reynolds, W. F. *Prog. Phys. Org. Chem.* **1983**, *14*, 165-203, (b) Epshtein, L. M.; *Russ. Chem. Rev.* **1979**, *48*, 854-867. (c) Thibaudeu, C.; Plavec, J.; Chattopadhyaya, J. *J. Org. Chem*. **1996**, *61*, 266-286.

28) (a) Kobayashi, Y.; Isemura, T. *Progress in Polymer science* **1992** Vol. 3, Okamura, S.; Takayanagi, M.; Eds., Kodnasha Limted, 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) Inouye, 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.

29) Feng, Y.; Melacini, G.; Taulane, J. P.; Goodmann, M. *J. Am. Chem.Soc*. **1996**, *118*, 10351-10358.

30) (a) Olsen, B. R.; Berger, A.; Sakakibara, S. Kishida, Y.; Prockop. D. J. *Arch. Biochim. Biophys*. **1971**, *57*, 589-595. (b) Miller, A.; Wray, J. S. *Nature* **1971**, *230*, 437-439. (c) Bruns, R. R.; Gross, J. *Biochemistry* **1973**, *12*, 808-815.

31) (a) Sakakibara, S.; Kishida, Y.; Okuyama, K.; Tanaka, N.; Ashida, T.; Kakudo, M. *J. Mol. Biol.* **1972**, *65*, 371-373. (b) Fraser, R. D. B.; Macrae, T. P.; Suzuki, E. *J. Mol. Biol.* **1979**, *129*, 463-481. (c) Bella, J.; Eaton, M.; Brodsky, B.; Berman, H. M. *Science* **1994**, *266*, 75-81.

32) (a) Von Hippel, P. H. in *Treaties of Collagen*, **1967**, Ramachandran, G. N. Ed., Academic Press, London & New York, pp. 258-338. (b) Piez, K. A. *in treaties on Collagen,* **1967,** Ramachandran, G. N. Ed., Academic Press, London & New York, pp. 207-252. (c) Engel, J. *Arch. Biochem. Biophys*. **1962**, *97*, 150-154. (d) Venugopal M. G.; Ramshaw, J. A. M.; Braswell, E.; Zhu, D.; Brodsky, D. *Biochemistry* **1994**, *33*, 7948-7956.

33) Franks, F. *Biophys. Chem.* **2002**, *96*, 117-127.

34) Kalinichev G. *in Molecular Modeling Theory: Applications in the Geosciences.* Ed. Cygan, R. T.; Kubikicki, J. D. *Rev. Mineralogy Geochem*. **2001**, *42*, 83-129.

35) (a) Bandyopadhyay, S.; Chakraborty, S.; Balasubramanian, S.; Bagchi, B. *J. Am. Chem. Soc.* **2005,** *127*, 4701-4075. (b) Bandyopadhyay, S.; Chakraborty, S.; Bagchi, B. *J. Am. Chem. Soc*. **2005**, *127*, 16660-16667.

36) Kurkal, V.; Daniel, R. M.; Finney, J. L.; Tehei, M.; Dunn, R. V.; Smith, J. C. *Chem. Phys.* **2005**, *137*, 267-273.

37) (a) Oleinikova, A.; Smolin, N.; Brovchenko, I.; Geiger, A.; Winter, R. *J. Phys. Chem. B* **2005**, *109*, 1988-1999. (b) Smolin, N.; Oleinikove, A.; Brovchenko, I.; Geiger, A.; winter, R. *J. Phys. Chem. B* **2005**, *109*, 10995-11005.

38) Nakasako, M. *Phil. Trans. R. Soc. Lond. B* **2004**, *359*, 1191-1206.

39) Creighton, T. E. *Proteins* 1993, W. h. Feeman & Co., NY, pp. 262-263.

40) (a) Von Hippel, P. H.; Wong, K. *Biochemistry* **1962**, *1*, 664-674. (b) Von Hippel, P. H.; Wong, K. *Biochemistry* **1963**, *2*, 1387-1398.

41) (a) Komsa-penkova, R.; Koynova, R.; Kostov, G.; Techov, B. G. *Biochim. Biophys. Acta* **1996**, *1297*, 171-181. (b) Zanaboni, G.; Rossi, A.; Onana, A. M. T.; Tenni, R. *Matrix Biol*. **2000**, *19*, 511-520.

42) Fernandez, F.; Scherega, H. A. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 113-118.

43) Ikura, T.; Urakubo, Y.; Nobutoshi, I. *Chem. Phys.* **2004**, *307*, 111-119.

44) Pacaroni, A.; Cinelli, S.; Cornicchi, E.; de Francesco, A.; Onori, G. *Chem. Phys. Lett.* **2005**, *410*, 400-403.

45) Cheung, M. S.; Garcia, A. E.; Onuchic, J. N. *Proc. Nat. Acad. Sci. USA* **2002**, *99* 685-690.

46) (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.

47) (a) Harrap, B. S. *Int. J. Pept. Protein Res*. **1969**, *1*, 2527-2532. (b) Gekko, K.; Koga, S. *J. Biochem*. **1983**, *94*, 199-205.

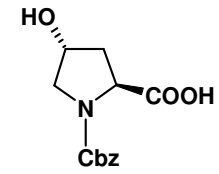
48) (a) Brown, F. R., III.; Carver, J. P.; Blout, E. R. *J. Mol. Biol.* **1996**, *39*, 307-313. (b) Brown, F. R., III.; Di Croato, A.; Lorenzi, G. P.; Blout, E. R. *J. Mol. Biol*. **1972**, *63*, 85-99. (c) Vuilleumier, S.; Mutter, M. *Biopolymers* **1993**, *33*, 389-400. (d) Choma, C. T.; Lear, J. D.; Nelson, M. J.; Dutton, P. L.; Roberson, D. E.; De Grado, W. F. *J. Am. Chem. Soc*. **1994**, *116*, 856-865.

Appendix I

Compound No.	Spectra	Page No
Compound **6** ……	^{1}HNMR,^{13}CNMR, DEPT…….	121
Compound **7** ……	^{1}HNMR,^{13}CNMR, DEPT…….	122
Compound **8**……	^{1}HNMR,^{13}CNMR, DEPT…….	123
Compound **9**……	^{1}HNMR,^{13}CNMR, DEPT…….	124
Compound **10**……	^{1}HNMR,^{13}CNMR, DEPT…….	125
Compound **12**……	^{1}HNMR,^{13}CNMR, DEPT…….	126
Compound **14**……	^{1}HNMR,^{13}CNMR, DEPT…….	127
Compound **15**……	^{1}HNMR,^{13}CNMR, DEPT…….	128
Compound **16**……	^{1}HNMR,^{13}CNMR, DEPT…….	129
Compound **18**……	^{1}HNMR,^{13}CNMR, DEPT…….	130
Compound **9**& Compound **15**…….. IR……		131
Peptide **19** ……	RP-18 HPLC and MALDI-TOF…….	132
Peptide **20** ……	RP-18 HPLC and MALDI-TOF…….	133
Peptide **21** ……	RP-18 HPLC and MALDI-TOF…….	134
Peptide **22** ……	RP-18 HPLC and MALDI-TOF…….	135
Compounds **12** and **18** ……. LC-MS ……..		137

Acquisition Time (sec)	2.5166	**Comment**	Amritraj patra ;1H		**Date**	19/05/2003 22:41:00	
Frequency (MHz)	500.13	**Nucleus**	1H	**Original Points Count** 16384	**Points Count**	16384	**Sweep Width (Hz)** 6510.42
Temperature (grad C)	0.000						

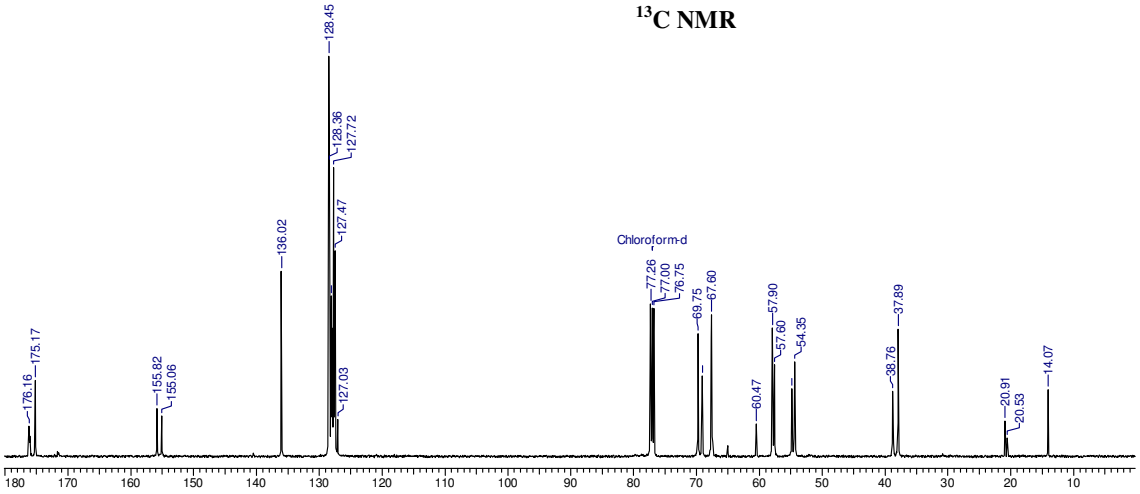
Compound 6



^{1}H NMR

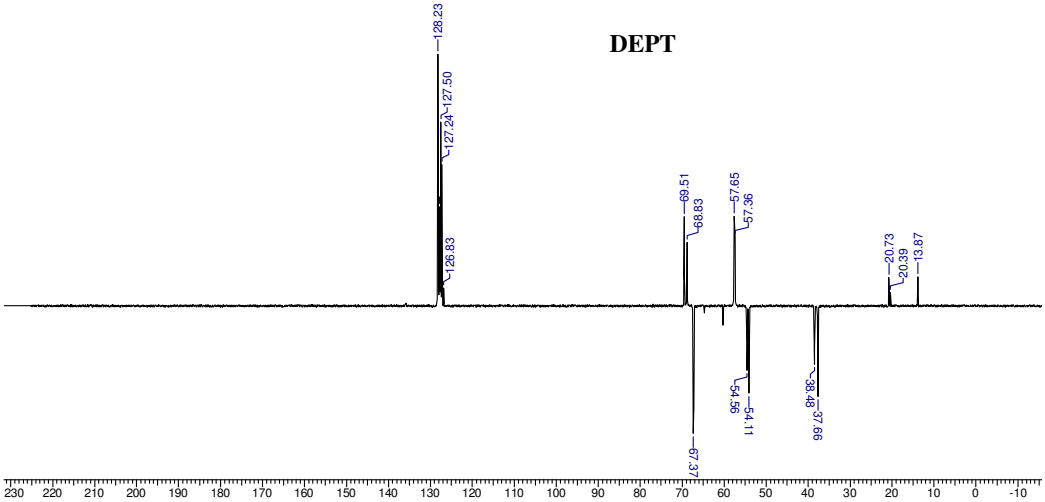
Acquisition Time (sec)	0.5210	**Comment**	Pallavi; 13C		**Date**	11/05/2003 14:34:00	
Frequency (MHz)	125.76	**Nucleus**	13C	**Original Points Count** 16384	**Points Count**	16384	**Sweep Width (Hz)** 31446.54
Temperature (grad C)	0.000						

^{13}C NMR



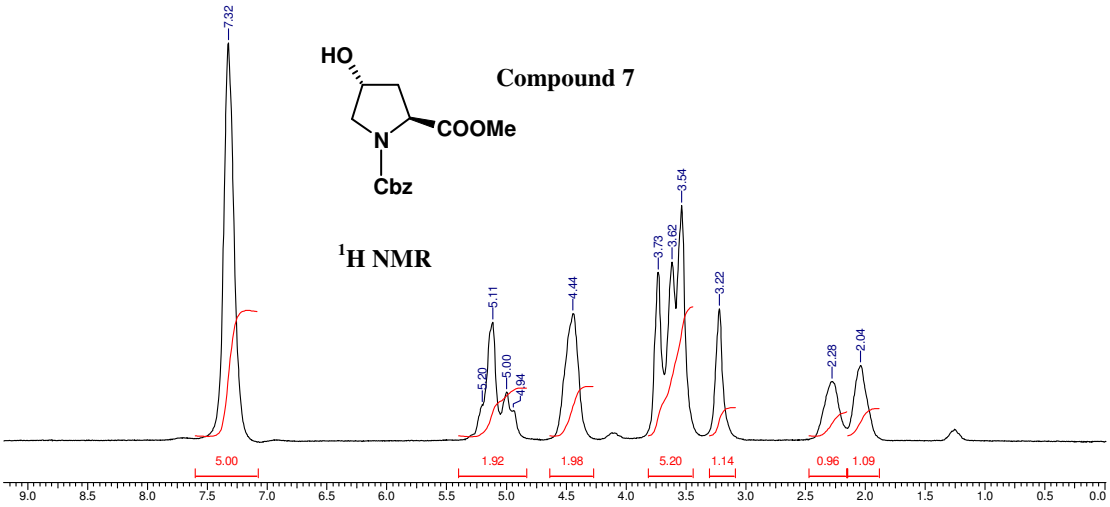
Acquisition Time (sec)	1.0420	**Comment**	13C Dept		**Date**	11/05/2003 14:34:00	
Frequency (MHz)	125.76	**Nucleus**	13C	**Original Points Count** 32768	**Points Count**	32768	**Sweep Width (Hz)** 31446.54
Temperature (grad C)	0.000						

DEPT



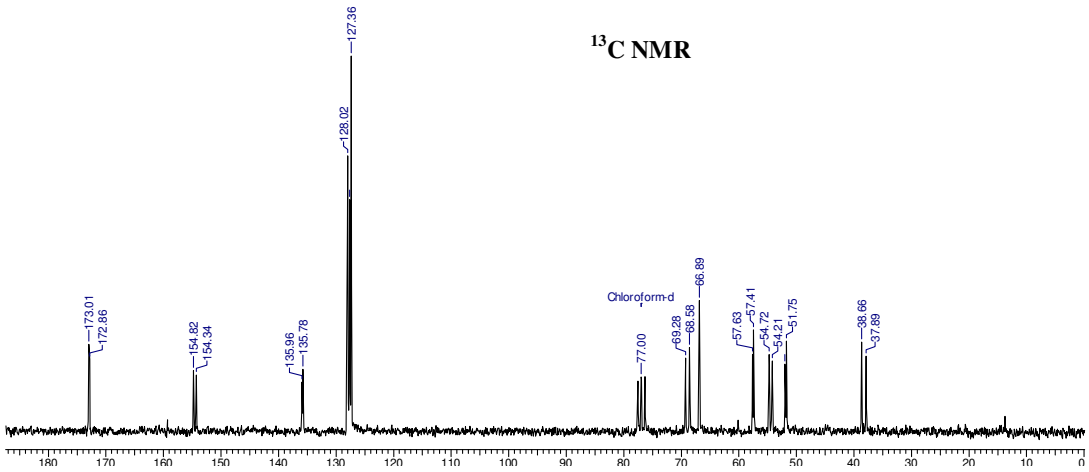
3 Jul 2006

Acquisition Time (sec)	2.0480	**Comment**	ABHI/BIA			**Date**	20/12/02 23:08:02
Frequency (MHz)	200.13	**Nucleus**	1H	**Number of Transients**	512	**Original Points Count**	8192
Sweep Width (Hz)	4000.00			**Temperature (grad C)**	24.000	**Points Count**	8192



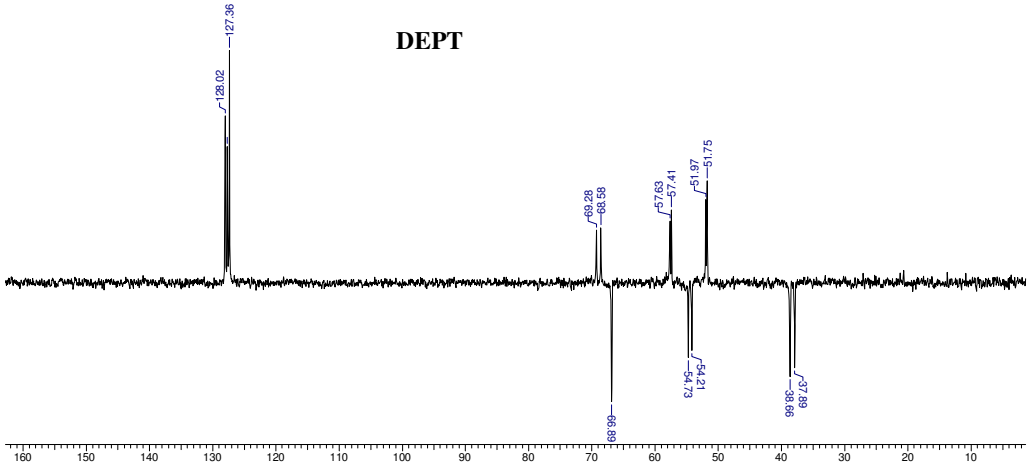
3 Jul 2006

Acquisition Time (sec)	0.5407	**Comment**	DR.R.A. JOSHI/C13/ME-2/**SVT			**Date**	21/12/02 15:10:10
Frequency (MHz)	50.32	**Nucleus**	13C	**Number of Transients**	20000	**Original Points Count**	8192
Sweep Width (Hz)	15151.52	**Temperature (grad C)**	24.000			**Points Count**	8192



3 Jul 2006

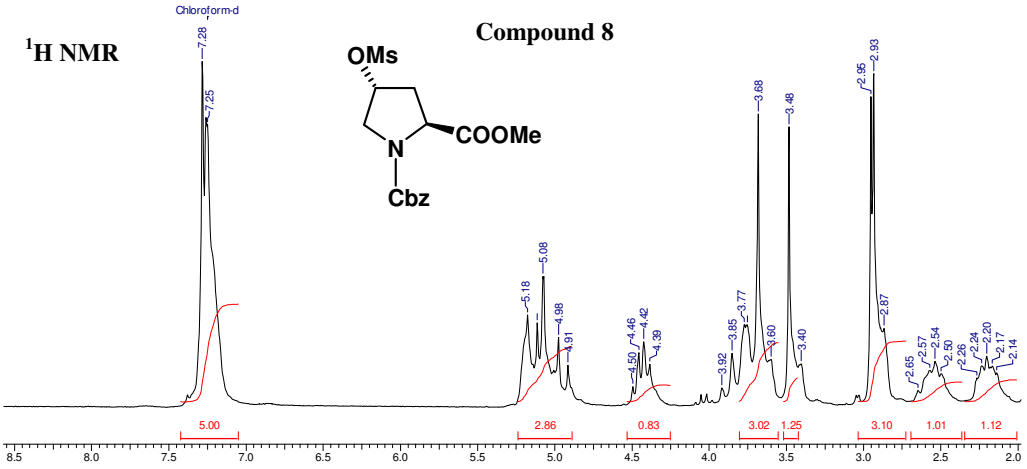
Acquisition Time (sec)	0.5407	**Comment**	D.K. MAHAPATRA/DEPT/**SVT			**Date**	30/04/02 10:18:27
Frequency (MHz)	50.32	**Nucleus**	13C	**Number of Transients**	2048	**Original Points Count**	8192
Sweep Width (Hz)	15151.52	**Temperature (grad C)**	24.000			**Points Count**	8192



Acquisition Time (sec) 2.0480	***Comment*** WR WED5AC.064			***Date*** 27/11/02 17:23:05			
Frequency (MHz) 200.13	***Nucleus*** 1H	***Number of Transients*** 512		***Original Points Count*** 8192	***Points Count*** 8192		
Sweep Width (Hz) 4000.00	***Temperature (grad C)*** 24.000						

^{1}H NMR

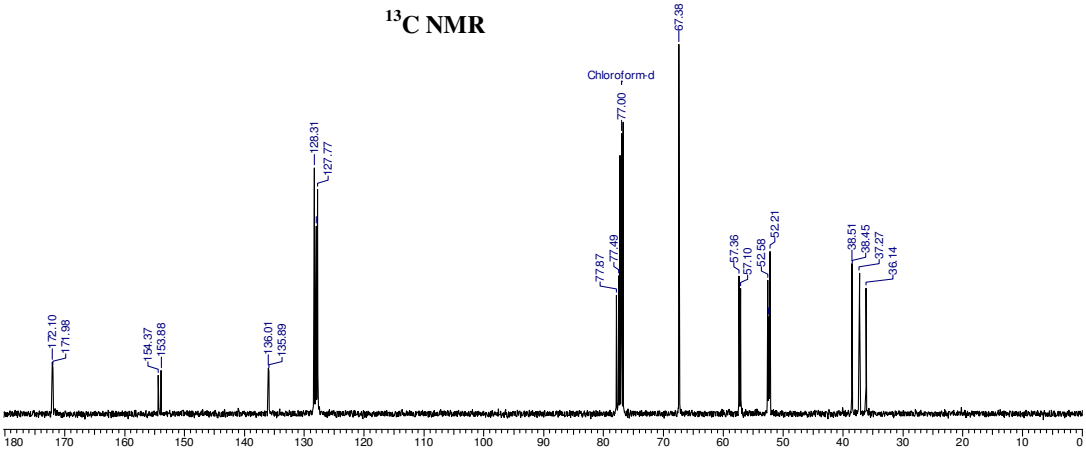
Compound 8



3 Jul 2006

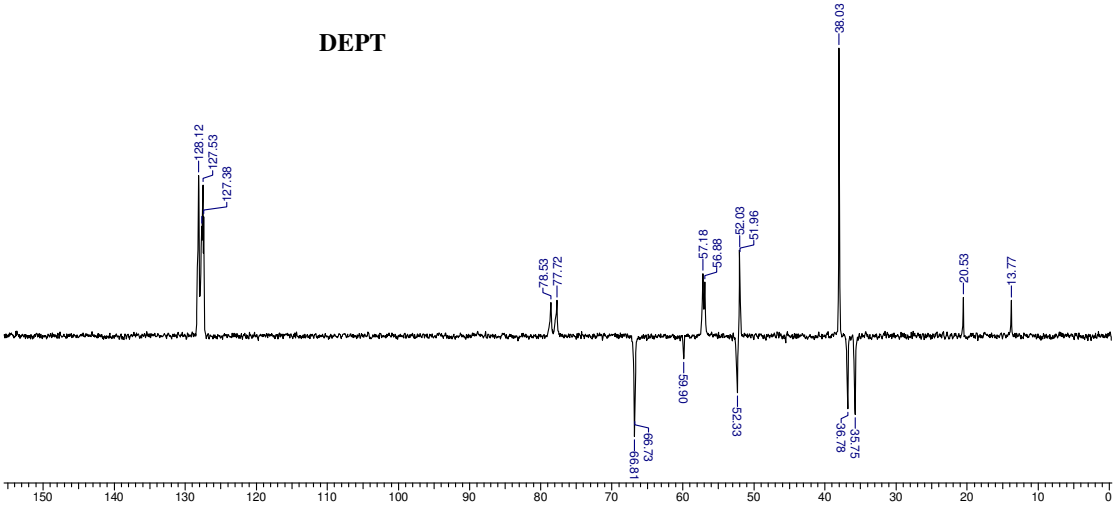
Acquisition Time (sec) 0.5210	***Comment*** UMASHANKA/13C//**SVT			***Date*** 14/05/2003 14:33:00			
Frequency (MHz) 125.76	***Nucleus*** 13C	***Original Points Count*** 16384		***Points Count*** 16384		***Sweep Width (Hz)*** 31446.54	
Temperature (grad C) 0.000							

^{13}C NMR



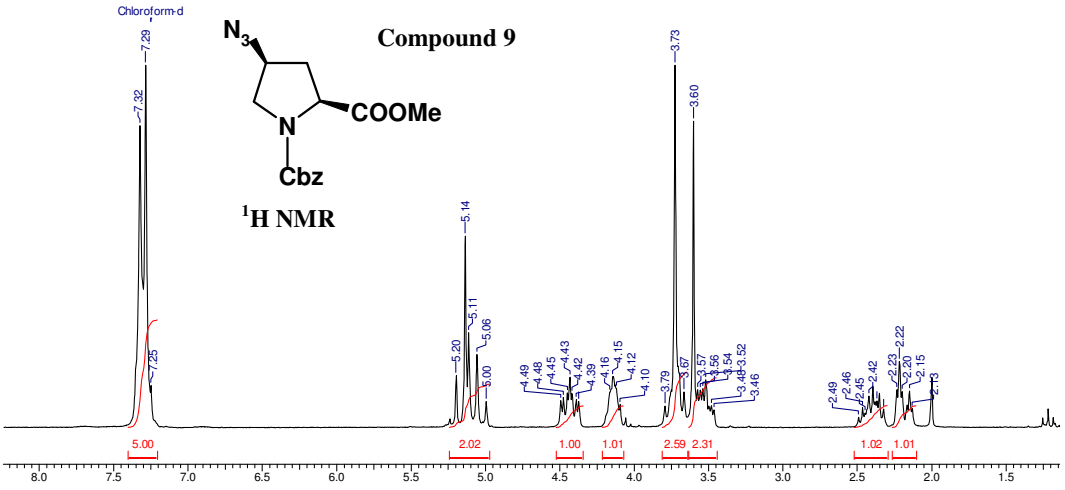
Acquisition Time (sec) 0.5407	***Comment*** UMA		***Date*** 24/04/02 08:54:30			***Frequency (MHz)*** 50.32	
Nucleus 13C	***Number of Transients*** 2048		***Original Points Count*** 8192		***Points Count*** 8192	***Sweep Width (Hz)*** 15151.52	
Temperature (grad C) 24.000							

DEPT



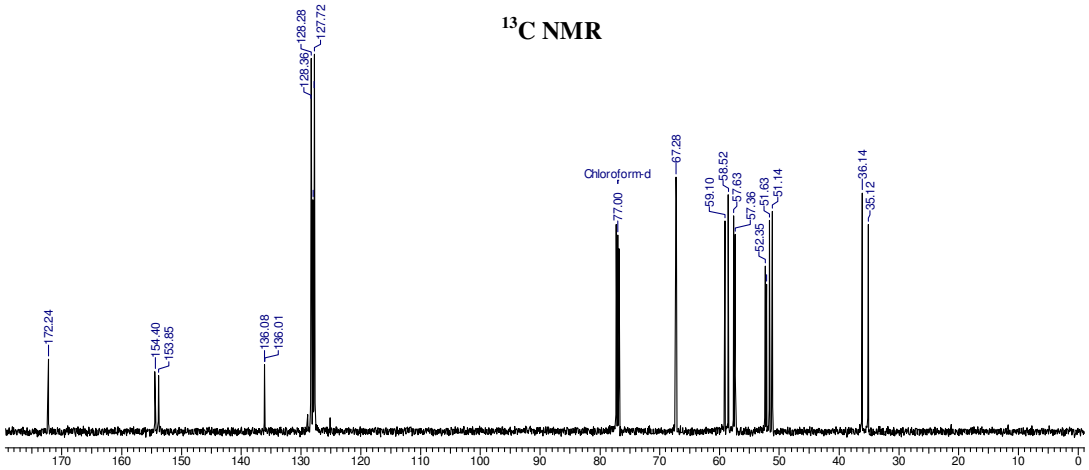
3 Jul 2006

Acquisition Time (sec) 7.9167	***Comment***	M. Umashankar		***Date***	24/12/2004 15:18:44		
Frequency (MHz) 200.13	***Nucleus***	1H	***Original Points Count*** 32768	***Points Count***	32768	***Sweep Width (Hz)***	4139.07
Temperature (grad C) 0.000							

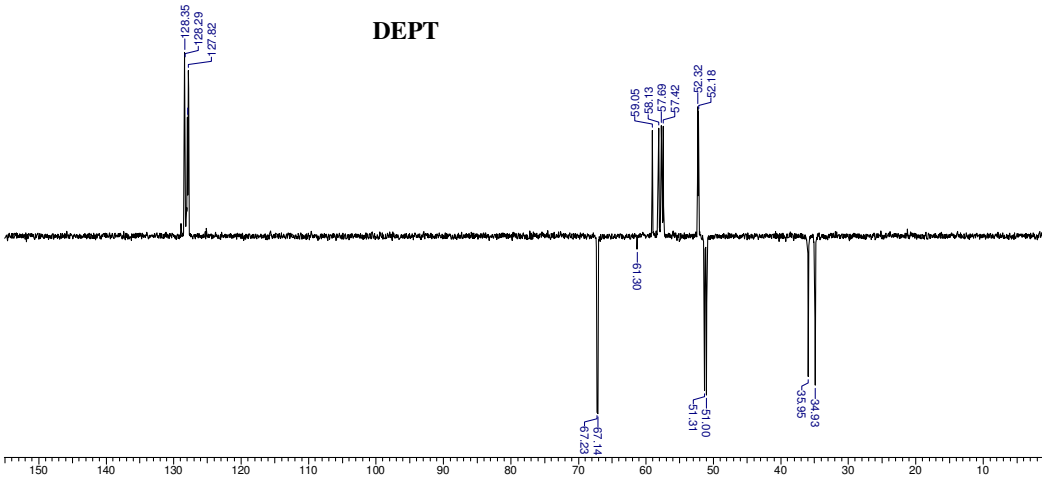


3 Jul 2006

Acquisition Time (sec) 0.5210	***Comment***	Umashankara;13C		***Date***	12/07/2003 15:03:00		
Frequency (MHz) 125.76	***Nucleus***	13C	***Original Points Count*** 16384	***Points Count***	16384	***Sweep Width (Hz)***	31446.54
Temperature (grad C) 0.000							

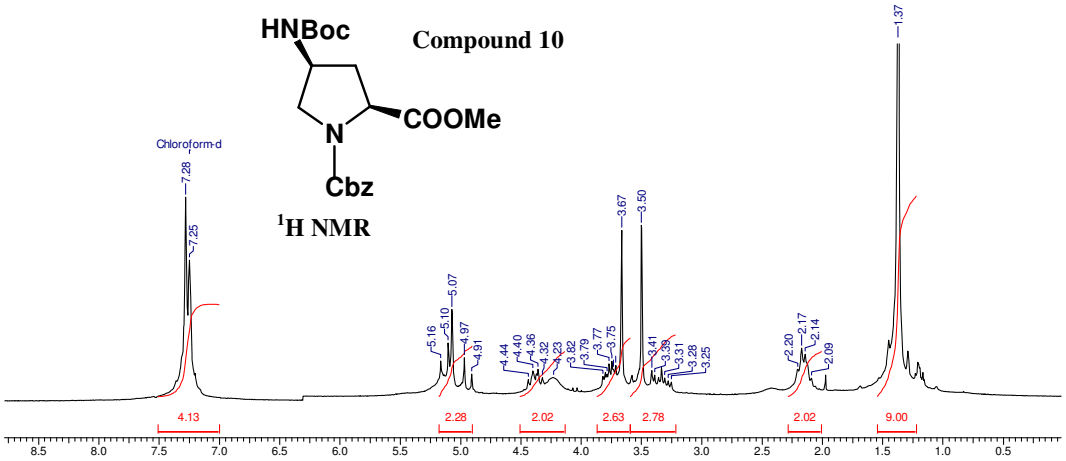


Acquisition Time (sec) 1.0420	***Comment***	13C Dept		***Date***	18/06/2003 21:36:00		
Frequency (MHz) 125.76	***Nucleus***	13C	***Original Points Count*** 32768	***Points Count***	32768	***Sweep Width (Hz)***	31446.54
Temperature (grad C) 0.000							



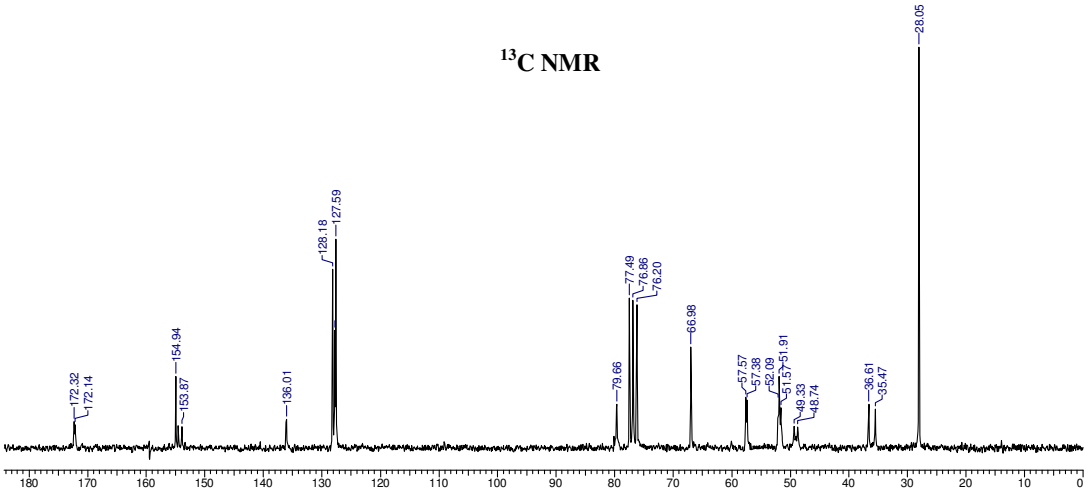
3 Jul 2006

Acquisition Time (sec)	2.0480	**Comment**	WR SUN4AC.028		**Date**	23/06/02 17:11:15	
Frequency (MHz)	200.13	**Nucleus**	1H	**Number of Transients** 512	**Original Points Count**	8192	**Points Count** 8192
Sweep Width (Hz)	4000.00	**Temperature (grad C)**	24.000				

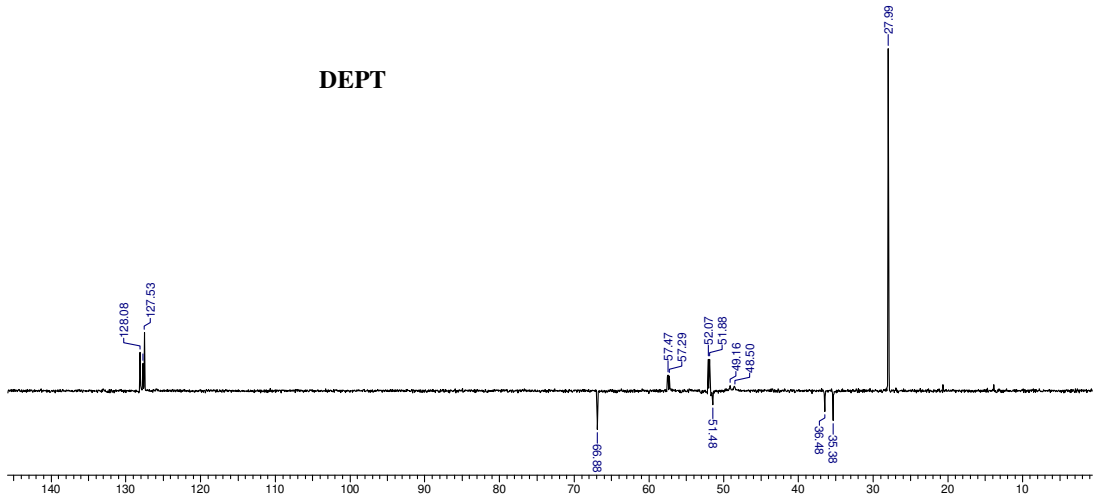


3 Jul 2006

Acquisition Time (sec)	0.5407	**Comment**	UMA SOMSEKHAR/C13/UMAC-031B/**SVT		**Date**	20/05/02 13:13:36	
Frequency (MHz)	50.32	**Nucleus**	13C	**Number of Transients** 2048	**Original Points Count**	8192	**Points Count** 8192
Sweep Width (Hz)	15151.52	**Temperature (grad C)**	24.000				

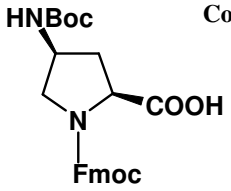


Acquisition Time (sec)	0.5407	**Comment**	UMAC-031B/DEPT		**Date**	20/05/02 13:46:41	
Frequency (MHz)	50.32	**Nucleus**	13C	**Number of Transients** 2048	**Original Points Count**	8192	**Points Count** 8192
Sweep Width (Hz)	15151.52			**Temperature (grad C)** 24.000			

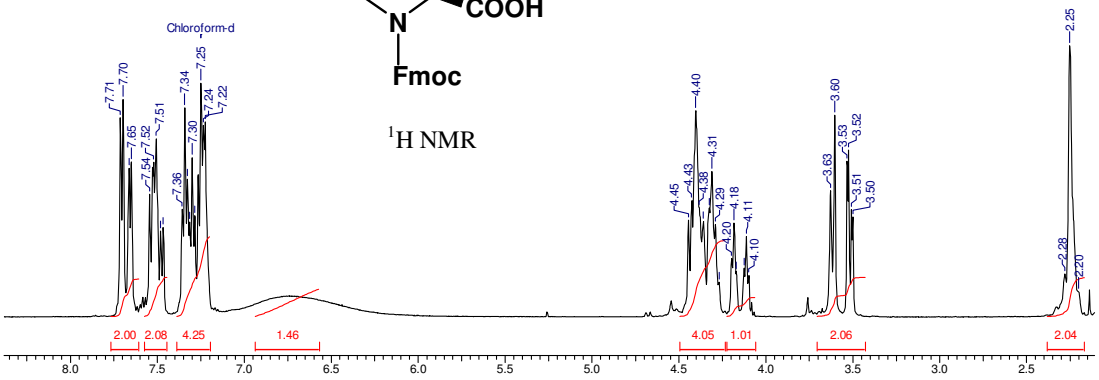


Acquisition Time (sec)	2.1823	***Comment***	1H ; CDCl3; Umashankar; OHphenylCOOH;			***Date***	05/07/2003 10:04:00		
Frequency (MHz)	500.13	***Nucleus***	1H	***Original Points Count***	16384	***Points Count***	16384	***Sweep Width (Hz)***	7507.51
Temperature (grad C)	0.000								

Compound 12

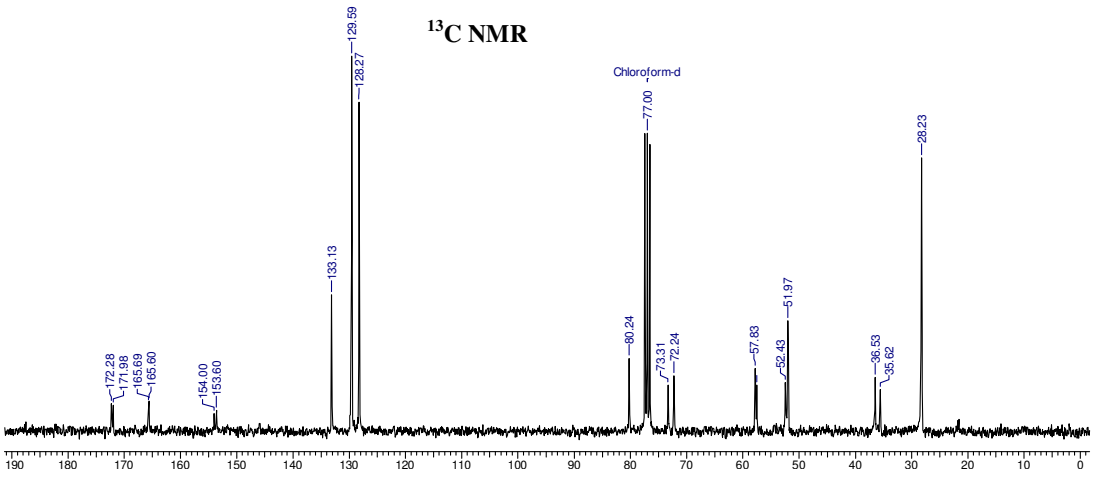


^{1}H NMR



Acquisition Time (sec)	0.4342	***Comment***	DO	***Date***	00/00/00 00:00:00			***Frequency (MHz)***	75.48
Nucleus	13C	***Number of Transients***	4000	***Original Points Count***	8192	***Points Count***	8192	***Sweep Width (Hz)***	18867.92
Temperature (grad C)	24.000								

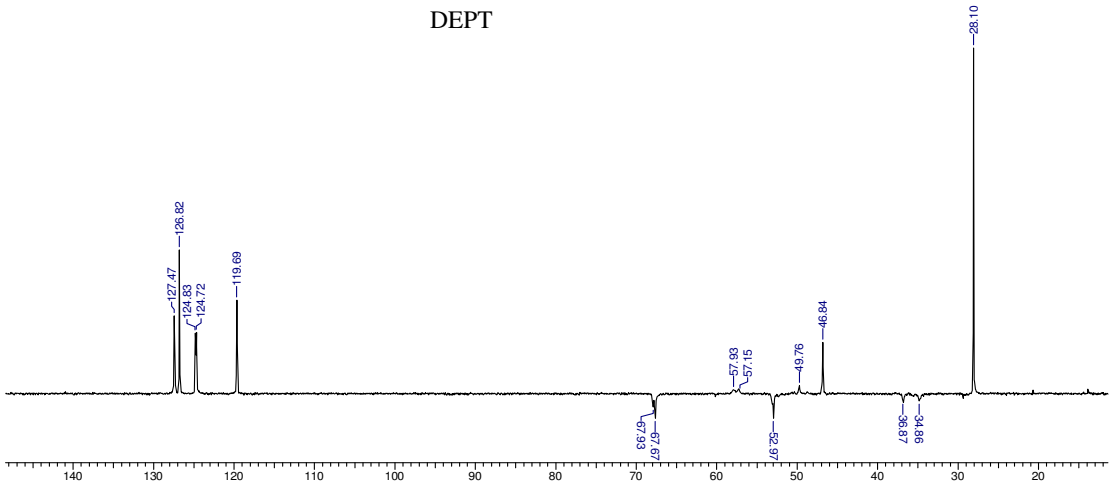
^{13}C NMR



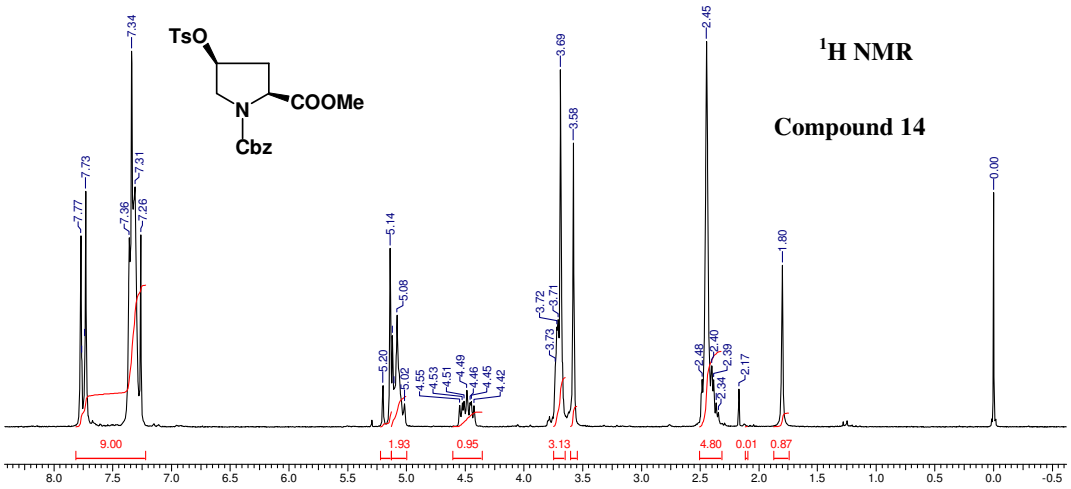
3 Jul 2006

Acquisition Time (sec)	1.0420	***Comment***	UMA SHANKAR//13C Dept//**SVT			***Date***	25/03/2003 15:28:00		
Frequency (MHz)	125.76	***Nucleus***	13C	***Original Points Count***	32768	***Points Count***	32768	***Sweep Width (Hz)***	31446.54
Temperature (grad C)	0.000								

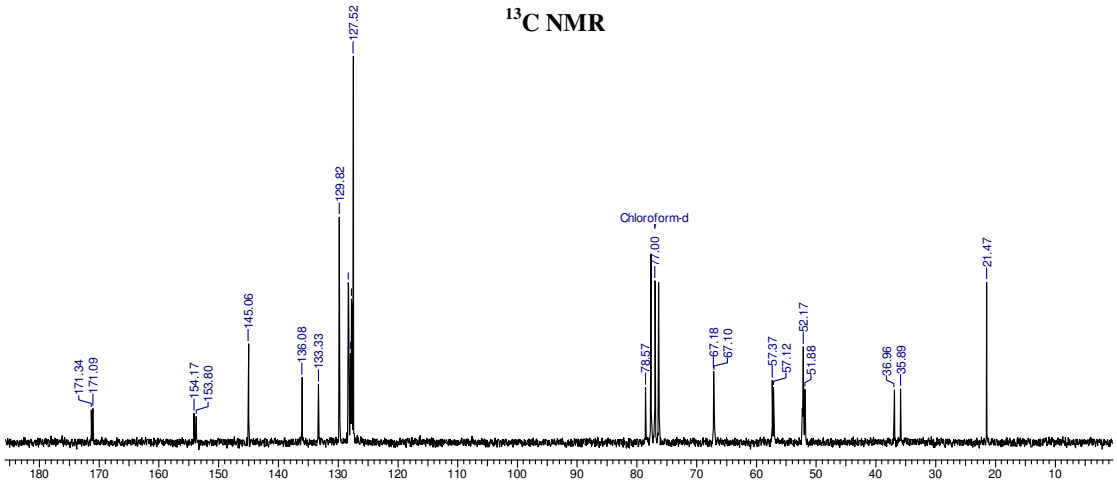
DEPT



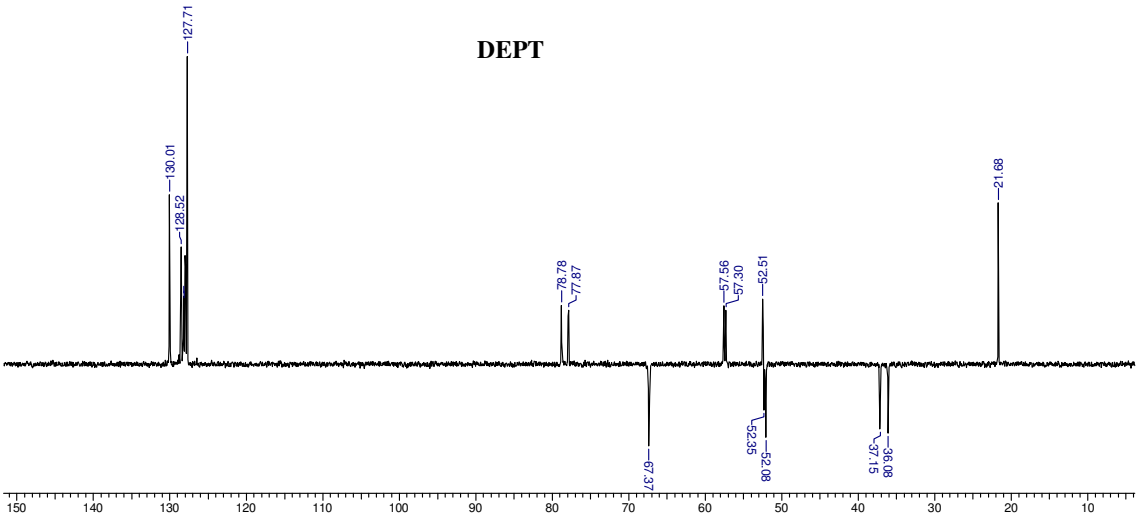
Acquisition Time (sec) 7.9167	***Comment***	Sonar Mahesh		***Date***	20/03/2006 17:34:40		
Frequency (MHz) 200.13	***Nucleus***	1H	***Original Points Count*** 32768	***Points Count***	32768	***Sweep Width (Hz)***	4139.07
Temperature (grad C) 0.000							



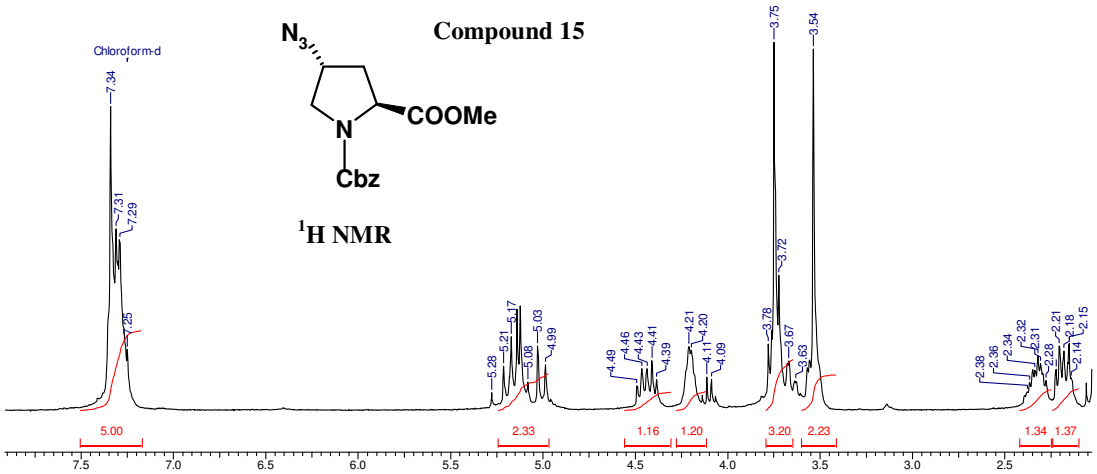
Acquisition Time (sec) 2.7329	***Comment***	Sonar mahesh		***Date***	21/03/2006 02:35:38		
Frequency (MHz) 50.32	***Nucleus***	13C	***Original Points Count*** 32768	***Points Count***	32768	***Sweep Width (Hz)***	11990.41
Temperature (grad C) 0.000							



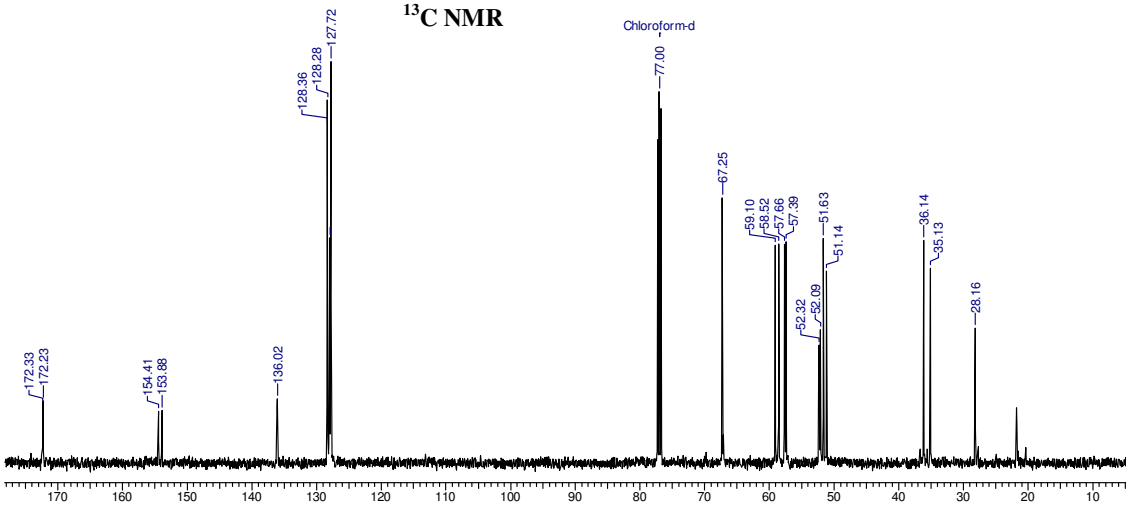
Acquisition Time (sec) 2.7329	***Comment***	Sonar mahesh		***Date***	21/03/2006 02:54:54		
Frequency (MHz) 50.32	***Nucleus***	13C	***Original Points Count*** 32768	***Points Count***	32768	***Sweep Width (Hz)***	11990.41
Temperature (grad C) 0.000							



Compound 15



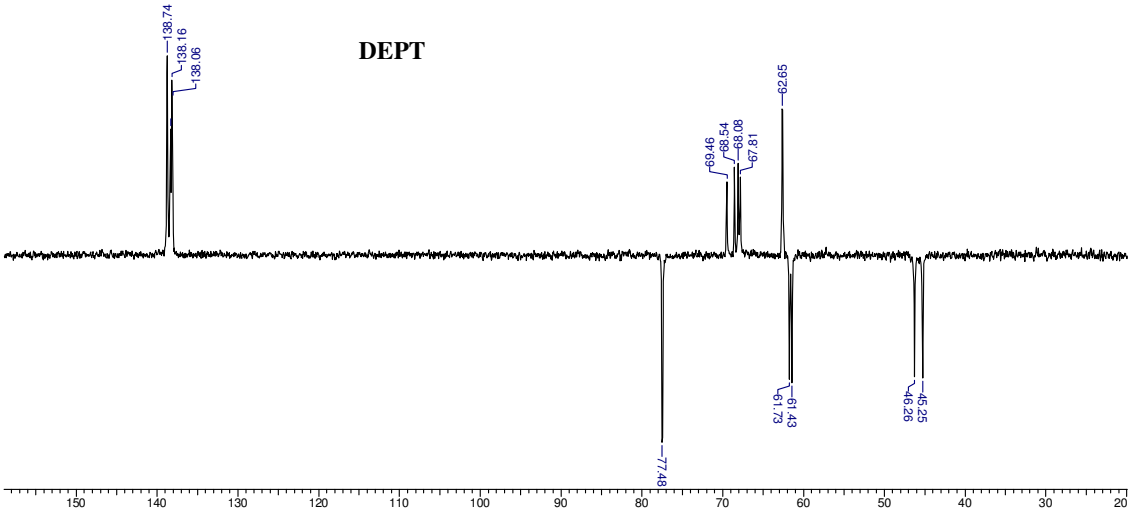
^{1}H NMR



^{13}C NMR

3 Jul 2006

Acquisition Time (sec)	0.4342	***Comment***	M UMASANKER			***Date***	00/00/00 00:00:00
Frequency (MHz)	75.48	***Nucleus***	13C	***Number of Transients***	512	***Original Points Count***	8192	***Points Count***	8192
Sweep Width (Hz)	18867.92			***Temperature (grad C)***	24.000		

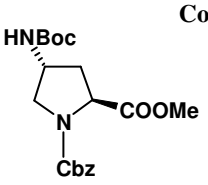


DEPT

3 Jul 2006

Acquisition Time (sec) 7.9167	***Comment*** M. Umashankara			***Date*** 30/06/2005 07:17:24			
Frequency (MHz) 200.13	***Nucleus*** 1H	***Original Points Count*** 32768		***Points Count*** 32768		***Sweep Width (Hz)*** 4139.07	
Temperature (grad C) 0.000							

Compound 16

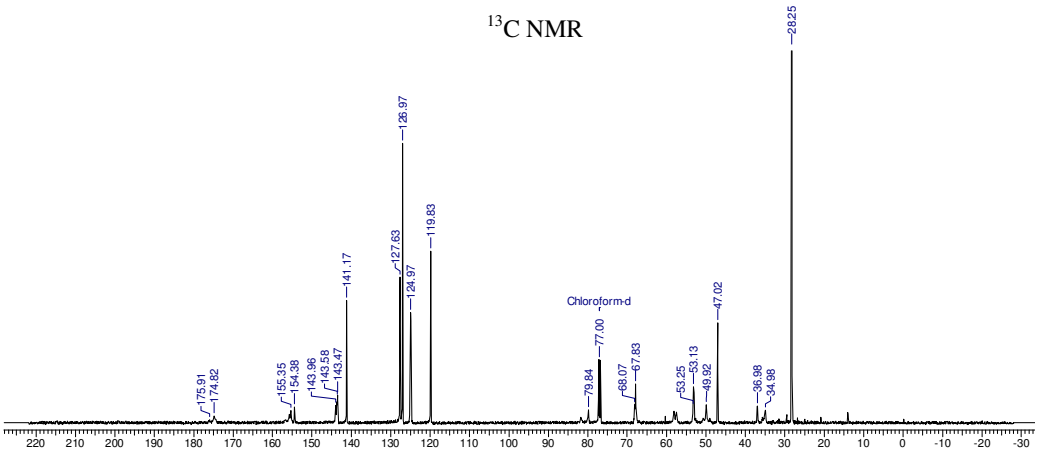


^{1}H NMR

3 Jul 200

Acquisition Time (sec) 0.5210	***Comment*** UMA SHANKAR/C13/CDCL3/**SVt			***Date*** 25/03/2003 15:21:00			
Frequency (MHz) 125.76	***Nucleus*** 13C	***Original Points Count*** 16384		***Points Count*** 16384		***Sweep Width (Hz)*** 31446.54	
Temperature (grad C) 0.000							

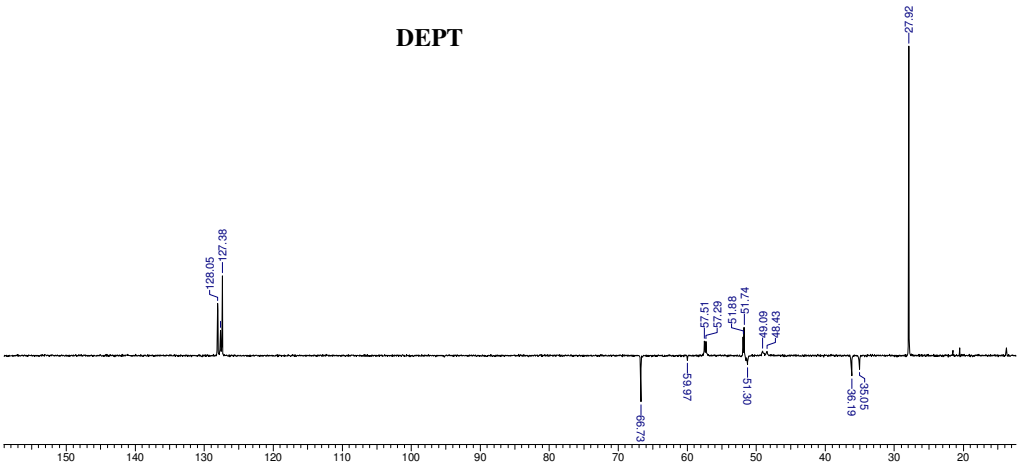
^{13}C NMR



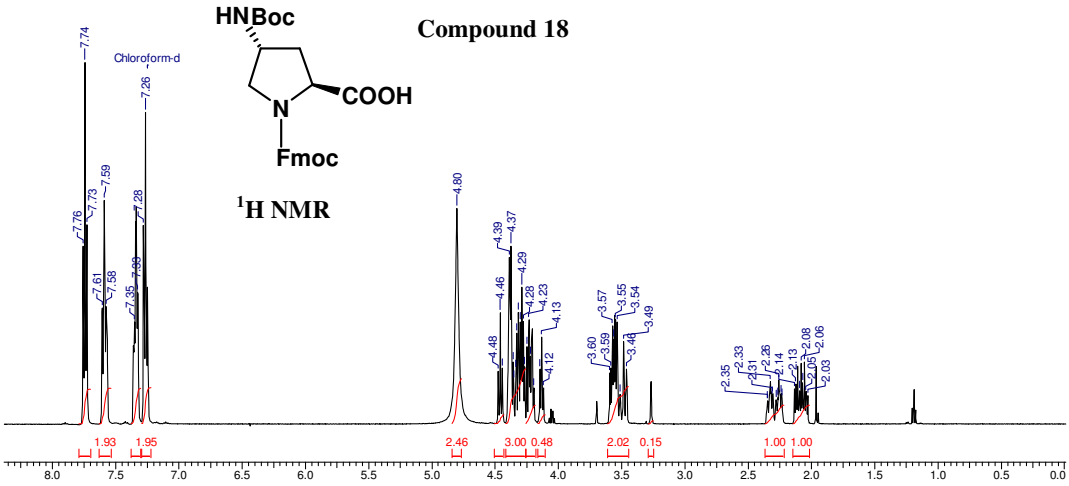
3 Jul 2006

Acquisition Time (sec) 0.5407	***Comment*** UMADEPT/CDCL3			***Date*** 21/05/02 13:15:33			
Frequency (MHz) 50.32	***Nucleus*** 13C	***Number of Transients*** 2048		***Original Points Count*** 8192		***Points Count*** 8192	
Sweep Width (Hz) 15151.52		***Temperature (grad C)*** 24.000					

DEPT

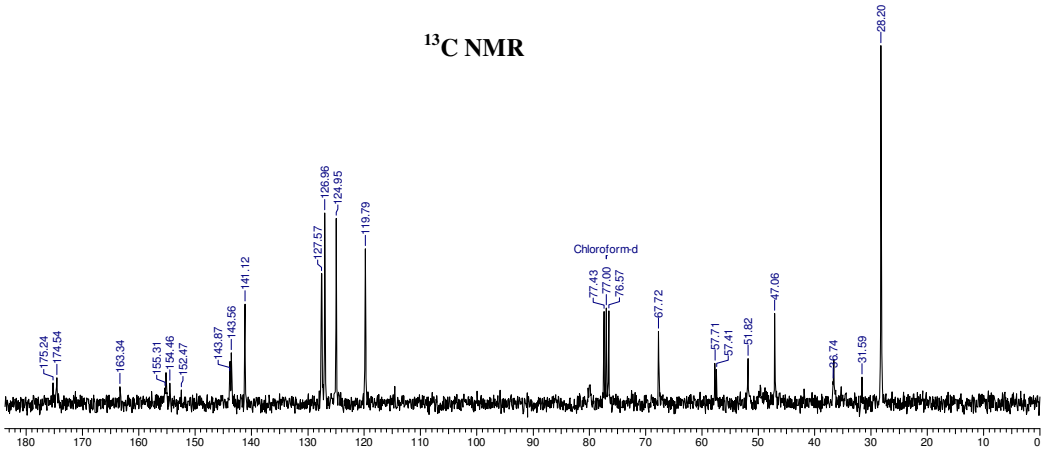


Acquisition Time (sec) 2.1823	**Comment**	UMASHANKAR/;1H//**SVT		**Date**	19/06/2003 10:13:00		
Frequency (MHz) 500.13	**Nucleus**	1H	**Original Points Count** 16384	**Points Count**	16384	**Sweep Width (Hz)**	7507.51
Temperature (grad C) 0.000							



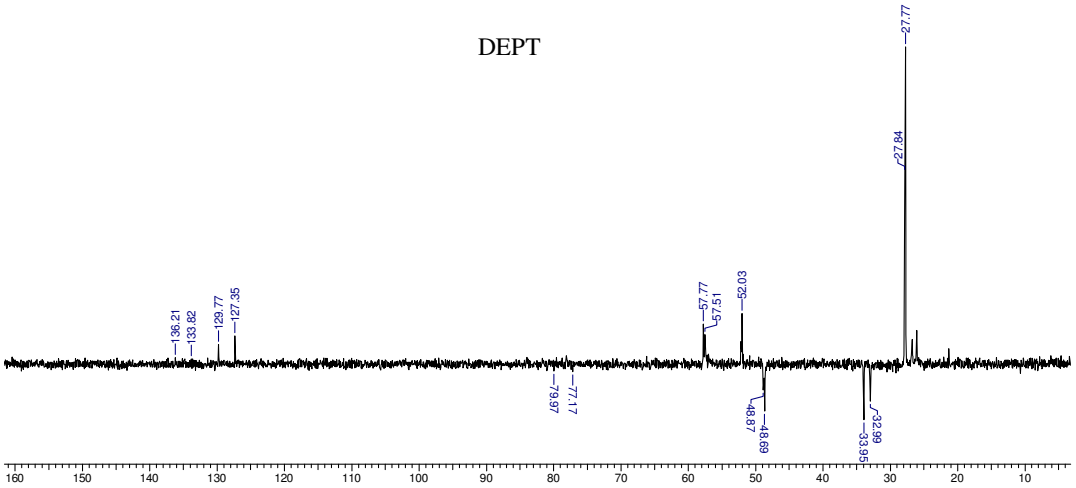
3 Jul 2006

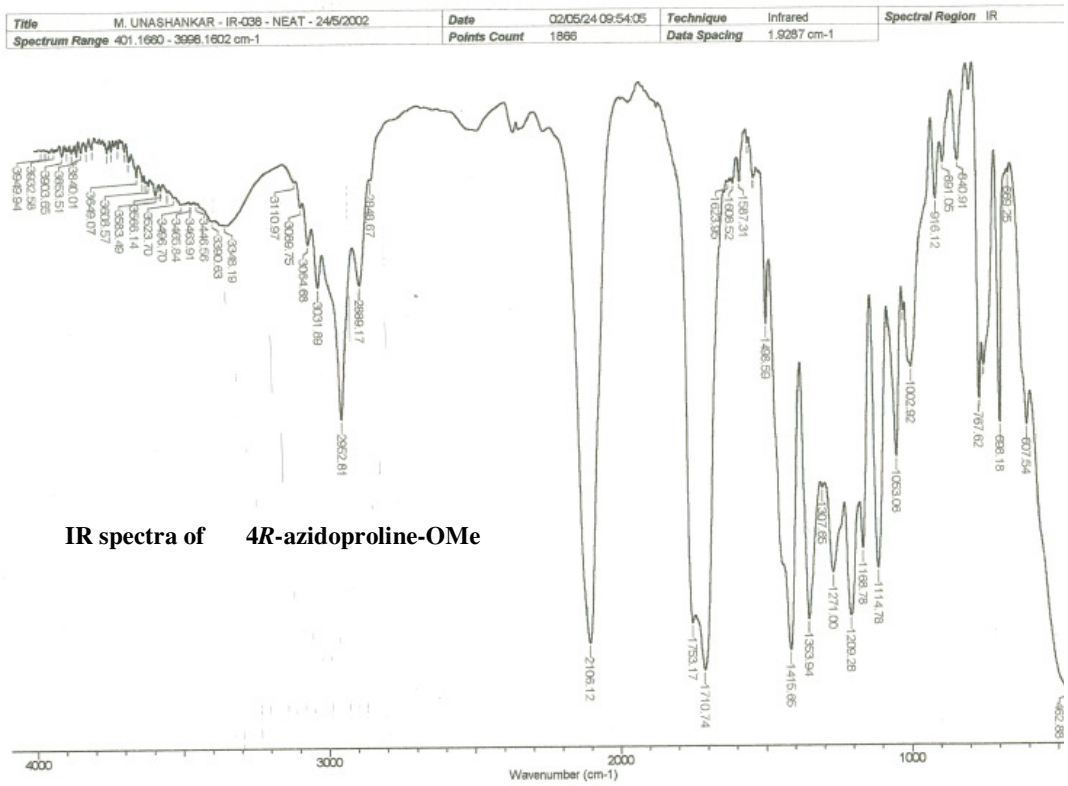
Acquisition Time (sec) 0.4342	**Comment**	UMASHANKAR/13C		**Date**	00/00/00 00:00:00		
Frequency (MHz) 75.48	**Nucleus**	13C	**Number of Transients** 2000	**Original Points Count**	8192	**Points Count**	8192
Sweep Width (Hz) 18867.92			**Temperature (grad C)** 24.000				



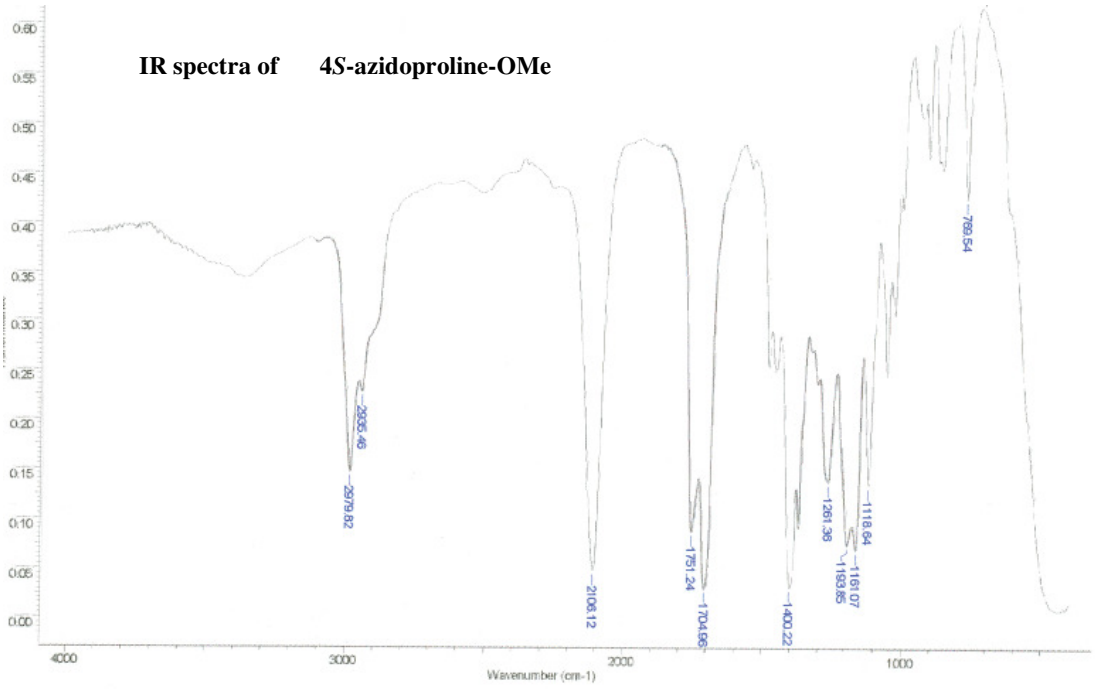
3 Jul 2006

Acquisition Time (sec) 0.5407	**Comment**	BHASKAR;ALCOHOL; 13C DEPT		**Date**	21/09/03 14:39:59		
Frequency (MHz) 50.32	**Nucleus**	13C	**Number of Transients** 1600	**Original Points Count**	8192	**Points Count**	8192
Sweep Width (Hz) 15151.52	**Temperature (grad C)** 24.000						

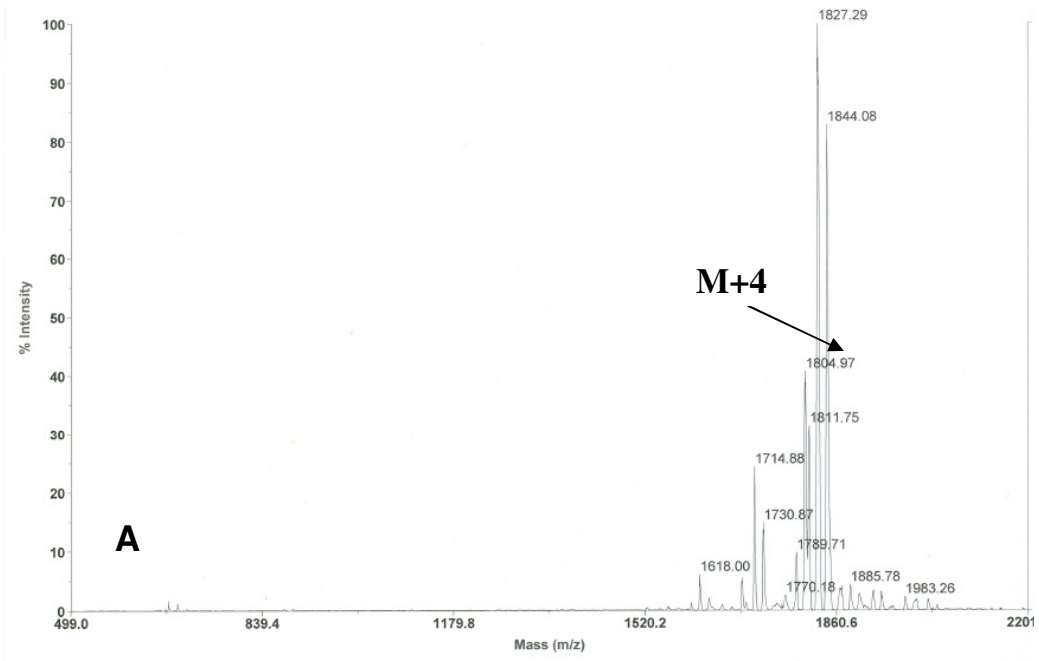


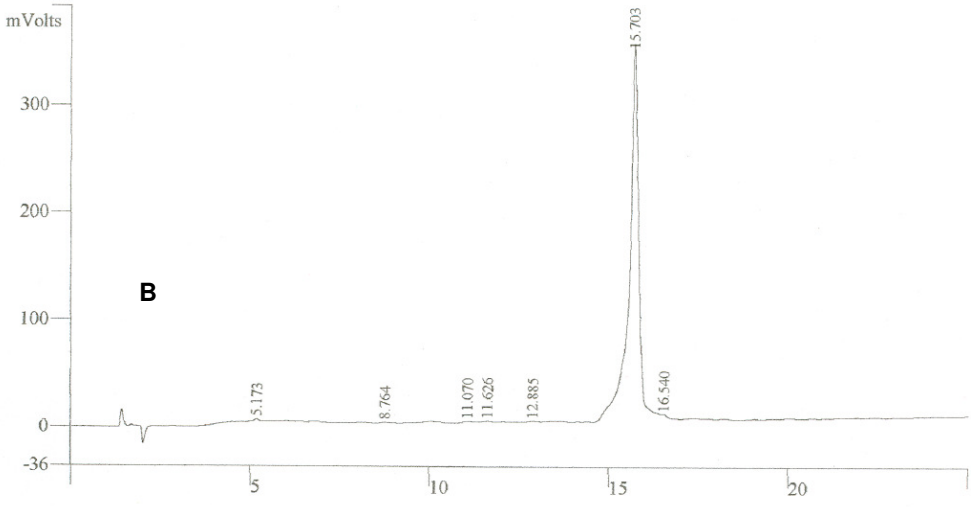


IR spectra of 4*R*-azidoproline-OMe

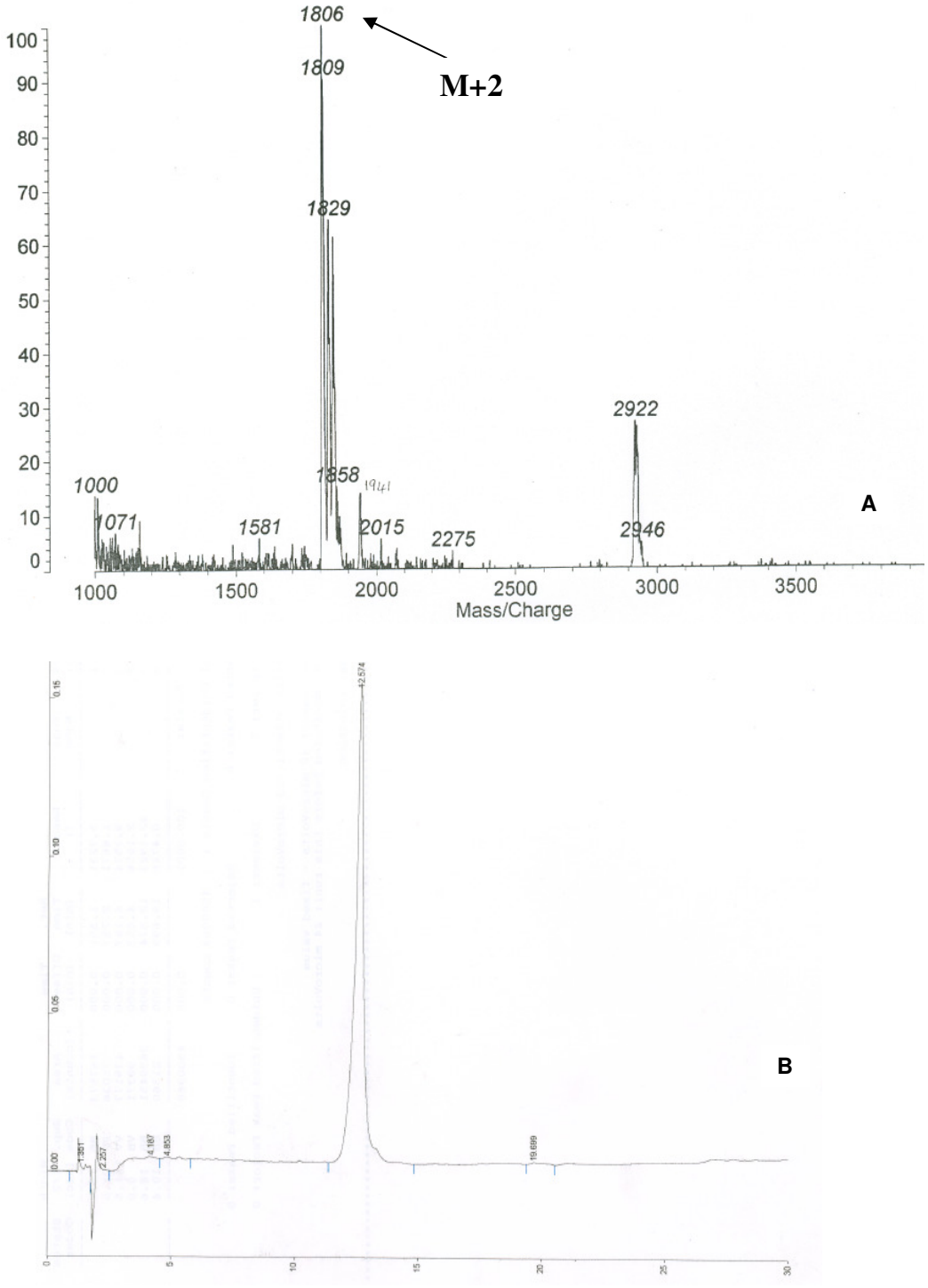


IR spectra of 4*S*-azidoproline-OMe

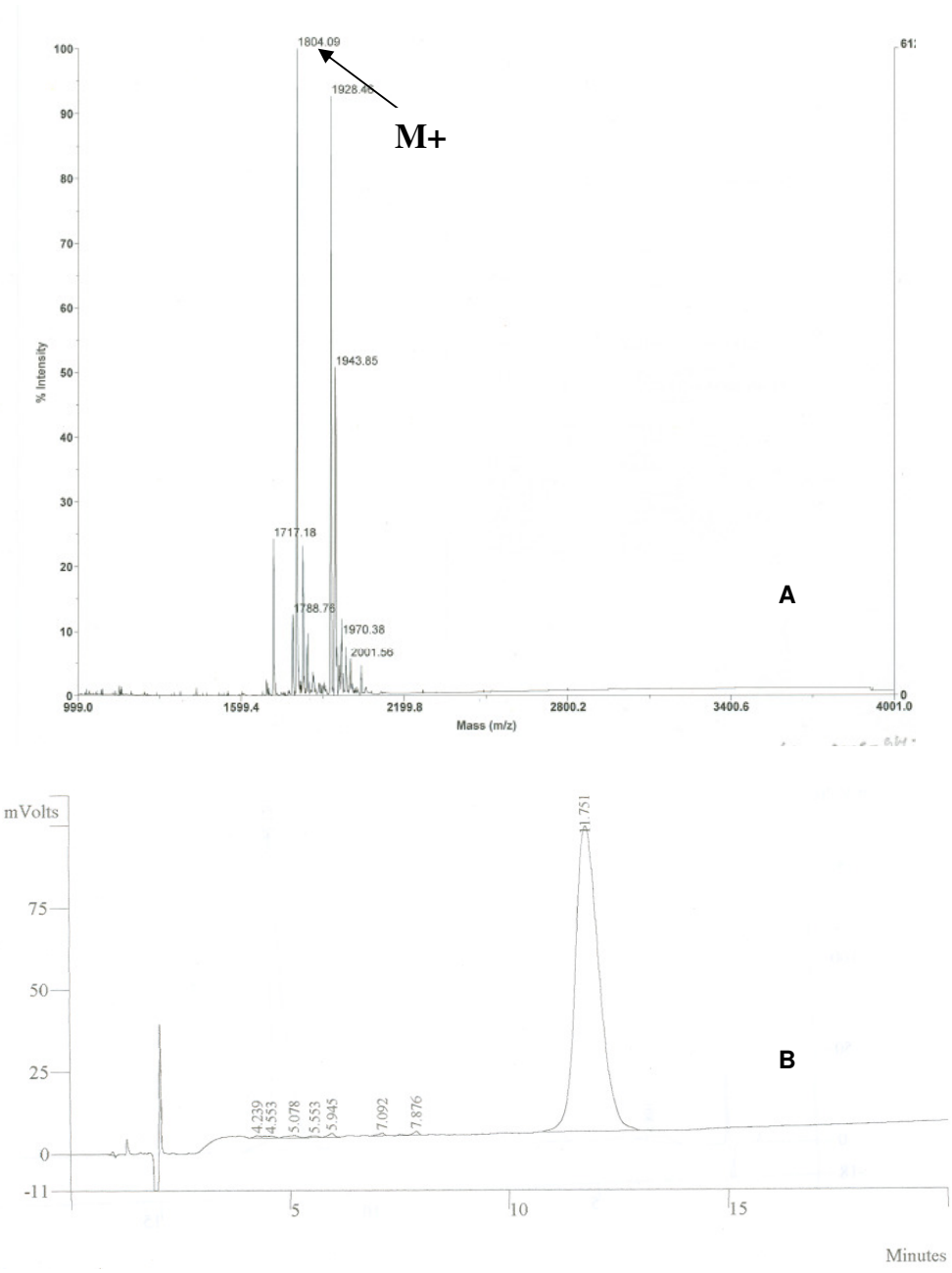




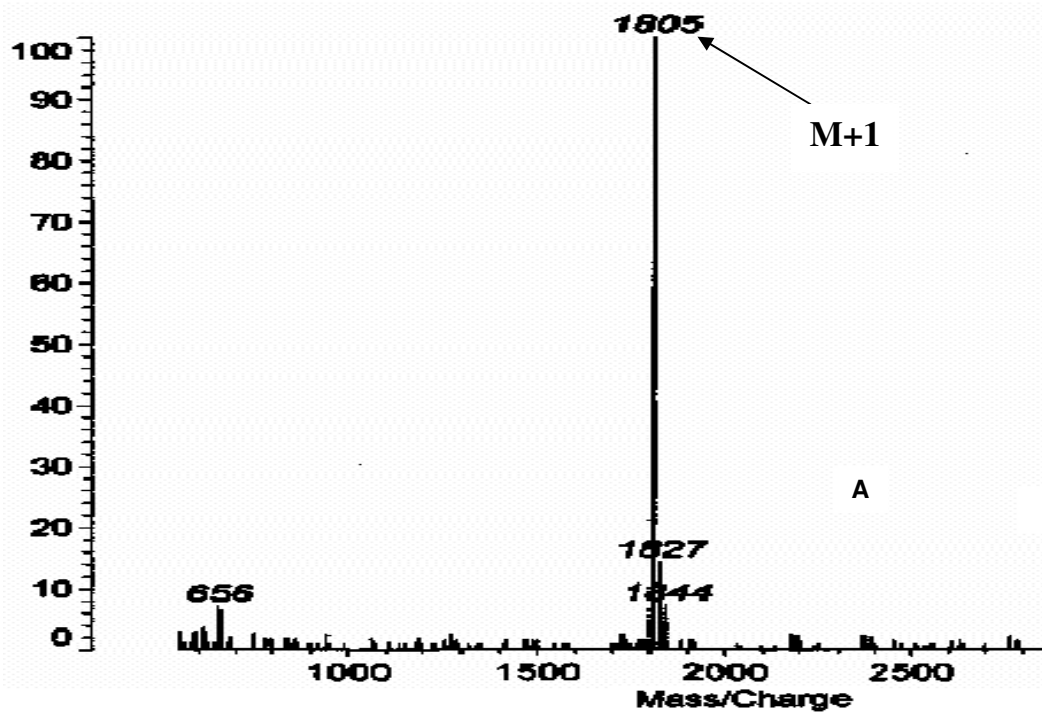
A; MALDI-TOF mass spectrum, **B**; RP-18 HPLC profile of **Peptide-19**. Ac-Phe(Pro-Amp-Gly)$_6$-NH_2. Calculated mass 1804 (observed 1804.97)

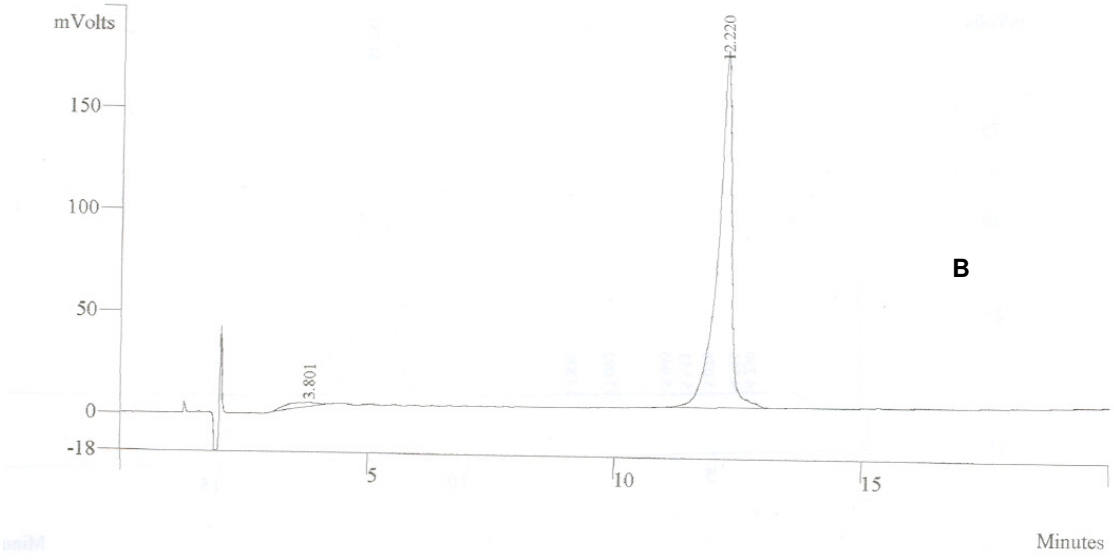


A; MALDI-TOF mass spectrum, **B**; RP-18 HPLC profile of **Peptide-20**. Ac-Phe(Amp-Pro-Gly)$_6$-NH_2. Calculated mass 1804 (observed 1806)

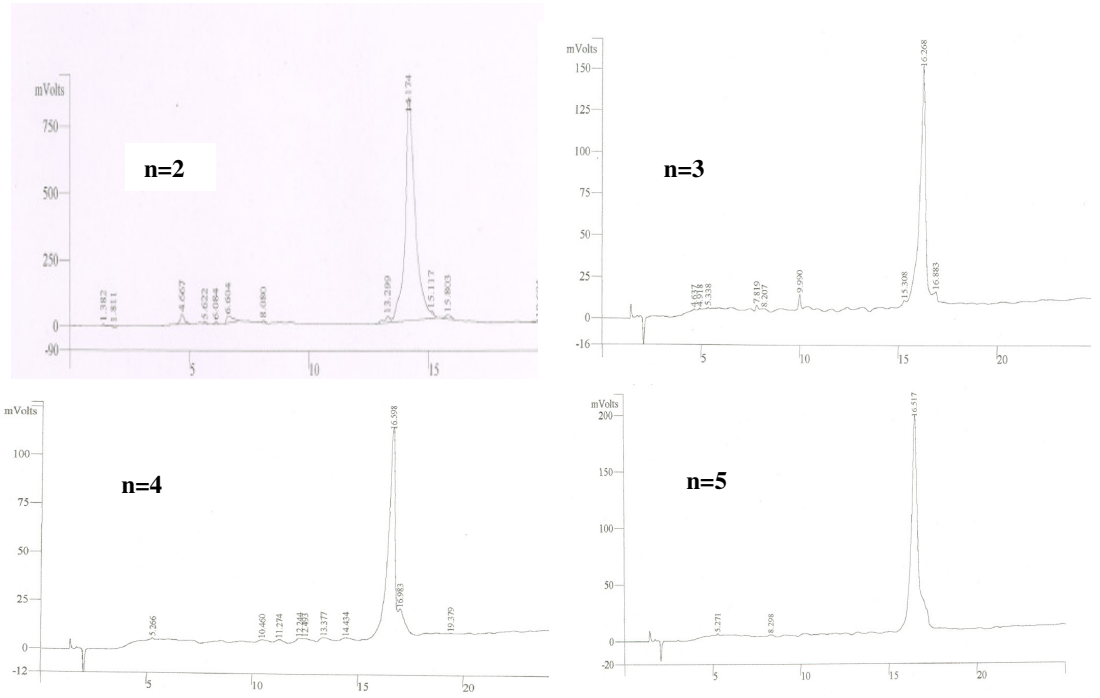


A; MALDI-TOF mass spectrum, **B**; RP-18 HPLC profile of **Peptide-21**. Ac-Phe(Pro-amp-Gly)$_6$-NH_2. Calculated mass 1804 (observed 1804.09)

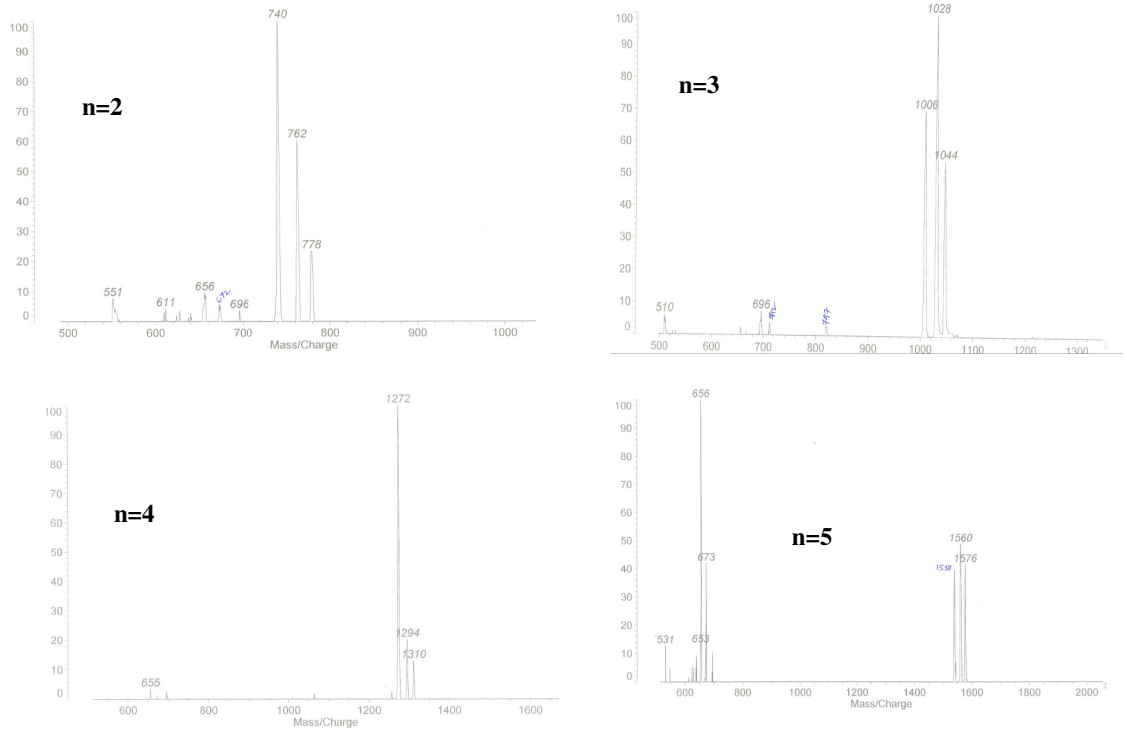




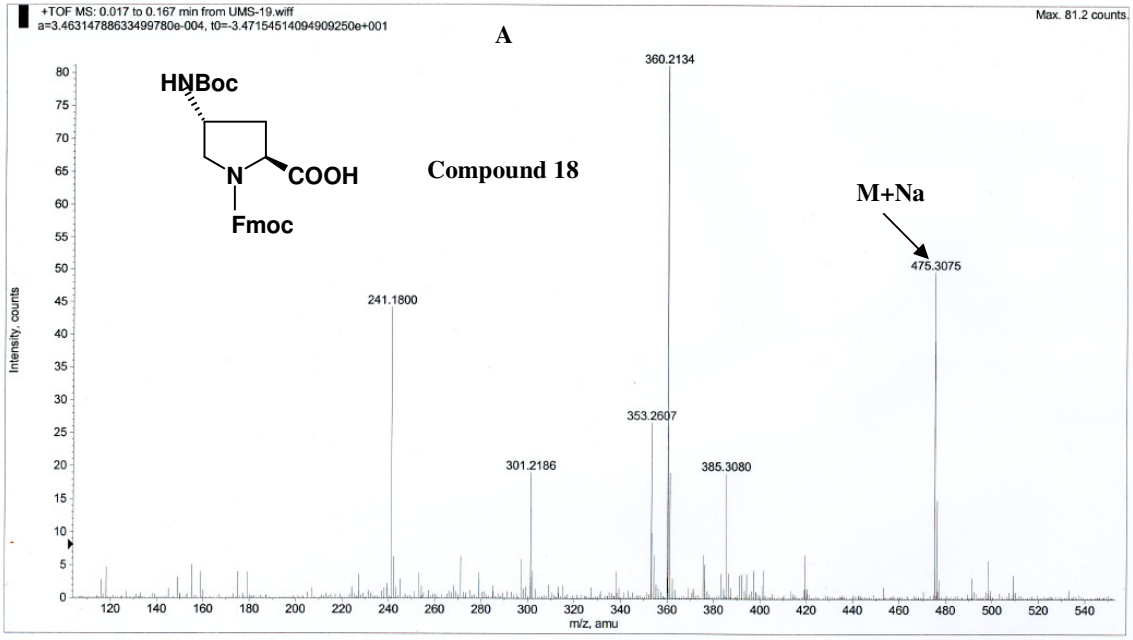
A; MALDI-TOF mass spectrum, **B**; RP-18 HPLC profile of **Peptide-22**. Ac-Phe(amp-Pro-Gly)$_6$-NH_2. Calculated mass 1804 (observed 1805)

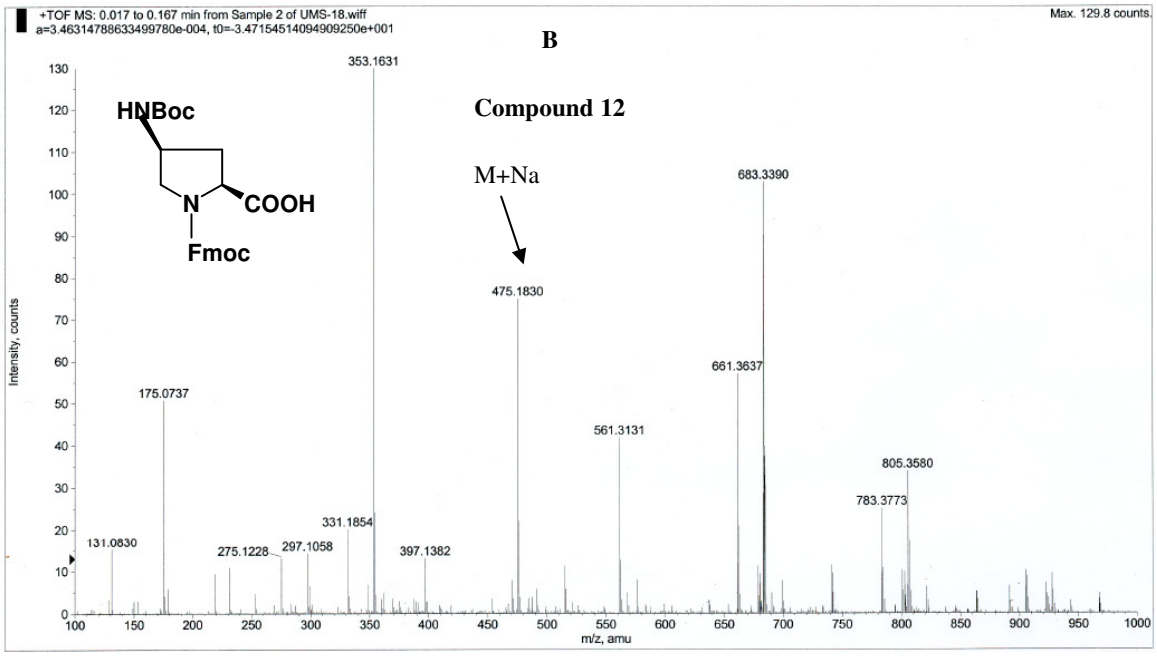


RP HPLC profiles of peptides AcPhe(Pro-Amp-Gly)$_n$-NH$_2$; (n=2-5)



MALDI-TOF mass spectra of peptides AcPhe(Pro-Amp-Gly)$_n$-NH$_2$; (n=2-5);





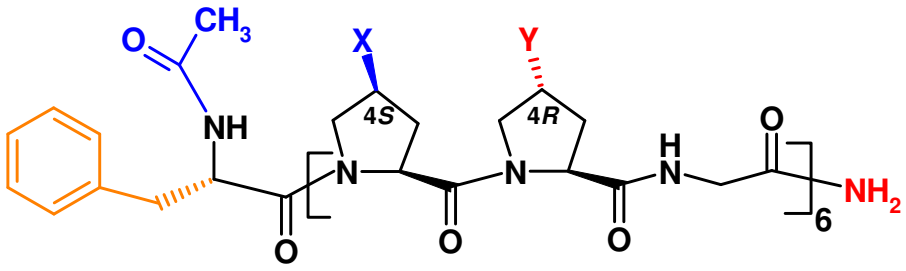
LC-MS mass spectra of **A**; compound **18** and **B**; compound **12**.

Chapter 3

TWO AMINOPROLINES WITH A DIFFERENCE:

Part A

Stereo electronic effects of 4-*R*/*S*-aminoproline on stability in collagen peptides [Pro(X)-Pro(Y)-Gly]$_n$.



3.A.1: Introduction

Single crystal X-ray diffraction structures of collagen model peptides (Pro. Pro. Gly)$_{10}$ and (Pro. Hyp. Gly)$_{10}$ have revealed unique conformational features of Pro and Hyp[1,2].

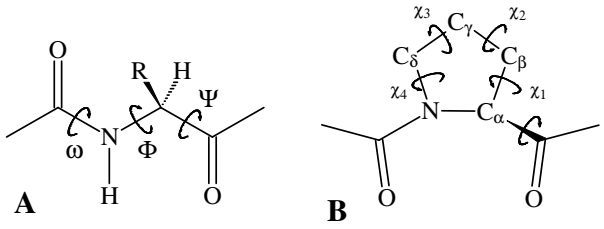


Figure 3.1: A Main chain torsion angles of amino acid residues in polypeptides; **B** Endocyclic torsion angles in imino acid proline (Pro).

The features such as Φ, Ψ and ω torsion angles together with idealized values for two currently accepted triple-helical model for collagen, namely Rich & Crick[3] and Okuyama models. A near planar amide bond with an average angle ω of ~180^{o}C has been observed in the crystal structure of collagenous peptides. Other two main chain dihedral angles Φ and Ψ vary depending on the residue and position of the residue. Proline can be accommodated in both X [Φ(X) = -71.1, Ψ (X) = 163. 8 in (Pro. Pro. Gly)$_{10}$ and Φ(X) = -72.7, Ψ (X) = -161.1 in (Pro. Hyp. Gly)$_{10}$] and Y [Φ(Y) = -62.4, Ψ (Y) = 154] positions of the triple-helix and the angles this residue adopts confirm closely to the idealized values of Okuyama model [(Φ(X) = -75.5, Ψ (X) = 152: Φ(Y) = -62.6, Ψ (Y) = 147.2). Hyp stabilizes collagen in the Y position and the main-chain torsional angles of this imino acid [Φ (Hyp) = -58.4; Ψ (Hyp) = 152.0] in the crystal structure of (Pro.Hyp.Gly)$_{10}$ are close to the idealized values of Okuyama mode (Table 1)[1]. Because of these differences in the torsion angels at each position, the preferred ring puckers are γ- *endo* for the X-position and γ-*exo* for the Y- position of the triple-helix.

Table 1. Triple-helical parameters from the X-ray structures of $(Pro.Pro.Gly)_{10}$[1c], $(Pro.Hyp.Gly)_{10}$[1d], together with the parameters for Okuyama (7/2)[1a-d] and Rick & Crick (10/3)[3] models.

	in $(Pro.Pro.Gly)_{10}$		In $(Pro.Hyp.Gly)_{10}$		7/2 model	10/3 model
Φ *Pro*	-77.1		-72.7		-75.5	-72.1
Ψ *Pro*	163.8	γ-*endo*‡	-161.1	γ-*endo*‡	152.0	164.3
ω *Pro*	179.1		179.6		-176.8	180.0
Φ *Y*†	-62.4		-58.4		-62.6	-75.0
Ψ *Y*	154.0	γ-*exo*‡	-152.0	γ-*exo*‡	147.2	155.8
ω *Y*	177.6		178.5		-172.8	180.0
Φ *Gly*	-76.7		-74.8		-70.2	-67.6
Ψ *Gly*	176.6		172.8		175.4	151.4
ω *Gly*	179.0		179.2		178.2	180.0

† Y residue is Pro in $(Pro.Pro.Gly)_{10}$ and Hyp in $(Pro.Hyp.Gly)_{10}$.
‡ ring-pucker preference of the respective residue

The extended nature of individual helices (polyproline II helix) in the triple-helix necessitates the *trans* (*Z*) geometry for the peptide bond. In most amino acids with the exception of proline, the *Z* conformation is greatly preferred over *E* (*cis*) by a ratio of 1000:1. In the imino acid proline, the steric repulsion with δ-mehylene decreases the energy difference between the two isomers. Consequently in Pro, the *Z* form is only slightly preferred over *E* form, generally by a ratio of 4:1[4]. The preference of Pro simultaneously in both X and Y positions decreases the triple-helical stability while Hyp in Y position stabilizes the collagen triple-helix[5].

More recently, a study of peptidyl-prolyl bond isomerization by Raines *et al*[6] has shown that the Z to E ratio ($K_{Z/E}$) increases with the increase in the electronegativity of

the 4-substituents. Hyp residue stabilizes collagen triple-helix by virtue of its inductive effect and not by hydrogen bonding. Also (2*S*,4*S*)-hydroxyproline (hyp) containing peptides $(hyp.Pro.Gly)_{10}$ and $(Pro.Hyp.Gly)_{10}$ don't form triple-helices[7] due to inadequate peptide bond conformation induced by this residue.

Table 2. Table of $K_{Z/E}^{25°C}$ values for 4-substituted imino acid residues and the triple helical thermal stabilities (T_m values) of peptide with these residues.

Imino acid residue	† $K_{Z/E}^{25°C}$	PeptideSequence	‡Triple-helical T_m (°C)
Pro	4.6	Phe.$(Pro.Pro.Gly)_6$	-nt-#
		$(Pro.Pro.Gly)_7$	6-7
Hyp	6.8	Phe.$(Pro.\textit{Hyp}.Gly)_6$	23
		$(Pro.\textit{Hyp}.Gly)_7$	36
		$(\textit{Hyp}.Pro.Gly)_{10}$	-nt-
Flp	7.0	$(Pro.\textit{Flp}.Gly)_7$	45
Amp	3.7	Phe.$(Pro.\textit{Amp}.Gly)_6$	42.5
hyp	2.4	$(Pro.\textit{hyp}.Gly)_{10}$	-nt-
		$(\textit{hyp}.Pro.Gly)_{10}$	-nt-
flp	2.5	$(Pro.flp.Gly)_7$	<2

† $K_{Z/E}^{25°C}$ values as measure in Ac-Xaa-OMe model compounds of Pro, Flp, flp and hyp from referenc 9; Hyp, Amp from *Chapter 9*.

‡ T_m values Phe.$(Pro.Pro.Gly)_6$, Phe.$(Pro.Hyp.Gly)_6$ and Phe.$(Pro.Amp.Gly)_6$ (at pH 3.0) from ref 9; $(Hyp.Pro.Gly)_{10}$, $(Pro.hyp.Gly)_{10}$ and $(hyp.Pro.Gly)_{10}$ (in 10% aq. acetic acid) from references 10 and 7; $(Pro.Pro.Gly)_7$ $(Pro.Flp.Gly)_7$ and $(Pro.Flp.Gly)_7$ (in 50 mM aq. acetic acid) from reference 8. In all the cases the T_m data is obtained at acidic pH.

no melting-transition was observed.

A similar destabilizing effect by (2*S*, 4*S*)-fluoroproline (flp) in the Y position was also observed. Table 2 summarizes the relationship between the $K_{Z/E}$ values of an imino acid and the triple-helical thermal stability of the collagenous peptides with these imino acids. $K_{Z/E}$ values at 25°C and T_m data obtained from literature sources are also given. The recent X–ray crystal structure,[11] and molecular modeling study of collagen model

peptide $[(Pro.Pro.Gly)_{10}]_3$ suggested that differential proline puckering in X and Y positions are sterically necessary for a favorable triple-packing arrangement. This fact was well supported by the results of previous chapter, that the 4*R*-aminoproline which prefers *up* conformation, stabilizes the collagen triple-helix when placed at Y position and destabilizes when placed at X position, according to the puckering requirement for close packing of triplex. The same effect was observed in case of 4S-aminoproline, which prefers down conformation stabilizing the triplex when placed at X position and destabilizing when placed at Y position.

Our results are further reinforced by the studies of many authors including Raines and Zagari, who have tried to explain the effect of substitution of Pro by Hyp or Flp on the stability of the triple-helical structure (Table 3). Using the result of X-ray analysis on various proline derivatives and collagen model peptides, it has been pointed out that gauche effect (a stereoelectronic effect which is induced by an electronegative element[12]) fixes the pyrrolidine ring puckering, such that Hyp and Flp prefer the *up* form, hyp and flp prefer the *down* form and Pro takes both forms.

Table 3: Stability of collagen triple helix with derived from positional dependence substitution of modified prolines and

peptides	Structure	T_m °C	reference
$(Pro\text{-}Pro\text{-}Gly)_{10}$	t	34	12
$(Pro\text{-}Hyp\text{-}Gly)_{10}$	t	61	5
$(Pro\text{-}hyp\text{-}Gly)_{10}$	s	<4	7
$(Hyp\text{-}Pro\text{-}Gly)_{10}$	s	<4	10
$(hyp\text{-}Pro\text{-}Gly)_{10}$	s	<4	7
$(Pro\text{-}Flp\text{-}Gly)_{10}$	t	91	13a
$(Flp\text{-}Pro\text{-}Gly)_{10}$	s	<4	14
$(Pro\text{-}flp\text{-}Gly)_{10}$	s	<2	13c
$(flp\text{-}Pro\text{-}Gly)_{10}$	s	58	14

t; triple-helical conformation.
s; single strand conformation.

As a general rule of the conformation of the pyrrolidine ring desirable for stabilizing the triple helix, Zagari and co-workers proposed that the puckering should be *down* at the X-position and *up* at the Y-position. However when $(hyp\text{-}Pro\text{-}Gly)_{10}$ did not follow this rule, they regarded this as an exception due to the steric interference between the hydroxyl group of hyp and the pyrrolidine ring of Pro in an adjacent chain[15] (Figure 3.2). The observation that the peptide $(flp\text{-}Pro\text{-}Gly)_{10}$ forms a stable triple-helix, contradicts these explanations and favors the earlier general rule of the conformation of the pyrrolidine ring proposed by Zagari and co-workers.

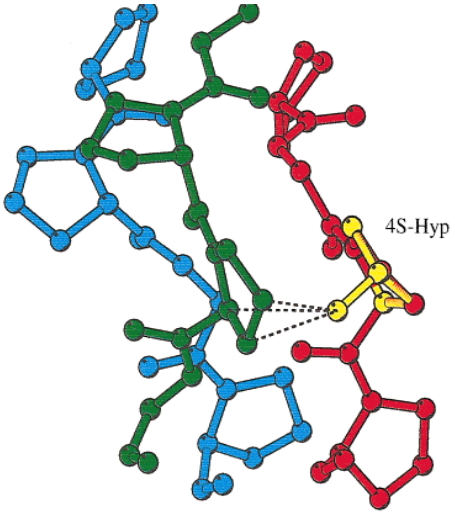


Figure 3.2: Stereoview showing the steric hindrance produced by the hydroxyl group of 4S-Hyp (in yellow) in X position of PPG structure. Interchain contacts at a distance lower than 2.9 Å are also shown.

3.A.1.1: Pyrrolidine ring stability by stereochemistry of 4-substituents.

As explained in Chapter **1**, the pyrrolidine ring conformation are mainly stabilized by the stereochemistry of the 4-substituent via gauche interaction between the 4-substituent and nitrogen atom of the pyrrolidine ring. For example, 4R-aminproline and 4S-aminoproline, are just diasteromers having same substituent at C4-position, with opposite stereochemistry. There are two minimal energy conformational structures

namely *anti* and *gauche* conformations possible for both isomers. In the 4*R*-substituted proline, the 4*R*-substituent may be expected to prefer pseudo-equatorial position (i.e. anti with respect to ring amide nitrogen) to minimize the steric repulsion between ring amide nitrogen and the 4-substituent (Figure 3.3A).

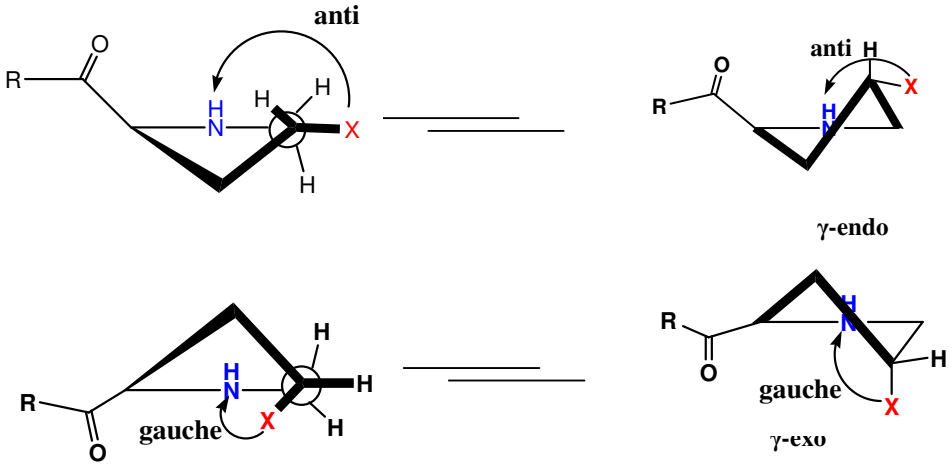


Figure 3.3A: Newman and saw-horse projections depicting the gauche effect and the resulting puckering preference in the 4*R*-substituted prolines.

However X-ray crystal structure and ^{1}H-^{1}H coupling constant analysis confirms the pseudo-axial position of 4*R*-substituted prolines resulting in γ-exo ringpucker.[16,17] This is due to the fact that in pseudo-axial position the 4*R*-substituent and ring nitrogen are stabilized by gauche interaction. In case of pseudo-equatorial position they, are anti to each other.

Similarly, in the 4*S*-substituted proline, the 4*S*-substituent may expected to prefer pseudo-axial position (i.e. anti with respect to ring amide nitrogen) to minimize the steric repulsion between the ring amide nitrogen and the 4-substituent (Figure 3.3B). However the gauche effect leads to a pseudo-axial preference for the 4*S*-substituted prolines resulting in γ-endo ringpucker. In this geometry, the 4S-substituent and the ring nitrogen are stabilized by gauche interaction. This suggest that stereochemistry of 4-

substituent in proline also has remarkable role in determining the ring-pucker along with electronegativity of the 4-substituent.

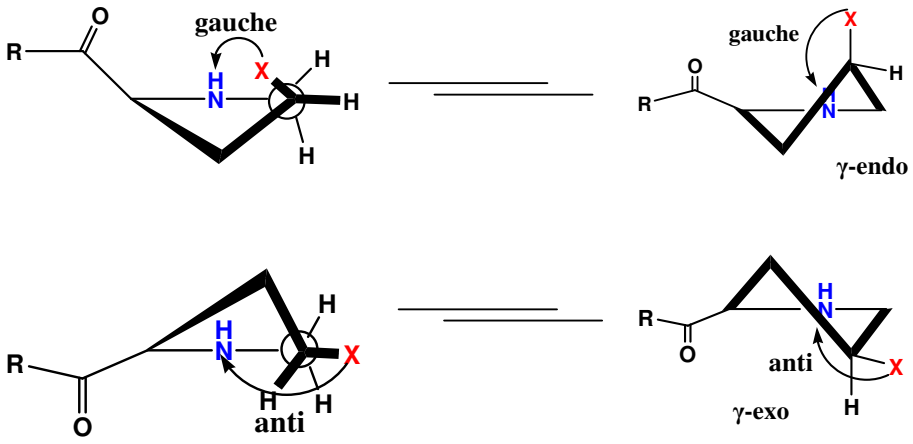


Figure 3.3B: Newman and saw-horse projections depicting the gauche effect and the resulting puckering preference in the 4*S*-substituted prolines.

3.A.2: Aim and rationale of the present work

In Chapter **2**, the study of triple helix forming abilities of the novel amino acids (2*S*,4*R*) and (2*S*,4*S*) aminoprolines in collagen mimetic peptides Ac-Phe(Pro-Amp-Gly)$_6$-NH$_2$, Ac-Phe(Amp-Pro-Gly)$_6$-NH$_2$, Ac-Phe(Pro-amp-Gly)$_6$-NH$_2$ and Ac-Phe(amp-Pro-Gly)$_6$-NH$_2$ were described. These peptides show highly positional preferences to form a more stable triple-helix. The 4*R*-aminoproline at theY position forms a more stable triple helix than other peptides, while 4*S*-aminoproline at X position forms a triple-helix, but not in Y position. Table 4 summarizes the positional preferences of various 4-substituted imino acids and their ability to stabilize collagen triple-helix.

Table 4: Conformational properties of 4-substituted imino acids and their corresponding triple- helix forming abilities.

Imino acid residue †	Ring-pucker preference	$K_{E/Z}^{25°C}$ #	Forms triple-helix?	
			X position	Y position
Pro	γ-*exo* (weak)	4.6	Yes	Yes[8]
Hyp	γ-*exo* (strong)	6.8	No	Yes[8]
Amp	-		No	Yes[9]
Amp^+ ‡	γ-*exo* (weak)	3.7	Yes	Yes[9]
Flp	γ-*exo* (strong)	7.0	No	Yes[8]
hyp	γ-*endo* (strong)	2.4	No	No[8]
amp^+ ‡	γ-*endo* (weak)	6.4	Yes	No[9]
flp	γ-*endo* (strong)	2.5	Yes	No[8]

† Hyp, Amp, Flp are 4*R*-OH, 4*R*-NH_2, 4*R*-F substituted prolines respectively, while hyp, amp and flp are 4*S*-substituent in the same order. In all cases, stereochemistry at C-*2* center is "*S*".

‡ Amp^+ and amp^+ (4*R*-NH_3^+ and 4*S*-NH_3^+) are protonated forms of Amp and amp respectively.

$K_{E/Z}^{25°C}$ values for the imino acids are as observed in their Ac-Xaa-OMe model compounds, where Xaa is the imino acid.

Proline is known to be accommodated in both X and Y positions of collagen triple-helix. This special ability of proline arises from the equal preference of both γ-*endo* and γ-*exo* conformations that are required in X and Y positions respectively. Such flexibility aids in the adaptation of main-chain dihedral angles Φ and Ψ according to the requirement of X and Y positions. In contrast, Hyp and Amp with highly favored γ-*exo* ring pucker can fit only in Y position of the collagen triple-helix. Interestingly amp residue in protonated form has higher preference for γ-*endo* ring pucker and is suitable for X position of the collagen triple-helix. Keeping the individual positional preferences of 4*R* and 4*S* aminoplrolines to stabilize the collagen triple-helix in mind, and in order to further understand the triplex stabilizing effects of 4-aminoprolines, it was thought to

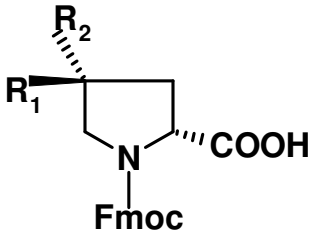
synthesise and study the positionally chimeric collagen peptides having different combination of 4*R*-Amp/Hyp and 4*S*-amp/hyp in Y and X positions of collagen triple-helix.

The objectives of this chapter are

1) Synthesis of (2*S*, 4*S*)-N^1-Fmoc-4-Hydroxyproline monomer (**11**).
2) Solid phase synthesis of collagen chimeric peptides using Amp/Hyp, and amp/hyp at Y and X position respectively.
3) Cleavage of peptides from the solid support, purification, and characterization of these peptides.
4) Triple-helix forming studies using temperature dependent CD spectrophotometry.

3.A.3: Results

Synthesis of (2*S*, 4*R*) (**13**) and (2*S*, 4*S*) (**14**) aminoproline monomers utilized for synthesis of chimeric collagen peptides was described in Chapter ***2***. (2*S*, 4*R*)-N^1-Fmoc-4-hydroxyproline (**12**) monomer was purchased form Nova-biochem, and used without protection of 4-OH group. Synthesis of N^1-Fmoc-4*S*-hydroxyproline (**11**) was achieved in six steps starting from commercially available trans-4*R*-hydroxyproline **1.**

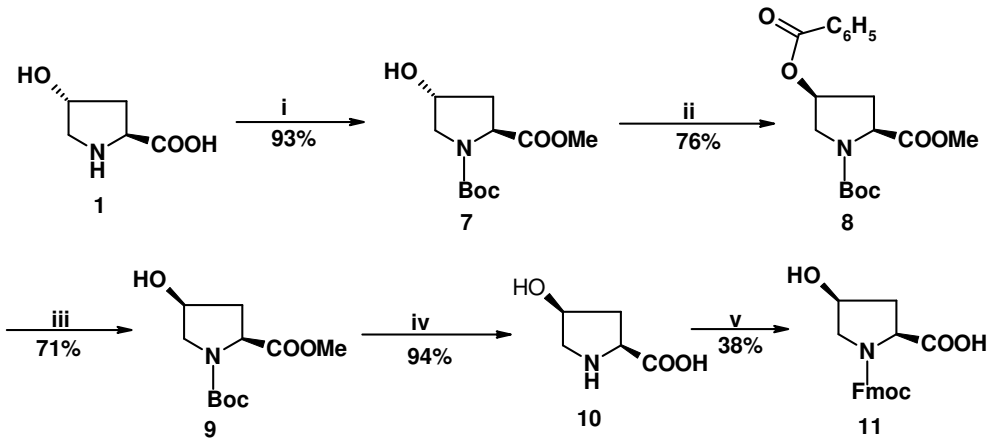


For **11** R_1= OH, R_2= H
For **12** R_1= H, R_2= O H
For **13** R_1= H, R_2= NH_2
For **14** R_1=NH_2, R_2=H

The nitrogen atom of **1** was protected by t-Boc, and the carboxyl group was protected as methyl ester **7** using dimethyl sulphate and K_2CO_3 in anhydrous acetone. The

methylester **7** was then converted to corresponding benzoylester **8** with inversion of stereocenter at C4 using mitsunobu reaction. The benzoylester was hydrolyzed to corresponding alcohol **9** using $NaHCO_3$, $MeOH/H_2O$. The compound **4**-t-Boc was deprotected with 50%TFA in dichloromethane, and the ester was hydrolyzed using 2N LiOH in MeOH to get compound **10**, followed by Fmoc protection using Fmoc-Cl, Na_2CO_3 to get compound **11** (Scheme 1).

SCHEME 1[a]



[a]**Reagents and conditions:** (i) a). NaOH, Boc anhydride, diaxane/water; b). K_2CO_3, dimethyl sulphate, acetone; (ii) DEAD, PPh_3, Benzoic acid, THF; (iii) $NaHCO_3$, $MeOH/H_2O$; (iv) a). 50% TFA in DCM; b). 2N LiOH, $MeOH/H_2O$: (v) Fmoc-Cl, Na_2CO_3, diaxane/water.

3.A.3.1: Solid phase synthesis of chimeric collagen peptides

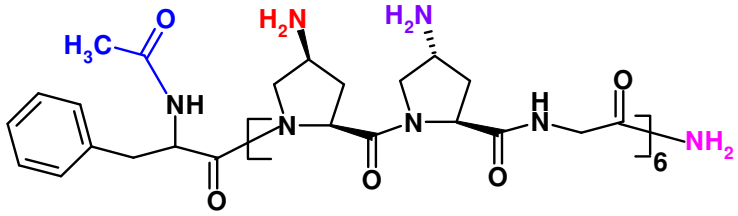
The 4*R*-amino (hydroxyl) prolines were incorporated manually by solid phase synthesis on Rink amide resin into the collagen model peptide Ac-Phe(X-Y-Gly)$_6$-NH_2 using Fmoc strategy. Successful incorporation of Hyp into peptides by solid phase method, without the protection of 4-hydroxyl group using carbodiimide for coupling has been reported[18]. Hence, 4-hydroxyproline containing peptides were synthesized using diisopropyl carbodiimide as a coupling reagent with HOBt as a catalyst and racemization-suppressent. The resin bound Fmoc group was first deprotected using 20% piperidine/DMF and the monomers were coupled as free acid using an *in situ* activation

procedure with DIPCDI as a coupling reagent and HOBt/DIPEA as catalyst. The deprotection and coupling reactions were monitored using qualitative ninhydrin (Kaiser) test for glycine and Chloranil test for imino acids. The amino acid Phe was included at the N-terminus for all peptides to enable accurate determination of the peptide concentrations by UV absorbance at 259nm (ε = 200$M^{-1}cm^{-1}$). The N, C-ends were capped for all peptides to eliminate the effects arising from charge-charge repulsion in terminally ionized forms, a phenomenon well established in collagen peptides. The various collagen chimeric peptides **23-27** (Fig 3.4) were synthesized by incorporating Amp, Hyp, amp and hyp at different positions and combinations. The peptides were cleaved from the resin using 50% TFA in DCM, with 1% TIS (Triisopropylsilane), followed by RP-C4 HPLC purification with water-acetonitrile gradient using 0.1% TFA as organic modifier, and characterized by mass spectrometry (MALDI-TOF). The purity of final peptides was ascertained by RP-C18 HPLC and the peptides were found be greater than 95% pure. The

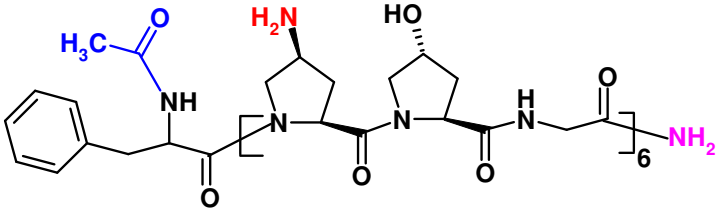
MALDI-TOF mass obtained for peptides agreed closely with calculated values (Table 5).

Table 5: Calculated and observed masses for peptides **23-26.**

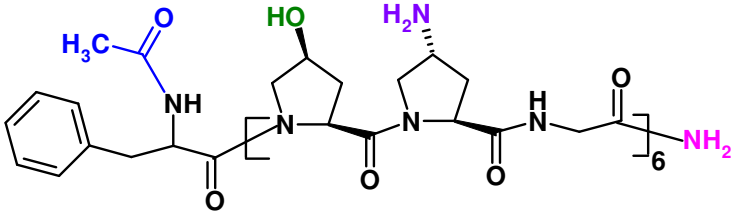
Peptide		Mol. formula	Mass (calc)	Mass (obs)
Ac-Phe(amp-Amp-Gly)$_6$-NH_2	**23**	$C_{83}H_{150}N_{32}O_{20}$	1914.21	1914.36
Ac-Phe(amp-Hyp-Gly)$_6$-NH_2	**24**	$C_{83}H_{144}N_{26}O_{26}$	1920.15	1920.67
Ac-Phe(hyp-Amp-Gly)$_6$-NH_2	**25**	$C_{83}H_{144}N_{26}O_{26}$	1920.15	1921.73
Ac-Phe(hyp-Hyp-Gly)$_6$-NH_2	**26**	$C_{83}H_{136}N_{20}O_{32}$	1926.42	1931.15



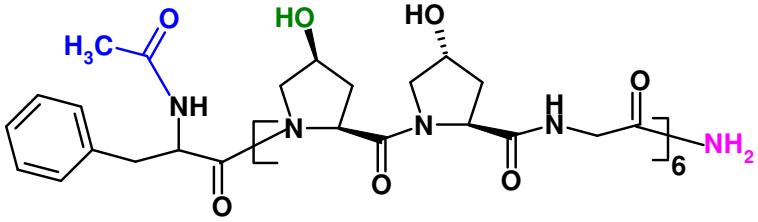
Peptide 23 **(amp-Amp-Gly)**



Peptide 24 **(amp-Hyp-Gly)**



Peptide 25 **(hyp-Amp-Gly)**



Peptide 26 **(hyp-Hyp-Gly)**

Figure 3.4: structure of chimeric peptides used in the present

3.A.4: Biophysical Studies

3.A.4.1: Concentration dependent triple-helix formation of peptides 23-26

As discussed in Chapter **2**, collagen like triple-helical and polyproline II like structures in solution exhibit fingerprint CD spectra that are characteristic by the presence of a large negative band around 200nm, a crossover at around 213nm, and a small positive band around 215-227nm.

In order to evaluate the concentration dependent conformational behavior of the chimeric peptides, CD spectra were recorded in the concentration range of 0.05mM-0.30mM at pH 3 (20mM acetate buffer, 0.1M NaCl). Figure 3.5 shows the CD spectra of chimeric peptides recorded at 25 ^{o}C in the concentration range of 0.05mM-0.30mM. CD spectra of peptides **23** (amp-Amp-Gly) and **24** (amp-Hyp-Gly) in the entire concentration range show identical positive and negative maxima at 223nm and 210nm respectively. Importantly, all the spectral traces pass through an isobestic point at 218nm. The plots of Rpn values derived from these spectra against concentration of the peptides show that Rpn values increase rapidly from 0.05mM through 0.10mM to reach a near saturation at 0.15mM and remain nearly constant thereafter. A critical triple-helical concentration of ~0.15mM is derived from these plots; hence, all further studies are performed at a concentration of 0.2mM. The concentration dependent CD spectra of peptide **25**, (hyp-Amp-Gly) in the entire concentration range show identical positive and negative maxima, at 230nm and 210nm respectively (usual range is 222-223nm). The plot of Rpn values derived form this spectra against concentration of the peptide **25** (hyp-Amp-Gly) was unusual and does not show any saturation like peptides **23** (amp-Amp-Gly) and **24**. (amp-Hyp-Gly). The CD spectra of peptide **26** (hyp-Hyp-Gly) does not show positive band.

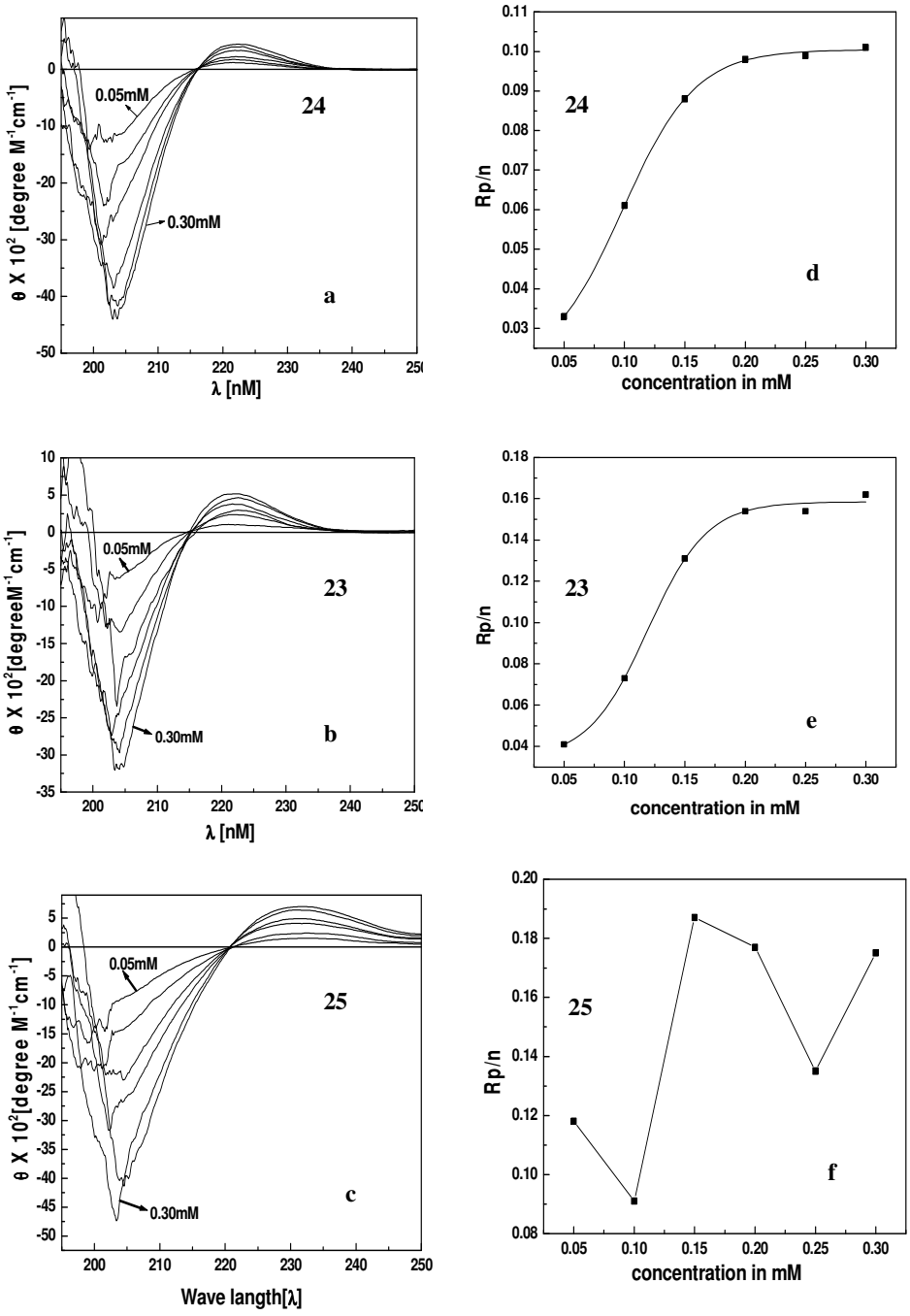


Figure 3.5: Concentration dependent CD spectra of **a** peptide **23**(amp-Amp); **b** peptide **24** (amp-Hyp) and **c** peptide **25** (hyp-Amp); **d**, **e** and **f** are plot of Rpn Vs concentration of corresponding peptides **23**, **24** and **25**

3.A.4.2: Characterization of triple-helical structure by CD spectroscopy

As we observed in the chapter 2, the triple-helix formation in aminoproline containing collagen peptides is dependent on pH because of protonating and nonprotonating ability of free amino group at acidic and basic conditions.

Figure 3.6 shows the CD spectra of 0.2mM solutions of peptides **23** (amp-Amp-Gly) recorded at 10 ^{o}C in acidic (20mM acetate, pH 3.0); neutral (20mM phosphate, pH 7.0) and alkaline (20mM borate, pH 9.0 and 12.0) buffer conditions in presence of 0.1M NaCl. In both alkaline and acidic conditions, the peptide **23** (amp-Amp-Gly) shows the CD spectra characteristic of collagen like triple-helical structure. The Rpn valves at different pH for peptide **23** are listed in Table 6. In all pH conditions, the Rpn values are in the range of triple-helical conformation. The maximum Rpn value is observed at pH 3.0 and minimum at pH 9.0.

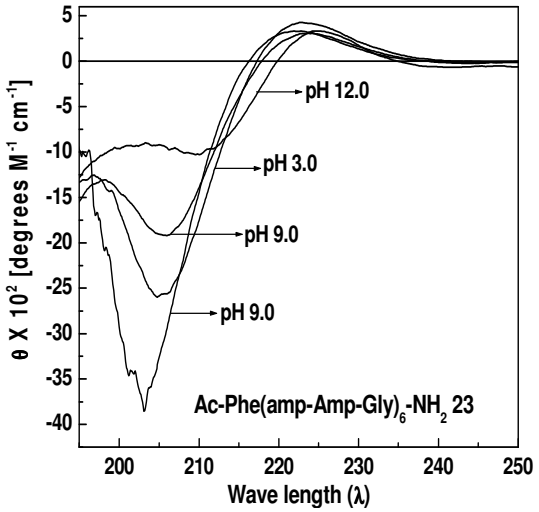


Figure 3.6: CD spectra of peptide **23**Ac-Phe(amp-Amp-Gly)$_6$-NH$_2$ taken at different pH conditions.

Table 6: Rpn values of peptide **23**(amp-Amp-Gly) at different pH conditions.

Peptide 23 Ac-Phe(amp-Amp-Gly)$_6$-NH$_2$			
pH	**+ve band(nm)**	**-ve band (nm)**	**Rpn value**
3.0	224	205	0.37
7.0	224	205	0.31
9.0	224	203	0.16
12.0	224	207	0.27

The appearance of collagen-like CD spectrum is neither a sufficient evidence for the presence of triple-helical structure, nor gives any information about the triple-helical strength. In order to determine the relative triple-helical strength of peptide **23** (amp-Amp-Gly), CD-thermal denaturation studies were carried out in acidic, neutral and basic conditions.

Figure 3.7 shows the thermal denaturation curves of peptide **23** (amp-Amp-Gly) obtained from temperature dependent ellipticity measurement at 225nm.

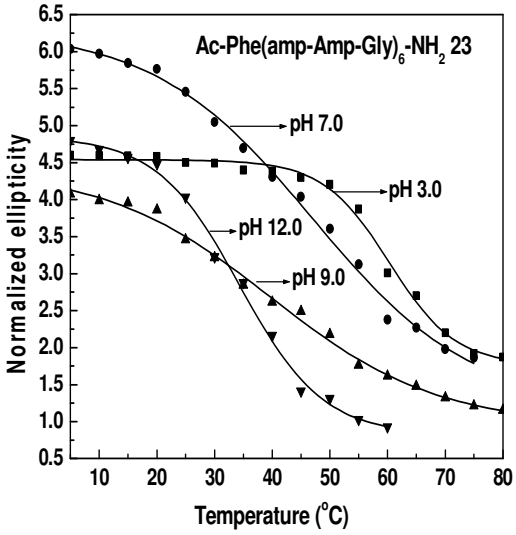


Figure 3.7: Thermal denaturation curves of peptide **23** Ac-Phe(amp-Amp-Gly)$_6$-NH$_2$ at different pH conditions. pH 3.0 (20mM acetate buffer), pH 9.0 (20mM phosphate buffer, pH 9 and 12 (20mM borate buffer). All buffers contain 0.1M NaCl.

Peptide **23** (amp-Amp-Gly) shows sigmoidal transition curves under all pH conditions, indicating that this peptide undergoes triple-helical to coil structure through out. The T_m values obtained form first derivatives of thermal denaturation curves (Table 7) decreases with increasing pH. At pH 3.0, the peptide **23** (amp-Amp-Gly) forms a stable triple-helical structure with T_m of 61 ^{0}C compared to T_m 46.6^{0}C at pH 7.0. At pH 9.0 (T_m 37 ^{0}C) and 12.0 (T_m 34 ^{0}C) the peptide **23** form a less stable triple-helices compared to acidic and neutral conditions. Interestingly at pH 12.0 peptide **23** shows a higher Rpn value (0.27) compared to pH 9.0 (0.16), but forms a less stable triple-helical structure at pH 12.0 compared to pH 9.0.

Table 7: CD-Tm values of peptide **23** (amp-Amp-Gly) at different pH conditions.

Peptide 23 Ac-Phe(amp-Amp-Gly)$_6$-NH$_2$	
pH	**T_m ^{o}C**
3.0	61
7.0	46.6
9.0	40.5
12.0	34

Figure 3.8 shows the CD spectra of 0.2mM solution of peptides **24** (amp-Hyp-Gly) recorded at 10 ^{o}C in acidic (20mM acetate, pH 3.0); neutral (20mM phosphate, pH 7.0) and alkaline (20mM borate, pH 9.0 and 12.0) buffer conditions in presence of 0.1M NaCl.

Table 8: Rpn values of peptide **24**(amp-Hyp-Gly) at different pH conditions.

Peptide 24 Ac-Phe(amp-Hyp-Gly)$_6$-NH$_2$			
pH	**+ve band(nm)**	**-ve band (nm)**	**Rpn value**
3.0	223	204	0.31
7.0	223	204	0.26
9.0	223	202	0.11
12.0	226	210	0.07

The peptide **24** (amp-Hyp-Gly) shows CD spectra characteristic of collagen like triple-helical structure at all pH conditions. The Rpn valves at different pH for peptide

24 decreases linearly with increasing pH (Table 8). At pH 3.0, 7.0 and 9.0, the Rpn values are within the range expected for triple-helical conformation.

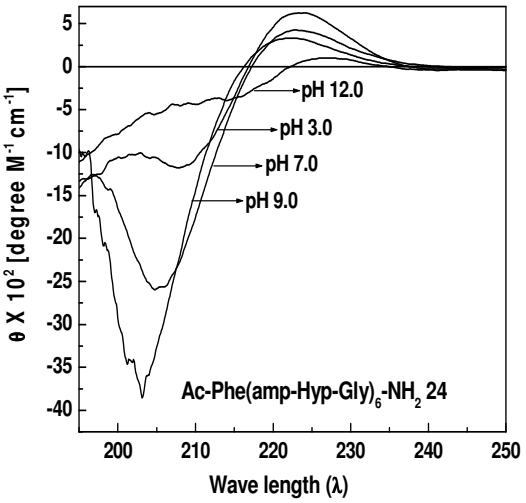


Figure 3.8: CD spectra of peptide **24** Ac-Phe(amp-Hyp-Gly)$_6$-NH$_2$ taken at different pH conditions.

Figure 3.9 shows the normalized ellipticity data at 225nm plotted against temperature for peptide **24** (amp-Hyp-Gly).

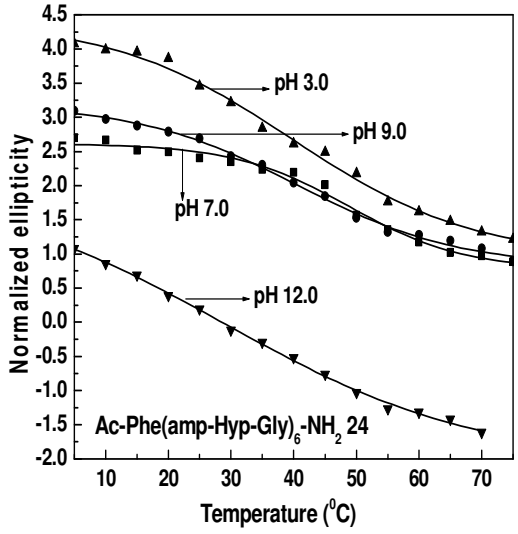


Figure 3.9: Thermal denaturation curves of peptide **24** Ac-Phe(amp-Hyp-Gly)$_6$-NH$_2$ at different pH conditions. pH 3.0 (20mM acetate buffer), pH 9.0 (20mM phosphate buffer, pH 9 and 12 (20mM borate buffer). All buffers contain 0.1M NaCl.

9.0, indicating that this peptide undergoes triple-helical to coil structure under these pH

conditions. At pH 12.0 it shows a linear decrease in ellipticity with temperature. The T_m values obtained form first derivatives of thermal denaturation curves (Table 9) decreases with increasing pH. Maximum T_m was observed at pH 3.0 (49.8^0C) which decreases to T_m 39.3°C at pH 7.0 and T_m 34°C at pH 9.0. No transition was observed at pH 12.0.

Table 9: CD-Tm values of peptide **23** (amp-Amp-Gly) at different pH conditions.

Peptide 24 Ac-Phe(amp-Hyp-Gly)$_6$-NH$_2$	
pH	**T_m °C**
3.0	49.8
7.0	39.3
9.0	34
12.0	-nt-

-nt- no transition

Figure 3.10 shows CD spectra of 0.2mM solutions of peptides **25** (hyp-Amp-Gly) recorded at 10 °C, in acidic (20mM acetate, pH 3.0); neutral (20mM phosphate, pH 7.0) and alkaline (20mM borate, pH 9.0 and 12.0) buffer conditions in presence of 0.1M NaCl. Large variations in the appearance of CD spectra with pH were observed for peptide **25** (hyp-Amp-Gly). At pH 3.0 it shows positive maxima at 233nm and negative minima at 202nm, at pH 7.0 positive maxima at 223nm and negative minima at 210nm. At pH 9.0, peptide **25** shows very small positive maxima at 223nm and negative minima at 205nm. There was no positive band observed at pH 12.0. The Rpn values at different pHs for peptide **25** (hyp-Amp-Gly) are listed in Table 10.

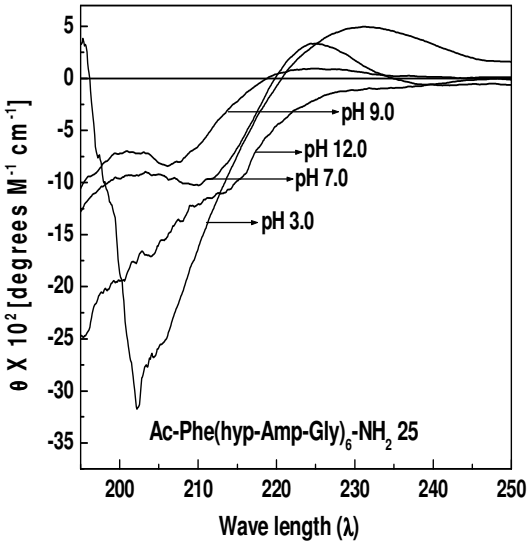


Figure 3.10: CD spectra of peptide **25** Ac-Phe(hyp-Amp-Gly)$_6$-NH_2 taken at different pH conditions.

Table 10: Rpn values of peptide **25** (hyp-Amp-Gly) at different pH conditions.

Peptide 25 Ac-Phe(hyp-Amp-Gly)$_6$-NH_2			
pH	**+ve band(nm)**	**-ve band (nm)**	**Rpn value**
3.0	233	202	0.16
7.0	223	210	0.11
9.0	223	205	0.09
12.0	-	-	

Figure 3.11 shows the thermal denaturation curves of peptide **25** (hyp-Amp-Gly) measure by monitoring the molar ellipticity at 225nm against temperature at acidic (20mM acetate, pH 3.0) neutral (20mM phosphate, pH 7.0) and alkaline (20mM borate, pH 9.0 and pH 12.0) conditions in presence of 0.1M NaCl. Under all pH conditions peptide **25** (hyp-Amp-Gly) displays linear decrease in ellipticity with increasing temperature. This indicates that this peptide does not associate to a triple-helical conformation under any pH conditions. The Replacement of amp to hyp in peptide **23** (amp-Amp-Gly) completely destabilizes the triple-helical conformation though both amp and hyp have same stereocenter at C4 (4*S*).

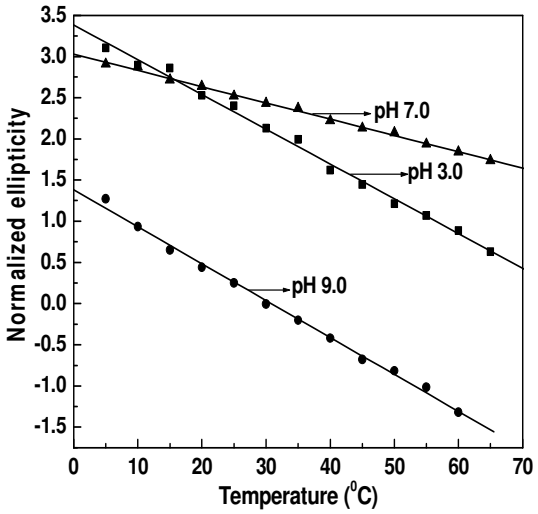


Figure 3.11: Thermal denaturation curves of peptide **25** Ac-Phe(hyp-Amp-Gly)$_6$-NH$_2$ at different pH conditions. pH 3.0 (20mM acetate buffer), pH 9.0 (20mM phosphate buffer, pH 9 and 12 (20mM borate buffer). All buffers contain 0.1M NaCl.

Figure 3.12 shows CD spectra of 0.2mM solution of peptides **26** (hyp-Hyp-Gly) recorded at 10 ^{o}C, in acidic (20mM acetate, pH 3.0); neutral (20mM phosphate, pH 7.0) and alkaline (20mM borate, pH 9.0 and 12.0) buffer conditions in presence of 0.1M NaCl. None of these match the CD spectra of collagen triplex, with the positive band completely absent in both acidic and basic conditions.

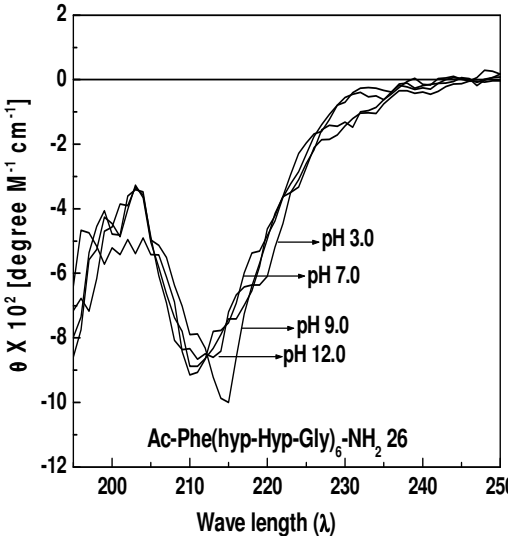


Figure 3.12: CD spectra of peptide **25** Ac-Phe(hyp-Hyp-Gly)$_6$-NH$_2$ taken at different pH conditions.

Interestingly peptide **26** (hyp-Hyp-Gly) displays similar kind of CD spectra at all pH conditions.

It was shown in the previous section that peptide **23** (amp-Amp Gly) and peptide **24** (amp-Hyp-Gly) show collagen-like CD spectra, while the peptide **25** (hyp-Amp-Gly) and peptide **26** (hyp-Hyp-Gly) do not. Peptides **25** (hyp-Amp-Gly), and **26** (hyp-Hyp-Gly) display linear decrease in ellipticity with increasing temperature under all pH conditions. Peptide **23** (amp-Amp-Gly) shows sigmoidal curves under all pH conditions, while peptide **24** (amp-Hyp-Gly) shows sigmoidal curve at acidic (pH 3.0), neutral (pH 7.0) and mild basic (pH 9.0) conditions but not at basic condition (pH 12.0). The T_m data obtained from the first derivative curves of sigmoidal fits of ellipticity data are shown in Table 11.

Table 11: CD-T_m values of peptides **23-26**

Ac-Phe(X--- Y--- Gly)$_6$-NH$_2$		**T_m values (^{o}C)**			
		pH 3.0	**pH 7.0**	**pH 9.0**	**pH 12.0**
Ac-Phe(amp-Amp-Gly)$_6$-NH$_2$	**23**	**61**	**46.6**	**40.5**	**34**
Ac-Phe(amp-Hyp-Gly)$_6$-NH$_2$	**24**	**49.8**	**39.3**	**37**	***ND***
Ac-Phe(hyp-Amp-Gly)$_6$-NH$_2$	**25**	***ND***	***ND***	***ND***	***ND***
Ac-Phe(hyp-Hyp-Gly)$_6$-NH$_2$	**26**	***ND***	***ND***	***ND***	***ND***

ND not detectable

Peptide **23** (amp-Amp-Gly) at pH 3.0 shows a T_m of 61^{0}C which is maximum and decreases through pH 7.0 (T_m = 46.6^{0}C), pH 9.0 (T_m = 40.5^{0}C) and pH 12.0 (T_m = 37^{0}C). A similar behavior is observed for peptide **24** (amp-Hyp-Gly) which shows maximum T_m 49.8^{0}C at pH 3.0, which decreases with increasing pH (pH 7.0, T_m = 39.3^{o}C, pH 9.0, T_m = 37^{0}C) and failed to form triple-helical structure at pH 12.0. There

was no transition observed for peptides **25** (hyp-Amp-Gly) and **26** (hyp-Hyp-Gly) at any of the pH conditions.

3.A.4.3: Effect of ethylene glycol on the stability of peptides 23-26 triple-helices

As explained in Chapter **2**, polyols and sugars are known to offer protection against thermal denaturation of most proteins, including collagen triple-helix. Ethylene glycol (EG) stabilizes the helical structure and therefore can be very useful to amplify and detect very weak triple-helical propensities.

Figure 3.13A shows the CD spectra of 0.2mM solution of the peptides **23-25** taken in a 3:1 mixture of EG:W. In comparison to aqueous solution, the spectrum of peptide **26** (hyp-Hyp) is characterized by the appearance of significant positive band at 225nm and a large decrease in molar ellipticity of the negative band at 200nm. The Rpn values obtained from these spectra are listed in Table 12.

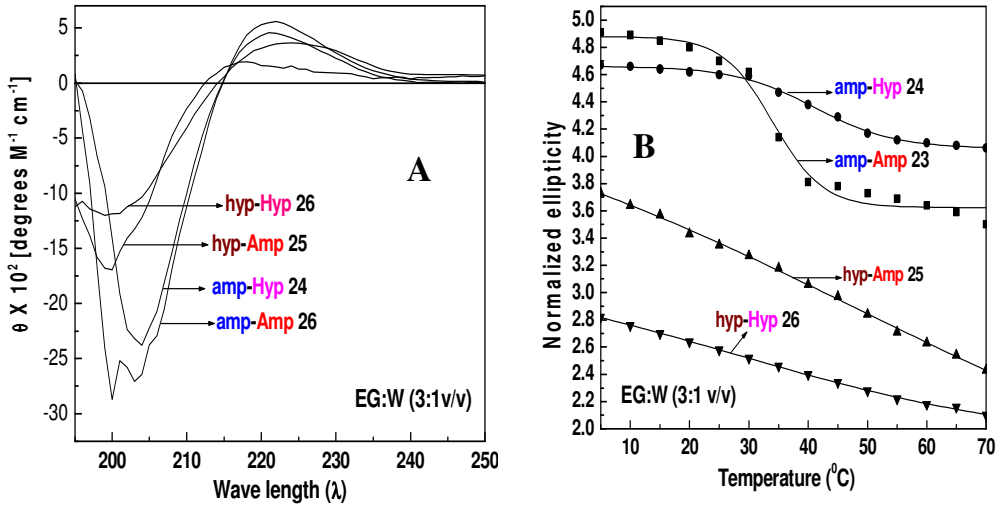


Figure 3.13: **A**, CD spectra at 10°C and **B,** CD-T_m curve of 0.2M solutions of peptides **23-26** in 3:1 (v/v) EG:W mixture

Table 12: Rpn values for peptides **23-26** calculated form the CD spectra measured at 10°C in 3:1 (v/v) EG:W mixture

Ac-Phe(X---Y---Gly)$_6$-NH$_2$		R_{pn} values	T_m ^0C
Ac-Phe(amp-Amp-Gly)$_6$-NH$_2$	**23**	0.18	37
Ac-Phe(amp-Hyp-Gly)$_6$-NH$_2$	**24**	0.20	42
Ac-Phe(hyp-Amp-Gly)$_6$-NH$_2$	**25**	0.17	*ND*
Ac-Phe(hyp-Hyp-Gly)$_6$-NH$_2$	**26**	0.083	*ND*

ND not detectable

Despite the significant appearance of positive and negative bands, peptide **26** (hyp-Hyp-Gly) and peptide **25** (hyp-Amp-Gly) do not associate in to triple-helical structure in EG:W system. The T_m of peptide **23** (amp-Amp) is 37^{0}C, which is lower compared Tm of 42^{0}C for to peptide **24** (amp-Hyp-Gly). This behavior of peptide **23** (amp-Amp-Gly) is opposite to its behavior in aqueous solutions where under all pH conditions it forms a more stable triple-helix than the peptide **24** (amp-Hyp-Gly). Interestingly, the order of magnitude in the Rpn values is matches that of T_m values. Peptide **23** (amp-Amp-Gly) with lower Rpn value shows a lower T_m and a higher Rpn value for peptide **24** (amp-Hyp-Gly) leads to higher T_m.

3.A.4.4: Salt effects on triple-helix stability of chimeric peptides 23-24.

In order to evaluate and elucidate the mechanism of triple-helix stabilization of chimeric peptides **23** (amp-Amp-Gly) and **24** (amp-Hyp-Gly), T_m measurements of both peptides were carried out at pH 7.0 with varying concentrations of NaCl. T_m data thus obtained is shown in Table 13.

Table 13: T_m values of peptides **23** and **24** at pH 7.0 (20mM phosphate buffer) with varying NaCl concentrations

Conc. of NaCl	T_m values (^{0}C)	
	23	**24**
No salt	20	40.5
50mM	43	43.0
100mM	61	49.8
150mM	63	48.2
200mM	53	39.4

Figure 3.14 shows the plot of triple-helical stability of chimeric peptides **23** (amp-Amp-Gly) and **24** (amp-Hyp-Gly) against concentration of NaCl. The T_m of peptide **23** shows a steady increase up to 150mM NaCl and decreases at higher salt concentration. In comparision to T_m of peptide **24,** under similar conditions uniformly raises up to 100mM shows NaCl and a slight decrease as salt concentration is increased.

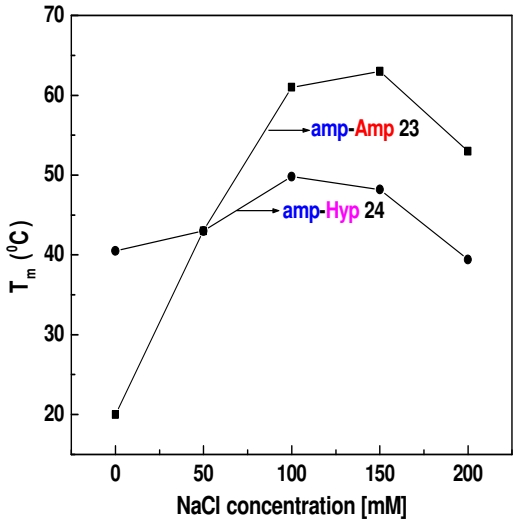


Figure 3.14: Plot of triple-helical stability of peptides **23** and **24** against concentration of NaCl.

3.A.5: Discussion

The chimeric peptides **23** (amp-Amp-Gly) and **24** (amp-Hyp-Gly) at pH 3.0 form a polyproline II like structure as indicated by a positive maxima at ~223nm and a strong negative band ~ 205nm in the CD spectra. The peak positions are nearly independent of the concentration for both peptides and most importantly all spectral traces pass through the same isobestic point at 217nm. These CD spectral properties indicate that within the concentration range measured i. e., 0.05mM-0.30mM, peptides **23** and **24** exist as triple-helices. Furthermore, the high Rpn values observed for peptide **23** (amp-Amp-Gly) also corresponds to a high T_m value.

The observed sigmoid transition in the variable temperature CD measurements provides additional evidence for the two state triple-helix to coil transition. T_m values of peptide **23** (amp-Amp-Gly) and **24** (amp-Hyp-Gly) decrease with an increase in pH. The thermal stability of peptide **23** is always greater than that of peptide **24** (amp-Hyp-Gly) over the entire pH range. The replacement of Pro to amp in the peptides Ac-Phe(Pro-Amp-Gly)$_6$-NH_2 **19** and Ac-Phe(Pro-Hyp-Gly)$_6$-NH_2 (control peptide, Chapter **2**) has remarkable effects on triple helical stability. The stability of peptide **24** (amp-Hyp-Gly) is always higher (ΔT_m 20.8^{o}C at pH 3.0, 11.3^{o}C at pH 7.0 and 10.0^{o}C at pH 9.0) compared to that of control peptide (Pro-Hyp-Gly)[9] under identical conditions. A similar trend is observed for the peptide **23** (amp-Amp-Gly) with amp at X and Amp at Y position, which forms a more stable triple helix compared to peptide **19** (Pro-Amp-Gly) with pro at X and Amp at Y position over the pH range 3.0-9.0 but form a less stable triple-helix at pH 12.0. This is due to the preferred puckering adopted by Hyp/Amp and amp at Y and X positions in accordance with the two prolines in natural collagen (Pro-Pro-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 change of ring pucker in going from acidic to basic pH for aminoproline may not be compatible for peptide **23** (amp-Amp-Gly) which leads to a decrease in thermal stability of triple-helix at pH 12.0. Peptides **25** (hyp-Amp-Gly) and **26** (hyp-Hyp-Gly) with hyp at X position and Amp and Hyp at Y position respectively, do not associate into the triple helix structure. This is evidenced by the appearance of unusual CD spectrum which is not like the CD spectrum of collagen triple helix, and showing a linear decrease in ellipticity with increasing temperature. This indicates that amp at X position has a positive stabilizing effect with Pro, Amp and Hyp in Y position, whereas hyp at X position destabilizes the triple-helix.

The presence of NaCl upto a concentration of 200mM has dramatic effects on the T_m values of both peptides **23** (amp-Amp-Gly) and **24** (amp-Hyp-Gly) triple-helices at pH 7.0. Higher concentration of salt stabilizes the triple-helical structure evidenced by very higher T_m values for peptide **23** (63^0C) at 150mM NaCl and 20^0C compared to T_m in the absence of salt concentration. This stabilization by NaCl arises from the screening of interstrand and intrastrand electrostatic repulsion caused by the protonated amino groups of 4-aminoprolines. As the pKa values are 9.3 for monomer amp and 10.5 for monomer Amp, the side chain 4-aminogroups in the peptide **23** (amp-Amp-Gly) would largely be protonated at pH 7.0. Collagen being a parallel triple-helix with one residue shift, these $-NH_3^+$ groups lie in close proximity to each other which results in large electrostatic repulsive interactions between the chains. The presence of salt counter-ions offers a shielding effect resulting in decreased electrostatic repulsion between the chains. In peptide **24** (amp-Hyp-Gly), where only interstrand electrostatic repulsion caused by protonated amino acid at X position is present, the stabilization by

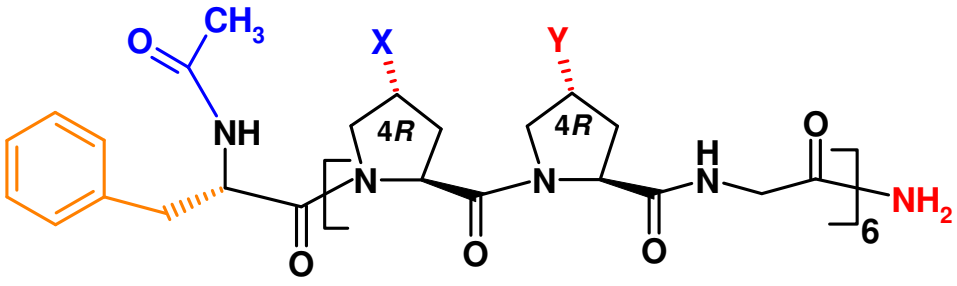
salt arises solely from the shielding of the interstrand repulsion between the chains. The amp units on adjacent strands are moved quite far due to effective repulsions. Beyond 200mM NaCl breaking of the hydrogen bond network, water-structure and poor solubility of peptides leads to decrease in the triple-helical stability and thus decrease in T_m values for a both peptides **23** (amp-Amp-Gly) and **24** (amp-Hyp-Gly).

In aqueous buffer conditions, the peptide **23** (amp-Amp-Gly) appears to offer higher stabilization under all pH. However, in 3:1 v/v mixture of EG:W, peptide **23** (amp-Amp-Gly) shows lower T_m than the peptide **24** (amp-Hyp-Gly). This reversal of stability in EG:W arises from the poor solvation of the EG for the charged and polar amino residue present at both X and Y positions in peptide **23** (amp-Amp-Gly). The additional interstrand repulsion arising from the charge on 4-NH_3^+groups in peptide **23** remains unscreened in less polar EG:W mixture. This repulsive interaction aggravated by the absence of salt may result in decreased thermal stability of triple-helix. Interestingly, despite a large increase in positive band intensity in EG:W mixture, the peptides **25** (hyp-Amp-Gly) and **26** (hyp-Hyp-Gly) still do not show a co-operative transition. This suggests that the triple-helical conformation is more favored for aminoproline containing peptides in aqueous condition compared those in organic solvents.

Part B

Dual compatibility of 4R-aminoproline

Hyperstable collagen peptides $(X\text{-}Y\text{-}Gly)_n$ with modified 4R-iminoacids both at X and Y positions.



3.B.1: Introduction

Since the first proposal of collagen models, the elucidation of collagen structure and the understanding on the molecular basis of its stability have been the subject of several studies, from both experimental and theoretical point of view. The availability of a high-resolution structure for the $(Pro\text{-}Pro\text{-}Gly)_{10}$ peptide[19] has made it possible to highlight the strong correlation between the position of the iminoacid in the chain and the puckering of the pyrrolidine ring. In collagen like polypeptides, prolines in X position adopt a down (more precisely C^{γ}-endo) puckering whereas those in Y position adopt an up (C^{γ}-exo) puckering. X-ray structure of $(Pro\text{-}Hyp\text{-}Gly)_n$ peptide shows the same X(down)-Y(up) (Fig 3.15) alternation.[19] These findings, together with the observation of a strong correlation between the adopted puckering and backbone dihedral angels suggests that the formation of collagen triple-helix requires the presence of down iminoacid in the X position and an up iminoacid in the Y position.

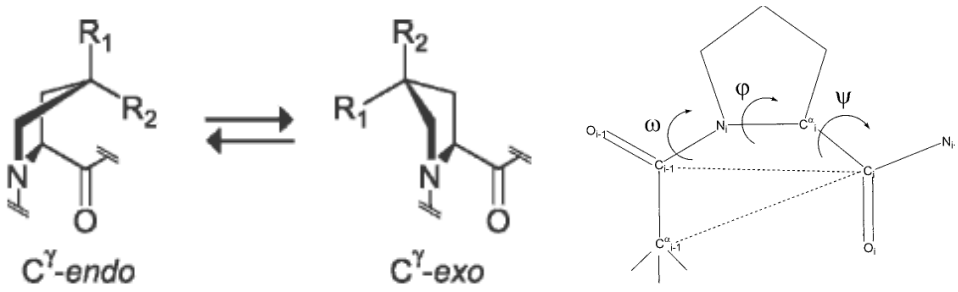


Figure 3.15: Pyrrolidine ring puckering and corresponding bond angles.

Notably these two conformers have values for the φ ($C`_{i-1}$-$N_i$-$C^{\alpha}_i$-$C`_i$) and ψ (N_i-C^{α}_i-$C`_i$-N_{i+1}) dihedral angles that correspond to those preferred by the residue in the X-position (where φ and ψ are -72.6 ± 7.6° and 163.8 ± 8.8° respectively[20]) and Y-position (where φ and ψ are -59.6 ± 7.3° and 149.8 ± 8.8° respectively[20]) of a triple-helix of $(Pro\text{-}Hyp\text{-}Gly)_n$ strands.[21] Hence, according to the principal of preorganization, nonnatural

proline derivatives that prefer a C^{γ}-exo or C^{γ}-endo conformation can produce more stable triple helices.

3.B.1.1: Preorganization of peptide back-bone angles

In the first part of this chapter, devoted to the preorganization of φ and ψ dihedral angles of the backbone by replacing 4*S*-amp at X position and 4*R*-Amp/Hyp at Y position, it was shown that (amp-Amp/Hyp-Gly)$_n$ peptide increase the triple-helix stability compared to reference peptides (Pro-Amp-Gly)$_n$ and (Pro-Hyp-Gly)$_n$. Raines et al[22] have observed that, 4-methylprolines peptide (Fig 3.16) (mep-Mep-Gly)$_7$ (T_m = 36^0C) forms stable triple-helix compared to peptides (mep-Pro-Gly)$_7$ (T_m = 13^0C) and (Pro-Mep-Gly)$_7$(T_m = 29^0C). These results indicate once again that triple-helical stability can be enhanced by preorganizing the backbone dihedral angles by incorporating the diasteromers suitably at X and Y positions of X-Y-Gly sequence.

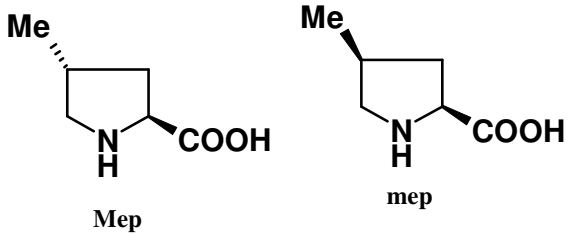


Figure 3.16: Structure of (2*S*,4R)-4-methylproline (Mep) and (2S,4s)-4-methylproline (mep)

3.B.1.2: Reciprocity of steric and steroelectronic effects

The preorganization of φ and ψ dihedral angles does not necessarily promote triple-helix formation. For example replacing Pro with 4*S*-Flp in the X position[23] leads to the increased stability of triple helix of peptide (4*S*-Flp-Pro-Gly)$_7$ whereas the peptide (4*S*-Flp-4*R*-Flp-Gly)$_7$, does not form stable triple-helix structure[24] in contrast to expectations. The CD spectrum of this peptide at 4^0C has weak maximum ellipticity at

226nm and a shallow minimum ellipticity at a longer wavelength. Such a spectrum is typical of an uncomplexed polyproline II helix. The modeling of peptide (4*S*-Flp-4*R*-Flp-Gly)$_n$ suggests that interstrand fluro-fluro interactions are averse to triple-helical stability (Fig 3.17).

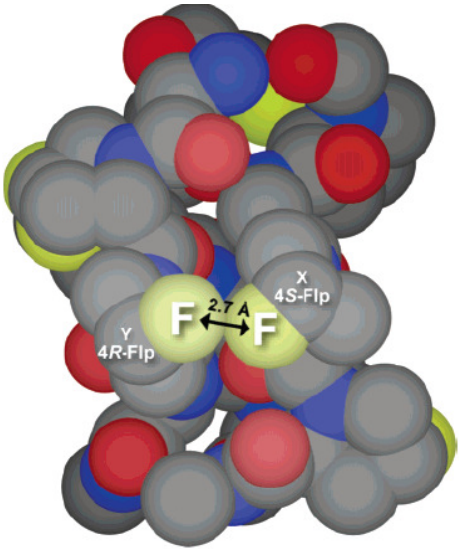


Figure 3.17: Model of a segment of a (4*S*-Flp-4*R*-Flp-Gly)*n* triple helix[24]. The model was constructed from the three-dimensional structure of a (Pro-4*R*-Hyp-Gly)*n* triple helix (PDB entry 1CAG) by replacing the appropriate H and OH on Pro and 4*R*-Hyp, respectively, with F. The interstrand F….F distance of 2.7 Å in one cross section is indicated explicitly.

Because fluoro groups at the 3-position of proline should not suffer from steric conflicts with adjacent strand, 3-fluroproline is an alternative residue to preorganize the main chain for triple helix formation. Since 3*S*-Flp also prefers a Cγ-endo conformation like 4S-Flp, it was found that the peptide (3*S*-Flp-4*R*-Flp-Gly)$_7$ is associated with triple-helical conformation. Its CD spectrum has a large maximum at 225nm at 4^{o}C and upon heating, a co-operative decrease in ellipticity was observed indicating triplex-denaturation. Apparently moving the fluoro group from the 4-position to the 3-position of a proline residue at X position eliminates the interstrand fluoro-fluoro steric interactions that prevented (4*S*-Flp-4*R*-Flp-Gly)$_7$ to fold into triple-helical conformation(Fig. 3.18). Nevertheless, the T_m of (3*S*-Flp-4*R*-Flp-Gly)$_7$ is 34^{o}C (Table-

14) which is substantially lower than that of (Pro-4*R*-Flp-Gly)$_7$ with Tm 45^{o}C.

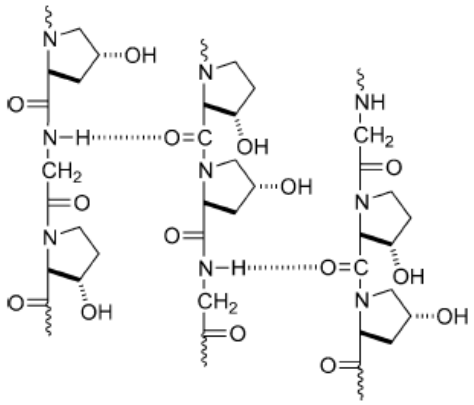


Figure 3.18: Putative interstrand hydrogen bonds in triple-helical[26] (3-Hyp-4-Hyp-Gly)*n*.

Table 14: Thermal stability of various collagen-like chimeric peptides

Peptide	T_m (^{o}C)	Reference
(4*S*-Flp-4*R*-Flp-Gly)$_7$	< 8	24
(3*S*-Flp-4*R*-Flp-Gly)$_7$	34	24
(4*S*-Mep-4*R*-Mep-Gly)$_7$	36	22
(4*S*-Hyp-4*R*-Hyp-Gly)$_{10}$	< 9	29
Ac(4*S*-Flp-4*R*-Hyp-Gly)$_5$-Pro-Cys(StBu)-(Gly)$_3$-NH_2	< 4	28
Ac(4*R*-Flp-4*R*-Hyp-Gly)$_5$-Pro-Cys(StBu)-(Gly)$_3$-NH_2	< 4	28
(Pro-Hyp-Gly)$_3$-3*S*-Hyp-4*R*-Hyp-Gly(Pro-Hyp-Gly)$_3$	32.7	26

With 4*R*-Flp in the Y position, replacement of Pro in the X position with 3*S*-Flp is destabilized even though 3*S*-Flp should preorganize the φ and ψ dihedral angles properly. This may due to the fact that in a triple-helix the main-chain C=O of each residue in the X position forms an interstrand hydrogen bond to the main chain N-H of glycine residue. These interstrand C=O….H-N hydrogen bonds are known to be critical to triple-helix stability.[25] An inductive effect from the proximal fluoro group in 3S-Flp would weaken these hydrogen bonds, as has been noted in triple-helices containing 3S-Hyp in the X position.[26] This effect was less problematic with 4*S*-Flp in the X

position,[27] because the fluoro group is separated from the main-chain C=O by an extra bond in 4*S*-Flp. Apparently, the weakening of these hydrogen bonds outweighs the benefit of preorganization. Moreover the polypeptide with repeating sequence (4*S*-Flp-4*R*-Hyp-Gly)$_{10}$ also does not form a triple-helix.[28,29]

3.B.1.3: Duality of 4*R*-hydroxyproline

Bächinger et al[30-32] have observed that a stable triple-helix can form with 4*R*-Hyp in the X position when the Y position is not proline, eventhough 4*R*-Hyp adopts the C^{γ}-exo conformation in accordance with unfavorable φ and ψ dihedral angles. The peptide Ac(4*R*-Hyp-Y-Gly)$_{10}$-NH_2 in which Y is threonine and valine, forms a very stable triple helix in water (Table 15) and with allo-threonine, alanine and serine forming a triple-helical structure in propanol.

Interestingly, the peptide Ac(4*R*-Hyp-4*R*-Hyp-Gly)$_{10}$-NH_2 which has 4*R*-Hyp in both X and Y positions, also successfully folds into a triple-helical conformation[33,34] in water at 4^{o}C. Given the preceding knowledge of the stability of the collagen triple-helix, the expectation is that the peptide Ac-(4*R*-Hyp-4*R*-Hyp-Gly)$_{10}$-NH_2 should form a triple-helix with a stability in-between that of Ac(Pro-4*R*-Hyp-Gly)$_{10}$-NH_2, which forms a stable triple-helix, and Ac(4*R*-Hyp-Pro-Gly)$_{10}$-NH_2, which does not form a triple-helix. Surprisingly peptide Ac(4*R*-Hyp-4*R*-Hyp-Gly)$_{10}$-NH_2 has a higher transition temperature than Ac(Pro-4*R*-Hyp-Gly)$_{10}$-NH_2.

Table 15: Thermodynamic data of the triple helix coil transition of the Ac-(Gly-4(R)-Hyp-Yaa)-NH_2 peptides in 1,2-propanediol.The thermodynamic parameters are expressed per mol of tripeptide units in a triple helix[31].

Peptide	T_m [a] °C	ΔH^o_{VH} KJ/mol	ΔS^o KJ/mol	T_m [b] °C	ΔH^o_{cal} KJ/mol
Ac-(Gly-4(R)-Hyp-Thr)$_{10}$-NH_2	48.2	-10.6	-29	50.1	-15.1
Ac-(Gly-4(R)-Hyp-Val)$_{10}$-NH_2	34.9	-6.3	-18	37.8	-7.0
Ac-(Gly-4(R)-Hyp-Ser)$_{10}$-NH_2	26.4	-10.4	-30	-[c]	-
Ac-(Gly-4(R)-Hyp-*allo*Thr)$_{10}$-NH_2	37.6	-7.9	-22	39.6	-7.2
Ac(4*R*-Hyp-4*R*-Hyp-Gly)$_{10}$-NH_2 ‖	64.6	-	-	-	-

a Melting temperatures (*Tm*) are concentration-dependent and are determined from the CD measurements (100 μM), together with ΔH^0_{VH} and ΔS^0.
b Melting temperatures (*Tm*) are determined by differential scanning calorimetry at the indicated concentration.
c The serine peptide was not soluble at the higher concentrations needed for differential scanning calorimetry.

Calorimetric studies have shown that the transition enthalpy of the trimer of the peptide Ac(4*R*-Hyp-4*R*-Hyp-Gly)$_{10}$-NH_2 is significantly smaller than that of Ac(Pro-4*R*-Hyp-Gly)$_{10}$-NH_2, and the amide proton exchange experiment also suggested that the triple-helical structure of Ac(4*R*-Hyp-4*R*-Hyp-Gly)$_{10}$-NH_2 appears to fluctuate more than that of the Ac(Pro-4*R*-Hyp-Gly)$_{10}$-NH_2 peptide. These studies showed distinct thermodynamics of folding for the Ac(4*R*-Hyp-4*R*-Hyp-Gly)$_{10}$-NH_2 peptide compared with the Ac(Pro-4*R*-Hyp-Gly)$_{10}$-NH_2 peptide, this indicates that these two peptides might have distinct unfolded states.[32]

3.B.1.4: Crystal structure of Ac(4*R*-Hyp-4*R*-Hyp-Gly)$_9$-NH_2

Bächinger et al[35] also reported the crystal structure of peptide Ac(4*R*-Hyp-4*R*-Hyp-Gly)$_9$-NH_2. Their structure shows that almost all 4*R*-Hyp residues in the X position take the up-pucker conformation. The φ and ψ dihedral angles of these residues are similar to those observed in residues in the X positions of other peptides but are slightly energetically unfavorable for up-puckered proline residues. The 4*R*-Hyp residues in the Y positions, although up-puckered similarly to residues in the Y position of all other

collagen structures, also adopts slightly unfavorable φ and ψ dihedral angles. However, these dihedral angles appear important in the formation of a collagen triple helix. Despite these differences between their structure and other collagen peptides, the overall triple helix of peptide conforms to 7/2 superhelix symmetry. This suggests that, although the slightly unfavorable φ and ψ dihedral angles for residues in the X and Y positions may play a minor role, the altered thermodynamic data for folding are most likely attributable to differences in the unfolded states of our peptide *versus* Pro-Pro-Gly and Pro-Hyp-Gly peptides, which all have predominantly down-puckered residues in their X positions (Fig. 3.19).

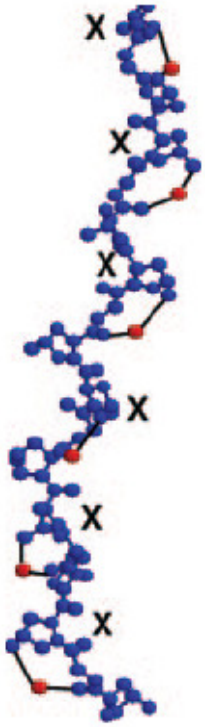


Figure 3.19: ***X 4R-Hyp-specific intra-chain water network.*** The intra-chain water network in which the hydroxyl moieties of the X 4*R*-Hyp residues hydrogen bond to a water molecule that is also hydrogen bonded to the carbonyl oxygen of the Y 4*R*-Hyp residue located one triplet NH_2-terminal to the X4*R*-Hyp. Shown in the figure is the central region of one collagen chain, which is colored *blue*, and the water molecules, which are colored *red*. Also labeled *X* are the X 4*R*-Hyp residues and the NH_2 and COOH termini[35].

Barone et al[30] calculate the importance of vicinal and long-range interresidue effects in determining the stability of the collagen triple-helix by quantum mechanical

(QM) and molecular mechanical (MM) computations on the model peptide Pro-Pro-Gly, taking into account the solvent effects by the polarizable continuum model (PCM). They analyzed all the four possible combinations of ring puckering i.e., Pro(down)-Pro(down)-Gly (hereafter **dd**), Pro(dow*n*)-Pro(up)-Gly (**du**), Pro(up)-Pro(down)-Gly (**ud**), and Pro(up)-Pro(up)-Gly (**uu**) (Fig. 3.20).

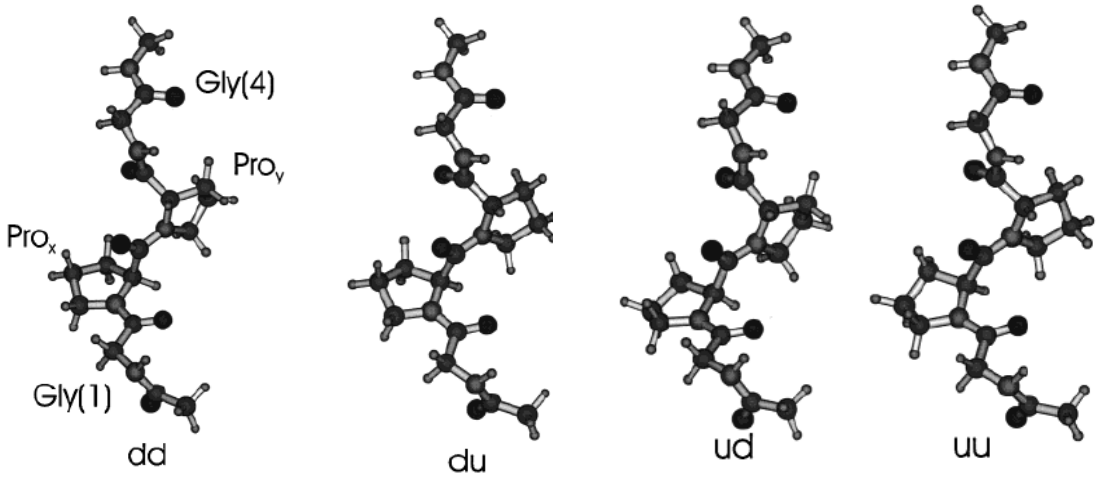


Figure 3.20: Minimum energy geometry (HF/6-31G(d) calculations)[36] of the all-PII isomers of **dd** **du**, **ud**, and **uu**. Φ and ψ dihedrals of the N-terminal glycine have been constrained to the average value of glycine residues in PPG10.

AMBER calculations predict that the triple helix formed by the **du** peptide is the most stable one. This result could be, in principle, due to two different effects: (i) the backbone dihedrals of the **du** peptides are the closest to those allowing the best packing of the triple helix, and (ii) the stability order is determined by the different ring-ring interactions of the four isomers. For example, the steric repulsions between two prolines with opposite puckering, as is the case in **du** and in **ud**, could be smaller than those between two prolines with equal puckering. The results of the AMBER geometry optimizations support the first hypothesis, since the stabilization due to the formation of the triple helix of the **uu** isomer is slightly larger than that of **du**, and, thus, **uu** is less stable than **du** just because the PII single helix is remarkably less stable for **uu** than for

du. The next task to tackle is to understand why the PPG_{10} triple helix tends to adopt backbone dihedrals typical of a down-up alternation. To gain some insights on this question, they performed some AMBER geometry optimizations constraining all the backbone dihedrals of **dd**, **du**, **ud**, and **uu** to their equilibrium values in the single helix. Interestingly, the **dd** bundle is now prevented from forming a regular triple helix, since one of the three chains is quite distant from the remaining two. The number of hydrogen bonds is remarkably smaller than normal triple helix, since all the hydrogen bonds involving the "distant" chain are actually lacking. The backbone dihedrals of the du chain seem to be the most favorable to the maximization of the interchain interactions, especially the electrostatic ones. As a matter of fact, when constrained to the single helix geometry, **du** recovers more than 92% (-196.1 kcal/mol of the total -212.8 kcal/mol) of the electrostatic interchain stabilization obtained by a full geometry optimization. On the other hand, the percentages for the remaining three isomers are smaller, going from 80% (-200.9 kcal/mol of the total -252.7 kcal/mol for **uu**) to 89% (-188.3 kcal/mol of the total -210.8 kcal/mol for **dd**). The Van der Waals interactions seem instead less dependent on the backbone dihedrals, since all the isomers recover more than 90% of the final van der Waals energetic stabilization when frozen at the insulated helix geometry. Not surprisingly, the **dd** chain, which does not exhibit a perfect triple helix, is the system which is less stabilized by Van der Waals interactions (-268.5 kcal/mol of the total -291.2 kcal/mol, i.e., ≈92%). Inclusion of solvent effects (by means of the CPCM) does not change the relative ordering of the different conformers, even if it decreases the energy gap between the **dd** conformer and the other three structures.

3.B.2: Aim and rationale of the present work

According to HF/6-31G(d) calculations discussed above, the **uu** conformer is 1 kcal/mol more stable than **du** and **ud** (exhibiting one up conformer), which are practically isoenergetic. Zagari and co-workers[34] concluded that 4*R*-Hyp stabilizes in the Y position independent of the residue type in X. Conversely, stabilization in the X occurs only when the residue located in Y is able to provide extra-stabilizing interactions. Hence it is tempting to believe that this is the reason polyhydroxylase almost exclusively acts upon proline residues at Y position of vertebrate collagens, with no need to control which residue is located in X. In Chapter **2** and first part of present chapter we have shown that 4*R*-Amp provide extra stabilizing interactions to stabilize the collagen triple helix when present at Y position. Keeping this in mind, in order to further understand the triplex stabilizing effects of 4*R*-aminoproline, it was thought to synthesise and study the chimeric collagen peptides having different combination of 4*R*-Amp/Hyp together in Y and X positions of collagen triple-helix.

The objectives of this chapter are

1) Solid phase synthesis of collagen chimeric peptides using 4*R*-Amp/Hyp, together in Y and X position respectively.
2) Cleavage of peptides from solid support, purification, and characterization of these peptides.
3) Triple-helix forming studies using temperature dependent CD spectrophotometry.

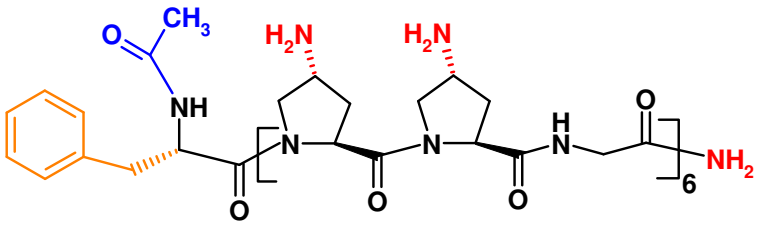
3.B.3: Synthesis of chimeric collagen peptides

The following N and C terminal capped peptides (Fig 3.21) were synthesized by incorporating 4*R*-amino(hydroxy) prolines manually by solid phase peptide synthesis on Rink amide resin into the collagen model peptide Ac-Phe(X-Y-Gly)$_6$-NH_2 using Fmoc

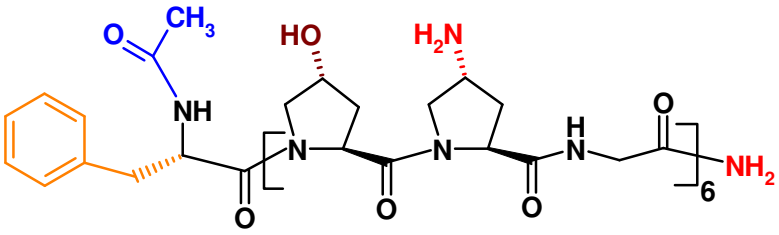
strategy. The peptides were cleaved from the resin using 50% TFA in DCM, with 1% TIS (triisopropylsilane), followed by RP-C4 HPLC purification with water-acetonitrile gradient using 0.1% TFA as organic modifier, and characterized by mass spectrometry (MALDI-TOF). The purity of final peptides was ascertained by RP-C18 HPLC and the peptides were found be greater than 95% pure. The MALDI-TOF mass obtained for peptides agreed closely with calculated values (Table 16).

Table 16: Calculated and observed masses for peptides **27-30.**

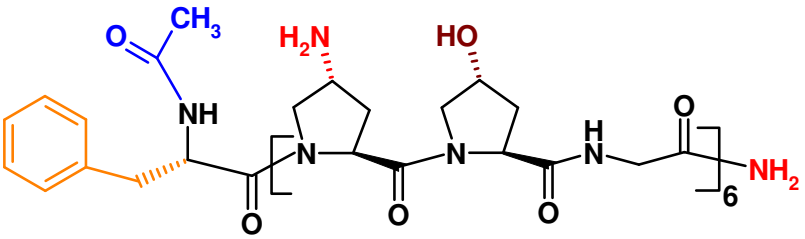
Peptide		Mol. formula	Mass (calc)	Mass (obs)
Ac-Phe(Amp-Amp-Gly)$_6$-NH$_2$	**27**	$C_{83}H_{150}N_{32}O_{20}$	1914.21	1916.58
Ac-Phe(Hyp-Amp-Gly)$_6$-NH$_2$	**28**	$C_{83}H_{144}N_{26}O_{26}$	1920.15	1921.73
Ac-Phe(Amp-Hyp-Gly)$_6$-NH$_2$	**29**	$C_{83}H_{144}N_{26}O_{26}$	1920.15	1923.03
Ac-Phe(Hyp-Hyp-Gly)$_6$-NH$_2$	**30**	$C_{83}H_{136}N_{20}O_{32}$	1926.42	1931.15



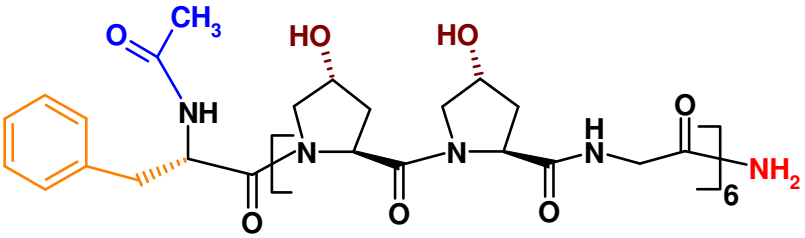
Peptide 27 **(Amp-Amp-Gly)**



Peptide 28 **(Hyp-Amp-Gly)**



Peptide 29 **(Amp-Hyp-Gly)**



Peptide 30 **(Hyp-Hyp-Gly)**

Figure 3.21: structure of chimeric peptides used in the present study.

3.B.4: Biophysical Studies

3.B.4.1: Concentration dependent triple-helix formation of peptides 27-30

In order to evaluate the concentration dependent conformation behavior of the chimeric peptides , CD spectra were recorded in the concentration range of 0.05mM-0.30mM at pH 3 (20mM acetate buffer, 0.1M NaCl). Figure 3.22 shows CD spectra of chimeric peptides **27**-**30** taken at 25 ^{o}C. Spectra of peptides **27** (Amp-Amp-Gly) and **28** (Hyp-Amp-Gly) in the entire concentration range show almost identical positive and negative maxima at 223nm and 205nm respectively. Importantly, all the spectral traces pass through an isobestic point at 216nm. The plot of Rpn values derived from these spectra against concentration of the peptides show that Rpn value increases rapidly from 0.05mM through 0.10mM to reach a near saturation at 0.15mM, and remains nearly constant thereafter. Though concentration dependent CD spectra of peptide **29** (Amp-Hyp-Gly) and **30** (Hyp-Hyp-Gly), in the entire concentration range show nearly identical positive and negative maxima, at 223 and 207nm respectively, they did not show single isobestic point. At lower concentration the cross over point appears at 219nm, shifted to usual 216nm at concentration above 1.5mM. The plot of Rpn values derived form these spectra against concentration also increases rapidly from 0.05mM through 0.10mM to reach a near saturation at 0.15mM, remains nearly constant thereafter. A critical triple-helical concentration of ~0.15mM is derived from these plots; hence, all further studies are performed at a concentration of 0.2mM.

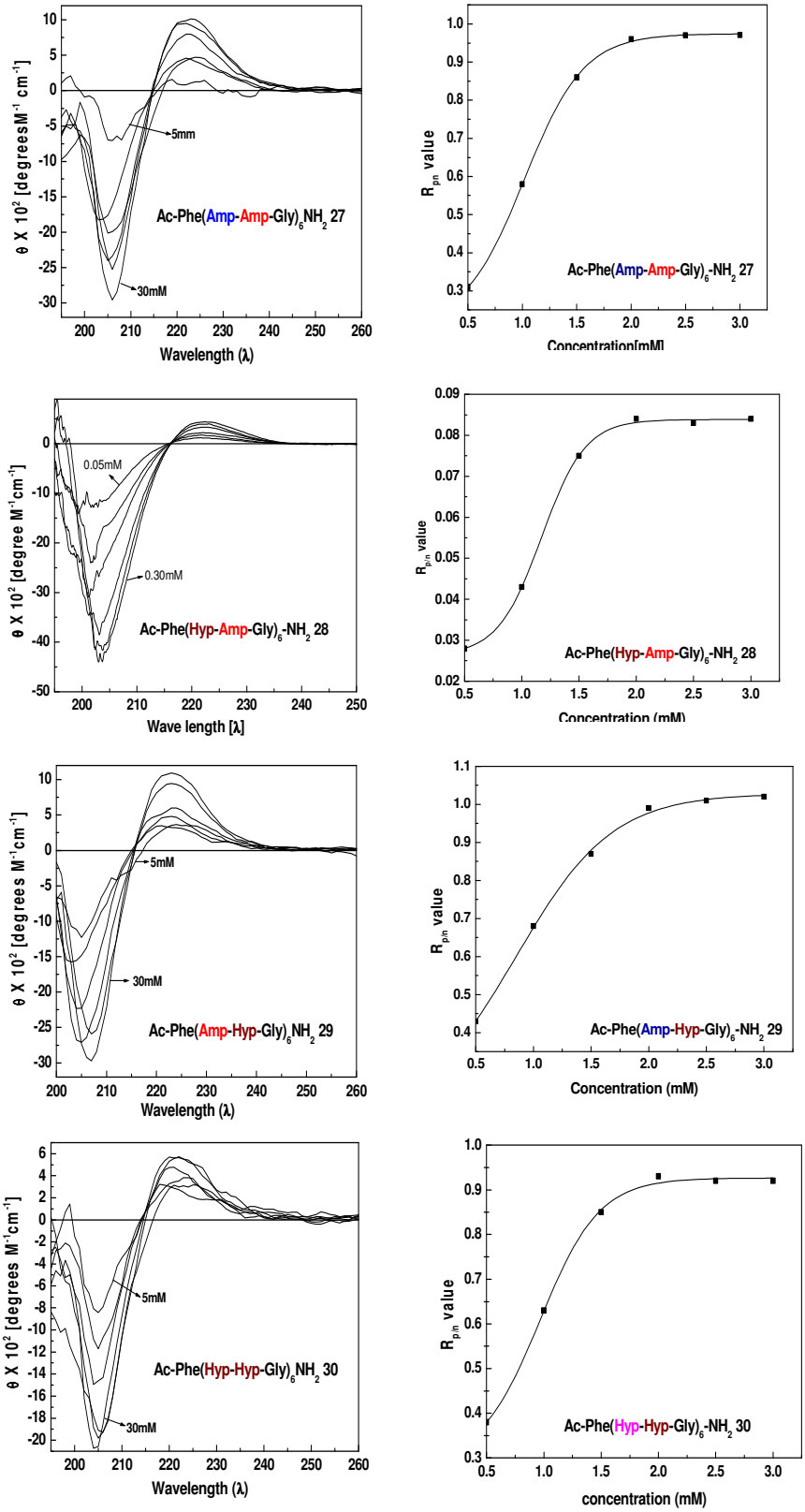


Figure 3.22: Concentration dependent CD spectra and plot of R_{pn} Vs concentrations for the peptides **27-30**

3.B.4.2: Characterization of triple-helical structure by CD spectroscopy

Figure 3.23 shows CD spectra of 0.2mM solutions of peptide **27** (Amp-Amp-Gly) taken at 10 ^{o}C in acidic (20mM acetate, pH 3.0) neutral (20mM phosphate, pH 7.0) and alkaline (20mM borate, pH 9.0 and 12.0) conditions in presence of 0.1M NaCl. Peptide **27** (Amp-Amp-Gly) shows the CD spectra which resemble the collagen triple-helix at pH 3.0, 7.0 and 9.0. Interestingly at pH 12 peptides **27** shows significant appearance of positive band but negative band completely disappeared.

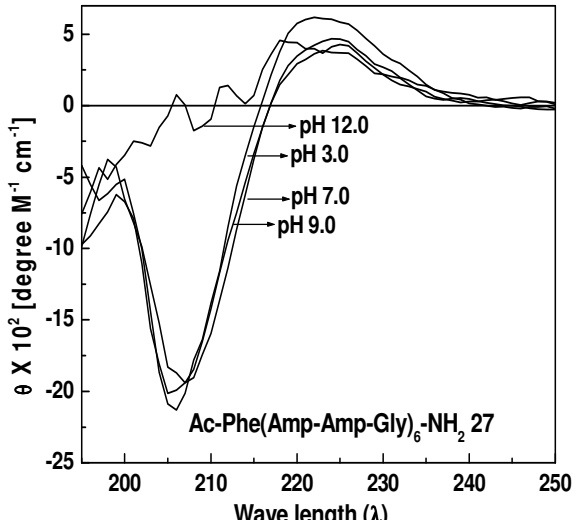


Figure 3.23: CD spectra of peptide **27**Ac-Phe(Amp-Amp-Gly)$_6$-NH$_2$ taken at different pH conditions.

Peptide **27** (Amp-Amp-Gly) shows negative minima at 205nm and positive maxima at 223 nm at pH 3.0, 7.0, and 9.0. The CD spectrum appearance does not change much in these pH conditions. The Rpn values of peptide **27** at different pHs are listed in Table 17. At pH 3.0 peptide shows maximum Rpn values 0.24 followed by at pH 7.0 (0.23) and at pH 9.0 (0.23). The Rpn values in these pH conditions are in the range proposed for triple-helical conformation.

Table 17: Rpn values of peptide **27** (Amp-Amp-Gly) at different pH conditions.

Peptide 27 Ac-Phe(Amp-Amp-Gly)$_6$-NH$_2$			
pH	**+ve band(nm)**	**-ve band (nm)**	**Rpn value**
3.0	223	205	0.24
7.0	223	205	0.23
9.0	223	205	0.1
12.0	223	-	-

Figure 3.23 shows the thermal denaturation curves of peptide **27** (Amp-Amp-Gly) measured by monitoring the molar ellipticity at 225nm against temperature at acidic (20mM acetate, pH 3.0) neutral (20mM phosphate, pH 7.0) and alkaline (20mM borate, pH 9.0 and 12.0) conditions in presence of 0.1M NaCl.

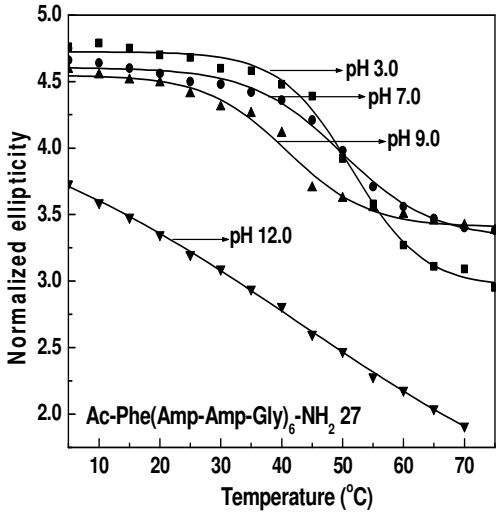


Figure 3.23: Thermal denaturation curves of peptide **27** Ac-Phe(Amp-Amp-Gly)$_6$-NH$_2$ at different pH conditions. pH 3.0 (20mM acetate buffer), pH 9.0 (20mM phosphate buffer, pH 9 and 12 (20mM borate buffer). All buffers contain 0.1M NaCl.

Except at pH 12.0, peptide **27** shows sigmoidal transition curve with increasing temperature at pH 3.0, 7.0 and 9.0. This indicates that peptide **27** is associated in triple-helical conformation at these pH. At pH 3.0 this peptide forms a more stable triplex structure with T_m 51.5^{0}C, which decreases with increasing pH (Table 18).

Table 18: CD-Tm values of peptide **27** (Amp-Amp-Gly) at different pH conditions.

~~**Peptide 28 Ac-Phe(Hyp-Amp-Gly)$_6$-NH$_2$**~~

Peptide 27 Ac-Phe(Amp-Amp-Gly)$_6$-NH$_2$

pH	**T_m ^oC**
3.0	51.5
7.0	40.8
9.0	25
12.0	-nt-

-nt- no transition

Figure 3.24 shows CD spectra of 0.2mM solutions of peptide **28** (Hyp-Amp-Gly) recorded at 10 ^oC in acidic (20mM acetate, pH 3.0) neutral (20mM phosphate, pH 7.0) and alkaline (20mM borate, pH 9.0 and 12.0) conditions in presence of 0.1M NaCl. In both alkaline and acidic conditions the peptide **28** (Hyp-Amp-Gly) shows the CD spectra characteristic of collagen like triple helical structure.

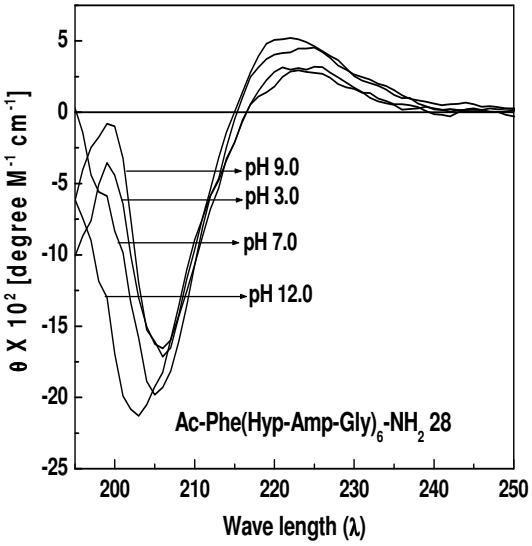


Figure 3.24: CD spectra of peptide **28** Ac-Phe(Hyp-Amp-Gly)$_6$-NH$_2$ taken at different pH conditions.

The Rpn values at different pH for peptide **28** are listed in Table 19. At all pH conditions the Rpn values are within the range of triple-helical conformation. The maximum Rpn value is observed at pH 3.0 and minimum at pH 9.0.

Table 19: Rpn values of peptide **28** (Hyp-Amp-Gly) at different pH conditions.

pH	+ve band(nm)	-ve band (nm)	Rpn value
3.0	223	205	0.35
7.0	223	205	0.21
9.0	223	205	0.18
12.0	223	202	0.33

The normalized ellipticity at 225nm plotted against temperatureis shown in Figure 3.25 for peptide **23** (amp-Amp-Gly).

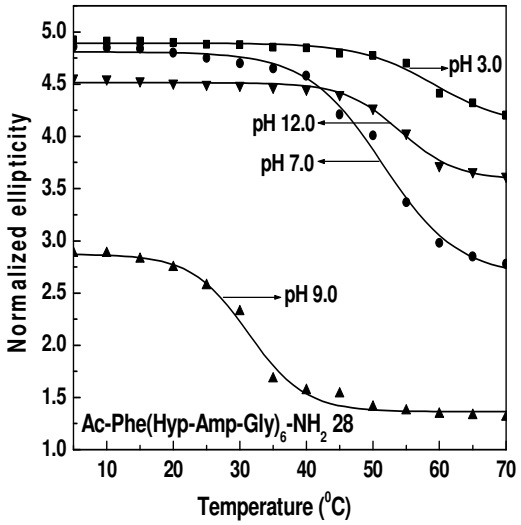


Figure 3.25: Thermal denaturation curves of peptide **28** Ac-Phe(Hyp-Amp-Gly)$_6$-NH$_2$ at different pH conditions. pH 3.0 (20mM acetate buffer), pH 9.0 (20mM phosphate buffer, pH 9 and 12 (20mM borate buffer). All buffers contain 0.1M NaCl.

Peptide **28** (Hyp-Amp-Gly) shows sigmoidal transition curves under all pH conditions, indicating that this peptide undergoes triple-helical to coil structure under all pH conditions. The T_m values obtained form first derivatives of thermal denaturation curves (Table 20) decreases with increasing pH form 3.0 -9.0 again increases at pH 12.0. At pH 3.0 peptide **28** (Hyp-Amp-Gly) forms a stable triple-helical structure with T_m of 55.8 ^{0}C compared to T_m of 51.5^{0}C at pH 7.0. At pH 9.0 it forms a least stable triplex structure (T_m 29 ^{0}C) and at pH 12.0 (T_m 53 ^{0}C) forms a stable triple-helix compared to pH 9.0.

Table 20: CD-Tm values of peptide **28** (Hyp-Amp-Gly) at different pH conditions.

Peptide 28 Ac-Phe(Hyp-Amp-Gly)$_6$-NH$_2$	
pH	**T_m ^oC**
3.0	55.8
7.0	51.5
9.0	29
12.0	53

Figure 3.26 shows CD spectra of 0.2mM solutions of peptide **29** (Amp-Hyp-Gly) recorded at 10 ^oC in acidic (20mM acetate, pH 3.0) neutral (20mM phosphate, pH 7.0) and alkaline (20mM borate, pH 9.0 and 12.0) conditions in presence of 0.1M NaCl.

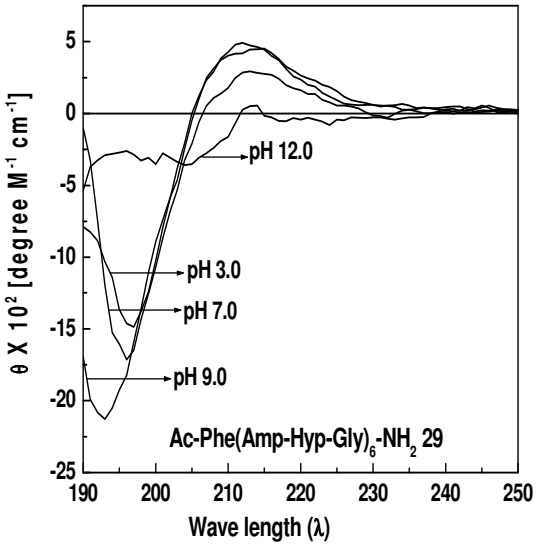


Figure 3.26: CD spectra of peptide **28** Ac-Phe(Amp-Hyp-Gly)$_6$-NH$_2$ taken at different pH

The peptide **29** (Amp-Hyp-Gly) shows the CD spectra characteristic of collagen likes triple-helical structure at pH 3.0, 7.0 and 9.0, conditions. Both positive and negative CD bands do not appears at pH 12.0.

Table 21: Rpn values of peptide **29** (Amp-Hyp-Gly) at different pH conditions.

Peptide 29 Ac-Phe(Amp-Hyp-Gly)$_6$-NH$_2$			
pH	**+ve band(nm)**	**-ve band (nm)**	**Rpn value**
3.0	223	195	0.32
7.0	223	195	0.28
9.0	223	193	0.17
12.0	-	-	-

The Rpn values at different pH for peptide **29** decreases with increasing pH (Table 21). At pH 3.0, 7.0 and 9.0, the Rpn values are in the range proposed for triple-helical conformation. The maximum Rpn value is observed at pH 3.0 (0.32) and minimum at pH 9.0 (0.17).

Figure 3.27 shows the thermal denaturation curves of peptide **29** (Amp-Hyp-Gly) measured by monitoring the molar ellipticity at 225nm against temperature at acidic (20mM acetate, pH 3.0) neutral (20mM phosphate, pH 7.0) and alkaline (20mM borate, pH 9.0 and 12.0) conditions in presence of 0.1M NaCl.

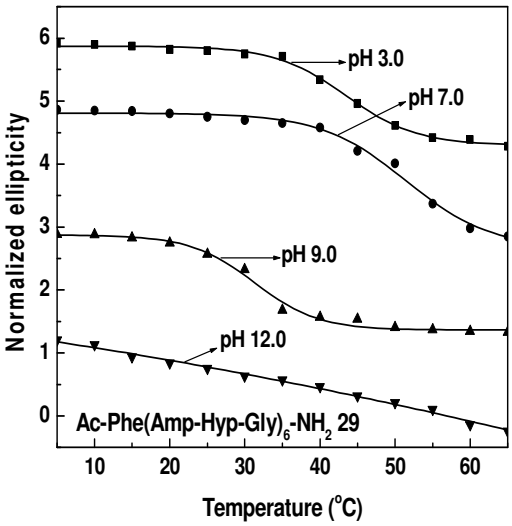


Figure 3.27: Thermal denaturation curves of peptide **28** Ac-Phe(Amp-Hyp-Gly)$_6$-NH$_2$ at different pH conditions. pH 3.0 (20mM acetate buffer), pH 9.0 (20mM phosphate buffer, pH 9 and 12 (20mM borate buffer). All buffers contain 0.1M NaCl.

Peptide **29** (Amp-Hyp-Gly) shows sigmoidal transition curves at pH 3.0, 7.0 and 9.0, indicating that this peptide undergoes triple-helical to coil structure under these pH conditions. At pH 12.0 it shows a linear decrease in ellipticity with temperature. The T_m values obtained form first derivatives of thermal denaturation curves (Table 22) decreases with increasing pH. Maximum T_m was observed at pH 3.0 (43.1^0C) which decreases trough pH 7.0 with T_m 36^{o}C and at pH 9.0 with T_m of 29^{o}C. No transition was observed at pH 12.0.

Table 22: CD-Tm values of peptide **29** (Amp-Hyp-Gly) at different pH conditions.

Peptide 29 Ac-Phe(Amp-Hyp-Gly)$_6$-NH$_2$	
pH	**T_m ^{o}C**
3.0	43.1
7.0	36
9.0	29
12.0	-nt-

-nt- no transition

Figure 3.28 shows CD spectra of 0.2mM solutions of peptides **30** (Hyp-Hyp-Gly) recorded at 10 ^{o}C in acidic (20mM acetate, pH 3.0); neutral (20mM phosphate, pH 7.0) and alkaline (20mM borate, pH 9.0 and 12.0) buffer conditions in presence of 0.1M NaCl. In both alkaline and acidic conditions the peptide **30** (Hyp-Hyp-Gly) shows the CD spectra characteristic of collagen like triple-helical structure. The Rpn values at different pHs for peptide **30** are listed in Table 23. At all pH conditions, the Rpn values are in the range of triple-helical conformation. The maximum Rpn values are observed at acidic and neutral conditions and minimum at basic conditions.

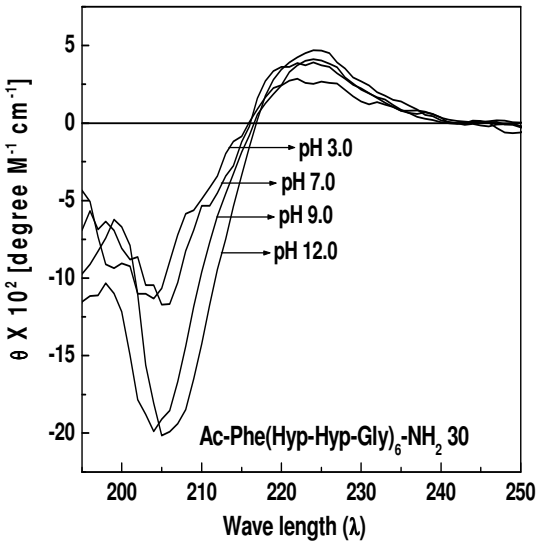


Figure 3.28: CD spectra of peptide **30** Ac-Phe(Hyp-Hyp-Gly)$_6$-NH$_2$ taken at different pH

Table 23: Rpn values of peptide **30** (Hyp-Hyp-Gly) at different pH conditions.

Peptide 30 Ac-Phe(Hyp-Hyp-Gly)$_6$-NH$_2$			
pH	**+ve band(nm)**	**-ve band (nm)**	**Rpn value**
3.0	223	205	0.4
7.0	223	205	0.35
9.0	223	205	0.21
12.0	223	205	0.19

Figure 3.29 shows the T_m curves of peptide **30** (Hyp-Hyp-Gly) measured by monitoring the molar ellipticity at 225nmagainst temperature. Peptide **30** (Hyp-Hyp-Gly) shows sigmoidal transition curves at all pH conditions, indicating that this peptide undergoes triple-helical to coil structure under all pH conditions. The T_m values obtained form first derivatives of thermal denaturation curves (Table 24) are almost constant under all pH conditions. This indicates that the triplex stability of the peptide **30** is independent of pH condition.

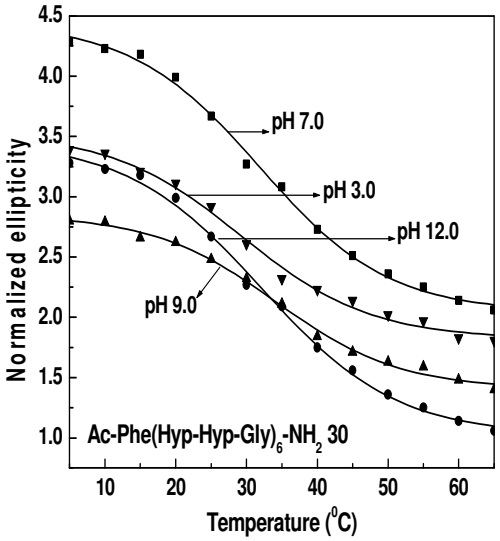


Figure 3.29: Thermal denaturation curves of peptide **30** Ac-Phe(Hyp-Hyp-Gly)$_6$-NH$_2$ at

Table 24: CD-Tm values of peptide **30** (Hyp-Hyp-Gly) at different pH conditions. NaCl.

Peptide 30 Ac-Phe(Hyp-Hyp-Gly)$_6$-NH$_2$	
pH	**T_m ^{o}C**
3.0	31.8
7.0	32
9.0	30.9
12.0	30

Peptide **28** (Hyp-Amp-Gly) under all pH conditions shows a maximum T_m compared to other peptides **27**, **29** and **30**. The T_m of 55.8^{0}C is observed at pH 3.0, which decreases through pH 7.0 (T_m = 51.5^{0}C), pH 9.0 (T_m = 31^{0}C) and again increases at pH 12.0 (T_m = 53^{0}C). Similar kind of behavior was observed for Peptide **27** (Amp-Amp-Gly) and peptide **29** (Amp-Hyp-Gly). Peptide **27** (Amp-Amp-Gly) which shows maximum T_m of 51.5^{0}C at pH 3.0 decreases with increasing pH (pH 7.0 T_m = 40.8^{o}C, pH 9.0 T_m = 35^{0}C) and similarly peptide **29** (Amp-Hyp-Gly) which shows maximum T_m of 43.1^{0}C at pH 3.0 also decreases with increasing pH (pH 7.0 T_m = 36^{o}C; pH 9.0 T_m = 29^{0}C). Both the peptides show linear a decrease in ellipticity, but do not form a triplex

at pH 12.0. The peptide **30** (Hyp-Hyp-Gly) shows almost constant T_m under all pH conditions (Table 25).

Table 25: CD-Tm values of peptides **27-30**

Ac-Phe(X--- Y--- Gly)$_6$-NH$_2$		T_m values (^oC)			
		pH 3.0	pH 7.0	pH 9.0	pH 12.0
Ac-Phe(Amp-Amp-Gly)$_6$-NH$_2$	27	51.5	40.8	25	-nt-
Ac-Phe(Hyp-Amp-Gly)$_6$-NH$_2$	28	55.8	51.5	31	53
Ac-Phe(Amp-Hyp-Gly)$_6$-NH$_2$	29	43.1	36	29	-nt-
Ac-Phe(Hyp -Hyp-Gly)$_6$-NH$_2$	30	31.8	32	30.9	30

3.B.4.3: Effect of ethylene glycol on the stability of peptides 27-30 triple-helices

Figure 3.30 shows the CD spectra of 0.2mM solution of the peptides **27-30** taken in a 3:1 mixture of EG: water. Like in aqueous solution, all peptides show CD spectra which resemble the CD spectra of collagen triple helix also in mixture of EG:W.

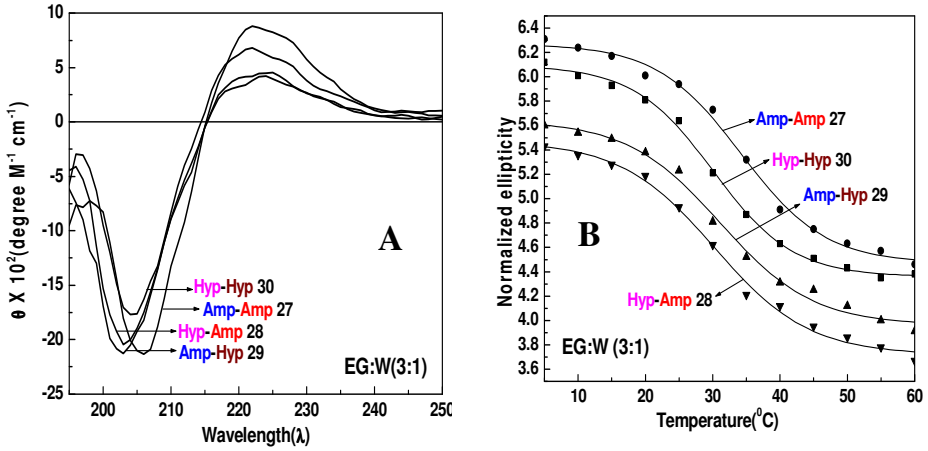


Figure 3.30: **A**, CD spectra at 10^{o}C and **B,** CD-T_m curve of 0.2M solutions of peptides **27-30** in 3:1 (v/v) EG:W mixture

In EG:W, the T_m of peptide **30** (Hyp-Hyp-Gly) is 29.8^{0}C which is lower compared to peptide **28** (Amp-Hyp-Gly) (T_m = 31.2^{0}C) (Table 26). This behavior of

peptide **28** (Hyp-Amp-Gly) is opposite to its behavior in aqueous solutions where under all pH conditions it forms more stable triple-helical structure than other peptides **27**, **29** and **30**.

Table 26: R_{pn} values and T_m of peptides **27-30** in 3:1 (v/v) EG:W mixture

Ac-Phe(X---Y---Gly)$_6$-NH$_2$		R_{pn} values	T_m ^0C
Ac-Phe(Amp-Amp-Gly)$_6$-NH$_2$	**27**	0.36	34
Ac-Phe(Hyp-Amp-Gly)$_6$-NH$_2$	**28**	0.33	31.2
Ac-Phe(Amp- Hyp-Gly)$_6$-NH$_2$	**29**	0.25	30.5
Ac-Phe(Hyp-Hyp-Gly)$_6$-NH$_2$	**30**	0.26	29.8

3.B.4.4: Salt effects on triple-helix stability of chimeric peptides

Figure 3.31 shows the plot of triple-helical stability of peptide **27** (Amp-Amp-Gly) and peptide **30** (Hyp-Hyp-Gly) against concentration of NaCl. The T_m values for peptide **27** shows a steady increase upto 100mM NaCl and a slight decrease beyond (no salt, 30 °C; 50mM, 45 °C; 100mM, 51.5 °C; 200mM 50 °C) (Table 27). As compared to that of peptide **30** under similar conditions (no salt M, 30 °C, 50mM, 31 °C; 100mM, 31.8 °C; 200mM, 30. 4°C), the T_m was almost uniform over the entire salt concentration range.

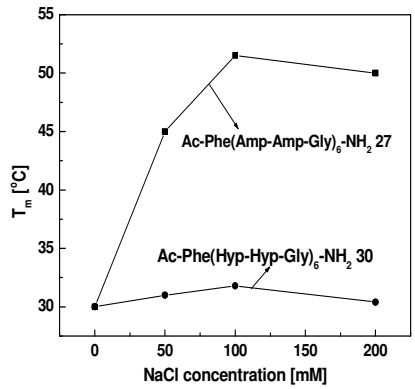


Figure 3.31: Plot of T_m Vs concentration of NaCl for Amp peptide **27** and Hyp peptide **30**.

Table 27: T_m values of peptides **27** and **30** at pH 7.0 (20mM phosphate buffer) with varying NaCl concentrations

Conc. of NaCl	T_m values (^{0}C)	
	27	**30**
0.1 M	30	30
0.6 M	45	31
1.0 M	51.5	31.8
2 .0 M	50	30.4

3.B.5: Discussion

The replacement of proline with either 4*R*-aminoproline or 4*R*-hydroxyproline in the X position of the peptides Ac-Phe(Pro-Amp-Gly)$_6$-NH_2 and Ac-Phe(Pro-Hyp-Gly)$_6$-NH_2 forms a polyproline-II like structure as indicated by a positive maxima at 223nm and strong negative band ~200nm in the CD spectra at pH 3.0. The peak positions are almost independent of concentration for all chimeric peptides **27-30**. All the spectral traces pass through the same isobestic point at 217nm for the peptides **27** (Amp-Amp-Gly) and **28** (Hyp-Amp-Gly) which contain Amp at Y position. The peptides **29** (Amp-Hyp-Gly) and **30** (Hyp-Hyp-Gly) contain Hyp at Y position and show two isobestic points. At lower concentration, the crossover point appears at 219nm, shifted to usual 216nm at concentration above 1.5mM. Within the concentration range measured i.e. 0.05mM-0.30mM all peptides **27**-**30** exist in the triple-helical state. A smooth and progressive increase in the Rpn values with the concentration also suggests a two state equilibrium.

The observed sigmoid transition for all chimeric peptides in the variable temperature CD measurements provide additional evidence for the two state (triple-helix to coil) transition. The stability (T_m values) of peptides **27** (Amp-Amp-Gly) and

29 (Amp-Hyp-Gly) decreases with increasing pH. The thermal stability of peptide **27** (Amp-Amp-Gly) is greater than that of peptide **29** (Amp-Hyp-Gly) at acidic and neutral pH conditions (ΔT_m 8.4^oC at pH 3.0 and 4.8^oC at pH 7.0), but the order is reversed at pH 9.0 in which peptide **29** (Amp-Hyp-Gly) forms more stable triple-helix (ΔT_m 4^oC). Both peptides did not associate to triple-helical conformation at pH 12.0 which was confirmed by the observed linear decrease in ellipticity for both in their thermal denaturation plot.

The replacement of Pro with Amp in the peptides Ac-Phe(Pro-Amp-Gly)$_6$-NH_2 **19** and Ac-Phe(Pro-Hyp-Gly)$_6$-NH_2 (control peptide[9] Capter **2**) has remarkable effect on triple helical stability. The stability of peptide **27** (Amp-Amp-Gly) is always lower compared to control peptide (Pro-Amp-Gly) under identical conditions, whereas the stability of peptide **29** (Amp-Hyp-Gly) is higher compared control peptide (Pro-Hyp-Gly). This may be due the intrastrand charge-charge repulsion arising from stereodisposition of the NH_3^+ groups of two aminoprolines which are very close to each other in peptide **27** (Amp-Amp-Gly). In case of peptide **29** (Amp-Hyp-Gly) which contains only one aminoproline, such a type of intrastrand charge-charge repulsion is negligible. The change in ring pucker with change of acidic to basic pH for aminoproline may not be compatible for both peptides **27** (Amp-Amp-Gly) and **29** (Amp-Hyp-Gly) this results in absence of triple-helix at pH 12.0. Peptide **30** (Hyp-Hyp-Gly) associates into the triple helix at all pH conditions. This is evidenced by the CD spectra which resemble collagen triple helix and the appearance of co-operative transition in ellipticity with increasing temperature. The presence of Hyp in the X position has positive effect in stabilizing the triple-helical structure since it forms a

stable triple-helix (ΔT_m 4.8°C at pH 3.0; 4.0°C at pH 7.0; 3.9°C at pH 9.0; 3.0°C at pH 12.0). Moreover the stability of triple-helix is independent of pH.

Peptide **28** (Hyp-Amp-Gly) forms a more stable triple-helical structure compared to other three peptides **27**, **29** and **30** under all pH conditions. The insertion of Hyp in place of Pro in the peptide (Pro-Amp-Gly) results in triple-helical structure under all pH conditions, it destabilizes the triple-helix at acidic and neutral conditions compared to (Pro-Amp-Gly) peptide (ΔT_m 4.2°C at pH 3.0; 2.9°C at pH 7.0) and stabilizes the triple helical structure at basic conditions (ΔT_m 5.0°C at pH 9.0; 4.0°C at pH 12.0).

The replacement of proline with either Amp, Hyp or amp in the X position of peptide Ac-Phe(X-Amp-Gly)$_6$-HN$_2$ has considerable effect on stability of triple-helical conformation. Replacement with amp (4*S*) [(amp-Amp-Gly) **23]** allows peptides to associate in triple-helical conformation under all pH conditions despite the decreased stability of the triple-helices with increasing pH (Fig. 3.32). Replacing with Amp (4*R*) [(Amp-Amp-Gly) **27**] allows peptide to associate triple-helical conformation only at pH 3.0, 7.0 and 9.0, and prevent at pH 12.0. The replacement of Hyp (4*R*) [(Hyp-Amp-Gly) 2**8**] also allows peptide to fold in triple-helical conformation under all pH conditions. The stability of triple-helical structure decreases from pH 3.0-9.0 followed by again increase at pH 12.0. Similar kind of behavior was observed for the peptide Ac-Phe(Pro-Amp-Gly)$_6$-NH$_2$.

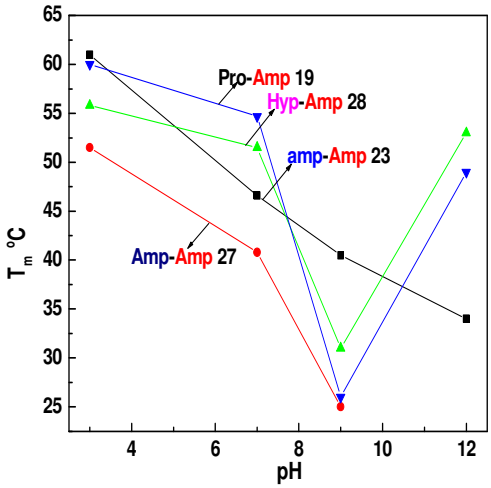


Figure 3.32: Comparison of T_m of peptides **19**, **23**, **27**, and **28**

The presence of NaCl in the concentration range 0-200mM has dramatic effects on the T_m values of peptide **27** (Amp-Amp-Gly) triple-helices at pH 7.0. Higher concentration of salt stabilizes the triple-helical structure as is evidenced by very high T_m values for peptide **27** (Amp-Amp-Gly) (51.5^0C) at 100mM NaCl and 30^0C at absence of NaCl. This stabilization by NaCl arises from the screening of interstrand as well as intrastrand electrostatic repulsion caused by protonated amino groups of 4-aminoprolines. The pKa value for Amp monomer is 10.5 indicating that the side chain 4-aminogroups in the peptide **27** (Amp-Amp-Gly) would largely be protonated at pH 7.0. Collagen is a parallel triple-helix with one residue shift, these $-NH_3^+$ groups lie in close proximity to each other causing large electrostatic repulsive interactions between the chains. The presence of salt counter-ions offers a shielding effect results decreased electrostatic repulsion between the chains. In peptide **30** (Hyp-Hyp-Gly) where electrostatic repulsion is absent, the concentration of salt did not have much effect on the stability of the triple-helix (Fig. 3.31). At higher salt concentration (200mM) of NaCl breaking of the hydrogen bond network, water-structure and poor solubility of

peptides lead to a decrease in the triple-helical stability consequently decreased the T_m value for both peptides **27** (Amp-Amp-Gly) and **30** (Hyp-Hyp-Gly).

3.B.6: Conclusion

The chimeric peptides obtained by replacing either Amp, Hyp or amp in X position of the collagen peptide sequence Ac-Phe(X-Amp(Hyp)-Gly)$_6$-NH_2 associate to form triple-helices. The triplex stability is highly dependent on the pH conditions. The chimeric peptides with Hyp at X [(Hyp-Amp-Gly) **28**, (Hyp-Hyp-Gly) **30**] position stabilizes the triple-helical structure better compared to Amp [(Amp-Amp-Gly) **27**, Amp-Hyp-Gly] **29**] at same position. Since the Hyp at the X position of (Hyp-HypGly)$_{10}$ allows the additional hydration in the single-coil state, resulting in a smaller enthalpy change than the (Pro-Hyp-Gly)$_{10}$. This additional hydration may not much effective by charged Amp at X position resulting in the destabilization of the triple-helix.

Actually the thermal stabilities of chimeric peptides were quite different from those predicted according to the conformational preference of pyrrolidine ring puckering at the X and Y positions of (X-Y-Gly)$_{10}$ when 4-aminoproline is present at both the positions. For example (amp-Amp-Gly) **23** expected to form much more stable triple-helical structure than the peptide (Pro-Amp-Gly) **19**, though it forms a stable tripple-helical stature the ΔT_m is 1^0C at pH 3.0. Similarly (Hyp-Hyp-Gly) peptide from more stable triple-helix compared to (Pro-Hyp-Gly). But (Amp-Amp-Gly) destabilizes triple-helical structure compared to (Pro-Amp-Gly) at pH 3.0. This may be due to the two $-NH_3^+$ groups positioned in close proximity to each other in triple-helical structure

causing electrostatic repulsion. The observed large salt effect in these two peptides confirms this fact.

The results in this chapter clearly demonstrate that 4*R*-Amp has dual compatibility in stabilizing the collagen triple-helix when present in X or Y-position, while 4S-amp only in X-site is compatible with Pro, Hyp and Amp in Y-site. It is concluded that the empirical rule about the effect of substitution on the thermal stability of the triple-helix structure is not at all additive with respect to the X and Y doubly-substituted collagen models, but requires another elaboration based on detailed structural analysis as well as precise thermodynamic investigations to address the various problems of the collagen triple-helix. The origin of triple-helix stabilization effects from Amp and amp seems to arise from a complex combination of stereoelectronic and electrostatic (electrosteric) effects, with the latter dominantly contributing to hyperstability of 4-aminocollagen triple-helices.

The conformational stability of the collagen triple helix is enhanced by proper preorganization of its individual polypeptide strands. The modifications that use sterolectronic effects to preorganize proline residue can also introduce adverse steric and inductive effects, thus complicating the design of more stable triple-helices. But a balance can be attained between the benefits of preorganization and the determinants of these adverse effects.

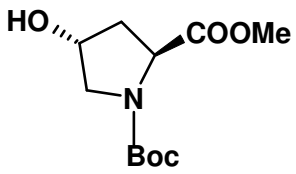
3.B.7: Experimental Section

General. All reagents were obtained from commercial sources and used without further purification except when necessary. ^{1}H and ^{13}C NMR spectra were recorded on a Bruker-AC400 spectrometer. All chemical shifts are with reference to TMS as an

internal standard and are expressed in δ-scale (ppm). Thin-layer chromatography (TLC) was carried out on pre-coated silicagel 60 F254 plates (E. Merck). TLC was visualized with UV and/or ninhydrin spray, followed by heating upto 100^0C.

3.B.7.1: Synthesis of compounds 7-11

(2S, 4R)-N^1-(t-butoxycarbonyl)-4-hydroxyproline methyl ester 7



A solution of compound **1** (10.9 g, 83.12mmol) in of 2N NaOH (50ml) and dioxan (50ml) was cooled to 0^0C and Boc anhydride (20ml 91.41mmol) was added dropwise using addition funnel. The reaction mixture was stirred at 0^0C for 1 h. Dioxane was removed under reduced pressure; aqueous layer was covered with EtOAc (150ml). The reaction mixture was neutralized by adding, solid $KHSO_4$ portion wise with vigorous stirring until pH 2. The organic layer was separated, aqueous layer was extracted with ethyl acetate (3 X 50ml), the combined organic layer was washed with water followed saturated brine solution and dried over anhydrous Na_2SO_4. Upon removal of ethyl acetate under reduced pressure, obtained a white solid which was recrystalized with EtOAc/Petroleum-ether.

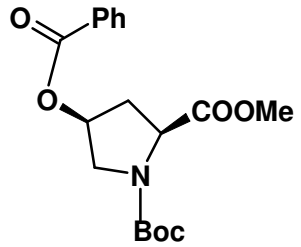
The mixture of white solid compound (19g 82.20mmol) obtained in above step and anhydrous K_2CO_3 (34g, 246.mmol) was dissolved in anhydrous acetone (250ml). The reaction mixture was stirred for 30min at room temperature and dimethylsulphate (9.4ml, 98.6mmol) was added and stirring was continued under reflux condition for 3h. Acetone was evaporated under reduced pressure, resulting solid was dissolved in ethyl acetate (300ml), washed with water (2 X 75ml), followed by saturated brine solution and dried over anhydrous Na_2SO_4. Ethyl acetate was concentrated under reduced

pressure to a pale yellow solid. The product was purified by silica gel chromatography, eluting with ethylacetate:petether (1:1) to yield compound **7** as a white solid. Yield 19.4g, 95.2% over the two steps. Mol Formula $C_{11}H1_9NO_5$; Mass (observed) 246.3 (245.13 Calc).

^{1}H NMR, **δ_H** 1.36 (9H, s), 2.03 (1H, m), 2.25 (1H, m), 3.05 (1H, m), 3.54 (2H, m), 3.69 (3H, s), 4.42 (2H, m).

^{13}C NMR, 28.0, 32.3, 38.9, 51.7, 54.4, 57.4, 57.8, 68.9, 69.5, 80.1, 153.9, 154.4, 173.5.

(2S, 4S)-N^1-(t-butoxycarbonyl)-4-benzoyloxyproline methyl ester 8

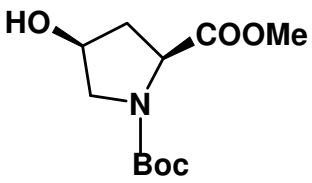


A solution of compound **7** (19g 77.5mmol), Benzoic acid (10.4g, 85.2mmol) and triphenylphosphine (22.36g, 85.25mmol) in dry THF, was cooled to 0^0C. DIAD (17.24ml, 85.2) was added dropwise with syringe under argon atmosphere. The reaction mixture was stirred for 8 h at RT, followed by evaporation to dryness under reduced pressure, resulting residue was purified by silica gel chromatography, eluting with ethyl acetate: petether (2:8) to yield compound **8** as a white solid. Yield 22g, 81%, Mol. Formula $C_{18}H_{23}NO_6$; Mass (observed) 350.2 (349.15 Calc).

^{I}H NMR; **δ_H** 1.41 (9H, s), 2.37-2.59 (2H, m), 3.63 (3H, s), 3.77 (2H, m), 4.91-4.59 (1H, m), 5.50-5.59 (1H, m), 7.36-7.56 (4H, m), 7.91-7.95 (2H, dd).

^{13}C NMR, 27.9, 35.3, 36.3, 51.7, 52.2, 57., 57.5, 72.0, 73.1, 79.9, 128.0, 129.3, 132.9, 153.3, 153.7, 165.3, 165.4, 171.7, 172.1.

(2S, 4S)-N^1-(t-butoxycarbonyl)-4-hydroxyproline methyl ester 9

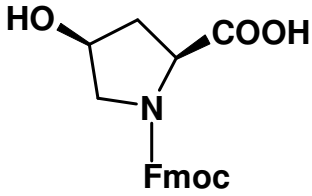


Compound **8** (20g 81.2mmol) was dissolved in methanol (50ml) to which saturated $NaHCO_3$ (20ml) was added. The reaction mixture was stirred for 5h at room temperature. Methanol was evaporated under reduced pressure and the aqueous layer was extracted with ethyl acetate (3 X 100ml). The combined organic layer was washed with water, followed by saturated brine solution and dried over anhydrous Na_2SO_4. Ethyl acetate was concentrated under reduced pressure to a pale yellow oil. The product was purified by silica gel chromatography, eluting with ethylacetate:petether (1:1) to a yield compound **9** as colorless oil. Yield 10g, 71%, Mol Formula $C_{11}H1_9NO_5$; Mass (observed) 245.2 (245.13 Calc).

^{1}H NMR δ_H 1.38 (9H, s), 2.0-2.08 (1H, m), 2.30-2.32 (1H, m), 3.50-3.62 (2H, m), 3.74 (3H, s), 4.23-4.31 (2H, m).

^{13}C NMR, 28.1, 37.7, 38.4, 53.1, 52.2, 54.8, 55.4, 57.7, 69.6, 70.6, 80.1, 153., 154.3, 174.7.

(2S, 4S)-N^1-(9-fluorenylmethyloxycaronyl)-4-hydroxyproline 11



A solution of compound **9** (8g 32.63mmol) in of 50% TFA in dichloromethane (10ml) was stirred at RT for 2 h and the solvent was evaporated under reduced pressure, the resulting residue was further evaporated with diethyl ether to get a white solid, which was dissolved in mixture of methanol (10ml) and 2N LiOH (10 ml). The reaction mixture was stirred for 1 h at RT and neutralized to pH 7 with $KHSO_4$. Methanol was removed under reduced pressure. The

aqueous layer was cooled to 0 ^{o}C, Saturated Na_2CO_3 (25ml) solution and dioxane (50 ml) were added. The reaction mixture was stirred at 0^{O}C for 15min, Fmoc-Cl (9.28g, 35.89mmol) was added in small portions, and the resulting mixture was stirred at room temperature for 8h, by maintaining at pH 8 throughout the reaction. The reaction mixture was then poured into separatory funnel with 50ml of H_2O and extracted with diethyl ether (3 X 25ml). The aqueous layer was covered with ethyl acetate (50ml) in a beaker, cooled to 0^{o}C, and acidified with $KHSO_4$ to pH 2 with stirring. The organic layer was separated and aqueous layer was extracted again with ethyl acetate (3 X 25ml). The combined ethyl acetate layers were dried over anhydrous Na_2SO_4 and the solvent was removed under reduced pressure to yield compound **11** as a white amorphous solid. Yield 5g, 43.37%, Mol. Formula $C_{20}H_{19}NO_5$, Mass (observed) 354.5 (353.13 Calc).

^{1}H NMR, δ_H 2.25-2.31 (2H, m), 3.50-3.67 (2H, m), 4.09-4.26 (1H, m), 3.34-4.49 (3H, m), 5.16-5.28 (2H, m), 7.29-7.42 (4H, m), 7.53-7.56 (2H, dd), 7.68-7.76 (2H, dd).

^{13}C NMR, 37.5, 38.5, 46.9, 56.0, 57.7, 58.0, 67.8, 70.5, 119.8, 124.9, 127.0, 127.6, 141.1, 143.4, 143.6, 155.5, 158.0, 175.7.

3.B.8: Refferences.

1) (a) Okuyama, K.; Tanaka, N.; Ashida, T.; Kakudo, M. *Bull. Chem. Soc. Jpn*. **1979**, *49*, 1850-1810. (b) Okuyama, K.; Tanaka, N.; Ashida, T.; Kakudo, M. *J. Mol. Biol*. **1981**, *152*, 427-443. (c) Nagarajan, V.; Kamitori, S.; Okuyama, K. *J. Biochem*. **1998**, *124*, 1117-1123. (d) Nagarajan, V.; Kamitori, S.; Okuyama, K. *J. Biochem*. **1999**, *125*, 310-318. (e) Kramer, R. Z.; Vitagliano, L.; Bella, J.; Berisio, R.; Mazzarella, L.; Brodsky, B.; Zagari, A.; Berman, H. M. *J. Mol. Biol*. **1998**, *280*, 623-628.

2) (a) Bella, J.; Eaton, M.; Brodsky, B.; Berman, H. M. *Science* **1994,** *266*, 75-81. (b) Bella, J. Brodsky, B.; Berman, H. M. *Structure* **1995**, *3*, 893-906.

3) (a) Rich, A.; Crick, F. H. C. *Nature* **1955**, *176*, 915-916. (b) Rich, A.; Crick, F. H. C. *J. Mol. Biol.* **1961**, *3*, 483-506.

4) Creighton, T. E. *Proteins 2nd edn*., W. H. Freeman & Co., **1993**, pp. 173-174.

5) Sakakibara, S.; Inouye, K.; Shudo, K. S.; Kishida, Y.; Kobayashi, Y.; Prockop, D. J. *Biophys. Biochem. Acta* **1973**, *303*, 198-202.

6) (a) Eberhardt, E. S.; Loh, S. N. P.; Raines, R. T. *Tetrahedron Lett*. **1993**, *33*, 3055-3056. (b) Eberhardt, E. S.; Panasik, N. Jr.; Raines, R. T. *J. Am. Chem. Soc*. **1996**, *118*, 12261-12266.

7) Inyoue, K.; Sakakibara, S.; Prockop, D. J. *Biochim. Bophys. Acta* **1976**, *420*, 133-142.

8) Bretsgher, L. E.; Jenkins, C. L.; Tylor, K. M.; DeRider, M. L.; Raines, R. T. *J. Am. Chem. Soc*. **2001**. *123*. 777-778.

9) Babu, I. R.; Ganesh, K. N. *J. Am. Chem. Soc*. **2001**, *123*, 2079-2080.

10) Inouye, K.; Kobayashi, Y.; Kyogoku, Y.; Kishida, Y.; Sakakibara, S.; Prockop, D. J. *Arch. Biochem. Biopsy*. **1982**, *219*, 198-203.

11) (a) Nagarajan, V.; Kamitori, S.; Okuyama, K. J. Biochem. 1999, 125, 310-318. (b) Berisio, R.; Vitagliano, L.; Mazzarella, L.; Zagari, A. *Protein Science* **2002**, *11*, 262-270.

12) (a) Kobayashi, Y.; Sakai, R.; Kakiuchi, K.; Isemura, T. *Biopolymers* **1970**, *9*, 415-425. (b) Uchiyama, S.; Kai, T.; Kajiyama, K.; Kobayashi, Y.; Tomiyama, T. *Chem. Phys. Lett*. **1997**, *281*, 92-96.

13) (a) Holmgren, S. K.; Taylor, K. M.; Brestscher, L. E.; Raines, R. T. *Nature* **1998**, *392*, 666-667. (b) Holmgren, S. K.; Brestscher, L. E.; Taylor, K. M.; Raines, R. T. *Chem. Biol*. **1999**, *6*, 63-70.

14) Doi, M.; Nishi, N.; Uchiyama, S.; Nishiuchi, Y.; Nakazawa, T.; Ohkubo, T.; Kobayashi, Y. *J. Am. Chem. Soc*. **2003**, *125*, 9922-9923.

15) Vitagliano, L.; Berisio, R.; Mazzarella, L.; Zagari, A. *Biopolymers* **2001**, *58*, 459-464.

16) Renner, C.; Alefeder, S.; Bae, J. H.; Budisa, N.;Huber, R.; Moroder, L. *Angew. Chem. Int. Ed***. 2001**, *40*, 923-935.

17) (a) Panasik, N. Jr.; Eberhardt, E. S.; Edison, A. S.; Powell, D. R.; Raines, R. T. *Int. J. Pept. Protein Res*. **1994**, *44*, 262-269. (b) Gerig, J. T.; McLeod, R.S. *J. Am. Chem. Soc.* **1973**, *95*, 5725-5729.

18) Stewart, J. M.; Ryan, J. W.; Brady, A. H. *J. Med. Chem*. **1974**, *17*, 537-539.

19) Berisio, R.; Vitagliano, L.; Mazzarella, L.; Zagari A. *Biopolymers* **2000**, *56*, 8.

20) Bella, J.; Eaton, M.; Brodsky, B.; Berman, H. M. *Science* **1994**, *266*, 75-81.

21) (a) Okuyama, K.; Arnott, S.; Takayanagi, M.; Kakubo, M. *J. Mol. Biol*. **1981**, *152*, 427-443. (b) Vitigliano, L.; Berisio, R.; Mazzarella, L.; Zagari, A. *Biopolymers* **2001**, *152*, 459-464. (c) DeRider, M. L.; Wilkens, S. J.; Waddell, M. J.; Bretscher, L. E.; Weinhold, F.; Raines, R. T.; Markley, J. L. *J. Chem. Soc.* **2002**, *124*, 2497-2505.

22) Shoulders, M. D.; Hodges, J. A.; Raines, R. T. *J. Am. Chem. Soc.* **2006**, *128*, 8112-8113.

23) Hodges, J. A.; Raines, R. T. *J. Am. Chem. Soc*. **2003**, *125*, 9262-9263.

24) Hodges, J. A.; Raines, R. T. *J. Am. Chem. Soc*. **2005**, *127*, 15923-15932.

25) Jenkins, C. L.; Vasbinder, M. M.; Miller, S. J.; Raines, R. T. *Org. Lett*. **2005**, *7*, 2619-2622.

26) Jenkins, C. L.; Bretscher, L. E.; Guzei, I. A.; Raines, R. T. *J. Am. Chem. Soc.* **2003**, *125*, 6422-6427.

27) Doi, M.; Nishi, Y.; Uchiyama, S.; Nishiuchi, Y.; Nakazawa, T.; Ohkubo, T.; Kobayashi, Y. *J. Am. Chem. Soc*. **2003**, *125*, 9922-9923.

28) Barth, D.;Milbradt, A. G.; Renner, C.; Moroder, L. *ChemBioChem* **2004**, *5*, 79-86.

29) Doi, M.; Nishi, Y.; Uchiyama, S.; Nishiuchi, Y.; Nakazawa, T.; Nishio, H.; Nakazava, T.; Ohkubo, T.; Kobayashi, Y. *J. Peptide Sci*. **2005**, *11*, 609-616.

30) Bann, J. G.; Bächinger, H. P. *J. Biol. Chem.* **2000**, *275*, 24466-24469.

31) Mizuno, K.; Hayashi, T.; Bächinger, H. P. *J. Biol. Chem*. **2003**, *278*, 32373-32379.

32) Mizuno, K.; Hayashi, T.; Peyton, D. H.; Bächinger, H. P. *J. Biol. Chem*. **2004**, *279*, 38072-38078.

33) Mizuno, K.; Hayashi, T.; Peyton, D. H.; Bächinger, H. P. *J. Biol. Chem*. **2004**, *279*, 282-287.

34) Berisio, R.; Granata, V.; Vitigliano, L.; Zagari, A. *J. Am. Chem. Soc*. **2004**, *126*, 11402-11403.

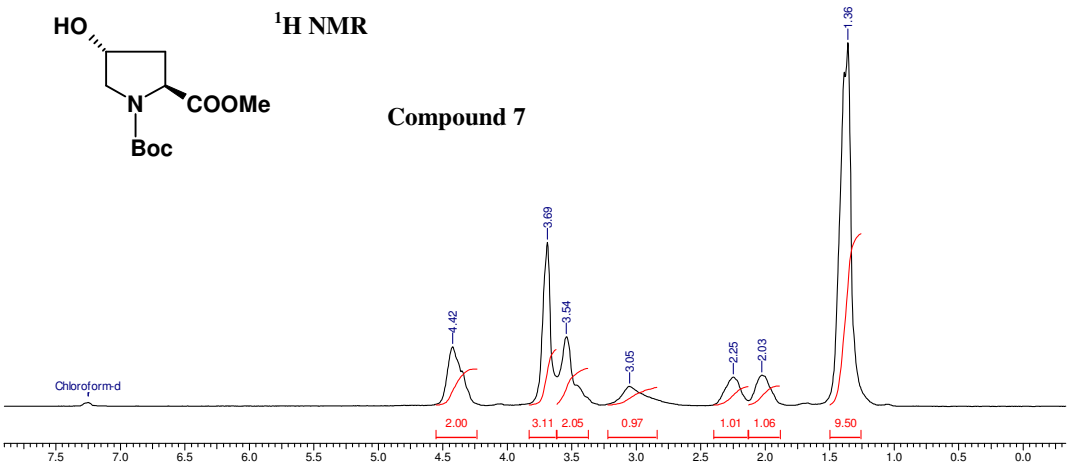
35) Schumacher, M.; Mizuno, K.; Bächinger, H. P. *J. Biol. Chem*. **2005**, *280*, 20397-20403.

36) Improta, R.; Mele, F.; Crescenzi, O.; Benzi, C.; Barone, V. *J. Am. Chem. Soc.* **2002,** *124,* 7857-7865.

Appendix II

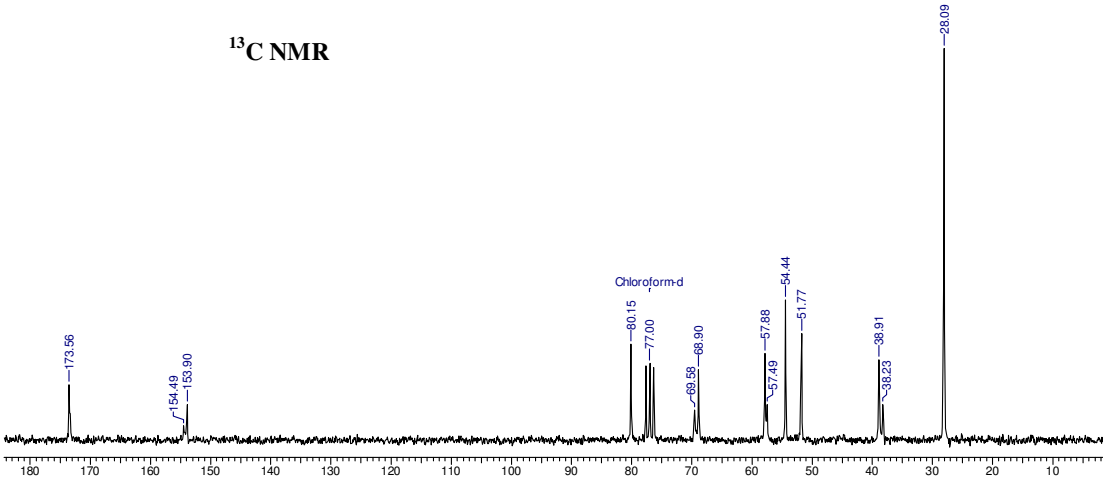
Compound No.	Spectra	Page No
Compound **7** ……	^{1}HNMR,^{13}CNMR, DEPT…….	205
Compound **8** ……	^{1}HNMR,^{13}CNMR, DEPT…….	206
Compound **9**……	^{1}HNMR,^{13}CNMR, DEPT…….	207
Compound **11**……	^{1}HNMR,^{13}CNMR, DEPT…….	208
Peptide **23**	RP-18 HPLC and MALDI-TOF ……..	209
Peptide **24**	RP-18 HPLC and MALDI-TOF ……..	210
Peptide **25**	RP-18 HPLC and MALDI-TOF ……..	211
Peptide **26**	RP-18 HPLC and MALDI-TOF ……..	212
Peptide **27**	RP-18 HPLC and MALDI-TOF ……..	213
Peptide **28**	RP-18 HPLC and MALDI-TOF ……..	214
Peptide **23**	RP-18 HPLC and MALDI-TOF ……..	215
Peptide **23**	RP-18 HPLC and MALDI-TOF ……..	216

Acquisition Time (sec)	1.1633	***Comment***	MRS. S.P. BAPAT/**SVT			***Date***	16/12/02 22:15:15
Frequency (MHz)	200.13	***Nucleus***	1H	***Number of Transients***	512	***Original Points Count*** 8192	***Points Count*** 8192
Sweep Width (Hz)	7042.25	***Temperature (grad C)***	24.000				

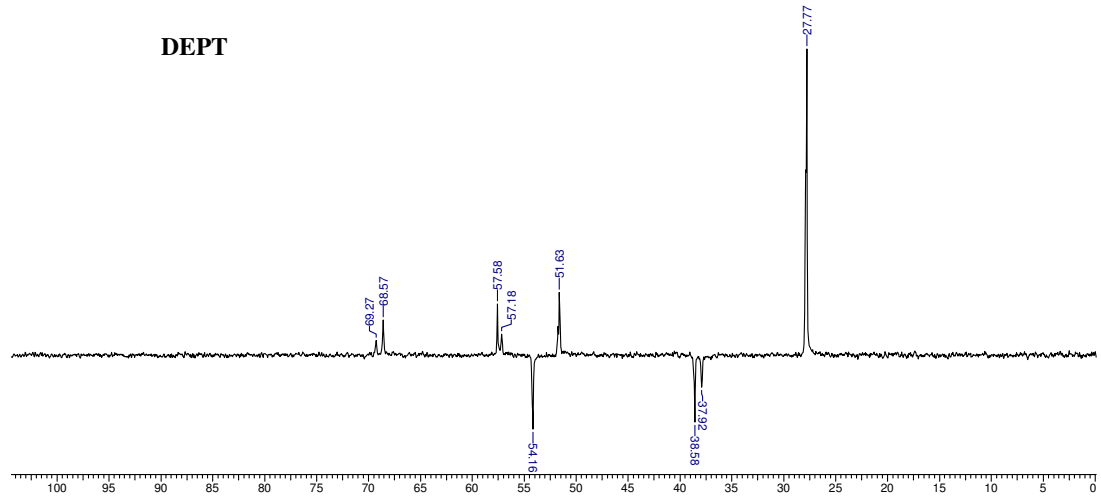


7 Jul 2006

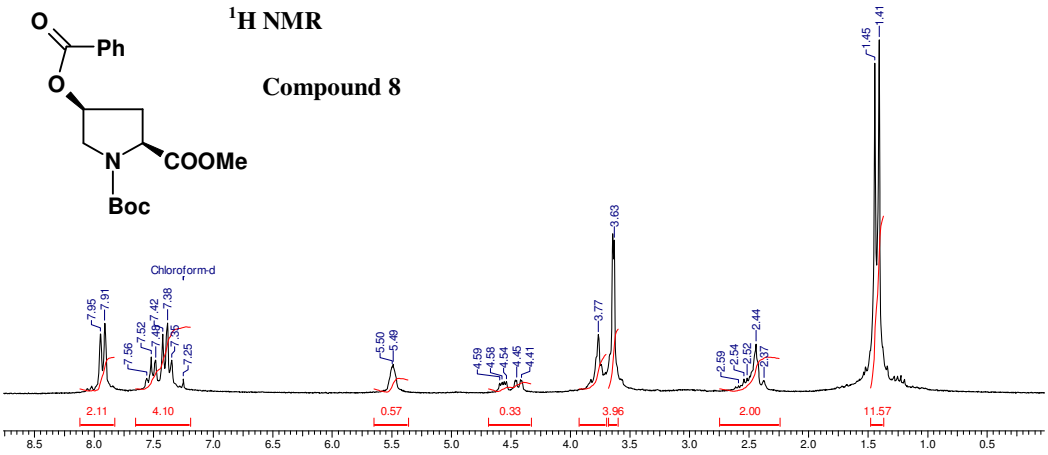
Acquisition Time (sec)	0.4096	***Comment***	DR.R.A. JOSHI/C13/ME-2/**SVT			***Date***	17/12/02 18:22:22
Frequency (MHz)	50.32	***Nucleus***	13C	***Number of Transients***	20000	***Original Points Count*** 8192	***Points Count*** 8192
Sweep Width (Hz)	20000.00	***Temperature (grad C)***	24.000				



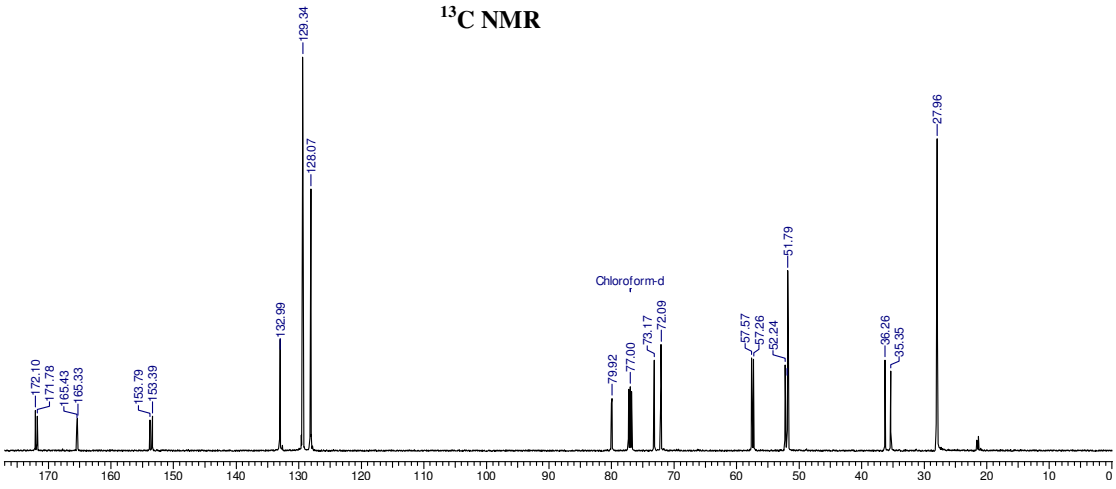
Acquisition Time (sec)	0.5407	***Comment***	D.K. MAHAPATRA/DEPT/**SVT			***Date***	17/12/02 18:29:28
Frequency (MHz)	50.32	***Nucleus***	13C	***Number of Transients***	2048	***Original Points Count*** 8192	***Points Count*** 8192
Sweep Width (Hz)	15151.52	***Temperature (grad C)***	24.000				



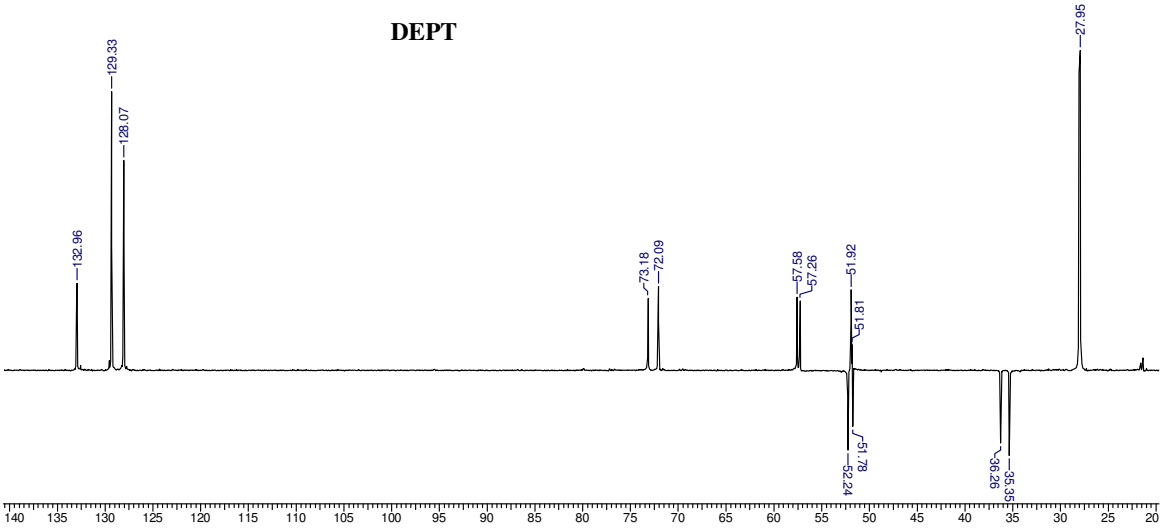
Acquisition Time (sec)	2.5559	***Comment***	UMA SHANKAR		***Date***	18/02/00 03:04:19	
Frequency (MHz)	200.13	***Nucleus***	1H	***Number of Transients*** 256	***Original Points Count***	8192	***Points Count*** 8192
Sweep Width (Hz)	3205.13	***Temperature (grad C)***	24.000				



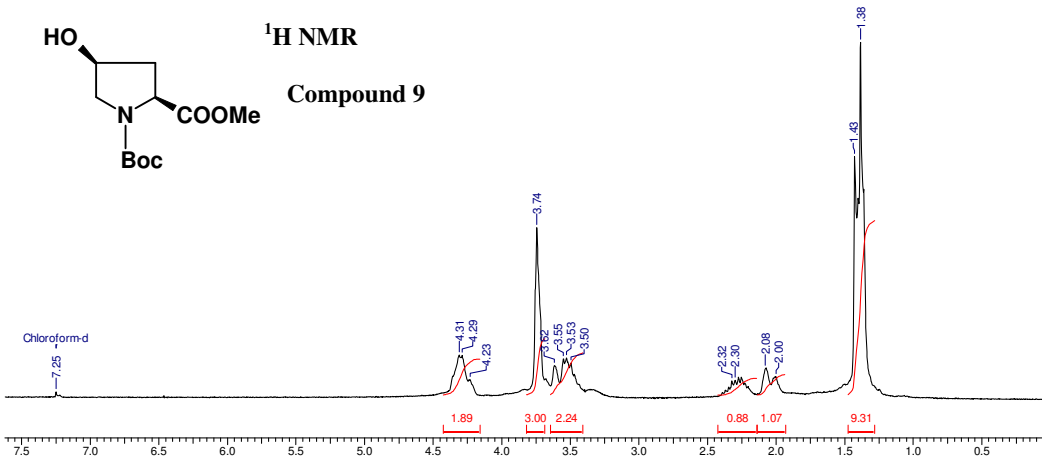
Acquisition Time (sec)	0.5210	***Comment***	Imported from UXNMR.		***Date***	21/07/2003 15:08:00	
Frequency (MHz)	125.76	***Nucleus***	13C	***Original Points Count*** 16384	***Points Count***	16384	***Sweep Width (Hz)*** 31446.54
Temperature (grad C)	0.000						



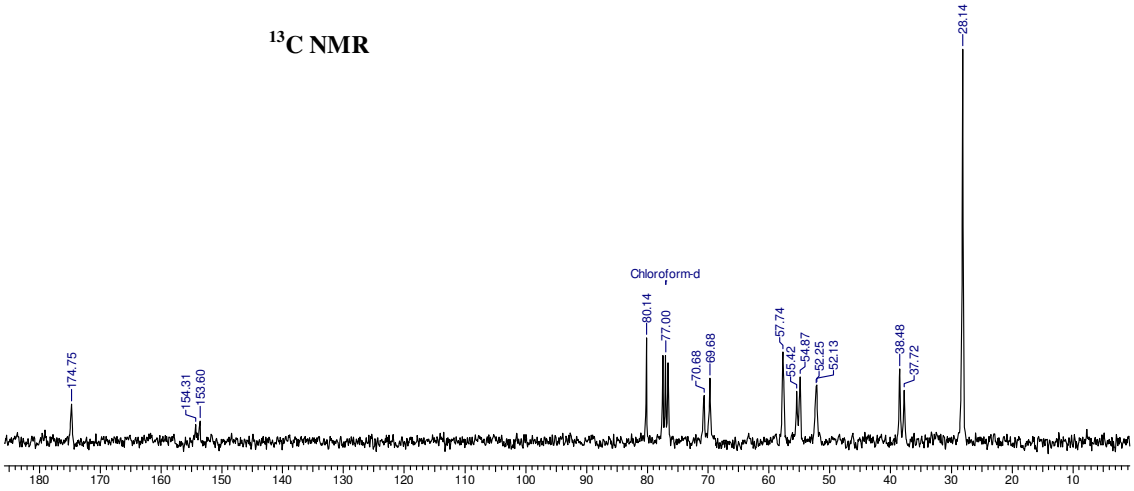
Acquisition Time (sec)	1.0420	***Comment***	Imported from UXNMR.		***Date***	21/07/2003 15:06:00	
Frequency (MHz)	125.76	***Nucleus***	13C	***Original Points Count*** 32768	***Points Count***	32768	***Sweep Width (Hz)*** 31446.54
Temperature (grad C)	0.000						



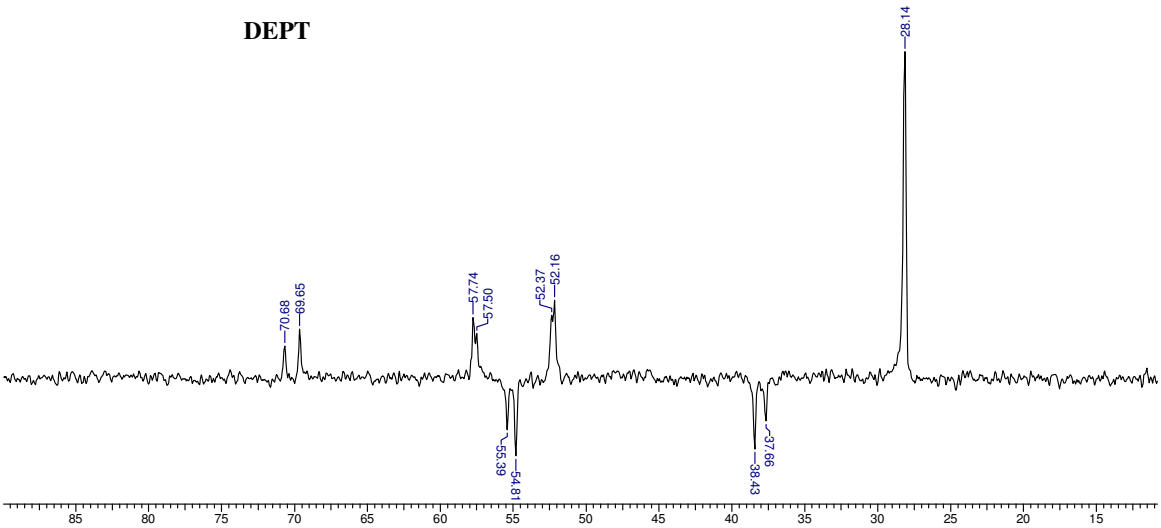
Acquisition Time (sec)	2.5559	**Comment**	SUNITA SATAV/1H/CDCL3			**Date**	06/05/00 21:24:13
Frequency (MHz)	200.13	**Nucleus**	1H	**Number of Transients**	12000	**Original Points Count**	8192	**Points Count**	8192
Sweep Width (Hz)	3205.13	**Temperature (grad C)**	24.000				



Acquisition Time (sec)	0.4342	**Comment**	DO	**Date**	00/00/00 00:00:00		**Frequency (MHz)**	75.48
Nucleus	13C	**Number of Transients**	400	**Original Points Count**	8192	**Points Count**	8192	**Sweep Width (Hz)**	18867.92
Temperature (grad C)	24.000								

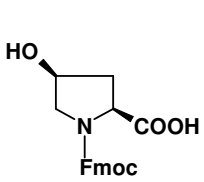


Acquisition Time (sec)	0.4342	**Comment**	DO	**Date**	00/00/00 00:00:00		**Frequency (MHz)**	75.48
Nucleus	13C	**Number of Transients**	512	**Original Points Count**	8192	**Points Count**	8192	**Sweep Width (Hz)**	18867.92
Temperature (grad C)	24.000								



Acquisition Time (sec)	7.9167	**Comment**	m umasanker		**Date**	02/01/2005 22:58:38	
Frequency (MHz)	200.13	**Nucleus**	1H	**Original Points Count** 32768	**Points Count**	32768	**Sweep Width (Hz)** 4139.07
Temperature (grad C)	0.000						

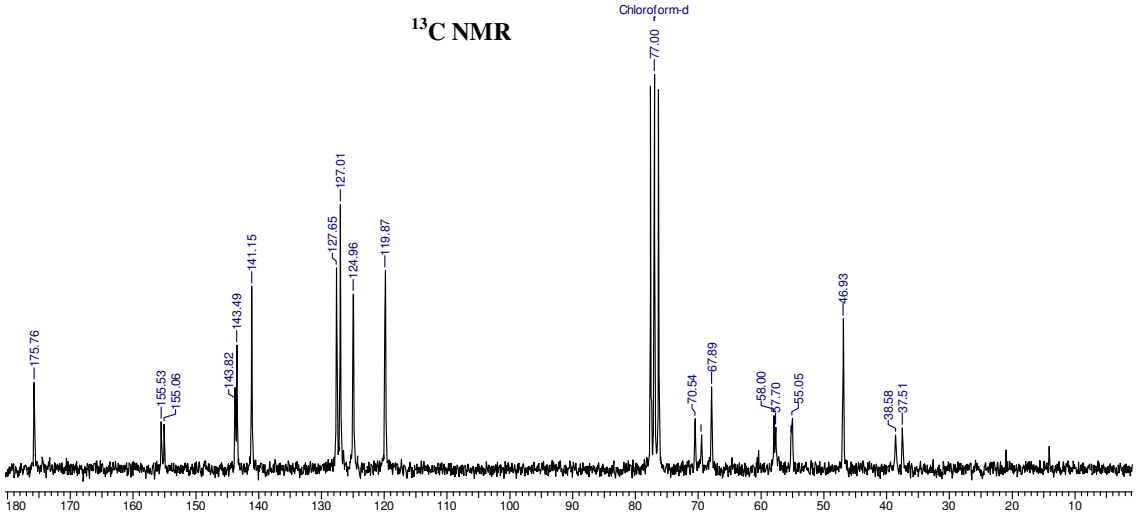
^{1}H NMR



Compound 11

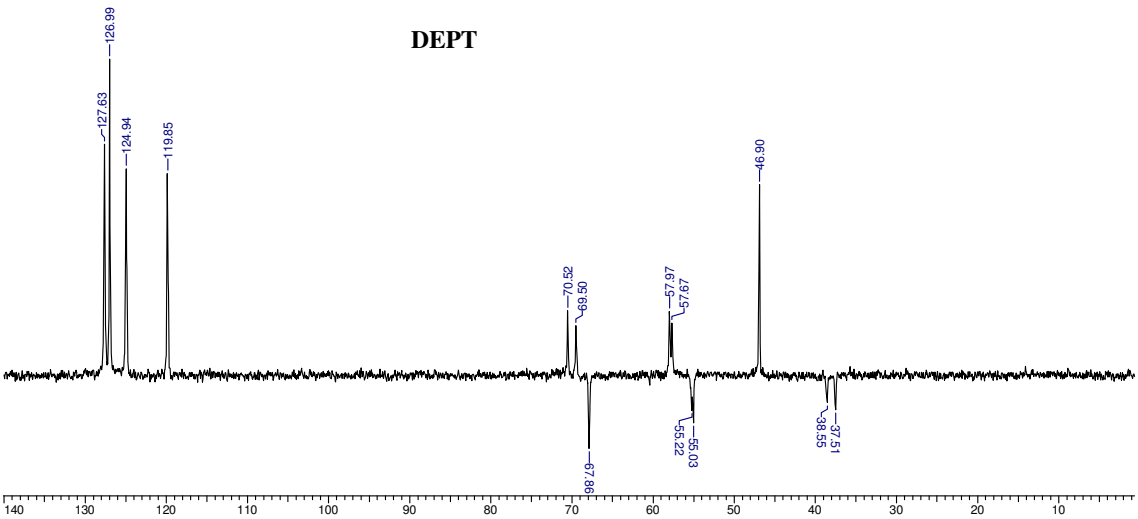
Acquisition Time (sec)	2.7329	**Comment**	m umashankara		**Date**	24/01/2005 03:27:26	
Frequency (MHz)	50.32	**Nucleus**	13C	**Original Points Count** 32768	**Points Count**	32768	**Sweep Width (Hz)** 11990.41
Temperature (grad C)	0.000						

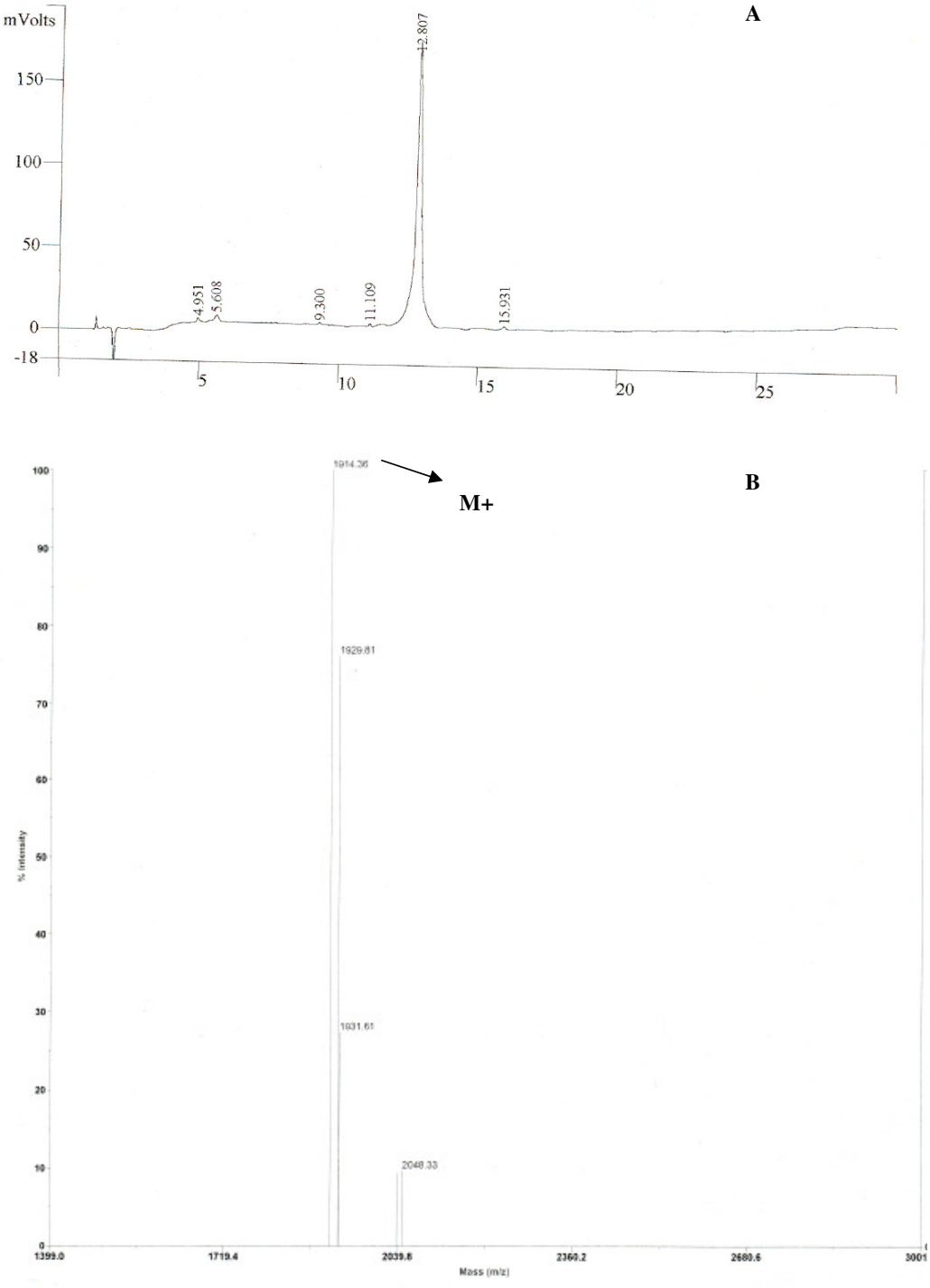
^{13}C NMR



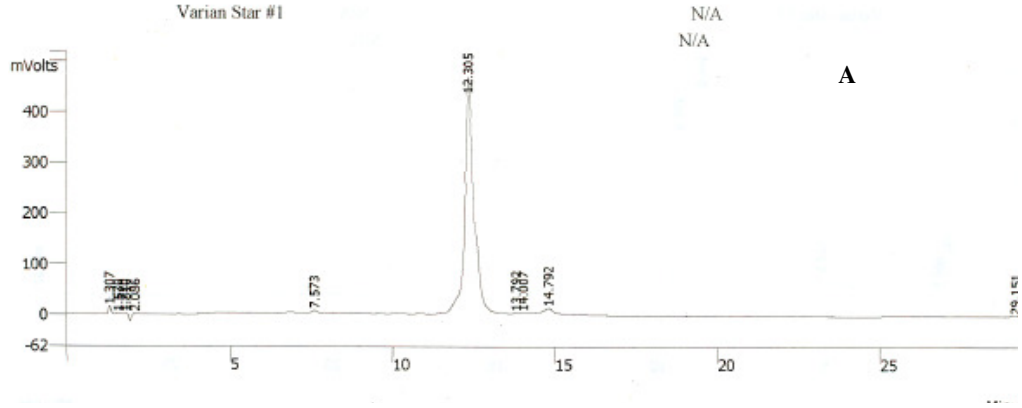
Acquisition Time (sec)	2.7329	**Comment**	m umashankara		**Date**	24/01/2005 03:55:30	
Frequency (MHz)	50.32	**Nucleus**	13C	**Original Points Count** 32768	**Points Count**	32768	**Sweep Width (Hz)** 11990.41
Temperature (grad C)	0.000						

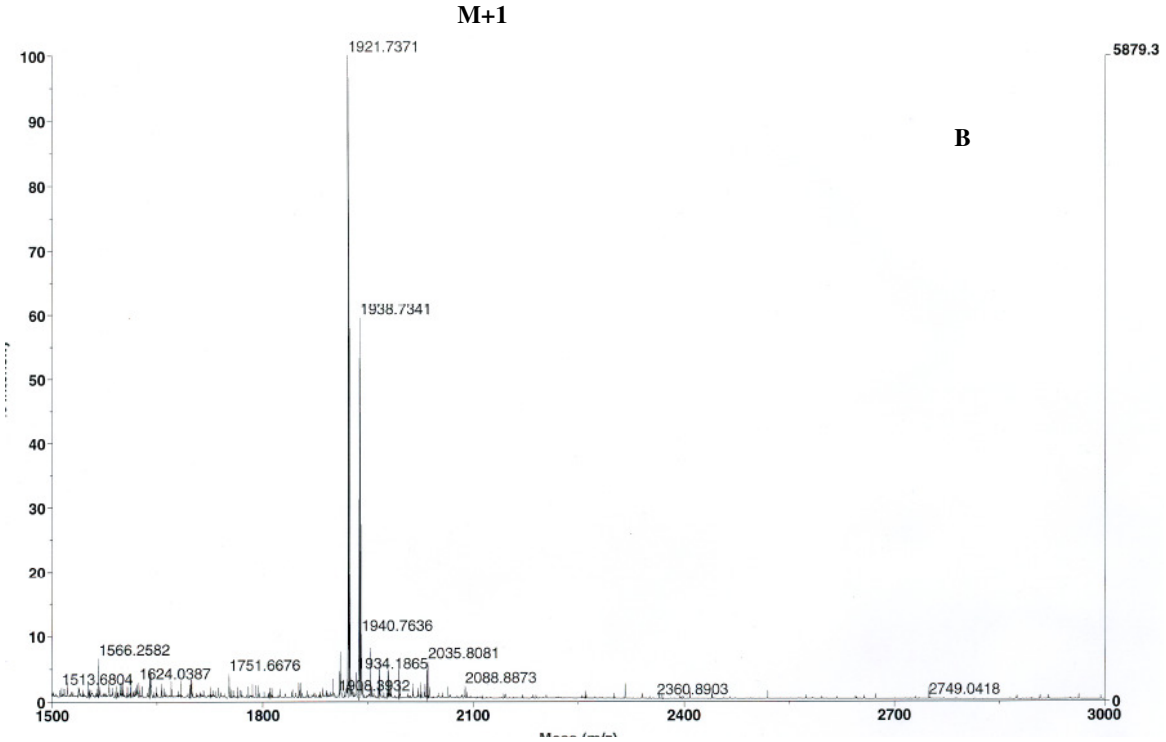
DEPT



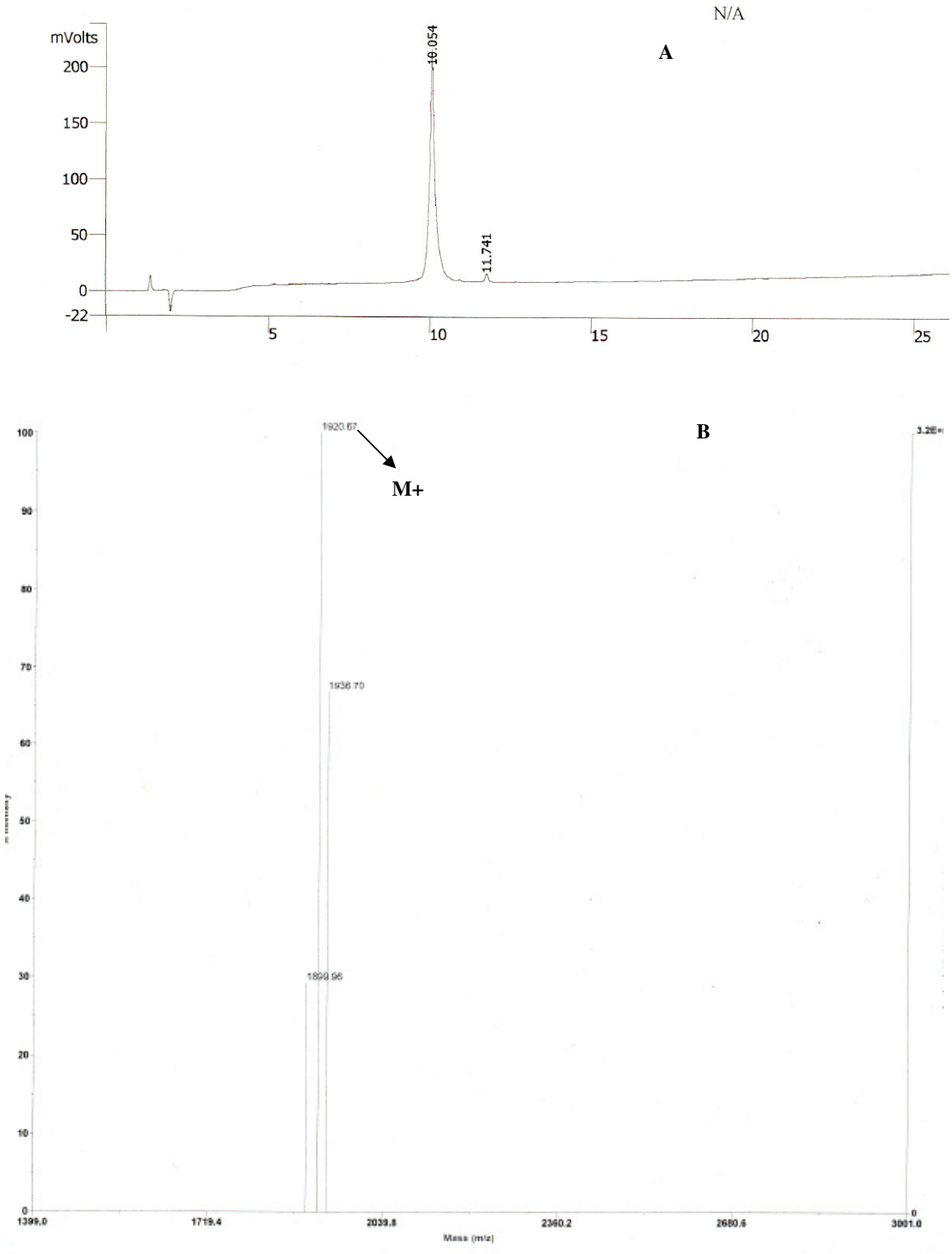


A; HPLC profile, **B;** MALDI-TOF mass spectra of peptide **23** Ac-Phe(amp-Amp-Gly)$_6$-NH$_2$

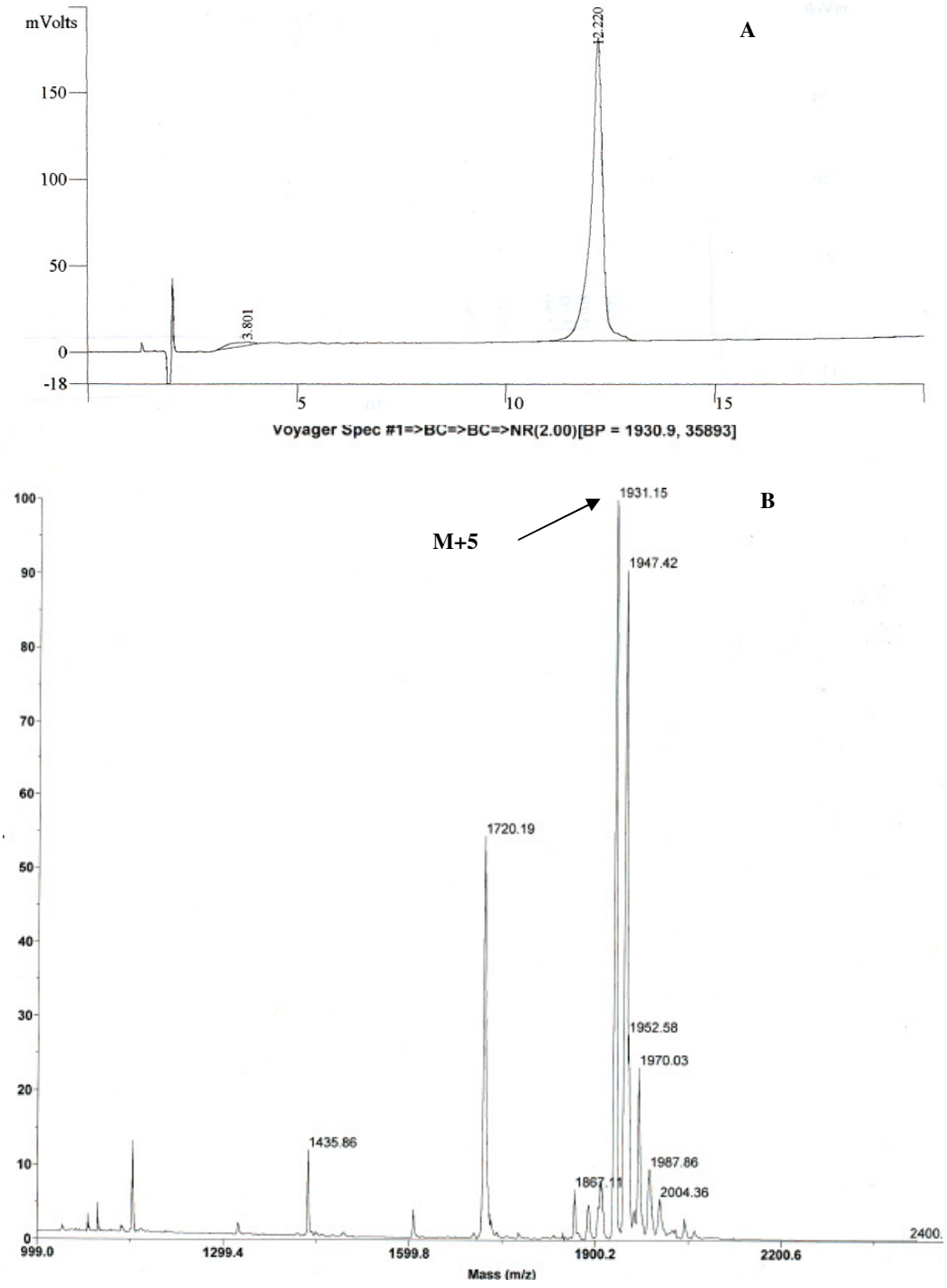




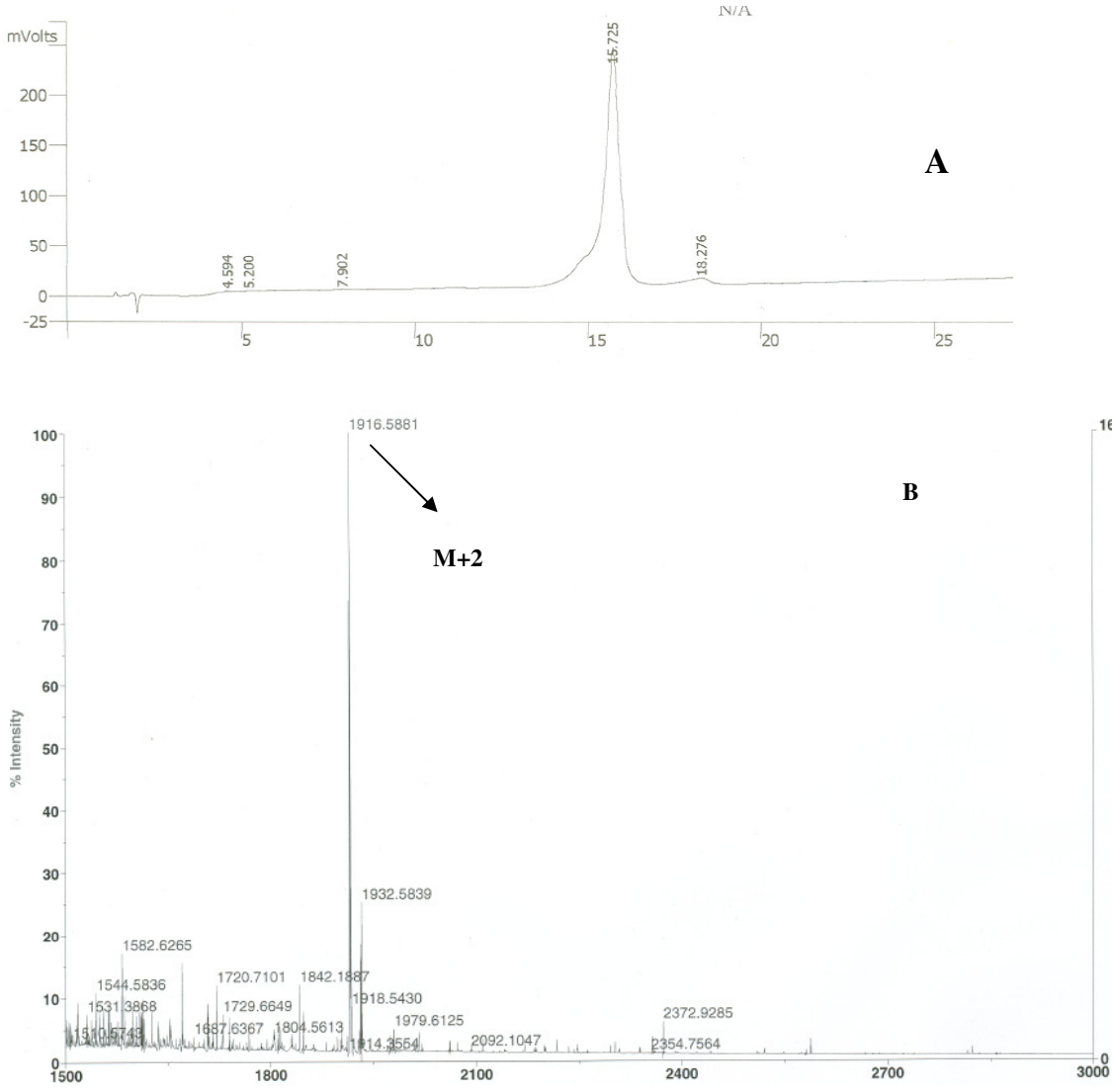
A; HPLC profile, **B** MALDI-TOF mass spectra of peptide **24;** Ac-Phe(amp-Hyp-Gly)$_6$-NH$_2$



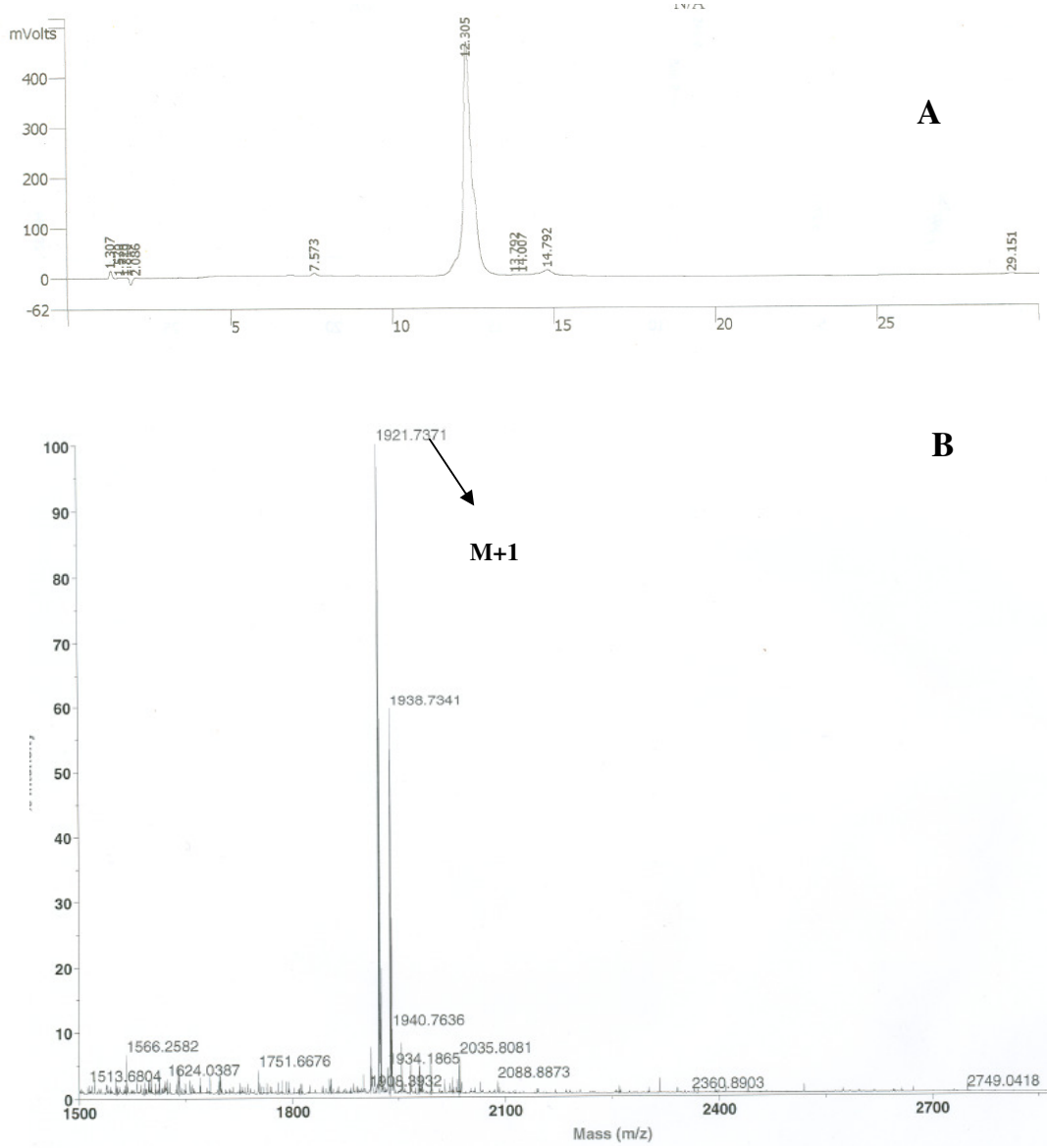
A; HPLC profile, **B;** MOLDI-TOF mass spectra of peptide **25** Ac-Phe(hyp-Amp-Gly)$_6$-NH$_2$



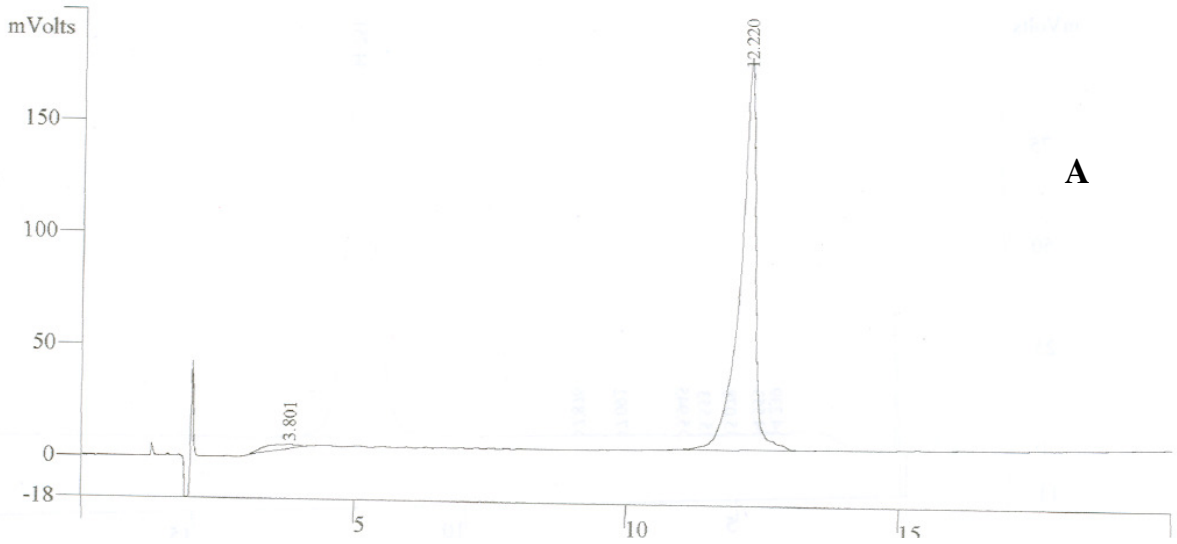
A; HPLC profile, **B;** MOLDI-TOF mass spectra of peptide **26** Ac-Phe(hyp-Hyp-Gly)$_6$-NH$_2$

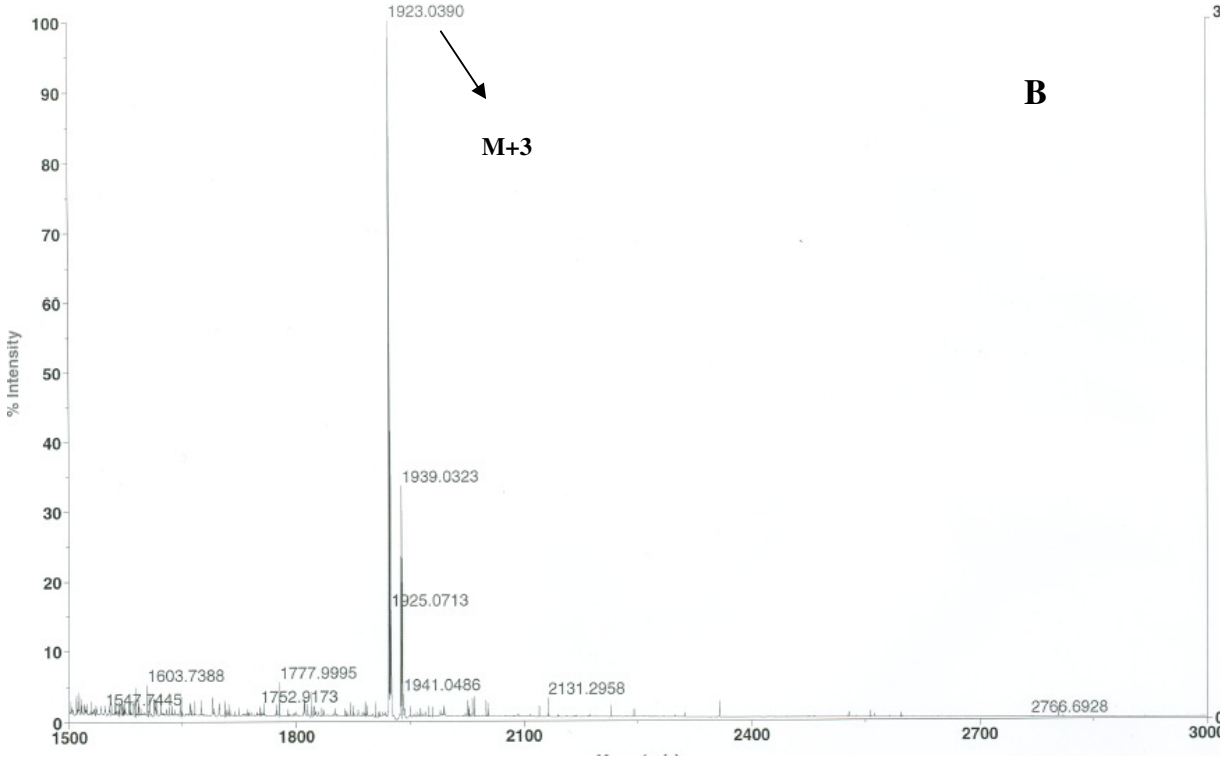


A; HPLC profile, **B;** MOLDI-TOF mass spectra of peptide **27** Ac-Phe(Amp-Amp-Gly)$_6$-NH$_2$

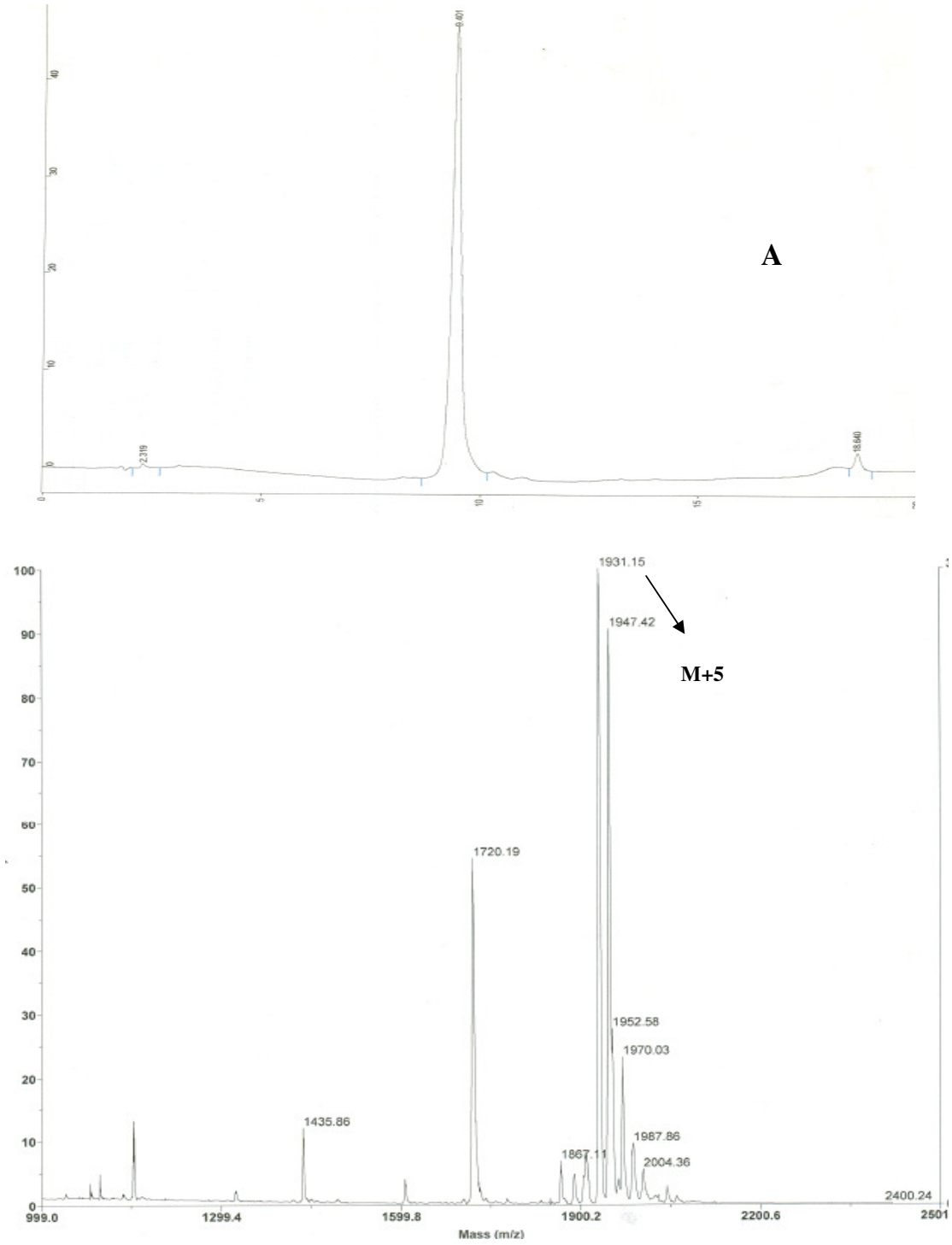


A; HPLC profile, **B;** MALDI-TOF mass spectra of peptide **28** Ac-Phe(Hyp-Amp-Gly)$_6$-NH_2





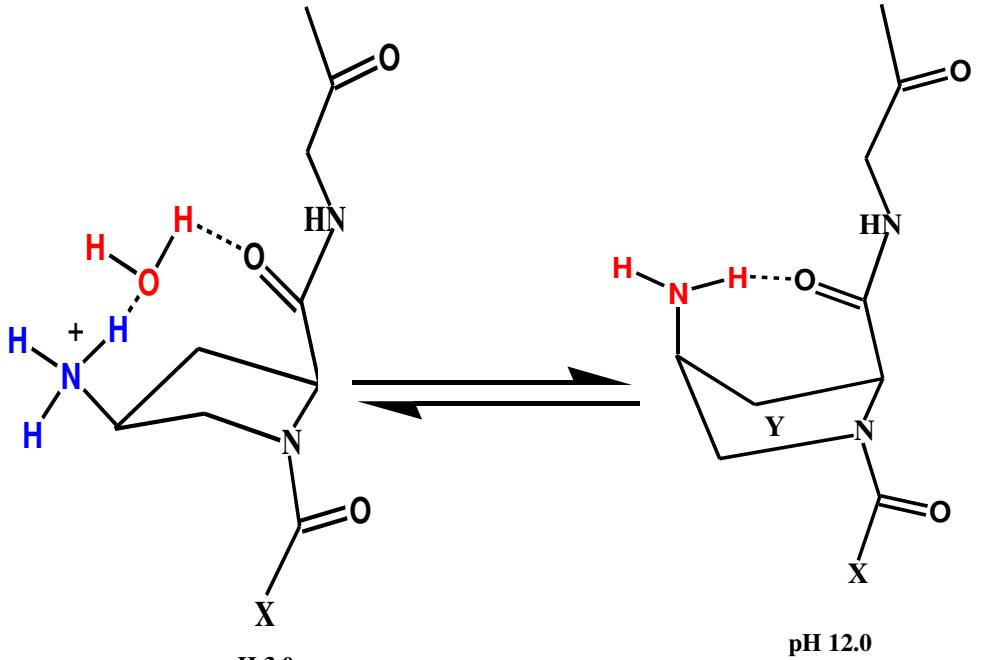
A; HPLC profile, **B;** MALDI-TOF mass spectra of peptide **29** Ac-Phe(Amp-Hyp-Gly)$_6$-NH$_2$



A; HPLC profile, **B;** MALDI-TOF mass spectra of peptide **30** Ac-Phe(Hyp-Hyp-Gly)$_6$-NH_2

Chapter 4

4-(N-Formyl)-aminoproline collagen peptides: probes for studying the role of protonated 4-amino group in triplex stability



4.1: Introduction

The previous two chapters (chapter **2** and chapter **3**) explain the formation of collagen like triple-helix structures by two diasteromers 4*R* and 4*S*-aminiprolines when present simultaneously with Pro, hyp, Hyp either in X or Y position of peptide X-Y-Gly.

Though 4-aminoprolines are compatible at both X and Y positions to form a triple-helical structure, the stability of the triple-helix containing 4-aminoprolines are highly pH dependent. For example when present at Y position, 4*R*-Amp forms a stable triple helix (Pro-Amp-Gly) in both acidic and basic conditions (pH 3.0-12.0), but while present at X position, triplex was formed at only acidic and neutral conditions (pH 3.0 and 7.0). This result suggests that the protonation of the 4-NH_2 group is a prerequisite for triplex formation in the X-position, while it is not so in the Y-position. 4*S*-amp also forms a stable triple helix at pH 3.0, 7.0 and 9.0 only at X position (amp-Pro-Gly), but in Y position 4*S*-amp (Pro-amp-Gly) declines the peptide to fold into triple-helix conformation.

4.1.1: Proline ring pucker and Amide bond geometry

As explained in introduction chapter, for simple 4-substituted proline monomer there are two energetically stable conformations namely Cγ-*exo* (up) and Cγ-*endo* (down) are possible. When 4-substitued proline is present in peptide there are 4 energetically stable conformations possible namely *trans*-up, *trans*-down, *cis*-up and *cis*-down (Figure 4.1). For unsubstituted proline both *trans*-down (Fig 4.1A) and cis-down (Fig 4.1C) are isoenergetic, but in peptide *trans*-down conformation is favored

compared to *cis*-down. The insertion of electron withdrawing 4*R*-substituent changes preferred conformation to *trans*-up (*exo* pucker) (Fig. 4.1B).

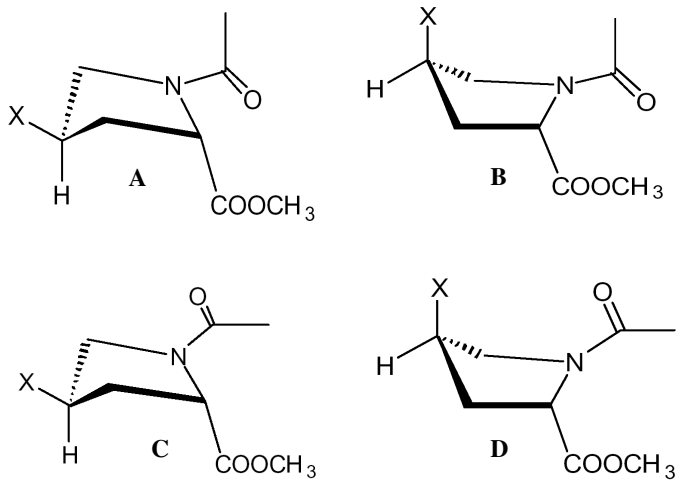


Figure 4.1: Pyrrolidine ring conformation. Illustration of the conformations adopted by pyrrolidine ring. A) *trans-down*; B) *trans*-up; C) *cis*-down; D) *cis*-up.

4.1.2 Conformational preferences by position of the pyrrolidine ring pucker.

Kollman, et. al.[1] calculated the conformational preferences by position of the pyrrolidine ring at different positions for $[(Pro\text{-}Pro\text{-}Gly)_{10}]_3$(hereafter PPG); $[(Pro\text{-}Hyp\text{-}Gly)_{10}]_3$ (PHG); $[(Pro\text{-}Flp\text{-}Gly)_{10}]_3$(PFG); $[Pro\text{-}Amp^{+}\text{-}Gly]_{10}]_3$(PMG) and $[(Pro\text{-}Amp\text{-}Gly)_{10}]_3$ (PNG) and found that the PMG peptide has the highest percentage Cγ-*exo* ring pucker (Figure 4.2) followed by the peptide PFG. Interestingly peptide PNG with neutral aminoproline at Y position has lowest percentage of Cγ-*exo* ring pucker compared to the peptide PPG. This work concluded that the protonation and nonprotonation of 4*R*-aminogroup has remarkable influence on the pyrrolidine ring conformation in the peptide $(Pro\text{-}Amp\text{-}Gly)_{10}$. They also calculated the preferred dihedral angles ψ, φ and ω in the folded trimer of PPG, PHG, PFG, PMG and PNG.

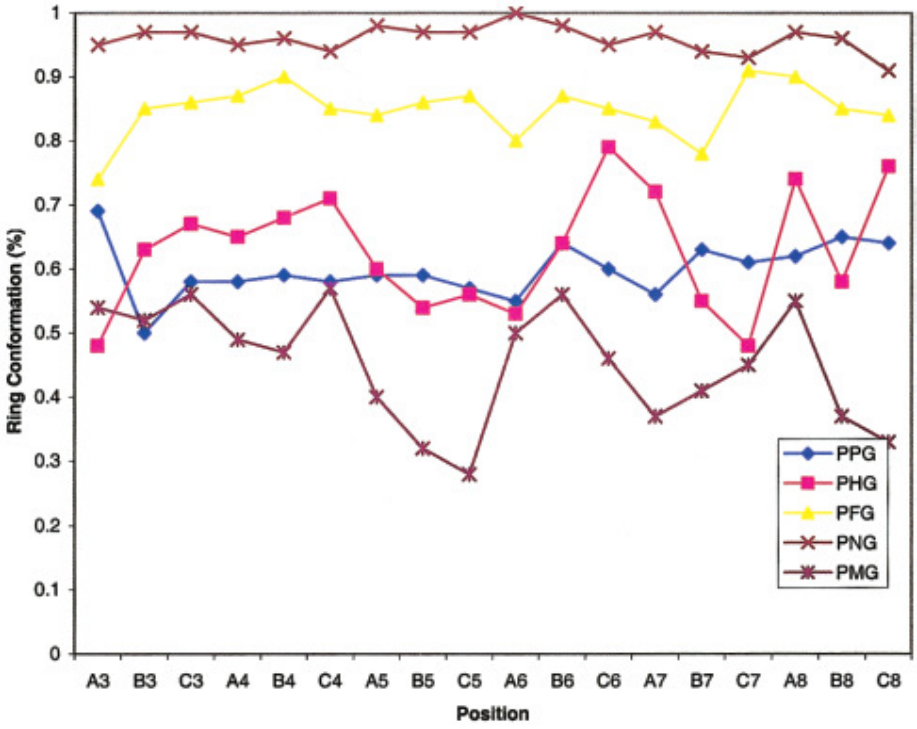


Figure 4.2: Conformational preferences by position of the pyrrolidine ring pucker. PPG is $[(Pro\text{-}Pro\text{-}Gly)_{10}]_3$; PHG is $[(Pro\text{-}Hyp\text{-}Gly)_{10}]_3$; PFG is $[(Pro\text{-}Flp\text{-}Gly)_{10}]_3$; PMG is $[Pro\text{-}Amp^{+}\text{-}Gly]_{10}]_3$; PNG is $[(Pro\text{-}Amp\text{-}Gly)_{10}]_3$. Y axis is in percent Cγ-exo and X axis is the residue number in the form of chain/triplet number (example: A3 is the third triplet in chain A).

Table 1: Average values of peptide dihedral angles

Dihedral angles	PPG	PHG	PFG	PMG	PNG	PHG X-Ray[3]	Collagen Ideal[5]
X φ	-70.13	-69.69	-69-64	-66.87	-69.60	-75.0	-72.1
X ψ	161.18	163.06	162.23	163.70	162.87	161.4	164.3
X ω	174.18	174.71	173.59	177.4	174.85	177.8	180.0
Y φ	-58.51	-58.34	-55.91	-59.33	-60.82	-61	-75
Y ψ	152.46	153.7	151.83	159.95	1753.94	153.3	155.8
Y ω	179.42	179.02	178.55	175.7	179.3	176.7	180
G φ	-74.60	-75.95	-75.16	-79.15	-76.45	-75.8	-67.6
G ψ	176.11	177.6	178.02	160.36	175.89	179.9	151.4
G ω	177.78	177.24	177.39	176.87	177.74	179.5	180

The values of φ at Y position obtained by simulation study for the trimers $(Pro\text{-}Pro\text{-}Gly)_{10}$ and $(Pro\text{-}Hyp\text{-}Gly)_{10}$ are in good agreement with the values reported from

crystallographic[2, 3] and NMR[4]studies. Surprisingly Amp^+ and Amp in the trimer peptide Pro-Y-Gly show slightly different ψ, φ and ω dihedral angles (Table-1). In a folded triple-helix, Cγ-*exo* conformation is preferred in proline, hydroxyproline, fluoroproline and aminoproline. Fluoroproline and charged aminoproline show a stronger gauche preference over both hydroxyproline and proline and have the greatest Cγ-exo preference, while hydroxyproline and uncharged aminoproline have less Cγ-exo preference. It is interesting to note that both the uncharged substituents with hydrogen-bonding potential (-OH and $–NH_2$) exhibit a wider range of observed conformations than any others (Fig 4.2). This may be due to environmental differences, solvent structure differences or neighbouring residue interaction differences.

4.1.3: Solvation of proline residue in triple-helix

The importance of solvent molecules is a matter of considerable debate within the collagen community. Many research groups have reported the presence of solvent molecules inside the triple-helical structure[6-9]. Kollman, et al.[1] have calculated the solvent/solute hydrogen bonds that are rapidly in exchange with each other in residues of trimers PPG, PHG, PFG, PMG and PNG. The fluoroproline shows the smallest number of water molecules per Y position residues and the charged aminoproline shows the largest number of solvent molecules (Table-2). This study indicates that the substituent effects by the F, OH and NH_2 substituents are site localized with the peptide while the effect of NH_3^+ is long range in nature.

Table 2: Preferred pucker conformation[a] (in percent Cγ-exo)

	PPG	PHG	PFG	PMG	PNG
Default parameter	61.9(4.8)	50.4(9.1)	68.8(7.7)	N/A	N/A
New dihedral parameters	59.8(4.3)	62.8(9.4)	84.8(4.2)	95.9(2.1)	45.3(9.2)
Three residue Folded	N/A	54.3(3.8)	84.0(7.5)	92.0(5.2)	31.0(22.3)

[a] Values are given in percent Cγ-exo. Standard deviations from each position are in parentheses.

4.2: Rationale and Objective of the present work.

Though 4*R*-aminoproline is compatible at both X and Y positions, the stability of the triple-helix of collagen peptide containing 4*R*-aminoproline is highly pH dependent. When present at Y position, 4*R*-Amp forms a stable triple helix in entire pH range, while at X position triplex is formed at only acidic pH. In comparison, 4*S*-amp also forms a stable triple helix only at acidic pH but not associated into triplex at pH 12. These results suggest that the protonation of the NH_2 group 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 pyrolidine ring is dependent on the protonation status of the 4-amino group and also influences the *cis-trans* amide rotameric equilibrium. This chapter examines the possible effect of protonation of 4-NH_2 group on the formation and stability of collagen triple helix. The 4-NH_2 group of 4-aminoproline was protected as amide via formyl derivative, which cannot be protonated easily. The 4-formyl protection group is stable at all pH conditions. Formyl group being small, any sterric interference of this group is only minimal during triplex formation with other strands.

The specific objectives of this chapter are

1) Synthesis of N^1-Fmoc-N^4-CHO (2*S*, 4*R*)-aminoproline and N^1-Fmoc-N^4-CHO (2*S*, 4*S*)-aminoproline.

2) Solid phase synthesis of collagen peptides by incorporating above monomers using Fmoc chemistry.

3) Cleavage of the peptide from the resin followed by purification and characterization of peptides.

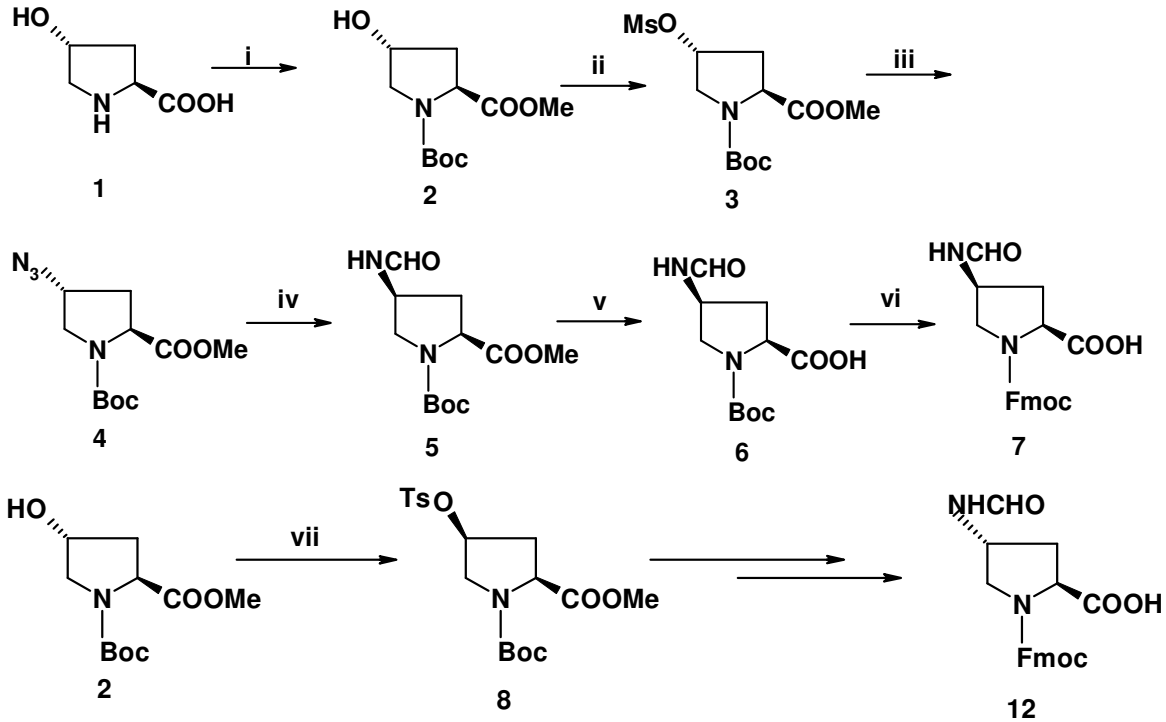
4) Study of triple-helical stability of peptides using CD spectrometry at different pH conditions.

4.3: Results

4.3.1: Synthesis of (2*S*,4*R*) and (2*S*,4*S*)-4N-formyl-amnioprolines

The N^1-Fmoc-N^4-formyl-4*R*-aminoproline (**12**), and N^1-Fmoc-N^4-formyl-4*S*-aminoproline (**7**) were synthesized from 4*R*-hydroxyproline in eight steps by sequential protection/deprotection strategies(Scheme 4.1)**.**

Scheme-4.1



i) a)Boc-anhydride, NaOH, Dioxane:Water; b) DMS, K_2CO_3, acetone; ii) Mesylchloride, Et_3N, DCM; iii) NaN_3, DMF, iv) a) H_2 Pd/C, MeOH, b) methylformate; v) LiOH, H_2O: MeOH; vi) a)TFA: DCM, b) Na_2CO_3, Fmoc-Cl, Dioxane: Water. Vii) PPh_3, DIAD, P-Ts-OMe, THF

The pyrrolidine ring nitrogen atom of 4*R*-hydroxyproline **1** was protected as N-Boc derivative using Boc-anhydride in presence of NaOH in dioxan water mixture. The carboxyl group of N-Boc-4*R*-hydroxyproline was protected as methyl ester **2** using dimethyl sulphate and K_2CO_3 in anhydrous acetone. The 4-OH group of methylester **2** was then converted to corresponding 4-O-mesyl derivative **3** using mesyl chloride in DCM in presence of triethylamine. The reaction of mesyl compound **3** with NaN_3 in DMF resulted in a S_N2 inversion of stereochemistry at C4-center, yielding the 4*S*-azide compound **4.** Catalytic hydrogenation of **4** with H_2-Pd/C, followed by protection of 4-NH_2 to formyl drerivative with methylformate yielded compound **5**, the methyl ester was hydrolyzed using 2N LiOH in MeOH to get compound **6**. Deprotection of Boc group with 50%TFA in dichloromethane, and then which was then Fmoc protected to get N^1-Fmoc-N^4-Formyl-4*S*-aminoproline **7** using Fmoc-Cl, Na_2CO_3, and dioxan/water (Scheme 4.1).

The 4-OH group of **2** was converted to 4*S*-*O*-tosylate **8** using Mitsunobu procedure with DEAD/PPh_3 and methyl-*p*-toluenesulfonate as nucleophile, resulting in an inversion of configuration at the C-4 center. Reacting the tosylate **8** under S_N2 conditions with NaN_3 in DMF at 55 °C resulted in a second inversion of stereochemistry at 4-center, yielding azide compound which was converted to N^1-Fmoc-N^4-Formyl-4*R*-aminoproline (**12**), using the same sequential step followed for compound **8.**

4.3.2: Solid phase synthesis of peptides 31-34

The following peptides were synthesized manually for the present study by incorporation of above synthesized monomers **7** and **12** at suitable position of the

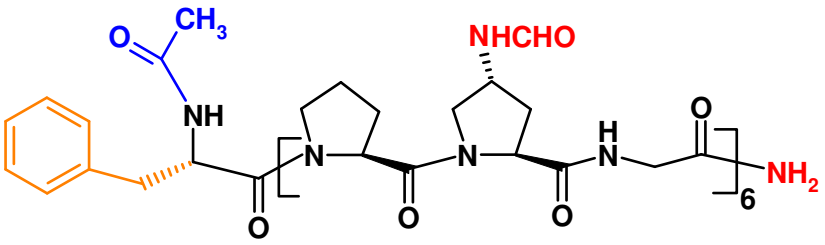
collagen peptide sequence $(X\text{-}Y\text{-}Gly)_n$ using standard Fmoc chemistry by solid phase peptide synthesis.

The monomer **12** (4*R*-fAmp) was incorporated into peptide **31** Ac-Phe(Pro-fAmp-Gly)$_6$-NH_2 in Y position and peptide **32** Ac-Phe(fAmp-Pro-Gly)$_6$-NH_2 in X position. The monomer **7** (4*S*-amp) is incorporated into peptide **33** Ac-Phe(famp-Pro-Gly)$_6$-NH_2 at X- position. In the peptide **34** Ac-Phe(famp-Amp-Gly)$_6$-NH_2, monomer **7** (4*S*-famp) was introduced at X position and the (4*R*-Amp) with free exo-cyclic amino group is introduced at Y position (Fig 4.3) as a control peptide to understand the effect of two charged aminoproline present side by side of the collagen peptide sequence.

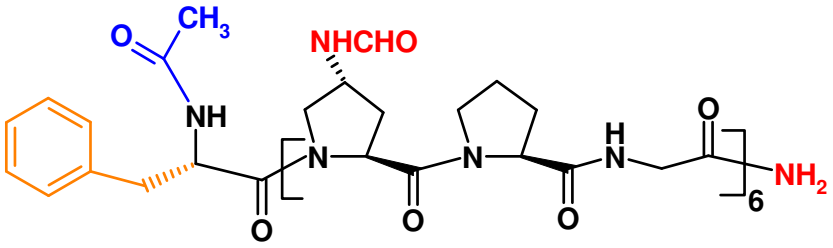
The peptides were cleaved form the resin using 5% TFA in DCM with 0.1% TIS (triisopropylsilane) as a scavenger. The peptides were purified by reverse phase HPLC on a Pharmacia Pro-RPC C8 column with water-acetonitrile gradient using 0.1% TFA as the organic modifier. The purity of final peptides as ascertained by RP-C18 HPLC were found to be greater than 97%. The structural integrity of the peptides was further confirmed by MALDI-TOF spectrometry which agreed closely with the calculated values (Table 3).

Table 3: Calculated and observed (MOLDI-TOF) masses of peptide **31-34**

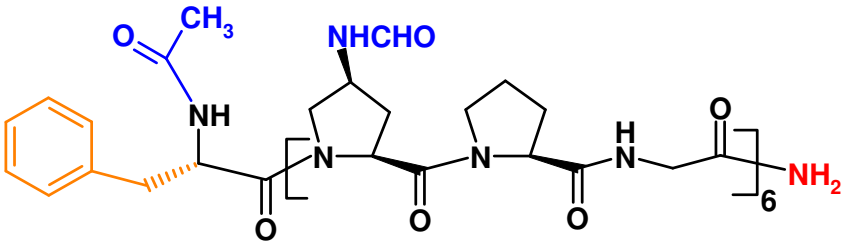
Peptide	Mol. formula	Mass(calculated)	Mass(observed)
Ac-Phe(Pro-fAmp-Gly)$_6$-NH_2	$C_{89}H_{130}N_{26}O_{26}$	1978.13	1981.53(M+3)
Ac-Phe(fAmp-Pro-Gly)$_6$-NH_2	$C_{89}H_{130}N_{26}O_{26}$	1978.13	1981.17(M+3)
Ac-Phe(famp-Pro-Gly)$_6$-NH_2	$C_{89}H_{130}N_{26}O_{26}$	1978.13	1981.26(M+3)
Ac-Phe(famp-Amp-Gly)$_6$-NH_2	$C_{89}H_{156}N_{32}O_{26}$	2088.27	2320.51(M+4+ $2CF_3COO^-$



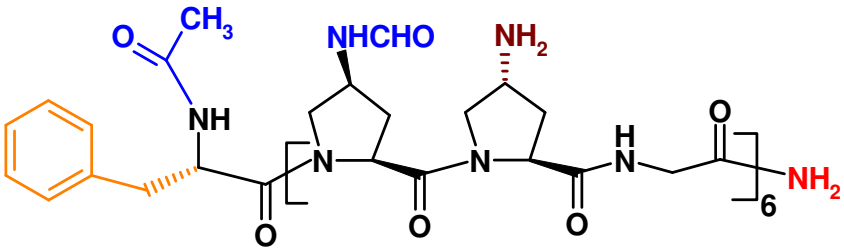
Peptide 31 (Pro-fAmp-Gly)



Peptide 32 (fAmp-Pro-Gly)



Peptide 33 (famp-Pro-Gly)



Peptide 34 (famp-Amp-Gly)

Figure 4.3: Structure of peptides **31-34** used for present study

4.4: Biophysical Studies

4.4.1: Characterization of triple-helical structure by CD spectroscopy

Natural collagen has a unique CD spectrum in which a small positive peak appears at about 220nm, a crossover near 213nm and a large trough at approximately 197nm.[10-12] This feature has been used as a reference to determine the presence of triple-helical like structure in synthetic peptides **31-34**.

4.4.2: Sample preparation

Peptide solutions (1ml of 0.2mM) at different pH (acidic; 20mM acetate buffer, neutral; 20mM phosphate buffer, Basic; 20mM borate buffer) containing 0.1M NaCl were prepared. After keeping a stock solution of a peptide at 85^{o}C for 1hr to complete dissociation in to single chains, the solution was gradually cooled to room temperature and equilibrated at 4^{o}C for 24hrs so that the triple-helices would form properly. The precise concentration of the peptides in solution was determined by recording UV absorbance of the N-terminal phenylalanine amino acid residue at 259nm (ε=200 $M^{-1}cm^{-1}$) as described in Chapter **2**.

CD measurements were carried out on JASCO J-715 spectropolarimeter using cylindrical jacketed cell of 1mm path length, which was connected to a Julabo-UC water bath circulator. Prior to spectroscopic analysis, annealed samples were allowed to incubate for ½ hr.

Figure 4.4 shows the CD spectra of 0.2mM solutions of peptide **31** taken in acidic (pH 3.0; 20mM acetate buffer) neutral (pH 7.0; 20mM phosphate buffer) pH 9.0 and 12.0 (both with 20mM borate buffer) conditions.

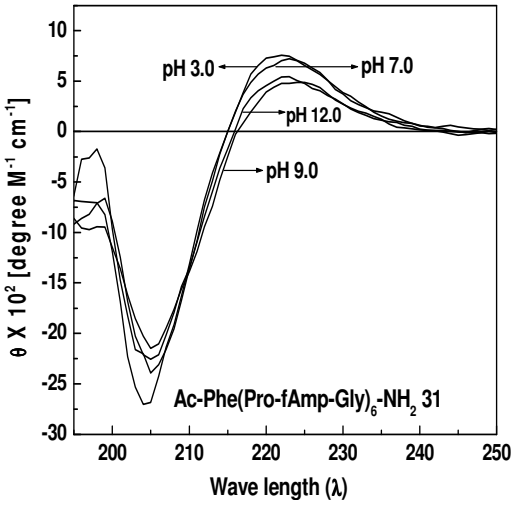


Figure 4.4: CD spectra of peptide **31** Ac-Phe(Pro-fAmp-Gly)$_6$-NH_2 taken at different pH conditions.

In both alkaline an shows the CD t pH for peptide **31** are listed in Table 4. The Rpn values are in the range of triple-helical conformation at both alkaline and acidic pH conditions.

Table 4: Rpn values of peptide **31** (Pro-fAmp-Gly) at different pH conditions.

Figure 4.5 shows the thermal denaturation curves of peptide **31** (Pro-fAmp-Gly) measured

Peptide 31 Ac-Phe(Pro-fAmp-Gly)$_6$-NH_2			
pH	**+ve band(nm)**	**-ve band (nm)**	**Rpn value**
3.0	223	205	0.27
7.0	223	205	0.24
9.0	223	203	0.24
12.0	223	205	0.25

monitoring the molar ellipticity at 225nm against temperature.

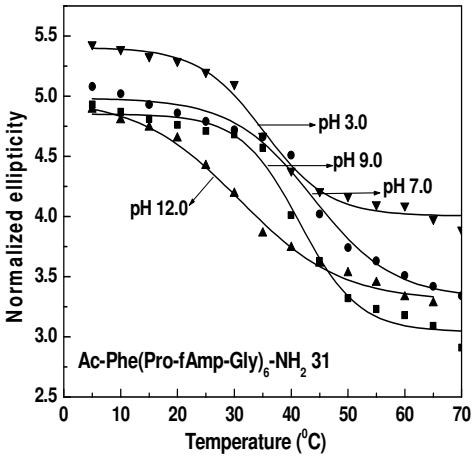


Figure 4.5: Thermal denaturation curves of peptide **31** Ac-Phe(Pro-fAmp-Gly)$_6$-NH$_2$ at different pH conditions. pH 3.0 (20mM acetate buffer), pH 9.0 (20mM phosphate buffer, pH 9 and 12 (20mM borate buffer). All buffers contain 0.1M NaCl.

Table 5: CD-T_m values of peptide **31** (Pro-fAmp-Gly) at different pH conditions.

Peptide 31 Ac-Phe(Pro-fAmp-Gly)$_6$-NH$_2$	
pH	**T_m ^{o}C**
3.0	42
7.0	41
9.0	40
12.0	40.5

The peptide **31** (amp-Amp-Gly) shows sigmoidal transition curves under all pH conditions, indicating that this peptide undergoes triple-helical to coil structure under all pH conditions. The T_m values obtained form first derivative of thermal denaturation curves (Table 5) almost invariable with pH.

The CD spectra of peptide **32** (fAmp-Pro) in all pH conditions exhibit profiles that are disordered as seen by the shallow trough and lack of positive peak (Fig.4.6) and which are not comparable to known collagen mimetic triple-helical conformations exhibiting natural collagen

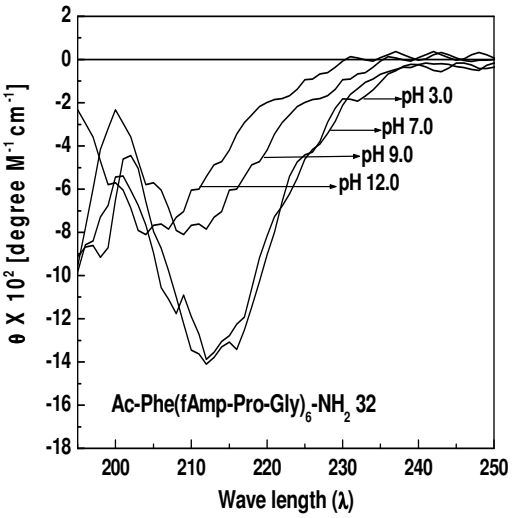


Figure 4.6: CD spectra of peptide **32** Ac-Phe(fAmp-Pro-Gly)$_6$-NH$_2$ taken at different pH conditions.

The CD spectra for peptide **33** (famp-Pro-Gly) under all pH conditions are indicative of triple-helical structures, where the crossovers are near 215nm, the positive peaks are at 225nm, and the trough around 205nm (Fig 4.7). The R_{pn} values[13,14] are listed in Table 6, which are comparable to known collagen mimetic triple-helical conformations exhibiting natural collagen and (Pro-Hyp-Gly)$_{10}$-NH$_2$.[15]

Table 6: Rpn values of peptide **33** (famp- Pro-Gly) at different pH conditions.

Peptide 33 Ac-Phe(famp- Pro-Gly)$_6$-NH$_2$			
pH	**+ve band(nm)**	**-ve band (nm)**	**Rpn value**
3.0	223	205	0.23
7.0	223	205	0.23
9.0	223	205	0.21
12.0	223	205	0.22

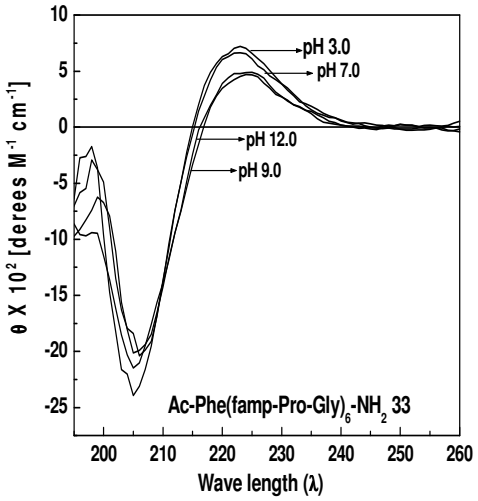


Figure 4.7: CD spectra of peptide **31** Ac-Phe(Pro-fAmp-Gly)$_6$-NH$_2$ taken at different pH conditions

Figure 4.8 shows the thermal denaturation curves of peptide **33** (famp-Pro-Gly) measured by plotting normalized ellipticity data at 225nm against temperature. Peptide **33** (famp-Pro-Gly) shows sigmoidal transition curves in all pH conditions, indicating that this peptide undergoes triple-helical to coil structure under all pH conditions. The T_m values obtained form first derivative of thermal denaturation curves (Table 7) is almost constant under all pH conditions. This indicates that the triplex stability of the peptide **33** is independent of pH conditions.

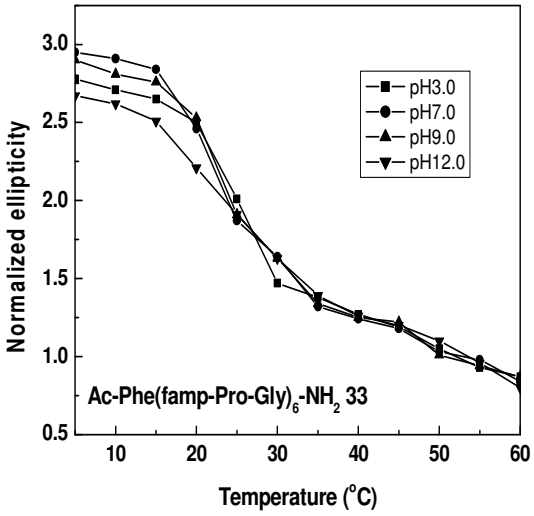


Figure 4.8: Thermal denaturation curves of peptide **33** Ac-Phe(famp- Pro-Gly)$_6$-NH$_2$ at different pH conditions. pH 3.0 (20mM acetate buffer), pH 9.0 (20mM phosphate buffer, pH 9 and 12 (20mM borate buffer). All buffers contain 0.1M NaCl.

Table 7: CD-T_m values of peptide **33** (famp- Pro-Gly) at different pH conditions.

Peptide 33 Ac-Phe(famp- Pro-Gly)$_6$-NH$_2$	
pH	**T_m ^oC**
3.0	23.2
7.0	21
9.0	21
12.0	20

Figure 4.9: shows the CD spectra of 0.2mM solution of peptides **34** (famp-Amp-Gly) in acidic (20mM acetate buffer pH 3.0) neutral (20mM phosphate buffer pH 7.0) and basic (20mM borate buffer pH 9.0 and pH 12) conditions in presence of 0.1M NaCl.

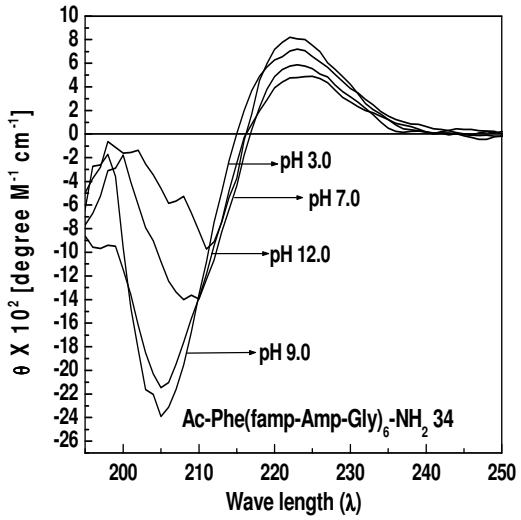


Figure 4.9: CD spectra of peptide **34** Ac-Phe(famp-Amp-Gly)$_6$-NH$_2$ taken at different pH conditions

The CD spectra of peptide **34** (famp-Amp-Gly) at both acidic and alkaline conditions resembles the CD spectra of native collagen. The Rpn values in both acidic and basic conditions are in the range of triple-helical conformation, though higher Rpn value is observed in acidic and neutral conditions compared to alkaline condition (Table 8). The CD spectrum at pH 9.0 resembles to the CD spectra of collagen peptide with respect to the appearance of positive maxima, negative minima and crossover point. But

the R_{pn} value, unlike at acidic and neutral conditions is not the range of triple-helical conformation.

Table 8: Rpn values of peptide **34** (famp-Amp-Gly) at different pH.

Peptide 34 Ac-Phe(famp-Amp-Gly)$_6$-NH$_2$			
pH	**+ve band (nm)**	**-ve band (nm)**	**Rpn**
3.0	223	206	0.37
7.0	223	206	0.31
9.0	223	204	0.22
12.0	223	204	0.26

In addition to CD spectroscopy, T_m measurements were also carried out to determine the triple-helical stability (Table 9) of the synthetic collagen-based peptide **34**. (famp-Amp-Gly). The temperature was varied in step of 2-5^oC and the spectra were recorded at each temperature. An equilibration period of 10min was allowed at each temperature. The melting transitions are consistent with the observations of CD spectroscopy (Table 9). The thermal denaturation curves show a decrease in molar ellipticity with increase in temperature under all pH conditions. Normalized ellipticity data at 225nm plotted against temperature is shown in Figure 4.10**.**

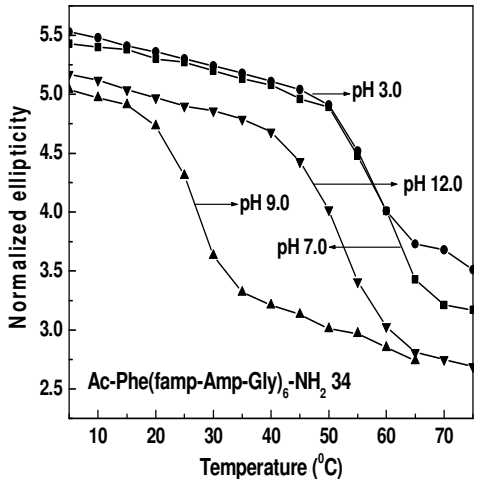


Figure 4.10: Thermal denaturation curves of peptide **34** Ac-Phe(famp-Amp-Gly)$_6$-NH$_2$ at different pH conditions. pH 3.0 (20mM acetate buffer), pH 9.0 (20mM phosphate buffer, pH 9 and 12 (20mM borate buffer). All buffers contain 0.1M NaCl.

Table 9: CD-T_m values of peptide **34** (famp- Amp-Gly) at different pH conditions.

pH	T_m ^oC
3.0	23.2
7.0	21
9.0	21
12.0	20

Among all the peptides, the peptide **34** (famp-Amp-Gly) shows a maximum T_m under all pH conditions (except at pH 9.0). The stability of triple-helix gradually decreases from pH 3.0-9.0, followed by an increase again at pH 12.0. Interestingly, the T_m values for peptides **31** (Pro-fAmp-Gly) and peptide **33** (famp-Pro-Gly) are almost invariant with pH. The peptide **34** (famp-Amp-Gly) which contain 4*S*-N-formyl substituent at X position and ionizable free 4*R*-aminegroup at Y position exhibits T_m gradually decreasing form pH 3.0-9.0 and again increasing at pH 12.0 (Table 10).

Table 10: CD-Tm values of peptides **31-34**

Peptides	T_m values (^oC)			
X---Y---Gly	**pH 3.0**	**pH 7.0**	**pH 9.0**	**pH 12.0**
Ac-Phe(Pro-fAmp-Gly)$_6$-NH$_2$ 31	**42**	**41**	**40**	**40.5**
Ac-Phe(fAmp-Pro-Gly)$_6$-NH$_2$ 32	**ND**	**ND**	**ND**	**ND**
Ac-Phe(famp-Pro-Gly)$_6$-NH$_2$ 33	**23.2**	**21**	**21**	**20**
Ac-Phe(famp-Amp-Gly)$_6$-NH$_2$ 34	**58**	**56**	**24**	**44**

4.4.3: Effect of ethylene glycol on the stability of peptides 31-34 triple-helices

The EG:W mixture is known to stabilize the helical structures and is useful to amplifying and detecting small amount of weak triple-helical collagen like structures. The CD spectra of the synthetic collagen mimetic peptides **31-34** in 3:1 v/v EG:W mixture is shown in Figure 4.10A. In comparison to aqueous solution, the spectrum of peptide **32** (fAmp-Pro-Gly) is characterized by the appearance of significant positive

band at 225nm, crossover at 220nm and large decrease in molar ellipticity of the negative band 210nm. The R_{pn} values obtained from this spectrum is listed in Table 11.

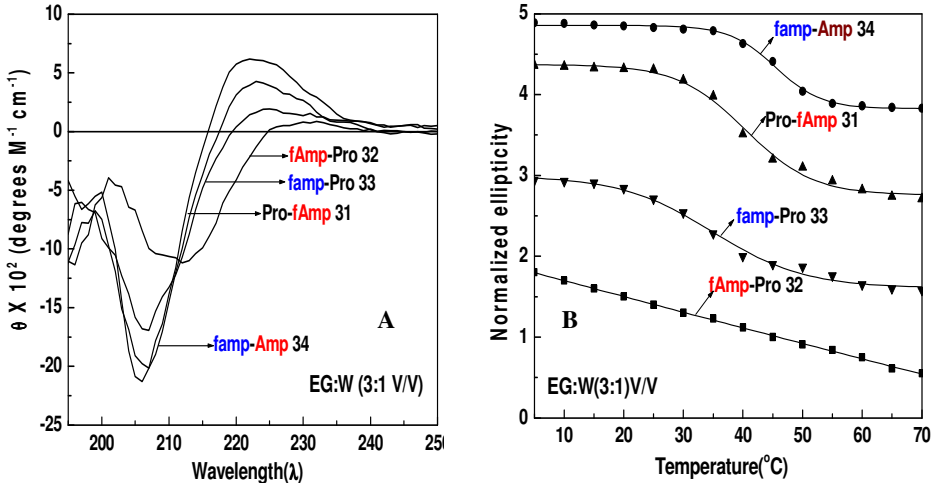


Figure 4.10: **A**, CD spectra at 10°C and **B,** CD-T_m curve of 0.2M solutions of peptides **31-34** in 3:1 (v/v) EG:W mixture

Table 11: Rpn and CD-T_m values for peptides **31-34.**

Peptide	Rpn values	T_m (°C)
Ac-Phe(Pro-fAmp-Gly)$_6$-NH$_2$ 31	0.35	40
Ac-Phe(fAmp-Pro-Gly)$_6$-NH$_2$ 32	0.012	ND
Ac-Phe(famp-Pro-Gly)$_6$-NH$_2$ 33	0.32	31
Ac-Phe(famp-Amp-Gly)$_6$-NH$_2$ 34	0.37	47

Interestingly in EG:W system all the three peptides exhibit different crossover points, peptide **33** (famp-Pro-Gly) exhibits a crossover at 215nm similarly peptide **31** (Pro-fAmp-Gly) exhibits crossover at 214nm and peptide **34** (famp-Pro-Gly) crossover at 213nm. Despite significant appearance of both positive and negative bands in the CD spectra for peptide **32** (fAmp-Pro) in EG:W system, it does not associate in to triple-helical structure. This is evidenced by the appearance of linear decrease in ellipticity with temperature in thermal denaturation study. The T_m of peptide **31** (Pro-fAmp-Gly)

is 40^0C two degrees lower compared to the T_m in aqueous solution (T_m 42°C). Similarly the peptide **34** (famp-Amp-Gly) forms 11^0C less stable triple-helix in EG:W system (T_m = 47^0C) compared to that in aqueous solution (T_m = 58^0C). Interestingly, the peptide **33** (famp-Pro-Gly) stabilizes the triple helical structure in EG:W (ΔT_m 7.8^0C) compared to that in aqueous conditions. The order of magnitude of the Rpn values is matches with T_m values for all peptides. Peptide **33** (famp-Pro) having lower Rpn value shows lower T_m and peptide **34** (famp-Amp) having higher Rpn value shows a higher T_m.

4.5: Discussion

The 4*R*-amnioproline with formyl protection of 4-amino group containing peptide **31** (Pro-fAmp-Gly) at Y position forms a polyproline II like structure as indicated by a positive band at ~223nm and a strong negative band at ~200nm in the CD spectra under all pH conditions. Peptide **32** (fAmp-Pro-Gly) with formyl protection of 4*R*-amino group at X position does not exhibit PP II like structure, this is evidenced by the shallow trough and lack of positive peak. All the 4-amino formyl protected peptides display almost identical CD spectra with matching peak intensities and crossover points under variable pH conditions, which is contrary to corresponding 4-amino free peptides (Chapter **2**), display changes in peak intensities and crossover points in their CD spectra under variable pH conditions. The observed sigmoid transition in the temperature dependent CD measurements provide additional evidence for the two state (triple-helix to coil) transition for the peptides **31** (Pro-Amp-Gly), **33** (famp-Pro-Gly) and **34** (famp-Amp-Gly). The melting transitions are consistent with the observations from CD spectroscopy. T_m values of peptide **34** (famp-Amp-Gly) with 4*S*-formyl protected aminoproline at X position and 4*R*-(free)aminoproline at Y position, decreases with an increase in pH over the range 3.0-9.0, followed by an increase again at pH 12. In the

corresponding chimeric peptide (Chapter **3**, part A) with 4*S*-amp at X and 4*R*-Amp at Y position (both containing free amines), the T_m values decreases with pH in the range 3.0-9.0, and it does not associate in triple-helical structure at pH 12.0. Peptide **31** with 4*R*-fAmp at Y position and peptide **33** with 4S-famp at X position associate to form triplex structure in both acidic and basic conditions, as indicated by the observed sigmoid transition in the variable temperature CD measurement. The T_m values of both peptides (**31** and **33**) are invariant with pH. The stability of the triple-helices of the corresponding peptides in which 4-formyl protected aminoproline is replaced by the 4-aminoproline with free amino group (Chapter **2**) are variable with varying pH (Fig 4.11).

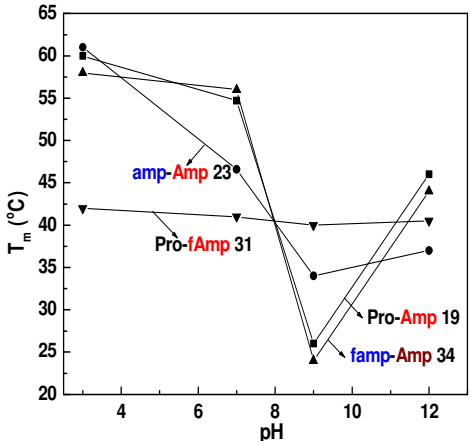


Figure 4.11: Comparison of T_m of peptides **19**, **23**,**31**, and **34**

This study reveals that protonation of 4*R*-aminogroup plays a vital role in formation and stability of the triple-helix containing 4*R*-aminoproline collagen peptides. Since all peptides were capped both at N and C terminals, the pH dependent triple-helix stability of 4-aminoproline containing collagen peptides is truly due to the protonation and nonprotonation effects of 4-amino group of aminoproline. Since the NH_3^+ and NH_2 groups have different electronegativity and the pyrrolidine ring puckering is dependent on the electronegativity of 4-substituent. It may be possible that 4-aminoproline exhibits

two different kinds of pyrrolidine ring puckering according to its protonated and nonprotonated forms. This may alter the difference in triple-helical stability under different pH conditions. In peptides **31** (Pro-fAmp-Gly) and **33** (famp-Pro-Gly) the 4-amino group protected as formyl derivative, avoid protonation under acidic conditions and hence the formyl protected aminoproline exists only in one conformation under all pH conditions resulting the formation of pH independent triple-helix structure. The peptide **34** (famp-Amp) which contains free aminoproline at Y position exits in protonated and nonprotonated forms in acidic and basic conditions respectively, forms pH dependent triple-helix. The 4-aminoformyl protected peptides form less stable triple-helix compared to the corresponding free aminoproline containing collagen peptides. This can be explained on the basis of preferred pyrrolidine ring conformation. Since the lone pair of electrons of the 4-aminogroup is in resonance with the formyl amide group, it may effect the gauche interaction with nitrogen of the α-amino atom, resulting changes in the magnitude of *exo-endo* conformation of the 4-aminoformyl protected aminoproline and also $K_{E/Z}$ value of peptide bond. The steric repulsion of 4-aminofomyl group with other strand in triple-helix structure can also be contributing factor in the destabilizing effect.

In EG:W (3:1) system, only peptide **33** with 4*S*-formyl protected aminoproline residue at X position (famp-Pro-Gly) stabilizes the triple-helical structure compared to that in aqueous condition (ΔT_m 7.8^{0}C). The other two peptides **31** (Pro-fAmp-Gly) and **34** (famp-Amp-Gly) form a less stable triple-helical structure in EG:W system compared to that in aqueous condition (ΔT_m 2.0^{o}C and 11.0^{o}C respectively). This may be due to the fact that in collagen triple-helix, unlike the Y-residue which is buried inside, the residue in X position is more exposed and is amenable to interactions with

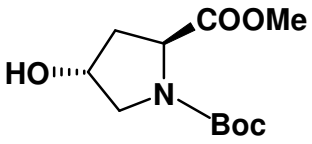
bulk solvent. As the famp residue is present in the X position in peptide **33** (famp-Pro-Gly), such higher order associations are possible which increase rapidly with increase in concentration. Hence reversal of stability in EG:W may arise from the difference in exposure to the solvent of residues at X and Y positions in triple-helix, and also the poorer solvating power of EG for charged and polar residues. Also the aminoproline containing peptides form less stable triple-helix in EG:W system compared to that in aqueous buffer conditions. The additional interstarand electrostatic repulsions arising from the 4-NH_3^+ groups in aminoproline peptides remains unscreened in EG:W mixture for peptide **34** which contain free aminoproline at Y position. Interestingly despite an increase in positive band intensity in EG:W for the peptide **32 (**fAmp-Pro-Gly**)** still it does not show any transition.

4.6: Experimental Section

4.6.1: General Procedures

Thin-layer chromatography was performed of gel GF250. Column chromatography was performed on silica gel, Merck grade 60-120 mesh. Reagents and chemicals were obtained from commercial suppliers, and reagents solvents were used without further purification. ^{1}H NMR and ^{13}C NMR spectra were recorded at 200MHz in $CDCl_3$. Both ^{1}H and ^{13}C NMR spectra were often complicated by the presence of carbamate conformers and pairs of ^{13}C NMR lines due to a single carbon. Chemical shifts are expressed in ppm relative to internal TMS (^{1}H) or solvent $CDCl_3$ (^{13}C).

(2S,4R)N-(t-Butyloxycabonyl)-4-HydroxyProlinemethylester **2.**



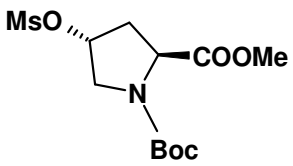
A solution of compound **1** (7 g, 53.3mmol) in 2N NaOH (50ml) and dioxan (50ml), was cooled to 0^0C and 12.8g (58.7mmol) Boc anhydride was added drop wise. The reaction mixture was stirred at 0^0C for 1 h. Dioxan was removed under reduced pressure; aqueous layer was covered with EtOAc (150ml). The reaction mixture was neutralized by adding, solid $KHSO_4$ portion wise with vigorous stirring until pH 2. Ethyl acetate layer was separated, aqueous layer was extracted with ethyl acetate (3 X 50ml), the combined organic layer was washed with water followed saturated brine solution, dried over anhydrous Na_2SO_4. Upon removal of ethyl acetate under reduced pressure, yielded a white solid which was recrystalized with EtOAc/Pet ether.

To the stirred solution of above white compound **(**12g, 51.88mmol) in 75ml of anhydrous acetone, 21.5g (155.6mmol) of anhydrous K_2CO_3 and dimethylsulphate 6ml (62.25mmol) in one shot were added. The mixture was then refluxed under nitrogen for 4hrs. The acetone was removed under vacuum; the resulting residue was dissolved in water then extracted with ethyl acetate. The combined organic layer was washed with water, followed by saturated brine solution, dried over Na_2SO_4 and concentrated under vacuum. The crude material was purified by silica gel chromatography (50% ethyl acetate/hexane elute) afford compound **2** as a white solid. Yield 12.1g (from two steps); 92.46%. Mol. Formula $C_{11}H_{19}NO_5$; Mass (observed) 246.5 (245.13 Calc).

^{1}H NMR δ_H 1.40 (s, 9H), 2.03-2.10 (m,1H), 2.27-2.39 (m, 1H), 3.51-3.65 (m, 2H), 3.77(s, 3H), 4.25-4.38 (m, 2H).

^{13}C NMR, 28.0, 38.3, 38.9, 51.4, 57.8, 68.9, 69.5, 80.1, 153.9, 154.4, 173.5.

(2S,4R)N-(t-Butyloxycabonyl)-4-O-Mesyl-Prolinemethylester **3.**

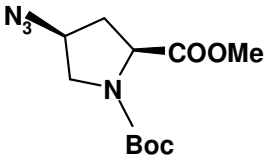


The solution of methylester **2** (12g 48.95mmol) and triethylamine 10.2ml (73.43mmol) in 100ml of dry dichloromethane was cooled to 0°C on ice bath under Argon. While stirring methanesulfonyl chloride 4.6ml, (58.74mmol) was added dropwise over a period of 3hrs at 0°C. The reaction mixture was washed with water, followed by saturated brine solution. The organic layer was dried over anhydrous Na_2SO_4 and concentrated under vacuum. The crude material purified by silica gel chromatography (30% ethyl acetate/hexane) afforded mesylated compound **3** as a white solid. Yield 15.5g; 97.9%; Mol. Formula $C_{12}H_{21}NO_7S$; Mass (observed) 324.4 (323.1 Calc).

^{1}H NMR δ_H 1.40 (s, 9H), 2.16-2.30 (m, 1H), 2.51-2.71 (m, 1H), 3.03 (s, 3H), 3.73 (s, 3H), 3.76-3.87 (m,2H), 4.34-4.48 (m, 1H), 5.24 (s, 1H). **^{13}C NMR,** 28.2, 31.1, 31.5, 32.5, 47.3, 47.8, 54.3, 58.3, 80.0, 153.5, 154.0, 154.8, 171.5.

(2S,4S)N-(t-Butyloxycabonyl)-4-Azido-Prolinemethylester **4.**

A solution of mesyl compound **3** (15.0g, 46.42mmol) and NaN_3 24.0g, (371.4mmol) in dry DMF (120ml) was stirred at 70°C for 8 hrs under argon. DMF was removed under vacuum and residue was dissolved in water. The aqueous layer was extracted with ethyl acetate. The combined organic layer was washed with water-followed by brine, dried over anhydrous Na_2SO_4 and concentrated under vacuum. The crude material was purified by silica gel chromatography (40% ethyl acetate/hexane) to afford azide compound **4** as a colorless

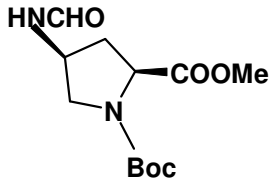


oil. Yield 12.0 g; 95.6%. Mol. Formula $C_{11}H_{18}N_4O_4$; Mass (observed) 271.3 (270.13 Calc);

^{1}H NMR δ_H 1.37 (s, 9H), 2.07-2.108 (m, 1H), 2.33-2.52 (m, 1H), 3.41-3.49 (m, 1H), 3.64-3.66 (m, 1H), 3.71 (s, 3H), 4.06-4.13 (d, 1H), 4.25-4.42 (m, 1H).

^{13}C NMR, 27.8, 34.7, 35.6, 50.4, 51.9, 57.0, 57.3, 57.9, 58.9, 80.1, 153.0, 153.6, 171.6, 171.9.

(2S,4S)N^1-(t-Butyloxycabonyl)-4-N-Formyl-aminoprolinemethylester **5.**

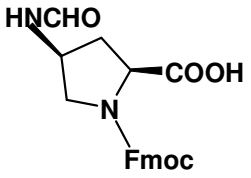


To a solution of azide (12g, 44.23mmol) in methanol (8ml) taken in hydrogenation flask was added Raney Nickel catalyst (2mol %). The reaction mixture was hydrogenated in a parr shaker apparatus for 4hrs at room temperature under H_2 pressure of 45-50 psi. The catalyst was filtered off, and then solvent was removed under reduced pressure to yield a residue of amine, which was stirred with methylformate at room temperature for 4hrs. The solvent was removed under reduced pressure and the crude product was purified by column chromatography (EtOAc/Petroleum ether [7:3]) to offer a pale yellow liquid of formylate compound **5**. Yield 11.8g, 78.%; Mol. Formula $C_{12}H_{20}N_2O_5$; Mass (observed) 272.7 (272.4 Calc);

^{1}H NMR δ_H 1.39 (s, 9H), 1.80-2.08 (m, 1H), 2.41-2.56 (m, 1H), 3.48-3.59 (m, 2H), 3.76 (s, 3H), 4.24-4.36 (m, 1H), 4.72-4.74 (d, 1H), 8.07 (s, 1H).

^{13}C NMR, 28.0, 35.1, 36.3, 47.9, 50.8, 52.0, 52.1, 57.3, 57.5, 80.5, 153.5, 154.1, 161.2, 172.5, 172.7.

(2S,4S)N^1-(Fmoc)-4S-N^4-Formyl-aminoproline **7.**



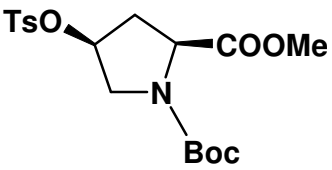
The compound **5** (5gm 18.35mmol) was stirred for 2hrs in 1:1 mixture of CH_2Cl_2 and TFA (8ml). The solvent was removed under reduced pressure, the resulting slurry was dissolved in diethyl ether, removal of ether yielded white compound of trifluoroacetate salt.

The above obtained salt was dissolved in a mixture of methanol (25ml) and 0.5M LiOH solution (20ml). The reaction mixture was stirred for 30min. The reaction mixture was then neutralized to pH 7, and methanol was removed under reduced pressure. The aqueous layer was cooled in ice-water bath, dioxan (50ml) and 2N sodium carbonate (50ml) was added to it. The solution of Fmoc-Cl (9.1gm 35.2mmol) in 10ml dioxane was added slowly over a period of 75min. the reaction mixture was stirred for 6hrs at room temperature, the pH of the reaction was maintained about 8-9 during the reaction. The dioxan was removed under reduce pressure, the aqueous layer was cooled and extracted with diethyl ether (2 X 25ml). Ether layer was discarded and aqueous layer was covered with ethylacetate layer and acidified with $KHSO_4$ solution to pH 2-3. The organic layer was separated and aqueous layer was extracted with ethylacetate. The combined organic layer was washed with water followed by saturated brine solution, dried over anhydrous Na_2SO_4. The ethylacetate was removed under reduced pressure, and the crude product was purified by column chromatography (EtOAc/Petroleum ether [9:1]) to offer a white amorphous solid of monomer **7**. Yield 4.5g, 64% over three steps. Mol. Formula $C_{21}H_{20}N_2O_5$; Mass (observed) 381.86 (380.14 Calc);

^{1}HNMR δ_H 2.03-2.13 (m,1H), 2.34-2.48 (m, 1H), 3.48-3.77 (m, 2H), 4.05-4.20 (t, 2H), 4.34-4.37 (m, 3H), 4.56-4.63 (d, 1H), 6.59 (b, 1H), 7.13-7.36 (m, 4H), 7.50 (dd, 2H), 7.67-7.73 (dd, 2H), 8.03 (s, 1H).

^{13}C NMR, 34.8, 46.8, 52.7, 58.5, 67.9, 119.9, 124.0, 127.0, 127.7, 141.1, 143.5, 143.8, 155.2, 161.8, 175.9.

(2S,4S)-N1-(t-butoxycarbonyl)-4-(p-toluenesulfonyloxy)prolinemethylester **8.**

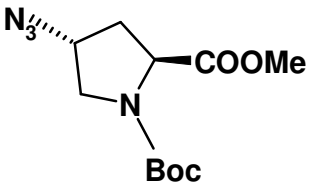


The mixture of compound **2,** (7g 28.55mmol), PPh_3, 8.24g (31.41mmol) and methyl-*p*-toluene sulfonate,4.74ml, (31.41mmol) dissolved in dry THF (100ml) was cooled to 0^oC on ice bath, under Argon. The mixture was stirred for 30min at 0^0C Diisopropylazodicarboxylate 4.74ml (31.41mmol) was added slowly with syringe. The reaction mixture was stirred at room temperature for 8hrs. Toluene was removed under vacuum. The resulting orange colored thick oil was dissolved in 150ml petroleum ether, by triturating with spatula; the resulting solution was kept overnight at room temperature. The white powder settled was filtered and the residue was washed with petroleum ether followed by diethylether, offered compound **8** as white power. Yield 8.7g (89%). Mol. Formula $C_{18}H_{25}NO_7S$; Mass (observed) 340.4 (339.14 Calc);

^{1}H NMR δ_H 1.36 (s, 9H), 2.30-2.51 (m, 2H), 2.42 (s, 3H), 3.66 (s, 3H), 3.57-3.73 (m, 2H), 4.27-4.43 (m, 1H), 5.0-5.03 (t, 1H), 7.31-7.34 (dd, 2H), 7,72-7.75 (dd, 2H).

^{13}C NMR, 21.4, 21.8, 28.0, 35.8, 36.7, 51.4, 52.0, 52.1, 56.8, 57.2, 62.9, 63.1, 69.7, 78.7, 80.4, 127.5, 129.8, 145., 153.1, 153.6, 171.4, 171.8.

(2S-4R)N-(t-Butyloxycabonyl)-4-Azido-Prolinemethylester **9.**

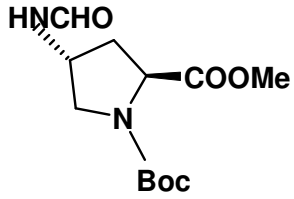


A solution of compound **8,** 8.5g, (25mmol) and NaN_3, 13g, (200.5mmol) in dry DMF (80ml) was stirred at 55-60^oC for 8hrs under Argon. DMF was removed under vacuum and the residue was dissolved in water. The aqueous layer was extracted with ethyl acetate. The combined organic layer was washed with water-followed by brine, dried over anhydrous Na_2SO_4 and concentrated under vacuum. The crude product obtained was purified by silica gel chromatography (40% ethyl acetate/hexane elute) afford **9** as a colorless thick oil. Yield 6.5g, 96%, Mol. Formula $C_{11}H_{18}N_4O_4$; Mass (observed) 271.5 (270.13 Calc);

^{1}H NMR δ_H 2.07-2.34 (m, 2H), 3.40-3.57 (m, 2H), 3.71 (s, 3H), 4.07-4.19 (m, 1H), 4.26-4.42 (m, 1H), 4.88-5.0 (t, 1H).

^{13}C NMR, 28.1, 35.2, 36.1, 51.1, 52, 52.2, 58.6, 59.1, 80, 153.8, 256.3, 172.8.

(2S,4R)N^1-(t-Butyloxycabonyl)-4-N-Formyl-aminoprolinemethylester **10.**

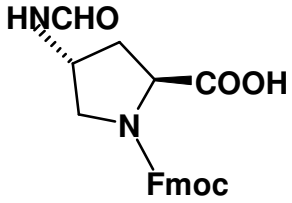


To a solution of azide (6gm, 22.2mmol) in methanol (6ml) taken in hydrogenation flask was added Raney Nickel catalyst (2mol %). The reaction mixture was hydrogenated in a Parr shaker apparatus for 4hrs at room temperature under H_2 pressure of 45-50 psi. The catalyst was filtered off, and then solvent was removed under reduced pressure to yield a residue of amine, which was stirred with methylformate at room temperature for 4hrs. The solvent was removed under reduced pressure and the crude product was purified by column chromatography (EtOAc/Petroleum ether [7:3]) to offer a pale yellow liquid of formylate compound**.** Yield 5.5g, 90.9% Mol. Formula $C_{12}H_{20}N_2O_5$**;** Mass (observed) 272.83 (272.4 Calc);

^{1}H NMR δ_H 1.38 (s, 9H), 2.25-2.30 (m, 2H), 3.30-3.43 (m, 1H), 3.63-3.78 (m, 1H), 3.71 (s, 3H), 4.11-4.57 (m, 2H), 8.12 (s, 1H).

^{13}C NMR, 28.0, 35.1, 36.2, 46.4, 50.8, 52.1, 57.3, 57.5, 80.50, 153.5, 154.1, 161.2, 172.5, 172.7.

N^{1}-(Fmoc)-4R-N-Formyl-aminoproline **12.**



This was prepared by same procedure as for compound **7** starting from compound **10**. Yield 5g 60%; Mol. Formula $C_{21}H_{20}N_2O_5$; Mass (observed) 381.86 (380.14 Calc);

^{1}H NMR δ_H 2.47-1.61 (q, 2H), 3.64-3.78 (dd, 1H), 3.98-4.14 (m, 1H), 4.32-4.52 (m, 1H), 4.54 -4.67 (m, 4H), 4.78-4.81 (m, 2H), 7.54-7.70 (m, 4H), 7.85-7.91(dd, 2H), 8.02-8.07 (d, 2H).

^{13}C NMR, 35.1, 36.2, 39.7, 45.9, 46.8, 46.9, 47.5, 51.3, 51.7, 58, 58.3, 67.2, 67.5, 106.4, 119.8, 120, 121.2, 125.2, 127.2, 127.4, 128.8, 137.6, 139.7, 140.8, 143.7, 152.2, 145.3, 161.3, 174.1, 174.4

4.7: Reference

1) Mooney, S. D.; Keein, T. E.; Kollman, P. A. *Biopolymers* **2002**, *64*, 63-71.

2) Vitagliano, L.; Berisio, R.; Mazzarella, L.; Zagari, A. *Biopolymers* **2001**, *58*, 459-464.

3) Karmer, R.; Vitagliano, L.; Bella, J.; Berisio, R.; Mazzarella, L.; Brodsky, B.; Zagari, A.; Berman, H. *J. Mol. Biol*. **1998**, *280*, 623-638.

4) Li. M.; Fan, P.; Brodsky, B.; Baum, J. *Biochemistry* **1993**, *32*, 13299-13309.

5) Fraser, R.; MacRae, T.; Suzuki, E. *J. Mol. Biol*. **1979**, *129*, 463-481.

6) Okuyama, K.; Arnott, S.; Takayangi, M.; Kakudo. M. *J. Mol. Biol*. **1981**, *152*, 427-443.

7) Migchels, C.; Berendse, H. *J. Chem. Phys*. **197**3, *59*, 296-305.

8) Bella, J.; Eaton, M.; Brodsky, B.; Berman, H. *Science* **1994**, *266*, 5182.

9) Bella, J.; Brodsky, B.; Berman, Beman, H. *Connective Tissues Res*. **1996**, *35*, 1-4.

10) Sakakibara, S.; Kishida, Y.; Okuyama, K.; Tanaka, N.; Ashida, T.; Kakudo, M. *J. Mol. Boil.* **1972,** *65*, 371-373.

11) Brown, F. R., III; Carver, J. P.; Bluot, E. R. *J. Mol. Boil.* **1969**, *39*, 307-313.

12) Brown, F. R., III; Di Corato, A.; Lorenzi, G. P.; Bluot, E. R. *J. Mol. Biol.* **1972**, *63*, 85-99.

13) Feng, Y.; Melacini, G.; Taulane, J. P.; Goodman, M. *J. Am. Chem. Soc.* **1996**, *118*, 10351-10358.

14) Feng, Y.; Melacini, G.; Taulane, J. P.; Goodman, M. *Biochemistry* **1997**, *36*, 8716-8724.

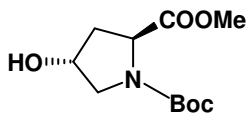
15) Venugopal, M. G.; Ramshaw, J. A. M.; Braswell, E.; Zhu, D.; Brodsky, B. *Biochemistry* **1994**, *33*, 7948-7956.

Appendix I

Compound No.	Spectra	Page No
Compound **2** ……	^{1}HNMR,^{13}CNMR, DEPT…….	249
Compound **3** ……	^{1}HNMR,^{13}CNMR, DEPT…….	250
Compound **4**……	^{1}HNMR,^{13}CNMR, DEPT…….	251
Compound **5**……	^{1}HNMR,^{13}CNMR, DEPT…….	252
Compound **7**……	^{1}HNMR,^{13}CNMR, DEPT…….	253
Compound **8**……	^{1}HNMR,^{13}CNMR, DEPT…….	254
Compound **9**……	^{1}HNMR,^{13}CNMR, DEPT…….	255
Compound **10**……	^{1}HNMR,^{13}CNMR, DEPT…….	256
Compound **12**……	^{1}HNMR,^{13}CNMR, DEPT…….	257
Peptide **31** ……	RP-18 HPLC and MALDI-TOF…….	258
Peptide **32** ……	RP-18 HPLC and MALDI-TOF…….	259
Peptide **33** ……	RP-18 HPLC and MALDI-TOF…….	260
Peptide **34** ……	RP-18 HPLC and MALDI-TOF…….	261
Compounds **12** and **7** …….	LC-MS ……..	262

7 Jul 2006

Acquisition Time (sec) 2.5559	***Comment*** DEPROTECTION			***Date*** 15/02/04 17:46:29			
Frequency (MHz) 200.13	***Nucleus*** 1H	***Number of Transients*** 256		***Original Points Count*** 8192		***Points Count*** 8192	
Sweep Width (Hz) 3205.13	***Temperature (grad C)*** 24.000						



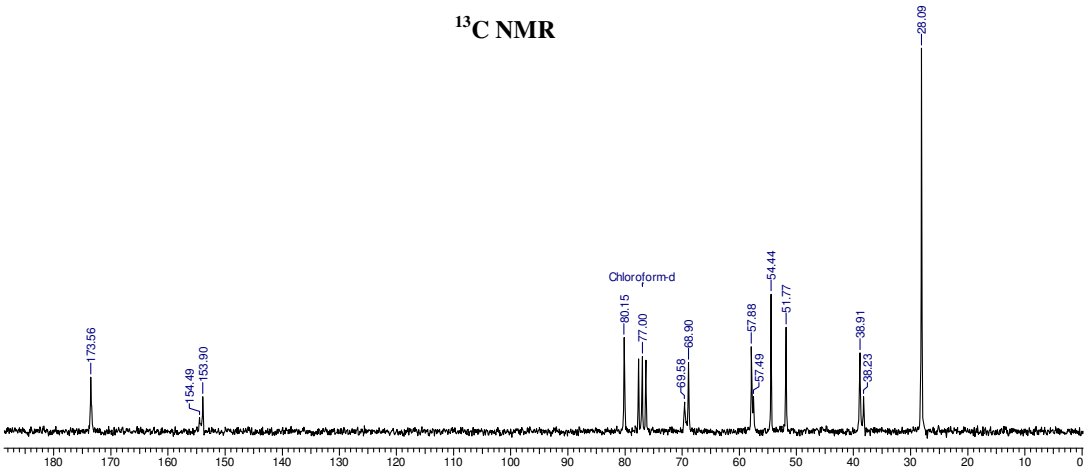
^{1}H NMR

Compound 2

7 Jul 2006

Acquisition Time (sec) 0.4096	***Comment*** DR.R.A. JOSHI/C13/ME-2/**SVT			***Date*** 17/12/02 18:22:22			
Frequency (MHz) 50.32	***Nucleus*** 13C	***Number of Transients*** 20000		***Original Points Count*** 8192		***Points Count*** 8192	
Sweep Width (Hz) 20000.00	***Temperature (grad C)*** 24.000						

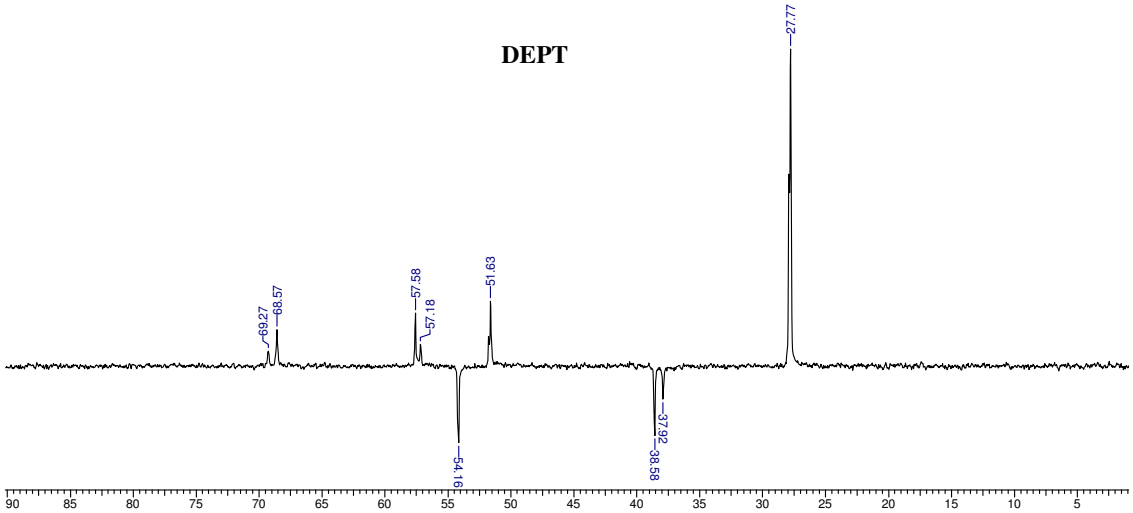
^{13}C NMR



7 Jul 2006

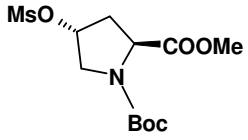
Acquisition Time (sec) 0.5407	***Comment*** D.K. MAHAPATRA/DEPT/**SVT			***Date*** 17/12/02 18:29:28			
Frequency (MHz) 50.32	***Nucleus*** 13C	***Number of Transients*** 2048		***Original Points Count*** 8192		***Points Count*** 8192	
Sweep Width (Hz) 15151.52	***Temperature (grad C)*** 24.000						

DEPT



Acquisition Time (sec)	2.5559	**Comment**	AMIT HORNE/1H/**SVT		**Date**	12/01/04 22:04:21	
Frequency (MHz)	200.13	**Nucleus**	1H	**Number of Transients** 256	**Original Points Count**	8192	**Points Count** 8192
Sweep Width (Hz)	3205.13	**Temperature (grad C)**	24.000				

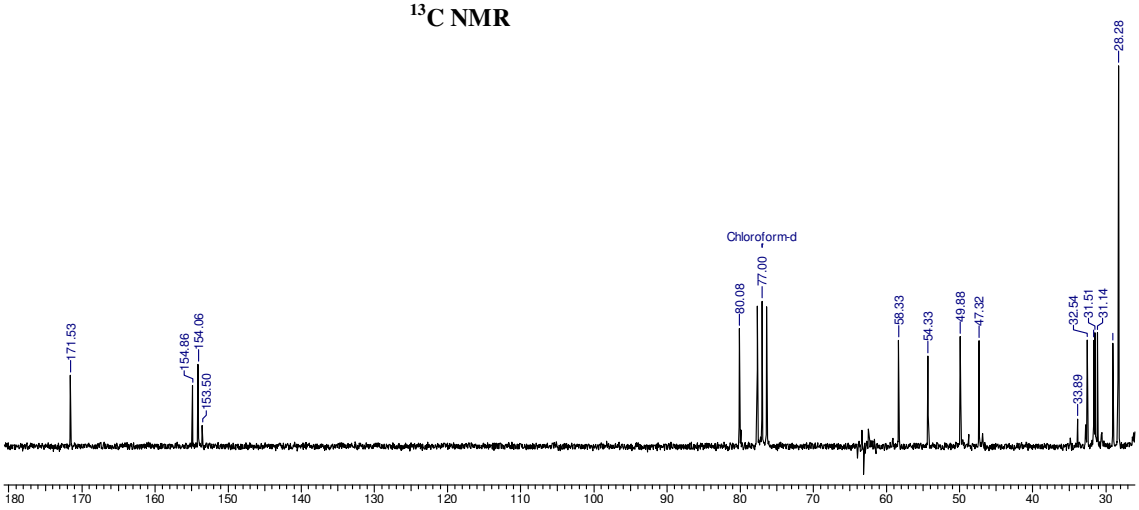
^{1}H NMR



Compound 3

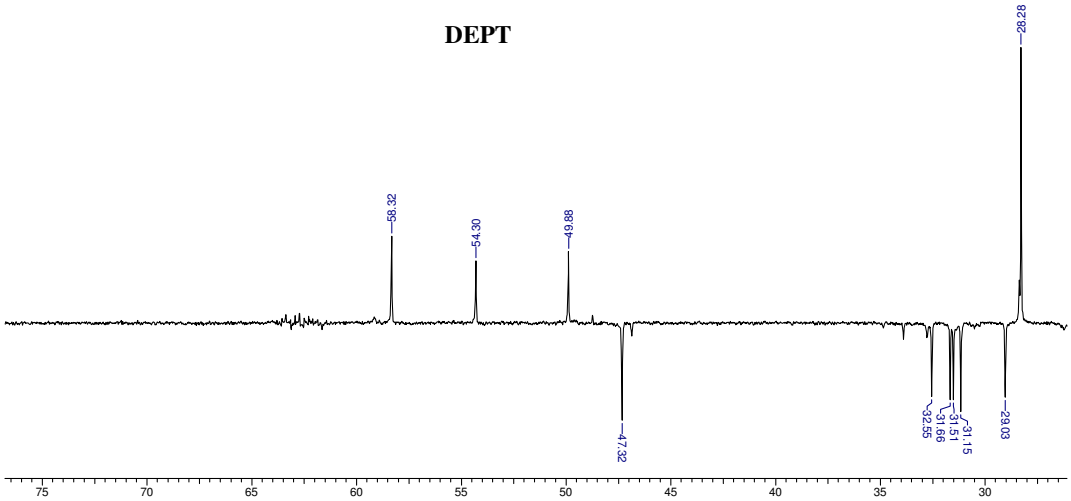
Acquisition Time (sec)	2.7329	**Comment**	M. Umasankar		**Date**	12/07/2006 23:24:38	
Frequency (MHz)	50.32	**Nucleus**	13C	**Original Points Count** 32768	**Points Count**	32768	**Sweep Width (Hz)** 11990.41
Temperature (grad C)	0.000						

^{13}C NMR



Acquisition Time (sec)	2.7329	**Comment**	M. Umasankar		**Date**	12/07/2006 23:44:02	
Frequency (MHz)	50.32	**Nucleus**	13C	**Original Points Count** 32768	**Points Count**	32768	**Sweep Width (Hz)** 11990.41
Temperature (grad C)	0.000						

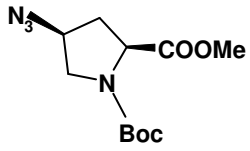
DEPT



7 Jul 2006

Acquisition Time (sec) 7.9167	***Comment***	M.Umashankara		***Date***	30/12/2004 13:12:16		
Frequency (MHz) 200.13	***Nucleus*** 1H		***Original Points Count*** 32768	***Points Count***	32768	***Sweep Width (Hz)***	4139.07
Temperature (grad C) 0.000							

^{1}H NMR

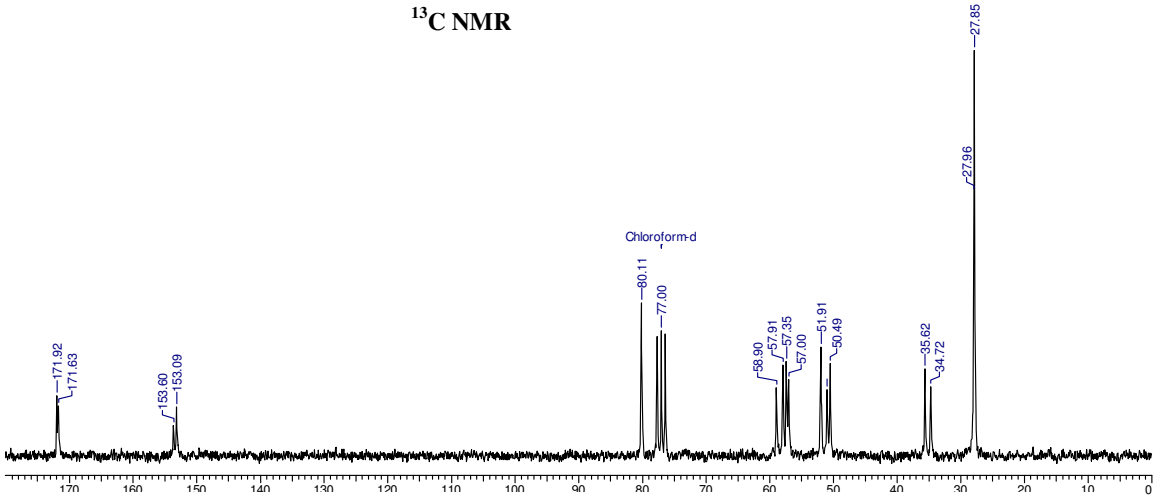


Compound 4

7 Jul 2006

Acquisition Time (sec) 2.7329	***Comment***	m umashankara		***Date***	09/01/2005 00:17:30		
Frequency (MHz) 50.32	***Nucleus*** 13C		***Original Points Count*** 32768	***Points Count***	32768	***Sweep Width (Hz)***	11990.41
Temperature (grad C) 0.000							

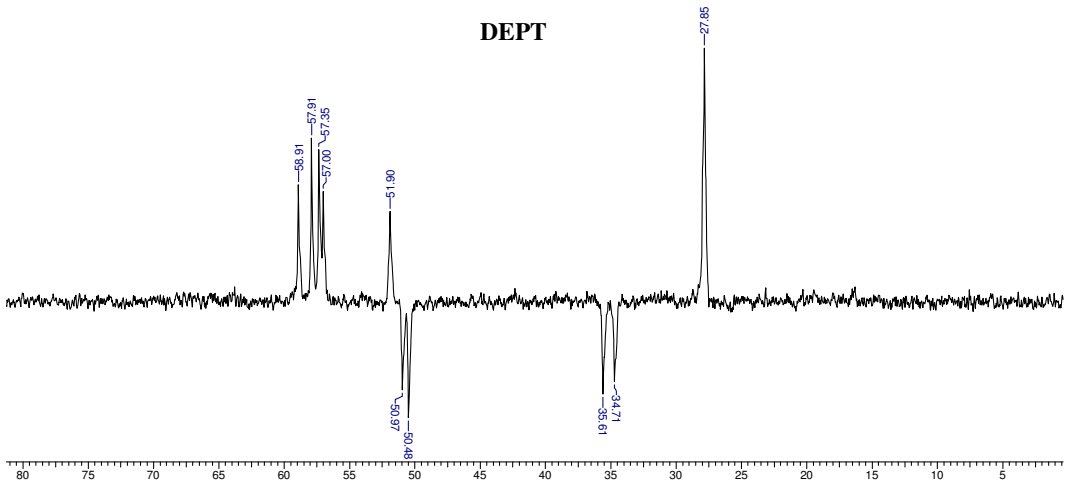
^{13}C NMR



7 Jul 2006

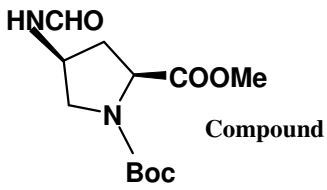
Acquisition Time (sec) 2.7329	***Comment***	m umashankara		***Date***	09/01/2005 00:24:02		
Frequency (MHz) 50.32	***Nucleus*** 13C		***Original Points Count*** 32768	***Points Count***	32768	***Sweep Width (Hz)***	11990.41
Temperature (grad C) 0.000							

DEPT



Acquisition Time (sec) 7.9167	***Comment***	M.Umasankara		***Date***	25/06/2005 21:53:50		
Frequency (MHz) 200.13	***Nucleus***	1H	***Original Points Count*** 32768	***Points Count***	32768	***Sweep Width (Hz)***	4139.07
Temperature (grad C) 0.000							

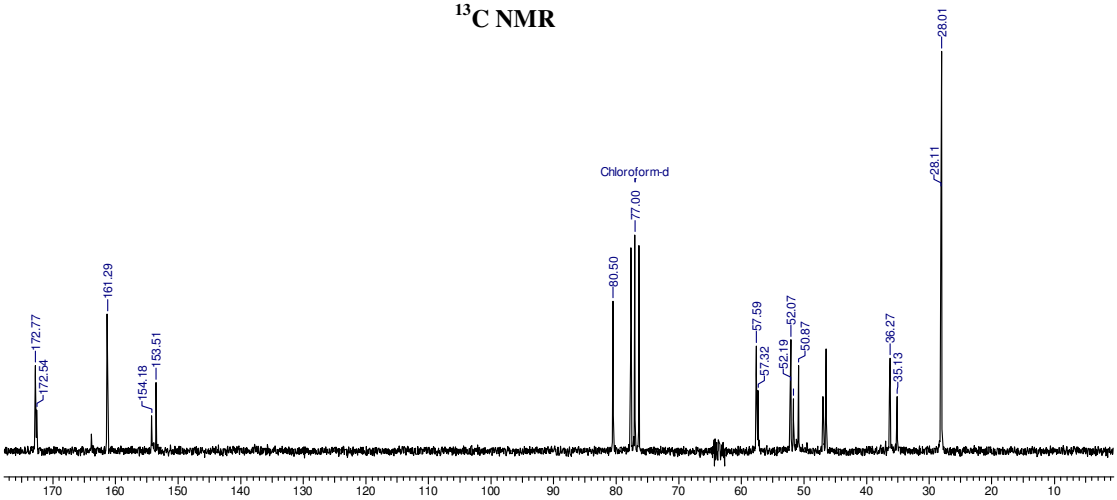
^{1}H NMR



Compound 5

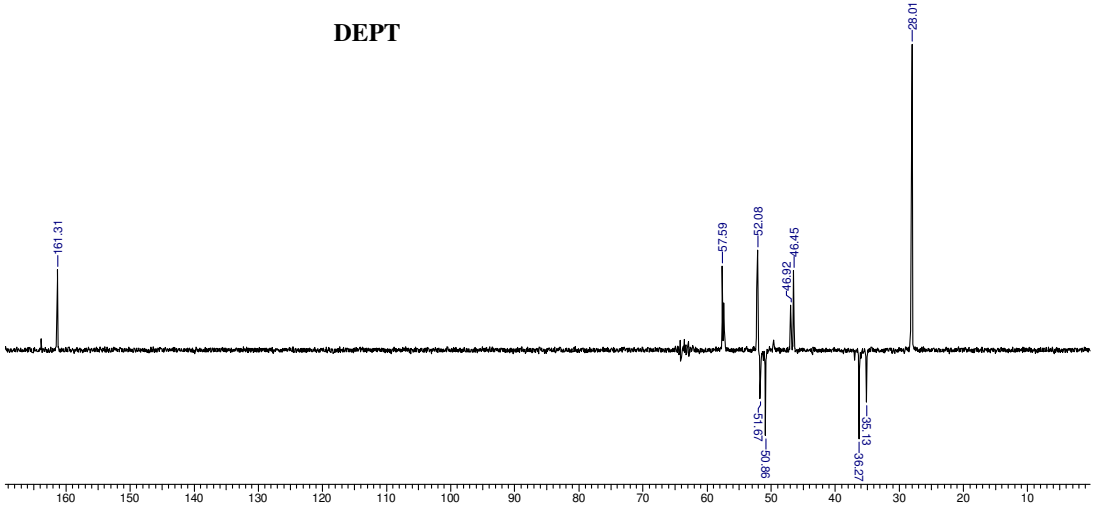
Acquisition Time (sec) 2.7329	***Comment***	m umasankara		***Date***	25/06/2006 04:43:38		
Frequency (MHz) 50.32	***Nucleus***	13C	***Original Points Count*** 32768	***Points Count***	32768	***Sweep Width (Hz)***	11990.41
Temperature (grad C) 0.000							

^{13}C NMR

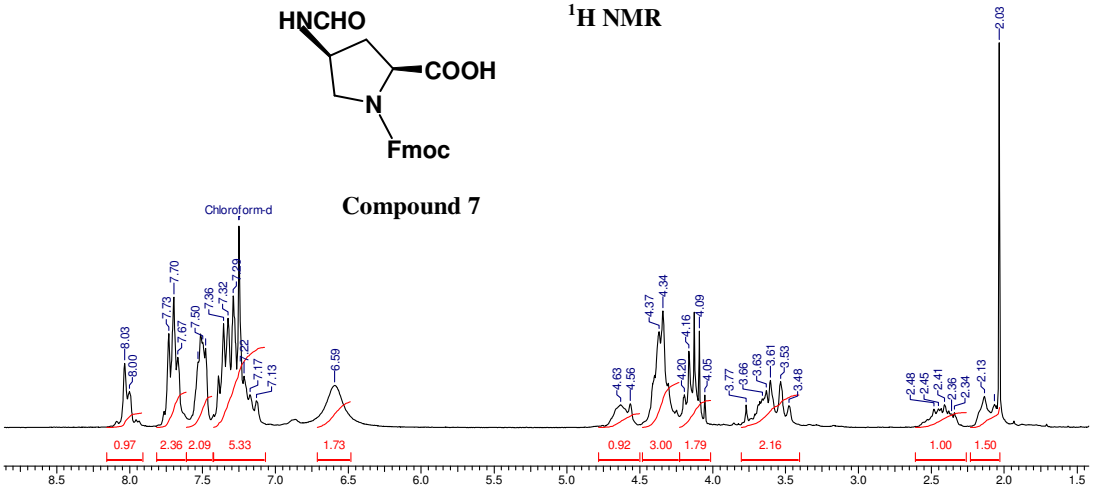


Acquisition Time (sec) 2.7329	***Comment***	m umasankara		***Date***	25/06/2006 05:02:54		
Frequency (MHz) 50.32	***Nucleus***	13C	***Original Points Count*** 32768	***Points Count***	32768	***Sweep Width (Hz)***	11990.41
Temperature (grad C) 0.000							

DEPT



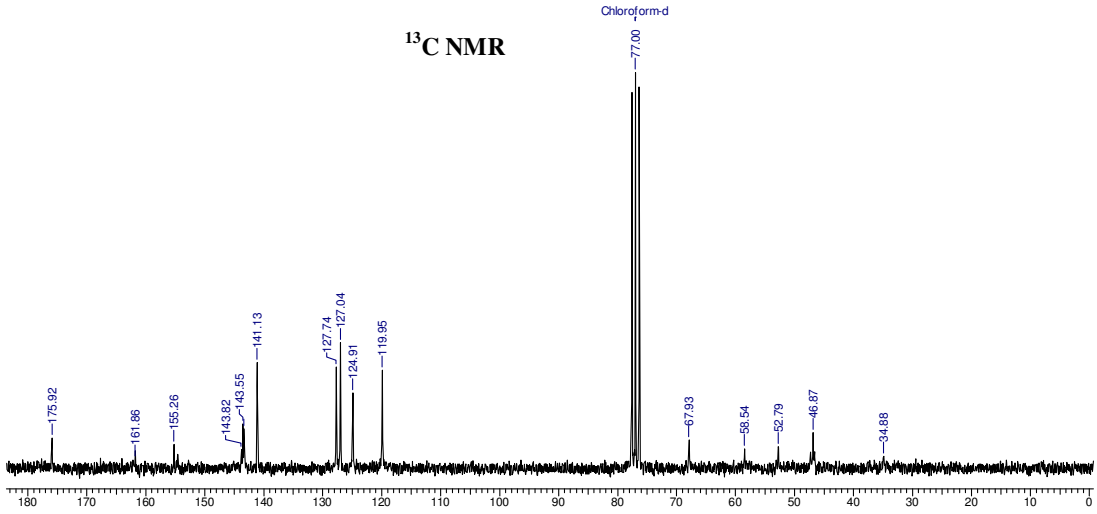
Acquisition Time (sec) 7.9167	***Comment***	M. UmaShankar		***Date***	04/10/2005 23:26:26		
Frequency (MHz) 200.13	***Nucleus***	1H	***Original Points Count*** 32768	***Points Count***	32768	***Sweep Width (Hz)***	4139.07
Temperature (grad C) 0.000							



^{1}H NMR

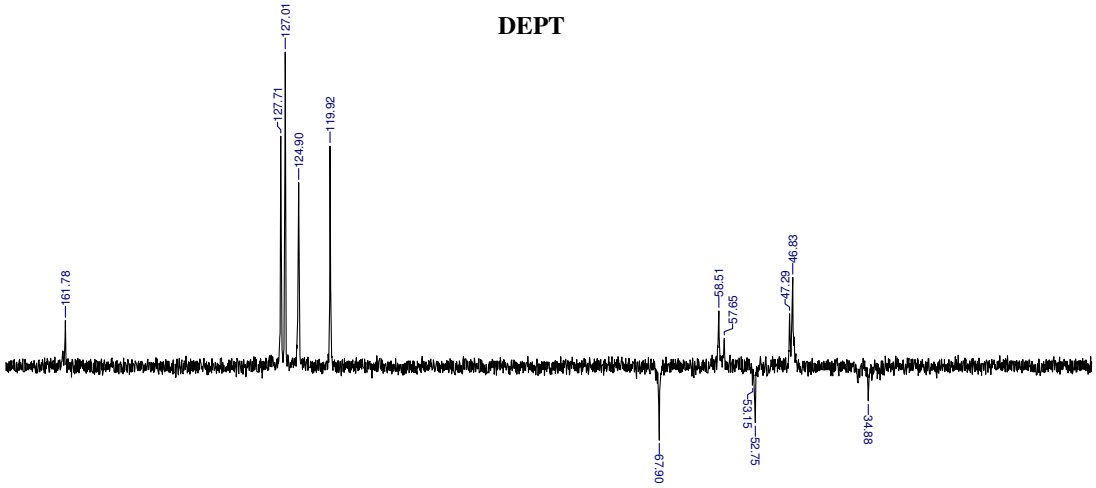
Compound 7

Acquisition Time (sec) 2.7329	***Comment***	M Umasankara		***Date***	20/06/2006 23:48:40		
Frequency (MHz) 50.32	***Nucleus***	13C	***Original Points Count*** 32768	***Points Count***	32768	***Sweep Width (Hz)***	11990.41
Temperature (grad C) 0.000							



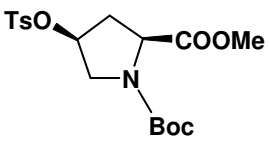
^{13}C NMR

Acquisition Time (sec) 2.7329	***Comment***	M Umasankara		***Date***	21/06/2006 00:26:08		
Frequency (MHz) 50.32	***Nucleus***	13C	***Original Points Count*** 32768	***Points Count***	32768	***Sweep Width (Hz)***	11990.41
Temperature (grad C) 0.000							



DEPT

Acquisition Time (sec)	2.0480	***Comment***	MRS. S.P. BAPAT/**SVT			***Date***	21/12/02 16:23:44
Frequency (MHz)	200.13	***Nucleus***	1H	***Number of Transients***	512	***Original Points Count***	8192
Sweep Width (Hz)	4000.00	***Temperature (grad C)***	24.000				

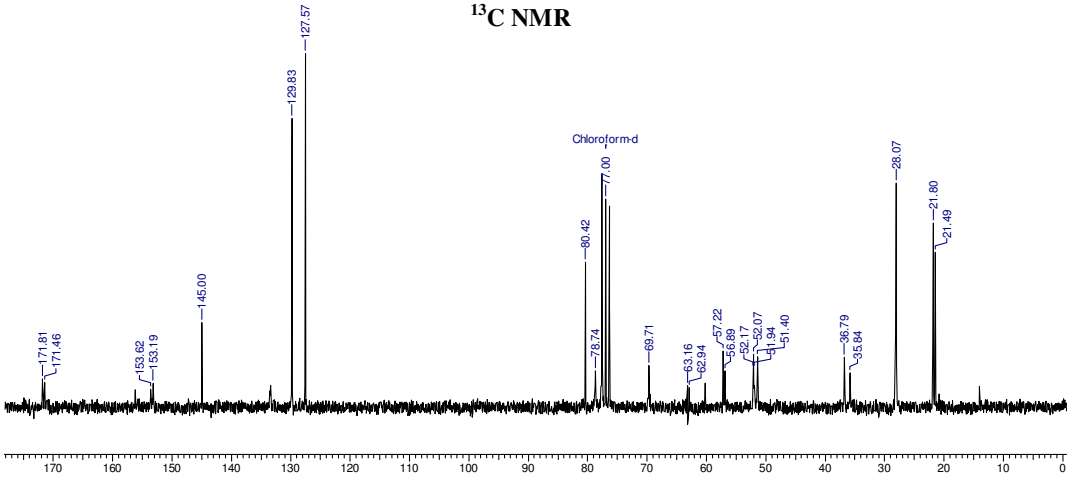


^{1}H NMR

Compound 8

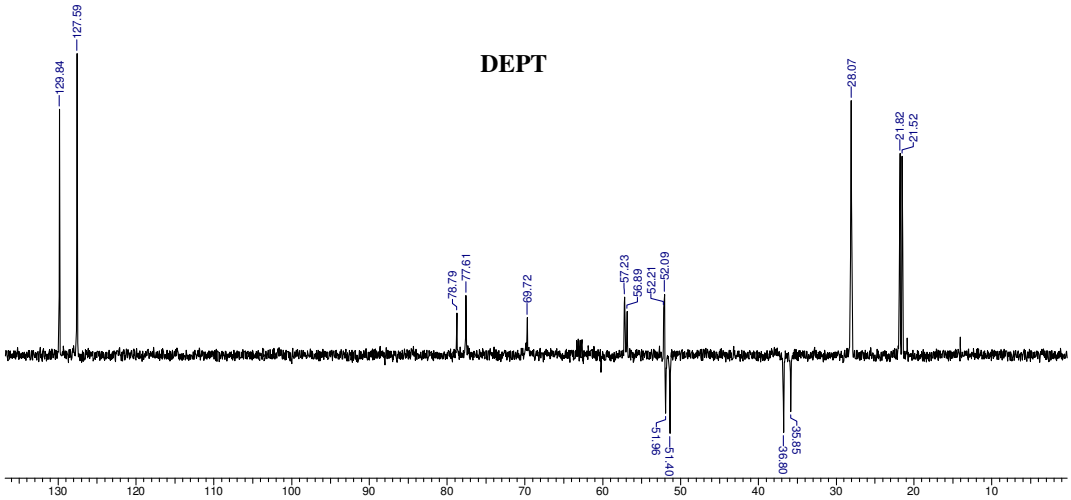
Frequency (MHz)	50.32	***Nucleus***	13C	***Original Points Count***	32768	***Points Count***	32768
Sweep Width (Hz)	11990.41	***Temperature (grad C)***	0.000				

^{13}C NMR



Acquisition Time (sec)	2.7329	***Comment***	M.Umashankara			***Date***	17/07/2006 09:50:50
Frequency (MHz)	50.32	***Nucleus***	13C	***Original Points Count***	32768	***Points Count***	32768
Sweep Width (Hz)	11990.41	***Temperature (grad C)***	0.000				

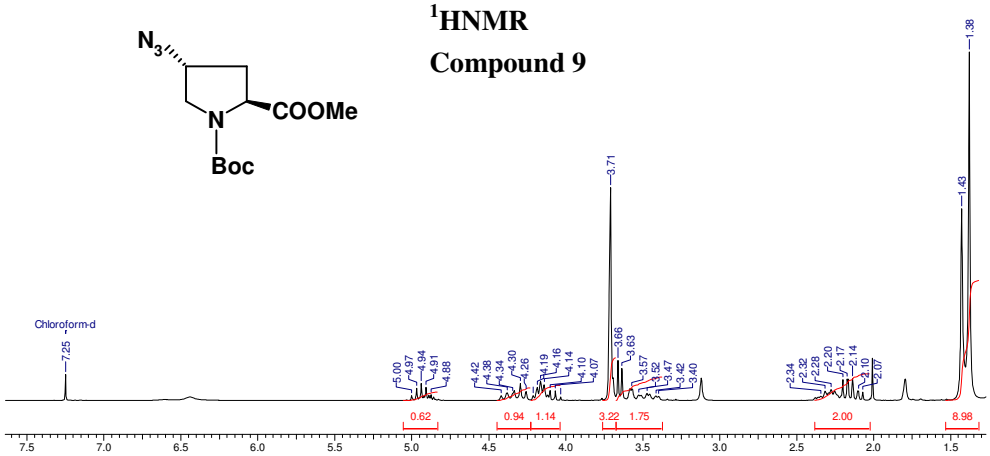
DEPT



Acquisition Time (sec)	7.9167	**Comment**	M Umashanakara		**Date**	13/08/2006 13:36:04	
Frequency (MHz)	200.13	**Nucleus**	1H	**Original Points Count** 32768	**Points Count**	32768	**Sweep Width (Hz)** 4139.07
Temperature (grad C)	0.000						

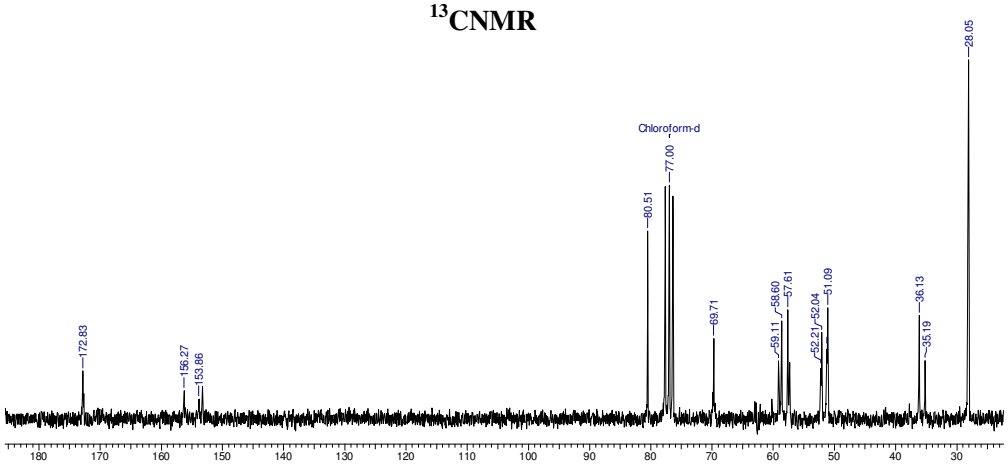
^{1}HNMR

Compound 9



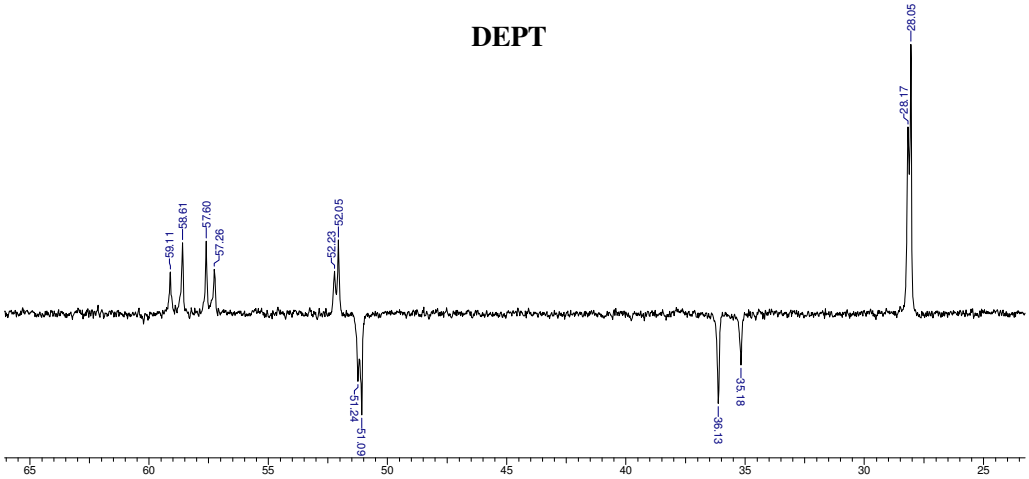
Acquisition Time (sec)	2.7329	**Comment**	M Umasankara		**Date**	15/08/2006 22:57:32	
Frequency (MHz)	50.32	**Nucleus**	13C	**Original Points Count** 32768	**Points Count**	32768	**Sweep Width (Hz)** 11990.41
Temperature (grad C)	0.000						

^{13}CNMR



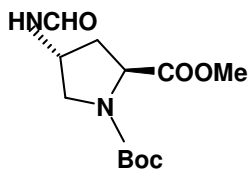
Acquisition Time (sec)	2.7329	**Comment**	M Umasankara		**Date**	15/08/2006 23:16:52	
Frequency (MHz)	50.32	**Nucleus**	13C	**Original Points Count** 32768	**Points Count**	32768	**Sweep Width (Hz)** 11990.41
Temperature (grad C)	0.000						

DEPT



Acquisition Time (sec) 1.3648	***Comment*** M UMASANKRA;			***Date*** 00/00/00 00:00:00			
Frequency (MHz) 300.13	***Nucleus*** 1H	***Number of Transients*** 256		***Original Points Count*** 8192		***Points Count*** 8192	
Sweep Width (Hz) 6002.40	***Temperature (grad C)*** 24.000						

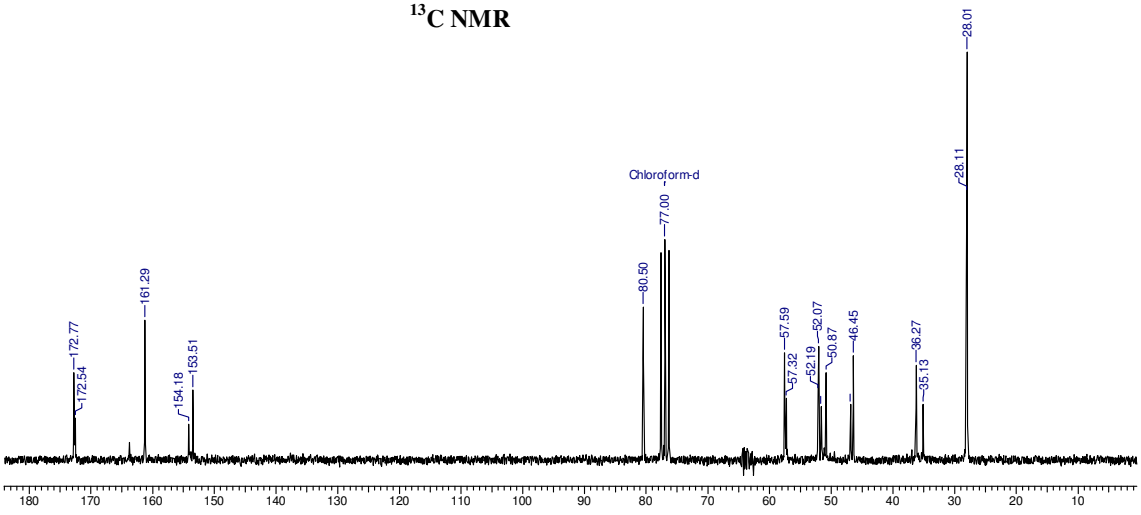
^{1}H NMR



Compound 10

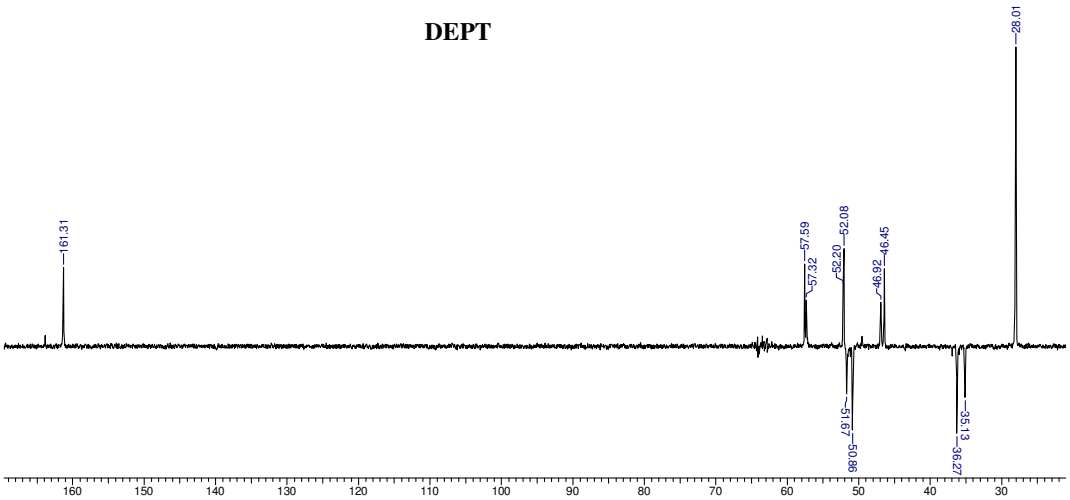
Acquisition Time (sec) 2.7329	***Comment*** m umasankara			***Date*** 25/06/2006 04:43:38			
Frequency (MHz) 50.32	***Nucleus*** 13C	***Original Points Count*** 32768		***Points Count*** 32768		***Sweep Width (Hz)*** 11990.41	
Temperature (grad C) 0.000							

^{13}C NMR



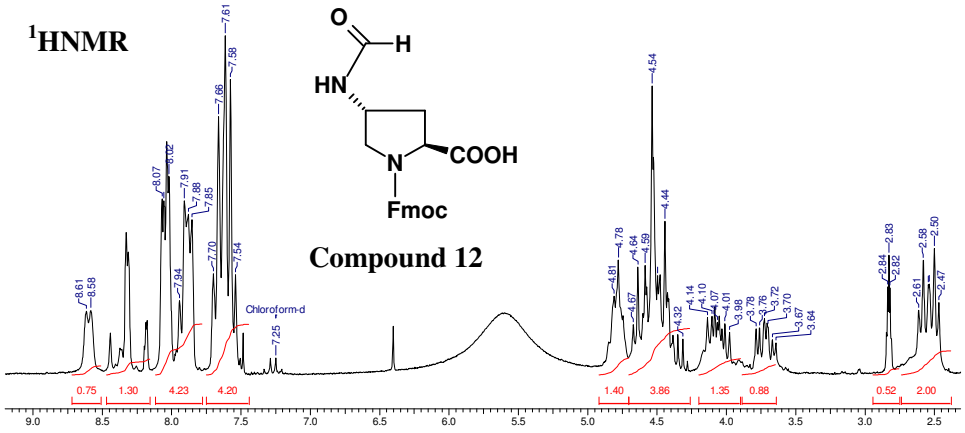
Acquisition Time (sec) 2.7329	***Comment*** m umasankara			***Date*** 25/06/2006 05:02:54			
Frequency (MHz) 50.32	***Nucleus*** 13C	***Original Points Count*** 32768		***Points Count*** 32768		***Sweep Width (Hz)*** 11990.41	
Temperature (grad C) 0.000							

DEPT



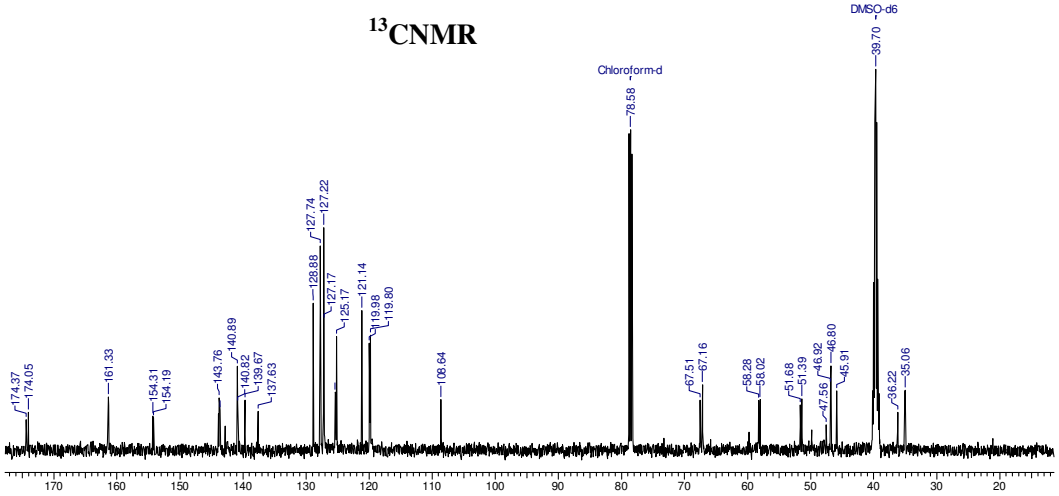
Acquisition Time (sec)	7.9167	**Comment**	M Umashankara		**Date**	20/08/2006 13:43:18	
Frequency (MHz)	200.13	**Nucleus**	1H	**Original Points Count** 32768	**Points Count**	32768	**Sweep Width (Hz)** 4139.07
Temperature (grad C)	0.000						

^{1}HNMR



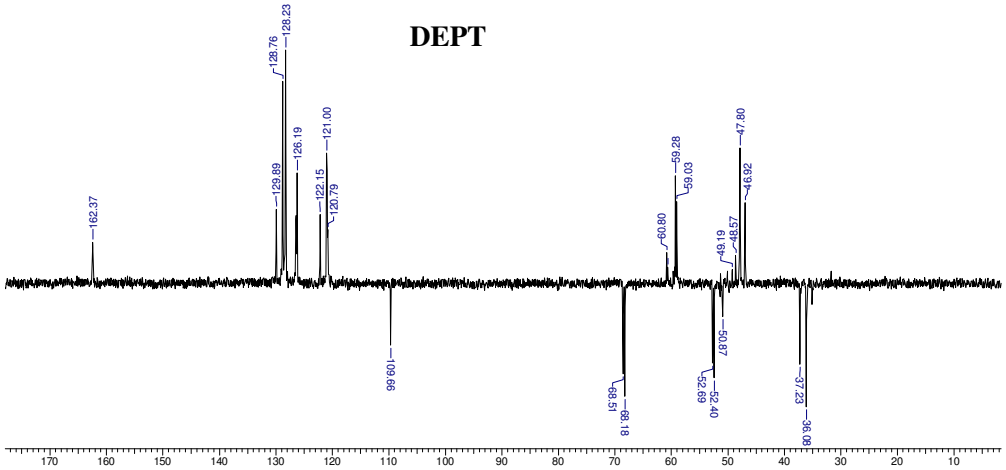
Acquisition Time (sec)	0.5210	**Comment**	13C	**Date**	30/08/2006 16:09:16	**Frequency (MHz)**	125.76
Nucleus	13C	**Original Points Count**	16384	**Points Count**	16384	**Sweep Width (Hz)**	31446.54
Temperature (grad C)	0.000						

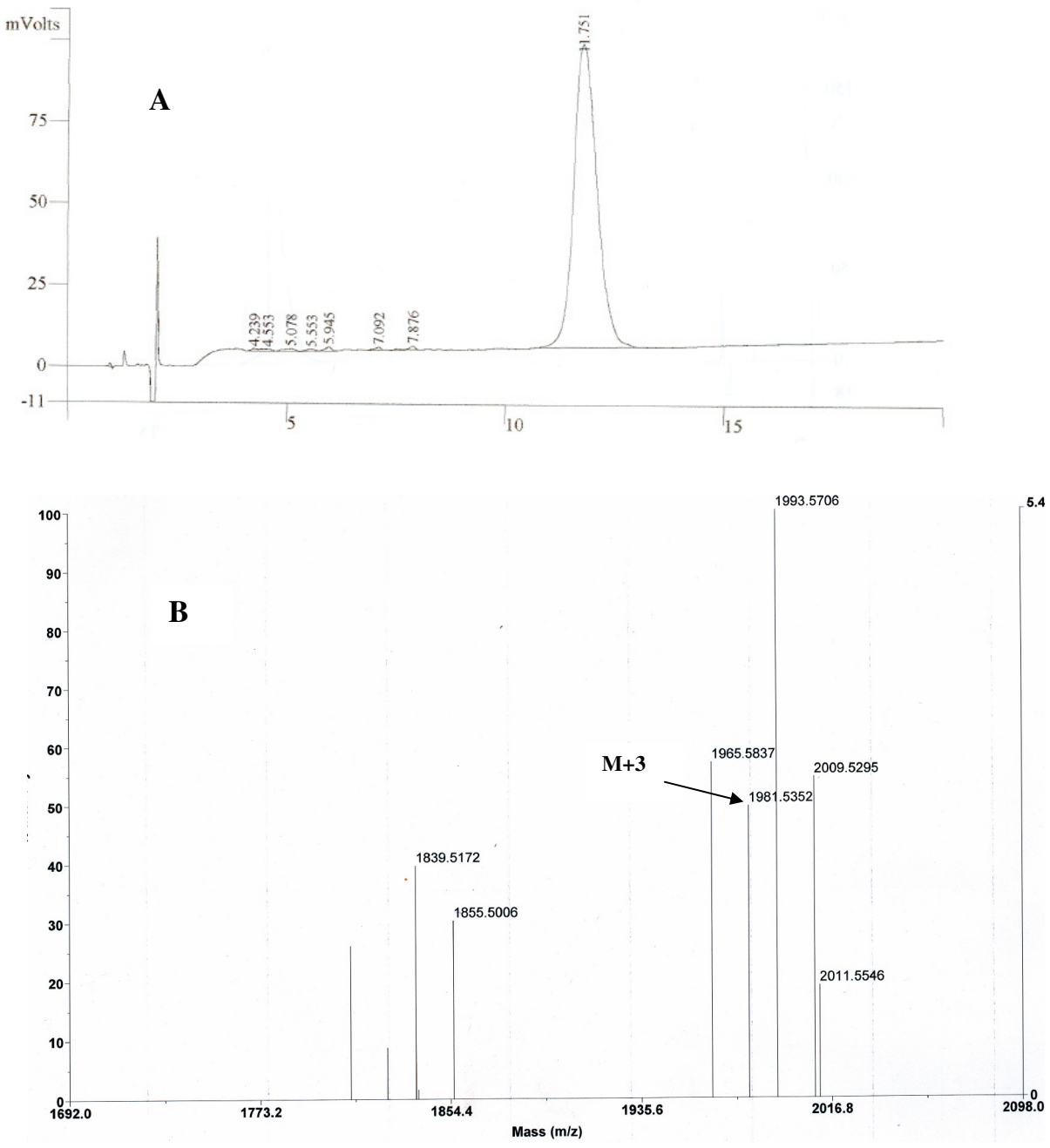
^{13}CNMR



Acquisition Time (sec)	1.0420	**Comment**	13C Dept		**Date**	30/08/2006 15:53:50	
Frequency (MHz)	125.76	**Nucleus**	13C	**Original Points Count** 32768	**Points Count**	32768	**Sweep Width (Hz)** 31446.54
Temperature (grad C)	0.000						

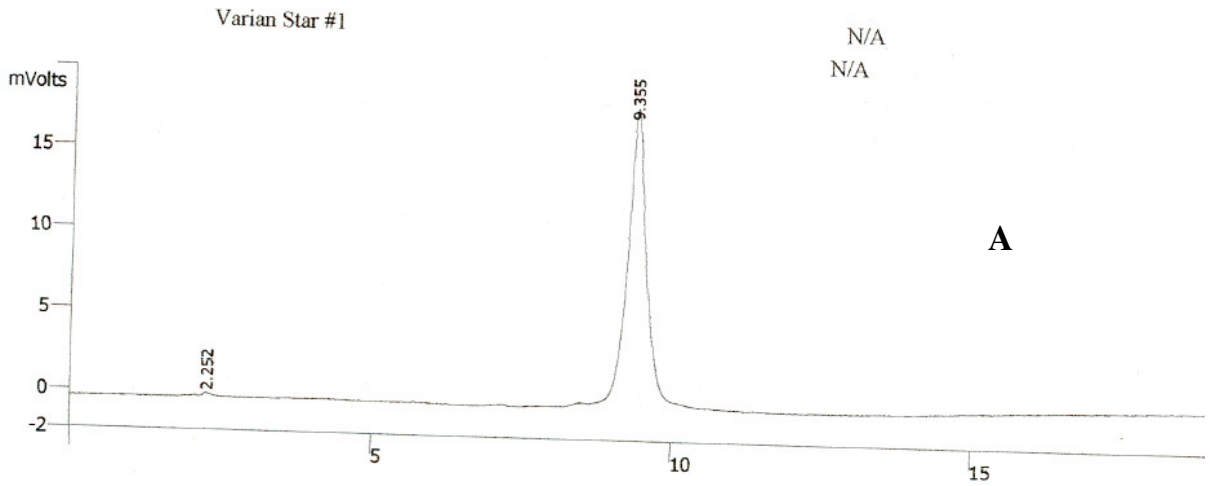
DEPT

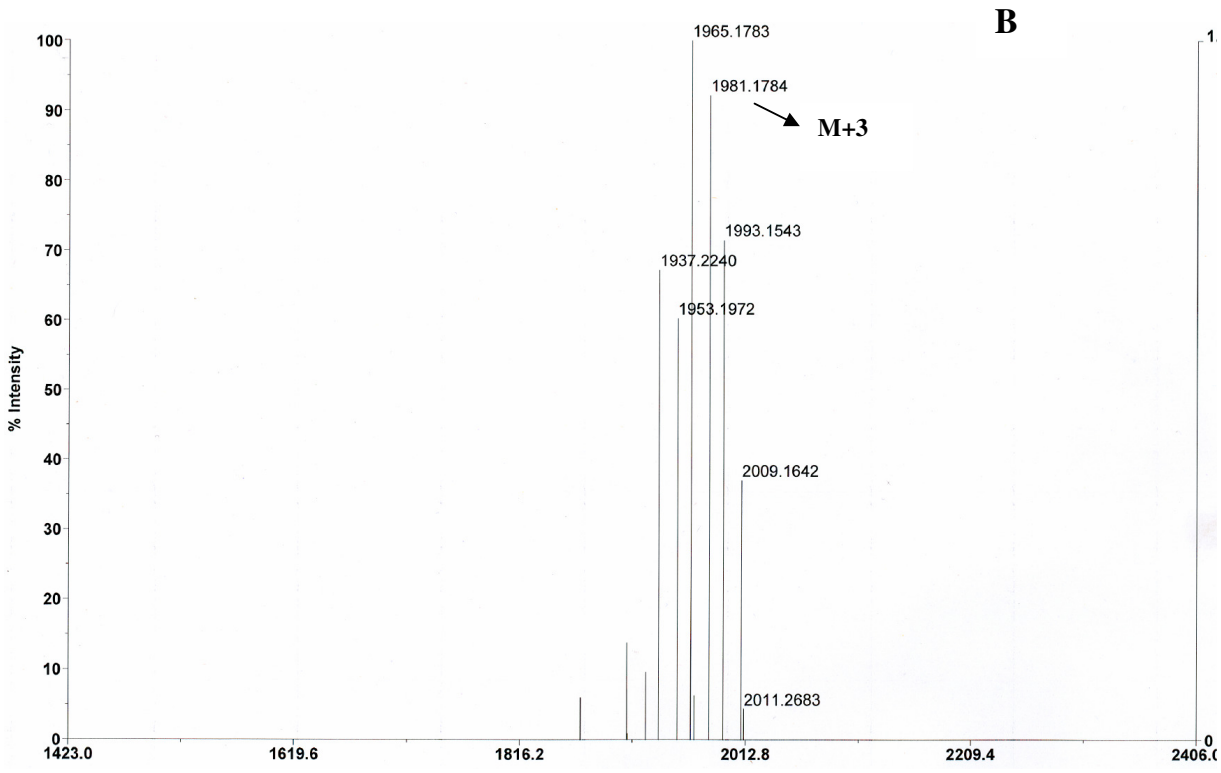




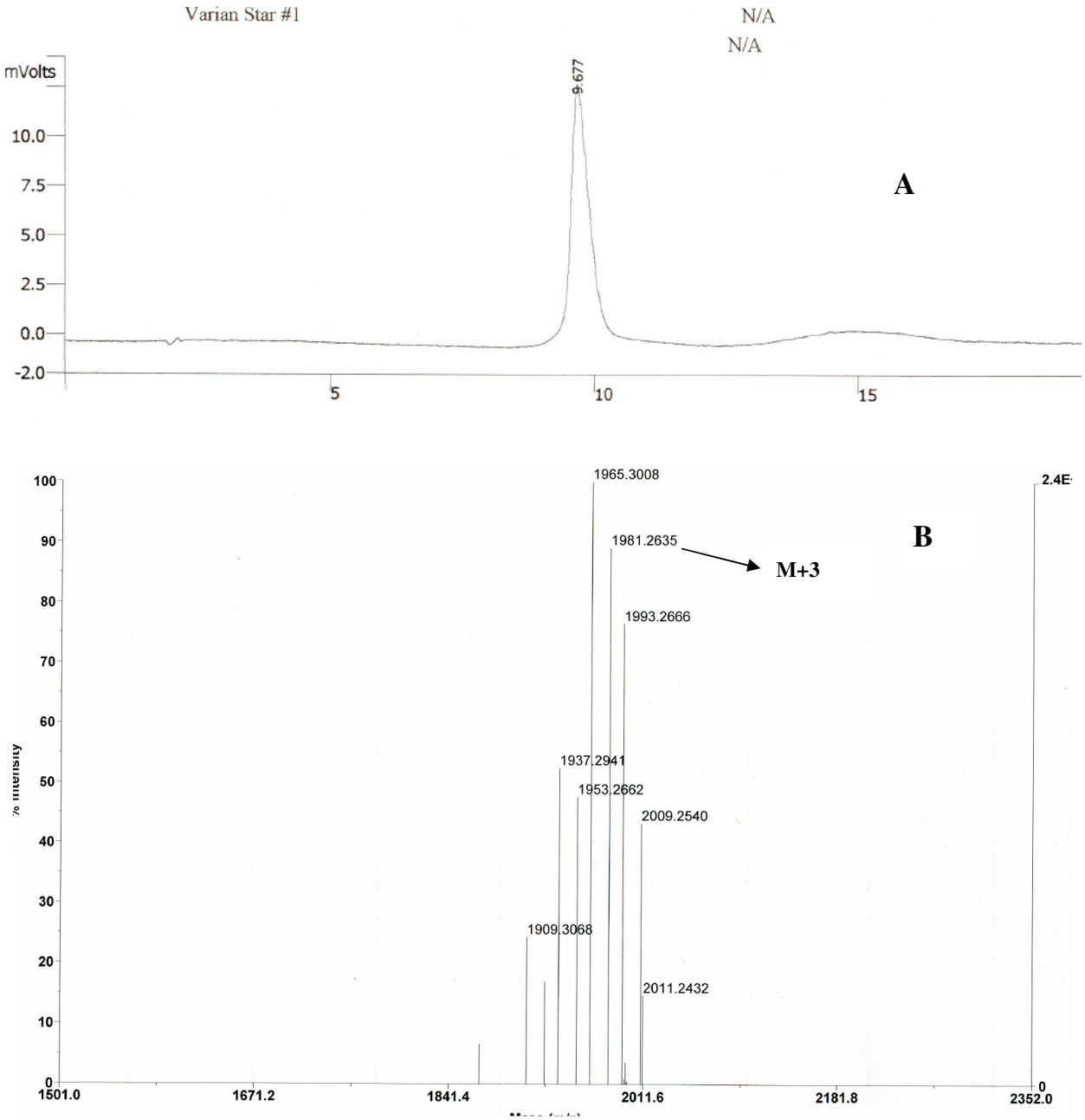
A; HPLC profile and **B**; MALDI-TOF mass spectrum of peptide **31**. Ac-Phe(Pro-fAmp-Gly)$_6$-NH$_2$

\

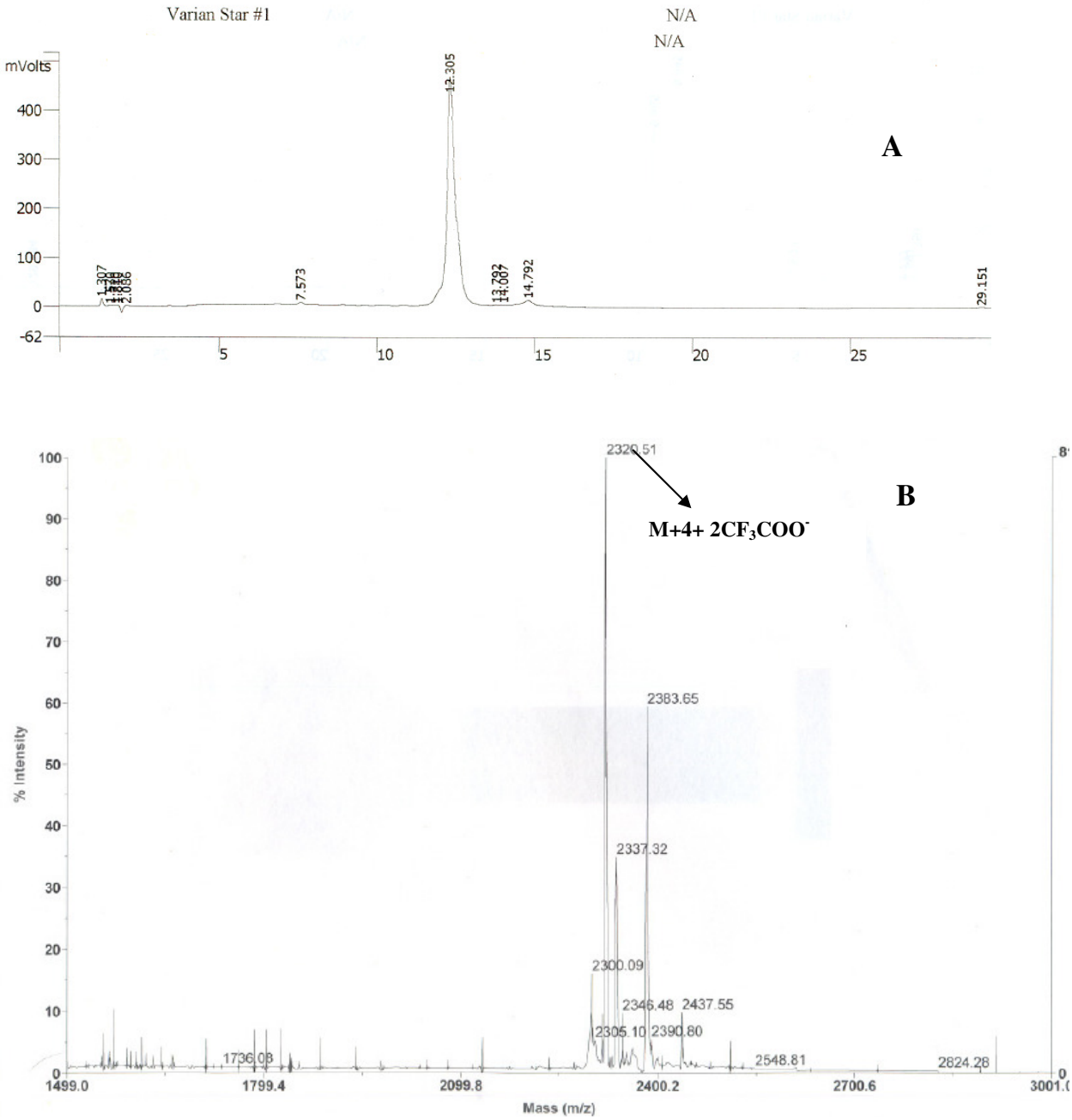




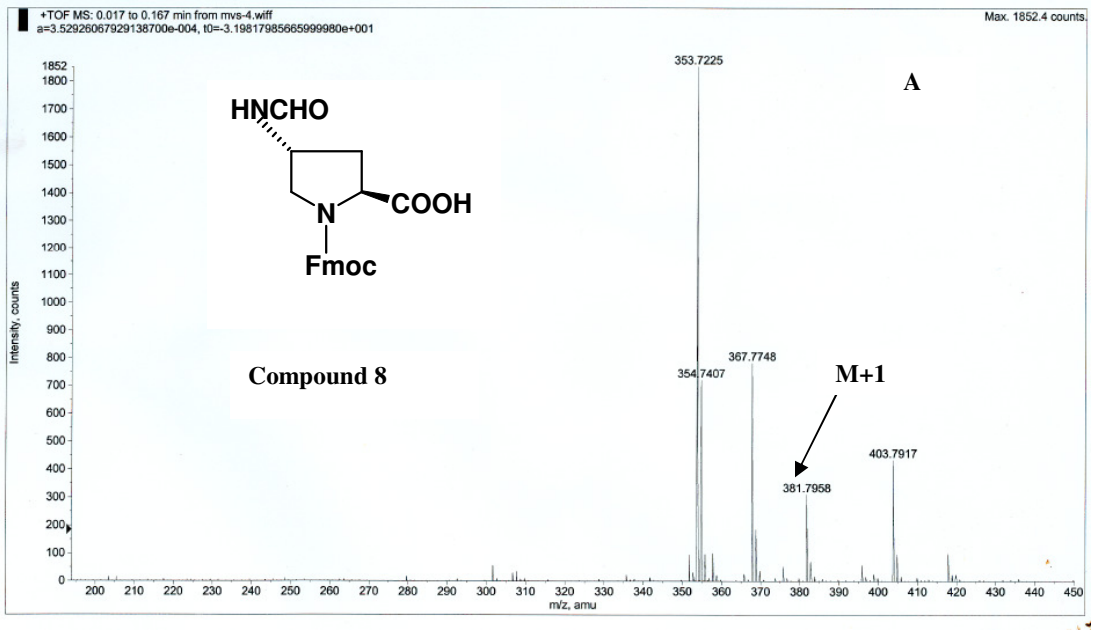
A; HPLC profile and **B**, MALDI-TOF mass spectrum of peptide **32**; Ac-Phe(fAmp-Amp-Gly)$_6$-NH$_2$

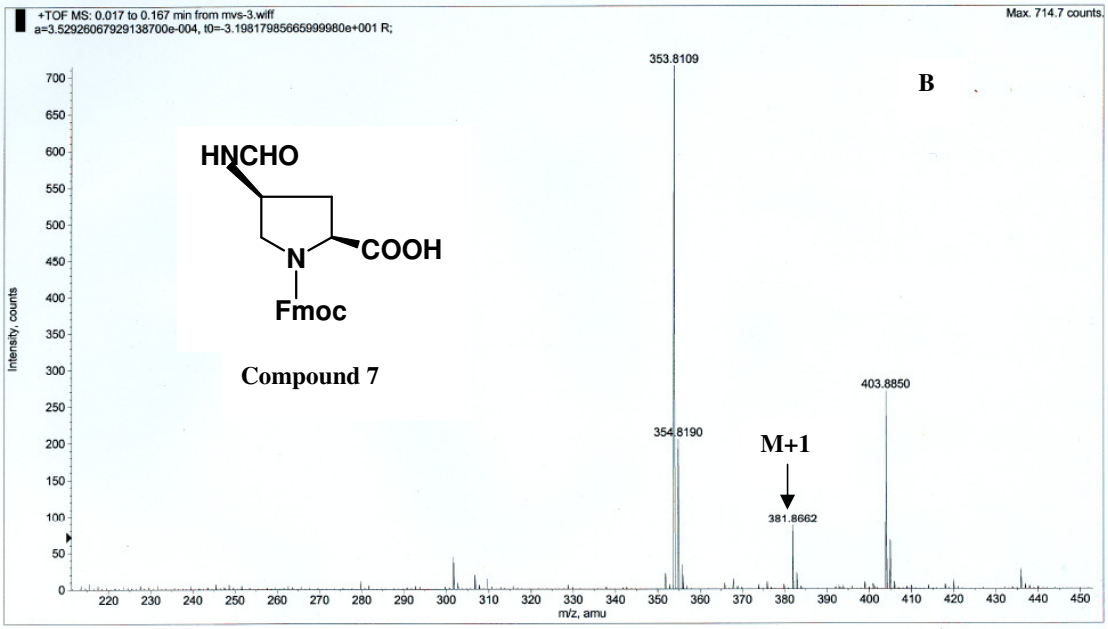


A; HPLC profile and **B**; MALDI-TOF mass spectrum of peptide **33**. Ac-Phe(famp-Pro-Gly)$_6$-NH$_2$



A, HPLC profile and **B**, MALDI-TOF mass spectrum of peptide **34**. Ac-Phe(famp-Amp-Gly)$_6$-NH$_2$

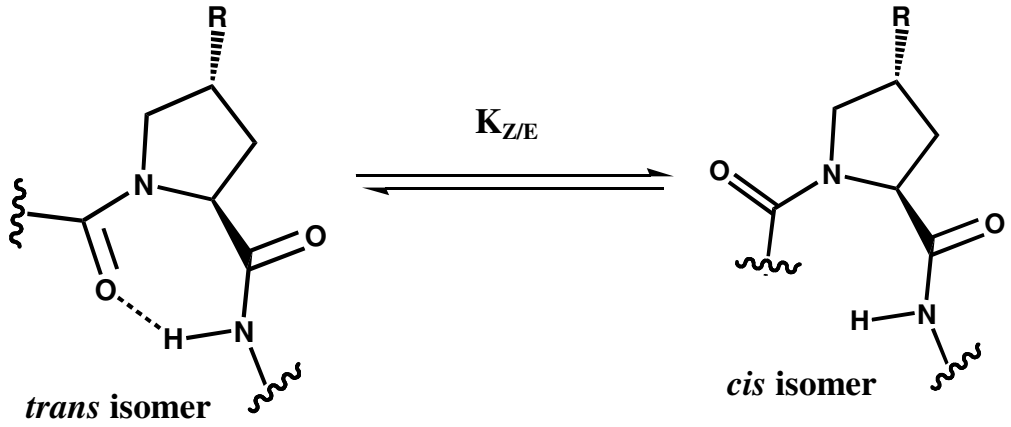




LC-MS mass spectra of **A**; compound **12** and **B**; compound **7**

Chapter 5

Prolyl- peptide- bond isomerization in 4*R*/*S* aminoproline: Effect of pH on the prolyl-peptide bond isomerization in aminoproline containing collagen peptides



trans isomer

cis isomer

5.1: Introduction

Proteins are long chain polypeptides that can adopt many different conformations; yet, the sequence of their aminoacid residues directs the folding to a particular native state conformation.[1] The loss of conformational entropy associated with such folding is expected to be unfavorable but is overcome by favorable non-covalent forces such as the hydrophobic effect, hydrogen bonds, electrostatic interactions and formation of covalent disulfide bonds.[2] Recently Raines et al[3]. have shown that an additional factor-the stereolectronic effect can contribute significantly to the conformational stability of the protein.

The structure and reactivity of an organic molecule also relies on the stereochemistry of its bonded and non-bonded electron pairs.[4] Stereolectronic 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.[5] For example, the multiple gauche effects (X-C-C-X) arising from 2` oxygen distinguishes stable conformations of RNA**.**RNA and RNA**.**DNA duplexes[6]. The anomeric effect (X-C-X) enhances the stability of the α-isomer of the glycosides.[6]

The common aminoacids L-serine and L-threonine, as well as 3(*S*)-hydroxy-L-proline and 5(*R*)-hydroxy-L-lysine (which are also found in collagen), all contain X-C-C-X systems that are subjected to a gauche effect. Suitable manipulation of these stereoelectronic effects with natural and non natural aminoacids enhances the stability of protiens.[7]

Study of pepdidyl-prolyl bond isomerization is intimately linked to the understanding of the folding and stability of the collagen. This knowledge has resulted in the design of novel aminoacids for use in collagen mimetics.[8] Incorporation of 4-fluroproline into *barstar* using a bacterial expression system is a demonstration of the importance of substituted prolines as tools for protein design and engineering.[9]

As demonstrated in earlier chapters, both the electronegativity and the stereochemistry of the C4 substituent of proline have a significant effect on T_m of collagen triple-helix. Larger T_m values of $(Pro\text{-}Yaa\text{-}Gly)_7$ triple-helices correlate with larger $K_{trans/cis}$ values of Ac-Yaa-OMe mimics (Table 5.1, section 5.02). It has been found that a C4 substituent can enhance the conformational stability by favoring the *trans* isomer, thereby preorganizing the individual strands to resemble more closely the strands in the triple-helix.

5.1.1: Factors affecting conformation of prolyl amide bond in proline-containing peptides.

Peptide bonds have partial double bond character which is explained by invoking the resonance between peptide carbonyl group and the lone pair electrons on nitrogen atom of amide group (Fig 5.1a). As a result of π-orbital overlap, and concomitant restricted rotation about the amide bond, there are two conformations which correspond to energy minima attained when the dihedral angles about the N-C=O bond (ω) is 0^o (*cis*) and 180^o (*trans*) (Fig 5.1b).

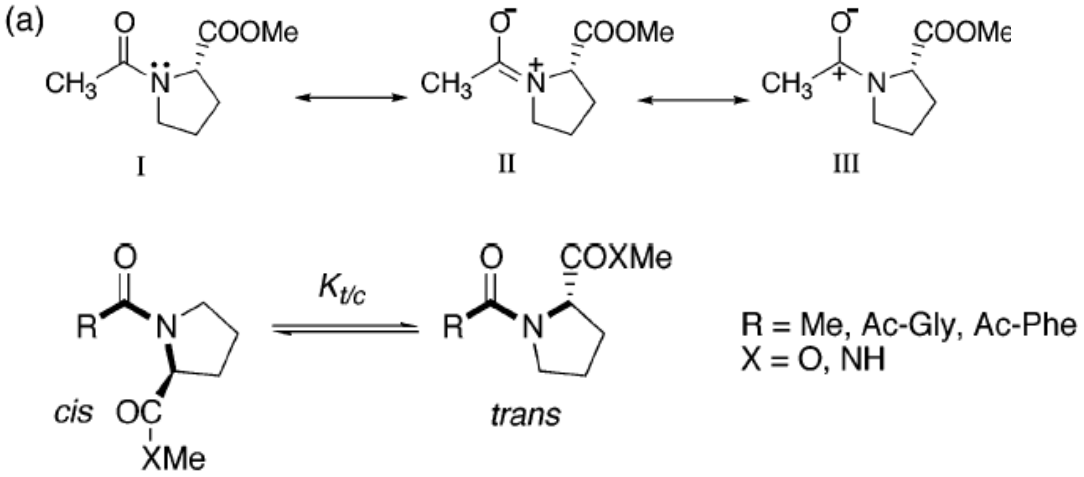


Figure 5.1: (a) Electron distribution in peptide bond; (b) *cis-trans* conformations of peptide bond.

protein structure and function. The cyclic nature of the side chain means that *cis* and *trans* conformations are closer in energy for prolyl amides than for other peptide bonds. For these two conformations to interconvert, the nitrogen must become transiently sp^3-hybridized, i.e., pyramidalized.[10,11]

Various factors are responsible for the *cis-trans* equilibrium of prolyl-peptide bond, and are studied by NMR, IR and theoretical calculations using quantum mechanical methods. Except in theoretical studies,[12,13] C-terminal amides (**2**) have rarely been used in conformational studies,[14] to avoid complications arising from hydrogen bonding. Evidence has been presented for hydrogen bonds in both *cis*[15] and *trans*[16] conformations in non-hydrogen-bonding solvents.[17] Such effects are less significant in an aqueous environment.

Zimmerman and Sheraga[18] have proposed that restricted rotation about the N-Cα bond leads to a favorable C=O----C=O electrostatic interaction. Meallum et. al. [19] have demonstrated that this is almost as strong (80%) as backbone hydrogen bond in a

protein. This stabilization involves n→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. Raines et al.[20] have recently provided evidence for the significance of this interaction, which they describe as quantum mechanical rather than electrostatic and estimate a contribution of 0.7 kcal mol^{-1} to the stability of the *trans* conformation at 298K.

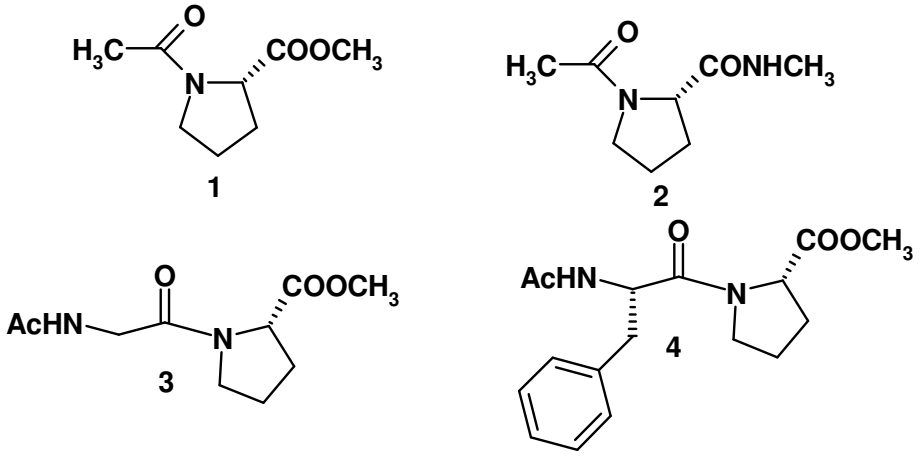


Figure 5.2: proline derivatives.

The addition of a glycine residue to the N-terminal proline residue (**3**)(Fig 5.2) leads to an increase in $K_{trans/cis}$ ratio.[21] This significant shift in favor of the *trans* conformation is a consequence of the increase in steric bulk. There is a high propensity for *cis* amide bond when proline is preceded by an aromatic residue.[22] Survey of crystallographic databases reveals that 5.7% of peptide bond is in the cis conformation in X-Pro dipeptide. When X is Phe (**4**) this percentage rises to 6.7% and leaps to 19.1% for Tyr.[23,24] This is attributed to a stabilizing Ar-Pro interaction in the cis conformation. Halab and Lubell[25] have suggested that the Ar-Pro interaction is cationic-pi in nature[26] (Fig 5.3). Evidence for this nonbonding interaction is found in the ^1H chemical shifts for Hα of the Pro residue. In the cis conformation, Hα is shielded by the aromatic ring and a dramatic upfield shift is observed.

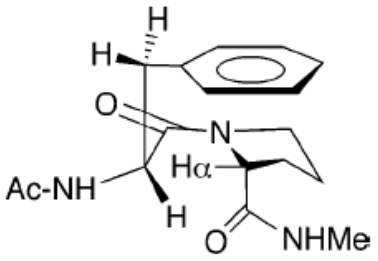
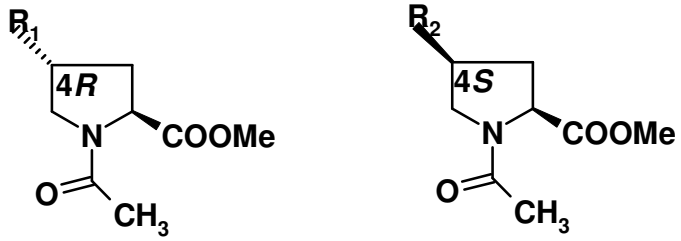


Figure 5.3: Ar-Pro interaction.

5.1.2: Conformational control of pyrrolidine ring

Among the proteinogenic amino acids, only proline is a secondary amine and has a saturated ring.[27] As a secondary amine, proline has a much greater propensity than other natural amino acids to form *cis* (i.e. *E*) peptide bonds.[28-30] In most peptide bonds, the trans (*Z*) isomer is favored over the *cis* (*E*) isomer. In contrast, the *trans* isomer of a prolyl peptide bond is only slightly favored over the *cis* isomer. The interconversion of *cis* and *trans* isomers about prolyl peptide bonds has been identified as the rate determining step in the protein folding pathways[31] including collagen.[32-34] Because all peptide bonds residing in the collagen triple-helix are in the *trans* conformation, variety of methods have been developed to control the *cis/trans* ratio of prolyl peptide bond, which includes buttressing 2, 3, and 5 positions of proline with a functional group,[35-40] replacing the peptide bond with an alkene isostere[41-46] including the amide in a ring system that is fused to the pyrrolidine ring[39] (Fig 5.4) and introducing bridge in the pyrrolidine ring.[47]

Among these, substitutions on C-4 of proline residues are known to have a large effect on the *trans/cis* ratio.[48-51] For example, the electronegative substituents in the 4*R*-position of the proline increase the stability of *trans* isomer, whereas electronegative substituents in the 4*S*-position decrease that stability (Table 5.1).



1 R_1= OH
2 R_1= F
3 R_1= CH_3
4 R_1= $C(CH_3)_3$

5 R_2= OH
6 R_2= F
7 R_2= CH_3
8 R_2= $C(CH_3)_3$

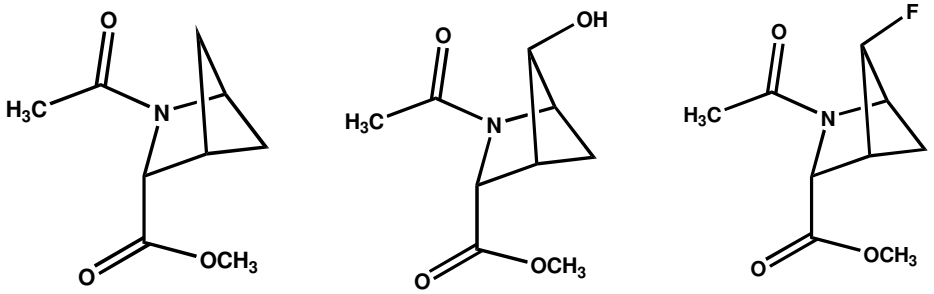


Figure 5.4: Proline derivatives

NMR analysis indicate that Ac-(2*S*,4*R*)-4-fluoroproline-OMe (**2**) resides predominantly (86%) in the C4-exo pucker in the solution, whereas Ac-(2*S*,4*S*)-4-fluroproline-OMe (**6**) is found almost exclusively (95%) in the C4-endo pucker. This dichotomy can be attributed to a conformation of the pyrrolidine ring that places the nitrogen and fluorine in a gauche orientation having important ramifications for the stability of collagen. It is important to mention that the magnitude of $K_{trans/cis}$ and the puckering preference of the pyrrolidine ring are mainly dependent on the elecronegativity and the stereochemistry of the 4-substituent.

Table 5.1: Effect of C4-substituent on $K_{trans/cis}$ of Ac-Xaa-OMe.

1	6.1	C4-*exo*
2	6.7	C4-*exo*
3	-	C4-*exo*
4	-	C4-*exo*
5	2.4	C4-*endo*
6	2.5	C4-*endo*
7	-	C4-*endo*
8	-	C4-*endo*
9	3.5	C4-*endo*
10	3.6	C4-*endo*
11	3.5	C4-*endo*

5.2: Rationale and Objectives of the present work

Previous chapters have explained the remarkable triple-helix forming abilities of 4*R*-aminoproline and 4*S*-aminoproline containing collagen model peptides. Interestingly 4-aminoproline containing collagen peptides form pH dependent triple-helical structure. The stability of triple-helix is much stronger in acidic condition compared to alkaline condition. This indicates that the protonation of 4-amino group of aminoproline has important role in stabilizing the triple-helix structure. This is also evidenced by the result of Chapter 4 in which the 4-formyl protected aminoprolines form less stable and pH independent triple-helical structures. It is well known that the electronegative 4-substituent on proline stabilizes the collagen triple-helix by preorganizing the favored *trans* conformation of peptide bond. In case of aminoproline, it is possible that the protonated and nonprotonated form of aminoprolines have different magnitude of $K_{trans/cis}$ ratio and this may result in the formation of pH dependent triple-helices, because the protonated and free anino group have different electronagativity. In order to understand the relative protonation nonprotonation, stereochemical, and electronic influences of 4-aminoproline on the *E-Z* equilibrium of prolyl-peptide bond and conformation of the pyrrolidine ring, it was thought to examine the 4-aminoproline

model compounds of the type Ac-Xaa-OMe. For this purpose amino and azido groups were chosen as 4-subsitituents on proline ring, in both 4*R* and 4*S* conformations.

This chapter describes the synthesis and comparative study of thermodynamics of *E-Z* isomerization of prolyl-peptide bond in N^1-acetyl-4*R*-aminoproline methylester (Ac-Amp-OMe **12**), N^1-acetyl-4*S*-aminoproline methylester (Ac-amp-OMe **13**), N^1-acetyl-4*R*-azidoproline methylester (Ac-Azp-OMe **14**) and N^1-acetyl-4*S*-azidoproline methylester (Ac-azp-OMe **15**) at both acetic and basic conditions. IR-spectroscopy and variable temperature (VT) ^{1}H NMR spectroscopy have been used to study the steric and electronic effect of $–NH_2$, $-NH_3^+$ and $–N_3$ on the *E-Z* equilibrium of prolyl-peptide bond.

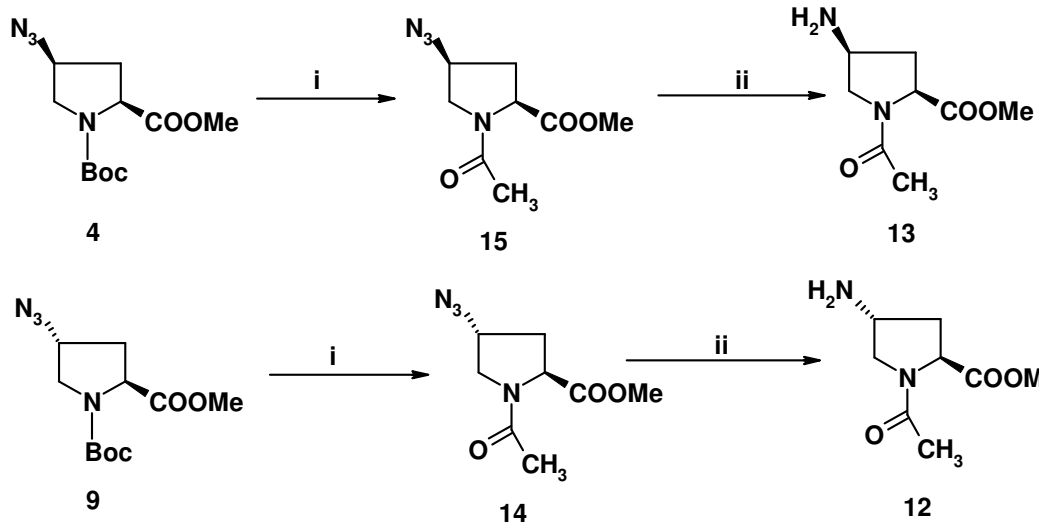
The specific objectives of this chapter are

1) Synthesis of 4-substituted proline model compounds **12-15**.
2) IR spectroscopic study of N-acetyl and C=O ester in protonated and nonprotonated 4-aminoprolines.
3) Variable temperature ^{1}H NMR study at acidic and basic conditions to estimate the pyrrolidine ring conformation and $K_{trans/cis}$ ratio.

5.3: Results

4-Substituted proline model compounds of type Ac-Xaa-OMe **12-15** were synthesized according to the Scheme 5.1**.** To avoid γ-turn formation with the nitrogen[52] (Fig 5.5) and to enable a direct comparison with the available literature data for similar model compounds,[53,54] methylester was chosen instead of N-methylamide for IR and ^{1}H NMR studies.

Scheme 5.1



Reagents and conditions: i) (a) TFA: DCM (1:1), (b) Acetic anhydride, DCM, Pyridine (1: 2: 2); ii) Pd/C H_2, MeOH.

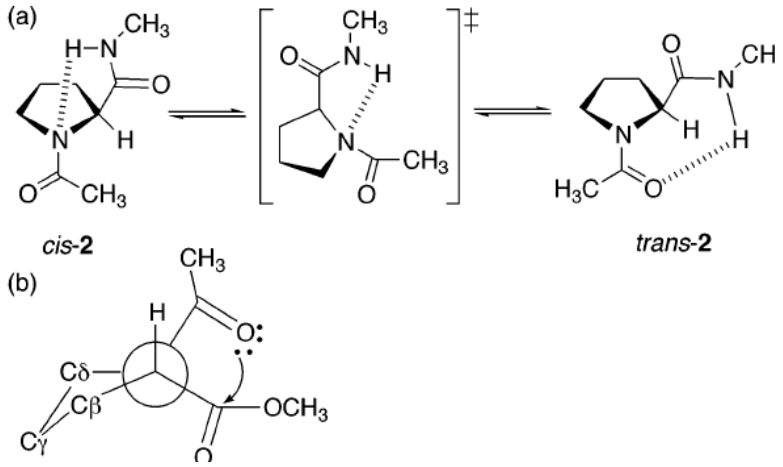


Figure 5.5: Intramolecular forces; (a) Hydrogen bonding and (b) n→π* interaction (looking down the Cα-N bond)

5.4: Study of conformational changes of model compounds 12-15

5.4.1: Inductive effect of 4-substituents on the prolyl-peptide bond-order in IR-Spectroscopic studies

As already mentioned, the gauche effect remarkably decreases the double bond character of the amide nitrogen of the proline ring in 4-substituted prolines. As a result,

the nitrogen acquires more sp^3 hybrid character which increases the amide C=O bond order.[55] Such an increase in the bond order results in a corresponding increase in the IR-vibration frequency of C=O ($\gamma_{amide\text{-}I}$). For example, when the 4-substituents are H, OH and F (Ac-Pro-OMe, Ac-Hyp-OMe and Ac-Flp-OMe) the observed amide-I vibrations were maximal at 1608.1cm^{-1}, 1613.1cm^{-1} and 1616.0cm^{-1} respectively,[54] which is same as their order of increasing electronegativities. Similar changes are also observed in the stretching vibration frequency of ester carbonyl[56] (γ_{ester}).

Figure 5.6 A and B show the amide-I and ester regions of IR-spectra of 0.1M aqueous solutions of compound **12** Ac-Amp-OMe at pH 2.0 and 12.0 respectively. The change in protonated to nonprotanated form remarkably affects the carbonyl vibration frequencies of both amide and ester carbonyls. The ester and amide carbonyl absorption frequencies of compound **12** (Ac-Amp-OMe) in protonated state appearing at 1747.39cm^{-1} and 1650.95cm^{-1} respectively is shifted to 1714.6 cm^{-1} and 1614.3 cm^{-1} for the nonprotonated state. Two well distinguishable absorption bands for the amide and ester carbonyl groups of compound **12** appear in the protonated form (Fig 5.6A). For the free amine (nonprotonated form) the amide carbonyl group absorption band is broad and merges with the ester carbonyl absorption band.

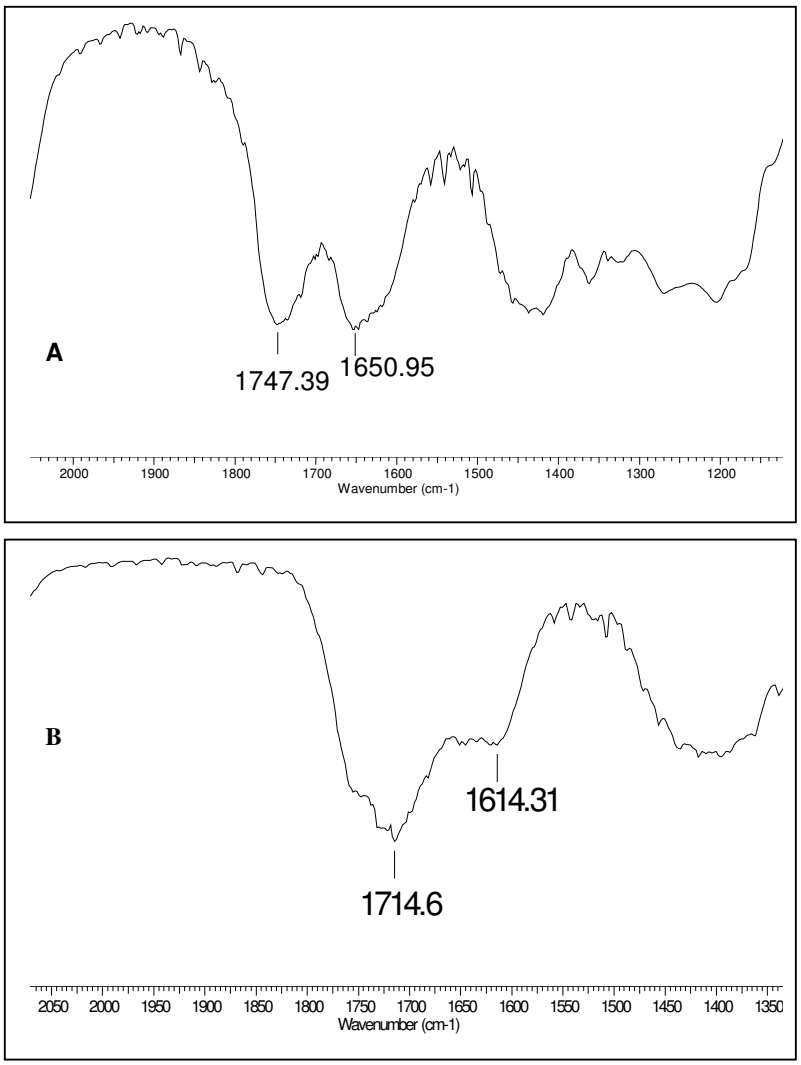


Figure 5.6: Amide-I and ester region of IR-spectra of compounds, **A**; Ac-Amp^{+}-OMe, **B**; Ac-Amp-OMe at pH 2.0 and 12.0 adjusted with acetic acid and NEt_3 respectively.

Figure 5.7 A and B show the amide-I and ester regions of IR-spectra of 0.1M solutions of compound **13** Ac-amp-OMe at pH 2.0 and 12.0 respectively. The ester and amide carbonyl absorption frequencies of compound **13** at protonated state appearing at 1747.4cm^{-1} and 1652.9cm^{-1} respectively is shifted to 1652.9 cm^{-1} and 1618.4 cm^{-1} for the nonprotonated state. Similar to the compound **12** the two well distinguishable absorption bands of amide and ester carbonyl groups appearing in the protonated form (Fig. 5.7A), merges completely lending to a single broad single absorption band in the nonprotonated form.

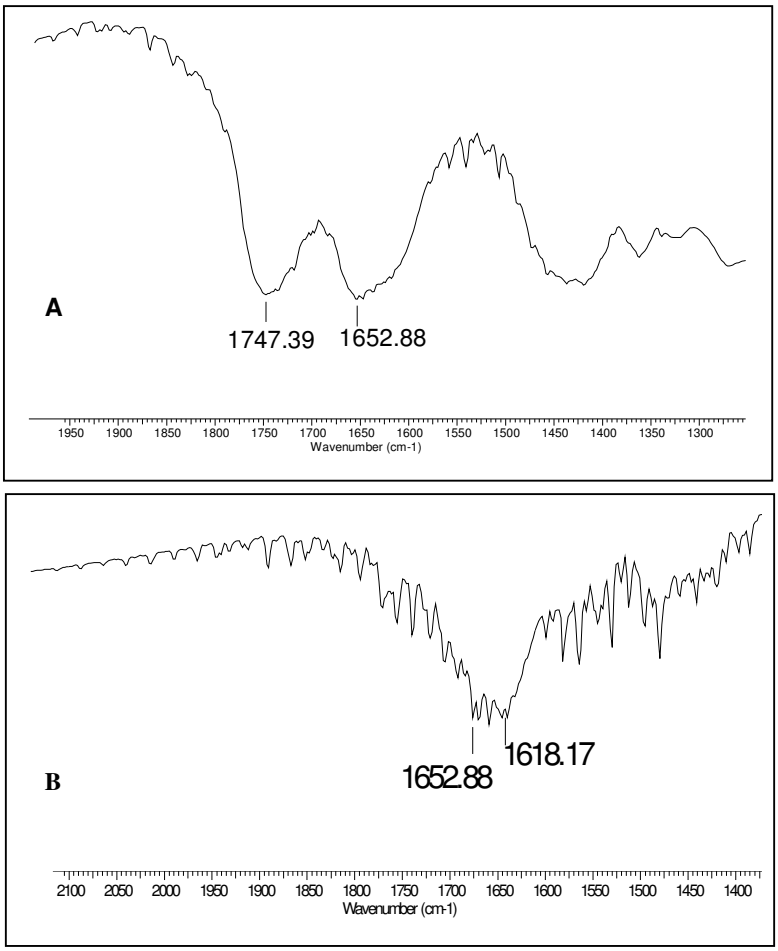


Figure 5.7: Amide-I and ester region of IR-spectra of compounds, **A**; Ac-amp^{+}-OMe, **B**; Ac-amp-OMe at pH 2.0 and 12.0 adjusted with acetic acid and NEt_3 respectively.

Figure 5.8 A-D show the amide-I and ester regions of IR-spectra of 0.1M aqueous solutions of compounds **14** Ac-Azp-OMe and **15** Ac-azp-OMe at pH 2.0 and 12.0 respectively. In constant to the IR-spectrum of compound **12** Ac-Amp-OMe and **13** Ac-amp-OMe the amide and ester carbonyl absorption frequencies of compound **14** Ac-Azp-OMe and **15** Ac-azp-OMe both at pH 2.0 and 12.0 appear at almost same wave numbers. The corresponding absorption frequencies of amide and ester carbonyl groups of compounds **12-15** along are listed in Table 5.2

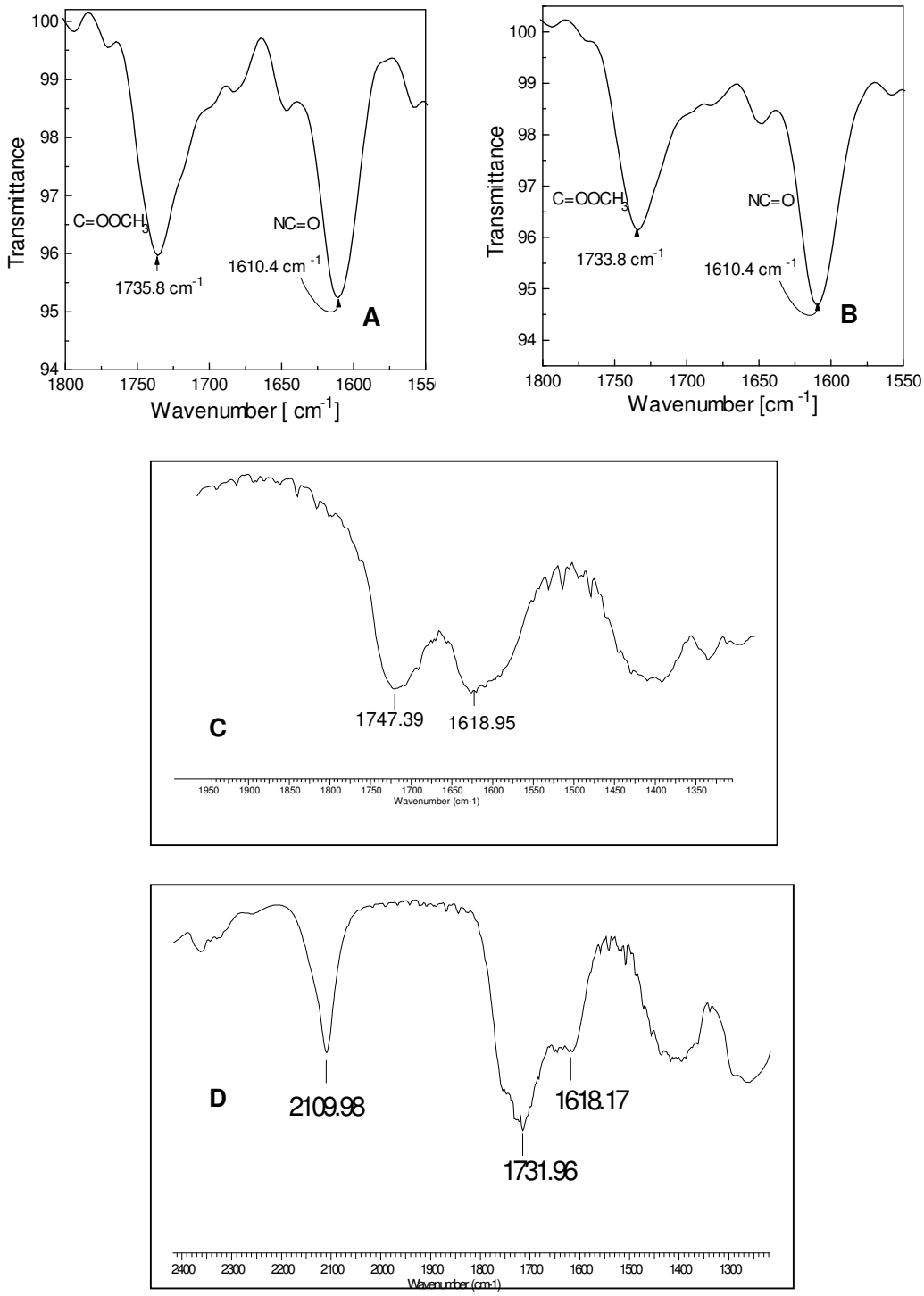


Figure 5.8: Amide-I and ester region of IR-spectra of Ac-Azp-OMe **A** and **D** and Ac-azp-OMe **B** and **C** at pH 2.0 and 12.0 respectively.

Table 5.2: Amide-I ($\gamma_{amide\text{-}I}$) and ester ($\gamma_{ester\text{-}I}$) vibrational frequencies of compounds **12-15** in Methanol at †; pH 2.0 and ‡; at pH 12.0.

Compounds	$\gamma_{amide\text{-}I}$ [cm^{-1}]	$\gamma_{ester\text{-}I}$ [cm^{-1}]
Ac-Hyp-OMe	1613.1*	-
Ac-Amp^{+}-OMe **12**†	1650.9	1747.4
Ac-Amp-OMe **12**‡	1614.3	1714.6
Ac-amp^{+}-OMe **13**†	1652.8	1747.4
Ac-amp-OMe **13**‡	1618.1	1652.8
Ac-Azp-OMe **14**†	1610.4	1735.8
Ac-Azp-OMe **14**‡	1618.1	1731.9
Ac-azp-OMe **15**†	1610.4	1733.8
Ac-azp-OMe **15**‡	1618.9	1747.3

* From reference 34, † at pH 2.0, ‡ at pH 12.0

5.4.2: Circular dichroism (CD) spectroscopic studies

When a molecule exists in more than one conformation in solution, each conformer will have its own ORD and CD curve and the sign and magnitude of the cotton effects will change with a shift in conformer population, caused either by altering solvent polarity or changing temperature. In the above IR study both amide and ester carbonyl vibrational frequencies shift with change in pH. This may due to change in the conformation of pyrrolidine ring as a function of pH. In order to study the pH effect on the conformation of the 4-aminoproline, CD spectral study was carried out.

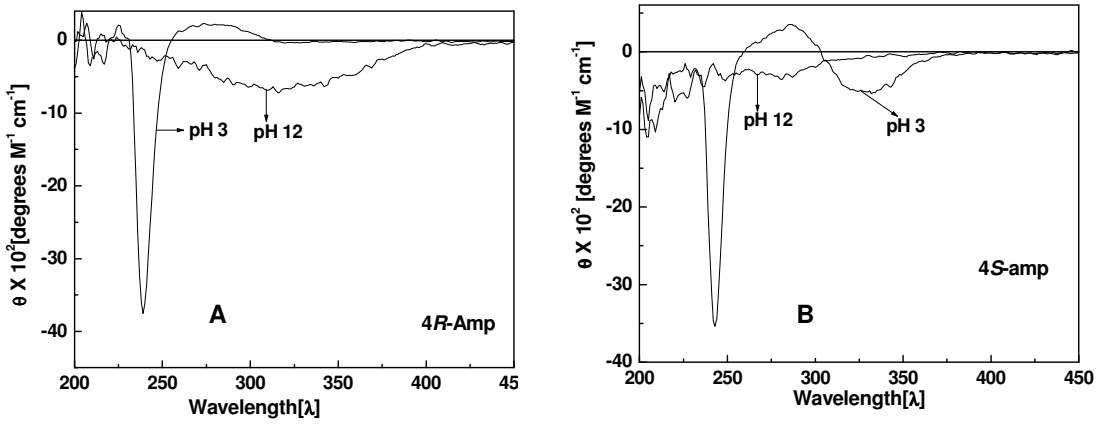


Figure 5.9: CD spectra of 0.1M soulutions of **A**; compound **12** (Ac-Amp-OMe), **B**; compound **13** (Ac-amp-OMe) taken at pH 2.0 (20mMacetate buffer) and at pH 12 (20mM borate buffer).

Figure 5.9 A-B show the CD spectra of 0.1M aqueous solutions of compound **12** Ac-Amp-OMe and **13** Ac-ame-OMe taken at pH 3.0 and 12.0 respectively. The CD spectrum of compound **12** Ac-Amp-OMe at pH 3.0 shows small positive peak at 275nm and a sharp and large negative at 245nm. At pH 12.0 this compound shows only a broad negative band with minima around 320nm. In contrast, compound **13** Ac-amp-OMe at pH 3.0 shows a small positive peak at 290nm and two negative peaks at 325nm (minor) and at 245nm (sharp and large). The CD spectral signal at pH 12.0 does not show any positive or negative bands.

5.4.3: NMR spectroscopy study

The identity of ring-pucker can be characterized in a qualitative manner by the distinct coupling pattern for $J_{H\text{-}H}$ in each type of ring-pucker (for numbering see the Fig 5.10). For example, in Cγ-*exo* pucker $J_{1\text{-}2}$ and $J_{1\text{-}3}$ corresponding to Cα-Cβ bond torsion angles are larger and nearly equal, resulting in a triplet like appearance for the H_1 signal in the ^{1}H NMR spectrum. Similarly for Cγ-*endo* pucker, $J_{1\text{-}2}$ and $J_{1\text{-}3}$ are small, resulting in a doublet of doublet for H_1 signal. An intermediate behavior can be identified from the relative magnitudes of $J_{1\text{-}2}$ and $J_{1\text{-}3}$ values.[57]

Figure 5.11 shows H_1 (Hα) region of the 400MHz ^{1}H-NMR spectra of compound **12** and **13** in D_2O at pH 2.0 and 12.0. For H_1 (Hα), two sets of signals major (*ma*) and minor (*mi*) arising from Z and E isomers are observed. The $J_{H\text{-}H}$ values for each isomer arise from their corresponding ring-pucker preferences, which are independent of the other isomer. H_1(*ma)* signal for compound **12** (Ac-Amp-OMe) at pH 2.0 (Fig 5.12A) shows a triplet like appearance, while that in at pH 12.0 (Fig 5.12B) shows quartet like appearance. Similarly H_1*ma* signals for a compound **13** (Ac-amp-OMe) at pH 2.0 (Fig 5.12C) show doublet of

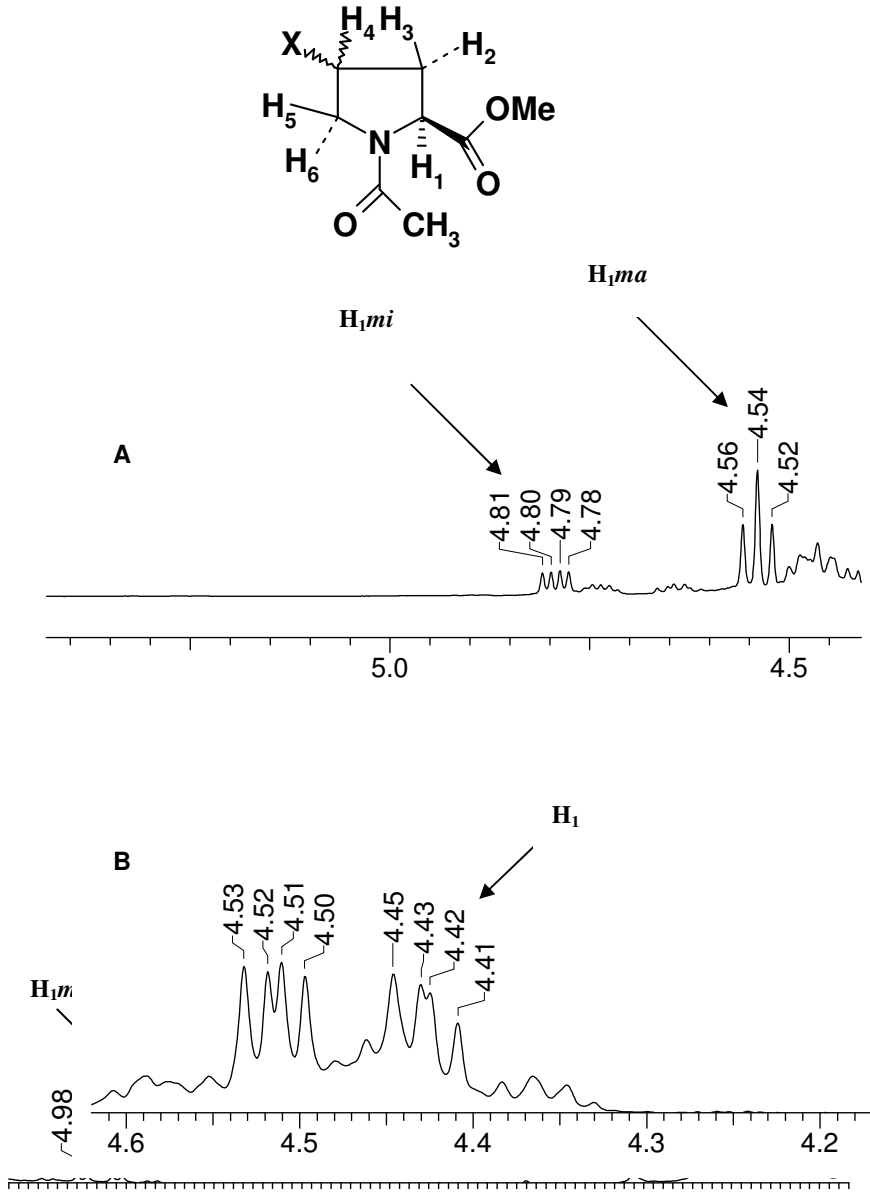


Figure 5.11: Expanded 400MHz ^{1}H-NMR spectrum of Ac-Amp-OMe **12** D_2O at, **A** ; pH 2.0 and **B**; at 12.0 respectively. *ma* is the major isomer and *mi* is the minor isomer.

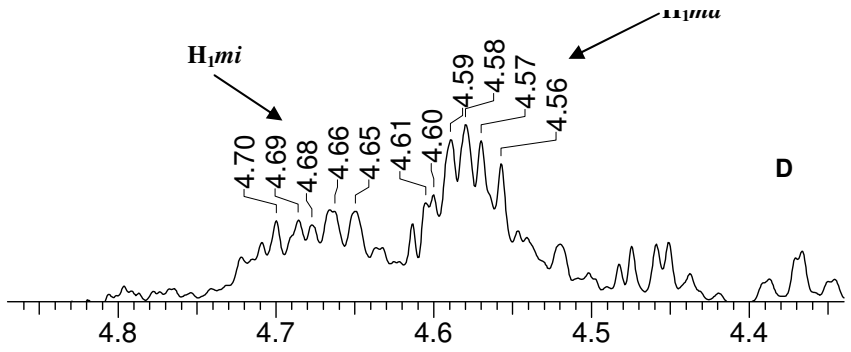


Figure 5.12: Expanded 400MHz ^{1}H-NMR spectrum of Ac-amp-OMe **13** in D_2O at, **C**; pH 2.0 and **D**; at 12.0 respectively. *ma* is the major isomer and *mi* is the minor isomer.

doublets (dd), while that in at pH 12.0 (Fig 5.12D) shows multiplet pattern. In contrast, H1(ma) signals for compounds **14** (Ac-Azp-OMe) (Fig 5.13) and compound **15** (Ac-azp-OMe) (Fig 5.14) at both pH 2.0 and pH 12.0 appear as triplet and doublet of doublets (dd) respectively.

The pattern of H_1 signals for compound **12** Ac-Amp-OMe and **13** Ac**-**amp-OMe which contain free NH_2 group at 4-position, changes with pH, while for compound **14** Ac-Azp-OMe and **15** Ac-azp-OMe the pattern of H_1 signal is independent of pH.

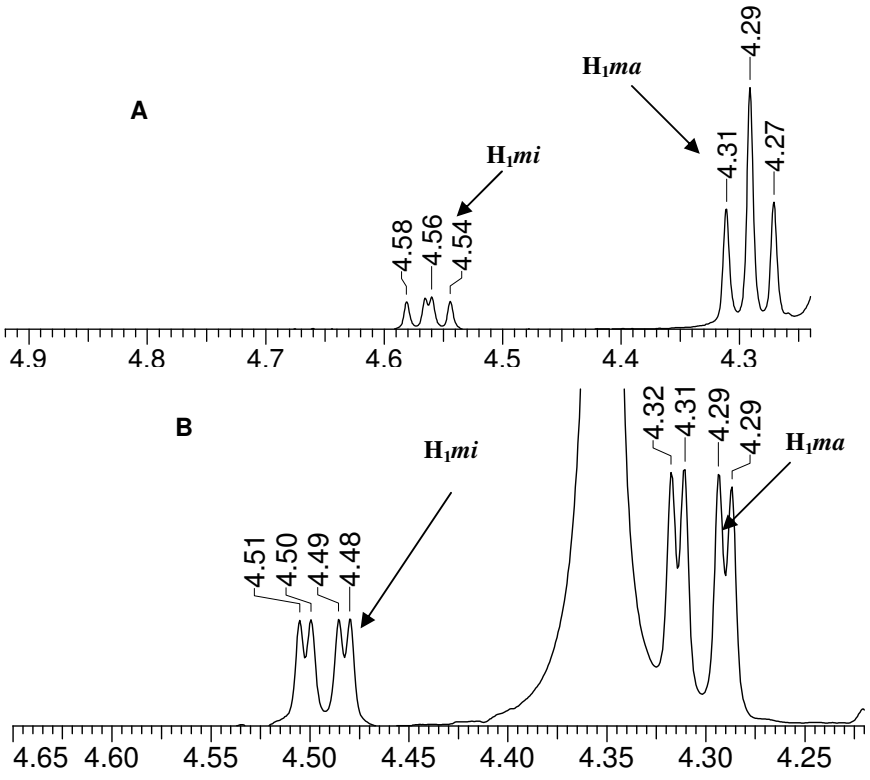


Figure 5.13: Expanded 400MHz ^{1}H-NMR spectrum of **A;** Ac-Azp-OMe **14** and **B;** Ac-azp-OMe **15** in D_2O at, pH 2.0. *ma* is the major isomer and *mi* is the minor isomer.

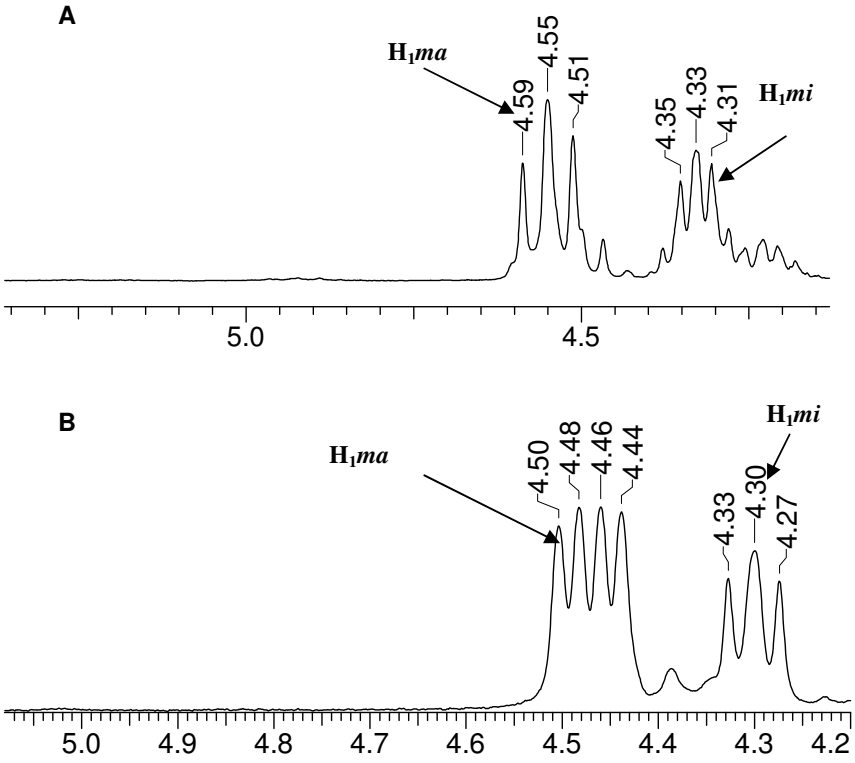


Figure 5.14: Expanded 400MHz ^{1}H-NMR spectrum of **A,** Ac-azp-OMe **14** and **B**, Ac-amp-OMe **15** in D_2O at 12.0. *ma* is the major isomer and *mi* is the minor isomer.

5.4.4: Thermodynamics of prolyl-peptide bond isomerization: Variable temperature NMR studies

The thermodynamics of *E/Z* isomerization of the prolyl-peptide bond can be effectively studied using variable temperature NMR spectroscopy which prerequisite well-separated signals for *E* and *Z* isomers in their ^{1}H and ^{13}C spectra. For the *E* and *Z* isomers, that are interconvertable, the ratio of their NMR-integrals (for the same nucleus in the both isomers) represents the equilibrium constant for the isomerization. An important consideration in such measurements is that a sufficient post-acquisition delay be allowed between the collections of FIDs to compensate for the difference in the relaxation times of nuclei. ^{1}H-NMR spectra of compounds **12-15** in D_2O at both acidic and basic conditions show two sets of signals corresponding to simultaneous preference

of a major and a minor isomer arising from prolyl-peptide bond isomerization Comparison of ^{1}H-NMR spectra of compounds **12-15** recorded at 400MHz using 5mm high-temperature probe revealed that the most suitable protons for integration are N-acetyl-C$\underline{H}_3$ ester-OC$\underline{H}_3$ and β-protons. The ^{1}H-NMR spectra were recorded at different temperatures and the equilibrium constant $K_{Z/E}$ at each temperature was calculated according to the following equation

$$K_{Z/E} = I_Z/I_E$$

where I_Z and I_E are the integrals of the ^{1}H-NMR signals corresponding to *Z* and *E* isomers respectively and $K_{Z/E}$ is the equilibrium constant.

Figure 5.15 and 5.16 show the ^{1}H-NMR signals for *E* and *Z* forms of compounds **12** Ac-Amp-OMe at pH 2.0 (ester O-CH_3 signals) and at pH 12.0 (H^{β} signals) and **13** Ac-amp-OMe at pH 2.0 (H^{β} signals) and at pH 12.0 (ester O-CH_3 signals) respectively. Similarly Figure 5.17 and Figure 5.18 show the NCO-CH_3 (N-acetyl-CH_3) ^{1}H-NMR signals for *E* and *Z* forms of compounds **14** (Ac-Azp-OMe) and **15** (Ac-amp-OMe) at pH 2.0 and 12.0 respectively.

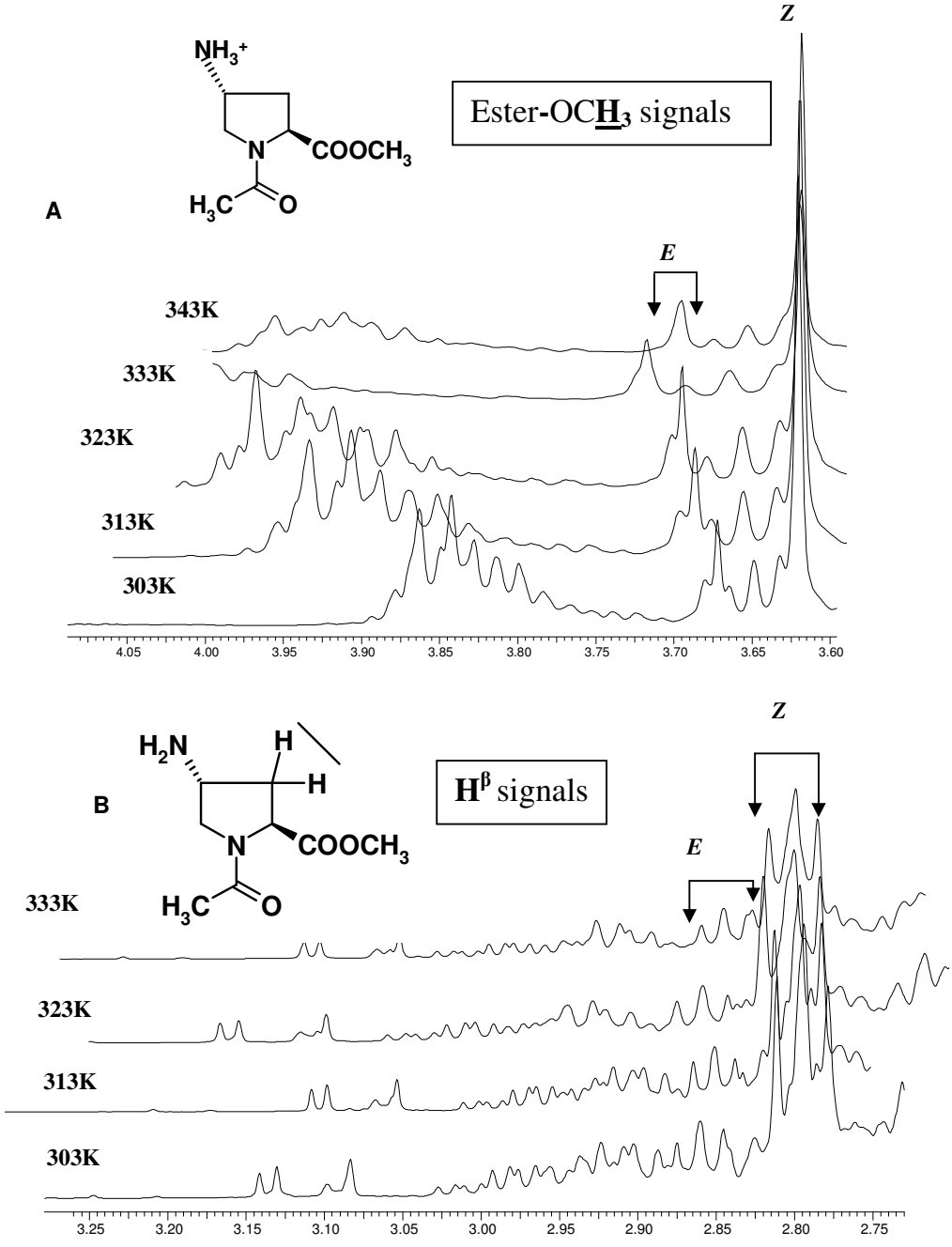


Figure 5.15: Temperature dependent 400MHz ^{1}H-NMR spectra of compound **12** Ac-Amp-OMe in D_2O, **A**; at pH 2.0 (ester O-CH_3 proton) and **B**; at pH 12.0. (β-proton).

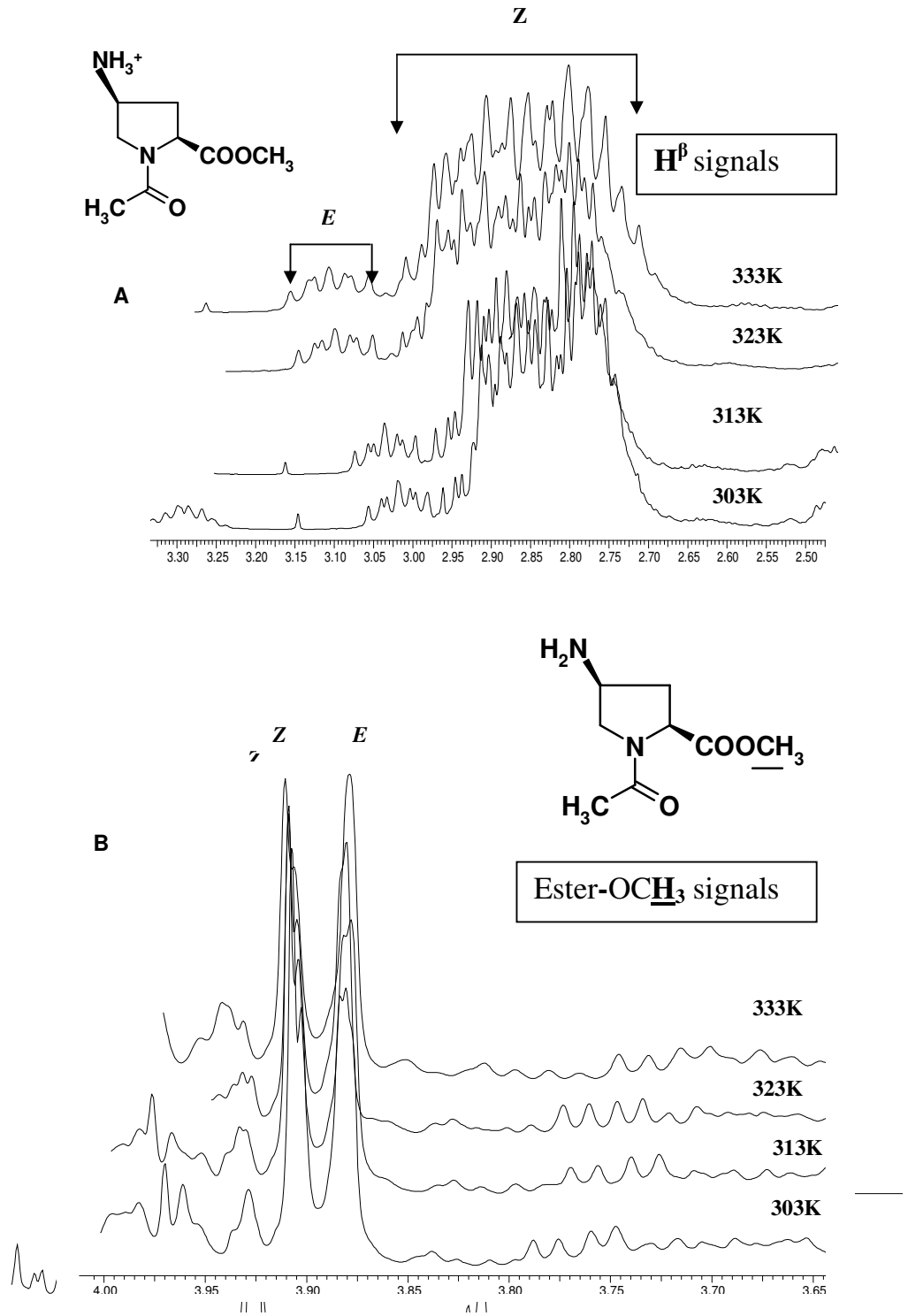


Figure 5.16: Temperature dependent 400MHz ^{1}H-NMR spectra of compound **13** Ac-amp-OMe in D_2O, **A**; at pH 2.0 (β-proton) and **B**; at pH 12.0. (ester O-CH_3 proton).

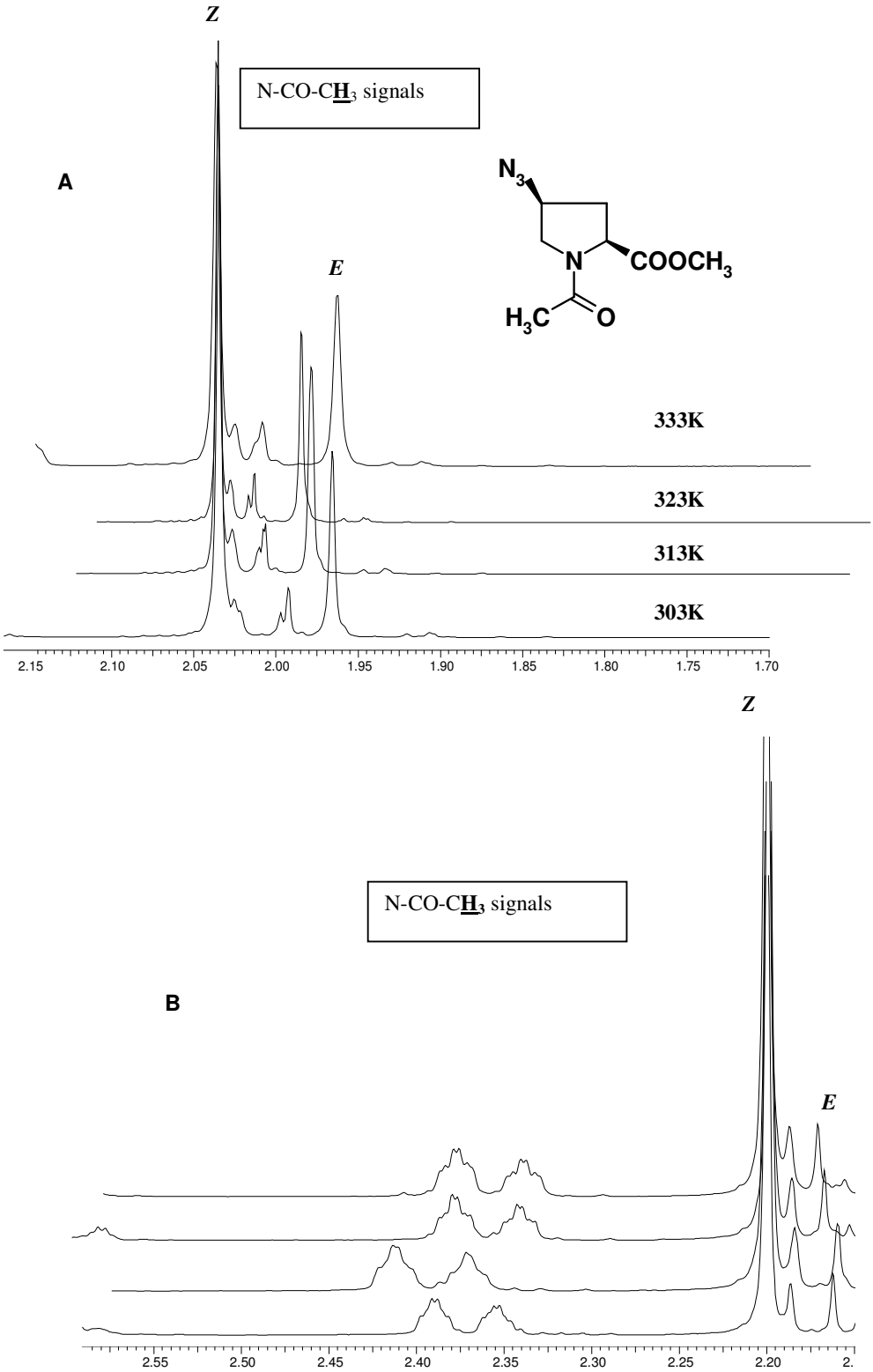


Figure 5.18: Temperature dependent 400MHz ^{1}H-NMR spectra of compound **15** Ac-azp-OMe in D_2O, **A**; at pH 2.0 (-NCO-CH_3 proton) and **B**; at pH 12.0. (-NCO-CH_3 proton).

The temperature dependent $K_{Z/E}$ values together with the relative integrations of ^1H signals derived from above spectra for the compounds **12-15** are listed in Table 5.3a-d.

Table 5.3: Temperature dependent equilibrium constant ($K_{Z/E}$) for Ac-Xaa-OMe model compounds **12-15** derived from VT-NMR data. I_Z and I_E are the relative integration of proton signals of the Z and E forms respectively.

(a) Ac-Amp-OMe 12

	pH 2.0			pH 12.0		
Temperature [K]	**I_Z**	**I_E**	**$K_{Z/E}$**	**I_Z**	**I_E**	**$K_{Z/E}$**
303	100	14.5	6.89	100	16.5	6.06
313	100	15.1	6.62	100	18.3	5.46
323	100	17.9	5.58	100	19.7	5.07
333	100	19.4	5.15	100	22.3	4.48

(b) Ac-amp-OMe 13

	pH 2.0			pH 12.0		
Temperature [K]	**I_Z**	**I_E**	**$K_{Z/E}$**	**I_Z**	**I_E**	**$K_{Z/E}$**
303	100	38	2.63	100	75	1.33
313	100	39.8	2.51	100	79	1.26
323	100	41.7	2.39	100	82	1.21
333	100	43	2.32	100	92	1.08

(c) Ac-Azp-OMe 14

	pH 2.0			pH 12.0		
Temperature [K]	**I_Z**	**I_E**	**$K_{Z/E}$**	**I_Z**	**I_E**	**$K_{Z/E}$**
303	100	15.3	6.53	100	15.6	6.41
313	100	16.7	5.98	100	16.8	5.95
323	100	19.5	5.12	100	19.3	5.18
333	100	20.7	4.83	100	20.8	4.80

(d) Ac-azp-OMe 15

	pH 2.0			pH 12.0		
Temperature [K]	**I_Z**	**I_E**	**$K_{Z/E}$**	**I_Z**	**I_E**	**$K_{Z/E}$**
303	100	29.1	3.43	100	30	3.33
313	100	31.5	3.17	100	33	3.03
323	100	36.5	2. 76	100	37	2.70
333	100	38.2	2.61	100	39	2.56

The enthalpy ΔH^o and the entropy ΔS^o of isomerization were calculated from the $K_{Z/E}$ data using Van't Hoff equation.

$$\ln K_{Z/E} = (-\Delta H^o/R)\,(1/T) + (\Delta S^o)/R$$

where $K_{Z/E}$ is the experimentally determined equilibrium constant and R is the ideal gas constant (1.987 cal deg^{-1} mol^{-1}). The value of ΔH^o and ΔS^o thus derived together with the calculated values of $K_{Z/E}$ at 30°C are listed in Table 5.4, assuming that the enthalpic and entropic difference between the *E* and *Z* isomers are independent of the temperature, that is, the change in heat capacity at constant pressure $\Delta C_p^o = 0$ for the isomerization. The observed linear Van't Hoff plot indicates that such an assumption is likely to be valid for the isomerization.

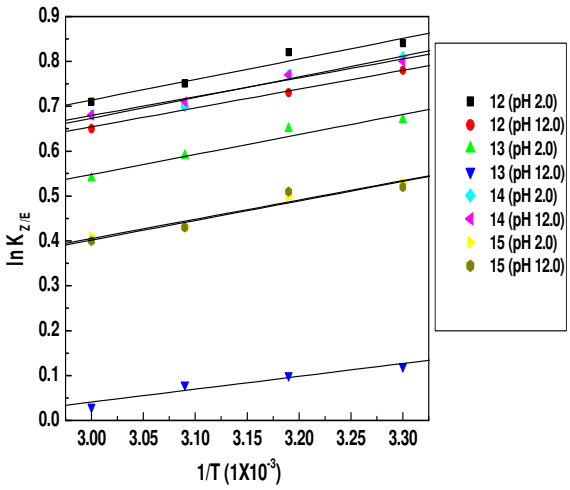


Figure 5.19: Van't Hoff plot of *Z-E* isomerization obtained from VT-NMR data for compounds **12-15**.

Table 5.4: Thermodynamic parameters for compounds **12-15** derived from VT-NMR data

Compound	$K_{Z/E}$ at 30 °C		ΔH^o Kcal mol^{-1}		ΔS^o Kcal mol^{-1}	
	pH 2.0	pH 12.0	pH 2.0	pH 12.0	pH 2.0	pH 12.0
Ac-Amp-OMe 12	6.89	6.06	-1.36	-1.21	-1.395	-1.237
Ac-amp-OMe 13	2.63	1.33	-0.94	-0.26	-1.071	-0.0655
Ac-Azp-OMe 14	6.53	6.41	-1.29	-1.27	-1.328	-1.313
Ac-azp-OMe 15	3.43	3.33	-0.69	-0.66	-0.782	-0.776

The VT-^{1}H-NMR spectra indicated that for all compounds at all temperatures measured, *Z* isomer is more favored than *E* isomer. This result is consistent with several measurements reported in the literature for similar kind of compounds. Also negative enthalpies for the *Z-E* isomerization (Table 5.4) indicate that the *Z* isomer becomes increasingly favored with decrease in temperature. An increase in temperature results in increase in ΔG which ultimately leads to a decrease in the $K_{Z/E}$ value.

5.4.5: Pyrrolidine ring conformation derived from J_{H-H} coupling constants of compounds 12-15

For a γ-*exo* conformation, the relative magnitudes of the vicinal coupling constants for each C-C bond fragment confirms to a pattern where $J_{1-2} \leq J_{1-3}$ (C^{α}-C^{β}), $J_{2-4} < J_{3-4}$ (C^{β}-C^{γ}) and $J_{4-5} > J_{4-6}$(C^{γ}-C^{δ}). Similarly, for γ-*endo* conformation, where the endocyclic torsion angles have an opposite sign, a reverse order in which $J_{1-2} > J_{1-3}$, $J_{2-4} > J_{3-4}$ and $J_{4-5} < J_{4-6}$ is expected.[57] Figure 5.20-5.27 show the complete 400MHz ^{I}H-NMR spectra of compounds **12-15** at pH 2.0 and 12.0.

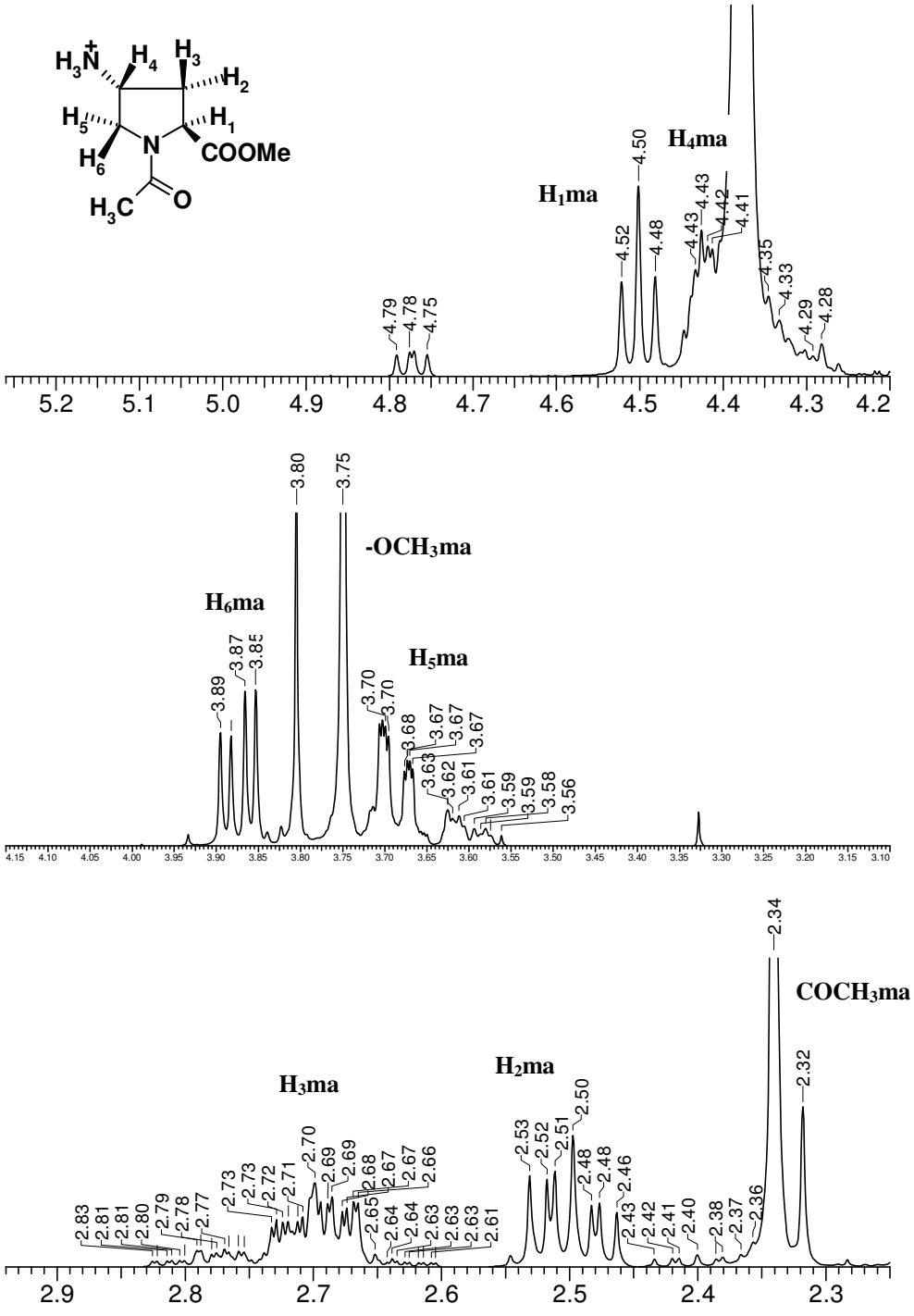


Figure 5.20: Expanded 400MHz ^{1}H-NMR spectrum of Ac-Amp-OMe **12** at pH 2.0 in D_2O. ma is the major (*Z*) isomer.

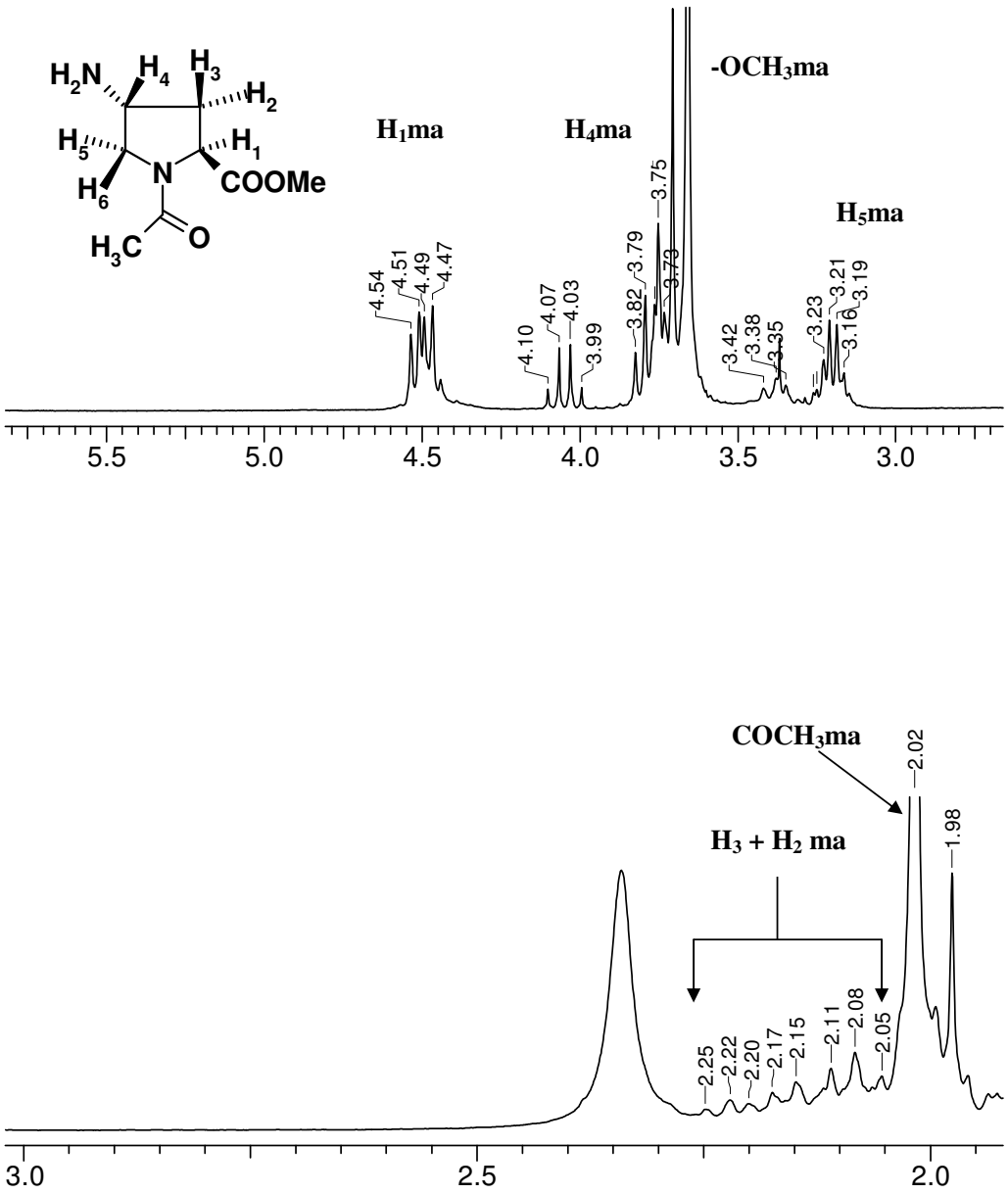


Figure 5.21: Expanded 400MHz ^{1}H-NMR spectrum of Ac-Amp-OMe **12** at pH 12.0 in D_2O. ma is the major (*Z*) isomer.

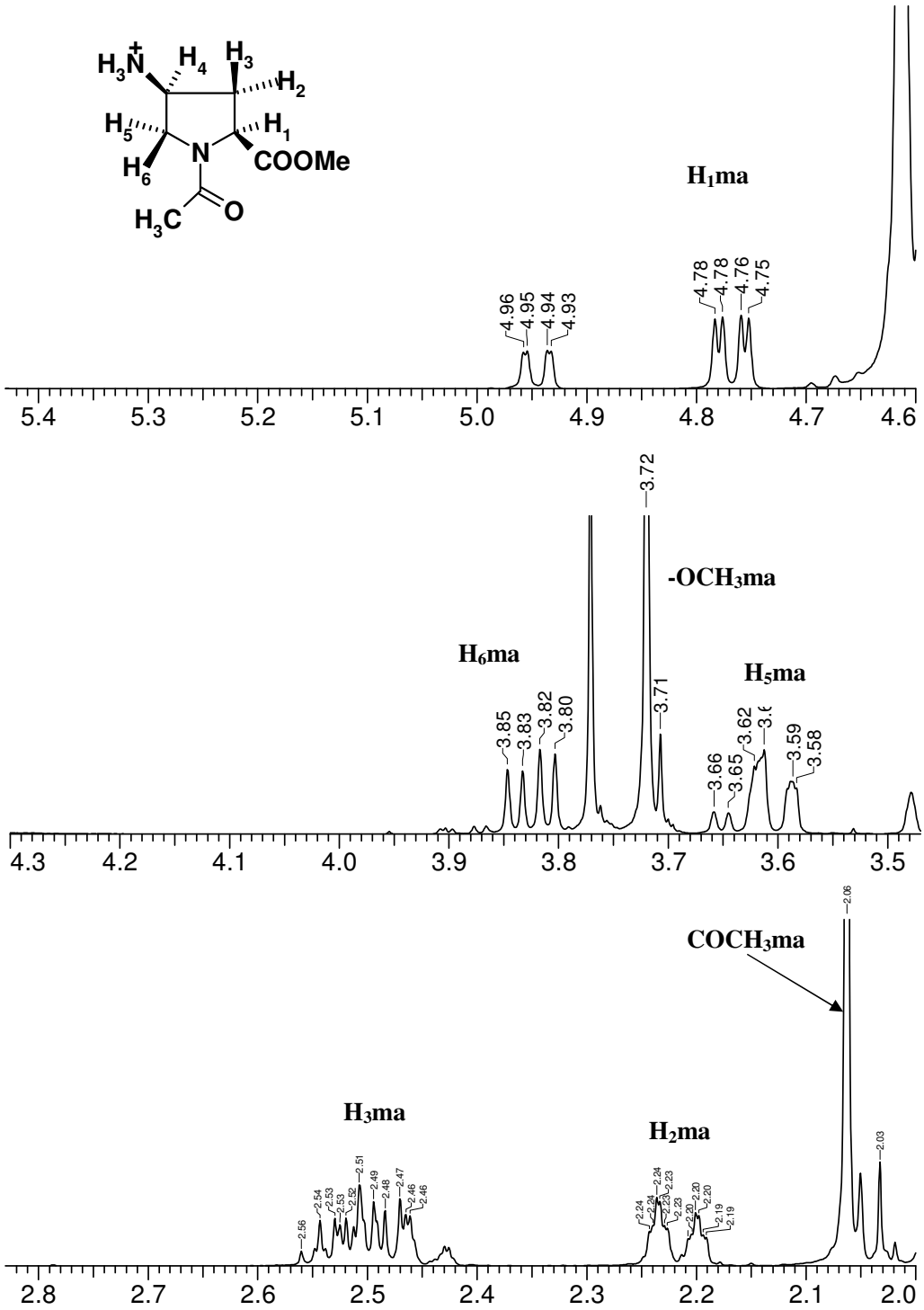


Figure 5.22: Expanded 400MHz ^{1}H-NMR spectrum of Ac-amp-OMe **13** at pH 2.0 in D_2O. ma is the major (*Z*) isomer.

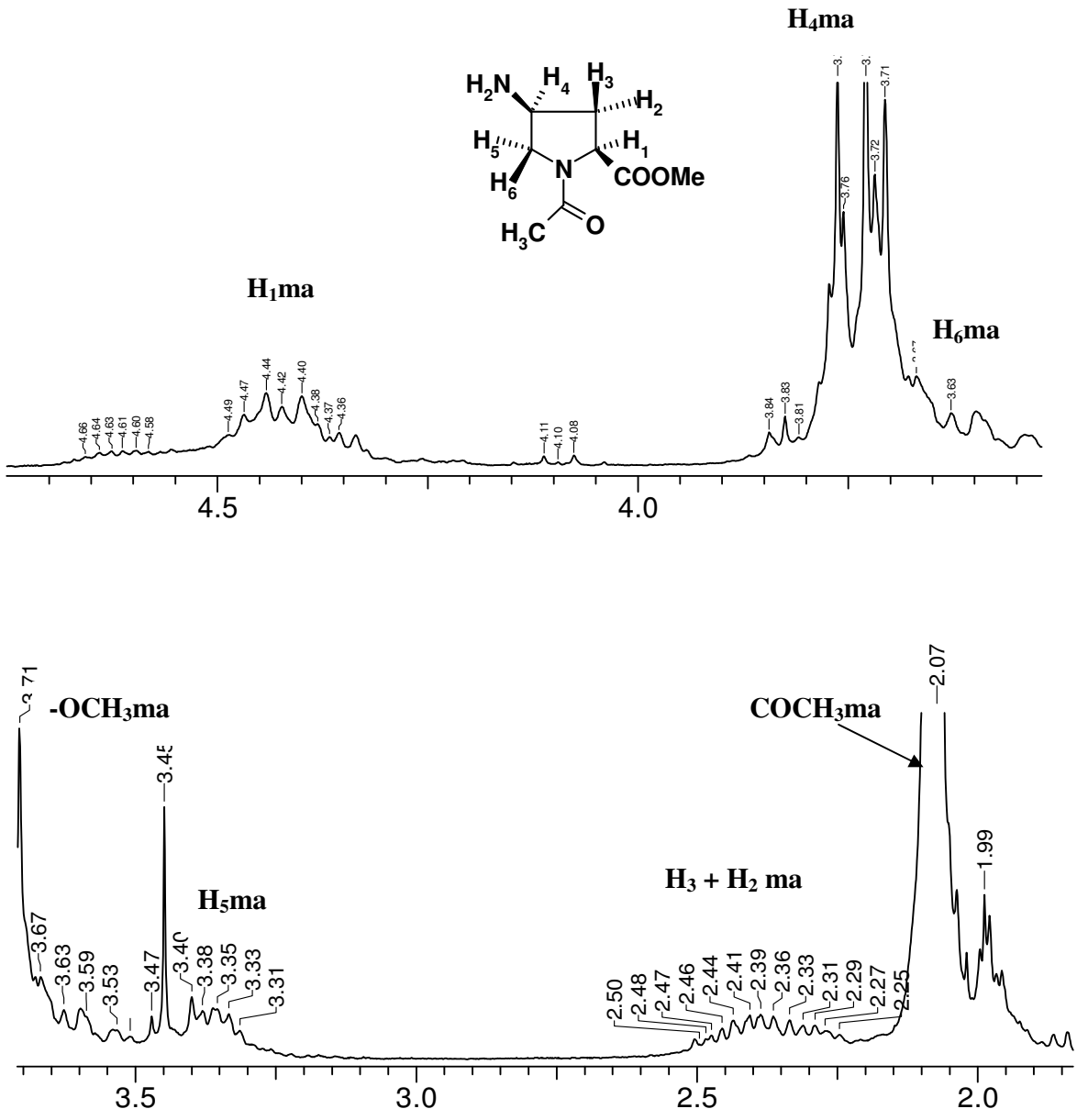


Figure 5.23: Expanded 400MHz ^{1}H-NMR spectrum of Ac-amp-OMe **13** at pH 12.0 in D_2O. ma is the major (*Z*) isomer.

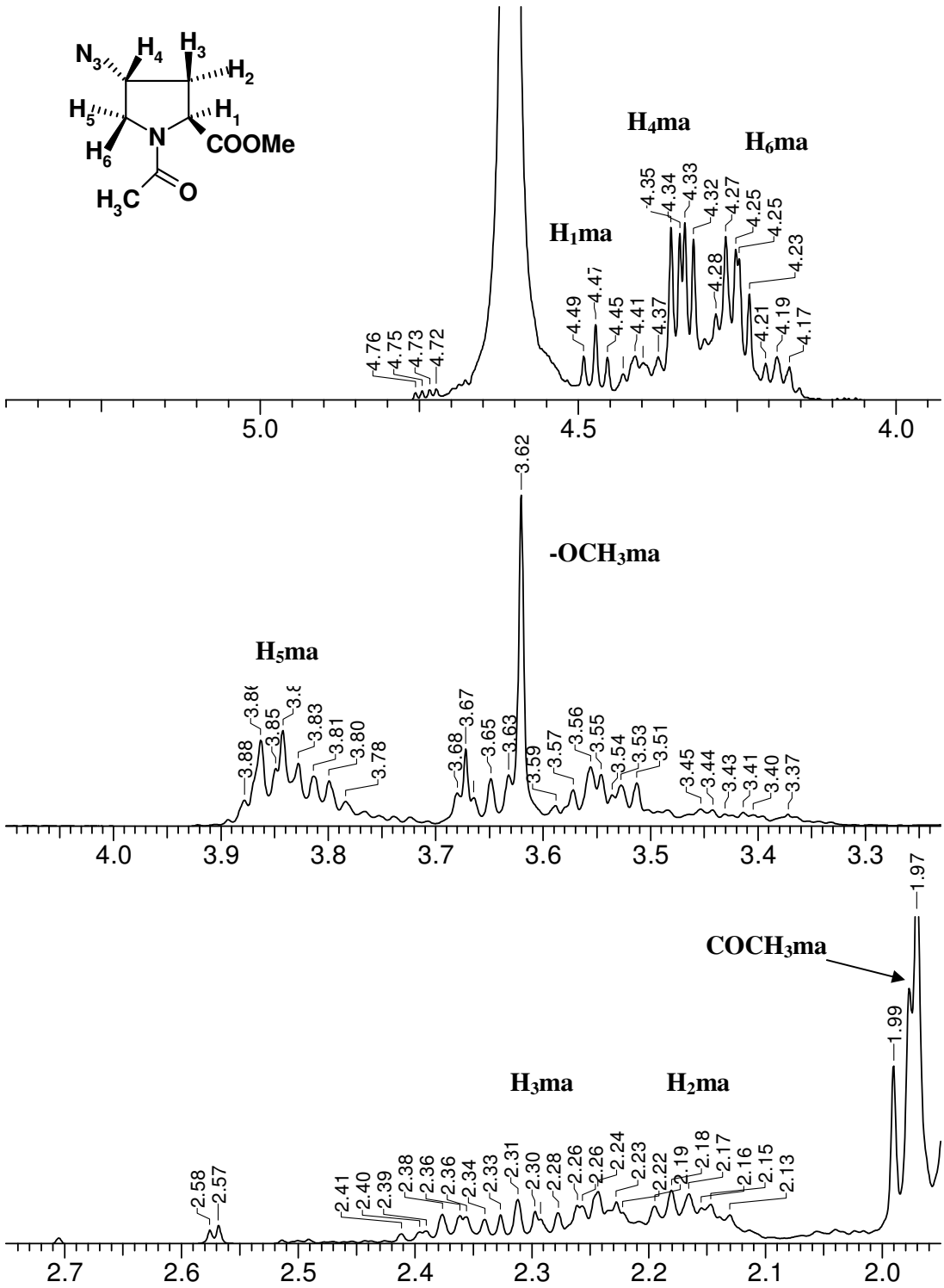


Figure 5.24: Expanded 400MHz ^{1}H-NMR spectrum of Ac-Azp-OMe **14** at pH 2.0 in D_2O. ma is the major (*Z*) isomer.

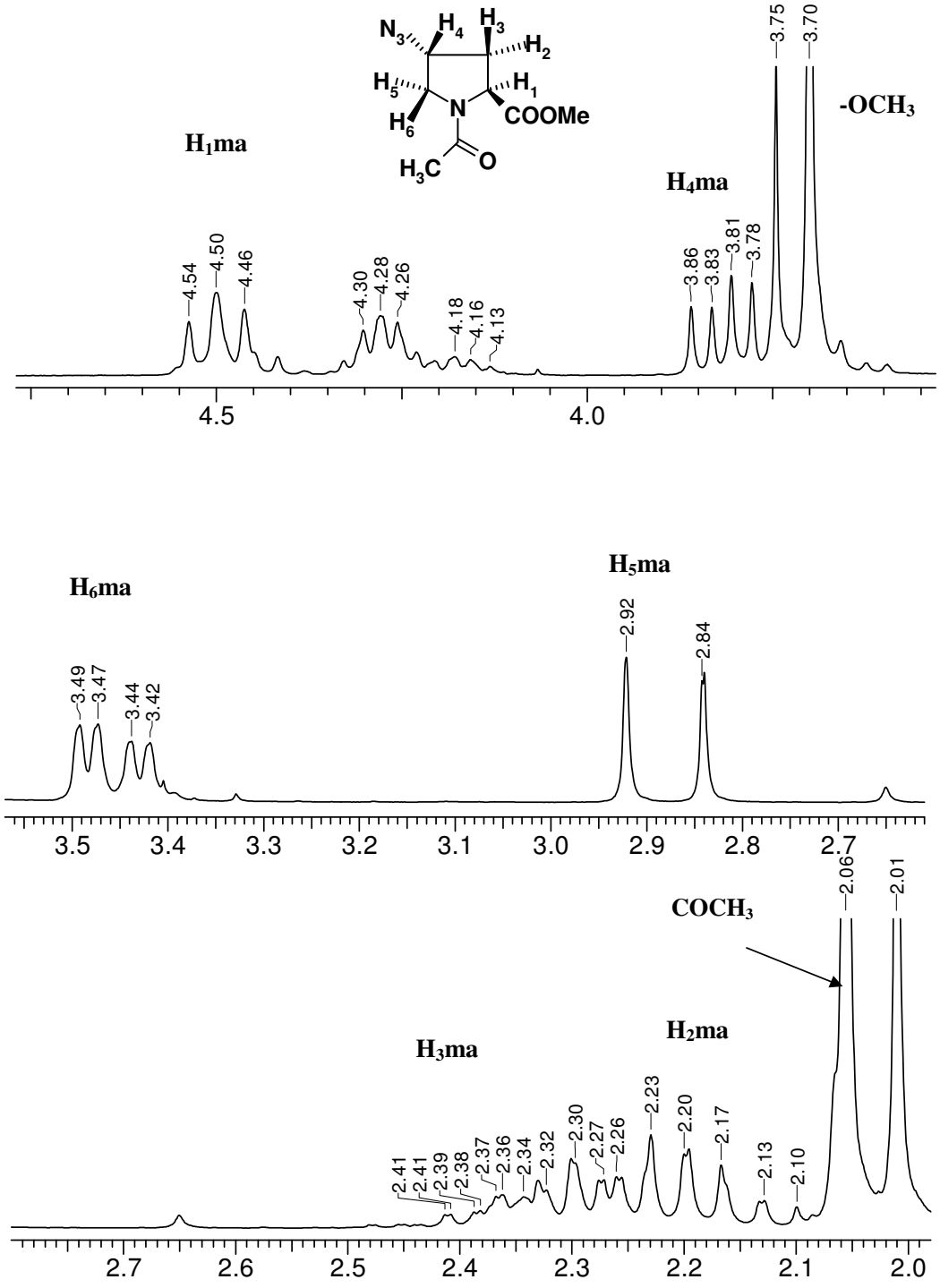


Figure 5.25: Expanded 400MHz ^{1}H-NMR spectrum of Ac-Azp-OMe **14** at pH 12.0 in D_2O. ma is the major (*Z*) isomer.

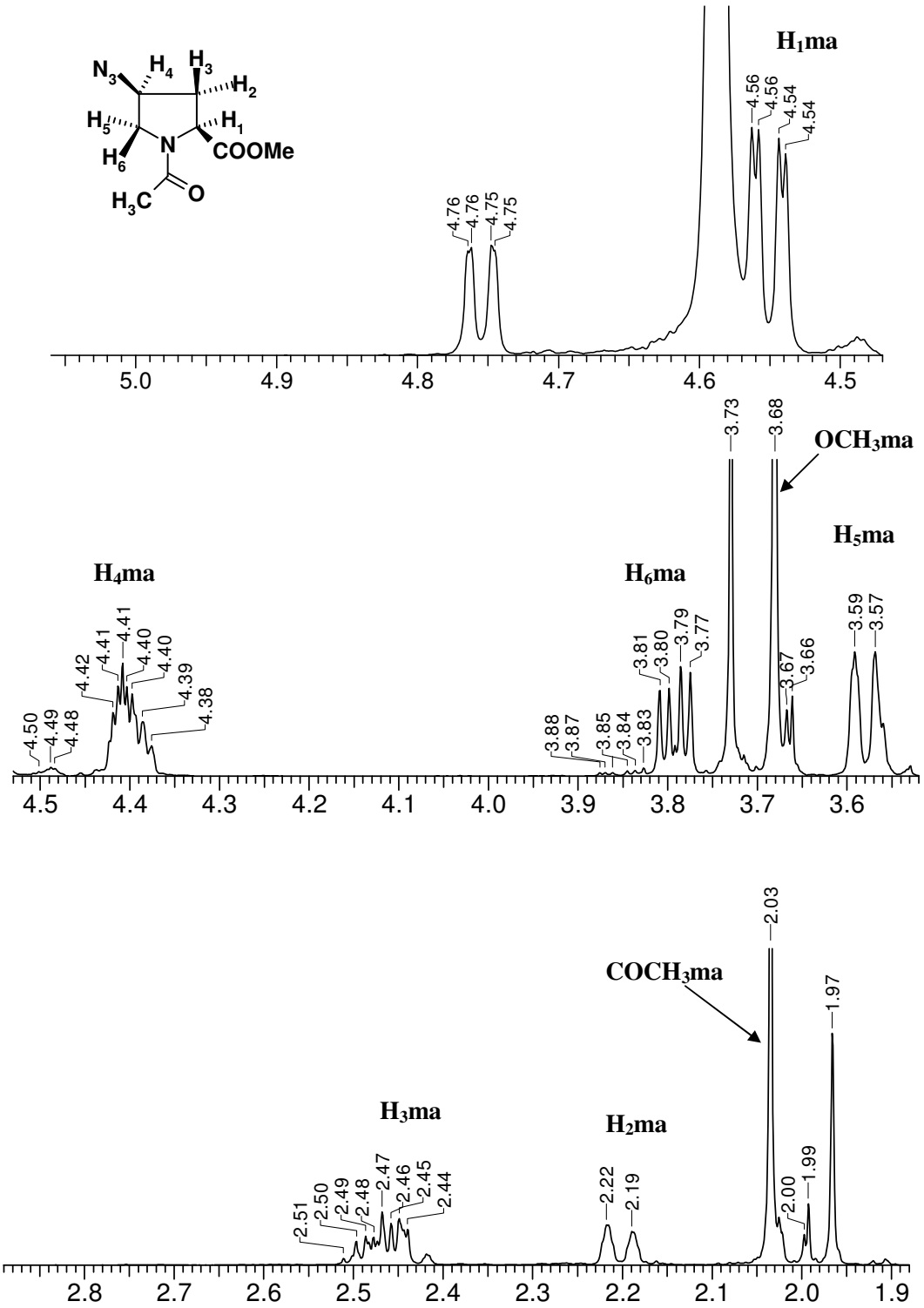


Figure 5.26: Expanded 400MHz ^{1}H-NMR spectrum of Ac-azp-OMe **14** at pH 2.0 in D_2O. ma is the major (*Z*) isomer.

Table 5.5: Vicinal coupling constant of compounds **12-15** extracted from 400MHz ^{1}H-NMR spectra at pH 2.0 and 12.0 and the derived pucker type of pyrrolidine ring.

Compound	Vicinal J_{H-H} values (Hz)						Derived pucker type
	J_{1-2}	J_{1-3}	J_{2-4}	J_{3-4}	J_{4-5}	J_{4-6}	
	pH 2.0						
J_{cal} for γ-exo pucker[1]	**8.10**	**10.36**	**0.93**	**3.82**	**3.26**	**0.62**	**γ-*exo***
J_{cal} for γ-endo pucker[1]	**10.50**	**2.80**	**385**	**1.10**	**0.81**	**3.42**	**γ-*endo***
Ac-Amp-OMe 12	6.5	6.26	3.54	8.81	3.30	4.54	**γ-*exo***
Ac-amp-OMe 13	8.93	6.03	6.54	5.48	3.51	6.52	**γ-*endo***
Ac-Azp-OMe 14	8.24	8.25	2.3	4.57	4.86	2.53	**γ-*exo***
Ac-azp-OMe 15	9.25	3.82	3.35	5.11	2.01	3.12	**γ-*endo***
	pH 12.0						
Ac-Amp-OMe 12	6.2	7.39	7.02	8.22	4.58	7.13	**γ-*exo***
Ac-amp-OMe 13	19.91	20.1	21.34	-	10.45	-	**γ-*endo***
Ac-Azp-OMe 14	8.22	8.26	2.34	4.77	4.86	2.81	**γ-*exo***
Ac-azp-OMe 15	9.32	3.33	3.4	5.42	3.10	6.88	**γ-*endo***

[1] from reference 57.

The ^{1}H-NMR spectral splitting pattern and J_{H-H} values of compound **12** Ac-Amp-OMe at pH 2.0 matches with J_{cal} pattern for γ-*endo* pyrolidine ring pucker (Table 5.5) , whereas at pH 12.0 J_{H-H} values matches with the J_{cal} pattern for γ-*exo* pyrrolidine ring pucker. Interestingly at pH 2.0 compound **12** shows a long range W-type coupling of J_{2-6} which is not observed at pH 2.0. This indicates that compound **12** adopts two different pyrrolidine ring conformations at acidic and basic conditions, even though a triplet is observed for H_1 proton in both acidic and basic pH conditions. In contrast compound **13** (Ac-amp-OMe) shows doublet of doublet (dd) splitting pattern for H_1 proton with J_{H-H} values, that match with the J_{cal} pattern for γ-*endo* pyrrolidine ring conformation, but at pH 12.0 the doublet of doublet (dd) splitting pattern of H_1 proton changes to mutiplet pattern which neither matches with γ-*endo* nor γ-*exo* patterns.

J_{H-H} values and ^{1}H-NMR splitting pattern of compound **14** Ac-Azp-OMe at both pH 2.0 and 12.0 resembles with the J_{cal} pattern of γ-*exo* pyrrolidine ring pucker and ^{1}H NMR compound **15** Ac-azp-OMe shows a pattern that resembles to γ-*endo* pyrrolidine

pucker. Unlike for compounds **12** and **13**, the magnitude of pyrrolidine ring puckers for compounds **14** and **15** are invariant to pH change.

5.5: Determination of pKa value of 4-aminoproline containing collagen peptides

The pK_a value(s) of a compound influences many characteristics of the compound such as its reactivity, solubility and structure correlated spectral properties (colour). In biochemistry, the pK_a values of proteins and amino acid side chains are of major importance for the activity of enzymes and the stability of proteins. Amino acid side chains play an important role in defining the pH-dependent characteristics of a protein. The pH-dependence of the activity displayed by enzymes and the pH-dependence of protein stability, for example, are properties that are determined by the pKa values of amino acid side chains.

5.5.1: pKa calculation methods

Some of the methods are based on solutions to the Poisson-Boltzmann equation (PBE), often referred to as FDPB-based methods (*FDPB* is for "finite difference Poisson-Boltzmann"). The PBE is a modification of Poisson's equation that incorporates a description of the effect of solvent ions on the electrostatic field around a molecule.

5.5.2: Molecular dynamics (MD)-based methods

Molecular dynamics methods of calculating pKa values involve computationally measuring the free energy difference between the protonated and deprotonated forms of the molecule. This free energy difference is measured using methods such as free energy perturbation, thermodynamic integration and the Bennett acceptance ratio. Molecular dynamics is typically a much more computationally expensive way to predict

pKa's than using the Poisson Boltzmann equation. Furthermore, it is currently much less accurate. This is because the currently used molecular force fields do not take polarisability into account. Polarisability is an important property for protonation energies.

5.5.3: Using pH titration curve

The pH value at half the equivalence point is equal to the pKa. Thus when the pH of a solution is at the pKa of an acid, half will be present in the deprotonated form and half will be protonated (HA). This is immediately evident from the Henderson Hasselbach equation. The pKa of side chain group in a protein can de determined by simply titrating the protein solutions of known concentration. Typically, the sample of interest was first brought to either high alkaline or low acidic conditions. It was then titrated with addition of fixed volume aliquots, with a strong acid in the former case, or strong base in latter case. After the addition of titrant, the pH was recorded. The pH of the solution changes rapidly at the pKa of the functional group.

The pKa values of ionizable residues play fundamental roles in the function and structure of protiens.[58] The pKa value itself is sensitive to a number of factors, including the change in dielectric constant of media,[59] H-bonding net-work[60] and the number of explicit water molecules inculded.[61] Recently, several reports have focused on the understanding the effects of conformational changes on the pKa values. For examples, a single torsion angle change in hen egg white lysozyme (HEWL) results in large pKa differences of active site residue.[62] Nielsen et.al.[63] have investigated the conformatinoal dependence of pKa values from 41 HEWL X-ray structures, focusing specifically on the ability to correctly identify proton donors and acceptors in two catalytic acids (Glu

35 and Asp 52). One interesting conculsion form their work relates to the origin of the conformational dependence of the pKa variability within these two positions.

The 4-amino group in 4-aminoproline containing collagen peptides gets protonated in acidic and neutral conditions and it pKa value determination may provide information about the triple-helix structure as a function of conformation of pyrrolidine ring and solvation of residues at X and Y positions etc. The prolines at X and Y position of the collagen sequence (X-Y-Gly) adopt different conformations and undergo solvation to different extents and all these factors affect the pKa value of 4-aminoproline.

pKa values of pepetides **19** Ac-Phe(Pro-Amp-Gly)$_6$-NH_2 and peptide **20** Ac-Phe(Amp-Pro-Gly)$_6$-NH_2 were measured to know whether the conformational and solvation changes in the aminoacid residue will affect the pKa value in single strand conformation and in the triple-helical structure. For this purpose, the pH tiration was carried out at concentrations corresponding to above (20mM) and below (10mM) the critical triplex cocentration respectively.

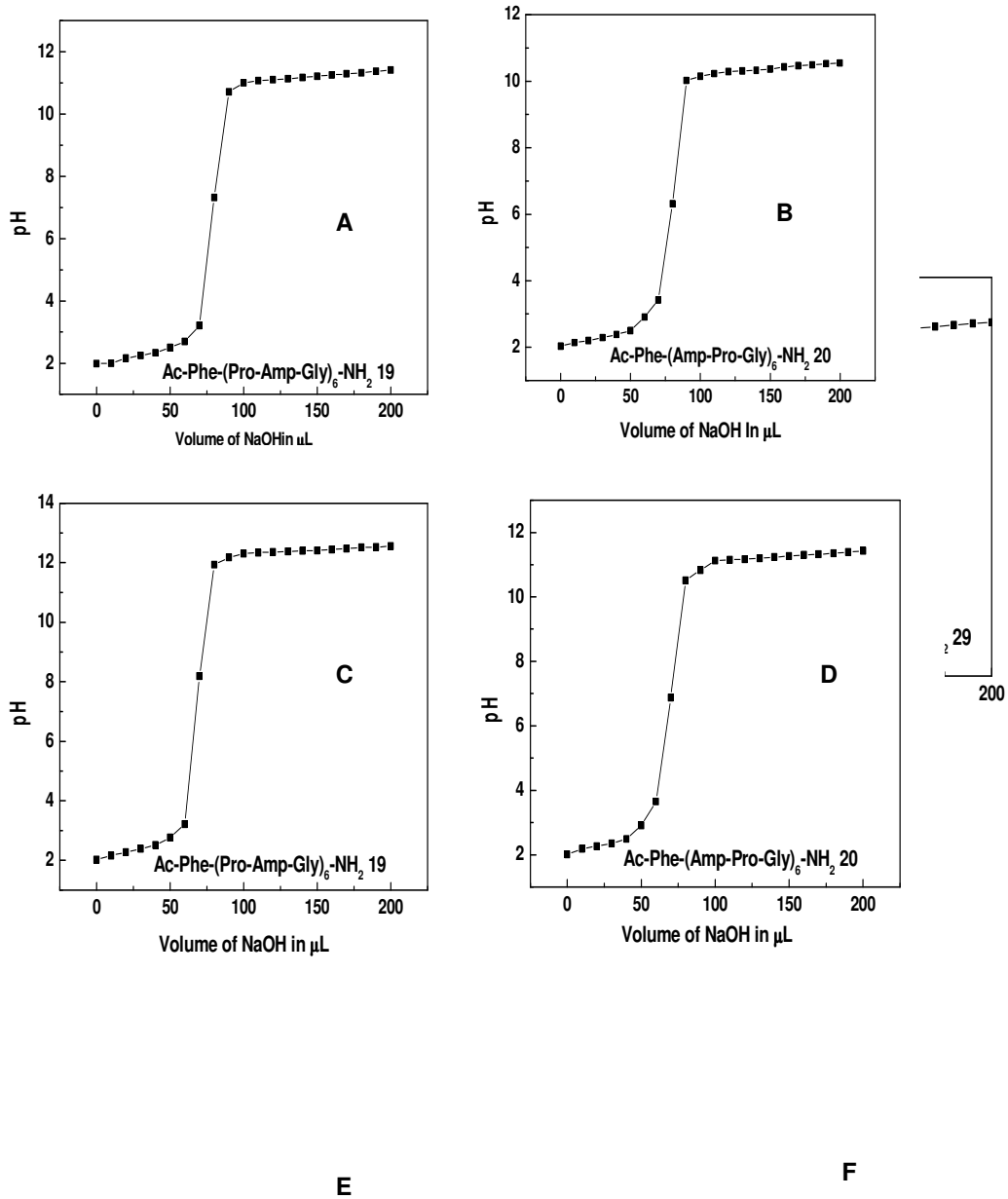


Figure 5.28: **A** and **B** pH titration curves of 0.10mM solutions of peptides **19** and **20** respectively; **C**, **D**, **E** and **F** pH titration curves or 0.20mM solutions peptides **19**, **20**, **28** and **29** respectively.

Table 5.6: pKa values of peptides **19** and **20** at 0.10mM and 0.20mM concentrations of the peptides

pKa values		
Concentrantion	**(Pro-Amp-Gly) 19**	**(Amp-Pro-Gly) 20**
0.10mM	7.9	6.72
0.20mM	8.30	7.86

Figure 5.28 A and B shows the pH titration curves of 0.10mM (below critical triplex concentration) solution of peptide **19** Ac-Phe(Pro-Amp-Gly)$_6$-NH$_2$ and peptide **20** Ac-Phe(Amp-Pro-Gly)$_6$-NH$_2$, simillarly Figure 5.28 C and D shows the pH titration curves of 0.20mM solution of peptide **19** and **20** repetively. In the single chain conformation pKa value of peptide **19** (Pro-Amp-Gly) is 7.9 and peptide **20** is 6.72 . Both peptides show higher pKa value in triple-helical conformation (Table 5.6).

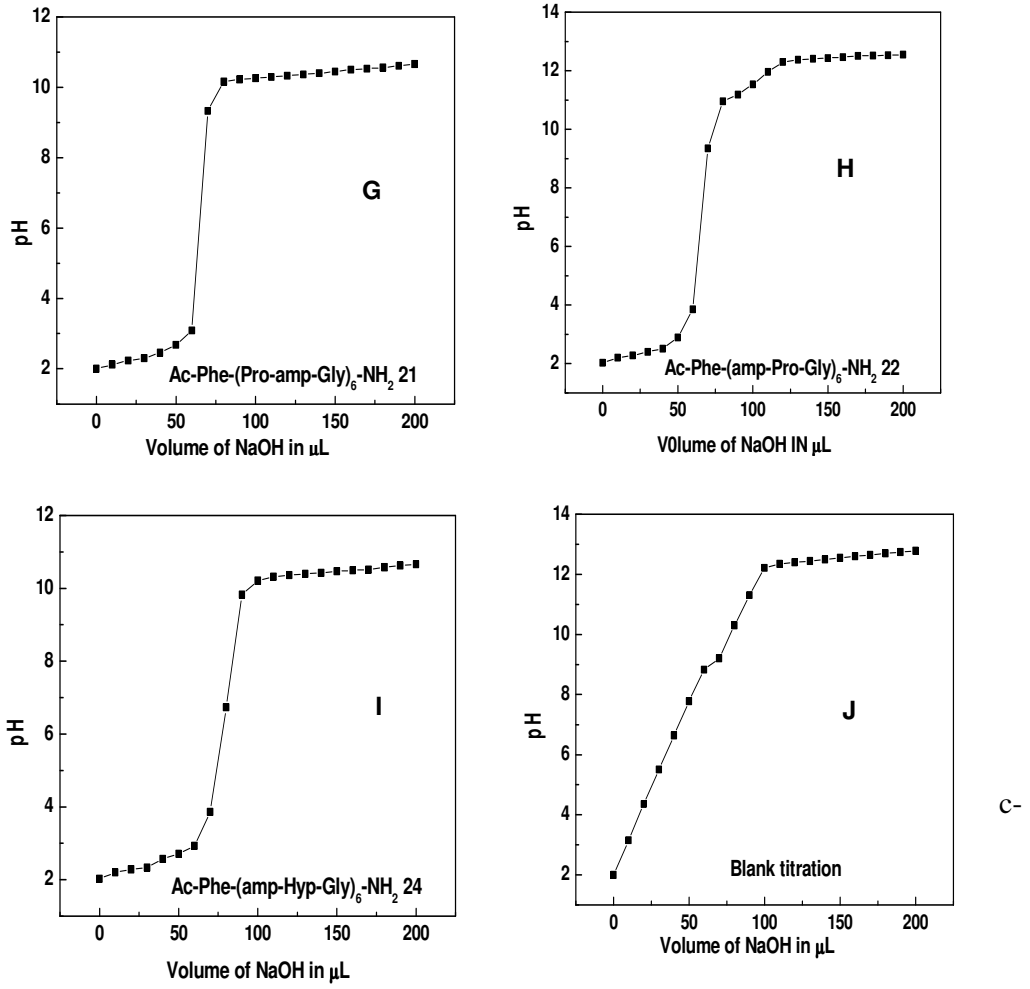
The replacement of Pro in the peptides **19** (Pro-Amp-Gly) and **20** (Amp-Hyp-Gly) to Hyp resulting in chimeric peptides **28** (Hyp-Amp-Gly) and **29** (Amp-Hyp-Gly) (Fig 5.28 E and F respectively, leads to higher pKa values of 9.06 (Hyp-Amp-Gly **28**) and 8.44 (Amp-Hyp-Gly **29**) compared to the respective model peptides (Pro-Amp-Gly **19**) and (Amp-Pro-Gly **20**).

Figure 5.29 G, H and I show the pH titration curves of the peptides **21** Ac-Phe(Pro-amp-Gly)$_6$-NH$_2$, **22** Ac-Phe(amp-Pro-Gly)$_6$-NH$_2$ and **24** Ac-Phe(amp-Hyp-Gly)$_6$-NH$_2$ respectively. The peptide **22** (amp-Pro-Hyp) shows two pKa values of 7.03 and 9.64, but the chimeric peptide **24** (amp-Hyp-Gly) in which Hyp was replaced in place of Pro of the peptide **22** (amp-Pro-Gly) shows only one pKa value of 6.73. Peptide **21** (Pro-amp-Gly) shows lower pKa value of 6.07 among three peptides (Table 5.7). Interestingly, the replcement of Pro in the peptide **22** (amp-Pro-Gly) to Hyp in the peptide **24** (amp-Hyp-Gly) decreases the pKa value. This observation is contrary to the

results of peptide **19** (Pro-Amp-Gly) and peptide **29** (Hyp-Amp-Gly) wherein the peptide **29** shows higher pKa value (8.44) than peptide **19**

(8.30 **Table 5.7:** pKa values of 0.20mM solutions of 4-aminoproline collagen peptides.

Peptide	pKa value
Ac-Phe(amp-Pro-Gly)$_6$-NH$_2$ **22**	7.03, 9.64
Ac-Phe(Pro-amp-Gly)$_6$-NH$_2$ **21**	6.07
Ac-Phe(amp-Hyp-Gly)$_6$-NH$_2$ **24**	6.73
Ac-Phe(Amp-Hyp-Gly)$_6$-NH$_2$ **29**	8.44
Ac-Phe(Hyp-Amp-Gly)$_6$-NH$_2$ **28**	9.06
Ac-Phe(Amp-Amp-Gly)$_6$-NH$_2$ **27**	8.93
Ac-Phe(amp-Amp-Gly)$_6$-NH$_2$ **23**	7.1, 11.2



c-

Figure 5.29: G, H and **I**; pH titration curves pH titration curves or 0.20mM solutions peptides **21**, **22** and **24** respectively; **J**; Blank titration curve.

monomer. Interestingly peptide **23** (amp-Amp-Gly) exhibit two pKa values of 7.1 and 11.2, whereas the peptide **27** (Amp-Amp-Gly) shows one pKa value of 8.93. Since all the peptides were capped at N and C termini, the observed sigmoidal curves and the determined pKa values are due to the protonation of 4-NH_2 groups in the 4-aminoproline containing peptides. Absence of sigmoidal transition in blank titration confirm this (Fig 5.29J).

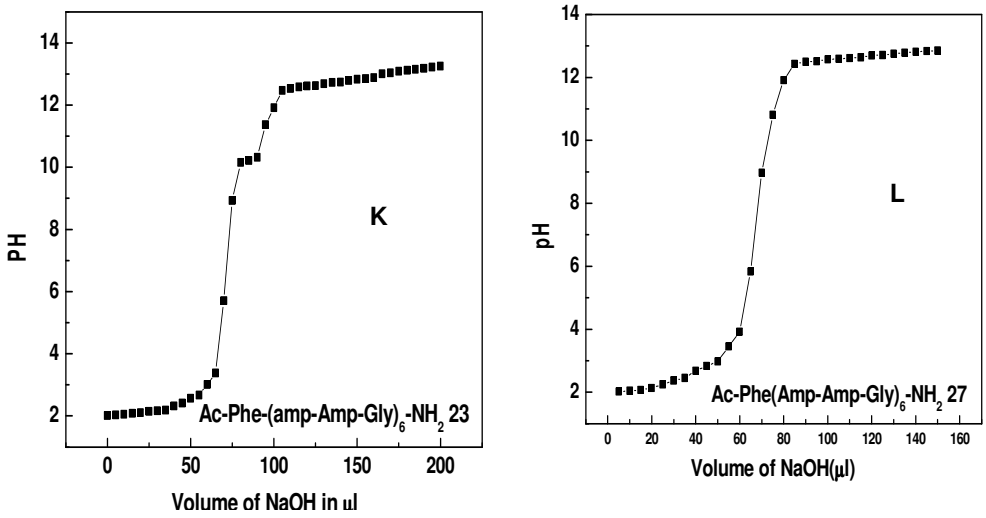


Figure 5.30: K and **L**; pH titration curves of 0.20mM solutions peptides **23** and **27.**

5.6: Discussion

An electron withdrawing group (-I group) in the 4-position decreases the bond order of the peptide bond that is, it weakens the partial double bond character, consequently increasing the *Z-E* isomerizaton rates and increase the preference for *Z* isomer. This preference arises from a decreased C^{γ}-C^{δ} bond length, pulling the C^{α} atom of proline away from the preceding residue thereby decreasing the steric hindrance. This structural manifestation of the inductive effect increases the stability of the *Z* (trans) isomer of the peptide bond.

5.6.1: Infrared spectroscopy

The $\gamma_{amide\text{-}I}$ may be used to estimate the relative inductive effect of the 4-substituent of proline on the prolyl-peptide bond. In H_2O, among the $–NH_2$, $-NH_3^+$ and $–N_3$ substituent groups, $–NH_3^+$ group exerts highest inductive effect as is evidenced by $\gamma_{amide\text{-}I}$ of compound **12** (Ac-Amp-OMe) and compound **13** (Ac-amp-OMe) at pH 2.0. Such a behavior is expected as the charged amine group has stronger inductive effect compared to free amine and azide groups. $–N_3$ and $–NH_2$ groups show almost the same absorption for amide groups.

The diastereomeric pair Ac-Azp-OMe **14** and Ac-azp-OMe **15** shows identical $\gamma_{amide\text{-}I}$ absorption both at acidic (1610.4cm^{-1}) and basic (1618.17cm^{-1}) conditions. Similar kind of behavior is observed for diastereomeric pair Ac-Amp-OMe **12** and Ac-amp-OMe **13** in acidic ($–NH_3^+$, 1650.95cm^{-1}) and basic ($–NH_2$, 1614.31cm^{-1}) conditions.

The 4-azido compounds **14** Ac-Azp-OMe and **15** Ac-azp-OMe show almost same absorption frequencies for both amide C=O and ester C=O groups at acid and basic conditions. In contrast to their corresponding free amine derivatives, compounds **12** Ac-Amp-OMe and **13** Ac-amp-OMe show drastic blue-shift (shorter wavelength) of absorption for amide C=O group when changing pH 2.0 (1650.9cm^{-1}) to pH 12.0 (1614.3cm^{-1}). This is due to the stronger inductive effect exerted by charged amine group increasing the amide C=O bond order. Such an increase in the bond order results in an equivalent increase in the IR-vibration frequency ($\gamma_{amide\text{-}I}$). Similar effect is also observed for γ_{ester} absorption.

5.6.2: ^{1}H-NMR spectroscopy

Compound **12** (Ac-Amp-OMe) and compound **14** (Ac-Azp-OMe) with 4-(2*S*,4*R)* stereochemistry show triplet like splitting pattern for H_1 proton in ^{1}H-NMR spectra in both acidic and basic conditions. This kind of triplet for H_α proton is observed for γ-exo pyrrolidine ring puckering. The compound **14** Ac-Azp-OMe at both acidic and basic conditions and compound **12** (Ac-Amp-OMe) only at basic condition (pH 12.0) exhibit long range M-type coupling between H_1 and H_6 protons which is characteristic of exclusive γ-*exo* ring puckering. Compound **12** (Ac-Amp-OMe) at pH 2.0 does not exhibit this kind of coupling Moreover it shows long range coupling between the H_3 and H_5 protons which is characteristic of exclusive γ-*endo* conformation. Hence compound **14** (Ac-Azp-OMe) at both pH 2.0 and 12.0 and compound **12** (Ac-Amp-OMe) only at pH 12.0 adopt γ-*exo* conformation, while compound **12** (Ac-Amp-OMe) at pH 2.0 adopts γ-*endo* conformation. This was further confirmed by $J_{H\text{-}H}$ coupling constant measurement. Similarly compound **15** (Ac-azp-OMe) with 2*S*,4*S* stereochemistry shows a doublet of doublet pattern for H_1 proton in ^{1}H-NMR spectra, both at acidic and basic conditions. This kind of triplet for H_1 is observed for γ-*endo* pyrrolidine ring puckering. Moreover it shows long range coupling between the H_3 and H_5 protons which is characteristic of exclusive γ-*endo* conformation of the pyrrolidine ring, The corresponding free amine derivative, compound **13** (Ac-amp-OMe) at acidic condition also shows doublet of doublet for H_1 and the long range coupling between the H_3 and H_5 protons in ^{1}H-NMR spectra indicates compound **13** (Ac-amp-OMe) to prefer γ-*endo* conformation. Surprisingly the splitting pattern of H_1 proton at pH 12.0 appears like multiplet at pH 12.0, which is contrary to the observed splitting pattern for H_1 in its diasteromer **12** (Ac-Amp-OMe), that shows a triplet at both acidic and basic conditions.

Unusual kind of splitting pattern for other protons (H_2, H_4, H_5 and H_6) is observed for compound **13** (Ac-amp-OMe) at pH 12.0, making it difficult to assign the preferred pyrrolidine ring conformation.

In general, compounds with free ionizable amino groups (**12** Ac-Amp-OMe and **13** Ac-amp-OMe) show two different pyrrolidine ring conformations at different pH (acidic pH 2.0 and basic pH 12.0) conditions, while compound with nonionizable substituent ($-N_3$) eg. **14** Ac-Azp-OMe and **15** Ac-azp-OMe a show preference of pyrrolidine ring conformation at both acidic and basic conditions. These compounds with 2*S*,4*R* stereochemistry prefer γ-exo pyrrolidine ring conformation and the compounds with 2*S*,4*S* stereochemistry prefer γ-endo pyrrolidine ring conformation. The magnitude of ring pucker changes with the nature of subtituents and the pH conditions. Compound **13** (Ac-amp-OMe) and **12** (Ac-Amp-OMe) change their preferred conformation with pH conditions. This is also evidenced by the $J_{H\text{-}H}$ coupling constant measurement of compounds.

5.6.3: Thermodynamic analysis

The slope of least square fits in the Van't Hoff plots (Fig. 5.19) for compound **12-15** at both acidic and basic pH conditions are different; this is due to the difference in enthalpies of isomerization of prolyl-peptide bonds. In general, the compounds with 4*R*-stereochemistry (**12** Ac-Amp-OMe and **14** Ac-Azp-OMe possess higher enthalpy of isomerization compared to the corresponding compounds with 4*S*-stereochemistry (**13** Ac-amp-OMe) and **15** Ac-azp-OMe). For all compounds, irrespective of the stereochemistry and pH conditions, the $K_{Z/E}$ values decrease with increasing temperature, at lower temperature *Z* (*trans*) form being preferred. Among the 4*R*-

substituted compounds, $-NH_3^+$ substituent shows highest $K_{Z/E}$ value (6.89), followed by $-N_3$ (6.53) and $-NH_2$ (6.06) substituents. The change of acidic to basic pH remarkably affects the $K_{Z/E}$ ($\Delta K_{Z/E}$ 0.83) values of compound **12** (Ac-Amp-OMe), even though *Z* isomer is considerably more stable under both pH conditions. Change of pH does not affect much the $K_{Z/E}$ value for compound **14** (Ac-Azp-OMe). Similarly among the 4*S*-substituted compounds $-N_3$ (azide) substituent has highest $K_{Z/E}$ (3.43) value followed by $-NH_3^+$ (2.63) and $-NH_2$ (1.33). Like for compound **12** (Ac-Amp-OMe), change of pH condition form acidic to basic pH drastically affects the $K_{Z/E}$ value of compound **13**. At pH 12.0, the $K_{Z/E}$ value for compound **13** (Ac-amp-OMe) is almost equal to 1 this indicates that both *E* and *Z* forms are present in 1:1 ratio like simple proline.

5.6.4: pKa of 4-amino group in 4-aminoproline collagen peptide

The stereochemical changes among 4*R* and 4*S* in X and Y positions of the collagen model peptides and chimeric peptides may result in differential solvent exposure experienced by the NH_2 group in different conformations adopted by the pyrrolidine rings at these positions in the triple helix. This coupled with protonation dependent puckering effects (stereoelectronic) and the necessity of alternating ring puckers in X and Y positions for effective triplex packing, may lead to the observed pH dependent triplex stability patterns for various 4-aminoproline containing collagen peptides. Mutually incompatible puckering of adjacent pyrrolidine rings results in non-detectable triplexes and the decrease in pKa values of neighbouring residue.

5.7: Implication for collagen stability

The effect of type and stereochemistry of 4-substituent on ring-puckering preferences of proline is very well understood. The change in the stereochemistry of 4-substitution alters the ring-puckering preference. The crystal structures[64] of the triple-

helical peptide $(Pro.Hyp.Gly)_{10}$, and $(Pro\text{-}Pro\text{-}Gly)_{10}$ clearly demonstrate that γ-*exo* pucker was favored for the Y residues (Hyp and Pro), and γ-*endo* pucker favored for the X residue Pro. For the cyclic imino acids Proline and 4-substituted prolines, change in the ring-pucker and thus the endocyclic torsion angles significantly alter the main-chain torsion angles Φ and ψ. It is well known that each of the X and Y residues of the polypeptide chain in the collagen triple-helix possess unique set of main-chain torsion angles. In case of an imino acid, these torsion angles are determined by the ring-pucker. Thus, γ-*endo* and γ-*exo* conformations in X and Y positions respectively are a necessary requirement for the formation of triple-helical structure.

It may seen that the observed $K_{Z/E}^{30°C}$ of 6.89 and the preferred γ-*endo* pyrrolidine ring conformation of Ac-Amp-OMe^+ **12** are incompatible since literature survey shows that the γ-*endo* pyrrolidine ring is always associated with the low $K_{Z/E}$ (around 2.0-3.5) value. The high preference of *trans* peptide bond can be explained by the through-space electrostatic interaction of the carbonyl group and ring nitrogen with the $–NH_3^+$ group. Hydrogen bonding between $–NH_3^+$ group with carbonyl group is feasible via water bridge when peptide bond is in *trans* conformation (Fig 5.31), leading to stabilization of the *trans-down* conformation.

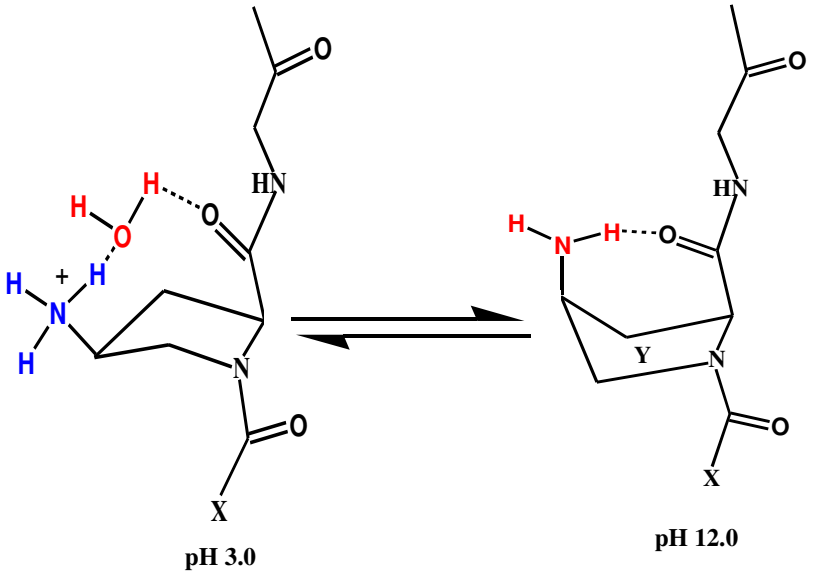


Figure 5.31: stabilizing interactions of 4*R*-aminoproline in trans-down and trans-up pyrrolidine ring pucker

The reasons for compatibility at X and Y positions and the `observed stabilization and destabilization of various 4-aminoproline containing collagen model peptides as discussed in previous chapters can be explained using the results of J_{H-H} coupling, and VT-NMR studies.

Table 5.8: Ring pucker and $K_{Z/E}$ value of peptide **19** at different pH conditions

Peptide 19 Ac-Phe(Pro(X)-Amp(Y)-Gly)$_6$-NH$_2$				
pH	**Puckering at Y**	**Puckering at X**	**$K_{Z/E}$ (Amp)**	**Triple helix**
3.0	Cγ-endo (c)	Cγ-endo(c)	6.83(c)	stabilizing
9.0	mix	mix	-	destabilizing
12.0	Cγ-exo(c)	Cγ-endo(c)	6.06(c)	satbilizing

c= compatible; mix= both Cγ-endo and Cγ-exo are present

*Peptide **19** Ac-Phe(Pro-Amp-Gly)$_6$-NH$_2$*

The high $K_{Z/E}$ (6.89) value observed for Amp monomer is suitable for Y position of the collagen peptide sequence $(X\text{-}Y\text{-}Gly)_n$, since the down–down ring puckering combination can also leads to triple-helical structures. The preferred γ-*endo* conformation at acidic condition (pH 3.0) results in the high triple-helicity of Amp containing peptide **19** Ac-Phe(Pro-Amp-Gly)$_6$-NH$_2$ (Chapter 2). The preferred γ-*exo* conformation and high $K_{Z/E}$ (6.06) value of Amp monomer at pH 12.0 is perfectly suitable for Y position resulting in the formation of stable triple-helical structure compared (Table 5.8) to the mixed puckering mode at the intermediate pH 9.0 of peptide **19** Ac-Phe(Pro-Amp-Gly)$_6$-NH$_2$

The different ring-puckering of Amp monomer at acidic and basic conditions suggests that this residue is compatible in the Y position of the triple-helix at both pH's without significantly affecting the main-chain torsion angles. It is also possible that the stability of Amp-collagen model peptides arises from mechanisms other than the increased *Z* isomer proportion and hydrogen bonding. The higher triple-helical stability

of collagenous peptides containing positively charged residues such as Arg (at Y position) also supports this view.

Table 5.9: Ring pucker and $K_{Z/E}$ value of peptide **20** at different pH conditions

Peptide 20 Ac-Phe(Amp(X)-Amp(Y)-Gly)$_6$-NH$_2$				
pH	**Puckering at Y**	**Puckering at X**	**$K_{Z/E}$ (Amp)**	**Triple helix**
3.0	Cγ-exo (c)	Cγ-endo(c)	6.83(ic)	destabilizing
12.0	Cγ-exo(c)	Cγ-exo(ic)	6.06(ic)	Not fold

*Peptide **20** Ac-Phe(Amp-Pro-Gly)$_6$-NH$_2$*

The preferred γ-*endo* conformation of Amp monomer at acidic condition is the reason for formation of triple-helix structure by this residue in the peptide **20** Ac-Phe(Amp-Pro-Gly)$_6$-NH$_2$ at pH 2.0, although its high $K_{Z/E}$ value is not suitable for X position (Table 5.9), and the extent of triple-helix structure at X position compared to Y in acidic condition is very low. The switching of conformation from γ-*endo* to γ-*exo* upon changing the pH from 2.0 to 12.0 and the high $K_{Z/E}$ value of Amp monomer are the main reasons for the peptide **20** Ac-Phe(Amp-Pro-Gly)$_6$-NH$_2$ to destabilizing the triplex conformation.

Peptide	**Puckering (Y)**	**Puckering (X)**	**$K_{Z/E}$[Y/X]**	**Triplex-helix**
23 [amp(X)-Amp(Y)-Gly]	Cγ-endo(c)	Cγ-endo(c)	6.83/2.63	stabilizing
24 [amp(X)-Hyp(Y)-Gly]	Cγ-exo (c)	Cγ-endo(c)	-/2.63	stabilizing

c= compatible; ic= incompatible

*Peptide **22** Ac-Phe(amp-Pro-Gly)$_6$-NH$_2$ and Peptide **21** Ac-Phe(Pro-amp-Gly)$_6$-NH$_2$*

Similarly the $K_{Z/E}^{30°C}$ of 2.63 and the preferred γ-*endo* pyrrolidine ring conformation in Ac-amp-OMe$^+$ **13** are well suited in X position but disfavored for Y position (Table 5.10) of the collagen peptide sequence (X-Y-Gly)$_n$. This results in the high triple-helicity of amp containing peptide **22** Ac-Phe(amp-Pro-Gly)$_6$-NH$_2$ and destabilization of triple-helix of peptide **21** Ac-Phe(Pro-amp-Gly)$_6$-NH$_2$. The

destabilization of triple-helix of peptide **22** Ac-Phe(amp-Pro-Gly)$_6$-NH_2 at pH 12.0 can be explained by unfavorable very low $K_{Z/E}^{30°C}$ value of 1.33

Similar reasons explain the stabilizing and destabilizing effects of Amp and amp molecules in chimeric peptides studied in Chapter 3.

*Peptide **23** Ac-Phe(amp-Amp--Gly)$_6$-NH_2 and peptide 24 Ac-Phe(amp-Hyp-Gly)$_6$-NH_2*

The preferred γ-*endo* pyrrolidine ring pucker and $K_{Z/E}$ value 2.63 of amp when present at X position are compatible with residues Amp and Hyp at Y position of the collagen peptide (X-Y-Gly)$_n$ at acidic pH 3.0 . These preorganized combination of amp at X position and Amp or Hyp at Y position enables the three strand to wind in triple-helical conformation with tight packing and hence the peptides **23** Ac-Phe(amp-Amp-Gly)$_6$-NH_2 and peptide **24** Ac-Phe(amp-Hyp-Gly)$_6$-NH_2 form more stable triple-helical structure (Table 5.11) compared to the corresponding proline peptides Ac-Phe(Pro-Amp-Gly)$_6$-NH_2 **19** and Peptide Ac-Phe(Pro-Hyp-Gly)$_6$-NH_2 (control pepide).

Table 5.11: Ring pucker and $K_{Z/E}$ value of peptides **23** and **24** at pH 3.0

Peptide 22 Ac-Phe(amp(X)-pro(Y)-Gly)$_6$-NH_2				
pH	**Puckering at Y**	**Puckering at X**	**$K_{Z/E}$ (Amp)**	**Triple helix**
3.0	Cγ-exo (c)	Cγ-endo(c)	2.63(c)	stabilizing
12.0	Cγ-exo(c)	-	1.33(ic)	desatbilizing

c= compatible;

The steric repulsion between the –OH group of 4*S*-hyp at X position with amide carbonyl group of adjacent chain as observed in the peptide (hyp-Pro-Gly)$_{10}$ makes peptide unsuitable to assume triple-helical conformation also in the peptides Ac-Phe(hyp-Amp-Gly)$_6$-NH_2 **25** and Ac-Phe(hyp-Amp-Gly)$_6$-NH_2 **26** and thereby resulting in destabilization.

*Peptide **28** Ac-Phe(Hyp-Amp--Gly)$_6$-NH$_2$*

In acidic condition, peptide **28** Ac-Phe(Hyp-Amp-Gly)$_6$-NH$_2$ exhibits up-down (ud) ring puckering (Table 5.12), and in basic condition, it exhibits up-up ring puckering. The up-up (uu) conformer is 1Kcal/mole more stable compared down-down (du) and hence the peptide **28** Ac-Phe(Hyp-Amp-Gly)$_6$-NH$_2$, stabilizes the triple-helical structure at pH 12.0 (Δ T$_m$ +7.0 ^{0}C) compared at pH 2.0 (ΔT$_m$ -4.2^{0}C) with respect to the Amp peptide **19** Ac-Phe(Pro-Amp-Gly)$_6$-NH$_2$.

Table 5.12: Ring puckering of peptides **28** at pH 3.0 and 12.0

pH	Puckering at Y	Puckering at X	Triple helix
3.0	Cγ-exo (c)	Cγ-endo(c)	stabilizing
12.0	Cγ-exo(c)	Cγ-exo(c)	satbilizing

c= compatible;

*Peptide **27** Ac-Phe(Amp-Amp--Gly)$_6$-NH$_2$*

The down-down (dd) conformation in the peptide **27** Ac-Phe(Amp-Amp-Gly)$_6$-NH$_2$ at pH 2.0 is favored to form triple-helix structure like the up-up (uu) conformation. The two charged aminoprolines in each strand of peptide **27** are in close proximity to each other while packing into triple-helix with large electrostatic repulsive interactions between the chains, consequently destabilizing the triple-helix (ΔT$_m$ -8.5^{0}C) with respect to the Amp peptide **19** Ac-Phe(Pro-Amp-Gly)$_6$-NH$_2$.

The peptides **27** Ac-Phe(Amp-Amp-Gly)$_6$-NH$_2$ and **28** Ac-Phe(Amp-Hyp-Gly)$_6$-NH$_2$ at pH 12.0 are expected to form a more stable triple-helix structure since both peptides exhibit up-up (uu) conformation at pH 12.0. Surprisingly the temperature dependent CD spectroscopic study showed that these peptides do not form triple-helical structures. In γ-*exo* conformation the 4–NH$_2$ group of Amp is in axial position. The extra hydrogen atom present in –NH$_2$ group may cause the steric repulsion with other strand, causing the peptide to destabilize the triple-helical structure, similar to –OH

group of 4*S*-hyp in the peptide (hyp-Pro-Gly)$_{10}$. Hence all the peptides with 4*R*-Amp at X position do not form triple-helical structure at pH 12.0.

The formation of highly stable collagen type triple helices at acidic pH by peptides with ionizable 4*R*-Amp and/or 4*S*-amp may be a consequence of factors beyond mere stereoelectronic and hydrogen bonding effects. Significant contributions can originate from electrostatic effects. The determination of pK_a values for ionizable NH_2 group in different 4-aminocollagen peptides and their triple-helix stabilities at different pH are shown in Table 5.13. The pK_a of 4-NH_2 at the monomeric level was significantly stereochemistry dependent, with 4*S*-amp (pK_a 9.3) inherently more acidic is significantly reduced to the range 6.0-9.0 in the triplexes and the pK_a of 4*S* and 4*R*-aminoprolines exhibited both stereochemistry and position dependence.

Table 5.13: Tm values at different pHs and pKa values of various collagen model peptides.

peptides	T_m ^{0}C				pKa value
	pH 3.0	pH 7.0	pH 9.0	pH 12.0	
Ac-Phe(Pro-Amp-Gly)$_6$-NH_2 **19**	60	54.7	26	46	8.30
Ac-Phe(Amp-Pro-Gly)$_6$-NH_2 **20**	36	36	-nt-	-nt-	7.86
Ac-Phe(amp-Pro-Gly)$_6$-NH_2 **22**	44	37	34	-nt-	7.03, 9.64
Ac-Phe(Pro-amp-Gly)$_6$-NH_2 **21**	-nt-	-nt-	-nt-	-nt-	6.07
Ac-Phe(amp-Hyp-Gly)$_6$-NH_2 **24**	49.8	39.3	37	-nt-	6.73
Ac-Phe(Amp-Hyp-Gly)$_6$-NH_2 **29**	43.1	36	29	-nt-	8.44
Ac-Phe(Hyp-Amp-Gly)$_6$-NH_2 **28**	55.8	51.5	31	53	9.06
Ac-Phe(Amp-Amp-Gly)$_6$-NH_2 **27**	51.5	40.8	25	-nt-	8.93
Ac-Phe(amp-Amp-Gly)$_6$-NH_2 **23**	61	46.6	40.5	34	7.1, 11.2

-nt- no transition was observed

The 4*R*-Amp peptides **19** (Pro-Amp-Gly) with pKa of 8.3 is in fully protonated form at pH 3.0 and 7.0 and non protonated at pH 12. Hence it forms more stable triple helix at both acidic and extreme basic pH conditions. At pH 9.0, a mixture of partially protonated and nonprotonated states of 4-aminogroup (since its pKa is 8.3) results non-

uniform pyrrolidine ring puckering, affecting the tight packing of three chains to form triple-helical structure. Consequently, a decrease in triplex staility at pH 9.0 is observed. Similarly peptide **20** (Amp-Pro-Gly) with pKa value of 7.86 is present in fully protonated form at pH 3.0 and 7.0 and fully free amine form at pH 9.0 and 12. In protonated state, the *γ-endo* pucker adopted by 4*R*-aminoproline at X position in the peptide **20** (Amp-Pro-Gly) stabilizes the triple-helical structure. The *γ-exo* pyrrolidine ring conformation of 4*R*-aminoproline at pH 9.0 and 12.0 results in incompatible puckering leading to its decline to form folded triple-helical structure. The presence of Hyp residue in the Y position of the peptide **28** (Amp-Hyp-Gly) increases the pKa value (8.44) compared to peptide **20** (Amp-Pro-Gly) (pKa 7.86). Similarly the pKa value of peptide **27** (Amp-Amp-Gly) is 8.93 significantly more than either Pro(X) or Hyp (Y) peptides.This indicates that Amp residue in peptides **27** and **28** is considerably protonated at pH 9.0 and hence these two peptides form triple-helical structure at pH 9.0. The protonation of 4*R*-aminoproline at X position is prerequisite to form triple-helical structure.

Peptide **22** (amp-Pro-Gly) shows two pKa values of 7.03 and 9.64. The lower pKa value may be the result of N-terminal amp which is highly exposed to solvent system and loosely packed in triple-helical structure and the internal amp residues which are packed inside the triple-helical structure showing a higher pKa value of 9.94. This suggests that amp residues in peptide **22** (amp-Pro-Gly) are always in protonated form at pH 3.0, 7.0 and 9.0. The *γ-endo* pucker and the $K_{Z/E}$ value of amp residue in protonated form are compatible at X position of the peptide **20** (amp-Pro-Gly) and hence this peptide associates to triple-helical conformation at these pHs. At pH 12.0, amp residues are completely in free amine form with incompatible ring puckering and a

very low $K_{Z/E}$ value (1.33) not suitable at X position, leading to destabilization of the triple-helix. Similarly peptide **23** (amp-Amp-Gly) shows two pKa values of 7.1 and 11.2. this suggested that some of the aminoproline residues in the peptide **21** (amp-Amp-Gly) are in protonated form even at pH 12.0 and hence this peptide forms stable triple-helices even at pH 12.0

5.8: Conclusion

The protonatable 4-aminoproline is shown to be a proline analogue which differs from known proline analogues for the following reasons. In case of other 4-substituted proline analogs (-F, -OH, $-N_3$, -Me etc….) the pyrrolidine ring conformation is dependent mainly on the stereochemistry and electronegativity of the 4-substituent. The pyrrolidine ring conformation of the 4-aminoproline is dependent not only on the stereochemistry but also on the protonation status of the 4-aminogroup.

The present study has demonstrated that the prolyl-peptide bond isomerization and the ring conformation of the proline are indeed affected by the inductive effect of the 4-substituent. The conformational equilibria such as *E-Z* isomerization and the ring-pucker of the imino acid proline are significantly affected by the 4-substituent. New and novel proline analogues such as 4*R*-azidoproline (Azp) and 4*S*-azidoproline (azp) have been studied. These, amino acids may provide new proline-hydoxylase inhibitors, and may be useful in the design of new class of collagen mimetics β-turn mimetic as biologically useful peptides.

5.9: Experimental

The chemicals used were of synthesis or reagent grade. D_2O was obtained from Aldrich USA. NMR spectra were obtained on a Bruker 400 MHz instrument equipped

with Aspect 2000 computer. The chemical shifts were measured with reference to TMS. Spectral plotting was carried out using ACD labs, Specviewer, freeware.

5.9.1: Synthesis compound 12-15.

*General procedure for synthesis of compound **14** and **15***

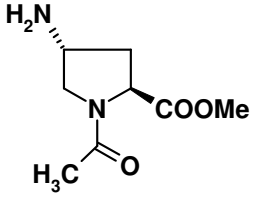
The azide (12g, 44.23mmol) compound was stirred for 2hrs in 8ml of 1:1 mixture of CH_2Cl_2 and TFA. The solvent was removed under reduced pressure; resulting slurry was dissolved in diethyl ether, removal of ether yielded white compound of trifluoroacetate salt. The salt obtained was stirred in 20ml of 1:1solution of acetic anhydride in anhydrous pyridine for 8hrs. Solvent was removed under reduced pressure. The resulting crude product was purified by column chromatography (EtOAc/petether 7:6) to offered color less liquid.

*General procedure for synthesis of compound **12** and **13***

To a solution of azide (12g, 44.23mmol) in methanol (8ml) taken in hydrogenation flask was added Pd/C catalyst (10 %). The reaction mixture was hydrogenated in a Parr shaker apparatus for 4hrs at room temperature under H_2 pressure of 45-50 psi. The catalyst was filtered off, and then solvent was removed under reduced pressure to yield a residue of amine, which was purified by column chromatography (MeOH/EtOAc 0.5:1) to pale yellow liquid.

Spectral characterization of compound 12-15

*(2S,4R)-N^1-acetyl-4-aminoproline methylester **12***



Mol. F: $C_8H_{14}N_2O_3$

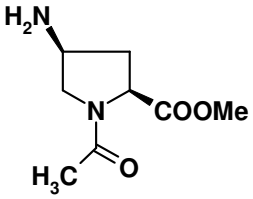
Mol. Mass: 186.1 (calculated), 187.17 (absorved) M+1.

IR: ($CHCl_3$) 1741.6, 1641.3, 1546.8, 1425.3 1207.4.

^{1}H NMR (D_2O, pD 12.0 400MHz) δ 2.07 (3H, s), 2.12-2.27 (2H, m), 3.20-3.31(1H, m), 3.17(3H, s) 3.78-3.38 (1H, m) 4.52-4.59(1H m).

^{13}C NMR: 23.1, 24,1, 36.2, 37.7, 50.9, 51.5, 52.2, 53.9, 55.7, 62.2, 63.8, 175.1, 176, 180.4.

(2S,4S)-N^{1}-acetyl-4-aminoprolin e methylester ***13***



Mol. F: $C_8H_{14}N_2O_3$

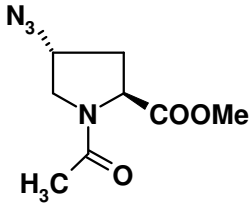
Mol. Mass: 186.1 (calculated), 187.16 (absorved) M+1.

IR: ($CHCl_3$) 1739.7, 1633.6, 1446, 1205.4.

^{1}H NMR (D_2O, pD 12.0 400MHz) δ 2.10(3H s), 2.02-2.07(1H d), 2.34-2.53(1H m), 3.34- 3.43(1H t), 3.76(3H s), 4.36-4.50(1H m).

^{13}C NMR: 23.9, 24.4, 35.6, 36.8, 51.4, 52, 52.7, 54.3, 56.2, 57.6, 60.5, 61.8, 63.2, 64.6. 175.9, 176.2, 181.8.

(2S,4R)-N^{1}-acetyl-4-azidoproline methylester ***14***



Mol. F: $C_8H_{12}N_4O_3$

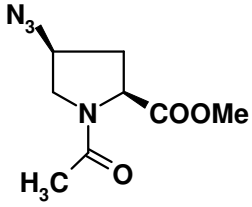
Mol. Mass: 212.09 (calculated), 213.17 (absorved) M+1.

IR: ($CHCl_3$) 2108.1, 1747.4, 1651.0

^{1}H NMR ($CDCl_3$, 400MHz) δ 2.11(3H s), 2.22-2.46(2H m) 3.47-3.54(1H dd), 3.75(3H S) 3.86-3.91(1H t) 4.26-4.38 (1H q), 4.51-4.59 (1H t).

^{13}C NMR: 21.8. 22, 34.4, 36.5, 51.2, 52.2, 52.3, 52.7, 57, 59.2, 169.4, 171.1.

(2S,4S)-N^{1}-acetyl-4-azidoproline methylester ***15***



Mol. F: $C_8H_{12}N_4O_3$

Mol. Mass: 212.09 (calculated), 213.15(absorved) M+1.

IR: ($CHCl_3$) 2108.1, 1747.4, 1652.9, 1417.6.

^{1}H NMR ($CDCl_3$, 400MHz) δ 2.11(3H s), 2.18-2.25(1H t), 2.41-1.55(1H q), 3.54-3.63(1H q), 3.75(3H s) 4.23-4.30(1H m), 4.59-4.66(1H dd).

^{13}C NMR: 21.3, 21.9, 34.5, 36.7, 52.4, 52.7, 57.1, 57.8, 59.3, 170.3, 174.4, 176.3.

5.9.2: NMR experiments

Variable temperature NMR

VT-NMR experiments were performed on a Bruker DRX-400 MHz instrument controlled by Bruker X-Win NMR software. Spectra were obtained using 5 mm high temperature probe and the variable temperature unit was calibrated with ethyleneglycol. Samples of **12**-**15** for NMR experiments were prepared at a concentration of 0.1 M in D_2O. For the preparation of NMR samples in pH 2.0, the respective compounds were dissolved in D_2O (0.1 M) and the solution was brought to pD 2.0 with $KHSO_4$ (direct pH meter reading). In order to remove any volatile impurities present, all the samples were freeze dried and the resulting solids were redissolved in D_2O. Similarly for the preparation of NMR samples in pH 12.0, the respective compounds were dissolved in borate buffer solution (prepared using 20mm boric acid solution in deionized water). The samples were freeze dried and the resulting solids were redissolved in D_2O.

The equilibrium constants for the *E*-*Z* interconversion of **12**-**15** were directly obtained from peak integration of ^{1}H signals for minor (*E*) and major (*Z*) isomers. A typical experiment was performed by first bringing the sample to the highest temperature of measurement followed by equilibration. The temperature was decreased in steps of 10 K and the spectra were collected only when a stable integration ratio of major/minor isomer was achieved. Typically equilibration times of 15-20 minutes were necessary. Spectra were collected at 0.1 Hz spectral-resolution and were an average of

150 scans. In order to achieve a reliable integration, a post acquisition delay of 5 seconds was also applied and the number of scans was kept constant throughout the experiment. Necessary phase corrections were applied and the peak integration was taken for any one set of well-separated signals. N^1-acetyl –C<u>H</u>$_3$ and H_β protons were found most suitable. Integration was also compared over more than one set of peaks. Curve fitting was carried out using Microcal Origin software and the standard errors were obtained from the software output.

5.9.3: Conformational analysis

Samples for ^{1}H-NMR spectroscopy were prepared at a concentration of 2 mg/mL. Spectra were collected at 400 MHz on a Bruker DRX-400 instrument equipped with a 5 mm QNP probe. Coupling constants were measured using Mestrec Freeware or Bruker Xwin-NMR software. Wherever necessary spectra were resolution enhanced using appropriate window functions such as, Lorenzian/Gaussian or exponential multiplication and were zero-filled.

5.10: Reference

1) Anfinsen, C. B. *Science* **1973**, *181*, 223-230.

2) Dill, K. A. *Biochemistry 1990*, *29*, 7133-7155.

3) Brestscher, L. E.; Jenkins, L. C.; Taylor, M. K.; DeRider, L. M.; Raines, R. T. *J. Am. Chem. Soc. 2001*, *123*, 777-778.

4) Deslongchamps, P. *Stereoelectronic Effects in Organic Chemistry*; Pergamon Press: New York, **1983**.

5) Gorenstein, D. G. *Chem. Rev*. **1987**, *87*, 1047-1077.

6) Plavec, J.; Thibaudeau, C.; Chattopadhyaya, J. *J. Am. Chem. Soc*. **1994**, *166*, 6558-6560.

7) Eberhardt, E. S.; Panasik, N., Jr.; Raines, R. T. J. Am. Chem. Soc. 1996, 118, 12261-12266.

8) (a) Holmgren, S. K.; Taylor, K. M.; Brestscher, L. E.; Raines, R. T.; *Nature* **1998**, *392*, 666-667. (b) Holmgren, S. K.; Brestscher, L. E.; Taylor, K. M.; Raines, R. T. *Chem. Biol*. **1999**, *6*, 63-70.

9) Renner, C.; Alefelder, S.; Bae, J. H.; Budisa, N.; Huber, R.; Moroder, L. *Angew. Chem. Int. Ed.* **2001**, *40*, 923-925.

10) Fischer, S.; Dunbrack, R. L., Jr.; Karplus, M. *J. Am. Chem. Soc*. **1994**, *116*, 11931-11937.

11) Kang, Y. K.; *THEOCHEM* **2002**, *585*, 209-221.

12) Kang, Y. K.; Jhon, J. S.; Han, S. *J. Peptide Res*. **1999**, *53*, 30-40

13) Zimmermann, S. S.; Scheraga, H. A. *Biopolymers* **1977**, *16*, 811-843.

14) (a) Halab, L.; Lubell, W. D. *J. Org. Chem*. **1999**, *64*, 3312-3321. (b) Halab, L.; Lubell, W. D. *J. Peptide Sci*. **2001**, *7*, 92-104.

15) Cox, C.; Lectka, T. J. Am. Chem. Soc. 1998, 120, 10660-10668.

16) (a) Mizushima, S.; Shimanouchi, T.; Tsuboi, M.; Susita, T.; Kurosaki, K.; Mataga, N.; Souda, R. *J. Am. Chem. Soc*. **1952**, *74*, 4639-4641. (b) Matsuzaki, T.; Iitaka, Y. *Acta Crystallogr*. **1971**, *B27*, 507-516.

17) Liang, G. B.; Rito, C. J.; Gellman, S. H. *Biopolymers* **1992**, *32*, 507-516.

18) Zimmerman, S. S.; Scheraga, H. A. Macromolecules 1976, 9, 408-416.

19) (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.

20) (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.

21) Taylor, C. M.; Hardre, R.; Edwards, P. J. B.; Park, J. H. *Organic letters* **2003**, *5*, 4413-4416.

22) (a) MacArthur, M. W.; Thornton, J. M. *J. Mol. Biol*. **1991**, *218*, 397-412. (b) Stewart, D. E.; Sarkar, A.; Wampler, J. E. *J. Mol. Biol*. **1990**, *214*, 253-260.

23) (a) Stimson, E. R.; Montenlione, G. T.; Meinwald, Y. C.; Rudolph, R. K. E.; Scheraga, H. A. *Biochemistry* **1982**, *21*, 5252-5262. (b) Juy, M.; Lam-Thanh, H.; Lintner, K.; Fermandjian, S. *Int. J. Peptide Protein Res*. **1983**, *22*, 437-449.

24) Poznanski, J.; Ejchart, A.; Wierzchowski, K. L.; Ciurak, M. *Biopolymers* **1993**, *33*, 781-795.

25) Halab, L.; Lubell, W. D. *J. Am. Chem. Soc*. **2002**, *124*, 2474-2484.

26) (a) Dougherty, D. A. *Science* **1996**, *217*, 163-168. (b) Gallivan, J. P.; Dougherty, D. A. *J. Am. Chem. Soc*. **2000**, *122*, 870-874.

27) Fischer, E.; *Ber. Deutsch. Chem. Ges*. **1906**, *39*, 530-610.

28) Stewart, D. E.; Sarkar, A.; Wampler, J. E. *J. Mol. Biol*. **1990**, *214*,253-260.

29) Weiss, M. S.; Jabs, A.; Hilgenfeld, R. *Nat. Struct. Biol*. **1998**, *5*, 676.

30) Jabs, A.; Weiss, M. S.; Hilgenfeld, R. *J. Mol. Biol.* **1999**, *286*, 291-304.

31) Brandts, J. F.; Halvorson, H. R.; Brennan, M. *Biochemistry* **1975**, *14*, 4853-4963.

32) Bächinger, H. P.; Bruckner, P.; Timpl, R.; Prockop, D. J.; Engel, J. *Eur. J. Biochem.* **1980**, *106*, 619-632.

33) Bächinger, H. P. *J. Biol. Chem*. **1987**, *262*, 17144-17148.

34) Liu, X.; Siegel, D. L.; Fan, P.; Brodsky, B.; Baum, J. *Biochemistry* **1996**, *35*, 4306-4313.

35) Dealaney, N. G.; Madison, V. *J. Am. Chem. Soc*. **1982,** *104*, 6635-6641.

36) Beausoleil, E.; Sharma, R. M.; Michnick, S. W.; Lubell, W. D. *J. Org. Chem*. **1998**, *63*, 6572-6578.

37) Magaard, V. W.; Sanchez, R. M.; Bean, J. W.; Moore, M. L. *Tetrahedron Lett*. **1993**, *34,* 381-384.

38) Aa, S.-S. A.; Lester, C. C.; Peng, J.-L.; Li, Y.-J.; Rothwarf, D. M.; Welker, E.; Tannhauser, T. W.; Zhang, L. S.; Tam, J. P.; Scheraga, H. A. *J. Am. Chem. Soc*. **1999**, *121*, 11558-11566.

39) Halab, L.; Gosselin, F.; Lubell, W. D. *Biopolymers* **2000**, *55*,101-122.

40) Arnold, U.; Hinderaker, M. P.; Koeditz, J.; Golbik, R.; Ulbrich-Hoffman, R.; Raines, R. T. *J. Am. Chem. Soc*. **2003**, *125*, 7500-7501.

41) Andres, C. J.; Macdonald, T. L.; Ocain, T. D.; Longhi, D. *J. Org. Chem.* **1993**, *58*, 6609-6613.

42) Welch, J. T.; Lin, J. *Tetrahedran* **1996**, *52*, 291-304.

43) Hart, S. A.; Sabt, M.; Etzkorn, F. A. *J. Org. Chem*. **1998**, *63*, 7580-7581.

44) Lin, J.; Toscano, P. J.; Welch, J. T. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 14020-14024.

45) Otaka, A.; Katagiri, F.; Kinoshita, T.; Odagaki, Y.; Oishi, S.; Tammamura, H.; Hamanaka, N.; Fujii, N. *J. Org. Chem*. **2002**, *67*. 6152-6161.

46) Wang, X. J.; Hart, S. A.; Xu, B.; Mason, M. D.; Goodell, J. R.; Etzkorn, F. A. *J. Org. Chem.* **2003**, *68*, 2343-2349.

47) Jenkins, C. L.; Lin, G.; Duo, J.; Rapolu, D.; Guzei, I. A.; Raines, R. T.; Krow, R. G. *J. Org. Chem*. **2004**, *69*, 8565-8573.

48) Eberhardt, E. S.; Panasik, N. J.; Raines, R. T.; *J. Am. Chem. Soc*. **1996**, *118*, 12261-12266.

49) Bretscher, L. E.; Taylor, K. M.; Raines, R. T. *In Peptides for the New Millennium: Proceedings of the Sixteenth American Peptide Symposium*; Dordrecht, The Netherlands, **2000**, pp 335-356.

50) Bretscher, L. E.; Jenkins, C. L.; Taylor, K. M.; DeRider, M. L.; Raines, R. T. *J. Am. Chem. Soc*. **2001**, *123*, 777-778.

51) Renner, C.; Alefelder, S.; Bae, J. H.; Budisa, N.; Huber, R.; Moroder, L. *Angew. Chem. Int. Ed.* **2001**, *40*, 923-925.

52) (a) Matzuzaki, T.; Iitaka, Y. *Acta Crystallogr*. **1971**. *B27*, 507-516. (b) Liang, G. B.; Rito, C. J. Gellman, S. H. *Biopolymers* **1992**, *32*, 293-301. (c) Flippen-Andersion, J. L.; Gilardi, R.; Karle, I. L.; Fery, M. H.; Opella, S. J.; Goodman, M.; Madison, V.; Deleney, N. G. *J. Am. Chem. Soc*. **1983**, *105*, 6609-6614.

53) Ranner, C.; Alefelder, S.; Bae, J. H.; Budisa, N.; Huber, R.; Moroder, L. *Angew. Chem. Int. Ed.* **2001**, *40*, 923-925.

54) (a) Eberhardt, E. S.; Loh, S. N. P.; Raines, R. T. *Tetrahedron Lett.* **1993,** *33*, 3055-3056. (b) Eberhardt, E. S.; Panasik, N, Jr.; Raines, R. T. *J. Am. Chem. Soc.* **1996**, *118*, 12261-12266.

55) (a) Jackso, M.; Matsch, H. H. *Crit. Rev. Biochem. Mol. Biol.* **1995**, *30*, 95-120. (b) Krimm, S.; Banderker, J. *In. Advances in Protein Chemistry*; Academic Press: NY, **1986**, Vol. *38*, pp 183-364.

56) Bretscher, L. E.; Jenkins, C. L.; Taylor, K. M.; DeRider, M. L.; Raines, R. T. J. Am. Chem. Soc. 2001, 123, 777-778.

57) (a) Pansik, N. Jr.; Eberhardt, E. S.; Edison, A. S.; Powell, D. R.; Raines, R. T. *Int. J. Pept. Protein Res.* **1994**, *44*, 262-269. (b) Gerig, J. T.; McLeod, R. S. J. *Am. Chem. Soc.* **1973**, *95*, 5725-5729.

58) Harris, T. K.; Turner, G. J. *IUBMB Life* **2002**, *53*, 85-98.

59) Antosiewicz, J.; McCammon, J. A.; Gilson, M. K. *J. Mol. Biol.* **1994**, *35*, 7819-7833.

60) Nielsen, J. E.; Vriend, G. *Proteins* **2001**, *43*, 403-412.

61) Gibas, C. J.; Subbramaniam, S. *Biophys. J.* **1996**, *71*, 138-147.

62) Kumar, S.; Nussinov, R. *Protein Eng*. **1999**, *12*, 657-662.

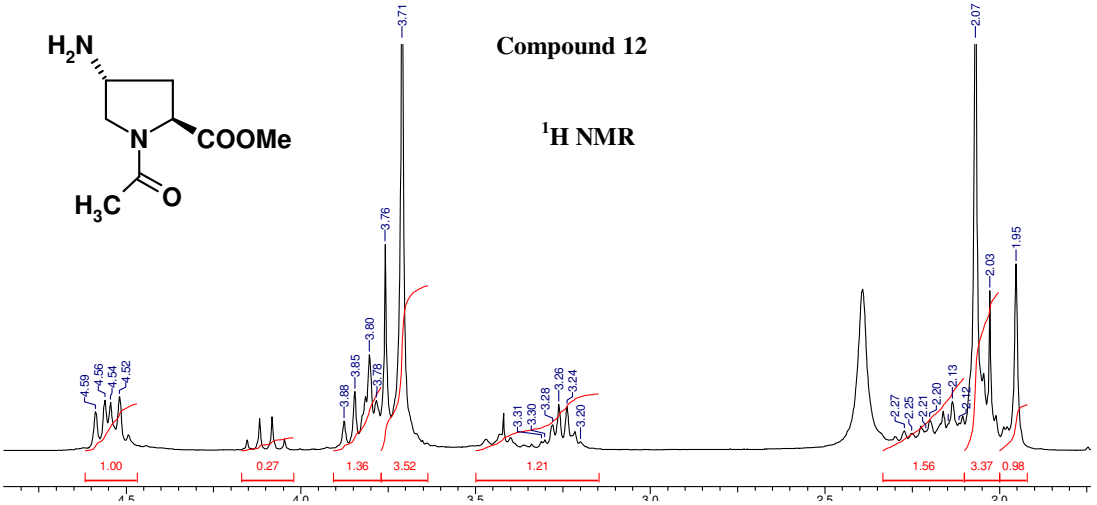
63) Nielsen, J. E.; McCammon, J. A. *Protein Sci*. **2003**, *12*, 313-326.

64) a) Nagarajan, V.; Kamitori, S.; Okuyama, K. *J. Biochem.* **1999**, *125*, 310-318. b) Berisio, R. ; Vitagliano, L.; Mazzarella, L.; Zagari, A. *Protein Science* **2002**, *11*, 262-270.

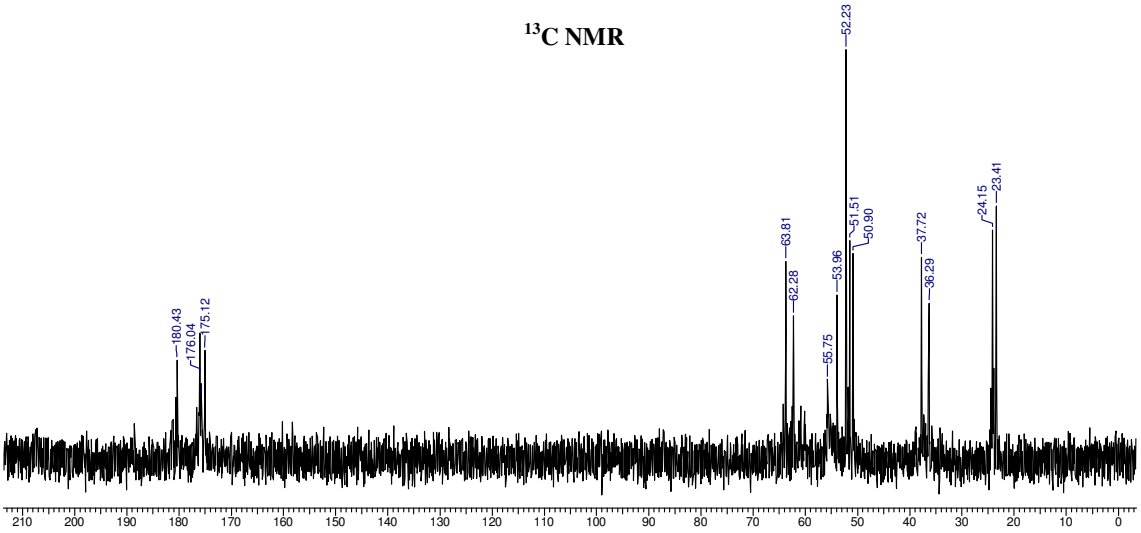
Appendix IV

Compound No.	Spectra	Page No
Compound **12** ……	^{1}HNMR,^{13}CNMR, DEPT…….	327
Compound **13** ……	^{1}HNMR,^{13}CNMR, DEPT…….	328
Compound **14**……	^{1}HNMR,^{13}CNMR, DEPT…….	329
Compound **15**……	^{1}HNMR,^{13}CNMR, DEPT…….	330
Compound **12** and **13**……	LC-MS …….	331
Compound **14** and **15**……	LC-MS …….	332

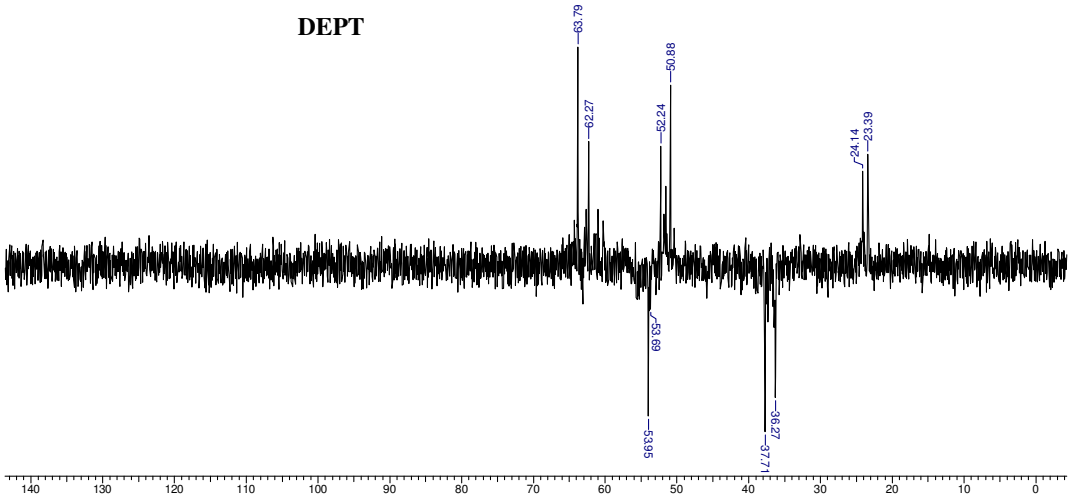
Acquisition Time (sec) 7.9167	***Comment*** M. Umashankar			***Date*** 30/03/2006 17:31:04			
Frequency (MHz) 200.13	***Nucleus*** 1H		***Original Points Count*** 32768	***Points Count*** 32768		***Sweep Width (Hz)*** 4139.07	
Temperature (grad C) 0.000							



Acquisition Time (sec) 2.7329	***Comment*** M UMASANKARA			***Date*** 04/09/2006 20:22:30			
Frequency (MHz) 50.32	***Nucleus*** 13C		***Original Points Count*** 32768	***Points Count*** 32768		***Sweep Width (Hz)*** 11990.41	
Temperature (grad C) 0.000							

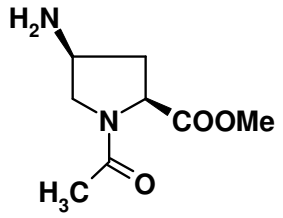


Acquisition Time (sec) 2.7329	***Comment*** M UMASANKARA			***Date*** 04/09/2006 20:41:52			
Frequency (MHz) 50.32	***Nucleus*** 13C		***Original Points Count*** 32768	***Points Count*** 32768		***Sweep Width (Hz)*** 11990.41	
Temperature (grad C) 0.000							



Acquisition Time (sec) 7.9167	*Comment*	M UMASANKARA		*Date*	01/04/2006 09:20:32		
Frequency (MHz) 200.13	*Nucleus*	1H	*Original Points Count* 32768	*Points Count*	32768	*Sweep Width (Hz)*	4139.07
Temperature (grad C) 0.000							

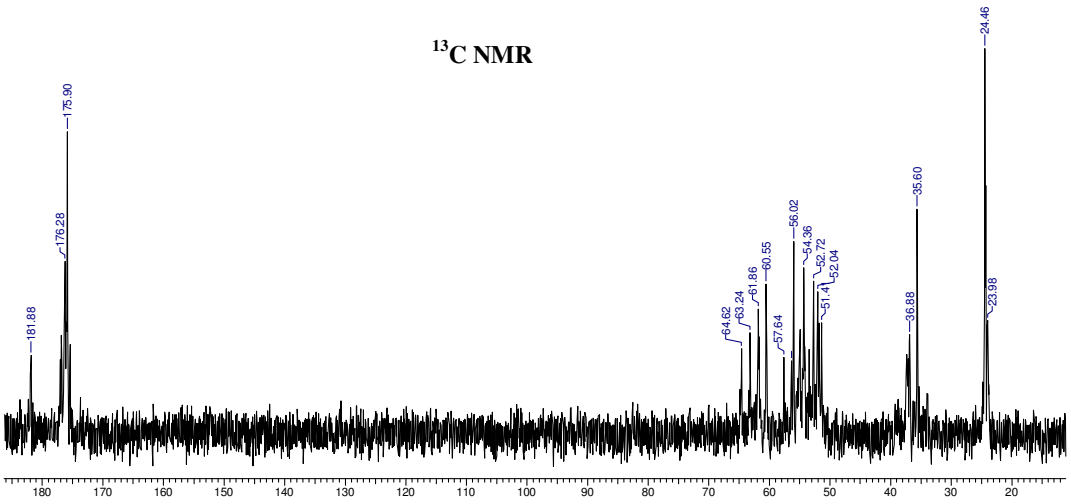
Compound 13



^{1}H NMR

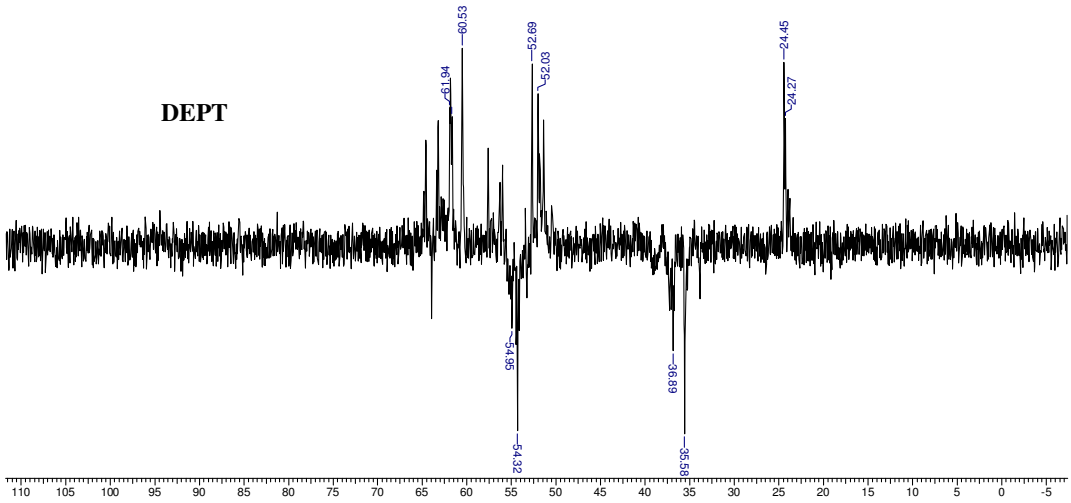
Acquisition Time (sec) 2.7329	*Comment*	M UMASANKARA		*Date*	04/09/2006 21:27:00		
Frequency (MHz) 50.32	*Nucleus*	13C	*Original Points Count* 32768	*Points Count*	32768	*Sweep Width (Hz)*	11990.41
Temperature (grad C) 0.000							

^{13}C NMR

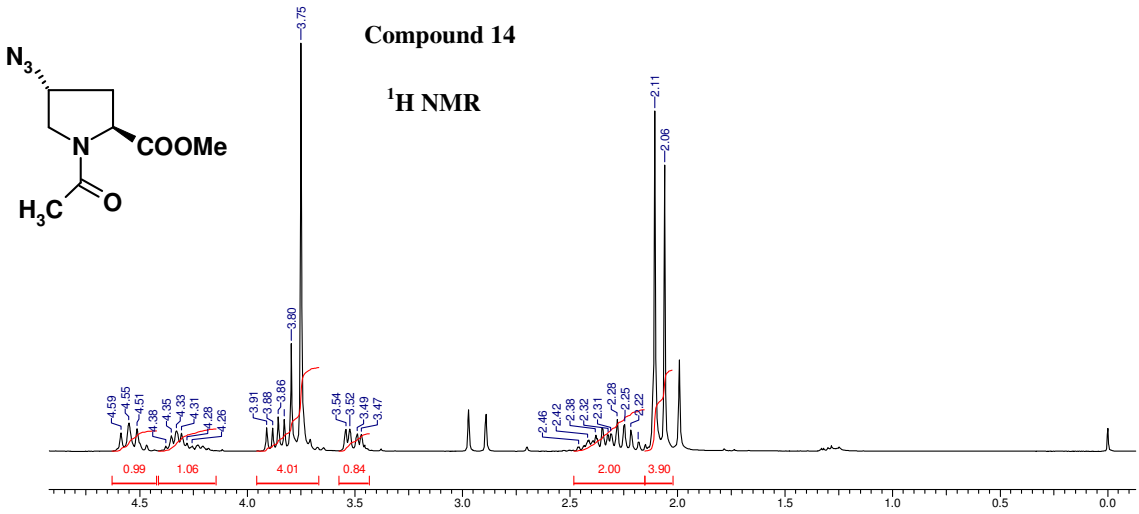


Acquisition Time (sec) 2.7329	*Comment*	M UMASANKARA		*Date*	04/09/2006 21:46:22		
Frequency (MHz) 50.32	*Nucleus*	13C	*Original Points Count* 32768	*Points Count*	32768	*Sweep Width (Hz)*	11990.41
Temperature (grad C) 0.000							

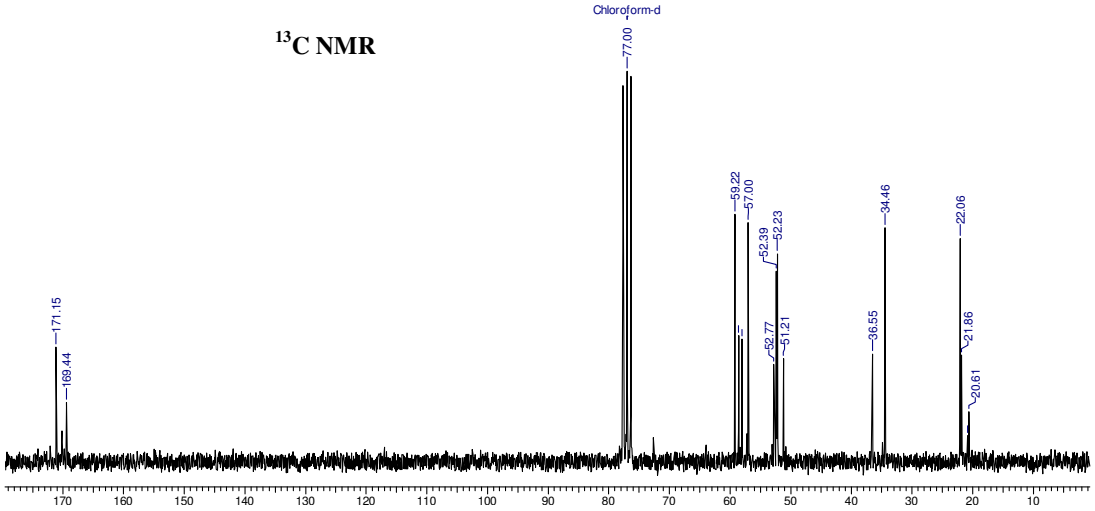
DEPT



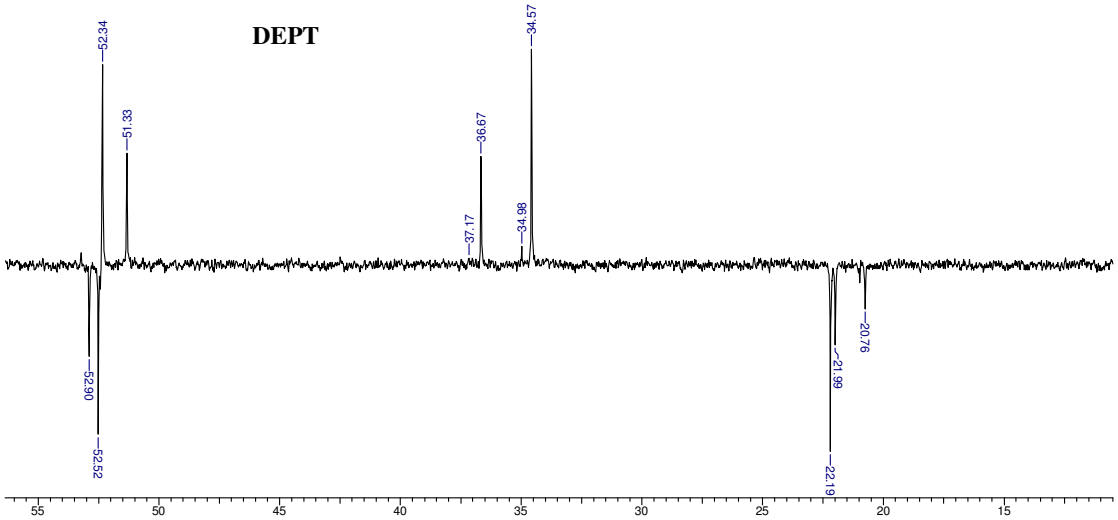
Acquisition Time (sec) 7.9167	***Comment***	Umashankar		***Date***	29/03/2006 22:22:22		
Frequency (MHz) 200.13	***Nucleus***	1H	***Original Points Count*** 32768	***Points Count***	32768	***Sweep Width (Hz)***	4139.07
Temperature (grad C) 0.000							



Acquisition Time (sec) 2.7329	***Comment***	M umashankara		***Date***	20/03/2006 04:31:22		
Frequency (MHz) 50.32	***Nucleus***	13C	***Original Points Count*** 32768	***Points Count***	32768	***Sweep Width (Hz)***	11990.41
Temperature (grad C) 0.000							



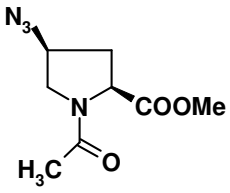
Acquisition Time (sec) 2.7329	***Comment***	M umashankara		***Date***	20/03/2006 04:50:44		
Frequency (MHz) 50.32	***Nucleus***	13C	***Original Points Count*** 32768	***Points Count***	32768	***Sweep Width (Hz)***	11990.41
Temperature (grad C) 0.000							



Acquisition Time (sec)	7.9167	**Comment**	M UMASHANKARA		**Date**	17/03/2006 16:25:52	
Frequency (MHz)	200.13	**Nucleus**	1H	**Original Points Count** 32768	**Points Count**	32768	**Sweep Width (Hz)** 4139.07
Temperature (grad C)	0.000						

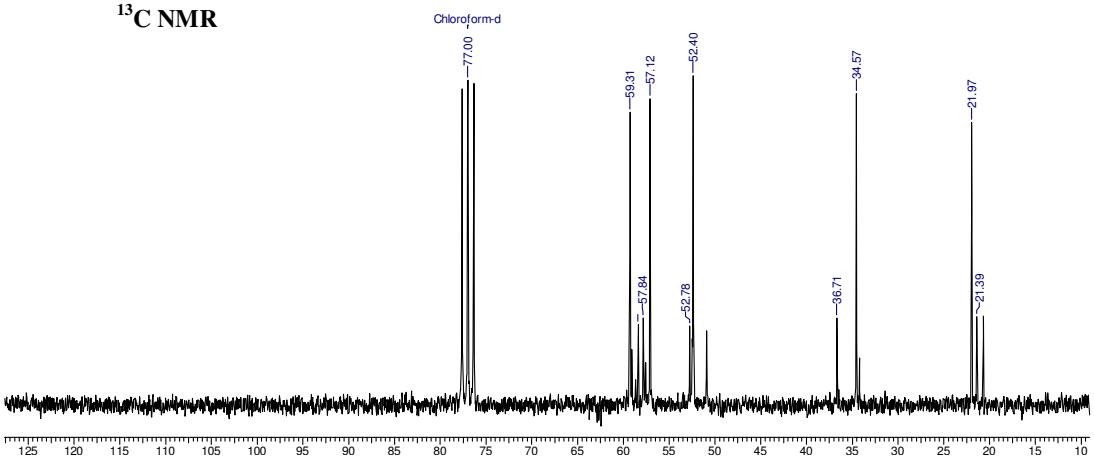
Compound 15

^{1}H NMR



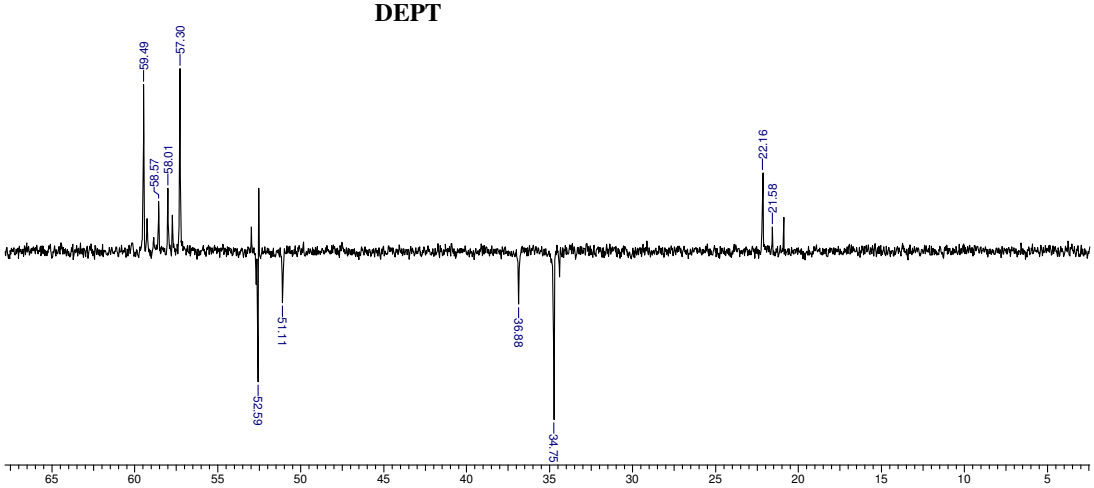
Acquisition Time (sec)	2.7329	**Comment**	M UMASANKARA		**Date**	05/09/2006 00:37:28	
Frequency (MHz)	50.32	**Nucleus**	13C	**Original Points Count** 32768	**Points Count**	32768	**Sweep Width (Hz)** 11990.41
Temperature (grad C)	0.000						

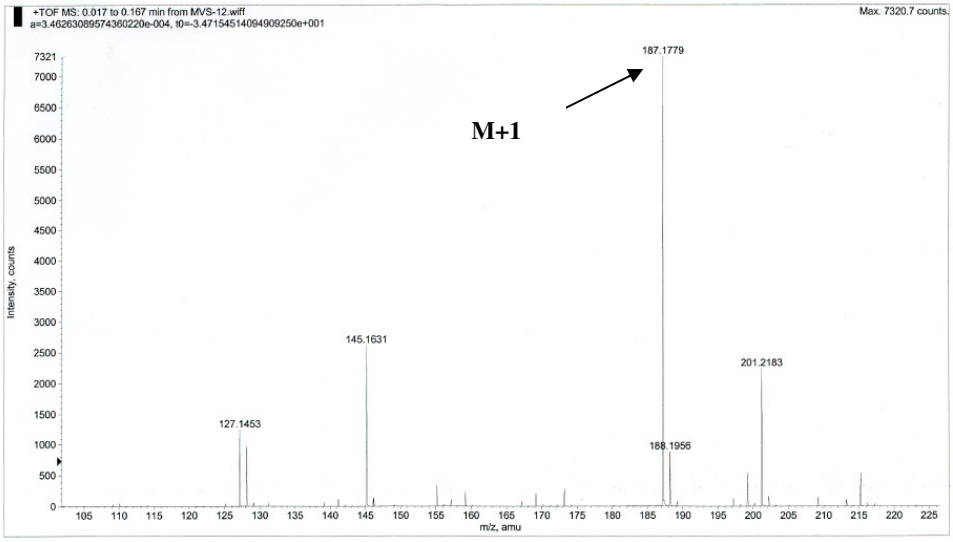
^{13}C NMR



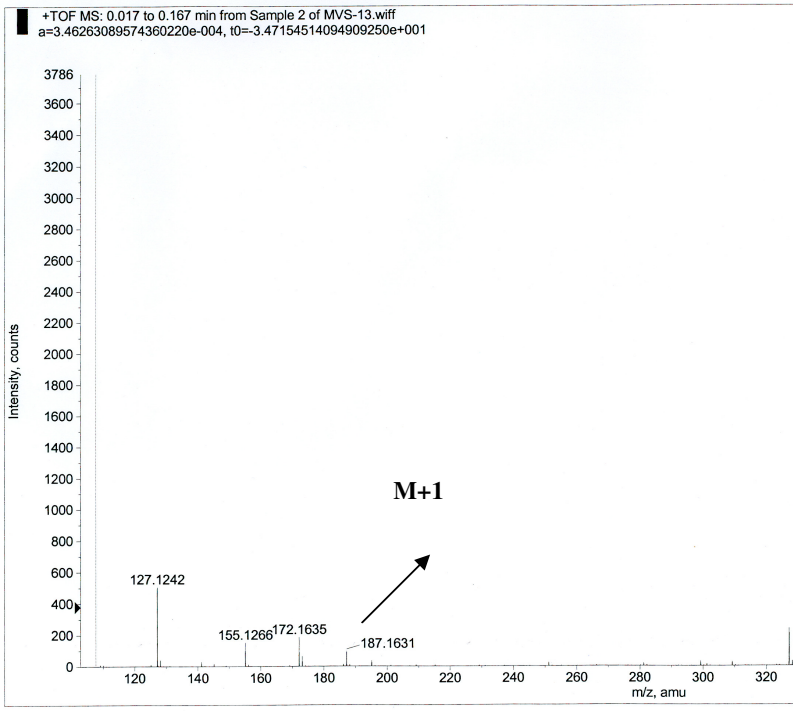
Acquisition Time (sec)	2.7329	**Comment**	M UMASANKARA		**Date**	05/09/2006 00:56:50	
Frequency (MHz)	50.32	**Nucleus**	13C	**Original Points Count** 32768	**Points Count**	32768	**Sweep Width (Hz)** 11990.41
Temperature (grad C)	0.000						

DEPT

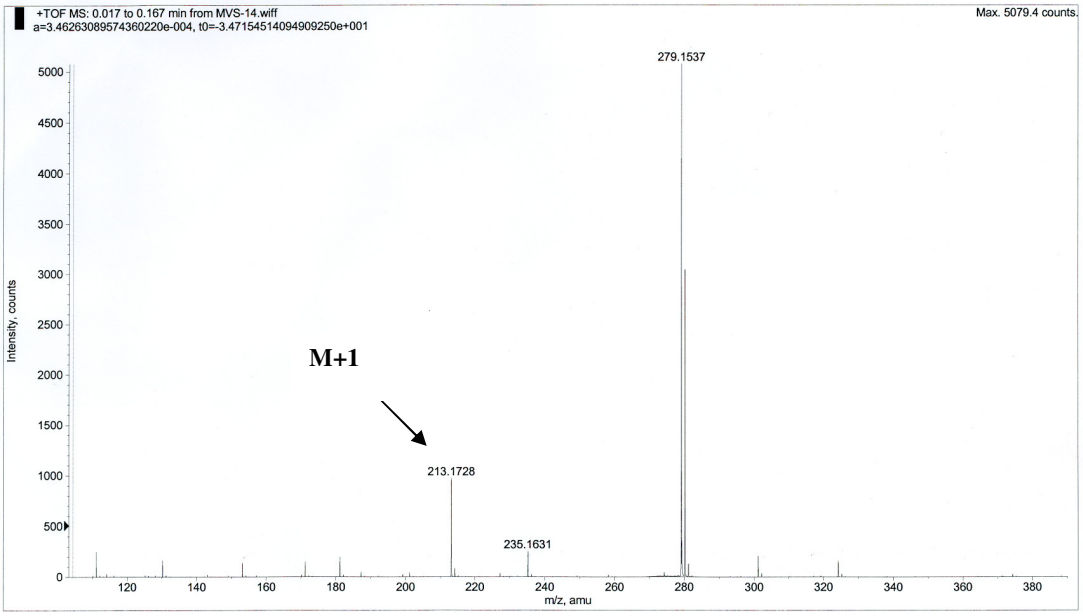




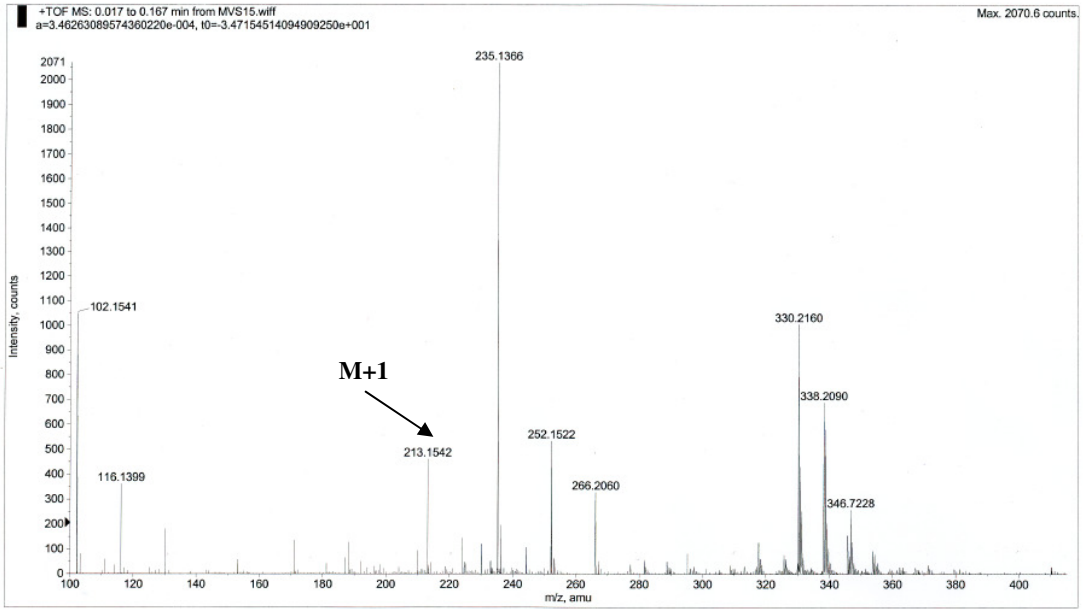
LC-MS mass spectra of compound **12**



LC-MS mass spectra of compound **13**



LC-MS mass spectra of compound **14**



LC-MS mass spectra of compound **15**