SYNTHESES AND CONFORMATION OF SYNTHETIC PEPTIDE SUBSTRATES OF PROTOCOLLAGEN LYSYL HYDROXYLASE

CENTRE FOR NEWFOUNDLAND STUDIES

TOTAL OF 10 PAGES ONLY
MAY BE XEROXED

(Without Author's Permission)

LYNN SODAI ZIJENAH

SYNTHESES AND CONFORMATION OF SYNTHETIC PEPTIDE SUBSTRATES OF PROTOCOLLAGEN LYSYL HYDROXYLASE.

BY

© Lynn Sodai Zijenah, B.Sc. (Hons.)

A thesis submitted to the School of Graduate Studies in partial fulfilment of the requirements for the degree of Master of Science.

Department of Biochemistry
Memorial University of Newfoundland

December, 1986

St. John's Newfoundland

Permission has been granted to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film.

The author (copyright owner) has reserved other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without his/her written permission.

L'autorisation a été accordée à la Bibliothèque nationale du Canada de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur (titulaire du droit d'auteur) se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation écrite.

ISBN 0-315-36996-5

Abstract

Hydroxylation of specific lysyl residues by lysyl hydroxylase is an important posttranslational modification process in collagen biosynthesis. The main objective of this work was to investigate the conformational requirement for the enzymic reaction.

Eight lysine-containing peptides which had amino acid sequences comparable to amino acid sequences around hydroxylysine or lysine in collagen were synthesized by solution-phase techniques. The peptides varied in length from three to seven amino acid residues. The structures of these peptides were investigated through circular dichroism (CD) and infrared (IR) spectroscopic methods.

Lysyl hydroxylase was partially purified from chicken embryos using the established procedures. Seven of the synthetic peptides were tested for their ability to act as substrates of partially purified lysyl hydroxylase. The hydroxylation reaction was assayed by a technique involving measurement of $^{14}CO_2$ released stoichiometrically from 2-[1-^{14}C]oxoglutarate and/or by a specific chemical procedure for hydroxylysine.

Five peptides with the -Lys-Gly- sequence were hydroxylated to varying degrees, the degree of hydroxylation increasing with increasing chain length. Examination of these hydroxylated peptides by CD and IR spectral measurements revealed that the tripeptides N^{α}tBocIleLysGlyOH and N^{α}tBocAlaLysGlyOH adopt a γ-turn in which lysine occupies the second position of this structure. The tetrapeptide (N^{α}tBocAlaLysGlySerOH) adopts both a β- and γ-turn and is more hydroxylated than the precursor tripeptide. This increase in the degree of hydroxylation may be attributed to the presence of the β-turn which may stabilize the γ-turn formed by the AlaLysGlyOH segment. The

hexapeptide (N^{α}tBocLeuHyPGlyAlaLysGlyOH) adopts a consecutive β- and γ-turn and is more hydroxylated than the tetrapeptide. This increase in hydroxylation may be attributable to the Gly^3-Ala^4 segment which may increase the binding of the enzyme to the substrate thereby enhancing hydroxylation. The heptapeptide (N^{α}tBocLeuHyPGlyAlaLysGlySerOH) is hydroxylated more than the precursor hexapeptide. CD and model building studies have shown that N^{α}tBocLeuHyPGlyAlaLysGlySerOH adopts two consecutive β-turns and a γ-turn. The second β-turn which is similar to that found in the tetrapeptide (N^{α}tBocAlaLysGlySerOH) may be responsible for the increase in hydroxylation in comparison with the hexapeptide.

All the hydroxylated peptides have one structural feature in common, namely the γ-turn with lysine in the second position. In contrast, two peptides (N^{α}tBocAlaGlyLysOH and N^{α}tBocAlaGlyLysHyPOH) which have the Gly-Lys sequence were not hydroxylated. Interestingly, both peptides adopt a γ-turn but the lysine is found in the third position of this structure. These data indicate that lysyl hydroxylase recognizes specific secondary structure(s) in its substrates. The nature of the amino acid around lysine and the chain length of the peptide may be the critical determinants in the synthesis of hydroxylysine by lysyl hydroxylase.

Acknowledgements

I would like to extend my appreciation to Dr. C. Loader of the Department of Chemistry and Dr. J.C. Orr for their excellent instruction in using the UNIX system for the preparation of this thesis.

Many thanks to Dr. S.K. Attah-Poku for his initial instruction in peptide synthesis.

Special thanks to my supervisor Dr. V.S. Ananthanarayanan for his guidance and financial assistance; to the members of the supervisory committee Dr. J.C. Orr and Dr. K. Keough; and to my colleagues: Mrs. P. Atreya and Mr. L. Taylor for their encouragement, assistance and friendship during the persuit of my graduate programme particularly when my supervisor was away on a one year sabbatical leave. I would also like to thank Dr. E. Barnsley for his suggestions and advice in the initial preparation of this thesis.

My thanks to Mr. D. Hall and his assistants for the amino acid analyses.

I would also like to thank Memorial University for giving me an opportunity to persue graduate studies and also for the financial support in the form of a bursary.

Last but not least, I would like to say thanks to my fiance, K.T.J. Walmsley for his moral support and motivation throughout my studies at this University.

Table of Contents

1. Introduction 1

1.1. Posttranslational modification of amino acids in proteins 1

1.2. Collagen 2

1.2.1. General structural features of the collagens 2

1.2.2. Types of collagen 3

1.3. Biosynthesis of collagen 4

1.3.1. Posttranslational modifications of peptide chains in collagen 6

1.3.2. Hydroxylation of prolyl residues 7

1.3.3. Hydroxylation of lysyl residues 9

1.3.4. Cross-linking in collagen 15

1.4. Importance of the hydroxylation reactions in collagen 22

1.5. Objectives and scope of present work. 24

2. Experimental 28

2.1 Materials 28

2.2. Methods 29

2.2.1. Syntheses of peptides 29

2.2.2. Analytical Reversed-Phase High Performance Liquid Chromatography 29

2.2.3. Preparative Reversed-Phase High Performance Liquid Chromatography 30

2.2.4. Silica gel Chromatography 30

2.2.5. Elemental Analysis 31

2.2.6. Amino Acid Analysis 31

2.2.7. Melting Points 32

2.2.8. Circular Dichroism 32

2.2.9. Infrared Spectroscopy 32

2.2.10. Partial Purification of Lysyl Hydroxylase. 33

2.2.11. Enzyme Assay 35

2.2.12. Other Assays 37

3. Synthesis 38

3.1. Synthesis of N^{α}tBocLeuHyPGlyAlaLysGlySerOH 38

3.1.1. Synthesis of N^{α}tBocLeuHyPGlyONSU 38

3.1.1.2. N^{α}tBocLeuHyPOMe 38

3.1.1.2. N^{α}tBocLeuHyPOH 42

3.1.1.3. N^{α}tBocLeuHyPONSU 42

3.1.1.4. N^{α}tBocLeuHyPGlyOH 42

3.1.1.5. N^{α}tBocLeuHyPGlyONSU 43

3.1.2. Synthesis of TFA.Ala(N^{ϵ}cbz)LysGlySerOH 43

3.1.2.1. N^{α}tBocAlaONSU 43

3.1.2.2. N^{α}tBocAla(N^{ϵ}cbz)LysOH 44

3.1.2.3. N^{α}tBocAla(N^{ϵ}cbz)LysGlyOMe 44

3.1.2.4. N^{α}tBocAla(N^{ϵ}cbz)LysGlyOH 45

3.1.2.5. N^{α}tBocAlaLysGlyOH 46

3.1.2.6. N^{α}tBocAla(Ncbz)LysGlySerOMe 46

3.1.2.7. N^{α}tBocAla(N^{ϵ}cbz)LysGlySerOH 47

3.1.2.8. N^{α}tBocAlaLysGlySerOH 47

3.1.2.9. TFA.Ala(N^{ϵ}cbz)LysGlySerOH 49

3.1.3. N^{α}tBocLeuHyPGlyAla($^{\epsilon}$cbz)LysGlySerOH 49

3.1.4. $N^{\alpha t}$tBocLeuHyPGlyAlaLysGlySerOH 49

3.2. Synthesis of N^{α}tBocAlaArgGlyIleLysGlyOH 50

3.2.1. Synthesis of N^{α}tBocAla(N^{ω}nitro)ArgGlyONSU 51

3.2.1.1. N^{α}tBocAla(N^{ω}nitro)ArgOMe 51

3.2.1.2. N^{α}tBocAla(N^{ω}nitro)ArgOH 51

3.2.1.3. N^{α}tBocAla(N^{ω}nitro)ArgGlyOMe 51

3.2.1.4. N^{α}tBocAla(N^{ω}nitro)ArgOH 51

3.2.1.5. N^{α}tBocAla(N^{ω}nitro)ArgGlyONSU 54

3.2.2. Synthesis of N^{α}tBocIle(N^{ϵ}cbz)LysGlyOMe 54

3.2.2.1. N^{α}tBocIleONSU 54

3.2.2.2. N^{α}tBocIle(N^{α}cbz)LysOH 54

3.2.2.3. NB$^{\alpha}$tBocIle(N^{ϵ}cbz)LysGlyOMe 54

3.2.2.4. N^{α}tBocIleLyusGlyOH (34) 55

3.2.2.5. TFA.Ile(N^{ϵ}cbz)LysGlyOMe 55

3.2.2.5. N^{α}tBocAla(N^{ω}nitro)ArgGlyIle(N^{ϵ}cbz)LysGlyOMe 55

3.2.4. N^{α}tBocAla(N^{ω}nitro)ArgGlyIle(N^{ϵ}cbz)LysGlyOH 56

3.2.2.7. N^{α}AlaArgGlyIleLysGlyOH (32) 56

3.3. Synthesis of N^{α}tBocAlaGlyLysOH·N^{α}tBocAlaGlyLysHyPOH 56

3.3.1. N^{α}tBocAlGlyLysOH 57

3.3.2 N^{α}tBocAlaGly(N^{ϵ}cbz)LysOMe 57

3.3.3. N^{α}tBocAlaGly(N^{ϵ}cbz)LysOH 57

3.3.4. N^{α}tBocAlaGlyLysOH 57

3.3.5. Synthesis of N^{α}tBocAlaGly(N^{ϵ}cbz)LysHyPOMe 60

3.3.6. N^{α}tBocAlaGly(N^{ϵ})LysHyPOH 60

3.3.7. NtBocAlaGlyLysHyPOH 60

3.4. Synthesis of N^{α}tBocLeuHyPGlyAlaLysGlyOH 61

3.4.1. N^{α}tBocLeuHyPGlyAlaOH 62

3.4.2. N^{α}tBocLeuHyPGlyAla(N^{ϵ}cbz)LysOMe 62

3.4.3. N^{α}tBocLeuHyPGlyAla(N^{ϵ}cbz)LysOH 62

3.4.4. N^{α}tBocLeuHyPGlyAla(N^{ϵ}cbz)LysGlyOMe 62

3.4.5. N^{α}tBocLeuHyPAla(N^{ϵ}cbz)LysGlyOH 62

3.4.6. N^{α}tBocLeuHyPGlyAlaLysGlyOH 64

4. Spectroscopic analysis of peptide structure 66

4.1. Infrared spectroscopy 66

4.1.1. Introduction to Infrared Spectroscopy 66

4.1.2. Tripeptides: N^{α}tBocAla(cbz)LysGlyOMe and N^{α}tBocIle(cbz)LysGlyOMe. 68

4.1.3. IR of Tetrapeptides: N^{α}tBocAla(cbz)LysGlySerOMe and N^{α}tBocAlaGly(cbz)LysHyPOMe. 76

4.1.4. IR of Hexapeptides: N^{α}tBocLeuHyPGlyAla(cbz)LysGlyOMe and N^{α}tBocAla(nitro)ArgGlyIle(cbz)LysGlyOMe. 82

4.2. Circular Dichroism Characterization 88

4.2.1. Introduction to Circular Dichroism. 88
4.2.2. CD spectra of the tripeptides: (BocAlaLysGlyOH, BocIleLys-GlyOH and BocAlaGlyLysOH) 96

4.2.3. CD spectra of Tetrapeptides: (BocAlaLysGlySerOH and BocAlaGlyLysHyPOH) 99

4.2.4. CD spectra of Hexapeptides: BocLeuHyPGlyAlaLysGlyOH and BocAlaArgGlyIleLysGlyOH. 103

4.2.5. CD spectra of Heptapeptide: BocLeuHyPGlyAlaLysGlySerOH. 108

5. Hydroxylation Data and Discussion 115

5.1. Introduction 115

6. Discussion and Conclusion 125

REFERENCES 129

List of Tables

Table 1-1: Structurally and Genetically Distinct Collagens 6

Table 2-1: HPLC parameters 30

Table 2-2: Partial Purification of lysyl hydroxylase from a chick-embryo $(NH_4)_2SO_4$ 34

Table 4-1: Infrared Data (ν_{NH} and ν_{CO}) for Synthetic peptides 69

Table 4-2: Maxima and minima of β-turn CD Spectral Classes as Defined by Woody (1974) 89

Table 5-1: (a). Amino acid sequences around HyL in collagen 117

Table 5-1: (b) Analysis of the frequency of occurrence of a.a.'s in the 117

Table 5-1: (c) Synthetic Peptides 119

Table 5-2: Examination of synthetic peptides as substrates for partially purified lysyl hydroxylase. 121

Table 6-1: Summary of hydroxylation data versus conformation 126

List of Figures

FIG. 1-1: The structure of collagen. 5

FIG. 1-2: Schematic presentation of the steps involved in the intracellular biosynthesis of procollagen and the assembly of collagen molecules into extracellular fibers. 8

FIG. 1-3: The structures of 4-*trans*-hydroxy-L-proline and 5-hydroxylysine and the equation describing the hydroxylation reactions. 10

FIG. 1-4: The structure of lysine, 4-HyL and allysine. 16

FIG. 1-5: The four principal cross-linking loci in molecules of types I, II and III collagens. 17

FIG. 1-6: Scheme outlining one of the two routes of cross-linking in collagen: Allysine-based cross-links. 19

FIG. 1-7: Scheme outlining one of two routes of cross-linking in collagen: Hydroxyallysine-based cross-links. 21

FIG. 1-8: Schematic representation of the proposed model substrate of prolyl hydroxylase. The arrow indicates the position of hydroxylation. 26

FIG. 2-1: Probable reactions for the oxidation of HyL by periodate. 37

FIG. 3-1: Flow chart summarizing the synthesis of N^{α}tBocAlaLysGlyOH, N^{α}tBocAlaLysGlySerOH and N^{α}tBocLeuHyPGlyAlaLysGlySerOH 40

FIG. 3-2: Analytical HPLC of (a) BocLeuHyPOMe and (b) BocLeuHyP-

GlyOH 41

FIG. 3-3: Analytical HPLC of (a) BocAla(N^{ϵ}cbz)LysGlyOH and (b) BocAlaLysGlyOH 45

FIG. 3-4: Analytical HPLC of (a) BocAla(N^{ϵ}cbz)LysGlySerOH and (b) BocAlaLysGlySerOH 48

FIG. 3-5: Analytical HPLC of BocLeuHyPGlyAlaLysGlySerOH 50

FIG. 3-6: Flow chart summarizing synthesis of N^{α}tBocIleLysGlyOH and N^{α}tBocAlaArgGlyIleLysGlyOH 52

FIG. 3-7: Analytical HPLC of (a)BocAla(nitro)ArgOMe and (b) BocAla(nitro)ArgGlyOMe 53

FIG. 3-8: Analytical HPLC of (a)BocIleLysGlyOH and (b) BocAlaArg-GlyIleLysGlyOH 58

FIG. 3-9: Flow chart summarizing the synthesis of N^{α}tBocAlaGlyLysOH and N^{α}tBocAlaGlyLysHyPOH 59

FIG. 3-10: Analytical HPLC of (a) BocAlaGlyLysOH and (b) BocAlaGlyLysHyPOH 61

FIG. 3-11: Flow chart summarizing synthesis of N^{α}tBocLeuHypGlyAlaLysGlyOH 63

FIG. 3-12: Analytical HPLC of BocLeuHyPGlyAlaLysGlyOH 65

FIG. 4-1: ν_{NH} bands of tripeptides in chloroform. (a) N^{α}tBocAla(cbz)LysGlyOMe and (b) N^{α}tBocIle(cbz)LysGlyOMe. 70

FIG. 4-2: Schematic representation of intramolecular hydrogen-bonded structures. (a) C_5; (b) C_7 or γ-turn and (c) C_{10} or β-turn 72

FIG. 4-3: ν_{CO} band of N^{α}tBocAla(cbz)LysGlyOMe in $CHCl_3$. 73

FIG. 4-4: ν_{CO} band of N^{α}tBocIle(cbz)LysGlyOMe in $CHCl_3$. 74

FIG. 4-5: Schematic Diagram of BocAla(cbz)LysGlyOMe, BocIle(cbz)LysGlyOMe 75

FIG. 4-9: Schematic Diagram of BocAlaLysGlySerOH 77

FIG. 4-6: ν_{NH} bands of tetrapeptides in $CHCl_3$. (a) N^{α}tBocAla(cbz)LysGlySerOMe and (b) N^{α}tBocAlaGly(cbz)LysHyPOMe. 80

FIG. 4-7: ν_{CO} of N^{α}tBocAla(cbz)LysGlySerOMe in $CHCl_3$. 79

FIG. 4-8: ν_{CO} band of N^{α}tBocAlaGly(cbz)LysHyPOMe. 81

FIG. 4-9: Schematic Diagram of BocAlaLysGlySerOH 77

FIG. 4-10: ν_{NH} band of N^{α}tBocLeuHyPGlyAla(cbz)LysGlyOMe and N^{α}tBocAla(nitro)ArgGlyIle(cbz)LysGlyOMe. 84

FIG. 4-11: ν_{CO} band of N^{α}tBocLeuHyPGlyAla(cbz)LysGlyOMe. 85

FIG. 4-12: ν_{CO} band of N^{α}tBocAla(nitro)ArgGlyIle(cbz)LysGlyOMe. 86

FIG. 4-13: Schematic Diagram of BocLeuHypGlyAlaLysGlyOH 94

FIG. 4-14: CD spectra of BocAlaLysGlyOH in MeOH, TFE, and H_2 or NH_4HCO_3. 95

FIG. 4-15: CD spectra of BocIleLysGlyOH in MeOH, TFE, H_2O and NH_4HCO_3. 96

FIG. 4-16: CD spectra of BocAlaGlyLysOH in MeOH, TFE, H_2O and NH_4HCO_3. 98

FIG. 4-17: CD spectra of BocAlaLysGlySerOH in MeOH, TFE, H_2O and NH_4HCO_3. 101

FIG. 4-18: CD spectra of BocAlaGlyLysHyPOH in MeOH, TFE, H_2O and NH_4HCO_3. 102

FIG. 4-19: CD spectra of BocLeuHyPGlyAlaLysGlyOH in MeOH, TFE, H_2O and NH_4HCO_3 104

FIG. 4-20: CD spectra of BocAlaArgGlyIleLysGlyOH in MeOH, TFE, and H_2O 107

FIG. 4-21: CD spectra of BocLeuHyPGlyAlaLysGlySerOH in MeOH, TFE, and H_2O. 110

FIG. 4-22: Schematic Diagram of BocLeuHyPGlyAlaLysGlySerOH 111

FIG. 4-23: CPK model of BocAlaLysGlyOH 112

FIG. 4-24: CPK model of BocAlaLysGlySerOH 113

FIG. 4-25: CPK model of BocLeuHyPGlyAlaLysGlySerOH 114

Abbreviations

Acn	acetonitrile
Aib	α-aminobutyric acid
Ala	alanine
Arg	arginine
Asn	asparagine
Asp	aspartic acid
Cbz	Carbobenzoxy
$CHCl_3$	chloroform
CD	circular dichroism
D-Ala	D-alanine
DCC	dicyclohexylcarbodiimide
DCU	dicyclohexylurea
D-Phe	D-phenylalanine
EDTA	ethylenediaminetetraacetic acid
Gly	glycine
Glu	glutamic acid
Gln	glutamine
His	histidine
Ile	isoleucine
HPLC	High Perfomance Liquid Chromatography
HyL	hydroxylysine
HyP	hydroxyproline
IR	infrared
Leu	leucine

Lys	lysine
MeOH	methanol
Met	methionine
Nle	norleucine
NSU	N-hydroxy succinimide
OSt	O-stearyl
O-Bzl	O-benzyl
Pro	proline
Ser	serine
*t*Boc	*tertiary*-butyloxycarbonyl
TFE	trifluoroethanol
THF	tetrahydrofuran
Tyr	tyrosine
Val	valine
UV	Ultraviolet

Chapter 1

1. Introduction

1.1. Posttranslational modification of amino acids in proteins.

The posttranslational modification of proteins is a well established phenomenon (Paik and Kim 1975). The modification often results in a significant alteration of the protein structure with a consequent change in its function. It can thus be regarded as one of nature's ways of converting one protein to another without *de novo* synthesis of the former polypeptide chain.

The covalent modification reactions may be subdivided into two distinct types. One of them involves peptide bonds as in zymogen activation, hormone activation, blood coagulation, immunoglobulin synthesis, complement fixation, collagen synthesis, assembly of virus particles, conversion of a precursor pancreatic protein to the secretory proteins, the removal of formylmethionine after the completion of the polypeptide *etc.* The second group of covalent modifications involves amino acid side chain residues and these include acetylation, disulfide bridge formation, hydroxylation, glycosylation, iodination, methylation, nucleotidylation, phosphorylation, carboxylation and thiolation.

In this thesis, we will be concerned with an important posttranslational covalent modification: the hydroxylation of lysyl residues, during the biosynthesis of collagen. Our aim will be to assess the conformational basis for the substrate specificity of the enzyme lysyl hydroxylase. A brief review of the general features of collagen and its posttranslational modifications with particular reference to lysine hydroxylation is given below.

1.2. Collagen.

The collagens are the major structural proteins of bone, cartilage, skin, tendon and tooth. Several major monographs have been published describing in detail the physical, chemical and biological properties of collagen (Ramachandran and Reddi, 1976; Walton and Blackwell, 1973; Fraser and MacRae, 1973; Ramachandran, 1967). In addition, several detailed review articles on the structure and function of collagen have appeared in the literature (Kuhn, 1969; Traub and Piez, 1971; Gallop *et al.*, 1972; Yannas, 1972; Piez and Miller, 1974; Bornstein and Traub, 1979).

1.2.1. General structural features of the collagens.

The collagen molecule owes its distinct structural and mechanical properties to its unusual amino acid composition, sequence and conformation. The basic collagen molecule common to all types of collagen (see 1.2.2. below) is composed of three polypeptide chains called the α-chains. Each of the α-chains is coiled in a left-handed helix which differs from the α-helix found in most other proteins in that it is more extended, the axial distance from one amino acid to the next being about 2.9 A instead of 1.5 A in an α-helix. The three α-chains are, in turn, coiled on each other with a right-handed twist, much like the strands of a rope, to form a triple helical structure (Ramachandran and Kartha, 1955; Rich and Crick, 1955). This unusual helical conformation gives the collagen molecule a rigid, rod-like shape.

The unique nature of the collagen triple helix is largely based on the unusual amino acid composition of the polypeptide chains. Glycine accounts for one-third of the total (about 1000) amino acids and it is evenly distributed throughout the molecule at every third residue (Gross, 1974; Fietzek *et al.*, 1972). Consequently, the polypeptide chains of collagen can considered to be made up of the repeating

triplet represented as (Gly-X-Y). The X and Y positions of this repeating triplet can be occupied by a variety of amino acids but most frequently the X position is occupied by proline and the Y position is occupied by 4-hydroxyproline (HyP) with these two imino acids forming about 22% of the total amino acid composition of the major types of collagen. Another characteristic of collagen is the presence of 4-hydroxylysine (HyL) in position Y of the triplet sequence Gly-X-Y; however, unlike hydroxyproline, this unusual amino acid constitutes on an average only about 0.10% of the total amino acid composition of collagen. The relatively high content of the imino acids and the characteristic distribution of glycine are necessary for the triple-helical conformation of the collagen molecule. This conformation is stabilized by interchain hydrogen bonds between the peptide carbonyl and amide groups. The exact details of the hydrogen bonding are still controversial (Ramachandran and Ramakrishina 1976; Bornstein and Traub, 1979). A schematic representation of the collagen helix is shown in Figure 1-1.

1.2.2. Types of collagen.

To date, 11 types of collagen have been recognized in vertebrate tissues and they have been characterized to varying degrees (Kresna and Miller, 1979). Recently Miller (1985) classified the collagens into three separate groups, based on their molecular weights and distribution. Table 1-1 shows the four major types of collagen.

Best known are the three interstitial collagen types I, II and III. Type I, the main constituent of skin, tendon, bone and vessel walls, synthesized by fibroblasts, smooth-muscles cells and osteoblasts consists of two identical $\alpha 1$ and $\alpha 2$ chains (i.e. $\alpha 1(I)_2 \alpha 2$, where the Roman numeral refers to the type of collagen). Type II collagen, the only collagen constituent of hyaline cartilage, is produced by the chondrocytes and consists of three identical, genetically distinct α-chains,

$\alpha 1(II)_3$. Skin, gastrointestinal and vascular connective tissue contain an additional genetically distinct type of collagen, type III. Type III collagen is composed of three identical α-chains, $\alpha 1(III)_3$ which display distinct chemical features such as relatively high contents of hydroxyproline and glycine and the presence of cystine.

Type V collagen is synthesized by foetal membranes (Bergeson *et al.*, 1976), smooth muscles, myoblasts and presumably also by fibroblasts (Kuhn and Glanville, 1980) and has a chain composition of $[\alpha B]_2\alpha A$. Basement membrane collagen type IV is not as well characterized as the other four types.

1.3. Biosynthesis of collagen.

It is now known that collagen is initially synthesized as a higher molecular weight precursor. This precursor molecule, procollagen, is soluble under physiological conditions in which collagen molecules form fibers. The procollagen polypeptides are synthesized on the membrane-bound ribosomes and the newly-synthesized chains are fed into the cisternae of the rough endoplasmic reticulum. The initially-synthesized polypeptides of procollagen called pro-α chains are larger than the α-chains and contain additional amino acids both at the amino- and carboxy terminal ends of the polypeptide chains. These extension peptides are rich in acidic amino acids, relatively poor in glycine, proline and HyP and they contain cysteine and tryptophan which are not present in type I or type II collagen (Martin *et al.*, 1975; Scholfield and Prockop, 1973). The major peptide extension in procollagen is located at the carboxyterminal of the polypeptide chain (Tanzer *et al.*, 1974; Park *et al.*, 1975) and contains interchain disulfide bonds which link three pro-α chains together (Byers *et al.*, 1973). The carboxy terminal peptide extension from the pro-α chains of type I or type II procollagen consists of approximately 200 to 300 amino acids while the amino terminal

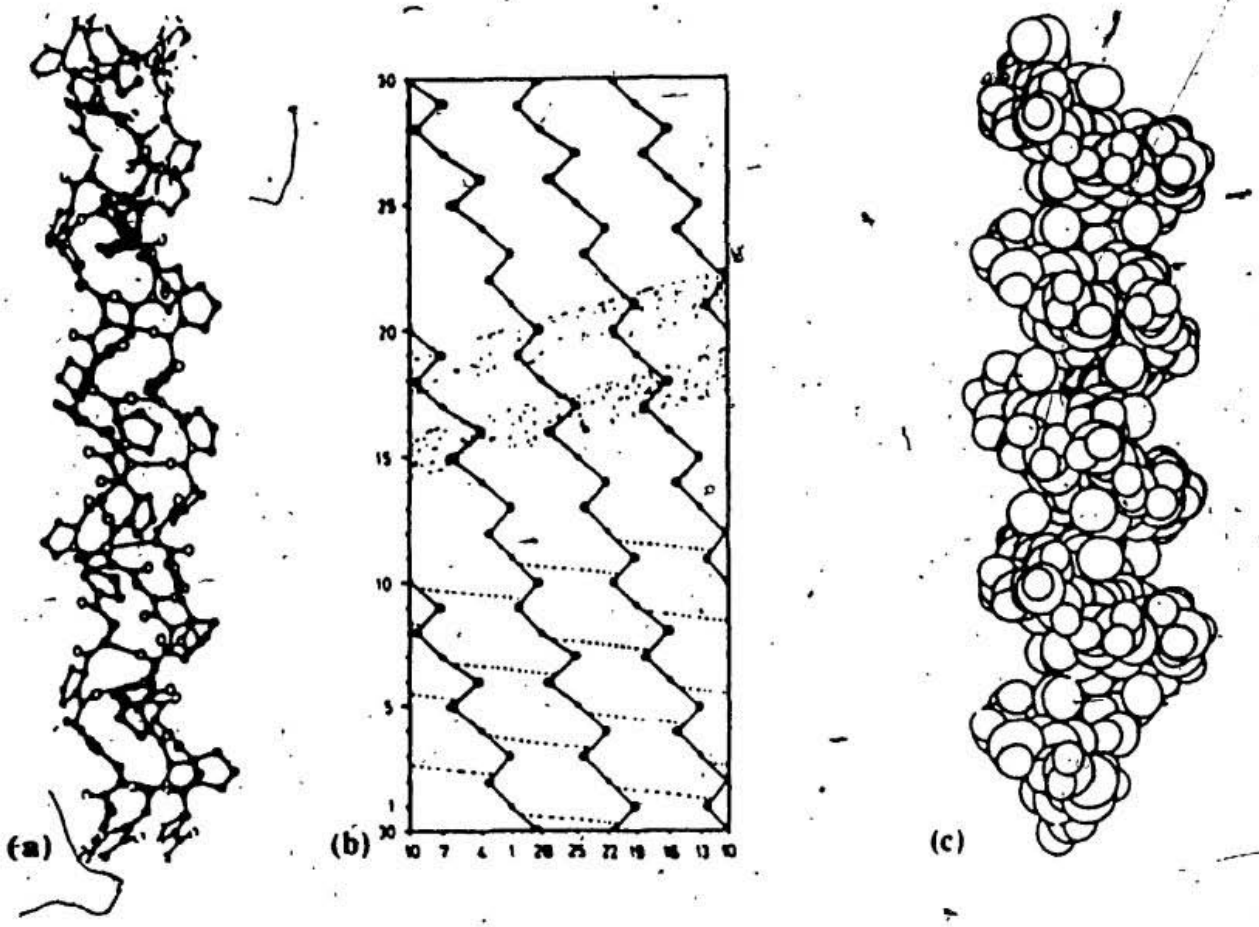

FIG. 1-1: The structure of collagen. (a) ORTEP-drawing of a collagen triple helix with the sequence $(Gly\text{-}Pro\text{-}Pro)_n$ and coordinates. Hydrogens are only inserted if taking part in bonds (thin lines, almost horizontal). (b) Cylindrical plot of a full repeat of the collagen triple helix. The chain course and the C_α-atoms are indicated. The three C_α-atoms in the sequence $(Gly\text{-}X\text{-}Y)_n$ are marked by (. . •). Hydrogen bonds are given as dashed lines. (c) ORTEP-drawing of a collagen triple helix corresponding to (a), except that atoms are given with their van der Waals radii and that all hydrogen atoms are included. The helical bulge of pro residues with a pitch of 9 A becomes visible. This bulge is also indicated by dotted stripes in (b).

Reproduced from from Schultz and Schimmer, 1979.

Table 1-1:

Structurally and Genetically Distinct Collagens

Type	Distribution	Chain composition	Distinctive features
I	Skin, bone, tendon, ligament, fascia, dentin, blood vessels, interstitial connective tissues	$[\alpha 1(I)]_2\alpha 2$	Hybrid of two chain types; low (15%) hydroxylation of lysine; low carbohydrate
II	Cartilage, nucleus pulposus	$[\alpha 1(II)]_3$	Intermediate (50%) hydroxylation of lysine; all hydroxylysines glycosylated
III	Same as type I, except bone and tendon; prominent in blood vessels, gastrointestinal tract, and fetal skin	$[\alpha 1(III)]_3$	Contains cysteine; 4-Hyp > Pro; Gly > ⅓ residues; low (15%) hydroxylation of lysine
IV	Basement membranes	$[\alpha 1(IV)]_3$[a]	High 3-Hyp (1%); contains cysteine; most lysines hydroxylated; all hydroxylysines glycosylated; low alanine; carbohydrate content not limited to glucose and galactose

Reproduced from Bornstein and Traub, 1979.

about 30 to 40% larger than the α-chain in its fully extended form.

1.3.1. Posttranslational modifications of peptide chains in collagen.

The biosynthesis of collagen is characterized by several posttranslational modifications which include (a) hydroxylation of selected prolyl residues (b) hydroxylation of selected lysyl residues; (c) attachment of sugars galactose and glucose into certain hydroxylysyl residues; (d) chain association, disulfide bond formation and triple helix formation; (e) proteolytic conversion of procollagen to collagen; (f) oxidative deamination of certain lysyl and hydroxylysyl residues and (g) intermolecular cross-link formation leading to fibrils and fibres.

1.3.2. Hydroxylation of prolyl residues.

One of the unique features of collagen is the presence of the unusual imino acid HyP. The biological role of HyP in collagen was unknown for a long time but it is now well established that a critical amount of HyP is required to stabilize the triple helical conformation at physiological temperatures (Berg and Prockop, 1973; Jiminez *et al.*, 1973; Rosenbloom *et al.*, 1973; Sakakibara *et al.*, 1973; Ward and Mason 1973; Fessler and Fessler, 1974; Uitto and Prockop, 1974). The triple helix in turn is required for the secretion of procollagen molecules out of the cells at an optimal rate, and the rigid triple-helical structure is a necessary requirement for the extracellular deposition of functionally adequate collagen fibers (Prockop *et al.*, 1975).

The enzyme prolyl hydroxylase (prolyl-glycyl-peptide, 2-oxoglutarate:oxygen 5 oxidoreductase E.C.1.14.11.2) which catalyses the hydroxylation of certain prolyl residues was first obtained in a relatively pure form using ion exchange and gel filtration chromatography, from chick embryos (Halmes *et al.*, 1970) and rat skin (Rhoads and Udenfriend, 1970). Subsequently, a rapid affinity chromatography procedure was developed using Ascaris collagen linked to agarose and eluting the enzyme from the column with a second polypeptide substrate having the structure $(Pro\text{-}Pro\text{-}Gly)_n$ (Berg and Prockop, 1973). Tuderman *et al.*, (1975) further modified the affinity column procedure and used polyproline linked to agarose for purifying prolyl hydroxylase from chick embryos and from human foetal tissues (Kuutti *et al.*, 1975). The affinity procedure was recently modified to include a DEAE ion-exchange chromatography step which enables the effective removal of the bound polyproline from the enzyme (Kerdesha and Berg, 1981). The chick embryo enzyme was judged by electron microscopy to be a tetramer of two 61,000 and two 64,000 molecular weight subunits forming two interlocking

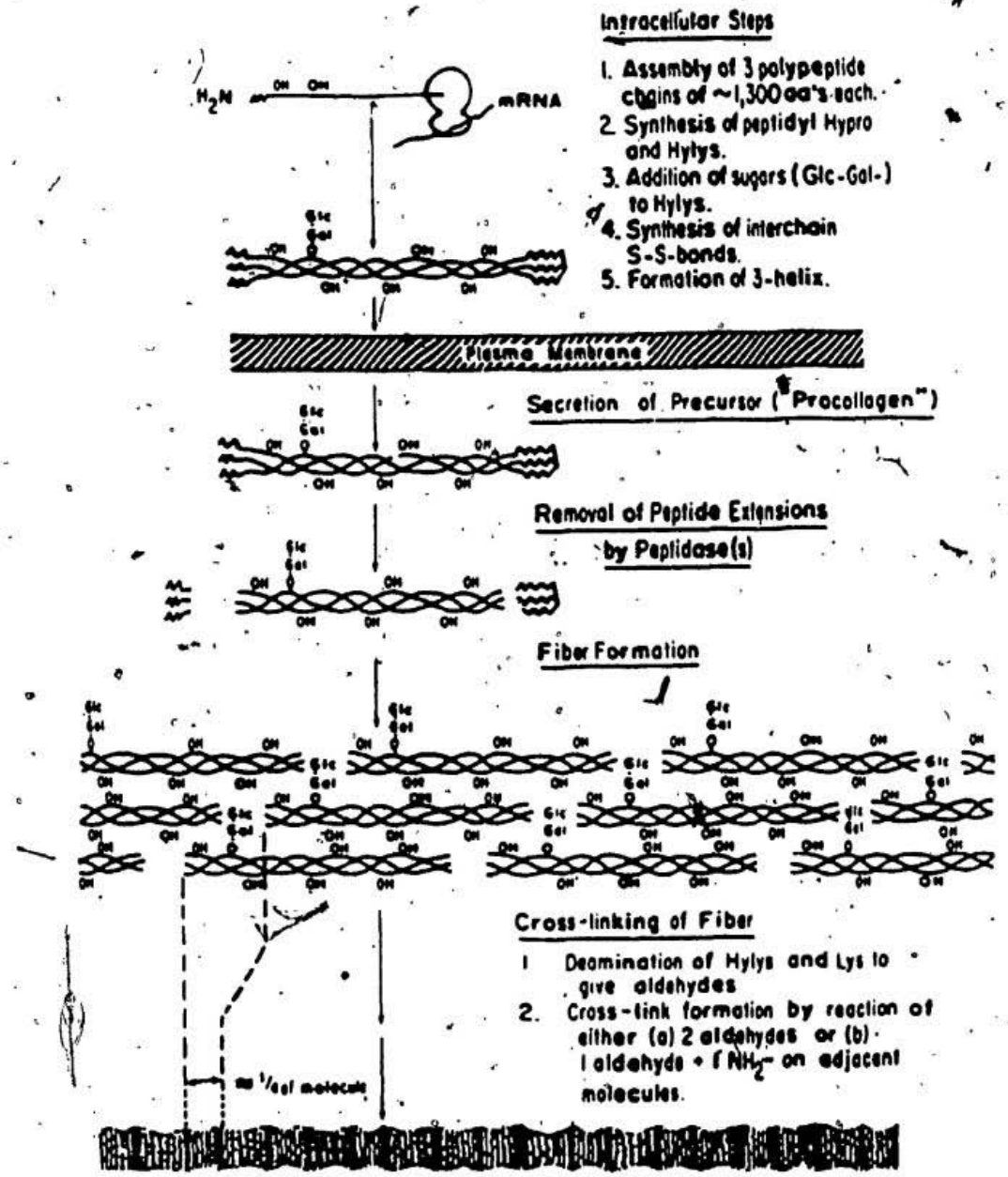

FIG. 1-2: Schematic presentation of the steps involved in the intracellular biosynthesis of procollagen and the assembly of collagen molecules into extracellular fibers. Abbreviations: mRNA, messenger RNA; aa's, amino acids; Hypro, hydroxyproline; Hylys, hydroxylysine; Glc-Gal-, glucosylgalactose attached to the hydroxyl group of hydroxylysine; Lys, lysine; NH_2-, amino-terminal ends of polypeptide chains, and also the ϵ-amino group on either hydroxylysine or lysine.
Reproduced from Uitto and Lichtenstein, 1976.

V-shaped dimers (Prockop *et al.*, 1976). Prolyl hydroxylase requires molecular oxygen and α-ketoglutarate as cosubstrates; non-heme iron and a reducing agent such as ascorbate are cofactors for the reaction. Prolyl hydroxylase has a very high substrate specificity and has been shown to hydroxylate only those prolyl residues which occur in the Y position of the repeating triplet sequence -Gly-X-Y of collagen (Prockop *et al.*, 1976). Not all the prolyl residues in the Y position are hydroxylated to the same extent (Bornstein and Traub, 1979).

1.3.3. Hydroxylation of lysyl residues.

HyL is another unusual amino acid characteristic of collagen. The presence of HyL has little known significance to the intracellular processing or secretion of the procollagen, but it plays a critical role in the formation of crosslinks which stabilize the extracellular collagen matrix (Bailey *et al.*, 1974).

The hydroxylation of selected lysyl residues in collagen is catalysed by lysyl hydroxylase (peptidyl-lysine, 2-oxoglutarate-oxygen 5-oxidoreductase, EC 1.14.11.4), (Figure 1-3). Lysyl hydroxylase has many similarities to prolyl hydroxylase in terms of the required cosubstrates, cofactors and in terms of the reaction mechanism (Bornstein and Traub 1979; Prockop *et al.*, 1976). Lysyl hydroxylase has been purified several hundred- to a thousand-fold from chick embryo extract (Kivirikko and Prockop 1972; Ryhanen, 1976), human fetal tissues and human placenta (Turpeenniemi-Hujanen *et al.*, 1981), and fetal porcine skin (Miller and Verner, 1979), by means of concanvalin A-Agarose and collagen-Agarose affinity columns with elution by ethylene glycol followed by gel filtration on a Biogel column. However, it is difficult to obtain a pure enzyme in view of its presence in relatively small amounts in the cell and because of the problems associated with its elution from the affinity column. The molecular weight of the chick embryo enzyme is about 200,000 by gel filtration and the active enzyme

appears to be a dimer consisting of only one type of subunit with a molecular weight of about 85,000 (Turpeeniemi-Hujanen *et al.*, 1980). There is an obvious discrepancy between the molecular weight of the active enzyme and that of its subunits. The molecular properties of lysyl hydroxylase from fetal calf skin differ distinctly from those of the chicken enzyme in that the former consists of two subunits with molecular weights of 70,000 and 115,000 (Miller and Varner, 1979).

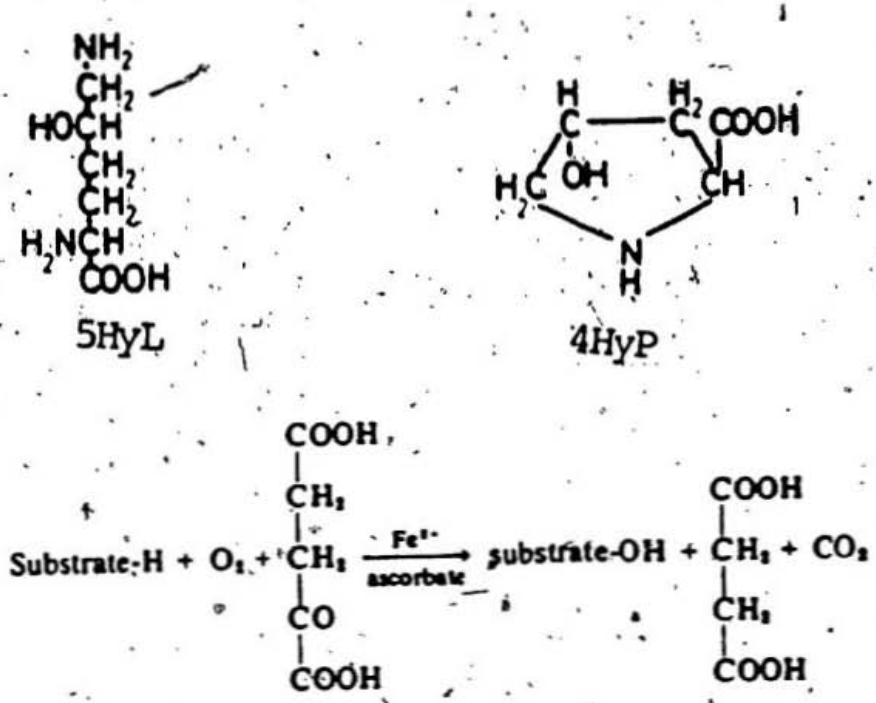

FIG. 1-3: The structures of 4-*trans*-hydroxy-L-proline and 5-hydroxy-lysine and the equation describing the hydroxylation reactions.

Lysyl hydroxylase, like prolyl hydroxylase, requires molecular oxygen, non-heme iron, α-ketoglutarate and ascorbate as cofactors and cosubstrates. The oxygen of the hydroxyl group is derived from molecular oxygen, the other atom of which is incorporated into the succinate (Kikuchi *et al.*, 1983). Fe^{2+} binds to

the enzyme but not firmly. It has been suggested that lysyl hydroxylase may bind 3 mol of Fe^{2+} at its maximum activity and that there is a positive co-operativity in this binding (Puistola *et al.*, 1980). The precise role of ascorbate in the hydroxylation reaction is not known with certainity. Although ascorbate is an almost specific requirement for highly purified lysyl hydroxylase (Puistola *et al.*, 1980), the vitamin is not, however, consumed stoichiometrically (Puistola *et al.*, 1980). Lysyl hydroxylase can catalyse its reaction in the complete absence of ascorbate at an essentially maximal rate for 5 to 10 seconds, corresponding to 15-30 reaction cycles (Myllyla *et al.*, 1978; Puistola *et al.*, 1980; De Jong *et al.*, 1982). It has been suggested (De Jong *et al.*, 1982) (from these findings together with kinetic and other data) that ascorbate is required to prevent oxidation of the enzyme-bound Fe^{2+}, and possibly some other groups on the enzyme molecules, during some catalytic cycles (Myllyla *et al.*, 1984). 2-oxoglutarate is an absolute and highly specific requirement for hydroxylation. The reaction involves the stoichiometric decarboxylation of 2-oxoglutarate (Kivirikko *et al.*, 1972) and the enzyme also catalyses an uncoupled decarboxylation in the absence of the peptide substrate (Puistola *et al.*, 1980). Dithiothreitol (DTT), bovine serum albumin (BSA) and catalase are also required for the enzymic reaction and all stimulate hydroxylation (Prockop, 1972; Miller, 1972; Popenoe and Aronson, 1972). The stimulation by DTT suggests that the catalytic site contains free thiol groups which are essential for the enzymic activity (Kivirikko and Myllyla, 1980). The effect of catalase is probably partly due to the destruction of peroxide, which is generated non-enzymatically by solutions of O_2, Fe^{2+} and ascorbate (Kivirikko and Prockop, 1967) and in part due to a non-specific "protein effect" (Rhoads and Udenfriend, 1970). The stimulatory effect of BSA maybe partly explained on the basis of a "protein effect" which has also been demonstrated

with other proteins and partly due to the presence on this protein of a number of free thiol groups (Rhoads *et al.*, 1967; Popenoe *et al.*, 1969). Even though hydroxylation of lysyl residues is initiated while the polypeptide is on the ribosomes, it seems that the formation of HyL continues for some time after the release of peptides from the ribosomes (Uitto and Prockop, 1974). Examination of the reported amino acid sequence of several collagens (Kivirikko *et al.*, 1972) reveals that the hydroxylation of lysine occurs in the Y position of the Gly-X-Y triplet, as was the case with HyP. The degree of hydroxylation of lysyl residues in the Y position of the Gly-X-Y sequences varies significantly among the collagens from different sources (Butler, 1968). In particular, collagens of type I and type III are frequently hydroxylated to a lower degree so that these collagens normally contain 4 to 8 HyL residues per 1000 residues. In type VI collagen most of the lysyl residues are hydroxylated (Dehm and Kefalides, 1978).

As in the case prolyl hydroxylase, the use of synthetic peptides has contributed to our understanding of the substrate specificity of lysyl hydroxylase and the nature of the enzymatic reaction (Kivirikko *et al.*, 1972). The studies on the specificity of lysyl hydroxylase using peptide substrates are, however, limited. They reveal that the sequence -Lys-Gly- is required for hydroxylation (Ryhanen, 1975). Unlike the case of prolyl hydroxylase, no conformational studies of lysine-containing synthetic peptide substrates have been reported to date. The substrate specificity of lysyl hydroxylase and the nature of the enzymatic reaction has been studied by Kivirikko *et al.*, (1972) using lysine-containing synthetic peptides (labelled as L-I, -II and -III) with amino acid sequences comparable to those found around glycosylated hydroxylysines in collagen. The peptides L-I had the sequence Ala-Arg-Gly-Ile-Lys-Gly-Ile-Arg-Gly-Phe-Ser-Gly, L-II had the sequence Ala-Arg-Gly-Met-Lys-Gly-His-Arg-Gly-(Pro-Pro-Gly)$_4$ and L-III had the sequence

$(Pro\text{-}Pro\text{-}Gly)_4$Ala-Arg-Gly-Met-Lys-Gly-His-Arg-Gly-$(Pro\text{-}Pro\text{-}Gly)_4$. L-I, -II and -III were all hydroxylated with a similar V_{max} but the K_m for L-I was twice that for L-II. There was no difference between the K_m for L-II and that for L-III. Occurence of hydroxylation was established by measurement of both the HyL synthesized and the $^{14}CO_2$ released from [1-^{14}C] α-ketoglutarate.

Free lysine and the tripeptide Lys-Gly-Pro were not hydroxylated (Kivirikko *et al.*, 1972). The tripeptide Ile-Lys-Gly showed a very slow rate of hydroxylation while the rate of hydroxylation of $(Ile\text{-}Lys\text{-}Gly)_2$ was over ten times greater. Lysine^8-vasopressin with the terminal sequence Cys-Pro-Gly-NH_2 was also hydroxylated. However, the kinetic measurements indicated that Ile-Lys-Gly and $(Ile\text{-}Lys\text{-}Gly)_n$ and lysine^8-vasopressin were poorer substrates compared to L-I, -II and -III(Kivirikko *et al.*, 1972). From these studies it was concluded that a single triplet of Gly-X-Y fulfills the minimum requirement for recognition by lysyl hydroxylase. The data also showed that although the chain length is critical with short peptides, increasing the chain length beyond seven triplets had no effect , or the chain must be extended with the triplets other than ProProGly in order to further improve the interaction with protocollagen lysyl hydroxylase (Kivirikko *et al.*, 1972). Recently, Glass *et al.*, (1985) synthesized a radiolabelled tridecapeptide, Gly-Leu-Hyp-Gly-Nle[4,5-^{3}H]-Lys-Gly-His-Arg-Gly-Phe-Ser-Gly which corresponds closely to residues 98 to 110 of the α-chain of type I collagen. This peptide was hydroxylated by human protocollagen lysyl hydroxylase. The non-radiolabelled analogue of this peptide was shown to inhibit the hydroxylation of [^{3}H]lysine-containing protocollagen by human lysyl hydroxylase. The IC_{50} values for the inhibition were in good agreement with the reported K_m values of the enzymes from either chick embryo (Kivirikko *et al.*, 1972; Puistola *et al.*, 1980) or human placenta (Turpeenniem-Hujanen *et al.*, 1981) using the substrate L-I.

Reduced and carboxymethylated Ascaris collagen does not serve as a substrate or as a competitive inhibitor of lysyl hydroxylase (Kivirikko *et al.*, 1972). This observation is surprising since Ascaris collagen contains 40 residues of lysine per 1000 residues (McBride and Harrington, 1967) and it is a good substrate for the synthesis of HyP by the proline hydroxylase (Fujimoto and Prockop, 1968). Perhaps in Ascaris collagen, all the lysyl residues are in the X position of the repeating triplets of Gly-X-Y, alternatively, the collagen contains lysyl residues but in sequences which prevent them from interacting with the enzyme (Weistein *et al.*, 1969).

In vitro studies by Ryhanen (1975) showed that lysyl residues in lysine-rich histones which contain Gly-X-Lys sequences were hydroxylated by lysyl hydroxylase and so were the lysyl residues in arginine-rich histones which did not have glycine following the lysyl residues. This is an important observation because although HyL invariably precedes glycine in the helical regions of collagen, the lysyl residues in the amino- and carboxy terminal telopeptide regions of the $\alpha 1$ chain, which are variably hydroxylated in some tissues, are in X-Lys-Ser and X-Lys-Ala sequences (Bornstein and Wolfe, 1979). Whether the partially purified lysyl hydroxylase contains more than one enzyme or whether a single enzyme is capable of these hydroxylations remains to be established. Recently, Royce and Barnes (1985), reported the failure of highly purified lysyl hydroxylase to hydroxylate the N-terminal non-helical telopeptide regions of chick type I collagen. The two telopeptide regions $\alpha 1$(I)-CB1 and (I)-CB1 (isolated from protocollagen following CNBr digestion) each containing a single lysyl residue in the sequences X-Lys-Ser and X-Lys-Ala, respectively, did not show any formation of HyL following incubation with highly purified lysyl hydroxylase. These results may suggest a requirement for a telopeptidyl-specific lysyl hydroxylase since the

same enzyme preparation used to test susceptibility of the telopeptide regions to hydroxylation was found to hydroxylate lysyl residues (with the Gly-X-Lys sequence) within the helical regions of the same protocollagen substrate (Royce and Barnes, 1985). Interestingly, some of the lysyl residues in lysine-rich histones shown to be hydroxylated by partially purified lysyl hydroxylase are in the X-Lys-Ser and X-Lys-Ala sequences (Ryhanen, 1975) apparent in the telopeptides. Ryhanen (1975) in the same study also reported that the same partially purified lysyl hydroxylase also hydroxylated lysyl residues in the X-Lys-Gly sequence. However, since a partially purified lysyl hydroxylase was used in these studies (Ryhanen, 1975), it is not possible to say whether more than one lysyl hydroxylase was present.

1.3.4. Cross-linking in collagen

Following the removal of extension peptides, the collagen molecules spontaneously align to form fibers. However, these fibers do not attain the necessary tensile strength until the molecules are linked by specific covalent bonds known as cross-links (Gross, 1974; Tanzer, 1973; Bailey *et al.*, 1974). The most common forms of cross-links in collagen are derived from lysine and HyL. Lysyl oxidase (protein lysine oxidase EC 1.4.3.13) is the only enzyme known to catalyse the oxidative deamination of specific ϵ-amino groups of lysyl and hydroxylysyl residues in collagen. Lysyl oxidase is a copper metalloenzyme that requires pyridoxal phosphate as a cofactor (Bird and Levene, 1982; 1983). This enzyme functions in the extracellular space and it has high activity with collagen which has been precipitated as native fibrils as compared to denatured collagen or isolated α-chains (Siegel and Pinnell, 1970) suggesting that the formation of aldehydes primarily occurs after the onset of fibril formation. Lysyl oxidase is irreversibly inhibited by the lathrogen β-aminopropionitrile, apparently because it covalently binds at

the active site, probably initially by aldimine addition to pyridoxol phosphate as in substrate binding but is not released as an oxidized product (Tang *et al.*, 1982). There are two pathways of cross-linking in the fibrillar collagens: one based on allysine,(see Figure 1-4) the lysine-derived aldehyde, the other on hydroxyallysine, the HyL-derived aldehyde (see Figures 1-6 and 1-7).

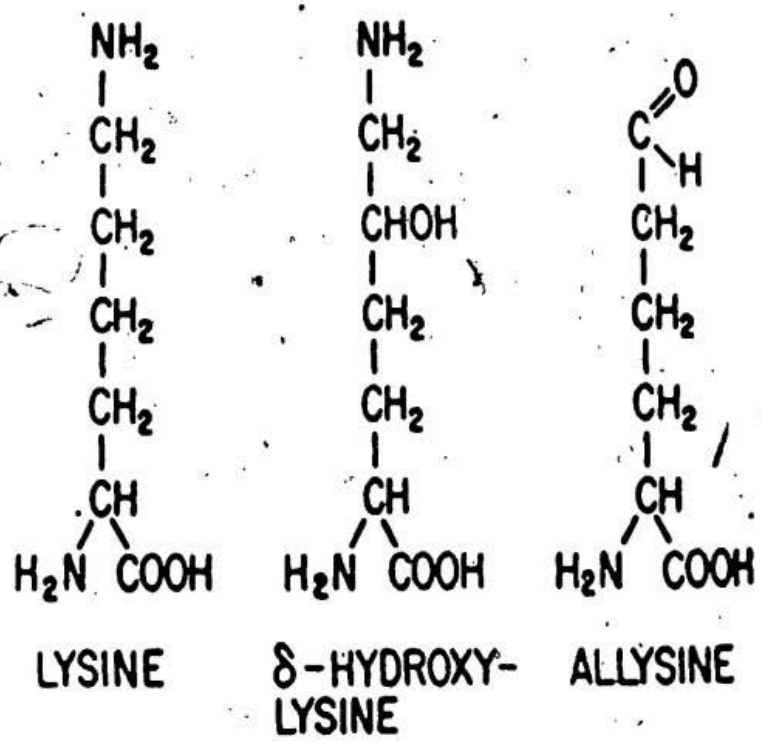

FIG. 1-4: The structure of lysine, 4-HyL and allysine.

Trifunctional, 3-hydroxypyridinium residues have been identified as the adult cross-links on the hydroxyallysine route, which predominates in connective tissues except skin (Mechanic *et al.*, 1971, Davis and Bailey, 1971). There are

two forms of 3-hydroxypyridinium cross-links, hydroxylysyl pyridinoline (HP) being the major one which embodies three hydroxylysines and a less abundant form, lysyl pyridinoline (LP) which embodies two hydroxylysines and one lysine (Eyre *et al.*, 1984). The mature cross-linking residues on the allysine pathway are still unknown, although complex structures capable of linking three or more molecules are suspected (see below). Histidine is probably a component on this pathway although there is not yet general agreement on this point (Tanzer *et al.*, 1973). Glycosylated HyL-derived cross-links have been reported in skin, bone and cartilage collagens (Eyre and Glimcher, 1973; Robins and Bailey, 1974). The significance of the sugar residues is unknown; for example whether they can facilitate or inhibit particular cross-linking reactions.

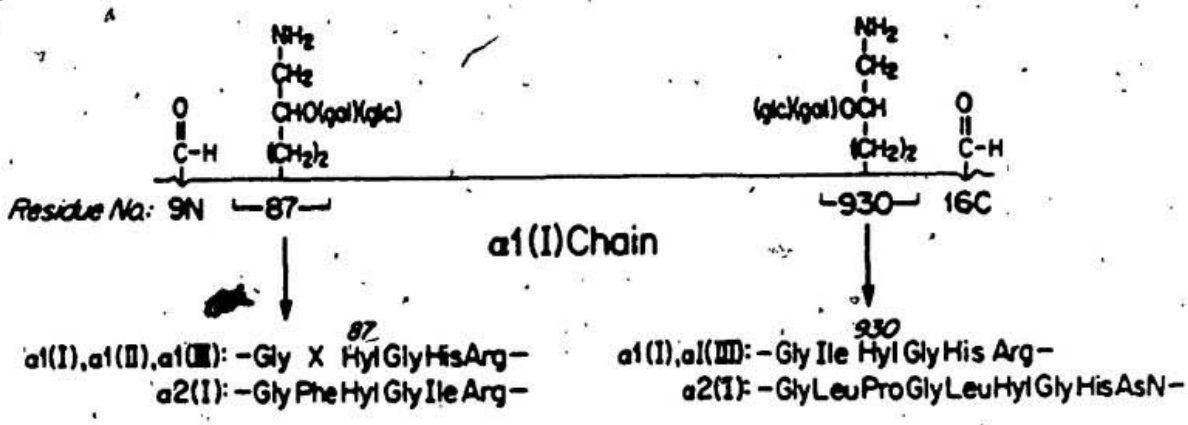

FIG. 1-5: The four principal cross-linking loci in molecules of types I, II and III collagens.

Reproduced from Eyre, 1984.

In collagen molecules of types I, II and III, there are four homologous loci of cross-linking and all have conserved characteristic amino acid sequences (Kuhn, 1982) (see Figure 1-5) Two of these are aldehyde-sites, one in each telopeptide region. The other two triple-helical sites (hydroxylysine-sites) are symmetrically placed at about 90 residues from each end of the molecule. These HyL sites in the helical region align and react with telopeptide aldehydes in adjacent molecules staggered axially by 4D periods (4x67nm) when collagen molecules pack into fibrils.

During *in vivo* maturation and also during *in vitro* incubation under physiological conditions, native collagen fibers display a progressive increase in tensile strength and insolubility (Sinex, 1968). These changes in the properties of collagen are paralleled by a progressive decrease in the content of the borohydride-reducible cross-links (Robins *et al.*, 1973). Studies of the reducible components presented in tissues covering a wide range of ages showed that no new compounds are formed during the aging process (Deshmukh *et al.*, 1971; Fujii and Tanzer, 1971). To account for these facts, it was proposed that the reducible cross-links act only as intermediates and are converted to more stable, non-reducible forms during the maturation process (Bailey *et al.*, 1974).

The nature of the mature cross-links has been investigated by several laboratories and is still a controversal issue (see Stimler and Tanzer, 1976, for review). It has now been widely accepted that pyridinoline is the mature cross link of collagen (Eyre and Oguchi, 1981; Kuboki *et al.*, 1981; Robins and Duncan, 198; Wu and Eyre, 1984). However, Light and Bailey (1985) failed to detect pyridinoline in skin, tendon bone and dentin of type I collagen. Instead, they isolated a polymeric complex which they termed polyα1CB6 which they propose must contain the mature cross-link.

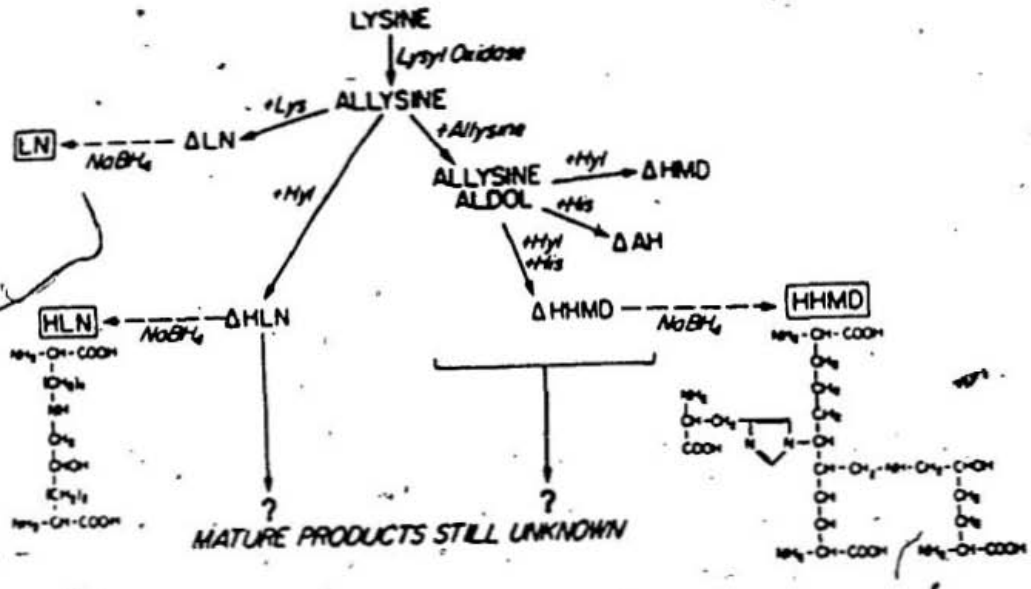

FIG. 1-6: Scheme outlining one of the two routes of cross-linking in collagen: Allysine-based cross-links. LN, lysinonorleucine; HMD, hydroxymerodesmosine; AH, aldolhistidine; HHMD, histidinohydroxymerodesmosine. The prefix Δ for dehydro signifies the natural, aldimine forms of the various compounds.

Reproduced from Eyre, 1984.

From the foregoing data, it is obvious that the hydroxylation of lysyl residues in the procollagen molecule forms the chemical basis on which subsequent cross-linking of the collagen triple helices is accomplished. The specific manner in which certain domains in the individual polypeptide chains are cross-linked seems to dictate the appropriate staggering of the triple helices to form functionally viable fibrillar aggregates of the collagen molecule in the tissues.

Because the cross-links of collagen provide the tensile strength required in a functioning tissue, a defect in the formation of these covalent bonds can lead to a disturbance in connective tissue. For example in Ehlers-Danlos syndrome type IV, characterized by a low HyL content (Pinnell *et al.*, 1972) due to lysyl hydroxylase deficiency (Krane *et al.*, 1972), results in an abnormal profile of reducible cross-links in skin, bone and cartilage collagens (Eyre and Glimcher, 1972).

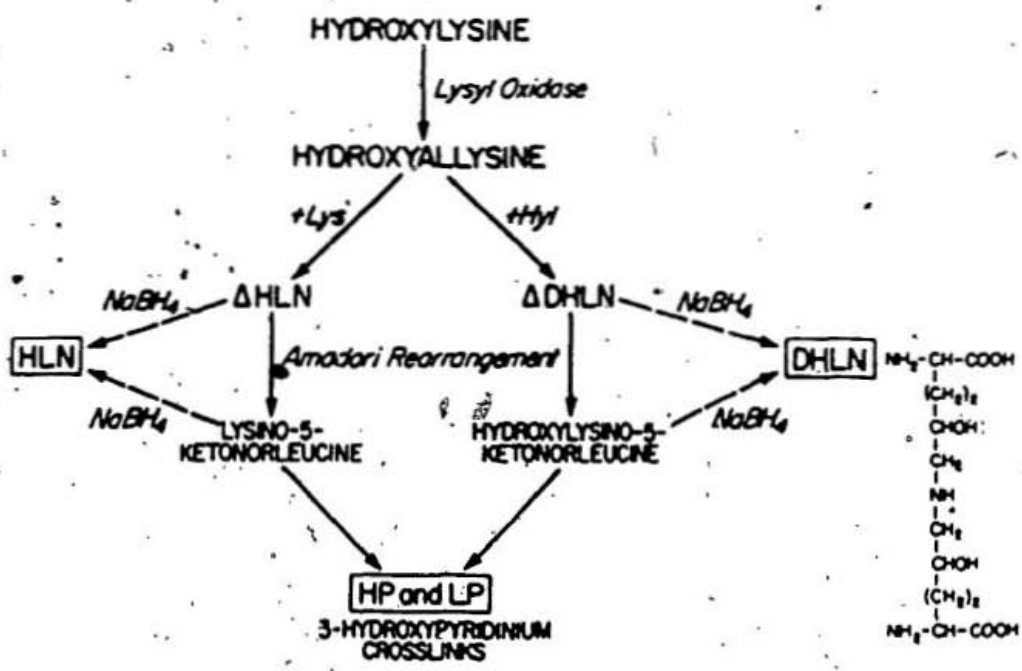

FIG. 1-7: Scheme outlining one of two routes of cross-linking in collagen: Hydroxyallysine-based cross-links. DHLN, dihydroxylysinonorleucine; HP, hydroxylysyl pyridinoline; LP, lysyl pyridinoline. The prefix Δ for dehyro, signifies the natural, aldimine forms of the various compounds.

Reproduced from Eyre, 1984.

1.4. Importance of the hydroxylation reactions in collagen.

Since the biosynthesis of collagen is characterized by several posttranslational modification processes, any defect in the normal mechanisms responsible for the synthesis and secretion of collagen molecules and the subsequent deposition of these molecules into extracellular fibers could result in abnormal fibrillogenesis. Such defects in turn could result in connective tissue disorders. As previously mentioned, HyP plays a critical role in the stabilization of the conformation of collagen. The triple helix, in turn, is required for the secretion and extracellular deposition of the collagen molecule. Therefore, in the absence of HyP, the collagen polypeptide chains would not acquire the critical triple helical structure under physiological conditions and no collagen fibers would appear in the extracellular space. As yet, no inherited disease in which collagen is deficient in HyP or in which there is deficiency in prolyl hydroxylase has been reported (Uitto and Lichtenstein, 1976). However, most of the clinical manifestations of scurvy which involve the connective tissue, could be explained on the basis of deficient hydroxylation of prolyl residues in collagen. Since prolyl hydroxylase requires a reducing agent such as ascorbic acid, a deficiency in ascorbic acid leads to decreased formation of collagen fibers (Cardinale and Udenfriend, 1974). This could explain some of the clinical manifestation such as poor wound healing and decreased tensile strength of connective tissues observed in scurvy (Barnes and Kodicek, 1972). An analogous situation may exist in tissues with relative anoxia since molecular oxygen is a specific requirement for the formation of the hydroxyl group in HyP. Uitto and Prockop (1974), Niinikoski (1969) and Silver (1973) using model animals have independently demonstrated that the healing of wounds is relatively poor under hypobaric conditions and in such a situation, the low O_2 levels may well limit the synthesis of HyP. This observation may well

explain the decreased healing tendency of wounds and ulcers in peripheral tissues which are anoxic due to relatively scant blood supplies.

HyL residues in collagen serve as sites of attachment of galactosyl and galactosyl glucose moieties (Butler and Cunningham, 1966; Spiro, 1969) and participate in extracellular interchain crosslink formation (Gallop and Blumenfeld, 1972; Tanzer, 1973). The suggestion that a normal content of HyL (and, by implication, of the glycosylated residue) was required for secretion of collagen was based on the observation that chick tibae, incubated in the presence of the lysine analog 4,5-dehydroxylysine failed to extrude the protein normally (Christner and Rosenbloom, 1971). Contrary to this hypothesis, Ramaley *et al.*, (1973) have shown that drastic inhibition of peptidyl hydroxylation by α,α'-dipyridyl retards but does not prevent secretion of procollagen by 3T6 fibroblasts in culture. Similarly, a clear-cut inhibition of crosslink formation by glycosylation of HyL residues has been called into question by the isolation of a glycosylated crosslink (Eyre and Glimcher, 1973). Nevertheless, it is likely that the modulation of lysyl hydroxylation and hydroxylysyl glycosylation plays an important role in modifying the structural and functional characteristics of collagen fibers in tissues. As an extreme example, a deficiency in lysyl hydroxylase leads to a severe connective tissue disorder, HyL-deficient collagen disease or the Ehlers-Danlos Syndrome type VI, in which the mechanical properties of tissues such as skin and ligaments are impaired (Pinnel *et al.*, 1972; Krane *et al.*, 1972). This syndrome is characterized by severe scoliosis, recurrent joint dislocation and hypertensible skin and joints (Pinnel *et al.*, 1972). Fibroblasts from individuals with HyL-deficient collagen are markedly deficient in lysyl hydroxyase activity (Krane *et al.*, 1972). The manifestations of this disorder probably result from inadequate cross linking of collagen, since lysyl aldehydes may not be as reactive, or lysyl-derived cross-

links as stable, as their HyL counterparts. The abnormal crosslink patterns obtained after reduction of affected tissues with sodium borohydride (Eyre and Glimcher, 1972) provide support for this suggestion.

Several other disease conditions have been associated with altered activities of prolyl and lysyl hydroxylase. The clinical and genetic aspects of some of these disorders have been collated by McKisick (1972). Among these are osteogenesis imperfecta, Marfan Syndrome, cardiac hypertrophy, sarcoma, pulmonary fibrosis, liver injury, arteriolosclerosis and hypertension.

1.5. Objectives and scope of present work.

In spite of its importance as an enzyme involved in the synthesis of the major protein collagen, (which accounts fo about 25%, by weight, of all proteins in vertebrates) lysyl hydroxylase has received scant attention for its general make-up and active site geometry in particular. It has been known for the past ten years that lysyl hydroxylase catalyses the 5-hydroxylation of lysyl residues in the Y position of the repeating triplet sequence -Gly-X-Y- but, to date no report has been made on the physicochemical properties of synthetic substrates for lysyl hydroxylase. It is also known that chain length and the amino acid in the position X of the repeating triplet sequence Gly-X-Y are critical determinants in the degree of lysine hydroxylation (Prockop *et al.*, 1976). Not much is known about the conformational basis, if any for these observations. The limited studies on the specificity of lysyl hydroxylase using a few synthetic peptides (Section 1.3.3) reveal that the sequence -Lys-Gly- is required for hydroxylation (Kivirikko and Prockop, 1972) but give no clues as to why such a sequence is necessary or why the Gly-Lys sequence does not permit hydroxylation. The reason for this lack of knowledge is mainly because no structural studies have been made on the synthetic substrates for lysyl hydroxylase.

Our laboratory has, for the past several years been involved in studying the conformational basis for specificity of prolyl hydroxylase (Brahmachari and Ananthanarayanan, 1979; Chopra and Ananthanarayanan, 1982; Ananthanarayanan, 1984; Ananthanarayanan *et al.*, 1985). To this end, proline-containing peptides corresponding to specific regions of th collagen sequence were synthesized and their physicochemical properties and interaction with pure prolyl hydroxylase studied. The earlier studies showed that prolyl hydroxylase recognizes the β-turn conformation formed at the Pro-Gly- segments of its substrates, the extent of this conformation and the extent of hydroxylation being governed by the nature of the adjoining residue X in the repeating sequence -Gly-X-Y-, where Y is the prolyl residue amenable to hydroxylation (Brahmachari and Ananthanarayanan, 1979). Recently, a further study of the interaction of prolyl hydroxylase with several peptides having the -Pro-Gly- and -Gly-Pro- sequences has led to the modification of the β-turn theory (Ananthanarayanan, 1984, 1985). These studies revealed that non-β-turn -Gly-Pro- peptides (and the β-turn -Pro-Gly- peptides) inhibited the hydroxylation of the standard synthetic substrate $(Pro\text{-}Pro\text{-}Gly)_5$. A study of the conformation of the Gly-Pro peptides showed that they contained an extended or rigid structure whose dihedral angles are similar to those of the polyproline-II (PP-II) helix whereas the Pro-Gly- sequence favoured the β-turn conformation. These findings led our laboratory to postulate that prolyl hydroxylation required both the PP-II-type structure (at the N-terminal side) and the β-turn, the hydroxylatable Pro being situated at the junction of these two conformations. While the PP-II arm seems to be necessary for the binding of the substrate in the active site of prolyl hydroxylase, the β-turn is essential at the catalytic site (Ananthanarayanan, 1984; Ananthanarayanan *et al.*, 1985) (see Figure 1-8).

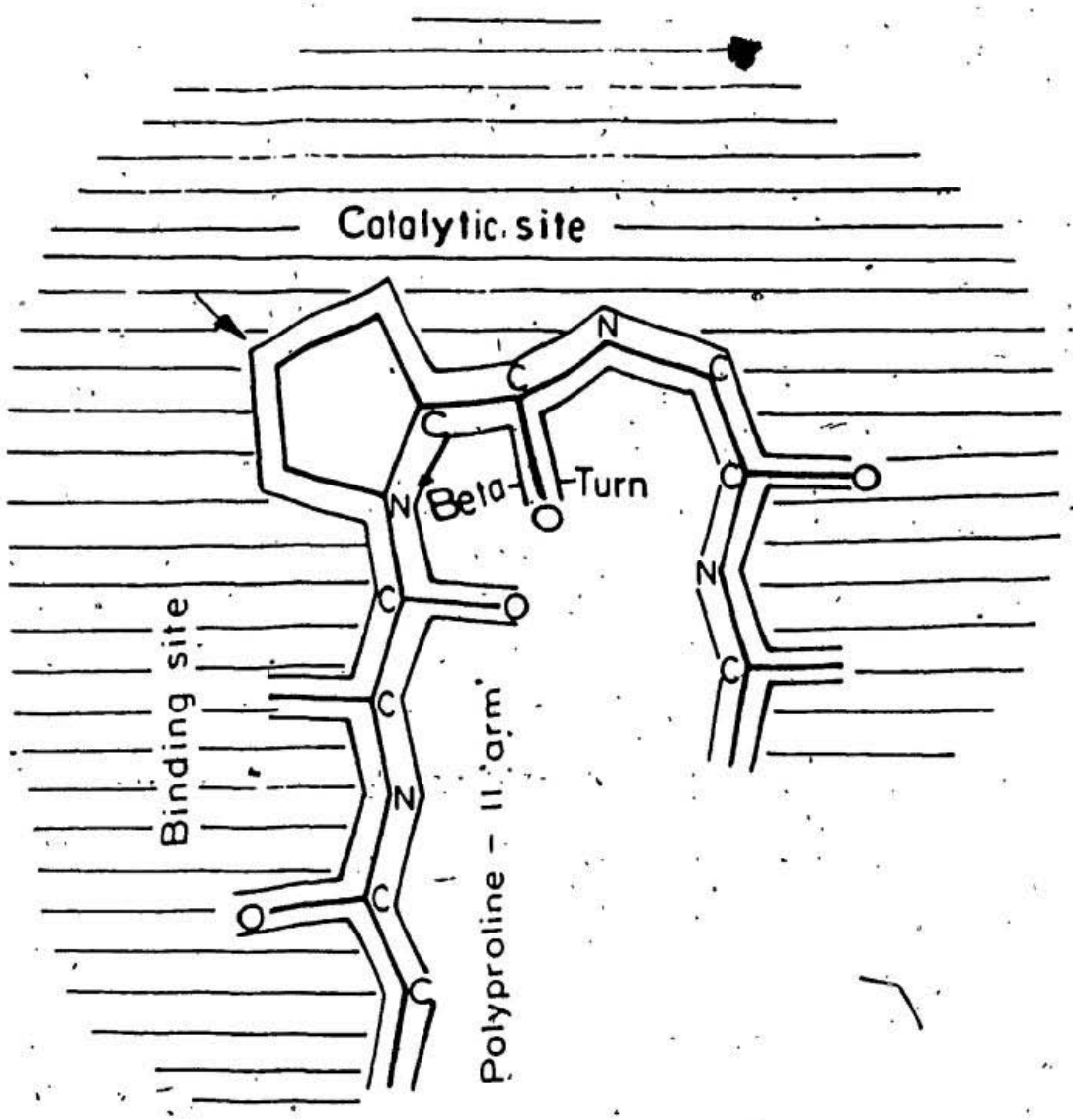

FIG. 1-8: Schematic representation of the proposed model substrate of prolyl hydroxylase. The arrow indicates the position of hydroxylation.

Reproduced from Ananthanarayanan *et al.*, 1985.

Since lysyl hydroxylase and prolyl hydroxylase share similar properties in terms of cofactors and mechanism of hydroxylation, we asked ourselves the following questions: (a) does lysyl hydroxylase recognize a β-turn? (b) does lysyl hydroxylase recognize the PP-II type structure or is the PP-II type structure a structural feature confined to the -Gly-Pro- peptides? In an attempt to get answers to some of these questions, it was decided that we follow the strategies of the prolyl hydroxylase work and study the interaction of suitably designed lysine-containing peptides with chick embryo lysyl hydroxylase.

The objectives of our work were: (a) to synthesize lysine-containing peptides found in the hydroxylated regions of the collagen, (b) to test the ability of these peptides to act as substrates or inhibitors of lysyl hydroxylase partially purified from chicken embryos, (c) to study the structures of these peptides and (d) to find if lysyl hydroxylase reconizes specific secondary structural feature(s) in the substrates.

The selection of peptide fragments for our synthesis, was based on the frequency of occurrence of certain sequences in collagen where lysyl residues were found to be either hydroxylated or unhydroxylated. In order to determine the minimal and optimal requirement for hydroxylation in terms of the chain length of the substrate and its amino acid sequence, peptides with different triplet sequence(X-Y-Gly) and varying chain length from three to seven residues peptides were synthesized (Chapter 3). The physicochemical properties of these peptides were studied by NMR, CD and IR (Chapter 4). The peptides were also tested for their ability to act as substrates or inhibitors of lysyl hydroxylase purified by us from chicken embryos (Chapter 5).

Chapter 2

2. Experimental

2.1 Materials

Amino acid derivatives and chemical reagents were commercial products which were used without further purification.

The following were purchased from Sigma Chemical Co. (Missouri, USA): N^{α}tBocAla, N^{α}tBocIle, N^{α}tBocLeu, N^{ω}nitro-L-ArgOMe HCL, N^{ϵ}cbz-L-LysOMe HCL, N^{ϵ}cbz-L-LysOH, 4-HyPOMe HCL, glycine, serine, dithiothreitol, glycerol, Trizma base, ammonium sulfate, manganese chloride, methyl-α-D-glucoside, ascorbic acid, ferrous sulfate, bovine serum albumin, catalase, α-ketoglutarate, N-N dicyclocarbodiimide, isobutylchloroformate, N-hydroxysuccinimide, trifluoroacetic acid, trifluoroethanol, anhydrous sodium sulfate, sodium chloride, sodium bicarbonate, potassium sulfate, 2,5-diphenyloxazole (PPO), 1,4-bis[2-(4-methyl-5phenyloxazoyl)] benzene and sodium metaperiodate.

The following were obtained from Fisher Scientific Co. (Nova Scotia, Canada): tetrahydrofuran, 1,4 dioxane, methanol (HPLC grade), NaOH, chloroform, dichloromethane, ethyl acetate (HPLC grade), petroleum ether, anhydrous ether, acetonitrile (HPLC grade), 2-propanol, p-dimethlyaminobenzaldehyde, isobutanol, perchloric acid, toluene and ethylene glycol monomethyl ether.

DL-5-hydroxylysine-HCL (mixture of DL and DL-allo) was obtained from Fluka Chemical Co. (Toronto, Canada). 2-[1-^{14}C]α-ketoglutarate and NCS tissue solubilizer were obtained from Amersham International (Ontario, Canada). Silica gel (60-200 mesh) was purchased from J.T. Baker Chemical Co. (Montreal, Canada). Concanvalin A-Sepharose 4B was obtained from Gibco Laboratories (Ontario, Canada). CNBr-activated Sepharose 4B was purchased from

Pharmacia (Quebec, Canada).

2.2. Methods.

2.2.1. Syntheses of peptides

Amino acids and amino acid derivatives were of L-configuration. All peptides were prepared by stepwise condensation of protected N^{α}tBoc-amino acids using either dicyclohexylcarbodiimide/N-hydroxysuccinimide or N-methylmorpholine/isobutylchloroformate as coupling reagents (Chapter 3).

2.2.2. Analytical Reversed-Phase High Performance Liquid Chromatography

The purity of peptides was checked by reversed-phase HPLC on a 5 micron HS C18 column (4.0mm I.D. x 6.25 mmO.D. x 15.0cm length, Perkin-Elmer) at ambient temperature. A Perkin-Elmer Model LC-4 HPLC instrument equipped with a microprocessor-controlled programmer was employed. Samples of approximately 1-5 mg were dissolved in 1-5 ml MeOH, filtered using HPLC nylon filters (3mm membrane, 0.45μm pore size) and applied to the column in volumes of 6-20μl. The eluate was monitored on a Perkin-Elmer LC-85B Spectrophotometric variable U.V. detector at 220nm and 230nm which was connected to a one-channel recorder REC-481 (Perkin-Elmer).

The stepwise gradients for elution were established with HPLC-grade solvents in reservoirs A, B, C, and D (see Table 2-1 for the solvent composition mixture and other parameters used for analytical HPLC). The stepwise gradients were chosen in such a way that peptides of different polarity could elute at different times. This was achieved by trial and error by increasing or decreasing the polarity of water. The solvent in A and B was acetonitrile, solvent in C was methanol and solvent in D was water. Prior to loading the sample the solvents were purged and the column was equilibrated with water for 10 minutes. For all HPLC work Hamilton microlitre syringes (Fisher Scientific) were used for loading

the samples.

2.2.3. Preparative Reversed-Phase High Performance Liquid Chromatography

For Preparative HPLC, the outlet was connected to the monitoring instruments and then to a Pharmacia fraction collector F-100. The HPLC parameters for purification were similar to those employed for analytical HPLC (Table 2-1) except for the flow rate (1 ml/minute) and time in each sector which was increased to 10 minutes. 500 mg of sample was dissolved in 2 ml MeOH and 250μl were applied to the Prep Synchropak RP-P C-18 column (250 x 10 mm, I.D. 300 A pores, 6.5μ particle size, Synchrom Inc.) at a flow rate of 1 ml/min. The fractions corresponding to the major peak were pooled and the solvent was evaporated on a rotary evaporator and re-chromatographed if still found not to be pure.

Table 2-1: HPLC parameters

SECTOR	TIME	FLOW MLS	CURVE	MeOH A%	Acn B%	Acn C%	H_2O D%	CURVE
EQUIL	10	0.8		0	0	0	100	
1	5	0.8		0	15	15	70	1
2	5	0.8		10	15	10	65	1
3	5	0.8		5	10	10	75	1
4	5	0.8		5	10	5	80	1
5	5	0.8		5	5	5	85	1

READY TIME 60min PMIN 0.2 MPa PMAX 42 MPa TEMPERATURE RT.

Chart Speed 2-5 (mm/ml). Abs. Scale 6.4 Wavelength(Nm) 220

2.2.4. Silica gel Chromatography

N- and C- terminal blocked tri- and hexapeptides were purified by chromatography on short and long silica gel columns respectively. For the tripeptides, about 100 ml silica gel (60-200 mesh) was suspended in petroleum ether and loaded onto the column. The sample (1.5-3.0g) was dissolved in 10-20 ml of

chloroform and the solution was applied to the column. The column was washed with 25% ethyl acetate in petroleum ether and the product was eluted with 5% MeOH in ethyl acetate. 50 ml fractions were collected and were analyzed by HPLC. The pure fractions were pooled and the solvent was removed on a rotary evaporator. The material was further purified by crystallization to afford the product.

For the hexapeptides, about 30 ml silica gel was suspended in petroleum ether and poured into a column. The peptide (1-1.5g) was dissolved in 5-10 ml chloroform and the solution was loaded onto the column where it was left for 30 minutes to 1 hour to allow retention of the more polar peptide(s) by the silica gel column. The column was washed successively with 100 ml of 50% ethyl acetate in petroleum ether and 100 ml ethyl acetate. The product was eluted with 20% isopropanol/5% MeOH in ethyl acetate. The 50 ml fractions collected were treated in the same manner as described for the tripeptides. The material was further purified by Prep. HPLC to afford the pure product.

2.2.5. Elemental Analysis

Microanalyses were carried out by Canadian Microanalytical Service Ltd. (British Columbia) and Guelph Chemical Laboratories Ltd. (Ontario).

2.2.6. Amino Acid Analysis

Amino acid analyses were carried out by Mr. D. Hall and his assistants in the Biochemistry Department Amino Acid Facility with a Beckman model 121 amino acid analyzer on samples that had been hydrolyzed using 6M HCl at 110C for 24 hours in evacuated sealed tubes.

2.2.7. Melting Points

Melting points were determined on a Thomas Hoover melting-point apparatus using unsealed capillaries and are uncorrected.

2.2.8. Circular Dichroism (CD) Measurement

CD spectra were recorded using a Jasco J-500A automatic recording spectropolarimeter equipped with a DP-500N data processor. All measurements were made in 0.1-1.00mm path-length quart cells (Jasco). Wavelength was calibrated with a neodymium glass and amplitude was calibrated using 0.6% D-10-camphor sulfonic acid in H_2O. Peptide concentration were 1-2mg/ml and spectra were recorded in HPLC grade trifluoroethanol, methanol, water, acetonitrile and 0.1M ammonium bicarbonate in water at ambient temperature. The spectra were averaged after thirty accumulations at 20 nm/min and a time constant of 2 seconds. The spectra were obtained by electronically subtracting the spectrum of solvent from that of the peptide in solvent. The ellipticity (Θ) values given in deg $cm^2/dmol$ were calculated on a mean residue weight basis. The mean molar ellipticity was calculated from the original spectrum using the formula:

$$[\Theta]_{res} = X \times S \times MRW \times (c \times l \times 10)^{-1}$$

Where X equals the pen deflection in cm, S is the scale used in millidegrees per cm, c is the concentration in mg ml^{-1}, and l refers to the path length in cm. MRW or mean residue weight is calculated from the molecular weight of the peptide divided by the number of peptide bonds.

2.2.9. Infrared Spectroscopy (IR)

The IR spectra were measured on Perkin-Elmer model 983G Spectrophotometer using 1.0mm BaF_2 cells (Buck Scientific). The wavelength calibration was checked using a 0.05 mm polystyrene film. The spectra were run in chloroform

freshly distilled over calcium chloride. The difference spectrum was obtained electronically by subtracting the $CHCl_3$ spectrum from the spectrum of the peptide dissolved in $CHCl_3$.

2.2.10. Partial Purification of Lysyl Hydroxylase.

All purification procedures described below were carried out at 4^oC. Lysyl hydroxylase was obtained from 14-day old whole chick embryos as described by Kivirikko and Myllya (1982) and modified by Royce and Barnes (1985). A total of 100 embryos were homogenized in batches of 30 in enzyme buffer (0.2M-NaCl/0.1M-glycine/10uM-dithiothreitol/1%-glycerol/ 20mM-Tris/HCl, pH 7.5 at 4°); supplemented with 0.1% Triton X-100 (1 ml of solution per gram of embryos) in a Waring blender at full speed twice for 30 seconds. The homogenate was allowed to stand with occasional stirring for 30 minutes and then centrifuged in volumes of 200 ml at 15,000*g* for 30 minutes using a Sorval Superspeed RC2-B automatic refrigerated centrifuge. Solid $(NH_4)_2SO_4$ was slowly stirred into the supernatant fraction to a final concentration of 17% saturation (97g/liter). The pellet obtained by centrifugation at 15,000*g* for 20 minutes was discarded, and solid $(NH_4)_2SO_4$ was slowly stirred into the supernatant to a final concentration of 55% saturation (244g/liter). The pellet obtained by centrifugation at 15000*g* for 20 minutes was dissolved in enzyme buffer, (0.2M-NaCl/0.1M-glycine/10μM-dithiothreitol/1%-glycerol/20mM-Tris/HCl, pH 7.5 at 4^oC) containing 3mM $MnCl_2$ and the solution was dialyzed against the same buffer for 4 and 12 hours. After removal of the insoluble material by centrifugation, the supernatant was added to concanvalin A-Sepharose previously equilibrated with enzyme buffer containing 3mM $MnCl_2$ and the mixture was stirred gently for 18 hours (Royce and Barnes,1985). The resin was subsequently washed with enzyme buffer containing 3mM $MnCl_2$ and 0.5M-methyl α-D-pyrannoside, and was then poured

into a column. Washing was continued until the absorbance of the eluate at 225nm was below 0.1 and the enzyme was subsequently eluted with enzyme buffer containing 0.3M-methyl α-D-glucoside and 60% ethylene glycol (v/v). The enzyme preparations were stored in 4 ml aliquots at -20°C until further use. Immediately before use, enzyme buffer was added until the ethylene glycol concentration was 6% following which the enzyme was concentrated in an Amicon ultrafiltration cell membrane with a YM-30 membrane. This enzyme fraction represented a 95-fold purification (see Table 2-2).

Table 2-2:

Partial Purification of lysyl hydrozylase from a chick-embryo $(NH_4)_2SO_4$

Enzyme fraction	Total protein (mg)	Total activity µmol	Specific activity µmol/mg	Purification -fold
A	8610	56	0.00065	2*
B	20	6.2	0.310	95

* assumed that the crude extract (15,000g) is the starting point i.e. purification-fold 1 (Kivirikko and Prockop; 1972, Ryhanen, 1976; Turpeenniemi-Hujanen *et al.*, 1977; Turpeenniemi-Hujanen; 1980).

A = $(NH_4)_2SO_4$

B = Concanvalin A-agarose

2.2.11. Enzyme Assay

Lysyl hydroxylase activity was assayed using synthetic peptide substrates by a technique involving the measurement of $^{14}CO_2$ released stoichiometrically from 2-[1-^{14}C]oxoglutarate (Kivirikko *et al* 1972). The incubations were carried out at 37°C in a water bath with shaking for 40-60 minutes in a final volume of 1-4 ml containing 1mg/ml peptide substrate, 1 ml enzyme containing approximately 0.18 mg of enzyme protein, 0.05mM $FeSO_4$, 0.1mM 2-[1-^{14}C]-oxoglutarate (55000 dpm), 1mM ascorbate, 0.1mM dithiothreitol, 0.1mg/ml catalase, 1mg/ml bovine serum albumin and 50mM Tris-HCl buffer, adjusted to pH 7.8 at 25°C (Kivirikko *et al* 1972). All the peptide substrates were heated to 100°C for 10 minutes and cooled to 0° just before addition to the incubation system. The $^{14}CO_2$ was collected onto filter papers as described by Rhoads and Udenfriend (1968) except that the reaction was stopped by injecting 1 ml of 1.0M potassium phosphate (pH 5.0) stored at 4°C. After the reaction had been stopped, the sealed tubes were shaken in a mechanical shaker for 2 hours and the filter papers were assayed in a solvent containing methyl Cellosolve and toluene (Prockop and Ebert, 1963). All values for disintegrations per minute of $^{14}CO_2$ (measured with a Beckman LS-330 liquid scintillation counter) released were corrected for the release of $^{14}CO_2$ observed with blank samples that did not contain the substrate.

HyL was also assayed with a specific chemical procedure described by Blumenkrantz and Prockop (1971). The method is based on periodate oxidation of HyL. Periodate oxidation of HyL produces glutamic semialdehyde and Δ^1-pyrrole-5-carboxylic acid which on further oxidation will form a colour with p-dimethylamino benzaldehyde as shown in Figure 2-1. Proline also gives rise to a similar chromophore in the assay and therefore a preliminary separation of samples by thin layer chromatography would be necessary. Since the synthetic

peptides used in the present study did not contain proline the thin layer chromatography step was omitted. The composition of the reaction mixture for the chemical method was as described above except that the 2-[1-^{14}C]oxoglutarate was replaced by 0.05mM oxoglutarate. The reaction was stopped by adding 6M HCl. The reaction mixture was then hydrolyzed for 24 hours at 120°C. The hydrolyzates were evaporated to dryness in a rotary evaporator. The dry residue was dissolved in 8 ml citrate phosphate (154 ml of 0.15M citric acid mixed with 346 ml 0.6M dibasic sodium phosphate, pH 7). The oxidation reaction was started by adding 0.3 ml of 0.3M sodium metaperiodate and the solution was stirred. Then 3.0 ml of extraction solution (250 ml of toluene mixed with 250 ml of isobutanol and 100 ml of n-propanol) was added and the tubes were stirred on a vortex mixer. The tubes were placed in a test tube rack; the rack was covered with aluminum foil and shaken on a horizontal shaking machine for 20 minutes. The tubes were centrifuged for 10 minutes at low speed in order to separate the aqueous and organic phases. 2 ml of the organic phase was placed in a test tube, and 0.5 ml of Ehrlich reagent (15 ml of isobutanol, 4 g of p-dimethylamino benzaldehyde and 4.5 ml of perchloric acid) solution was added. The tubes were immediately stirred vigorously. The colour was allowed to develop for 15 minutes at room temperature and the absorbance was read at 565 nm. The amount of HyL was estimated from the standard curve of δ-HyL-HCl (mixture of DL and DL-allo).

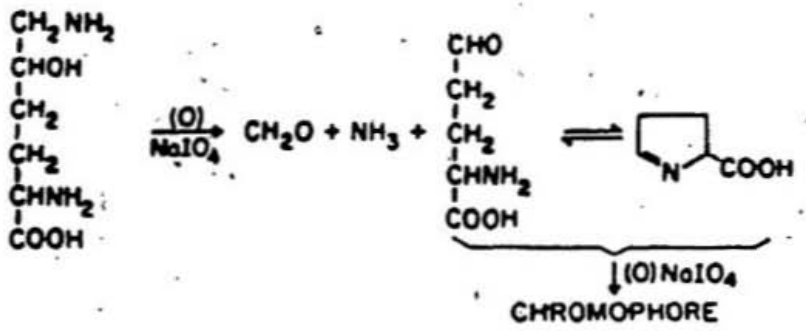

FIG. 2-1: Probable reactions for the oxidation of HyL by periodate. The first reaction product is glutamic semialdehyde, which is probably in equilibrium with Δ^1-pyrroline-5carboxylic acid. Δ^1-pyrroline-5-carboxylic acid did not give colour with Ehrlich reagent unless it was further oxidized to an unidentified product by periodate.

Reproduced from Blumenkrantz and Prockop, 1971.

2.2.12. Other Assays

HyL was also assayed by an amino acid analyzer just for one tetrapeptide to confirm the results of the other two assays (2.2.11.) After the reaction mixture was hydrolyzed for 24 hours at 110°C, the sample was analyzed with a Beckman model 121 amino acid analyzer. δ-Hyl-HCl (a mixture of DL and DL-allo) was used as the standard.

Chapter 3

3. Synthesis

The synthetic steps of all the peptides synthesized are summarized in the flow charts presented in Figs. 3-1, 3-6, 3-9 and 3-10. Significant details of the synthetic procedures are described below with a number identifying the step in the flow chart.

Each peptide was synthesized by two different methods; the mixed anhydride method using N-methylmorpholine and isobutyl chloroformate and the active ester method using N-hydroxysuccinimide (NSU).

3.1. Synthesis of N^{α}tBocLeuHypGlyAlaLysGlySerOH (16)

This heptapeptide was synthesized by the fragment condensation of N^{α}tBocLeuHyPGlyONSU and TFA.AlaLysGlySerOH as described below (see flow chart in Figure 3-1).

3.1.1. Synthesis of N^{α}tBocLeuHypGlyONSU(6)

3.1.1.2. N^{α}tBocLeuHyPOMe (2)

N^{α}tBocLeuOH (5.00 g, 21.65 mmol) in 50 ml THF was cooled in a salt-ice mixture and equivalent molar quantities of N-methylmorpholine (2.38 ml) and isobutylchloroformate (2.81 ml) were added; the reaction mixture was stirred for 15 minutes. 4-hydroxy-L-proline methyl ester HCl (4.34 g, 23.82 mmol) and N-methylmorpholine (3.54 ml, 23.82 mmol) were added and the reaction mixture was stirred for 24 hours; the temperature was allowed to come to ambient. The solution was evaporated to dryness on a rotary evaporator, the dry residue was redissolved in ethyl acetate and washed successively with cold 2N HCl, saturated sodium bicarbonate, and saturated NaCl. The organic phase was dried with

anhydrous sodium sulfate, filtered and concentrated on a rotary evaporator. The material was crystallized from hot ethyl acetate and gave 6.50 g of product (18.16 mmol, 83.9%). Analysis of the product by HPLC gave a single peak with a small shoulder.

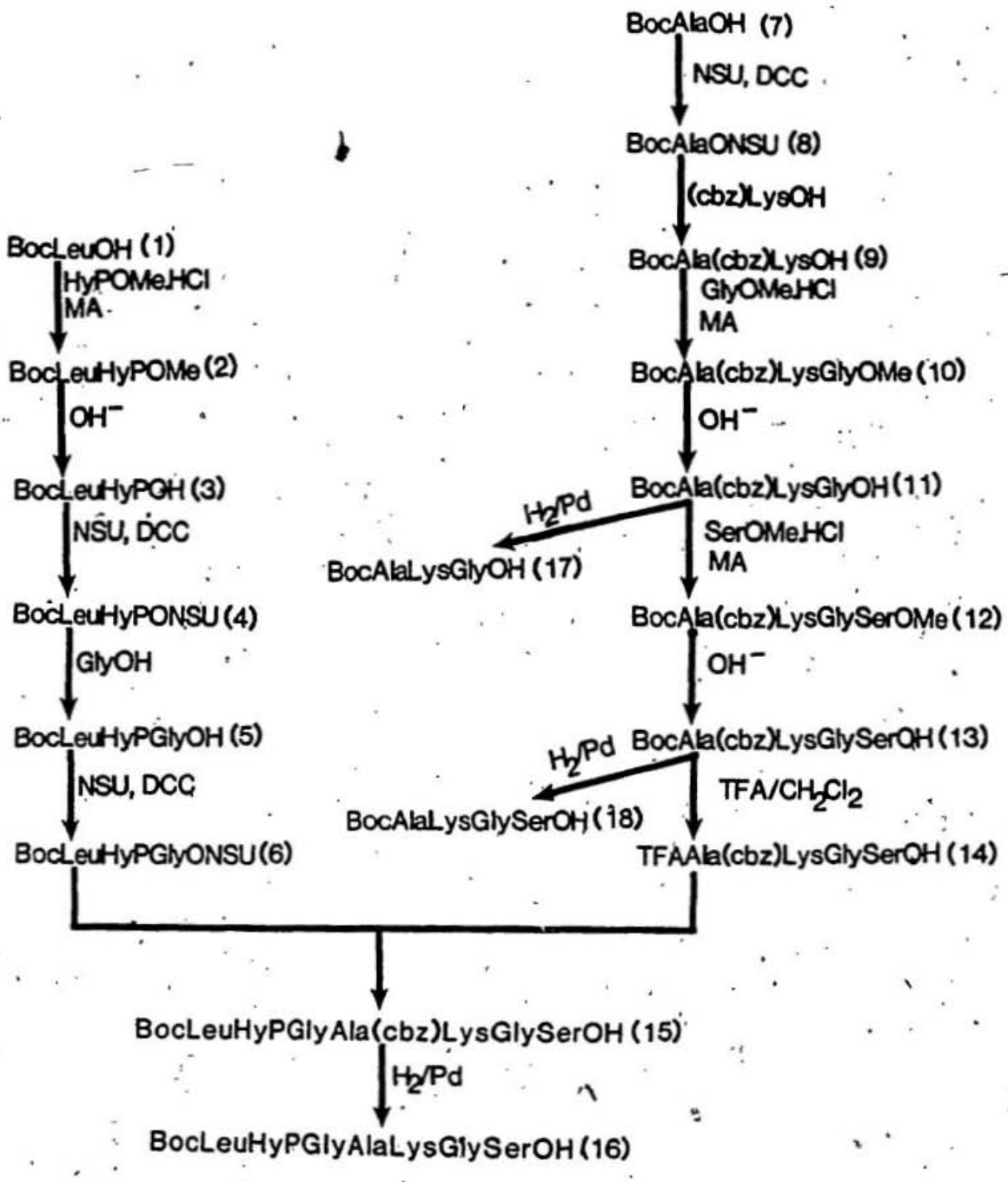

FIG. 3-1: Flow chart summarizing the synthesis of N^{α}tBocAlaLysGlyOH, N^{α}tBocAlaLysGlySerOH and N^{α}tBocLeuHyPGlyAlaLysGlySerOH

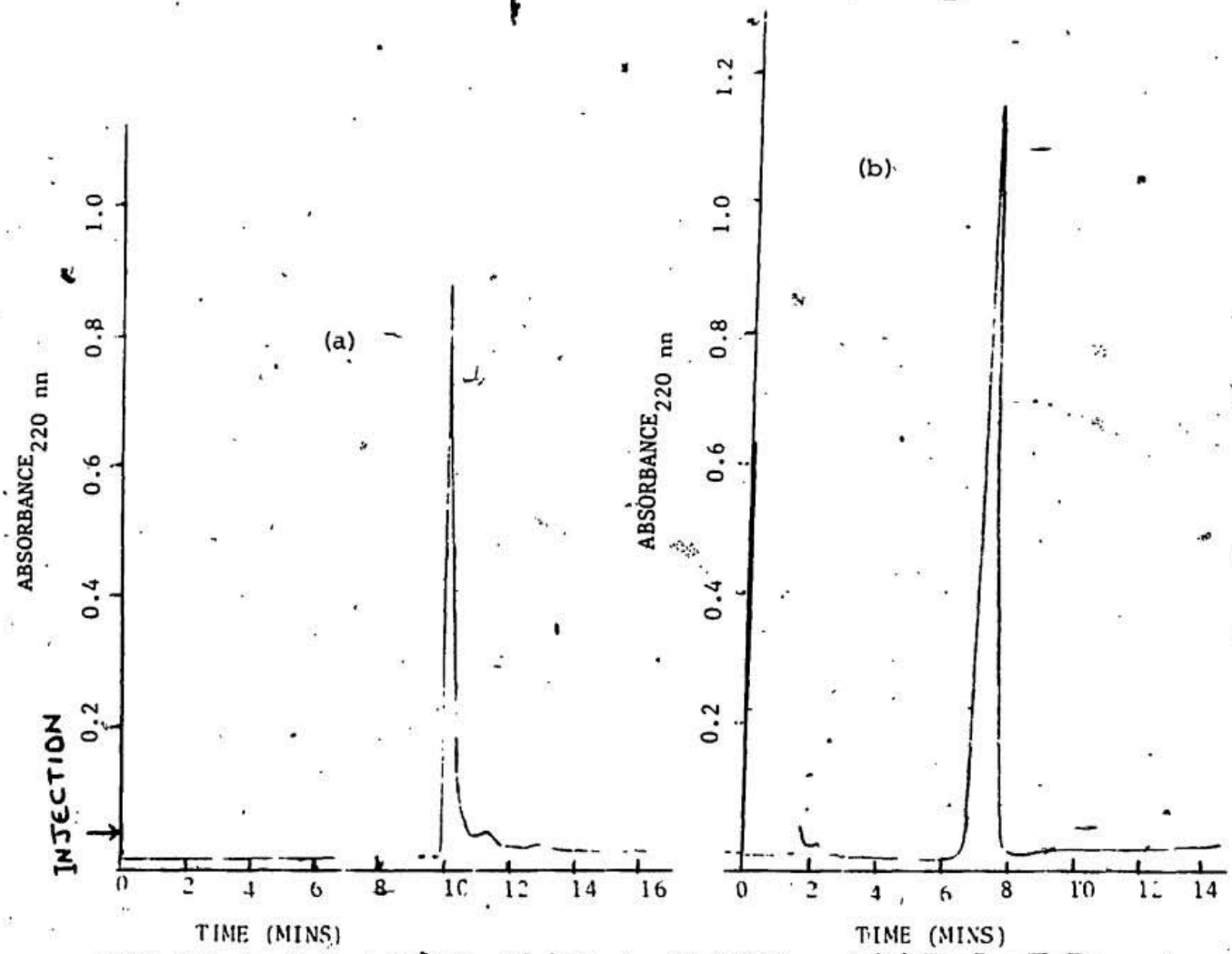

FIG. 3-2: Analytical HPLC of (a) BocLeuHyPOMe and (b) BocLeuHyP-GlyOH

3.1.1.2. N^{α}tBocLeuHypOH (3)

N^{α}tBocLeuHypOMe (6.30 g, 17.60 mmol) was dissolved in MeOH (8.80 ml) and the solution was stirred at room temperature for 5 minutes. One molar equivalent of 2N NaOH (8.80 ml, 17.60 mmol) was added, and the reaction mixture was stirred for 24 hours. The solvent was evaporated and the dry residue was dissolved in 100 ml water, and the solution was extracted with 50 ml ethyl acetate. The aqueous layer was concentrated to 50 ml and adjusted to pH 2. The solution was slowly saturated with NaCl and then extracted twice with chloroform. The chloroform layers were dried with anhydrous sodium sulfate, filtered and concentrated on a rotary evaporator. The material was crystallized from hot ethyl acetate and gave 4.05 g of product (11.77 mmol, 66.9%). Analysis of the product by HPLC gave a single peak.

3.1.1.3. N^{α}tBocLeuHypONSU (4)

A solution containing equimolar amounts of N^{α}tBocLeuHypOH (4.05 g, 11.77 mmol) and NSU (1.35 g, 11.77 mmol) in dioxane containing 25% THF was cooled to $0^{o}C$ in an ice-bath. DCC (2.43 g, 11.77 mmol) was added and the mixture was stirred for 18 Hours at 0-5°C. The reaction mixture was filtered to remove the DCU by-product, and the solvent was removed on a rotary evaporator. The material was crystallized from a chloroform-petroleum ether mixture and gave 5.20 g of product (10.44 mmol, 99.6%).

3.1.1.4. N^{α}tBocLeuHypGlyOH (5).

N^{α}tBocLeuHypONSU (5.00 g, 11.34 mmol) was dissolved in 75 ml dioxane. Glycine (0.94 g, 12.47 mmol) and sodium bicarbonate (1.05 g, 12.47 mmol) dissolved in 75 ml water were added to the first solution, and the reaction mixture was stirred for 36 hours at room temperature. The solvent was evaporated and the dry residue was dissolved in 100 ml water and the solution was extracted

with chloroform. The aqueous layer was concentrated to 50 ml and adjusted to pH 2. The solution was slowly saturated with NaCl and then extracted twice with ethyl acetate. The ethyl acetate layers were dried over anhydrous sodium sulfate, filtered and concentrated on a rotary evaporator. The material was crystallized from an ethyl acetate-petroleum ether mixture and gave 11.00 g of product (24.39 mmol, 46.8%). The product was analyzed by HPLC and showed a single peak (retention time, 6.4 minutes) (see Figure 3-2 (b).

MP, 89-91°C; *Anal.* calcd. for $C_{22}H_{33}O_7N_3$: C 53.87 H 7.73 O 27.93 N 10.47; found: C 53.78 H 7.61 O 28.10 N 10.51.

3.1.1.5. N^{α}tBocLeuHypGlyONSU (6)

A solution containing equimolar amounts of N^{α}tBocLeuHypGlyOH (4.20 g, 10.47 mmol) and NSU (1.21 g, 10.47 mmol) in dioxane containing 25% THF was cooled to 0°C in an ice-bath. DCC (2.16 g, 10.47 mmol) was added, and the mixture was stirred for 24 hours at 0-5°C. The reaction mixture was filtered to remove the DCU by-product, and the solvent removed on a rotary evaporator. The material was crystallized from a chloroform-petroleum ether mixture and gave 5.20 g of product (10.44 mmol, 99%).

3.1.2. Synthesis of TFA.Ala(N^{ϵ}cbz)LysGlySerOH (14)

3.1.2.1. N^{α}tBocAlaONSU (8)

A solution containing equimolar amounts of N^{α}tBocAlaOH (10.00 g, 52.92 mmol) and NSU (6.08 g, 52.92 mmol) in dioxane containing 25% THF was cooled to 0°C in an ice-bath. DCC (10.00 g, 52.92 mmol) was added, and the mixture was stirred for 14 hours at 0-5°C. The reaction mixture was filtered to remove the DCU by-product, and the solvent was removed on a rotary evaporator. The material was crystallized from a chloroform-petroleum ether mixture and gave 14.90 g of product (52.10 mmol, 98.5%).

3.1.2.2. N^{α}tBocAla(N^{ϵ}cbz)LysOH (9)

N^{α}tBocAlaONSU (14.90 g, 52.10 mmol) was dissolved in dioxane containing 25% THF. N^{ϵ}cbzLysOH (16.06 g, 57.31 mmol) and sodium bicarbonate (4.82 g, 57.31 mmol) dissolved in 75 ml water were added to the first solution, and the reaction mixture was stirred for 60 hours at room temperature. The solvent was evaporated and the dry residue was dissolved in 100 ml water and the solution was extracted with ethyl acetate. The aqueous layer was concentrated to 50 ml and adjusted to pH 2. The solution was slowly saturated with NaCl and then extracted twice with ethyl acetate. The ethyl acetate layers were washed with saturated NaCl dried over anhydrous sodium sulfate, filtered and concentrated on a rotary evaporator. The material was crystallized from a chloroform-petroleum ether mixture and gave 11.00 g of product (24.39 mmol, 46.8%). *Anal.* calcd. for $C_{22}H_{33}O_7N_3$: C 58.54 H 7.32 O 24.83 N 9.31; found: C 58.46 H 7.29 O 25.14 N 9.11 .

3.1.2.3. N^{α}tBocAla(N^{ϵ}cbz)LysGlyOMe (10)

N^{α}tBocAla(N^{ϵ}cbz)LysOH (11.00g. 24.39 mmol) in 100 ml was cooled in a salt-ice mixture while stirring. One equivalent molar quantities of N-methylmorpholine (2.68 ml) and isobutylchloroformate (3.16 ml) were added, and the reaction mixture was stirred for 15 minutes. Glycine methyl ester HCl (3.37 g, 26.83 mmol) and N-methylmorpholine (2.95 ml, 26.83 mmol) were added, and the reaction mixture was stirred for 48 hours and the temperature was allowed to come ambient. The solution was evaporated to dryness on a rotary evaporator, the dry residue redissolved in ethyl acetate and washed successively with cold 2N HCl, saturated sodium bicarbonate and saturated NaCl. The organic phase was dried with anhydrous sodium sulfate, filtered and concentrated. The material was crystallized from a chloroform-petroleum ether mixture and gave 7.76 g of

product (14.87mmol, 61%). Further purification of the product by chromatography on silica gel afforded a pure product (7.50 g, 14.76 mmol, 96.6%). *Anal.* calcd. for $C_{25}H_{38}O_8N_4$: C 57.46 H 7.28 O 24.52 N 10.73; found: C 57.36 H 7.25 O 24.74 N 10.65.

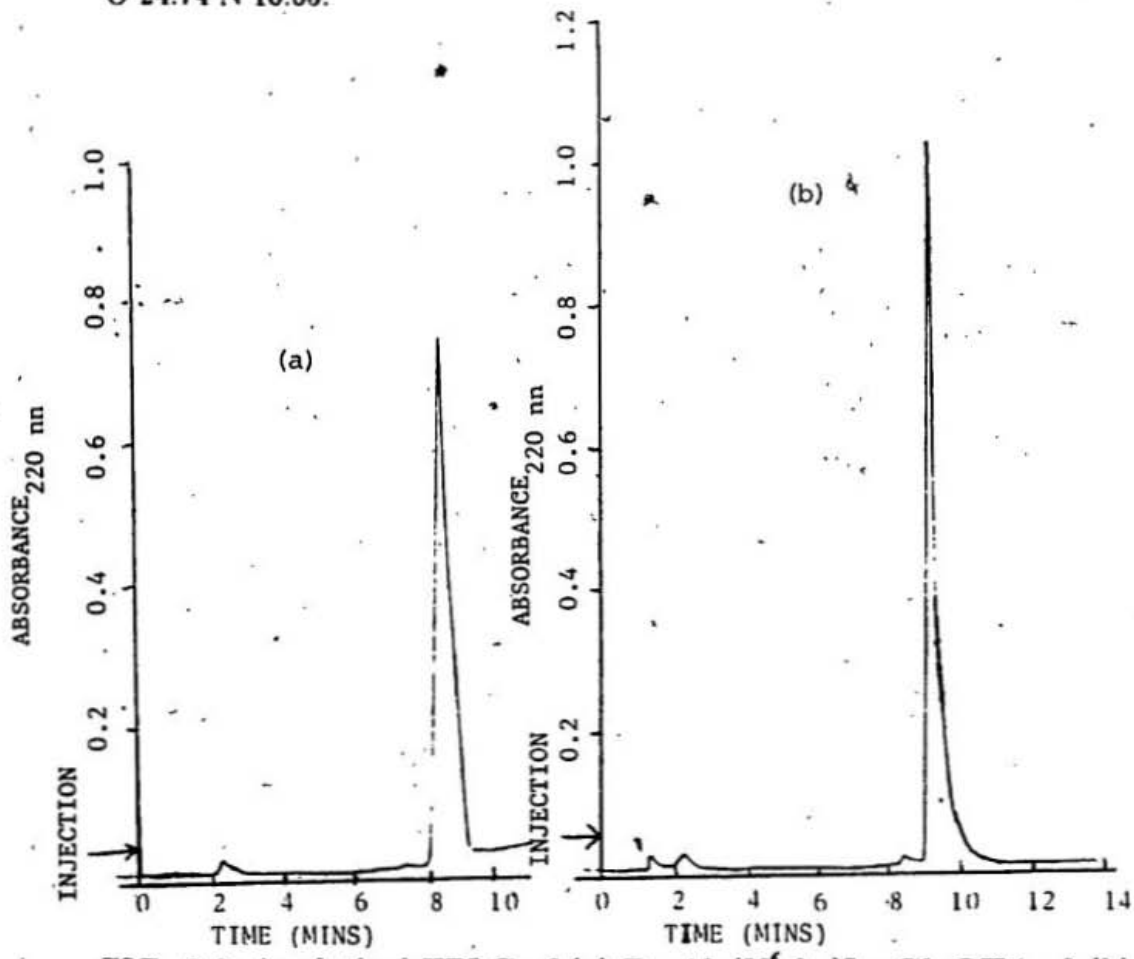

FIG. 3-3: Analytical HPLC of (a) BocAla(N$^{\epsilon}$cbz)LysGlyOH and (b) BocAlaLysGlyOH

3.1.2.4. N$^{\alpha}$tBocAla(N$^{\epsilon}$cbz)LysGlyOH (11)

N$^{\alpha}$tBocAla(N$^{\epsilon}$cbz)LysGlyOMe (7.70 g, 14.75 mmol) was dissolved in MeOH (7.38 ml) and the solution was stirred at room temperature for 5 minutes. One molar equivalent of 2N NaOH (7.38 ml) was added, and the reaction mixture was

stirred for 48 hours. The solvent was evaporated and the dry residue was dissolved in 100 ml water, and the solution was extracted with 50 ml chloroform. The aqueous layer was concentrated to 50 ml and adjusted to pH 2. The solution was slowly saturated with NaCl and then extracted twice with ethyl acetate. The ethyl acetate layers were dried with anhydrous sodium sulfate, filtered and concentrated on a rotary evaporator. The material was crystallized from a chloroform-petroleum ether and gave 7.27 g of product (14.31 mmol, 97.1%). The product was analyzed by HPLC (retention time, 8 minutes) (see Figure 3-3 (a)). *Anal.* calcd. for $C_{24}H_{36}O_8N_4$: C 56.69 H 7.09 O 25.20 N 11.02; found: C 56.62 H 7.52 O 24.71 N 11.15.

3.1.2.5. N^{α}tBocAlaLysGlyOH (18)

To a solution of N^{α}tBocAla(N^{ϵ}cbz)LysGlyOH (1.50 g, 2.95 mmol) in 25 ml MeOH was added 150 mg of palladium (10%) on carbon suspended in 20 ml MeOH. The reaction mixture was hydrogenated at room temperature and 1 atmosphere with stirring for 24 hours. The reaction was filtered to remove the catalyst, and the solvent was removed on a rotary evaporator. The residue was dissolved in MeOH and diethyl ether added to precipitate the product. The product was filtered and dried in vacuo (0.89 g, 2.38 mmol, 80.9%). The product was analyzed by HPLC (retention time, 9.1 minutes) (Figure 3-3 (b)).

MP, 115°C (decomp). *Anal.* calcd. for $C_{16}H_{30}O_6N_4$: C 51.34 H 8.02 N 14.97 O 25.67; found: C 51.18 H 7.92 O 26.14 N 14.76

3.1.2.6. N^{α}tBocAla(N^{ϵ}cbz)LysGlySerOMe (12)

N^{α}tBocAla(N^{ϵ}cbz)LysGlyOH (5.77 g, 11.36 mmol) in 75 ml THF was cooled in a salt-ice mixture and equivalent molar quantities of N-methlymorpholine (1.25 ml) and isobutylchloroformate (1.47 ml) were added, and the reaction mixture was stirred for 15 minutes. Serine Methyl Ester HCl (1.94 g, 12.50 mmol) and N-

methylmorpholine (1.37 ml, 12.50 mmol) were added and the reaction mixture was stirred for 48 hours and the temperature was allowed to come to ambient. The solvent was evaporated and the dry residue was redissolved in ethyl acetate and washed successively with 2N HCl, saturated sodium bicarbonate and saturated NaCl. The organic phase was dried with anhydrous sodium sulfate, filtered and concentrated. The material was crystallized from an ethyl acetate-petroleum ether mixture and gave 6.27 g of product (10.30 mmol, 90%). The product was further purified by HPLC give 5.00 g of final pure product (8.21 mmol, 79.7%). *Anal.* calcd. for $C_{28}H_{43}O_{10}N_5$: C 55.17 H 7.06 O 26.27 N 11.49; found: C 55.04 H 7.08 O 26.29 N 11.59.

3.1.2.7. N^{α}tBocAla(N^{ϵ}cbz)LysGlySerOH (13)

N^{α}tBocAla(N^{ϵ}cbz)LsyGlySerOMe (6.00 g, 9.85 mmol) was dissolved in MeOH (4.95 ml) and the solution was stirred at room temperature for 5 minutes. One molar equivalent of 2N NaOH (4.93 ml) was added, and the reaction mixture was stirred for 48 hours. The solvent was evaporated, and the dry residue was dissolved in 100 ml water, and the solution was extracted with 50 ml ethyl acetate. The aqueous layer was concentrated to 50 ml and adjusted to pH 2. The solution was slowly saturated with NaCl and then extracted twice with ethyl acetate. The ethyl acetate layers were dried with anhydrous sodium sulfate, filtered and concentrated on a rotary evaporator. The material was crystallized from a chloroform-petroleum ether mixture and gave 4.72 g, of product (7.95 mmol, 80.5%). Analysis of the product by HPLC gave a single peak (retention time, 8.2 minutes) (see Figure 3-4 (a)).

3.1.2.8. N^{α}tBocAlaLysGlySerOH (19)

To a solution of N^{α}tBocAla(N^{ϵ}cbz)LysGlySerOH (1.39 g, 2.33 mmol) in 20 ml MeOH was added 150 mg of palladium (10%) on carbon suspended in 20 ml

MeOH. The stirred reaction was hydrogenated at room temperature and 1 atmosphere for 24 hours. The catalyst was removed by filtration, and the solvent was removed on a rotary evaporator. The residue was dissolved in MeOH, and diethyl ether was added to precipitate the material. The material was filtered and dried in vacuo (0.79 g, 1.72 mmol, 73.8%). The material was further purified by HPLC to give 0.59 g of product (1.28 mmol, 74.7%). The product was analyzed by HPLC and gave a single peak (retention time, 9 minutes (see Figure 3-3 (b)).

MP, 115 °C (decomp.). *Anal.* calcd. for $C_{19}H_{35}O_8N_5$: C 49.45 H 7.59 N 15.18 O 27.76; found: C 49.43 H 8.29 N 15.20 O 27.08.

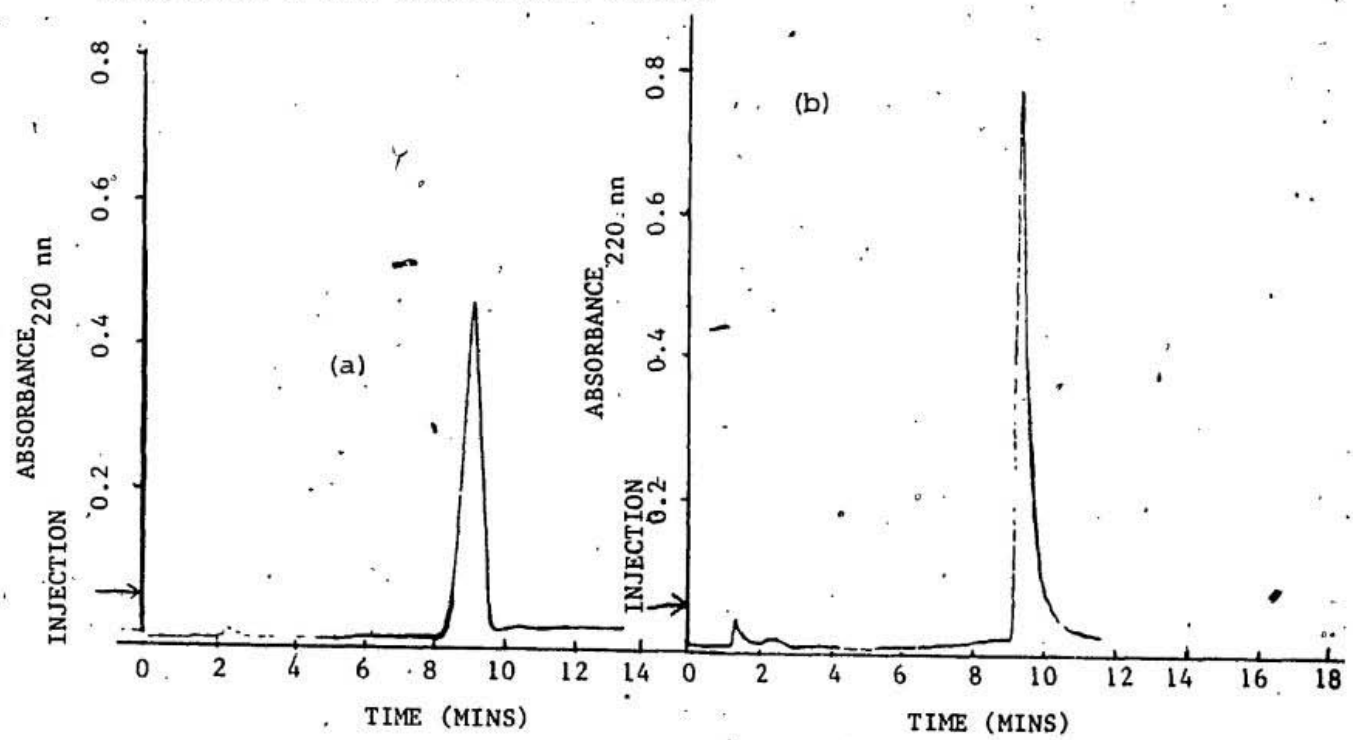

FIG. 3-4: Analytical HPLC of (a) BocAla(N^{ϵ}cbz)LysGlySerOH and (b) BocAlaLysGlySerOH

3.1.2.9. TFA.Ala(N^{ϵ}cbz)LysGlySerOH (14)

A solution of N^{α}tBocAla(N^{ϵ}cbz)LysGlySerOH (4.00 g 6.72 mmol) in 30 ml of 30% TFA in methylene chloride was stirred at 0-5°C for 30 minutes and then at room temperature for 1 hour and 30 minutes. The solvent was evaporated in a rotary evaporator. The residue was triturated with diethyl ether and dried in vacuo to give a white powder (4.04 g, 6.63 mmol). Analysis of the product by HPLC gave a single product.

3.1.3. N^{α}tBocLeuHypGlyAla(N^{ϵ}cbz)LysGlySerOH (15)

N^{α}tBocLeuHypGlyONSU (3.00 g, 6.02 mmol) was dissolved in 50 ml dioxane containing 50% THF. TFAAla(N^{ϵ}cbz)LysGlySerOH (4.04 g, 6.63 mmol) and sodium bicarbonate (0.56 g, 6.63 mmol) were added to the first solution and the reaction mixture was stirred for 36 hours at room temperature. The solvent was evaporated and the dry residue was dissolved in 100 ml water, and the solution was extracted with chloroform. The aqueous layer was concentrated to 50 ml and adjusted to pH 2. The solution was slowly saturated with NaCl, and then extracted three times with ethyl acetate. The ethyl acetate layers were dried with anhydrous sodium sulfate, filtered, evaporated to dryness in a rotary evaporator and gave 2.39 g of product (2.71 mmol, 46.8%). Analysis of the product by HPLC gave two major peaks, but it could not be purified because of solubility problems. The product was used in the following steps without further purification.

3.1.4. N^{α}tBocLeuHypGlyAlaLysGlySerOH (16)

N^{α}tBocLeuHypGlyAla(N^{ϵ}cbz)LysGlySerOH (2.2 g, 3.08 mmol) was suspended in 25 ml MeOH and hydrogenated at room temperature, 1 atmosphere, using 10% palladium on carbon (250 mg). The catalyst was removed by filtration, and the solvent was removed with a rotary evaporator. The residue was redissolved in MeOH, and diethyl ether was added to precipitate the material. The

material was filtered and dried in vacuo (0.59 g, 0.79 mmol, 45.8%). The material was further purified by HPLC and gave 0.33 g of pure product (0.44 mmol, 55.3%). The purity of the product was checked by analytical HPLC (retention time , 8.4 minutes) (see Figure 3-5). MP, 145-146°C. *Anal.* calcd. for $C_{31}H_{54}O_{11}N_8$: C 51.41 H 7.90 O 25.70 N 14.99; found: C 51.11 H7.67 O 25.88 N 15.34

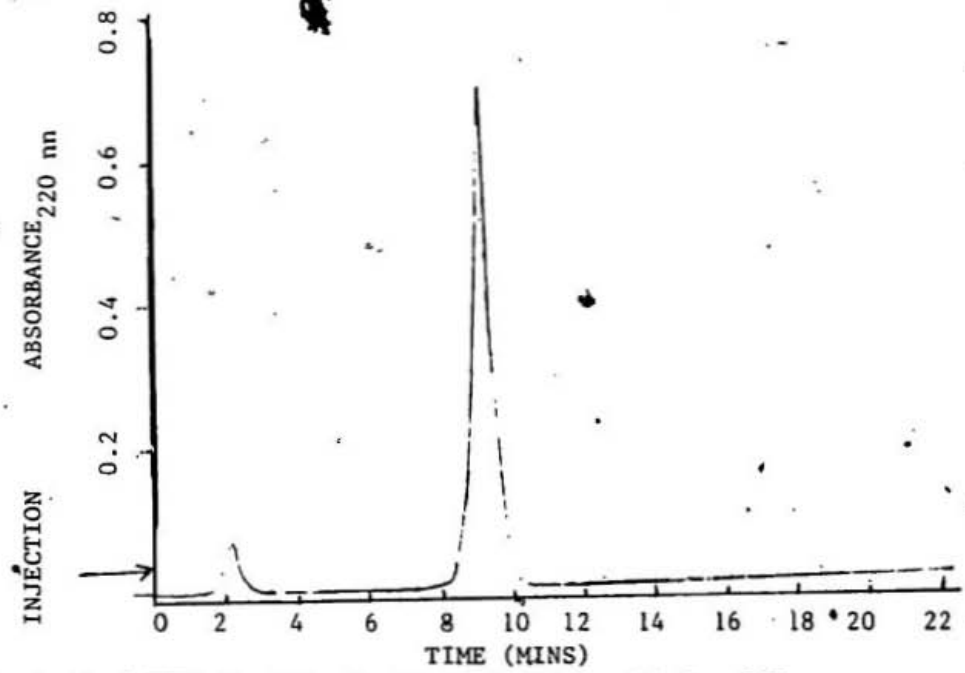

FIG. 3-5: Analytical HPLC of BocLeuHyPGlyAlaLysGlySerOH

3.2. Synthesis of N^{α}tBocAlaArgGlyIleLysGlyOH (32)

3.2.1. Synthesis of N^{α}tBocAla(N^{ω}nitro)ArgGlyONSU (24)

3.2.1.1. N^{α}tBocAla(N^{ω}nitro)ArgOMe (20)

This was obtained from N^{α}tBocAlaOH (5.00 g, 26.43 mmol) and (N^{ω}nitro) arginine methyl ester HCl (7.84 g, 29.07 mmol) by the mixed anhydride method as described for the preparation of (2). The material was crystallized from an ethyl acetate-petroleum ether mixture and gave 10.10 g of product (24.88 mmol, 94%). Analysis of the product by HPLC gave a single peak.

3.2.1.2. N^{α}tBocAla(N^{ω}nitro)ArgOH (21)

To a solution of N^{α}tBocAla(N^{ω}nitro)ArgOMe HCl (10.00 g, 24.63 mmol) in MeOH (12.32 ml) was added 2N NaOH (12.32 ml), 24.63 mmol) and the reaction mixture was stirred for 24 hours. The reaction was treated in the same manner as described for the preparation of (5). The material was crystallized from an ethyl acetate-petroleum ether mixture and gave 5.64g of product (14.39mmol, 58.4%). The product was analyzed by HPLC (retention time, 6.8 minutes)

3.2.1.3. N^{α}tBocAla(N^{ω}nitro)ArgGlyOMe (22)

This was obtained from N^{α}tBocAla(N^{ω}nitro)ArgOH (5.40 g, 13.78 mmol) and glycine methyl ester HCl (1.90 g, 15.16 mmol) by the mixed anhydride method as described for the preparation of (2). The material was crystallized from an ethyl acetate-petroleum ether mixture and gave 3.25 g of product (7.03 mmol, 51.1%). The product was further purified by column chromatography on silica gel to give 3.00 g of fine white crystals (6.49 mmol, 92.3%). Analysis of the product by HPLC gave a single peak. *Anal.* calcd. for $C_{17}H_{32}O_8N_7$: C 44.16 H 6.93 O 27.71 N 21.21; found: C 44.10 H 6.76 O 27.66 N 21.40.

3.2.1.4. N^{α}tBocAla(N^{ω}nitro)ArgGlyOH (23)

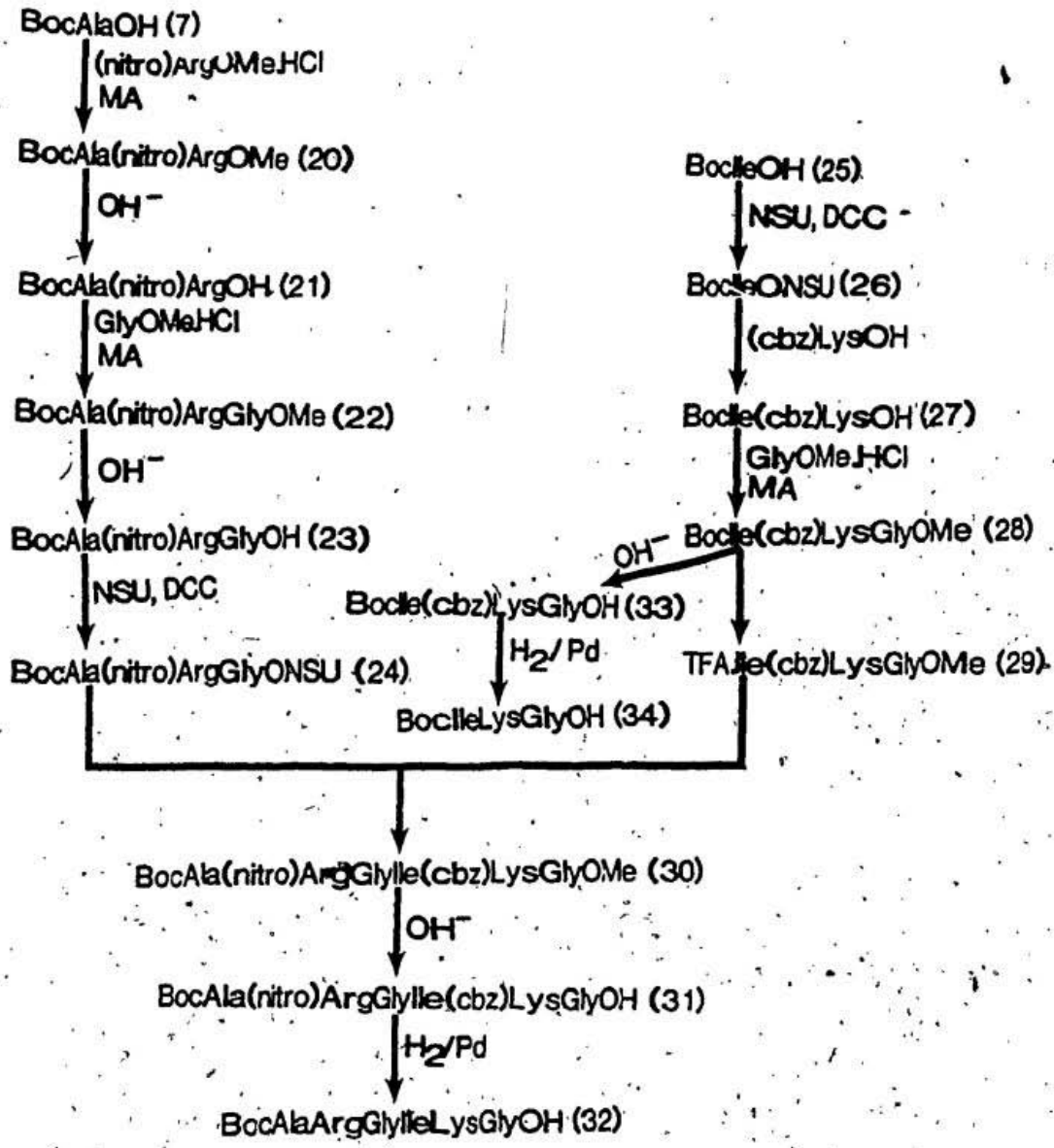

FIG. 3-6: Flow chart summarizing synthesis of N^{α}tBocIleLysGlyOH (34) and N^{α}tBocAlaArgGlyIleLysGlyOH (32)

To a solution of N^{α}tBocAla(N^{ω}nitro)ArgGlyOMe (3.00 g, 6.49 mmol) in MeOH (3.25 ml) was added 2N NaOH (3.25 ml, 6.49 mmol) and the reaction mixture was stirred for 24 hours. The reaction mixture was treated in the same way as described for the preparation of (2). The material was crystallized from a chloroform-petroleum ether mixture and gave 2.76 g of product (6.16 mmol, 95.1%).

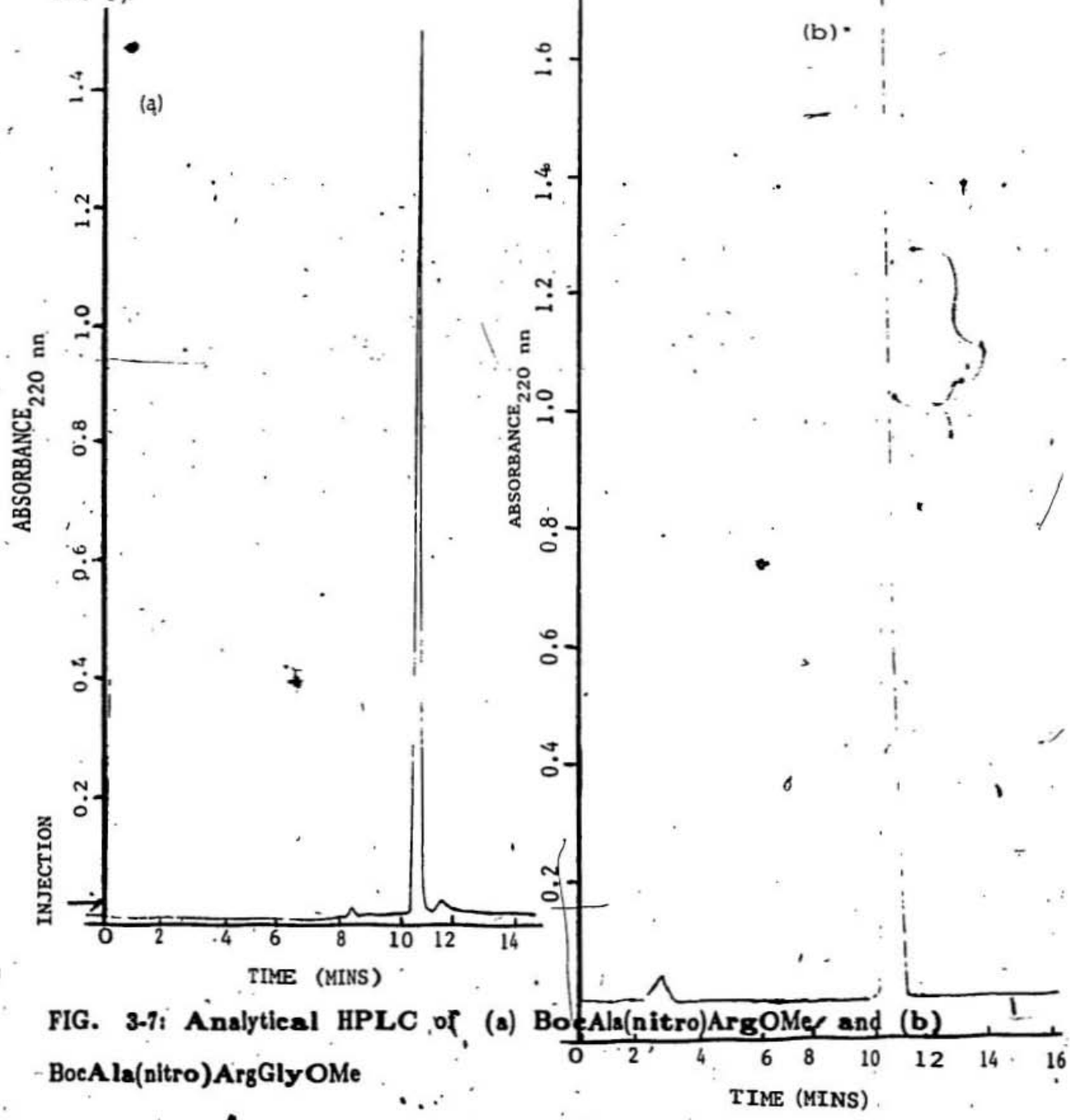

FIG. 3-7: Analytical HPLC of (a) BocAla(nitro)ArgOMe and (b) BocAla(nitro)ArgGlyOMe

3.2.1.5. N^{α}tBocAla(N^{ω}nitro)ArgGlyONSU (24)

This was obtained from N^{α}tBocAla(N^{ω}nitro)ArgGlyOH (4.00 g, 10.86mmol) and NSU (1.25 g, 10.86 mmol) in the same manner as described for the preparation of (4). The material was crystallized from an ethyl acetate-petroleum ether mixture and gave (8.81 mmol, 98.8%).

3.2.2. Synthesis of N^{α}tBocIle(N^{ϵ}cbz)LysGlyOMe (28)

3.2.2.1. N^{α}tBocIleONSU (26)

This was obtained from N^{α}tBocIleOH (5.00 g, 21.65 mmol) and NSU (2.49 g, 21.65 mmol) in the same manner as described for the preparation of (4). The material was crystallized from a dichloromethane-petroleum ether mixture and gave 5.53 g of product (16.86 mmol, 77.9%).

3.2.2.2. N^{α}tBocIle(N^{α}cbz)LysOH (27)

This was obtained from N^{α}tBocIleONSU (7.00 g, 21.34 mmol) and N^{ϵ}cbzLysOH (6.58 g, 23.48mmol) in the same manner as described for the preparation of (5). The material was dried in vacuo and gave 3.02 g of product (6.13mmol, 28.7%).

3.2.2.3. N^{α}tBocIle(N^{ϵ}cbz)LysGlyOMe (28)

This was obtained from N^{α}tBocIle(N^{ϵ}cbz)LysOH (3.00 g, 6.09mmol) and Glycine methyl ester HCl (0.84 g, 6.70 mmol) by the mixed anhydride method as described for the preparation of (2). The material was dried in vacuo and was further purified on a silica gel column to give 2.27 g of product (4.03 mmol, 66.2%). Analysis of the product by HPLC gave one major peak and a smaller one. The product was used in the following steps without further purification. *Anal.* calcd. for $C_{28}H_{44}O_8N_4$: C 59.57 H 7.80 O 22.70 N 9.92; found: C 59.53 H 8.10 O 22.35 N 9.59.

3.2.2.4. N^{α}tBocIleLyusGlyOH (34)

To a solution of N^{α}tBocIleLysGlyOH (0.6 g, 1.09 mmol) in 10 ml MeOH was added 80 mg of palladium (10%) on carbon suspended in 10 ml MeOH. The reaction was hydrogenated at room temperature and 1 atmosphere with stirring for 24 hours. The reaction mixture was filtered to remove the catalyst, and the solvent was removed on a rotary evaporator. The residue was dissolved in MeOH and diethyl ether added to precipitate the product. The product was filtered and dried in vacuo (0.33 g, 0.79 mmol, 73.3%). Purity of the peptide was checked by analytical HPLC (retention time, 8.6 minutes) (see Figure 3-8 (a)).

MP, 115°C (decomp.). *Anal.* calcd. for $C_{19}H_{36}O_6N_4$: C 54.81 H 8.65 O 23.08 N 13.46; found: C 54.53 H 8.47 O23.13 N 12.97

3.2.2.5. TFA.Ile(N^{ϵ}cbz)LysGlyOMe (29)

A solution of N^{α}tBocIle(N^{ϵ}cbz)LysGlyOMe (3.80 g, 6.74 mmol) in 30 ml of 30% TFA in methylene chloride was stirred at 0-5°C for 30 minutes and then at room temperature for one hour and 30 minutes. The reaction mixture was treated in the same manner as described for the preparation of (14) to give 3.85 g of product (6.70 mmol, 99%).

3.2.3. N^{α}tBocAla(N^{ω}nitro)ArgGlyIle(N^{ϵ}cbz)LysGlyOMe (30)

This was obtained from N^{α}tBocAla(N^{ω}nitro)ArgGlyONSU (3.5 g, 6.42 mmol) and TFA.Ile(N^{ϵ}cbz)LysGlyOMe (3.71 g, 6.42 mmol) in the same manner as described for the preparation of (5). The material was crystallized from a chloroform-petroleum ether mixture and gave 3.20 g, of impure product (3.58mmol, 55.8%). The impure product was purified by chromatography on a silica gel column and gave 2.15 g of product (2.41 mmol, 71.7%).

3.2.4. N^{α}tBocAla(N^{ω}nitro)ArgGlyIle(N^{ϵ}cbz)GlyOH (31)

To a solution of N^{α}tBocAla(N^{ω}nitro)ArgGlyIle(N^{ϵ}cbz)LysGlyOMe (2.00 g, 2.24 mmol) in MeOH (1.12 ml) was added 2N NaOH (1.12 ml 2.24 mmol) and the reaction mixture treated in the same manner as described for the preparation of (3). The material was dried in vacuo and gave 1.55 g of product (1.76 mmol, 78.7%). The product was further purified by HPLC to give 0.90 g of fine white crystals (1.02 mmol, 40.7%). The purity of the peptide was checked by HPLC (retention time, 6.5 minutes) (see Figure 3-8 (b)).

MP, 135°C (decomp.).

3.2.2.7. N^{α}AlaArgGlyIleLysGlyOH (32)

N^{α}tBocAla(N^{ω}nitro)ArgGlyIle(N^{ϵ}cbz)LysGlyOH (1.5 g, 1.73 mmol) was dissolved in MeOH (20 ml) and hydrogenated at room temperature, 1 atmosphere, using 10% palladium on carbon (150mg). The reaction was treated in the same manner as described for the preparation of (16) to yield 0.79 g of material (1.13 mmol, 66.4%). The material was further purified by HPLC to give 0.27 g of fine shiny crystals (0.27 g, 34.2%). *Anan.* calcd. for $C_{30}H_{57}O_9N_{10}$: C 51.43 H 8.14 O 20.57 N 20.00; found: C 51.43 H 8.14 O 20.57 N 20.00

3.3. Synthesis of N^{α}tBocAlaGlyLysOH(36) and N^{α}tBocAlaGlyLysHyPOH(39)

Figure 3-9, summarizes the syntheses of these peptides.

3.3.1. N^{α}tBocAlaGlyOH (33)

This was obtained from N^{α}tBocAlaONSU (4.5 g, 15.73 mmol) and Glycine (1.30 g, 17.30 mmol) in the same manner as described for the preparation of (4). The material was crystallized from an ethyl acetate-petroleum ether mixture and gave 1.80 g of product (7.32 mmol, 46.5%).

3.3.2. N^{α}tBocAlaGly(N^{ϵ}cbz)LysOMe (34)

This was obtained from N^{α}tBocAlaGlyOH (1.80 g, 7.32 mmol) and NcbzLysOMe Ester HCl (2.66 g, 8.05 mmol) by the mixed anhydride method as described for the preparation of (2). The material was evaporated in vacuo and further purified by chromatography on silica gel and gave 3.00 g of hygroscopic product (5.76 mmol, 78.7%).

3.3.3. N^{α}tBocAlaGly(N^{ϵ}cbz)LysOH (35)

To a solution of N^{α}BocAlaGly(N^{ϵ}cbz)LysOMe (3.00 g, 5.76 mmol) in MeOH (2.9 ml) was added 2N NaOH (2.9 ml, 5.80 mmol) and the solution was stirred for 36 hours. The reaction mixture was treated in the same manner as described for the preparation (3). The material was crystallized from an ethyl acetate-petroleum ether mixture and gave 2.72 g of product (5.36 mmol, 93.2%).

3.3.4. N^{α}tBocAlaGlyLysOH (36)

N^{α}tBocAlaGly(N^{ϵ}cbz)LysOH (1.50 g, 2.95 mmol) was dissolved in 25 ml MeOH and was hydrogenated at room temperature, 1 atmosphere, using 150 mg of palladium (10%) on carbon suspended in 20 ml MeOH. The reaction mixture was treated in the same manner as described for the preparation of (17) and gave 0.65 g of material (1.74 mmol, 58.9%) which was further purified by HPLC to give 0.5 g of product (1.34 mmol, 83.3%). Purity of the product was checked by HPLC (retention time, 8 minutes) (see Figure 3-10 (a)).

MP. 110°C (decomp). *Anal.* calcd. for $C_{16}H_{30}O_0N_4$: C 51.34 H 8.02 O 25.67 N 14.97; found: C 51.13 H 7.92 O 25.94 N 15.01

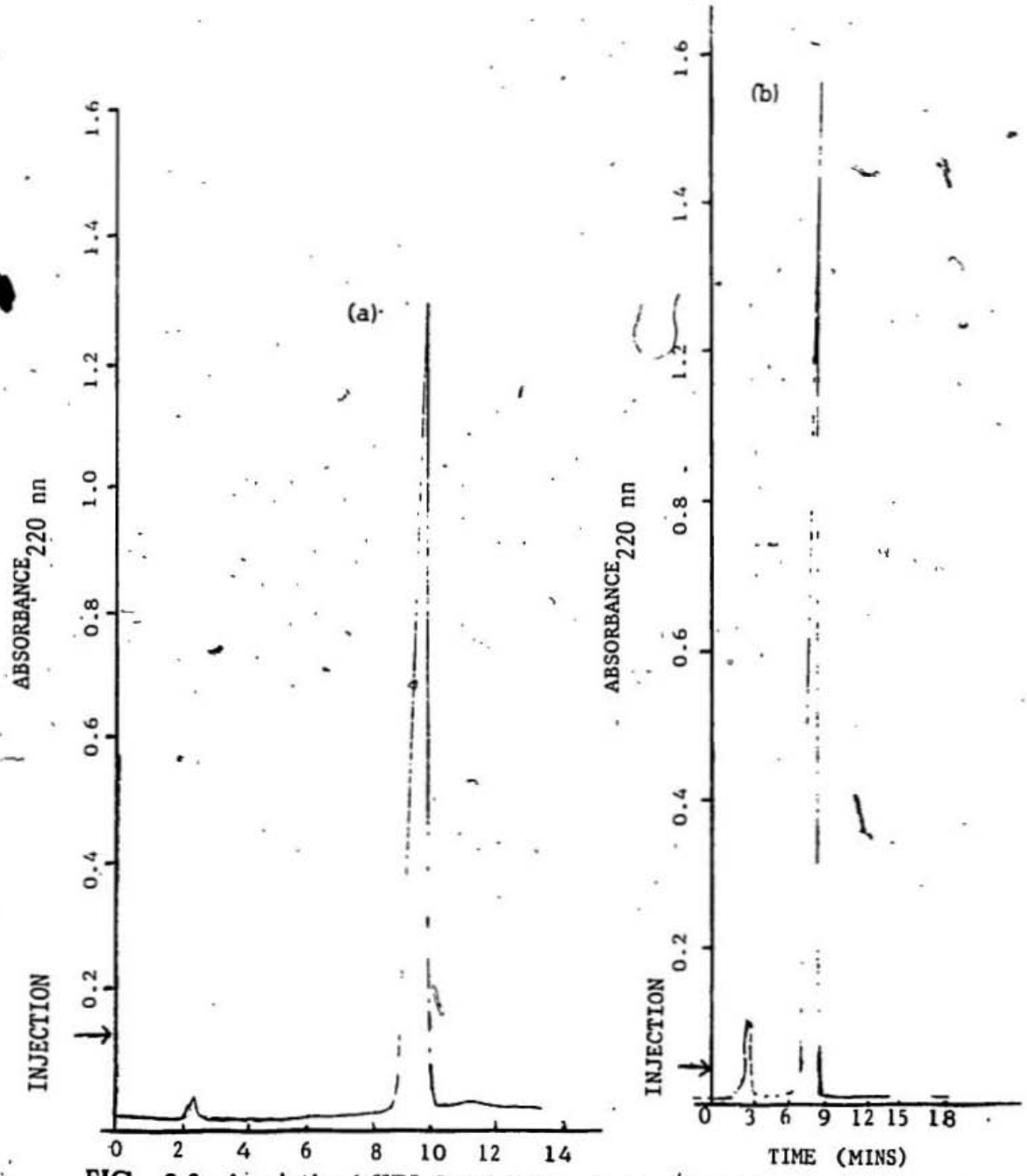

FIG. 3-8: Analytical HPLC of (a)BocIleLysGlyOH and (b) BocAlaArg-GlyIleLysGlyOH

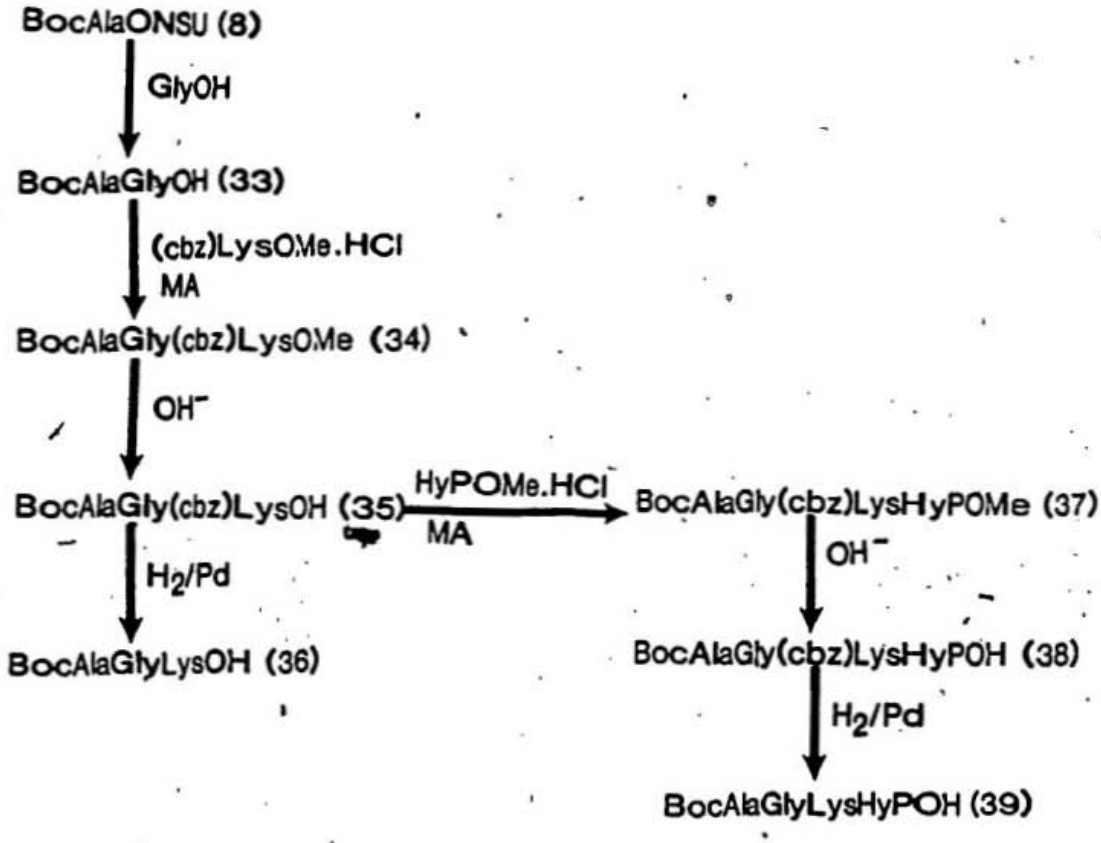

FIG. 8-9: Flow chart summarizing the synthesis of N^{α}tBocAlaGlyLysOH (36) and N^{α}tBocAlaGlyLysHyPOH (39)

3.3.5. Synthesis of N^{α}tBocAlaGly(N^{ϵ}cbz)LysHyPOMe (37)

This was obtained from N^{α}BocAlaGly(N^{ϵ}cbz)LysOH (1.20 g, 2.36 mmol) and HyPOMe HCl (0.47 g, 2.60 mmol) by the mixed anhydride method as described for the preparation of (2). The material was crystallized from an ethyl acetate-petroleum ether mixture and gave 1.14 g of product (1.80 mmol, 81.4%).

3.3.6. N^{α}tBocAlaGly(N^{ϵ}cbz)LysHyPOH (38)

To a solution of N^{α}BocAlaGly(N^{ϵ}cbz)LysHyPOMe (1.10 g, 1.73 mmol) in 0.87 ml MeOH was added 2N NaOH (0.87 ml, 1.74 mmol) and the solution was stirred for 36 hours. The reaction mixture was treated in the same as described for the preparation of (2). The material was dried in vacuo and gave 0.82 g of product (1.32 mmol, 75.9%).

3.3.7. N^{α}tBocAlaGlyLysHyPOH (39)

N^{α}tBocAlaGly(N^{ϵ}cbz)LysHyPOH (0.80 g, 1.29 mmol) was dissolved in 15 ml MeOH and was hydrogenated at room temperature, 1 atmosphere, using 80 mg palladium (10%) on carbon suspended in 10 ml MeOH. The reaction mixture was treated in the same manner as described for the preparation of (17) and gave 0.50 g of material (1.03 mmol, 79.4%) which was further purified by HPLC to give 0.35 g of product (0.72 mmol, 70%). Purity of the product was checked by analytical HPLC (retention time, 8 minutes) (see Figure 3-10 (b)).

MP, 115°C (decomp.).

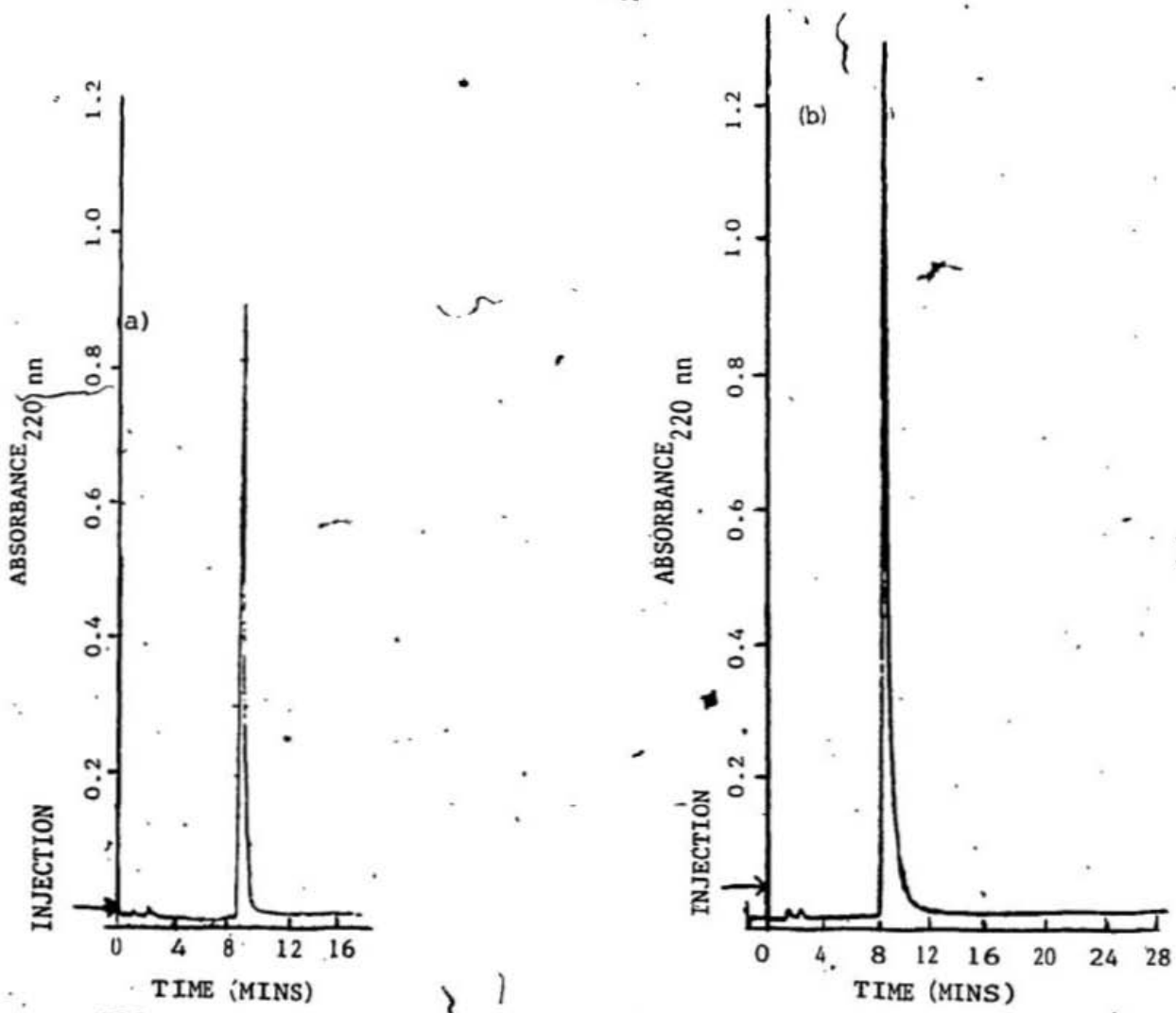

FIG. 3-10: Analytical HPLC of (a) BocAlaGlyLysOH and (b) BocAlaGlyLysHyPOH

3.4. Synthesis of N^{α}tBocLeuHyPGlyAlaLysGlyOH (45)

3.4.1. N^{α}tBocLeuHyPGlyAlaOH (40)

This was obtained from N^{α}tBocLeuHyPGlyONSU (2.58 g, 5.18 mmol) and Alanine (0.51 g, 5.70 mmol) in the same manner as described for the preparation of (5). The material was crystallized from an ethyl acetate-petroleum ether mixture and gave 1.82 g of product (3.86 mmol, 74.6%).

3.4.2. N^{α}tBocLeuHyPGlyAla(N^{ϵ}cbz)LysOMe (41).

This was obtained from N^{α}tBocLeuHyPGlyAlaOH (2.58 g, 5.18 mmol) and N^{ϵ}cbzLysOMe Ester HCl (1.39 g, 4.19 mmol) by the mixed anhydride method as described for the preparation of (2). The material was dried in vacuo and was further purified by chromatography on silica gel to give 2.52 g of product (3.37 mmol, 87.7%).

3.4.3. N^{α}tBocLeuHyPGlyAla(N^{ϵ}cbz)LysOH (42)

To a solution of N^{α}tBocLeuHyPGlyAla(N^{ϵ}cbz)LysOMe (2.50 g, 3.35 mmol) in MeOH (1.67 ml) was added 2N NaOH (1.67 ml, 3.34 mmol) and the solution was stirred for 36 hours. The reaction mixture was treated in the same manner as described for the preparation of (2). The material was crystallized from an ethyl acetate-petroleum mixture and gave 2.35 g of product (3.21 mmol, 95.9%).

3.4.4. N^{α}tBocLeuHyPAla(N^{ϵ}cbz)LysGlyOMe (43)

This was obtained from N^{α}tBocLeuHyPGlyAla(N^{ϵ}cbz)LysGlyOH (2.32 g, 3.17 mmol) and GlyOMe HCl (0.44 g, 3.49 mmol) by the mixed anhydride method as described for the preparation of (2). The material was dried in vacuo to yield 2.15 g of product (2.63 mmol, 83.0%), which was further purified by chromatography on silica gel and gave 1.52 g of product (1.86 mmol, 70.7%).

3.4.5. N^{α}tBocLeuHyPGlyAla(N^{ϵ}cbz)LysGlyOH (44)

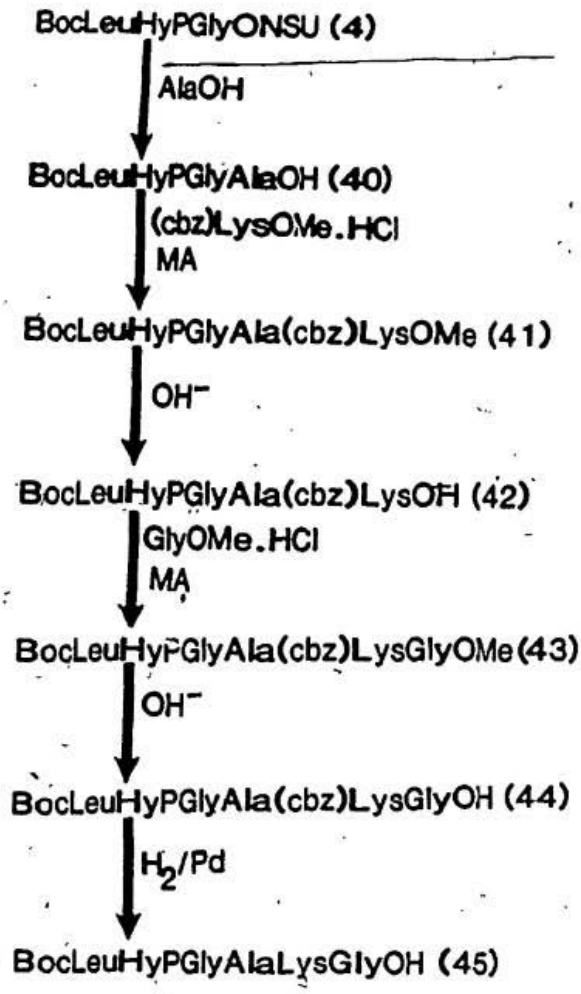

FIG. 3-11: Flow chart summarizing synthesis of N^{α}tBocLeuHypGlyAlaLysGlyOH (45)

To a solution of N^{α}tBocLeuHyPGlyAla(N^{ϵ}cbz)LysGlyOMe (1.50 g, 1.84 mmol) in MeOH (0.92 ml) was added 2N NaOH (0.92 ml, 1.84 mmol) and the solution was stirred for 36 hours. The reaction mixture was treated in the same manner as described for the preparation of (3). The material was crystallized from an ethyl acetate-petroleum ether mixture and gave 1.24 g of material (1.52 mmol, 83.8%) which was purified by chromatography on a silica gel column to give 1.00 g of product (1.22 mmol, 80.6%).

3.4.6. N^{α}tBocLeuHyPGlyAlaLysGlyOH (45)

N^{α}tBocLeuHyPGlyAla(N^{ϵ}cbz)LysGlyOH (1.00 g, 1.25 mmol) was dissolved in 20 ml MeOH and was hydrogenated at room temperature, 1 atmosphere using 100 mg of palladium (10%) on carbon suspended in 20 ml MeOH. The reaction mixture was treated in the same manner as described for the preparation of (17) and gave 0.64 g of material (0.96 mmol, 77.1%), which was further purified by HPLC to give 0.44 g of product (0.66 mmol, 68.8%). Puritdy of the peptide was checked by analytical HPLC (retention time, 11.5 minutes) (see Figure 3-12).

MP, 140°C (decomp.). *Anal.* calcd. for $C_{29}H_{51}O_{10}N_7$: C 52.97 H 7.76 O 24.35 N 14.92; found: C 52.89 H 7.75 O 24.37 N 14.97

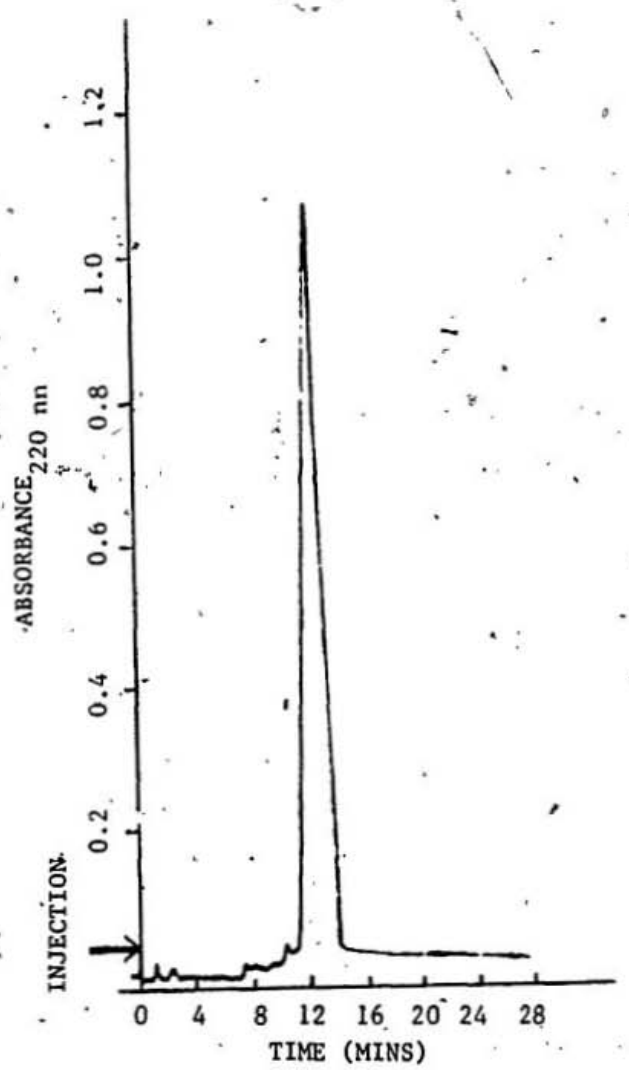

FIG. 3-12: Analytical HPLC of BocLeuHyPGlyAlaLysGlyOH

Chapter 4

4. Spectroscopic analysis of peptide structure

4.1. Infrared spectroscopy

4.1. Introduction to Infrared Spectroscopy

Vibrational spectroscopic techniques, namely, infrared and Raman when used in conjunction with other physical methods such as CD, NMR and X-ray crystallography are useful for identifying and characterizing secondary structures in peptides and proteins. IR provides a useful means for identifying hydrogen bonded and non-hydrogen bonded conformation of peptides and proteins. It is a particularly useful technique for peptide structure determination.

The IR absorption spectrum of a molecule results from transitions involving vibrational and rotational energy levels. When a molecule absorbs radiation, its energy increases in proportion to the energy of the photon as expressed by the relation

$$\Delta E = h\nu = hc/\lambda$$

where h is the Planck's constant, ν and λ are the frequency and the wavelength of the radiation respectively and c is the velocity of light. Changes in vibrational energy involve small quanta and the changes in rotational energy involve quanta even smaller than those of vibrational energy. If a molecule absorbs radiation in the far IR region (650 to 100 cm^{-1}), only its rotational energy will change, no matter which vibrational or electronic state it is in. If the radiation is in the medium IR region (4000 to 650cm^{-1}), both the vibrational and rotational energies of the molecule will change. IR frequencies are determined mainly by the mechanical motions in the molecule while the intensities are related to the electrical properties.

The IR spectra of peptides and proteins exhibit characteristic bands associated with the peptide (-CO-NH-) groups (Miyazawa, 1967). Since these amide bands show variations in frequencies and intensities, conclusions can be drawn regarding the different conformations of a peptide or protein from its IR spectrum. There are nine distinct amide bands in proteins: amide A, amide B and amide I to amide VII. Amide bands A and B, which have the highest frequency, primarily originate from NH stretching vibration in the region 3200-3500 cm^{-1}. In the amide A region, the appearance of a band between 3200-3400 cm^{-1} is expected for conformations with both intra- and intermolecular hydrogen bonds (Miyazawa and Blout, 1967; Stimson *et.al.*, 1977). For a non-hydrogen-bonded NH group, the NH vibration amide B band is observed between 3400-3500 cm^{-1} The amide I band 1600-1700 cm^{-1} is due to CO stretching and its appearance at 1690 cm^{-1} may be indicative of hydrogen-bonded conformation (Miyazawa, 1967;Kawai and Fasman 1978). The amide II (1480-1575 cm^{-1}), III (1230-1430 cm^{-1}) and V (640-800 cm^{-1}) frequencies are much more sensitive to the presence of an intramolecular hydrogen bond than amide I (Maxfield *et al.*, 1981; Hollosi *et al.*, 1985). However, these are difficult to access experimentally due to solvent absorption and therefore difficult to interpret correctly. IR has been used extensively for determining the presence of reverse turns in peptides by several investigators including our own laboratory (Ananthanarayanan and Shyamasundar, 1981 Brahmachari *et al.*, 1982; Rao *et al.*, 1980, see Smith and Pease, 1980 for a review). Krimm and Bandekar (1979, 1980) have used normal vibrational analysis calculations to suggest that the positions of the amide I and III bands in the IR and Raman spectra of proteins and peptides may provide criteria for the identification of the presence of reverse-turn structure. These workers have defined amide I, II and III frequencies for type I and type II β-turns that should

be useful in identifying these structures. These results have been applied successfully to determine chain conformations of a number of fibrous proteins (Krimm and Bandekar, 1980). The 1700-1754 cm^{-1} region due to ester and urethane carbonyl stretching is also important in determining β-turn structures. The rest of the amide bands are less frequently used for structural characterization purposes.

The final deblocked peptide products used in the present study were not soluble in chloroform or other non-polar solvents, hence the IR spectra were measured using blocked precursors of these peptides. These precursors have the ϵ-NH of the lysine residue blocked with the carbo-benzoxy (cbz) group while the carboxylic end is blocked with a methyl group. The side chain guanidine group of the arginine residue was blocked by the nitro group. It is assumed that the conformation of the peptide backbone will not be significantly affected by such modifications of the side chains and the carboxylic group.

Table 4-1:

Infrared Data (ν_{NH} and ν_{CO}) for Synthetic peptides

Peptide	ν_{NH} (cm^{-1})	ν_{CO}(cm^{-1})
N^{α}tBocAla(cbz)LysGlyOMe	3443,3356	1745,1703, 1677
N^{α}tBocIle(cbz)LysGlyOMe	3443,3360	1746, 1678,1707
N^{α}tBocAla(cbz)LysGlySerOMe	3441,3343	1747,1700,1678
N^{α}tBocAlaGly(cbz)LysHyPOMe	3444,3336	1744,1704,1648
N^{α}tBocLeuHyPGlyAla(cbz)LysGlyOMe	3443,3324	1737,1700,1679,1634
N^{α}tBocAla(nitro)ArgGlyIle(cbz)LysGlyOMe	3441,3349	1746,1702,1657

4.1.2. Tripeptides: N^{α}tBocAla(cbz)LysGlyOMe and N^{α}tBocIle(cbz)LysGlyOMe

Figure 4-1 shows the NH region of the above two tripeptides in chloroform. Two bands are seen to be present, one at about 3443 cm^{-1} and another at either 3356 cm^{-1} in the case of N^{α}tBocAla(cbz)LysGlyOMe or 3360 cm^{-1} in N^{α}tBocIle(cbz)LysGlyOMe. From the information available on IR of peptides presented in the introduction, we can conclude that, in these two tripeptides, both non-hydrogen bonded ("free") NH group(s) and hydrogen-bonded NH group(s) are present, the former appearing at 3443 cm^{-1} and the latter at 3356 cm^{-1} or 3360 cm^{-1}. The proportion of the hydrogen-bonded conformer in both the tripeptides appears to be smaller compared to the hydrogen- bonded species as judged from the intensities of the bands.

The carbonyl regions of these tripeptides are shown in Figs. 4-3 and 4-4. The ester carbonyl is found at 1740 cm^{-1}. Rao *et al.*, (1980) have observed that for small dipeptides that contain the Boc-X group (where X= Leu, Ile, Val) and show no evidence for any hydrogen bonded conformation, the IR band of the urethane group (involving the Boc carbonyl) appears in the region 1700-1720 cm^{-1}. The dominant band seen in our peptides at 1702-1710 cm^{-1} may therefore be assigned to the urethane carbonyl which is not involved in hydrogen bonding. The weak peptide CO band observed at 1670 cm^{-1} in both the tripeptides indicates a small population of hydrogen-bonded conformer(s). The hydrogen-bonded conformers indicated by the IR data of the two tripeptides in the NH and CO stretching regions should, therefore, be visualized without the involvement of the Boc group and this leads us to expect the presence of structures stabilized by five-atom (C_5) and/or seven-atom (C_7) hydrogen-bonded rings (Figure 4-2). The C_5 conformer would form a hydrogen bond between the Ala^1 CO or Ile^1 CO and Lys^2 NH. The C_7 conformer would form a hydrogen bond between Ala^1 CO or Ile^1 CO and Gly^3 NH. The IR band positions in the NH region for the hydrogen

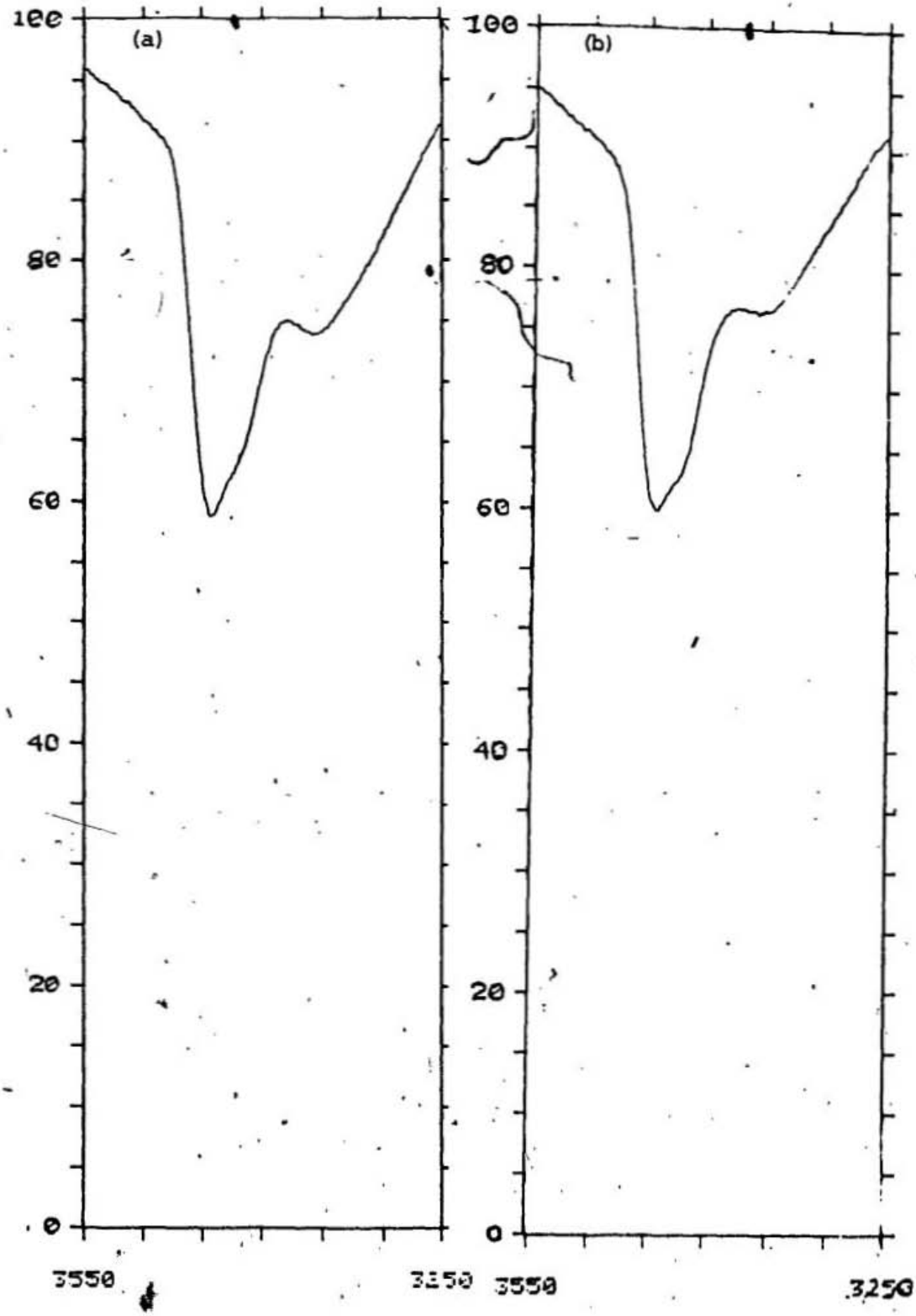

FIG. 4-1: ν_{NH} bands of tripeptides in chloroform. (a) N^{α}tBocAla(cbz)LysGlyOMe and (b) N^{α}tBocIle(cbz)LysGlyOMe.

bonded species in these peptides (3356 and 3360cm^{-1}) appear, however, to correspond more to the presence of C_7 rather than the C_5 conformation. (The band position for the C_7 conformer is found near 3390-3400 cm^{-1} in model peptides (Rao *et al.*, 1980)). It should, in addition, be pointed out that the C_5 conformation is rather a restrictive ring structure and is favoured only under extreme conditions such as in peptides that contain residues whose presence limits the conformational freedom such as the Aib residue (Aib = α-amino isobutyric acid), and/or when the peptides are dissolved in highly non-polar solvents like CCl_4 (Rao *et.al.*, 1980). We therefore can consider the probability of the presence of the C_5 conformer to be very small in our peptides.

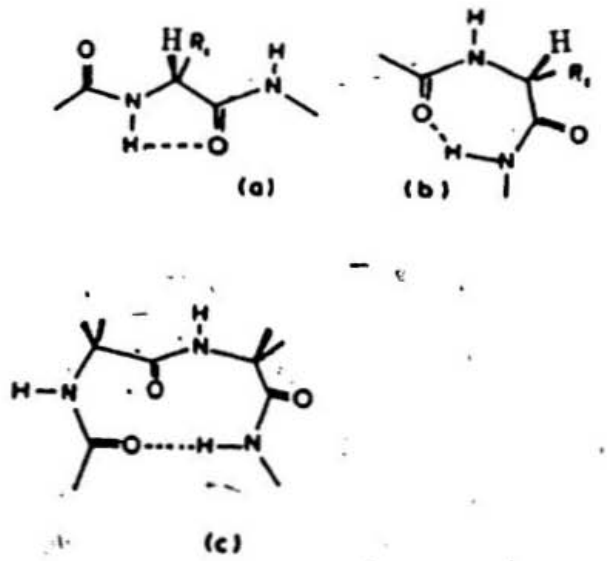

FIG. 4-2: Schematic representation of intramolecular hydrogen-bonded structures. (a) C_5; (b) C_7 or γ-turn and (c) C_{10} or β-turn.

It is interesting to note that while the C_{10} (Figure 4-2) i.e. β-turn (characterized by the reversal by nearly 180° in the direction of the peptide chain) (Zimmerman and Scheraga, 1977; Crawford *et al.*, 1973; Smith and Pease 1980), is a possible conformation for these peptides (involving the urethane CO and Gly NH groups), the IR data rule this conformation out. This is in conformity with our expectations from data on the positional preferences of amino acids for the β-turn (Chou and Fasman, 1977), whereby Lys does not prefer to occupy the 3rd position of the turn. The C_7 conformation of these peptides is shown in Figure (4-4). This conformation could also be conceived of as a "partial β-turn" as found in the tripeptides such as BocProDAlaAlaOH (Ananthanarayanan and Cameron, 1986) and BocProGlyAlaOH (Benedetti, 1977). As will be shown below, the CD data on the tripeptides support the conclusions from IR.

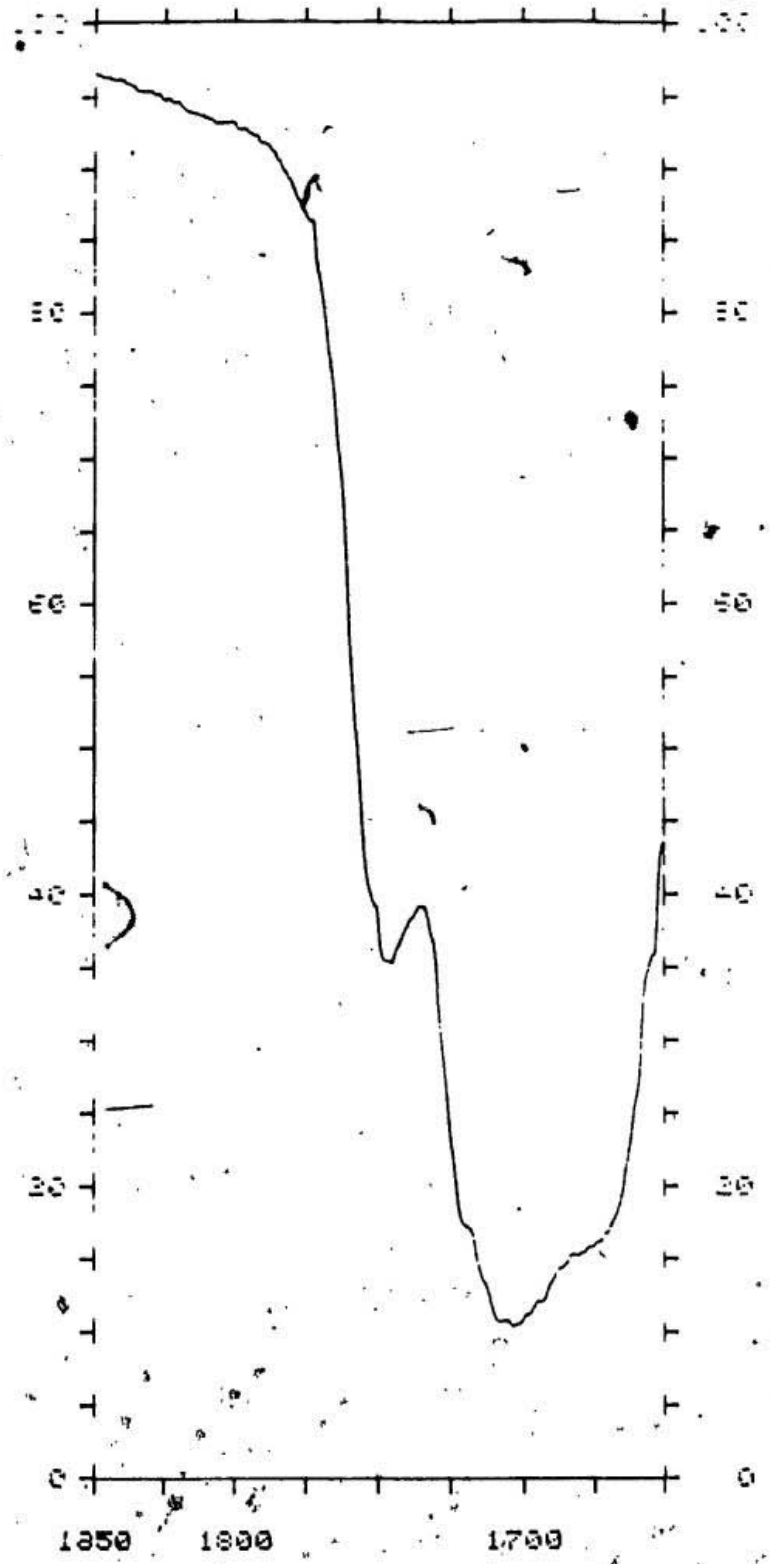

FIG. 4-3: ν_{CO} band of N^{α}tBocAla(cbz)LysGlyOMe in $CHCl_3$.

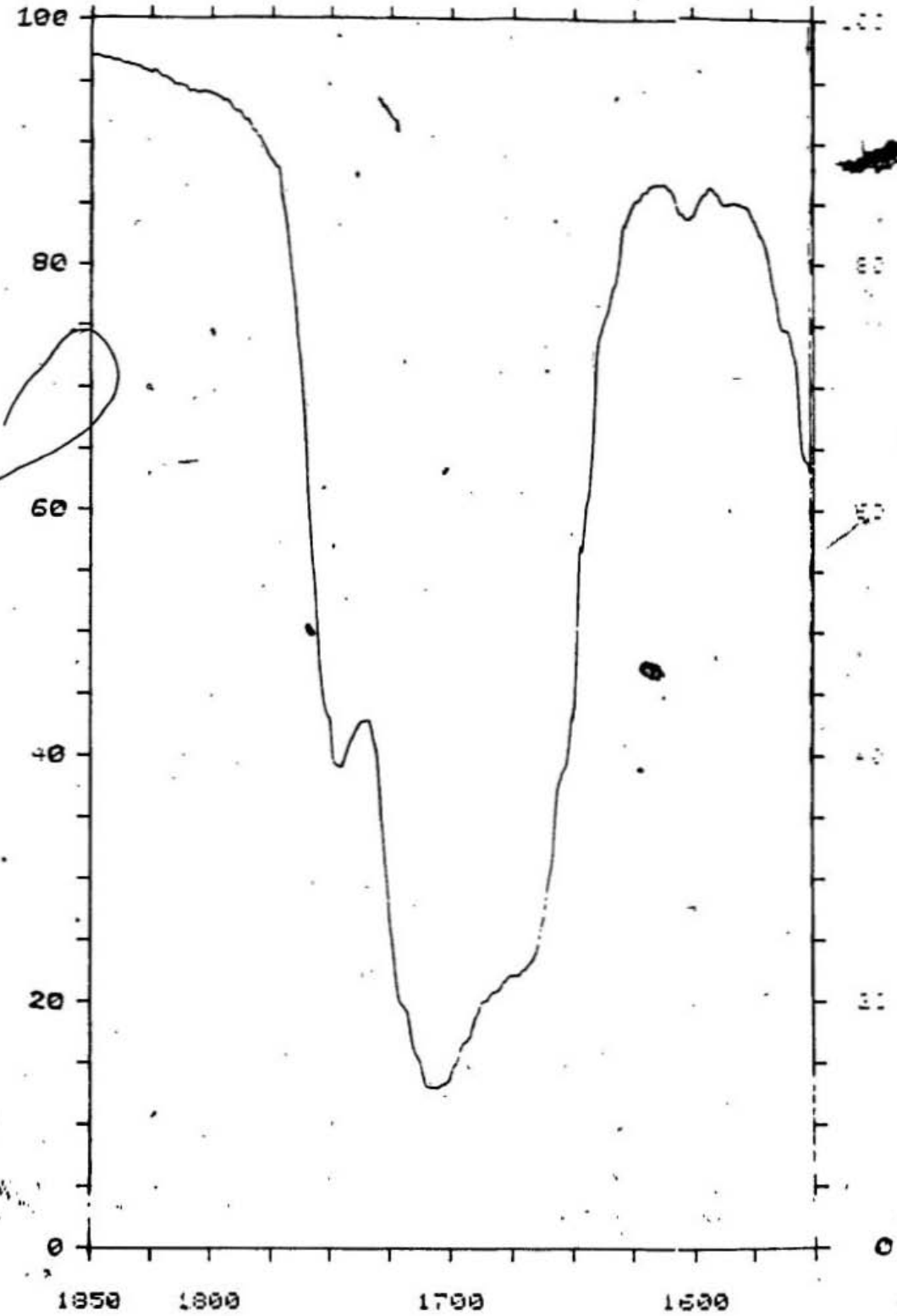

FIG. 4-4: ν_{CO} band of N^{α}tBocIle(cbz)LysGlyOMe in $CHCl_3$.

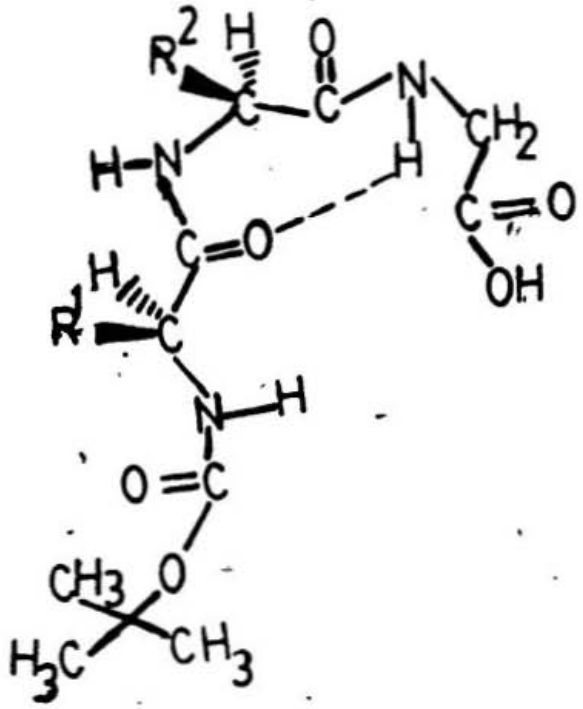

FIG. 4-5: Schematic Diagram of BocAla(cbz)LysGlyOMe, BocIle(cbz)LysGlyOMe

R^1 = Ala or Ile

R^2 = Lys

4.1.3. IR of Tetrapeptides: N^{α}tBocAla(cbz)LysGlySerOMe and N^{α}tBocAlaGly(cbz)LysHyPOMe.

The NH regions of the IR spectra N^{α}tBocAla(cbz)LysGlySerOMe and N^{α}tBocAlaGly(cbz)LysHyPOMe are shown in Figure 4-6. The data may be interpreted in an analogous manner to the tripeptide data presented above. The free (non-hydrogen bonded) NH band is observed at 3441cm^{-1} while the hydrogen bonded NH band occurs at 3343cm^{-1} or 3336cm^{-1}. The intensities of the IR bands of the hydrogen-bonded NH and non-hydrogen bonded NH are found to be almost the same for N^{α}tBocAla(cbz)LysGlySerOMe while, in N^{α}tBocAlaGly(cbz)LysHyPOMe, the intensity of the hydrogen-bonded NH band at 3336 cm^{-1} is weaker compared to the non-hydrogen-bonded NH band at 3441 cm^{-1}.

The carbonyl regions of the IR of these two tetrapeptides are shown in Figures 4-7 and 4-8. The ester carbonyl is found to appear at 1747 cm^{-1} or 1744 cm^{-1} while the urethane occurs around 1700-1705 cm^{-1}. The later indicates that, as in the case of the tripeptides, the Boc group does not seem to be involved in the hydrogen bonded conformer(s). On the basis of the IR data in the NH region, we can, however, expect the presence of hydrogen-bonded conformers.

In N^{α}tBocAla(cbz)LysGlySerOMe, the amide I band due to peptide CO appears at 1679 cm^{-1}. Its position is attributable to its involvement in a hydrogen-bonded conformer (Rao *et.al.*, 1980) as has been indicated earlier by the NH bands of this peptide. These could be the C_7 and/or C_{10} conformers. The C_7 conformer would involve a hydrogen bond between Ala^1 CO and Gly^3 NH. The C_{10} (i.e. the β-turn) conformation would involve hydrogen bonding between Ala^1 CO and Ser^4 NH. The ratio of the intensity of the hydrogen bonded NH and non-hydrogen-bonded NH (1:1) suggests the participation of two out of the

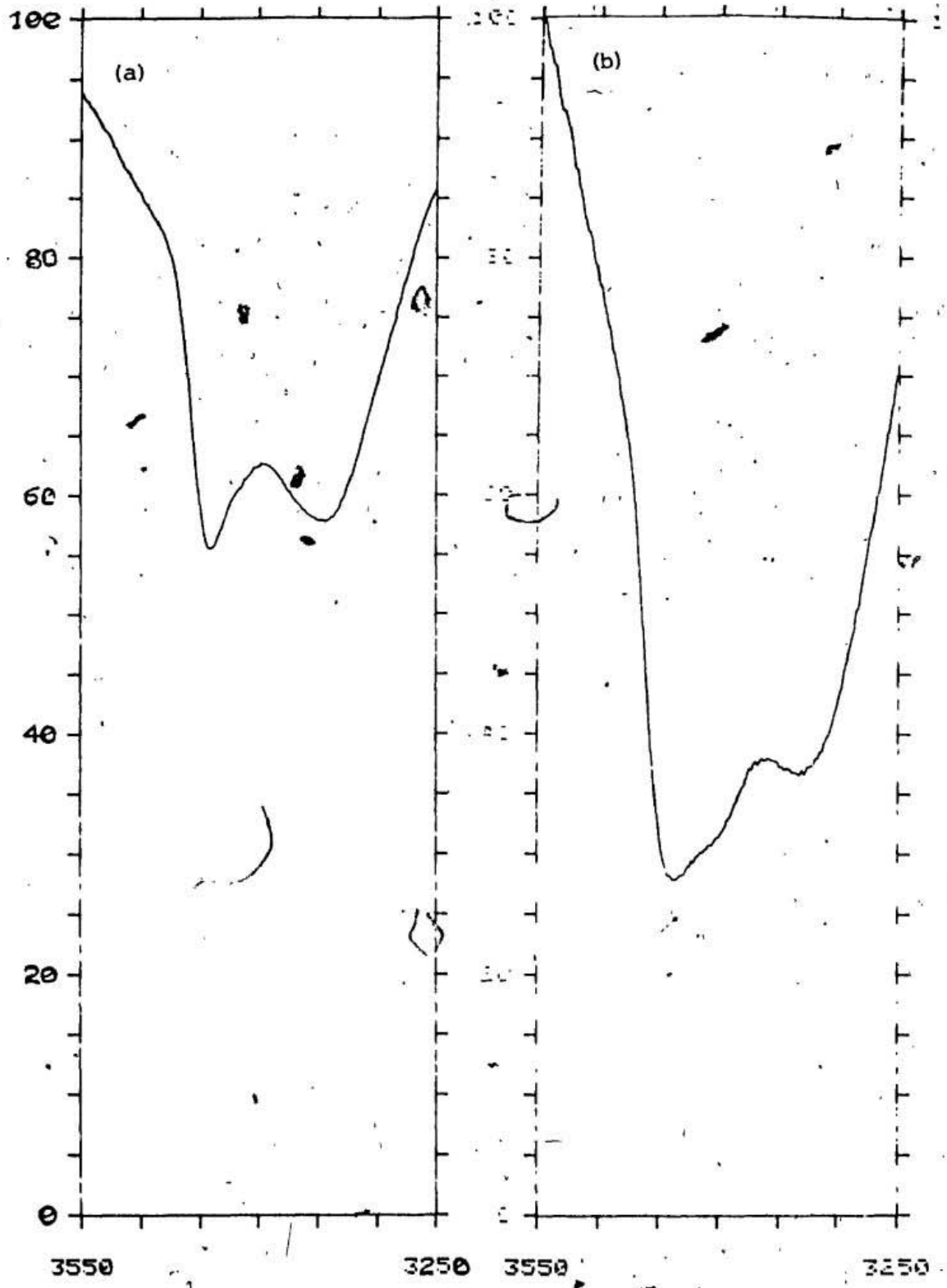

FIG. 4-6: ν_{NH} bands of (a) N^{α}tBocAla(cbz)LysGlySerOMe and (b) N^{α}tBocAlaGly(cbz)LysHyPOMe in $CHCl_3$.

four NHs in hydrogen bonding in N^{α}tBocAla(cbz)LysGlySerOMe. This would be possible if this peptide could form both a γ-turn (C_7) and a β-turn (C_{10}) as shown in Figure 4-6. The two hydrogen bonds would involve Ala^1 CO and Gly^3 NH (γ-turn) and Ala^1 CO and Ser^4 NH (β-turn). A bifurcated hydrogen bond would thus result. The occurrence of such bifurcated hydrogen-bonding has been suggested by the recent theoretical computations of Paul and Ramakrishnan (1985) and has been reported by Hollosi *et al.*, (1985) for the peptide cbzGlyProAsp(OBu)GlyOSt; this peptide is found to have a conformation which is be stabilized by a 2--->1 hydrogen bond (in a C_5 structure) between urethane CO and Gly^1 and a 5--->1 hydrogen (in a C_{13} conformation), bond between urethane CO and Gly^4.

In the case of N^{α}tBocAlaGly(cbz)LysHyPOMe, the amide I band of the peptide CO is found to be unresolved from the urethane group (Figure 4-8), indicating that the peptide CO groups are possibly in the non-hydrogen-bonded conformation (which have bands at higher frequencies than those of the hydrogen-bonded conformers). The band due to the tertiary amide of the HyP residue is seen at 1648 cm^{-1}. In this peptide, the amide proton of the pyrrolidine ring is not available for hydrogen bonding and therefore the 4--->1 hydrogen bonded β-turn cannot be formed in this peptide (recalling that the Boc carbonyl does not participate in hydrogen bonding). In conjunction with the IR data in the NH region, we conclude the absence of significant amounts of hydrogen bonded conformers in this peptide other than a relatively small proportion of the γ-turn conformation, with a hydrogen bond formed between Ala^1CO and Lys^3 NH.

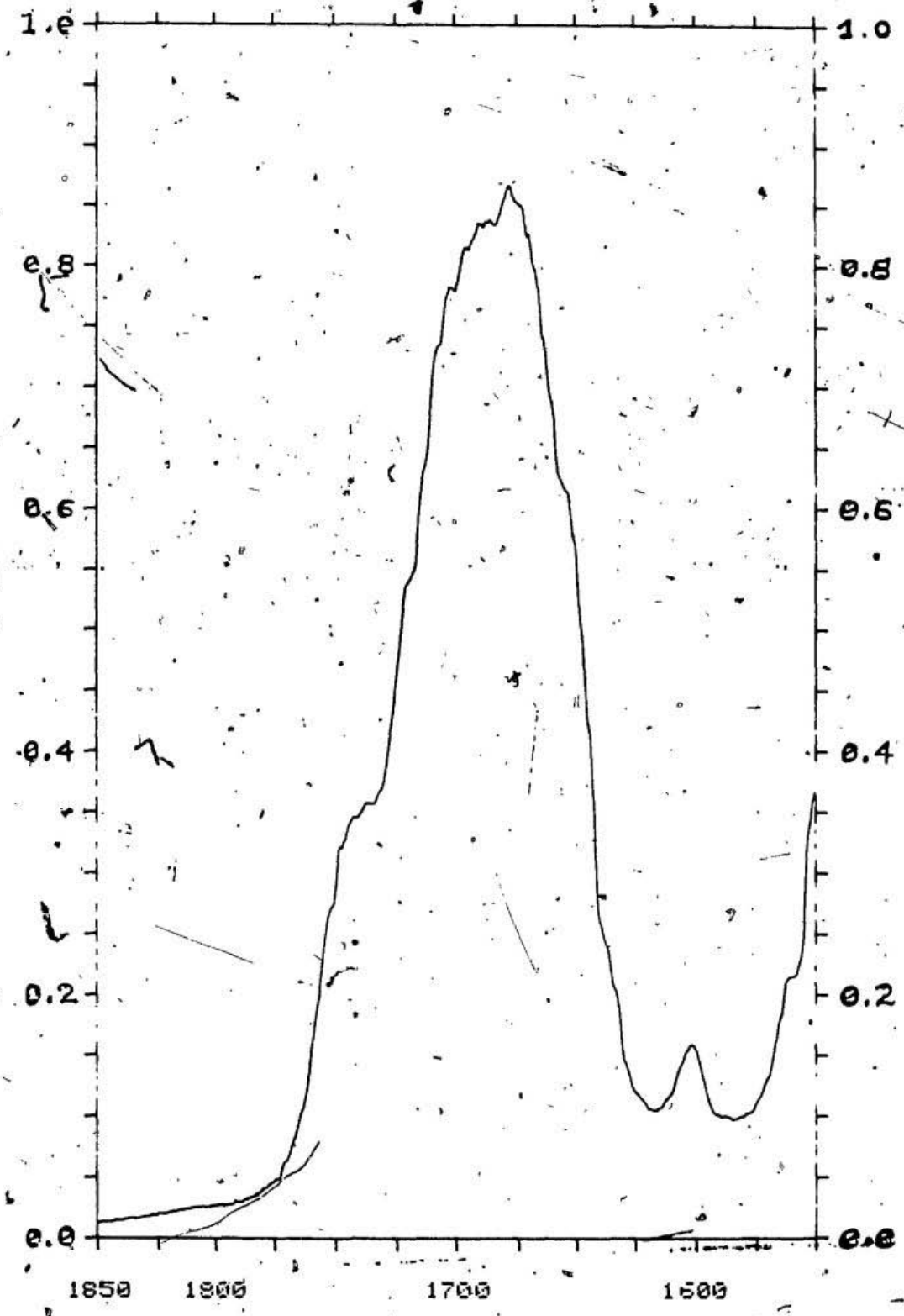

FIG. 4-7: ν_{CO} of N^{α}tBocAla(cbz)LysGlySerOMe in $CHCl_3$.

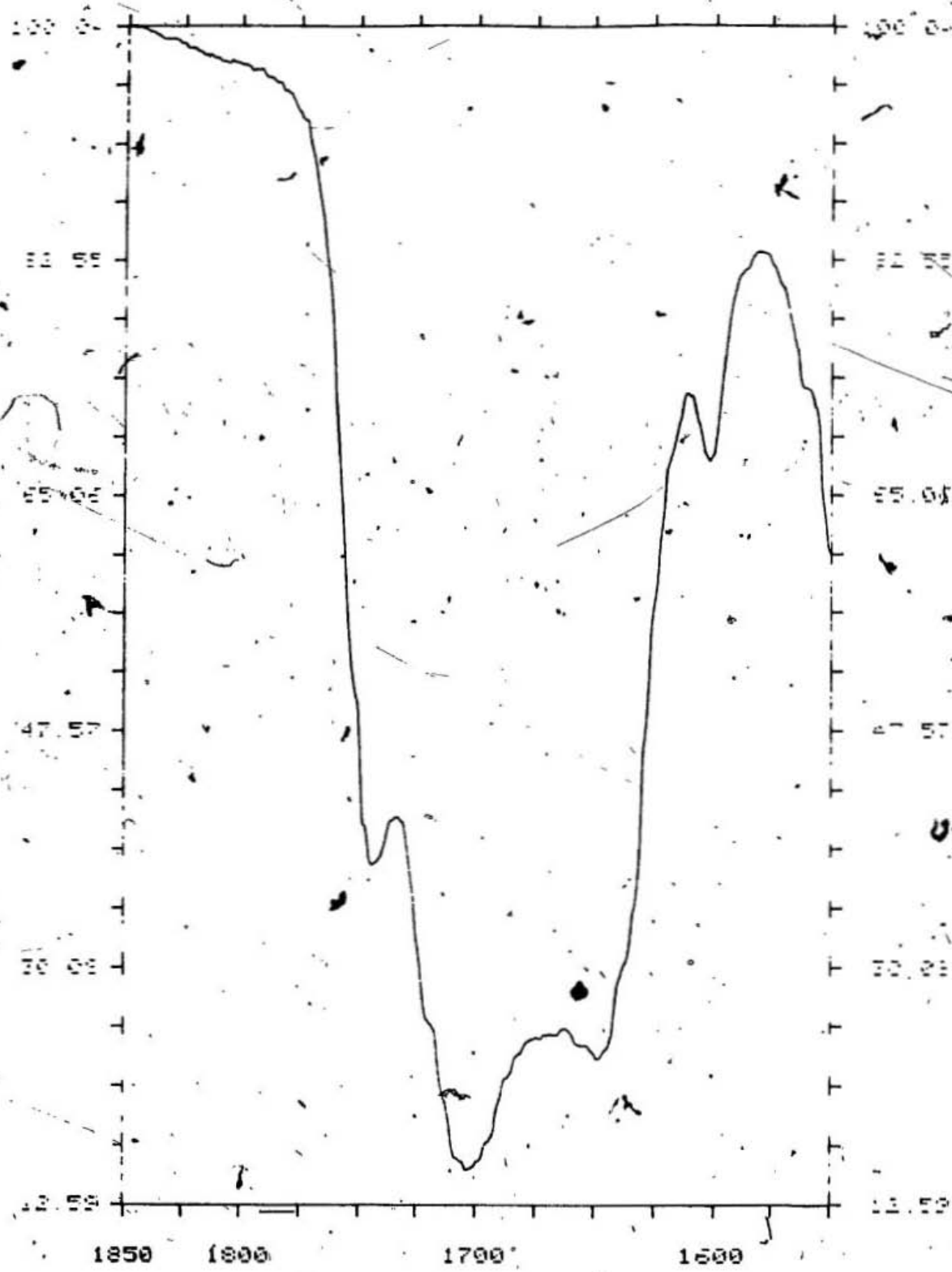

FIG. 4-8: ν_{CO} band of N$^{\alpha}$tBocAlaGly(cbz)LysHyPOMe.

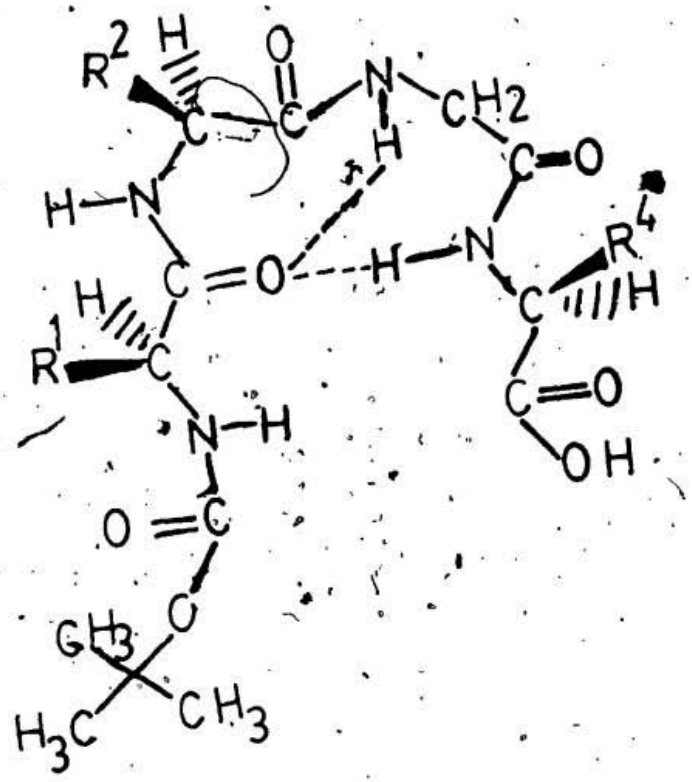

FIG. 4-9: Schematic Diagram of BocAlaLysGlySerOH

R^1 = Ala

R^2 = Lys

R^4 = Ser

4.1.4. IR of Hexapeptides: N^{α}tBocLeuHyPGlyAla(cbz)LysGlyOMe and N^{α}Ala(nitro)ArgGlyIle(cbz)LysGlyOMe

Figure 4-10 shows the NH region of the two hexapeptides in $CHCl_3$. Two bands are seen to be present in both the peptides. The non-hydrogen-bonded NH band occurs at 3443 cm^{-1} (N^{α}tBocLeuHyPGlyAla(cbz)LysGlyOMe) or 3441 cm^{-1} (N^{α}tBocAla(nitro)ArgGlyIle(cbz)LysGlyOMe) while the respective hydrogen-bonded NH band appears at 3324 cm^{-1} or 3349 cm^{-1}. From Figure 4-10 (a), it can be seen that N^{α}tBocLeuHyPGlyAla(cbz)LysGlyOMe appears to exist predominantly in the hydrogen-bonded conformer as judged by the relatively larger intensity of the hydrogen-bonded NH band compared to the non-hydrogen-bonded NH band. In contrast, N^{α}tBocAla(nitro)ArgGlyIle(cbz)LysGlyOMe appears to exist predominantly in the non-hydrogen-bonded conformer as shown by the relatively weaker intensity of the NH band in the hydrogen-bonded conformation (Figure 4-9 (a)).

The carbonyl region of the IR of these two hexapeptides is shown in Figures 4-11 and 4-12. The ester carbonyl is found to appear at 1737 cm^{-1} or 1746 cm^{-1} while the urethane carbonyl appears between 1700-1702 cm^{-1}. Similar to the tri- and tetrapeptides, the urethane carbonyl of the tBoc group does not seem to be involved in the hydrogen bonded conformer(s). Like in the case of N^{α}tBocAla(cbz)LysGlySerOMe, the carbonyl region of N^{α}tBocLeuHyPGlyAla(cbz)LysGlyOMe shows an intense amide I band at 1678 cm^{-1} attributable to the involvement of the peptide CO group in the hydrogen-bonded conformer. The position of the tertiary amide band at 1634 cm^{-1} also indicates the presence of a hydrogen bonded conformer (Rao *et al.*, 1980). The possible hydrogen bonded conformers for N^{α}tBocLeuHyPGlyAla(cbz)LysGlyOMe are the C_7 and/or C_{10} structures. The C_7 conformer would involve a hydrogen

bond between HyP^2 CO and Ala^4 NH or between Ala^4 CO and Gly^6 NH while the C_{10} conformer involves a hydrogen bond between Leu^1 CO and Ala^4 NH. From the large intensity of the hydrogen-bonded peptide NH and CO bands, it is likely that N^{α}tBocLeuHyPGlyAla(cbz)LysGlyOMe adopts both a γ- or a β-turn as shown in Figure 4-13 involving a bifurcated hydrogen bond as in the case of N^{α}tBocAla(cbz)LysGlySerOMe.

In the case of N^{α}tBocAla(nitro)ArgGlyIle(cbz)LysGlyOMe the amide I band is not resolved from the urethane band, instead, a band appears at 1657 cm^{-1} which is attributable to the free tertiary amide carbonyl (Rao *et al.*, 1980). Together, the IR data in the NH and CO regions of this peptide point out to the presence of a weakly hydrogen-bonded conformer, possibly a γ-turn involving a hydrogen bond between Ile^4 CO and Gly^6 NH, as in the case of the tripeptide N^{α}tBocIle(cbz)LysGlyOMe.

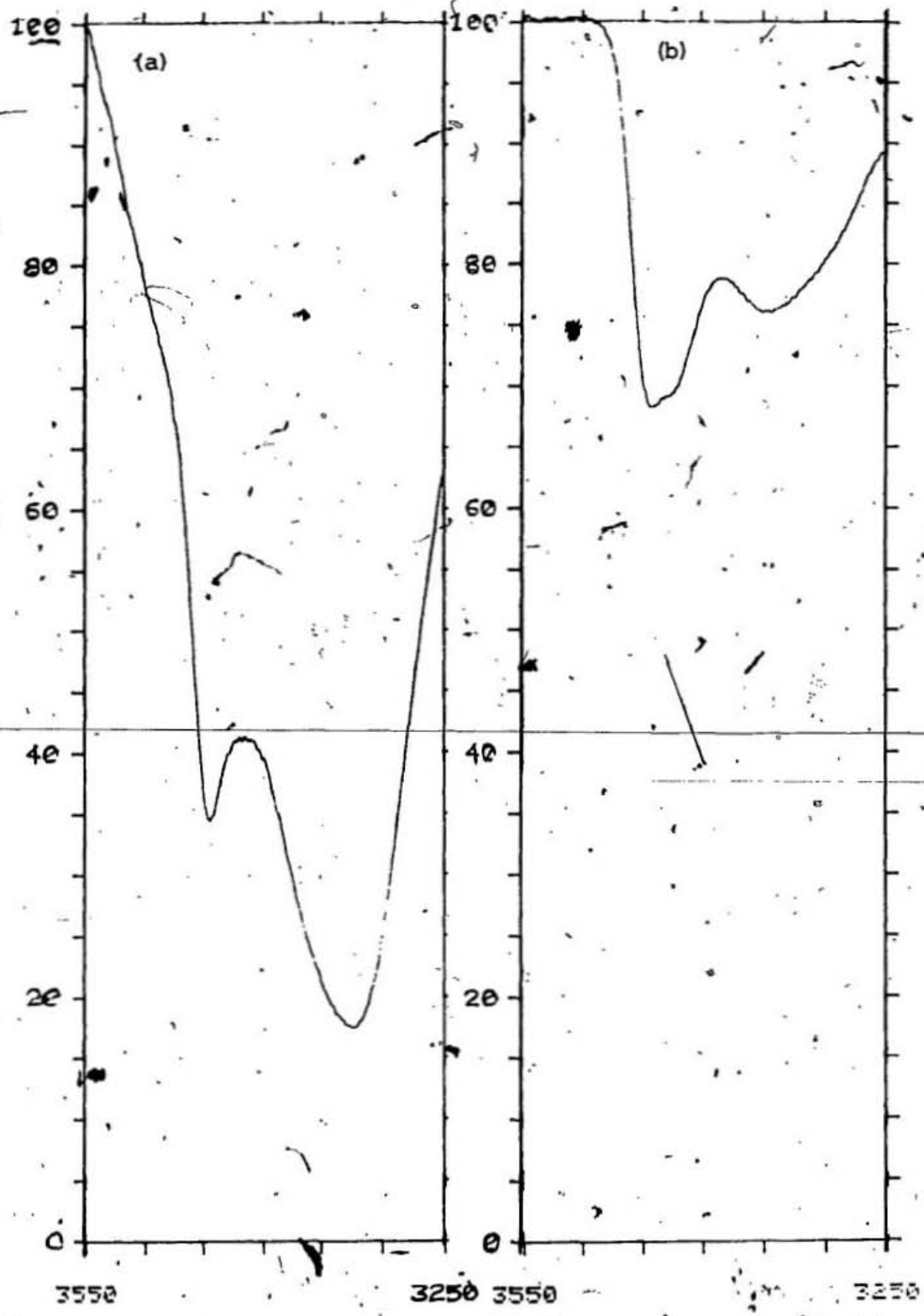

FIG. 4-10: ν_{NH} band of N^{α}tBocLeuHyPGlyAla(cbz)LysGlyOMe and N^{α}tBocAla(nitro)ArgGlyIle(cbz)LysGlyOMe in $CHCl_3$.

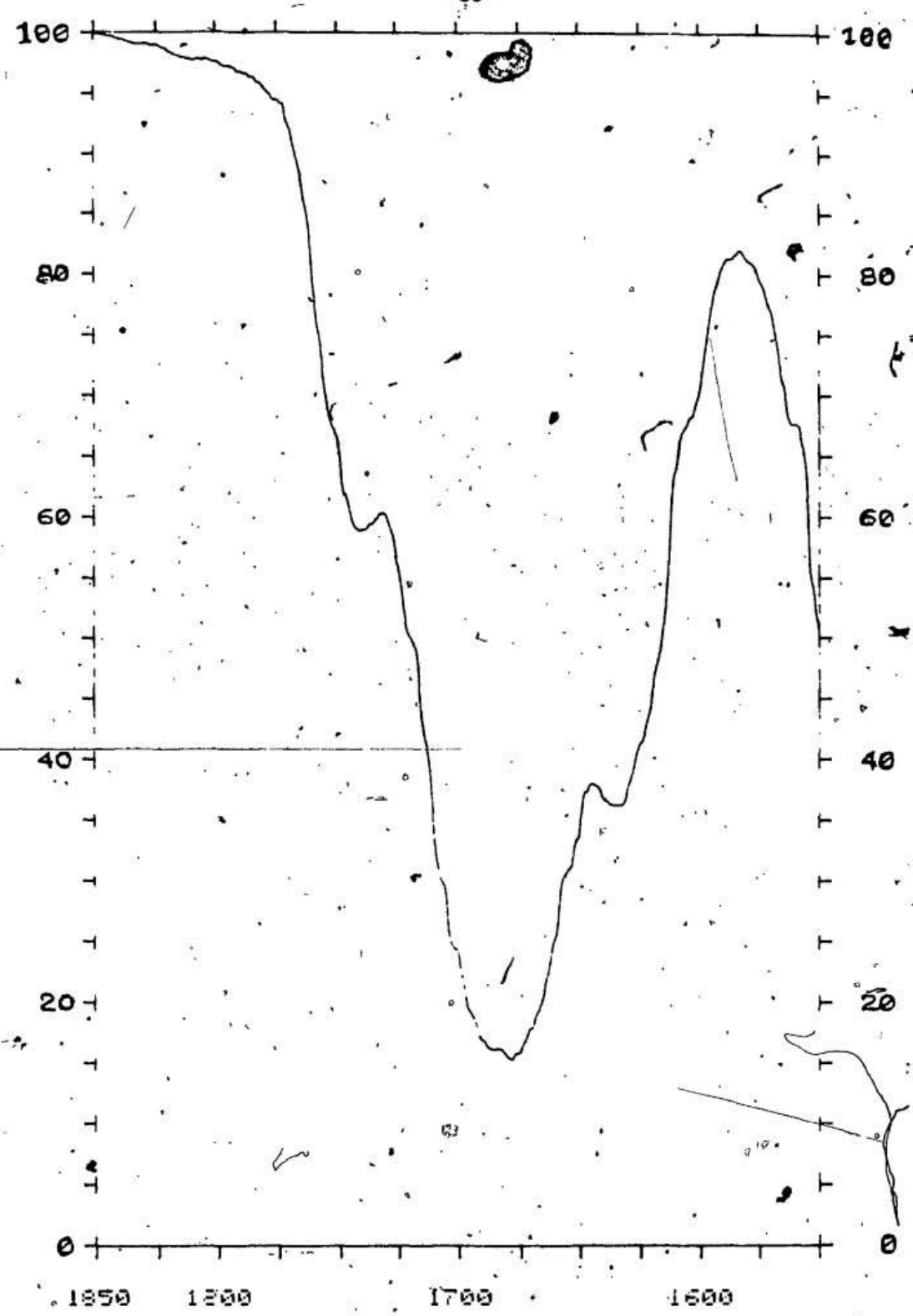

FIG. 4-11: ν_{CO} band of N^{α}tBocLeuHyPGlyAla(cbz)LysGlyOMe.

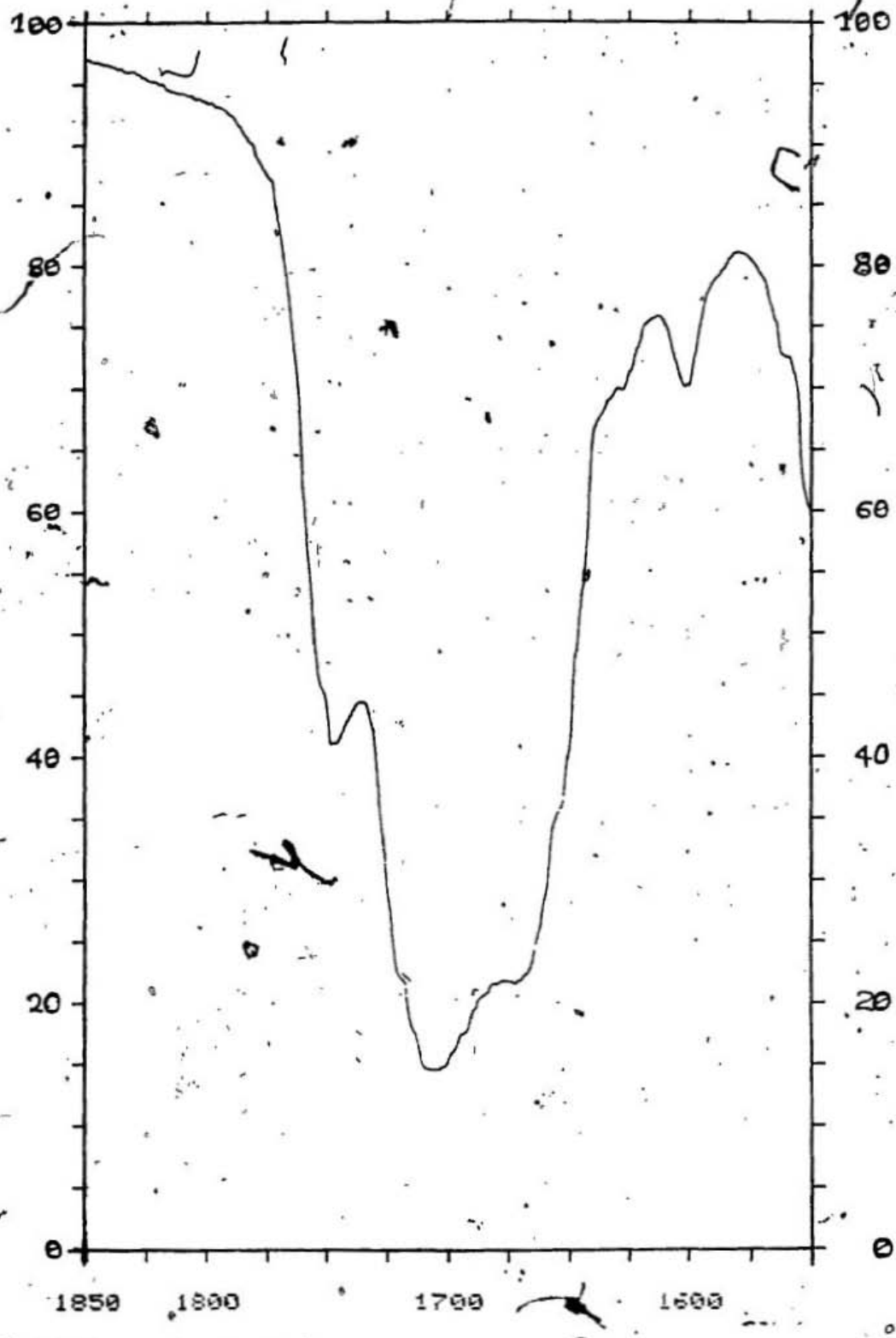

FIG. 4-12: ν_{CO} band of N^{α}tBocAla(nitro)ArgGlyIle(cbz)LysGlyOMe.

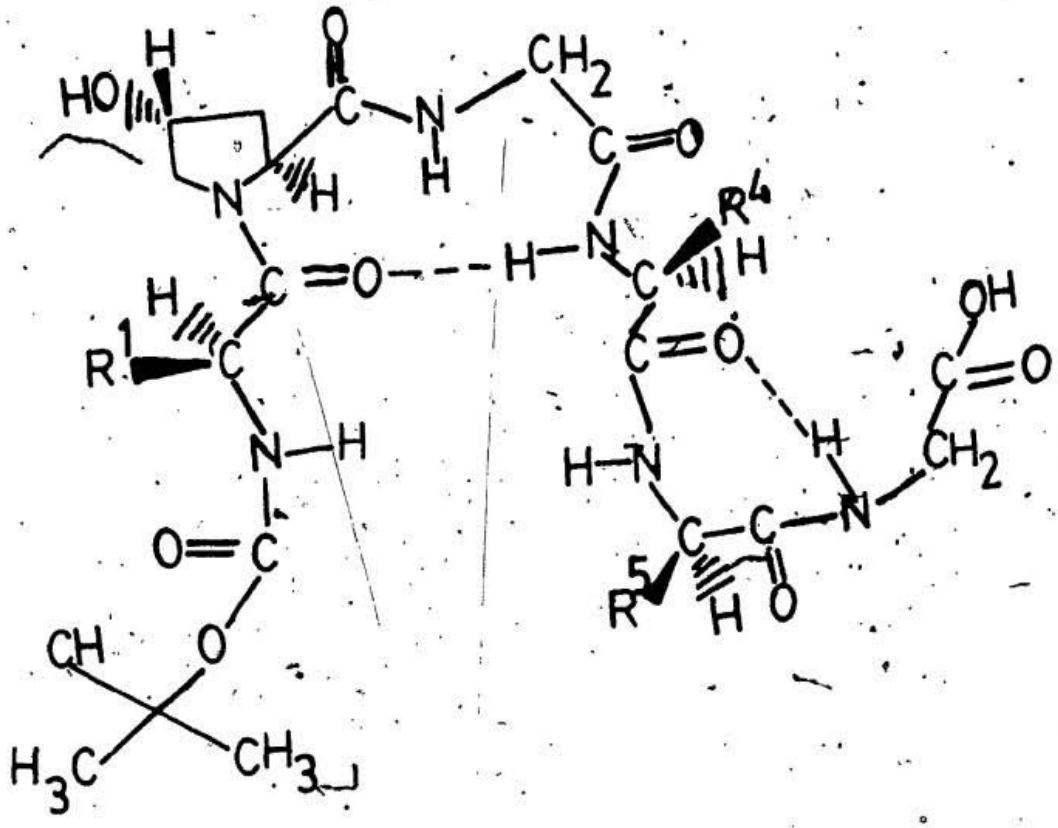

FIG. 4-13: Schematic Diagram of BocLeuHypGlyAlaLysGlyOH

R^1 = Leu

R^4 = Ala

R^5 = Lys

4.2. Circular Dichroism characterization

4.2.1. Introduction to Circular Dichroism

Circular dichroism (CD) measures the differential absorption of right- and left-circularly polarized light by optically active asymmetric chromophores or by non-asymmetric ones that have been asymmetrically oriented (Bell, 1981). The unequal absorptivity for the right- and left-circularly polarized light is defined as $\epsilon_l - \epsilon_r$ where ϵ_l and ϵ_r, respectively, are the molecular extinction coefficients for the left and right components (Haschemeyer and Haschemeyer, 1973).

All amino acids except glycine have at least one asymmetric carbon and therefore the backbone of a peptide molecule itself is an assembly of optically active chromophores whose contributions are found to be additive in the random coil structure. A CD spectrum of a given chromophore such as the peptide bond is made up of the individual spectral contributions of the various conformations present within a given peptide molecule (Chang *et al.*, 1978). This may be expressed as:

$$[\Theta]^{obs} = f_1[\Theta]^1 + f_2[\Theta]^2 + \cdots$$

where $[\Theta]^{obs}$ is the observed (mean residue)-ellipticity at a specified wavelength, f_1, f_2 etc. refer to the fractions of particular secondary structures present and $[\Theta]^1$, $[\Theta]^2$ etc. are the (mean residue) ellipticity values for these secondary structures.

In peptides and proteins, the CD spectra arise from electronic transitions of the peptide bond. These transitions are the n-π* (representing the excitation of an electron from a nonbonding orbital to an excited antibonding π orbital) and the π-π^* transitions (arising from excitation from a bonding to an antibonding π orbital) (Haschemeyer and Haschemeyer, 1973).

Table 4-2:

Maxima and minima of β-turn CD Spectral Classes as Defined by Woody (1974)

Class	Sign	Wavelength Range (nm)
A	-	210-220
	+	195-200
	-	<190
B	-	>220
	+	210-210
	-	<190
C	-	200-210
	+	180-190
D	-	>225
	+	210-220
	-	190-200
	+	<190

CD is exquisitely sensitive to peptide conformation and has been widely used for studying the conformation of peptides (Adler et al., 1973; Sears and Beychok, 1973; Balasubramanian and Kumar, 1976 Brahmachari *et al.*, 1982). While the α-helix and the β-structure have relatively well-defined spectra as gathered from the CD data on model polypeptides and proteins of known structures (see Fox et al., 1981, for a review), such reference spectra are not readily available for β- and γ-turns. Prior to the theoretical CD spectra of isolated β-turns by Woody (1974, see later), the CD curves were analyzed in terms of the relative contribution of α-helical, β-sheet and random coil structures only without including the γ- or β-turns (Greenfield and Fasman, 1969).

Woody (1974) calculated the contributions of the various types of β-turns to the CD of proteins and polypeptides by computing n-π^* and π-π^* rotational strengths for tripeptides in these β-turn conformations. He defined four classes of of CD spectra (A, B, C and D). Table 4-1 shows the criteria for his classification. Type I β-turn gives a CD spectrum similar to that of class B. For type II β-turns about 80% of the variants (i.e. subclasses within this type) give class B CD spectrum while about 20% give an inverted α-helix pattern (class C). About two thirds of the type III β-turns give class B CD patterns while about 15% give class A and another 20% give class C spectrum. Class D spectrum is most likely to occur with an open β-turn (i.e β-turn without the 4-->1 hydrogen bond) and class C' (the mirror image of class C) spectra are likely to be associated either with type II β-turns or open β-turns. Conversely, class C spectra can also be associated with type II' β-turns (as in the case of gramicidin S) (Woody, 1974).

One of the difficulties in the interpretation of Woody's findings is that an unequivocal assignment of a β-turn conformation to a peptide cannot be made on the basis of the spectral class comparison alone without corroborating

information obtained from another physical technique (e.g. IR, X-ray diffraction or NMR). Despite this setback, there is now adequate experimental evidence indicating that the occurrence of these spectral classes is related to the β-turn conformation in a given peptide, if one is indeed present.

Expected CD spectra have also been calculated for γ-turns (Madison and Schellman, 1970). Model peptide data (Ananthanarayanan and Shyamasundar, 1981; Bush *et al.*, 1978; Brahmachari *et al.*, 1979; Gierasch *et al.*, 1981; Madison and Kopple, 1980; Bandekar *et al.*, 1982) have given support to certain of the calculated β- and γ-turns spectral types and have raised doubts about others. The experimental data of Hollosi *et al.*, (1985); Rao *et al.*, (1980); Pease and Watson, (1984) and Ananthanarayanan *et al.*, 1976, are useful guides for interpretation of CD data on β-turns. Experimental CD data on γ-turns are rare but at present it appears that γ-turns yield CD bands around 230 nm (Rose *et al.*, 1985). In what follows, we shall attempt to interpret the CD data of our peptides, within the limitation of theoretical understanding of the (basis of) the CD of β-turn types, experimental errors and lack of precise experimental data on CD of each type.

The CD studies were carried out in both polar solvents (water and aqueous 0.1M ammonium bicarbonate) and in less polar solvents (methanol and trifluoroethanol). Polar solvents are known to disrupt ordered structures by competing for peptide hydrogen bonds while relatively non-polar or less polar solvents such as methanol and trifluoroethanol promote structure by stabilizing the intramolecular hydrogen-bonded conformation. In small linear peptides of the type used here, such hydrogen-bonded conformations may be C_5, C_7 (γ-turn) and C_{10} (β-turn) structures. Of these the C_5 is sterically more restrictive and difficult to be detected by CD and is ignored in our analysis of CD spectra.

The CD spectra of eight lysine-containing peptides were measured in different solvents and the data are presented below. Our interpretation of these spectra would be based on the information gathered from the IR data on several of these peptides together with the available literature data on CD of other linear peptides. As mentioned above, an ordered hydrogen-bonded structure, such as the β-turn, in a given peptide is usually sustained better in solvents like TFE, MeOH than in H_2O or aqueous NH_4HCO_3 which tend to break intramolecular hydrogen bonds. We can therefore infer the presence of an ordered structure in a peptide if the CD spectra in the former set of solvents are different from those in the latter. Of course, the ordered structure(s) will be in equilibrium with the disordered or "open" conformations in any solution of a peptide in a solvent so that the observed CD spectrum arises from an average of these structures.

Table 4-3:
Summary of CD data for peptides

Peptide	conc. mg/mL	λ nm	$[\Theta]^a$	λ nm	$[\Theta]^a$	λ nm	$[\Theta]^a$
N^{α}tBocAlaLysGlyOH							
methanol	2			206	-10000	190	+6400
trifluoroethanol	2			202	-5600	182	+11000
Water	2			200	-8000	185	+5000
0.1M ammonium bicarbonate	2			200	-5100	187	+2800
N^{α}tBocAlaGlyLysOH							
methanol	2	220	-880	195	+7300		
trifluoroethanol	2	220	-690	197	+6400	185	-4200
water	2	214	-1100	198	+440	188	-3300
0.1M ammonium bicarbonate	2	210	-1200			189	-3100
N^{α}tBocIleLysGlyOH							
methanol	2			210	-8400	190	+7000
trifluoroethanol	2			210	-6300	185	+9900
water	2			206	-4500	186	+2000
0.1M ammonium bicarbonate	2			204	-4400	188	+3600
N^{α}tBocAlaLysGlySerOH							
methanol	2	218	-3000	196	-2000	186	3300
trifluoroethanol	2	218	-2900	194	-1700	184	4400
water	2			193	-6200		
0.1M ammonium bicarbonate	2			193	-4100		
N^{α}tBocAlaGlyLysHyPOH							
methanol	2			214	-4400	194	+5000
trifluoroethanol	2			214	-3400	193	+3900
water	2			210	-3400	195	+120
0.1M ammonium bicarbonate	2			210	-3800	195	-120
N^{α}tBocLeuHyPGlyAlaLysGlyOH							
methanol	1			208	+4600	192	-4900
trifluoroethanol	1			204	+9200	184	-9900
water	1			216	+1700	194	-7700
0.1M ammonium bicarbonate	1			216	+1300	193	-5700
N^{α}tBocAlaArgGlyIleLysGlyOH							
methanol	2			204	-4200	188	+10000
trifluoroethanol	2			204	-5700	185	+7100
water	2			204	-3200	185	+3000
N^{α}tBocLeuGlyHyPGlyAlaLysGlySerOH							
methanol	2	238	-380	218	+960	198	-4100
trifluoroethanol	2	230	-700	212	+940	188	-3200
water	2			218	+1900	194	-5100

$[\Theta]^a$ = deg $cm^2 dmol^{-1}$ (molar ellipticity)

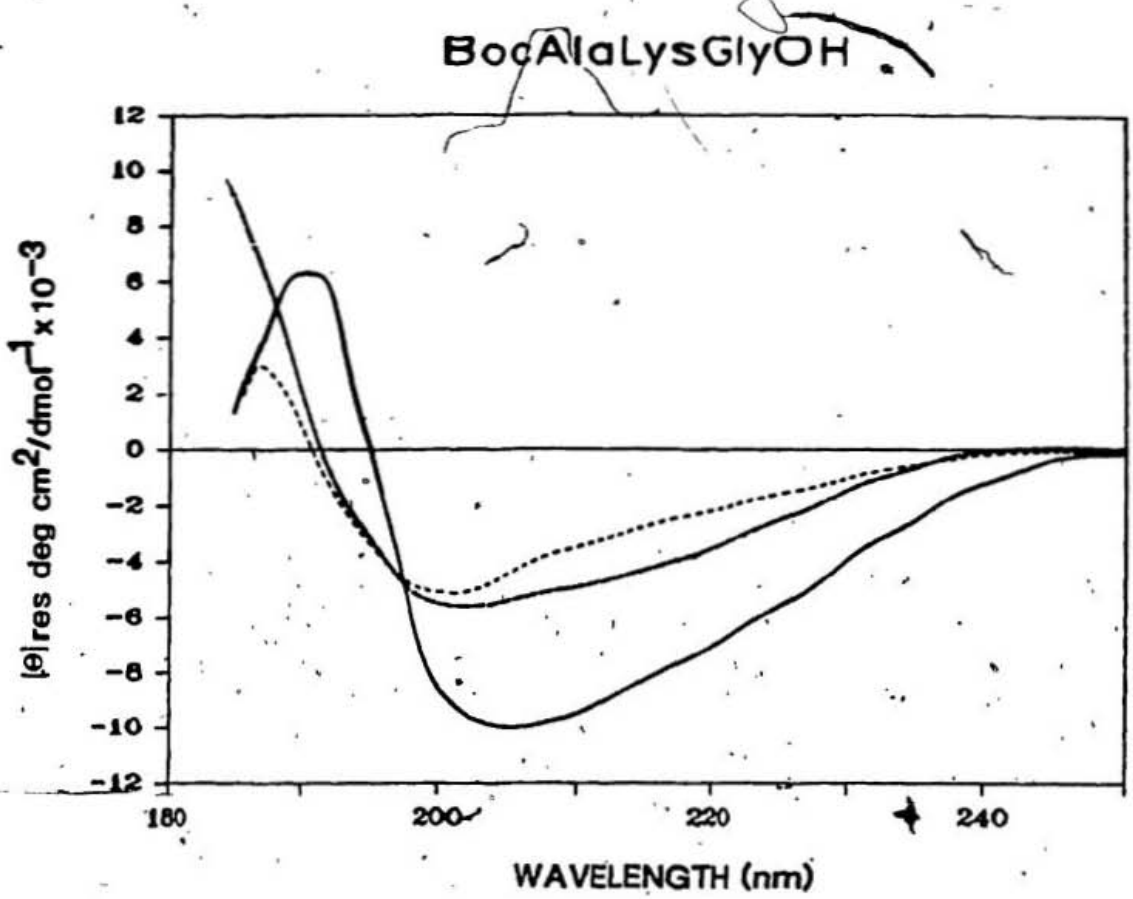

FIG. 4-14: CD spectra of BocAlaLysGlyOH in MeOH, (——); TFE, (- - -); and H_2O (· –) or NH_4HCO_3.

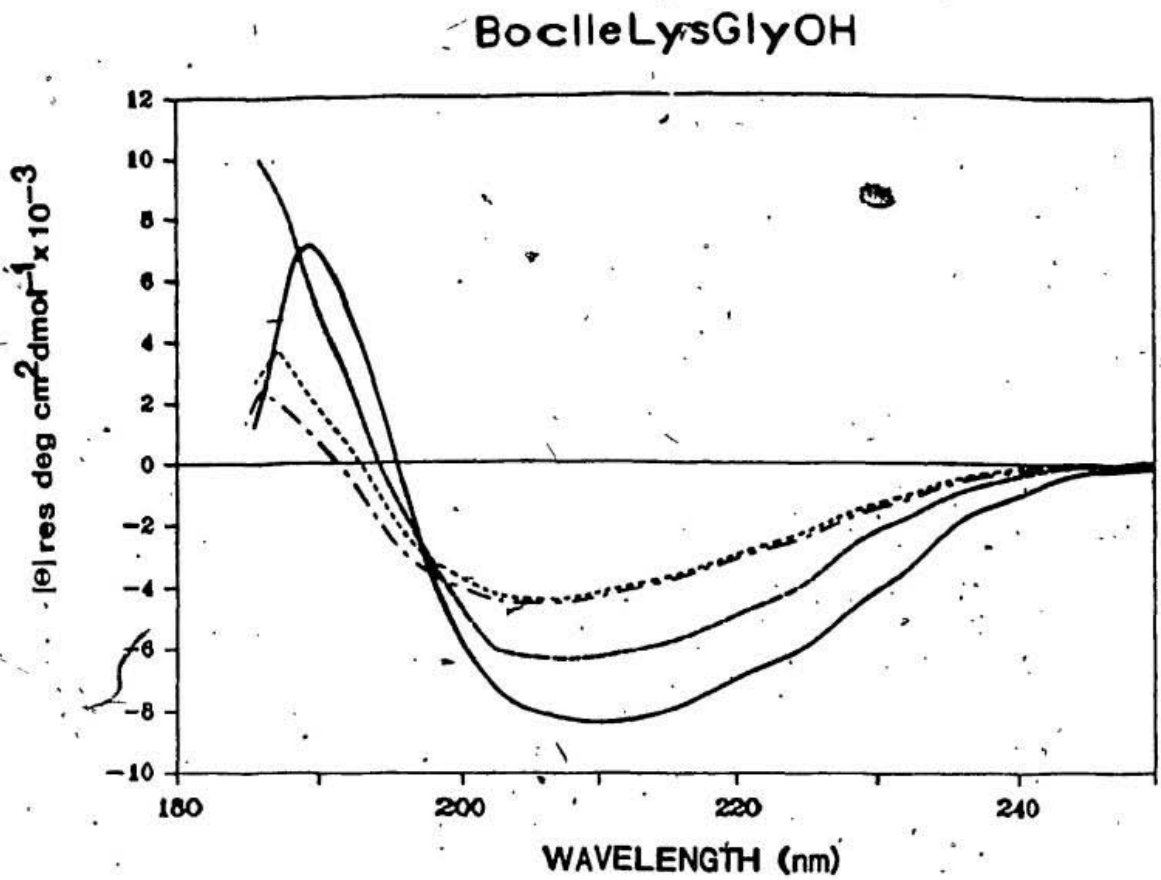

FIG. 4-14: CD spectra of BocIleLysGlyOH in MeOH, (——); TFE, (——); H_2O (- —); and NH_4HCO_3 (----).

4.2.2. CD spectra of the tripeptides (BocAlaLysGlyOH, BocIleLysGlyOH and BocAlaGlyLysOH).

The CD spectra of these three peptides in MeOH, TFE, water and aqueous ammonium bicarbonate are shown in Figures 4-14, 4-15 and 4-16. It can be seen that the spectra of BocAlaLysGlyOH and BocIleLysGlyOH are similar in their overall characteristics while that of BocAlaGlyLysOH is markedly different. In TFE and MeOH, both BocAlaLysGlyOH and BocIleLysGlyOH exhibit a minimum at 202-210 nm and a maximum around 182-185 nm. One could also notice a shoulder at a longer wavelength (above 215 nm) which makes the curves non-symmetrical (i.e. non-Gaussian) in the 195-240 nm region. Such a shoulder around 225 nm and a minimum around 212 nm has indeed been observed in cyclo(ProGly)$_3$ in acetonitrile by Madison *et al.*, (1979, see Table II of this paper). who have characterized this cyclic peptide extensively and found it to have a γ-turn (i.e a C_7 structure). Another peptide containing a γ-turn and which exhibits CD spectra closely similar to our tripeptides is N-Acetyl-alanine-N-methyl amide reported by Crippen and Yang (1974) and by Madison and Kopple (1980). The latter authors have found that in water, the γ-turn of this peptide is fully destabilized giving rise to a CD spectrum with a single minimum around 195 nm. The CD spectra of the tripeptides in H_2O and NH_4HCO_3 in Fig. 4.1 and 4.2 do not, however, show a negative extremum below 200 nm, although the magnitude of the positive band around 185-188 nm is seen to be diminished. This would indicate that the γ-turn is not completely disrupted in the polar solvents but is admixed with significant amounts of the disordered conformation. The presence of the γ-turn in these peptides in relatively less polar solvents was also indicated by the IR data presented in the earlier section (4.1.2.).

In contrast to BocAlaLysGlyOH and BocIleLysGlyOH, the CD spectra of BocAlaGlyLysOH in MeOH and TFE are characterized by a shallow minimum at

220 nm and a maximum between 197-200 nm. A second minimum at a shorter wavelength (185 nm) is observable in TFE. In the more polar solvents, a broad minimum centered around 214 nm is seen and another minimum appears at a shorter wavelength (188-189 nm). The difference in magnitudes of the extrema between the less polar and polar solvents is significant and therefore indicates the presence of an ordered structure in the alcoholic solvents. It is difficult to assess the nature of the ordered conformation since we do not have any IR data available on this peptide. (The IR of BocAlaGly(cbz)LysOMe could not be measured because the peptide was hygroscopic and could not therefore be accurately weighed). However, on the basis of the IR data on an analogous tetrapeptide namely BocAlaGly(cbz)LysHyPOMe (see Section 4.1.3) and on the basis of the similarity of the CD data of the tetrapeptide BocAlaGlyLysHyPOH (see Section 4.2.3), we might assume that the carbonyl of the Boc group is not involved in hydrogen-bonding. This leaves us with the γ-turn (C_7) conformer as the possible ordered conformation present in the alcoholic solvents. The positions of the minimum and maximum relate well with those of the tetrapeptide BocAlaGlyLysHyPOH. In this context, it is worth noting that the ratio (R) of the magnitude of the π-π^* CD band around 200 nm to that of the n-π^* band (in the 215-225 nm) has been correlated with the extent of ordered structure in β-turn containing tripeptides (Brahmachari *et al.*, 1982). A value of R=12, for example, is indicative of a nearly 100% β-turn while R=4 indicates a nearly random structure. If one were to extrapolate this observation to the γ-turn, we might consider the tripeptide with an R value of 8 (i.e. ratio of $[\Theta]_{195}/[\Theta_{220}]$) to have more γ-turn than the tetrapeptide with an R value of about 1. The minimum observed at 220 nm is indicative of the presence of a γ-turn in the alcoholic solvents which is significantly reduced in H_2O and aqueous 0.1M

NH_4HCO_3 (Crippen and Yang, 1974).

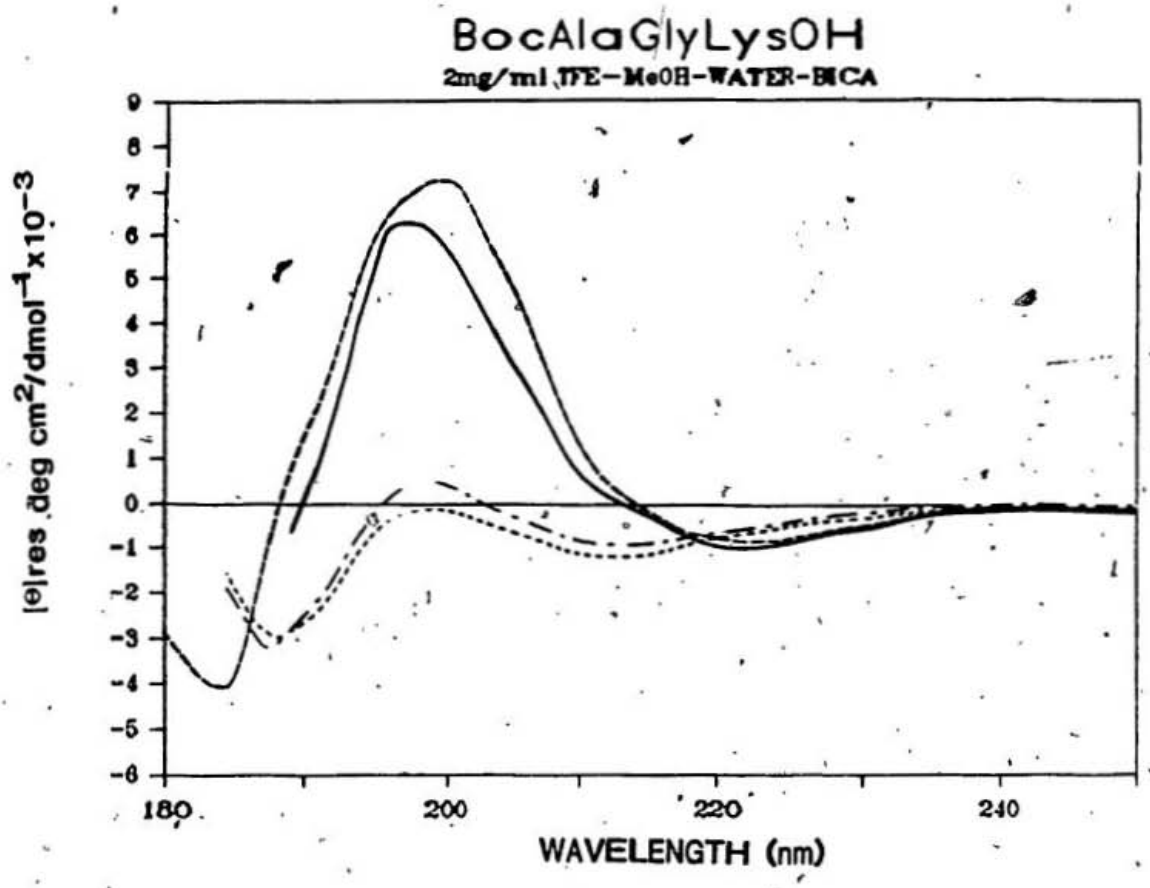

FIG. 4-16: CD spectra of BocAlaGlyLysOH in MeOH, (——); TFE, (----); H_2O (– · –) and NH_4HCO_3 (-----).

4.2.3. CD spectra of Tetrapeptides: (BocAlaLysGlySerOH and BocAlaGlyLysHyPOH)

Figures 4-17 and 4-18 show the CD spectra of BocAlaLysGlySerOH and BocAlaGlyLysHyPOH in MeOH, TFE, H_2O and aqueous 0.1M NH_4HCO_3. The spectral features of the two tetrapeptides are found to be markedly different from each other (see below).

In MeOH and TFE, the CD spectra of BocAlaLysGlySerOH are characterized by two minima at about 218 nm and 194-196 nm and a maximum at 184-186 nm. On the other hand, in H_2O and aqueous 0.1M NH_4HCO_3, only a single negative band at 193 nm is observed. Assuming that the structures of this peptide in the latter set of (more polar) solvents is a disordered one, we might conclude that the peptide has an ordered structure in the alcohols which, however, may be admixed, with the disordered structure in equilibrium. It is, however, difficult to judge the nature of the ordered structure from the CD spectral features of the peptide in TFE and MeOH, since the spectra do not fully resemble any of the published CD spectra of linear or cyclic peptides in ordered conformations such as the γ- or β-turn (Madison and Kopple, 1980; Rose *et al.*, 1985; Kawai and Fasman, 1978; Hollosi *et al.*, 1985; Crisma *et al.*,). We could, however, arrive at a plausible structure for this peptide if we assume that the observed spectrum in TFE (or MeOH) is the resultant of the contributions of the γ- and β-turn (and, possibly, the disordered structure). Thus, the trough around 218 nm can be attributed to the Type I β-turn (Rose *et al.*, 1975). The broad shallow nature of this trough could be taken to indicate the presence of the γ-turn as in the precursor peptide BocAlaLysGlyOH which exhibits a broad minimum band in the 202-206 nm region. The minimum seen around 185 nm may arise from the contribution of disordered structure which is found in the

aqueous solvents (Figure 4-17).

This assessment of the peptide structure would be compatible with the IR data of the (blocked) peptide in $CHCl_3$ (Section 4.1.4) which indicated the presence of both the γ- and β-turn in $CHCl_3$. In the relatively less non-polar solvents like TFE and MeOH used in the CD experiments, one could also expect the presence of some disordered structure as well. It is interesting to note that conformational energy calculations of a similar peptide, Ac-Ala-Lys-Gly-Ser-$NHCH_3$ (made in collaboration with Professor Scheraga's group, 1986), have shown that this peptide exists in solution as an ensemble of conformations, without a uniquely preferred low-energy structure. However, these workers predicted that a type II β-turn is one of the major components of the ensemble of conformations.

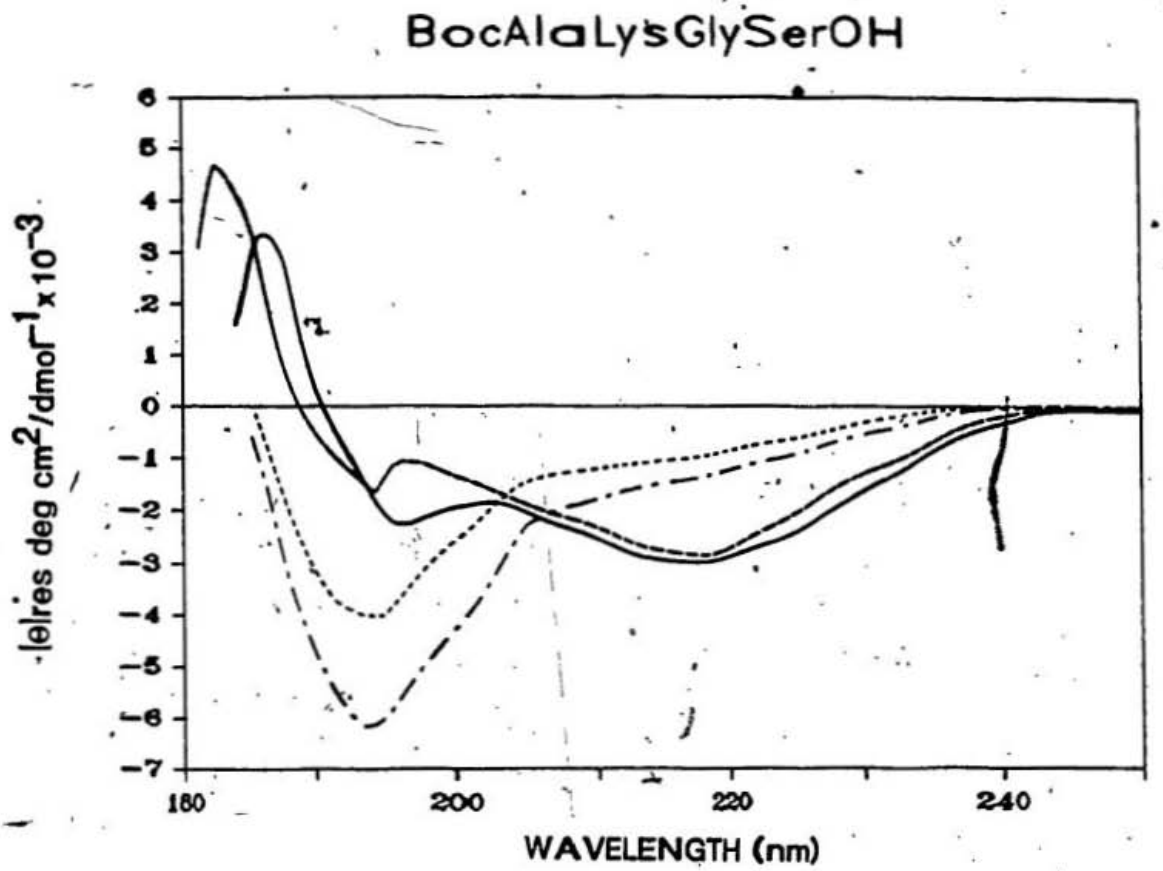

FIG. 4-17 CD spectra of BocAlaLysGlySerOH in MeOH, (——); TFE, (– – –); H_2O (– · –) and NH_4HCO_3 (-----).

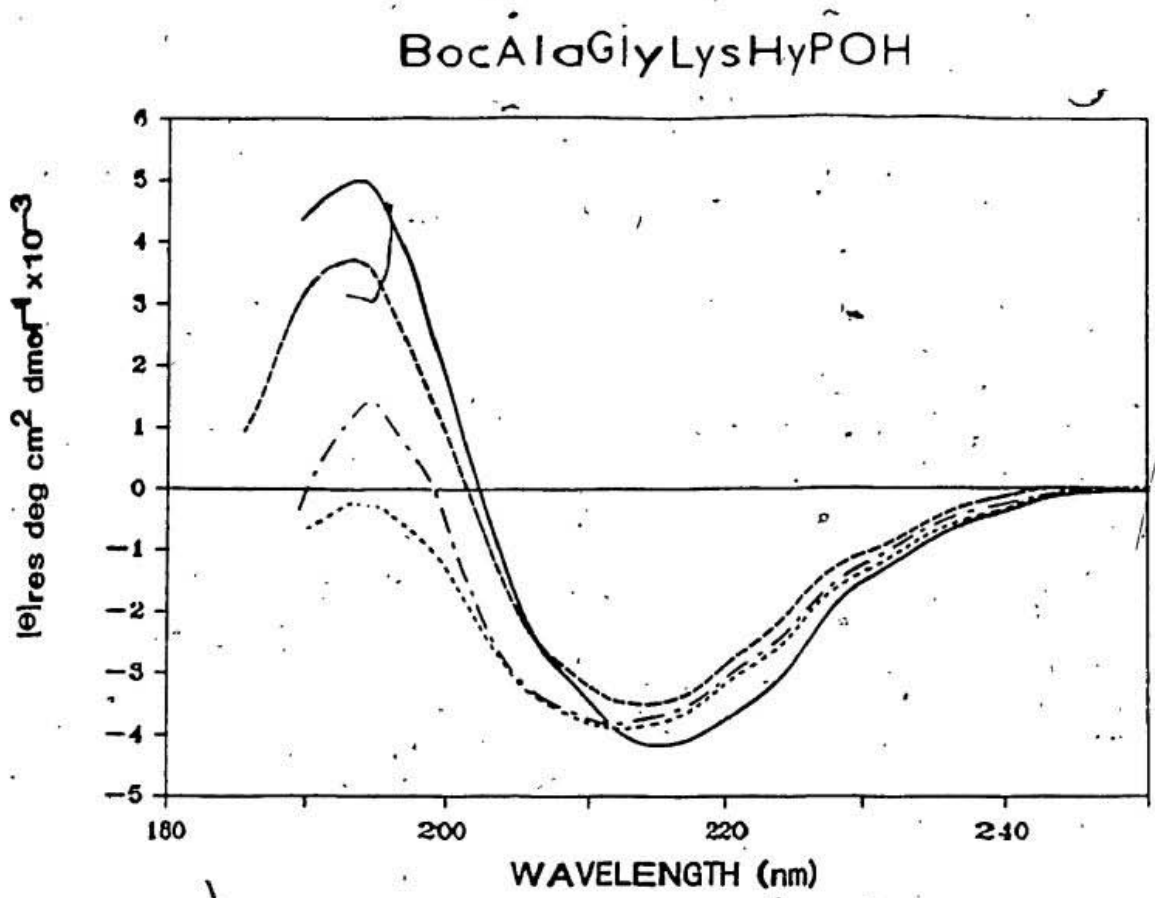

FIG. 4-18: CD spectra of BocAlaGlyLysHyPOH in MeOH, (—); TFE, (----); H_2O (— ·) and NH_4HCO_3 (·····).

The CD spectra of BocAlaGlyLysHyPOH (Figure 4-18) show a minimum at 214 nm and a maximum at 193 nm in MeOH and TFE. In MeOH, a weak shoulder around 220 nm is also visible. In H_2O and aqueous 0.1M NH_4HCO_3, the minimum is shifted to shorter wavelength (210 nm) while the maximum is shifted to longer wavelength (195 nm). Comparing these CD spectral features of this tetrapeptide with those observed for the tripeptide BocAlaGlyLysOH (Figure 4-16), we find that the change from less-polar to more polar solvents causes a more pronounced change (with decreased ellipticities at the extrema) in the spectra of the tripeptide than it does in the tetrapeptide. This would indicate that the ordered structure, if any, in the tetrapeptide is not as populated in the less polar solvents as in the case of the tripeptide. In agreement with this, the R value of the tetrapeptide (i.e. the ratio of $[\Theta_{194}/\Theta_{214}]$ is only 1 compared to the R value of 8 for the tripeptide. Ignoring the relatively minor differences in the positions of the extrema between the tri- and tetrapeptides, we may conclude that the tetrapeptide BocAlaGlyLysHyP has a smaller population of the γ-turn conformation relative to the tripeptide BocAlaGlyLysOH. This presence of the γ-turn (not involving the Boc carbonyl group) is in accord with the IR results (Section 4.1.3).

4.2.4. CD spectra of Hexapeptides: BocLeuHyPGlyAlaLysGlyOH and BocAlaArgGlyIleLysGlyOH

The CD spectra of the above two hexapeptides in the two sets of solvents are very different from each other (Figures 4-19 and 4-20).

In the alcoholic solvents, the CD spectra of BocLeuHyPGlyAlaLysGlyOH show a maximum between 204-208 nm with a broad shoulder in the 210-220 nm region and a minimum between 184-192 nm. In the more polar solvents, the maximum at 216 nm is significantly reduced in magnitude while the magnitude of the minimum is larger than that in MeOH but smaller than that in TFE. The

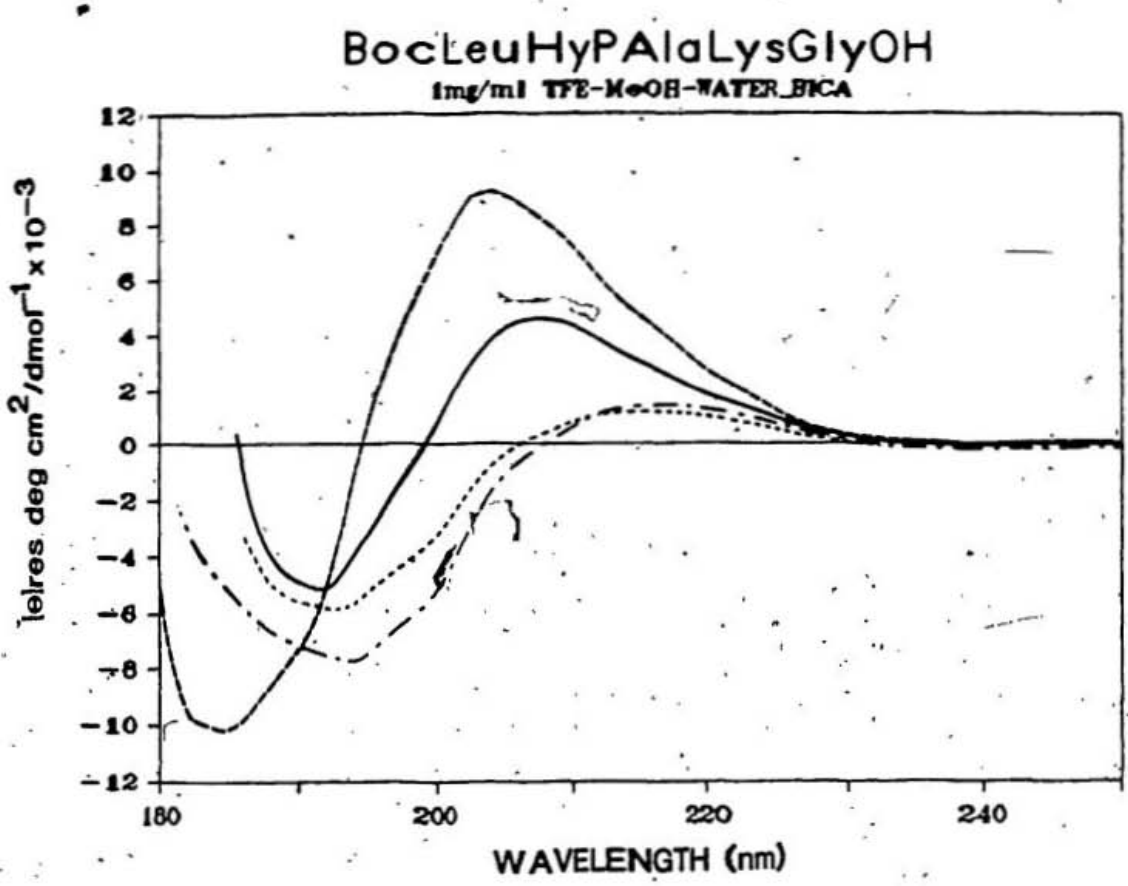

FIG. 4-19: CD spectra of BocLeuHyPGlyAlaLysGlyOH in MeOH, (——); TFE, (— — —); H_2O (– · –) and NH_4HCO_3 (·····).

spectra of BocLeuHyPGlyAlaLysGlyOH do not fully resemble any of the spectra reported in the literature. However, the CD spectral features of this peptide are somewhat comparable to those of cyclo-(L-Ala-L-Pro-Gly)$_2$ (Pease, 1975) which has been shown to adopt a β-turn conformation and exhibits two minima at 223 nm and 188 nm and a maximum at 196 nm in TFE. Bush *et al.*, (1978) have also reported the characteristic CD bands for a type II β-turn contained in the cyclic peptide (cyclo-(Gly-L-Pro-Gly-D-Ala-L-Pro) which showed a minimum at about 232-234 nm and a maximum at 208-210 nm in TFE and MeOH. These comparisons would tend to indicate that the hexapeptide might contain a β-turn (exhibiting the positive π-π* band near 192 nm). This is compatible with the IR data on this peptide (4.1.3.). However, the presence of the γ-turn which is expected to display a shallow negative n-π* band near 230 nm (as in the cyclic peptides cited above) is not immediately obvious from the CD spectra shown in Fig. 4-19. It is likely that this band is masked by the positive n-π* contribution of the β-turn which is found in peptides like cyclo(ProGly)$_3$ (Madison *et al.*, 1974). We therefore have to rely on the IR data and model building which, in conjunction with the CD data presented here, indicate a conformation shown in Fig (4-9), where a β-turn in the LeuHyPGlyAla sequence and a γ-turn in the AlaLysGly sequence (as it occurs in the tripeptides) are found to be present.

In contrast to BocLeuHyPGlyAlaLysGlyOH, BocAlaArgGlyIleLysGlyOH exhibits a broad minimum at 204 nm with a shoulder at about 235 nm and a maximum at 185-188 nm in TFE and MeOH. Nearly the same CD spectral features are maintained in H_2O except for the decrease in magnitudes of the minimum and maximum. The CD behaviour of this hexapeptide in the three solvents is very similar to that of the tripeptides BocAlaLysGlyOH or BocIleLysGlyOH. We can therefore interpret the CD data of BocAlaArgGlyIleLysGlyOH

in an analogous fashion to that of the tripeptides. We thus conclude that BocAlaArgGlyIleLysGlyOH adopts a γ-turn in the alcohols which appears to be weakened in H_2O. This conclusion is in good agreement with the IR data of the blocked hexapeptide in $CHCl_3$.

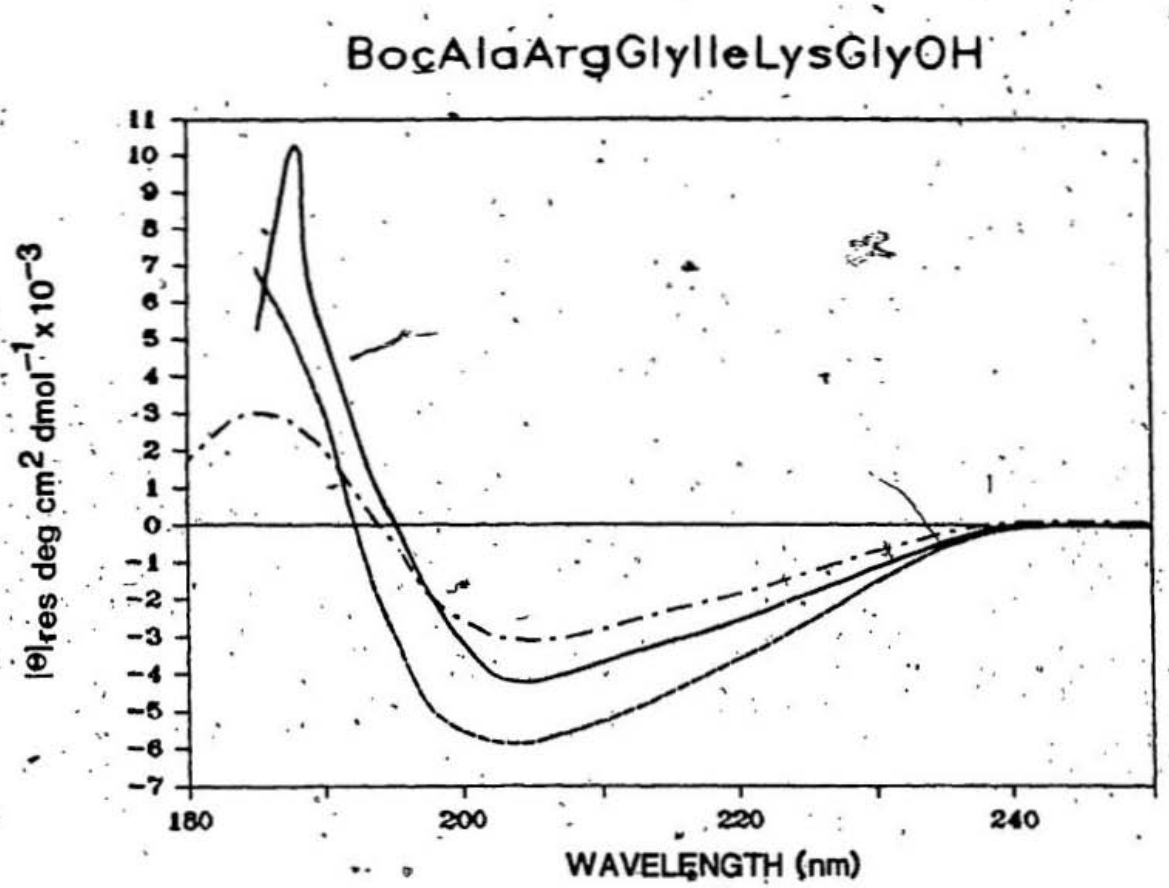

FIG. 4-20: CD spectra of BocAlaArgGlyIleLysGlyOH in MeOH, (——); TFE, (- - -); and H_2O (– · –).

4.2.5. CD spectra of Heptapeptide: BocLeuHyPGlyAlaLysGlySerOH

Fig. 4-21 shows the CD spectra of BocLeuHyPGlyAlaLysGlySerOH in MeOH, TFE and H_2O. In the alcohols, two minima are observed at 230-238 nm and 188-198 nm and a maximum occurs at 212-218 nm. In H_2O, a maximum is observed at 218 nm while the minimum appears at 194 nm. It may be noted that the CD spectra of this peptide in the three solvents are distinctly different from each other. This is unlike the case of the other peptides described so far in which the spectra in TFE and MeOH are quite similar to each other and can be distinguished from those in the more polar aqueous solvents. One may therefore infer the presence of a larger population of conformers in equilibrium in the heptapeptide than found in the other peptides (Crisma *et al.*, 1984). In TFE and MeOH, the presence of the negative n-π^* band around 230-238 nm is indicative of the the presence of a γ-turn which appears to be disrupted in H_2O. Since no IR data is available for this heptapeptide, our interpretation of the CD data of this peptide is based on the IR data of the hexapeptide BocLeuHyPGlyAla(cbz)LysGlyOMe and the available CD data of related peptides. We shall thus assume that the carbonyl of the Boc group is not involved in the hydrogen bonding. Although the CD spectra of the heptapeptide do not bear any immediate resemblance to those of the hexapeptide, a closer look reveals some similarities between the two. The CD spectral features of the heptapeptide in TFE are similar to those of the hexapeptide (BocLeuHyPGlyAlaLysGlyOH) in MeOH (Figure 4-19) except that the long-wavelength band (230 nm) observed in the former peptide is not seen in the hexapeptide. Also, the magnitudes of the extrema in BocLeuHyPGlyAlaLysOH are significantly larger than those observed in BocLeuHyPGlyAlaLysGlySerOH. The maximum at 212 nm in the heptapeptide, like in the case of the hexapeptide can be assigned to the presence of a β-

turn but, as indicated by the magnitude of this band, the β-turn conformation is not as dominant over other structures (such as random or γ-turn) as it is in the case of the hexapeptide where it is so dominant that it overshadows the presence of a γ-turn. Again, the CD spectral features in MeOH and H_2O of BocLeuHyP-GlyAlaLysGlySerOH resemble those of BocLeuHyPGlyAlaLysGlyOH in H_2O except that the 238 nm band seen in MeOH in the heptapeptide is not observed in the H_2O spectra of both peptides. In view of our earlier analysis of the CD spectra of the hexapeptide in H_2O (4.2.4), we can conclude that the β-turn (at the HyP-Gly segment) in the heptapeptide is significantly destabilized in MeOH and H_2O (more so in TFE). Finally, in H_2O, the γ-turn (at the Lys-Gly segment) is also destabilized in comparison with MeOH and TFE. It is interesting to note that the CD spectra of BocLeuHyPGlyAlaLysGlySerOH can be regarded as β-turn in the class D category (negative extremum above 225 nm and between 190-200 and a positive band between 210-220 nm) according to the classification of Woody (1974).

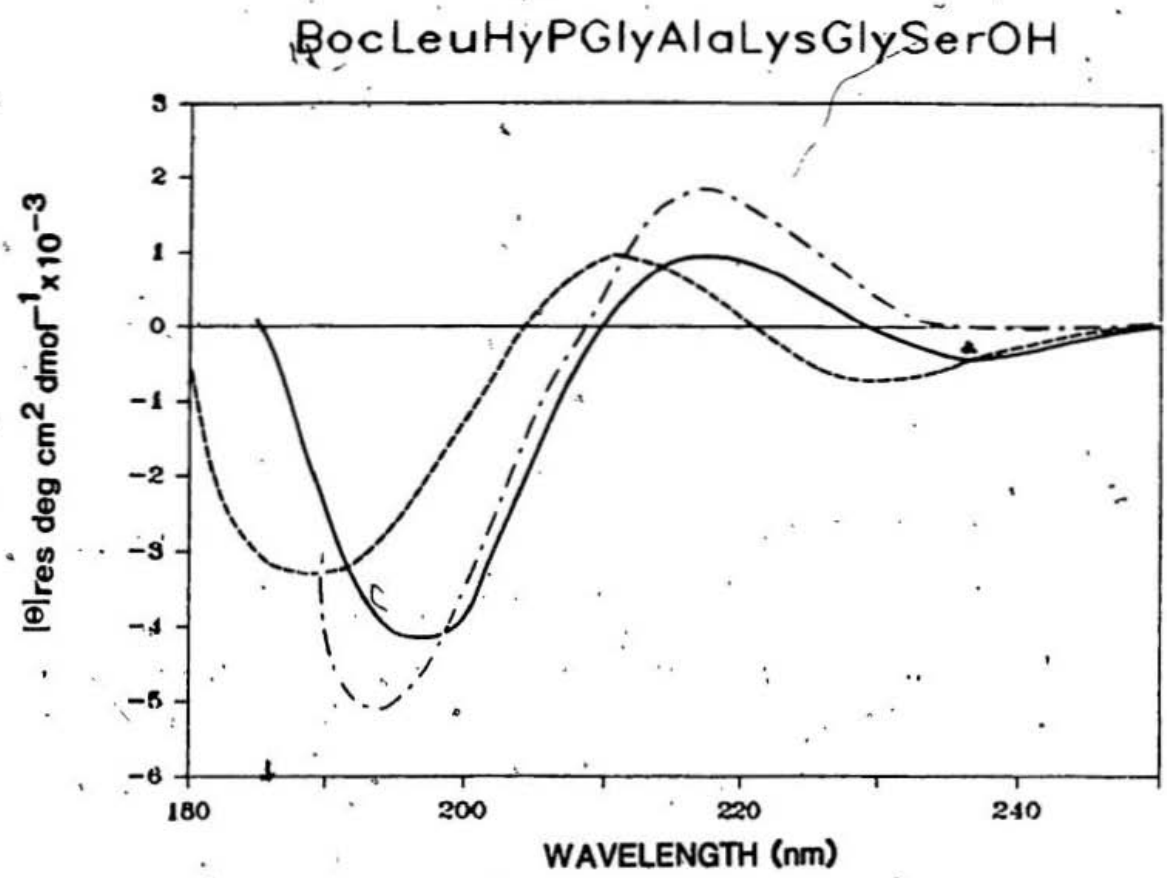

FIG. 4-21: CD spectra of BocLeuHyPGlyAlaLysGlySerOH in MeOH, (——); TFE, (----); and H_2O (·—)

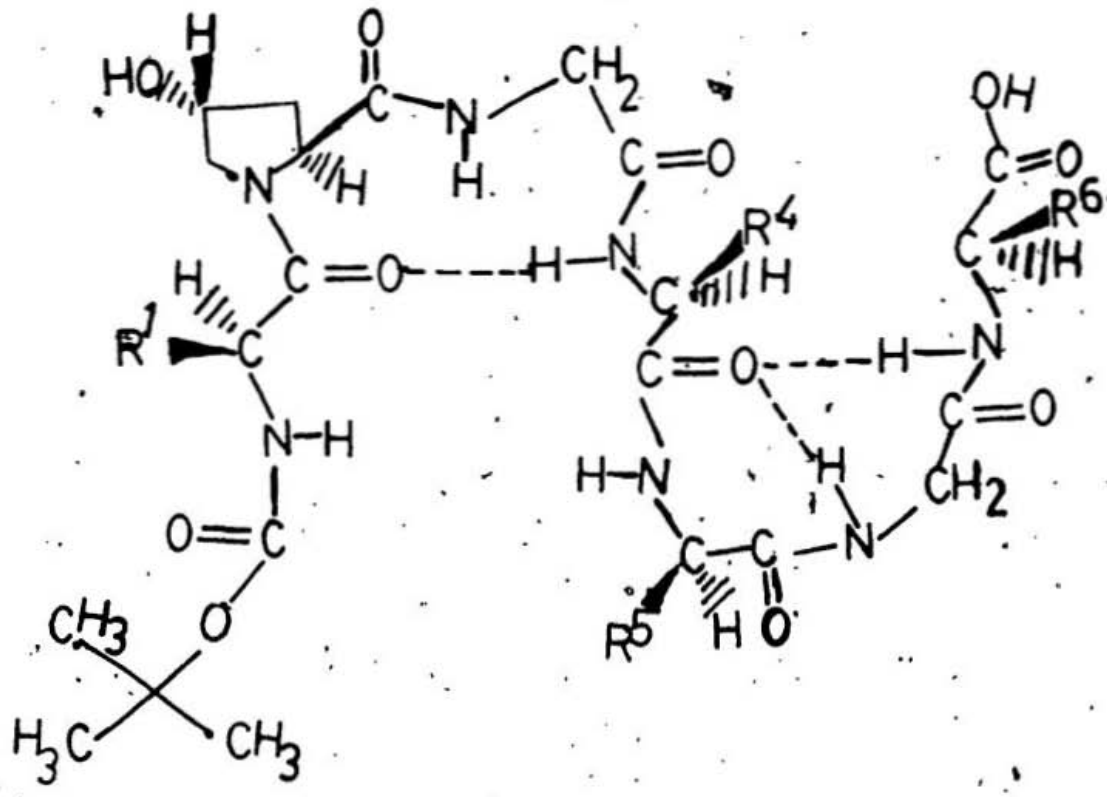

FIG. 4-22: Schematic Diagram of BocLeuHyPGlyAlaLysGlySerOH

Hollosi *et al.*, (1985) have reported the CD structural features, of cbzGlyProSer(OBu^t)GlyOSt which are qualitatively similar to those of our heptapeptide, *viz* a weak negative band at 231 nm, a positive band at 216 nm and a second negative band at 195 nm in cyclohexane. In acetonitrile, the positive band appears at 216 nm, the long-wavelength negative band is absent while the short-wavelength band is shifted below 195 nm with decreased intensity. It must be pointed out that the magnitudes of the maxima and minima of this peptide are different from those of our heptapeptide. Hollosi and co-workers (1985) assigned their tetrapeptide as adopting a β-turn conformation with CD features of the class D spectra of Woody (1974). In view of these data, we suggest that the heptapeptide (BocLeuHyPGlyAlaLysGlySerOH) adopts a β-turn (formed by the

Leu-HyP-Gly-Ala segment) followed by a β-turn and γ-turn (both formed by the Ala-Lys-Gly-Ser segment) (Figure 4-21).

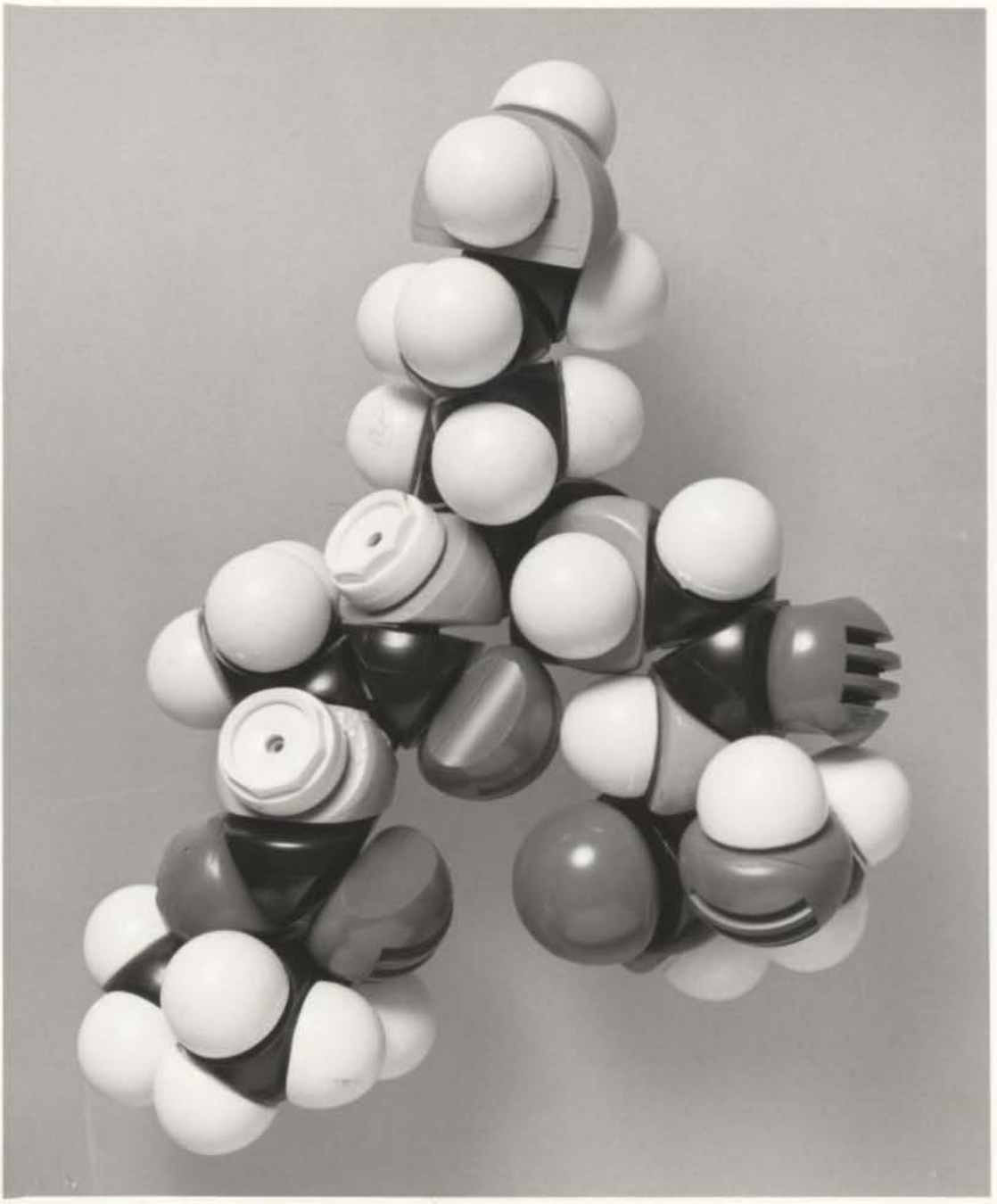

FIG. 4-23: CPK model of BocAlaLysGlyOH

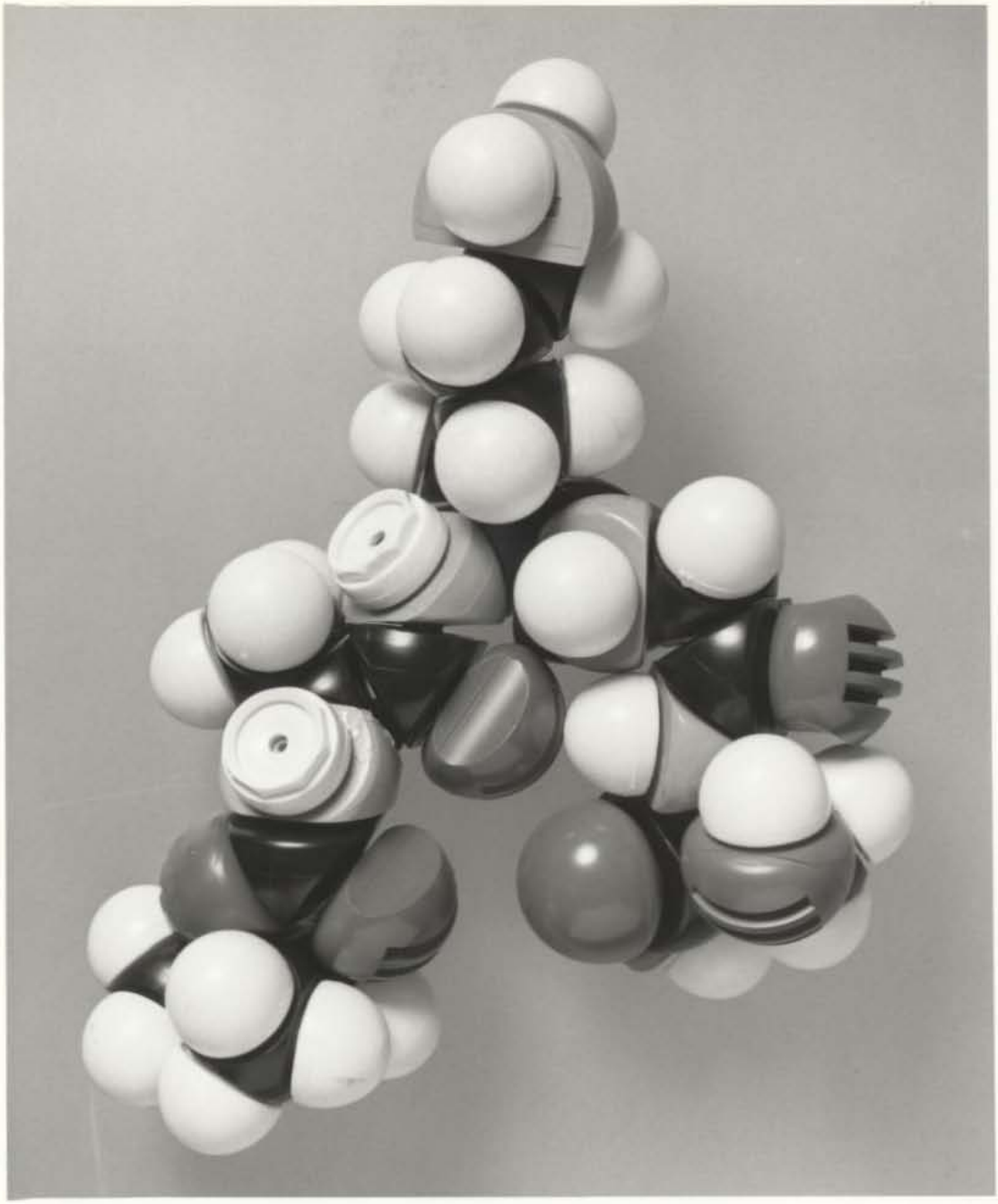

FIG. 4-24: CPK model of BocAlaLysGlySerOH

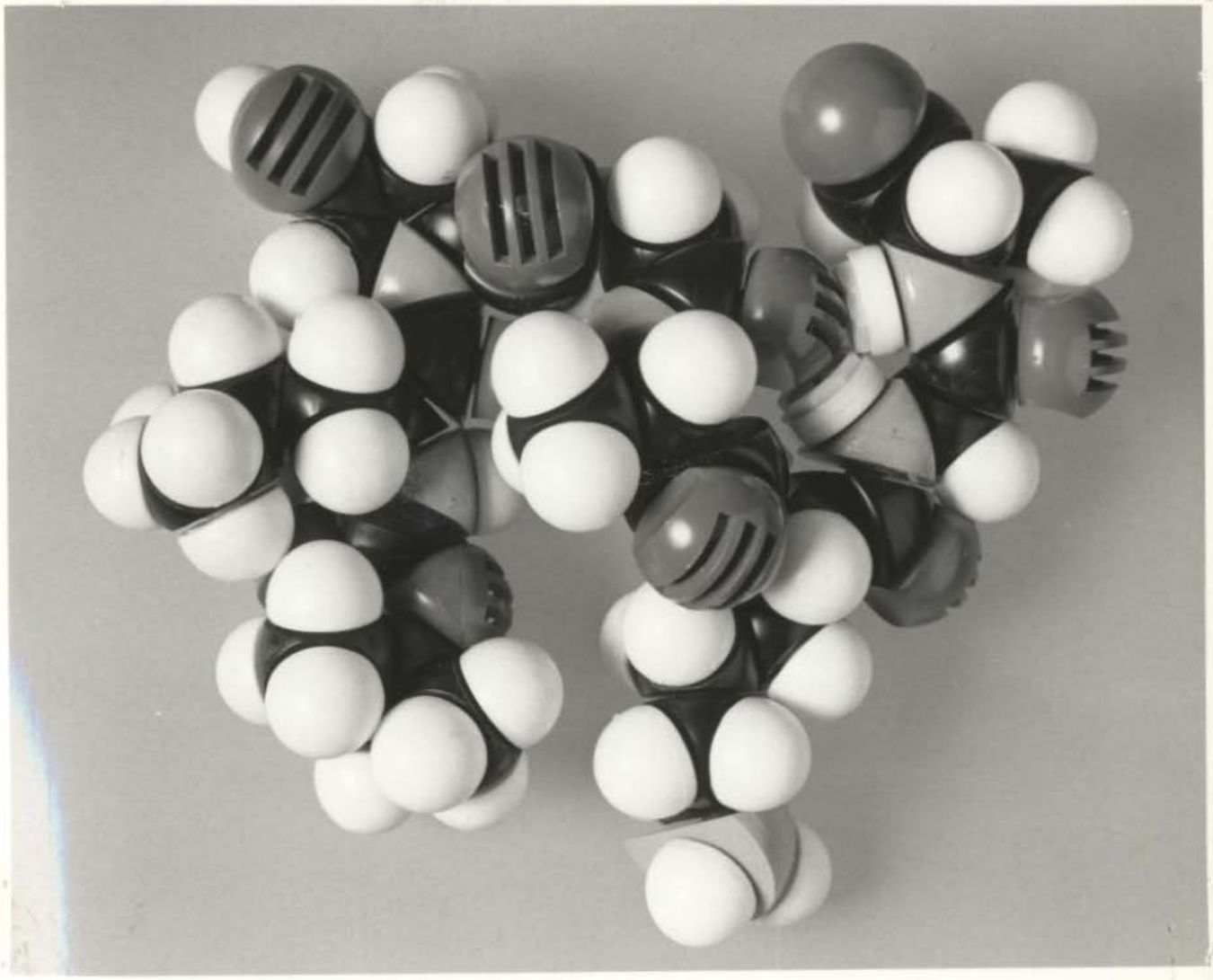

FIG. 4-25: CPK model of BocLeuHyPGlyAlaLysGlySerOH

Chapter 5

5. Hydroxylation Data and Discussion

5.1. Introduction

In order to arrive at a possible conformational basis of the substrate specificity of lysyl hydroxylase, the synthetic lysine-containing peptides were studied for their interaction with the enzyme isolated from chick embryo. The choice of these peptides for our study was made on the basis of the following considerations.

First, in the collagen triplet repeat sequence -Gly-X-Y-, HyL is found only in position Y. Where Lys occurs in the X position, as in the sequence 57-61 in collagen from calf, chicken and rat (types $\alpha 1$(I), $\alpha 1$(II), $\alpha 1$(III) and α(2), it is not hydroxylated (Bornstein and Traub, 1979; see Table 5.1a below). An analysis of the hexa-peptide segments -Y-Gly-X-HyL-Gly-X′- in collagen from the above sources showed (Tables 5.1 a and b) that HyP is the most frequently occurring residue in the Y position of this sequence while Ala occurred most frequently in the X position (i.e. on the N-terminal side of HyL). In collagen sequences where the HyL is subsequently glycosylated, the X position is usually occupied by Ile or His (Kivirikko et al., 1972). In view of these data, the synthetic peptides to be used as substrates for lysyl hydroxylase were designed such that the amino acid preceding Lys (i.e. in the X position) was Ala in six peptides and Ile in two (Table 5.1c). In one of the two hexapeptides, Arg was selected at the Y position while in the other hexapeptide and in the heptapeptide, Hyp was selected at this position. As for the X′ position, Ser was selected in one of the tetrapeptides as well as in the heptapeptide following the oligopeptide sequences used by Kivirikko *et al.* (1972).

One tripetide and one tetrapeptide were synthesized with the Gly-Lys sequence, i.e. with Lys in the X position.

Since the procedure for obtaining a highly purified lysyl hydroxylase is rather tedious and yields relatively very low quantities of the pure enzyme (Kivirikko *et al*, 1980), we have chosen to use a partially purified preparation of the enzyme as discussed in Section 2.2.7. The use of the partially purified

-- p. 116

enzyme (which may contain prolyl hydroxylase as an impurity) in conjunction with synthetic peptides is permissible since, unlike the protocollagen substrate (which contains both susceptible lysyl and prolyl residues) the peptides do not contain any prolyl residues. It is, however, necessary to assume that the assay method employed is specific for HyL. Four different assay methods have been used in earlier studies on lysyl hydroxylase. (a) One method is the radiochemical procedure using [^{14}C]lysine-labelled protocollagen as a substrate (Ryhanen, 1976; Kivirikko *et al.*, 1972; Ryhanen and Kivirikko 1974); the major disadvantages of this method are the cost of and the impurity of the labelled protocollagen. (b) The second assay is the chemical method for measuring HyL formed by a peptide substrate (Blumenkrantz and Prockop 1969) is rather tedious and insensitive requiring large amounts of enzyme. (c)The third is an assay based on the stoichiometric decarboxylation of 2-[1-^{14}C]oxoglutarate during the reaction (Kivirikko *et al.*, 1972; Turpeenniemi-Hujanen *et al.*, 1980, 1981; Myllya *et al.*, 1984; Kivirikko *et al.*, 1973; Ryhanen, 1976) using synthetic peptides as substrates; this assay is convenient but is most useful for purified or at least partially purified lysyl hydroxylase. In less pure extracts, the $^{14}CO_2$-release assay is relatively nonspecific due to the decarboxylation of α-ketoglutarate either by the more active protocollagen prolyl hydroxylase (assuming a suitable cosubstrate is present) or by a number of other decarboxylases able to act on α-ketoglutarate (Prockop *et al.*, 1976). (d) The fourth method involves the measurement of tritium released to water during the hydroxylation reaction using L-[4,5-^{3}H]lysine-labelled protocollagen (Puistola *et al.*, 1980; Puistola, 1982; Miller and Verner, 1979; Ryhanen, 1976; Glass, *et al.*, 1985), or labelled synthetic peptides (Glass *et al.*, 1985) as substrates. As pointed by Glass (*et al.*, 1985), the use of a labelled synthetic substrate has an additional advantage over non-labelled synthetic

peptides; its more specific since HyL is measured in the form of tritium released to water during the hydroxylation reaction. Because of the specificity of the measurement of HyL, the labelled peptide can be used for assaying lysyl hydroxylase in relatively crude extracts (Glass *et al.*, 1985).

Table 5-1: (a). *Amino acid sequences around HyL in collagen*

Type	Residue no.	Sequence (Y-Gly-X-HyL-Gly-X')
α1(I)	84-90	-HyPGlyMetHyLGlyHis-
α1(II)	84-90	-HypGlyValHyLGlyHis-
α1(III)	84-90	-HypGlyMetHyLGlyHis-
α2	84-90	-HyPGlyPheHyLGlyIle-
α1(II)	96-102	-AspGlyAlaHyLGlyGlu-
α2	216-222	-HyPGlyAlaHyLGlyAla-
α2	105-111	-HyPGlyValHyLGlyGlu-
α1(II)	169-175	-AlaGlyAlaHyLGlyGlu-
α1(II)	215-221	-HyPGlyAlaHyLGlySer-
α2	261-267	-AlaGlySerHyLGlyGlu-
α1(II)	249-254	-LeuGlyProHyLGlyG-
α1(II)	405-411	-HyPGlyProHylGlyAla-
α1(II)	417-423	-AlaGlyGluHyLGlyLeu-
α1(II)	601-606	-XxxGlyGluHyLGlyGlu-
α1(II)	651-657	-AlaGlyAlaHyLGlyAsp-
α1(II)	642-648	-ProGlyAlaHyLGlyGlu-
α2	641-647	-HyPGlyAlaHyLGlyGlu-
α1(I)	678-684	-HyPGlyProHyLGlyAla-
α1(I)	912-918	-ArgGlyAspHyLGlyGlu-

α1(I)	923-930	-ArgGlyIleHyLGlyHis-
α1(II)	528-534	-AspGlyProHyLGlyAla-
α1(II)	561-567	-AlaGlyProHyLGlyAsp-
[α1(II)	56-61	-GluAlaGlyLysHyPGly-] non-hydroxylated

Taken from Bornstein and Traub, 1979. (Xxx = unknown a.a.)

Table 5-1 (b) *Analysis of the frequency of occurrence of a.a.s in -Y-Gly-X-HyL-Gly-X' in collagen*

	X	Y	X'
HyP	0	11	0
Ala	7	5	4
Arg	0	2	0
Asp	1	2	2
Pro	5	1	0
Leu	0	1	1
Val	2	0	0
Met	2	0	0
Phe	1	0	0
Ser	1	0	3
Glu	1	0	8
Ile	1	0	1
His	0	0	4

The use of synthetic peptides as substrates has advantages over the use of crude preparations of either non-labelled or [^{14}C]lysine- or[^{3}H]lysine-labelled protocollagen (Kivirikko and Myllya, 1982; Blumenkrantz and Prockop, 1971; Miller, 1972; Peterkosky and Diblasio, 1975). These include greater stability and purity of the peptides as compared to crude preparations of protocollagen and the availability in larger quantities of the peptides. Since synthetic peptides are well defined substrates with known site(s) of hydroxylation, they are useful in mechanistic and conformational studies.

Table 5-1: (c) **Synthetic Peptides**

N$^{\alpha}$tBocAlaLysGlyOH

N$^{\alpha}$tBocIleLysGlyOH

N$^{\alpha}$tBocAlaGlyLysOH

N$^{\alpha}$tBocAlaLysGlySerOH

N$^{\alpha}$tBocAlaGlyLysHyPOH

N$^{\alpha}$tBocAlaArgGlyIleLysGlyOH

N$^{\alpha}$tBocLeuHyPGlyAlaLysGlyOH

N$^{\alpha}$tBocLeuHyPGlyAlaLysGlySerOH

5.1.2. Hydroxylation

The main objective of this thesis was to assess whether a given peptide was hydroxylated or not and to relate this information to the conformation of the peptide. Because of the experimental uncertainity involved in our procedures, we cannot unequivocally quantify our hydroxylation data in terms of the amount of HyL produced by the various peptides unless there is a marked difference between the amount of HyL produced by the different peptides. Table 5-2 shows the results of the enzymic hydroxylation of the synthetic peptides. As can be seen from these results all the peptides with the Lys-Gly sequence were hydroxylated to varying degrees while the two peptides (BocAlaGlyLysOH and BocAlaGlyLysHyPOH) with the Gly-Lys sequence were not hydroxylated. Extension of the tripeptide BocAlaLysGlyOH on the C-terminal by one residue to BocAlaLysGlySerOH, facilitates more hydroxylation by a factor of about two. Similarly, extension of the same tripeptide on the N-terminal by three residues to BocLeuHyPGlyAlaLysGlyOH or by four residues to BocLeuHyPGlyAlaLysGlySerOH significantly enhances the degree of hydroxylation by a factor of six and ten, respectively. Kivirikko *et al.*, (1972) also observed that the tripeptide ProGlyLys was not hydroxylated. These workers also showed that the tripeptide IleLysGly was hydroxylated and extension of the same tripeptide on the C-terminal by three residues to IleLysGlyIleLysGly resulted in a ten-fold increase in the rate of hydroxylation.

We will now attempt to interpret the hydroxylation data in relation to the structures of both the hydroxylated and unhydroxylated peptides. Our interpretation will be based on our observations as well as on our knowledge of the conformational requirement of prolyl hydroxylase which as indicated earlier (Section 1.3.3.) shares common properties with lysyl hydroxylase in terms of cofactors and cosubstrates requirement as well as the reaction mechanism.

Table 5-2:

Examinction of synthetic peptides as substrates for partially purified lysyl hydroxylase.

Substrate	HyL synthesized (μmol/ml/hour)		
	$^{14}CO_2$ method	chemical method	Amino acid analysis
N^{α}tBocAlaLysGlyOH	0.005 ±0.0007	0.006 ±0.0014	†
N^{α}tBocAlaGlyLysOH	0 ± 0.001	†	†
N^{α}tBocIleLysGlyOH	0.007 ±0.0017	0.008 ±0.0014	†
N^{α}tBocAlaLysGlySerOH	0.011 ±0.0016	0.014 ±0.002	0.014
N^{α}tBocAlaGlyLysHyPOH	0 ± 0.001	†	†
N^{α}tBocLeuHyPGlyAlaLysGlyOH	0.034 ±0.0044	†	†
N^{α}tBocLeuHyPGlyAlaLysGlySerOH	0.056 ±0.001	†	†

† not determined

Incubations were carried out as described in methods. The μmoles of HyL ($^{14}CO_2$ method) were calculated from the observed dpm of $^{14}CO_2$ trapped and the specific activity of the [1-^{14}C]oxoglutarate. Values are corrected for negative controls which contained no substrates.

The error ranges in the values of HyL synthesized were calculated from averaging three duplicate experiments.

We will now attempt to interpret the hydroxylation data in relation to the structures of both the hydroxylated and unhydroxylated peptides. Our interpretation will be based on our observations as well as on our knowledge of the conformational requirement of prolyl hydroxylase which as indicated earlier (Section 1.3.3.) share common properties with lysyl hydroxylase in terms of cofactors and cosubstrates requirement as well as the reaction mechanism.

The IR and CD data (Chapter 4) show that the tripeptides BocAlaLysGlyOH, BocIleLysGlyOH and BocAlaGlyLysOH all adopt a γ-turn (C_7) conformation in relatively non-polar solvents. In the first two tripeptides, lysine is found in the second position of the γ-turn while in BocAlaGlyLysOH lysine is in the third position of this structure (Figure 4.10). These results may provide the basis for the specificity of lysyl hydroxylase. Although all the tripeptides adopt the same conformation, the enzyme seems to be able to distinguish between lysine in the 2nd position and lysine in the 3rd position of the γ-turn, and will hydroxylate lysine only when it occurs in the second position of the γ-turn (see below). This finding lead us to suggest that the conformational minimal requirement for lysine hydroxylation is a γ-turn in which lysine occupies the second position of this structure i.e. lysine has to be followed by glycine in the repeating triplet sequence X-Y-Gly for enzymic hydroxylation to occur.

Extension of the tripeptide BocAlaLysGlyOH on the C-terminal to BocAlaLysGlySerOH drastically changed both the peptide structure and the degree of hydroxylation. The IR and CD data showed that the tetrapeptide BocAlaLysGlySerOH takes up mainly a β-turn conformation in non-polar solvents. Within the overall geometry of this 4---$>$1 hydrogen-bonded β-turn, a γ-turn with a 3---$>$1 hydrogen-bonding between Ala^1 CO and Gly^3 NH is also possible (Figure 4-10). Thus, the γ-turn found in the the precursor tripeptide

(BocAlaLysGlyOH) is maintained in the tetrapeptide (BocAlaLysGlySerOH) as well. The hydroxylation data shows that BocAlaLysGlySerOH is more hydroxylated (by a factor of two) than the precursor tripeptide BocAlaLysGlyOH. The presence of a β-turn in this tetrapeptide which is not observed in the tripeptide, may account for the observed increase in the degree of hydroxylation but how this is achieved remains to speculation. One plausible explanation is that the β-turn with the 4-->1 hydrogen bond between Ala^1 CO and Ser^4 NH stabilizes the γ-turn and thereby increase the degree of hydroxylation. The β-turn formed by the AlaLysGlySer segment can be envisaged as locking the γ-turn which should result in a greater stability of this γ-turn compared to that present in the tripeptides BocAlaLysGlyOH and BocIleLysGlyOH (Section 4.2.1.).

In order to test the effect of chain length and the possible change in conformation that could result therefrom, the tripeptide BocAlaLysGlyOH was extended at the N-terminal to BocLeuHyPGlyAlaLysGlyOH and BocLeuHyPGlyAlaLysGlySerOH. Both peptides were hydroxylated six to ten times more than the tripeptide with the heptapeptide showing about twice more hydroxylation than the hexapeptide. (Table 5-2). The conformational preference of the hexapeptide in non-polar solvents is a consecutive β- (formed by the LeuHyPGlyAla segment) and γ-turn (formed by the AlaLysGly segment) according to the IR and CD data (Chapter 4). In BocAlaLysGlySerOH, lysine is found in the second position of the β-turn (and hence in the γ-turn that coexists with the β-turn as described above) while in the case of the hexapeptide, lysine is involved in the γ-turn where it also occupies the second position. In the hexapeptide (BocLeuHyPGlyAlaLysGlyOH), however, lysine is not part of a β-turn, but is in the γ-turn present in the AlaLysGly segment. Since the γ-turn is maintained in both the tetra- and hexapeptides, the observed increase in the degree of

hydroxylation may lie in the G^3A^4 segment of the LeuHyPGlyAla. This segment is very close to the γ-turn formed by the AlaLysGly segment and may increase the binding of the enzyme to the substrate. The β-turn formed by the LeuHyP-GlyAla segment may also stabilize the γ-turn in the overall context of a compact ordered structure for the whole peptide but we do not expect this stabilization to be as good as the one afforded by the β-turn of the tetrapeptide which "locks" the γ-turn formed by the AlaLysGly segment resulting in greater stability. This assumption would also explain why the heptapeptide (BocLeuHyPGlyAlaLys-GlySerOH) is more hydroxylated than its precursor hexapeptide BocLeuHyP-GlyAlaLysGlyOH. The heptapeptide, as well as having the G^3A^4 segment of the LeuHyPGlyAla β-turn which as we have suggested above increases the binding of the substrate to the enzyme, has a second β-turn formed by the AlaLysGlySer segment as in the case of the tetrapeptide BocAlaLysGlySerOH which, as discussed above, stabilizes the γ-turn. Our theory of the AlaLysGlySer segment (present in both BocAlaLysGlySerOH and BocLeuHyPGlyAlaLysGlySerOH) as the stabilizer of the γ-turn in the AlaLysGlyOH segment of the heptapeptide is supported by the CD data of this peptide in relatively less polar solvents (4.2.5). While the CD spectral features of the hexapeptide (BocLeuHyPGlyAlaLysGlyOH) (4.2.4), show that the γ-turn is masked by the dominant β-turn, the γ-turn in the heptapeptide is quite evident from the CD spectra.

Chapter 6

6. Discussion and Conclusion

The main objectives of this thesis were four-fold: (a) synthesis of lysine-containing peptides found in the hydroxylated or unhydroxylated region of collagen (b) to test the ability of these peptides to act as substrates of lysyl hydroxylase partially purified from chick embryos (c) to study the structures of these peptides and (d) to find if lysyl hydroxylase recognized specific secondary features in the substrates.

The first objective was accomplished by synthesizing eight lysine-containing peptides which mimic specific segments of the sequences of the collagen α-chains. These peptides vary in length from three to seven amino acid residues long (Chapter 3). Seven of these were tested for their ability to act as substrates of lysyl hydroxylase partially purified from chicken embryos (2.2.10). All the peptides with Lys-Gly sequence were hydroxylated to varying degrees while the two peptides with the Gly-Lys were not. The CD and IR data showed that all the peptides adopt ordered structures in relatively less polar solvents (Chapter 4). In the hydroxylated peptides, a common structure is apparent; namely a γ-turn (C_7) conformation with lysine occupying the second position of this structure. The unhydroxylated peptides (BocAlaGlyLysOH and BocAlaGlyLysHyPOH) also show the presence of the γ-turn, but in this case, lysine is found to occur in the third position of this structure. This is a very interesting finding and provides a clue to the basis of the substrate specificity of lysyl hydroxylase. Lysyl hydroxylase must be very highly specific, our preliminary results seem to indicate that the enzyme can distinguish between a lysyl residue in the second position of the γ-turn and a lysyl residue in the third position of the same structure.

As the peptides were elongated both at the N- and C- terminal a second ordered structure, namely the β-turn became apparent. The appearance of this structure is correlated with an increased degree of hydroxylation. The β-turn may be stabilizing the γ-turn thereby increasing the degree of hydroxylation. It is interesting to note that in the case of the hexapeptide (LeuHyPGlyAlaLysGly), the β-turn formed by the LeuHyPGlyAla segment may not be directly involved in the stabilization of the γ-turn (formed by the AlaLysGly segment), but part of it (Gly^3Ala^4) may increase the binding of the enzyme to the substrate. The correlation between the conformation of the peptide with its susceptibility for hydroxylation is shown in Table 6-1.

Table 6-1:

Summary of hydroxylation data versus conformation

Peptide		Hydroxylation
BocAlaGlyLys	No	γ-turn; lys in position 3 of the γ-turn
BocAlaGlyLysHyP	No	γ-turn; lys in position 3 of the γ-turn
BocAlaLysGly	Yes	γ-turn; lys in position 2 of the γ-turn
BocIleLysGly	Yes	γ-turn; lys in position 2 of the γ-turn
BocAlaLysGlySer	Yes	γ-turn (+β-turn); lys in position 2 of the γ-turn
BocLeuHyPGlyAlaLysGly	Yes	γ-turn (+β-turn); lys in position 2 of the γ-turn
BocLeuHyPGlyAlaLysGlySer	Yes	γ-turn (+2 β-turns); lys in position 2 of the γ-turn

In order to rank the peptides for their degrees of hydroxylation, one really needs to determine their K_m and k_{cat} values. This, however, was not done due to limitation in the time available.

In conclusion, our results suggest that the specific structures, the γ-turn and β-turn may be required for recognition by the enzyme. Though these results

strongly suggest the requirement for specific structures(s), we do not have any proof that these are the structures that bind to the enzyme. Obviously more work could have been done to substantiate some of our speculations but this was, however, beyond the scope of this thesis. For example synthesis of more γ- and/or β-turn containing peptides with lysine in either the X or Y position of the repeating triplet sequence X-Y-Gly would help in the understanding of the substrate specificity of the enzyme. The CD and IR data of γ-turn containing peptides is very scarce. The availability of such data on model peptides would be helpful in the understanding of the specific γ-turn recognized by lysyl hydroxylase. In the case of β-turns, the positional preference as well as the structural preference (e.g. certain amino acids prefer to be in the β-turn conformation rather than random) of amino acids as well as the effect of the amino acids on the stability of β-turns are known. This kind of information is not available for γ-turns.

For future work, it would also be interesting to synthesize tri-, tetra- and hexapeptides where replacements are made to prevent formation of the critical hydrogen bond, for example N-methyl amino acids or replacement of carbonyl by a CH_2 group, *etc.*

Our hydroxylation data with BocAlaLysGlyOH and BocIleLysGlyOH seem to indicate that Ala or Ile occupying the X position preceding HyL has no effect on the degree of hydroxylation. However, because of experimental uncertainity we cannot accurately quantify our hydroxylation data in terms of the amount of HyL produced. It would be very interesting to synthesize more tripeptides of the X-Lys-Gly nature where X is any of the twenty-one amino acids then study their conformation as well as their interaction with lysyl hydroxylase. Such a study should be able to establish the effect, if any of the residue in the X position preceding HyL on the degree of hydroxylation.

One of the major setbacks in lysyl hydroxylase research is the lack of a purification procedure which gives high yields of the pure enzyme. In this regard a synthetic peptide substrate (including for example the heptapeptide used in this study), with a very high affinity for lysyl hydroxylase could be worth trying to use for purifying the

enzyme. This peptide could be coupled to Agarose and the enzyme could be eluted with an excess of a lower affinity substrate followed by a DEAE ion-exchange chromatography step which ~~should~~ afford a pure enzyme.

Our experiments if pursued further could have a medical application. Deficiency in lysyl hydroxylase is known to cause Ehlers-Danlos syndrome type VI. A suitable substrate for lysyl hydroxylase could be synthesized with the lysine labelled and used to diagnose this deficiency even in relatively crude extracts since as mentioned earlier (Section 5 2), the use of a labelled synthetic peptide substrate has an advantage over a non-labelled peptide or protocollagen as HyL produced during the reaction is most accurately quantified in the form of tritium released to water.

These results provide a basis for expanded efforts to evaluate the chemical and conformational specificity of lysyl hydroxylase.

REFERENCES

1. ADLER, A. J., N. GREENFIELD and G. D. FASMAN, *Methods Enzymol.* 27: 675-735, 1973.
2. ANANTHANARAYANAN, V. S., *J. Biomol. Struct. Dyn.* 1: 843-855, 1983.
3. ANANTHANARAYANAN, V. S., S. K. ATTAH-POKU, P. L. MUKKAMALA and P. H. REHSE, *Proc. Internatl. Symp. Biomol. Struct. Interactions, Suppl. J. Biosci.* 8: 209-221, 1985.
4. ANANTHANARAYANAN, V. S. and T. S. CAMERON, *In Press*.
5. ANANTHANARAYANAN, V. S. and N. SHYAMASUNDAR, *Biochem. Biophys. Res. Comm.* 102: 295-301, 1981.
6. ANANTHANARAYANAN, V. S., R. S. RAPAKA, S. K. BRAHMACHARI and R. S. BHATNAGAR, *Biopolymers* 15: 707-716, 1976.
7. ANTINNEN, H., E. O. JARVENSIVU and E. SAVOLAINEN, *Clin. Chim. Acta* 3: 3-8, 1985.
8. ASHWELL, G. and A. MORELL, *Adv. Enzymol.* 41: 99-128, 1974.
9. BAILEY, A. J., S. P. ROBINS and G. BALIAN, *Nature* 251: 105-109, 1974.
10. BAILEY, A. J. and M. SHIMOKOMAKI, *FEBS. Lett.* 16: 86-88, 1971.
11. BAILEY, A. and S. P. ROBINS, *Nature* 251: 105-109, 1974.
12. BALASUBRAMANIAN, D. and C. KUMAR, *Appl. Spectroscopy Rev.* 11: 223-286, 1976.
13. BANDEKAR, J. and S. KRIMM, *Proc. Natl. Acad. Sci. USA* 76: 774-777, 1980.
14. BARNES, M. J. and E. KODICEK, *Vit. Horm.* 30: 1-64, 1972.
15. BARNES, M. J., B. J. CONSTABLE, L. F. MORTON and P. M. ROYCE, *Biochem. J.* 139: 461-468, 1974.
16. BAUSE, E., H. HETTKAMP and G. LEDGER, *Biochem. J.* 203: 761-768, 1982.
17. BAUSE, E., *Biochem. J.* 209: 331-336, 1983.
18. BEELY, J. G., *Biochem. Biophys. Res. Comm.* 76: 1051-1055, 1977.
19. BENEDETTI, E., In: *Peptides*, vol. Proceedings of the fifth American Peptide Symposium, edited by G. M. and M. J.. New York: John Wiley and Sons, 1978, p. 257.
20. BERG, R. A. and D. J. PROCKOP, *J. Biol. Chem.* 248: 1175-1182, 1973.
21. BERZOFSKY, J. A., G. K. BUCKNMEYRE, G. HICKS, F. R. N. GURD, R. J. FELDMAN and J. MINNA, *J. Biol. Chem.* 257: 3189-3198, 1982.
22. BHATNAGAR, R. S. and R. S. RAPAKA, In: *Biochemistry of Collagen*, edited by G. N. RAMACHANDRAN and A. H. REDDI. New York London: Plenum Press, 1976, p. 479-523.
23. BHATNAGAR, S. J. and R. S. RAPAKA, *Adv. Exptl. Med. Biol.* 86: 491-507, 1977.
24. BLUMENKRANTZ, N. and D. J. PROCKOP, *Anal. Biochem.* 30: 377-385, 1969.
25. BLUMENKRANTZ, N. and D. J. PROCKOP, *Anal. Biochem.* 39: 59-64, 1971.
26. BORNSTEIN, P. and W. TRAUB, In: *The Proteins*, vol. IV, edited by L. W. CUNNINGHAM and D. W. FREDERIKSEN. New York: Academic Press, 1979, p. 412-632.
27. BORNSTEIN, P., *Ann. Rev. Biochem.* 43: 567-604, 1974.
28. BOUSSARD, G., M. MARRAUD and M. NEEL, *J. Chem. Phys.* 71: 1081-1091, 1974.
29. BRAHMACHARI, S. K., R. S. RAPAKA, R. S. BHATNAGAR and V. S. ANANTHANARAYANAN, *Biopolymers* 21: 1107-1125, 1982.
30. BRAHMACHARI, S. K., V. S. ANANTHANARAYANAN, S. BRAHMS, J. BRAHMS, R. S. RAPAKA and R. S. BHATNAGAR, *Biochem. Biophys. Res. Comm.* 86: 605-612, 1979.
31. BRAHMACHARI, S. K. and V. S. ANANTHANARAYANAN, *Proc. Natl. Acad. Sci. USA* 76: 5119-5123, 1979.
32. BRAHMS, S. and J. BRAHMS, *J. Mol. Biol.* 138: 149-178, 1980.
33. BURGESON, R. T., F. A. EL-ADLI, I. I. KAILILA and D. W. HOLLISTER, *Proc. Natl. Acad. Sci. USA* 73: 2579-2582, 1976.

34. BURGESS, A. W. and H. A. SCHERAGA, *Biopolymers* 12: 2177-2183, 1973.

35. BUSH, C. A., S. K. SARKAR and K. D. KOPPLE, *Biochem.* 17: 4951-4954, 1978.

36. BUTLER, W. T., *Science* 161: 796-797, 1968.

37. BUTLER, W. T. and L. W. CUNNINGHAM, *J. Biol. Chem.* 241: 3882-3888, 1966.

38. BYERS, P. H., E. M. CLICK, E. HARPER and P. BORNSTEIN, *Proc. Natl. Acad. Sci. USA* 72: 3009-3013, 1973.

39. CARBONE, F. R. and S. J. LEACH, *Internatl. J. Peptide Protein Res.* 26: 225-235, 1985.

40. CARDINALLE, G. J. and S. UDENFRIEND, *Adv. Enzymol.* 42: 255-290, 1974.

41. CHOPRA, R. S., V. S. ANANTHANARAYANAN and P. L. MUKKAMALA, *Biochem. Internatl.* 7 (4): 415-421, 1983.

42. CHOPRA, R. S. and V. S. ANANTHANARAYANAN, *Proc. Natl. Acad. Sci. USA* 79: 7180-7181, 1982.

43. CHOU, P. Y. and G. D. FASMAN, *J. Mol. Biol.* 26: 135-175, 1977.

44. CHRISTNER, P. J. and J. ROSENBLOOM, *J. Biol. Chem.* 246: 7551-7556, 1971.

45. CRAWFORD, J. L., W. N. LIPSCOMB and C. G. SCHELLMAN, *Proc. Natl. Acad. Sci. USA* 76: 5119-5123, 1973.

46. CRAWFORD, J. L., W. N. LIPSCOMB and C. G. SCHELLMAN, *Proc. Natl. Acad. Sci. USA* 70: 538-542, 1973.

47. CRIPPEN, G. M. and J. T. YANG, *J. Phys. Chem.* 78: 1127-1131, 1974.

48. CRISMA, M., G. D. FASMAN, H. BALARAM and P. BALARAM, *Internatl. J. Peptide Protein Res.* 23: 411-419, 1984.

49. DANIELS, J. R. and G. H. CHU, *J. Biol. Chem.* 250: 3531-3537, 1975.

50. DAVIS, N. R. and A. J. BAILEY, *Biochem. Biophys. Res. Comm.* 45: 1416-1422, 1971.

51. DEHM, P. and N. A. KEFALIDES, *J. Biol. Chem.* 253: 6680-6686, 1978.

52. DESLAURIERS, R., D. J. EVANS, Y. C. MEINWALD, E. MINASIAN, G. NEMETHY, I. D. RAE, H. A. SCHERAGA, R. L. SOMORJAI, E. R. STIMSON, J. W. V. NISPEN and R. W. WOODY, *Macromolecules* 14: 985-966, 1981.

53. ENGEL, J., J. KURTZ and E. BERGER, *J. Mol. Biol.* 17: 255-272, 1972.

54. EVANS, D. J., I. D. RAE, E. MINASIAN, G. NEMETHY, H. A. SCHERAGA and S. J. LEACH, *J. Protein Chem.* 2: 77-99, 1983.

55. EYRE, D. R. and M. J. GLIMCHER, *Proc. Natl. Acad. Sci. USA* 69: 2594-2598, 1972.

56. EYRE, D. R. and M. J. GLIMCHER, *Biochem. Biophys. Res. Comm.* 52: 663-671, 1973.

57. EYRE, D. R. and M. A. PAZ, P.M. Gallop. *Ann. Rev. Biochem.* 53: 717-748, 1984.

58. EYRE, D. R. and M. J. GLIMCHER, *Proc. Natl. Acad. Sci. USA* 69: 2594-2598, 1972.

59. FESSLER, L. I. and J. H. FESSLER, *J. Biol. Chem.* 249: 7637-7646, 1974.

60. FIETZEK, P. P., F. W. REXRODT, P. WENDT, M. STARK and K. KUHN, *Eur. J. Biochem.* 30: 163, 1972.

61. FOX, J. A., A. T. TU, V. J. HRUBY and H. I. MOSBERG, *Arch. Biochem. Biophys.* 211: 628-631, 1981.

62. FRASER, R. D. B. and T. P. MACRAE, *Conformation in Fibrous Proteins.* New York: Academic Press, 1973.

63. FREYTAG, J. W., M. OHNO and B. G. HUDSON, *Biochem. Biophys. Res. Comm.* 72: 796-802, 1976.

64. FUJII, K. and M. L. TANZER, *FEBS. Lett.* 43: 300-302, 1974.

65. GALLOP, P. M., O O BLUMEFELD and S SEIFTER, *Ann. Rev. Biochem.* 41: 617-672, 1972.

66. GLASS, D B. and J. M. MAY, *Collagen Rel. Res.* 4: 63-74, 1984.

67. GLASS, D. B. and S. B. SMITH, *J. Biol. Chem.* 258: 14797-14803, 1983.

68 GLASS, D. B., P. P. DEMBURE, J. H. PRIEST and L. J. ELSAS, *Biochim. Biophys. Acta* 810: 143-152, 1985.

69. GONZALEZ, E., A. HAMABATA and M. ROIKIND, *Collagen Rel. Res.* 4: 351-364, 1984.

70. GREENFIELD, N. and G. D. FASMAN, *Biochemistry* 8 (10): 4108-4116, 1969.

71. GROSS, J., *Harvey Lecture* 68: 351-432, 1974.

72. HAGAISHI, O. and K. UEDA, *Ann. Rev. Biochem.* 46: 95-116, 1977.

73 HALME, J., K. I. KIVIRIKKO and K. SIMONS, *Biochim. Biophys. Acta* 198: 460-470, 1970.

74. HALME, J., K. I. KIVIRIKKO and K. SIMONS, *Biochim. Biophys. Acta* 198: 460-470, 1970.

75. HOLLOSI, M., M. KAWAI and G. D. FASMAN, *Biopolymers* 24: 211-242, 1985.

76. IIME, A., A. NERLICH, U. FELDMANN, J. RAUTERBERG, R. W. GLANVILLE, G. EDEL, and P. K. MILLER, *J. Invest. Dermatol.* 83: 161-165, 1984.

77. JIMINEZ, S. A., M. HARSCH and J. ROSENBLOOM, *Biochem. Biophys. Res. Comm.* 52: 106-114, 1973.

78. KAWAI, M. and G. D. FASMAN, *J. Amer. Chem. Soc.* 100: 3630-3632, 1978.

79. KEDERSHA, N. L. and R. A. BERG, *Collagen Rel. Res.* 1: 345-353, 1981.

80. KEFALIDES, N. A., *Int. Rev. Connect. Tissue* 6: 63-104, 1973.

81. KEFALIDES, N. A., *Biochem. Biophys. Res. Comm.* 45: 226-234, 1972.

82. KIKUCHI, Y., Y. SUZUKI and N. TAMIYA, *Biochem. J.* 213: 507-512, 1983.

83. KIVIRIKKO, K. I. and D. J. PROCKOP, *Biochim. Biophys. Acta* 258: 380-379, 1972.

84. KIVIRIKKO, K. and R. MYLLYLA, *Methods Enzymol.* 82: 245-346, 1980.

85. KIVIRIKKO, K. I., K. SHUDO, S. SAKAKIBARA and D. J. PROCKOP, *Biochemistry* 11: 122-129, 1972.

86. KIVIRIKKO, K. I., L. RYHANEN, H. ANTINNEN, P. BORNSTEIN and D. J. PROCKOP, *Biochemistry* 12: 4696-4970, 1973.

87. KOPPLE, K. D., A. GO and D. R. PILIPAUSKAS, *J. Amer. Chem. Soc.* 97: 6830-6838, 1975.

88. KRANE, S. M., S. R. PINNEL and R. W. ERBE, *Proc. Natl. Acad. Sci. USA* 69: 2899-2903, 1972.

89. KRANE, S. M., S. R. PINNEL and R. W. ERBE, *Proc. Natl. Acad. Sci. USA* 69: 2899-2903, 1971.

90. KRESNA, K. F. and E. J. MILLER, *Biochemistry* 4: 3089-3097, 1979.

91. KRIMM, S. and J. BANDEKAR, *Biopolymers* 19: 1-36, 1980.

92. KUHN, K., In: *Immuhochemistry of the extracellular matrix*, edited by H. FURTHMAYR. Florida: CRC Press: 1982, p. 1-29.

93. KUHN, K., *Essays in Biochemistry* 5: 59-63, 1969.

94. KUUTTI, E. R., L. TUDERMAN and K. I. KIVIRIKKO, *Proc. Natl. Acad. Sci. USA* 57: 181-188, 1975.

95. LEWIS, P. N., F. A. MOMANY and A. H. SCHERAGA, *Proc. Natl. Acad. Sci. USA* 70: 538-, 1971.

96. LEWIS, P. N., F. A. MOMANY and H. A. SCHERAGA, *Biochim. Biophys. Acta* 303: 211-229, 1973.

97. LIGHT, N and A. J BAILEY, *FEBS Lett.* 182: 503-508, 1985.

98. LUNNEY, J. and G. ASHWELL, *Proc. Natl. Acad. Sci. USA* 73: 341-343, 1976.

99. MADISON, V. and K. D. KOPPLE, *J. Amer. Chem. Soc.* 102: 4855-4863, 1980.

100. MARRAUD, M. and A. AUBURY, *Internatl J. Peptide Protein Res.* 23: 123-138, 1981.

101. MARTIN, G. R., P. H. BYERS and K. A PIEZ, *Adv. Enzymol.* 42: 167-191, 1975

102. MCKUSICK, V A., *Heritable disorders of connective tissue.* St. Luois, Mosby 1972.

103. MECHANIC, G. L., P. M. GALLOP and M. L. TANZER, *Biochem. Biophys Res. Comm* 45: 644-653, 1971

104. MILLER, R. L., *Anal. Biochem* 45. 202-210, 1972.

105. MILLER, R. L. and H. H. J. VARNER, *Biochemistry* 18 (26). 5928-5933, 1979.

106. MIYAZAWA, T., In: *Poly-α-Amino acids*, edited by G. D. FASMAN. New York: Marcel Dekker, 1967, p. 70-103.

107. MUTTER, M. and K. H. ALTMANN, *Internatl. J. Peptide Protein Res.* 26: 373-380, 1985.

108. MYLLYEA, R. L., E. KUUTTI and K. I. KIVIRIKKO, *Biochem. Biophys. Res. Comm.* 83: 441-448, 1978.

109. NEMETHY, G. and M. P. PRINTZ, *Proc. Natl. Acad. Sci. USA* 76: 755-758, 1972.

110. NIETFELD, J. J. and A. KEMP, *Biochim. Biophys. Acta* 657: 157-167, 1981.

111. NIINIKOSKI, J., *Acta Physiol. Scand. Suppl.* 334: 1-72, 1969.

112. PAIK, W. K. and S. KIM, *Adv. Enzymol.* 42: 227-286, 1975.

113. PANKALAINEN, M., H. ARO, K. SIMONS and K. I. KIVIRIKKO, *Biochem. Biophys. Acta* 221: 559-565, 1970.

114. PARK, E., R. L. CHURCH and M. L. TANZER, *Immunology*, 781-790, 1975.

115. PARKER, F. S., *In Applications of Infrared and Resonance Raman Spectroscopy in Biochemistry.* New York: Plenum Press, 1983.

116. PAUL, P. K. C. and C. RAMAKRISHNAN, *J. Biomol. Struct. Dyn.* 2: 879-898, 1985.

117. PEASE, L. G., Ph.D Thesis, Harvard University, 1975.

118. PEREJDA, A. J., E. J. ZARAGOZA, E. ERIKSEN and J. UITTO, *Collagen Rel. Res.* 4: 428-438, 1984.

119. PIEZ, A. and A. MILLER, *J. Supramol. Structure* 2: 121-137, 1974.

120. PINNEL, S. R., S. M. KRANE, J. E. KENZORA and M. J. GLIMCHER, *New Engl. J. Med.* 286: 1013-1020, 1972.

121. PINNELL, S. R. and G. R. MARTIN, *Proc. Natl. Acad. Sci. USA* 61: 708-713, 1968.

122. POPENOE, E. A. and R. B. ARONSON, *Biochim. Biophys. Acta* 258-: 380-386, 1972.

123. POPENOE, A., R. B. ARONSON and D. D. SLYKE, *Arch. Biochem. Biophys.* 133: 286, 1967.

124. POPENOE, E. A. and R. B. ARONSON, *Biochim. Biophys. Acta* 258: 380-386, 1972.

125. POSER, J. W. and N. RAMAN, *Proc. Natl. Acad. Sci. USA* 73: 3374-3375, 1976.

126. PRIEST, J. H., *Medical Genetics & Cell Culture.* Philadelphia: Lea & Febiger, 1977, p. 163-195.

127. PROCKOP, D. G., R. A. BERG, K. I. KIVIRIKKO and J. UITTO, In: *Biochemistry of Collagen*, edited by G. N. RAMACHANDRAN and A. H. REDDI. New York: Plenum Press, 1976, p. 163-273.

128. PROCKOP, D. J., K. JUVA and J. ENGEL, *Z. Physiol. Chem.* 348: 553-560, 1967.

129. PUISTOLA, U., T. M. TURPEENNIEMI-HUJANEN, R. MYLLYLA and K. I. KIVIRIKKO, *Biochim. Biophys. Acta* 611: 40-50, 1980a.

130. PUISTOLA, U., T. M. TURPEENNIEMI-HUJANEN, R. MYLLYLA and K. I. KIVIRIKKO, *Biochim. Biophys. Acta* 611: 51-60, 1980b.

131. QUINN, S. R. and S. M. KRANE, *Biochim. Biophys. Acta* 585: 589-598, 1979.

132. RAMACHANDRAN, G. N. and A. H. REDDI, In: *Biochemistry of Collagen*, edited by G. N. RAMACHANDRAN and A. H. REDDI. New York: Plenum Press, 1976, p. 45-84.

133. RAMACHANDRAN, G. N., In: *Treatise on Collagens*, vol. I, edited by G. N. RAMACHANDRAN. New York: Academic Press, 1967, p. 103-183.

134. RAMALEY, P. B., S. A. JIMENEZ and J. ROSENBLOOM, *FEBS. Lett.* 33: 187-191, 1973.

135. RAO, C. P., R. NAGARAJ, C. N. R. RAO and P. BALARAM, *Biochemistry* 19: 425-431, 1980.

136. RAO, P. C., R. NAGARAJ, C. N. R. RAO and P. BALARAM, *Biochemistry* 19: 425-431, 1980.

137. RAO, B. N. N., A. KUMAR, H. BALARAM, A. RAVI and P. BALARAM, *J. Amer. Chem. Soc.* 105: 7423-7428, 1983.

138. RHOADS, R. E. and S. UDENFRIEND, *Biochim. Biophys. Acta* 198: 460-470, 1970.

139. ROBINS, S. P., M. SHIMOKOMAKI and A. J. BAILEY, *Biochem. J.* 131: 771-780, 1973.

140. ROBINS, S. P. and A. J. BAILEY, *Biochem. J.* 131: 771-780, 1973.

141. ROBINS, S. P. and A. J. BAILEY, *FEBS. Lett.* 38: 334-336, 1974.

142. ROSE, G. D., L. M. GIERASCH and J. A. SMITH, *Adv. Protein Chem.* 37: 1-109, 1985.

143. ROSENBLOOM, J., M. HARSH and S. JIMINEZ, *Arch. Biochem. Biophys.* 158: 478-484, 1973.

144. RUBIN, C. S. and O. M. ROSEN, *Ann. Rev. Biochem.* 44: 831-887, 1975.

145. RYHANEN, L., *Biochim. Biophys. Acta* 438: 71-89, 1976.

146. SAKAKIBARA, S., Y. SHIMONISHI, Y. KISHIDA, M. OKADA and H. SUGIHARA, *Bull. Chem. Soc. Jap.* 40: 2164-2167, 1976.

147. SAKAKIBARA, S., K. INOUYE, K. SHUDO, Y. KOBAYASHI and D. J. PROCKOP, *Biochim. Biophys. Acta* 303: 198-202, 1973.

148. SATHYANARAYANA, B. K. and J. APPLEQUIST, *Internatl. J. Peptide Protein Res.* 27: 86-94, 1986.

149. SCHOFIELD, J. D. and D. J. PROCKOP, *Clin. Orthop.* 97: 175-195, 1973.

150. SEARS, D. W. and S. BEYCHOK, In: *Physical Principles and Techniques of Protein Chemistry*, vol. Part C, edited by L. S.J.. New York: Academic Press, 1973, p. 445.

151. SHATO, T. and R. G. SPIRO, *J. Biol. Chem.* 251: 4062-4070, 1976.

152. SHIELDS, J. E., S. T. MCDOWELL, J. PAULOS and G. R. GRAY, *J. Amer. Chem. Soc.* 90: 1049-1060, 1968.

153. SIEGEL, R. C., S. R. PINNELL and G. R. MARTIN, *Biochemistry* 9: 4486-4492, 1970.

154. SIEGEL, R. C., *Proc. Natl. Acad. Sci. USA* 71: 4726-4830, 1974.

155. SILVER, I. A., In: *Biology of fibroblasts*, edited by E. KULONEN and J. PIKKARAINEN. London: Academic Press, 1973, p. 507-519.

156. SINEX, F. M., In: *Treatise on collagen*, edited by G. N. RAMACHANDRAN. New York: Academic Press, 1968, p. 410-448.

157. SMITH, F. A. and L. G. PEASE, *Crit. Rev. Biochem.* 8: 315-399, 1980.

158. SPIRO, R. G., *J. Biol. Chem.* 244: 602-612, 1969.

159. STADTMAN, E. R., *The Enzymes* 8: 2-49, 1973.

160. STENFLO, J. and J. W. SUTTIE, *Ann. Rev. Biochem.* 46: 157-172, 1977.

161. STIMLER, N. P. and M. L. TANZER, In: *Protein cross-linking*, vol. 2, edited by M. FRIEDMAN. New York and London: Plenum Press, 1976, p. 675-697.

162. STIMSON, E. R., S. S. ZIMMERMAN and H. A. SCHERAGA, *Macromolecules* 10: 1049-1060, 1977.

163. SUDA, T. S. and P. BALARAM, *FEBS. Lett.* 134: 32-36, 1981.

164. SUDHAKAR, V., M. VIJAYAN, R. S. RAPAKA, R. S BHATNAGAR and V. S. ANANTHANARAYANAN, *J. Amer. Chem. Soc.* 103: 1703-1708, 1981.

165. TAJIMA, S., S. MURAD and S PINNELL, *Proc. Natl. Acad. Sci. USA* 82: 3091-3095, 1985.

166. TAKAHASHI, K., G. NEMETHY, H. A. SCHERAGA and V. S. ANANTHANARAYANAN, *In Press*.

167. TANZER, M. L., *Science* 180: 561-566, 1973.

168. TANZER, M. L., R. L. CHURCH, J. A. YAEGER, D. J. WAMPLER and E. PARKER, *Proc. Natl. Acad. Sci. USA* 71: 3009-3031, 1974.

169. THOMAS, G. J. J. and Y. KYOGUKU, In: *Infrared and Raman Spectroscopy*, vol. Part C, edited by E. G. BRAME and J. G. GRASSELLI. New York: Marcel Dekker, 1977, p. 743-775.

170. TIMPL, R., G. R. MARTIN, P. BRUCKNER, G. WICK and H. WIEDMANN, *Eur. J. Biochem.* 84: 43-52, 1978.

171. TRAUB, W. and K. A. PIEZ, *Adv. Protein Chem.* 25: 243-352, 1971.

172. TRAUB, B., R. FLUCKIGER and K. H. WINTERHALTER, *Collagen Rel. Res.* 4: 239-251, 1984.

173. TURPEENNIEMI-HUJANEN, T. and R. L. MYLLYLA, *Biochim. Biophys. Acta* 800: 59-65, 1984.

174. TURPEENNIEMI-HUJANEN, T. M., U. PUISTOLA and K. I. KIVIRIKKO, *Biochem. J.* 189: 247-253, 1980.

175. UITTO, J. and D. J. PROCKOP, *Arch. Biochem. Biophys.* 164: 210-217, 1974.

176. UITTO, J. and D. J. PROCKOP, *Eur. J. Biochem.* 221-230: 43, 1974.

177. UITTO, J. and J. R. LICHTENSTEIN, *J. Invest. Dermatol.* 66: 59-79, 1976.

178. UITTO, J. and D. J. PROCKOP, *Biochem. Biophys. Res. Comm.* 60: 414-423, 1974.

179. URRY, D. W., M. M. OHNISHI and M. JACOBS, *Biochem. Biophys. Res. Comm.* 61: 1427-1433, 1974.

180. URUSHIZAKI, Y. and S. SEIFTER, *Proc. Natl. Acad. Sci. USA* 82: 3091-3095, 1985.

181. VARNER, H. H. J. and R. L. MILLER, *Prep. Biochem.* 10: 1-9, 1980.

182. VENKATACHALAMI, C. M., *Biopolymers* 6: 1425-1436, 1968.

183. WALTON, A. G. and J. BLACKWELL, *Biopolymers* New York, 1973.

184. WARD, A. W. and P. MASON, *J. Mol. Biol.* 79: 431-435, 1973.

185. WOODY, R. W., In: *Peptides, Polypeptides and Proteins*, edited by E. R. BLOUT, F. A. BOVEY and N.LOTAN. New York: John Wiley and Sons, 1974, p. 338-350.

186. YANNAS, I. V., *J. Macromol. Sci. Revs. Macromol. Chem.* C7 (1): 49, 1972.

187. ZIMMERMAN, S. S. and H. A. SCHERAGA, *Proc. Natl. Acad. Sci. USA* 74: 4126-4129, 1977.

188. ZIMMERMAN, S. S. and H. A. SCHERAGA, *Biopolymers* 16: 811-843, 1977.

189. ZUSSMAN, J., *Acta Cryst.* 4: 72-93, 1951.