Synthesis and Structural Studies of Prion Peptides

by

Raúl Horacio Zambrano

B.A.
Columbia College, 1993

SUBMITTED TO THE DEPARTMENT OF CHEMISTRY IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN CHEMISTRY
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

SEPTEMBER 1995

Signature of Author: _____
Department of Chemistry
August 4, 1995

Certified by: _____
Peter T. Lansbury
Professor of Chemistry
Thesis Supervisor

Accepted by: _____
Dietmar Seyferth
Chairman, Departmental Committe on Graduate Students

Synthesis and Structural Studies of Prion Peptides

by

Raúl H. Zambrano

Submitted to the Department of Chemistry
on August 4, 1995 in Partial Fulfillment of the Requirements for the Degree of Master of Science in Chemistry

ABSTRACT

The ubiquitously ^{13}C-labeled dipeptide tyrosyl lysine was synthesized for use as a model compound in examining radio frequency driven recoupling solid state nuclear magnetic resonance spectroscopy as a possible tool in elucidating the structure of the prion protein.

One, two, and three repeats of the octarepeat region (proline-histidine-glycine-glycine-glycine-tryptophan-glycine-glutamine) of the prion protein were synthesized and preliminary circular dichroism (CD) studies were performed. A slight concentration dependence was observed and the observed CD absorbances did not correspond to any known structure.

Thesis Supervisor: Peter T. Lansbury Jr

Title: Associative Professor of Chemistry

Table of Contents

Chapter 1

Introduction

The class of diseases known as the transmissible spongiform encephalopathies (TSE's) are hypothesized to be caused by infective pathogens known as prions (Chesebro, 1991). TSE's affect both humans and animals. Human TSE's include kuru, Creutzfeldt-Jakob disease (CJD), fatal familial insomnia (FFI), and Gerstmann-Sträussler-Scheinker syndrome (GSS). Scrapie, bovine spongiform encephalopathy (BSE), and transmissible mink encephalopathy are some of the animal TSE's. The TSE's, also known as prion diseases, are all fatal. A unique feature of the prion diseases is that they are inheritable and transmissible (Stahl, & Prusiner, 1991).

It is postulated that prions, proteinaceous infectious particles, are composed of a host derived protein, PrP^{C}. The protein consists of 254 amino acids (Figure 1).

SHaPrP MANLSYWILALFVAMWTDVGLCKKRPKPGGWNTGGS RYPG

QGSPGGGNRYPPQGGGTWGQPHGGGWGQPHGGGGWGQPHGGGWGQ

PHGGGWGQGGGTHNQWNKPSKPKTNMKHMAGAAAAGAVVGGLGG

YMLGSAMSRPMMHFGNDWEDRYYRENMNRYPNQVYYRPVQYNNQN

NFVHDCVNITIKQHTVTTTTKGENFTETDIKIMERVVEQMCTTQYQKESQ

AYYDGRRSSAVLFSSPPVILLISFLIFLMVG

Figure 1: Sequence of the prion protein in Syrian hamster.

The protein is made up of a signal peptide cleaved during biosynthesis, five octapeptide and two hexapeptide repeats rich in Gly and Pro, a transmembrane like sequence, a possible amphipathic helix, and a hydrophobic terminal sequence (Figure 2).

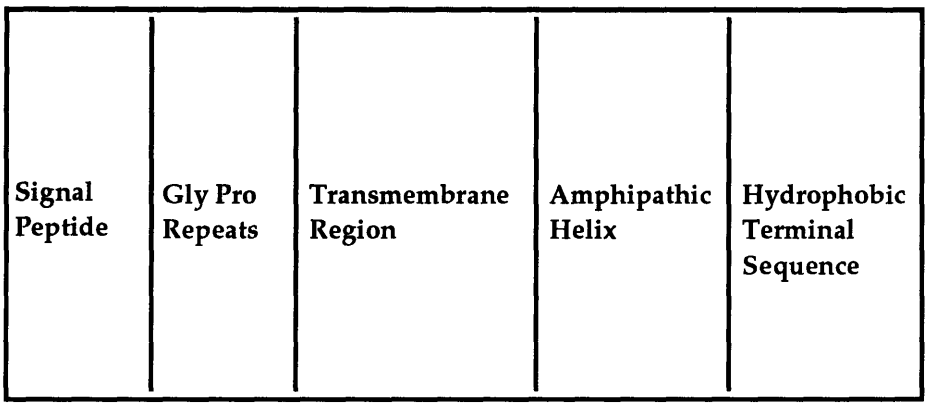


Figure 2: Schematic of PrP

The infectious form of the prion protein, PrP^{Sc}, seems to be covalently identical to PrP^{C}. Although there does not seem to be any covalent difference between PrP^{C} and PrP^{Sc}, there are some physical differences between the two isoforms. PrP^{C} is significantly more soluble than PrP^{Sc}. PrP^{C} is more sensitive to protease digestion than PrP^{Sc}. PrP^{Sc} is only partially degraded by proteinase K (PK) under conditions that completely degrade PrP^{C}. The PK treatment of PrP^{Sc} removes about 67 amino acids from the amino terminus of PrP. It should be noted that cleavage at the N-terminus does not affect infectivity (Prusiner, 1991).

The function of PrP in cells is as of yet unknown. However, it is known that PrP^{Sc} is derived from PrP^{C}. PrP mRNA concentration is highest in the neurons and does not increase over the course of infection (Chesebro, et al., 1985). PrP^{Sc} is formed after PrP^{C} is anchored to the cell membrane by its glycosylphosphatidylinositol (GPI) anchor (Caughey, & Raymond, 1991).

Certain mutations in the PrP gene have been linked with GSS and CJD. The open reading frame of the PrP gene is found on a single exon on the short arm of human chromosome 20 (Sparkes, et al., 1986). A cytosine to thymine substitution in the second position of codon 102 causes a proline to leucine change which segregates with GSS (Hsiao, et al., 1990). A mutation at

codon 200 (glutamic acid to lysine) has also been linked to CJD in Libyan Jews (Goldfarb, et al., 1992). Finally the insertion of several additional octapeptide repeats has been found to segregate with GSS (Owen, et al., 1991).

Structure of PrP

One structural study of PrP^C has been reported using Fourier Transform Infrared Spectroscopy (FTIR) and Circular Dichroism (CD) (Pan, et al., 1993). Using FTIR two structural studies of PrP^{Sc} have appeared (Caughey, et al., 1991; Gasset, Baldwin, Fletterick, & Prusiner, 1993) These studies both correlated β structure with infectivity. Implying that the difference between the two isoforms of the prion may be conformational. Recently, PrP^C has been converted to PrP^{Sc} in vitro supporting the idea that the non- and infectious forms differ only in their 3-dimensional structure (Kocisko, et al., 1994). Therefore, in order to understand the mechanism of interconversion, which may be essential to the process of infection, it is necessary to elucidate the structures of PrP^C and PrP^{Sc}.

Caughey, B., & Raymond, G. J. (1991). The Scrapie-associated Form of PrP Is Made from a Cell Surface Precursor That Is Both Protease-and Phospholipase-Sensitive. The Journal of Biological Chemistry, 266(27), 18217-18223.

Caughey, B. W., Dong, A., Bhat, K. S., Ernst, D., Hayes, S. F., & Caughey, W. S. (1991). Secondary Structure Analysis of the Scrapie-Associated Protein PrP 27-30 in Water by Infrared Spectroscopy. Biochemistry, 30, 7672-7680.

Chesebro, B., Race, R., Wehrly, K., Nishio, J., Bloom, M., Lechner, D., Bergstrom, S., Robbins, K., Mayer, L., Keith, J. M., Garon, C., & Haase, A. (1985). Identification of scrapie prion protein-specific mRNA in scrapie-infected and uninfected brain. Nature, 315, 331-333.

Chesebro, B. W. (1991). Transmissible Spongiform Encephalopathies Scrapie, BSE and Related Disorders. New York: Springer-Verlag, 288.

Gasset, M., Baldwin, M. A., Fletterick, R. J., & Prusiner, S. B. (1993). Perturbation of the secondary structure of the scrapie prion protein under conditions that alter infectivity. Proceedings of the National Academy of Sciences, 90, 1-5.

Goldfarb, L. G., Petersen, R. B., Tabaton, M., Brown, P., LeBlanc, A. C., Montagna, P., Cortelli, P., Julien, J., Vital, C., Pendelbury, W. W., Haltia, M., Wills, P. R., Hauw, J. J., McKeever, P. E., Monari, L., Schrank, B., Swergold, G. D., Autilio-Gambetti, L., Gajdusek, D. D., Lugaresi, E., & Gambetti, P. (1992). Fatal Familial Insomnia and Familial Cretzfeldt-Jakob Disease: Disease Pheno type Determined by a DNA Polymorphism. Science, 258, 806-808.

Hsiao, K. K., Scott, M., Foster, D., Groth, F. D., DeArmond, S. J., & Prusiner, S. B. (1990). Spontaneous Neurodegeneration in Transgenic Mice with Mutant Prion Protein. Science, 250, 1587-1590.

Kocisko, D. A., Come, J. H., Priola, S. A., Chesebro, B., Raymond, G. J., Lansbury Jr, P. T., & Caughey, B. (1994). Cell-free formation of protease-resistant prion protein. Nature, 370, 471-474.

Owen, F., Poulter, M., Collinge, J., Leach, M., Shah, T., Lofthouse, R., Chen, Y., Crow, T. J., Harding, A. E., Hardy, J., & Rossor, M. N. (1991). Insertions in the Prion Protein Gene in Atypical Dementias. Experimental Neurology, 112, 240-242.

Pan, K., Baldwin, M., Nguyen, J., Gasset, M., Serban, A., Groth, D., Mehlhorn, I., Huang, Z., Fletterick, R. J., Cohen, F. E., & Prusiner, S. B. (1993). Conversion of α-helices into β-sheets features in the formation of the scrapie proteins. Proceedings of the National Academy of Sciences. USA, 90, 10962-10966.

Prusiner, S. B. (1991). Novel Properties and Biology of Scrapie Prions. In B. W. Chesebro (Ed.), Transmissible Spongiform Encephalopathies Scrapie, BSE and Related Disorders (pp. 233-257). New York: Springer-Verlag.

Sparkes, R. S., Simon, M., Cohn, V. H., Fournier, R. E. K., Lem, J., Klisak, I., Heinzmann, C., Blatt, C., Lucero, M., Mohandas, T., DeArmond, S. J., Westawat, D., Prusiner, S. B., & Weiner, L. P. (1986). Assignment of the human and mouse prion protein genes to homologous chromosomes. Proceedings the National Academy of Sciences, 83, 7358-7362.

Stahl, N., & Prusiner, S. B. (1991). Prions and prion proteins. The FASEB Journal, 5, 2799-2807.

Chapter 2

Solid State Nuclear Magnetic Resonance Spectroscopy

Radio Frequency Driven Recoupling (RFDR)

In trying to elucidate the structure of PrP^{Sc} only low resolution methods have been applied, as PrP^{Sc} forms an insoluble, non-crystalline, aggregate that is classified as amyloid(Kelly, & Lansbury, 1994) . As a result, traditional methods of structure elucidation such as X-ray crystallography and solution nuclear magnetic resonance spectroscopy cannot be applied. Therefore, solid state nuclear magnetic resonance spectroscopy (SSNMR) has been considered for application to this problem (Griffiths, & Griffin, 1993). In particular, a new SSNMR method called radio frequency driven recoupling (RFDR), developed by Robert Griffin at the Massachusetts Institute of Technology, is being examined for application to the prion(Griffiths, et al., 1993).

RFDR is the solid state NMR equivalent of 2D Nuclear Overhauser effect spectroscopy (NOESY) and total correlation spectroscopy (TOCSY). In NOESY and TOCSY proton to proton spatial and connective information is acquired (Derome, 1987). In RFDR carbon-13 to carbon-13 or carbon-13 to nitrogen-15 spatial and connective information is acquired. RFDR is an improvement over the Rotational Resonance (R^2) technique previously employed, as R^2 measures one distance between nuclei at a time and the spinning speed of the sample must be maintained within stringent limits for distances to be acquired (Griffiths, et al., 1993).

To examine the effectiveness and accuracy of the RFDR method a model compound whose intermolecular carbon to carbon and carbon to nitrogen distances were known was needed. The dipeptide tyrosine lysine (Figure 3) was chosen as this model compound as its crystal structure was known.(Urpi, Coll, & Subirana, 1988). Therefore, a synthesis of tyrosine lysine from the commercially available ubiquitously carbon-13 labeled amino acids tyrosine and lysine was developed.

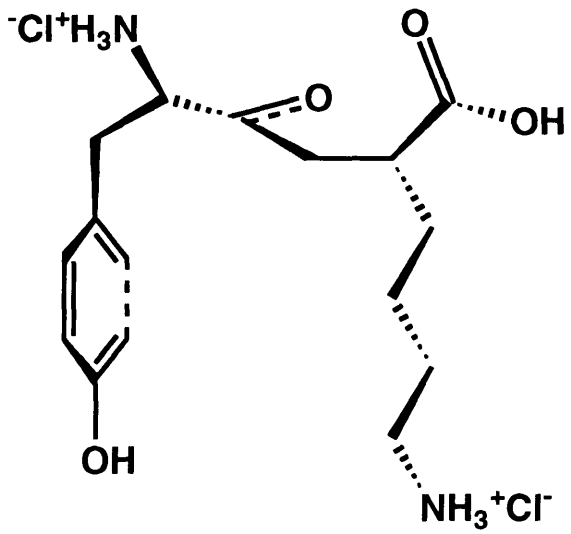
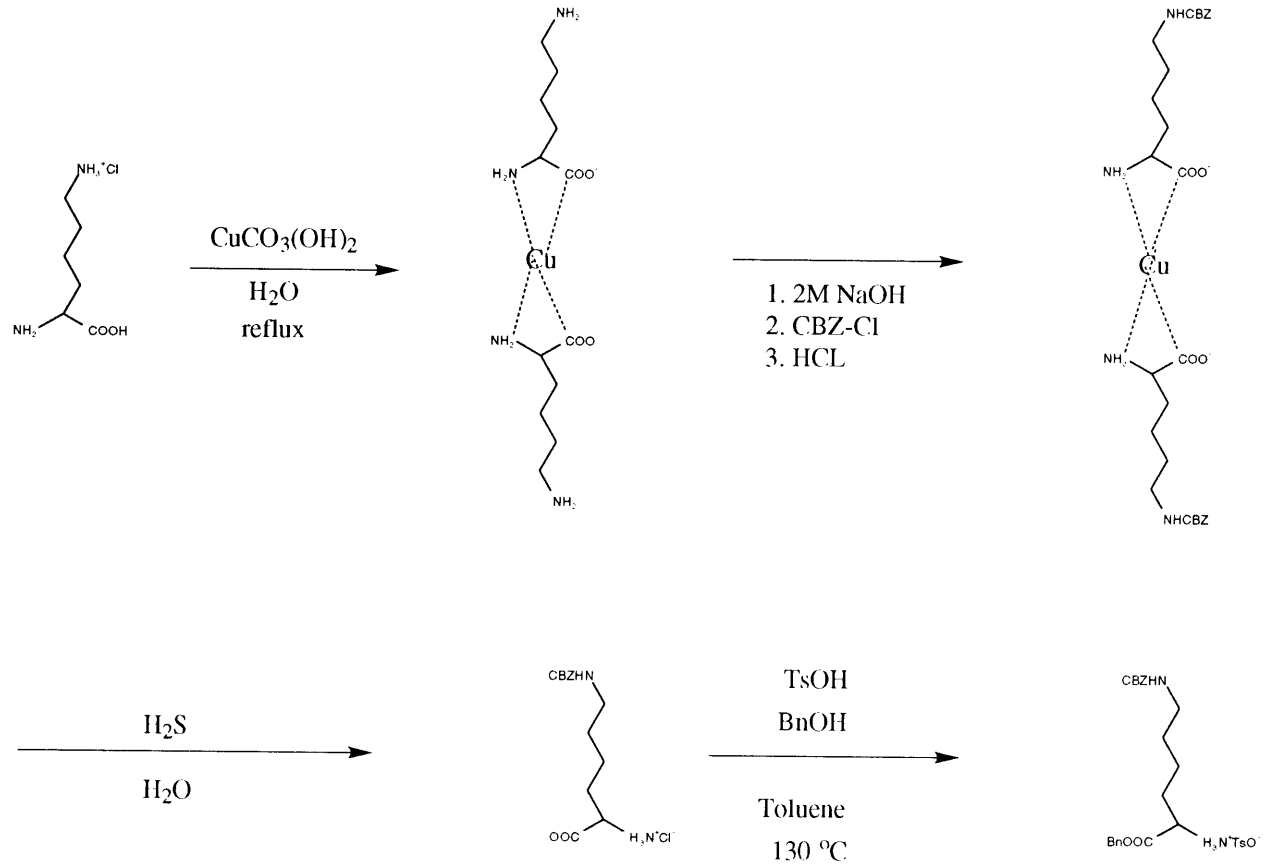
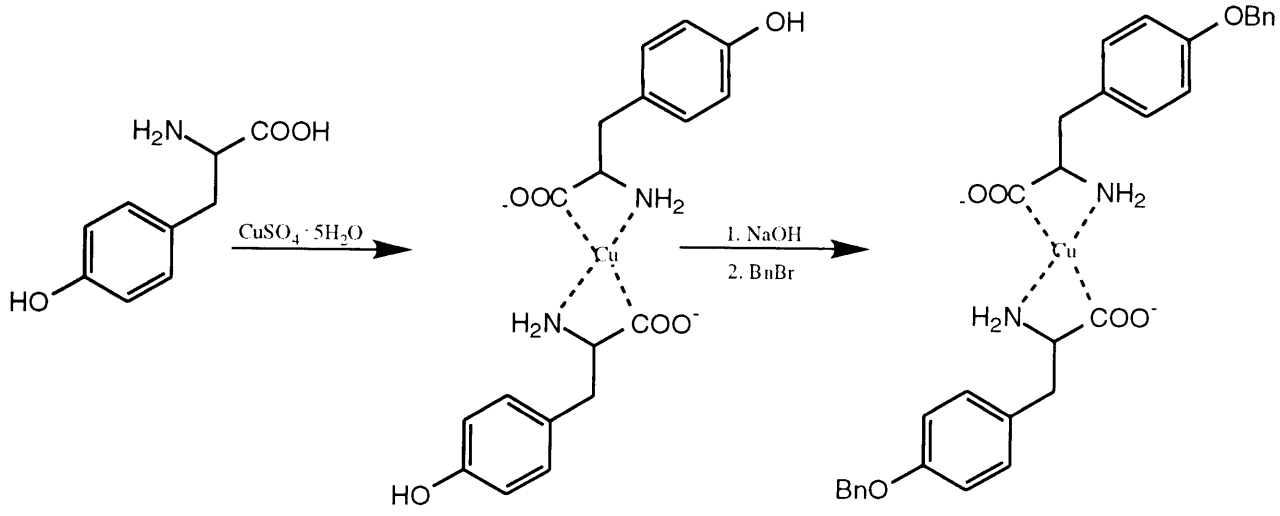


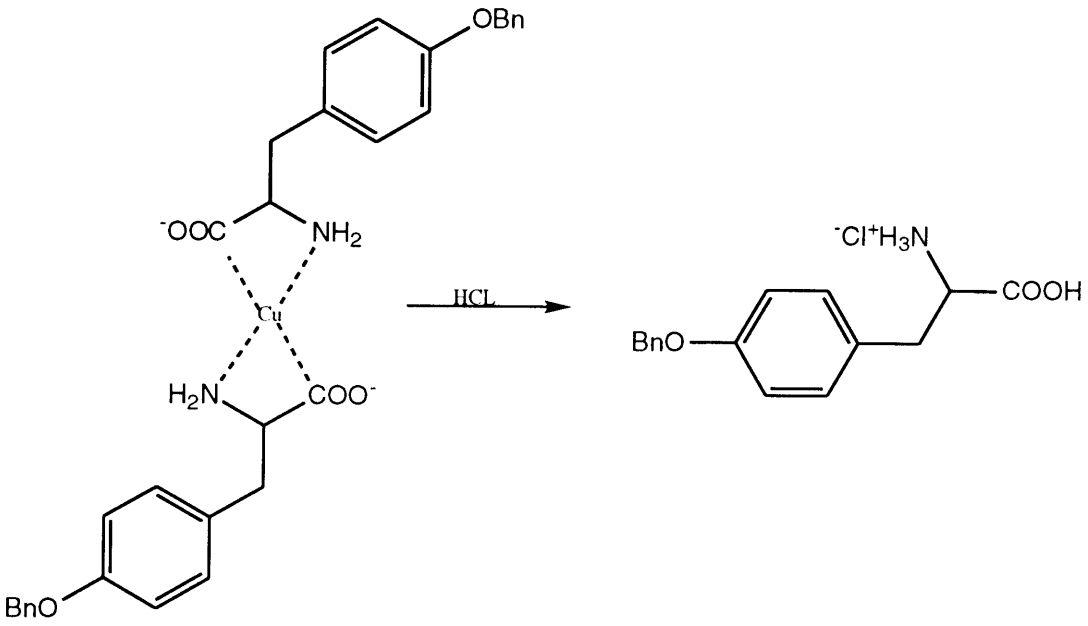
Figure 3 : Dipeptide Tyrosine Lysine (Figure drawn to represent three dimensional crystal structure)

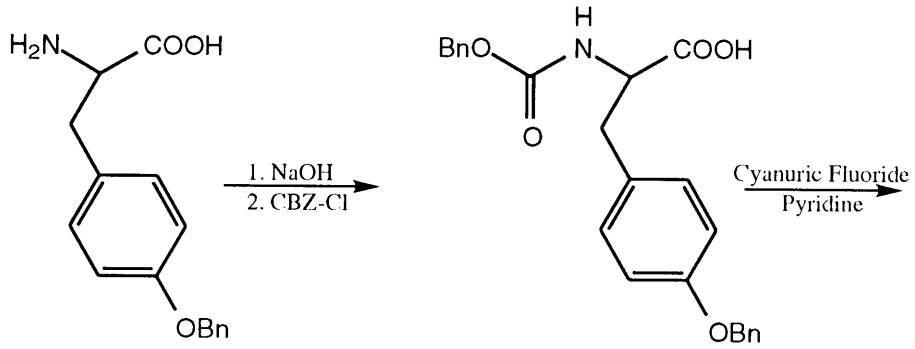
Synthesis of the Dipeptide Tyrosyl Lysine

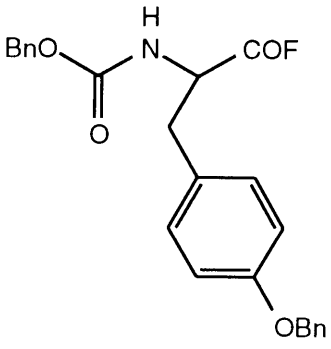
The following synthetic scheme was developed using known chemistry to synthesize carbon-13 labeled tyrosine lysine.

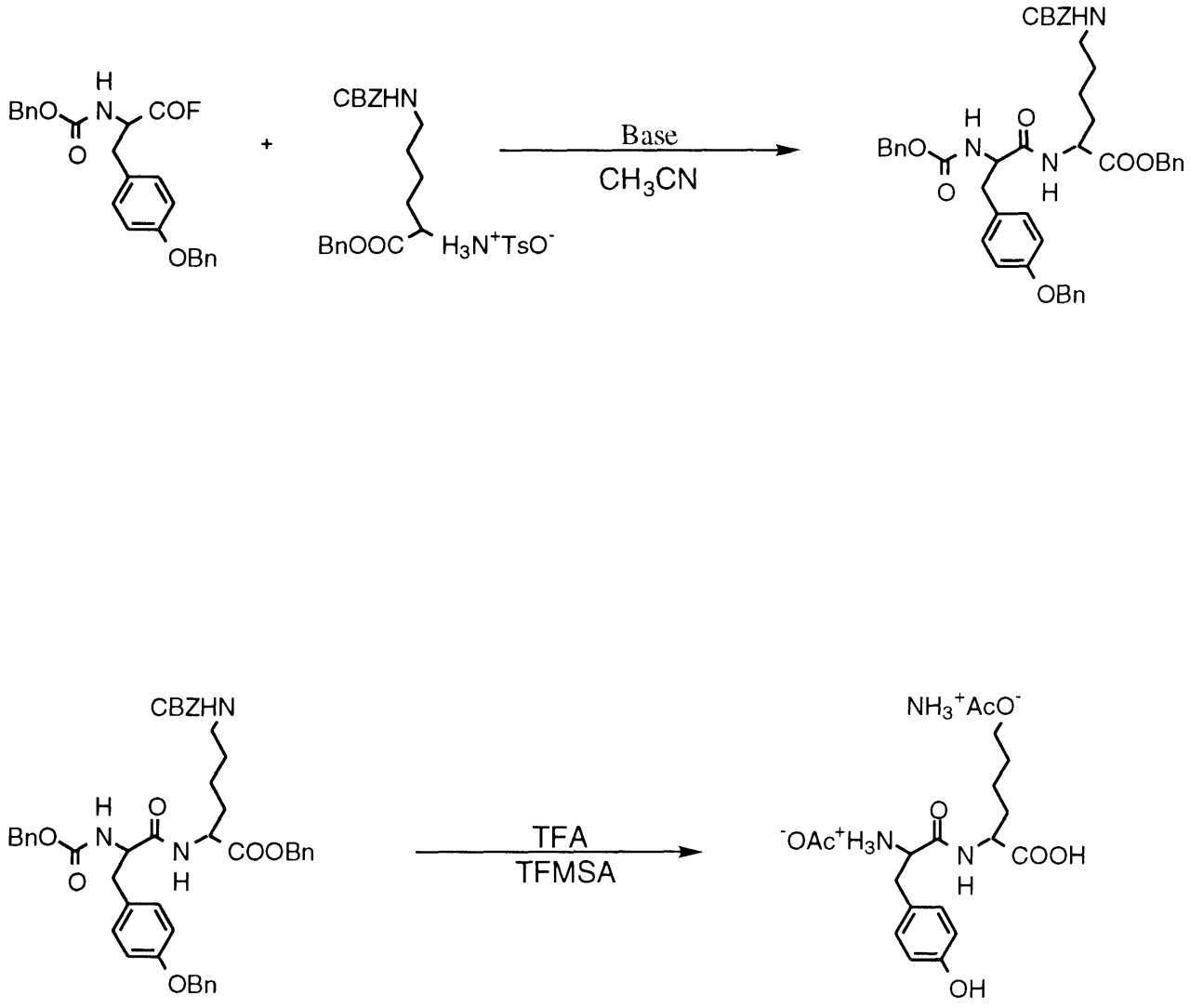

CuCO3(OH)2
H2O
reflux
1. 2M NaOH
2. CBZ-Cl
3. HCL
H2S
H2O
TsOH
BnOH
Toluene
130 °C


CuSO4 · 5H2O
1. NaOH
2. BnBr


OBn
-OOC
NH2
Cu
H2N
COO-
BnO
HCl
-Cl+H3N
COOH
BnO


H2N
COOH
OBn
1. NaOH
2. CBZ-Cl
H
N
BnO
O
COOH
OBn
Cyanuric Fluoride
Pyridine


H
N
BnO
O
COF
OBn



The synthesis of tyrosyl lysine (YK) was much more challenging than first expected. The protecting groups chosen for the synthesis of YK were all benzyl carbamates, esters, and ethers. These protecting groups were chosen with the purpose of a simple final reaction and purification. It was hoped that the

protected YK (Figure 3A) could be deprotected in one step through catalytic palladium hydrogenation.

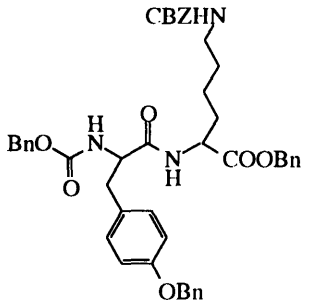


Figure 3A: Protected YK

As it turned out hydrogenation as a deprotection method was not possible as the protected YK was not soluble in any of the typical hydrogenation solvents. However, as it was later discovered trifluoromethanesulfonic (TFMSA) acid cleavage was very effective in the deprotection of YK.

Protection of the individual amino acids in YK also proved to be very challenging as both Y and K have three functionalities. First attempts at protecting the individual amino acids involved orthogonal protection of the amino and carboxy moieties of the amino acid followed by protection of either the amino or hydroxy amino acid side chain. These attempts generated very low yielding and impure products as multiple reactions and purifications were required. Eventually, it was decided to use copper chemistry to protect the amino and carboxy termini of the amino acids as the procedure reduced both the number of reactions and purifications.

The coupling reaction between the tyrosine and lysine produced side products that were inseparable from the protected dipeptide. The solution to this problem turned out to be the use of the acyl fluoride. A method which produces a very easily purified product as the products of the reaction are the product, starting materials, and salt.

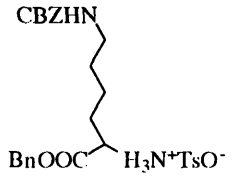
Conclusion

Tyrosine lysine will now be recrystallized over three months and used in studies of the effectiveness of RFDR.

Experimental*

* Note the carbon spectra that follow are proton decoupled, however, as the compound is ^{13}C ubiquitously labeled splitting of the carbon signal occurs. Also proton NMR's will be very broad as the carbons will be coupling to the hydrogens in addition to the normal hydrogen to hydrogen coupling.

ε-N-CBZ-Lysine-O-Benzyl Ester p-toluenesulfonic Acid Salt



2.0 grams (0.0098 mol, 1 eq) of labeled lysine was dissolved in 15 ml of water and 2.6 grams (0.0108 mol, 1.2 eq) of $CuCO_3(OH)_2$ added. The solution was heated to between 85-90 degrees Celsius and refluxed for one hour. The solution was then filtered to remove excess copper carbonate. The filter was washed with hot water until the filtrate was colorless. The solution was then cooled to room temperature. 14.4 ml (0.0288 mol, 4 eq) of 2M NaOH was added to the solution and it was cooled in an ice bath. 2.06 ml (0.0144 mol, 2eq) Benzyloxycarbonyl chloride was then added dropwise over ten minutes with stirring. After two hours the solution was filtered and 2.5g (0.00334 mol) of a blue solid was collected by filtration. The solid was dried under vacuum overnight and ground into a fine powder with a mortar and pestle. The solid was taken up in 200 ml of water and stirred vigorously as H_2S was bubbled through for thirty minutes. The solution was boiled and filtered hot. The filtrate was allowed to cool to room temperature at which time ε-N-CBZ-Lysine precipitates out. The ε-N-CBZ-Lysine was recrystallized multiple times from fresh water. 296 mg (9.3% yield) of the compound was recovered (Neuberger, & Sanger, 1943).

ε-N-CBZ-Lysine (296 mg, 0.00091 mol, 1 eq) was suspended in a mixture of 2 ml (20 eq) of freshly distilled benzyl alcohol and 2 ml (10 eq) of toluene. 190 mg (0.001 mol, 1.1 eq) of p-toluenesulfonic acid monohydrate was added and the solution was refluxed overnight. The solution was then added to cold

ether (200ml). A precipitate immediately formed and was collected through filtration. 403 mg (99% yield) of the ε-N-CBZ-Lysine-O-benzyl ester p-toluenesulfonic Acid Salt was collected (Bodanszky, & Bodanszky, 1984).

The product was identified through NMR (Figure 4 & 5; ^{1}H NMR ($CDCl_3$, 250 MHz) δ 0.9, 1.5, 2.0 (6H, Cβ-H,Cδ-H,Cγ-H), δ 2.3 (s, 3H, CH_3 p-toluenesulfonic acid), δ 2.6, 3.2 (s, 2H, Cε-H), δ 3.7, 4.3 (s, 1H, Cα-H), δ 5.0 (s, 4H, benzyl protons), δ 7.0, 7.7 (m, 4H, p-toluenesulfonic acid Ar), δ 7.3 (s, 10H, Phenyl H); ^{13}C NMR (CDCl3, 60 MHz) δ 21 (m, β C), δ 31 (m, δ C), δ 42 (m, γ C), δ 52, 53 (m, α,ε C), δ 171 (d, COOBn C).

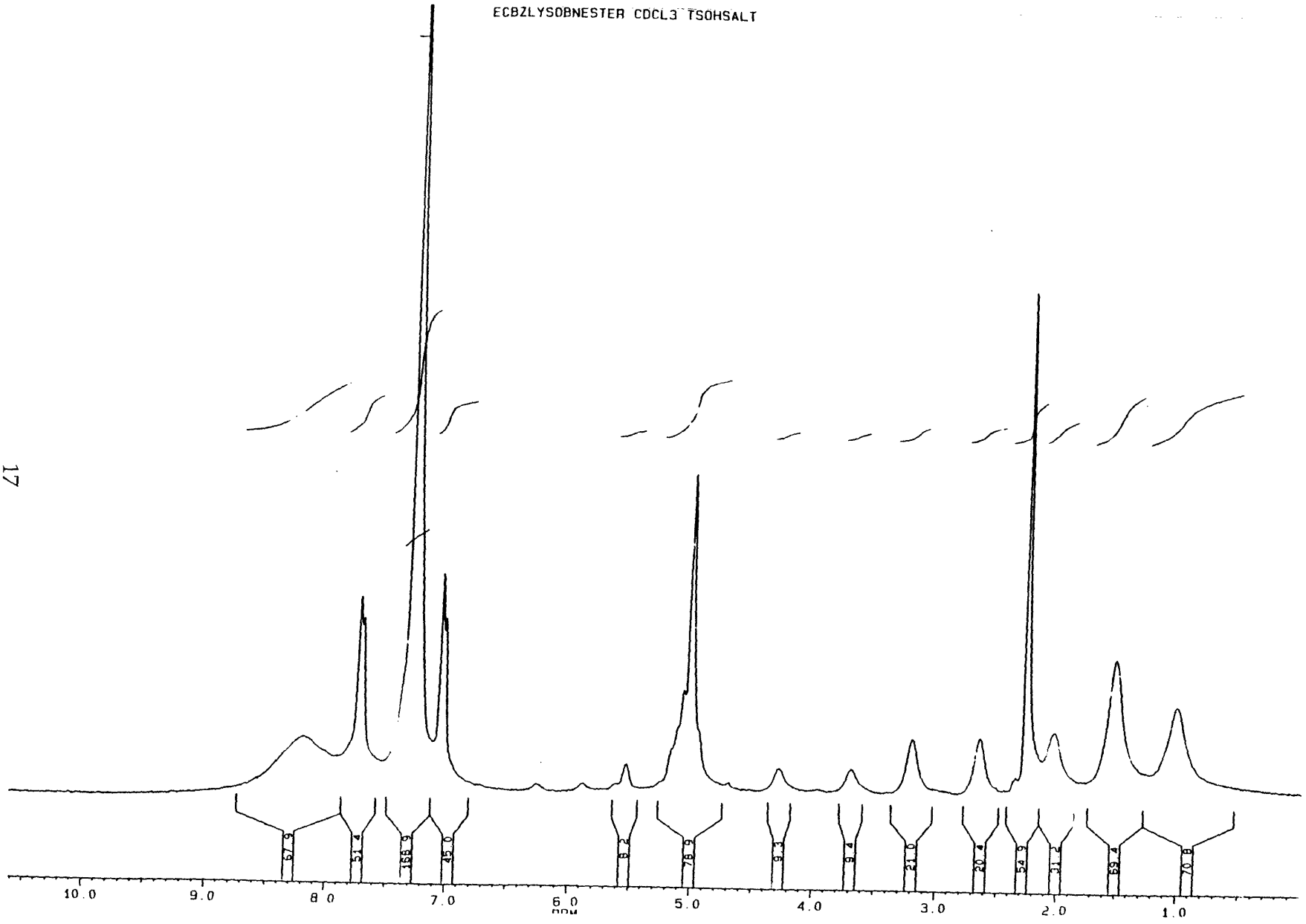


Figure 4: Proton NMR of e-N-CBZ-Lysine-O-Benzyl Ester p-toluenesulfonic Acid Salt

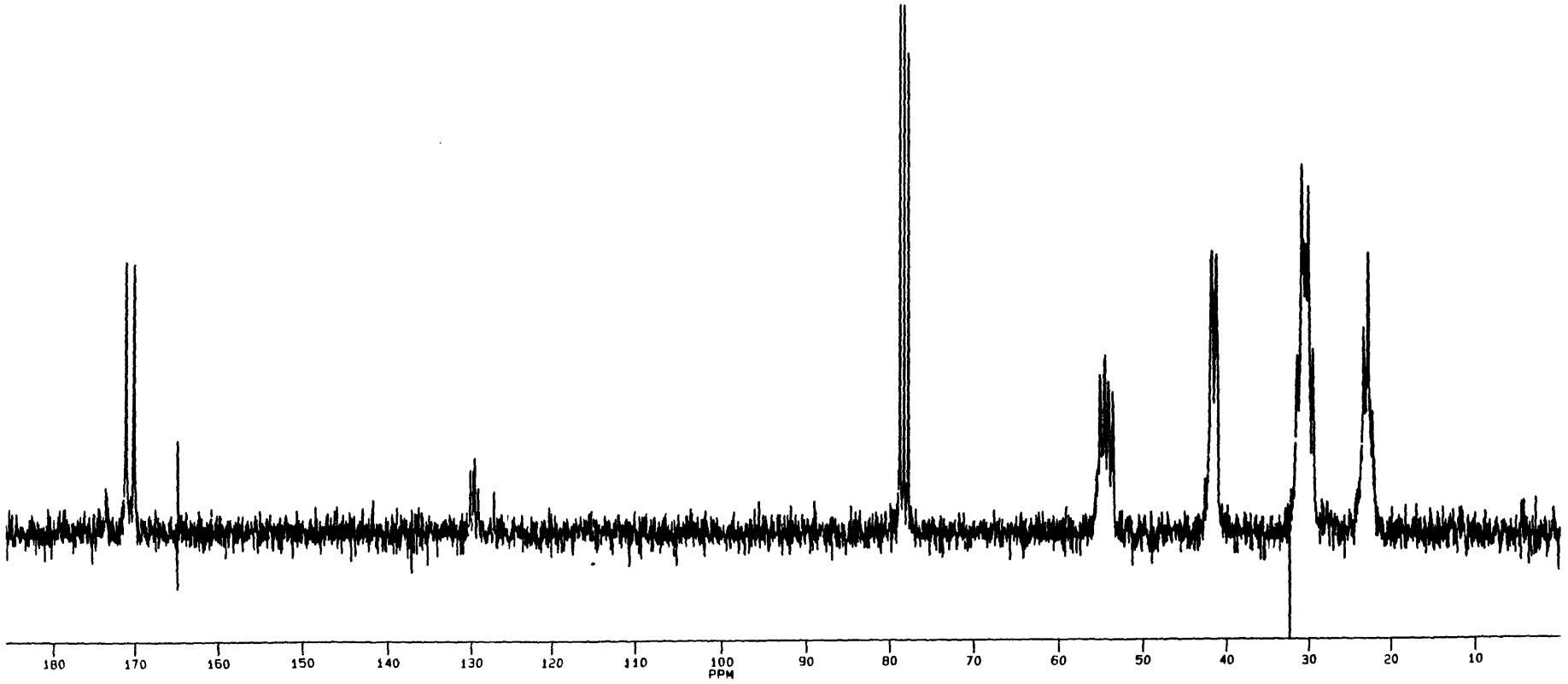
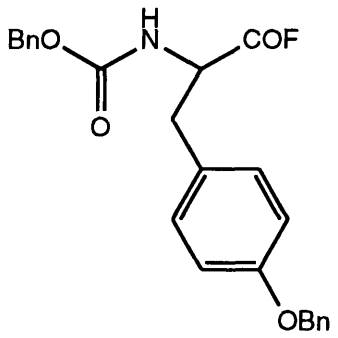


Figure 5: Carbon NMR of e-N-CBZ-Lysine-O-Benzyl Ester p-toluenesulfonic Acid Salt

N-CBZ-Tyrosine-O-Benzyl Ether Acyl Fluoride



0.5 grams (0.00262 mol, 1 eq) of tyrosine was dissolved in 2M NaOH (2.61 ml) and a solution of cupric sulfate pentahydrate (0.655 g, 0.00262 mol, 2 eq) in 1.32 ml of water was added. The mixture was heated to 60 ^{o}C and cooled to room temperature. The mixture was diluted with methanol (9.2ml) and made more alkaline with 393 microliters of 2M NaOH. 0.373 ml of benzyl bromide was added and the reaction mixture stirred for three hours. The blue precipitate was collected on a filter and washed with methanol and water. The compound was dried and ground into a fine powder. 965 mg (0.00155 mol, 60% yield) of the copper complex was obtained. The compound was then washed repeatedly in a glass sintered funnel with 1M HCl, water, and 1.5 M NH_4OH. 200 mg (0.000714 mol, 27% yield) of the tyrosine-O-benzyl ether was obtained.

The tyrosine-O-benzyl ether was taken up in 1.5 ml of 2M NaOH and chilled in an ice bath. To this solution 0.254 ml (0.0021 mol, 3 eq) of CBZ chloride was added and to stirred for two hours. Reaction was monitored by TLC (10% acetic acid/toluene; rf product=0.4; rf starting material= 0). 89. 3 mg (0.00021 mol, 30% yield) of N-CBZ-O-benzyl ether was recovered after purification on a silica column (10% AcOH/Toluene) (Bodanszky, et al., 1984).

89.3 mgs (0.00021 mol, 30% yield) of N-CBZ-O-benzyl ether was dissolved in 5 ml of acetonitrile. To this solution 0.021 ml (1.2 eq) of pyridine was added along with 0.018 ml (0.00021 mol, 1.0 eq) of cyanuric fluoride. The reaction was stirred for three hours and monitored by TLC (10% AcOH/$CHCl_3$; rf starting material= 0.2, rf product= 0.4. The reaction mixture was quenched with ice water and the cyanuric acid filtered off. The filtrate was extracted with $CHCl_3$ and dried with $MgSO_4$ and the material was dried in

vacuo. 88 mg (0.00021 mol, 98% yield) was recovered.(Bertho, Loffet, Reuther, & Sennyey, 1991)

The product was identified through NMR (Figure 6 and 7); ^{1}H NMR ($CDCl_3$, 250 MHz) δ 2.9, 3.5 (2H, Cβ-H), δ 4.0, 4.4 (1H, Cα-H), δ 4.7, 5.0 (4H, benzyl-H), d 6.3 (2H, 1,6 tyrosine aromatic H), δ 7.0 (10H, phenyl H); ^{13}C NMR ($CDCl_3$, 60 MHz) δ 38 (m, βC), δ 58 (m, αC), δ 115, 130 (m, tyrosine aromatic ring C) δ 158 (m, 4 tyrosine aromatic ring C), δ 172 (d, COF C).

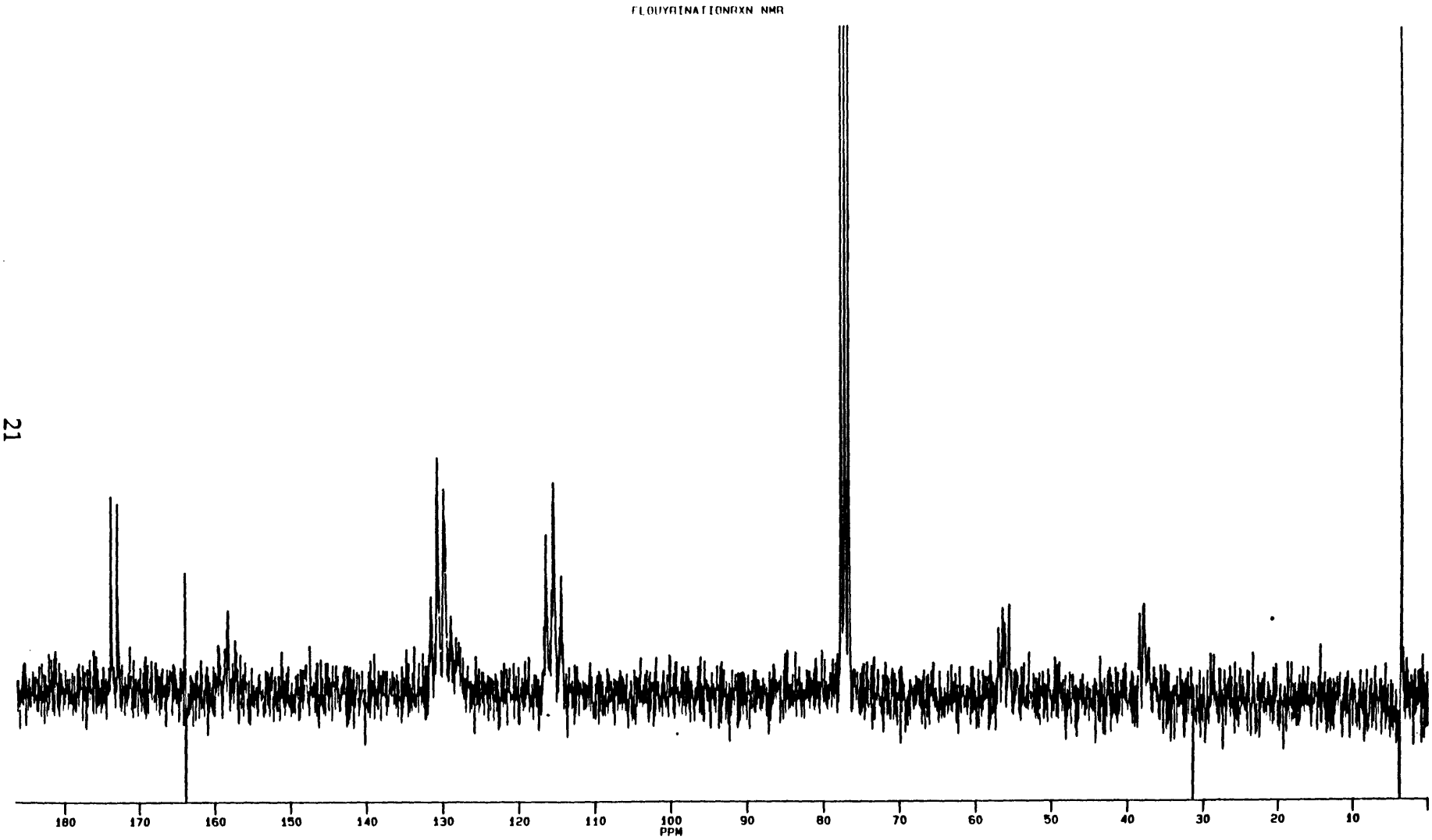


Figure 6: Carbon NMR of N-CBZ-Tyrosine-O-Benzyl Ether Fluoride Ester

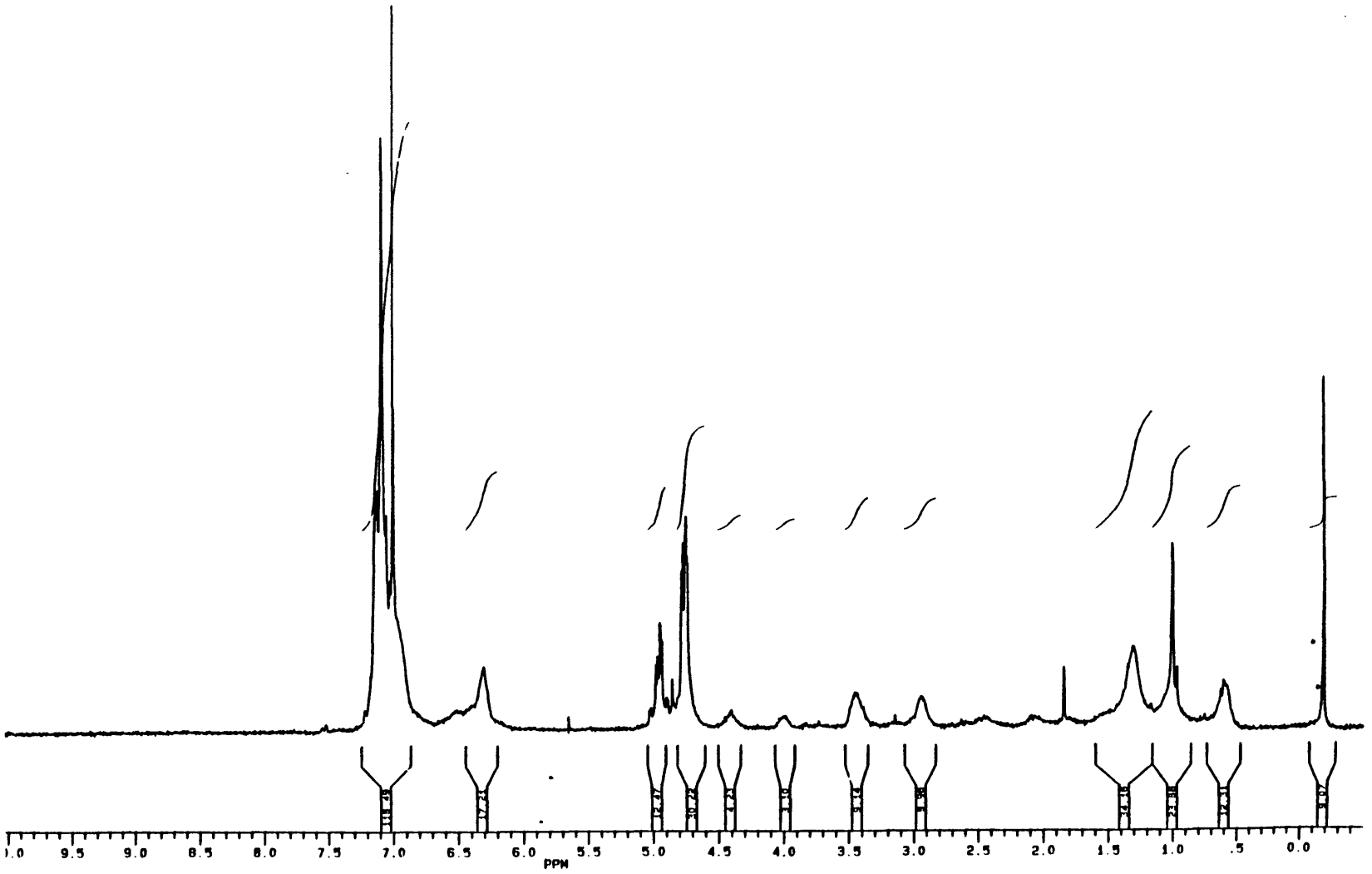


Figure 7: Proton NMR of N-CBZ-Tyrosine-O-Benzyl Ether Fluoride Ester

Tyrosine-Lysine

89.3 mg (0.00021 mol, 1 eq) of N-CBZ-Tyrosine-O-benzyl ether fluoride ester was dissolved in 3 ml of CH_3CN. 142 mg (0.00025 mol, 1.2 eq) of ε-N-CBZ-Lysine-O-benzyl ester p-toluenesulfonic acid salt and 0.047 ml (0.00035 mol) of 4-methyl-morpholine were added to the solution. The ε-N-CBZ-Lysine-O-benzyl ester p-toluenesulfonic acid salt had been previously neutralized with 0.047 ml (0.000355 mol) of 4-methyl-morpholine in 3 ml of CH_3CN and filtered. The reaction was monitored by TLC (10%AcOH/toluene; rf of N-CBZ-Tyrosine-O-benzyl ether acyl fluoride = 0.4, rf of ε-N-CBZ-Lysine-O-benzyl ester p-toluenesulfonic acid salt= 0.0, and rf of protected tyrosine lysine=0.6). The reaction was stirred for three hours. The product was purified on preparatory TLC (1000 microns; 10%AcOH/toluene; rf of N-CBZ-Tyrosine-O-benzyl ether fluoride ester=0.4, rf of ε-N-CBZ-Lysine-O-benzyl ester p-toluenesulfonic acid salt=0.0, and rf of protected tyrosine lysine=0.6). 28.9 mg (0.000038 mol, 5.6% yield) of protected tyrosine lysine was recovered.

28.9 mg (0.0000381) of protected-tyrosine lysine was then submitted to a trifluoromethanesulfonic (TFMSA) acid cleavage. Protected tyrosine lysine was dissolved in 750 µl of thioanisole/ethanedithiol. The mixture was then chilled in an ice bath and 5 ml of trifluoroacetic acid (TFA) was added and stirred for 5-10 minutes. 500 µl of TFMSA was then added dropwise. The reaction was stirred for 30 minutes. The TFA was removed under a stream of argon and the remaining solution was added dropwise to cold ether. Tyrosine-lysine immediately precipitated and was collected by centrifugation. The resulting precipitate was washed with cold ether several times to remove the scavengers. The resulting solid was purified using high performance liquid chromatography under isocratic conditions (100% water /0.1% TFA). Purified yield was 15 mg (0.0000376, 98.7% yield)

The product was identified through NMR (Figure 8 and 9); ^1H NMR (D_2O, 250 MHz) δ 1.0, 1.5, 2.0 (6H, Cβ-H,Cδ-H,Cγ-H of the lysine), δ 3.1,3.4, 2.9 (2H, Cβ-H of tyrosine), δ 6.6, 6.9, 7.3, 7.4 (4H, tyrosine aromatic H); ^{13}C NMR (D_2O, 60 MHz) δ 11 (m, Cβ lysine), δ 18 (m, Cδ lysine), δ 21 (m, Cε lysine), δ 30 (m, Cβ lysine), δ 48 (m, Cα lysine and tyrosine), δ 109, 121 (m, C tyrosine aromatic ring), δ 151 (m, C 6 tyrosine aromatic ring), δ 169 (m, COOH).

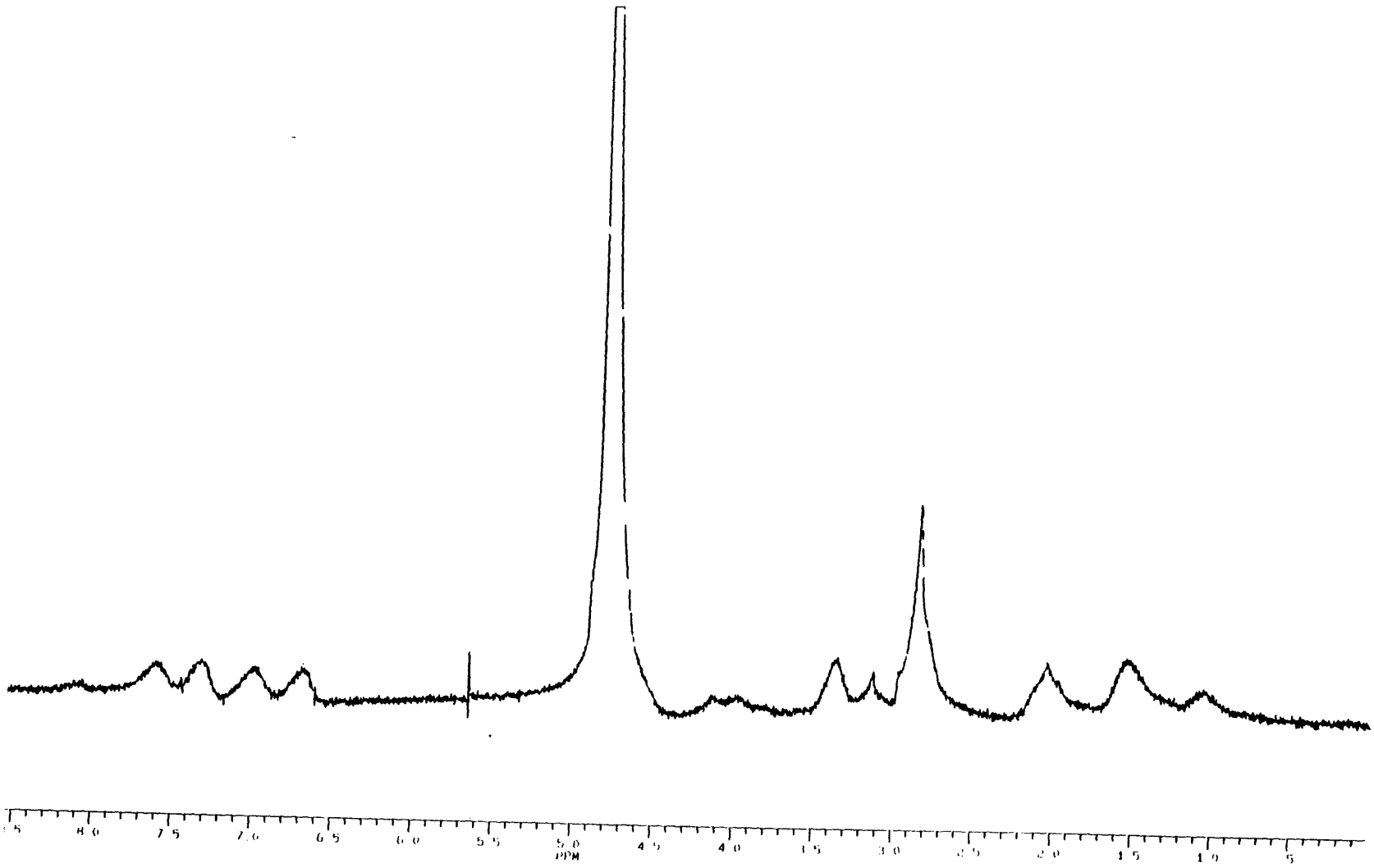


Figure 8: Proton NMR of Tyrosine Lysine

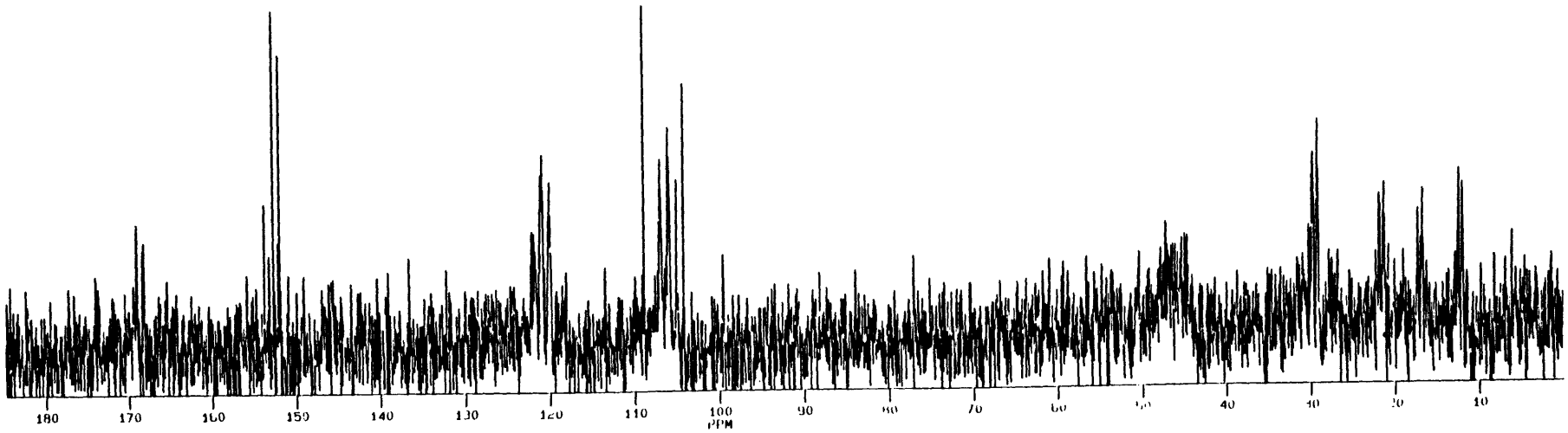


Figure 9: Carbon NMR of Tyrosine Lysine

Bertho, J., Loffet, A., Reuther, F., & Sennyey, G. (1991). Amino Acid Fluorides: Their Preparation and use in Peptide Synthesis. Tetrahedron Letters, 32(10), 1303-1306.

Bodanszky, M., & Bodanszky, A. (1984). The Practice of Peptide Synthesis . New York: Springer-Verlag.

Derome, A. E. (1987). Modern NMR Techniques for Chemistry Research . New York: Pergamon Press.

Griffiths, J. M., & Griffin, R. G. (1993). Nuclear magnetic resonance methods for measuring dipolar couplings in rotating solids. Analytica Chimica Acta, 283, 1081-1101.

Kelly, J. W., & Lansbury, P. T. (1994). A chemical approach to elucidate the mechanism of transthyretin and β-protein amyloid fibril formation. Amyloid: The International Journal of Experimental and Clinical Investigation, 1, 186-205.

Neuberger, A., & Sanger, F. (1943). The availability of the Acetyl Derivatives of Lysine for Growth. Biochemical Journal, 37, 518.

Urpi, L., Coll, M., & Subirana, J. A. (1988). Crystal and molecular structure of the dipeptide L-TYR-L-Lys. International Journal of Biological Macromolecules, 10, 55-59.

Chapter 3

Octarepeat Region of the PrP Protein

The PrP protein contains within it a repeat with the sequence PHGGGWGQ (Figure 10).

SHaPrP MANLSYWILALFVAMWTDVGLCKKRPKPGGWNTGGS RYPG

QGSPGGGNRYPPQGGGTWGQ**PHGGGWGQPHGGGWGQPHGGGWGQ**

PHGGGWGQGGGTHNQWNKPSKPKTNMKHMAGAAAAGAVVGGLGG

YMLGSAMSRPMMHFGNDWEDRYYRENMNRYPNQVYYRPVQYNNQN

NFVHDCVNITIKQHTVTTTTKGENFTETDIKIMERVVEQMCTTQYQKESQ

AYYDGRRSSAVLFSSPPVILLISFLIFLMVG

Figure 10: Syrian Hamster PrP protein with the octamer repeat region highlighted

It was discovered that an insertion of about 150 bp in the open reading frame of the PrP gene segregated with Creutzfeldt-Jakob disease(Owen, et al., 1989). The insertion was later found to be an extra 6 repeat sequences in addition to the 5 repeats normally found in the open reading frame of the human PrP gene (Owen, et al., 1990). Since it has been hypothesized that the only difference between PrP^{C} and PrP^{Sc} is structural it seemed appropriate to examine the ocatreapeat region and then possibly look at the effect of the addition of extra octarepeat regions. A simple yet powerful approach to examining the secondary structure of the repeat region was chosen, circular dichroism spectroscopy (CD) (Johnson, 1990). CD would permit the examination of the gross secondary structure of the PrP repeat region and any

changes that might occur upon addition of more repeats. Therefore, it was decided that an 8mer, 16mer, and 24mer would be synthesized (See Figure 11) and examined by CD.

PHGGGWGQ

PHGGGWGQPHGGGWGQ

PHGGGWGQPHGGGWGQPHGGGWGQ

Figure 11: 8mer, 16mer, and 24mer

Experimental

The three peptides were made using standard fluorenylmethoxycarbonyl (Fmoc) chemistry on the Rink Amide MBHA resin.

The Fmoc protecting group was cleaved by treatment with 50% piperidine in dimethylformamide (DMF) for 20 minutes. The resin was washed 3 times with DMF, 3 times with methanol (MeOH), and 3 times with methylene chloride (CH_2Cl_2). A Kaiser test was performed. The resin was then washed 3 times with DMF. Then 2 equivalents of amino acid and benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium hexafluorophosphate (PyBOP) were added with 3.3 eq of diisopropylethylamine (DIEA). The coupling was allowed to proceed for 1 hour at which time the resin was washed 3 times with DMF, 2 times with CH_2Cl_2, 2 times with MeOH, and 2 times with CH_2Cl_2. A Kaiser test for free amines then performed. The resin was acetylated after couplings by treatment with 5 equivalents acetic anhydride and 5 equivalents of DIEA in CH_2Cl_2.

The resin was cleaved and deprotected using the reagent K cocktail (82% TFA, 5% H_2O, 5% thioanisole, 5% phenol, and 2.5% ethanedithiol). The resin was allowed to stir for two hours. The mixture was filtered and washed 4 times with neat TFA. The resulting solution was evaporated by blowing argon over it. Ether was added to the remaining solution. The resulting solid was triturated with ether several times. The solid was then air dried and

submitted to RPHPLC. All three peptides were purified under the following conditions. 0.1% TFA CH_3CN and H_2O. 90% water : 10% CH_3CN for ten minutes then 90% water : 10% CH_3CN to 100% CH_3CN over 20 minutes on a C18 RPHPLC column. Analytical RPHPLC were performed under the same conditions unless otherwise stated.

All peptides were characterized by reverse phase high performance liquid chromatography (RPHPLC) (Compounds were monitored at 225 and 254 nm) and mass spectra.

NH_2-PHGGGWGQ-OH Plasma desorption mass spectra MH^+ calculated = 795.8; experimental = 794.5. (Figure 11, 12) Peak at 817 is sodium salt of MH^+. 80 mg Crude. 9.6 mgs Pure.

NH_2-PHGGGWGQPHGGGWGQ-OH Plasma desorption mass spectra MH^+ calculated = 1572.6; experimental = 1572. (Figure 13, 14) Peaks at 1595.2 and 1730.2 are the sodium salts of MH^+ peaks. 80 mg Crude. 3.8 mg Pure.

NH_2-PHGGGWGQPHGGGWGQPHGGGWGQ-OH Plasma desorption mass spectra MH^+ calculated = 2349.4; experimental = 2348.5. (Figure 15, 16) Peak at 817 is sodium salt of MH^+. 600 mg Crude. 8,2 mg Pure.

It should be noted that for the 16mer and 24mer the peaks at MH^+ 137 are an impurity containing an extra Histidine residue. This was due to the use of incompletely protected Histidine. The Fmoc Histidine was only partially protected on the amino group with Fmoc allowing the addition of an unwanted histidine to the peptide Therefore, the 16mer and 24mer contain the corresponding 17mer and 25mer impurity that contains an extra histidine. Removal of this impurity through the use of RPHPLC under isocratic conditions was not possible. This impurity may affect the CDs of the 16mer and 24mer but in what manner is not clear.

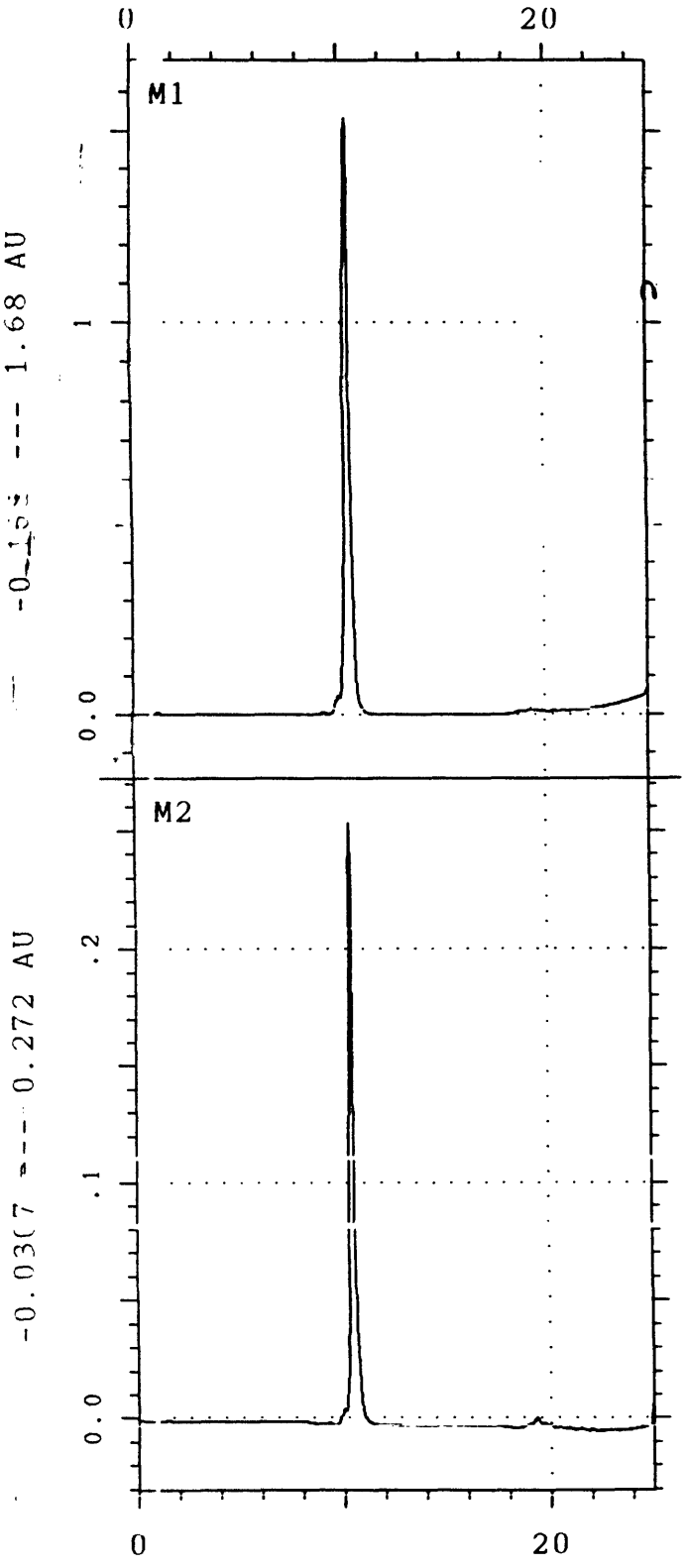
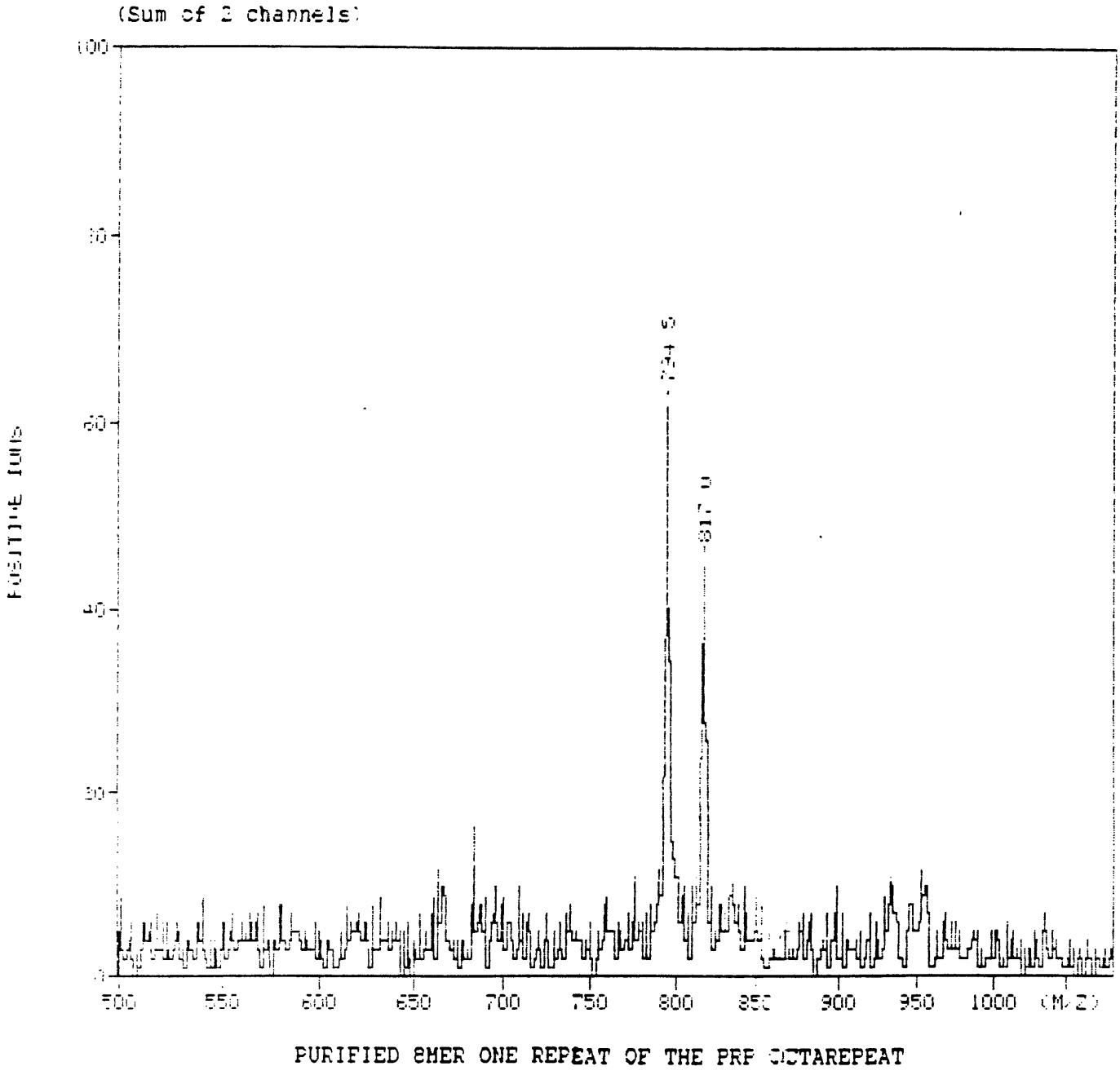


Figure 11: Analytical RPHPLC of 8mer
(0.1% TFA CH_3CN and H_2O. 90% water : 10% CH_3CN for ten minutes then 90% water : 10% CH_3CN to 100% CH_3CN over 20 minutes on a C18 RPHPLC column.)



Figrue 12: PDMS of 8mer

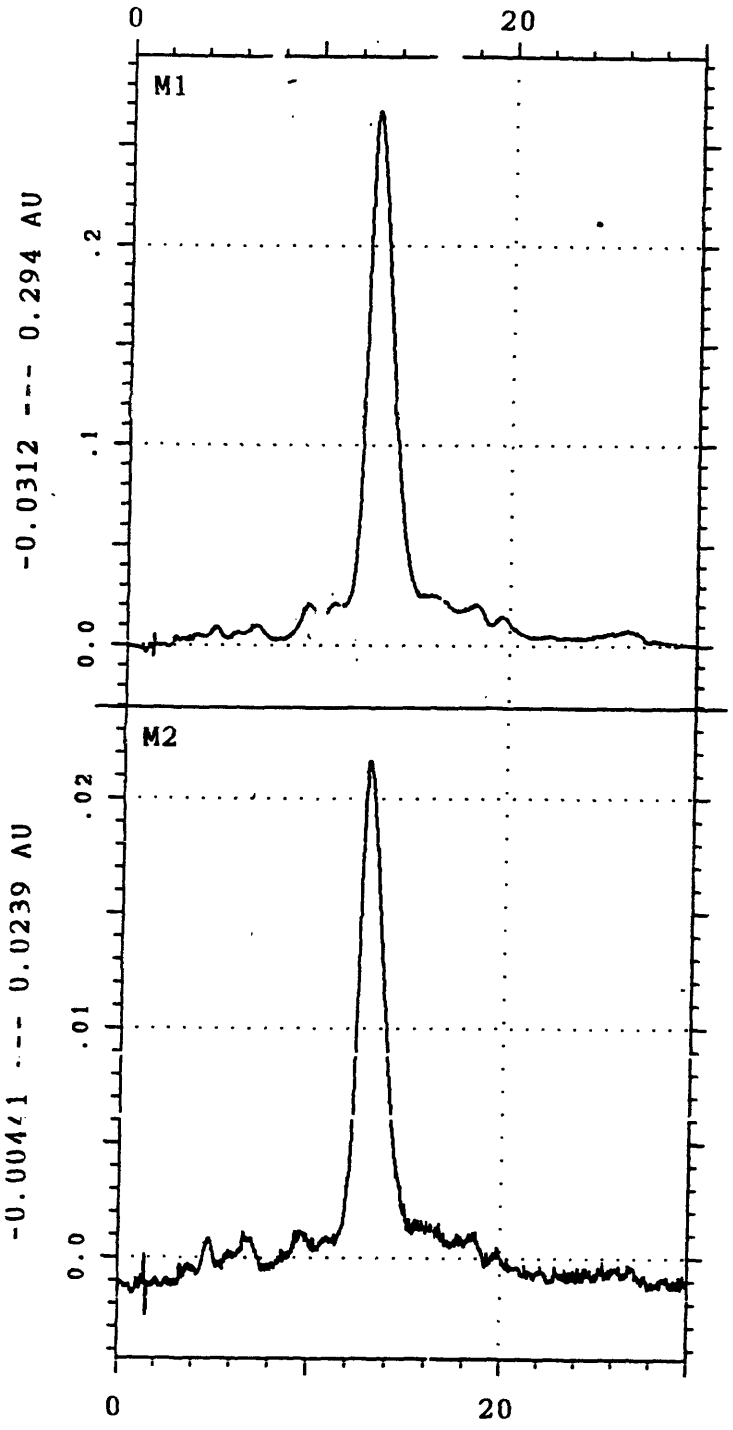
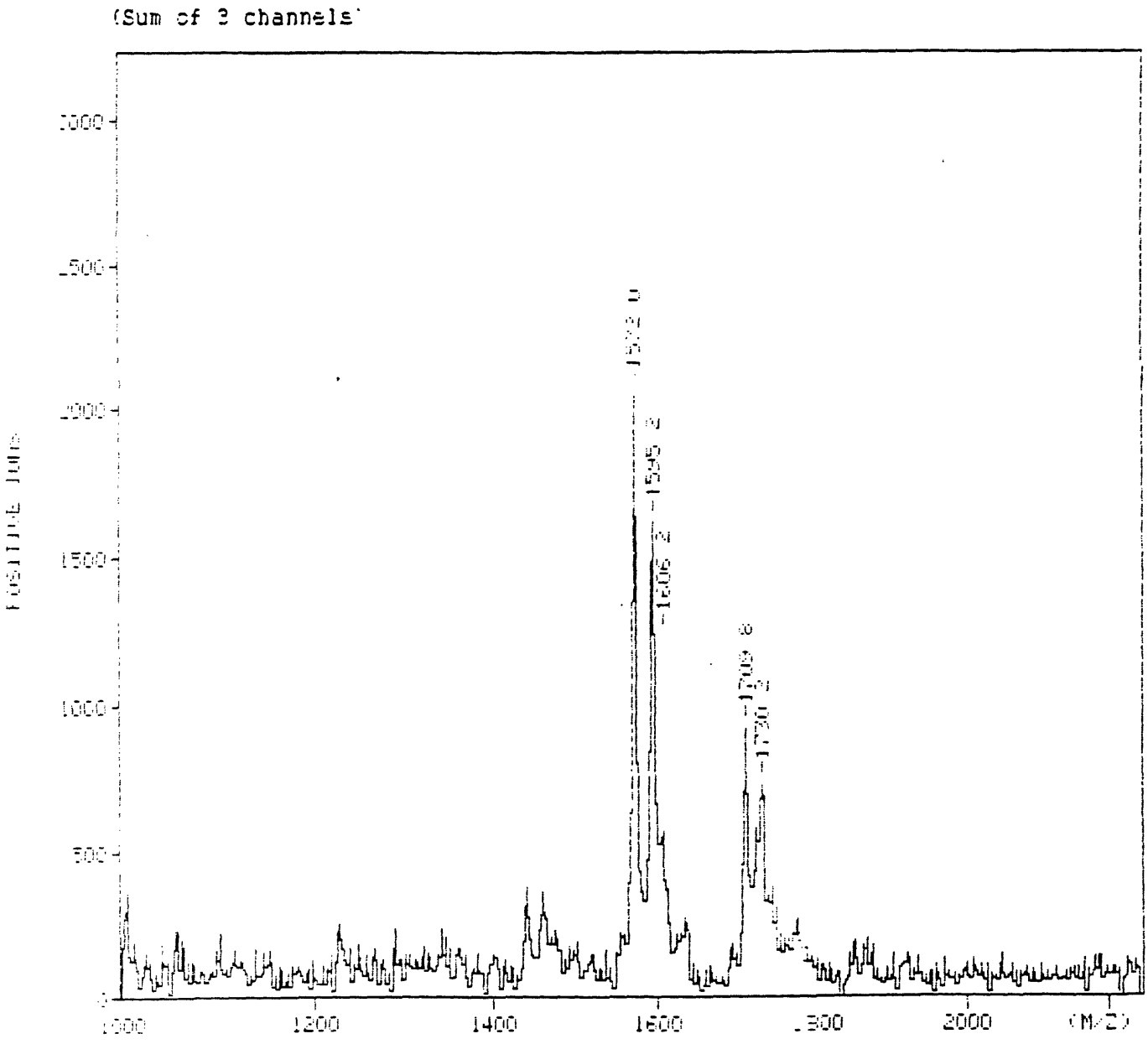


Figure 13: Analytical RPHPLC of 16mer
(Isocratic Conditions 86% H_2O : 14% CH_3CN)
(0.1% TFA CH_3CN and H_2O)



Figrue 14: PDMS of 16mer

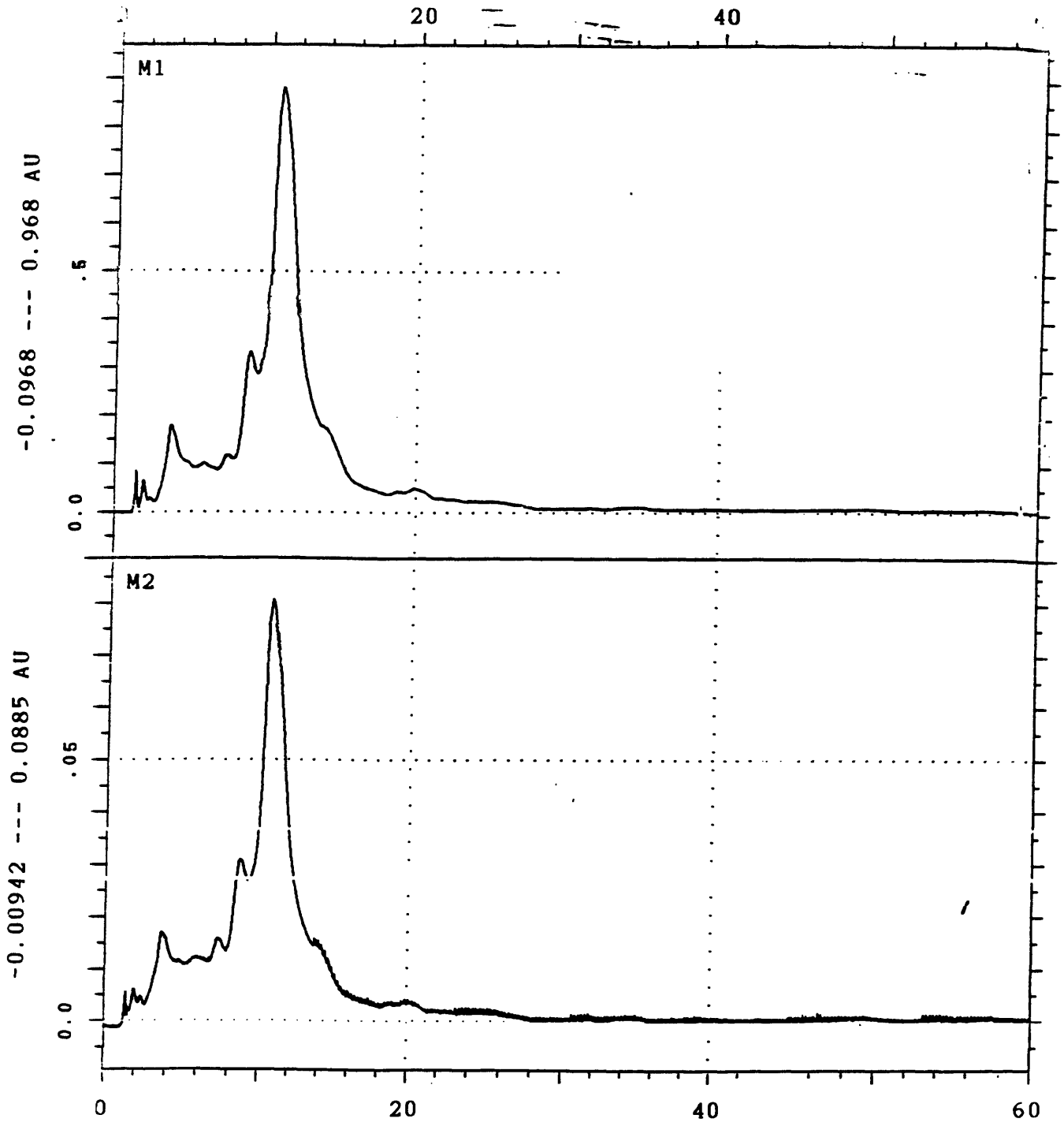
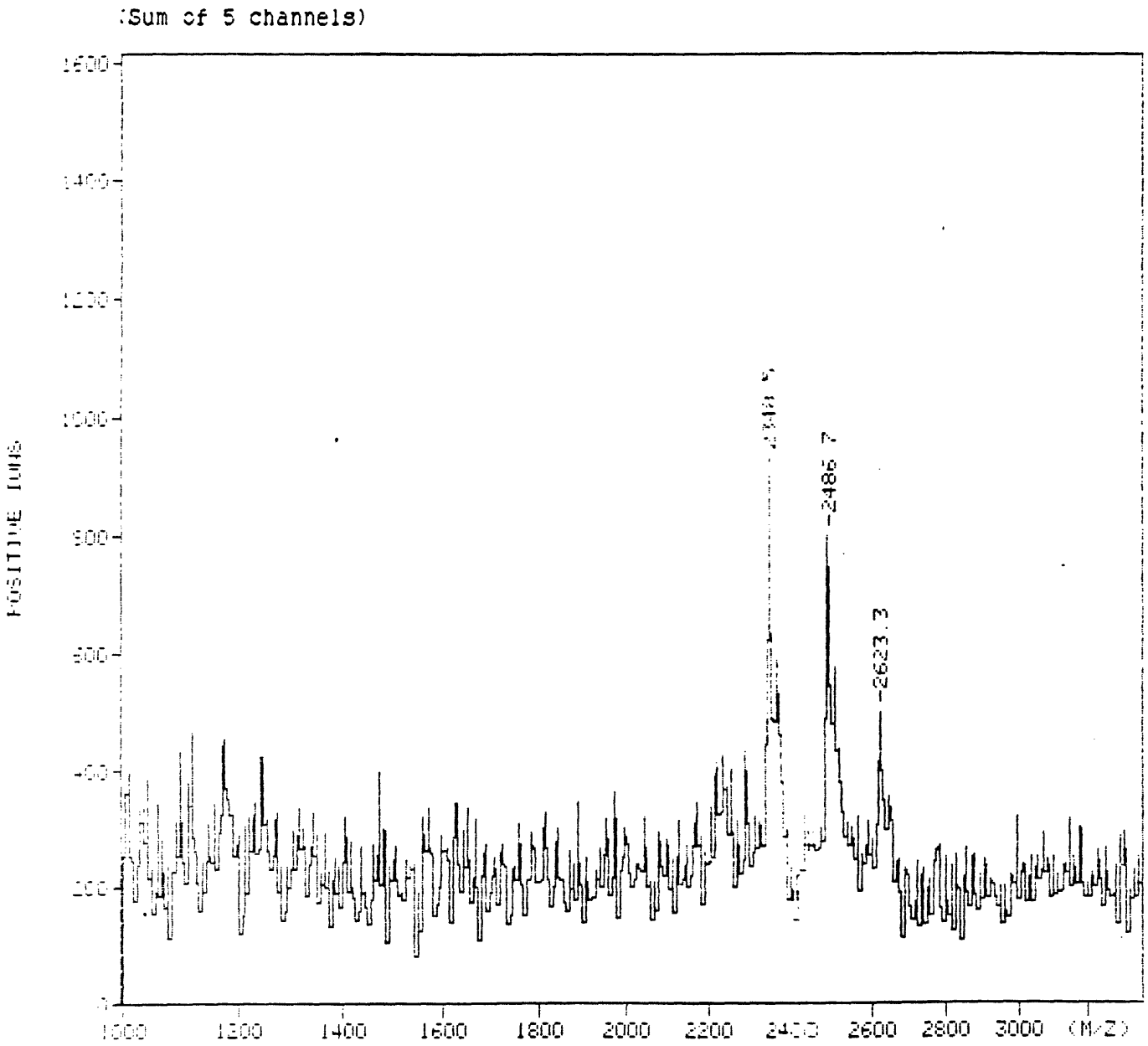


Figure 15: Analytical RPHPLC of 24mer
(Isocratic Conditions 83% H_2O : 17% CH_3CN)
(0.1% TFA CH_3CN and H_2O)



Figrue 16: PDMS of 24mer

Circular Dichroism

CD was run in an aqueous buffer that was 50 mM in triethylamine and 50 mM in sodium acetate at a pH of 7.4.

All spectra were taken on a Jasco J500 spectropolarimeter between 190 and 300 nm at a scan speed of 100 nm per minute with a sensitivity of 10 mdeg. A cell of path-length 0.1 cm was used. Three runs were averaged for every spectra

Results

The 8mer, 16mer, and 24mer were synthesized and purified and a CD study of concentration was performed. CD's of 8mer, 16mer, and 24mer at 5mg/ml, 4mg/ml, 3mg/ml, 2mg/ml, 1mg/ml, 0.5mg/ml, 0.1mg/ml were taken (Figure 17-19).

mdeg = ellipticity in millidegrees

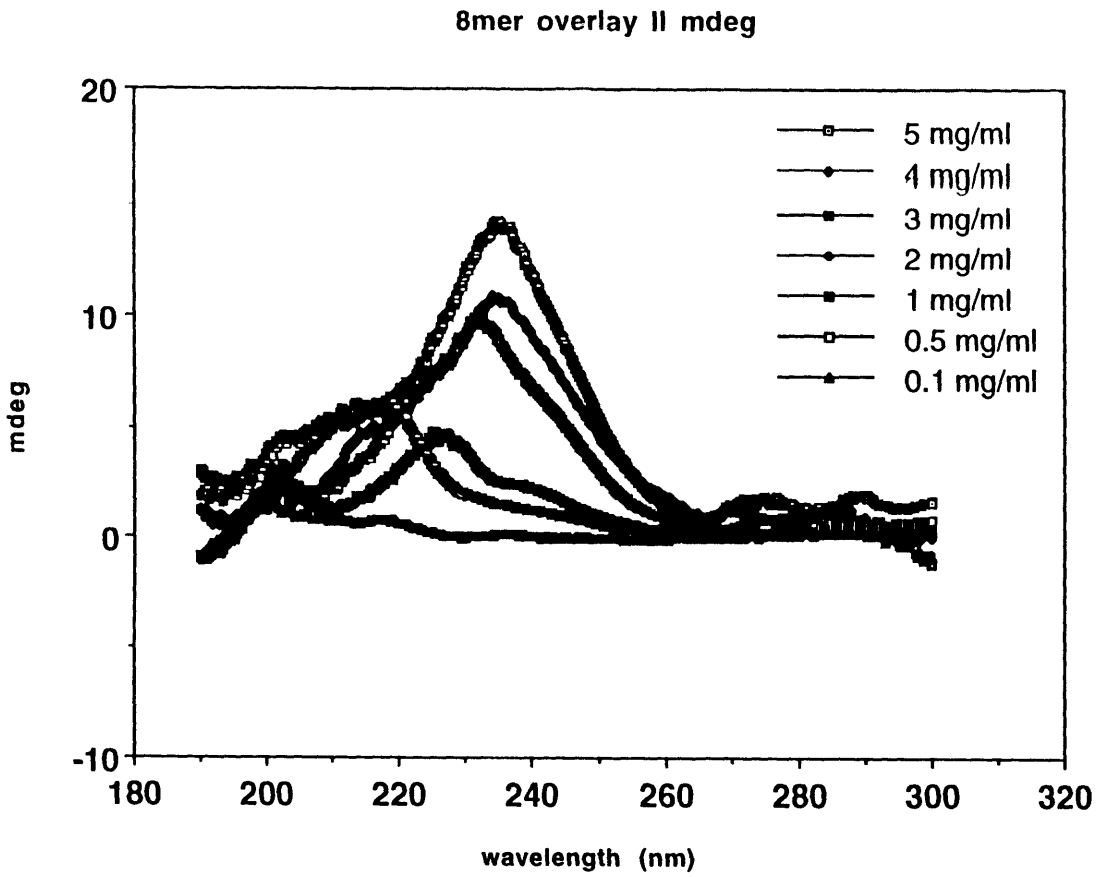


Figure 17: Overlay of 8 mer CD's at varying concentrations

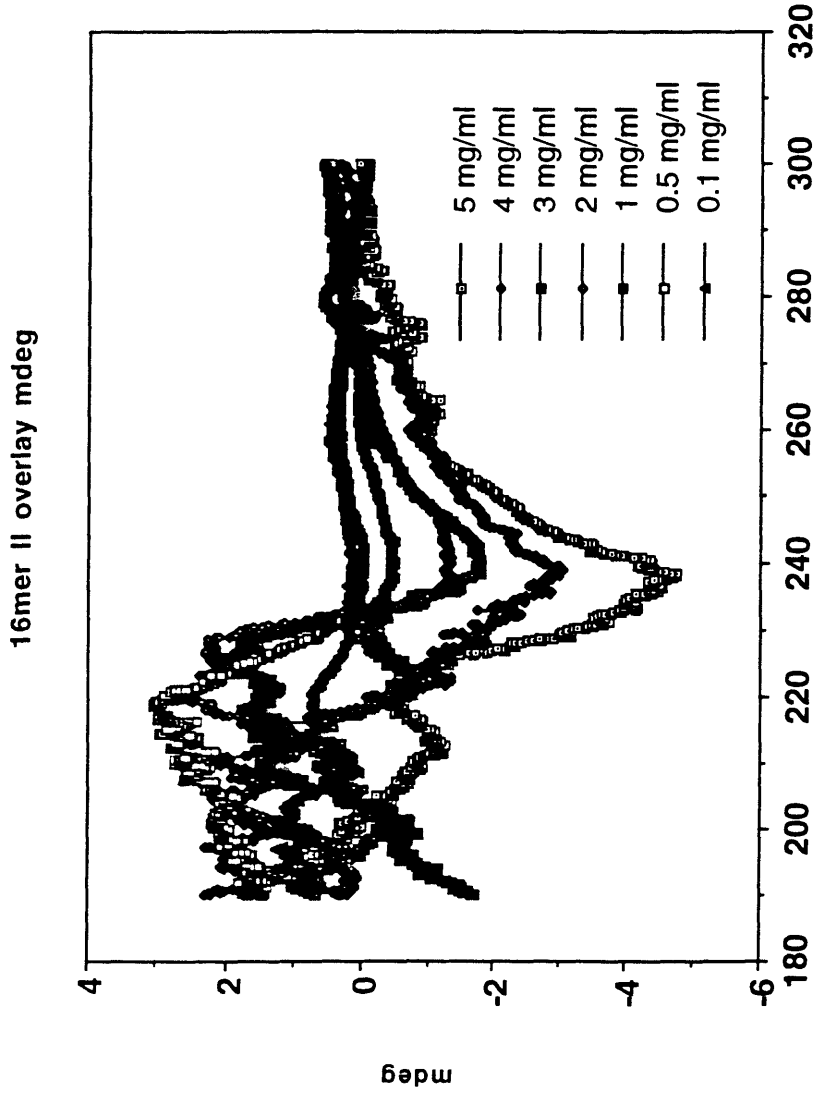


Figure 18: Overlay of 16 mer CD's at varying concentrations

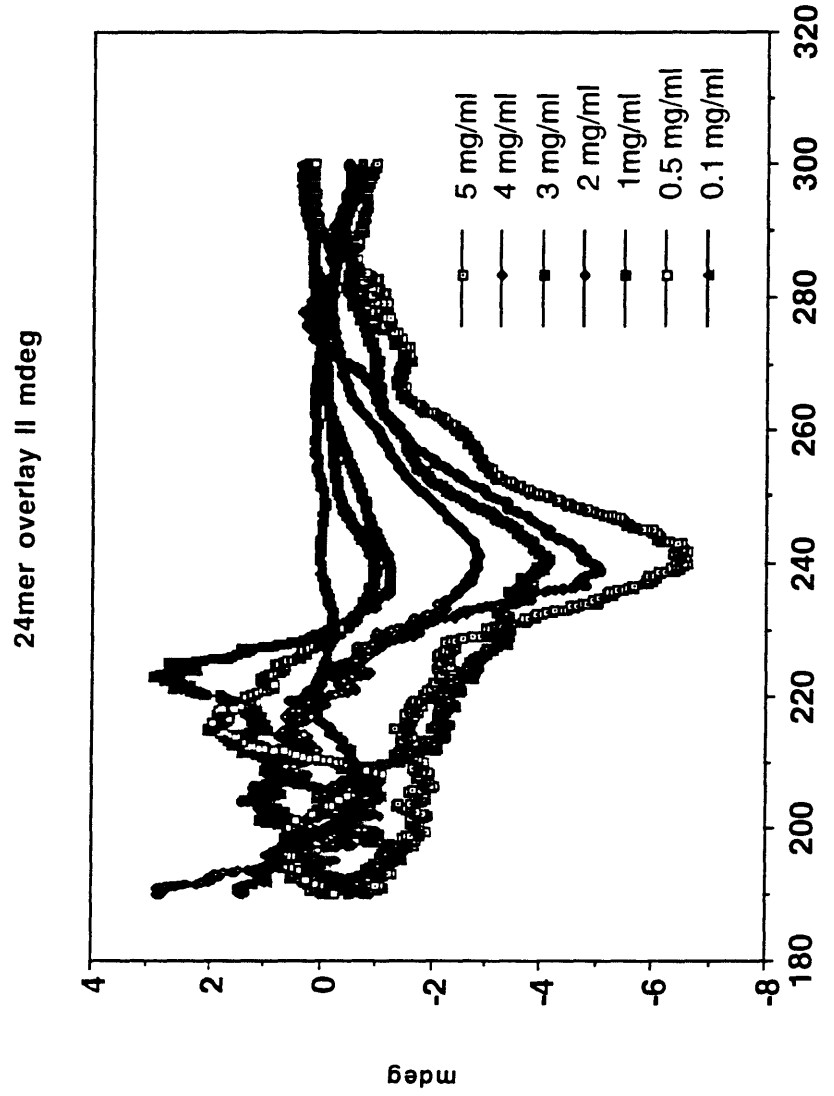


Figure 19: Overlay of 24 mer CD's at varying concentrations

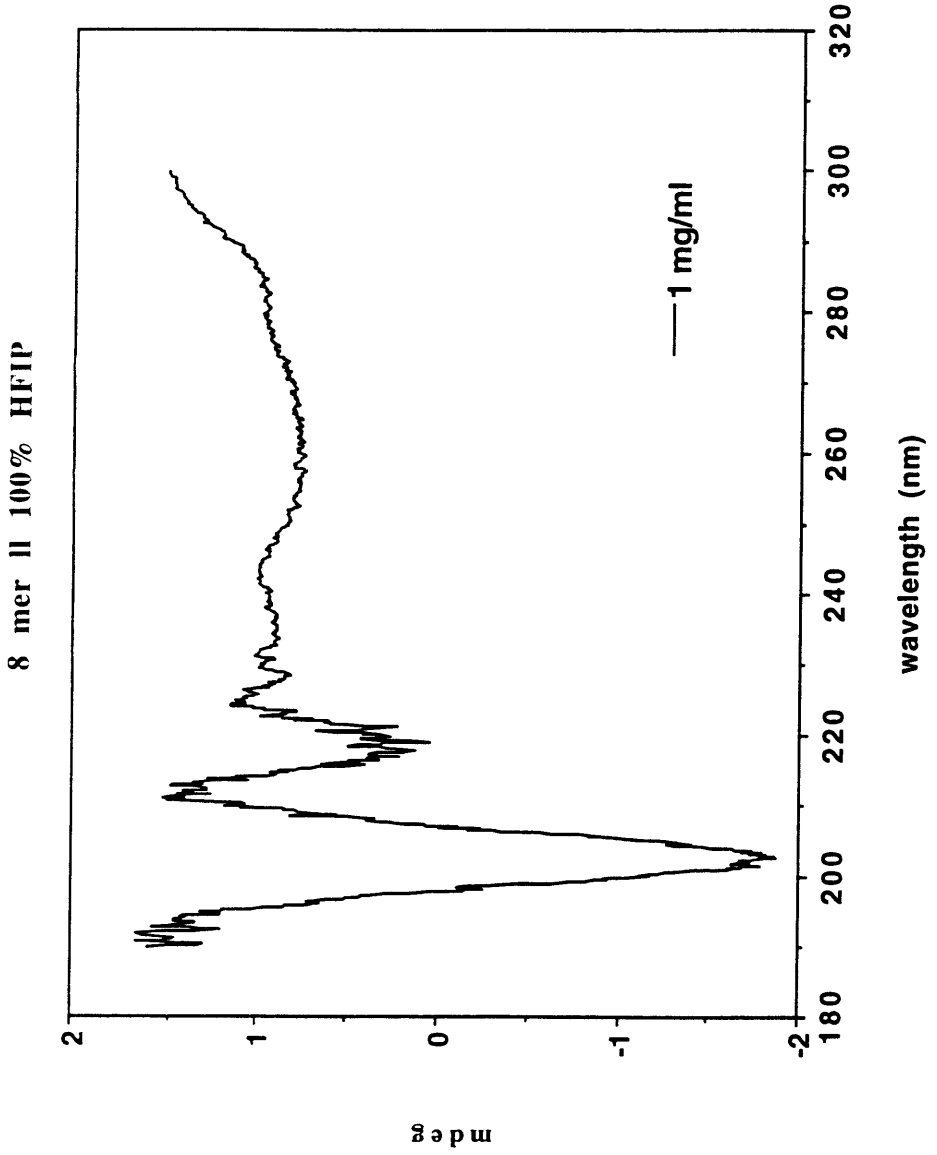


Figure 20: CD of 8 mer in 100% HFIP

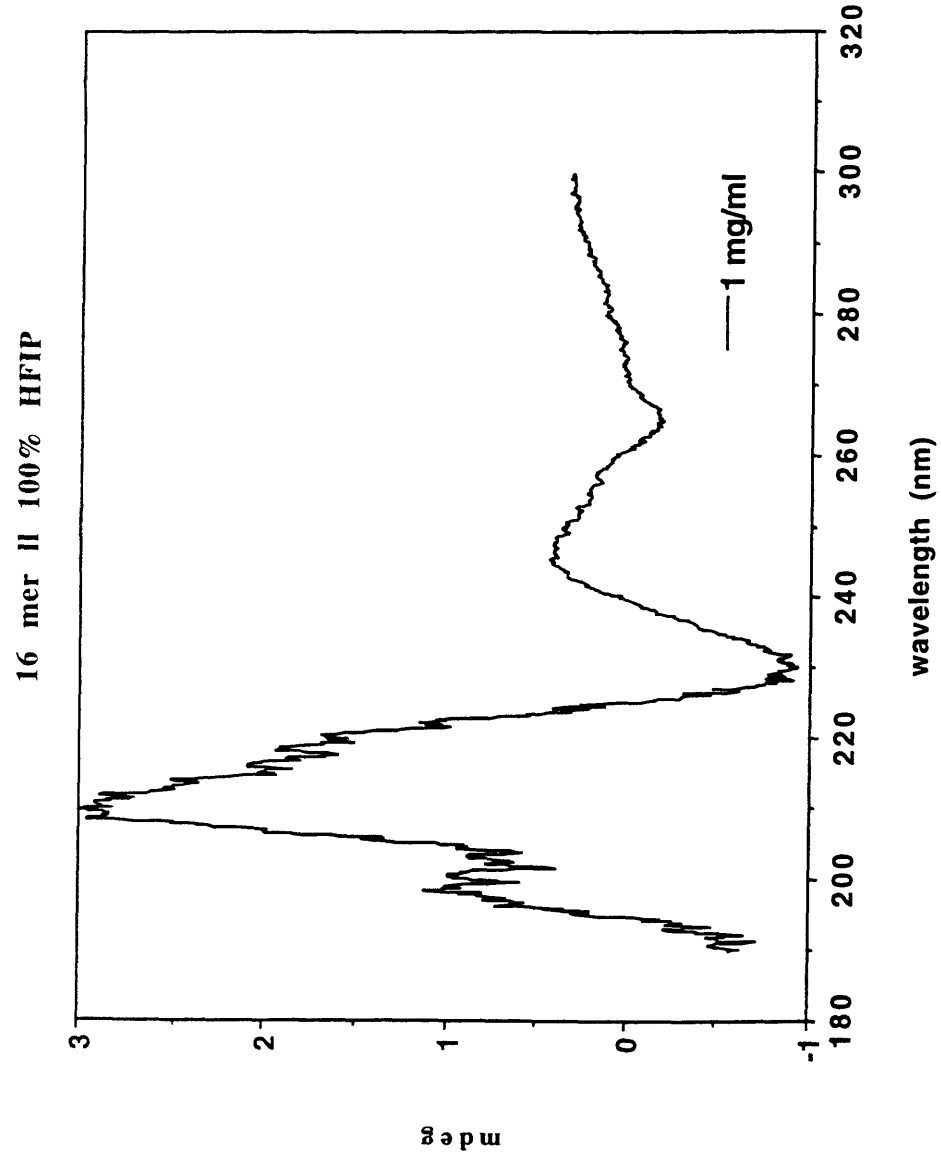


Figure 21: CD of 16 mer in 100% HFIP

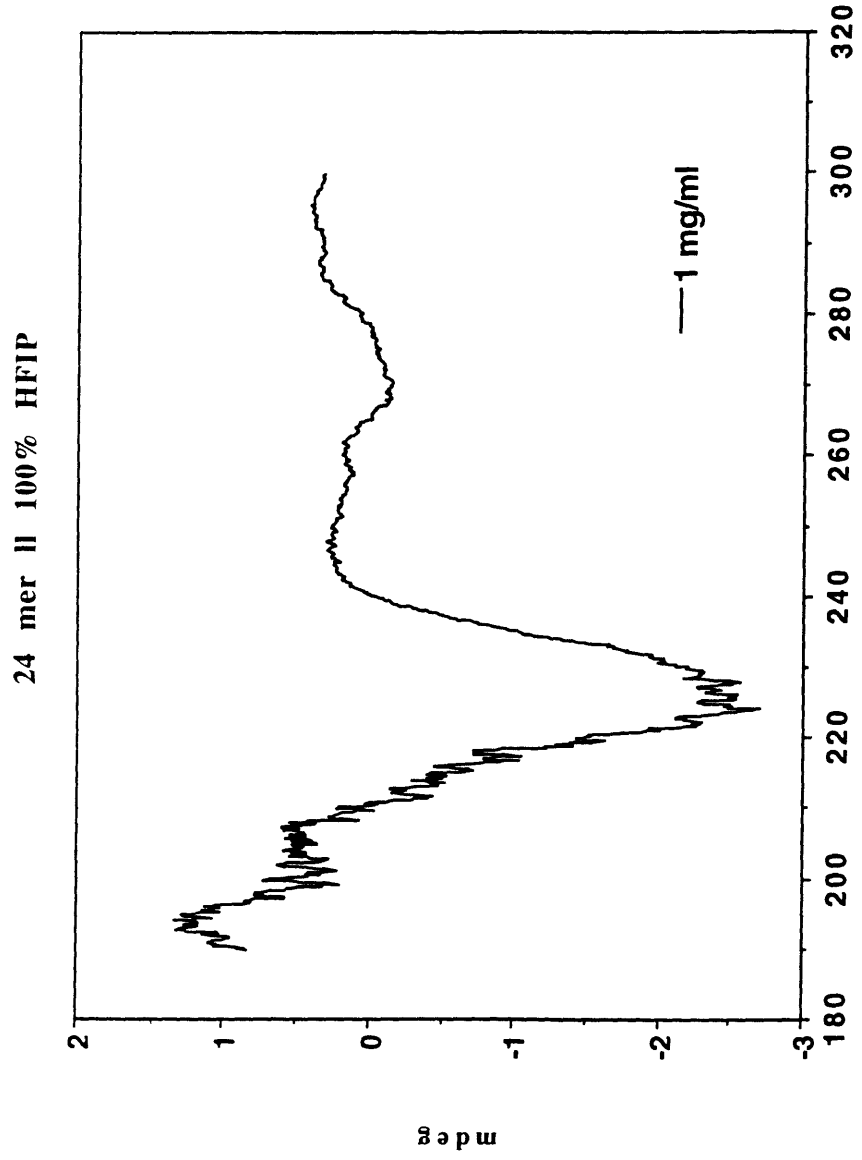


Figure 22: Figure CD of 24mer in 100% HFIP

Discussion

CD typically gives a good indication of four types of secondary structure: α helices, β sheets, β-tuns, and random coil (Figure 23).

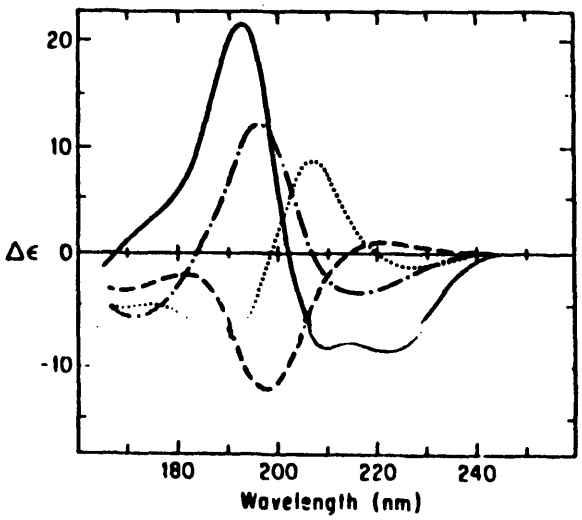


Figure 23: CD spectra for the α-helix (solid line), β-sheet (dots and dashes), β-turn (dotted line), and random coil (dashed line) (Johnson Jr., 1988)

The resulting CD's indicate no classic or known structure. The absorbance of the 8 mer is positive in a region where typically there is no positive absorbance. If the absorbance were at a lower wavelength the compound might be said to be composed of β-turn. It is interesting to note that as the concentration decreases the absorbance maximum decreases toward the β-turn region. This slight movement in the absorbance maximum indicates a concentration dependence, implying that there may be some interaction between the individual 8 mer units at higher concentrations. However,

predictions about 8 mer structure must be made carefully as small peptides normally do not take on well defined structures.

If Figures 18 and 19 are examined it can be seen that the CD spectra given by the 16 mer and 24 mer are very different from the 8 mer spectra. Although, the absorbance is centered around 238 nm as with the 8 mer it is negative. This negative absorbance, as far as the author is aware, does not correspond to any known structure. As with the 8 mer there seems to be a slight concentration dependence In most respects the spectra of the 16 mer and the 24 mer are similar. The largest difference between spectra occurs at the absorbance maximum. In the case of the 16 mer the maximum absorbance is centered around 238 nm, the 24 mer is centered around 240 nm. Since, the 16 mer and 24 mer spectra were so unusual an attempt was made to confirm that the spectra were a result of peptide conformation and not just simply some artifact. It was decided to attempt the induction of a conformational change in the peptides to a known structure using the α-helix inducing solvent hexafluoroisopropanol (HFIP). As Figures 20-22 display the results of this test produced some interesting results of their own.

Once again the 8 mer spectra is different from the 16 mer and 24 mer. The spectra of the 8 mer has absorbances that correspond to both β-sheet (210 nm) and β-turn (222 nm).

The 16 mer and 24 mer interestingly have an absorbance that roughly corresponds to β-turn(Brahmachari, Rapaka, Bhatnagar, & Ananthanarayanan, 1982). In these papers linear and cyclic peptides that formed β-turns were examined. In the Brahmachari paper the peptides contain proline and glycine which when found in a peptide have a propensity to form β-turns. Examining the sequence of the repeats (PHGGGWGQ) it is not hard to believe that the HFIP induced the 16 mer and 24 mer to form β-turn instead of α-helix.

In conclusion these preliminary studies on the octarepeat have produced some very interesting results. There are spectra whose absorbances are as of the time of this writing still unknown. There is a small concentration dependence which might imply some intermolecular interaction. Finally, there is still the yet unanswered and unexplored question of how this data affects prion pathology.

Brahmachari, S. K., Rapaka, R. S., Bhatnagar, R. S., & Ananthanarayanan, V. S. (1982). Proline-Containing β-Turns in Peptides and Proteins. II. Physicochemical Studies on Tripeptides with the Pro-Gly Sequence. Biopolymers, 21, 1107-1125.

Johnson Jr., C. W. (1988). Secondary Structure of Proteins Through Circular Dichroism Spectroscopy. Ann. Rev. Biophys. Biophys. Chem.

Johnson, W. C. (1990). Protein Secondary Structure and Circular Dichroism A Practical Guide. PROTEINS: Structure, Function, and Genetics, 7, 205-214.

Owen, F., Poulter, M., Lofthouse, R., Collinge, J., Crow, T. J., Risby, D., Baker, H. F., & Ridley, R. M. (1989). Insertion in Prion Protein Gene in Familial Creutzfeldt-Jakob Disease. Lancet, i, 51-52.

Owen, F., Poulter, M., Shah, T., Collinge, J., Lofthouse, R., Baker, H., Ridley, R., McVey, J., & Crow, T. J. (1990). An in-frame insetion in the prion protein gene in familial Creutzfeldt-Jakob disease. Molecular Brain Research, 7, 273-276.