AN ABSTRACT OF THE THESIS OF

Chartchai Krittanai for the degree of Doctor of Philosophy in Biochemistry and Biophysics presented on April 29, 1997.

Title: Helical Propensity of Amino Acids Changes with Solvent Environment.

Abstract Approved: Redacted for Privacy

W. Curtis Johnson, Jr.

The propensity of amino acids to form an α-helix is investigated using circular dichroism (CD) spectroscopy. A peptide model of sequence acetyl-Y-VAXAK-VAXAK-VAXAK-amide, where X is substituted with one of the 20 naturally occurring amino acids, was synthesized by a solid phase method and titrated from phosphate buffer (pH 5.5) to 90% methanol. The CD spectra indicate a conformational change for peptides from random coil to α-helix when the concentration of methanol is increased. Assuming a two-state model, a free energy of α-helix formation, representing the helical propensity, is calculated for each amino acid as a function of methanol content. The free energy of these amino acids are compared in various solvent conditions. The plot of free energy demonstrates changes in relative order of helical propensity upon changing the solvent environment. The order of helical propensity in 63% methanol solution agrees well with the propensity order obtained by using statistical calculation

from the X-ray crystallography database. It suggests that the 63% methanol/ buffer is similar to an average environment found in proteins.

A method is proposed to correct the peptide CD for a contribution from absorbing side-chains in the far ultraviolet (UV) range. The method performs a Singular Value Decomposition (SVD) analysis to extract the common basis vectors and the singular values from related peptides without absorbing side-chains. The common basis vectors and their singular values are fitted to a portion of CD spectra of peptide being corrected in the range that is free from side-chain contribution. The coefficients obtained from the fitting are then used to regenerate the corrected CD spectra from the common basis vectors. The corrections show a positive contribution for the aromatic side-chains of W, Y, F, and a negative contribution for the sulfur-containing side-chain of M and C, in the 225-230 nm region.

Helical Propensity of Amino Acids Changes with Solvent Environment.

by

Chartchai Krittanai

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented April 29, 1997
Commencement June 1997

Doctor of Philosophy thesis of Chartchai Krittanai presented on April 29, 1997.

APPROVED:

Redacted for Privacy

Major Professor, representing Biochemistry and Biophysics

Redacted for Privacy

Chair of Department of Biochemistry and Biophysics

Redacted for Privacy

Dean of Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Redacted for Privacy

Chartchai Krittanai, Author

ACKNOWLEDGMENTS

First of all, I would like to thank my major professor, Dr. W. Curtis Johnson, Jr. for his guidance and encouragement. He has taught me not only how to do excellent research but also how to be a good scientist. This work would never be possible without him.

I would like to thank my graduate committee members, Dr. Pui Shing Ho, Dr. Victor L. Hsu, Dr. Joseph W. Nibler, and Dr. Carol J. Rivin for their help and discussion in guiding my graduate program.

I would like to thank Dr. Araz Toumadje, Dr. Vincent Waterhous, and Jeannine R. Lawrence for their helpful discussion and making the laboratory a fun place to work.

I would also like to thank the Development and Promotion of Science Talents Project (DPST) of the Royal Thai Government for their support and belief in science.

This dissertation is dedicated to my parents for their love, and to all my past teachers.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF TABLES

PREFACE

This thesis contains three parts based on manuscripts written for publication. These manuscripts result from a learning process of using circular dichroism and linear dichroism spectroscopy to study the secondary structure of nucleic acids and polypeptides. The first two manuscripts deal with the correcting of CD spectra for the contribution from absorbing side-chains of amino acids, and the investigating of helical propensity of amino acids in a changing solvent environment, respectively. These two manuscripts have been submitted for publication. The third manuscript has been published in *Biospectroscopy* (1995) **1**, 247-254. It deals with base inclinations of left-handed DNA and RNA as revealed by linear dichroism. Since this research is not related directly to the first two manuscripts, it is included as an Appendix. The Appendices also include the experimental method of peptide synthesis and peptide aggregation study, in detail. The format of the three manuscripts has been restyled slightly to meet the standard requirement of the university thesis format.

HELICAL PROPENSITY OF AMINO ACIDS CHANGES WITH SOLVENT ENVIRONMENT

CHAPTER I

INTRODUCTION

Historical Background

The study of protein structure began in the 1950's with proposed models for the α-helix and β-pleated sheets, the fundamental structural elements that constitute a protein molecule (Pauling et al., 1951; Pauling & Corey, 1951). These predicted structural elements were validated by the then new breakthrough, X-ray crystallography, a powerful technique for determining protein structure (Kendrew et al., 1960; Blake et al., 1965; Perutz et al., 1965; Kartha et al., 1967; Matthews et al., 1967).

In 1961, the refolding experiment of denatured ribonuclease by Anfinsen et al. demonstrated that under the right conditions a denatured protein can refold into a native structure and resume its functional activity. This work suggested that the information encoded within an amino acid sequence is the key factor in dictating the protein native structure. The concept has been widely accepted because a distinct native structure for a protein is normally found to correspond to a particular amino acid sequence. In the hierarchic point of view, it is important to examine the information from amino acid sequence since the folding process begins by the formation of secondary structure from the amino acid

sequence. Since then, researchers have put a lot of effort into studying amino acid sequence, mostly by looking at the structures solved by X-ray crystallography and trying to find a correlation with the sequences. The ultimate goal of this research has been to be able to predict the native structure of a protein from its amino acid sequence.

Early attempts to predict a secondary structure from a primary sequence has concentrated on the α-helix (Guzzo, 1965; Prothero, 1966; Periti et al., 1967; Low et al., 1968). More recent attempts included the β-sheet (Kuntz, 1972; Rose & Seltzer, 1977; Chou & Fasman, 1979). There are many predicting approaches available now, based on sequence homology (Ponger & Szaley, 1985; Levin et al., 1986; Nishikawa & Ooi, 1986; Sweet, 1986; Zvelebil, 1987), linear optimized predictors (Edelmand & White, 1989), double code methods (Shestopalov, 1990), nearest neighbors (Salzberg & Cost, 1992; Yi & Lander, 1993), segment-based approaches (Presnell et al., 1992), neural networks (Qian & Sejnowsky, 1988; Holley & Karplus, 1989; Stolorz et al., 1992; Rost & sander, 1993), and statistical calculations for amino acid preferences (Chou & Fasman, 1974a,1974b, 1978; Burgess et al., 1974; Lim, 1974; Robson & Suzuki, 1976; Garnier et al., 1978; Gibrat et al., 1987). Other predicting methods still appear.

Even with the increasing power of today's computing and the expanding number of structures in the X-ray database, these predicting methods still have a limited success of about 70%. A good discussion on the prediction limit was published by Palau et al. (1982) and Rao et al. (1993). The reason for the imperfection of predicting is generally believed to be the effect of non-local

interactions within proteins. Dill (1990) has proposed a nonhierarchical model for protein folding in which the driving force comes from the hydrophobicity of amino acid sequence. In this model, a hydrophobic collapse occurring in the earlier stage provides specific local interactions for a particular sequence to be folded into secondary structures, and then stabilized by other forces.

We believe that the correct folded secondary structure is the result of the surrounding solvent environment as well as information from amino acid sequence. The hydrophobic collapse in Dill's model may be seen as a means of providing a micro-solvent environment for each particular sequence. It has been widely recognized that solvent can influence the formation of secondary structure of peptides and proteins. For example, gramicidin structure can be varied depending on its solvent environment (Short et al., 1987; Killian et al., 1988); many signal peptides are α-helices in lipid monolayer but are β-strands at the surface (Briggs et al., 1986; Cornell et al., 1989). Amino acids with a high preference for α-helix can be induced to form β-strands (Brahm et al., 1977). There also are peptide sequences that can be manipulated to form an α-helix, random coil and β-sheet by changing the solvent environment, regardless of the prediction from their amino acid sequences (Zhong & Johnson, 1992; Waterhous & Johnson, 1994). Several laboratories have investigated the helical propensity of amino acids using different peptide and protein systems (Lyu et al., 1990; Merutka et al., 1990; O'Neil & DeGrado, 1990; Padmanabhan et al., 1990; Chakrabartty et al., 1991; Gans et al., 1991; Kemp et al., 1991; Scholtz et al., 1991; Rohl et al., 1992; Stellwagen et al., 1992; Bell et al., 1992; Serrano et al.,

1992; Blaber et al., 1993). The result from these works are in disagreement. Instead of searching for the best set of helical propensities under one selected condition, this work determines the helical propensity of the 20 amino acids as a function of the changing solvent environment. We find that the relative order of propensity for α-helix formation among amino acids changes in various concentration of methanol. The finding suggests an important role of environment to influence the formation of secondary structure. We believe that the 70% limitation of sequence-based predicting will be improved by including the environment factor. Agreement between the relative order of propensity in 63% methanol and the propensity derived from the X-ray database suggests that the 63% methanol condition is similar to an average environment found in real proteins.

Circular Dichroism Spectroscopy

Circular dichroism (CD) is a widely-used technique for investigating the secondary structure of peptides and proteins. Left and right circularly polarized light is absorbed by a sample molecule in solution, and the difference in absorption by both types of light is recorded (A_L-A_R). The light used is normally in the ultraviolet (UV) region, corresponding to the transitions of electrons-in the case of proteins, these are in the peptide bonds. CD spectra are expressed through Beer's Law as the difference between ε_L and ε_R .

$$\Delta A = A_L - A_R$$

$$= \varepsilon_R lc - \varepsilon_R lc$$

$$= (\varepsilon_L - \varepsilon_R) lc$$

$$= \Delta\varepsilon \quad \text{(Where } l = 1 \text{ cm, and } c = 1 \text{ Molar)}$$

A routine measurement for peptides and proteins is in the range from 300 to 180 nm. The near-UV CD around 280 nm provides an information on the flexibility and surrounding environment for amino acids with aromatic side-chains. This information is rather limited, and generally used to monitor the folding-unfolding stages of proteins. On the other hand, the far-UV CD from 260 to 180 nm can provide more information on protein secondary structure. Researchers can analyze the CD in this region for the content of each secondary structures within protein. There are many computing algorithms available for this type of analysis, and their results are recognized to be in good agreement the structure from the x-ray database (Hennessey & Johnson, 1981; Provencher & Glockner, 1981; Compton & Johnson, 1986; Yang et al.,1986; Manavalan & Johnson, 1987; Johnson, 1990; Perczel et al., 1991; Bohm et al., 1992; Sreerama & Woody, 1993, 1994). The reason for the popularity of using CD to study molecular conformation is simple. The technique takes a small amount of sample in solution, and the measurement can be performed rapidly. This work makes use of CD spectroscopy as a tool to monitor a conformational change, and to quantify the content of secondary structure.

Among the several types of protein secondary structure, the α-helix is the most well-defined element that accounts for almost one-third of the residues in

globular proteins (Barlow & Thornton, 1988). The structure is a right-handed helix with 3.6 residue per turn, stabilized mainly by the hydrogen bonding between the carbonyl oxygen of residue *i* to the amide hydrogen of residue *i* + 4. The CD spectrum of an α-helix has its own characteristics with an intense positive band at 190 nm and two negative bands of moderate size around 208 and 222 nm. The 222 nm band is due to the n-π* transition (Schellman & Oriel, 1962; Woody & Tinoco, 1967), while the negative 208 and positive 192 nm bands are the result of exciton splitting of the π-π* transition (Moffitt, 1956; Tinoco et al., 1963; Woody & Tinoco, 1967; Mandel & Holzwarth, 1972).

The estimation of helical content for peptides can be performed rapidly using CD intensity of the 222 nm band. This CD band is present only for the helix. The intensity of the 222 nm band -from 0 to -10$\Delta\varepsilon$ (M^{-1} cm^{-1})- is generally used to describe the helical content from 0 to 100%. However, some corrections may have to be included to deal with the effect from chain length for short peptides (Chen et al., 1974; Manning & Woody, 1991). The characteristic CD for random coil conformation of peptides is simple. It contains a single negative band just below 200 nm. But some proteins may also give a small band within the 210 to 230 nm range.

Correcting CD Spectra of Peptides for Absorbing side-chains

The far-UV CD in a 180-260 nm region is primarily due to the transition of the amide bond. Sometimes a contribution from absorbing side-chains of an amino acid is found to be present also in the far-UV. There are reports for such

contributions in both proteins and peptides (Manning & Woody, 1989; Vuilleumier et al., 1993; Chakrabartty et al., 1993; Freskgard et al., 1994; Boren et al., 1996). By comparing the normal absorption spectra for the 20 amino acids in the far-UV, it is possible to identify the amino acid side-chains that contribute to the CD spectra. We find that most of amino acids have only one intense band around 190 nm, while the aromatic, (tryptophan, tyrosine, phenylalanine) and the sulfur-containing (cysteine, methionine) amino acids give a additional band in the 200 to 230 nm region (Wetlaufer, 1962)).

We have proposed a method of correcting the CD spectra of peptide for those absorbing side-chains using Singular Value Decomposition (SVD) (Forsythe et al., 1977; Noble & Daniel, 1977; Johnson, 1992). The method is based on using SVD to extract common basis vectors from related peptides without absorbing side-chains, and fit them to a part of the CD of peptide with absorbing side-chain, in the region that is free from side-chain contributions. The fitting coefficients will enable us to reconstruct a corrected CD spectrum from the common basis vectors. The method is tested by using a set of CD spectra obtained from 19 related peptides in helix and random coil conformation. The result demonstrates an identical CD spectra for those peptide without absorbing side-chain, but gives a difference in CD for peptides containing tryptophan, tyrosine, phenylalanine, cysteine and methionine. The difference CD of those five peptides indicate a positive CD band contribution from the aromatic side-chain amino acid, and the negative CD band from the sulfur-containing ones.

CHAPTER II

CORRECTING THE CD SPECTRA OF PEPTIDES FOR ABSORBING SIDE-CHAINS

Chartchai Krittanai and W. Curtis Johnson, Jr.

Department of Biochemistry and Biophysics,

Oregon State University, Corvallis, OR 97331-7305

Submitted for publication

February 1997

Abstract

The aromatic and sulfur-containing side-chains Trp, Tyr, Phe, Cys and Met contribute to the CD spectra of peptides and proteins in the amide region, interfering with the analysis for secondary structure. We propose a method to correct the CD spectra of peptides due to absorbing side-chains using Singular Value Decomposition. The method uses the common basis vectors obtained from an analysis of the CD spectra of related peptides without the aromatic and sulfur-containing amino acids. The common basis vectors are fitted to a portion of the CD spectrum of the peptide being corrected, in the range that is not affected by its side-chain contributions. Then the resulting coefficients from the fitting are used along with the common basis vectors to regenerated the entire corrected spectrum. The method is illustrated for the CD spectra of the peptide sequence acetyl-Y-VAXAK-VAXAK-VAXAK-amide, where X is substituted with the natural occurring 20 amino acids. This peptide model adopts a random coil conformation in 2 mM sodium phosphate buffer pH 5.5, and becomes an alpha helix in methanol/buffer solutions. The difference between the original and corrected spectra indicates a fair contribution from the aromatic and sulfur-containing side-chains.

Introduction

Circular dichroism (CD) spectroscopy has become the method of choice for monitoring the secondary structure of peptides and proteins as a function of

solvent. Researchers measure the CD spectra and analyze for secondary structure using several available computer programs (Hennessey & Johnson, 1981; Provencher & Glockner, 1981; Compton & Johnson, 1986; Yang etal.,1986; Manavalan & Johnson, 1987; Perczel et al., 1991; Bohm et al., 1992; Sreerama & Woody, 1993, 1994). The analyses require the measurement of a CD spectrum into the far-UV, where the CD is dominated by the amide chromophores. However, it is known that some amino acid side-chains absorb light in the far-UV region, causing interference with the amide CD (Woody, 1978; Manning & Woody, 1989; Vuilleumier et al., 1993). The CD studies of several proteins in the far-UV have demonstrated the unusual CD spectra due to aromatic side-chains (Beychok et al., 1966; Green & Melamed, 1966; Yang et al., 1968; Simon & Blout, 1968; Baba et al., 1969; Timasheff & Nernardi, 1970; Maeda et al., 1973; Manez et al., 1976; Cameron & Tu, 1976; Chen et al., 1977). In most cases, the tryptophan and tyrosine side-chains are found to be the source of interference. Some workers have carried out theoretical studies and calculated the rotational strength resulting from the interaction of these side-chain with the electronic transitions of the peptide bond (Hooker & Schellman, 1970; Bush & Gibbs, 1972; Goux & Hooker, 1975; Snow et al., 1977; Woody, 1978)

The contribution from amino acid side-chains becomes a problem when a CD spectrum is interpreted and analyzed for secondary structure. Several groups have investigated the CD spectra of peptides and proteins containing different types and numbers of aromatic side-chains (Chakrabartty et al., 1993;

Freskgard et al., 1994; Boren et al., 1996). The CD spectra for proteins or peptides of similar secondary structure but with different numbers of aromatic side-chains are found to be different in the far-UV. The estimate of structural content of these proteins by several algorithms yields information that is unreliable (Boren et al., 1996). The error in structural estimates exists for both methods using a wavelength range and those that rely on a single wavelength. Recently, the contribution from aromatic side-chains was investigated in a series of short peptides (Chakrabartty et al., 1993). The experiments showed that the contribution from an aromatic tag located at the peptide terminal is between 207 and 240 nm, and that the contribution can be eliminated by inserting several glycines between the aromatic tag and the peptide.

Here we investigate the contribution of absorbing side-chains to the CD of a different peptide model. We synthesized the model acetyl-Y-VAXAK-VAXAK-VAXAK-amide, with one of the 20 amino acids substituted at the X position. We measured their CD spectra as a function of solvent beginning with the random coil in buffer and ending with the α-helix in methanol. Our peptides contain a tyrosine tag at the N-terminal which obviously contributes to the CD spectra. However we concentrate on the contribution arising from the three substitution sites, assuming that the N-terminal tyrosine should have the same effect on all peptides with the same amount of helical structure. The contribution from the absorbing side-chains is observed not only for aromatic side-chains, but also for the sulfur-containing side-chains. We propose a method to correct the spectra for the side-chain contribution using the Singular Value Decomposition (SVD)

theorem. The method derives the common basis vectors for the amides only from the SVD of the CD spectra for related peptides without absorbing side-chains. The common basis vectors are fitted to a portion of the CD spectra for the peptides with absorbing side-chains in the wavelength region that is free from the side-chain contribution, yielding a set of fitting coefficients. The fitting coefficients are then used along with the common basis vectors to generate the corrected spectra for the entire range.

Materials and Methods

Peptide Synthesis

The peptide models Ac-Y-VAXAK-VAXAK-VAXAK-amide, where X is substituted with the 20 natural occurring amino acids, were synthesized manually using a modified solid phase method (Atherton & Sheppard, 1989). The synthesis strategy employed an active ester coupling of 9-fluorenylmethoxycarbonyl (FMOC) amino acids (Bachem California Inc.) to a Rink resin (4-[2',4'-dimethoxyphenyl-FMOC-aminomethyl]-phenoxy on 1% cross-linked divinylbezene-styrene) inside a mesh bag. The amino protecting group, FMOC, was removed using 30% piperidine/N, N-dimethylformamide and the coupling of the next FMOC-amino acid was performed. The deprotecting-coupling cycle was repeated in a polyethylene bottle on a shaker until reaching the desired length of the peptide. The N-terminal of a newly synthesized peptide was acetylated with acetic anhydride before the peptide was cleaved from the

solid resin. Cleaving with 0.5 % trifluoroacetic acid/H_2O yielded a peptide with C-terminal amide. Crude peptides were purified by reverse-phase high performance liquid chromatography using a Vydac C-18 reverse-phase semipreparative column. The hydrophobic gradient was 0.1% trifluoroacetic acid/water and 0.1% trifluoroacetic acid/acetonitrile. The fractions were collected according to the absorption of the amide bond at 214 nm and a tyrosine tag at 275 nm. The molecular weight for all peptides was confirmed by Fast Atom Bombardment or Electrospray mass spectroscopy. The purity of peptides used in the research was above 95% .

Circular Dichroism Measurements

The CD measurements were performed at 25 °C using a Jasco J-720 spectropolarimeter (Jasco Inc.). The instrument was purged with nitrogen gas at 20 L/min for 20 minutes before and during measurements. A two-point calibration with (+)-10-camphorsulfonic acid (CSA) gave a ratio of 2.0 to 2.1 between the magnitude of the 192.5 and 290.5 nm bands. Methanol/buffer solvent was prepared by mixing 20 µL of 100 mM sodium phosphate buffer pH 5.5, spectroscopic grade methanol, and 18-MΩ distilled water together, yielding 1 ml of solvent containing 2 mM sodium phosphate buffer pH 5.5. A series of these 1 ml methanol/buffer solutions were prepared fresh from 0 to 95 % methanol with a 5% increment by varying the amount of methanol and water. These mixed solutions were kept in closed microcentrifuge tubes to avoid evaporation. The top of each tube was pierced with a syringe needle and then

the stock peptide in water was gradually added into the tubes on a vortexer. The sample was then transferred to a 1-mm rectangular quartz cell (from Hellma Cells, Inc.), waiting for 5 minutes to reach equilibrium before making a measurement. The spectra were recorded using a 2 nm slit-width, 20 nm/min scan speed, two millisecond response time, and three accumulations. Because of peptide solubility and light absorption of methanol, a sample concentration of 1.0 to 2.0 mM amide and a wavelength range of 260 to 195 nm were chosen to minimize the noise and error during the measurement. Each spectrum was corrected by a baseline measured with the same solvent in the same cell. The intensity of the CD spectra are expressed as $\Delta\varepsilon$ per amide bond.

Ultraviolet Absorption Measurements

The UV absorption measurements were performed at room temperature using a Cary-15 spectrophotometer interfaced to a computer. The machine was flushed with nitrogen gas for 20 minutes before and during the measurement. The absorption spectra for the amides were recorded from 340 to 185 nm in a 50-μm cylindrical quartz cell. The absorption spectra for the aromatics were recorded from 340 to 240 nm, using a 1-cm rectangular quartz cell.

Amide Concentration Measurements

Amide concentration can be calculated from its absorption at 190 nm using the extinction coefficient of the amide bonds, ε_{190}. We first measured the

UV absorption of random coil peptides in water at 280 nm using 1.0 cm cells. The absorption at 280 nm and the known ε_{280} of the tyrosine tag (1,280 M^{-1} cm^{-1}) were used to calculate the molar concentration of peptide. The peptide concentration was converted to the amide bond concentration by multiplying by the number of peptide bonds. The absorption at 190 nm was measured for the stock peptide as a random coil in water using a 50-μm cylindrical cell, and the amide concentration was used to calculate the ε_{190} value (9,000 to 13,000 M^{-1} cm^{-1}) for each peptide. These ε_{190} were used to quantify the peptides in the same cell used for the CD measurement.

Results and Discussion

CD Measurement on the Peptides

The CD spectra of the 20 peptides in a series of solvents consisting of methanol and 2 mM phosphate buffer pH 5.5 show a transition between a random coil and an alpha helix. All peptides (except the one with proline substitution which is not included in the analysis) adopted the random coil conformation in buffer and became more helical when the methanol content was increased, as shown for the alanine substituted peptide in Figure 2.1. The CD spectra show a negative band about 198 nm for the random coil, two negative bands at 222 and 208 nm and an intense positive band about 193 nm for the

helix. The isosbestic point is located about 203 nm, indicating two dominant states, the random coil and alpha helix.

Singular Value Decomposition Analysis

Singular Value Decomposition (SVD) is a mathematical theorem that can be applied to spectral analysis (Forsythe et al., 1977; Noble & Daniel, 1977; Henry & Hofrichter, 1992; Johnson, 1992). The theorem states that a data set written as a matrix, A, can be decomposed into three matrices, **U**, **S** and $\mathbf{V}^T$.

$$\mathbf{A} = \mathbf{USV}^T$$

If the spectra are column vectors in the matrix **A**, the **U** matrix consists of orthonormal basis vectors for the data space. **S** is a matrix containing a set of numbers called singular values, which weigh the significance of each basis vector in matrix **U**. $\mathbf{V}^T$ contains the orthonormal coefficients that fit matrix **US** to the data matrix **A**.

We analyzed our data matrix **A**, which contains all the CD spectra of 19 peptides in all different methanol/buffer solutions (0 to 95% with a 5% increment), for the matrices **U**, **S** and $\mathbf{V}^T$. The analysis gave us two large singular values in the **S** matrix. This means that there are only two significant basis vectors in the data space, supporting the observation of two dominant species in the transition for all peptides.

Contribution from Absorbing Side-Chains

Researchers often observe the CD bands arising from absorbing amino acid side-chains around 280 nm in the near UV region. These CD bands are usually recognized as an aromatic band resulting from Trp, Tyr or Phe. These absorbing side-chains also contribute to the CD in the far UV region as has been reported by several laboratories (Woody, 1978; Manning & Woody, 1989; Vuilleumier et al., 1993; Chakrabartty et al., 1993; Freskgard et al., 1994; Boren et al., 1996).

Here our CD spectra (the dashed lines in Figures 2.2-2.6) show the contribution from aromatic and sulfur-containing side-chains between 210 to 240 nm range, adding a distortion to the typical amide CD spectra for helix and coil. We confirm that the contribution is real by analyzing the CD spectra of the 19 peptides using SVD (the proline substituted peptide is excluded), and then regenerating the spectra using only two significant basis vectors from the analysis. The regenerated spectra fit their original spectra very well for 14 peptides, but poorly for the ones that have Phe, Trp, Tyr, Cys and Met substitutions (not shown). The original and the regenerated spectra for those peptides with aromatic and sulfur side-chains are different between 210 and 240 nm, but are the same from 195 to 210 nm.

Supporting evidence for the existence of the CD band for aromatic and sulfur side-chains comes from the UV absorption spectra of the peptides. In general, a CD band is found corresponding to the absorption band of each

electronic transition. The UV absorption spectra of amino acids, Phe, Tyr, Trp, Cys, and Met have an absorption band located in the 210 to 240 nm range.

Correcting CD Spectra for Absorbing Side-Chains

Our primary goal of correcting the spectra is to eliminate the bands that belong to the amino acid side-chains. The spectra in our data set are recorded for each peptide along the transition from coil to helix. We assume that all the spectra, including ones with the side-chain contributions, share the common basis vectors for the amides undergoing the structural transition. If the correct fitting coefficients are known, one can reconstruct any of our spectra from those common basis vectors. To obtain the common basis vectors, SVD is performed on the spectra measured for peptides with no side-chain absorption. The exclusion of spectra measured from peptides with absorbing side-chains guarantees that the resulting common basis vectors are due to amides and free from side-chain contributions. After the analysis, SVD gives us an **S** matrix with two large singular values indicating only two important common basis vectors. These two basis vectors can be used to regenerate the amide spectra for all peptides, regardless of their side-chains.

The next problem is to find the set of coefficients that fit the common basis vectors to the CD spectra of the peptides with absorbing side-chains. We obtain these coefficients by fitting the two common basis vectors to the CD spectra in the range that is free from side-chain contributions. Chakrabartty and coworkers (1993) demonstrated that the interfering CD band from aromatic side-

chains is located around 208 to 250 nm. Here we chose the 195 to 207 nm range for the fitting, which appears to be a good range for sulfur-containing side-chains as well. The fitting coefficients obtained for each peptide with absorbing side-chains are then used with the two common basis vectors to regenerate full corrected spectra from 195 to 260 nm, seen as the solid lines in Figures 2.2-2.6.

To verify the procedure, we obtain the coefficients for the 14 peptides without absorbing side-chains over the 195 to 207 nm range. The regenerated CD spectra for these 14 peptides are identical with the original spectra over the entire wavelength range. In contrast, the regenerated spectra for five peptides with Phe, Tyr, Trp, Cys and Met substitution are different from the original spectra, as expected (Figures 2.2-2.6). A good fit between the regenerated and original spectra is found in the 195-207 nm region, with the difference showing up from 207 to 240 nm.

Using 14 peptides to get the common basis vectors may not be practical. We repeated the correction using common basis vectors deriving from one peptide. The peptides with an aliphatic amino acid substitution are expected to be a good choice. We find the common basis vectors for an aliphatic amino acid are almost identical to those derived from the 14 peptides.

When the corrected CD spectra are subtracted from the original spectra, the difference spectra indicate CD bands arising from each absorbing side-chain. We find a positive band around 220, 224, and 226 nm, representing the CD for Phe, Trp and Tyr side-chains respectively, as shown in Figure 2.7. The

difference CD found for Cys and Met side-chains show a negative CD band at about 225 and 230 nm, respectively (Figure 2.8).

The difference CD spectra change somewhat with methanol concentration. In all cases, the side-chain contribution for random coil is found to be smaller than for the helical structure. This phenomenon may be due to a higher constraint of the side-chains in the helix. The intensity of the CD spectra in our experiments is expressed per amide chromophore. The difference spectra show the average CD intensity per absorbing side-chain at the three substitution sites. We also show the total side-chain contribution to the CD of the peptide on a per amide basis. The calculation indicates a single side-chain contribution of about 5 %, or up to 15 to 20 % for three side-chains per peptide.

The use of this method to correct the CD spectra of proteins might be feasible, but is far more complicated. Unlike peptides, most proteins do not have a single secondary structure. Sometimes two proteins with about the same type and content of secondary structures have different CD spectra. This makes finding the common basis vectors far more difficult. The method presented here is expected to be useful to correct the CD of most peptides.

Conclusions

The aromatic and sulfur-containing side-chains show a fair contribution to the CD spectra of peptides. Their contributions are found not only around 280 nm, but also in the amide region. In most cases, a CD spectrum below 260 nm is governed by the amide chromophore, and the side-chain contribution is

negligible. When the side-chain contribution is more profound, it adds an error to the estimates of peptide secondary structure.

We suggest a method to correct the CD spectra for side-chain contributions, using Singular Value Decomposition. The method analyzes the CD spectra of peptides without absorbing amino acid side-chains for the common basis vectors, fits the common basis vectors to a portion of the CD spectrum measured for peptides with absorbing amino acid side-chains, and then uses the fitting coefficients and the common basis vectors to generate the entire corrected spectrum.

Difference CD spectra show a positive interfering band around 224 nm for the aromatic side-chains, and a negative band around 228 nm for the sulfur-containing side-chains. When the side-chain contribution to the CD is calculated, their interference is as much as 20% for our peptides.

Figure 2.1: The typical CD spectra of a Y-(VAXAK)$_3$ peptide measured in various methanol/buffer concentrations.

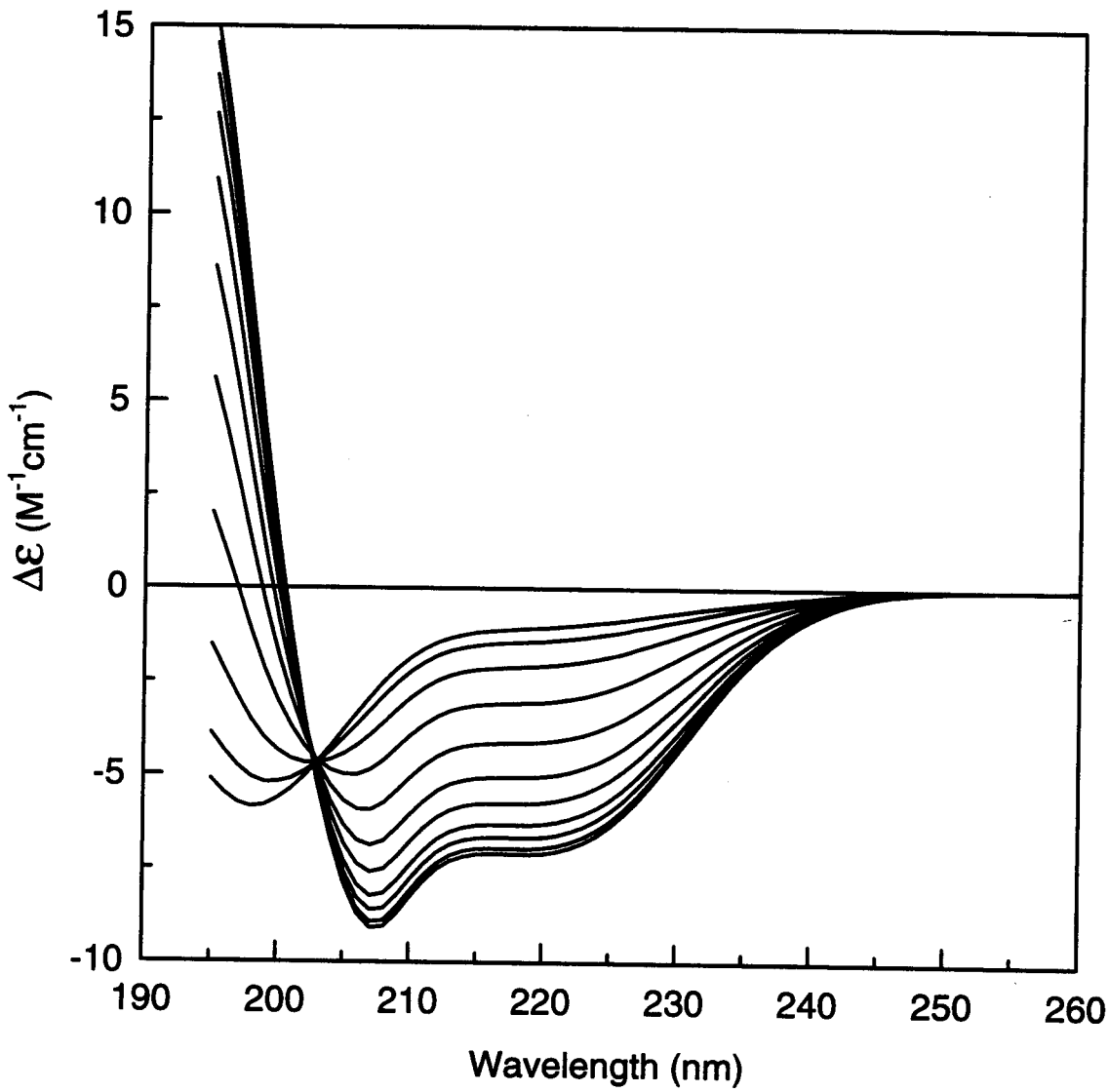


Figure 2.2: The corrected (——) and original (------) CD spectra for Y-$(VAWAK)_3$.

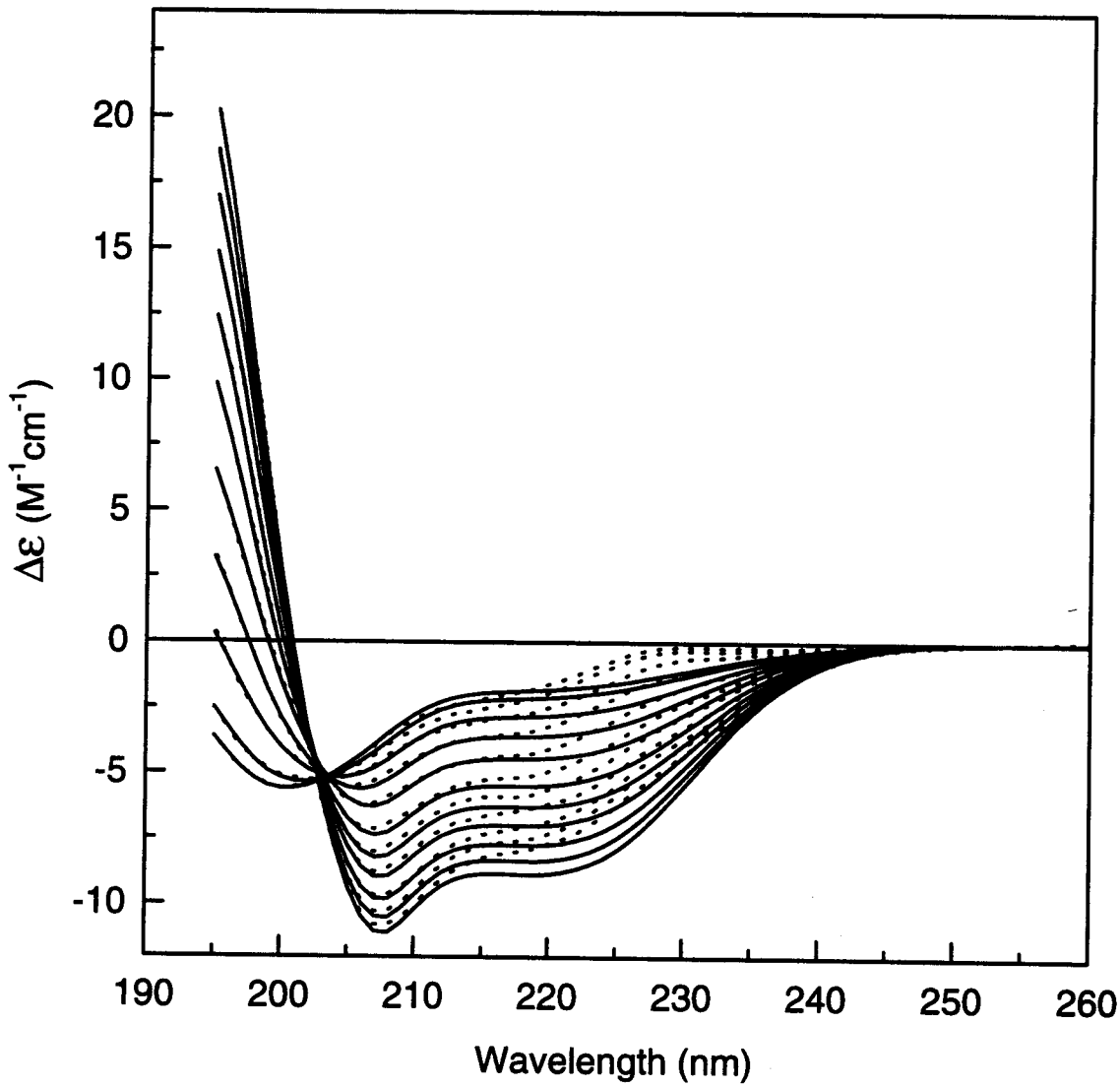


Figure 2.3: The corrected (——) and original (------) CD spectra for Y-$(VAFAK)_3$.

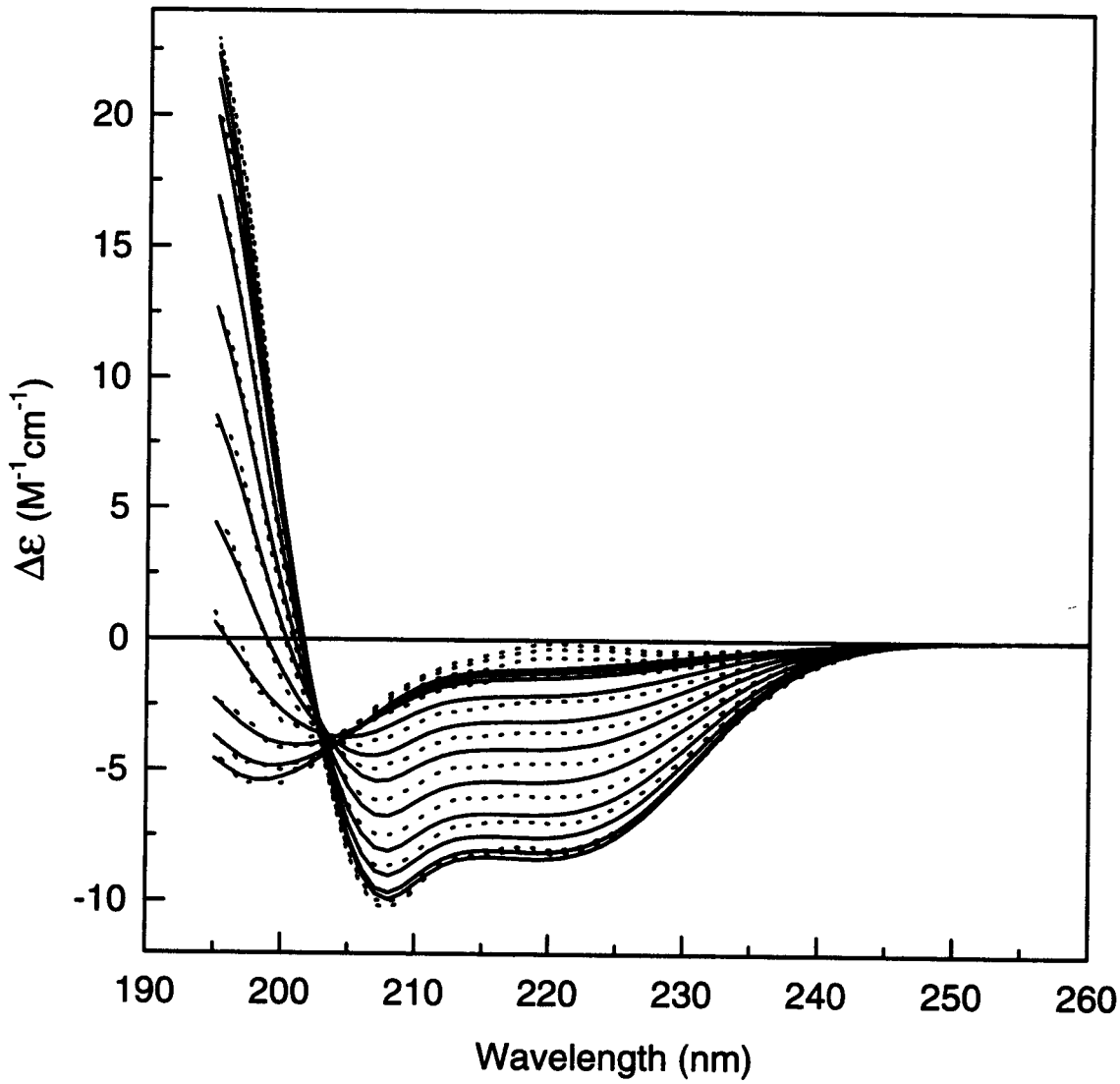


Figure 2.4: The corrected (——) and original (------) CD spectra for Y-$(VAYAK)_3$.

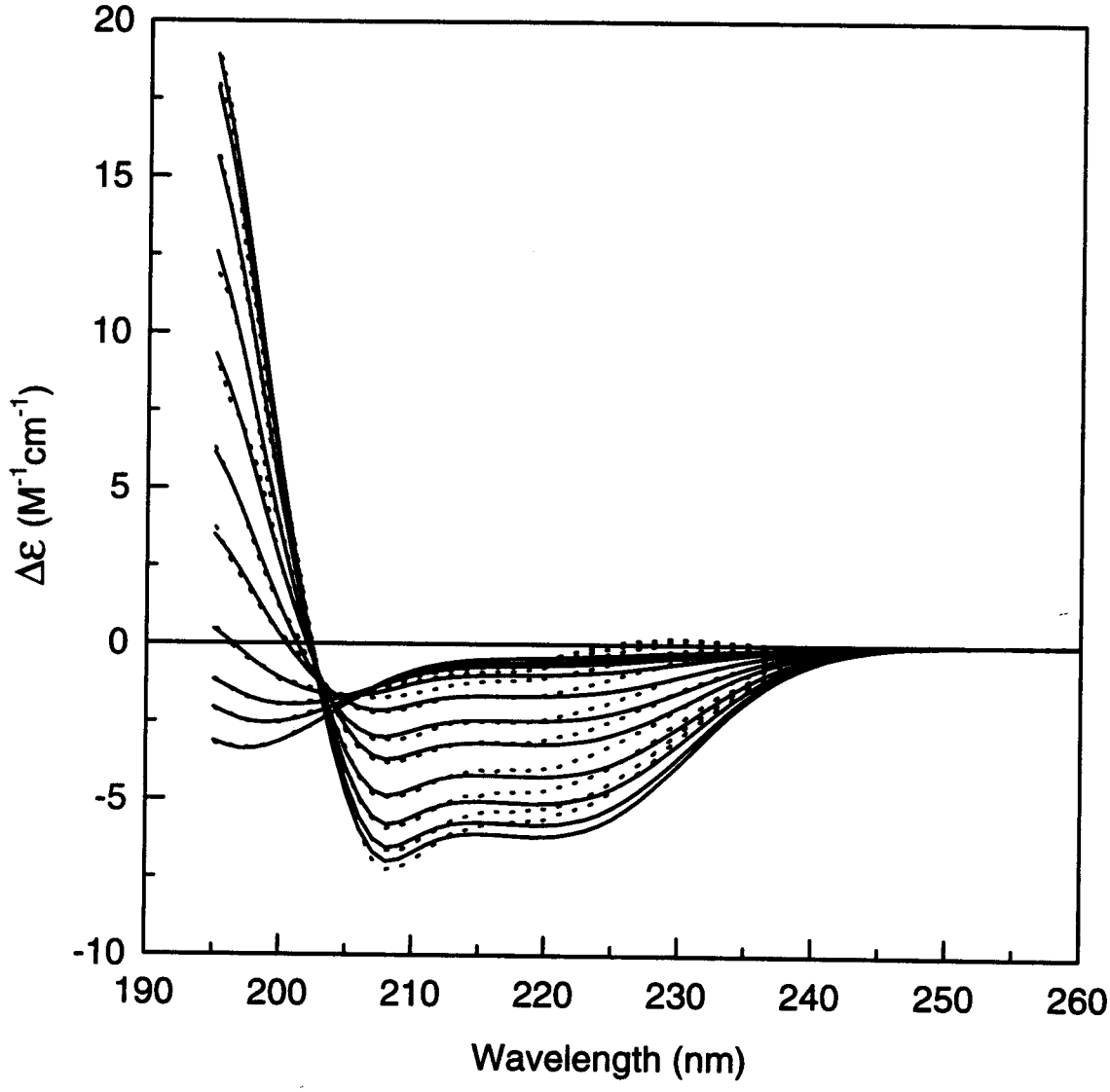


Figure 2.5: The corrected (——) and original (------) CD spectra for Y-$(VAMAK)_3$.

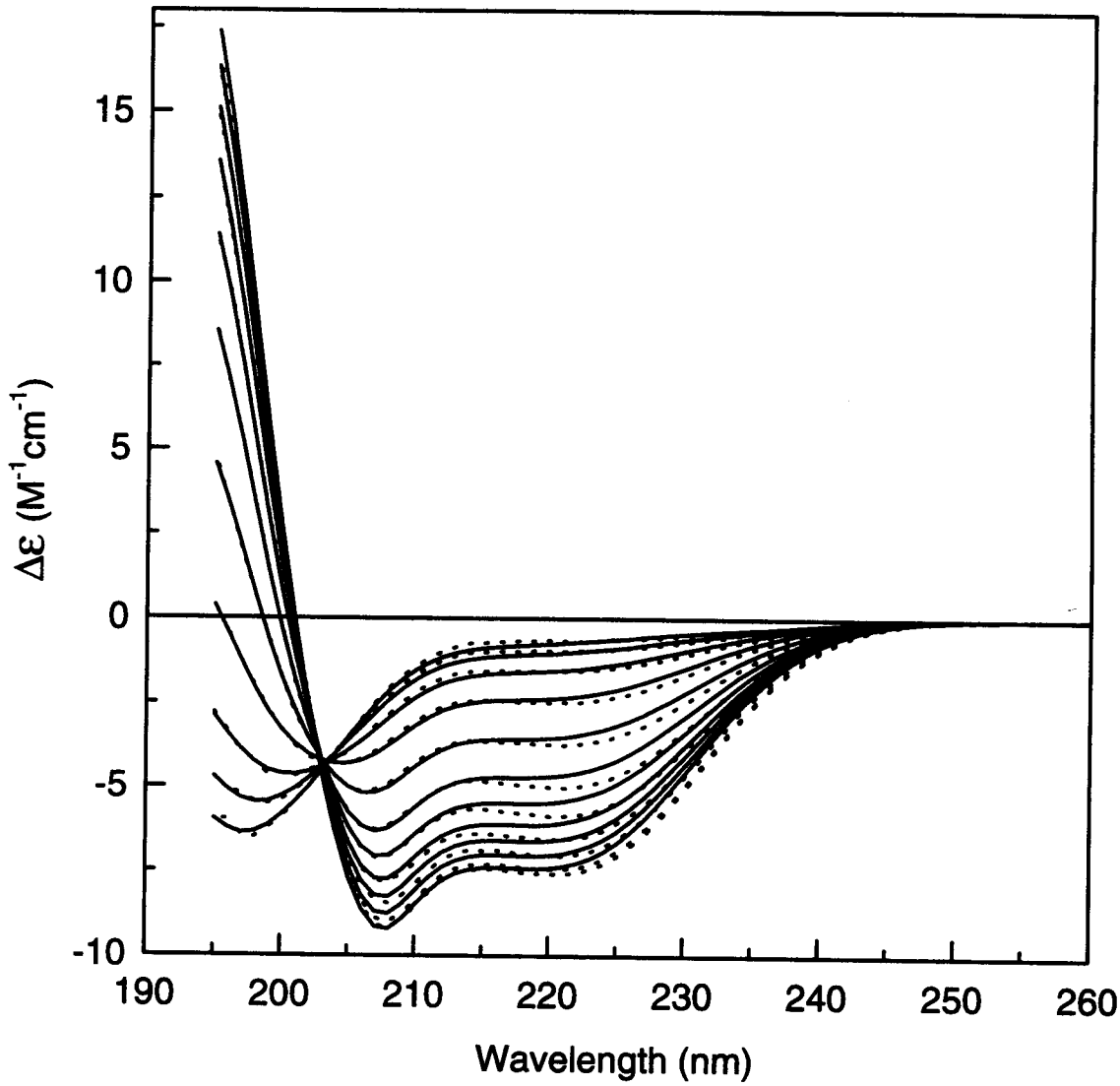


Figure 2.6: The corrected (——) and original (------) CD spectra for Y-$(VACAK)_3$.

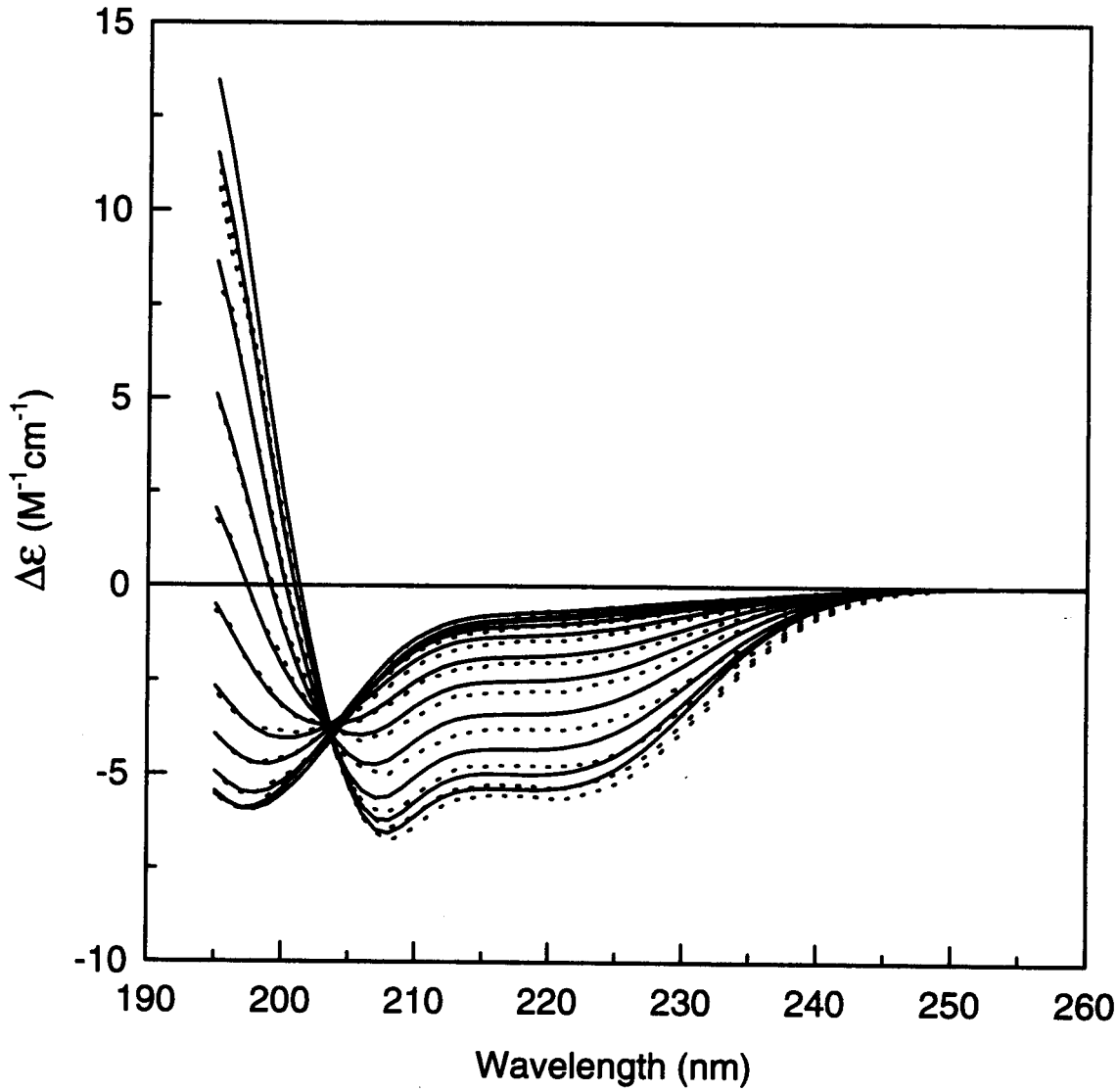


Figure 2.7: The difference CD spectra between corrected and original Y-$(VAXAK)_3$ where X = Trp (a), Phe (b), and Tyr (c). The largest contribution is observed for the α-helix and the smallest contribution for the random coil.

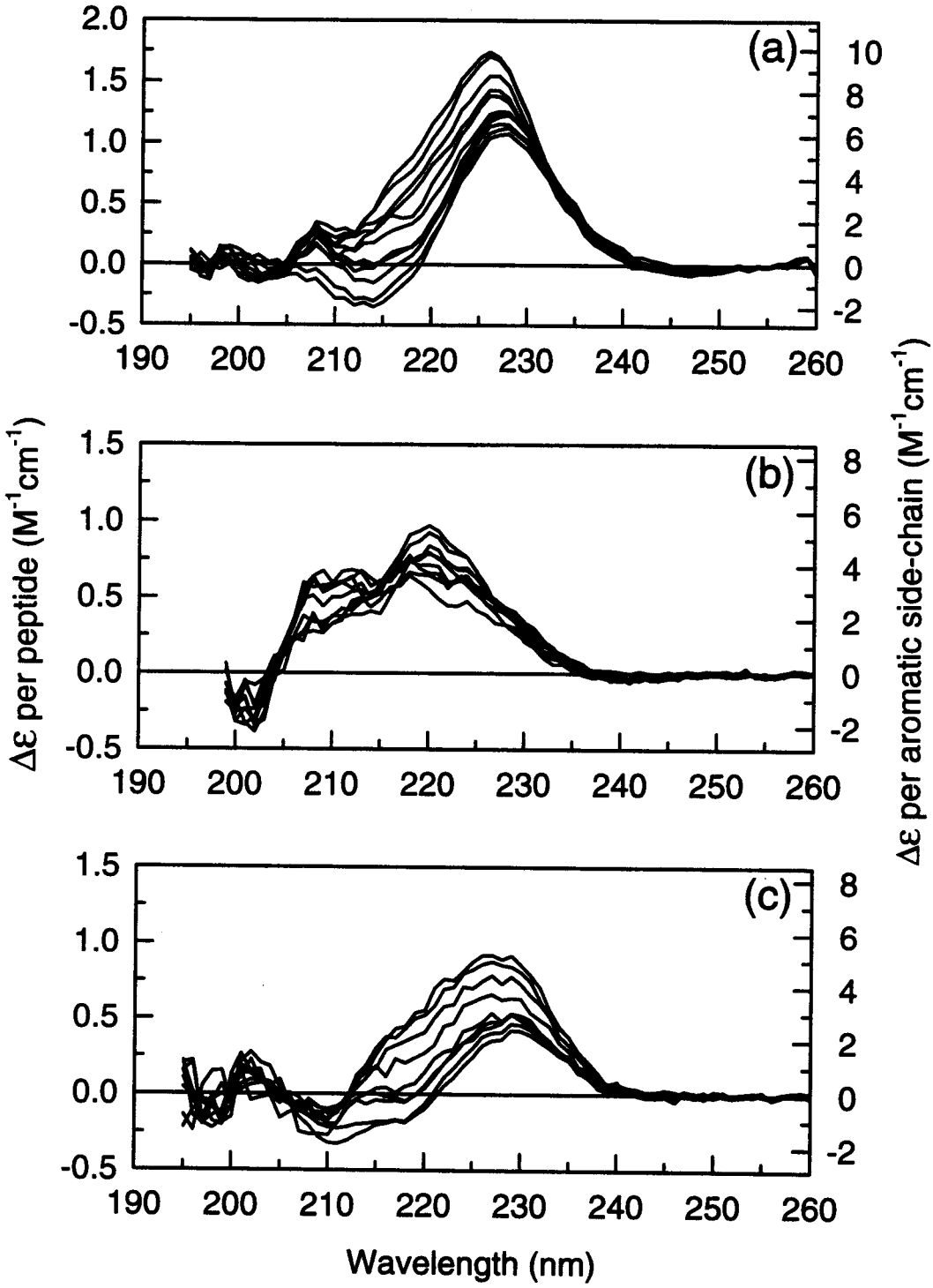
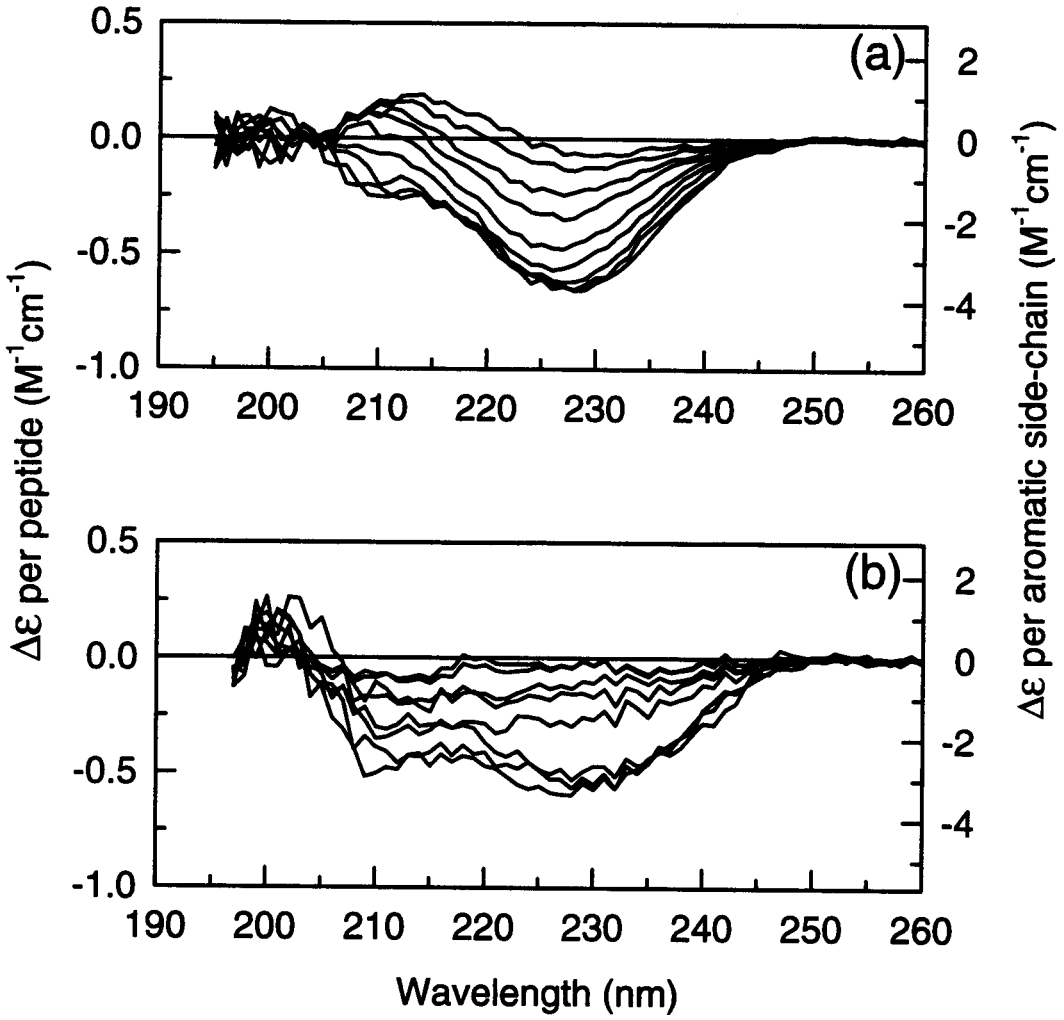


Figure 2.8: The difference CD spectra between corrected and original Y-$(VAXAK)_3$ where X = Met (a) and Cys (b). The largest contribution is observed for the α-helix and the smallest contribution for the random coil



Acknowledgments

This work was supported by a grant from the National Institutes of Health (GM 21479) and the Development and Promotion of Science Talents Project (DPST) by the Royal Thai Government.

References

Atherton, E., & Sheppard, R. C. (1989) in *Solid Phase Peptide Synthesis,* IRL Press at Oxford University Press, Oxford, UK.

Baba, M., Hamaguchi, K., & Ikenaka, T. (1969) *J. Biochem. 65*, 113-121.

Beychok, S., Armstrong, J. M., Lindblow, C., & Edsall, J. T. (1966) J. Biol. Chem. 241, 5150-5160.

Bohm, G., Muhr, R., & Jaenicke, R. (1992) *Protein Eng. 5*, 191-195.

Boren, K., Freskgard, P. O., & Carlsson, U. (1996) *Prot. Sci. 5*, 2479-2484.

Bush, C. A., & Gibbs, D. E. (1972) *Biochemistry 11*, 2421-2427.

Cameron, D. L., & Tu, A. T. (1977) *Biochemistry 16*, 2546-2553.

Chakrabartty, A., Kortemme, T., Padmanabhan, S., & Baldwin, R. L.(1993) *Biochemistry 32*, 5560-5565.

Chen, Y. H., Lo, T., B., & Yang, J. T. (1977) *Biochemistry 16*, 1826-1830.

Compton, L. A., & Johnson, W. C., Jr. (1986) *Anal. Biochem. 155*, 155-167.

Forsythe, G. E., Malcolm, M. A., & Moler, C. B. (1977) in *Computer Methods for Mathematical Computations.* pp. 192-236, Prentice-Hall Inc., Englewood Cliffs, NJ.

Freskgard, P. O., Martensson, L. G., Jonasson, P., Jonsson, B. H., & Carlsson, U. (1994) *Biochemistry 33*, 14281-14288.

Goux, W. J., & Hooker, T. M., Jr. (1975) *J. Am. Chem. Soc. 97*, 1605-1606.

Green, N. M., & Melamed, M. D. (1966) *Biochem. J. 100*, 614-621.

Hennessey, J. P., Jr., & Johnson, W. C., Jr. (1981) *Biochemistry 20*, 1085-1094.

Henry, E. R., & Hofrichter, J. (1992) in *Methods in Enzymology* (Brand, L., & Johnson, M. L., Eds.) pp. 129-192, Academic Press, San Diego, CA.

Hooker, T. M. Jr., & Schellman, J. A. (1970) *Biopolymers 9*, 1319-1348.

Johnson, W. C., Jr. (1992) in *Methods in Enzymology* (Brand, L., & Johnson, M. L., Eds.) pp. 426-447, Academic Press, San Diego, CA.

Maeda, H., Shiraishi, H., Onodera, S., & Ishida, N. (1973) *Int. J. Peptide Protein Res. 5*, 19-26.

Manavalan, P., & Johnson, W. C., Jr. (1987) *Anal. Biochem. 167*, 76-85.

Manez, A., Bouet, F., Tamiya, N., & Fromageot, P. (1976) *Biochim. Biophys. Acta 453*, 121-132.

Manning, M. C., & Woody, R. W. (1989) *Biochemistry 28*, 8609-8613.

Noble, B. & Daniel, J. W. (1977) in *Applied Linear Algebra*. pp. 323-342, Prentice-Hall Inc., Englewood Cliffs, NJ.

Perczel, A., Hollosi, M., Tusnady, G., & Fasman, G. D. (1991) *Protein Eng. 4*, 669-679.

Provencher, S. W., & Glockner, J., (1981) *Biochemistry 20*, 33-37.

Simon, E. R. & Blout, E. R. (1968) *J. Biol. Chem. 243*, 218-221.

Snow, J. W., Hooker, T. M. Jr., & Schellman, J. A. (1977) *Biopolymers 16*, 121-142.

Sreerama, N., & Woody, R. W. (1993) *Anal. Biochem. 209*, 32-44.

Sreerama, N., & Woody, R. W. (1994) *J. Mol. Biol. 242*, 497-507.

Timasheff, S. N., & Nernardi, G. (1970) *Arch. Biochem. Biophys. 141*, 53-58.

Vuilleumier, S., Sancho, J., Loewenthal, R., & Fersht, A. R. (1993) *Biochemistry 32*, 10303-10313.

Woody, R. W. (1978) *Biopolymers, 17*, 1451-1467.

Yang, C. C., Chang, C. C., Hayashi, K., Suzuki, T., Ikeda, K., & Hamaguchi, K. (1968) *Biochim. Biophys. Acta 168*, 373-376.

Yang, J. T., Wu, C. C., & Martinez, H. M. (1986) *Methods Enzymology. 130*, 208-269.

CHAPTER III

RELATIVE ORDER OF HELICAL PROPENSITY OF AMINO ACIDS CHANGES WITH SOLVENT ENVIRONMENT

Chartchai Krittanai and W. Curtis Johnson, Jr. *

Department of Biochemistry and Biophysics,

Oregon State University, Corvallis, Oregon 97331

Submitted for publication

May 1997

Abstract

A peptide model of sequence Ac-Y-VA**X**AK-VA**X**AK-VA**X**AK-amide, where **X** is substituted with one of the 20 amino acids, was synthesized and titrated with methanol to study helix propensity as a function of environment. The CD spectra of these peptides demonstrate the conformational change from a random coil in 2 mM sodium phosphate buffer (pH 5.5) to an α-helix with the increase in methanol content. The CD spectra were corrected for the absorbing side-chains of W, Y, F, C, M, and then analyzed using singular value decomposition. The spectral analysis confirms a two-state transition between random coil and helix conformation. The free energy of helix formation, representing the helical propensity for each amino acid, was calculated from the amount of helix and coil, and plotted as a function of methanol concentration. The relative order of helical propensity among the amino acids is found to be dependent on solvent environment. A comparison of the free energy in buffer at 63% methanol with the helical preference derived from the X-ray database is in a good agreement, indicating that 63% methanol represents the average environment found in proteins.

Introduction

Since Anfinsen's refolding experiment (Anfinsen *et al.*, 1961) on ribonuclease in the early 1960's, the amino acid sequence has been known to play a crucial role in determining the ultimate three-dimensional structure of

proteins. A number of studies have focused on how to predict the structure of a protein using information extracted from the amino acid sequence. There are many approaches developed for such a prediction over the last two decades based on sequence homology (Pongor & Szaley, 1985; Sweet, 1986; Nishikawa & Ooi, 1986; Levin, 1986; Zvelebil et al, 1987), linear optimization of predictors (Edelman & White, 1989), neural networks (Qian & Sejnowski, 1988; Holley & Karplus, 1989, Stolorz et al., 1992; Rost and Sander, 1993), statistical methods (Lim, 1974; Burgess *et al.*, 1974; Chou & Fasman, 1974a,b, 1978; Garnier *et al.*, 1978), and the nearest neighbor method (Salzberg & Cost, 1992; Yi & Lander, 1993). The accuracy of these approaches is currently limited to about 70%, despite attempts to improve the algorithms and the greatly enlarged database of protein structure (Hayward & Collins, 1992).

There are many cases where a sequence is predicted to form an α-helix but is found to be a beta strand, or vice versa. Our previous work suggested that solvent environment can greatly influence the formation of protein secondary structure (Zhong & Johnson, 1992; Waterhous & Johnson, 1994). We find it possible to manipulate the secondary structure simply by changing the solvent environment. Here we determine the relative propensity for each of 19 amino acids to form an α-helix as a function of methanol concentration. We show that the order of helical propensity for the amino acids, represented by the free energy (ΔG°) of helix formation, can change when the solvent environment changes. Our findings indicate that a 63% methanol solvent is similar to the *average* environment found in proteins. Since the helical propensity derived

from the protein structure database averages the various environments of each amino acid as it is located in various part of protein molecule, these findings could explain why the prediction of secondary structure from amino acid sequence alone remains about 70% successful.

Materials and Methods

Peptide Synthesis

The peptide models of sequence acetyl-Y-VAXAK-VAXAK-VAXAK-amide, where X is substituted with one of the 20 naturally occurring amino acids, were synthesized manually using the solid phase method (Atherton & Shepard, 1989). The synthesis was performed using Rink resin (4-[2',4'-dimethoxyphenyl-FMOC-aminomethyl]-phenoxy on 1% cross-linked divinylbezene-styrene) sealed inside a permeable polypropylene bag. The mesh bag allowed diffusion of amino acids in solution through the bag while keeping the resin and the growing peptide chain inside during the synthesis. The first N^{α} -[9-fluorenyl(methoxycarbonyl)] (FMOC) amino acid was attached to the resin using 1.0:0.9:1.5/HOBT:TBTU:DIPEA. The FMOC protecting group was removed using 30% piperidine/DMF before the next FMOC-amino acid was added. The deprotecting-coupling cycle was performed until the desired length of peptide was reached. The N-terminus of each newly synthesized peptide was then acetylated using 8:1:1/DMF:DIPEA:acetic anhydride for 2 hours before they were cleaved from the resin. The identity and

purity of all peptides were confirmed by high performance liquid chromatography (HPLC) and fast atom bombardment mass spectrometry (FAB-MS).

Circular Dichroism Measurements

The CD spectra were recorded on the Jasco J-720 spectropolarimeter (Jasco Inc.) at 25 °C. The instrument was calibrated daily using (+)-camphorsulfonic acid giving $\Delta\varepsilon_{192.5}/\Delta\varepsilon_{290.5}$ of -2.08 (Chen & Yang, 1977). A 20 L/min nitrogen gas was purged through the optics and sample compartment 20 min before and during the measurement. The data were collected at a 20 nm/min scanning speed, 2 millisecond response time, 2 nm bandwidth, and 3 accumulations using a 1-mm rectangular cell. Samples were prepared by adding an equal amount of stock peptide into a series of microcentrifuge tubes containing 2 mM phosphate buffer (pH 5.5) and methanol from 0 to 90% in 5% increments. All samples were well mixed and reached equilibrium before the measurement began. The amide concentration of peptides used in the measurement was 1 to 2 mM.

Determination of Peptide Concentration

UV absorption spectra of the peptides were measured in buffer from 185 to 400 nm. The molar concentration was calculated using the intensity of the 280 nm band (Elwell, 1976) and the known extinction coefficients (ε_{280}) of some absorbing residues within the sequences (Y = 1,280 $M^{-1}cm^{-1}$, W = 5,690 $M^{-1}cm^{-1}$,

C = 120 $M^{-1}cm^{-1}$). The molar concentration was converted to an amide concentration by using the number of the peptide bonds, and then the extinction coefficient at 190 nm (ε_{190}) was calculated. The ε_{190} was used to determine the amide concentration of a particular peptide before the CD measurement.

Correcting the CD spectra for Absorbing Side-Chains

The CD spectra of peptides in the far UV are normally due to the amide chromophore, so it is routine to analyze a CD spectrum for the amount of each secondary structure. However, amino acids with aromatic and other absorbing side-chains can affect the CD spectra of peptides by the contribution from the transition of such side-chains (Woody, 1978; Chakrabartty et al., 1993). We corrected the CD spectra of our peptides containing amino acids W, Y, F, C, and M for the contribution of their side-chains by singular value decomposition as described previously (Krittanai & Johnson, to be published). In brief, the CD of the 15 peptides without absorbing side-chain are analyzed for their common basis vectors. For peptides with absorbing side-chains, the common basis vectors are fitted to a portion of the CD spectrum that is not affected by the absorbing side-chains, giving us a set of fitting coefficients for each peptide. The fitting coefficients and the common basis vectors are then used to construct the corrected spectra that are free from the contribution of their absorbing side-chains.

Peptide Aggregation Study

If the peptides aggregate, then the environment seen by the amino acids is not simply the solvent. To investigate whether our peptides are aggregated, we measured the CD intensity of all peptides at 222 nm over a large range of concentrations (Lyu et al., 1990). The experiment was performed in 88% methanol solution, which is the solvent most likely to facilitate the aggregation, and the peptide concentration was varied from 2 μM to 10 mM amide. The CD intensity of our peptides were found to be independent of the concentration. Since the concentration used in our titration experiment is about 1-2 mM, we are certain that no aggregation occurs in this work.

Data Analysis

The CD spectra obtained for the 20 peptides at various methanol concentrations were smoothed simultaneously using singular value decomposition (SVD) (Johnson, 1992; Henry & Hofrichter, 1992). This inter-spectral smoothing technique decomposes a matrix containing all CD spectra into three matrices for singular values, basis vectors and fitting coefficients. The basis vectors corresponding to singular values at the noise level were eliminated and the two remaining vectors were used to regenerate the smoothed spectra.

Assuming a two-state transition between random coil and α-helix for the titration (that is, no other significant components), we can estimate the amount of an α-helix for each peptide using the CD intensity at 222 nm ($\Delta\varepsilon_{222}$). The

amount of random coil and helix can be used to determine an equilibrium constant (K_{eq}) for the transition:

$$K_{eq} = [Helix]/[Coil]$$

$$= (\Delta\varepsilon_{obs}-\Delta\varepsilon_{coil})/(\Delta\varepsilon_{helix}-\Delta\varepsilon_{obs})$$

where $\Delta\varepsilon_{helix}$ is the $\Delta\varepsilon_{222}$, representing 100 % α-helix, $\Delta\varepsilon_{coil}$ is the $\Delta\varepsilon_{222}$, representing 100 % random coil, and $\Delta\varepsilon_{obs}$ is the $\Delta\varepsilon_{222}$, for peptide in each condition. The K_{eq} then gives us a free energy of helix formation (ΔG°_{X}) at 25 °C for each peptide X.

$$\Delta G^{\circ}_{X} = -RT \ln [(\Delta\varepsilon_{obs}-\Delta\varepsilon_{coil})/(\Delta\varepsilon_{helix}-\Delta\varepsilon_{obs})]$$

Where R is a gas constant (8.31 Kcal/Mol K) and T is an absolute temperature (K). The ΔG°_{X} (Kcal/Mol) is a measurement of the helical propensity for each of the 19 amino acids. These ΔG°_{X} values are calculated for each peptide in the various methanol concentrations and then plotted in an energy profile as a function of methanol in the solvent. The order of helical propensity in various methanol concentrations is then compared with the helical preference derived from the X-ray database by referring to the free energy of glycine.

$$\Delta\Delta G^{\circ}_{X} = \Delta G^{\circ}_{X} - \Delta G^{\circ}_{Gly}$$

Results and Discussion

Titration of Peptides with Methanol

The CD spectra observed for all 20 peptides in 2 mM sodium phosphate buffer (pH 5.5) have the characteristics of the random coil conformation, showing an intense negative band around 198 nm. When the methanol concentration is increased, the CD indicates a change in peptide conformation from a random coil toward an α-helix (Figure 3.1). The CD spectra of peptides at high methanol concentration have the characteristics of an α-helix with an intense positive band at 193 nm and two negative bands around 208 and 222 nm. The transition between the two characteristic CD spectra can be distinctly observed with every 5% increment of methanol. The plot of CD spectra for the transition has an isosbestic point around 203 nm, supporting the assumption of only two significant components.

There is an exception for the CD spectra for the peptide with X = proline. Unlike others, this peptide does not demonstrate a conformational change from a random coil to an α-helix. Instead, the CD of this peptide in buffer solution has the characteristics of a 3_1 (polyproline II) helix, and becomes a random coil when methanol concentration is increased (data not shown). Since proline has a reputation of being an α-helix breaker in most sequences, this unique outcome would not be a surprise to us. A goal of this work is to compare a relative helical propensity among amino acids substituted in our sequence model, therefore the data obtained from the peptide with X = proline is excluded from our analysis.

We expect the distribution between helix and coil to vary with the peptide as well as with the solvent environment. At high methanol concentration some peptides are more helical than others; in buffer, some peptides have a helical contribution.

Correcting the CD spectra for absorbing side-chains and spectral smoothing

The contribution from amino acids with absorbing side-chains can interfere with the estimate for the α-helix content in many peptides including the ones used in our experiment. We corrected the CD spectra of peptides that contain W, Y, F, C, and M substitution as described in the Materials and Methods section. The corrected and the raw spectra for those peptides are found to be different in the range from 210 to 240 nm but identical elsewhere (Krittanai & Johnson, to be published).. All of our 20 peptides, Ac-Y-VA**X**AK-VA**X**AK-VA**X**AK-amide, are labeled with one tyrosine on the N-terminal. The contribution from this tyrosine is clearly present (Chakrabartty *et al.*, 1993). We assume that the N-terminal tyrosine on each peptide has the same effect for the same percentage of α-helix, and will not affect the comparison among these 19 peptides.

The singular value decomposition analysis of the CD spectra obtained from all peptides gave us two significant singular values. This proves that the system contains only two significant components during our titration. We

reconstruct a spectrum for each peptide using the two singular values and the corresponding basis vectors, and get the smoothed spectra.

Calculation of Free Energy for Helix Formation

The titration curves for this experiment can be derived by two different procedures. One is to make a plot of the intensity of the CD spectra at 222 nm as a function of methanol added, and the other method is by performing SVD on all the spectra from the titration and plotting the V coefficients. The latter method has the advantage of utilizing more data points in the analysis. The titration curves obtained from the intensity at 222 nm are similar to a plot from the V coefficients for the 205 to 250 nm spectral range. Here we use the titration curves derived from the simplest method, the traditional CD at 222 nm.

We can estimate the amount of α -helix for peptides in each condition along the titration by using the intensity of their CD spectra at 222 nm. Researchers usually use a scale of $\Delta\varepsilon_{222}$ value ranging from 0 to -10, to represent the amount from 0 to 100% α-helix in most proteins. Here we use a slightly different scale of $\Delta\varepsilon_{222}$ values, based on the lowest (-0.10) and the highest (-8.75) $\Delta\varepsilon_{222}$ values found among our 19 peptides. We assume the two values are corresponding to a completely random coil and a 100% helix conformation, respectively. However, we find that the two estimating scales (0 to -10 and -0.10 to -8.75) give us only a small difference in free energy, which does not affect the relative order of helical propensity among our 19 amino acids.

Singular value decomposition analysis and an isosbestic point prove there are only two significant components and allow us to simply analyze the data using a two-state model. The free energy (ΔG^{o}) calculated in this work is described as the free energy required by one mole of substituted amino acid to adopt an α-helix conformation. This free energy represents the propensity of a particular amino acid to be found in a helical conformation. When these ΔG^{o}_{X} are profiled as a function of methanol concentration, some of them demonstrate a cooperative sigmoidal curve for the transition (Figures 3.2-3.6). By browsing through the patterns for the energy profiles of all 19 amino acids, we notice some common features that allow us to categorize them into the following groups.

a. Amino Acids with Aliphatic Side-Chains:

Each amino acid in this group has a nonpolar hydrocarbon side-chain of different hydrophobicity. Their relative propensity for α-helix over a wide range of solvent conditions is found to be correlated with the hydrophobicity of their side-chains, L > A > I > V > G (Figure 3.2). While L and A demonstrate the highest propensity to adopt an α-helix, G is found to the poorest helix former (except for P, which is constrained by a ring structure). We find that the helical propensity of L, A, and I are come very close together at high methanol content.

b. Amino Acids with Aromatic Side-Chains:

The hydrophobicity of aromatic side-chains is also found to be correlated well with their relative propensity, which can be ranked: W > F > Y (Figure 3.3). The

side-chain of Y and F are very similar in their structure, but their propensity profiles are significantly different. This is explained by the significant difference between the hydrophobicity of their side-chains, F > Y.

c. Amino Acids with Acidic and Amino Side-Chains:

The amino acids that fall into this group are D, N, E, and Q. The similarity of their side-chain structure and their hydrophobicities have a great influence in dictating the pattern of their propensity profiles. The propensity profiles of D and N are almost identical, and as are E and Q (Figure 3.4). The helical propensity of E and Q are higher than those for D and N at high methanol content. This is consistent with the difference in hydrophobicity due to the length of their side-chains.

d. Amino Acids with Hydroxy and Sulfur-Containing Side-Chains:

S, T and C can be classified into the same group according to the hydrophilic character of their side-chains (OH and SH). The pattern of their propensity profiles is about the same. The exception is found in M, which also has a sulfur atom in its side-chain (Figure 3.5). The hydrophillic character of the sulfur on the side-chain of M should be weak due to its long hydrophobic carbon chain. Thus the hydrophobicity of its side-chain would be virtually the same as those of aliphatic side-chains. This explanation is confirmed by the similar propensity profile for M, L, and A.

e. Amino Acids with Bulky, Charged Side-Chains:

The amino acids showing a similar propensity profile in this group are K, R, and H. These amino acids have a long and bulky side-chain with a positive charge. This unique character of their side-chains may be responsible the similar pattern for their propensity profiles (Figure 3.6).

Changes in Relative Order of Helical Propensity

When the propensity profiles for all 19 amino acids are compared together on the same plot, we find several crossovers among their free energy (Figure 3.7, shows ten profiles). These crossovers indicate changes in the order of helical propensity as the solvent environment changes. The relative propensity for α-helix conformation of a specific amino acid is relatively high in one environment and becomes low in different environment, or vice versa. For instance, H, R and K are poor helix formers comparing to D and N in 15 % methanol, but they are significantly better in 80% methanol. Such a switching in propensity order is also found in several other amino acids (Figure 3.8). When we compare the relative propensity for all 19 amino obtained at various methanol concentrations with the helical preferences($P\alpha$) derived for internal amino acids from statistical calculation from 212 proteins in the X-ray database (Williams et al., 1987), we find that the relative order of propensity in 63% methanol solvent correlates well with those preferences (Figure 3.9). This correlation indicates that a 63% methanol environment is similar to an *average* environment found in proteins. The changing of relative propensity found in this work undoubtedly

happens in protein structures, where the 20 amino acids are found in various locations. The local environment that is seen by an amino acid may play a role in dictating secondary structure. Here we mimic the environment of protein interior and exterior by using the various concentrations of methanol, and these solvent condition affect the relative order for helical propensity.

A number of laboratories have determined the helical propensity from the structure of proteins in the X-ray database, and used them to predict the secondary of proteins. These sequence-based predictions are limited to around 70% accuracy, regardless of the algorithm and the size of the database. Our findings can be part of the explanation for the limitation of these predicting methods. Most of these methods concentrate on the information from amino acid sequence alone. These sequences come from the whole length of polypeptide chain that make up a protein molecule, meaning that various types of environments are included and averaged. Therefore, the calculated helical propensity for the 20 amino acids is only good for amino acids in an average environment, while the relative helical propensity of an amino acid depends in fact on its local environment. Understanding the changes in relative propensity for protein secondary structure as a function of environment should improve the prediction of protein secondary structure.

Many laboratories have measured the helical propensity of amino acids in buffer. These measurements have been performed using various peptide models (Lyu et al., 1990; Merutka et al., 1990; O'Neil & DeGrado, 1990; Padmanabhan et al., 1990; Chakrabartty et al., 1991; Gans et al., 1991; Kemp et

al., 1991; Scholtz et al., 1991; Rohl et al., 1992; Stellwagen et al., 1992; Park et al., 1993; Chakrabartty et al., 1994). The order of helical propensities for all amino acids obtained among those systems do not agree very well. When we compared the propensity from this work to those data, they do not correlate well with any particular one either. However, there is some universal agreement among those and our data, mostly for the order of amino acids with aliphatic side-chains.

Figure 3.1: Typical CD spectra of a Ac-Y-VAXAK-VAXAK-VAXAK-amide during a methanol titration (shown here for X = Alanine).

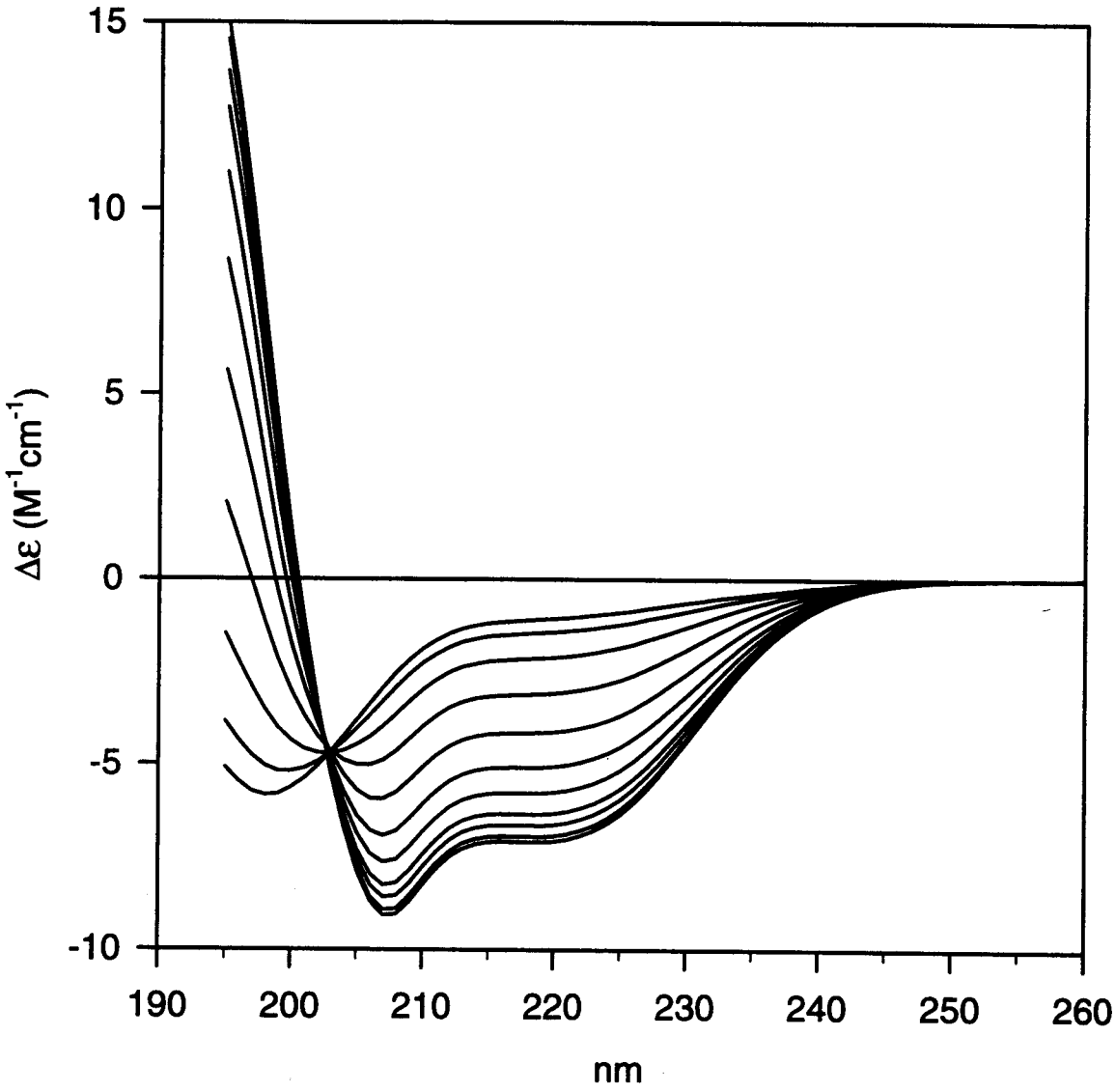


Figure 3.2: Profiles of free energy (ΔG^{o}) for α-helix formation by amino acids with nonpolar aliphatic side-chains, G, A, V, I, L.

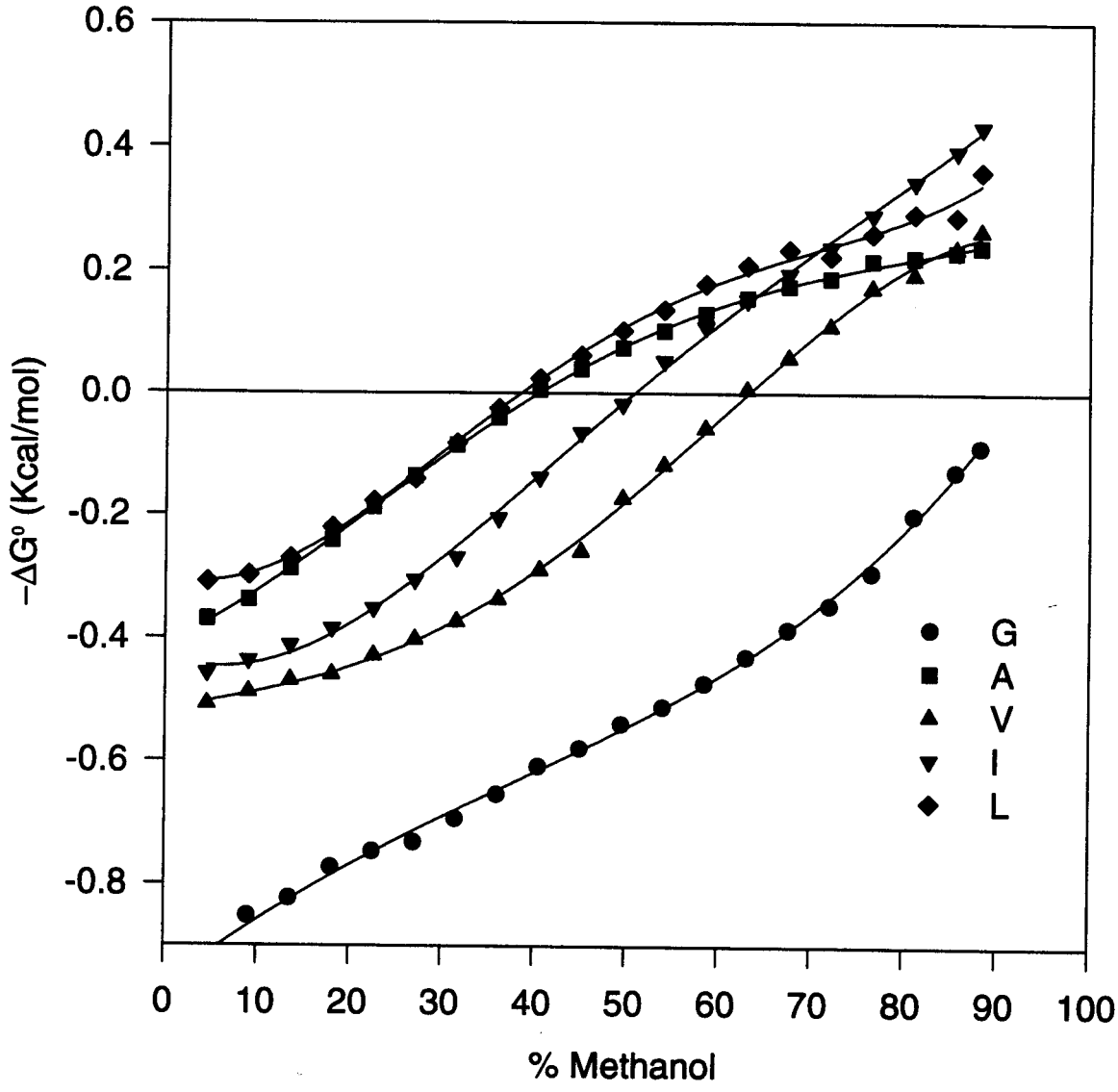


Figure 3.3: Profiles of free energy (ΔG^{o}) for α-helix formation by amino acids with aromatic side-chains, F, W, Y.

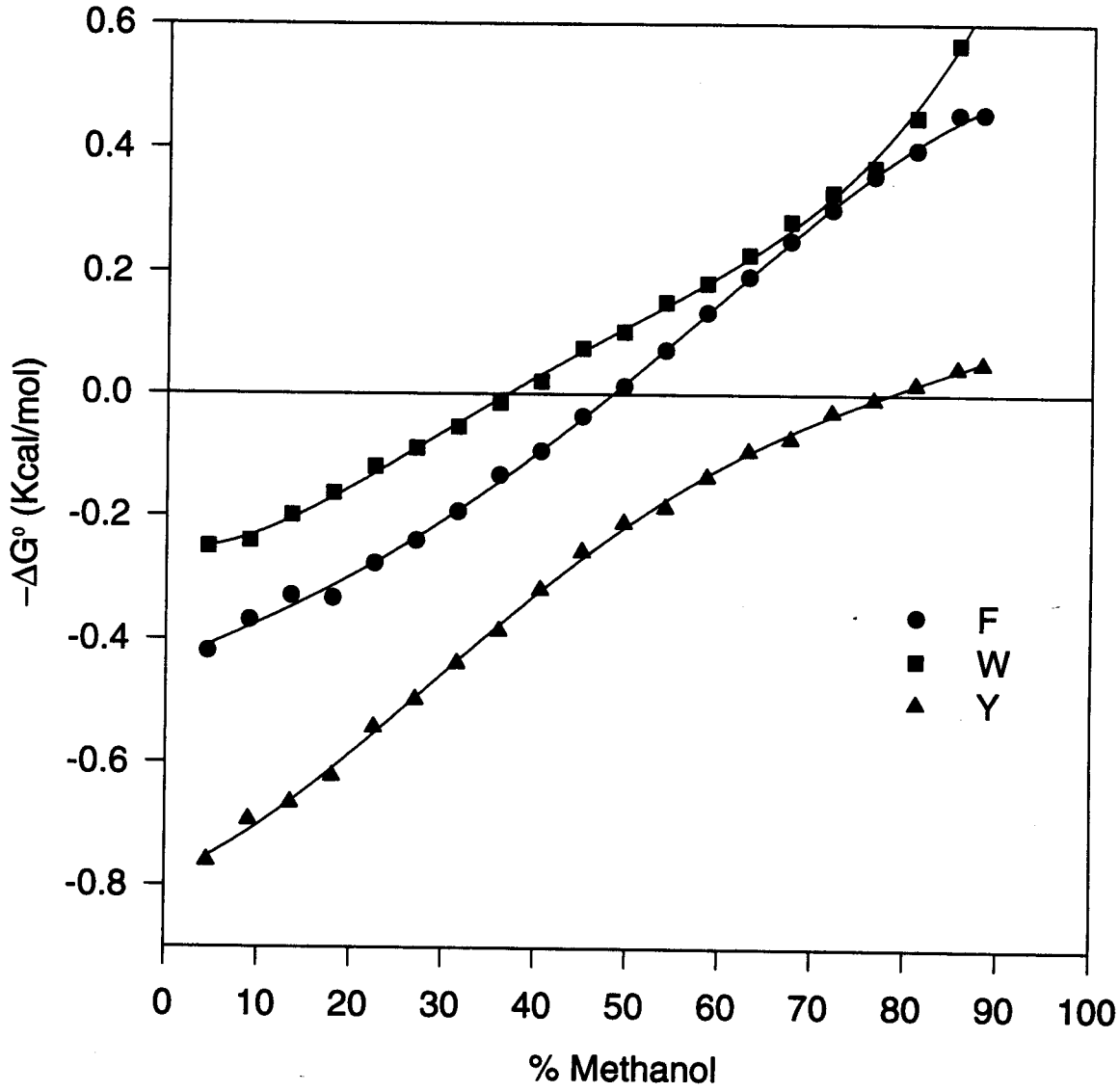


Figure 3.4: Profiles of free energy (ΔG^{o}) for α-helix formation by amino acids with acidic and amino side-chains, D, N, E, Q.

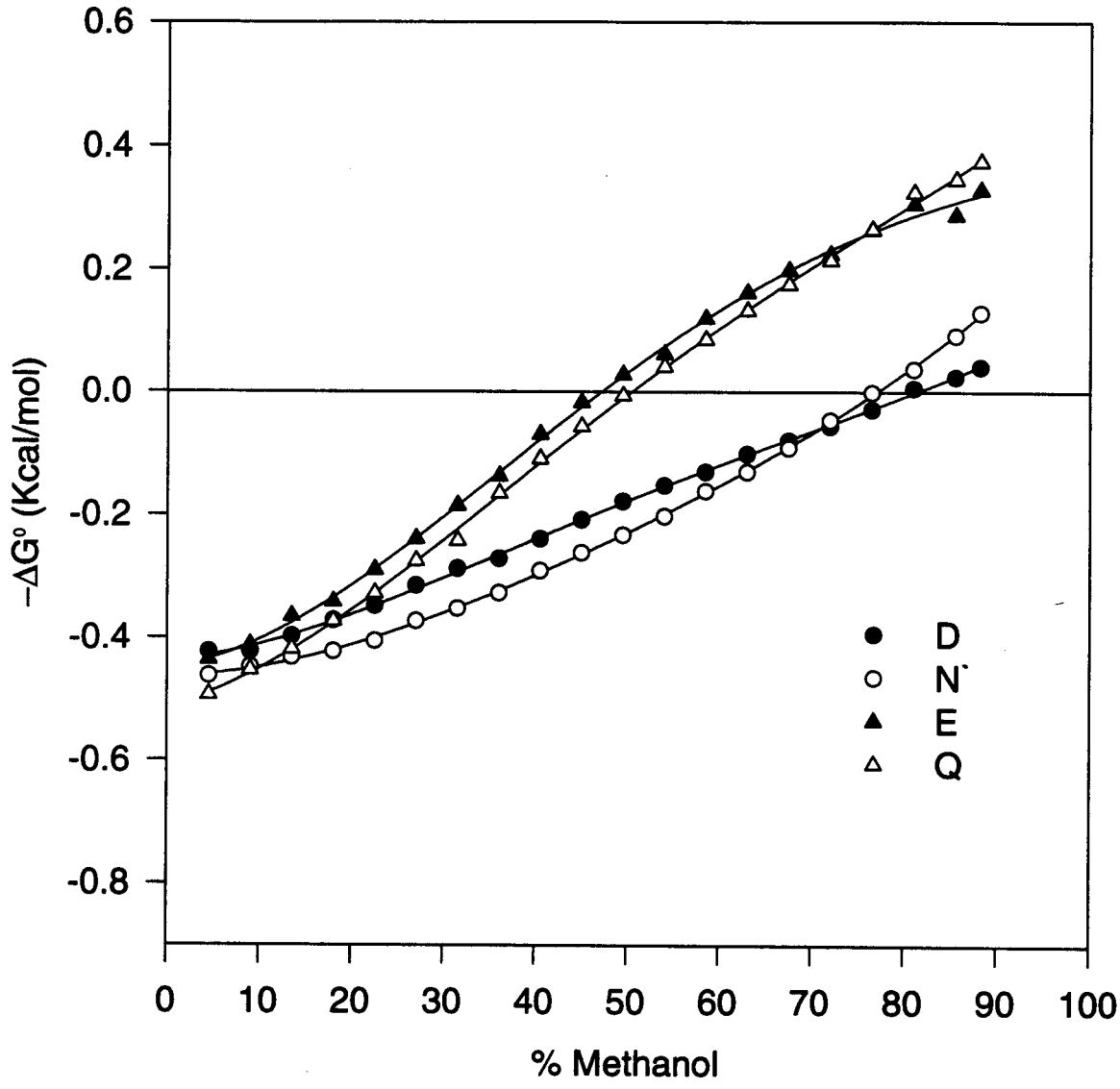


Figure 3.5: Profiles of free energy (ΔG°) for α-helix formation by amino acids with hydroxy and sulfur-containing side-chains, S, T, C, M.

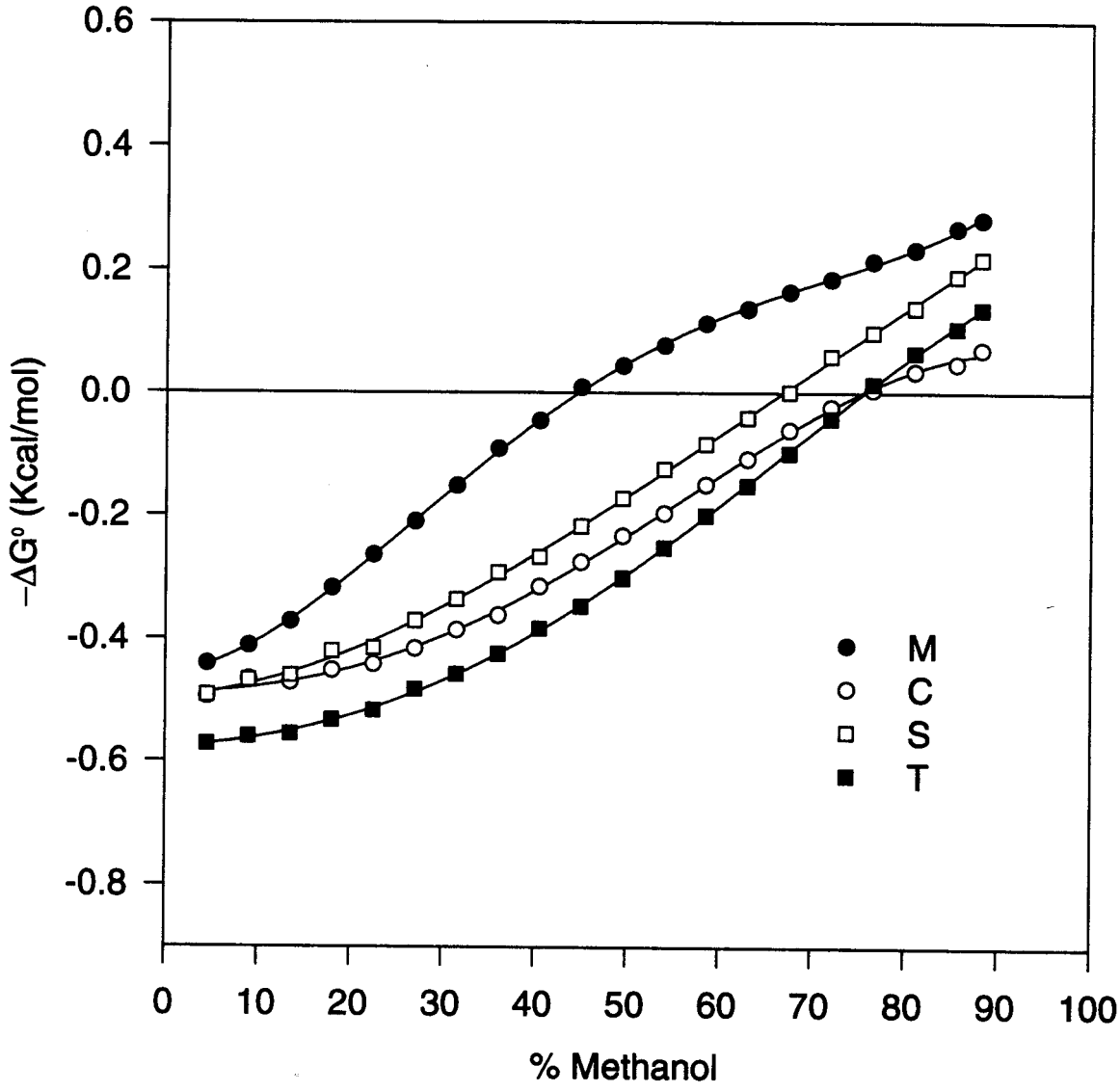


Figure 3.6: Profiles of free energy (ΔG^{o}) for α-helix formation for amino acid with bulky, charged side-chains, H, K, R.

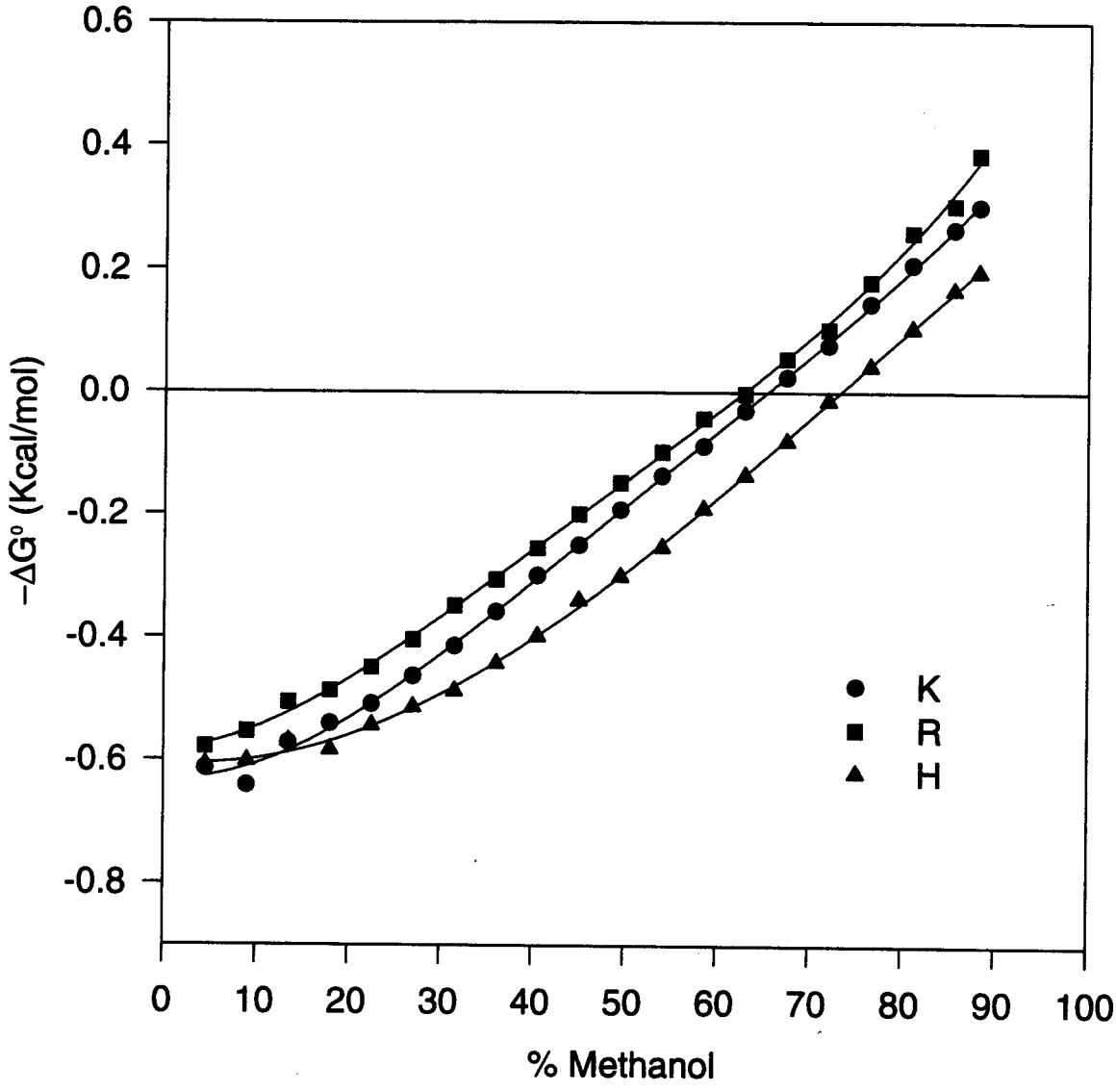


Figure 3.7: Profiles of free energy (ΔG°) for α-helix formation compared among some amino acids.

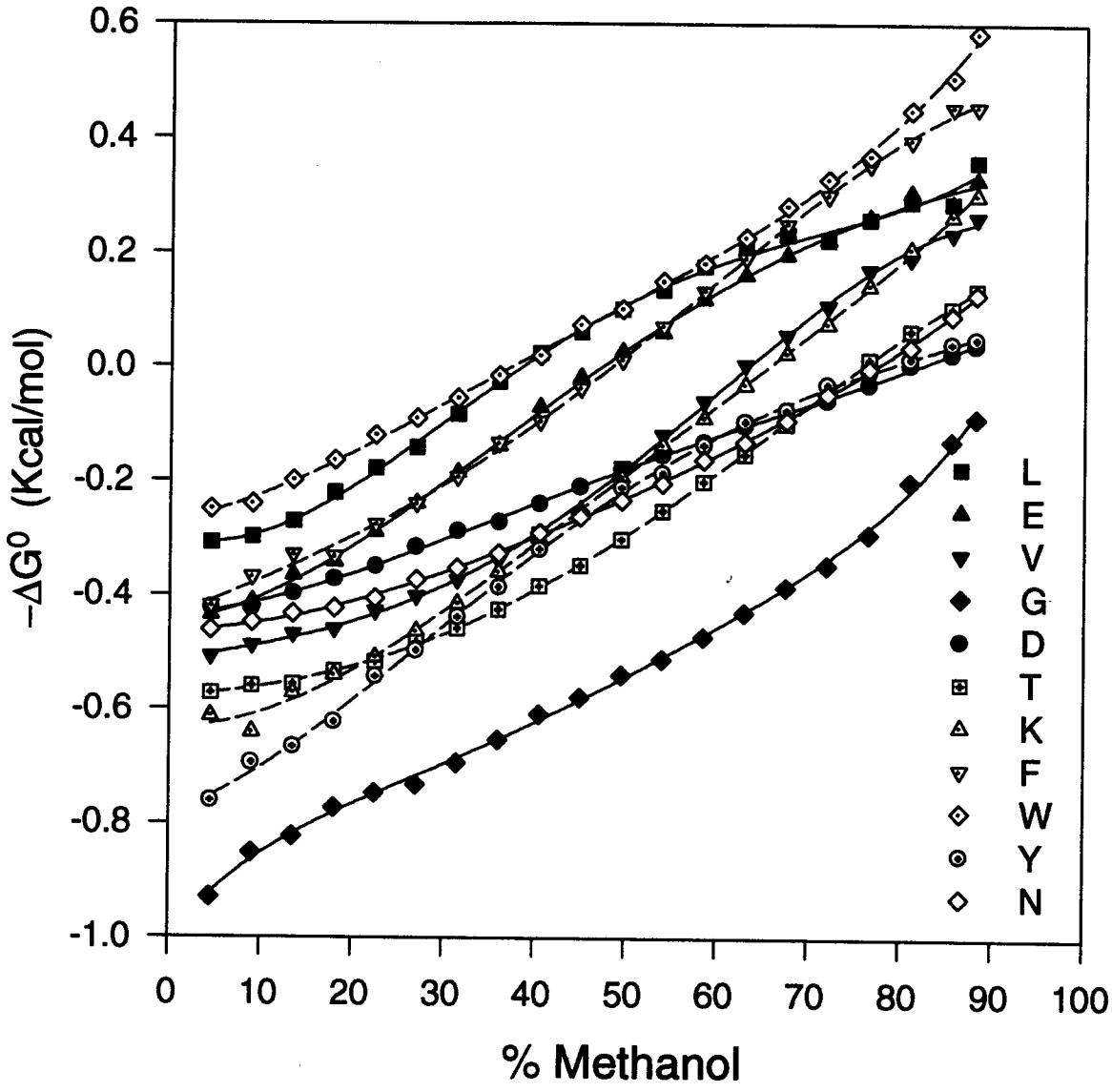


Figure 3.8: A bar-chart showing the comparison of the relative propensity among conditions with 13, 63, and 85% methanol content.

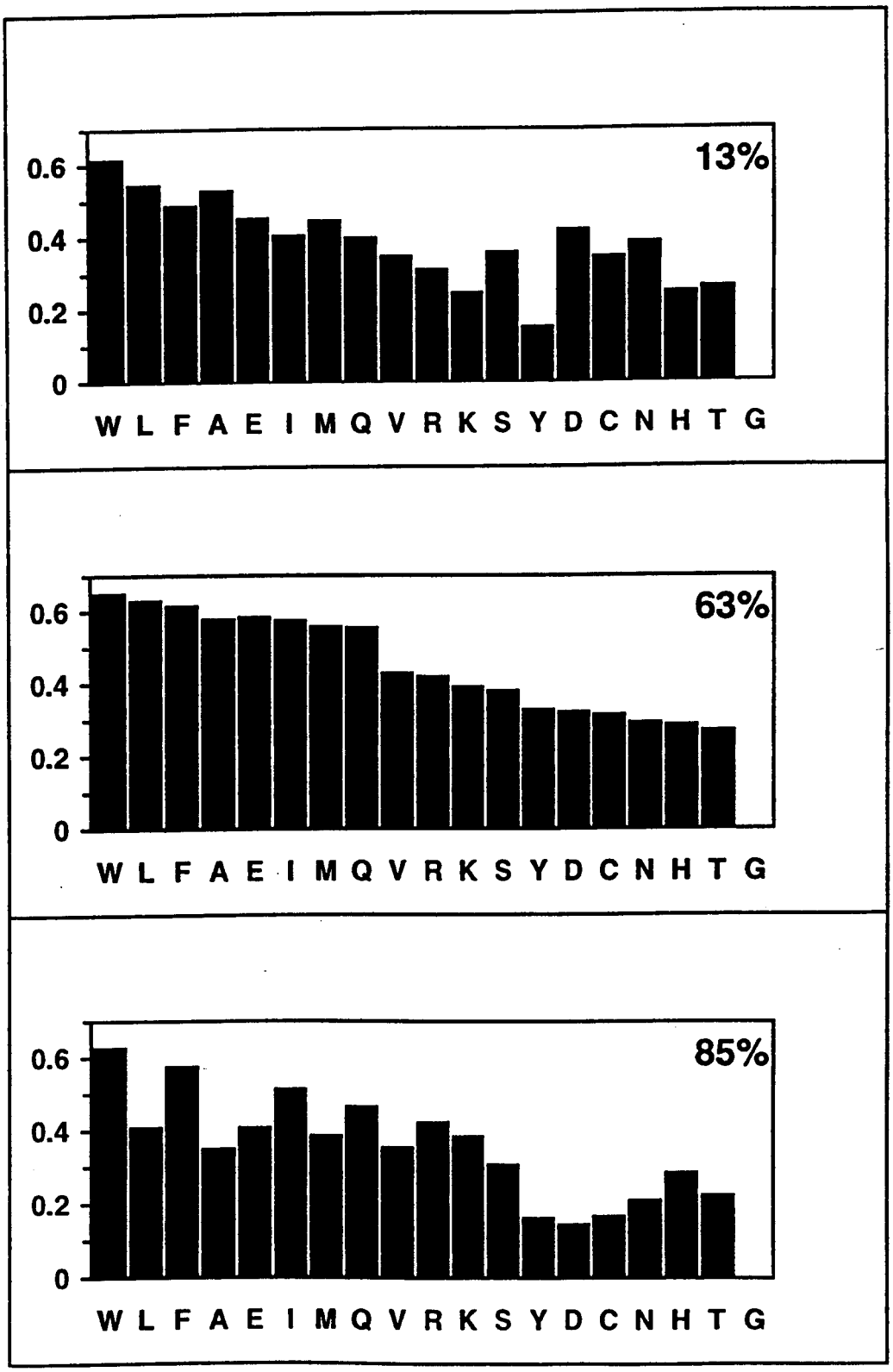
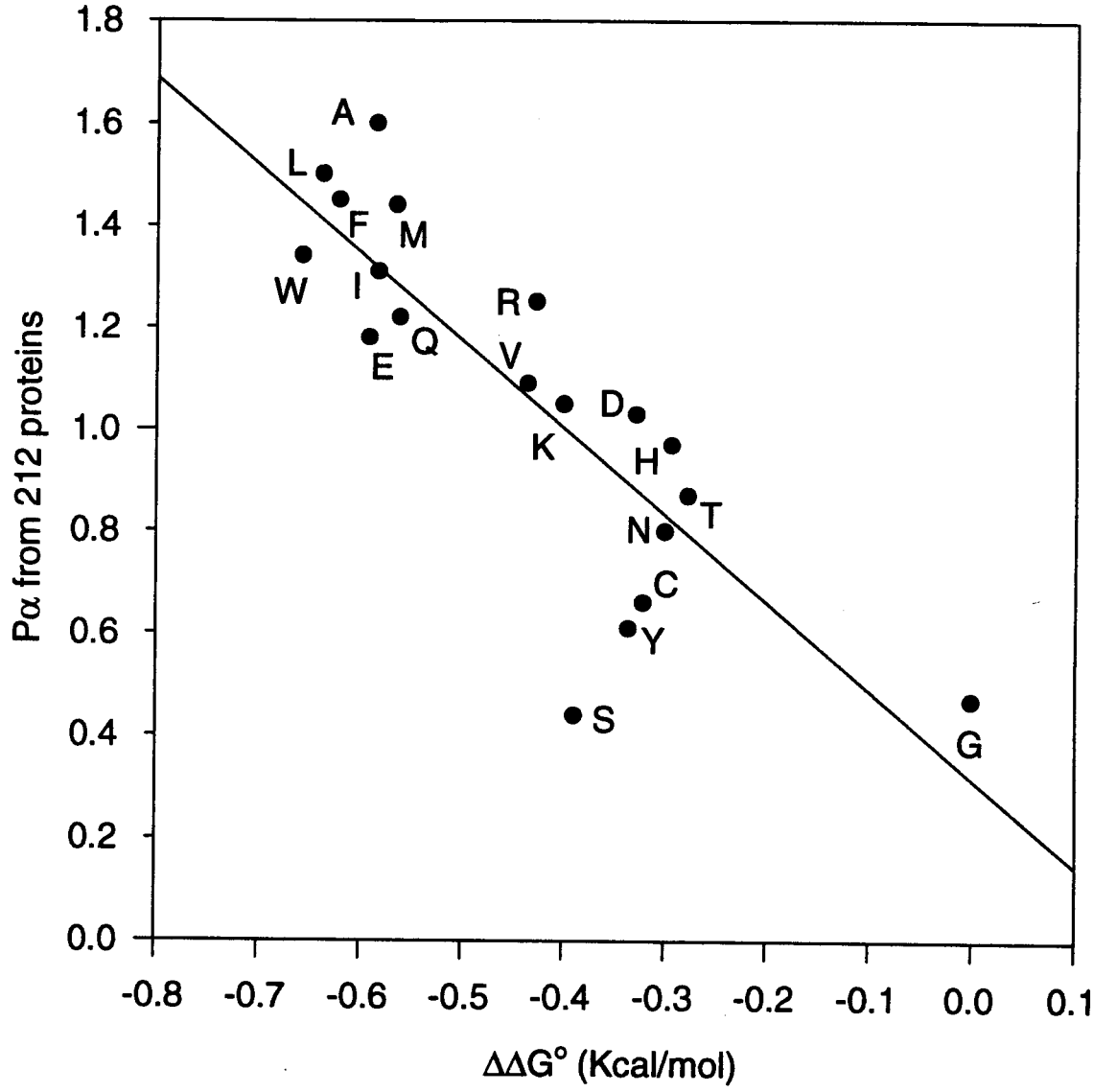


Figure 3.9: A plot of $\Delta\Delta G^{\circ}$ for amino acids in 63% methanol against the helical preference parameter (Pα) derived from 212 proteins.



Conclusions

We have measured the free energy of helix formation, which represents the helical propensity of 19 amino acids as a function of methanol. CD measurements indicate a two-state transition from random coil to α-helix, when its solvent environment has increased methanol content. The profiles of relative α-helical propensity plotted as a function of solvent environment are found be to correlated well with the hydrophobicity of amino acid side-chains, making it possible to classify the propensity profiles according to their hydrophobic similarity. The comparative analysis of these profiles has demonstrated changes of the relative helical propensity of the amino acids according to their solvent environment. We believe this propensity change can occur within the structure of protein molecules, resulting in the environment-dependent propensity. The relative propensity order measured for the peptides in 63% methanol correlates well with the propensity calculated from X-ray data, indicating that our 63% methanol is similar to an average environment in proteins. This work suggests the importance of environment in assigning of structural propensity which is used for determining the secondary structure of proteins.

Acknowledgments

This work was supported by PHS Grant GM 21479 from the National Institutes of Health.and the Development and Promotion of Science Talents Project (DPST) by the Royal Thai Government.

References

Atherton, E., & Sheppard, R. C. (1989) in *Solid Phase Peptide Synthesis*, IRL Press at Oxford University Press, Oxford, UK.

Anfinsen, C. B., Haber, E., Sela, M., & White, F. H. (1961) *Proc. Natl. Acad. Sci. USA. 47*, 1309-1314.

Burgess, A. W., Ponnuswamy, P. K., & Scheraga, H. A. (1974) *Israel J. Chem. 12*, 239-286.

Chakrabartty, A., Schellman, J. A., & Baldwin, R. L. (1991) *Nature 351*, 586-588.

Chakrabartty, A., Kortemme, T., Padmanabhan, S., & Baldwin, R. L. (1993) *Biochemistry 32*, 5560-5565.

Chakrabartty, A., Kortemme, T., & Baldwin, R. L. (1994) *Protein Sci. 3*, 843-852.

Chen, G. C., & Yang, J. T. (1977) *Anal. Letts 10*, 1195-1270.

Chou, P. Y., & Fasman, G. D. (1974a) *Biochemistry 13*, 211-222.

Chou, P. Y., & Fasman, G. D. (1974b) *Biochemistry 13*, 222-245.

Chou, P. Y., & Fasman, G. D. (1978) *Adv. Enzymology. 47*, 45-148.

Edelman, J., & White, S. P. (1989) *J. Mol. Biol. 210*, 195-209.

Elwell, M. L. (1976) Ph.D. thesis, Department of Chemistry, University of Oregon.

Gans, P. J., Lyu, P. C., Manning, M. C., Woody, R. W., & Kallenbach, N. R. (1991) *Biopolymers 31*, 1605-1614.

Garnier, J., Osguthorpe, D. J., & Robson, B. (1878) *J. Mol. Biol. 120*, 97-120.

Hayward, S., & Collins, J. F. (1992) *Proteins: Struct., Funct., Genet. 14*, 372-381.

Henry, E. R., & Hofrichter, J. (1992) in *Methods in Enzymology* (Brand, L., & Johnson, M. L., Eds.) pp.426-447, Academic Press, San Diego, CA.

Holley, L. H., & Karplus, M. (1989) *Proc. Natl. Acad. Sci. USA. 86*, 152-156.

Johnson, W. C., Jr. (1992) in *Methods in Enzymology* (Brand, L., & Johnson, M. L., Eds.) pp.426-447, Academic Press, San Diego, CA.

Kemp, D. S., Boyd, J. G., & Muendel, C. C. (1991) *Nature 352*, 451-454.

Levin, J. M., Robson, B., & Garnier, J. (1986) *FEBS Lett. 205*, 303-308.

Lim, V. I. (1974) *J. Mol. Biol. 88*, 857-894.

Lyu, P. C., Liff, M. I., Marky, L. A., & Kallenbach, N. R. (1990) *Science 250*, 669-673.

Merutka, G., Lipton, W., Shalongo, W., Park, S. H., Stellwagen, E. (1990) *Biochemistry 29*, 7511-7515.

Nishikawa, K., & Ooi, T. (1986) *Biochim. Biophys. Acta. 871*, 45-54.

O' Neil, K. T., & DeGrado, W. F. (1990) *Science 250*, 646-651.

Padmanabhan, S., Marqusee, S., Ridgeway, T., Laue, T. M., & Baldwin, R. L. (1990) *Nature 344*, 268-270.

Park, S. H., Shalongo, W., & Stellwagen, E. (1993) *Biochemistry 32*, 7048-7053.

Pongor, S., & Szaley, A. A. (1985) *Proc. Natl. Acad. Sci. USA. 82*, 366-370.

Qian, N., & Sejnowski, T. J. (1988) *J. Mol. Biol. 160*, 865-884.

Rohl, C. A., Scholtz, J. M., York, E. J., Stewart, J. M., & Baldwin, R. L. (1992) *Biochemistry 31*, 1263-1269.

Rost, B., & Sander, C. (1993) *J. Mol. Biol. 232*, 584-599.

Salzberg, S., & Cost, S. (1992) *J. Mol. Biol. 227*, 371-374.

Scholtz, J. M., Qian, H., York, E. J., Stewart, J. M., & Baldwin, R. L. (1991) *Biopolymers 31*, 1463-1470.

Stellwagen, E., Park, S. H., Shalongo, W., & Jain, A. (1992) *Biopolymers 32*, 1193-1200.

Stolorz, P., Lapedes, A., & Xia, Y. (1992) *J. Mol. Biol. 225*, 363-377.

Sweet, R. M. (1986) *Biopolymers 25*, 1565-1577.

Waterhous, D. V., & Johnson, W. C., Jr. (1994) *Biochemistry 33*, 2121-2128.

Williams, R. W., Chang, A., Juretic, D., & Loughran, S. (1987) *Biochim. Biophys. Acta. 916*, 200-204.

Woody, R. W. (1978) *Biopolymers 17*, 1451-1467.

Yi, T. M., & Lander, E. S. (1993) *J. Mol. Biol. 232*, 1117-1129.

Zhong, L., & Johnson, W. C., Jr. (1992) *Proc. Natl. Acad. Sci. USA. 89*, 4462-4465.

Zvelebil, M. J., Barton, G. J., Taylor, W. R., & Sternberg, M. J. E. (1987) *J. Mol. Biol. 195*, 957-961.

CHAPTER IV

CONCLUSIONS

1. Methanol induces a conformational change in the peptide sequence acetyl-Y-VAXAK-VAXAK-VAXAK-amide, where X is one of the 19 amino acids excluding P, from a random coil to α-helix.

2. The sequence acetyl-Y-VAPAK-VAPAK-VAPAK-amide is a left-handed 3_1 (polyproline II) helix in sodium phosphate buffer and becomes a random coil in methanol solution.

3. In addition to an amide CD in the far-UV spectra, there is an extra contributing CD band found in a peptide that contains amino acids with absorbing side-chains. These amino acids are tryptophan, tyrosine, phenylalanine, cysteine, and methionine.

4. The side-chain contribution to the CD of peptide can be corrected by using our method based on Singular Value Decomposition (SVD) theory.

5. The common basis vectors derived from a peptide of related structure and sequence is required as basis elements in correcting the CD spectra for side-chains.

6. The contribution to the CD spectra of peptide acetyl-Y-VAXAK-VAXAK-VAXAK-amide is positive when X = tryptophan, tyrosine, phenylalanine, and negative when X = cysteine, and methionine.

7. A free energy for helix formation of our peptides is favorable with the solvent condition of high methanol content.

8. The relative order of helical propensity for the 19 amino acids, derived from a peptide in different methanol contents, are not the same.

9. The changes of helical propensity for amino acids are the result of environment provided by different concentrations of methanol.

10. The 63% methanol solution represents an environment similar to the average environment found in proteins.

BIBLIOGRAPHY

Anfinsen, C. B., Haber, E., Sela, M., & White, F. H. (1961) *Proc. Natl. Acad. Sci. USA. 47*, 1309-1314.

Atherton, E., & Sheppard, R. C. (1989) in *Solid Phase Peptide Synthesis,* IRL Press at Oxford University Press, Oxford, UK.

Baba, M., Hamaguchi, K., & Ikenaka, T. (1969) *J. Biochem. 65*, 113-121.

Barlow, D. J., & Thornton, J. M. (1988) *J. Mol. Biol. 201*, 601-619.

Bell, J. A., Becktel, W. J., Sauer, U., Baase, W. A., & Matthews, B. W. (1992) *Biochemistry 31*, 3590-3596.

Beychok, S., Armstrong, J. M., Lindblow, C., & Edsall, J. T. (1966) J. Biol. Chem. 241, 5150-5160.

Blaber, M., Zhang, X. J., & Matthews, B. W. (1993) *Science 260*, 1637-1640.

Blake, C. C. F., Koeing, D. F., Mair, G. A., North, A. C. T., Phillips, D. C., & Bush, C. A., & Gibbs, D. E. (1972) *Biochemistry 11*, 2421-2427.

Bohm, G., Muhr, R., & Jaenicke, R. (1992) *Protein Eng. 5*, 191-195.

Boren, K., Freskgard, P. O., & Carlsson, U. (1996) *Prot. Sci. 5*, 2479-2484.

Brahms, S., Brahms, J., Spach, G., & Brack, A. (1977) *Proc. Natl. Acad. Sci. USA.74*, 3208-3212.

Briggs, M. S., Cornell, D. G., Dluhy, R. A., & Gierasch, L. M. (1986) *Science 233*, 206-208.

Burgess, A. W., Ponnuswamy, P. K., & Scheraga, H. A. (1974) *Israel J. Chem. 12*, 239-286.

Bush, C. A., & Gibbs, D. E. (1972) *Biochemistry 11*, 2421-2427.

Cameron, D. L., & Tu, A. T. (1977) *Biochemistry 16*, 2546-2553.

Chakrabartty, A., Kortemme, T., & Baldwin, R. L. (1994) *Protein Sci. 3*, 843-852.

Chakrabartty, A., Kortemme, T., Padmanabhan, S., & Baldwin, R. L.(1993) *Biochemistry 32*, 5560-5565.

Chakrabartty, A., Schellman, J. A., & Baldwin, R. L. (1991) *Nature 351*, 586-588.

Chen, G. C., & Yang, J. T. (1977) *Anal. Letts 10*, 1195-1270.

Chen, Y. H., Lo, T., B., & Yang, J. T. (1977) *Biochemistry 16*, 1826-1830.

Chen, Y. H., Yang, J. T.,& Chan, K. H. (1974) Biochemistry 13, 3350-3359.

Chou, P. Y., & Fasman, G. D. (1974a) *Biochemistry 13*, 211-222.

Chou, P. Y., & Fasman, G. D. (1974b) *Biochemistry 13*, 222-245.

Chou, P. Y., & Fasman, G. D. (1978) *Adv. Enzymol. 47*, 45-148.

Chou, P. Y., & Fasman, G. D. (1979) *Biophys. J. 26*, 367-400.

Compton, L. A., & Johnson, W. C., Jr. (1986) *Anal. Biochem. 155*, 155-167.

Cornell, D. G., Dluhy, R. A., Briggs, M. S., McKnight, C. J., & Gierasch, L. M. (1989) *Biochemistry 28*, 2789-2797.

Dill, K. A. (1990) *Biochemistry 29*, 7133-7155.

Edelman, J., & White, S. P. (1989*) J. Mol. Biol. 210*, 195-209.

Elwell, M. L. (1976) Ph.D. thesis, Department of Chemistry, University of Oregon.

Forsythe, G. E., Malcolm, M. A., & Moler, C. B. (1977) in *Computer Methods for Mathematical Computations*. pp. 192-236, Prentice-Hall Inc., Englewood Cliffs, NJ.

Freskgard, P. O., Martensson, L. G., Jonasson, P., Jonsson, B. H., & Carlsson, U. (1994) *Biochemistry 33*, 14281-14288.

Gans, P. J., Lyu, P. C., Manning, M. C., Woody, R. W., & Kallenbach, N. R. (1991) *Biopolymers 31*, 1605-1614.

Garnier, J., Osguthorpe, D. J., & Robson, B. (1878) *J. Mol. Biol. 120*, 97-120.

Gibrat, J. F., Garnier, J., & Robson, B. (1987) J. Mol. Biol. 198, 425-443.

Goux, W. J., & Hooker, T. M., Jr. (1975) *J. Am. Chem. Soc. 97*, 1605-1606.

Green, N. M., & Melamed, M. D. (1966) *Biochem. J. 100*, 614-621.

Guzzo, A. V. (1965) *Biophys. J. 5*, 809-822.

Hayward, S., & Collins, J. F. (1992) *Proteins: Struct., Funct., Genet. 14*, 372-381.

Hennessey, J. P., Jr., & Johnson, W. C., Jr. (1981) *Biochemistry 20*, 1085-1094.

Henry, E. R., & Hofrichter, J. (1992) in *Methods in Enzymology* (Brand, L., & Johnson, M. L., Eds.) pp. 129-192, Academic Press, San Diego, CA.

Holley, L. H., & Karplus, M. (1989) *Proc. Natl. Acad. Sci. USA. 86*, 152-156.

Hooker, T. M. Jr., & Schellman, J. A. (1970) *Biopolymers 9*, 1319-1348.

Johnson, W. C., Jr. (1990) *Proteins 7*, 205-214.

Johnson, W. C., Jr. (1992) in *Methods in Enzymology* (Brand, L., & Johnson, M. L., Eds.) pp. 426-447, Academic Press, San Diego, CA.

Kartha, G., Bello, J., & Harker, D. (1967) *Nature (London) 213*, 862-865.

Kemp, D. S., Boyd, J. G., & Muendel, C. C. (1991) *Nature 352*, 451-454.

Kendrew, J. C., Dickerson, R. E., Strandbverg, B. E., Hart, R. G., Davies, D. R., Phillips, D. C., & Shore, V. C. (1960) *Nature (London) 185*, 422-427.

Killian, J. A., Prasad, K. U., Hains, D., & Urry, D. W. (1988) *Biochemistry 27*, 4848-4855.

Kuntz, I. D. (1972) *J. Am. Chem. Soc. 94*, 4009-4012.

Levin, J. M., Robson, B., & Garnier, J. (1986) *FEBS Lett. 205*, 303-308.

Lim, V. I. (1974) *J. Mol. Biol. 88*, 857-894.

Low, B. W., Lovell, F. M., & Rudko, A. D. (1968) *Proc. Natl Acad. Sci. USA. 69*, 1519-1526.

Lyu, P. C., Liff, M. I., Marky, L. A., & Kallenbach, N. R. (1990) *Science 250*, 669-673.

Maeda, H., Shiraishi, H., Onodera, S., & Ishida, N. (1973) *Int. J. Peptide Protein Res. 5*, 19-26.

Manavalan, P., & Johnson, W. C., Jr. (1987) *Anal. Biochem. 167*, 76-85.

Mandel, R., & Holzwarth, G. (1972) *J. Chem. Phys. 57*, 3469-3477.

Manez, A., Bouet, F., Tamiya, N., & Fromageot, P. (1976) *Biochim. Biophys. Acta 453*, 121-132.

Manning, M. C., & Woody, R. W. (1989) *Biochemistry 28*, 8609-8613.

Manning, M. C., & Woody, R. W. (1991) *Biopolymers 31*, 569-586.

Matthews, B. W., Siegler, P. B., Henderson, R., & Blow, D. M. (1967) *Nature (London) 214*, 652-656.

Merutka, G., Lipton, W., Shalongo, W., Park, S. H., Stellwagen, E. (1990) *Biochemistry 29*, 7511-7515.

Moffitt, W. (1956) *J. Chem. Phys. 25*, 467-478.

Nishikawa, K., & Ooi, T. (1986) *Biochim. Biophys. Acta. 871*, 45-54.

Noble, B. & Daniel, J. W. (1977) in *Applied Linear Algebra*. pp. 323-342, Prentice-Hall Inc., Englewood Cliffs, NJ.

O' Neil, K. T., & DeGrado, W. F. (1990) *Science 250*, 646-651.

Padmanabhan, S., Marqusee, S., Ridgeway, T., Laue, T. M., & Baldwin, R. L. (1990) *Nature 344*, 268-270.

Palau, H., Argos, P., & Puigdomenech, P. (1982) *Int. J. Peptide Protein Res. 19*, 394-401.

Park, S. H., Shalongo, W., & Stellwagen, E. (1993) *Biochemistry 32*, 7048-7053.

Pauling, L., & Corey, R. B. (1951) *Proc. Natl. Acad. Sci. USA. 37*, 235-285.

Pauling, L., Corey, R. B., & Branson, H. R. (1951) *Proc. Natl. Acad. Sci. USA. 37*, 205-211.

Perczel, A., Hollosi, M., Tusnady, G., & Fasman, G. D. (1991) *Protein Eng. 4*, 669-679.

Periti, P. F., Quagliarotti, G., & Liquori, A. M. (1967) *J. Mol. Biol. 24*, 313-322.

Perutz, M. F., Kendrew, J. C., & Watson, M. J. (1965) *J. Mol. Biol. 13*, 669-678.

Pongor, S., & Szaley, A. A. (1985) *Proc. Natl. Acad. Sci. USA. 82*, 366-370.

Presnell, S. R., Cohen, B. E., & Cohen, F. E. (1992) *Biochemistry 31*, 983-993.

Prothero, J. W. (1966) *Biophys. J. 6*, 367-370.

Provencher, S. W., & Glockner, J., (1981) *Biochemistry 20*, 33-37.

Qian, N., & Sejnowski, T. J. (1988) *J. Mol. Biol. 160*, 865-884.

Rao, S., Zhu, Q. L., Vajda, S., & Smith, T. (1993) *FEBS 322*, 143-146.

Robson, B., & Suzuki, E. (1976) *J. Mol. Biol. 107*, 327-356.

Rohl, C. A., Scholtz, J. M., York, E. J., Stewart, J. M., & Baldwin, R. L. (1992) *Biochemistry 31*, 1263-1269.

Rose, G. D., & Seltzer, J. P. (1977) *J. Mol. Biol. 113*, 153-164.

Rost, B., & Sander, C. (1993) *J. Mol. Biol. 232*, 584-599.

Salzberg, S., & Cost, S. (1992) *J. Mol. Biol. 227*, 371-374.

Sarma, V. R. (1965) *Nature (London) 206*, 757-763.

Schellman, J. A. & Oriel, P. (1962) *J. Chem. Phys. 37*, 2114-2124.

Scholtz, J. M., Qian, H., York, E. J., Stewart, J. M., & Baldwin, R. L. (1991) *Biopolymers 31*, 1463-1470.

Serrano, L., Sancho, J., Hirshberg, M., & Firsht, A. R. (1992) *J. Mol. Biol. 227*, 544-559.

Shestopalov, B. V. (1990) *Mole-Kuly. Biol. 24*, 1117-1125.

Short, K. W., Wallace, B. A., Myers, R. A., Fodor, S. P. A., & Dunker, A. K. (1987) *Biochemistry 26*, 557-562.

Simon, E. R. & Blout, E. R. (1968) *J. Biol. Chem. 243*, 218-221.

Snow, J. W., Hooker, T. M. Jr., & Schellman, J. A. (1977) *Biopolymers 16*, 121-142.

Sreerama, N., & Woody, R. W. (1993) *Anal. Biochem. 209*, 32-44.

Sreerama, N., & Woody, R. W. (1994) *J. Mol. Biol. 242*, 497-507.

Stellwagen, E., Park, S. H., Shalongo, W., & Jain, A. (1992) *Biopolymers 32*, 1193-1200.

Stolorz, P., Lapedes, A., & Xia, Y. (1992) *J. Mol. Biol. 225*, 363-377.

Sweet, R. M. (1986) *Biopolymers 25*, 1565-1577.

Timasheff, S. N., & Nernardi, G. (1970) *Arch. Biochem. Biophys. 141*, 53-58.

Tinoco, I., Jr., Woody, R. W., & Bradley, D. F. (1963) *J. Chem. Phys. 38*, 1317-1324.

Vuilleumier, S., Sancho, J., Loewenthal, R., & Fersht, A. R. (1993) *Biochemistry 32*, 10303-10313.

Waterhous, D. V., & Johnson, W. C., Jr. (1994) *Biochemistry 33*, 2121-2128.

Wetlaufer, D. B. (1962) *Adv. Protein Chem. 17*, 375-378.

Williams, R. W., Chang, A., Juretic, D., & Loughran, S. (1987) *Biochim. Biophys. Acta. 916*, 200-204.

Woody, R. W. (1978) *Biopolymers, 17*, 1451-1467.

Woody, R. W., & Tinoco, I., Jr. (1967) *J. Chem. Phys. 46*, 4927-4945.

Yang, C. C., Chang, C. C., Hayashi, K., Suzuki, T., Ikeda, K., & Hamaguchi, K. (1968) *Biochim. Biophys. Acta 168*, 373-376.

Yang, J. T., Wu, C. C., & Martinez, H. M. (1986) *Methods Enzymol. 130*, 208-269.

Yi, T. M., & Lander, E. S. (1993) *J. Mol. Biol. 232*, 1117-1129.

Zhong, L., & Johnson, W. C., Jr. (1992) *Proc. Natl. Acad. Sci. USA. 89*, 4462-4465.

Zvelebil, M. J., Barton, G. J., Taylor, W. R., & Sternberg, M. J. E. (1987) *J. Mol. Biol. 195*, 957-961.

APPENDICES

Appendix A

COMPARISON OF BASE INCLINATIONS FOR LEFT-HANDED DNA AND RNA

Chartchai Krittanai, Hunseung Kang, Xiaokui Jin,

Jeannine H. Riazance-Lawrence and W. Curtis Johnson, Jr.

Department of Biochemistry and Biophysics

Oregon State University, Corvallis, Oregon 97331

Published in *Biospectroscopy* (1995) **1**, 247-254.

Abstract

The Z'-form of poly(dG-m^{5}dC)·poly(dG-m^{5}dC) in 85% ethanol/MOPS and the Z-form poly(rG-rC)·poly(rG-rC) in 4.8 M $NaClO_4$, 20% ethanol are studied by flow linear dichroism. Simultaneous analysis of the isotropic absorption and linear dichroism data yields the angle of inclination of the base normal relative to the helix axis, and the orientation of the axis around which the bases incline. For the Z'-form of poly(dG-m^{5}dC)·poly(dG-m^{5}dC) the inclination angles are 31.3° for guanine, and 27.9° for cytosine; for the Z-form of poly(rG-rC)·poly(rG-rC) they are 25.0° for guanine, and 23.3° for cytosine. The inclination angles for the Z' form DNA are similar to those for the Z-form, and the angles for Z-form RNA are somewhat smaller. The axes of inclination for guanine and cytosine are similar in all three left handed forms.

Introduction

The structure of nucleic acids has been extensively studied since Watson and Crick elucidated the helical structure of classical right-handed B-form DNA in the early 1950s. Much of the information on the conformation of nucleic acids comes from X-ray diffraction of fibers and crystals, nuclear magnetic resonance, and other spectroscopic methods. These studies have revealed that the secondary structure of DNA is polymorphic. The variant conformations can be classified by handedness, helical parameters, and backbone geometry into, for instance, the A-, B-, and Z-forms. The ability of a nucleic acid to adopt any form

is dependent to its sequence (Bram & Tougard, 1972; Leslie et al., 1980; Dickerson & Drew, 1981; Sarai et al., 1988), cation type, (Fuller & Wilkins, 1965; Bram & Tougard, 1972; Ivanov et al., 1973; Leslie et al., 1980; Taillander et al., 1984; Taboury & Taillander, 1985; Adam et al., 1986; Devarajan & Shafer, 1986; Riazance & Johnson, 1992; Harder & Johnson, 1990), temperature and solvent conditions (Girod & Johnson, 1973; Pohl, 1976; Riazance et al., 1987).

The left-handed Z-form of DNA has been found to exist in four variant forms in crystals, called Z, Z', Z_I and Z_{II} (Drew et al., 1980; Wang et al., 1981; Hall & Maestre, 1984; Saenger, 1984; Gessner et al., 1989; Zhong & Johnson, 1990). These Z-type structures have the common characteristics of a 12-fold double helix with axial rise per dinucleotide repeat unit of about 7.4 A° and a -60° rotation. The Z_I and Z_{II} forms observed for the d(GpC) hexamer are distinguished according to their coordination of phosphate to water or to magnesium ions (Wang et al., 1981; Gessner et al., 1989). For the d(CpG) tetramer the Z' variant is found to be different from the Z-form with sugar puckering C_1'-exo rather than C_3'-endo (Drew et al., 1980). In solution a variant of the Z-form DNA has been detected for poly(dG-dC)·poly(dG-dC) and poly(dG-m^5dC)·poly(dG-m^5dC), called the Z'-form. The Z'-form in solution has been related to the Z_{II}-form in crystals (Harder & Johnson, 1990). The transition from Z- to Z'-form is a function of divalent and alcohol concentration (Pohl, 1976; Hall & Maestre, 1984, Zhong & Johnson, 1990). These two forms in solution have distinct characteristics that can be observed by using circular dichroism (CD) spectroscopy.

For RNA, poly(rG-rC)·poly(rG-rC) is found to undergo a transition from the right-handed A-form in moderate salt to the left-handed Z-form in high salt solution (Hall & Maestre 1984; Davis et al., 1987). This Z-form RNA is confirmed by CD, NMR and Raman scattering (Cruz et al., 1986). Thus the A- to Z-form transition indicates that RNA is also polymorphic.

In this study, we investigate the orientation of the bases in the Z'-form DNA, poly(dG-m^5dC)·poly(dG-m^5dC), and in the Z-form RNA, poly(rG-rC)·poly(rG-rC), in solution by using flow linear dichroism (FLD) spectroscopy. Our objective is to compare the inclination angle and axis of inclination for the Z'-DNA and Z-RNA to the related polymers in the A- and Z-forms.

Materials and Methods

Sample preparation

Poly(dG-m^5dC)·poly(dG-m^5dC) was purchased from Pharmacia Biotech (lot no. 3047938021). 25 OD units of sample were dissolved in 3 ml of 1 mM MOPS, pH 7.0, and kept in the refrigerator overnight before making a dilution. The polynucleotide solution was dialyzed using SpectroPor tubing (1000 molecular weight cutoff) against 1 L 0.5 M NaCl, 10 mM EDTA, pH 8.0 for 24 hr to remove divalent impurities and then against 1 mM MOPS, pH 7.0 to yield a stock solution. To obtain the Z'-form, 99.5% pure ethanol (USI Chemical Co.) was added to the stock solution in a closed centrifuge tube using a 1 ml syringe to minimize the evaporation of ethanol during the mixing process. The mixing

was begun by adding ethanol dropwise into the tube to get a concentration of 30 % v/v, and 1/2 hour was allowed for the slow transition from B to Z form. The mixing then proceeded until reaching a concentration of 85% v/v ethanol/MOPS. The sample was confirmed to be in Z'-form by circular dichroism (CD) measurement.

Poly(rG-rC)·poly(rG-rC) is not commercially available. It was synthesized by transcription catalyzed with T7 RNA polymerase (Jin & Johnson, 1995). The optimized reaction conditions were 40 mM Tris buffer, pH 8.0, 15 mM $MgCl_2$, 5 mM dithiothritol, 0.5 mg/ml BSA (nuclease free), 2 mM rGTP, 2 mM rCTP, 0.1 mM poly(dI-dC)·poly(dI-dC), and 2.8 units/µl T7 RNA polymerase in a total reaction volume of 1 ml. Reaction time was about one hour at 37 °C, and the mixture was then heated to 70 °C for about 10 minutes to melt synthesized RNA from the template. After cooling to ambient temperature, DNase I was added at the concentration of 3 units/µl to digest DNA in the presence of 5 mM $CaCl_2$ at 25 °C for one hour. The protein was removed with two phenol extractions and three chloroform:isoamyl alcohol (24:1) extractions. The sample was dialyzed three times against 20 ml of autoclaved 10 mM sodium phosphate buffer, pH 7.0, to remove the digested DNA and change the solvent system. The synthesis yield was about 1.25 mg of poly(rG-rC)· poly(rG-rC) per ml of reaction mixture with the length of 100 to 500 base pairs.

To obtain the Z-form, poly(rG-rC)·poly(rG-rC) was dialyzed against 4.8 M sodium perchlorate, 20 % ethanol. The final concentration of the sample was

about 5.5 OD units/ml. The CD spectrum confirmed that the sample was indeed in the Z-form in this solvent system.

Spectral measurements

Measurement of linear dichroism (LD), CD and isotropic absorption (A) have been described in detail previously (Causeley & Johnson, 1982; Edmondson & Johnson, 1985). Briefly, the isotropic absorption spectra were recorded at ambient temperature on a Cary 15 spectrometer that has been interfaced with a computer. Data were collected every 1 nm for the sample in a 200 μm cylindrical cell for poly(rG-rC)·poly(rG-rC) or in a 1 mm square cell (because of limited solubility) for poly(dG-m^5dC)·poly(dG-m^5dC). The instrument was equipped with a Hamamatsu Super Quiet deuterium lamp and flushed with nitrogen 30 minutes before and during the measurements. The CD spectra were recorded on a JASCO J720 spectrometer using the cells with the same path length as for the isotropic absorption. The CD spectrum was monitored before, during and after recording each LD spectrum, to ensure that the form of the sample remained unchanged during the entire experiment. The CD spectrometer was calibrated with (+) -10-camphorsulfonic acid (CSA), giving the ratio of $\Delta\varepsilon$ at 290.5 nm : $\Delta\varepsilon$ at 192.5 nm = 2.08.

Flow linear dichroism (FLD) spectra were recorded on a vacuum UV LD spectrometer with a stress plate modulator as a half-wave retarder. The signal was detected by a phase-sensitive lock-in amplifier connected to a computer. The flow system was a stainless-steel Micro Flow-Thro cell (from Barnes) with

two quartz windows separated by a 30 μm Teflon spacer. The sample was aligned by pumping the sample through the cell at about 4 ml/minute by a Master-Flex pump (from Cole-Parmer) using 0.8 mm silicone tubing. The flow system required a sample volume of about 600 μl. A single quartz plate tilted at 15° with respect to incident light was used to calibrate the LD spectrometer, as described by Norden and Seth (1985). Calibration were carried out at 0°, 90°, 180° and 270° rotation of quartz plate relative to the linear polarization. The wavelength dependence of this calibration was within 2% of the theoretical curve (Norden & Seth, 1985), which was used to determine the LD scale factor in LD/volt. Stability of the instrument with time was also 2%.

Data analysis

In general, LD is the difference in absorption for linearly polarized light passing through the sample parallel and perpendicular to the alignment axis.

$$LD = A_{\parallel} - A_{\perp}$$

For a polynucleotide, the helical axis is assumed to cause alignment. Norden (Norden, 1978) has showed that the LD can be related to the orientation of bases inside the helix by :

$$LD(\lambda) = 3S \Sigma \varepsilon_{ij}(\lambda) [3 \sin^2 \alpha_i \sin^2(\chi_j - \delta_{ij}] / 2$$

where LD at wavelength λ is the sum of the contributions from each transition dipole (*j*) of each base (*i*) weighted by the extinction coefficient (ε) for its isotropic absorption. **S** is an orientation factor ($0 \leq S \leq 1$) of the molecule, α_i is the angle between normal of base *i* and the helical axis (the result of twist, tilt, roll and buckle), χ_j is the orientation of the inclination axis relative to the C4-C5 bond for purines or the C6-C5 bond for pyrimidines, and δ_{ij} is the orientation of transition dipole *j* of base *i* measured in the same way. All the angles are shown in Figure 4.1. Only the shape of the LD is used for our analysis, and we do not have to extrapolate our LD data to perfect orientation. We measure the LD for many transitions, and since the direction of the transition dipoles are known, we can analyze the wavelength dependence of $A(\lambda)$ and $LD(\lambda)$ for α_i and χ_j. Any reasonable tertiary structure of RNA and DNA that would affect the ability to extrapolate to a perfect alignment is contained in the orientation factor S, but does not affect the wavelength dependence of the LD.

The data are analyzed with a sophisticated computer program that uses the Levenberg-Marquardt algorithm and minimizes the variance of all variables, rather than fitting our imperfect data perfectly (Chou & Johnson, 1993). The absorption spectra of the bases have been investigated in detail by Clark and coworkers (Clark, 1977; Zaloudek et al., 1985; Clark, 1994), who determined the number of bands for each absorption and their transition dipole directions. The individual bands are represented as log-normal functions with band position (μ), area (ζ), band width(σ), and band skewness (ρ). The absorption spectrum of the

monomers has been resolved by the algorithm for the individual band parameters (Chou & Johnson, 1993), and these parameters are given in Table 1. These monomer parameters serve as the initial input for analysis of data. The algorithm simultaneously decomposes the absorption and the linear dichroism into individual bands to generate the α and χ angles. It is possible to fit the data perfectly, but the data are imperfect. All the parameters are varied, and we look for the iteration that minimizes the variance of these parameters and also gives a reasonable fit with stability, rather than looking for an unreasonably good fit with unstable parameters.

We do not know the absolute values of the variances for this non-linear system. To estimate the error, we randomly changed each transition dipole direction in the range of ± 10 degree, and fit the data in this manner for 100 times. The standard deviation for these 100 fittings is the estimated error in the Tables.

Figure 4.1: Diagram showing the angles: base inclination, α; inclination axis, χ; transition dipole direction, δ.

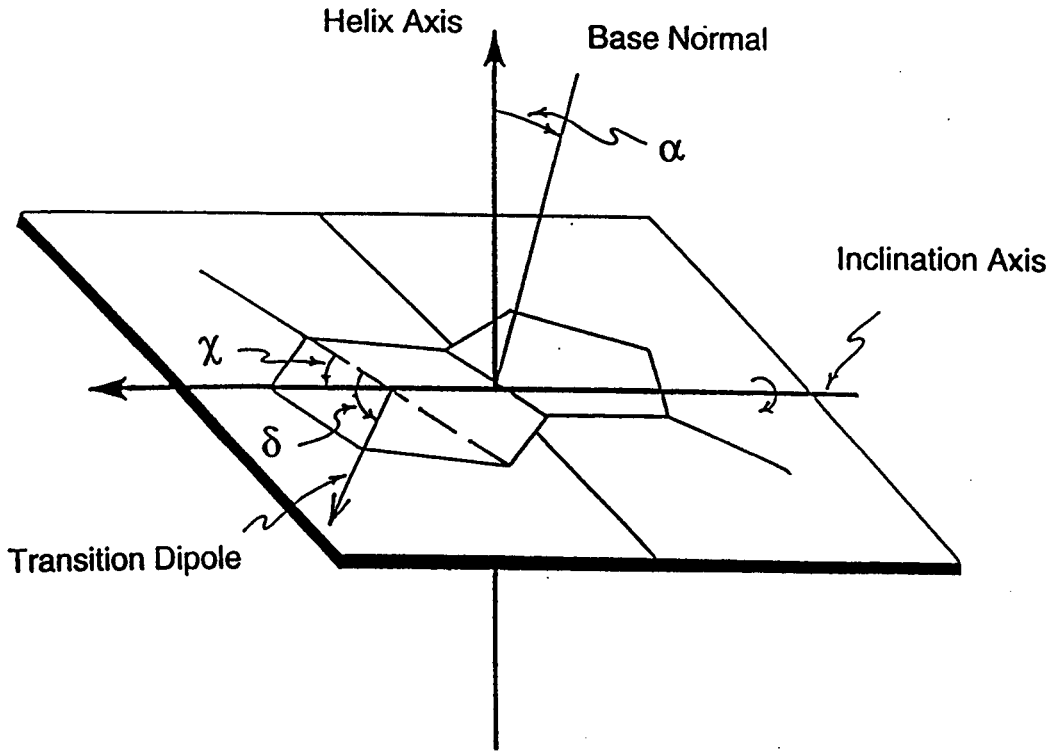


Table 1: Spectral decomposition of monomer absorption.

Base	μ (nm)[a]	ζ x10^{-3}[b]	σ(nm)[c]	ρ[d]	δ (deg)[e]
Guanine	274.5	288.7	16.7	1.50	-4
	248.5	309.5	13.9	1.10	-75
	198.8	471.2	11.6	1.03	-71
	183.2	449.4	11.6	1.50	41
Cytosine	269.0	301.2	15.3	1.12	6
	228.1	319.2	19.8	1.31	-35
	211.6	86.8	7.1	1.00	76
	196.5	403.1	9.9	1.43	86
	170.1	94.0	12.4	1.03	0

[a] wavelength of decomposed bands

[b] band intensity

[c] half bandwidth at half height

[d] band skewness

[e] transition dipole direction (refs.31-33)

Results and Discussion

The CD for Z-DNA, Z'-DNA and Z-RNA are depicted in Figure 2. Our samples of poly(rG-rC)·poly(rG-rC) in 4.8 M $NaClO_4$, 20% ethanol and poly(dG-m^5dC)·poly(dG-m^5dC) in 85% ethanol gave the expected CD for the left handed Z-form RNA and Z'-form DNA. CD confirmed that we are making A and LD measurements on the correct form of the samples throughout the experiment. The LD and A for poly(rG-rC)·poly(rG-rC) and poly(dG-m^5dC)·poly(dG-m^5dC) are shown in Figures 3 and 4. The LD spectrum has been normalized to the same area as A and its sign is reversed for ease in comparison. Although we analyzed LD and A simultaneously, the reduced linear dichroism, L' = LD/A is also plotted. Since the L' varies with wavelength, the bases in both samples are not perpendicular to the helical axis of the polymers. We start the fitting using the band parameters from monomers in Table 1, and let the program run until we find the iteration with the lowest variance and a reasonable fit. The parameters defining the component bands obtained from the analyses are given in Tables 2 and 3.

The transition dipole directions we used in previous work were those measured by Clark and coworkers (Clark, 1977, 1994). The bases are stacked in the crystals, and resonance (exciton) interaction can mix transitions between bases and change the transition dipole directions. Base stacking in the polymers can also change the transition dipole directions. An assumption in our previous work is that this effect is small, or at least the same in the polymer as it is in crystal. Recently, Clark's laboratory has considered resonance coupling in their

crystals of guanine, and corrected their transition dipoles to the true directions for noninteracting bases (Zaloudek et al., 1985). We used these slightly different transition dipole directions for G, and found that the α and χ angles obtained from this analysis are almost the same. This result indicates that the angles are not sensitive to reasonable uncertainty in transition dipole directions. In this report, we determine the standard deviation for the α and χ angles by repeating each fitting for 100 times with the transition dipole direction randomly varied within ±10 degree.

Z-form RNA

The absorption spectra are more difficult to measure than the LD, and the A of Z-form poly(rG-rC)·poly(rG-rC) could be obtained only to 200 nm. Therefore, its spectrum was analyzed utilizing the first 3 bands from G and the first 4 bands from C giving the spectra decomposition in Figure 5. Although the L' spectrum shown in Figure 3 is not particularly varied, the analysis predicts inclination angles of 25° for rG and 23° for rC. The angles are compared with related polymers in Table 4. Guanine in our Z-form RNA has essentially the same inclination as found in both A-form poly(rG-rC)·poly(rG-rC) (25°) and in Z-form poly(dG-dC)·poly(dG-dC) (27°). Furthermore, its inclination axis has almost the same orientation for both Z and Z'-form DNA. Cytosine has a lower inclination from the other related polymers, but at 23° is about the same as the

guanine base (25°). The inclination axis for C is about the same as those of the Z-form and Z' form DNA.

Z'-form DNA

The poly(dG-m^5dC)·poly(dG-m^5dC) in this work adopted the Z-form in 30% v/v of ethanol solution and changed to the Z'-form at 85%. The transition from B to Z and then Z' form was monitored by CD. The CD spectra of Z and Z' are plotted together in Figure 2. The Z-form poly(dG-m^5dC)·poly(dG-m^5dC) in 30% ethanol has a characteristic CD with negative band at about 292 nm. This negative band disappears when the polymer adopts the Z'-form in 85% ethanol, giving a positive band at about 279 nm. The characteristics of the Z'-form are similar to those of the Z-form RNA. All the Z-type polymers shown in the figure have an intense band in the far UV region at about 195-196 nm, which is a hallmark for all the left-handed helices (Sutherland et al., 1981).

The LD and A spectra for the Z'-form were analyzed only to 188 nm because of the limitation of the absorption measurement in 85% ethanol solution. Both LD and A spectra are normalized and plotted with L' in Figure 4. We performed the fitting with 4 bands from G and 4 bands from C showing the spectral decomposition in Figure 6. The results from the analysis for the Z'-form give a large inclination of 31° for dG and 28° for dC. Comparing to other polymers in Table 3, the α and χ angles are quite closed to those of Z-form poly(dG-dC)·poly(dG-dC) , but with the angles for each type of base reversed. The inclination axes are found to be similar for all the Z and Z'-form polymers

which are 130-140° for G and 7-20° for C. The comparison among the related polymers in the table also indicates that inclination angles in RNA is smaller than those in DNA.

If we compare the shape of the CD for nucleic acids, their intense CD bands located in the far UV region which indicate the helical handedness, are similar for each of the left-handed or right-handed polymers. When the buffer conditions surrounding these left-handed polymers are changed from more to less hydrated, polymorphism occurs and the characteristic CD is changed. This change in CD is found only in the near UV region leaving the far UV band denoting handedness the same. The transition from Z- to Z'-form for poly(dG-m^5dC)·poly(dG-m^5dC) occurs by changing the buffer conditions from more hydrated (30% ethanol) to less hydrated (85% ethanol) giving a CD for the dehydrated Z'-DNA similar to the CD of Z-RNA. The similarity of CD for the left-handed DNA in the dehydrated form to the CD of the left-handed RNA is very interesting. This phenomenon is founded not only for the left-handed polymers, but also for the right-handed polymers. The B to A transition of the right-handed polymers yields a dehydrated A-form with a CD similar to the CD of the right-handed A-RNA. This indicates that the structure of DNA in the dehydrated environment is similar to the structure of RNA, for both right and left handed structures.

Figure 4.2: The CD of poly(dG-m^{5}dC)·poly(dG-m^{5}dC) as the Z-form in 30% ethanol (——), poly(dG-m^{5}dC)·poly(dG-m^{5}dC) as the Z'-form in 85% ethanol (----), and poly(rG-rC)·poly(rG-rC) as the Z-form in 4.8 $NaClO_4$, 20% ethanol (- - -).

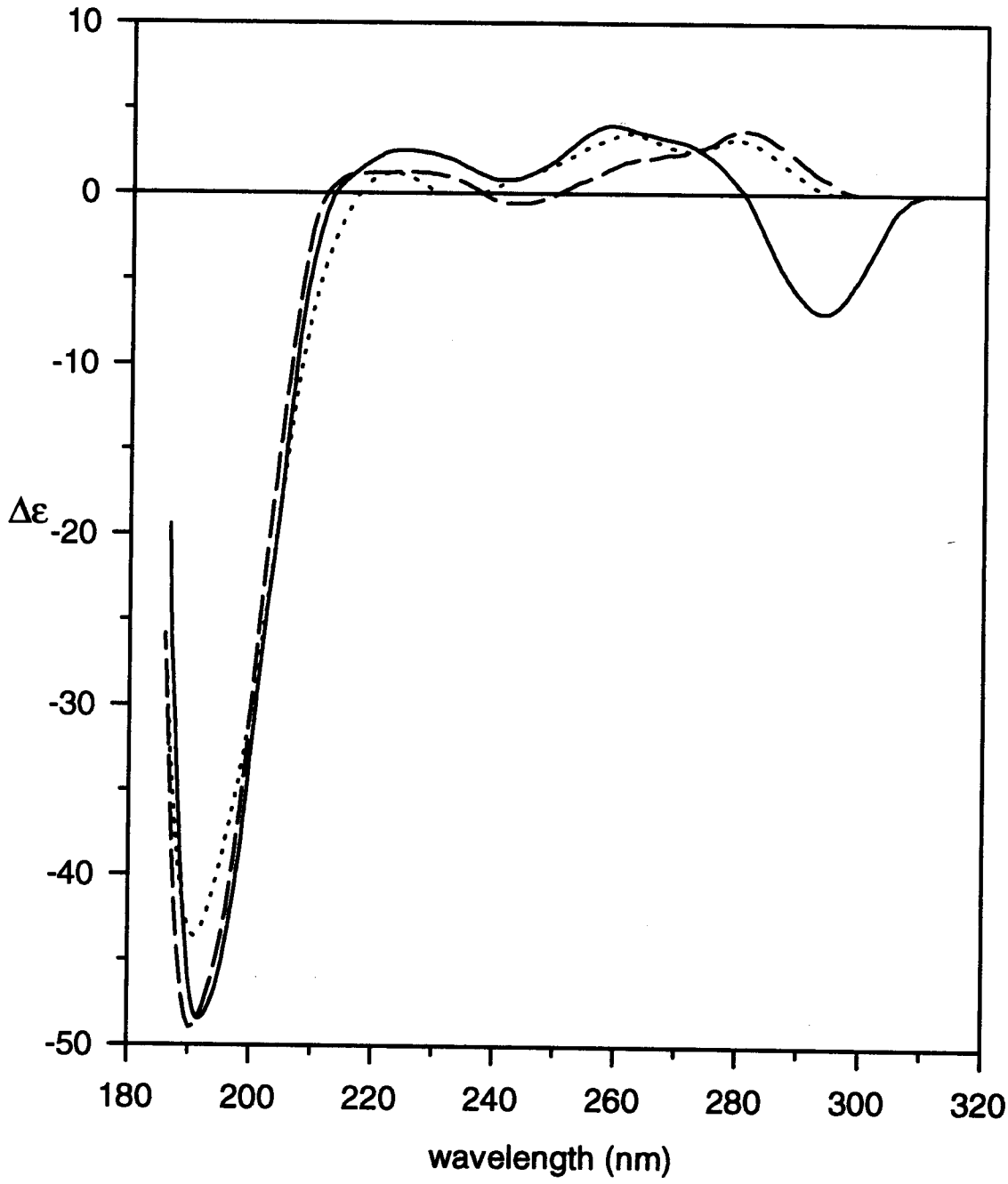


Figure 4.3: Poly(rG-rC)·poly(rG-rC) as the Z-form in 4.8 $NaClO_4$, 20% ethanol : normalized A (—), normalized LD with sign reversed (----), and L' (- - -).

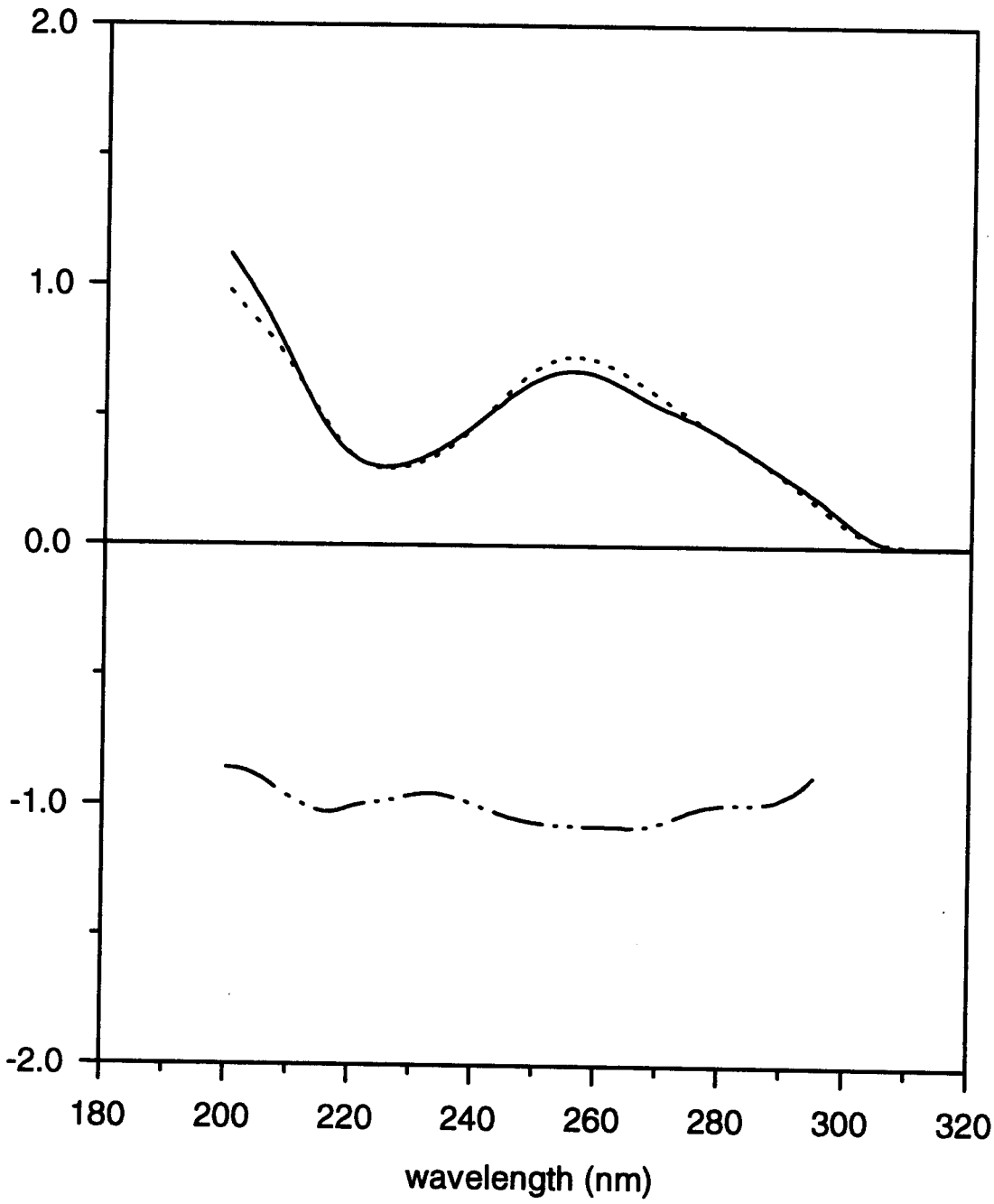


Figure 4.4: Poly(dG-m^{5}dC)·poly(dG-m^{5}dC) as the Z'-form in 85% ethanol : normalized A (——), normalized LD with sign reversed (----), and L' (- - -).

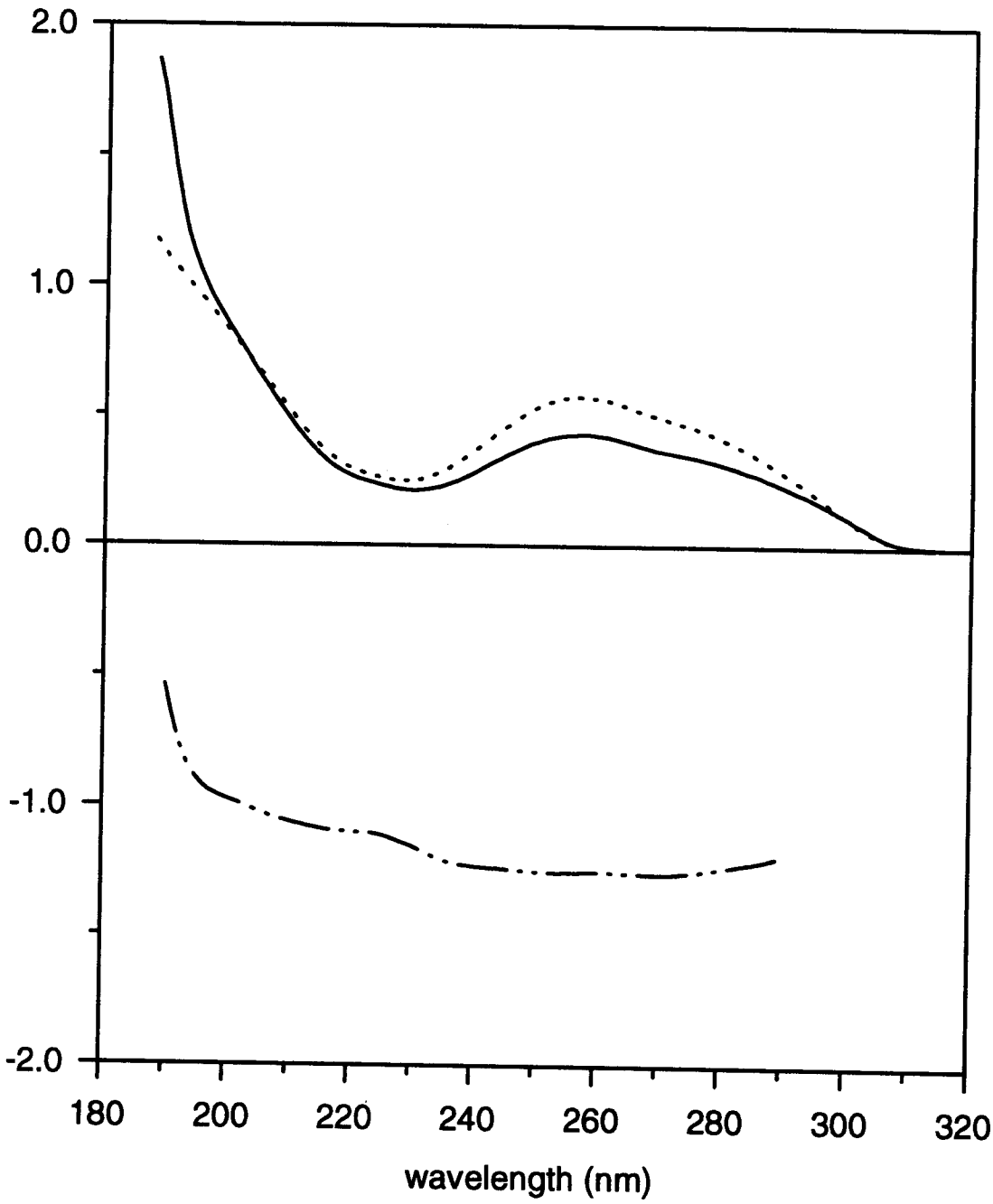


Table 2: Spectral decomposition of absorption and LD for Z-form poly(rG-rC)·poly(rG-rC).

Base	μ(nm)	$\zeta x10^{-3}$	σ (nm)	ρ	α (deg)	χ (deg)
rG	282.7	107.7	17.5	1.50	25.0±3.1	131.2±6.7
	249.4	113.6	12.9	1.13		
	203.2	151.1	15.3	1.45		
rC	265.1	92.5	14.1	1.00	23.3±1.0	7.3±4.7
	222.2	85.7	18.5	1.20		
	205.7	26.3	7.9	1.42		
	193.1	140.5	9.3	1.10		

Table 3: Spectral decomposition of absorption and LD for Z'-form poly(dG-m^5dC)·poly(dG-m^5dC).

Base	μ(nm)	$\zeta x10^{-3}$	σ (nm)	ρ	α (deg)	χ (deg)
dG	284.2	80	19.3	1.50	31.3±0.5	135.2±3.6
	250.7	80	14.9	1.07		
	193.4	110	14.1	1.01		
	183.4	220	7.4	1.40		
dC	270.0	70	19.6	1.24	27.9±1.7	12.7±4.1
	214.6	70	15.4	1.11		
	206.2	20	6.3	1.03		
	191.6	130	9.3	1.12		

Figure 4.5A: Spectral decomposition of A for the Z-form of poly(rG-rC)·poly(rG-rC): measured spectrum (○), fitted spectrum (—), guanine bands (----), and cytosine bands (- - -).

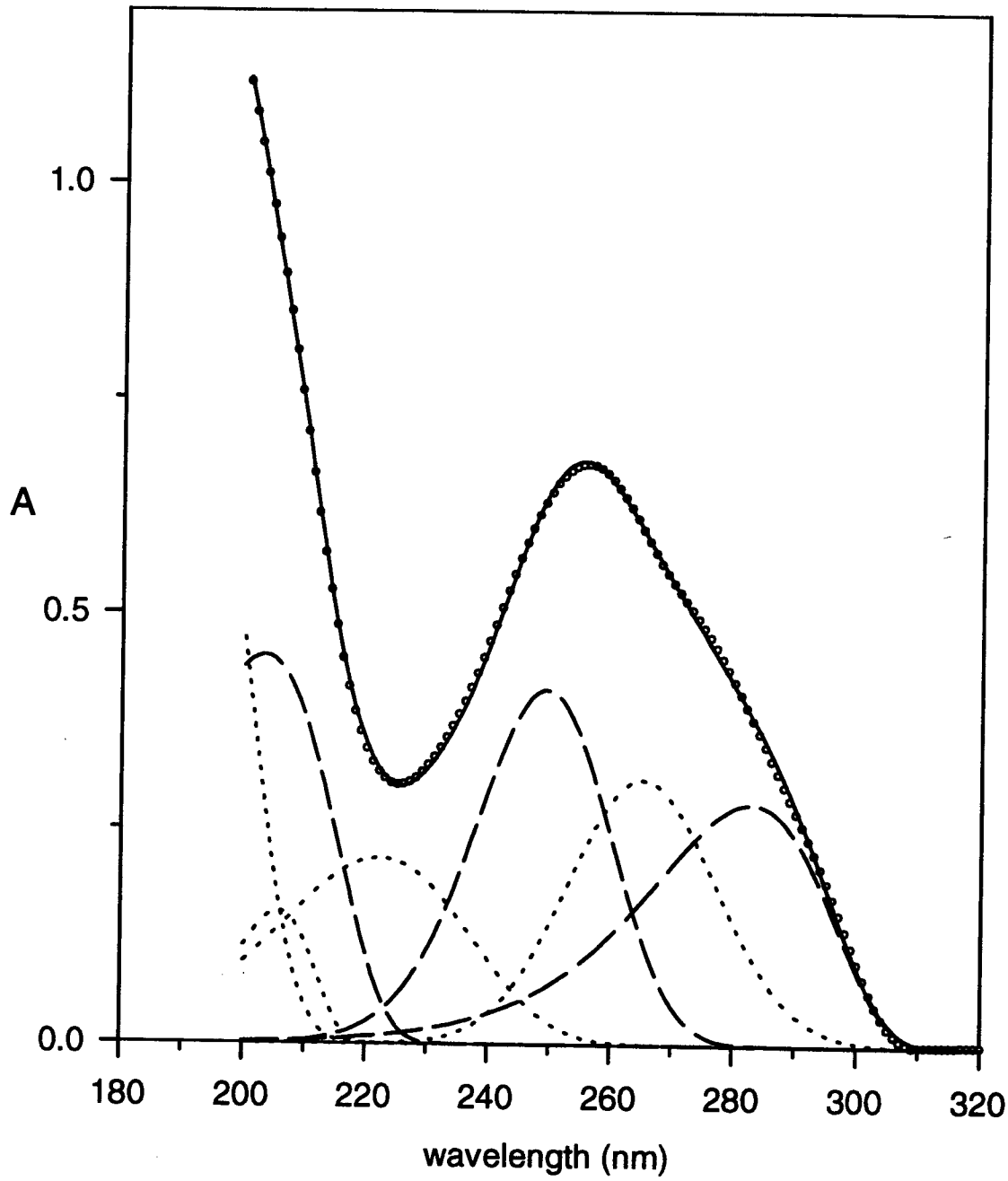


Figure 4.5B: Spectral decomposition of LD for the Z-form of poly(rG-rC)·poly(rG-rC): measured spectrum (○), fitted spectrum (—), guanine bands (----), and cytosine bands (- - -).

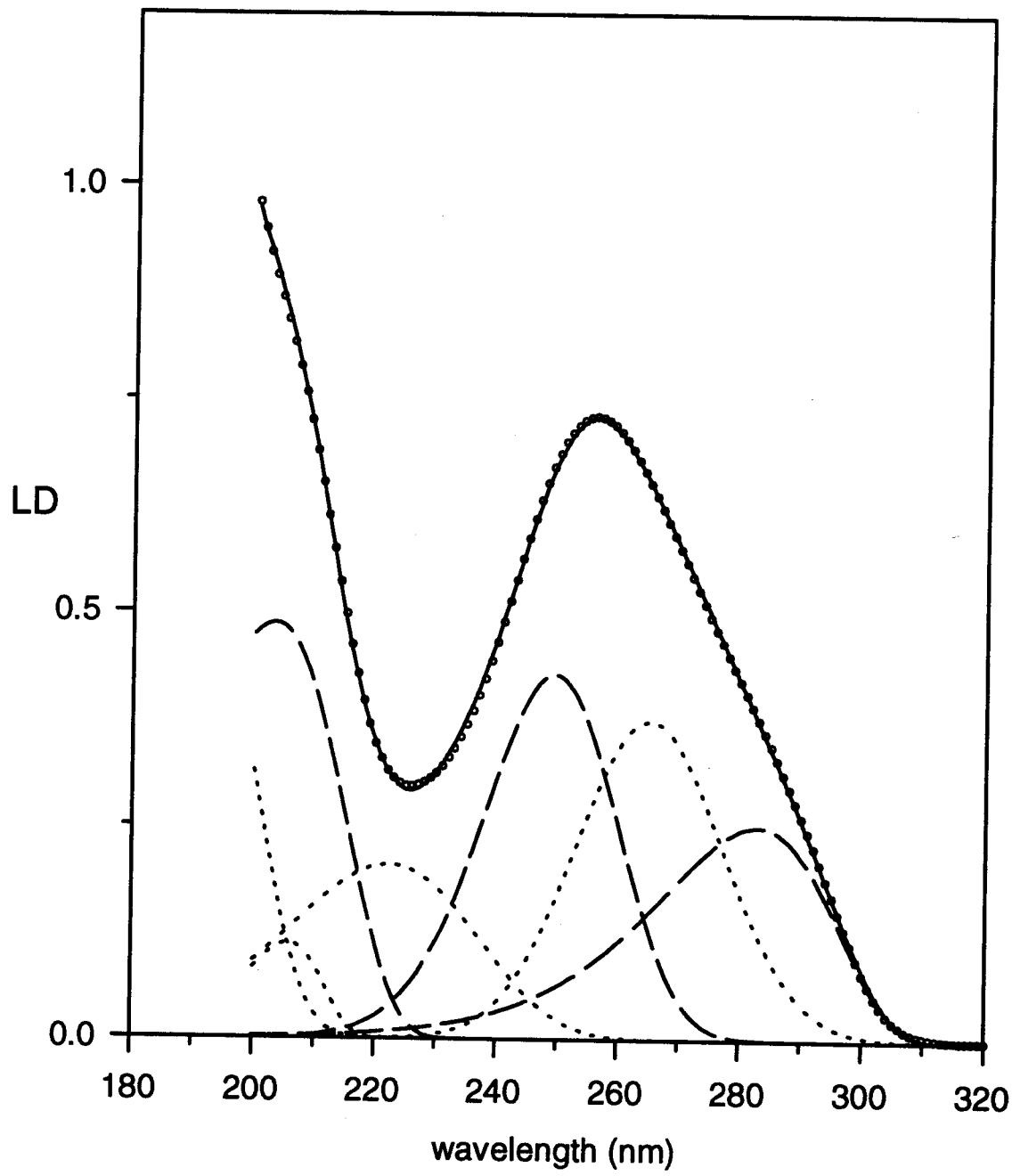


Table 4: Comparison of base inclination in related polymers as revealed by flow LD.

Base	Polymer	α (deg)	χ (deg)	Ref.
G	A-form poly(rG-rC)·poly(rG-rC)	25.3±1.3	48.8±5.0	*(i)*
	Z-form poly(dG-dC)·poly(dG-dC)	27.1±1.1	137.6±3.6	*(ii)*
	Z-form poly(rG-rC)·poly(rG-rC)	25.0±3.1	131.2±6.7	*this work*
	Z'-form poly(dG-m^5dC)·poly(dG-m^5dC)	31.3±0.5	135.2±3.6	*this work*
C	A-form poly(rG-rC)·poly(rG-rC)	33.4±0.8	64.7±2.7	*(i)*
	Z-form poly(dG-dC)·poly(dG-dC)	32.1±1.7	21.5±2.8	*(ii)*
	Z-form poly(rG-rC)·poly(rG-rC)	23.3±1.0	7.3±4.7	*this work*
	Z'-form poly(dG-m^5dC)·poly(dG-m^5dC)	27.9±1.7	12.7±4.7	*this work*

(i) Jin & Johnson, 1995.

(ii) Chou & Johnson, 1993.

Figure 4.6A: Spectral decomposition of A for the Z'-form of poly(dG-m^5dC)·poly(dG-m^5dC): measured spectrum (○), fitted spectrum (—), guanine bands (----), and cytosine bands (- - -).

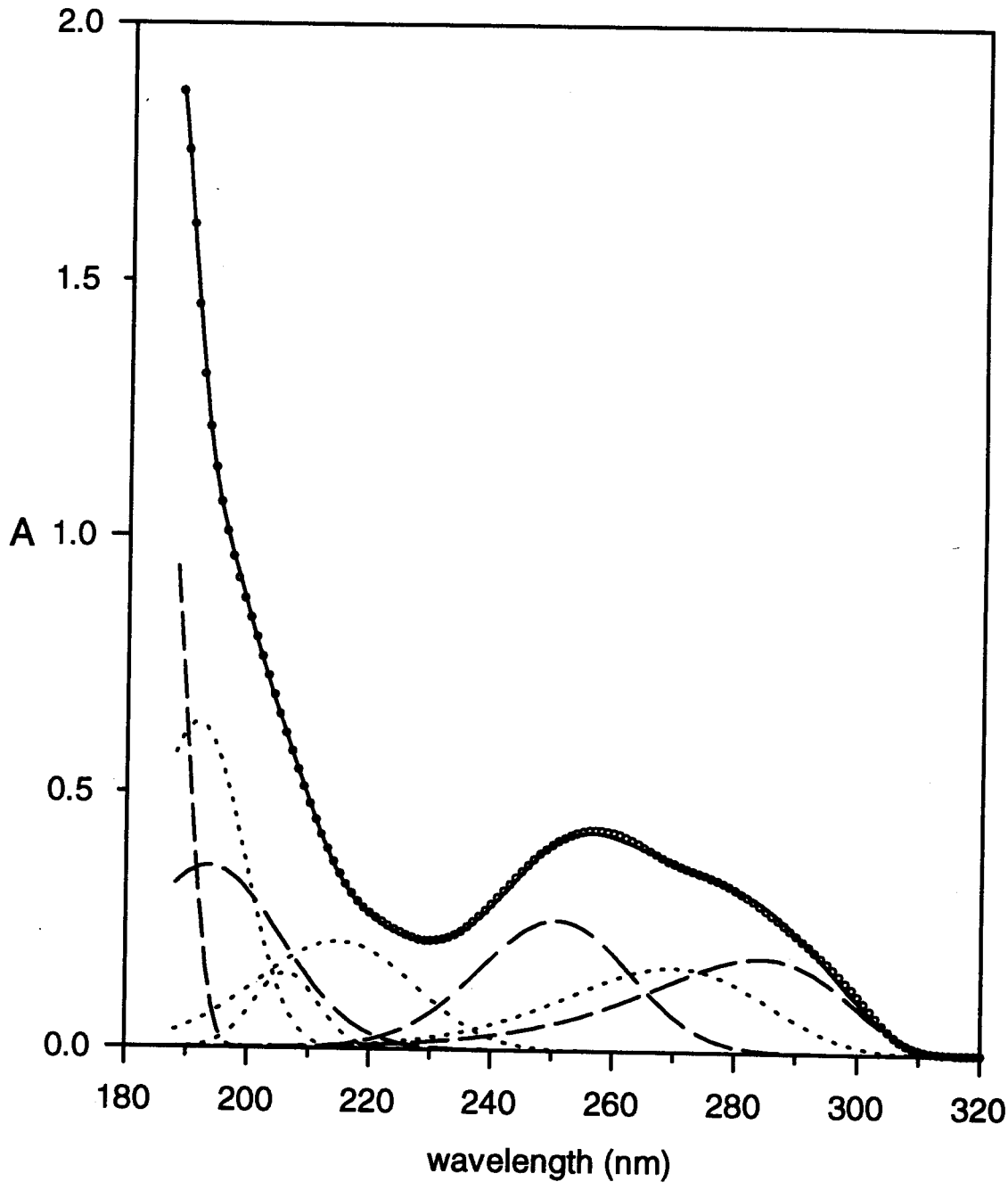
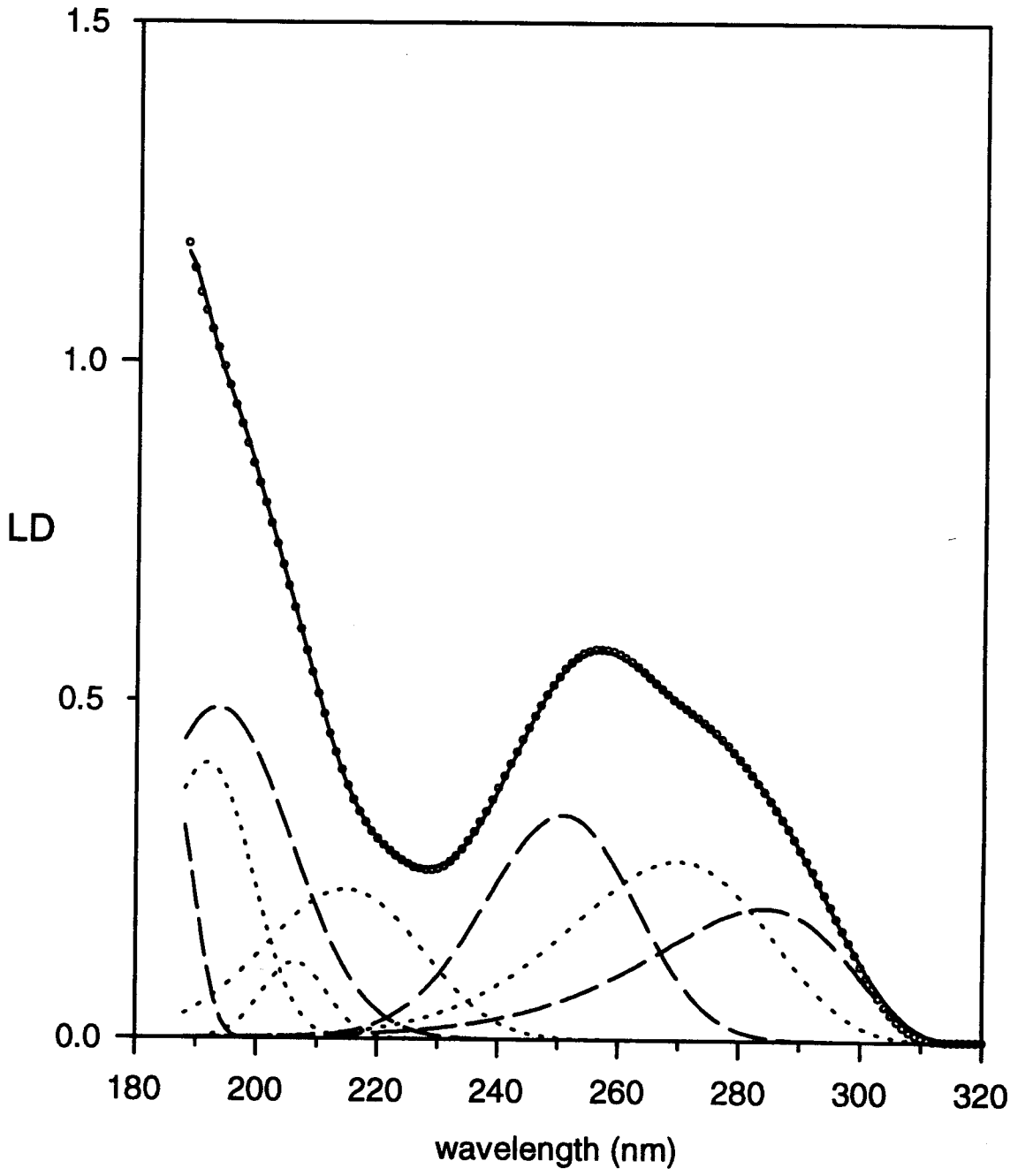


Figure 4.6B: Spectral decomposition of LD for the Z'-form of poly(dG-m^5dC)·poly(dG-m^5dC): measured spectrum (○), fitted spectrum (——), guanine bands (----), and cytosine bands (- - -).



Acknowledgment

This work was supported by PHS grant GM 43133 and GM 21479 from the National Institutes of Health.

References

Adam, S., Bourtayre, P., Liquier, J., & Taillandier, E. (1986) *Nucleic Acids Res. 14*, 3501-3513.

Bram, S., & Tougard, P. (1972) *Natural New Biol. 239*, 128-131.

Causley, G. C., & Johnson. W. C. Jr. (1982) *Biopolymers 21*, 1763-1780.

Chou, P. J., & Johnson, W. C. Jr. (1993) *J. Am. Chem. Soc. 115*, 1205-1214.

Clark, L. B. (1977) *J. Am. Chem. Soc. 99*, 3934-3938.

Clark, L. B. (1994) *J. Am. Chem. Soc. 116*, 5265-5270.

Cruz, P., Hall, K., Puglisi, J., Davis, P.,Hardin, C. C.,Trulson, M.O., Mathies, R. A., Tinoco, I. Jr., Johnson. W. C. Jr., & Neilson, T. (1986) in *Biomolecular Stereodynamics IV*, Adenine Press, pp. 179-200.

Davis, P. W., Hall, K., Cruz, P., Tinoco, I. Jr., & Neilson, T. (1987) *Nucleic Acids Res. 14*, 1279-1291.

Devarajan. S., & Shafer, R. H. (1986) *Nucleic Acids Res. 14*, 5099-5109.

Dickerson, R. E., & Drew, H. (1981) *J. Mol. Biol. 149*, 761-786.

Drew, H., Takano, T., Tanaka, S., Itakura, K., & Dickerson, R. E. (1980) *Nature 286*, 567-573.

Edmondson, S. P., & Johnson, W. C. Jr. (1985) *Biopolymers 24*, 825-841.

Fuller, W., & Wilkins, M. H. F. (1965) *J. Mol. Biol. 12*, 60-80.

Gessner, R.V., Frederick, C. A., Quigley,G. J., Rich, A., & Wang, A. H. (1989) *J. Biol. Chem. 264*, 7921-7935.

Girod, J. C., Johnson, W. C. Jr., Huntington, S. K., & Maestre, M. F. (1973) *Biochemistry 12*, 5092-5096.

Hall, K., Cruz, P., Tinoco, I. Jr., Jovin, T. M., & van de Sande, J. H. (1984) *Nature 311*, 584, (1984).

Hall, K. B., & Maestre, M. F. (1984) *Biopolymers 23*, 2127-2139.

Harder, M. E., & Johnson, W. C. Jr. (1990) *Nucleic Acids Res. 18*, 2141-2148.

Ivanov, V. I., Minchenkova, L. E., Schyolkina, A. K., & Poletayev, A. I. (1973) *Biopolymers 12*, 89-110.

Jin, X., & Johnson. W. C. Jr. (1995) *Biopolymers* 36, 313-322.

Leslie, A. G. W., Arnott, S., Chandrasekaran, R., & Ratliff, R. L. (1980) *J. Mol. Biol. 143*, 49-72.

Norden, B. (1978) *Appl. Spectros. Rev. 14*, 157-248.

Norden, B., & Seth, S. (1985) *Appl. Spectros. 39*, 647-655.

Pohl, F. M. (1976) *Nature 260*, 365-366.

Riazance, J. H., Johnson, W. C. Jr., McIntosh. L. P., & Jovin, T. M. (1987) *Nucleic Acids Res. 15*, 7627-7636.

Riazance, J. H., & Johnson, W. C. Jr. (1992) *Biopolymers 32*, 271-276.

Saenger, W (1984) in *Principles of Nucleic Acid Structure*, Chapter 12, ed. By Cantor, C. R., Springer-Verlag, New York, pp. 283-297.

Sarai, A., Mazur, J., Nussinov, R., & Jernigan, R. L. (1988) *Biochemistry 27*, 8498-8502.

Sutherland, J. C., Griffin, K. P., Keck, P. C., & Takacs, P. Z. (1981) *Proc. Natl. Acad. Sci. 78*, 4801-4804.

Taboury, J. A., & Taillandier, E. (1985*) Nucleic Acids Res. 13*, 4469-4483.

Taillandier, E., Taboury, J. A., Adam, S., & Liquier, J. (1984) *Biochemistry 23*, 5703-5706.

Wang, A. H., Quigley, G. J., Kolpak,F. J., Van der Marel, G., Van Boom, J. H., & Rich, A. (1981) *Science 211*,171-176.

Zaloudek, F., Novros, J. S., & Clark, L. B. (1985) *J. Am. Chem. Soc. 107*, 7344-7351.

Zhong, L., & Johnson, W. C. Jr. (1990) *Biopolymers 30*, 821-828.

Appendix B

Peptide Synthesis

Introduction

This protocol procedure is a modification of the original solid phase peptide synthesis by Merrifield (1963) in which the solid support resin is sealed in a permeable polypropylene bag. The mesh bag allows diffusion of the amino acid in solution through without loss of the solid support and the growing peptide chain. The so-called T-bag method was original described by Houghten et al. (1986) as a simple rapid method for Simultaneous Multiple Peptide Synthesis (SMPS). It was originally established by BOC chemistry, using tert-butyloxycarbonyl (tBOC) protecting group on the amino terminal of the growing peptide chain. Although tBOC protected amino acids are inexpensive, the cleaving of the individual amino acid blocking group and the peptide product use trifluoroacetic acid (TFA) and hydrofluoric (HF), respectively. These highly toxic chemicals and their waste make it more difficult for handling.

We adapted SMPS using the 9-fluorenylmethoxycarbonyl (FMOC) protected amino acids as described by Beck-Sickinger et al.(1991). The method was slightly modified in terms of the number of solvents used and the number of washes by different solvents during several steps. The outline of our final modification was mostly similar to a procedure published by Alewood et al. (1990), but we use only a single solvent throughout the procedure.

Materials

Synthesis bag

The design of the polypropylene bag is originally described by Houghten (1986). The original method described a 20 x 20 mm bag containing 100 mg resin; our synthesis use a 40 x 40 mm bag with 200 mg resin. The bag is made from propyltex polypropylene monofilament screening, with about 160 mesh/inch (74 μm pore size). It is cut into a 90 x 45 mm patch, folded into half length, and sealed on both edges using an impulse heat sealer made for sealing polyethylene bags. The bag is carefully filled with 200 mg solid resin and then the opening is sealed. The solid resin used in our experiment is called a Rink resin, containing 4-(2',4'-dimethoxyphenyl-FMOC-aminomethyl)-phenoxy on 1% cross-linked divinylbezene-styrene on 1% cross-linked divinylbenzene-styrene. The activity of resin is 0.35 meq/g.

The bag containing solid resin needs to be checked to make sure that it will remain sealed under the shaking used during the synthesis. The test is performed by placing the bag inside a polypropylene bottle containing 10 ml of dimethylformaide (DMF), closing the cap, and shaking vigorously on a shaker for 2-3 minutes. The DMF is then decanted into a small glass breaker for a visual inspection. If there is no leaking resin particle found in the DMF, the bag is dried under the hood before use. If there is resin found in the DMF, discard the DMF and retest the bag again with another 10 ml of fresh DMF. The second test can

confirm whether the bag is actually leaking, or the resin seen in the first test comes from spills on the outside of the bag.

Amino Acids

Amino acids are purchased from Bachem California. The amino terminals of these amino acids are protected by 9-fluorenylmethoxycarbonyl (FMOC), and the active side-chains are protected by different blocking group. The following are the amino acids used for our synthesis:

N-FMOC-Glycine	N-α -FMOC-L-Tryptophan
N-FMOC-L-Alanine	N-FMOC-Proline
N-FMOC-L-Valine	N-FMOC-L-Aspartic acid-β-t-Butyl ester
N-FMOC-L-Isoleucine	N-FMOC-L-Glutamic acid-γ-t-Butyl ester
N-FMOC-L-Leucine	N-α-FMOC-N-β-Trityl-L-Asparagine
N-FMOC-L-Methionine	N-α-FMOC-N-γ-Trityl-L-Glutamine
N-FMOC-L-Phenylalanine	N-α-FMOC-N-im-Trityl-L-Histidine
N-FMOC-O-t-Butyl-L-Tyrosine	N-α-FMOC-N-ε-Boc-L-Lysine
N-FMOC-O-t-Butyl-L-Serine	N-FMOC-S-Trityl-L-Cysteine
N-FMOC-O-t-Butyl-L-Threonine	N-α-FMOC-N-mtr*-L-Arginine
	N-α-FMOC-N-pmc**-L-Arginine

*Mtr: 4-Methoxy-2,3,6-trimethylbenzenesulfonyl

** Pmc: 2,2,5,7,8-Pentamethylchroman-6-sulfonyl

The amount of the amino acid added into each synthesis cycle should be 4 to 10 times excess to the molar activity of solid resin.

Coupling Reagent (TBTU/HOBT)

A mixed reagents of 2-(1H-Benzotriaole-1-yl)-1,1,3,3- Tetramethyluronium tetrafluoroborate (TBTU) and 1-Hydroxbenzotriaole hydrate (HOBT) is made up in DMF. The concentration of TBTU and HOBT in solution is calculated to get a molar ratio of 1:0.95:1 for HOBT:TBTU:amino acid added, with 10 ml volume of mixture per one synthesis bag.

Deblocking Reagent

A deblocking reagent is prepared from 30% (v/v) piperidine in DMF. One should avoid contact with piperidine and work under a fume hood.

Inducing Reagent

A 1 molar solution of N,N-Diisopropylethylamine (DIPEA) is prepared in DMF. The amount of this reagent added is calculated to get a molar ratio of 1.5:1 for DIPEA:amino acid.

Synthesis Procedures

The synthesis is performed by growing the peptide chain by one residue from C- to N- terminal for each synthesis cycle. All the reactions for removing the FMOC protecting group, coupling with the new amino acid, and the washing in between, occur in a shaking polypropylene bottle. If many peptides are synthesized simultaneously, many bags can be put together in one bottle for a synthesis of their common sequences, and are separated for the different couplings. The following steps demonstrate the synthesis of a single peptide using only one bag.

1. Place the resin-containing bag in a 250 ml polypropylene bottle.
2. Add 10 ml of deblocking agent (30% piperidine/DMF) to the bottle, screw on the cap, and shake on a shaker at 300 RPM for 10 min.
3. Decant off the deblocking agent, add 10 ml of DMF and shake 2 min.
4. Decant off DMF from the first wash and repeat another wash for 2 min.
5. Decant off the DMF. Weigh out an amino acid at 4 to 10X excess of the resin activity, dissolve in 10 ml coupling agent (TBTU/HOBT mixture).
6. Calculate the volume of 1M. DIPEA to be added to obtain 1.5 times of mole of amino acid in (5).
7. Add DIPEA into the bottle, and immediately add the amino acid solution prepared from (5), close bottle and put on a shaker for 30 min.
8. Decant off solutions, and wash with 10 ml DMF for 2 min. on shaker.

9. Repeat washing with fresh DMF three more times.
10. Continue the next cycle by repeating steps (2) to (9) until reaching the desired length of peptide chain.

The amount of reagents used in the procedure may be varied, but the molar ratio should be kept as suggested. The synthesis with multiple bags in the same bottle will need more volume of reagents, which is a multiplication of the suggested volume by the number of bags.

C-terminal Acetylation

After the synthesis of peptide chain, an acetyl group caps the peptide to help stabilize the α-helix conformation. The procedure is simple, with the coupling of the acetyl group to the N-terminal of the newly synthesized peptide which is attached to the solid support. The following steps demonstrate how to acetylate the N-terminal for one peptide.

1. After the synthesis of the peptide chain, remove the FMOC protecting group from the N-terminal of peptide by steps (2) to (4) of the synthesis cycle.
2. Wash the bag twice with 10 ml of isopropanol for 2 min. and then twice with 10 ml of diethyl ether.
3. Remove the bag from the bottle, and dry it well on a paper towel.
4. Cut open the bag on one corner, pour the resin into a 1.5 ml Eppendorf tube, and add 1 ml of acetyling cocktail containing 8:1:1 (v/v) of DMF:DIPEA:acetic anhydride.

5. Close the tube tightly, vortex well on a Styrofoam holder for 2-3 hrs, remove supernatant with Pasteur pipette, and wash with 1 ml of DMF for 1 min.
6. Wash again with 1 ml of isopropanol and then with 1 ml of diethyl ether for 1 min. each, and remove the supernatant.
7. Open the cap, and dry the peptide in the Eppendorf tube under a fume hood. The peptide should be now acetylated.

Cleaving from Solid Support

Before the peptide product is purified and used, it needs to be cleaved from the solid support. Since the chemicals will cleave the C-terminal of peptide from the solid support and remove various blocking groups from amino acid side-chains, different chemical cocktails are selected, depending on the amino acid residues and their blocking groups within a peptide. The following are the cocktails used for cleaving our peptides.

Cocktail A: for peptide containing no methionine (M), arginine (R), typtophan (W), or the Trt protecting group.

950 µl TFA

50 µl H_2O

Cocktail B: for peptide containing methionine (M) or Arginine (R).

75 mg crystalline phenol

25 µl EDT

50 µl thioanisole

50 µl H_2O

1000 µl TFA

Cocktail C: for peptide containing typtophan (W) or Trt group

25 µl EDT

25 µl H_2O

950 µl TFA

The following steps demonstrate the cleaving of a single peptide from the solid support.

1. Add 1 ml of a cleaving cocktail into an Eppendorf tube containing the peptide attached to the resin.
2. Close the tube tightly, and vortex on a Styrofoam holder for 2-3 HRS (6 HRS for a peptide with arginine).
3. Place 30 ml of ice-cold ether into a 250 ml vacuum flask equipped with a fritted-glass funnel.
4. Pour the reaction cocktail into a fritted-glass funnel and apply a vacuum slowly. The white precipitate in the cold ether is the peptide product.
5. Pour the precipitate into a second fritted glass funnel and apply the vacuum again. Now the peptide will stay on the funnel.
6. Add 3 ml of 18 MΩ metal-free water to the funnel. Wait for 5 min. for the peptide dissolves in the water.
7. Change the vacuum flask, and slowly apply a vacuum to suck the peptide through the flask. The solution is then collected as fraction #1.

8. Add another 3 ml of water to dissolve the peptide for 5 min. and suck it down into the flask again. Now the solution is labeled fraction #2.
9. Dissolve more peptide by repeat step (8) for 3 to 4 times, until the optical density at 190 nm of the peptide fraction is below 0.5 in a 200 μm cell.

Peptide Purification

The purification of crude peptide is performed by using Reverse Phase High Performance Liquid Chromatography (RP-HPLC). A solution of crude peptide from the cleaving step is filtered through a 0.22 μm Millipore membrane and injected into a Hewlett-Packard Ti-series HPLC. The reverse phase column is a Vydac semi-preparative C_{18}, running through with 2.5 ml/min of mobile phase. The gradient system contains solution A, 0.1% TFA in 18 MΩ-water, and solution B, 0.1 %TFA in acetonitrile. The gradient is varied from 10% to 50% of B in 20 min., 50% to 80% in 2 min., stays at 80% for 5 min. and then goes back to 10% in 2 min. before injecting the next sample.

Since all of our peptides contain a tyrosine tag on the N-terminal, the fraction is collected based upon the absorption observed at 280 nm. The purified peptide is then lyophilized on a speed rotor under vacuum. This lyophilized peptide is now ready to use, and can be stored in refrigerator for a long period of time.

Peptide Identification

A confirmation of peptide identity is performed by analyzing its mass. Fast Atom Bombardment Mass Spectrometry (FAB-MS) is a technique that takes less than a μg of sample, and it is fast. The mass of molecular ion (M+) observed in mass spectra is compared with the calculated mass of the peptide. The result from the analysis not only confirms the identity of a peptide, but also provide us with information on its purity.

References

Alewood, P. F., Croft, M., Schnolzer, M., & Kent, S. B. H. (1990) in *Peptides: Proceedings of the 21st European Peptide Symposium* (Giralt, E., & Andreu, D. Eds.) pp.174-175, ESCOM Science Publishers B. V.

Beck-Sickinger, A. G., Durr, H., & Jung, G. (1991) *Peptide Res. 4*, 88-94.

Houghten, R. A., Degraw, S. T., Bray, M. K., Hoffman, S. R., & Frizzell, N. D. (1986) *BioTechniques 4*, 522-528.

Merrifield, R. B. (1963) *J. Am. Chem. Soc. 85*, 2149-2153.

Appendix C

Peptide Aggregation Study

Introduction

It is known that protein and peptide aggregation can complicated an analysis of CD spectra. For most cases, one can detect a large aggregate by visual inspection of the solution or by seeing an increase in absorption spectrum of sample in the 300-400 nm region where there should be no absorption. The scanning method is simply based on a detection of light scattering caused by an aggregate, resulting in a shifting of the absorption spectrum of sample from the absorption of its solvent baseline.

In some cases, sample molecules form an aggregate containing intermolecular association of molecules, and the effect of light scattering monitored from a shifting of the absorption spectrum is too small to be detected. Based on the assumption that the CD spectrum measured for an association of peptides is different from one of a unimolecular peptide, we can investigate whether the sample is aggregated by monitoring the CD intensity of the peptide sample over a range of concentrations. If aggregation occurs in the system, it would be dependent to the sample concentration resulting in changes in CD intensity.

Our experiment generally measures the CD spectrum using a sample containing 1-2 mM of amide, with 0 to 88% methanol present in the system. The peptide sample is sometimes found to form an aggregate during the experiment

at high methanol content. We expect 88%, the highest content of methanol used in the experiment, to be the condition most likely to find an aggregate. This condition is used throughout the aggregation study with the CD measurement at 222 nm over the peptide concentration range from 2 μM to 10 mM.

Experimental Procedures

1. If the stock peptide is already dissolved in water, calculate and add a pure methanol into stock peptide in water, to obtain a 88% methanol content (for example, adding 352 μl of methanol into 48 μl of stock peptide in water to get 400 μl of peptide sample in 88% methanol).
2. Scan baseline CD of 88% methanol solution using 1-mm cell, and clean the cell.
3. Measure the OD of the sample from (1) using a 1 mm cell. The sample alone without subtracting baseline should give an OD at 222nm between 0.5 to 1.0. If not, the peptide concentration can be adjusted either by adding a freshly made 88% methanol solution, or by adding an appropriate amount of lyophilized peptide.
4. After adjusting, label the solution from (3) as a 1X solution. Put 140 μl of this solution in a 1 mm rectangular cell and measure CD spectra from 260 to 220 nm.
5. Subtract the baseline CD from (2) from the CD from (4) and note the CD intensity at 222 nm.

6. Add 140 µl of 88% methanol solution into the sample from step (4). Mix it well by flipping cell several times. This sample is called a 2X solution.
7. Scan the CD spectra for the 2X solution, subtract by the baseline from (2), and note its intensity at 222 nm.
8. Before cleaning the cell, use a long pipette tip to recover 100 µl of the 2X sample, put it in a 10 ml volumetric flask, and make up the volume to 10 ml with 88% methanol solution.
9. Label the sample from (8) as a 200X sample. Put 6 ml of this sample in a 2-cm cylindrical cell, and measure its CD.
10. Clean the cell from (9) Scan the baseline with 88% methanol.
11. Subtract the baseline CD from (10) from the CD from (9). Note the CD at 222 nm.
12. Add 100 µl of 1X sample into a 1 ml volumetric flask, and make up volume to 1 ml with 88% methanol solution.
13. Label the sample from (12) as a 10X sample, and measure its CD in a 1 mm cell.
14. Subtract the baseline CD from (2) from the CD from (13). Note the CD at 222 nm.
15. Determine the OD at 190 of the 1X sample in 200 µm cell, subtract baseline, and note the OD_{190}.

After these steps, we now have the CD intensity at 222 nm for a peptide in 88% methanol solution with the peptide concentration at 1X, 2X, 10X and 200X

dilution. The result is then shown by plotting the CD at 222 nm versus the log of peptide concentration. The independence of the CD intensity to the sample concentration will indicate that the peptide sample does not form an aggregate. The aggregation study is performed for all of our peptide samples, and we confirm that the peptides under our experiment conditions are a peptide monomer (Figures 5.1- 5.3).

Figure 5.1: Plot of CD intensity at 222 nm versus concentration of peptide ac-Y-VAXAK-VAXAK-VAXAK-amide in 88% methanol solution, where X = A, T, I, W, Y, G, E.

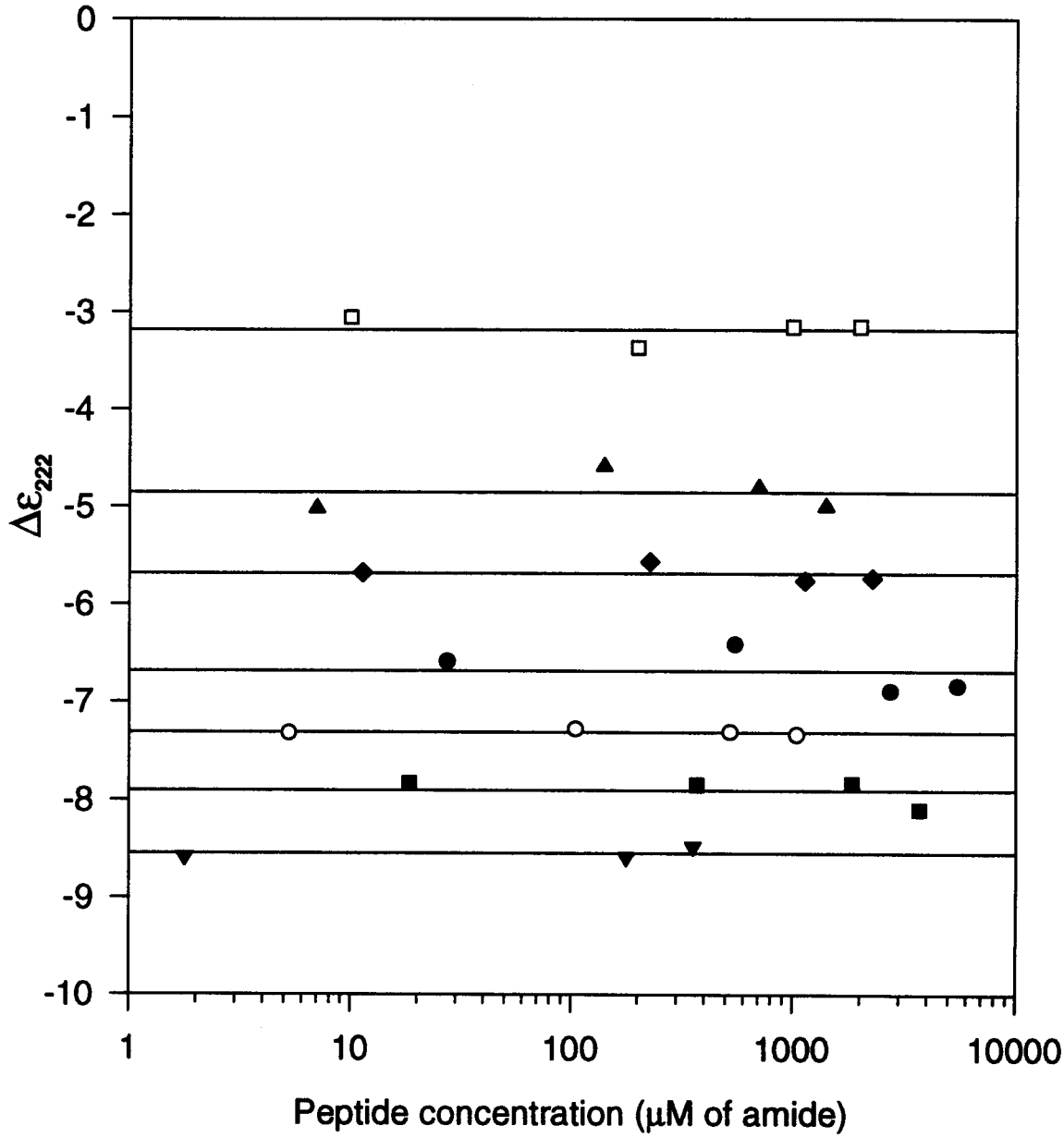


Figure 5.2: Plot of CD intensity at 222 nm versus concentration of peptide ac-Y-VAXAK-VAXAK-VAXAK-amide in 88% methanol solution, where X = M, L, H, C, F.

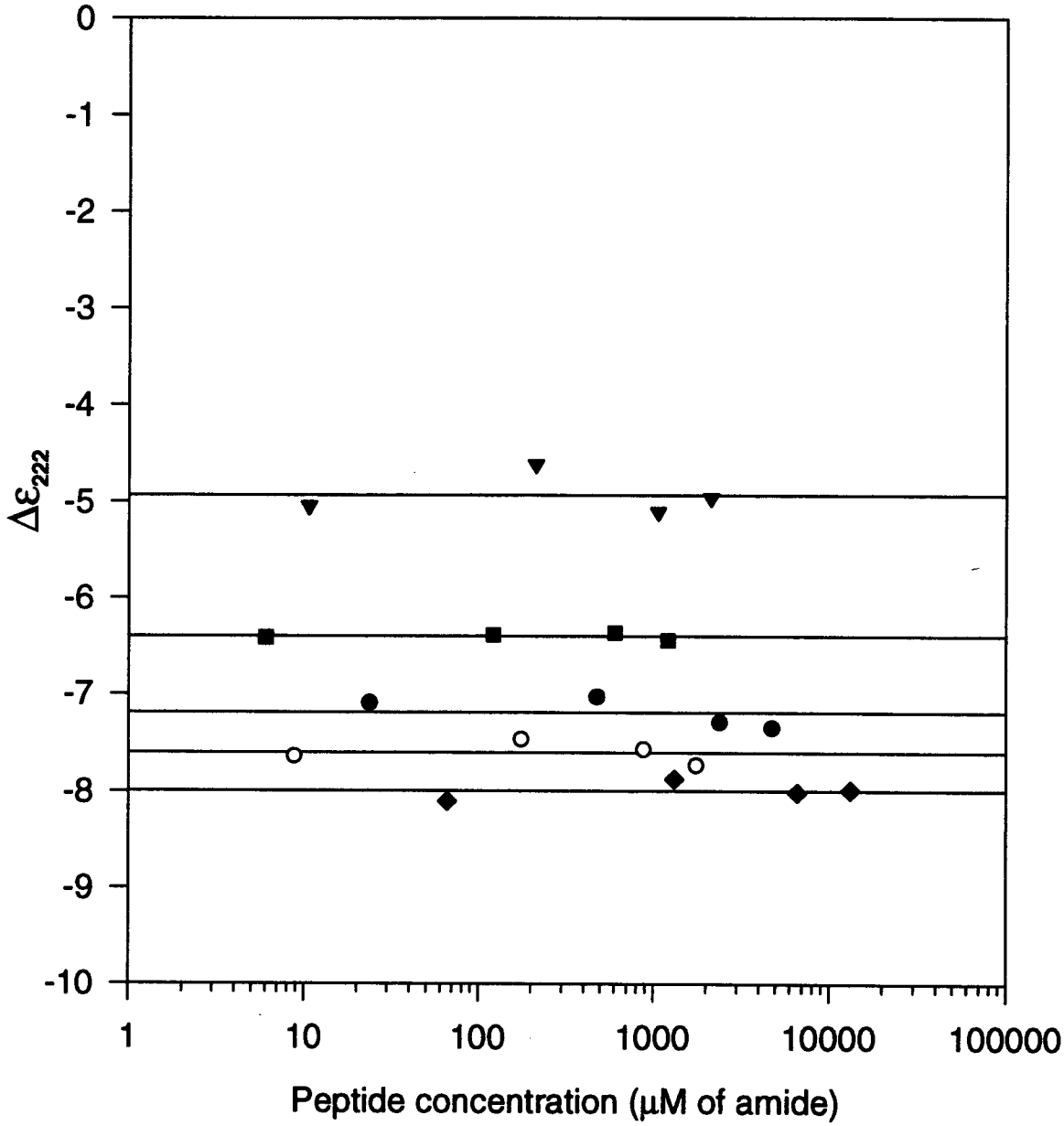
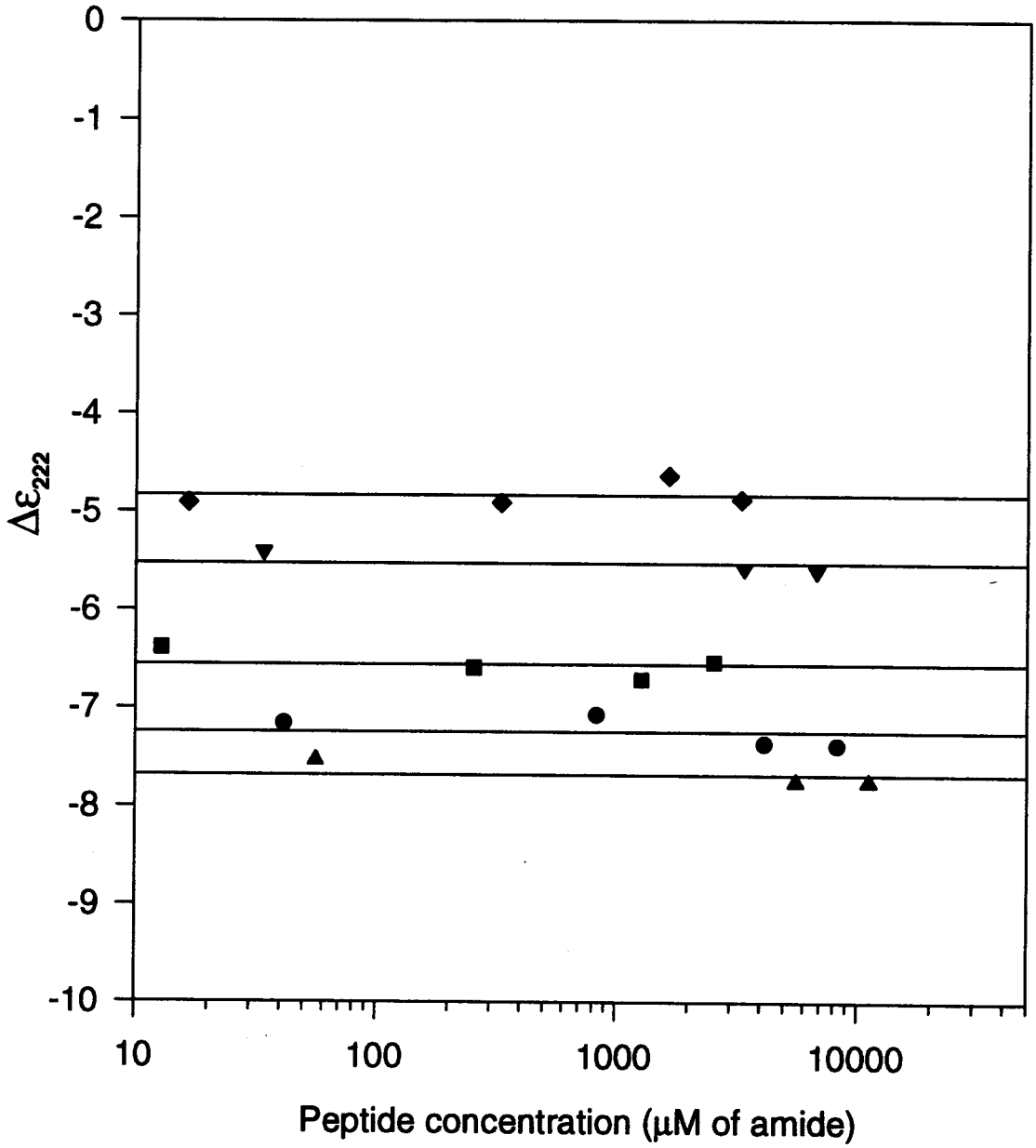


Figure 5.3: Plot of CD intensity at 222 nm versus concentration of peptide ac-Y-VAXAK-VAXAK-VAXAK-amide in 88% methanol solution, where X = K, S, Q, N, D.



Appendix D

Circular Dichroism Spectra from Methanol Titration

Figure 6.1: Circular dichroism spectra of Ac-Y-$(VAAAK)_3$-amide in 2mM sodium phosphate buffer pH 5.5 with 0, 9, 18, 27, 36, 45, 54, 63, 72, 81, 88.2% methanol.

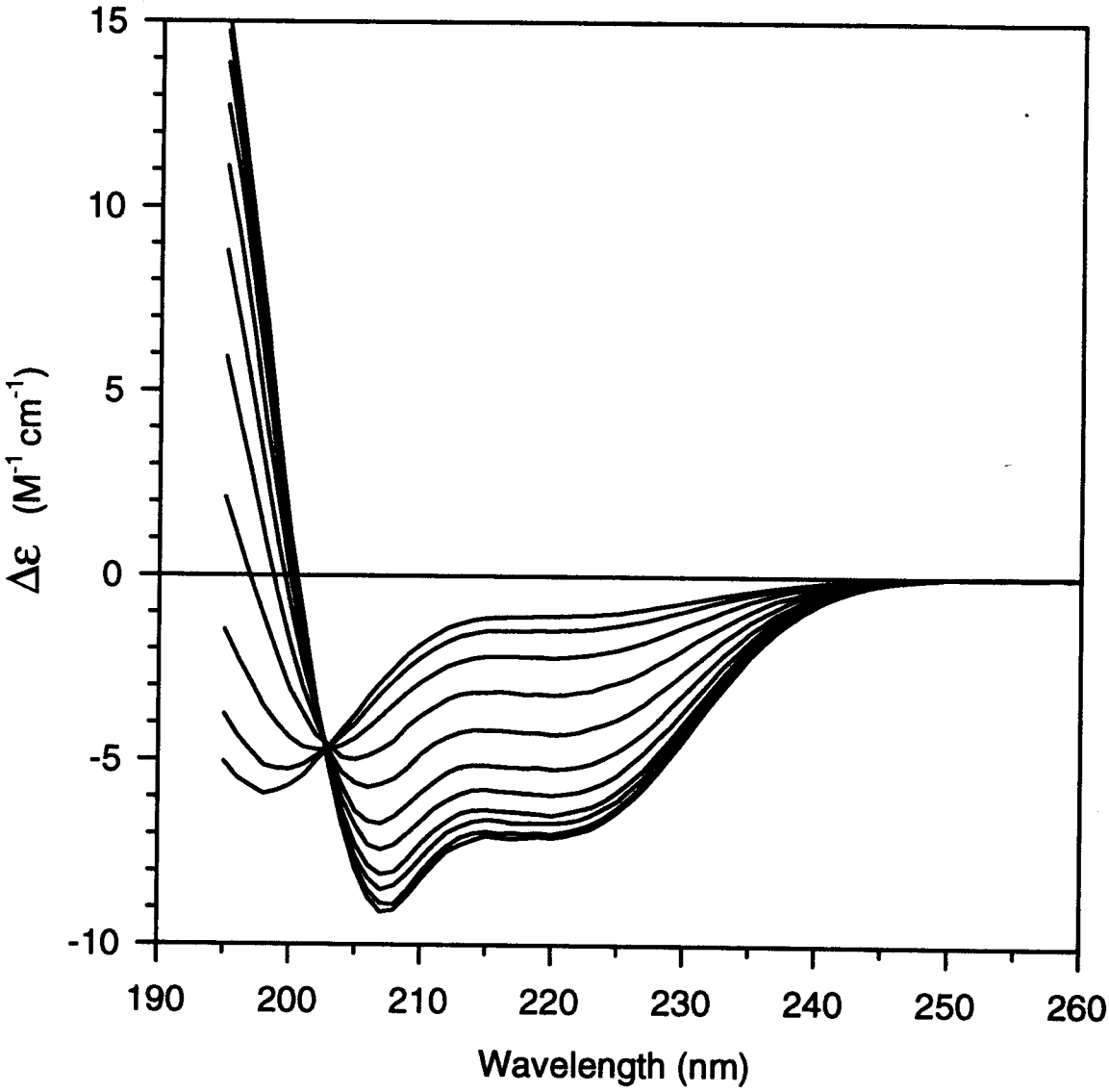


Figure 6.2: Circular dichroism spectra of Ac-Y-$(VALAK)_3$-amide in 2mM sodium phosphate buffer pH 5.5 with 0, 9, 18, 27, 36, 45, 54, 63, 72, 81, 88.2% methanol.

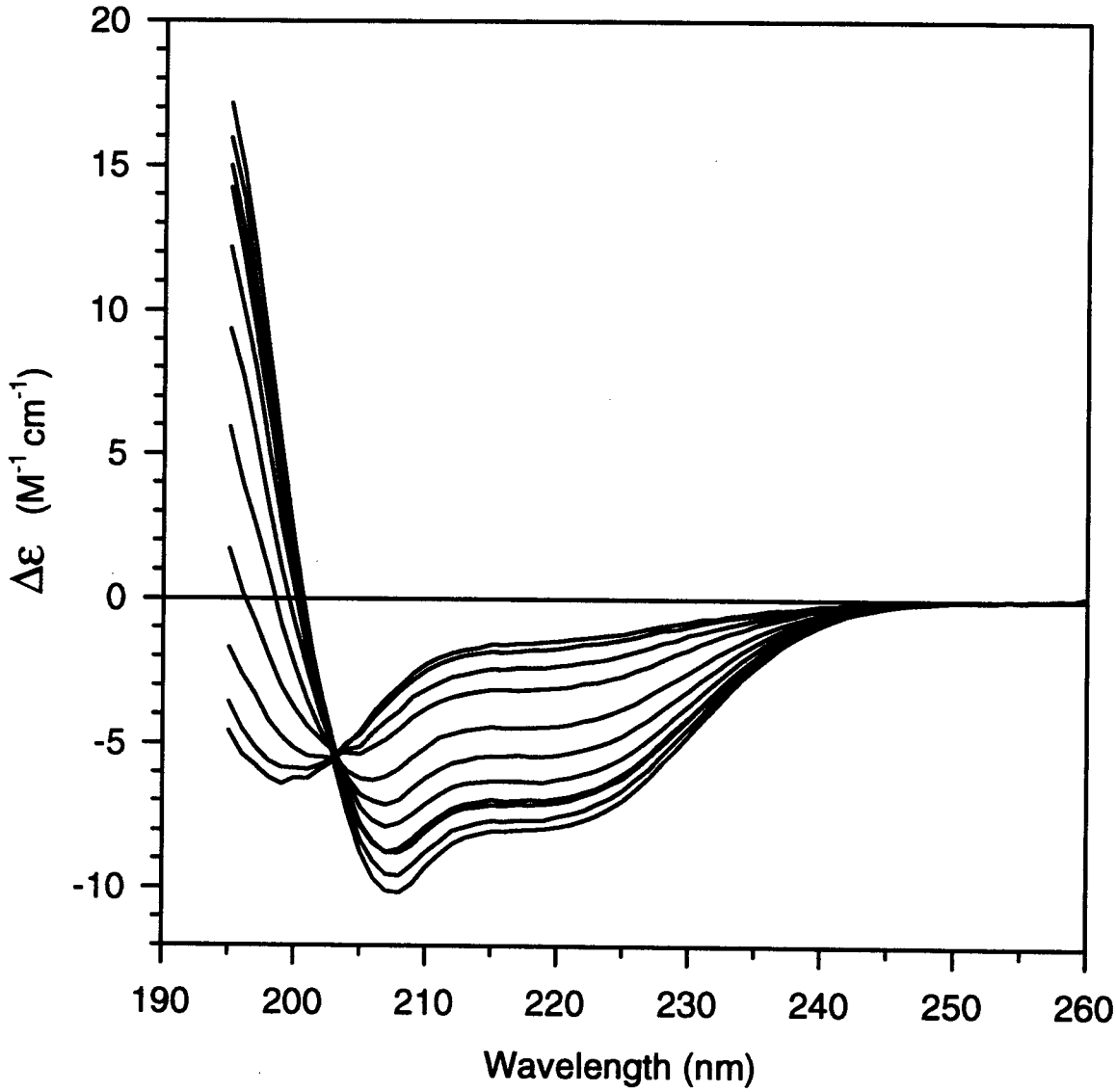


Figure 6.3: Circular dichroism spectra of Ac-Y-$(VAVAK)_3$-amide in 2mM sodium phosphate buffer pH 5.5 with 0, 9, 18, 27, 36, 45, 54, 63, 72, 81, 88.2% methanol.

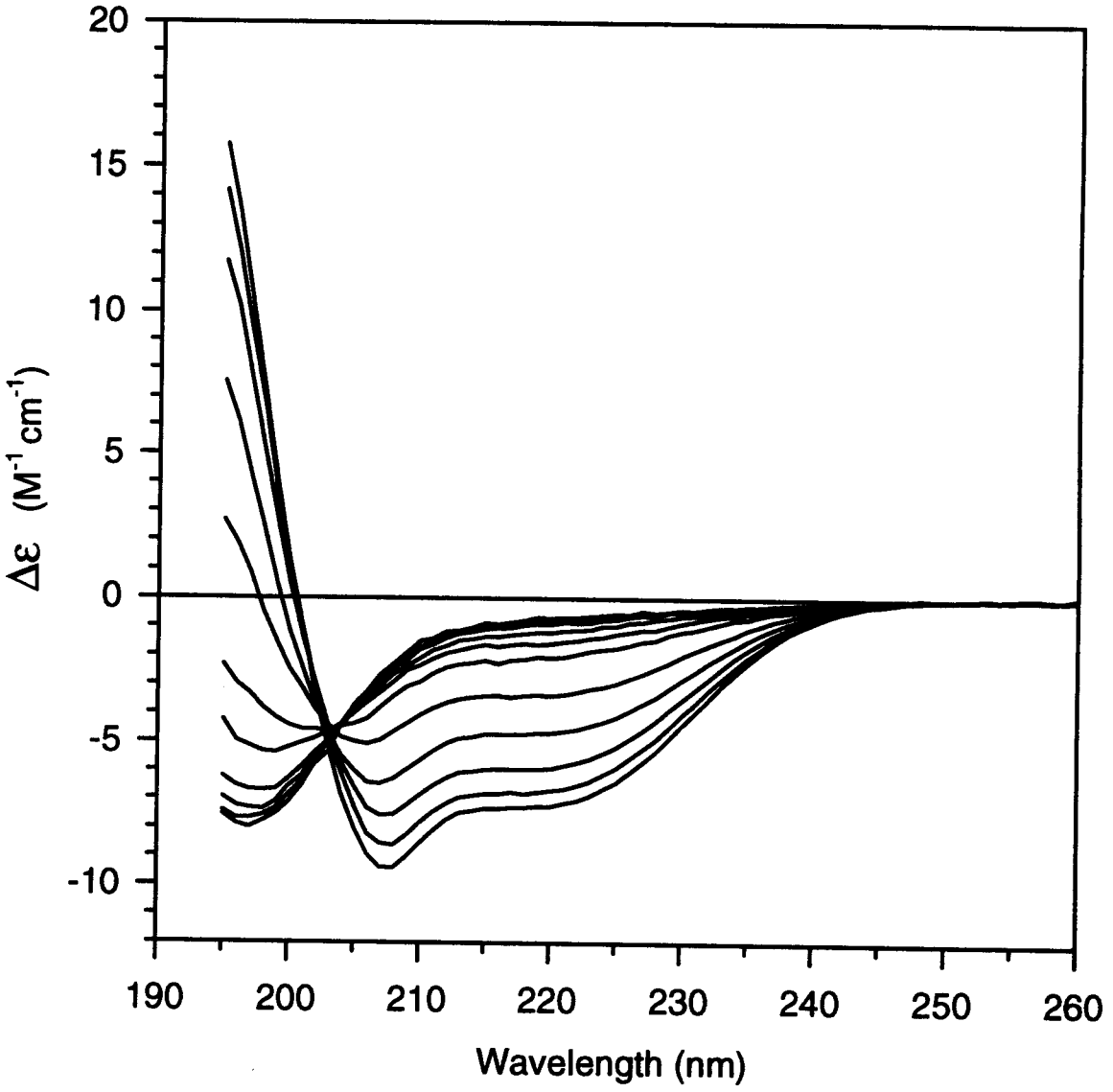


Figure 6.4: Circular dichroism spectra of Ac-Y-(VAIAK)$_3$-amide in 2mM sodium phosphate buffer pH 5.5 with 0, 9, 18, 27, 36, 45, 54, 63, 72, 81, 88.2% methanol.

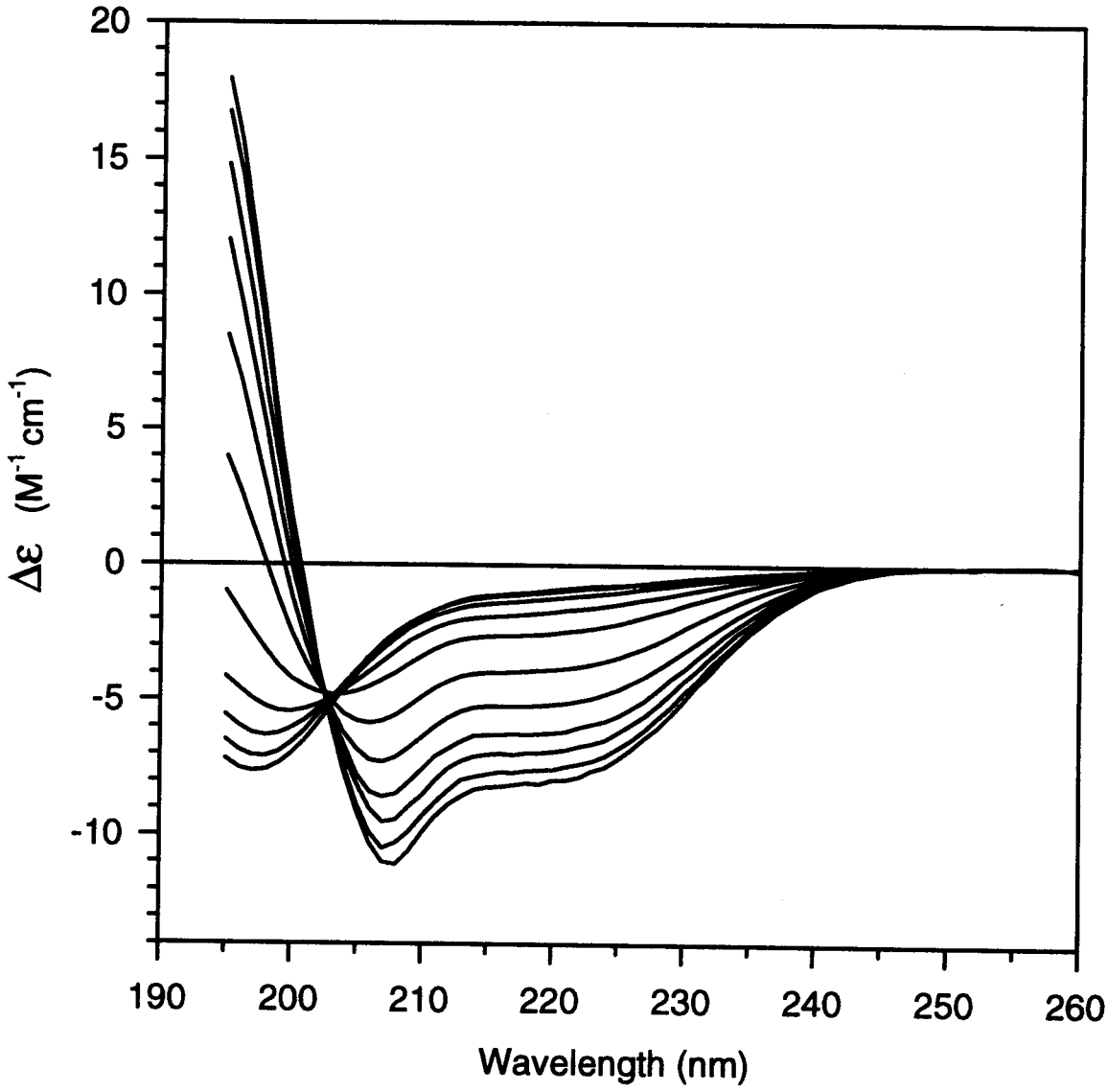


Figure 6.5: Circular dichroism spectra of Ac-Y-(VAGAK)$_3$-amide in 2mM sodium phosphate buffer pH 5.5 with 0, 9, 18, 27, 36, 45, 54, 63, 72, 81, 88.2% methanol.

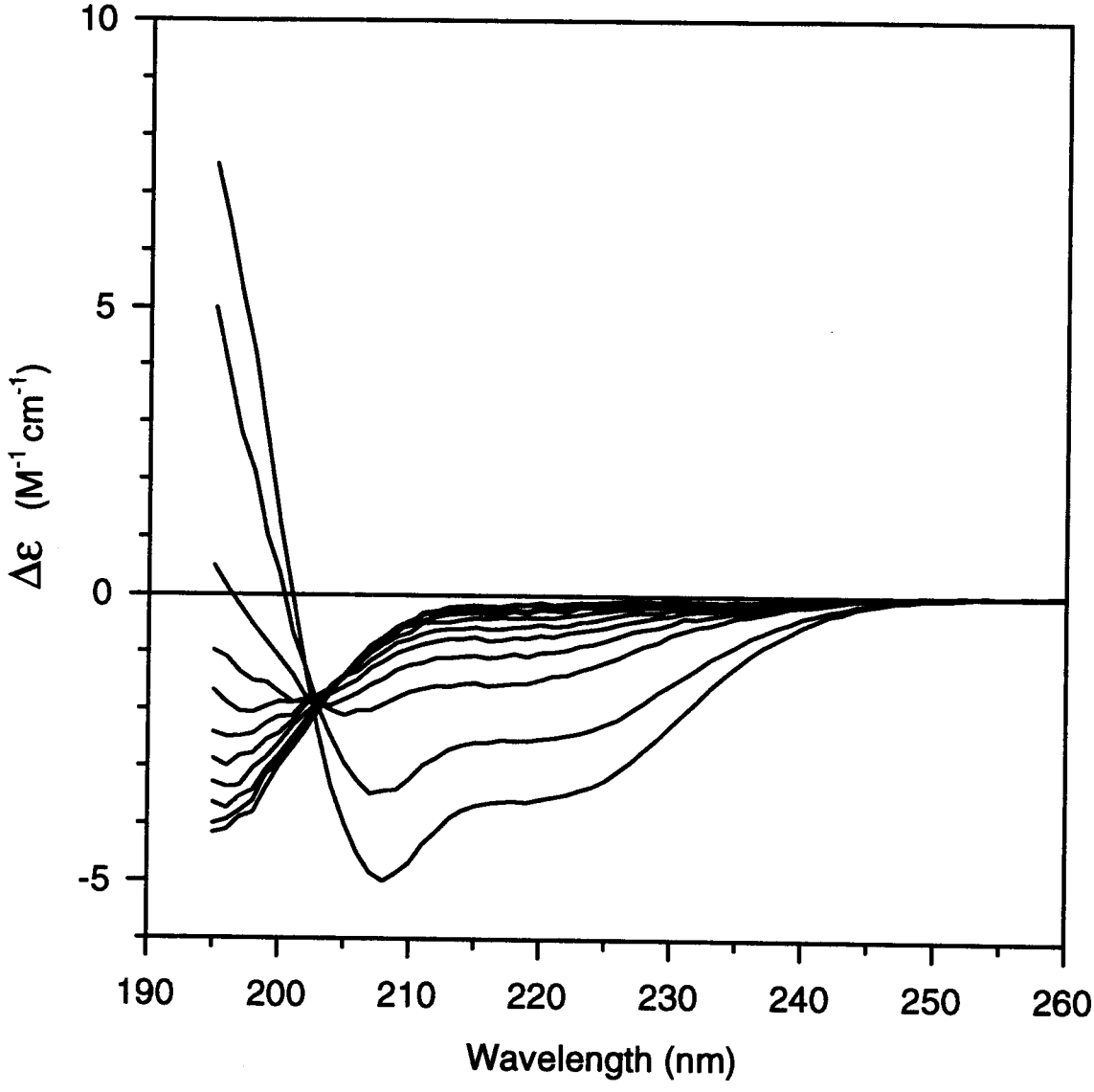


Figure 6.6: Circular dichroism spectra of Ac-Y-(VAFAK)$_3$-amide in 2mM sodium phosphate buffer pH 5.5 with 0, 9, 18, 27, 36, 45, 54, 63, 72, 81, 88.2% methanol.

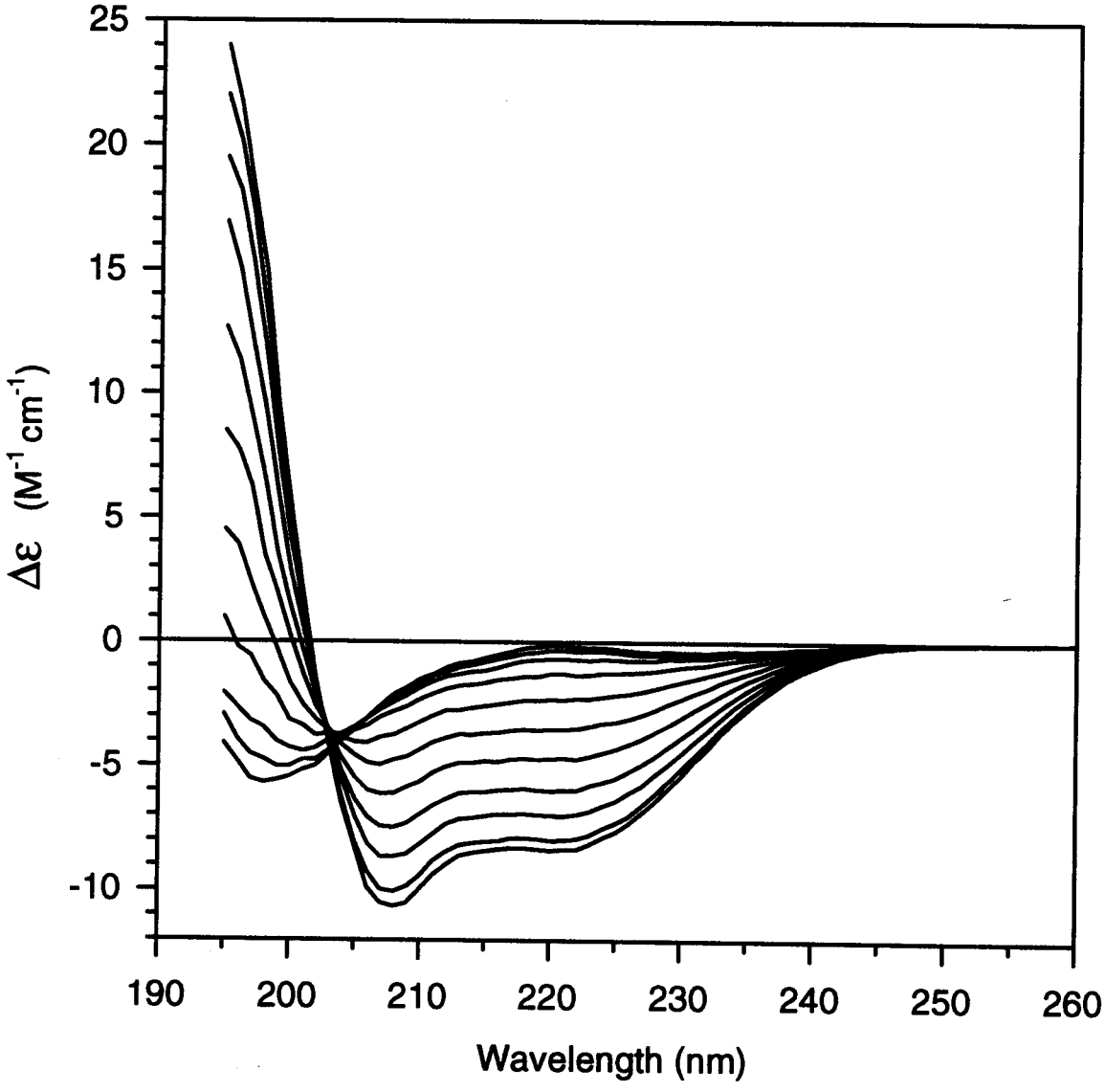


Figure 6.7: Circular dichroism spectra of Ac-Y-$(VAWAK)_3$-amide in 2mM sodium phosphate buffer pH 5.5 with 0, 9, 18, 27, 36, 45, 54, 63, 72, 81, 88.2% methanol.

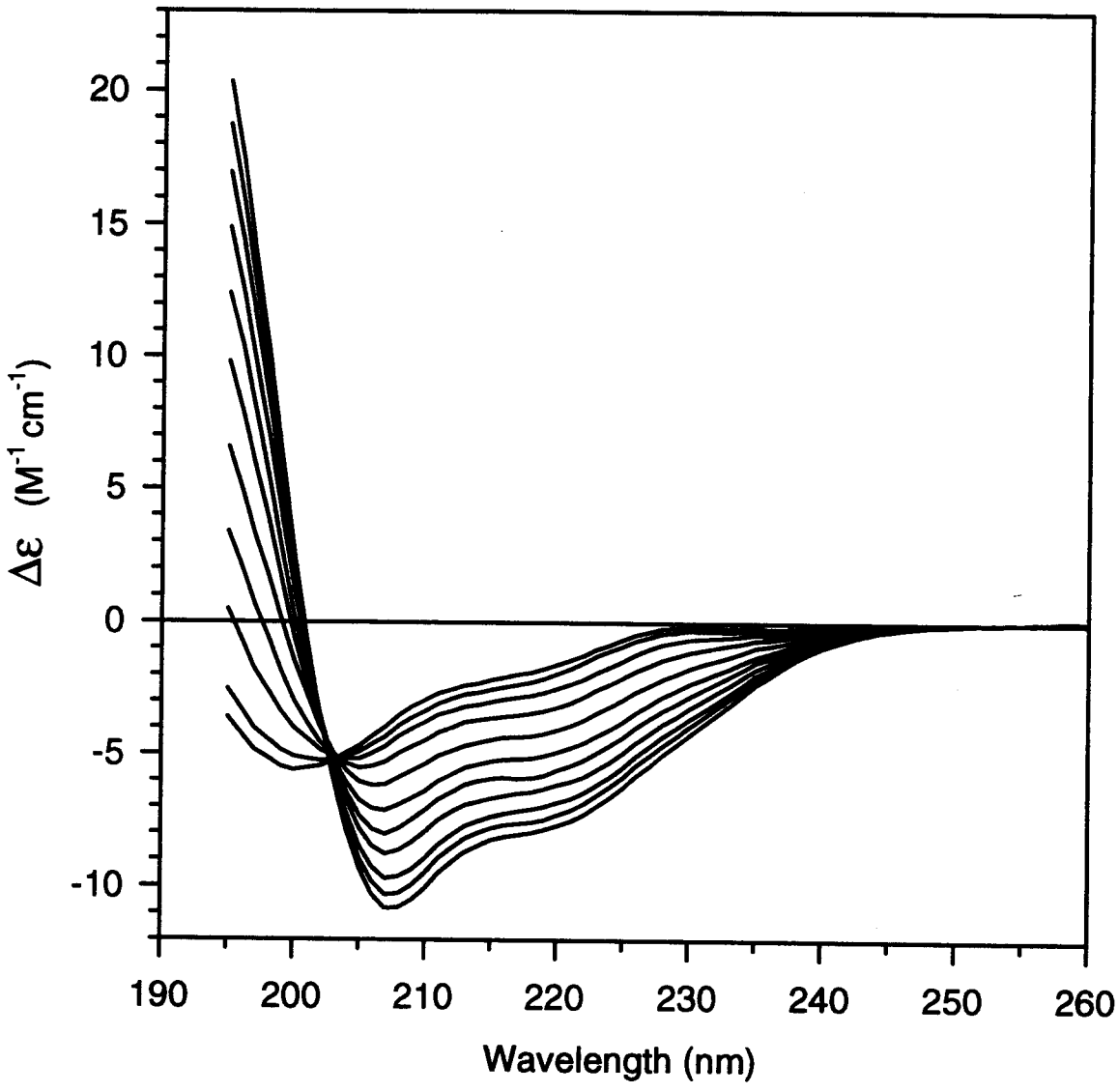


Figure 6.8: Circular dichroism spectra of Ac-Y-$(VAYAK)_3$-amide in 2mM sodium phosphate buffer pH 5.5 with 0, 9, 18, 27, 36, 45, 54, 63, 72, 81, 88.2% methanol.

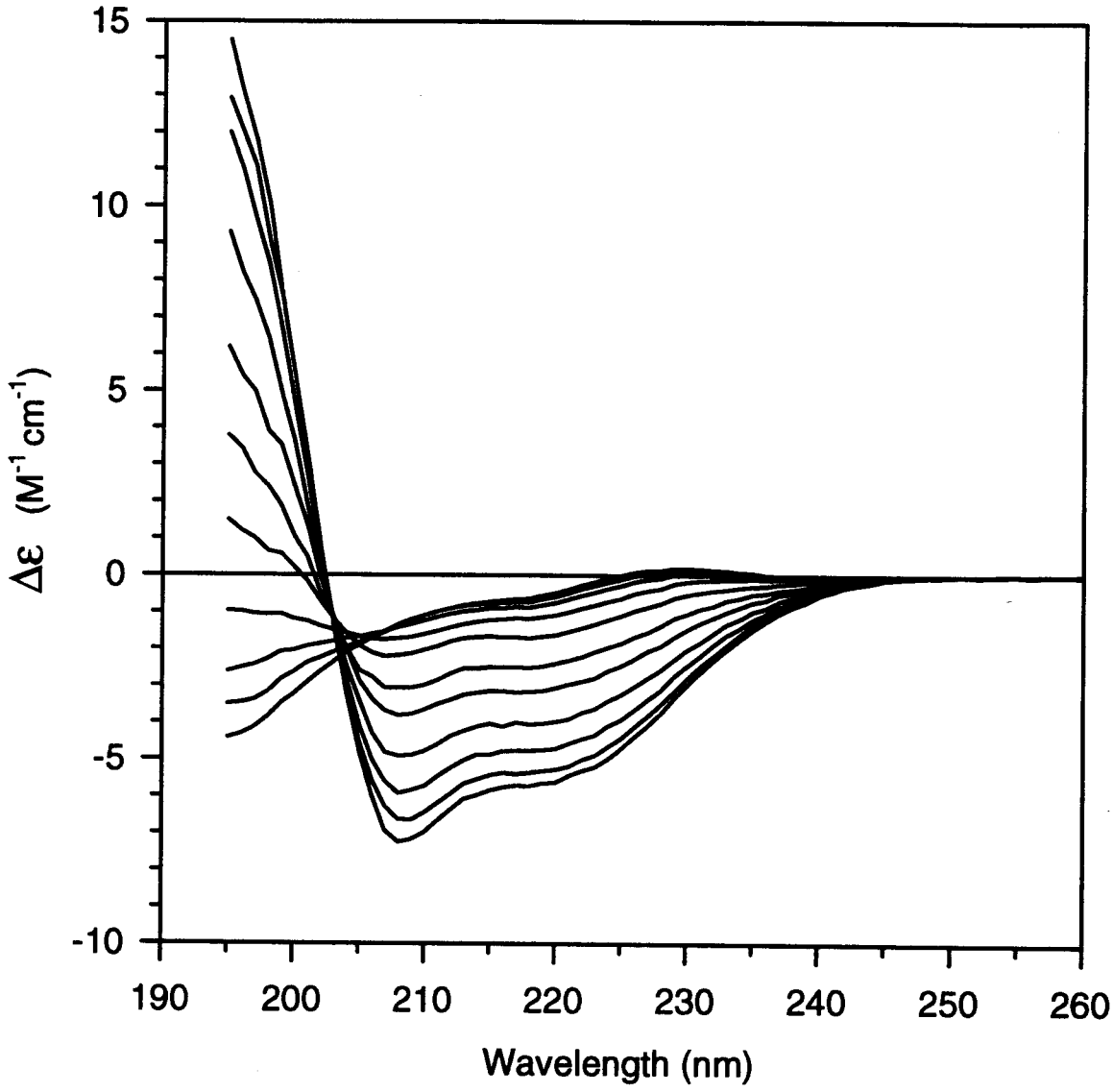


Figure 6.9: Circular dichroism spectra of Ac-Y-$(VADAK)_3$-amide in 2mM sodium phosphate buffer pH 5.5 with 0, 9, 18, 27, 36, 45, 54, 63, 72, 81, 88.2% methanol.

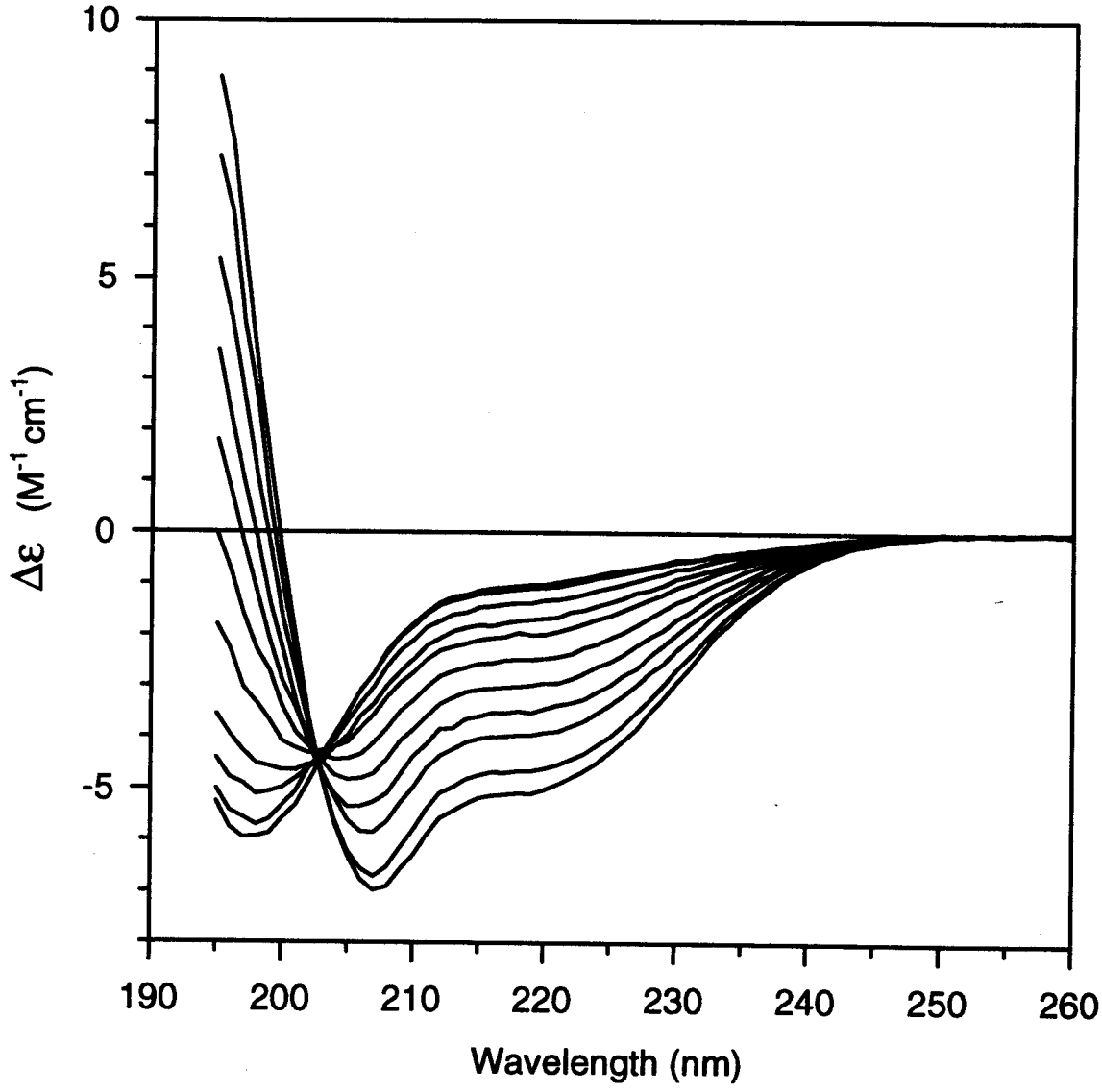


Figure 6.10: Circular dichroism spectra of Ac-Y-(VANAK)$_3$-amide in 2mM sodium phosphate buffer pH 5.5 with 0, 9, 18, 27, 36, 45, 54, 63, 72, 81, 88.2% methanol.

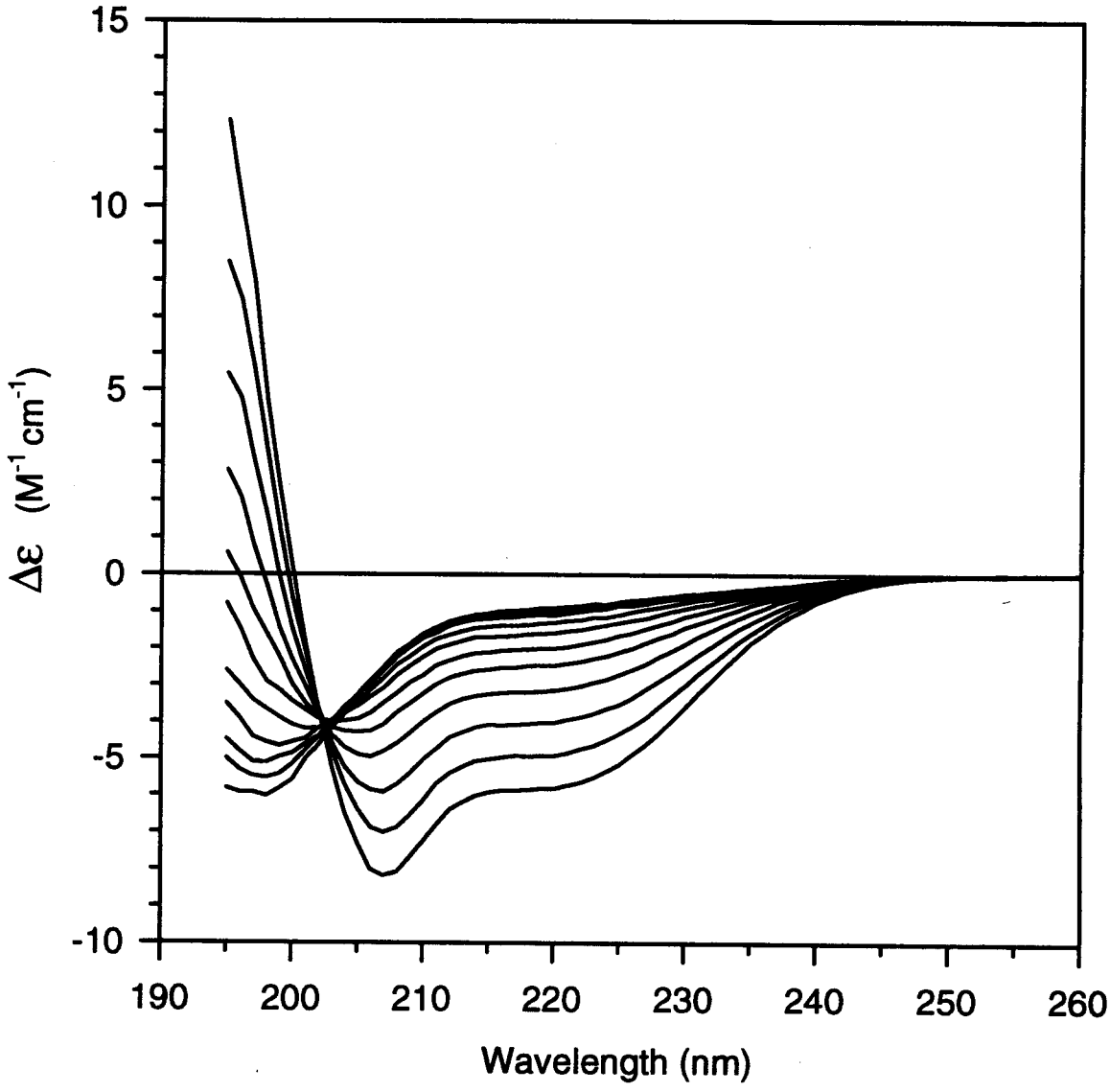


Figure 6.11: Circular dichroism spectra of Ac-Y-(VAEAK)$_3$-amide in 2mM sodium phosphate buffer pH 5.5 with 0, 9, 18, 27, 36, 45, 54, 63, 72, 81, 88.2% methanol.

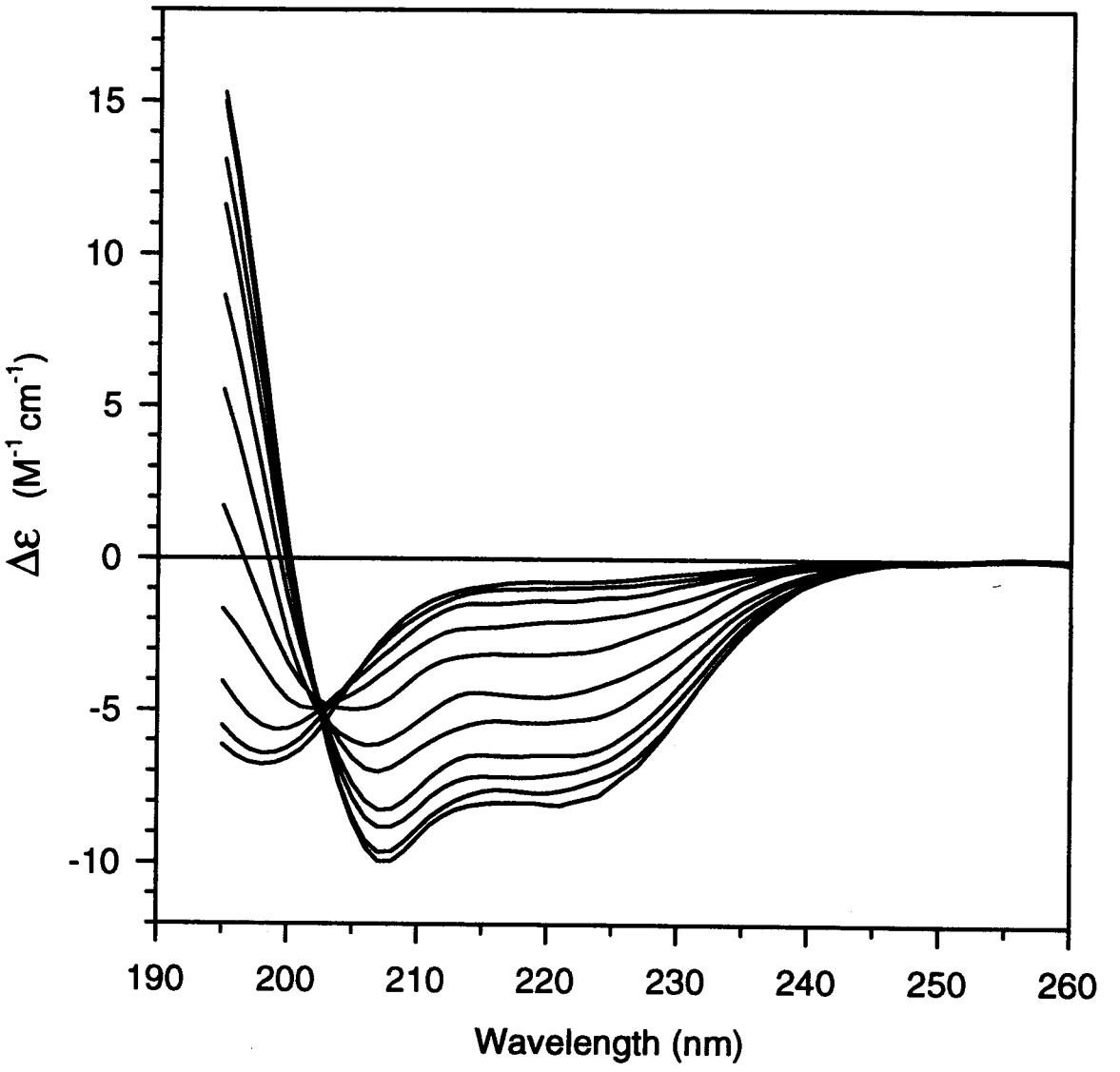


Figure 6.12: Circular dichroism spectra of Ac-Y-(VAQAK)$_3$-amide in 2mM sodium phosphate buffer pH 5.5 with 0, 9, 18, 27, 36, 45, 54, 63, 72, 81, 88.2% methanol.

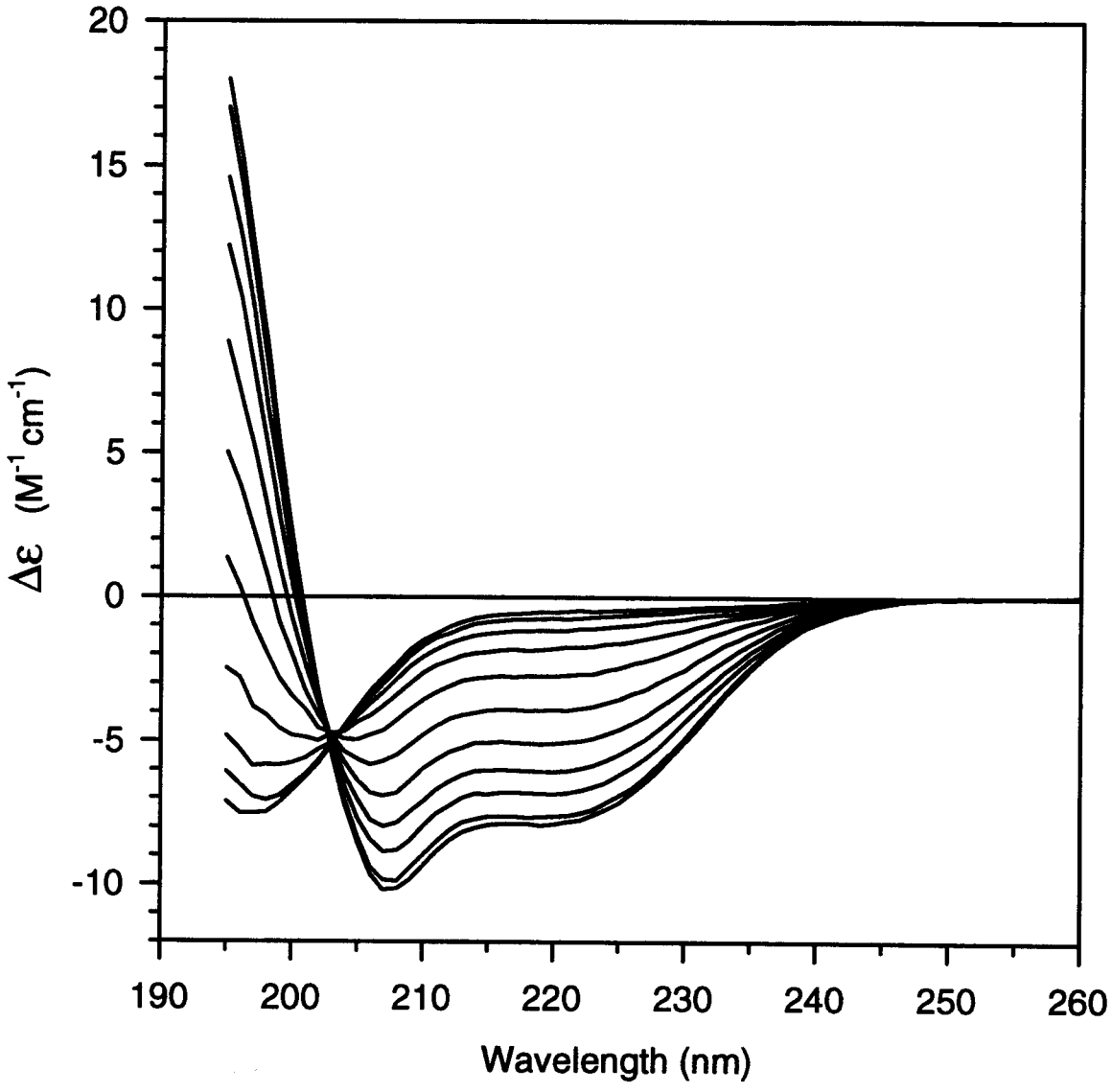


Figure 6.13: Circular dichroism spectra of Ac-Y-(VAKAK)$_3$-amide in 2mM sodium phosphate buffer pH 5.5 with 0, 9, 18, 27, 36, 45, 54, 63, 72, 81, 88.2% methanol.

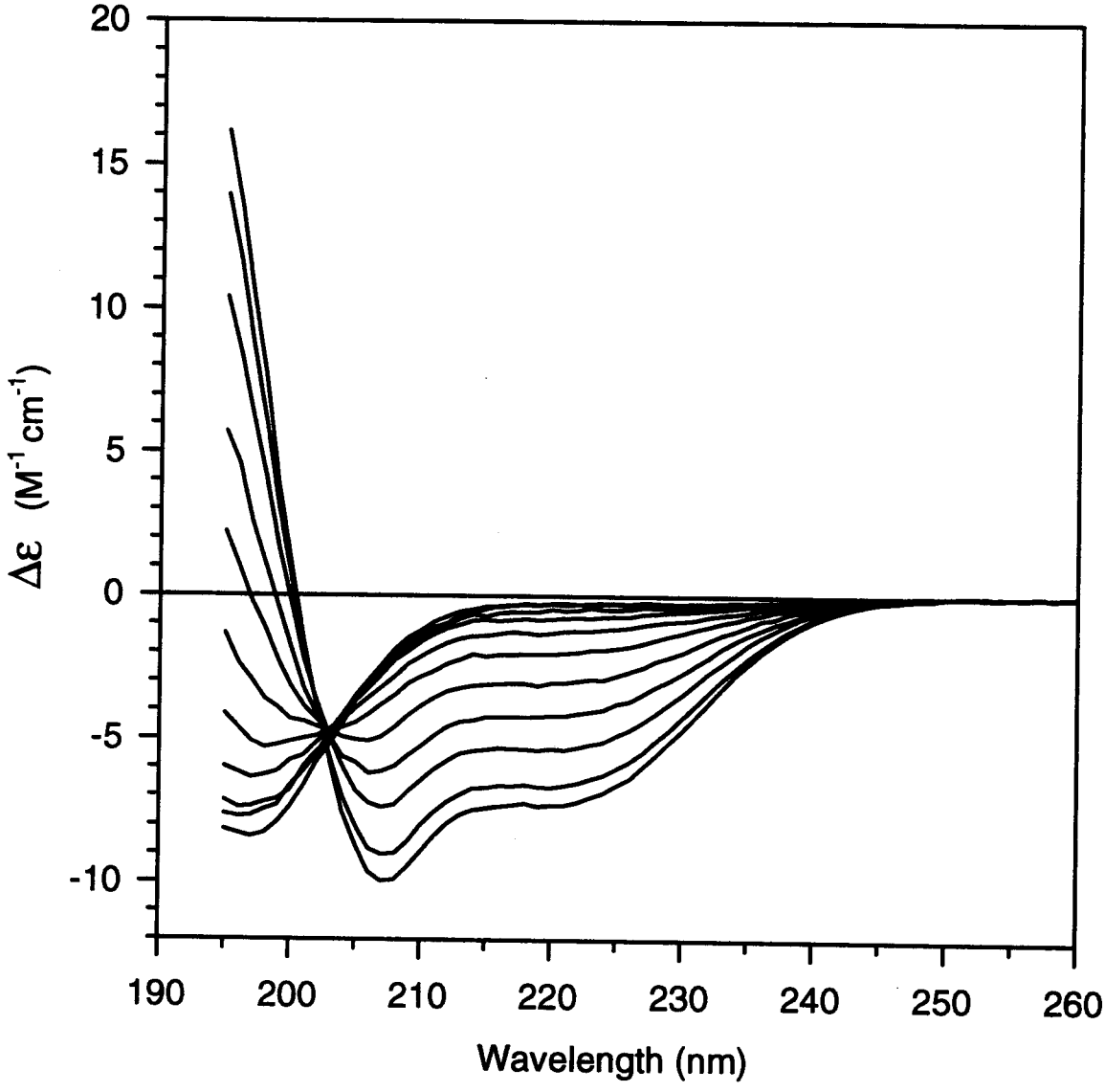


Figure 6.14: Circular dichroism spectra of Ac-Y-(VAHAK)$_3$-amide in 2mM sodium phosphate buffer pH 5.5 with 0, 9, 18, 27, 36, 45, 54, 63, 72, 81, 88.2% methanol.

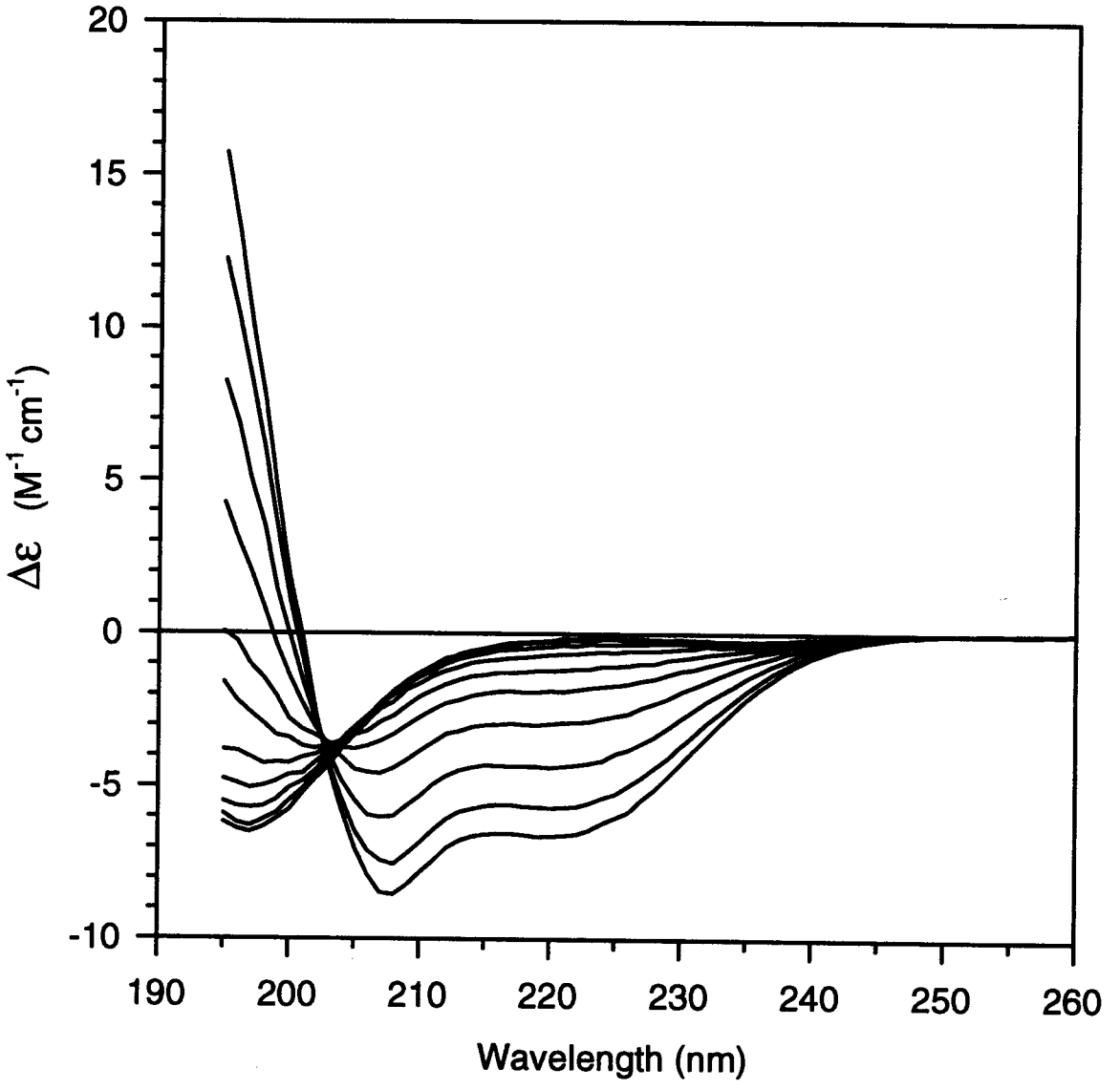


Figure 6.15: Circular dichroism spectra of Ac-Y-$(VARAK)_3$-amide in 2mM sodium phosphate buffer pH 5.5 with 0, 9, 18, 27, 36, 45, 54, 63, 72, 81, 88.2% methanol.

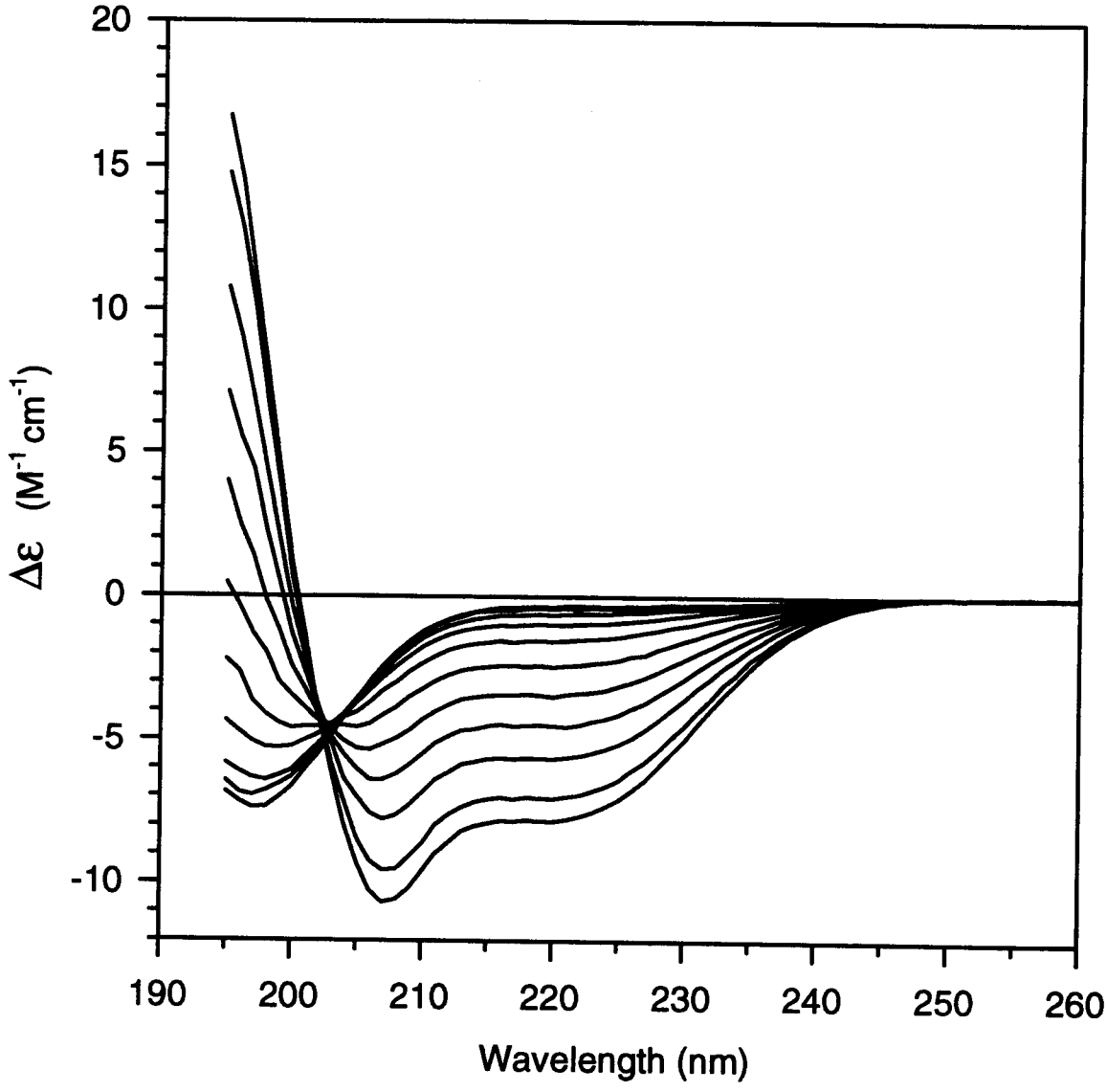


Figure 6.16: Circular dichroism spectra of Ac-Y-(VAMAK)$_3$-amide in 2mM sodium phosphate buffer pH 5.5 with 0, 9, 18, 27, 36, 45, 54, 63, 72, 81, 88.2% methanol.

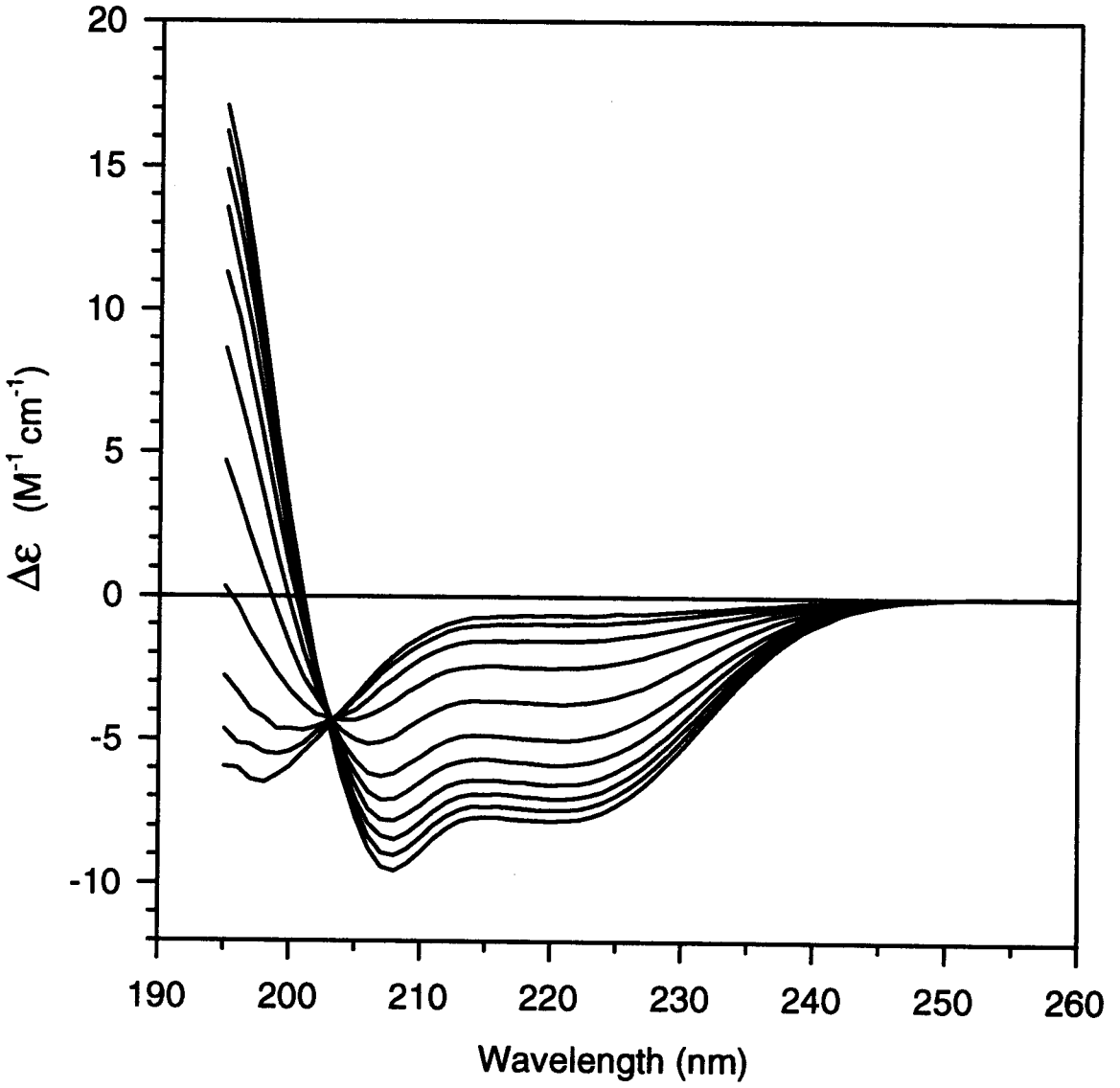


Figure 6.17: Circular dichroism spectra of Ac-Y-(VACAK)$_3$-amide in 2mM sodium phosphate buffer pH 5.5 with 0, 9, 18, 27, 36, 45, 54, 63, 72, 81, 88.2% methanol.

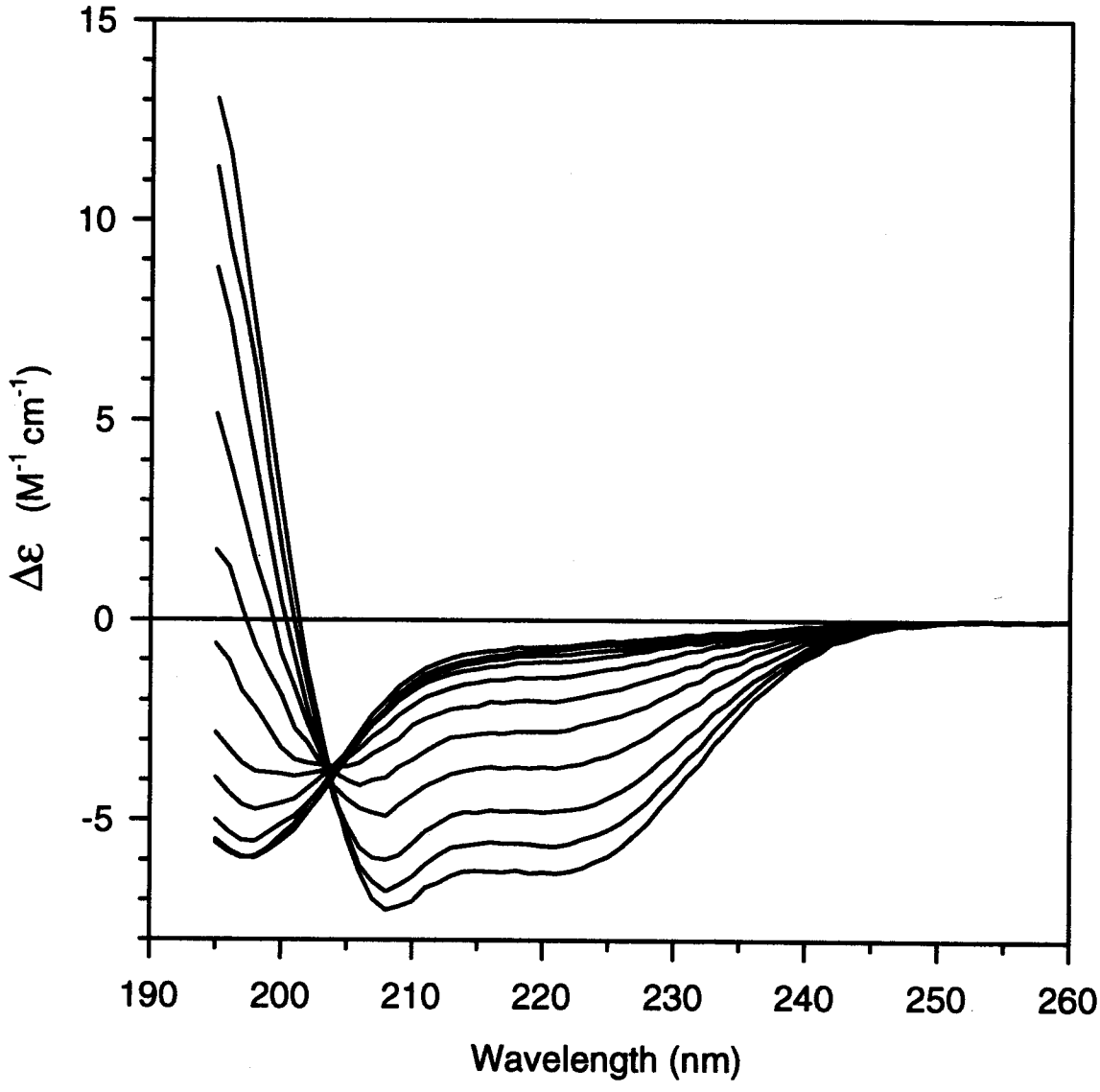


Figure 6.18: Circular dichroism spectra of Ac-Y-$(VASAK)_3$-amide in 2mM sodium phosphate buffer pH 5.5 with 0, 9, 18, 27, 36, 45, 54, 63, 72, 81, 88.2% methanol.

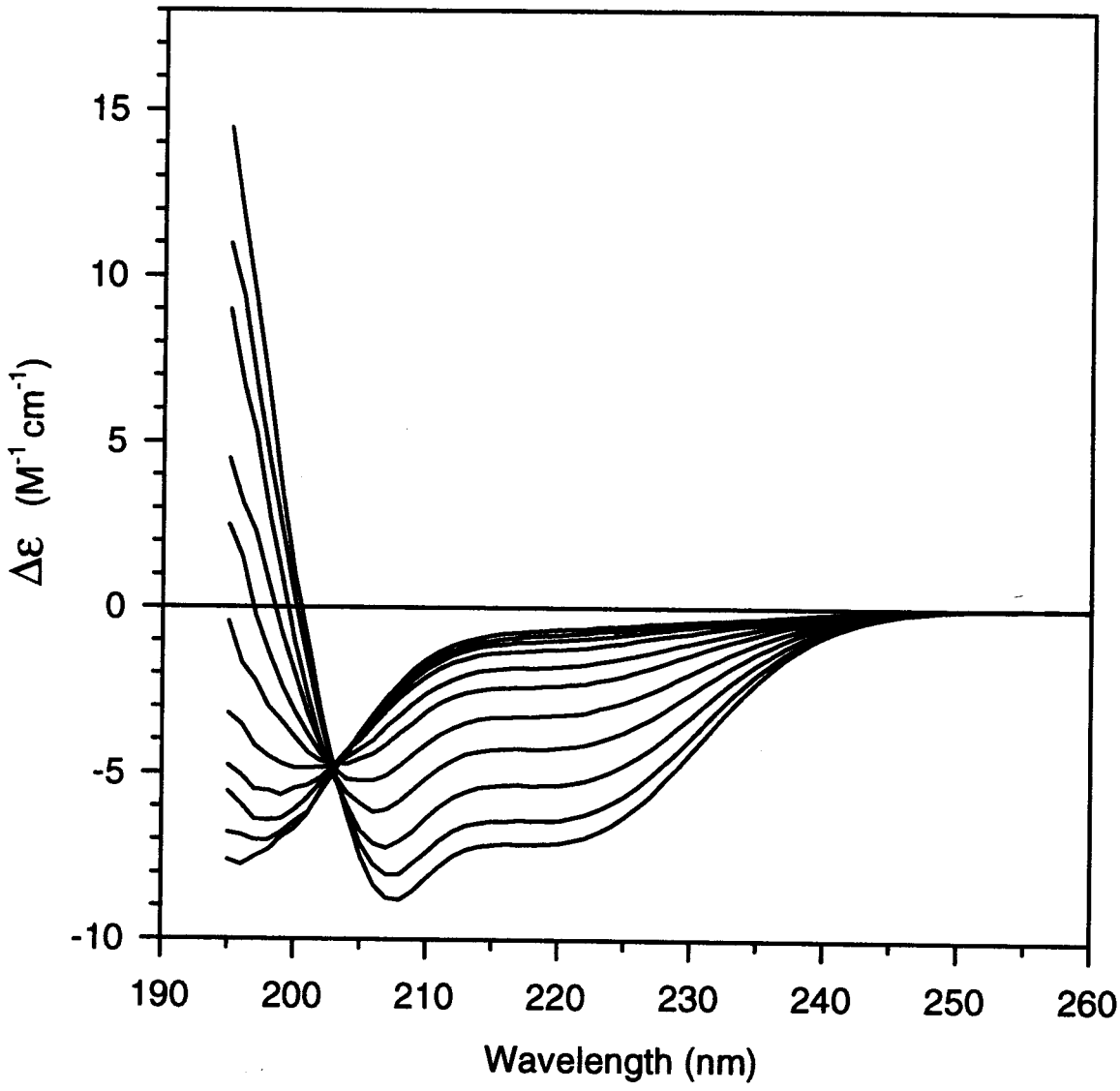


Figure 6.19: Circular dichroism spectra of Ac-Y-$(VATAK)_3$-amide in 2mM sodium phosphate buffer pH 5.5 with 0, 9, 18, 27, 36, 45, 54, 63, 72, 81, 88.2% methanol.

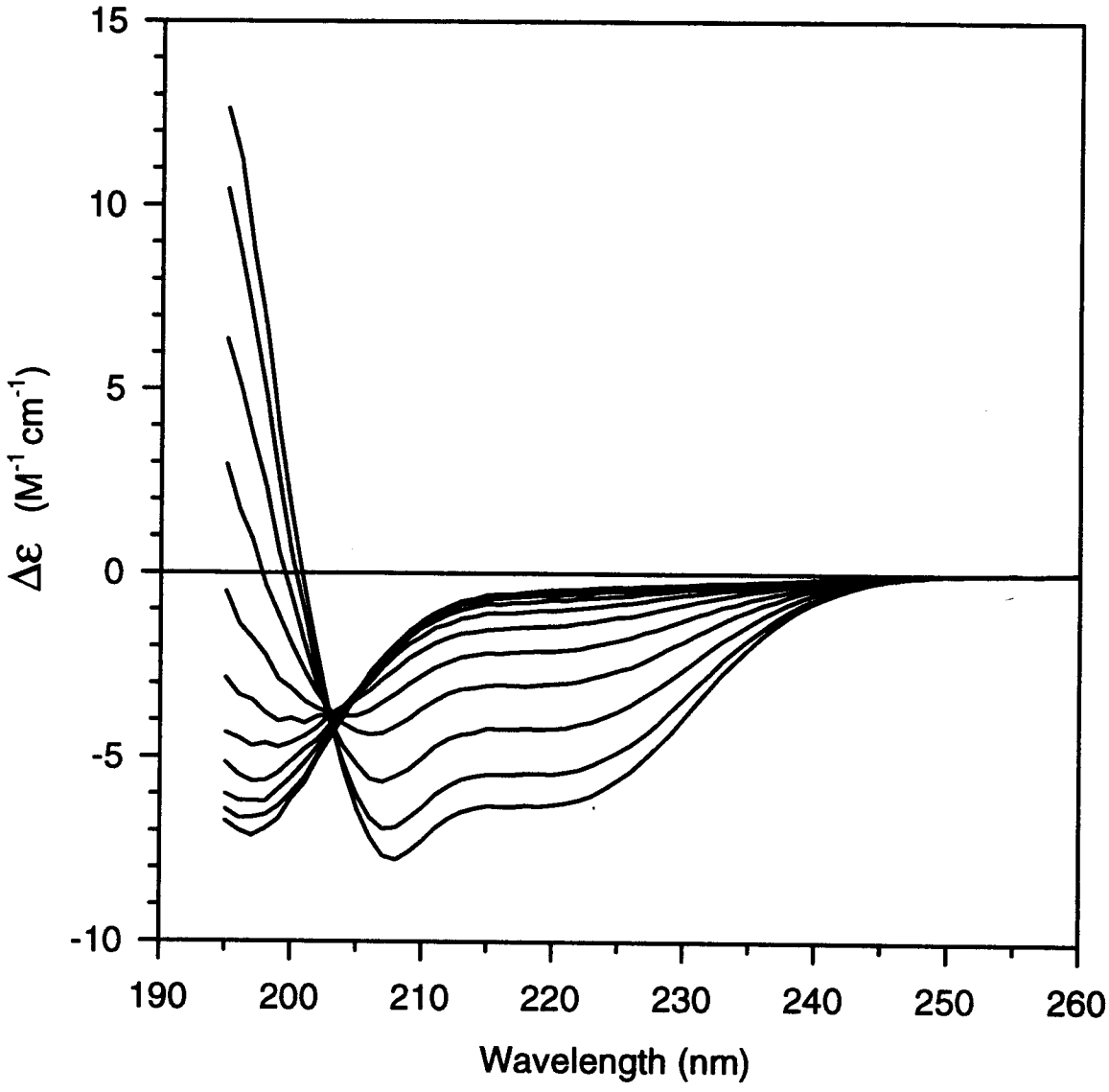
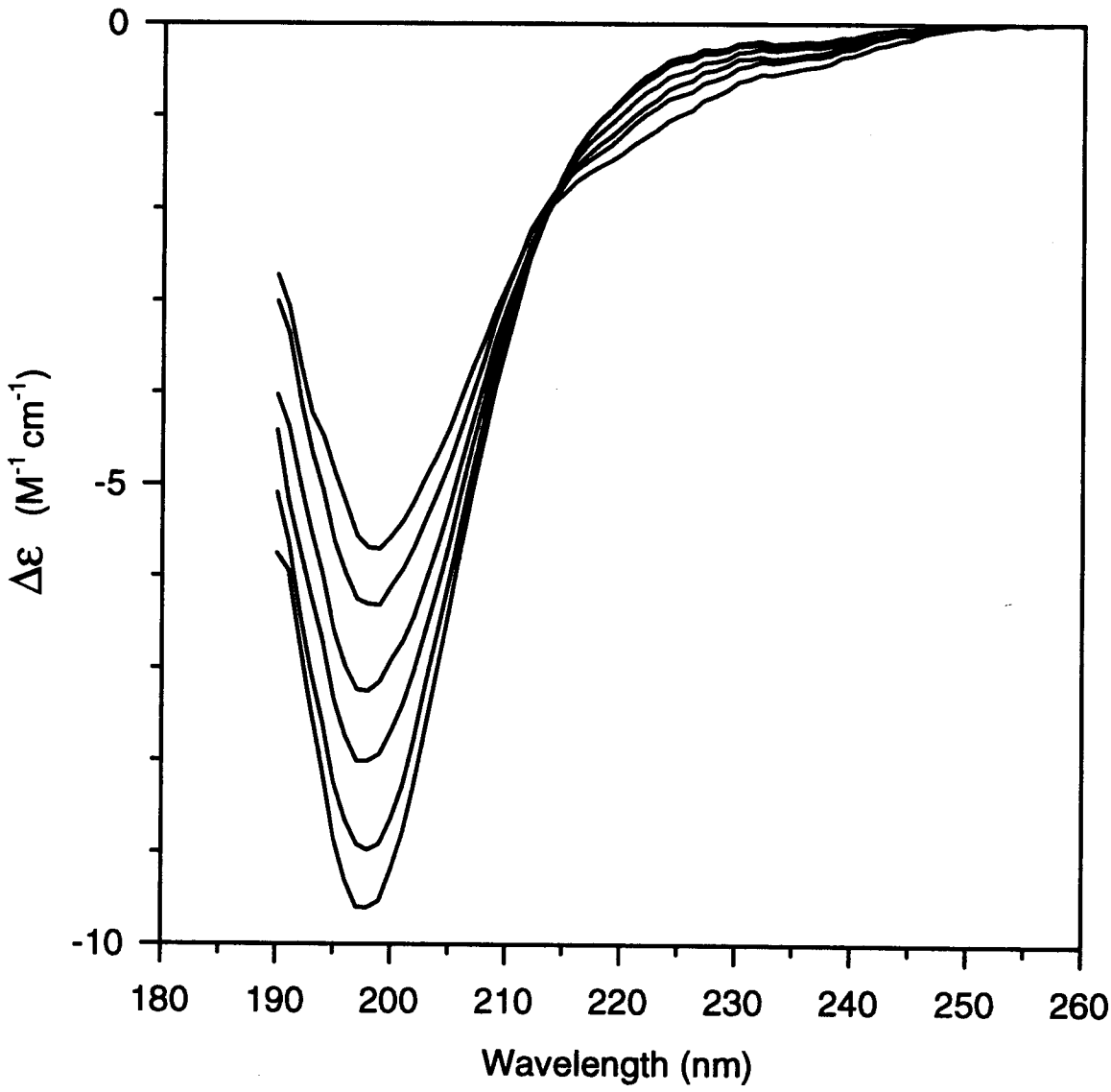


Figure 6.20: Circular dichroism spectra of Ac-Y-$(VAPAK)_3$-amide in 2mM sodium phosphate buffer pH 5.5 with 0, 18, 36, 54, 72, 88.2% methanol.



Appendix E

Extinction Coefficients of 20 peptides

Extinction coefficients at 190 nm (ε_{190}) are determined for all 20 peptide sequences using the UV absorption bands at 190 and 280 nm, and the known extinction coefficients at 280 nm (ε_{280}) of some absorbing residues (Y = 1,280 $M^{-1}cm^{-1}$, W = 5,690 $M^{-1}cm^{-1}$, C = 120 $M^{-1}cm^{-1}$) within the sequences(Elwell, 1976).

Peptide Sequences	ε_{190} (M^{-1} cm^{-1})
Ac-Y-$(VAAAK)_3$-Amide	9,000
Ac-Y-$(VALAK)_3$-Amide	9,900
Ac-Y-$(VAVAK)_3$-Amide	9,300
Ac-Y-$(VAIAK)_3$-Amide	9.800
Ac-Y-$(VAGAK)_3$-Amide	9,150
Ac-Y-$(VAFAK)_3$-Amide	22,350
Ac-Y-$(VAWAK)_3$-Amide	16,400
Ac-Y-$(VAYAK)_3$-Amide	14,200
Ac-Y-$(VADAK)_3$-Amide	9,900
Ac-Y-$(VANAK)_3$-Amide	10,400
Ac-Y-$(VAEAK)_3$-Amide	9,850
Ac-Y-$(VAQAK)_3$-Amide	11,450
Ac-Y-$(VAKAK)_3$-Amide	9,700

Ac-Y-$(VAHAK)_3$-Amide	11,150
Ac-Y-$(VARAK)_3$-Amide	12,400
Ac-Y-$(VAMAK)_3$-Amide	9,900
Ac-Y-$(VACAK)_3$-Amide	9,400
Ac-Y-$(VASAK)_3$-Amide	10,300
Ac-Y-$(VATAK)_3$-Amide	8,400
Ac-Y-$(VAPAK)_3$-Amide	9,950

References

Elwell, M. L. (1976) Ph.D. thesis, Department of Chemistry, University of Oregon.