Understanding Structural Stability of Pharmaceutically Relevant Macromolecular

Complexes: A Biophysical and Biochemical Approach

By

Reza Esfandiary

B.S. University of Kansas, 2005
M.S. University of Kansas, 2008

Submitted to the Department of Pharmaceutical Chemistry and the faculty of the Graduate School of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Dissertation Committee:

_____
Chairperson

_____

_____

_____

_____

Dissertation defended on July 31, 2009

The Dissertation Committee for Reza Esfandiary certifies that
this is the approved version of the following dissertation:

Understanding Structural Stability of Pharmaceutically Relevant Macromolecular

Complexes: A Biophysical and Biochemical Approach

Dissertation Committee:

_____
Chairperson

_____

_____

_____

_____

Dissertation approved: July 31, 2009

Abstract

Macromolecules, due to their large size and complexity are prone to a variety of physical and chemical degradations. Development of stable formulations that retain the macromolecule's stability and activity over its designated shelf life is therefore a crucial step in the production of such complexes as safe and effective therapeutics. One rapid and systematic formulation tool is the utility of the empirical phase diagrams (EPDs) which have enhanced our understanding of the response of the structure and stability of proteins to environmental perturbations. In the case of more complex macromolecular systems (*e.g.*, multi-domain fusion proteins, large recombinant proteins, and viruses), however, the measured stability is presumably the sum of all components. This makes interpretations at the molecular level difficult. Herein, we have investigated the structural stability of three macromolecular systems of varying complexity to see if their structural stability can at least be partially understood in terms of the behavior of their individual domains and components. We have examined the effect of the observed structural alterations on the losses in activity. We have discussed how this information can be used in designing high throughput screening assays for identification of solution stabilizers for development of optimal formulations. The utility of the obtained information in interpretation of the biological functions of these systems in vivo is also evaluated.

Furthermore, EPDs constructed based on techniques sensitive to transitions due to alterations of protein motions have been shown to provide information above and beyond that obtained by the static approach. Therefore, integration of techniques that detect extensive and subtle conformational alterations of proteins, as well as structural fluctuations into an EPD appears to be crucial. Herein, we show that plots of the temperature dependent 2^{nd} derivative peak positions of aromatic residues have measurable slopes prior to protein unfolding and that these slopes are sensitive to the dielectric properties of the surrounding microenvironment. We show that such data hold valuable information regarding protein dynamic fluctuations that can be integrated into dynamic-based EPDs.

Dedicated to:

My parents,

Mansour & Farah Esfandiary

Acknowledgements

I would first like to thank my graduate advisor, Russell Middaugh. Russ you have been a great mentor and teacher and I thank you for giving me the opportunity to work in your lab. I thank you for your patience and sincerity at all times. I really appreciate your integrity, depth of knowledge, and dedication to science. I would like to thank the past and present members of the Middaugh group for your friendship and all the scientific discussions that helped me enormously along the way. I especially like to thank Dr. Sangeeta Joshi for being an outstanding lab manager, a great scientific resource, and a good friend during my stay in the Middaugh lab.

I would like to thank Drs. Cory Berkland, Teruna Siahaan, Jennifer Laurence, and Sunil David for serving on my dissertation defense committee. Thank you all for taking the time out of your busy schedule to evaluate my graduate work. I specially would like to thank Drs. Berkland and Siahaan for reading my thesis. Thank you for all your constructive comments and suggestions. I would like to thank Dr. Gerald Lushington, the director of the Molecular Structures Group at University of Kansas for all the computational studies on the temperature dependent UV peak shifts of the aromatic amino acids which helped me tremendously in explaining the origin of such phenomenon. I would like to thank Dr. David Moore, the director of the Imaging Lab at University of Kansas for all the work on the rotavirus EM studies.

I would like to thank my family without whom I would have never made it. I would like to thank my father Mansour, my mother Farah, and my sister Rana for your undying support and for all the sacrifices you have always made for me. You are the most selfless, thoughtful, loving family I know and I thank God everyday for blessing me with angels like you. I would also like to thank my uncle Ramin and my aunt Jennifer. You have blessed me with your love and support for the past eight years and no words can describe my appreciation for everything you have done for me. So I simply say thank you to you all because I wouldn't have made it without you.

Sincerely, Reza Esfandiary

Table of Contents

Chapter 1: Introduction

Chapter 2: Structural stability of vault particles

Chapter 3: Structural stability of a novel fusion cytotoxin: Application of a high throughput excipient screening assay for development of a stable liquid formulation

Chapter 6: Conlcusion & future work

Index of Figures

35 °C. Model compounds used were N-acetyl-L-tyrosine ethyl ester and N-acetyl-L-phenylalanine ethyl ester. (n=3)

Index of Tables

List of Abbreviations

A_{280nm}	Absorbance at 280 nm
ANS	8-Anilino-1-Naphthalene Sulfonate
CD	Circular Dichroism
CR	Congo Red
CryoEM	Cryoelectron Microscopy
DLS	Dynamic Light Scattering
DMSO	Dimethyl Sulfoxide
DNA	Deoxy Ribonucleic Acid
EM	Electron Microscopy
EPD	Empirical Phase Diagram
FDA	Food and Drug Administration
FTIR	Fourier Transform Infrared
GBM	Glioblastoma Multiforme
GRAS	Generally Regarded as Safe
H/D	Hydrogen / Deuterium
HIV	Human Immunodeficiency Virus
IL-4	Interleukin 4
IL-13	Interleukin 13
IA	Inhibition of Aggregation

kDa	Kilo Dalton
MDa	Mega Dalton
MG	Molten Globule
MSD	Multi-Size Distribution
MVP	Major Vault Protein
NMR	Nuclear Magnetic Resonance
NNLS	Non-Negatively Constrained Least Square
OD_{350}	Optical Density at 350 nm
PE	Pseudomonas Exotoxin
PE4E	A Derivative of Pseudomonas Exotoxin
PHE	Phenylalanine
pI	Isoelectric Point
PTI	Photon Technology International
RGB	Red / Green / Blue
RNA	Ribonucleic Acid
RP-HPLC	Reversed Phase High Performance Liquid
TEM	Transmission Electron Microscopy
	Chromatography
TEP1	Telomerase-Associated Protein 1
ThT	Thioflavin T
T_m	Midpoint of Thermal Transition Temperature
TP40	Fusion Protein Composed of Transforming Growth

	Factor-α and a Derivative of Pseudomonas Exotoxin
TRP	Tryptophan
TYR	Tyrosine
UV	UltraViolet
VLP	Viral-Like Particles
VPARP	Vault Poly (ADP Ribose) Polymerase
vRNA	Vault RNA

Chapter 1

Introduction

1.1. Overview

Due to their large size and complexity, macromolecular-based therapeutics are prone to a variety of different physical and chemical degradations often leading to loss of their biological activity and altered immunogenicity[1]. Physical degradation primarily occurs due to protein structural changes, aggregation, and undesired adsorption to surfaces while chemical degradation refers to modification of covalent bonds such as oxidation, deamidation, and disulfide bond shuffling[2]. Both degradation pathways may occur simultaneously during different stages of manufacturing, storage, and shipment of protein drugs, introducing significant difficulties in their development into safe and efficacious therapeutics. Retaining storage stability is often a major challenge and a bottleneck in protein development due to the commonly required long protein shelf lives of 18-24 months[3].

Given the high degree of structural complexity and associated degradation pathways, a robust formulation (liquid or lyophilized) that retains the protein's stability and activity over its entire shelf life is essential. The FDA requires that the stability and activity of potential protein therapeutics be demonstrated in real time under the proposed labeled storage conditions. Considering the long protein shelf life, however, these types of studies demand significant resources be dedicated to them for extended periods of time. One time and cost effective approach is to conduct accelerated stability studies to screen for potential formulations with properties that lead to enhanced stability under more moderate storage conditions. Such accelerated studies are typically conducted under a variety of different stressed conditions (*e.g.*,

elevated temperatures, suboptimal pH, high or low ionic strengths, *etc*.) to facilitate protein degradation in a short period of time. Herein, we have employed a rapid and systematic approach for biophysical characterization of complex macromolecular systems under accelerated degradation conditions. This approach, referred to as the "Empirical phase diagram" (EPD approach) utilizes advanced algebraic procedures to integrate a large library of data obtained from a variety of biophysical techniques into an easy-to-interpret color coded map to provide a global picture of protein structural alterations under a wide range of solution conditions.

Employing this approach, we have characterized the stability of three complex macromolecular systems; a large recombinant ribonucleoprotein (chapter 2), a novel multi-domain fusion cytotoxin (chapter 3), and a live-attenuated double stranded RNA virus (chapter 4). We not only demonstrate the ability of this approach to provide a thorough in vitro characterization of such complex macromolecular systems but we also discuss the details obtained from such analyses that could be used to interpret the biological function of these systems in vivo. In certain cases, we use complementary techniques such as electron microscopy imaging and activity assays to further explain the importance of the structural transitions associated with such complex systems and highlight the biological significance and physiological relevance of the observed transitions.

1.2. Biophysical Characterization of Macromolecules

Development of a stable formulation that retains protein stability over its designated shelf life is a crucial step in the production of such complexes as safe and effective therapeutics. Rigorous design of such formulations requires an understanding of the macromolecule's behavior under a variety of stressed solution conditions. Demanding structural stability for proper functioning of protein therapeutics and optimization of conditions that maintain and/or enhance the stability of macromolecules is the main objective of pre-formulation studies. The information obtained from the EPD approach serves such a purpose by providing a basis for the development of high-throughput excipient screening assays due to its utility in identification of boundary regions over which degradation pathways are accelerated. Empirical phase diagrams as will be described in more detail later in this chapter are constructed based on a variety of biophysical (*i.e.*, often spectroscopic and calorimetric) techniques, a few of which are frequently used in these investigations are described below.

1.2.1. Circular Dichroism

Normal absorption spectroscopy of electronic transitions is measured as absorbance, A, according to Beer-Lambert law:

$$A = \mathrm{Log}_{10}\,(I_0/I) = \varepsilon c l$$

Where A is the measured absorbance signal, l is the length of the light pass, ε is the molar extinction coefficient, c is concentration, and I_0 and I are the intensity of the

incident and transmitted light respectively. Circular dichroism is a variation of electronic absorption spectroscopy that uses circularly polarized light rather than normal isotropic light and is defined as the difference in absorption of the left and right handed circularly polarized light due to the presence of optically active chromophores[4].

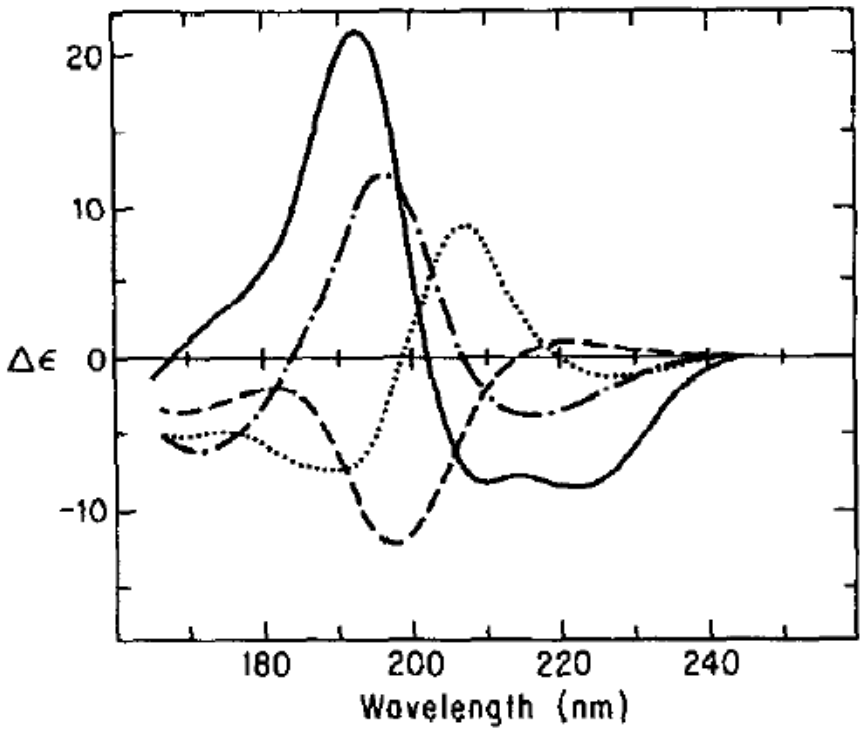


Figure 1.1.The CD spectra for the α-helix (solid line), β-sheet (dots and dashes), β-turn (dotted line), and random coil (dashed line).

The amide bond of the peptide backbone is the main chromophore used when analyzing CD signals due to sensitivity to various types of protein secondary structure. Different protein secondary structural elements exhibit distinct spectra in the range of 190-260 nm. For example, α-helices exhibit a strong double minima at 222 and 208-210 nm and a stronger maximum at 191-193 nm. β-sheet structures typically manifest a weaker single minimum between 215 and 217 nm and a stronger positive maximum between 195 and 200 nm[5]. Figure 1.1 shows distinct CD spectra of a variety of key secondary structural elements with α-helix (solid line), β-sheet (dots and dashes), β-turn (dotted line), and random coil (dashed line) illustrated. Alterations of protein secondary structure as a function of a variety of stressed solution variables can therefore be studied by monitoring changes in CD signals with often a loss of signal intensities or distortion of spectra and shifts in peak position as indicative of protein secondary structural alterations.

1.2.2. Intrinsic Fluorescence Spectroscopy

Intrinsic fluorescence spectroscopy is often used to study protein tertiary structural alterations. The three aromatic amino acids phenylalanine, tyrosine, and tryptophan all can contribute to protein fluorescence spectra. Phenylalanine has a very low quantum yield of ~ 0.03 and does not contribute significantly under most conditions. If excited below 270 nm, it exhibits a weak emission peak at ~ 282 nm if it is not obscured by the emission of tyrosine and tryptophan. Tyrosine has a quantum yield of ~ 0.14 (very close to that of tryptophan with a value of ~ 0.13) exhibiting an

emission maximum at ~ 303 nm. Unlike tryptophan, however, the emission spectra of both phenylalanine and tyrosine are rather insensitive to the polarity of their surrounding environment. Free tryptophan manifests an emission peak at longer wavelengths at 350-355 nm, and is very sensitive to solvent polarity. It accounts for most of a protein's fluorescence with excitation wavelengths above 290 nm[6]. The emission spectra of indole can reveal the location of the tryptophan residue in protein since the emission from an exposed tryptophan residue occurs at higher wavelengths than that of a buried one. Monitoring of such emission maxima is therefore utilized in studies of protein conformational alterations. Upon unfolding of proteins, the tryptophan residues become exposed to the polar solvent, resulting in shifts of their emission maximum to higher wavelengths (*i.e.*, red shifts). On the other hand, the burial of tryptophan residues due to the protein fold or from oligomerization or aggregation can cause blue shifts in the emission maximum.

The presence of multiple tryptophans, such as is often the case for complex macromolecular systems such as those examined here, can complicate the quantitative interpretation of the intrinsic fluorescence data since the microenvironment of each tryptophan residue is distinct but the emission of different tryptophan residues extensively overlap. This makes it impossible to separate the contribution of each individual residue. Although quantitative analyses of such systems appears to be information poor, qualitative studies based on the average peak shifts from a collective number of tryptophan residues are indicative of protein conformational alterations.

It is also possible to monitor light scattering data simultaneously with emission data by monitoring the scattered light seen at the excitation wavelength. This can be most easily achieved through the use of a second photomultiplier located at 180^{o} to the fluorescence detector. Increases in scattering signals are often indicative of aggregation or the presence of some form of higher order species contributing significantly to the scattered signals observed. Decreases are also seen due to the decrease in scattering produced by product of reduced refractive index.

1.2.3. Extrinsic Fluorescence Spectroscopy

In addition to the naturally occurring tryptophan chromophore, a variety of extrinsic probes can also be used to monitor protein structural alterations due to the sensitivity of their emission bands to the polarity of the surrounding microenvironments. The most commonly used extrinsic probe in protein unfolding studies is 8-anilino-1-naphthalenesulfonate (ANS). The fluorescence of the hydrophobic probe ANS is highly quenched in aqueous solution. It, however, exhibits strong fluorescence in less polar environments. Upon protein structural alterations, ANS may bind to exposed apolar patches on proteins causing blue shifts of its emission maximum and an increase in its fluorescence intensity[7,8]. Caution should be taken in interpretation of such data, however, since interaction between the negatively charged sulfonate group of ANS and positively charged residues on the protein can contribute to alterations of the ANS fluorescence signals observed[7,8].

Other types of extrinsic fluorescence probes are also commonly employed. For example, Thioflavin T and Congo red (CR) have been shown to bind to protein structures rich in intermolecular β-sheet content. Therefore, these dyes have been used as a tool for detection of amyloidogenesis during amyloid fibril formation in Alzheimer's disease[9,10]. Congo red selectively binds to proteins with extended intermolecular β-sheet conformations, a defining characteristic of amyloid fibrils. Inouye *et al.* have shown that binding of Congo red to amyloid fibrils occurs through both ionic and non-ionic processes. In the former, the negatively charged sulfonate groups of CR bind to positively charged histidine residues in a protein, and therefore, is a pH dependent event. The latter was proposed to be due to van der Waals interactions of the dye π system with the peptide[9,10]. Other macromolecules such as RNA-polymerase[11], dehydrogenases[12], cellulose[13], elastin[14], chitin[15], and HIV-1 protease[16] also bind to Congo red, making it an effective tool for the study of their conformational alterations.

Thioflavin T (ThT) extinction coefficient and the position of its long wavelength absorption band depend on solvent polarity[17]. Monitoring ThT absorbance in a series of water-glycerol mixtures shows red shifts of the long wavelength absorption band of ThT from 412 nm in water to 421 nm in 99% glycerol. The fibril-bound ThT exhibits a long-wavelength absorption band at ~ 450 nm, significantly red shifted compared to its position in water; therefore, it can be used to detect fibril formation, or perhaps other structures with highly rich β-sheet content.

1.2.4. Dynamic Light Scattering

Dynamic light scattering is a useful technique for determining the hydrodynamic size of the particles. Particles undergo Brownian motion in solution and therefore change their relative position as a function of time. Smaller particles diffuse quickly, generating scattering signals that fluctuate rapidly. Larger particles generate signals that fluctuate more slowly. The time dependence of these fluctuations can be characterized by an intensity autocorrelation function which is a function of the diffusion coefficient. This allows direct measurements of the latter. Assuming a uniform spherical shape for the particles, the Stokes-Einstein equation is then used for calculations of the hydrodynamic radius (R_h) based on the experimentally determined diffusion coefficient (D_T):

$$R_h = k_B.T / 6\pi\eta D_T$$

In which k_B is the Boltzman constant, T is the absolute temperature, and η is the solvent viscosity. The calculation of the R_h can be carried out using commercially available software packages equipped with multiple analyses tools. A common approach is the method of cumulants which is based on a polynomial fit to the log of the autocorrelation function[18]. For narrow size distributions, the autocorrelation function is adequately analyzed by the method of cumulants. For polydisperse or multimodal size distributions, however, the data analysis is much more complex[19]. The hydrodynamic diameter can also be calculated in a multi-size distribution (MSD) mode, using a non-negatively constrained least square (NNLS) algorithm in which the autocorrelation function is deconvoluted, generating resolved values comparable to

the actual particle size[19]. The choice of the analysis method depends on the nature and heterogeneity of the macromolecular system examined.

As mentioned previously, the Stokes-Einstein equation assumes spherical particles and therefore deviations in shape (*e.g.*, rod shape particles) could introduce significant deviations in obtaining the absolute size of the particles in solution. This technique, however, can be used qualitatively in studies of protein self-association and aggregation in which such processes often result in an increase in average particle size as observed by this technique. Employing the MSD approach involving deconvolution of multi-size distributions provides a very powerful quantitative tool in detection of very small amounts (less than a percent) of higher order species (*i.e.*, dimers, tetramers, *etc*.) as well as aggregates since small amount of larger size particles scatter light more effectively than a large amount of smaller size particles.

1.3. Construction of Empirical Phase Diagrams

A variety of biophysical techniques such as those described above, each sensitive to a certain structural feature of a protein, are used to obtain a complete picture of protein behavior. One of the potential drawbacks is that the interpretation of such large and complex data sets can be quite difficult. Therefore, integration of such large data sets into a unified picture is beneficial by providing the pharmaceutical scientist the opportunity to predict the state of the protein under all conditions of interest. An empirical phase diagram (EPD), based on a combination of a large data set provides an easy analysis by following the color-mapped image of a

protein's global behavior under a wide range of solution conditions. Employing such an approach, a macromolecule or its complexes can be represented by a vector in a highly dimensioned experimental space. The vector varies as a function of alteration in solution parameters such as pH, temperature, ionic strength, and macromolecule concentrations.

In brief, the macromolecule is represented as an *n* dimensional vector in an experimental space in which the n dimensions are constructed at any two combinations of stress conditions (*i.e.*, pH, temperature, ionic strength, protein concentration, *etc*.); *n* refers to the number of different data sets collected (*i.e.*, CD intensity, Trp peak position, ANS intensity, hydrodynamic diameter, *etc*.). Experimental data at each combination of stress variables (*e.g.*, pH/ temperature) serve as one component of an n dimensional vector. The data is normalized between -1 and 1 and are integrated into an *n* x *n* density matrix containing all the individual vectors. The *n* sets of eigenvalues and eigenvactors of the density matrix are calculated, and the three largest eigenvectors corresponding to the three largest eigenvalues are identified. The density matrix is then truncated and re-expanded into three dimensions with the three largest sets of eigenvectors corresponding to its basis. The resultant three dimensional vectors are then converted into a color-coded map in which each vector corresponds to a color using an arbitrary red/green/blue (RGB) color system. All calculations are performed using Matlab software (The MathWorks, Natick, MA). Details of the mathematical theory and calculation process can be found elsewhere[20].

Empirical phase diagrams can be constructed based on a single or multiple techniques. The first successfully utilized EPD was generated based on the single technique of second derivative absorbance peak positions of these aromatic residues[20]. Derivative analysis was used to deconvolute the raw data into multiple high resolution absorption peaks that could be used as semi-independent data sets in the EPD. In this case, integration of such data sets into an EPD is highly informative due to dispersion of the three aromatic residues throughout protein three dimensional structures and the high sensitivity of the derivative absorption peak positions to alterations in the surrounding microenvironment. Phenylalanine residues are often buried in the core of proteins while tyrosines are interfacial and tryptophan residues are more widely dispersed. Employing this approach provides a global picture of a protein's behavior in solution.

The EPD approach has been successfully employed to examine a wide variety of macromolecular entities including peptides, proteins, recombinant protein based vaccines, viruses and virus like particle (VLP) based vaccines, DNA, bacterial vaccines, and gene delivery vectors[20-35]. Employing multiple techniques provides a more detailed picture by integration of more independent data sets, each sensitive to a certain feature of protein structure. As implemented in the three case studies presented in this thesis, a combination of circular dichroism, intrinsic and extrinsic fluorescence, and static and dynamic light scattering techniques generates an EPD reflecting protein secondary, tertiary, and quaternary structural alterations as well as aggregation tendencies. In either case, when analyzing an EPD, regions of continuous

color define distinct physical states of the macromolecular system under the stress conditions investigated. Abrupt changes in color, however, are indicative of some form of protein structural alteration and correspond to transitions between various physical states of the macromolecule. It should also be emphasized that empirical phase diagrams are not thermodynamic phase diagrams in which equilibrium exists between different phases. At least partially irreversible aggregation or conformational change prevents any equilibrium among different phases.

1.4. High Throughput Screening Assays

The abrupt changes in color in the EPD represent apparent phase boundaries over which protein shows marginal stability. High throughput screening assays to identify potential stabilizer(s) of macromolecules can therefore be developed over such regions in which protein degradation is accelerated. A large library of potential excipients is screened and more promising stabilizers are identified. The compounds tested are preferably selected from a large library of GRAS (Generally Regarded as Safe) stabilizers approved by the FDA for therapeutic use. Non-GRAS stabilizers such as polyanions and other low molecular weight di- and multi-ions are also screened. Although applications of the latter are not yet used in FDA approved protein-based drug formulations, studying the mechanisms by which they affect protein stability provides insight into protein stabilization strategies. Examined excipients are often tested over a reasonable range of concentration (*e.g.,* 1-20%) and in triplicate to ensure reproducibility of the results. Identification of such apparent

phase boundaries then serves as a basis for development of high throughput screening assays to identify potential stabilizers.

A rational high-throughput screening assay is often developed as the first step of multi-iterative optimization of protein formulation. The preliminary screening step usually monitors macromolecule aggregation kinetics in which the inhibitory effect of a given excipient on protein aggregation is examined as a function of time over some marginally stable condition identified in the EPD (*e.g.*, extremes of pH and elevated temperatures). In the second stage, stabilizers that show early promise as aggregation inhibitors are then examined further over a wider range of concentrations as well as in combination with other promising excipients to see if there is an additive or synergistic stabilizing effect of these compounds. Optimization of the selected excipients is then verified in stage three by testing their ability to stabilize protein conformation. This step is crucial since not all excipients which are effective inhibitors of protein aggregation necessarily stabilize its conformation. This is observed in these studies on the fusion cytotoxin (see chapter 3) in which polyanions extensively inhibit aggregation of the protein but lead to significant destabilization of its tertiary structure. In this last step, a protein's apparent thermal transition temperature(s) (T_m) in the presence and absence of excipients is determined where an effective excipient results in shifting of the T_m values to higher temperatures. The stabilizer combination and their final concentrations are further optimized to provide a suitable liquid formulation that is compatible with the macromolecule's therapeutic application and its route of administration. In addition, the ionic strength, osmolarity,

and viscosity of the protein formulation are also adjusted as a result of presence of such excipients.

Although such accelerated stability studies provide a rapid and systematic approach for identification of potential stable protein formulations under artificially induced degradation conditions, the true ability of such formulations to provide real time stability should be further examined in actual long-term stability studies. This is due to the possible presence of different degradation mechanisms under accelerated versus more moderate storage conditions[36]. The chemical stability and biological activity of potential formulations from preliminary accelerated studies should be tested to ensure that there is no degradation of the protein or the additives as well as no undesired effects of the excipients on the biological activity of the protein during long term storage. The high throughput screening approach described here has been successfully used for the formulation of a variety of macromolecules such as peptides, proteins, viruses, and virus like particles[1,37-39].

1.5. Protein Dynamics

In the past, proteins were often perceived as relatively static structures based on their X-ray analysis. It is now well accepted, however, that proteins are highly dynamic molecules exhibiting a wide variety of different types of motions in which both their static and dynamic properties as well as the nature of the surrounding environment affects their structure and function. Central questions as to how protein dynamics affect their structure and function, and how environmental factors affect

protein dynamical-structural-functional properties have engendered great interest. Various protein functions such as substrate binding, product release, regulation, allosteric behavior, as well as contractile and motor functions are heavily controlled by collective motions within proteins[40]. Fluctuations of the surrounding solvent molecules have been shown to control protein dynamic motions, engendering the use of the term "slaving" to describe the significance of the dominant role of solvent in controlling protein dynamics and function[41].

The relationship between protein dynamics and stability, although still somewhat unclear, is a point of great interest with extensive studies linking the two[42]. Such studies suggest that analyses of protein flexibility along with its static properties are necessary for successful characterization of protein structural stability. Therefore, the availability of techniques that detect extensive and subtle conformational alterations of proteins, as well as structural fluctuations appears to be crucial. Herein, we have presented the EPD approach based on a variety of biophysical techniques that are sensitive to conformational alterations and aggregation (*i.e.*, static properties) of macromolecules. A variety of biophysical methods have been used to measure various aspects of protein structure from which dynamic behavior can be inferred. These techniques include isotope (H/D) exchange monitored by infrared spectroscopy, mass spectrometry, and nuclear magnetic resonance (NMR) spectroscopy as well as other methods such as red edge shift spectroscopy, solute fluorescence quenching, ultrasonic spectroscopy, pressure perturbation calorimetry and molecular dynamics simulations[42].

Recent efforts have employed data from spectroscopic techniques sensitive to protein dynamic properties for construction of empirical phase diagrams sensitive to transitions due to alterations of protein motions. Such a dynamic-based approach revealed that dynamic EPDs provide information above and beyond that obtained by the static approach[43]. As part of efforts to utilize available spectroscopic tools in a better understanding of protein dynamic properties, for the first time we here examine the utility of the temperature dependent UV absorption spectroscopy of aromatic residues in protein pre-transition regions to show that information obtained from such analyses hold valuable information regarding protein dynamic fluctuations. Details will be found in chapter 5 but here a brief overview of UV absorption spectroscopy of proteins is presented.

1.5.1. Ultraviolet Absorption Spectroscopy of Proteins

1.5.1.1. Protein Intrinsic Chromophores

Proteins display a broad peak in the 250-300 nm region of their ultraviolet spectrum composed of multiple overlapping bands from the aromatic residues phenylalanine, tyrosine, and tryptophan. These peaks are due primarily to $\pi \rightarrow \pi^*$ transitions involving the electrons of their aromatic rings[44]. The indole side chain of tryptophan absorbs strongly (ε_{280nm} = 5540) with a maximum peak at ~ 280 nm and a less intense transition, usually observed as a shoulder at ~ 292 nm. Alterations of tryptophan absorption spectra could occur due to oxidation of the indole moiety, which contributes to the change in the wavelength[45,46]. The absorption of tyrosine is less intense (ε_{280nm} = 1480) with maximum absorbance at ~ 276 nm and two small

shoulders. Tyrosine is even more susceptible to structural modifications with the phenol side chain susceptible to ionization[47], iodination[48], chlorination[48], and oxidation[46]. Finally, the phenylalanine side chain, benzene, exhibits the weakest transition (ε_{258nm} = 197) in the 250-270 nm range, usually appearing as multiple subtle inflection points (*i.e.*, bumps) in the near-UV region with a peak center at ~ 260 nm. When analyzing the zero-order spectra, phenylalanine weak absorbance is often obscured in the presence of tyrosine and tryptophan residues demanding derivative analysis for recognition of its contribution.

The most common chromophore of proteins is perhaps the amide group of the peptide backbone. It contains two major transitions with the first strongly absorbing at 195 nm ($\pi \rightarrow \pi^*$) and a second weaker transition occurring at ~ 220 nm ($n \rightarrow \pi^*$). With α-helix, β-sheet, and random coil structures each exhibiting distinct absorbance characteristics in the far UV region, the peptide group was in the past used to study protein secondary structure[49]. The associated data interpretation is, however, difficult due to interferences by inorganic ions and dissolved oxygen absorbing below 200 nm[48] and requires purging of spectrophotometers with nitrogen. Availability of other sensitive techniques (*e.g.*, circular dichroism and FTIR spectroscopy) to protein secondary structural alterations has also shifted the utility of the ultraviolet absorption spectroscopy to studies of protein tertiary structure in the 250-300 nm region where the aromatic residues show distinct absorbance characteristics.

A non-aromatic amino acid contributing appreciably in the 250-300 nm region of the protein absorption spectrum is the disulfide-bridge form of cysteine (*i.e.*,

cystine)[50] and the presence of such should be considered. Other residues such as methionine and histidine absorb weakly in the lower wavelength regions but their contributions are usually hampered by the much stronger absorption bands of the peptide backbone and aromatic side chains[48].

1.5.1.2. Data Analysis

The application of UV absorption spectroscopy to proteins was initiated more than half a century ago by analysis of their "zero order spectra" at relatively low resolution[51]. Zero order spectra represent the raw data collected by the spectrophotometer. In the case of protein studies, it typically consists of a strong transition below 210 nm from the amide group of the peptide backbone followed by a weaker transition in the 250-300 nm region, composed of multiple overlapping bands from the aromatic residues phenylalanine, tyrosine, and tryptophan[44]. As mentioned above, the utility of the amide backbone peak has been of limited use due to interferences from inorganic ions and dissolved oxygen. It is now well understood, however, that peak shifts of the aromatic residues correlate with protein structural perturbations due to their sensitivity to the polarity of the local environment of these side chains. Until early 80s, the utility of the aromatic residues was quite limited due to the extensive overlap of their absorption peaks and only major structural changes such as extensive unfolding events were routinely studied.

Derivatization of the zero-order data in the 250-300 nm region of the UV spectrum resolves the overlapping peaks of phenylalanine, tyrosine, and tryptophan to

allow qualitative analysis of the contribution of each residue to the overall protein spectrum. The limitations of early spectrophotometers, however, allowed only detection of ~1 nm shifts of the absorbance derivative peaks[52]. Following advances in instrumentation, in particular the availability of diode array detectors and supplementing the derivative analysis with interpolation algorithms[53] the resolution of the absorption bands of each of the three aromatic residues were greatly enhanced. Our experience shows that resolution as high as 0.01 nm can be obtained when employing a nine-data point filter and fifth degree Savitzky-Golay polynomial, following by fitting to a cubic function, with 99 interpolated points per raw data point. This provides a powerful tool for probing protein conformational alterations by monitoring the derivative peak shifts of the aromatic residues as a function of variety of stressed conditions. Furthermore, the intensity of these derivative peaks has been employed as a quantitative probe of the number of phenylalanine and tryptophan residues[54], as a measure of Tyr/Trp ratios, and as a qualitative measure of solvent polarity surrounding the aromatic side chains[55]. Because derivative analysis is not sensitive to broad spectral components, light-scattering effects do not dramatically affect determination of the derivative peak positions[56]. Figure 1.2 illustrates the high resolving power of the derivative analysis (right panel) in which the broad, low resolution zero-order absorption peaks of the three aromatic residues (left panel) have been deconvoluted to their comprising peaks exhibited as minima in the 2nd derivative plots.

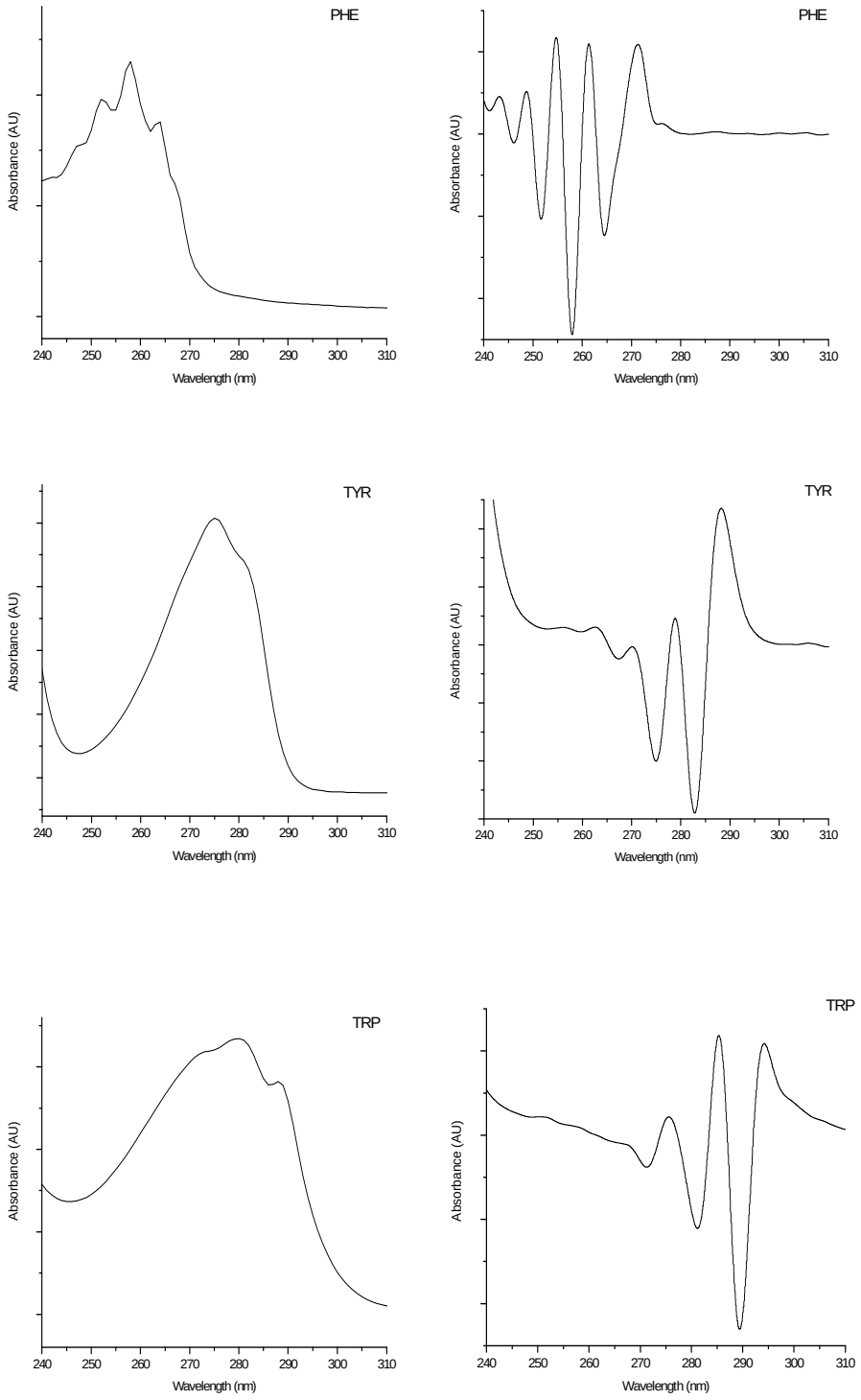


Figure 1.2.Zero-order and second derivative absorption spectra of N-acetyl-X-ethyl ester derivatives of and phenylalanine (PHE), tyrosine (TYR), and tryptophan (TRP). The ordinate axis is presented in arbitrary units.

1.5.2. Studies of Protein Conformational Alterations (Static Properties)

With the availability of diode array detectors and employing high resolution 2nd derivative analysis supplemented with interpolation algorithms, the absorption bands of each of the three aromatic residues can be deconvoluted with a resolution of approximately 0.01 nm. This provides a very sensitive tool with which to probe protein conformational alterations in which the polarity of the microenvironment surrounding the aromatic side chains and their level of exposure to solvent are detectably altered upon protein structural changes. Therefore, by monitoring the individual shifts of the derivative peak position of these residues, fairly detailed information can be obtained regarding the conformational alterations of proteins as a function of variety of solution conditions such as pH, temperature, ionic strength, presence of additives, *etc*. In general, peak shifts to lower wavelengths (*i.e.*, blue shifts) are indicative of the aromatic side chains exposure to a more polar solvent environment whereas shifts to longer wavelengths (*i.e.*, red shifts) manifest burial of these residues. Employing this technique, shifts as small as 0.1 nm and as large as 6 nm have been correlated with protein structural alterations[57,58].

1.5.3. Studies of Protein Dynamic Properties

Recent efforts have attempted to extend the utility of the 2nd derivative UV absorbance spectroscopy of the aromatic residues beyond protein static studies into the world of protein dynamics. This initial work was initiated by examination of the pioneering idea of monitoring cation interactions with the π systems of the aromatic

side chains employing high resolution 2nd derivative UV absorbance spectroscopy. By analogy to solute-based fluorescence quenching of proteins, Lucas *et al.* performed UV-spectroscopy dynamic-based analyses by employing the spectral shifts induced by cation (Na^+, Li^+, Cs^+)-π interactions as a function of increasing cation concentration[45]. Small cations were found to be more effective at inducing spectral shifts due to their ability to diffuse through a protein's matrix and make contact with the aromatic side chains. In some cases, interpretation of dynamic effects was, however, difficult due to specific interactions between the cations (and perhaps accompanying counter anions, Cl^-) with the proteins[44]. Thus, a method that does not require the presence of potentially perturbing solutes, but rather involves a simple intrinsic effect such as temperature seems desirable. This is the motivation for the studies presented in chapter 5 on the origin of the temperature dependent 2nd derivative peak shifts of the aromatic residues and integration of such information to qualitatively probe protein dynamic motions.

1.6. Bibliography

(1) Ausar SF, Espina M, Brock J, Thyagarayapuran N, Repetto R, Khandke L, Middaugh CR. High-throughput screening of stabilizers for respiratory syncytial virus: identification of stabilizers and their effects on the conformational thermostability of viral particles. Human Vaccines (2007), 3, 94-103.

(2) Chi EY, Krishnan S, Randolph TW, Carpenter JF. Physical stability of proteins in aqueous solution: Mechanism and driving forces in nonnative protein aggregation. Pharm. Res. (2003), 20, 1325-1336.

(3) Randolph TW, Carpenter JF. Engineering challenges of protein formulations. AlChE J. (2007), 53, 1902-1907.

(4) Johnson WC. Secondary structure of proteins through circular dichroism spectroscopy. Annu. Rev. Biophys. Biomol. Struct. (1988), 17, 145-66.

(5) Venyaminov SY, Yang JT. Determination of protein secondary structure in Circular dichroism and the conformational analysis of biomolecules (Fasman, G. D., Ed.) Plenum Press, New York (1996), pp. 69-107

(6) Lakowicz JR. Principles of Fluorescence Spectroscopy. 2nd ed, Kluwer Academic / Plenum Publisher (New York, NY) (1999), 698 pp.

(7) Matulis D, Baumann CG, Bloomfield VA, Lovrien RE. 1-Anilino-8-Naphthalene Sulfonate as a protein conformational tightening agent. Biopolymers (1999), 49, 451-458.

(8) Matulis D, Lovrien RE. 1-Anilino-8-Naphthalene Sulfonate anion-protein binding depends primarily on ion pair formation. Biophys. J. (1998), 74, 422-429.

(9) Inouye H, Kirschner DA. Alzheimer 's β-amyloid: Insights into fibril formation and structure from congo red binding. Subcellular Biochemistry (2005), 38, 203-224.

(10) Inouye H, Kirschner DA. Aβ Fibrillogenesis: Kinetic parameters for fibril formation from Congo Red binding. J. Struct. Biol. (2000), 130, 123-129.

(11) Woody A, Young M, Reisbig RR, Woody RW. Spectroscopic studies of Congo Red binding to RNA polymerase. Biochim. Biophys. Acta. Nucleic Acids Protein Synth. (1981), 655, 82-8.

(12) Edwards RA, Woody RW. Spectroscopic studies of Cibacron Blue and Congo Red bound to dehydrogenases and kinases. Evaluation of dyes as probes of the dinucleotide fold. Biochemistry (1979), 18, 5197-204.

(13) Quenin I, Henrissat B. Precipitation and crystallization of cellulose doped with dyes. Makromolekulare Chemie, Rapid Commun. (1985), 6, 737-41.

(14) Tan OTian, Faris B, Franzblau C. Aortic elastin fluorescence after in vivo labeling with Congo Red. Journal of Fluorescence (1991), 1, 147-51.

(15) Bartnicki-Garcia S, Persson J, Chanzy H. An electron microscope and electron diffraction study of the effect of calcofluor and Congo red on the biosynthesis of chitin in vitro. Arch. Biochem. Biophys. (1994), 310, 6-15.

(16) Ojala WH, Ojala CR, Gleason WB. The x-ray crystal structure of the sulfonated azo dye Congo Red, a non-peptidic inhibitor of HIV-1 protease which also binds to reverse transcriptase and amyloid proteins. Antiviral Chem. Chemother. (1995), 6, 25-33.

(17) Maskevich AA, Stsiapura VI, Kuzmitsky VA; Kuznetsova IM, Povarova OI, Uversky VN, Turoverov KK. Spectral properties of Thioflavin T in solvents with different dielectric properties and in a fibril-incorporated form. J. Proteome Res. (2007), 6, 1392-1401.

(18) Brown JC, Pusey PN. Photon correlation study of polydisperse samples of polystyrene in cyclohexane. J. Chem. Phys. (1975), 62, 1136-44.

(19) Morrison ID, Grabowski EF, Herb CA. Improved techniques for particle size determination by Quasi-Elastic light scattering. Langmuir (1985), 1, 496-501.

(20) Kueltzo LA, Ersoy B, Ralston J, Middaugh CR. Derivative absorbance spectroscopy and protein phase diagrams as tools for comprehensive protein characterization: A bGCSF case study. J. Pharm. Sci. (2003), 92, 1805-1820.

(21) Fan H, Vitharana SN, Chen T, O'Keefe D, Middaugh CR. Effects of pH and polyanions on the thermal stability of fibroblast growth factor 20. Mol. Pharm. (2007), 4, 232-240.

(22) Fan H, Kashi RS, Middaugh CR. Conformational lability of two molecular chaperones Hsc70 and GP96: effects of pH and temperature. Arch. Biochem. Biophys. (2006), 447, 34-45.

(23) Fan H, Ralston J, DiBase M, Faulkner E, Middaugh CR. Solution behavior of IFN-β-1a: An empirical phase diagram based approach. J. Pharm. Sci. (2005), 94, 1893-1911.

(24) Fan H, Li H, Zhang M; Middaugh CR. Effects of solutes on empirical phase diagrams of human fibroblast growth factor 1. J. Pharm. Sci. (2007), 96, 1490-1503.

(25) Jiang G, Joshi SB, Peek LJ, Brandau DT, Huang CR, Ferriter MS, Woodley WD, Ford BM, Mar KD, Mikszta JA, Hwang CR, Ulrich R, Harvey NG, Middaugh CR, Sullivan VJ. Anthrax vaccine powder formulations for nasal mucosal delivery. J. Pharm. Sci. (2006), 95, 80-96.

(26) Peek LJ, Brey RN, Middaugh CR. A rapid, three-step process for the preformulation of a recombinant ricin toxin A-chain vaccine. J. Pharm. Sci. (2007), 96, 44-60.

(27) Brandau DT, Joshi SB, Smalter AM, Kim S, Steadman B, Middaugh CR. Stability of the Clostridium botulinum type A neurotoxin complex: An empirical phase diagram based approach. Mol. Pharm. (2007), 4, 571-582.

(28) Ausar SF, Rexroad J, Frolov VG, Look JL, Konar N, Middaugh CR. Analysis of the thermal and pH stability of human respiratory syncytial virus. Mol. Pharm. (2005), 2, 491-499.

(29) Ausar SF, Foubert TR, Hudson MH, Vedvick TS, Middaugh CR. Conformational stability and disassembly of Norwalk virus like particles: Effect of pH and temperature. J. Biol. Chem. (2006), 281, 19478-19488.

(30) Harn N, Allan C, Oliver C, Middaugh CR. Highly Concentrated Monoclonal Antibodies: Direct Analysis of Structure and Stability. J. Pharm. Sci. (2007), 96, 532-546.

(31) Peek LJ, Brandau DT, Jones LS, Joshi SB, Middaugh CR. A systematic approach to stabilizing EBA-175 RII-NG for use as a malaria vaccine. Vaccine (2006), 24, 5839-5851.

(32) Rexroad J, Martin TT, McNeilly D, Godwin S, Middaugh CR. Thermal stability of Adenovirus type 2 as a function of pH. J. Pharm. Sci. (2006), 95, 1469-1479.

(33) Rexroad J, Evans RK, Middaugh CR. Effect of pH and ionic strength on the physical stability of adenovirus type 5. J. Pharm. Sci. (2006), 95, 237-247.

(34) Ruponen M, Braun CS, Middaugh CR. Biophysical characterization of polymeric and liposomal gene delivery systems using empirical phase diagrams. J. Pharm. Sci. (2006), 95, 2101-2114.

(35) Thyagrajapuram N, Olsen D, Middaugh CR. The structure stability and complex behavior of recombinant human gelatins. J. Pharm. Sci. (2007), 96, 3363-3378.

(36) Yoshioka S, Stella, VJ. Stability of drugs and dosage forms, Kluwer Academic/Plenum Publishers, New York (2000), 187-199.

(37) Jiang G, Joshi SB, Peek LJ, Brandau DT, Huang CR, Ferriter MS, Woodley WD, Ford BM, Mar KD, Mikszta JA, Hwang CR, Ulrich R, Harvey NG, Middaugh CR, Sullivan VJ. Anthrax vaccine powder formulations for nasal mucosal delivery. J. Pharm. Sci. (2006), 95, 80-96.

(38) Peek LJ, Brey RN, Middaugh CR. A rapid, three-step process for the preformulation of a recombinant ricin toxin A-chain vaccine. J. Pharm. Sci. (2007), 96, 44-60.

(39) Joshi SB; Kamerzell TJ, McNown C, Middaugh CR. The interaction of heparin/polyanions with bovine, porcine, and human growth hormone. J. Pharm. Sci. (2008), 97, 1368-1385.

(40) Berendsen HJ, Hayward S. Collective protein dynamics in relation to function. Curr. Opin. Struct. Biol (2000), 10, 165-169.

(41) Fenimore PW, Frauenfelder H, McMahon BH, Parak FG. Slaving: Solvent fluctuations dominate protein dynamics and functions. Proc. Nat. Acad. Sci. U.S.A. (2002), 99, 16047-51.

(42) Kamerzell TJ, Middaugh CR. The complex inter-relationships between protein flexibility and stability. J. Pharm. Sci. (2008), 97, 3494-3517.

(43) Ramsey JD, Gill ML, Kamerzell TJ, Price ES, Joshi SB, Bishop SM, Oliver CN, Middaugh CR. Using empirical phase diagrams to understand the role of intramolecular dynamics in immunoglobulin G stability. J. Pharm. Sci. (2009), 98, 2432-2447.

(44) Lucas LH, Ersoy BA, Kueltzo LA, Joshi SB, Brandau DT, Thyagarajapuram N, Peek LJ, Middaugh CR. Probing protein structure and dynamics by second-derivative ultraviolet absorption analysis of cation-π interactions. Protein Sci. (2006), 15, 2228-2243.

(45) Zhang ZH, Smith DL. Smith JB. Human β-crystallins modified by backbone cleavage, deamidation and oxidation are prone to associate. Exp. Eye Res. (2003), 77, 259-272.

(46) Ramachandran LK, Witkop B. Selective cleavage of C-tryptophyl peptide bonds in proteins and peptides. J. Am. Chem. Soc. (1959), 81, 4028-32.

(47) Creighton TE. Proteins: Structures and Molecular Properties, 2nd ed. W.H. Freeman and company (New York, NY) (1993), 507 pp.

(48) Wetlaufer DB. Ultraviolet spectra of proteins and amino acids. Adv. Protein Chem. (1962), 17, 303-90.

(49) Mach H, Volkin DB, Burke CJ, Middaugh CR. Ultraviolet absorption spectroscopy. Methods Mol. Biol (Totowa, NJ, United States) (1995), 40, 91-114.

(50) McGlynn SP, Chaudhuri JN, Good M. Possible effect of charge transfer complexation on the dihedral angle of dialkyl disulfides. J. Am. Chem. Soc. (1962), 84, 9-12.

(51) Beaven GH, Holiday ER. Ultraviolet absorption spectra of proteins and amino acids. Adv. Protein Chem (Academic Press Inc., New York, N.Y.) (1952), 7, 319-86.

(52) Ichikawa T, Terada H. Estimation of state and amount of phenylalanine residues in proteins by second derivative spectrophotometry. Biochimica et Biophysica Acta, Protein Struct. (1979), 580, 120-8.

(53) Mach H, Sanyal G, Volkin DB, Middaugh CR. Applications of ultraviolet absorption spectroscopy to the analysis of biopharmaceuticals. ACS Symposium Series (1997), 675, 186-205.

(54) Balestrieri C, Colonna G, Giovane A, Irace G, Servillo L. Second-derivative spectroscopy of proteins. A method for the quantitative determination of aromatic amino acids in proteins. Eur. J. Biochem. (1978), 90, 433-40.

(55) Servillo L, Colonna G, Balestrieri C, Ragone R, Irace G. Simultaneous determination of tyrosine and tryptophan residues in proteins by second-derivative spectroscopy. Anal. Biochem. (1982), 126, 251-7.

(56) Mach H, Thomson JA, Middaugh CR, Lewis RV. Examination of phenylalanine microenvironments in proteins by second-derivative absorption spectroscopy. Arch. Biochem. Biophys. (1991), 287, 33-40.

(57) Kueltzo LA, Middaugh CR. Ultraviolet absorption spectroscopy. Biotechnology: Pharmaceutical Aspects (2005), 3, 1-25.

(58) Lange R, Balny C. UV-visible derivative spectroscopy under high pressure. Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol. (2002), 1595, 80-93.

Chapter 2

Structural Stability of Vault Particles

2.1. Introduction

Some twenty years ago, Kedersha and Rome purified a large ribonucleoprotein particle from rat liver homogenates and named it the "vault" particle on the basis of its morphological resemblance to vaulted ceilings in medieval cathedrals.[1-2] Cryoelectron microscopy (CryoEM) single particle reconstructions and X-ray crystallography show vaults to posses a hollow, barrel-like structure with two protruding caps and an invaginated waist.[3-4] With a molecular weight of 12.9±1 MDa and dimensions of ~ 420 x 420 x 750 Aº, vaults are the largest ribonucleoproteins known.[2] Conservation of vaults among eukaryotes as diverse as mammals, amphibians, avians, sea urchins, and slime molds[5] points to an important functional role for these unique particles. Although the natural function of the vault complex is still unclear, their morphology and subcellular localization suggest some type of role in intracellular (*e.g.*, nucleocytoplasmic) transport.[6] The presence of these particles has also been linked to multidrug resistance in tumor cells since high level expression of the major vault protein is observed in certain transformed cell lines that are resistant to xenobiotics.[7-10] In addition, vaults may play a role in scaffolding different cell signaling pathways[11] and in protection against some forms of infectious disease.[12]

Vaults are composed of multiple copies of three proteins: 96 copies of the 97 kDa major vault protein (MVP), which accounts for more than 70% of the particle mass, 2 copies of the 290 kDa telomerase-associated protein 1 (TEP1) and 8 copies of the 193 kDa poly (ADP ribose) polymerase (VPARP). In addition, at least 6 copies of an untranslated small RNA (vRNA) are present.[10] The potential functions of vault

proteins [2, 10,13] and their interactions with one another are discussed in detail elsewhere.[6,11]

In vitro, expression of MVP alone in Sf9 insect cells employing a bac-to-bac baculovirus expression system results in the production of recombinant particles with the characteristic vault morphology.[5,13] The most reproducible form of recombinant vaults and the model system employed in this study is the Cys-rich tagged (CP-MVP) vault which contains 96 copies of the major vault protein modified at the N-terminus with a 12 amino acid peptide tag (MAGCGCPCGCGA) from the metal binding protein, metallothionine.[14] The CP-MVP form of vaults has been used to generate high resolution CryoEM images (16 Aº) and appears to be more stable than other recombinant forms, presumably due to the Cys residues of the tag forming disulfide bridges at the vault's waist, stabilizing the overall vault's structure.[14]

Recent findings concerning the dynamic structure of the vault exterior shell[15] and dissociation of the vault particle into halves at low pH[14] along with successful attempts to target and sequester biologically active materials within the vault cavity[3] have engendered interest in employing recombinant vaults as nanocapsules for the delivery of biomolecules. Since one of the major issues facing the field of bionanotechnology is cellular compatibility, using the cell compatible, naturally occurring vault nanocapsules as drug delivery devices could potentially have unique therapeutic applications.

Characterization of the physical stability of vault particles under a variety of solution conditions has the potential to provide important information concerning

vault's structural integrity and their potential use as drug delivery vehicles. With this goal in mind, we have subjected recombinant vaults to a wide range of pH and temperature employing a variety of spectroscopic techniques and electron microscopy (EM). This data has been integrated into an empirical phase diagram (EPD) to provide a global view of vault's secondary and tertiary conformational alterations over a wide range of experimental conditions.[16-17]

2.2. Materials and Methods

2.2.1. Vault Purification

Recombinant vaults formed from CP-MVP were purified from baculovirus infected Sf9 cells, as previously described.[5,13] A 12-residue cysteine rich motif (MAGCGCPCGCGA) on the N-terminus of CP-MVP was previously shown to help stabilize the recombinant vaults.[4,13] Purified vault particles were resuspended in 20 mM citrate-phosphate buffer pH 6.5.

2.2.2. Sample Preparation

Vault solutions at different pH values were prepared in 20 mM isotonic citrate/phosphate buffer by dialyzing stock solutions into buffers ranging from pH 3 to 8, at one unit pH intervals. The isotinicity was maintained using sodium chloride. For buffer exchange, vault samples were dialyzed at refrigerator temperatures using Slide-A-Lyzer® Dialysis Cassettes, 10 kDa MWCO (Pierce, Rockford, IL). Vault solutions were studied at a concentration of 100 μg/ml with the exception of ANS

fluorescence studies where 50 μg/ml vault samples were used. The concentrations of vault samples after dialysis were measured by UV absorption spectroscopy (A_{280nm}, $E_{1cm}^{0.1\%} = 1.039$).[14] Three independent samples were evaluated to ensure reproducibility of the measurements.

2.2.3. Transmission Electron Microscopy

Transmission electron microscopy of uranyl acetate-stained vaults was carried out to analyze structural alterations due to changes in pH and/or temperature. Briefly, purified vaults (0.5-1 mg/ml) were dialyzed into isotonic 20 mM citrate phosphate buffer at pH 3, 5, 6, or 8 for about 16 h at 4 °C. The dialyzed vaults were recovered and conformational changes were analyzed under variable temperatures by negatively staining with uranyl acetate and viewed with a JEM1200-EX transmission electron microscope (JEOL, Tokyo, Japan). Vault samples were absorbed onto carbon coated grids and after incubating for 5 min at the appropriate temperature and pH, the EM grids with absorbed vaults were blotted on filter paper and floated on 1% uranyl acetate at the appropriate temperature for staining. They were then blotted and dried on filter paper. TEM images were captured with a BioScan 600W digital camera (Gatan Inc., Pleasanton, CA) using Gatan's DigitalMicrograph (version 3.7.1).

2.2.4. Far-UV Circular Dichroism (CD) Spectroscopy

CD spectra were acquired using a Jasco J-810 spectropolarimeter equipped with a 6-position Peltier temperature controller. The CD spectra were obtained from

260-190 nm with a scanning speed of 20 nm/min, a 2 sec response time and an accumulation of 3. To study thermal transitions (melting curves) of the vaults, the CD signal at 222 nm was monitored in 0.1 cm pathlength cuvettes every 0.5 °C over a 10 to 85 °C temperature range employing a temperature ramp of 15 °C/hr.

2.2.5. Intrinsic Tryptophan (Trp) Fluorescence Spectroscopy

Fluorescence spectra were acquired using a Photon Technology International (PTI) spectrofluorometer (Lawrenceville, NJ) equipped with a turreted 4-position Peltier-controlled cell holder. An excitation wavelength of 295 nm was used to primarily excite Trp residues and the emission spectra were collected from 310 to 400 nm with a step size of 1 nm and a 1 sec integration time. Excitation slits were varied over the pH range examined from 3 to 5 nm. Light scattering was also monitored at 295 nm using a separate photomultiplier placed at 180º to the fluorescence detector. Both the fluorescence intensity and light scattering data were normalized with respect to the initial reading at 10 °C to permit direct comparison of the data. Emission spectra employing 4 nm slit widths were collected every 2.5 °C with a 3 min equilibration time over a temperature range of 10 to 85 °C. A buffer baseline was subtracted from each raw emission spectrum.

Peak positions of the emission spectra were obtained from polynomial fits using Origin software using a “center of spectral mass” method. Due to the nature of center of spectral mass analysis, the reported emission peak position values do not correspond to the actual peak positions, but do accurately reflect the changes. The

actual peak positions determined by derivative analysis of the native protein are shifted approximately 10-14 nm from their center of mass values.

2.2.6. ANS Fluorescence Spectroscopy

Accessibility of apolar sites on vaults monitored by fluorescence emission of the extrinsic probe 8-Anilino-1-naphthalene sulfonate (ANS). Each sample contained a 10-fold molar excess of ANS to vault, an optimal ratio determined in preliminary experiments. The ANS was excited at 385 nm and emission spectra were collected from 425-550 nm with a step size of 1 nm and 1 sec integration time. Emission spectra were collected every 2.5 °C with 3 min of equilibration over a temperature range of 10 to 85 °C. Both excitation and emission slits were varied over the pH range examined from 6 nm for pH 3 and 4 to 9 nm for pH 5-8; the intensity is quantitatively dependent on the excitation source intensity, which is controlled by the amount of light passed through the slits. The more open the slits, the higher the excitation source intensity and therefore the higher the fluorescence intensity. Thus, data were normalized with respect to the initial reading at 10 °C to permit direct comparison. The ANS-buffer baseline containing almost no fluorescence at each corresponding pH was subtracted from raw emission spectra.

2.2.7. Empirical Phase Diagram (EPD)

Empirical phase diagrams were constructed employing CD mean residue ellipticity, intrinsic Trp fluorescence peak position and intensity, static light scattering,

and ANS fluorescence intensity data. All calculations were performed using Matlab software (The MathWorks, Natick, MA). In brief, the normalized experimental data at each coordinate (*i.e.*, at a specific temperature and pH combinations) are first converted into an *N*-dimensional vector, where *N* refers to the number of variables included (*i.e.*, number of different types of data). The complex data sets from all measurements are now defined as multi-dimensional vectors sitting in a temperature/pH phase space. Projectors of each individual vectors are then calculated and summed into an $N \times N$ density matrix. By definition, an $N \times N$ matrix has *N* sets of eigenvalues and eigenvectors. The individual vectors at each coordinate were truncated into 3-dimensional vectors and re-expanded into a new basis set consisting of the three eigenvectors corresponding to the three largest eigenvalues. The resultant 3-dimensional vectors were then converted into a color plot with each vector component corresponding to a color using an arbitrary RGB (red, green, blue) color system.[18] Details of the mathematical theory and calculation process can be found elsewhere.[17]

2.3. Results

2.3.1. Far-UV Circular Dichroism (CD) Spectroscopy

Alterations in vault secondary structure were studied by monitoring the changes in the CD signal over a wide range of pH (3 to 8) when subjected to thermal gradients from 10 to 90 ºC. The CD signal is expressed in units of mean residue ellipticity which are independent of molecular weight, since the average molecular

weight of vault particles in solution vary across the pH range due to conformational alterations. Using MALLS, Goldsmith *et al.* showed that the average molecular weight of CP-MVP changes from 9.3 MDa at pH 6.5 to ~ 6.7 MDa at pH 3.4.[14]

The CD spectra of vaults at 10 °C are highly pH dependent and display distinctive characteristics at different pH values (Fig 2.1).

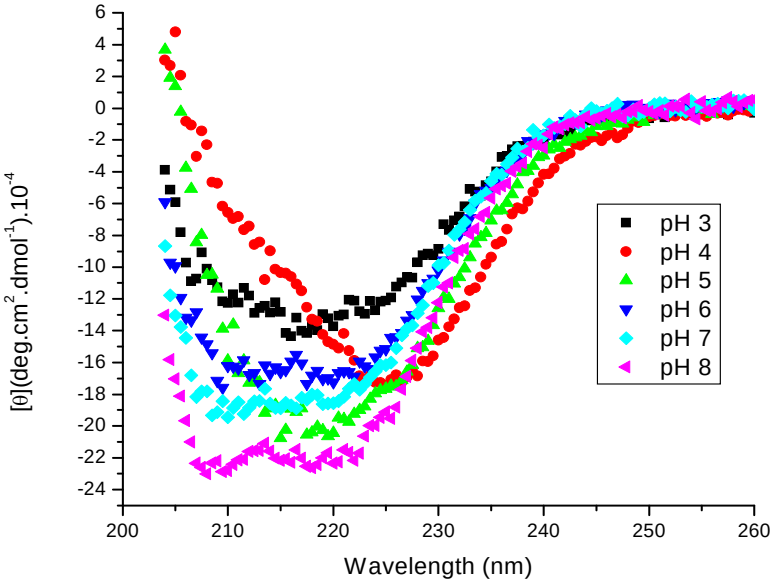


Figure 2.1.CD spectra of Vaults at various pH values. CD spectra were recorded at 10 °C from 190 to 260 nm at each of the indicated pH values (*n* =3).

At pH 3, a broad minimum is observed in the region of 208-225 nm with a shoulder at 208 nm, indicative of a mixture of α-helical and β-sheet structures. At pH 4, a sharp minimum at 225 nm suggests the presence of type II β-turn structure or same type of distorted β-sheet. A minimum at 218 nm at pH 5 indicates predominantly β-sheet content. Two distinct minima at 208 and 222 nm observed at pH 6-8 suggest predominantly α-helical structure in the MVP. Poor quality data below 200 nm due to solute interference prevented an unambiguous deconvolution of spectra for estimation of secondary structural content.

Figure 2.2 emphasizes the dramatic effect of temperature on both the loss of secondary structure and changes in secondary structural content of the vaults over the pH range examined. An overall loss of secondary structure at high temperature is observed across the entire pH range; at pH 6, 7, and 8 the secondary structure of the vaults goes from predominantly α-helical to β-sheet. This can be explained as due to formation of aggregates rich in β-sheet upon heating the vault solutions (see below).

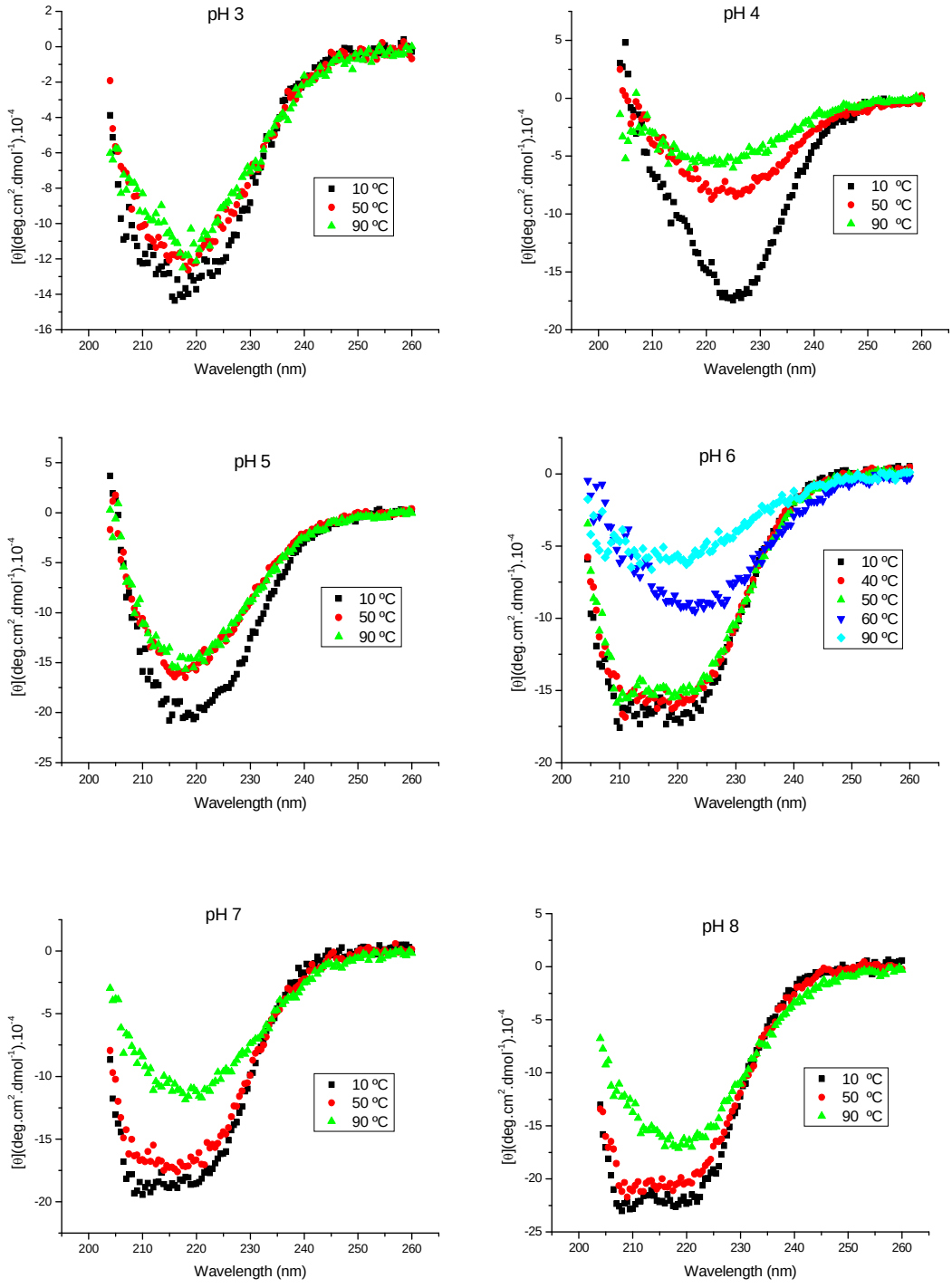


Figure 2.2.CD spectra of Vaults at various pH values collected at the indicated temperatures (n=3).

Secondary structural alterations of vaults were further studied by CD thermal unfolding (T_m) studies in which the mean residue ellipticity at 222 nm was continuously monitored upon heating (Fig 2.3). The negative intensity of the CD signals at 222 nm decreased significantly upon heating suggesting at least partial unfolding of MVPs. The midpoint of these thermal transitions (T_m) is highly pH dependent and in general was found to increase with increasing pH with the exception of pH 4 and 5. Thermal transitions occur at ~ 35.6 °C at pH 3, ~ 44.9 °C at pH 4, ~ 41.2 °C at pH 5, and ~ 56.2 °C at pH 6. (The transition temperatures were determined by fitting the data to a non-linear sigmoidal function defined by the Boltzman equation). The T_m shifts to higher temperatures at increasing pH suggesting a more stable vault assembly at neutral compared to low pH values. The transitions at pH 7 and 8 are very broad, hampering accurate determination of the T_m values. Heating of the vaults at the pH extremes of 3 and 8 leads to smaller overall loss as well as more complex behavior (Fig 2.3) suggesting additional subtle conformational alterations, the presence of folding intermediates, changes in oligomerization state, multiple domains, or some combination of the above.

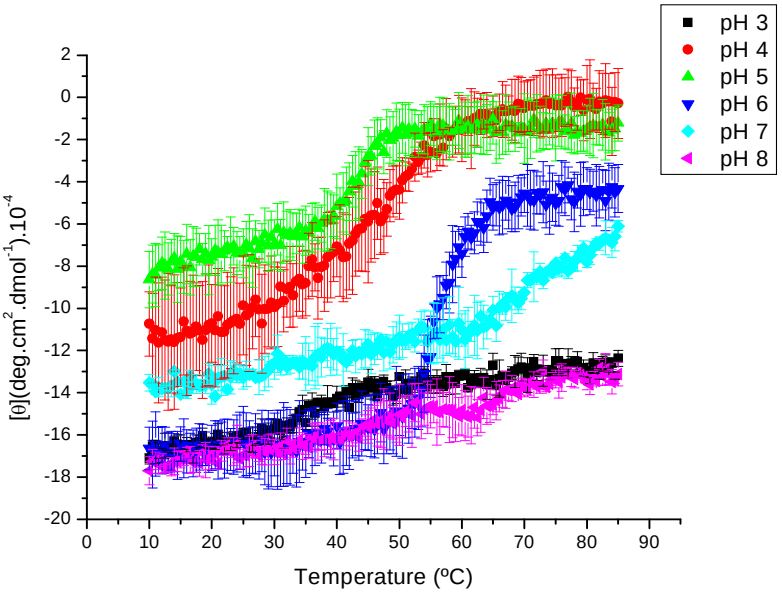


Figure 2.3.The effect of temperature on the secondary structure of Vaults at different pH values. Mean residue molar ellipticity at 222 nm was measured as a function of temperature over the pH range 3–8 (*n* = 3).

2.3.2. Intrinsic Trp Fluorescence Spectroscopy

Tertiary structure alterations in vaults were studied by monitoring changes in the wavelength of the Trp emission peak position and fluorescence intensity as a function of pH and temperature (Fig 2.4a and 2.4b, respectively). Initial peak positions near 342-343 nm suggest only partial burial (on average) of the indole side chains. Trp emission maxima manifest a red shift as the temperature is increased over the pH range examined suggesting enhanced exposure of indole side chains to solvent upon conformational alterations. Complete unfolding of the vaults at high pH, however, seems unlikely due to the small red shifts observed and failure to reach emission peak positions above 350 nm.[16] The small red shifts at low pH could also reflect the presence of half vaults as partially open, structurally altered assemblies.

The temperature-dependent Trp fluorescence intensity data produced estimates of transition temperatures of ~ 45 ºC at pH 4, ~42.5 °C at pH 5, ~50 °C at pH 6, ~52.5 °C at pH 7, and ~52.5 °C at pH 8 (Fig.4b) At pH 3, no obvious transitions are observed with only a slight decrease in intensity seen.

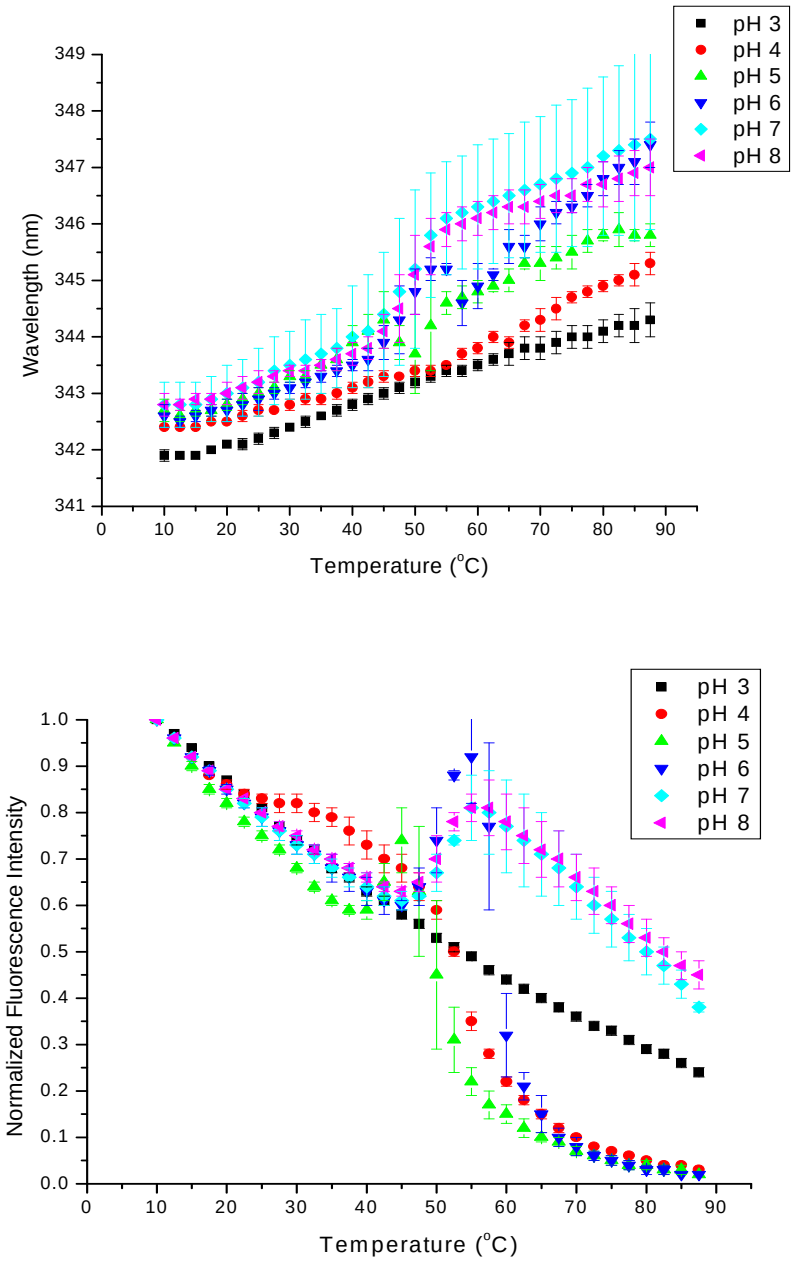


Figure 2.4.a**)** Tryptophan emission peak position of vaults as a function of pH and temperature. Vault suspensions at pH 3–8 were heated from 10 to 85 °C, and the fluorescence emission maxima were determined (*n*=3) after excitation at 295 nm. b) Tryptophan emission fluorescence intensity of vaults as a function of pH and temperature. Vault suspensions at pH 3–8 were heated from 10 to 85 °C, and the fluorescence intensity was monitored (*n*=3).

2.3.2.1. Static Light Scattering

The aggregation behavior of the vaults was simultaneously examined by monitoring the scattered light at the wavelength of excitation (Fig 2.5). The onset of aggregation is pH dependent and occurs at ~ 40 °C at pH 3 (please see inset graph in Fig 2.5), ~ 40 °C at pH 4, ~ 42.5 °C at pH 5, ~ 48 °C at pH 6, ~ 56 °C at pH 7 and ~ 63 °C at pH 8. The onset of aggregation is shifted to higher temperatures at increasing pH suggesting a more stable vault complex at higher pH in agreement with both the CD and intrinsic fluorescence data.

The decrease in light scattering after the initial increase at the observed temperature seen in some cases is presumably due to precipitation of insoluble aggregates. This is observed over the pH range of 3-6 with a less intense effect seen at pH 3. This phenomenon is not observed at pH 7 and 8 with insoluble aggregates not observed at these pH values. Thus, the slight increase in light scattering signal at pH 7 and 8 may be due to the formation of soluble oligomers. The aggregation behavior of the vaults was further confirmed by optical density measurements (OD_{350}) as a function of pH and temperature (data not shown) in which the OD_{350} values decreased continually over the pH range from 0.26±0.022 at pH 3 to 0.14±0.012 at pH 7, again indicating the increasing stability of vault structure as a function of increasing pH.

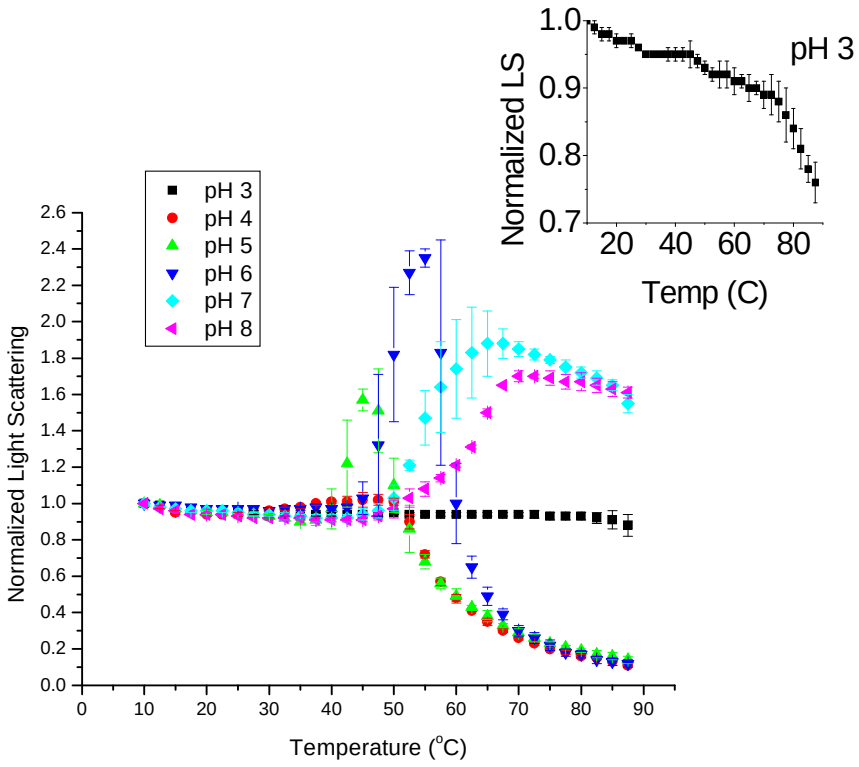


Figure 2.5.Static light scattering as a function of pH and temperature. Vault suspensions at pH 3–8 were heated from 10 to 85 °C, and the fluorescence scattering intensity was monitored at 295 nm (*n*=3).

2.3.3. ANS Fluorescence Spectroscopy

Alterations in the tertiary structure of vaults were further evaluated in the presence of the extrinsic fluorescence probe, 8-Anilino-1-naphthalene sulfonate (ANS). The fluorescence of ANS is highly quenched in aqueous solution but can increase dramatically in non-polar environments. In addition, the emission maximum of the probe is usually blue-shifted upon binding to apolar regions of proteins.[16] In this experiment, the excitation and emission slit widths were set at 6 nm for pH 3 and 4; 9 nm for pH 5-8 to permit quantitative comparisons. Comparing the absolute fluorescence intensity data obtained at 10 ºC (Please see inset graph in Fig 2.6) at low and high pH (*i.e.*, pH 3 and 7) generated comparable values of the same order of magnitude; Considering the smaller slit widths used for the low pH studies, it can be concluded that ANS is increasingly bound to vault complexes at pH 3 compared to pH 7 at 10 ºC. This suggests a greater exposure of previously buried apolar regions at low compared to high pH. Greater ANS binding may be associated with the presence of half vaults and the opening of vault petals into flower-like structures as subunit interfaces are exposed.[14]

Rat MVP (20) containing an identical amino acid sequence to recombinant MVP vaults but lacking the unchanged Cys-rich peptide tag has a calculated pI of 5.3.[2] The strong interactions between vaults and ANS at low pH could also be partially due to electrostatic interactions between the positively charged vaults and the negative charged sulfonate group of ANS.[20-21]

In the pH range 4-8, ANS binding occurs in a multi-step manner. In addition to the initial binding, the intensity of the fluorescence signal rises at some point during the thermal gradient suggesting increased exposure of buried hydrophobic regions. The temperature at which the ANS binds, however, is again strongly pH dependent (Fig 2.6), with T_m values of ~ 32.5 ºC at pH 4, 47.5 ºC at pH 5, and 55 ºC at pHs 6-8. These results are also supported by the blue shifts observed in the ANS emission peak maxima at 485 nm (data not shown). The delayed partial unfolding of the vaults at higher pH again suggests a more stable vault conformation at high pH in agreement with the circular dichroism and intrinsic fluorescence data. At pH 3, however, a continuous decrease in fluorescence intensity is observed with no transitions seen throughout the thermal gradient. This again suggests that at pH 3, the majority of vault particles are already present as partially unfolded half vaults, in which the apolar regions are already more exposed. Continuous decreases in fluorescence intensity are presumably due to the intrinsic effect of temperature on the fluorophore.

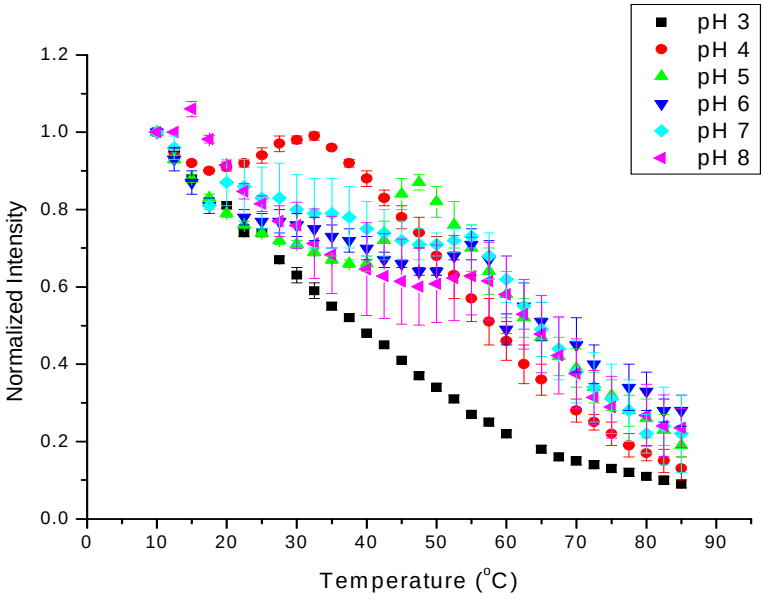
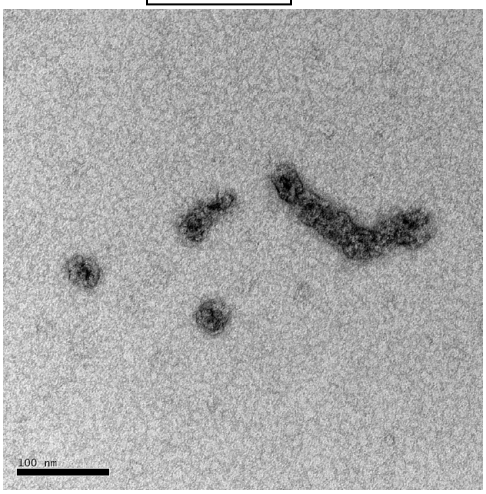


Figure 2.6.Binding of ANS to Vaults as a function of pH and temperature. Vault suspensions in the presence of 10x ANS were excited at 385 nm, and the fluorescence intensity at 485nm was monitored as a function of temperature at each indicated pH (*n*=3).

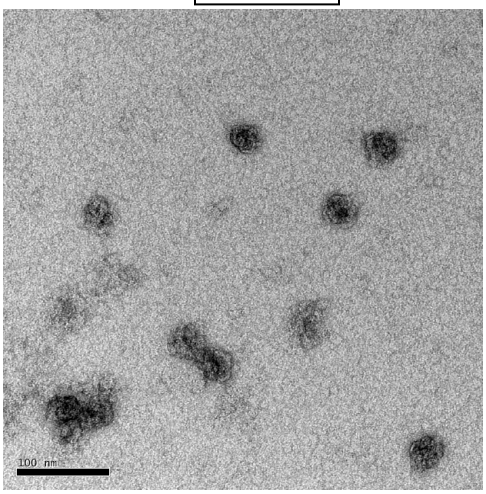
2.3.4. Transmission Electron Microscopy

As an additional tool to confirm the spectroscopic data, EM studies were employed as a function of pH and temperature. The rationale for the selection of the temperature and pH combinations for the EM images was to cover as many distinct conformational phases identified in the EPD as possible. As shown in figure 2.7, at low pH and temperature (pH 3, 4 ºC) vaults are present in the form of half vault monomers, clusters (aggregates) of half vaults, open flower-like structures, and some intact vaults that are in the process of coming apart (present as two half vaults in very close proximity). The extent of aggregation increases as the temperature is increased to 52 ºC and 65 ºC at this pH. At pH 5 and 25 ºC, vaults are primarily intact with some minor aggregates and half vaults observed. Increasing temperature to values as high as 70 ºC results in highly perturbed and aggregated vault particles. At pH 6 and 4 ºC, the vaults appear in the native, assembled form. As the temperature increases to 52 ºC, however, there seems to appear some irregularities in vault structure while still primarily maintaining intact forms. The spectroscopic data also suggest a unique intermediate (molten globule like, see below) conformational state under this solution condition. At pH 8 and 25 ºC, the vaults are intact although considerable aggregation is observed as the temperature increases to 75 ºC. In summary, the EM analysis confirmed the transitions observed by the spectroscopic methods and provide morphological correlates.

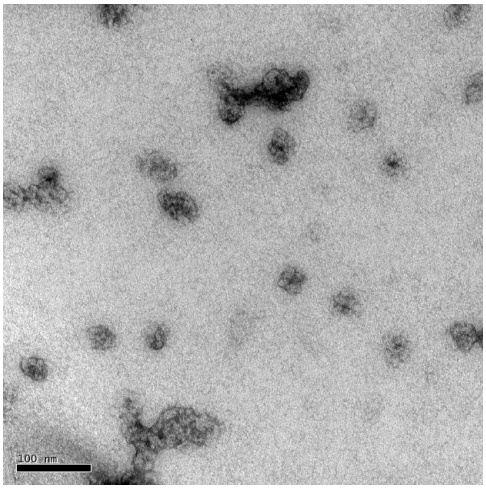
pH 3 / 4 ºC



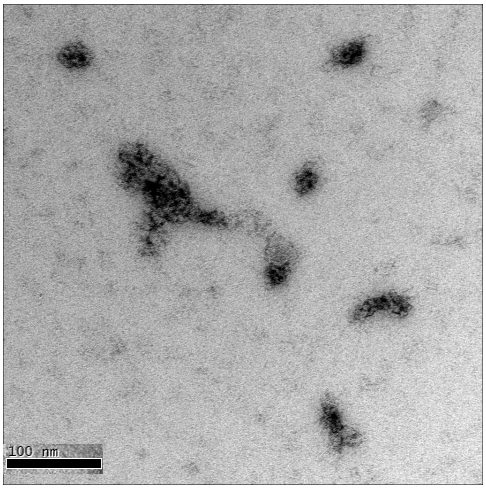
pH 3 / 4 ºC



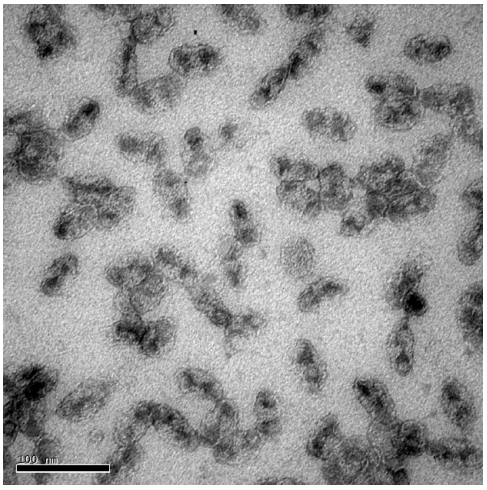
pH 3 / 52 ºC



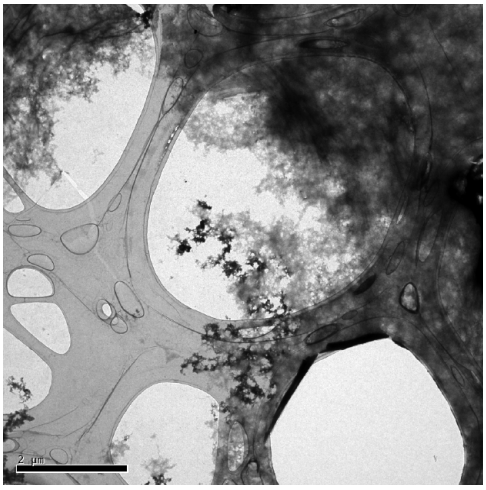
pH 3 / 65 ºC



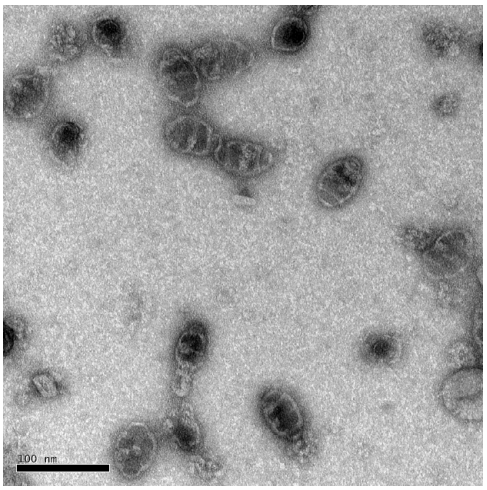
pH 5 / 25 ºC



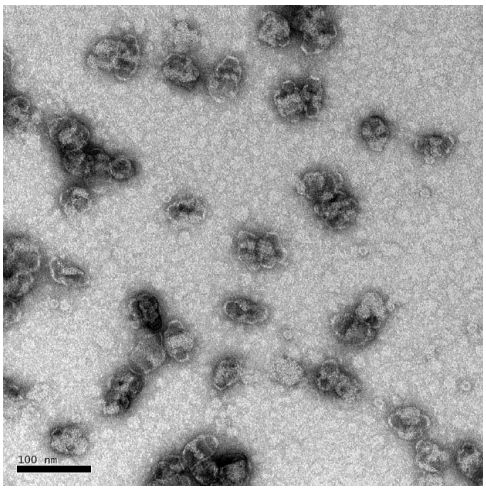
pH 5 / 70 ºC



pH 6 / 4 ºC



pH 6 / 52 ºC



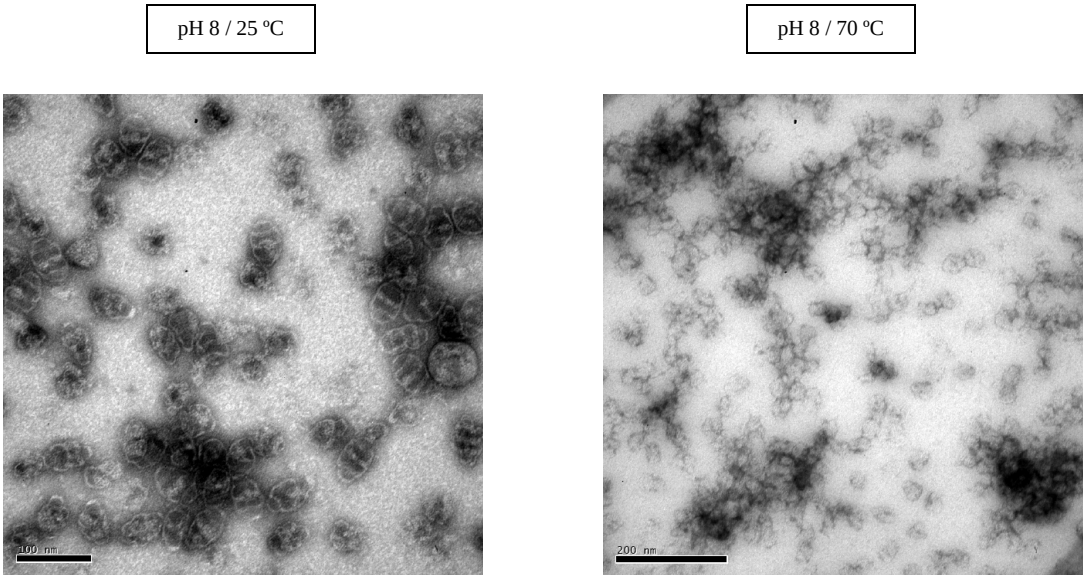


Figure 2.7.Electron microscopy images of vault particles as a function of pH and temperature

2.3.5. Empirical Phase Diagram (EPD)

To provide a more global picture of vault behavior under the temperature and pH conditions examined, much of the spectroscopic data were integrated into an empirical phase diagram (Fig 2.8). We emphasize that empirical phase diagrams (EPD) should not be confused with thermodynamic phase diagrams in which equilibrium exists between different phases. At least partially irreversible aggregation as seen here prevents any such analysis. The EPD was constructed from a variety of spectroscopic techniques, sensitive to both secondary and tertiary structural changes of the vaults. Regions of continuous color define uniform structural states within the limit of resolution of the techniques employed. More importantly, abrupt changes in color identify alterations in the physical state of the vaults over the conditions examined.

The EPD of the vaults clearly shows a large number of physical states over the pH and temperature range studied. The properties of the vaults within each phase can be at least partially established by referring to the individual measurements. Inspection of the EPD reveals ten distinct phases in the EPD. The blue region labeled P1 is the region in which the majority of the complexes are present in the form of half vaults. This is confirmed by the enhanced ANS binding as described earlier. The dark green phase labeled P2 is also a state containing half vaults with an increased level of aggregates as shown in EM studies. The light blue phase (P3) is similar to P1, the only difference being a loss of secondary structure as shown by the CD thermal unfolding data. The dark blue P4 region is the state of maximum stability based on

the predominantly intact vaults. The light green region labeled P5 is also a state of high stability with intact vaults as seen in P4 with somewhat decreased secondary structure based on CD measurements. The pinkish region designated P6 is similar to P5. This pH 5 region, however, is where the earliest transitions in the vault secondary structure start. The red region labeled P7, is another transition region like P6 containing some molten globule-like characteristics based on changes in tertiary structure preceding those in secondary structure. The green region (P8) is the state of minimum stability based upon the magnitude of the spectral changes and the presence of aggregated vaults as shown by both EM images and light scattering. The purple region (P9) is characterized by the presence of soluble aggregates (oligomers). The red region (P10) is very similar to P9 in terms of tertiary structural content as shown by intrinsic and ANS fluorescence data, with the difference between the two a loss of secondary structural content compared to the P10 region as reflected in CD data.

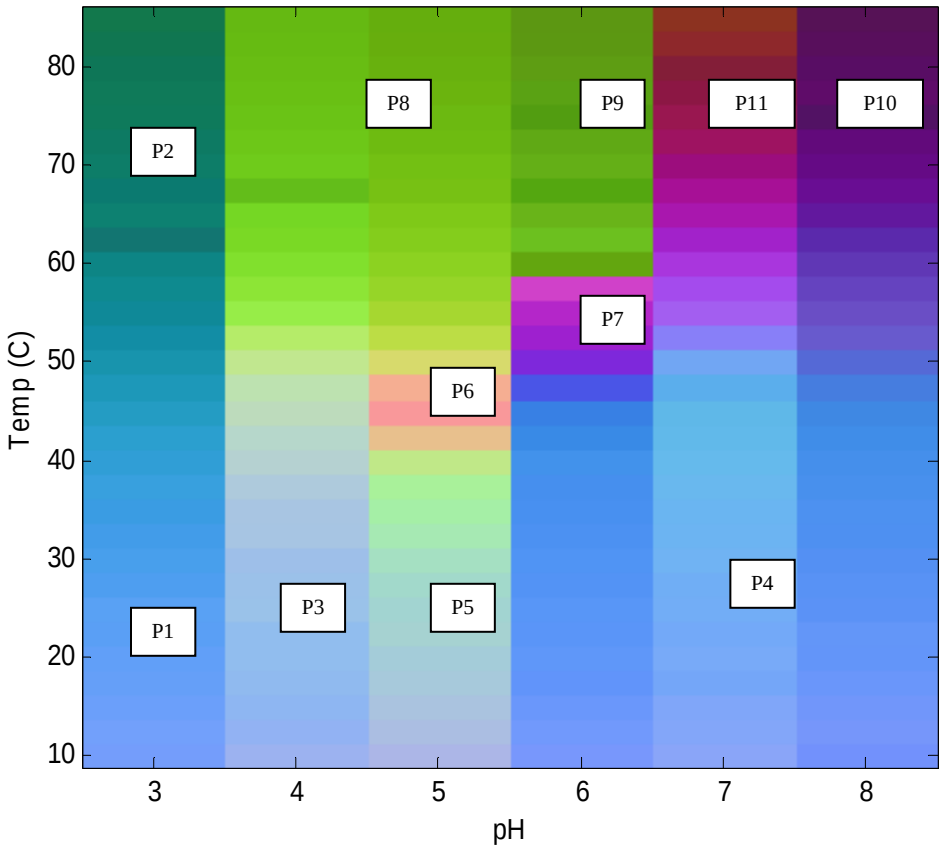


Figure 2.8. Temperature/pH empirical phase diagram of vaults based on intrinsic and extrinsic fluorescence, light scattering and CD results. Six distinct phases (*P*) of the vaults were observed; *P1*, half vaults + aggregates; *P2*, half vaults + higher level of aggregates compared to P1; *P3*, half vaults (less secondary structural content compared to P1); *P4*, intact stable vault assembly; *P5*, intact vault assembly (less secondary structural content compared to P4); *P6*, similar to P5, even less secondary structural content compared to P5; *P7*, intermediate, molten globule like state; *P8*, highly perturbed, aggregated vaults; *P9*, intact but highly plumped vault structure; *P10*, aggregates + soluble oligomers of vaults; *P11*, similar to P10 (less secondary structural content compared to P10)

2.4. Discussion

Employing a variety of techniques, we have extensively characterized the behavior of vaults over a wide range of solution variables (i.e. pH and temperature). The pH range examined covers the extent of physiological pH in different cellular compartments and organs and should therefore provide a basis for further interpretation of vault behavior *in vivo*. In addition, the combination of pH and temperature effects on the structural integrity of vaults should aid in formulation of these nanocapsules for drug delivery since one mechanism by which therapeutic agents entrapped in vaults could be released is an opening of the intact vault structure.

The overall secondary structure of vaults was shown to be highly pH dependent since CD showed changes in the secondary structural content of vaults from a mixture of α-helical and β-sheet at low pH to predominantly β-sheet at intermediate pH to increased α-helix at high pH. Thermal unfolding CD studies over the pH range of 3-8 showed disruption and loss of secondary structure at higher temperatures. Transitions were shifted to higher temperatures at increasing pH suggesting a more stable, intact vault complex at neutral pH.

The induced stability of vault tertiary structure as a function of increasing pH was supported by ANS binding experiments. The onset of ANS binding to previously buried apolar regions was shifted to higher temperatures by increasing pH again suggesting more stable vault assembly at neutral compared to low pH, in agreement with the CD thermal unfolding data.

Intrinsic fluorescence studies demonstrate a 1-5 nm red shift in tryptophan emission maxima as a function of temperature across the pH range, suggesting only limited unfolding of the vault complex. The transition temperatures observed exactly match those obtained from the ANS studies, supporting the presence of a thermally induced event that at least partially opens up the vault's intact structure. Larger red shifts, however, were observed at high pH compared to low, presumably due to the presence of intact vaults at neutral pH as oppose to partially unfolded, half vaults under acidic conditions.

Zon et al. demonstrated that MVP molecules interact with each other via a coiled coil domain in their C-terminal halves forming MVP-MVP subunits that are the basis for vault assembly.[6] Formation of vault like particles by expression of only MVP in insect cells in the absence of other minor vault proteins and RNA further suggests the dominant role of MVP in the assembly of vaults. The secondary and tertiary structural changes seen here could severely alter the integrity of vault assembly by disrupting MVP-MVP interactions. With the goal of developing vaults into drug delivery vehicles, however, some of the observed alterations leading to opening of intact vault structure are advantageous for release of therapeutic agents from vault's cavity.

Physical degradation pathways, particularly aggregation, could severely jeopardize the vault's integrity and disrupt their utility as a delivery device. The tendency of vaults to aggregate was studied by multiple techniques including static light scattering and OD_{350} measurements. The pH dependent aggregation behavior of

vaults observed from static light scattering finds the onset of aggregation shifted to higher temperatures as the pH is increased from 5 to 8, suggesting more stable vaults near neutral pH. A unique phase, labeled P7 in the EPD, is distinguishable at pH 6. The spectroscopic data suggest the presence of an intermediate, molten globule-like state in this region based on the finding that tertiary structure changes precede secondary structure alterations at this pH. This is seen to a lesser extent at pH 5 (P6). Molten globule states are characterized by the presence of extensive secondary structure, minimal tertiary interactions (extensive solvent exposure of previously buried apolar moieties), and a tendency to aggregate.[22] As described above, the current results suggest such a state for vaults under these limited solution conditions based on the CD results (Fig 2.2, pH 6), exposure of the previously buried indole side chains (Fig 2.4a and Fig 2.6), and the tendency of vaults to aggregate (Fig 2.5) in these regions (P6 and P7). Referring to the EM images obtained at pH 6 and 52 ºC, the vaults appear to be less plump and a bit disordered and irregular in this region which may also represent these states. Further, dynamic light scattering (DLS) shows a significant increase in hydrodynamic diameter of vault particles at pH 6 starting at ~ 55 ºC (data not shown). Nevertheless, deconvolution of the DLS data in this region provides evidence for the presence of extensive quantities (over 90%) of monomeric particles further supporting the presence of less expanded vault structure (in agreement with EM analysis). The less expanded structure of vaults under these conditions may also reflect an increase in the particle's density resulting in increased scattering (scattering intensity is proportional to $(dn/dc)^2$).

2.5. Conclusion

Understanding the effects of pH and temperature on vault stability should aid in the ultimate goal of utilizing these particles as potential drug delivery devices. Our studies show that vault conformation was altered as a function of decreasing pH in which intact particles at neutral pH open into flower-like structures at acidic pH. Opening of intact vaults could be potentially developed as an intracellular controlled-release device for drug delivery.

With the identification of at least ten distinct apparent phases in the empirical phase diagram of vaults, it is clear that they can assume a wide variety of different structural forms, consistent with a highly dynamic nature. Comparison of the effects of temperature and pH show that vaults are significantly more stable at high pH and below 40 ºC as encompassed by the blue P4 phase in the EPD. The most unstable, conformationally perturbed structures are present at pH 4 and 5 above 60 ºC identified by the green P8 region in the EPD.

In vivo, naturally occurring vault nanocapsules posseses a dynamic structure and appear to be highly interactive with their surrounding environment.[15] Construction of an EPD based on techniques such as hydrogen/deuterium exchange, red edge shift spectroscopy, time correlated single photon anisotropy measurements, pressure perturbation calorimetry, and high resolution ultrasonic spectroscopy that are more sensitive to the dynamic of vault structure should help to clarify this aspect of vault behavior.

2.6. Bibliography

(1) Kedersha NL, Rome LH. Isolation and characterization of a novel ribonucleoprotein particle: Large structures contain a single species of small RNA. J. Cell. Biol. (1986), 103, 699-709.

(2) Suprenant KA. Vault ribonucleoprotein particles: Sarcophagi, gondolas, or safety deposit boxes? Biochemistry (2002) 41, 14447-14454.

(3) Kickhoefer VA, Garcia Y, Mikyas Yeshi, Johansson Erik, Zhou JC, Raval-Fernandes S, Minoofar P, Zink JI, Dunn B, Stewart PL, Rome LH. Engineering of vault nanocapsules with enzymatic and fluorescent properties. Proc. Nat. Acad. Sci. U.S.A. (2005) 102, 4348-4352.

(4) Anderson DH, Kickhoefer VA, Sievers SA, Rome LH. Eisenberg David. Draft crystal structure of the vault shell at 9-A° resolution. PLoS biology (2007), 5, e318

(5) Stephen AG, Raval-Fernandes S, Huynh T, Torres M, Kickhoefer VA, Rome LH. Assembly of vault-like particles in insect cells expressing only the major vault protein. J. Biol. Chem. (2001), 276, 23217-23220.

(6) Van Zon A, Mossink MH, Schoester M, Scheffer GL, Scheper RJ, Sonneveld P, Wiemer EA. Structural domains of vault proteins: A role for the coiled coil domain in vault assembly. Biochem. Biophys. Res. Commun. (2002) 291, 535-541.

(7) Dalton WS, Scheper RJ. Lung resistance-related protein: Determining its role in multidrug resistance. Journal of the National Cancer Institute (1999), 91, 1604-5.

(8) Kickhoefer VA, Rajavel KS, Scheffer GL, Dalton WS, Scheper RJ, Rome LH. Vaults are up-regulated in multidrug-resistant cancer cell lines. J. Biol. Chem. (1998), 273, 8971-8974.

(9) Mossink MH, van Zon A, Scheper RJ, Sonneveld P, Wiemer EA. Vaults: A ribonucleoprotein particle involved in drug resistance? Oncogene (2003), 22, 7458-7467.

(10) Steiner E, Holzmann K, Elbling L, Micksche M, Berger W. Cellular functions of vaults and their involvement in multidrug resistance. Current Drug Targets (2006), 7, 923-934.

(11) Kozlov G, Vavelyuk O, Minailiuc O, Banville D, Gehring K, Ekiel I. Solution structure of a two-repeat fragment of major vault protein. J. Mol. Biol. (2006), 356, 444-452.

(12) Kowalski MP, Dubouix-Bourandy A, Bajmoczi M, Golan DE, Zaidi T, Coutinho-Sledge YS, Gygi MP, Gygi SP, Wiemer EA, Pier GB. Host resistance to lung infection mediated by major vault protein in epithelial cells. Science (2007), 317, 130-132.

(13) Mikyas Y, Makabi M, Raval-Fernandes S, Harrington L, Kickhoefer VA, Rome LH, Stewart PL. Cryoelectron microscopy imaging of recombinant and tissue derived vaults: Localization of the MVP N Termini and VPARP. J. Mol. Biol. (2004), 344, 91-105.

(14) Goldsmith LE, Yu M, Rome LH, Monbouquette HG. Vault nanocapsule dissociation into halves triggered at low pH. Biochemistry (2007), 46, 2865-2875.

(15) Poderycki MJ, Kickhoefer VA, Kaddis CS, Raval-Fernandes S, Johansson E, Zink JI, Loo JA, Rome LH. The vault exterior shell is a dynamic structure that allows incorporation of vault -associated proteins into its interior. Biochemistry (2006), 45, 12184-12193.

(16) Ausar SF, Foubert TR, Hudson MH, Vedvick TS, Middaugh CR. Conformational stability and disassembly of norwalk virus- like particles: Effect of pH and temperature. J. Biol. Chem. (2006), 281, 19478-19488.

(17) Kueltzo LA, Ersoy B, Ralston JP, Middaugh CR. Derivative absorbance spectroscopy and protein phase diagrams as tools for comprehensive protein characterization: A bGCSF case study. J. Pharm. Sci. (2003), 92, 1805-1820.

(18) Fan H, Li H, Zhang M, Middaugh CR. Effects of solutes on empirical phase diagrams of human fibroblast growth factor 1. J. Pharm. Sci. (2007), 96, 1490-1503.

(19) Kickhoefer VA, Rome LH. The sequence of a cDNA encoding the major vault protein from rattus norvegicus. Gene (1994), 151, 257-60.

(20) Matulis D, Baumann CG, Bloomfield VA, Lovrien RE. 1-Anilino-8-Naphthalene Sulfonate as a protein conformational tightening agent. Biopolymers (1999), 49, 451-458.

(21) Matulis D, Lovrien RE. 1-Anilino-8-Naphthalene Sulfonate anion-protein binding depends primarily on ion pair formation. Biophys. J. (1998), 74, 422-429.

(22) Kuwajima K. The molten globule state as a clue for understanding the folding and cooperativity of globular-protein structure. Proteins (1989), 6, 87-103.

Chapter 3

Structural Stability of a Novel Fusion Cytotoxin: Application of a High Throughput Excipient Screening Assay for Development of a Stable Liquid Formulation

3.1. Introduction

Malignant tumors of the central nervous system are the third leading cause of cancer-related deaths in adolescents and adults between the ages of 15 and 34[1]. Glioblastoma multiforme (GBM) is the deadliest form of malignant brain tumor with ~ 13,000 annual deaths in the United States alone[2]. Available therapies (i.e., surgery, radiation and chemotherapy) result in ~ 20% survival rates with an average of two and a median of about one year survival periods[3]. This has resulted in extensive efforts toward the development of novel treatments for malignant brain tumors, one of which involves the use of anti-brain tumor cytotoxins with high specificity towards cancer cells[3].

It has been previously reported that fusion of Interleukin-13 with a bacterial toxin, Pseudomonas Exotoxin A (PE) generates a proteinaceous complex with very high specificity towards brain tumor cells[4]. The IL-13 cytokine, containing a high degree of structural homology to IL-4, performs a variety of immunomodulatory functions with both pro- and anti-inflammatory immune actions[5,6]. On untransformed tissue cells, IL-13 and IL-4 compete for binding to IL13/4R, a heterodimeric receptor complex composed of IL13Rα1 and IL4Rα. In contrast, IL13Rα2 which is a highly specific binding target of IL-13 alone is exclusively present on the surface of cancer cells[7]. The novel fusion cytotoxin of this study contains a genetically engineered form of IL-13 in which the glutamic acid residue at position 13 has been replaced by a lysine residue (IL13.E13K). Position 13 has been recognized as a critical hot spot in the IL-13 wild type sequence involved in binding to IL13/4R. The mutated version,

however, neither binds nor activates IL13/4R, but retains specificity towards glioma-restricted IL13Rα2[7]. The bacterial toxin portion of the fusion protein, PE4E, is a derivative of Pseudomonas Exotoxin A. It is composed of three major domains; domain I binds to the cell surface receptor (i.e., α_2-macroglobulin), domain II catalyzes the translocation of the toxin out of endosomes into the cytosol, while domain III contains the ADP-ribosylation activity which inactivates elongation factor 2 resulting in cell death[8,9]. The glutamic acid residues of PE4E at positions 57, 246, 247, and 249 have been substituted with basic amino acids to inhibit binding to its cell surface receptor on normal cells[10,11]. Therefore, fusion of Pseudomonas Exotoxin A with IL-13 essentially generates a cytotoxic entity targeted specifically towards cancer cells.

To develop this novel fusion cytotoxin as a safe and efficacious therapeutic, a robust formulation (liquid or lyophilized) that retains the protein's stability and activity over its entire shelf life is required. Proteins are, however, frequently thermolabile and susceptible to a variety of physical and chemical degradations, often leading to their altered immunogenicity and loss of biological activity. This can introduce significant challenges to the development of safe, efficacious, and stable protein formulations[12].

Given this high degree of structural complexity and associated degradation pathways, the FDA requires that the stability and activity of potential protein therapeutics be demonstrated in real time under the proposed labeled storage conditions. These types of studies, however, demand significant resources be

dedicated to them for extended periods of time since the common shelf life of protein drugs are usually at least 18-24 months[13]. One time and cost effective approach is to conduct accelerated stability studies to screen for potential formulations with properties that lead to enhanced stability under more moderate storage conditions. Such accelerated studies typically involve stressed conditions (e.g., elevated temperatures, suboptimal pH, high or low ionic strengths) to facilitate protein degradation in a short period of time.

Herein, we have employed a variety of spectroscopic techniques (i.e., circular dichroism, intrinsic and extrinsic fluorescence, and static light scattering) to characterize the thermal stability of the fusion protein in solution under accelerated degradation conditions. The results from these analyses served as a basis to develop a high throughput screening assay to identify potential GRAS (Generally Regarded as Safe) stabilizers that inhibited aggregation and enhanced conformational stability of the fusion protein. We also show that the protein's thermal stability can be understood in terms of the behavior of its individual domains.

3.2. Materials & Methods

3.2.1. Materials

Highly purified bulk protein samples are stored frozen in 25 mM succinate buffer (200 mM urea, 150 mM NaCl, pH 4.8). The purity of the bulk samples was confirmed by RP-HPLC studies in which a homogeneity of >99.4% was established.

All other chemicals were of analytical grade and were purchased from Sigma (St. Louis, MO) and Fisher Scientific (Pittsburgh, PA).

3.2.2. Sample Preparation.

Frozen protein solution was thawed at room temperature and was buffer exchanged into 20 mM isotonic citrate phosphate buffer ranging from pH 3 to 8, at one pH unit intervals (the isotonicity was maintained using sodium chloride). For buffer exchange, protein samples were dialyzed at refrigerator temperatures using Spectra/Por molecular porous membrane tubing (SpectrumLab, Rancho Dominguez, CA) with a 15KDa MW cutoff. Samples were examined at a concentration of 0.1 mg/ml for all studies with the exception of CD and excipient screening studies in which samples at 0.2 and 0.35 mg/ml were used, respectively. The concentration of protein samples after dialysis was determined by UV absorption spectroscopy (A_{280nm}) using E_{1cm} 0.1% of 1.2. The fusion protein's extinction coefficient was calculated using the program Protparam on the Expasy server (http://www.expasy.ch/tools/protparam.html). For all studies, three independent samples were evaluated to ensure reproducibility of the measurements. Values are presented as averages with accompanying standard deviations.

3.2.3. Far-UV Circular Dichroism (CD) Spectroscopy

CD spectra were acquired using a Jasco J-810 spectropolarimeter equipped with a 6-position sample holder and a Peltier temperature controller. The CD spectra

were obtained from 260-190 nm with a scanning speed of 20 nm/min, a 2 sec response time and an accumulation of 3. To study thermal transitions (melting curves) of the fusion protein, the CD signals at 222 nm were monitored in 0.1 cm pathlength cuvettes every 0.5 °C over a 10 to 85 °C temperature range employing a temperature ramp of 15 °C/hr.

3.2.4. Intrinsic Tryptophan (Trp) Fluorescence Spectroscopy

Fluorescence spectra were acquired using a Photon Technology International (PTI) spectrofluorometer (Lawrenceville, NJ) equipped with a turreted 4-position Peltier-controlled cell holder. An excitation wavelength of 295 nm was used to primarily excite Trp residues and the emission spectra were collected from 310 to 400 nm with a step size of 1 nm and a 1 sec integration time. Excitation and emission slits were set at 3 and 1 nm, respectively. Light scattering was also monitored at 295 nm using a separate photomultiplier placed at 180º to the fluorescence detector. Both the fluorescence intensity and the light scattering data were normalized with respect to the initial reading at 10 °C. Emission spectra were collected every 2.5 °C with a 3 min equilibration time over a temperature range of 10 to 85 °C. A buffer baseline was subtracted from each raw emission spectrum.

Peak positions of the emission spectra were obtained from polynomial fits using Origin software using a “center of spectral mass” method. Due to the nature of center of spectral mass analysis, the reported emission peak position values do not correspond to the actual peak positions, but do accurately reflect the changes and produce more reproducible results than alternative methods. The actual peak positions

determined by derivative analysis of the native protein are shifted approximately 10-14 nm from their center of mass values.

3.2.5. ANS Fluorescence Spectroscopy

Accessibility of apolar sites on the fusion protein was monitored by fluorescence emission of the extrinsic probe 8-Anilino-1-naphthalene sulfonate (ANS). Each sample contained a 20-fold molar excess of ANS to protein, an optimal ratio determined in preliminary experiments. The ANS was excited at 385 nm and emission spectra were collected from 425-550 nm with a step size of 1 nm and 1 sec integration time. Emission spectra were collected every 2.5 °C with 3 min of equilibration over a temperature range of 10 to 85 °C using a Photon Technology International (PTI) spectrofluorometer (Lawrenceville, NJ) equipped with a turreted 4-position Peltier-controlled cell holder. Slits were set at 5 nm excitation and emission in the pH range of 5-8, 3 nm excitation/2 nm emission at pH 3, and 4 nm excitation and emission at pH 4; data were normalized with respect to the initial reading at 10 °C to permit direct comparison. The ANS-buffer baseline with almost no fluorescence at each corresponding pH was subtracted from raw emission spectra.

3.2.6. Empirical Phase Diagram (EPD)

To facilitate analysis of the complex data sets, an empirical phase diagram was constructed employing CD signals at 222 nm, intrinsic Trp fluorescence peak position, static light scattering, and ANS fluorescence intensity data. All calculations

were performed using Matlab software (The MathWorks, Natick, MA). In brief, the normalized experimental data at each coordinate (i.e., at specific temperature and pH combinations) are first converted into an *N*-dimensional vector, where *N* refers to the number of variables included (i.e., number of different types of data). The complete data sets from all measurements are now defined as multi-dimensional vectors in a temperature/pH phase space. Projectors of each individual vectors are then calculated and summed into an $N \times N$ density matrix. By definition, an $N \times N$ matrix has *N* sets of eigenvalues and eigenvectors. The individual vectors at each coordinate are truncated into 3-dimensional vectors and re-expanded into a new basis set consisting of the three eigenvectors corresponding to the three largest eigenvalues. The resultant 3-dimensional vectors are then converted into a color plot with each vector component corresponding to a color using an arbitrary RGB (red, green, blue) color system. Details of the mathematical theory and calculation process can be found elsewhere[14].

3.2.7. High Throughput Excipient Screening

Protein samples (0.35 mg/ml) in 20 mM isotonic citrate phosphate buffer, pH 5.0 were incubated in the presence and absence of excipients at 45 °C for 4 hrs and aggregation kinetics were monitored by measuring the optical density ($O.D._{350nm}$) every 2 min using a SpectraMax-M5 96 well microplate reader equipped with a temperature controller (Molecular Devices, Sunnyvale CA). A buffer baseline containing only excipient in buffer was subtracted from data obtained from the

excipient solution in the presence of protein. The optimal protein concentration of 0.35 mg/ml was chosen from initial concentration dependent aggregation studies (data not shown). Concentrated stock solutions of excipients were prepared by dissolution of excipients in 20 mM isotonic citrate phosphate buffer, pH 5. Upon dissolution, the pH was adjusted to 5 by addition of HCl or NaOH.

The percent inhibition of aggregation (%IA) was calculated using:

$$\%IA = 100- [OD_{350(E)} / OD_{350(NE)}] \times 100$$

in which $OD_{350(E)}$ and $OD_{350(NE)}$ are the changes in the optical density of protein in the presence and absence of excipients after 240 min (time needed to a reach a plateau), respectively.

3.3. Results

3.3.1. Far-UV Circular Dichroism (CD) Spectroscopy

Alterations of the fusion protein secondary structure were studied by monitoring the changes in the CD signal over a wide range of pH (3 to 8) when subjected to thermal gradients from 10 to 90 ºC. The CD spectra at 10 °C exhibit distinct pH dependency (Fig 3.1a). At pH 3 and 4, a broad minimum is observed in the region of 208-225 nm with a shoulder at 208 nm, indicative of a mixture of α-helical and β-sheet structures. Two distinct minima at 208 and 222 nm observed in the pH range of 5-8 suggest the presence of predominantly α-helical structure. The

spectrum at pH 5 is much less intense compared to other pH values suggesting some form of unique structure under this pH condition. Poor quality data below 200 nm prevented deconvolution of spectra for quantitative estimation of secondary structural content.

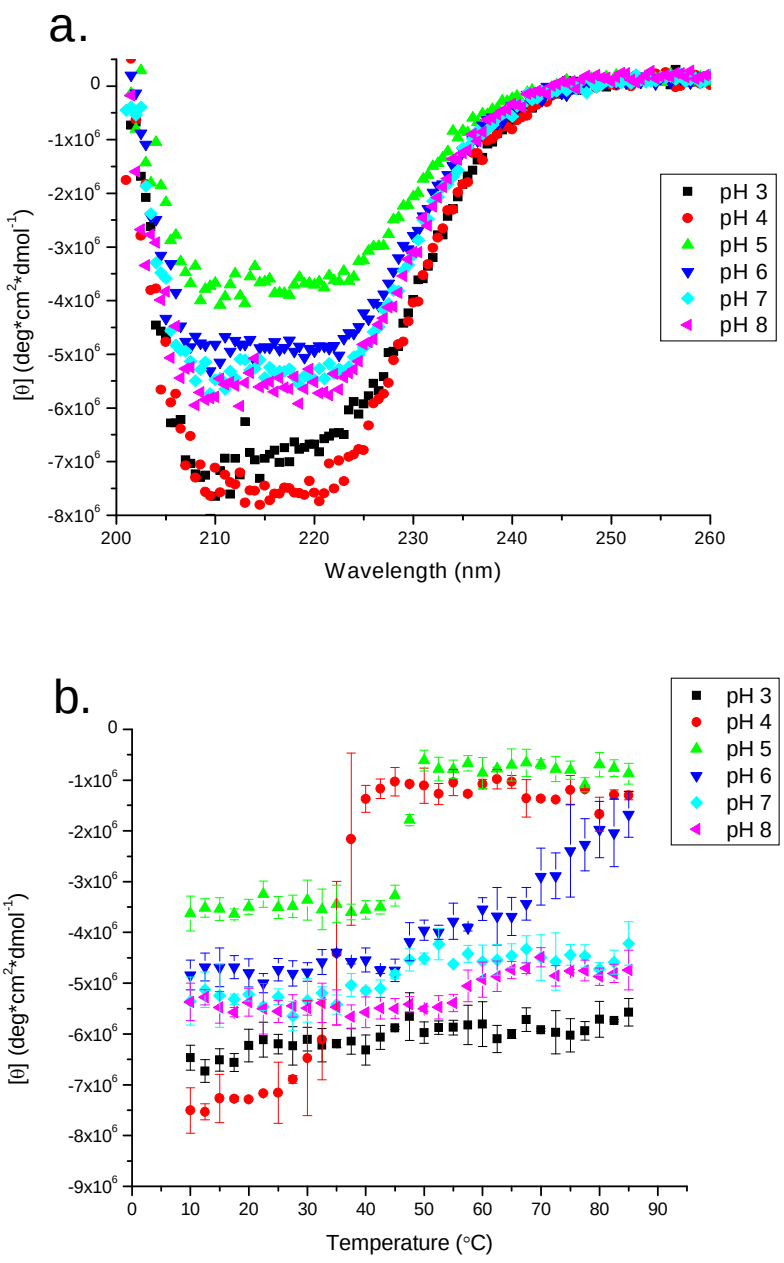


Figure 3.1.**a)** CD spectra of fusion protein at various pH values. CD spectra were recorded at 10 ° C from 190 to 260 nm at each of the indicated pH values (*n* =3). Error bars are not shown to enhance clarity. **b)** The effect of temperature on the secondary structure at different pH values. Molar ellipticity at 222 nm was monitored as a function of temperature over the pH range 3–8 (*n* = 3).

The effect of temperature on protein secondary structure was further investigated employing thermal melt analysis by monitoring the CD signals at 222 nm (Fig 3.1b). An overall loss in the secondary structure is observed as a function of increasing temperature across the pH range examined. Sharp transitions are observed in the pH range of 4-6 with an approximate onset transition temperatures of ~ 25 °C at pH 4, ~ 42.5 °C at pH 5, 6, 7, and ~ 52.5 °C at pH 8. The transitions at extreme pH values, however, are more subtle presumably suggesting either the presence of already altered and partially unfolded conformational states at pH 3 (see below) or more stable secondary structure at pH 7 and 8 over the temperature range examined.

The dramatic decrease in the transition temperatures by ~ 25 °C going from pH 8 to 4, suggests reduced stability of the fusion protein secondary structure at acidic pH values. In the pH range of 5-7, similar onset transition temperatures are observed, although, the extent to which the secondary structure is altered is somewhat different; less extensive perturbations are observed at higher pH values suggesting more stable secondary structure at neutral and basic pH values.

3.3.2. Intrinsic Trp Fluorescence Spectroscopy

Alterations in fusion protein tertiary structure were studied by monitoring the changes in its Trp emission peak position and fluorescence intensity as a function of pH and temperature. An overall red shift in the Trp emission maxima is observed as a function of increasing temperature over the pH range examined (Fig 3.2a) suggesting enhanced exposure of the indole side chains to solvent upon conformational

alterations[15]. The examined fusion protein contains a total of 12 tryptophan residues, out of which 11 belongs to the toxin domain suggesting its dominant role in the tryptophan peak position shifts of the fusion protein.

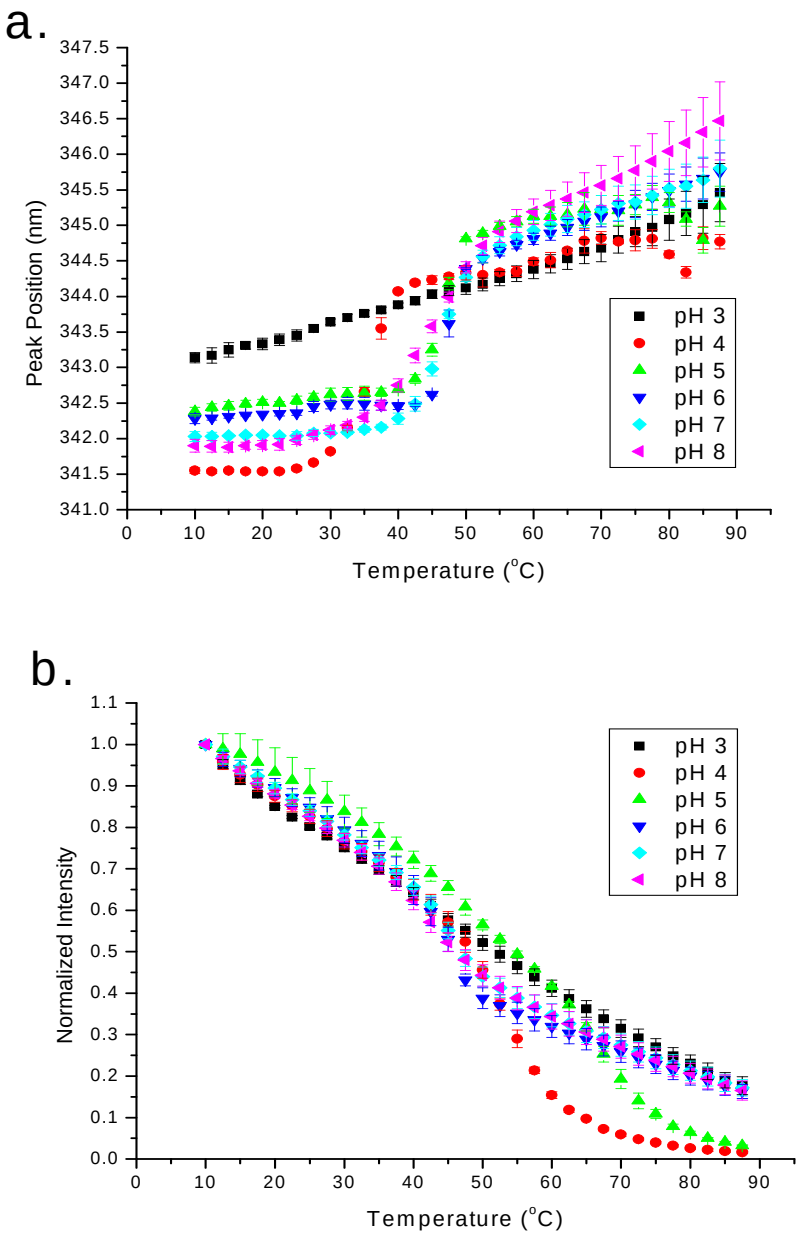


Figure 3.2.**a)** Tryptophan emission peak position as a function of pH and temperature. Protein solutions at pH 3–8 were heated from 10 to 85 °C, and the fluorescence emission maxima were monitored (*n*=3) upon excitation at 295 nm. **b)** Tryptophan emission fluorescence intensity as a function of pH and temperature. Protein solutions at pH 3–8 were heated from 10 to 85 °C, and the fluorescence intensity was monitored (*n*=3).

The midpoints of the thermal transitions occur at ~ 36 °C at pH 4, ~ 46 °C at pH 5, ~ 50 °C at pH 6, and ~ 48 °C at 7 and 8 suggesting that protein at neutral pH values is the most stable. No clear transitions were observed at pH 3 (The transition temperatures were determined by fitting the data to a non-linear sigmoidal function defined by the Boltzman equation). The extent of exposure of indole side chains to solvent on average is somewhat similar across the pH range with red shifts of ~ 3-4.5 nm in magnitude. The small red shifts in peak position indicate partial unfolding of the protein with increasing temperature (Fig 3.2a). The initial peak positions near 330 nm (calculated based on a direct peak picking method) suggest that on average, the indole side chains of the 12 Trp residues in the fusion protein are buried in a more apolar environment. The temperature-dependent Trp fluorescence intensity data exhibited weak but distinct transitions across the pH range of 3-8 over a similar range seen in the peak shift studies (Fig 3.2b). The sharp decrease in fluorescence intensity at pH values 4 and 5 is presumably due to precipitation of protein at these pH values at higher temperatures (see below).

3.3.2.1. Static Light Scattering

The aggregation behavior of the fusion protein was simultaneously examined by monitoring the scattered light at the wavelength of excitation, 295 nm.

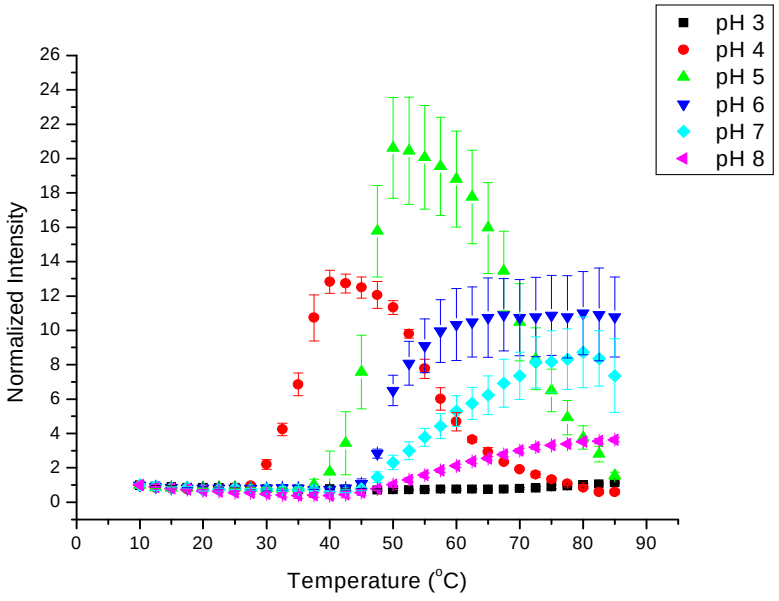


Figure 3.3.Static light scattering of fusion protein as a function of pH and temperature. Protein solutions at pH 3–8 were heated from 10 to 85 °C, and the scattering intensity was monitored at 295 nm (*n*=3).

The onsets of aggregation are highly pH dependent and occur at ~ 30 °C at pH 4, ~40 °C at pH 5, and ~45 °C at pH 6, 7, and 8 (Fig 3.3). Overall, the onset of aggregation is shifted to lower temperatures with decreasing pH suggesting a less stable protein at more acidic pH values in good agreement with both the CD and intrinsic fluorescence data. The extent to which the fusion protein aggregates is also highly pH dependent with the greatest extent of aggregation observed at pH 4 and 5; a decrease in light scattering after the initial increase at higher temperatures is presumably due to precipitation of insoluble aggregates in agreement with the intrinsic fluorescence intensity data. This phenomenon is not observed in the pH range of 6-8. A slight increase in light scattering signal at pH 7 and 8 is due to formation of soluble oligomers. No clear transitions were observed at extreme pH 3, presumably due to the presence of already altered and partially unfolded protein as already suggested by the CD and Trp fluorescence results.

3.3.3. ANS Fluorescence Spectroscopy

Alterations in the tertiary structure of the fusion protein were further studied in the presence of an extrinsic fluorescence probe, 8-Anilino-1-naphthalene sulfonate (ANS). The fluorescence of ANS is highly quenched in aqueous solutions but increases dramatically in non-polar environments[16,17](Fig 3.4a). In addition, the emission maximum of the probe is usually blue-shifted upon binding to apolar regions of proteins (Fig 3.4b). Although the negative charge on ANS can occasionally cause it to bind to positively charged sites on proteins, it is generally accepted as a

probe of apolar sites. The accessibility of the ANS probe to hydrophobic regions of the fusion protein was examined in the presence of a 20 molar excess of ANS to protein.

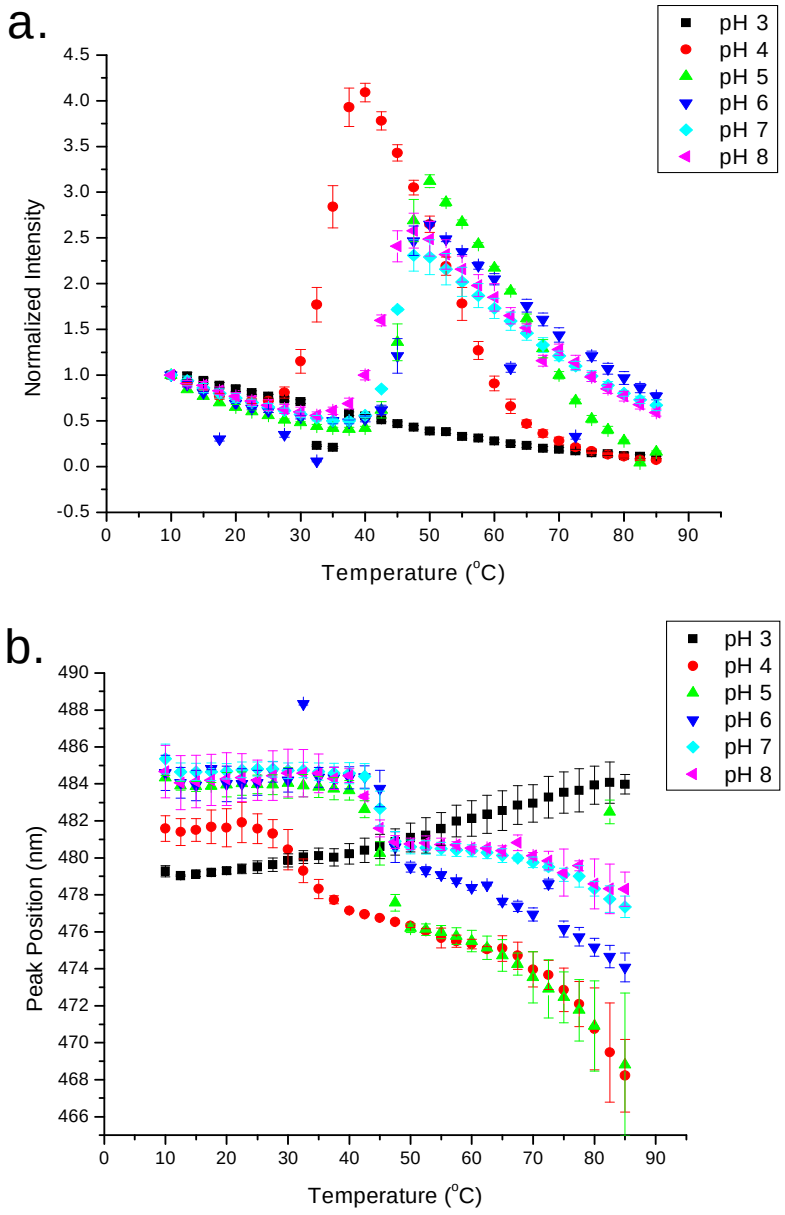


Figure 3.4.**a)** Binding of ANS to fusion protein as a function of pH and temperature. Protein solutions in the presence of a 20 molar excess of ANS were excited at 385 nm, and the fluorescence intensity at 485 nm was monitored as a function of temperature at each indicated pH (*n*=3). **b)** ANS emission peak position as a function of pH and temperature (calculated using the "center of spectral mass" method). Protein solutions at pH 3–8 were heated from 10 to 85 °C, and the fluorescence emission maxima were monitored (*n*=3) upon excitation at 385 nm.

The temperature at which the ANS binds to the protein is highly pH dependent with an approximate onset of thermal transitions at ~ 25 °C at pH 4, ~ 40 °C in the pH range 5-7, and ~ 35 °C at pH 8 (Fig 3.4a). The earlier unfolding of the fusion protein at pH 4 compared to other pH values suggests a less stable protein conformation at this pH in agreement with both the circular dichroism and light scattering data. At pH 3, however, a continuous decrease in fluorescence intensity is observed with no transitions apparent throughout the thermal gradient. This suggests that at pH 3, the protein is substantially unfolded (note that data at each pH has been normalized with respect to the initial value at 10 °C). The enhanced binding of ANS to apolar regions of the fusion protein above pH 3 as a function of increasing temperature is supported by the blue shifts of the ANS emission peak maxima at ~ 485 nm (Fig 3.4b). This is not observed at pH 3 where ANS is already bound to the partially unfolded protein at the initial temperature of 10 °C. The early blue shifts observed at pH 4 confirm the presence of less stable and altered forms of protein at this pH as suggested by the intensity data.

3.3.4. Empirical Phase Diagram (EPD)

The spectroscopic data were integrated into an empirical phase diagram to provide a more global picture of the fusion protein behavior under the temperature and pH conditions examined (Fig 3.5). We emphasize that empirical phase diagrams should not be confused with thermodynamic phase diagrams in which equilibrium exists between different phases. At least partially irreversible aggregation as seen here

prevents any such analysis. Regions of continuous color define uniform structural states while abrupt changes in color identify alterations in the physical state of the protein over the conditions examined. The properties of the fusion protein within each phase can be established by referring to the individual measurements.

The EPD of the fusion protein clearly shows a number of distinct conformational states over the pH and temperature range examined. The green region labeled as P1 is the region of maximum stability with protein present in what can be defined as its native-like and folded conformation. The purple region labeled P2 is a state of perturbed tertiary structure and contains soluble oligomers and aggregates of the fusion protein. The region labeled P3 is the state of minimum stability with extensive aggregation and perturbed secondary and tertiary structures. The orange region P4 is very similar to P1, the only difference being a loss of secondary structure as shown by the CD spectra and thermal unfolding data. The P5 green region contains extensively unfolded structure. The dark green P6 region is similar to P5, with an even greater extent of unfolding seen in the P6 region as implied in the tryptophan fluorescence peak position data.

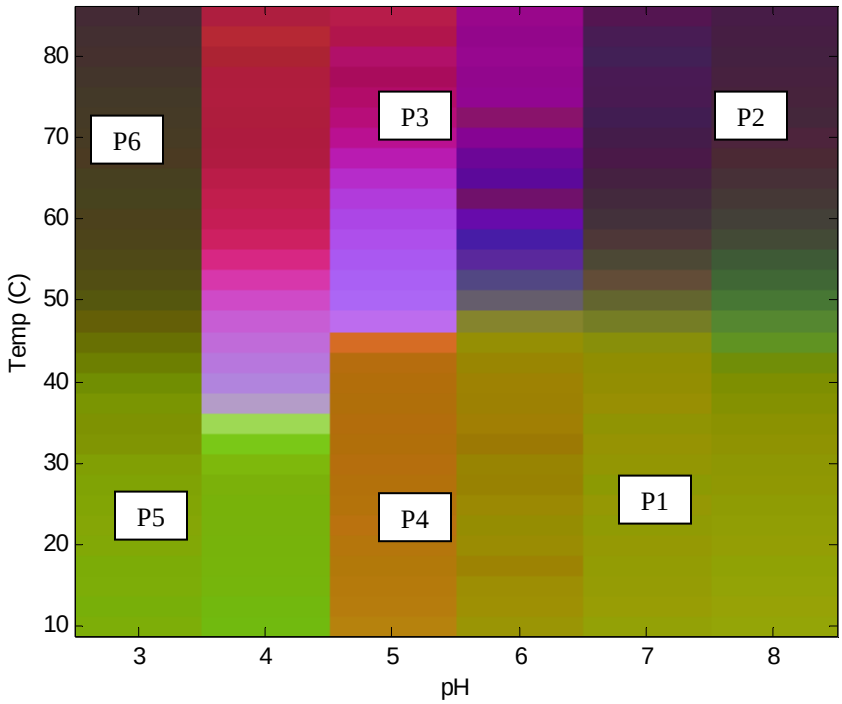


Figure 3.5.Temperature/pH empirical phase diagram of the fusion protein based on intrinsic and extrinsic fluorescence, light scattering and CD thermal melt data. For a description of the different regions (P1-P6), see the text.

3.3.5. High Throughput Screening for Inhibitors of Protein Aggregation

The pH and temperature conditions under which the fusion protein showed marginal stability were identified by analysis of the raw data and identification of apparent phase boundaries (i.e. abrupt changes in color) in the empirical phase diagram. One such boundary was identified at pH 5 and ~ 45 °C based on extensive aggregation of protein (Fig 3.3 and 3.5). Therefore, aggregation kinetics under this condition was selected as a tool to screen for potential stabilizers. A high throughput screening assay was developed (as described in the method section) to examine the ability of a library of Generally Regarded as Safe (GRAS) excipients to inhibit aggregation of the fusion protein. Figure 6 shows representative aggregation kinetic traces of protein in the presence of selected excipients. The left panel includes excipients that inhibited the aggregation of protein whereas the ones in the right panel proved to be ineffective. The presence of a lag time implies the existence of a critical nucleation event.

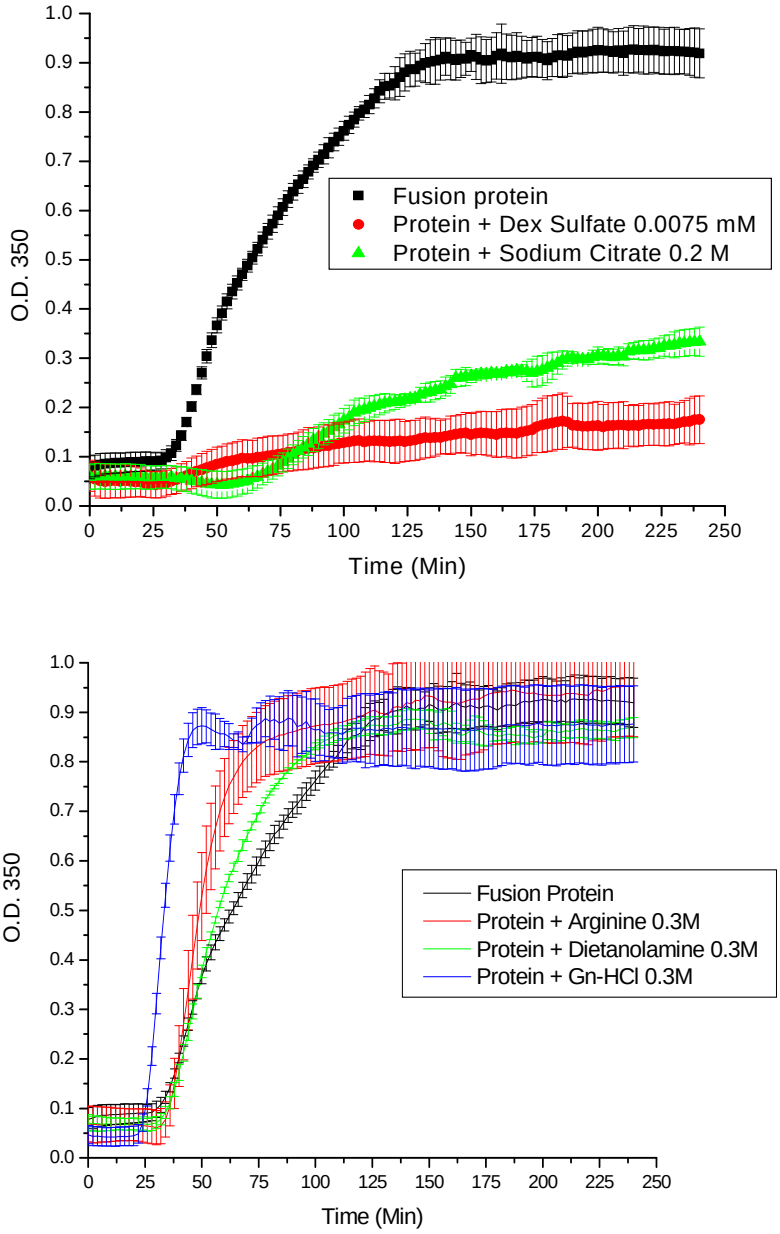


Figure 3.6.Aggregation kinetic traces of the fusion protein in the presence of selected excipients. Protein samples (0.35 mg/ml) in 20 mM isotonic citrate phosphate buffer, pH 5.0 were incubated in the presence and absence of excipients at 45 °C for 4 hr and aggregation kinetics were monitored by measuring the optical density (OD_{350nm}) every 2 min.

A complete list of excipients examined and their inhibitory effect on protein aggregation represented by percent inhibition of aggregation (%IA) is shown in descending order in Table 3.1.

Excipient (conc.)	%IA	Excipient (conc.)	%IA
Dextran Sulfate (0.0075mM)	81	Dietanolamine (0.3M)	4
Dextran Sulfate (0.003mM)	78	Guanidine-HCl (0.3M)	4
Sodium Citrate (0.2M)	64	Tween 80 (0.05% w/v)	4
Dextran Sulfate (0.02mM)	63	Lactic acid (0.15M)	3
Dextrose (20% w/v)	61	Bridge 35 (0.01% w/v)	3
Dextran Sulfate (0.0003mM)	53	Sucrose (10% w/v)	3
Sorbitol (20% w/v)	48	Trehalose (10% w/v)	3
Trehalose (20% w/v)	42	Bridge 35 (0.05% w/v)	0
Glutamic acid (0.15M)	40	Arginine (0.3M)	-3
Lactose (10% w/v)	37	2-OH propyl ğ-CD (10% w/v)	-5
Sucrose (20% w/v)	34	Gelatin (5% w/v)	-6
Sodium Citrate (0.1M)	33	Glycerol (10% w/v)	-6
Tween 80 (0.1% w/v)	33	Histidine (0.3M)	-8
Tween 20 (0.1% w/v)	28	Phytic Acid (5:1 w/w)*	-13
Bridge 35 (0.1% w/v)	24	Calcium Chloride (0.015M)	-14
Mannitol (10% w/v)	23	Phytic Acid (1:1 w/w)*	-15
Sorbitol (10% w/v)	22	2-OH propyl γ-CD (5% w/v)	-16
Tween 20 (0.05% w/v)	22	α-Cyclodextrin (2.5% w/v)	-19
Glycerol (20% w/v)	18	Arginine/Glutamic Acid (25mM each)	-21
Dextrose (10% w/v)	17	2-OH propyl β-CD (5% w/v)	-22
Tween 80 (0.01% w/v)	15	Phytic Acid (10:1 w/w)*	-23
Glycine (0.3M)	12	2-OH propyl β-CD (10% w/v)	-26
Proline (0.3M)	11	Arginine/Glutamic Acid (50mM each)	-27
Tween 20 (0.01% w/v)	9	Pluronic F-68 (0.05% w/v)	-29
Aspartic Acid (0.075M)	9	Pluronic F-68 (0.1% w/v)	-29
Pluronic F-68 (0.01% w/v)	7	Albumin from Human Serum (5% w/v)	-48
Malic Acid (0.15M)	4		

Table 3.1.Percent inhibition of the fusion protein aggregation (%IA) in the presence of variety of excipients after 240 min. The average standard deviation is ± 5% (n=3).

* Excipient to protein ratio

A number of excipients manifested the ability to significantly inhibit the aggregation of the fusion protein (Table 3.1). Selected concentrations of polyanions (dextran sulfate), polysaccharides (sucrose, dextrose, lactose, and trehalose), polyols (sorbitol and glycerol), salt (sodium citrate), and detergents (Tween 20 and 80) were among the most effective stabilizers as shown by their inhibition of the protein aggregation by ~ 25 to 80%. The negative %IA values are indicative of induced aggregation of protein in the presence of the indicated excipients.

3.3.6. Optimization of Excipient Combinations and Concentrations

Based on the results obtained from preliminary aggregation kinetic-based screening studies, combinations of some of the more promising stabilizers were chosen to see if there was an additive or synergistic effect of these compounds on the inhibition of protein aggregation. A complete list of all combinations examined is presented in Table 3.2. The combination of sodium citrate (≥ 0.1M) with sugars or polyols (≥10% w/v) inhibited the aggregation of fusion protein dramatically.

Excipient (conc.)	% IA	Excipient (conc.)	% IA
Sodium citrate 0.1 M + 20% Dextrose	97	Sodium citrate 0.15 M + 5% Trehalose	81
Sodium citrate 0.3M + 20% Sorbitol	96	Sodium citrate 0.05 M + 20% Trehalose	81
Sodium citrate 0.15M + 20% Sorbitol	96	Sodium citrate 0.15 M + 10% Dextrose	80
Sodium citrate 0.3M + 20% Dextrose	95	Sodium citrate 0.3M + 10% Lactose	77
Sodium citrate 0.1 M + 20% Sorbitol	94	Tween 80 + 20% Dextrose	77
Sodium citrate 0.15M + 20%Dextrose	92	Sodium citrate 0.05 M + 10% Sorbitol	66
Sodium citrate 0.05 M + 20% Dextrose	92	Sodium citrate 0.05 M + 10% Trehalose	65
Sodium citrate 0.1 M + 20% Trehalose	90	Sodium citrate 0.05 M + 10% Dextrose	62
Sodium citrate 0.1 M + 10% Sorbitol	90	Tween 80 + 20% Trehalose	61
Sodium citrate 0.1 M + 10% Trehalose	89	Sodium citrate 0.1 M + 10% Dextrose	61
Sodium citrate 0.3M + 20% Trehalose	88	Sodium citrate 0.15M + 10% Lactose	57
Sodium citrate 0.05 M + 20% Sorbitol	87	Tween 80 + 20% Sorbitol	49
Sodium citrate 0.15 M + 10% Sorbitol	85	Sodium citrate 0.15 M + 5% Dextrose	45
Sodium citrate 0.15M + 20%Trehalose	84	Sodium citrate 0.1 M + 5% Lactose	40
Sodium citrate 0.15 M + 5% Sorbitol	84	Tween 80 + 10% Lactose	19
Sodium citrate 0.15 M + 10% Trehalose	83		

Table 3.2.Percent inhibition of the fusion protein aggregation (%IA) in the presence of combination of variety of excipients after 240 min. The average standard deviation is ± 5% (n=3).

3.3.7. Effect of Selected Excipients on the Conformational Stability of Fusion Protein

To further explore if the inhibitors of aggregation were also stabilizing the structure of the protein, the conformational stability of the fusion protein was examined in the presence of the aforementioned combinations. Effects of combinations of sodium citrate with sugars (trehalose and dextrose) and a polyol (sorbitol) on the tertiary structure of the fusion protein were examined employing tryptophan fluorescence spectroscopy. Enhanced exposure of indole side chains to solvent upon conformational alterations (i.e. unfolding) is manifested by shifts in Trp emission maxima to higher wavelengths (see Fig 3.2a). The stabilizing agents should therefore delay or inhibit such temperature dependent peak shifts.

These studies were performed at protein concentration of 0.1 mg/ml in histidine buffer, pH 6.5. All of the preliminary screening studies, however, were performed at the suboptimal pH 5 at which the degradation of protein was accelerated. The choice of pH 6.5 for the final stage of excipient screening was based on the desire to examine the stability of the fusion protein in the presence of stabilizers at a pH condition close to the targeted formulation pH. Combinations of sodium citrate with dextrose, trehalose, or sorbitol all delayed the unfolding of protein by shifting the T_m values to higher temperatures by ~ 3 to 6 °C. Table 3.3 summarizes all of the combinations examined and their stabilizing effect on protein tertiary structure.

Exipient (conc.)	T_m (°C)	ΔT_m (°C)
Fusion protein	47.3	
0.1M sodium citrate + 20% Dextrose	53.8	6.5
0.05M sodium citrate + 20% Sorbitol	53.6	6.3
0.1M sodium citrate + 20% Sorbitol	53.5	6.2
0.1M sodium citrate + 20% Trehalose	53.4	6.0
0.05M sodium citrate + 20% Dextrose	53.2	5.9
0.05M sodium citrate + 20% Trehalose	53.1	5.7
0.1M sodium citrate + 15% Trehalose	52.8	5.5
0.1M sodium citrate + 15% Dextrose	52.5	5.2
0.1M sodium citrate + 15% Sorbitol	52.3	5.0
0.05M sodium citrate + 10% Sorbitol	52.1	4.7
0.05M sodium citrate + 10% Trehalose	52.0	4.7
0.1M sodium citrate + 10% Trehalose	51.7	4.3
0.1M sodium citrate + 10% Dextrose	51.5	4.2
0.1M sodium citrate + 10% Sorbitol	50.7	3.4
0.05M sodium citrate + 10% Dextrose	50.3	3.0

Table 3.3.Midpoint of protein thermal transitions in the presence of combination of selected excipients. The average standard deviation is less than ± 1 °C (n=3).

3.4. Discussion

The novel fusion anti-brain tumor cytotoxin of this study is composed of mutated versions of Pseudomonas Exotoxin A and Interleukin-13 in which the mutations collectively result in inhibition of the cytotoxin binding to untransformed cells, further enhancing its specificity towards glioma cells. Here, we have characterized the thermal stability of the fusion protein over a wide range of pH (3-8) and temperature (10-85 °C) employing a variety of spectroscopic techniques in search of a stable liquid formulation to permit the protein's therapeutic use.

The CD spectra at 10 °C exhibit predominantly α-helical structure over the pH range of 5-8 (Fig 3.1a), in agreement with the highly rich α-helical contents of both domains. IL-13 is known to be folded into four α-helical bundles[18]. The Pseudomonas Exotoxin A is also rich in α-helical content, specifically in its domain II which contains six α-helices[9]. The CD spectra of both the wild-type toxin and cytokine exhibit minima at 208 and 222 nm very similar to the results obtained herein for the fusion protein[19,20].

Mere et. al. have shown that an acid-triggered unfolding event is an essential step in the retrograde cycling of the exotoxin to escape from endosomal compartments[21]. Tryptophan 305 was identified as a key membrane anchor to initiate the translocation of the exotoxin out of the endosome. This residue is, however, buried in the hydrophobic core of PE at neutral pH, masked by helix F of domain II (i.e., the translocation domain). Upon acidification in the low pH environment of the endosome, helix F rearranges to expose the tryptophan residue to the surrounding

environment allowing it to insert its side chain into endosomal membranes. Our studies also show that the fusion protein α-helical content is highly pH dependent with reduced amounts of α-helix as the pH is lowered (Fig 3.1a). This is consistent with a similar acid triggered α-helical disruptive mechanism as seen in the native toxin.

The effect of temperature on secondary structure was further investigated by monitoring the CD signals at 222 nm. Previous studies with PE show such transitions occurring at ~ 45-50 °C at pH 7.4[22]. The fusion protein's transitions occur in the same temperature range at ~ 45 and 57 °C at pH 7 and 8 respectively (Fig 3.1b) suggesting that the overall secondary structural stability of the protein may be significantly influenced by its toxin domain with minimal contributions from IL-13. This is consistent with the size of the two domains in which the toxin domain constitutes ~ 84% of the total fusion protein's mass.

Alterations in the fusion protein tertiary structure were studied employing intrinsic tryptophan and extrinsic ANS fluorescence. An overall red shift of the Trp emission maximum was observed as a function of increasing temperature over the pH range examined (Fig 3.2a). The midpoint of the thermal transitions was highly pH dependent and in general increased as a function of increasing pH occurring at ~ 36 °C at pH 4 and ~ 48 °C in the neutral pH range of 7-8. Pseudomonas Exotoxin A shows similar pH dependency with T_m values of ~ 36 °C at pH 4 and ~ 56 °C at pH 7[23]. Identical transitions at acidic pH suggest a similar role for the PE structural transitions in acid triggered unfolding of the fusion protein.

Alanine scanning mutagenesis of α-helix D of IL-13 shows that presence of positively charged residues at the three key binding positions of 105, 106, and 109 enhanced the avidity of IL-13 towards its receptor, IL13Rα2[18]. It is therefore plausible that a complementary region in the IL13Rα2 receptor has a negatively charged character, raising the possibility of attractive electrostatic interactions between the two. In addition, pH dependent insertion/membrane binding studies of PE finds the transition pH increasing by about 0.8 units in the presence of anionic lipids in a model membrane system, presumably due to destabilization and enhanced unfolding of PE[23]. In general, polyanions are known to interact non-specifically with a variety of proteins[24-32] and play a role in protein-protein interactions, protein folding and stabilization[33,34]. Our preliminary screening studies showed that the polyanion dextran sulfate was very effective in inhibiting protein aggregation but destabilized the fusion protein's tertiary structure (reducing the T_m by ~ 10 °C, data not shown) suggesting such interactions to be in effect. The fusion protein behaved similarly in the presence of heparin in which the polyanion inhibited its aggregation by ~ two fold but destabilized its tertiary structure by reducing the T_m by ~ 7 °C (data not illustrated). The destabilizing effect of polyanions on the fusion protein is in agreement with the destabilizing effect of anionic lipids of the endosomal membranes on the toxin domain. Therefore, the effect of polyanions on conformational stability of this novel fusion cytotoxin should be studied further since such interactions have also proven to be essential for the proper functioning of a number of other macromolecular complexes[35-40].

One similarity between the fusion protein and its toxin domain is the lack of any clear transitions in temperature dependent tryptophan peak shifts at pH values below 4. This is observed at pH 3 for the fusion protein (Fig 3.2a, 3.4a, and 3.4b) and at pH values below 3.7 for PE[23]. Jiang *et al.* show that PE undergoes a two step conformational alteration in the acidic pH range exhibiting two distinct conformational states[23]; an intermediate low pH folded state is observed in the pH range of 3.7-5.4 with very weak or no hydrophobicity followed by a second low pH unfolded-like conformation below pH 3.7 containing strong apolarity. The pH-dependent behavior of the fusion protein suggests a similar highly disrupted conformation at acidic pH values. The fusion protein shows enhanced hydrophobicity as a function of decreasing pH in which the onset of ANS binding to apolar regions of protein shifts from ~ 40 °C at pH 5 to ~ 25 °C at pH 4. No clear transitions are observed at pH 3 presumably due to the presence of already structurally disrupted protein, in agreement with the behavior of the toxin domain.

The tryptophan emission maximum of PE occurs at ~ 336 nm below pH 3.7, 331 nm between pH 3.7 and 5.4, and 334 nm above pH 5.4[23]. The fusion protein follows a similar trend in which Trp emission maxima were observed at ~ 332 nm at pH 3, ~ 330 nm at pH 4, and ~ 331 nm above pH 4 (calculated based on a direct peak picking method). The similar pH dependent emission peak shifts again suggest similar tertiary structural alterations between the two proteins. The magnitudes of the shifts are, however, smaller for the fusion protein, presumably due to the presence of IL-13 and the differences in stability between wild type PE and its derivative, PE4E.

Both circular dichroism and fluorescence spectroscopy studies show more thermally stable protein at pH values at or above neutral pH (Fig 3.1, 3.2, and 3.3) in agreement with optimal toxin enzymatic activity at more basic pH values[20]. Thus, the results presented here strongly suggest that the thermal stability of the fusion protein can be understood at least qualitatively as a summation of the transitions of its individual domains.

Analysis of the secondary and tertiary structural alterations of the fusion protein suggests the presence of intermediate molten globule (MG) states across the pH range 5-8. Comparing the onset transition temperatures obtained from CD studies (Fig 3.1b) to those of tryptophan peak position (Fig 3.2a) and ANS intensity (Fig 3.4a) shows that tertiary structure changes precede secondary structure alterations, a key characteristic of MG states. MG sates are often characterized by their extensive secondary structure, minimal tertiary structure, and a tendency to aggregate[41]. The importance of an acid-triggered unfolding event in membrane penetration ability of the toxin and considering the similarities in structural alterations between the fusion protein and its toxin domain presented here, it is plausible that qualitatively equivalent MG states play a role in the protein's biological activity. Presence of such states might be essential for endosomal membrane penetration and translocation across such membranes. Studies of Sanyal and Gress[42,43] on TP40, a fusion protein composed of transforming growth factor-α and a derivative of pseudomonas exotoxin, showed presence of similar intermediate molten globule states. Although the secondary structural alterations of the fusion protein herein are more extensive

compared to those of TP40 and therefore nature of the molten globule states appear to be weakened, the results contribute to emerging view of involvement of MG states in facilitating membrane translocation of fusion cytotoxins.

By further integration of the spectroscopic data into an empirical phase diagram, we were able to identify structural changes which served as a basis for screening of effective stabilizers. It is well known that a variety of potential additives such as sugars, amino acids, surfactants, salts and polyols can protect proteins against chemical and physical degradation in both liquid and lyophilized formulations[44-47]. Our studies indicate that combinations of sodium citrate with dextrose, trehalose, or sorbitol are effective in both inhibiting the fusion protein aggregation as well as stabilizing its tertiary structure under accelerated degradation conditions (Tables 3.2&3.3). The aggregation kinetic trace of the fusion protein (Fig 3.6) showed a characteristic lag time implying the existence of a critical nucleation-dependent aggregation event which has frequently been observed for a variety of other proteins [47-50]. These promising stabilizers proved to be effective in inhibiting the extent of aggregation but none showed the ability to lengthen the nucleation lag phase (not illustrated).

Final concentration and combination of more promising stabilizers identified herein should be further optimized to provide a suitable liquid formulation compatible with the protein's therapeutic application and its route of administration. Future studies are thus focused on evaluating and optimizing parameters such as ionic strength, osmolarity, and viscosity of the protein formulation and its integrity over its

shelf life upon addition of such excipients. The success of such potential formulations should be further examined in long-term stability studies which will provide a further test of the formulation’s stabilizing ability.

3.5. Bibliography

(1) Debinski W. Molecular targeting of brain tumors with cytotoxins: Novel bacterial toxin-containing anti-brain tumor therapeutics. Cellular and Molecular Mechanisms of Toxin Action (2002), 4, 222-246.

(2) Husain SR, Puri RK. Interleukin-13 receptor-directed cytotoxin for malignant glioma therapy: From bench to bedside. Journal of neuro-oncology (2003), 65, 37-48.

(3) Debinski W. Anti-brain tumor cytotoxins. Science & Medicine (Philadelphia) (1998), 5, 36-42.

(4) Debinski W. Recombinant cytotoxins specific for cancer cells. Annals of the New York Academy of Sciences (1999), 886, 297-299.

(5) Minty A, Chalon P, Derocq JM, Dumont X, Guillemot JC, Kaghad M, Labit C, Leplatois P, Liauzun P, Miloux B. Interleukin-13 is a new human lymphokine regulating inflammatory and immune responses. Nature (1993), 362, 248-50.

(6) McKenzie AN, Culpepper JA, de Waal Malefyt R, Briere F, Punnonen J, Aversa G, Sato A, Dang W, Cocks BG. Interleukin-13, a T-cell derived cytokine that regulates human monocyte and B-cell function. Proc. Nat. Acad. Sci. U.S.A. (1993), 90, 3735-9.

(7) Madhankumar AB, Mintz A, Debinski W. Interleukin 13 mutants of enhanced avidity toward the glioma-associated receptor, IL13Rα2. Neoplasia (Ann Arbor, MI, United States) (2004), 6, 15-22.

(8) Wedekind JE, Trame CB, Dorywalska M, Koehl P, Raschke TM, McKee M, FitzGerald D, Collier RJ, McKay DB. Refined crystallographic structure of Pseudomonas aeruginosa Exotoxin A and its Implications for the molecular mechanism of toxicity. J. Mol. Biol. (2001), 314, 823-837.

(9) Allured VS, Collier RJ, Carroll SF, McKay DB. Structure of exotoxin A of Pseudomonas aeruginosa at 3.0-Angstrom resolution. Proc. Nat. Acad. Sci. U.S.A. (1986), 83, 1320-4.

(10) Debinski W, Puri RK, Kreitman RJ, Pastan I. A wide range of human cancers express interleukin 4 (IL4) receptors that can be targeted with chimeric toxin composed of IL-4 and Pseudomonas exotoxin. J. Biol. Chem. (1993), 268, 14065-70.

(11) Chaudhary VK, Jinno Y, Gallo MG, FitzGerald D, Pastan I. Mutagenesis of Pseudomonas exotoxin in identification of sequences responsible for the animal toxicity. J. Biol. Chem. (1990), 265, 16306-10.

(12) Ausar SF, Espina M, Brock J, Thyagarayapuran N, Repetto R, Khandke L, Middaugh CR. High-throughput screening of stabilizers for respiratory syncytial virus: Identification of stabilizers and their effects on the conformational thermostability of viral particles. Human Vaccines (2007), 3, 94-103.

(13) Randolph TW, Carpenter JF. Engineering challenges of protein formulations. AIChE J. (2007), 53, 1902-1907.

(14) Kueltzo LA, Ersoy B, Ralston JP, Middaugh CR. Derivative absorbance spectroscopy and protein phase diagrams as tools for comprehensive protein characterization: A bGCSF case study. J. Pharm. Sci. (2003), 92(9), 1805-1820.

(15) Lakowicz JR. Principles of Fluorescence Spectroscopy. 2nd ed, Kluwer Academic / Plenum Publisher (New York, NY) (1999), 698 pp.

(16) Matulis D, Baumann CG, Bloomfield VA, Lovrien RE. 1-Anilino-8-Naphthalene Sulfonate as a protein conformational tightening Agent. Biopolymers (1999), 49, 451-458.

(17) Matulis D, Lovrien RE.1-Anilino-8-Naphthalene Sulfonate anion-protein binding depends primarily on ion pair formation. Biophys. J. (1998), 74, 422-429.

(18) Madhankumar AB, Mintz A, Debinski W. Alanine-scanning mutagenesis of α-Helix D segment of interleukin-13 reveals new functionally important residues of the cytokine. J. Biol. Chem. (2002), 277, 43194-43205.

(19) Cannon-Carlson S, Varnerin J, Tsarbopoulos A, Jenh CH, Cox MA, Chou CC, Connelly N, Zavodny P, Tang JC. Expression, purification, and characterization of recombinant human interleukin-13 from NS-0 cells. Protein Expression and Purif. (1998), 12, 239-248.

(20) Beattie BK, Merrill AR. In vitro enzyme activation and folded stability of Pseudomonas aeruginosa Exotoxin A and Its C-terminal peptide. Biochemistry (1996), 35, 9042-9051.

(21) Mere J, Morlon-Guyot J, Bonhoure A, Chiche L, Beaumelle B. Acid-triggered membrane insertion of Pseudomonas Exotoxin A involves an original

mechanism based on pH-regulated tryptophan exposure. J. Biol. Chem. (2005), 280, 21194-21201.

(22) McKee ML, FitzGerald DJ. Reduction of Furin-nicked Pseudomonas Exotoxin A: An unfolding story. Biochemistry (1999), 38, 16507-16513.

(23) Jiang JX, London E. Involvement of denaturation-like changes in Pseudomonas exotoxin A hydrophobicity and membrane penetration determined by characterization of pH and thermal transitions. Roles of two distinct conformationally altered states. J. Biol. Chem. (1990), 265, 8636-41.

(24) Jones LS, Yazzie B, Middaugh CR. Polyanions and the proteome. Molecular and Cellular Proteomics (2004), 3, 746-769.

(25) MacLean DS, Qian Q, Middaugh CR. Stabilization of proteins by low molecular weight multi-ions. J. Pharm. Sci. (2002), 91, 2220-2229.

(26) Tao J, Frankel AD. Electrostatic interactions modulate the RNA-binding and transactivation specificities of the human immunodeficiency virus and simian immunodeficiency virus Tat proteins. Proc. Nat. Acad. Sci. U.S.A. (1993), 90, 1571-5.

(27) Salamat-Miller N, Fang J, Seidel CW, Assenov Y, Albrecht M, Middaugh CR. A network-based analysis of polyanion-binding proteins utilizing human protein arrays. J. Biol. Chem. (2007), 282, 10153-10163.

(28) Volkin, David B; Tsai, P. K.; Dabora, Jonathan M; Gress, Jacqueline O; Burke, Carl J; Linhardt, Robert J; Middaugh, C. Russell. Physical stabilization of acidic fibroblast growth factor by polyanions. Arch. Biochem. Biophys (1993), 300(1), 30-41.

(29) Volkin DB, Verticelli AM, Marfia KE, Burke CJ, Mach H, Middaugh CR. Sucralfate and soluble sucrose octasulfate bind and stabilize acidic fibroblast growth factor. Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol. (1993), 1203, 18-26.

(30) Fang J, Dong Y, Salamat-Miller N, Middaugh CR. DB-PABP: a database of polyanion-binding proteins. Nucleic Acids Research (2008), 36, D303-D306.

(31) Joshi SB, Kamerzell TJ, McNown C, Middaugh CR. The interaction of heparin/polyanions with bovine, porcine, and human growth hormone. J. Pharm. Sci. (2008), 97, 1368-1385.

(32) Salamat-Miller N, Fang J, Seidel CW, Smalter AM, Assenov Y, Albrecht M, Middaugh CR. A network-based analysis of polyanion-binding proteins utilizing yeast protein arrays. Molecular and Cellular Proteomics (2006), 5, 2263-2278.

(33) Hianik T, Ponikova S, Bagel'ova J, Antalik M. Specific volume and compressibility of human serum albumin-polyanion complexes. Bioorganic & Medicinal Chemistry Letters (2006), 16, 274-279.

(34) Shalova IN, Asryants RA, Sholukh MV, Saso L, Kurganov BI, Muronetz VI, Izumrudov VA. Interaction of polyanions with basic proteins, 2a: Influence of complexing polyanions on the thermoaggregation of oligomeric enzymes. Macromol. Biosci. (2005), 5, 1184-1192.

(35) Dormitzer PR, Nason EB, Venkataram PB, Harrison SC. Structural rearrangements in the membrane penetration protein of a non-enveloped virus. Nature (London, United Kingdom) (2004), 430, 1053-1058.

(36) Schneider S, Sharp KH, Barker PD, Paoli M. An induced fit conformational change underlies the binding mechanism of the heme Transport Proteobacteria-protein HemS. J. Biol. Chem. (2006), 281, 32606-32610.

(37) Hayashi H, Mizuguchi H, Miyahara I, Nakajima Y, Hirotsu K, Kagamiyama H. Conformational change in aspartate aminotransferase on substrate binding induces strain in the catalytic group and enhances catalysis. J. Biol. Chem. (2003), 278, 9481-9488.

(38) Mayer C, RajBhandary UL. Conformational change of Escherichia coli initiator methionyl-tRNAfMet upon binding to methionyl-tRNA formyltransferase. Nucleic Acids Res. (2002), 30, 2844-2850.

(39) Dudas KC, Scouten SK, Ruyechan WT. Conformational change in the Herpes simplex single-strand binding protein induced by DNA. Biochem. Biophy. Res. Commun. (2001), 288, 184-190.

(40) Hanein D, Matsudaira P, Derosier DJ. Evidence for a conformational change in actin induced by fimbrin (N375) binding. J. Cell. Biol. (1997), 139, 387-396.

(41) Mach H, Ryan JA, Burke CJ, Volkin DB, Middaugh CR. Partially structured self-associating states of acidic fibroblast growth factor. Biochemistry (1993), 32, 7703-11.

(42) Sanyal G, Marquis-Omer D, Gress JO, Middaugh CR. A transforming growth factor-α-Pseudomonas exotoxin hybrid protein undergoing pH-dependent

conformational changes conducive to membrane interaction. Biochemistry (1993), 32, 3488-97.

(43) Gress JO, Marquis-Omer D, Middaugh CR, Sanyal G. Evidence for an equilibrium intermediate in the folding-unfolding pathway of a transforming growth factor-α-Pseudomonas Exotoxin hybrid protein. Biochemistry (1994), 33, 2620-7.

(44) Wong D, Parasrampuria J. Pharmaceutical excipients for the stabilization of proteins. BioPharm (Eugene, Oregon) (1997), 10, 52-56, 58, 60-61.

(45) Timasheff SN, Arakawa T. Stabilization of protein structure by solvents. Protein Structure (2nd Edition) (1997), 349-364.

(46) Arakawa T, Kita Y, Carpenter JF. Protein-solvent interactions in pharmaceutical formulations. Pharm. Res. (1991), 8, 285-91.

(47) Kim YS, Cape SP, Chi EY, Raffen R, Wilkins-Stevens P, Stevens FJ, Manning MC, Randolph TW, Solomon A, Carpenter JF. Counteracting effects of renal solutes on amyloid fibril formation by immunoglobulin light chains. J. Biol. Chem. (2001), 276, 1626-1633.

(48) Hamada D, Dobson CM. A kinetic study of β-lactoglobulin amyloid fibril formation promoted by urea. Protein Sci. (2002), 11, 2417-2426.

(49) Come JH, Fraser PE, Lansbury PT. A kinetic model for amyloid formation in the prion diseases: Importance of seeding. Proc. Nat. Acad. Sci. U.S.A. (1993), 90, 5959-63.

(50) Zurdo J, Guijarro JI, Jimenez JL, Saibil HR, Dobson CM. Dependence on solution conditions of aggregation and amyloid formation by an SH3 domain. J. Mol. Biol. (2001), 311, 325-340.

Chapter 4

Biophysical Characterization of Rotavirus Serotypes G1, G3, and G4

4.1. Introduction

Acute infectious dehydrating diarrhea is the most common cause of morbidity and mortality among children resulting in more than five million annual deaths worldwide[1]. Viruses are recognized as the most important cause of this disorder[2] with rotavirus as the most fatal etiologic agent in this category. It is alone responsible for more than half a million deaths annually[3]. Essentially all children under the age of five regardless of sanitation, socioeconomics, or geography are infected. The rate of mortality, however, differs dramatically between developed and developing countries with 85% of the deaths occurring in the poorest countries of Africa and Southeast Asia[4].

The epidemiologic characteristics associated with rotavirus recommend vaccination as the primary public health intervention. This is based on the observed similar morbidity rates in both developed and developing countries. Moreover, existing high hospitalization rates suggests that neither improved sanitation nor the available oral rehydration therapies can effectively prevent rotavirus-related gastroentritis. Extensive studies, however, suggest that mild asymptomatic infections effectively protect against subsequent severe rotavirus infections indicating the use of a vaccine which produces mild infections[5].

The first rotavirus vaccine, Rotashield by Wyeth-Lederle, was introduced in the United States in 1998[6]. A serious complication (i.e. intussusception, a fatal bowel obstruction) associated with this tetravalent human-bovine reassortant vaccine, however, resulted in its withdrawal from the market less than a year later[7]. The next

generation of rotavirus vaccines, the GlaxoSmithKline's monovalent vaccine, RotaRix[8], and Merck's pentavalent human-bovine reassortant vaccine, RotaTeq[9], were introduced in 2005 and 2006 respectively. These vaccines have already had a dramatic effect on rotavirus induced disease and mortality in the developed world. The complications and high costs associated with the distribution of these vaccines in developing world, however, have limited their use in the poorest countries of Africa and Southeast Asia where the highest rates of rotavirus mortality reside. Therefore, a need for a third generation of rotavirus vaccines targeted towards the developing world still exists.

One major complication in the development of vaccines for mass immunization, particularly in the developing world, is creation of a vaccine formulation that can be stored and remained biologically active over the entire required shelf life of 24 months at as high a temperature as possible. The main antigenic component of vaccine formulations is some form of the infective virus, either killed or live but attenuated. Virus particles are large macromolecular complexes[10] with inherent thermolability which is responsible for lack of optimal vaccine stability[11-16]. Storage at refrigerated temperatures or as lyophilized formulations is generally required for retaining the stability and efficacy of vaccines. This can, however, introduce significant difficulties in their storage, shipment, and delivery. Due to insufficient technological support including lack of refrigerators and frequent power failures, vaccines can be exposed to elevated temperatures for extended periods of time. This has traditionally resulted in wastage of approximately

half of the supplied vaccines[17]. Thus, there is a need for thermally stable vaccine formulations.

The traditional approach to evaluating vaccine efficacy has been by means of measuring the production of antibodies in animal models in response to vaccine administration. Although this method provides information regarding the efficacy of the vaccine formulation, it forces an essentially empirical approach to improve stability. Biophysical studies on the other hand have the potential to provide insight into the mechanisms of inactivation of a viral vaccine as a function of a variety of environmental conditions. Herein, we have employed a variety of spectroscopic techniques (i.e., circular dichroism, fluorescence spectroscopy, and static and dynamic light scattering) for a comprehensive examination of the thermal stability of three live-attenuated human-bovine reassortant rotavirus strains (G1, G3, and G4) over a wide pH range (i.e. 5-8). The three serotypes examined, classified based on the molecular composition of their outer coat viral glycoprotein, VP7 [18], are ideal vaccine candidates. These strains are the most infectious in their family accounting for ~ 70% of all human rotavirus infections worldwide[19]. In addition, the already marketed rotavirus vaccine, Merck's RotaTeq, includes these three strains. This vaccine has been shown to provide a synergistic immunological response effective against each individual strain[9]. The major issue we address here is whether the methods we employ which produce results that are averaged over the entire virus structure, can detect destabilizing events which can be used to produce more stable formulations.

4.2. Materials and Methods

4.2.1. Sample Preparation

The purified bulk rotavirus samples were stored at -80 °C in 11 mM potassium phosphate buffer, pH 7.2, containing 7.22% sucrose for improved long term storage. Samples were thawed at room temperature and were concentrated ten fold employing an Amicon stirred ultrafiltration cell (model 8400) for spectroscopic studies. The samples were then buffer exchanged into 20 mM isotonic citrate phosphate buffer ranging from pH 5 to 8, at one pH unit intervals (the isotonicity was maintained using sodium chloride). Exchange of the viral stock solutions into citrate phosphate buffer below pH 5 resulted in aggregation of the viral particles and therefore was not used in our studies. For buffer exchange, rotavirus samples were dialyzed at refrigeration temperatures using a Spectra/Por molecular porous membrane tubing (SpectrumLab, Rancho Dominguez, CA) with a 15KDa MW cutoff. The concentration of the samples after dialysis was determined by BCA assay (Pierce, Rockford, IL). Samples at a concentration of 100 μg/ml were used for all spectroscopic studies with the exception of CD and dynamic light scattering in which concentrations of 200 and 50 μg/ml were employed, respectively. For all spectroscopic studies, three independent samples were evaluated to ensure reproducibility of the measurements. A virus concentration of 60 μg/ml was used for the electron microscopy imaging.

4.2.2. Far-UV Circular Dichroism (CD) Spectroscopy

CD spectra were acquired using a Jasco J-810 spectropolarimeter equipped with a 6-position Peltier temperature controller. The CD spectra were obtained from 260-190 nm at a scanning speed of 20 nm/min, a 2 second response time and an accumulation of 3. To study thermal transitions (melting curves) of the rotavirus, the CD signal at 222 nm was monitored in 0.1 cm pathlength cuvettes every 0.5 °C over a 10 to 85 °C temperature range employing a temperature ramp of 15 °C/hr.

4.2.3. Intrinsic Tryptophan (Trp) Fluorescence Spectroscopy

Fluorescence spectra were acquired using a Photon Technology International (PTI) spectrofluorometer (Lawrenceville, NJ) equipped with a turreted 4-position Peltier-controlled cell holder. An excitation wavelength of 295 nm was used to excite Trp and the emission spectra were collected from 300 to 400 nm with a step size of 1 nm and 1 sec integration time. Excitation slits were varied over the pH range. Therefore, both the fluorescence intensity and light scattering data were normalized with respect to the initial reading at 10 °C for comparison purposes. Emission spectra were collected every 2.5 °C with a 3 min equilibration time over a temperature range of 10 to 85 °C. A buffer baseline was subtracted from each raw emission spectrum.

4.2.4. Dynamic Light Scattering

The hydrodynamic diameter of the viral particles was analyzed using a dynamic light scattering instrument (Brookhaven Instrument Corp., Holtzille, NY)

equipped with a 50 mW diode-pumped laser (λ = 532 nm). The scattered light was monitored at 90° to the incident beam, and autocorrelation functions were generated using a digital auto-correlator (BI-9000AT). Data were collected continuously for five 30-seconds intervals for each sample and averaged. The hydrodynamic diameter was calculated from the diffusion coefficient by the Stokes-Einstein equation using the method of cumulants[20]. The instrument was equipped with a temperature-controlled circulating water bath (RTE111, Neslab, Newington, NH) and the hydrodynamic diameter was monitored over a temperature range of 10 to 85 °C.

4.2.5. Transmission Electron Microscopy

Electron microscopy imaging of uranyl acetate-stained rotavirus strains was performed to analyze structural alterations due to changes in pH and temperature. Briefly, rotavirus samples were dialyzed into 20 mM isotonic citrate phosphate buffer at pH 5 and 7 for about 16 h at refrigeration temperature. The dialyzed samples were recovered and conformational changes were analyzed under variable temperatures by negatively staining with uranyl acetate and viewed with a Tecnai G^2 transmission electron microscope (FEI Company, Hillsboro, OR). Rotavirus samples at 60 μg/ml were placed into small vials and incubated in a water bath. Vials were heated from 10 to 90 °C at a ramp rate of 15 °C / hr. At 4, 40 and 70 °C, virus samples (20 μl) were collected and were immediately placed onto the carbon-coated grids (Electron Microscopy Sciences, Hatfield, PA) for 2 min. After washing with three consecutive drops of distilled water, the grids were then placed into 2% uranyl acetate for 1 min.

The extra stain was removed by blotting with tissue paper. The wet grids were allowed to air-dry for at least 48 hours prior to being examined by the TEM.

4.2.6. Fluorescence Focus Assay

The assay plates were prepared with a monolayer of MA-104 cells in flat-bottomed 96-well assay plates at 10,000 cells per well (200 µL cell suspension per well). The plates were stored in a humidified 37 °C incubator, 5% CO_2 for 3-5 days until a confluent monolayer was reached (i.e., 4th day post plating). The old media was replaced with fresh MEM on the day of the assay. The virus samples were activated with 5 µg/mL trypsin (final concentration) for one hour in a humidified 37 °C incubator (5% CO_2). Serial dilutions (1:4) of the virus samples were made in duplicate in a separate 96-well dilution block allowing one well as cell control for each sample. 50 µL of the diluted samples were transferred to the 96-well assay plate containing 150 µL MEM per well. The assay plates were stored in a humidified 36 °C incubator (5% CO_2) for 18 hours. The cell monolayer was washed with MEM post-incubation and fixed with 80% Acetone for 20 minutes at -20 °C. 50 µL of the diluted primary (i.e., `rabbit polyclonal`) antibody was added to each well of the assay plate and incubated at 37 °C for 60 min. The secondary (i.e., `goat anti-rabbit IgG (H + L), biotinylated`) antibody at 50 µL/well was added after washing the cell monolayer with Tween PBS (`0.5% Tween 20 in PBS`). The plates were incubated at 37 °C for another 60 min and then washed with TPBS. The Alexa Fluor 488 Streptavidin conjugate was added at 50 µL/well and incubated

at 37 °C for 60 minute and finally washed with Tween PBS. To avoid bleaching, the stained assay plates were wrapped with foil and stored in a 4 °C refrigerator. Fluorescent cells were imaged under a microscope with an appropriate filter and lamp. The fluorescent cell forming units per mL were calculated based on the count of the fluorescent cells, the dilution used, the magnification, and the total surface area. The titer is an average of the two duplicate assays performed on one sample.

4.2.7. Empirical Phase Diagrams (EPDs)

Empirical phase diagrams were constructed employing CD, intrinsic Trp fluorescence peak position, and static and dynamic light scattering data. All calculations were performed using Matlab software (The MathWorks, Natick, MA). In brief, the normalized experimental data at each coordinate (i.e., at specific temperature and pH combinations) are first converted into an *N*-dimensional vector, where *N* refers to the number of variables included (i.e., number of different types of data). The complex data sets from all measurements are now defined as multi-dimensional vectors in a pH/temperature phase space. The projectors of each individual vectors are then calculated and summed into an $N \times N$ density matrix. By definition, an $N \times N$ matrix has *N* sets of eigenvalues and eigenvectors. The individual vectors at each coordinate are truncated into 3-dimensional vectors and re-expanded into a new basis set consisting of the three eigenvectors corresponding to the three largest eigenvalues. The resultant 3-dimensional vectors are then converted into a color plot with each vector component corresponding to a color using an

arbitrary RGB (red, green, blue) color scheme. Details of the mathematical theory and calculation process can be found elsewhere[21].

4.3. Results

4.3.1. Far-UV Circular Dichroism (CD) Spectroscopy

Alterations of rotavirus secondary structure were studied by monitoring the changes in the CD signals as a function of pH and temperature. Each serotype shows a distinct pH dependence of its secondary structure content (Fig 4.1). All of the spectra display the well recognized signature of α-helix, a double minima at 208 and 222 nm. With the exception of G4 at pH 8, the intensity of the CD ellipticity at 222 nm is greater or equal to the signal at the lower wavelength. This is consistent with a significant content of β-structure in almost all cases. What is most distinctive about the results is that the CD spectrum of each of the 3 serotypes responses differently to pH. This is most clearly seen at pH 6 where G1 demonstrates the least intense spectrum in contrast to the other pH values. Because of the low quality of the data below 205 nm, it is not possible to make quantitative estimates of secondary structure content.

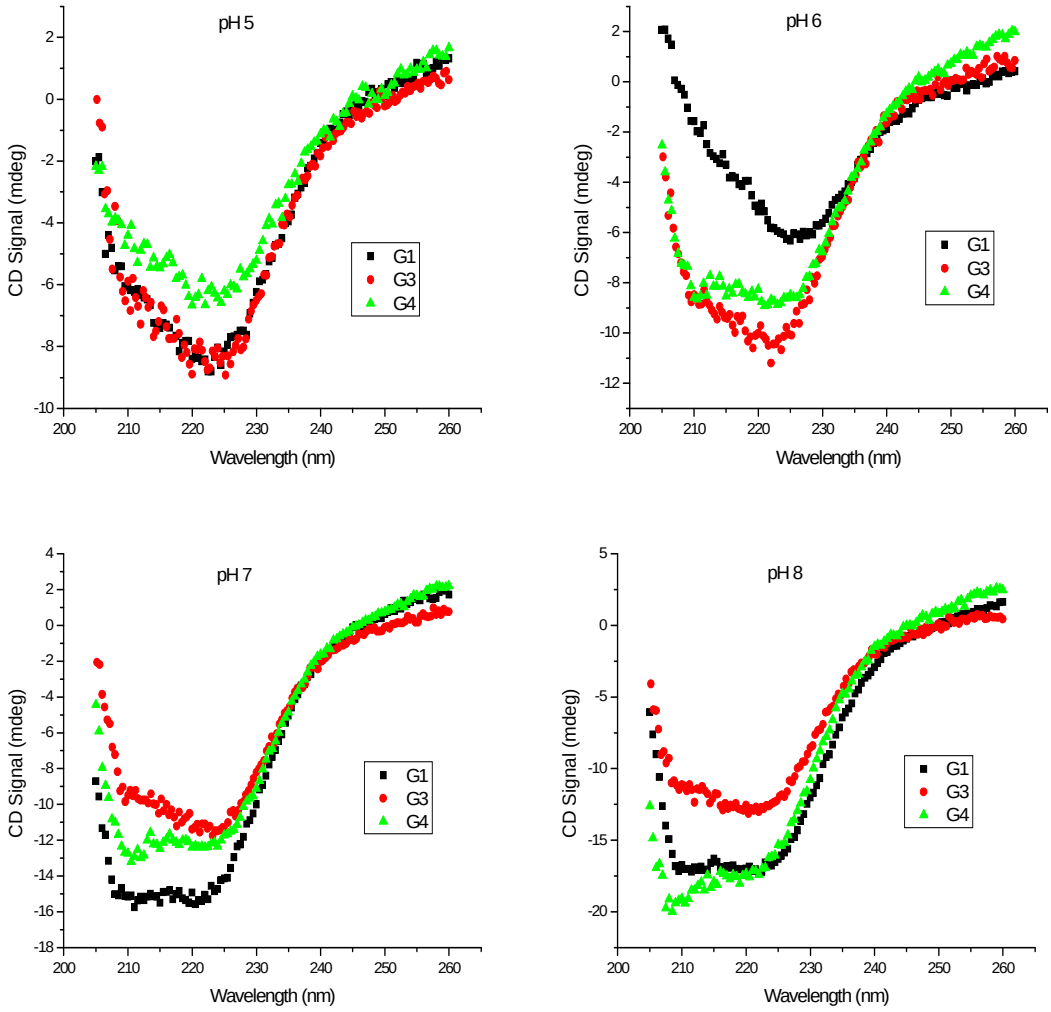


Figure 4.1.CD spectra of three rotavirus strains at various pH values. CD spectra were recorded at 10 ° C from 190 to 260 nm at each of the indicated pH values (*n* =3).

The effect of temperature on the rotavirus secondary structure was further investigated employing thermal melt analysis (Fig 4.2). The destabilizing effect of increasing temperatures on rotavirus secondary structural stability is indicated by an overall loss of the secondary structure content of all three strains over the pH range examined. The transitions at pH 5 are all very gradual. Following this initial gradual transition, strains G1 and G3 exhibit a much sharper transition at higher temperature centered near 75 °C. At pH 6, G1 and G4 show biphasic behavior with the first transition onset temperatures at ~ 37.5 and 50 °C, respectively. A later transition is observed at ~ 75 °C for both strains. At this pH, the G3 strain shows a gradual transition at lower temperatures followed by a sharp transition which begins at ~ 70 °C. All three strains show biphasic behavior at pH 7 with the earlier transitions occurring at ~ 40, 45, and 47.5 °C for G1, G3, and G4, respectively. Again, a later transition at ~ 70 °C for G1 and G3 and ~ 75 °C for G4 is observed. The G3 strain exhibits biphasic behavior at pH 8 with a gradual transition at low temperatures followed by a later transition with an onset at ~ 70 °C. Early transitions at ~ 35 and 40 °C for the G1 and G4 strains is followed by a second transition at ~ 65 and 70 °C respectively. Overall, the G4 strain shows higher thermal stability compared to the other two strains in terms of thermally induced changes in CD signals. The onset of the early transitions is lowered as the pH is increased over the range of 6 to 8, suggesting reduced thermal stability as a function of increasing pH.

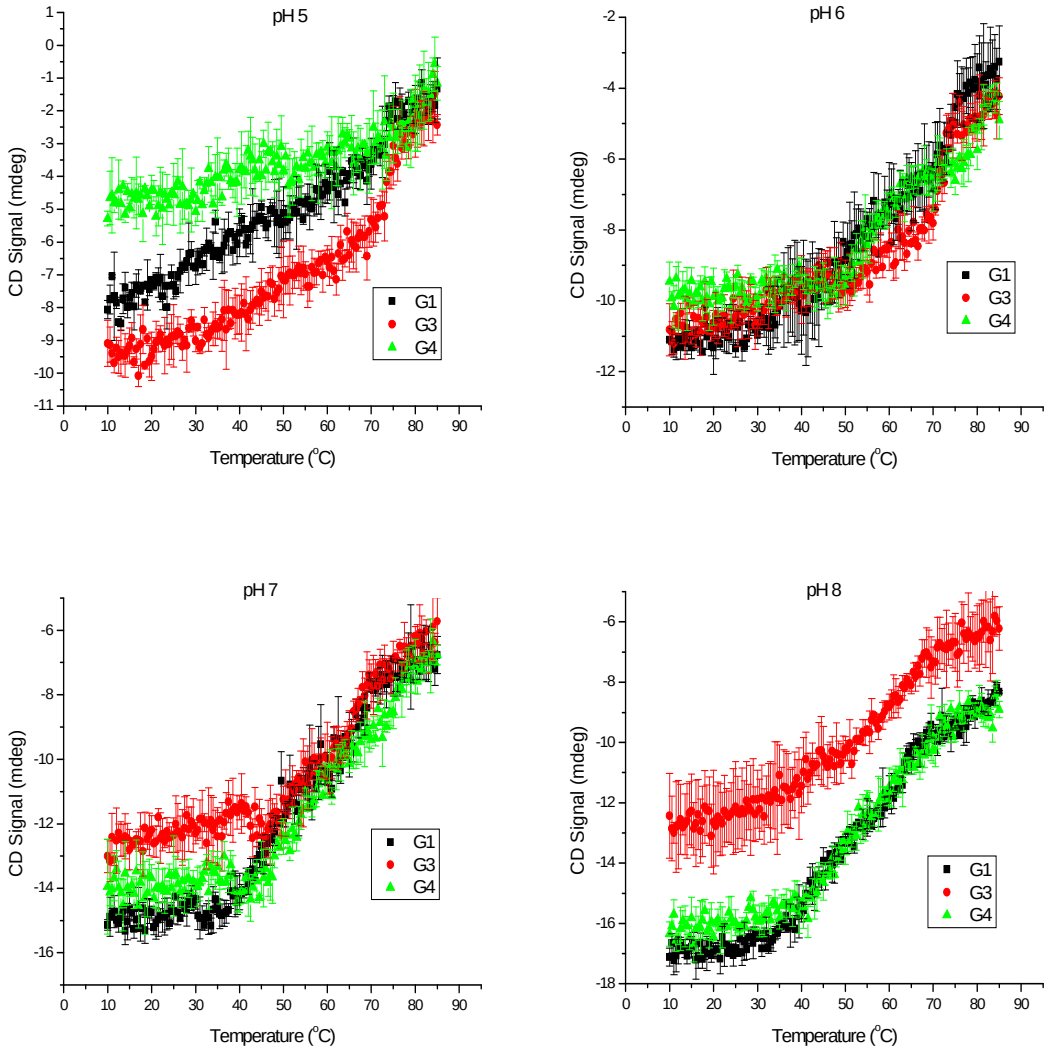


Figure 4.2.The effect of temperature on the secondary structure of rotavirus strains at different pH values. Signal at 222 nm was monitored as a function of temperature over the pH range 5–8 (n = 3).

4.3.2. Intrinsic Tryptophan (Trp) Fluorescence Spectroscopy

Alterations in the rotavirus tertiary structure were studied by monitoring changes in the Trp emission peak position as a function of pH and temperature. An overall red shift in the Trp emission maximum is observed as a function of increasing temperature over the pH range examined (Fig 4.3) suggesting enhanced exposure of the indole side chains to solvent upon conformational alterations. The magnitude of the red shifts at pH 5 is smaller for all the strains compared to higher pH. Over the pH range of 6-8, a gradual red shift at low temperatures is observed for all the strains followed by sharper transitions at higher temperature. The onsets of the higher temperature transition are similar among the three strains and occur at ~ 67.5 °C at pH 6, 62.5 °C at pH 7, and 55 °C at pH 8, indicating decreased tertiary structure stability as a function of increasing pH and in direct agreement with the CD studies. G3 exhibits the largest red shifts of ~ 9-10 nm over the pH range of 6 to 8 suggesting more pronounced unfolding events for this strain. The extent of temperature-induced exposure of indole side chains to solvent is similar for the G1 and G4 strains over this pH range with red shifts of ~ 4-5 nm in magnitude. The initial peak positions near 330 nm suggest that on average, the indole side chains of the Trp residues in the viral particle are buried in a fairly apolar environment. The temperature dependent Trp fluorescence intensities exhibited no distinct transitions from pH 5 to 8 and only a curvilinear decrease due to the intrinsic thermal quenching of the fluorescence signals was observed (data not shown).

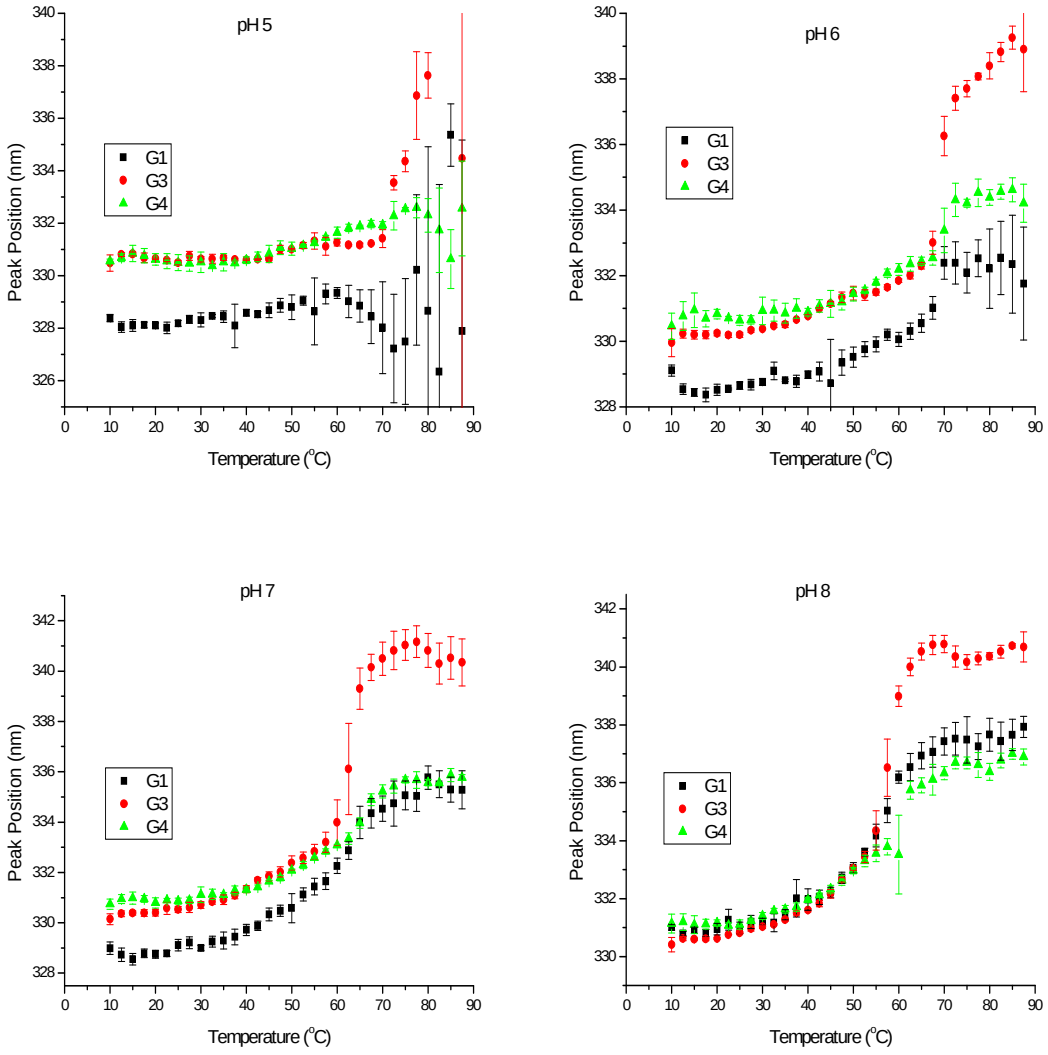


Figure 4.3.Tryptophan emission peak position as a function of pH and temperature. Rotavirus solutions at pH 5–8 were heated from 10 to 85 °C, and the fluorescence emission maxima were monitored (*n*=3) upon excitation at 295 nm.

4.3.2.1. Static Light Scattering

The aggregation behavior of the rotavirus strains was simultaneously examined by monitoring the scattered light at the wavelength of fluorescence excitation (Fig 4.4). Complex, multi-phasic behavior is observed for all strains across the pH range examined. At pH 5, the G3 and G4 strains show multiple weak transitions possibly due to the formation of micro-aggregates and soluble oligomers. A much sharper transition is observed for the G1 strain at ~ 40 °C. At pH 6, both the G3 and G4 strains exhibit an early transition at ~ 42.5 °C and a later one at ~ 67.5 and 70 °C respectively. The G1 strain manifests a significant transition at ~ 35 °C. At pH 7, G3 shows no distinct early transitions, and only a weak one at ~ 70 °C while G4 shows multiple transitions at ~ 45 and 70 °C. The G1 strain shows similar behavior to that at pH 6 with the onset of the transition observed at ~ 35 °C. At pH 8, the G1 and G4 strains show a weak transition at ~ 60 °C while G3 shows no significant transitions (Fig 4.4). Overall, the G1 strain shows more pronounced transitions suggesting a greater propensity for the formation of soluble oligomers at lower temperatures. This can lead to more extensive aggregation as a function of increasing temperatures as suggested by the dynamic light scattering data (see below). Decreases in scattering intensity probably reflect particle swelling since the refractive index increment (dn/dc) is a square function of the scattering intensity.

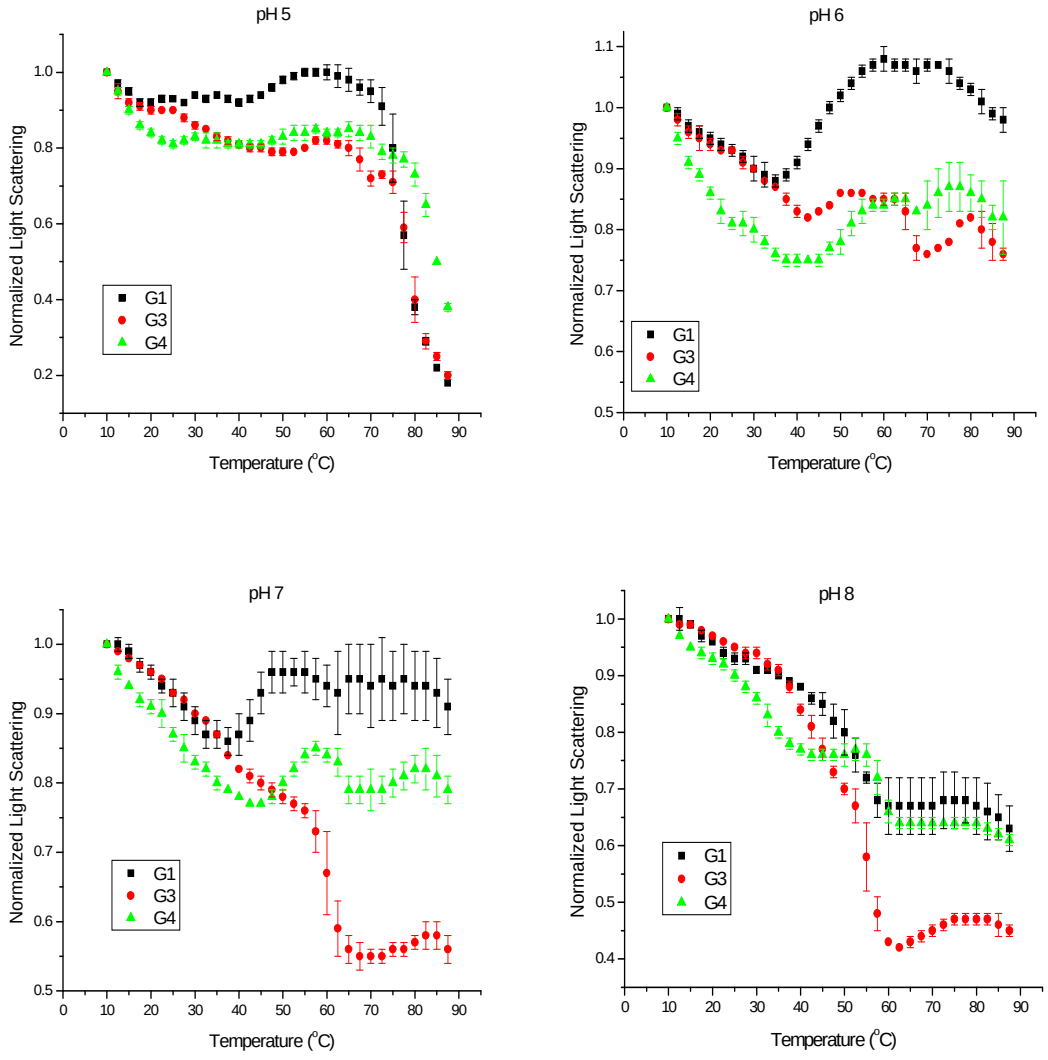


Figure 4.4.Static light scattering as a function of pH and temperature. Rotavirus solutions at pH 5–8 were heated from 10 to 85 °C, and the scattering intensity was monitored at 295 nm (*n*=3).

4.3.3. Dynamic Light Scattering

Alterations in the quaternary structure of rotavirus were studied by monitoring its DLS measured hydrodynamic diameters as a function of pH and temperature (Fig 4.5). The hydrodynamic diameters were calculated from diffusion coefficients using the Stokes-Einstein equation and the method of cumulants[23]. The onset of transition temperatures occur at ~ 45 °C for all three strains at pH 5. The rate of increase in the hydrodynamic diameter, however, appears to occur more rapidly for G1 compared to the other two strains. At pH 6, all three strains exhibit biphasic transitions in size with early onsets at ~ 45, 47.5, and 50 °C for G1, G3, and G4, respectively. The second transition occurs at ~ 67.5 °C for G1 and G3 and 72.5 °C for the G4 strain. The G1 and G4 strains also exhibit biphasic behavior at pH 7 in which early transition onsets are observed at ~ 50 and 55 °C followed by a second transition at ~ 77.5 °C. The G3 strain shows an onset transition at ~ 65 °C at pH 7. At pH 8, this biphasic behavior continues in which early onset transitions occur at ~ 47.5, 57.5, and 55 °C followed by later transitions at 72.5, 75, and 77.5 °C for the G1, G3, and G4 strains, respectively. Overall, the onset of size increase is similar for a given strain over the pH range examined with the exception of G3 which exhibits a detectably higher onset transition temperature at pH 7. The G1 strain exhibits higher aggregation rates over the pH range examined as manifested by more rapidly increasing rates of its average size.

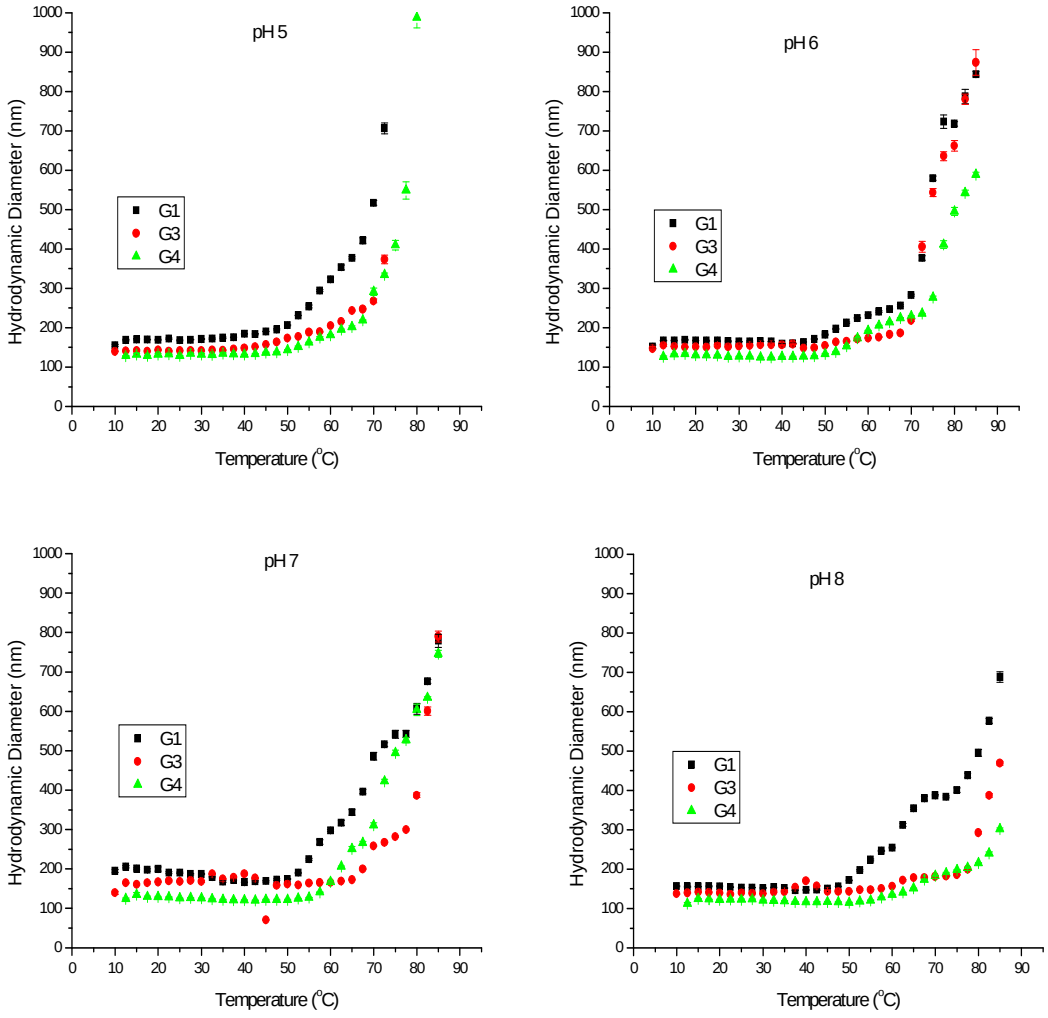


Figure 4.5.Analysis of the size of rotavirus strains as a function of pH and temperature. Data analysis was performed using the method of cumulants. The data presented are an average of five consecutive scans.

The hydrodynamic diameters based on the method of cumulants are larger than the typical rotavirus size of ~ 70-100 nm (Fig 4.5). This discrepancy can be further examined using a non-negatively constrained least square (NNLS) analysis of the DLS data. This algorithm is often used when studying heterogeneous samples with multimodal size distributions in which the autocorrelation function is deconvoluted to yield size distributions[22]. For example, the NNLS analysis of the higher order species (< 1%) at pH 7 and 10 °C reveals particles approximately fifteen, four, and three times larger than the monomeric species for G1, G3, and G4, respectively (data not shown). The presence of the higher molecular weight species clearly accounts for the elevated sizes observed by the cumulants method and also for the differences observed in size among the three strains at 10 °C. In addition, the NNLS analyses of the three strains at 10 °C reveal that more than 99% of the viral particles possess a size of ~ 70-120 nm, consistent with the expected rotavirus size. The variations in size (i.e., the wide 50 nm range) observed when using the NNLS method are presumably due to heterogeneity of the samples (polydispersity index values > 0.2).

4.3.4. Activity Measurements

The activity of the individual strains at pH / temperature combinations corresponding to the distinct physical states identified in the EPD (see below) was examined using a Fluorescence Focus assay (Table 4.1). At 4 °C, the G1 strain exhibits the highest activity at both pH 5 and 7 followed by the G3 and G4 strains in

which the latter appears to be the most prone to pH inactivation at refrigeration temperatures. At pH 5, all of the strains retain their activity when the temperature is increased to 40 °C. At pH 7, however, the activity is lost suggesting a damaging effect of higher pH values under thermally induced degradation conditions. Finally, at 70 °C all three strains lose their activity regardless of pH.

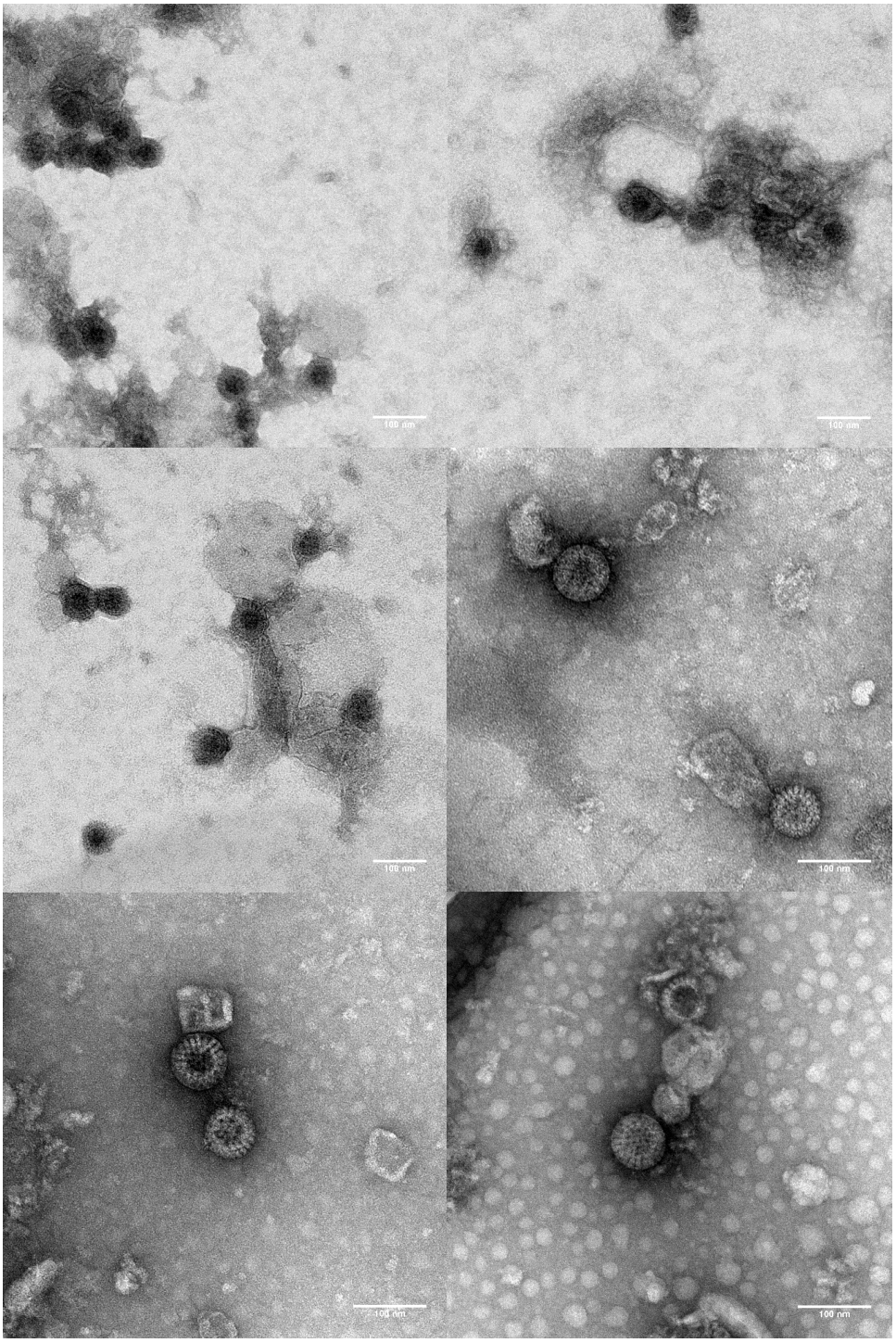
Strain	G1		G3		G4	
Temp / pH	pH 5	pH 7	pH 5	pH 7	pH 5	pH 7
4 °C	7.19 ± 0.015	7.44 ± 0.012	7.16 ± 0.052	7.14 ± 0.108	6.81 ± 0.107	7.13 ± 0.003
40 °C	7.09 ± 0.046	bdl*	7.28 ± 0.081	bdl	6.66 ± 0.051	bdl
70 °C	bdl	bdl	bdl	bdl	bdl	bdl

Table 4.1. Rotavirus strain activities upon exposure to different pH / temperature conditions. Activity is measured using a fluorescence focus assay (see methods) and is in log FFU/ml.

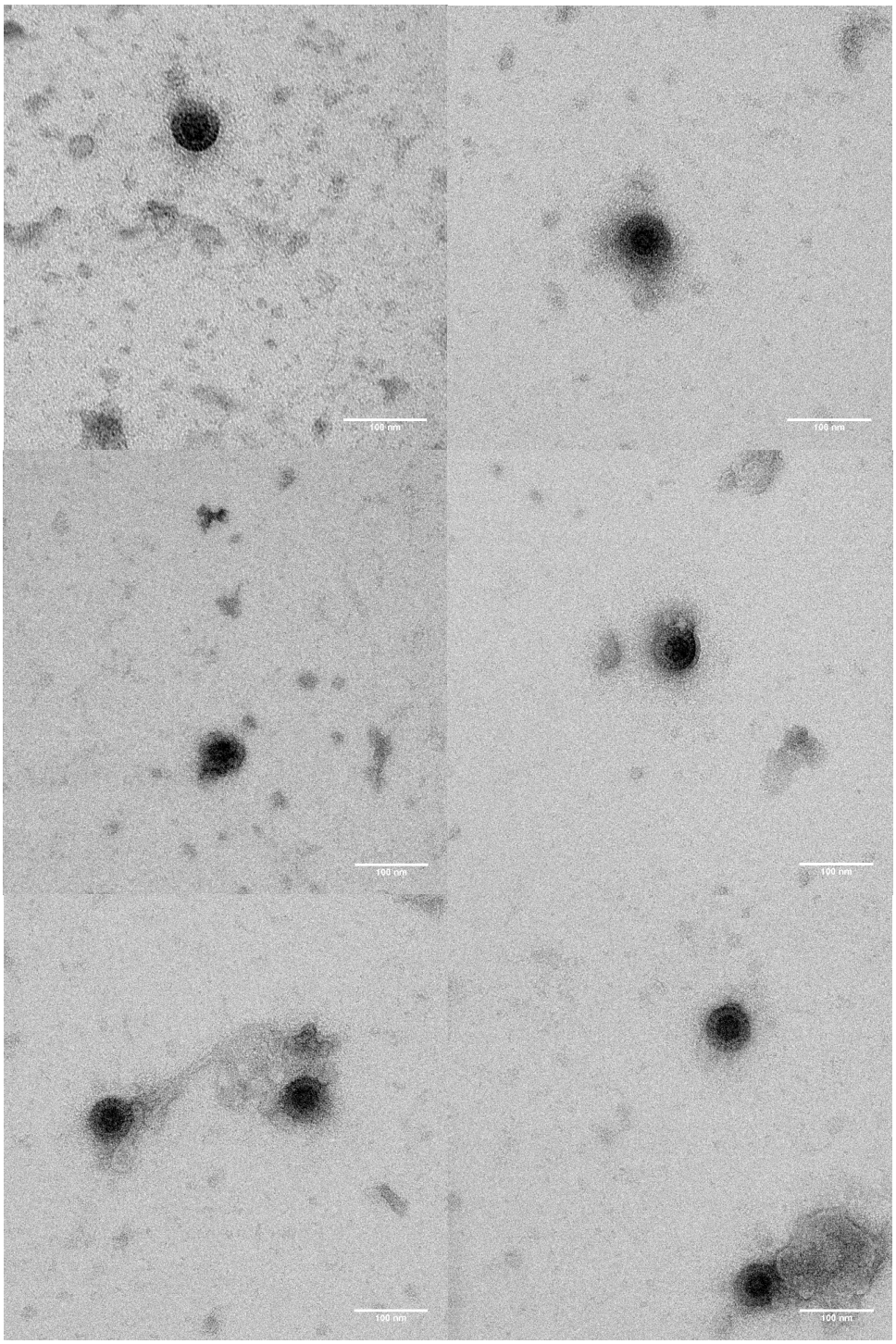
4.3.5. Transmission Electron Microscopy

As an additional tool to enhance interpretation of the spectroscopic results, EM studies were employed as a function of pH (5 and 7) and temperature (4, 40, 70 °C) to cover a range of distinct physical states identified in the EPDs of the three strains (Fig 4.6). The strains are used interchangeably here since the EM analysis did not recognize any detectable differences among the three strains at different pH / temperatures combinations, in agreement with the EPD results (see below). At 4 °C, both pH 5 and 7 exhibit dominantly infectious, triple layer particles along with a small population of particles with disrupted outer layer capsids (Fig 4.6a & 4.6d, respectively). At pH 5 and 40 °C, triple layer particles still appear to be dominant (Fig 4.6e). At pH 7 and 40 °C, however, most particles appear to have disrupted outer layers (Fig 4.6b) which can explain the loss of activity at this pH / temperature combination. At pH 7 and 70 °C (Fig 4.6c), EM images show primarily clusters of particles and aggregated sheets in agreement with the scattering data (the data at pH 5 shows similar behavior at elevated temperatures and is not shown).

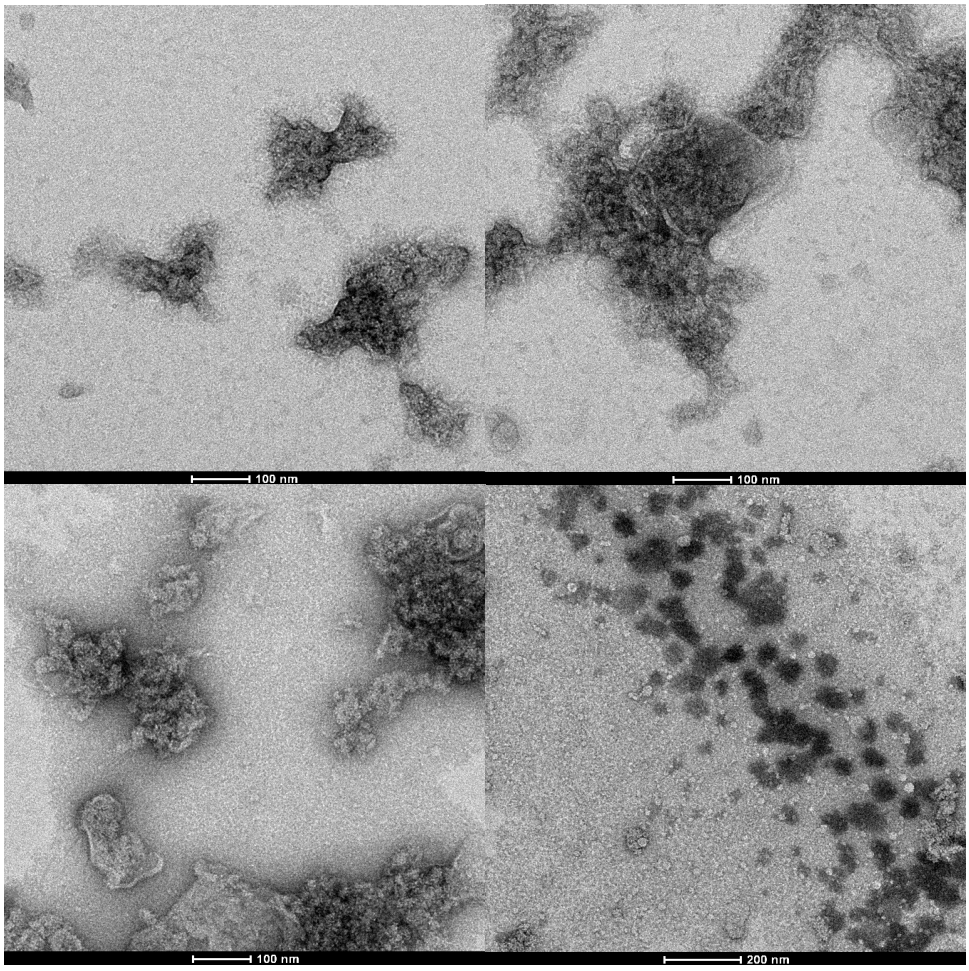
a) pH 7 / 4 °C



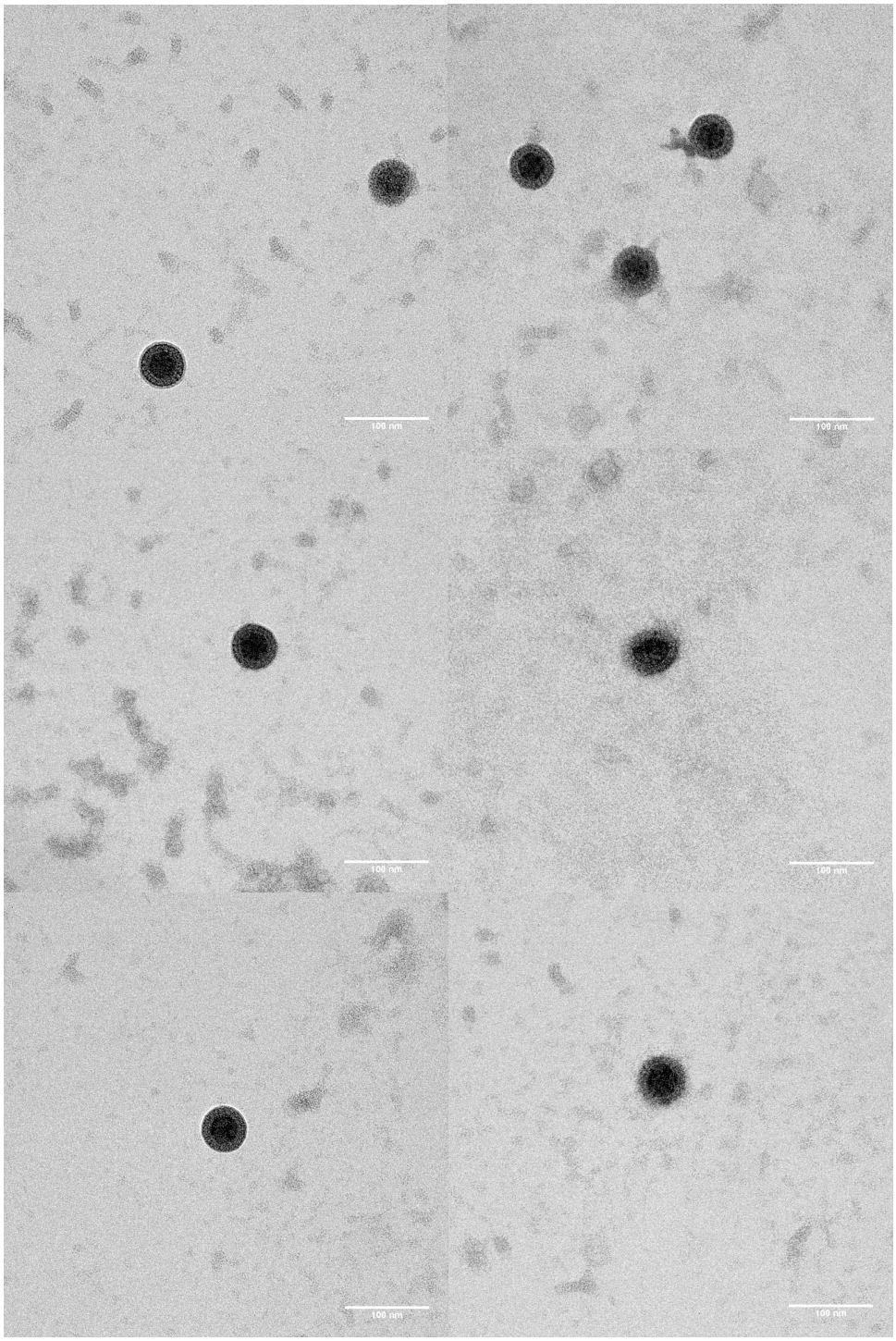
b) pH 7 / 40 °C



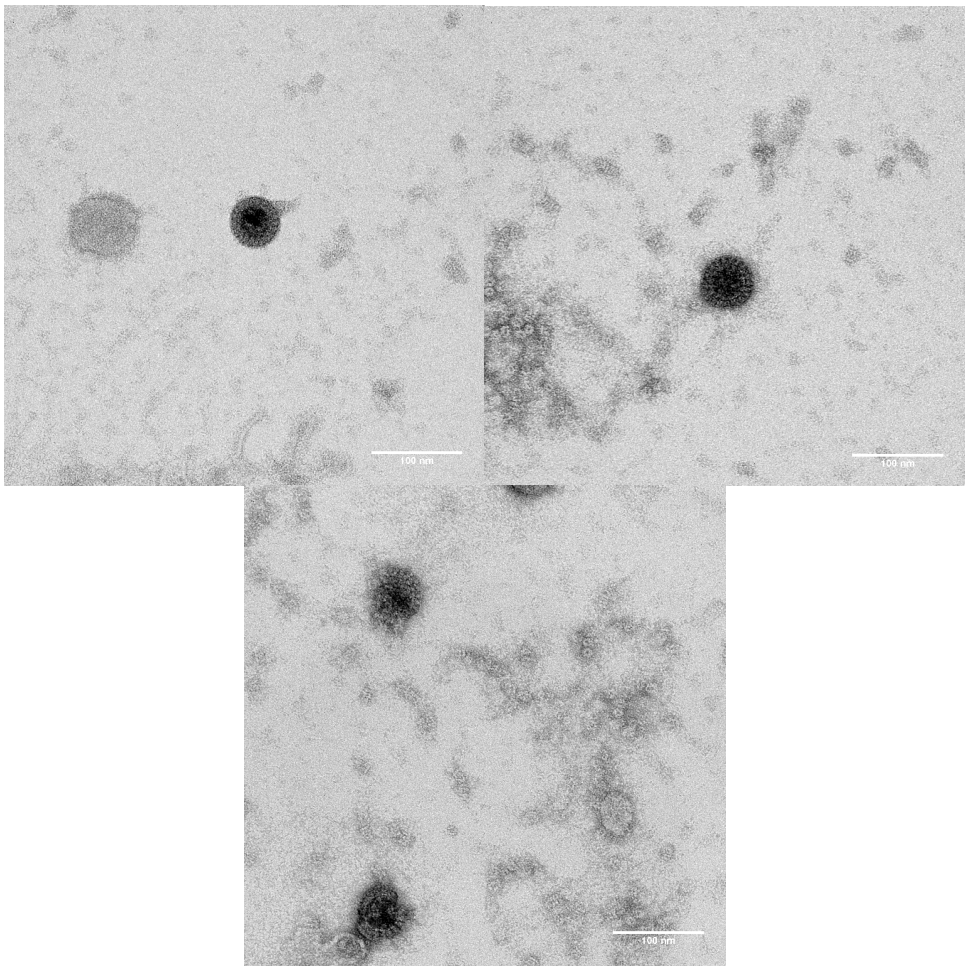
c) pH 7 / 70 °C



d) pH 5 / 4 °C



pH 5 / 4 °C cont…



e) pH 5 / 40 °C

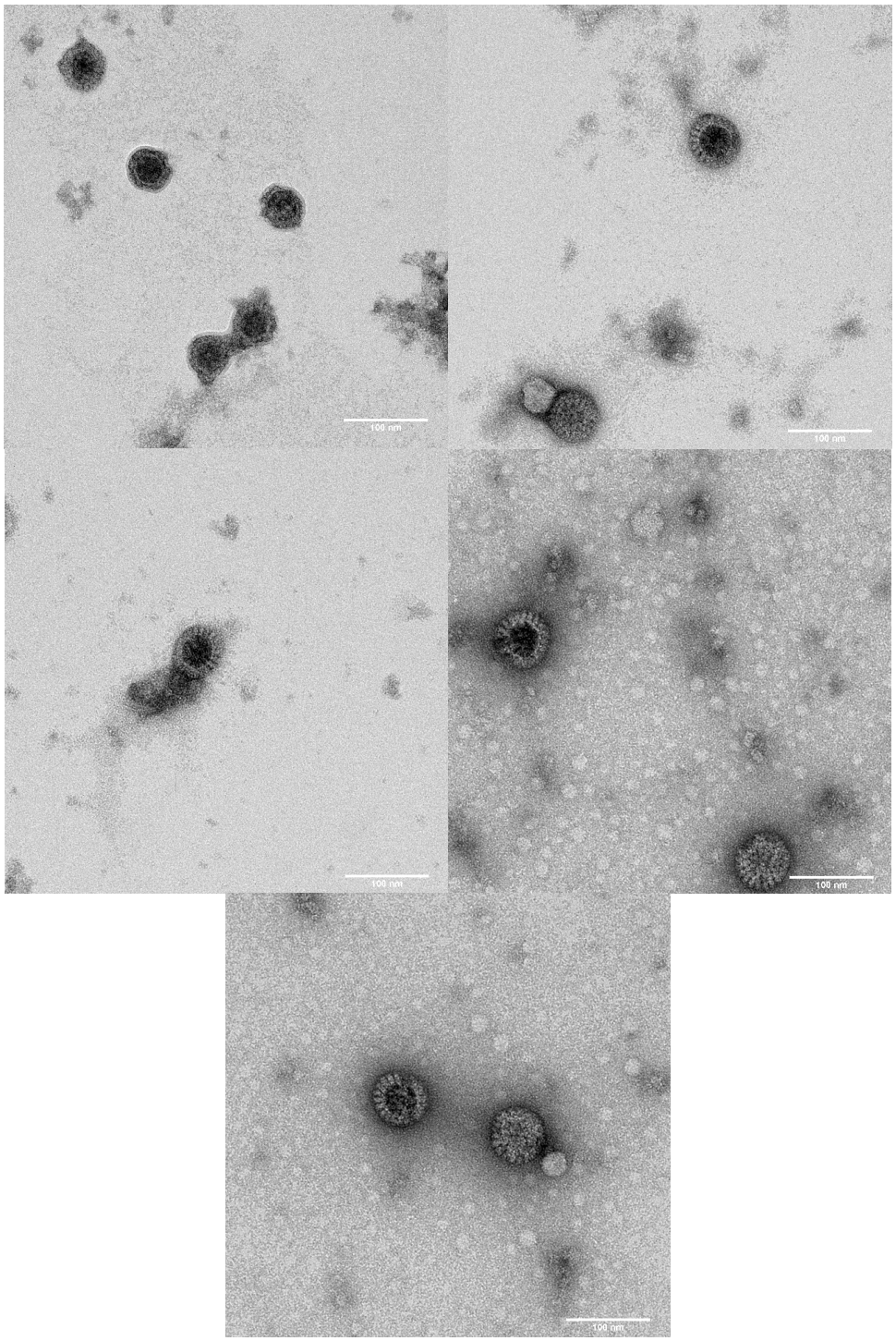


Figure 4.6.Electron microscopy images of rotavirus as a function of pH and temperature.

4.3.6. Empirical Phase Diagrams (EPDs)

The spectroscopic data for each strain were integrated into an empirical phase diagram to provide a more global picture of the virus response to temperature and pH (Fig 4.7). In the EPD, regions of continuous color define uniform structural (non-thermodynamic) states while abrupt changes in color identify alterations in the physical state of the rotavirus over the conditions examined. The properties of the rotavirus strain within each phase can be established by reference to the individual measurements. We emphasize that the EPDs should not be confused with classic thermodynamic phase diagrams in which equilibrium exists between different phases. At least partially irreversible aggregation as seen here prevents any such analysis. The EPD of each rotavirus strain clearly shows a number of distinct physical states over the pH and temperature range examined.

G1. The green P1 phase across the pH range of 5 to 8 below ~ 40 °C is the region of maximum viral integrity with the virus particles dominantly in their intact triple layer conformations. The purple and brown phases labeled P2 at pH 5 and 6 between 60 and 70 °C respectively contain some aggregates as manifested by the static and dynamic light scattering data at pH 5 and the dynamic light scattering results at pH 6. This phase at pH 6 also contains altered protein tertiary structure as indicated by shifts of the tryptophan peak position. The blue / purple region labeled P3 over pH 5 and 6 above 70 °C contain further alterations of protein secondary structure compared to the P2 phase. The brown region labeled P4 contains aggregated particles. The pinkish P5 region contains particles with perturbed secondary and tertiary structures

along with more extensive aggregation. The brown region labeled P6 contains particles of altered secondary structure according to the CD thermal melt studies.

G3. The brownish P1 phase across the pH range of 5 to 8 below ~ 40 °C is again the region of maximum stability. The bright green P2 phase is a region of minimum stability with perturbed secondary and tertiary structures and the presence of extensive aggregates. The intermediate P3 phase contains particles with perturbed tertiary structures. The light blue region labeled P4 at pH 5 contains extensive aggregates.

G4. The green P1 phase across the pH range of 5 to 8 below ~ 40 °C is the region of maximum stability with the virus particles dominantly in their intact triple layer form. The dark green phase labeled P2 at pH 5 contains some aggregates as manifested by the DLS data. The brown P3 phase contains particles with altered secondary structure and some aggregates. The purple and red regions at pH 5 and 6 above ~ 70 °C contain more extensive aggregates with perturbed secondary and tertiary structures. The brownish region labeled P5 at pH 7 and 8 contains some aggregates as well as perturbed tertiary structure. The pinkish region labeled P6 contains protein with altered secondary and tertiary structure as well as extensive aggregates.

Overall, the EPDs of the three strains show very similar behavior in terms of the strains' global responses to variations in pH and temperature as suggested by the similar physical states identified. The EPDs of the G1 and G4 strains manifest sharper and clearer intermediate states, absent in the EPD of the G3 strain. Examination of the EPD of the three strains reveals distinct conformational and aggregation phases

across the pH and temperature range examined. The spectroscopic data suggest that all three strains are stable in the pH range of 5 to 8 below ~ 40 °C (the P1 phase in all three EPDs). This is supported by EM image analyses at low temperatures which reveal dominantly intact, triple layer infectious particles. All three strains show a region of minimum stability with perturbed protein secondary and tertiary structures along with the presence of aggregates in the pH range of 5 to 8 at elevated temperatures above 55 to 60 °C (P3&P5 phases in the G1 EPD, the P2&P4 phases in G3 EPD, and the P4&P6 phases in G4 EPD). The EM images further exhibit extensive aggregation at the more elevated temperatures of 70 °C. The onset temperatures at which the structural alterations occur are lowered in a stepwise fashion as a function of increasing pH, suggesting reduced virus thermal stability at increasing pH. This is supported by the stabilizing effect of lower pH on virus activity upon increasing temperatures from 4 to 40 °C (Table 4.1). At more moderate temperatures between 40 and 60 °C, the spectroscopic analyses show subtle and gradual alterations of the virus's secondary and tertiary structures along with formation of micro-aggregates / soluble oligomers detected by the static light scattering data (the P4 and P6 phases in G1 EPD, P3 in G3 EPD, and P3 and P5 in G4 EPD). The EM analyses at pH 7 and 40 °C further show intact, triple layer rotavirus particles with disrupted capsids.

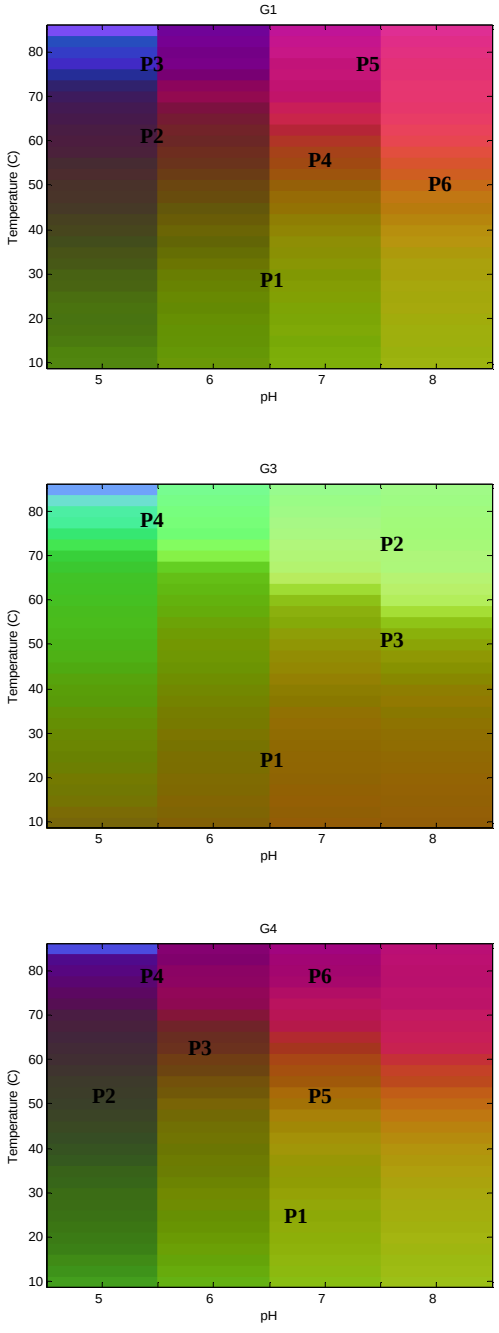


Figure 4.7. Temperature / pH empirical phase diagrams of rotavirus strains G1, G3, and G4 based on intrinsic fluorescence, CD thermal melts, and static and dynamic light scattering data. A detailed description of the origin of the individual phases identified are found in the text. Note that the color itself is arbitrary.

4.3.7. Effect of Formulation Buffer

Studies by Peterson *et al.* on the individual strains (G1, G2, G3, G4, P1) of the pentavalent human-bovine reassortant rotavirus vaccine, RotaTeq, suggest that the formulation buffer has major impact on virus activity. They suggest that increasing concentrations of citrate cause disaggregation of the virus particles presumably due to a carboxylic-acid mediated interaction with the surface charges of the virus [23]. Dynamic light scattering studies showed a decrease in virus size as a function of increasing citrate concentration. Figure 4.8 is the initial portion of the DLS thermal curves over the temperature range of 10-30 °C in the absence of any major structural alterations as suggested by the collective spectroscopic data. Detectably smaller particle sizes are observed at pH 5 compared to pH 7 for the G1 and G3 strains in agreement with the studies of Peterson *et al*. This can be explained as due to the presence of citrate ions at three fold higher concentrations at pH 5 compared to pH 7. Although G4 exhibits a smaller size compared to the other two strains, it does not show any pH dependent citrate-mediated effect.

The two surface proteins, VP4 and VP7 are presumably involved in electrostatic interactions with the buffer ions. Estimation of the charge on the human G1 strain's VP7 protein (PDB # ABI63532) shows a net negative charge of ~ -0.6 and -8.1 at pH values 5 and 7, respectively. Similar analysis of the VP4 protein of the human G1 Wa strain[24] finds VP4 net charges of ~ 20.6 and -1.2 at pH values 5 and 7, respectively. The net negative charges on both proteins at pH 7 are repulsive for electrostatic interactions with the buffer anions in agreement with the less pronounced

citrate mediated mechanism observed. At pH 5, however, the VP4 protein possesses a highly net positive charge, suggesting its dominant role in the proposed electrostatic interactions. Such charge behavior, however, cannot explain the much weaker citrate effects on G4 since it possesses a very similar surface net charge as the other two strains (data not shown) suggesting involvement of other factors in this process.

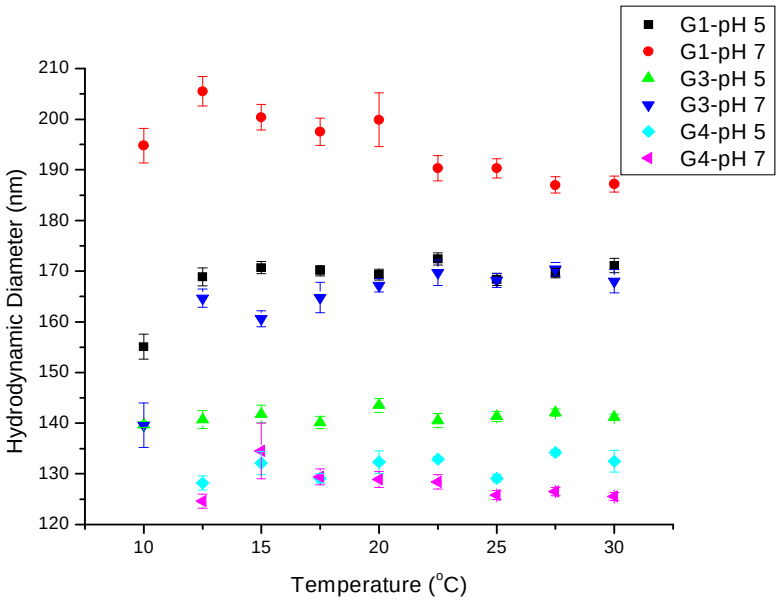


Figure 4.8.Analysis of the size of rotavirus strains as a function of pH and temperature. Data analysis was performed using the method of cumulants. The data presented are an average of five consecutive scans.

4.4. Discussion

Viruses, the main macromolecular ingredients of many vaccine formulations are thermolabile and are often destroyed either due to excessive heat (e.g., measles, polio) or freezing (e.g., tetanus, diphtheria). Therefore, development of thermostable vaccine formulations resistant to temperature excursions has major benefits. These include reducing vaccine wastage, increasing vaccine effectiveness, reducing dependency on the cold chains and related equipments, and enabling vaccine delivery to remote populations[25].

The typical vaccine shelf life of 24 months typically requires long and expensive stability studies to determine the effect of different formulation parameters on vaccine stability. Accelerated stability studies are therefore routinely employed to screen formulations for their potential to meet or exceed stability goals. Herein, we have employed such an accelerated approach using a variety of spectroscopic techniques (i.e., circular dichroism, intrinsic fluorescence, and static and dynamic light scattering) to characterize the thermal stability of three live attenuated human-bovine reassortant rotavirus strains as potential vaccine candidates. Liquid formulation of live-attenuated viruses, as is the case here, is often considered to be the most desirable type of vaccine formulation. This is due to both the ability to stimulate long lasting humoral and cellular immune responses[16] as well as reducing complicating factors involved with lyophilized vaccine formulations [26-28].

Studies by Meng *et al.* on the human serotype 1 (FH432) and simian agent 11 (SA11) rotavirus strains show a destabilizing effect of pH and temperature variations

on virus activity[29]. At refrigeration temperatures, the activity of both strains is retained in the pH range of 3-10 while higher or lower values result in partial or complete loss of virus activity. These results are in agreement with the accelerated studies at low temperatures (*i.e.*, 10 °C) in which the virus activity is retained at both pH values of 5 and 7. Studies of Meng *et al.* showed that both strains lost activity upon very short incubations (i.e., a few minutes) at temperatures at or above 50 °C consistent with our accelerated stability studies at elevated temperatures (examined at 70 °C) in which the activity of all strains is completely lost regardless of the pH. The extensive secondary and tertiary structural alterations as well as formation of aggregates observed by both the spectroscopic and EM studies are presumably responsible for the loss of activity observed. Variations in pH have a major impact on virus activity when moderately raising the temperatures (*i.e.*, to 40 °C) in which all of the strains retain activity at pH 5 but not at pH 7. Considering that the two antigenic surface proteins, VP4 and VP7 are responsible for the attachment and entry of the virus into the cells[1,30], the increasing number of virus particles with partially disrupted outer layers in the EM images can explain such activity losses.

The two antigenic proteins, VP7 and VP4 are believed to dominate the differences in the observed spectroscopic signals among the three strains. This is presumably due to both their spatial arrangement on the most outer layer of the virus and their abundance (2^{nd} and 3^{rd} respectively) in the rotavirus particle[34]. Although the intermediate layer protein, VP6 is the most abundant protein, it is part of the bovine core shared among all the three reassortants. The genome comprises about 15% of the

virus's total mass and as observed in Figure 1, does not contribute detectably to the signals observed (no significant absorbance at ~ 260 nm).

It can therefore be proposed that the structural alterations observed herein should directly impact virus activity. At pH 5, the G4 strain with the least amount of secondary structure shows reduced activity compared to the other two strains. At pH 7, G1 shows the highest activity in agreement with its more extensive secondary structure content. At this pH, the G3 and G4 strains possess similar amounts of secondary structure and activity values. Activity data suggest that G3 is the least sensitive strain to pH variations in agreement with its rather pH insensitive secondary structural alterations compared to the G1 and G4 strains (Figure 9).

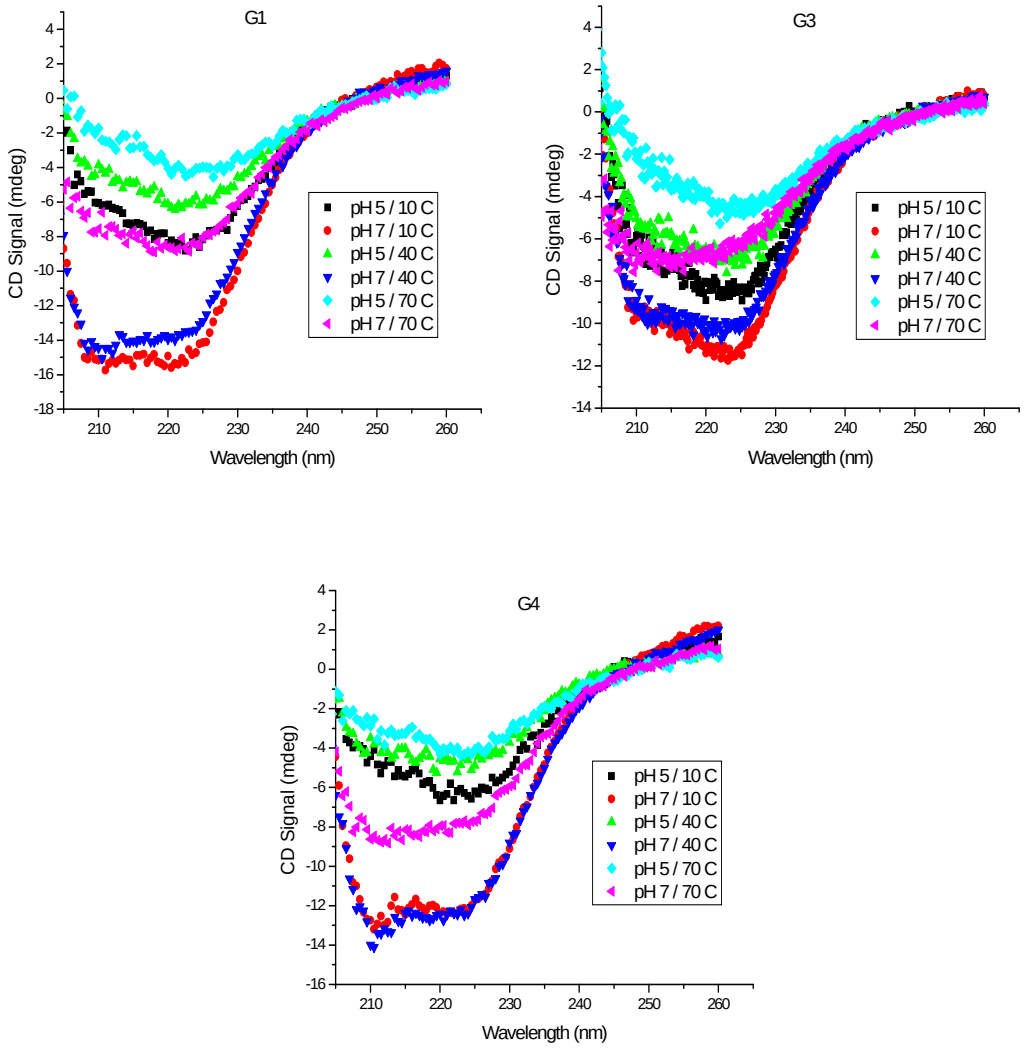


Figure 4.9.CD spectra of rotavirus strains at pH 5 and 7. CD spectra were recorded at 10, 40, and 70° C from 190 to 260 nm at each of the indicated pH values (*n* =3).

The effect of tertiary structure alterations on virus activity appear to be less significant with subtle alterations in the low to intermediate temperature range exhibited by the small red shifts in tryptophan peak position (Figure 3).

Both of the currently available rotavirus vaccines are administered orally, exposing the virus to the acidic pH of the stomach (i.e., pH 2 to 6). An optimal rotavirus vaccine formulation should therefore be able to protect the virus against such harsh acidic environments. The EPD analysis of the three rotavirus strains shows induced thermal stability as a function of decreasing pH. Moreover, the stabilizing effect of lowering pH on virus activity is observed in which increasing temperatures from 4 to 40 °C resulted in loss of virus activity at pH 7 but not at pH 5. These observations encourage formulations at lower pH values suitable for oral administrations. Dialysis of the virus bulk samples into citrate phosphate buffer below pH 5, however, resulted in aggregation of viral particles suggesting the need for a stabilizing agent in the vaccine formulation. The commercially available RotaTeq formulation using similar citrate phosphate buffers contains sucrose as a stabilizing agent.

Analyses of the individual spectroscopic techniques reveal detectable differences in the thermal stability of the three strains. Analyzing the EPDs, however, suggest that the global response of the strains to variations in pH and temperature is similar as manifested by the similar apparent pseudo-phase boundaries in the EPDs. The EPD approach therefore suggests similar stabilization strategies for all the strains and the feasibility of inclusion of multiple strains in a single rotavirus vaccine

formulation as has already been demonstrated in the commercial vaccines. The EPD approach has been successfully employed in the past for formulation of vaccines due to its ability to identify regions of marginal stability which can serve as a basis for the design and development of high throughput excipient screening assays[15,32,33]. The EPD approach has been further used for the biophysical characterization of a variety of other macromolecular systems and as a tool for the stabilization of such complexes[21, 34-38].

4.5. Conclusion

The complexity associated with viruses and rotavirus in particular, may appear to decrease the utility of biophysical measurements in the analyses of such complicated systems. Although the information obtained are complex and interpretations at the molecular level is difficult, the results can still be used to detect stability changes due to the distinctive and highly reproducible signals produced. Rotavirus exhibits well defined size and shape by dynamic light scattering. In addition, spectroscopic techniques such as CD and fluorescence generated reproducible results across the pH and temperature range examined. Thus, these studies suggest that the EPD approach is potentially an effective tool in the early development of thermostable vaccine formulations based on the observation that alterations in selected physical parameters reflect losses in virus activity, a key parameter in design and development of such vaccines.

4.6. Bibliography

(1) Arias CF, Isa P, Guerrero CA, Mendez E, Zarate S, Lopez T, Espinosa R, Romero P, Lopez S. Molecular biology of rotavirus cell entry. Arch. Medical Res. (2002), 33, 356-361.

(2) Wilhelmi I, Roman E, Sanchez-Fauquier A. Viruses causing gastroenteritis. Clinical microbiology and infection (2003), 9, 247-62.

(3) Glass RI, Parashar UD, Bresee JS, Turcios R, Fischer TK, Widdowson MA, Jiang B, Gentsch JR. Rotavirus vaccines: Current prospects and future challenges. Lancet (2006), 368, 323-332.

(4) Roberts L. Rotavirus vaccines ' second chance. Science (Washington, DC, United States) (2004), 305, 1890-1893.

(5) Bernstein DI. Rotavirus overview. The Pediatric infectious disease Journal (2009), 28, S50-3.

(6) Matson DO. RotaShield: the ill-fated rhesus-human reassortant rotavirus vaccine. Pediatric annals. (2006), 35, 44-50.

(7) Bines J. Intussusception and rotavirus vaccines. Vaccine (2006), 24, 3772-3776.

(8) Bernstein DI. Live attenuated human rotavirus vaccine, Rotarix. Seminars in pediatric infectious diseases (2006), 17, 188-94.

(9) Offit PA, Clark HF. RotaTeq: A pentavalent bovine-human reassortant rotavirus vaccine. Pediatric annals (2006), 35, 29-34.

(10) Sheridan C. The business of making vaccines. Nat. Biotech. (2005), 23, 1359-1366.

(11) Zaffran M. Vaccine transport and storage: environmental challenges. Developments in biological standardization (1996), 87, 9-17.

(12) Peetermans J. Factors affecting the stability of viral vaccines. Developments in biological standardization (1996), 87, 97-101.

(13) Milstien JA, Zaffran M. Thermostability of vaccines. World Health Organization: Geneva, Switzerland (1998).

(14) Newman JF, Tirrell S, Ullman C, Piatti PG, Brown F. Stabilizing oral poliovaccine at high ambient temperatures. Developments in biological standardization (1996), 87, 103-11.

(15) Ausar SF, Espina M, Brock J, Thyagarayapuran N, Repetto R, Khandke L, Middaugh CR. High-throughput screening of stabilizers for respiratory syncytial virus: identification of stabilizers and their effects on the conformational thermostability of viral particles. Human Vaccines (2007), 3, 94-103.

(16) Burke CJ, Hsu TA, Volkin DB. Formulation, stability, and delivery of live attenuated vaccines for human use. Critical reviews in therapeutic drug carrier systems (1999), 16, 1-83.

(17) Brandau DT, Jones LS, Wiethoff CM, Rexroad J, Middaugh CR. Thermal stability of vaccines. J. Pharm. Sci. (2003), 92, 218-231.

(18) Hoshino Y, Kapikian AZ. Classification of rotavirus VP4 and VP7 serotypes. Arch. Virol (Suppl) (1996), 12, 99-111.

(19) Gentsch JR, Woods PA, Ramachandran M, Das BK, Leite JP, Alfieri A, Kumar R, Bhan MK, Glass RI. Review of G and P typing results from a global collection of rotavirus strains: implications for vaccine development. The Journal of infectious diseases (1996), 174, 1, S30-6.

(20) Brown JC, Pusey PN. Photon Correlation Study of Polydisperse Samples of Polystyrene in Cyclohexane. J. Chem. Phys. (1975), 62, 1136-44.

(21) Kueltzo LA, Ersoy B, Ralston JP, Middaugh CR. Derivative absorbance spectroscopy and protein phase diagrams as tools for comprehensive protein characterization: a bGCSF case study. J. Pharm. Sci. (2003), 92, 1805-1820.

(22) Morrison ID, Grabowski EF, Herb CA. Improved techniques for particle size determination by quasi-elastic light scattering. Langmuir (1985), 1, 496-501.

(23) Peterson SE; Wang S, Ranheim T, Owen KE. Citrate-mediated disaggregation of rotavirus particles in RotaTeq vaccine. Antiviral Res. (2006), 69, 107-115.

(24) Lopez S, Lopez I, Romero P, Mendez E, Soberon X, Arias CF. Rotavirus YM gene 4: analysis of its deduced amino acid sequence and prediction of the secondary structure of the VP4 protein. J. Virol. (1991), 65, 3738-45.

(25) Chen D, Kristensen D. Opportunities and challenges of developing thermostable vaccines. Exp. Rev. Vaccines (2009), 8, 547-557.

(26) Gupta RK, Rost BE, Relyveld E, Siber GR. Adjuvant properties of aluminum and calcium compounds. Pharm. Biotechnol. (1995), 6, 229-48.

(27) Newman MJ, Powell MF. Immunological and formulation design considerations for subunit vaccines. Pharm. Biotechnol. (1995), 6, 1-42.

(28) Zapata MI, Feldkamp JR, Peck GE, White JL, Hem SL. Mechanism of freeze-thaw instability of aluminum hydroxycarbonate and magnesium hydroxide gels. J. Pharm. Sci. (1984), 73, 3-8.

(29) Meng Z, Birch C, Heath R, Gust I. Physicochemical stability and inactivation of human and simian rotaviruses. Applied and Environmental Microbiology (1987), 53, 727-30.

(30) Lopez S, Arias CF. Multistep entry of rotavirus into cells: A Versaillesque dance. Trends. Microbiol. (2004), 12, 271-278.

(31) Fields BN, Knipe DM, Howley PM. Editors. Fields Virology: 4th ed, volume 2, Lippincott Williams&Wilkins (2001), 1445 pp.

(32) Peek LJ, Brandau DT, Jones LS, Joshi SB, Middaugh CR. A systematic approach to stabilizing EBA-175 RII-NG for use as a malaria vaccine. Vaccine (2006), 24, 5839-5851.

(33) Kissmann JM, Ausar SF, Rudolph A, Braun C, Cape SP, Sievers RE, Federspiel MJ, Joshi SB, Middaugh CR. Stabilization of measles virus for vaccine formulation. Human vaccines (2008), 4, 350-9.

(34) Salnikova MS, Joshi SB, Rytting JH, Warny M, Middaugh CR. Preformulation studies of Clostridium difficile toxoids A and B. J. Pharm. Sci. (2008), 97, 4194-4207.

(35) Ausar SF, Rexroad J, Frolov VG, Look JL, Konar N, Middaugh CR. Analysis of the thermal and pH stability of human respiratory syncytial virus. Mol Pharm. (2005), 2, 491-499.

(36) Ausar SF, Foubert TR, Hudson MH, Vedvick TS, Middaugh CR. Conformational stability and disassembly of Norwalk virus-like particles. Effect of pH and temperature. J. Biol. Chem. (2006), 281, 19478-19488.

(37) Kissmann JM, Ausar SF, Foubert TR, Brock J, Switzer MH, Detzi EJ, Vedvick TS, Middaugh CR. Physical stabilization of Norwalk virus-like particles. J. Pharm. Sci. (2008), 97, 4208-4218.

(38) Ruponen M, Braun CS, Middaugh CR. Biophysical characterization of polymeric and liposomal gene delivery systems using empirical phase diagrams. J. Pharm. Sci. (2006), 95, 2101-2114.

Chapter 5

Temperature Dependent 2nd Derivative Absorbance Spectroscopy of Aromatic Amino Acids as a Probe of Protein Dynamics

5.1. Introduction

In the past, proteins were often perceived as relatively static structures based on their X-ray analysis. It is now well accepted, however, that proteins are highly dynamic molecules exhibiting a wide variety of different types of motions. Protein dynamic behavior plays a crucial role in their structural stability, folding, and biological functions[1] and such motions are strongly coupled to solvent mobility[2].

A variety of biophysical methods have been used to measure various aspects of protein structure from which dynamic behavior can be inferred. These techniques include isotope (H/D) exchange monitored by infrared spectroscopy, mass spectrometry, and nuclear magnetic resonance (NMR) spectroscopy as well as other methods such as red edge shift spectroscopy, solute fluorescence quenching, ultrasonic spectroscopy, pressure perturbation calorimetry and molecular dynamics simulations. Here, for the first time, we examine the utility of the temperature dependent 2nd derivative absorbance signals of protein aromatic amino acids as a qualitative probe of intramolecular dynamics.

The application of UV absorption spectroscopy to proteins was initiated more than half a century ago at relatively low resolution[3]. Proteins display a broad peak in the 250-300 nm region of the ultraviolet spectrum composed of multiple overlapping bands from the aromatic residues phenylalanine, tyrosine, and tryptophan due primarily to $\pi \rightarrow \pi^*$ transitions involving the electrons of their aromatic rings[4]. Due to the extensive overlap of these absorption peaks, the utility of UV absorption spectroscopy in protein analysis was, however, quite limited. Recent advances in

instrumentation, in particular the availability of diode array detectors and computer based derivative absorbance[5], it is now possible to resolve the absorption bands of each of the three aromatic residues into multiple peaks with a resolution of approximately 0.01 nm. This provides a very sensitive tool with which to probe protein conformational alterations.

Upon protein structural changes, the polarity of the microenvironment surrounding the aromatic side chains and their level of exposure to the surrounding solvent can be detectably altered. By monitoring the individual shifts of the derivative peak position of these residues, fairly detailed information can be obtained regarding the conformational alterations of proteins as a function of variety of conditions such as pH, temperature, ionic strength, etc. Using high resolution 2nd derivative UV spectroscopy, shifts in peak positions as small as 0.5 nm and as large as 6 nm have been monitored and correlated with protein structural alterations[5-8].

The temperature dependence of 2nd derivative peak shifts of the aromatic residues has been widely employed as a tool to probe protein thermal unfolding. In this case, shifts to lower wavelength (i.e., blue shifts) as a function of increasing temperature are often indicative of enhanced exposure of aromatic residues to solvent. In contrast, peak shifts in pre-transition regions below detectable unfolding events have received no attention. Here, we demonstrate that plots of the temperature dependent 2nd derivative peak positions of the aromatic residues have measurable slopes in protein pre-transition regions. We have investigated the nature of these temperature dependent spectral alterations as a function of solvent physical properties

and provide evidence that such alterations can be used qualitatively to probe protein dynamics.

It is worth mentioning that we have previously attempted related analyses by employing the spectral shifts induced by cation (Na^+, Li^+, Cs^+)-π interactions as a function of increasing cation concentration[4]. By analogy to solute-based fluorescence quenching of proteins, it was generally found that small cations were more effective at reducing spectral shifts due to their ability to diffuse through a protein's matrix and make contact with the aromatic side chains. In some cases, however, specific interactions between the cations (and perhaps accompanying counter anions, Cl^-) made interpretation of the data difficult in terms of dynamic effects. Thus, a method that does not require the presence of potentially perturbing solutes, but rather involves a simple intrinsic effect such as temperature seems desirable.

5.2. Materials and Methods

5.2.1. Materials

Ribonuclease T1 purified from Aspergillus oryzae was obtained from Epicentre (Madison, Wisconsin). *N*-acetyl-L-phenylalanine ethyl ester, *N*-acetyl-L-tryptophan ethyl ester, *N*-acetyl-L-tyrosine ethyl ester, melittin, Substance P, Leucine Enkephaline, and all other proteins were obtained from Sigma (Saint Louis, Missouri).

5.2.2. Sample Preparation

To investigate the intrinsic effect of pH and temperature on UV spectral peaks, model amino acids *N*-acetyl-L-phenylalanine ethyl ester, *N*-acetyl-L-tyrosine ethyl ester and *N*-acetyl-L-tryptophan ethyl ester were dissolved in 20 mM citrate phosphate buffer, pH range of 3-8 at one pH unit intervals at final concentrations of 510, 146, and 37 µM, respectively. All peptide and protein samples were prepared by dialysis against 20 mM citrate phosphate buffer, pH 7.0 at refrigerated temperatures. The final pH of the protein samples were confirmed to ensure that the pH was within 7.0 ± 0.05.

The effect of solvent dielectric constant and viscosity (both of which are temperature dependent) were also explored as potential sources of the spectral alterations of the aromatic side chains. Solutions of aromatic amino acid analogs were prepared in a number of different organic solvent-water mixtures in which titration of organic solvents in water was used to mimic thermally induced changes of water dielectric and viscosity values over the temperature range of 25-60 °C. The dielectric constants and viscosity of the mixtures were determined as the sum of the mole fraction of each component times the dielectric constant and viscosity of the neat liquid, respectively. Prior to absorbance measurements at 25 °C, all samples were incubated at that temperature for five minutes, sufficient to reach thermal equilibrium.

5.2.3. Absorbance Measurements and Data Analysis

High resolution absorbance spectra were obtained over a temperature range (10-60 °C) employing an Agilent 8453 UV-visible spectrophotometer equipped with a

diode array detector and Peltier temperature controller (HP 89090A). All samples were analyzed in a quartz cuvette with a 1-cm path length. The temperature was allowed to equilibrate for 5 min prior to the acquisition of each spectrum, sufficient for equilibrium to be obtained. An integration time of 25 sec was used to obtain spectra with a high degree of precision and samples were analyzed over the wavelength range of 200-400 nm. Second-derivative spectra were calculated using a nine-point data filter and third-order Savitzky–Golay polynomial with Chemstation software (Agilent). A spline function was applied to the resulting spectra using 99 interpolated points between each raw data point. This approach permits a resolution of 0.01 nm to be obtained under optimal conditions[5-7]. All second-derivative spectra were exported to Origin software to determine peak positions. Results are reported with error bars representing the standard deviation of the mean from three independent measurements.

5.2.4. Fluorescence Acrylamide Quenching

Fluorescence spectra were acquired using a Photon Technology International (PTI) spectrofluorometer (Lawrenceville, NJ) equipped with a turreted 4-position Peltier-controlled cell holder. An excitation wavelength of 295 nm was used to primarily excite Trp residues and the emission spectra were collected from 310 to 400 nm with a step size of 1 nm and a 1 sec integration time. Excitation and emission slits were set at 5 nm. Emission spectra were collected every 2.5 °C with a 3 min equilibration time over a temperature range of 10 to 40 °C. A buffer baseline was

subtracted from each raw emission spectrum. The fluorescence of proteins were monitored at their emission maximum and quenched by the progressive addition of small aliquots of an acrylamide stock solution prepared in 20 mM citrate phosphate buffer, pH 7.0. The ratio of the tryptophan fluorescence intensity in the absence and presence of acrylamide was plotted as a function of increasing acrylamide concentration to estimate extent of quenching. No corrections were necessary for filter effects under these conditions.

5.2.5. Computational Analysis

Computational assessment of solvent effects on the ultraviolet absorption peaks of the three model aromatic residues was performed employing quantum chemical calculations. The molecular structure of the amino acids was sketched in SYBYL[9] and refined to default convergence thresholds via molecular mechanics optimization using the Tripos Molecular Forcefield[10] and Gasteiger-Marsili charges[11]. These structures were then subjected to a quantum chemical optimization in Gaussian 03[12] using the B3-LYP[13,14] hybrid density functional method and the split valence 6-31G(p,d) basis set[15,16] (all convergence criteria left at default values). During all geometry optimizations, the molecules were presumed to be in their ground electronic state, which implies neutral charge and singlet spin configuration.

The ultraviolet spectrum of each of these systems was predicted computationally by the time-dependent density functional method as implemented in Gaussian 03[17], again using B3-LYP functionals and 6-31G(p,d) basis sets. Temperature-dependent

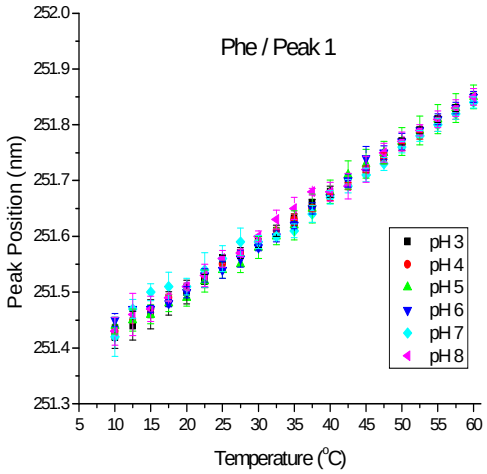
solvent perturbations of these spectra were modeled with the IEF-PCM method[18]. Default values for all solvation parameters were used, except for the following:

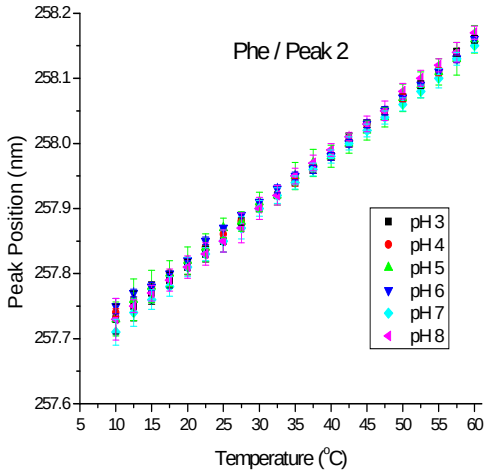
- Pauling radii were specified for all atoms
- Solvent temperatures were explicitly specified: for water, a distinct calculation was performed at each temperature from 0 – 100 ºC in 5 degree increments;
    - Temperature dependent dielectric constants were specified for each solvent / T instance: values for water were obtained from the studies of Fernandez *et al.*[19]
    - Temperature dependent solvent densities were specified for each solvent / T instance: all were derived from the measurements of Lide and Kehiaian[20].

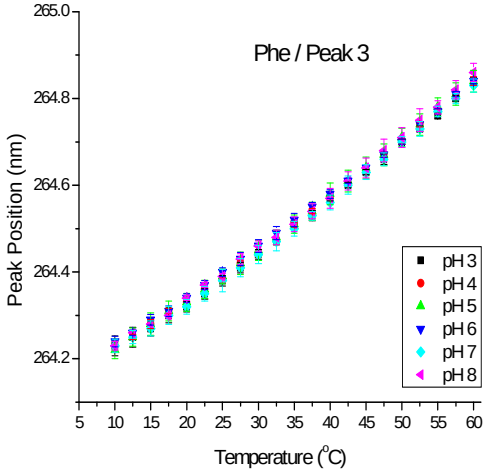
5.3. Results

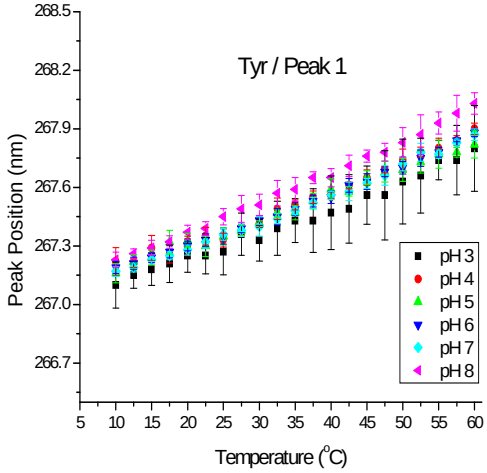
5.3.1. Temperature dependent 2nd Derivative UV Shifts of Model Aromatic Residues

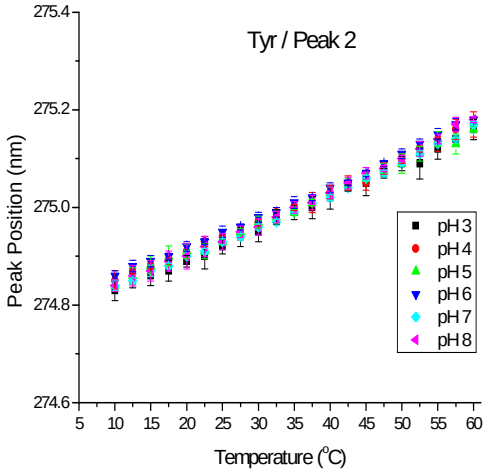
Figure 5.1 displays the intrinsic effect of pH and temperature on the 2nd derivative ultraviolet absorption spectra of the model aromatic amino acid analogs *N*-acetyl-L-phenylalanine ethyl ester, *N*-acetyl-L-tyrosine ethyl ester and *N*-acetyl-L-tryptophan ethyl ester over a wide range of pH (3-8) and temperature (10-60 °C). The N-acetylated C-ethyl esterified analogs were selected to eliminate and/or reduce undesired electrostatic interactions of the termini with solvent molecules. Moreover, the rigidity introduced to the molecules due to presence of such bulky substitutes provides a more realistic simulation of the aromatic side chains as part of protein peptide backbones.

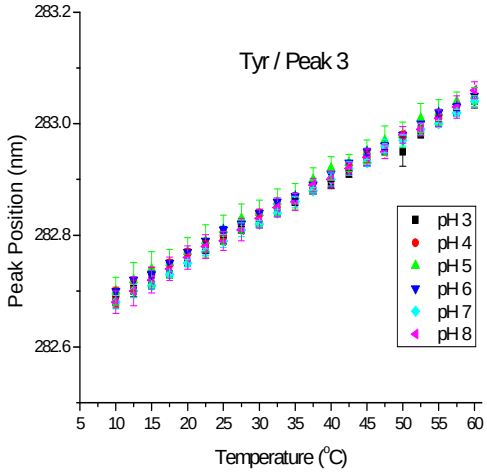

Phe / Peak 1
Peak Position (nm)
Temperature (°C)
pH 3
pH 4
pH 5
pH 6
pH 7
pH 8

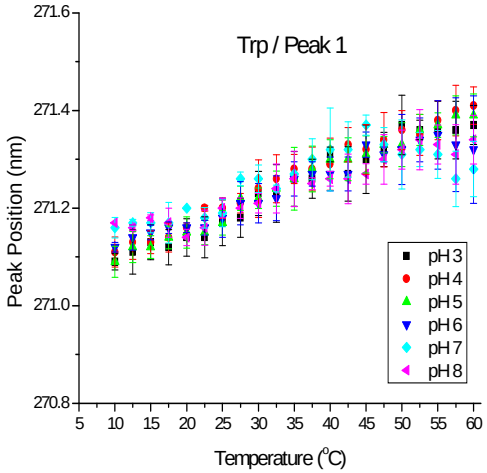

Phe / Peak 2
Peak Position (nm)
Temperature (°C)
pH 3
pH 4
pH 5
pH 6
pH 7
pH 8

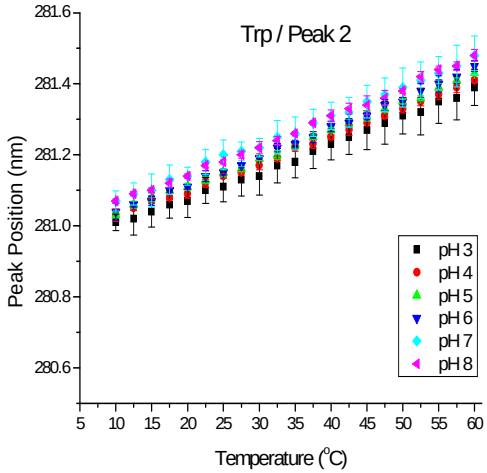

Phe / Peak 3
Peak Position (nm)
Temperature (°C)
pH 3
pH 4
pH 5
pH 6
pH 7
pH 8


Tyr / Peak 1
Peak Position (nm)
Temperature (°C)
pH 3
pH 4
pH 5
pH 6
pH 7
pH 8


Tyr / Peak 2
Peak Position (nm)
Temperature (°C)
pH 3
pH 4
pH 5
pH 6
pH 7
pH 8


Tyr / Peak 3
Peak Position (nm)
Temperature (°C)
pH 3
pH 4
pH 5
pH 6
pH 7
pH 8


Trp / Peak 1
Peak Position (nm)
Temperature (°C)
pH 3
pH 4
pH 5
pH 6
pH 7
pH 8


Trp / Peak 2
Peak Position (nm)
Temperature (°C)
pH 3
pH 4
pH 5
pH 6
pH 7
pH 8

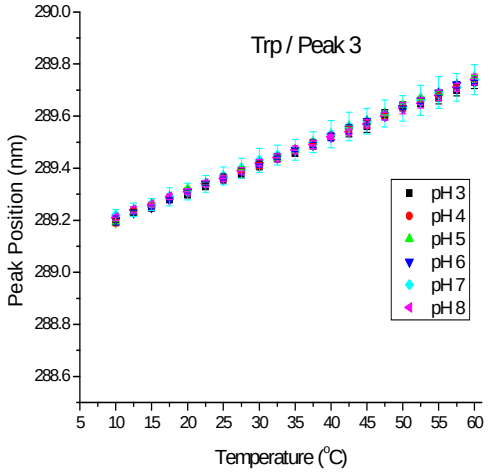


Figure 5.1.The pH and Temperature dependence of the derivative absorbance peaks of model amino acids. Spectra were collected at 2.5 °C intervals after a 5 min equilibration over the temperature range of 10-60 °C. The model compounds used were N-acetyl-L-phenylalanine ethyl ester, N-acetyl-L-tyrosine ethyl ester and N-acetyl-L-tryptophan ethyl ester. (n=3)

The 2nd derivative spectrum of each model aromatic residue displays three distinct absorption peaks. The phenylalanine analog exhibits peaks at ~ 251, 257, and 264 nm; tyrosine at ~ 267, 274, and 282 nm and tryptophan at ~ 271, 281, and 289 nm. All derivative peaks show a general linear / quasi linear shift to higher wavelengths (i.e., red shifts) as a function of increasing temperature. No significant pH dependence is observed over the range examined. To estimate the slopes of the temperature dependent peak shifts, the initial linear portion of the data in the range of 10-35 °C was fitted to a straight line to minimize the slight curvature of the data in the non-linear regions at higher temperatures. Since there is no detectable pH dependency, data at pH 7.0 was selected for comparison purposes in Table 5.1. The data suggests that each of the three aromatic residues and their three deconvoluted derivative peaks exhibit different dependencies on solvation effects as manifested by different temperature dependent slopes. The magnitudes of the shifts are significant since shifts as low as 0.5 nm have been indicative of protein conformational alterations when employing this technique[5-8].

		Model Compound		
		N-acetyl-L-phenylalanine ethyl ester	N-acetyl-L-tyrosine ethyl ester	N-acetyl-L-tryptophan ethyl ester
Peak 1	Initial λ (nm)	251.42±0.04	267.13±0.03	271.16±0.02
	Slope x 10^3 (nm/°C)	7.0 (R^2=0.95)	12.4(R^2=0.99)	4.8(R^2=0.81)
Peak 2	Initial λ (nm)	257.71±0.02	274.84±0.01	281.07±0.03
	Slope x 10^3 (nm/°C)	9.1(R^2=0.99)	6.0(R^2=0.99)	7.8(R^2=0.99)
Peak 3	Initial λ (nm)	264.23±0.02	282.68±0.01	289.22±0.02
	Slope x 10^3 (nm/°C)	10.8(R^2=0.99)	7.3(R^2=0.99)	10.3(R^2=0.99)

Table 5.1.Linear fits to derivative absorbance plots between 10-35 °C.

5.3.2. Origin of the Temperature Dependent Peak Shifts

A probable source of the observed temperature induced spectral alterations in aqueous environments involves the interaction of the aromatic side chains with solvent molecules. Observation of the effect of solvent on electronic spectra dates back to more than a century ago[21]. Since then, extensive experimental and computational analyses have revealed a variety of solute-solvent interactions responsible for alterations of the electronic spectra of absorbing molecules in solution. Dipolar (i.e., dielectric effects), dispersive (i.e., van der Waals forces), and short-range specific interactions (e.g., hydrogen bonding) have all been shown to contribute to spectral alterations[22]. The known temperature dependence of such interactions suggests that thermal alterations of solvent physical properties and their effects on the solute electronic spectra could potentially be a source of the peak shifts observed. Association or dissociation of solute molecules can also affect their electronic spectra[22]. The possibility of such solute effects is, however, unlikely here due to the low micro-molar concentrations of the residues employed and the pH independence of the temperature dependent data (Fig 5.1).

5.3.3. Computational Analysis

The origin of the observed spectral alterations was investigated employing quantum chemical calculations according to the computational analysis described in the method section. For each of the three amino acids, quantum chemical excited state calculations resolve the $\pi \rightarrow \pi^*$ transition band as a number of discrete states, each of

which exhibits a roughly linear dependence on solvation effects, but with significant variation in slopes from one state to the next (Fig 5.2). Our calculations do tend to systematically overestimate vertical excitation energies (e.g., the leading edge of the absorption band for the tryptophan analog is computed to be at 269.98 nm), however this is consistent with prior observations on cyclic/polycyclic aromatic compounds[23]. To estimate the effective shift of the observed composite band at a given temperature, the weighted average shift was calculated according to:

$$\overline{\Delta\lambda} = \sum_{i} \frac{\Delta\lambda_i F_i}{\mathbf{F}}$$

where *i* indexes specific optically active peaks (i.e., $F_i \neq 0$) within the $\pi \rightarrow \pi^*$ band, $\Delta\lambda_i$ is the shift of peak *i* at a given T and expressed relative to the peak position for the specific solvent of interest at T = 0 °C, F_i is the oscillator strength of peak *i* at a given T, and **F** is the sum of all oscillator strengths of all peaks within the $\pi \rightarrow \pi^*$ band. For the purpose of this computation, the list of peaks in the $\pi \rightarrow \pi^*$ band included all of those between (and including) the $\pi \rightarrow \pi^*$ peak maximum and the spectral minimum in the trough between the $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ peaks.

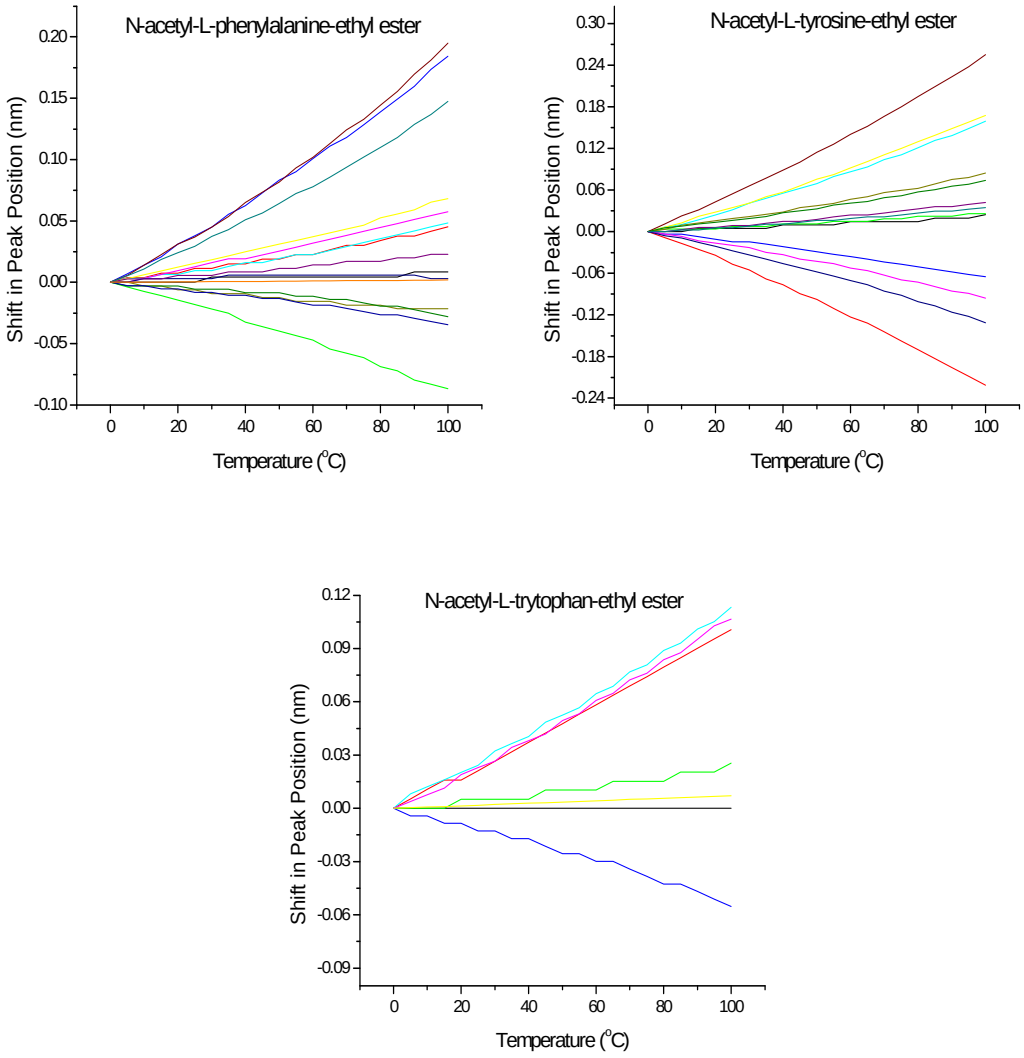


Figure 5.2.Different dependence of the multiple discrete states within the $\pi \rightarrow \pi^*$ transition band on solvation effects manifested by their different temperature dependencies. Model compounds used were N-acetyl-L-phenylalanine ethyl ester, N-acetyl-L-tyrosine ethyl ester, and N-acetyl-L-tryptophan ethyl ester.

Quantum chemical calculations provide a detailed characterization of the multiple excited states within the $\pi \rightarrow \pi^*$ band (Fig 5.2) that collectively lead to the overall red shift observed both experimentally (Fig 5.1) and computationally (Fig 5.3). The trends observed in Figure 5.3 correspond to the temperature dependent solvent shifts averaged over all states, weighted according to the oscillator strength, and are within the $\pi \rightarrow \pi^*$ band of the spectrum. The trends (i.e., red shifts as a function of increasing temperature) obtained from the computational analysis are in agreement with the experimental results.

The magnitudes of the shifts of the absorbance peak positions, however, are smaller than the experimental ones. A variety of effects may account for this discrepancy, of which the greatest probably arises from the differences between the heterogeneous interaction profiles in a rigorously explicit solvation system versus the time averaged implicit solvent model employed herein. In an explicitly solvated environment, it is reasonable to expect that the solute will influence the relative alignments of surrounding solvent molecules in ways that favor and thus enhance the polarization trends suggested in the implicit system. Moreover, the vibrational effects may also induce a greater solvent effect than is suggested by our static model; rigorously testing this prospect would be computationally very challenging due to the large number of vibrational modes accessible to these species at room temperature.

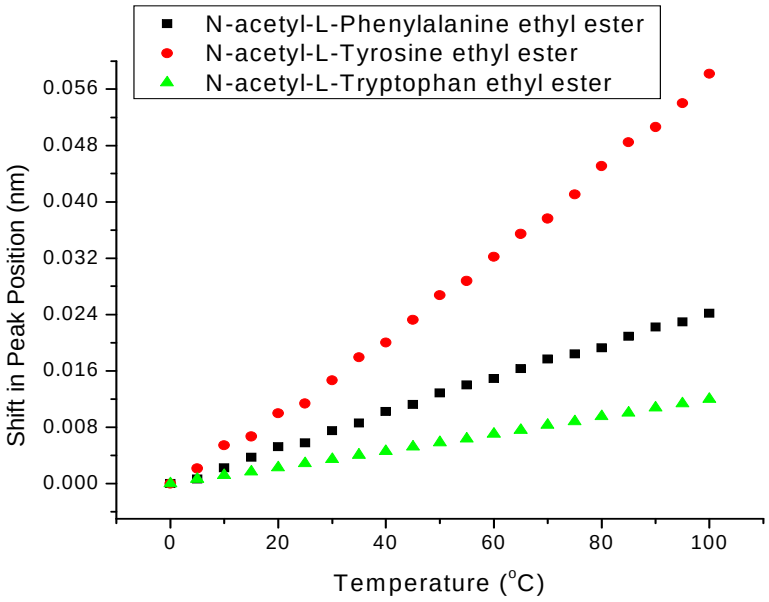


Figure 5.3.Temperature dependence of the absorbance peaks of model amino acids obtained from quantum mechanical analysis over the temperature range of 0-100 °C. Model compounds used were N-acetyl-L-phenylalanine ethyl ester, N-acetyl-L-tyrosine ethyl ester, and N-acetyl-L-tryptophan ethyl ester.

Temperature dependent density and dielectric values were the only explicit solvent parameters defined in the computational model. Since the temperature dependent density alterations showed minor effects (data not shown), the computational analysis suggests the temperature dependent dielectric alterations to be a major source of the spectral changes observed.

The relative shifts of the component state vary significantly (some in fact show blue shifts) as a result of structural anisotropy within these model systems (Fig 5.2). Since each state can be theoretically decomposed into relative contributions arising from specific atoms, it could be possible to use incumbent state-specific information to predict differences in the solvent effects arising from differences in the solvent exposure profiles of amino acids in different environments.

5.3.4. Other Solvent Properties Contributing to the Spectral Alterations Observed

According to Coulomb's law, the potential energy between two charges (here solute-solvent dipoles) is altered as a function of solvent dielectric constant[24]. It is reasonable to expect that the temperature induced alterations of the solvent dielectric constant will be at least a partial source of the spectral alterations observed, as already suggested by the computational analysis. The implicit nature of the computational analysis, however, fails to reflect differences in the spectroscopic effects arising from other solvent physical properties. Thus, essentially all of the effects on the solute spectra are assigned to temperature dependent dielectric alterations. A variety of organic solvents with a range of physical properties were selected to experimentally

investigate other solvent properties that might be contributing to the spectral alterations observed (Table 5.2)[24].

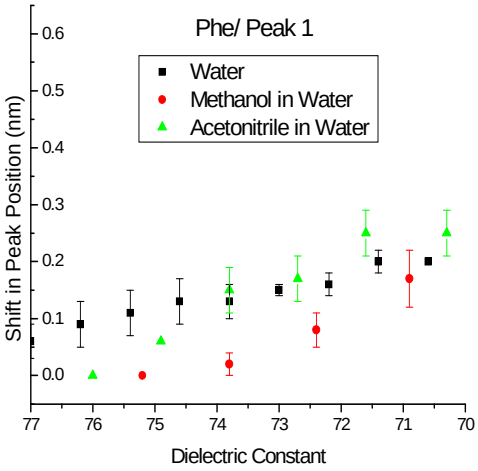
ε is solvent dielectric constant, π* is the solvent polarizability, α is solvent hydrogen donating ability, and β is solvent hydrogen accepting ability. DMSO with a high proton accepting ability (large β value) is the model solvent of choice for investigating short-range specific interactions (i.e., hydrogen bonding). Methanol and acetonitrile are polar organic solvents with considerably lower dielectric properties compared to water and are employed here for examining dipolar interactions and hexane is a nonpolar solvent suitable for investigating the effect of van der Waals interactions on the temperature dependent spectral alterations observed.

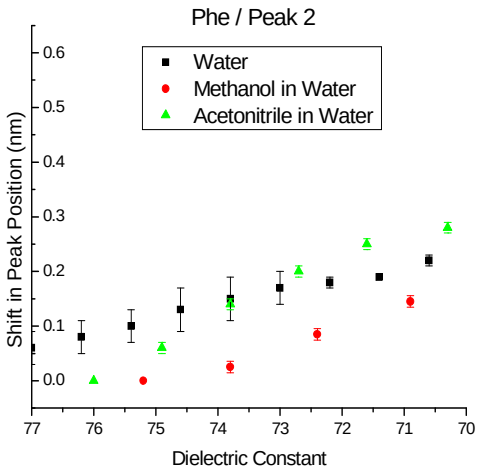
Solvent	ε	π	α	β
DMSO	47	1.0	0.00	0.76
Methanol	33	0.6	0.93	0.66
Water	78	1.1	1.17	0.47
Acetonitrile	36	0.75	0.19	0.40
Hexane	2	-0.04	0.00	0.00

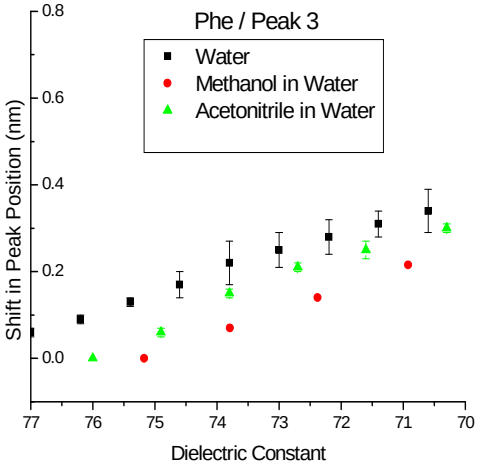
Table 5.2.Various solvent scales. ε is solvent dielectric constant, π* is the solvent polarizability, α is solvent hydrogen donating ability, and β is solvent hydrogen accepting ability.

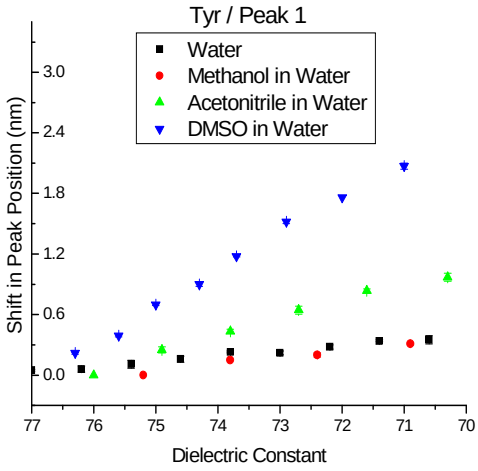
To examine whether temperature dependent solvent dielectric changes are the dominant contributor to the spectral alterations observed as suggested by the computational analysis, mixtures of organic solvents in water were used to mimic changes in the dielectric constant of water induced thermally. The dielectric constants of the mixtures were estimated as the sum of the mole fraction of each component times the dielectric constant of the neat liquid[20, 25-27].

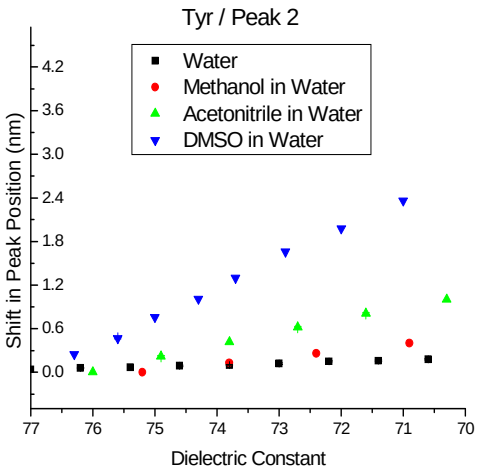
If the spectral alterations observed are due to alterations of the solvent dielectric properties alone, then all binary mixtures should produce similar peak shifts (i.e., slopes). This was not seen to be the case here, however (Fig 5.4). Mixtures of methanol and acetonitrile in water produced similar magnitude of peak shifts for phenylalanine and tryptophan residues with the exception of the tyrosine analog where larger red shifts were observed as a function of increasing acetonitrile in water. The latter is presumably due to the presence of the well known π-π interactions between the acetonitrile cyanide group π orbitals and the phenol ring of the tyrosine side chain. The large electron donating ability of the phenol hydroxyl group enhances the electron density of the aromatic ring and contributes to the enhanced π-π interactions observed. The difference in stabilization energy provided by the presumed π-π interactions on the tyrosine ground and excited state explains the red shifts observed; the electron donating capability of oxygen to the benzene ring of phenol is significantly higher in the excited state compared to the ground state further strengthening the π-π interactions in the excited state[28]; the larger stabilization energy of the excited state, according to Planck's law, results in the red shifts observed.

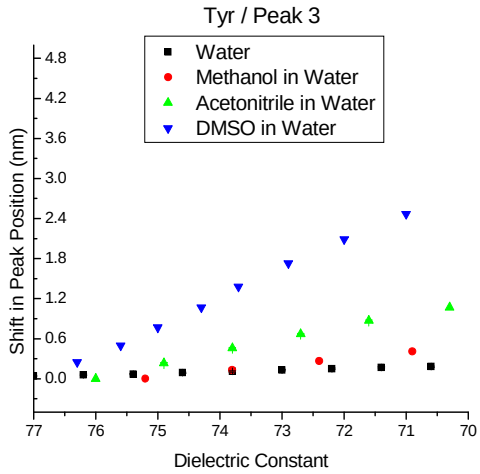

Phe/ Peak 1
Water
Methanol in Water
Acetonitrile in Water
Shift in Peak Position (nm)
Dielectric Constant

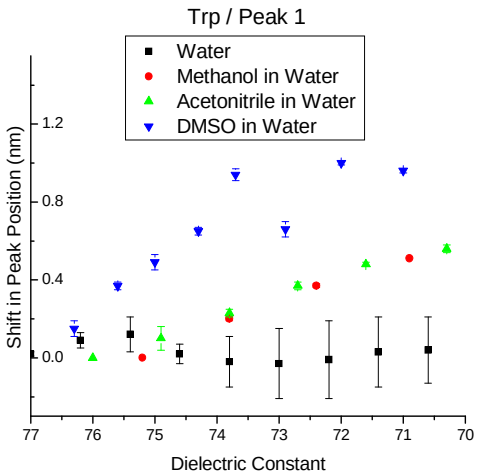

Phe / Peak 2
Water
Methanol in Water
Acetonitrile in Water
Shift in Peak Position (nm)
Dielectric Constant

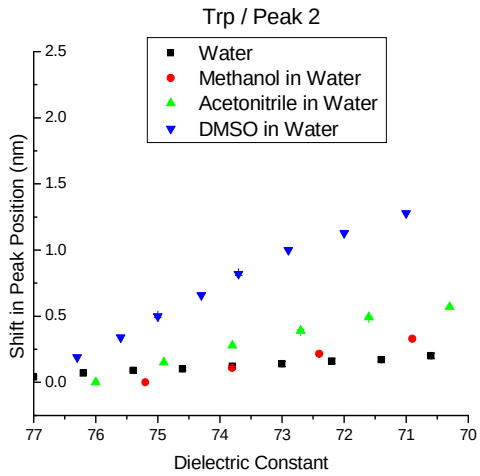

Phe / Peak 3
Water
Methanol in Water
Acetonitrile in Water
Shift in Peak Position (nm)
Dielectric Constant


Tyr / Peak 1
Water
Methanol in Water
Acetonitrile in Water
DMSO in Water
Shift in Peak Position (nm)
Dielectric Constant


Tyr / Peak 2
Water
Methanol in Water
Acetonitrile in Water
DMSO in Water
Shift in Peak Position (nm)
Dielectric Constant


Tyr / Peak 3
Water
Methanol in Water
Acetonitrile in Water
DMSO in Water
Shift in Peak Position (nm)
Dielectric Constant


Trp / Peak 1
Water
Methanol in Water
Acetonitrile in Water
DMSO in Water
Shift in Peak Position (nm)
Dielectric Constant


Trp / Peak 2
Water
Methanol in Water
Acetonitrile in Water
DMSO in Water
Shift in Peak Position (nm)
Dielectric Constant

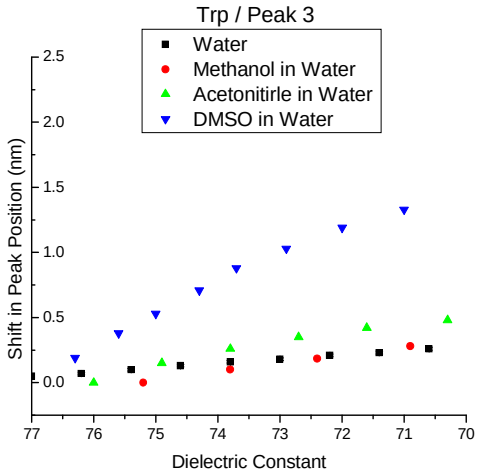


Figure 5.4.Shifts of the derivative absorbance peaks of model amino acids as a function of solvent dielectric constant. Model compounds used were N-acetyl-L-phenylalanine ethyl ester, N-acetyl-L-tyrosine ethyl ester, and N-acetyl-L-tryptophan ethyl ester. (n=2)

The largest deviations are observed for the DMSO mixtures presumably due to its extensive proton accepting ability (Table 5.2). We were unable to collect reliable data for the phenylalanine analog in DMSO-water mixtures since DMSO absorbs in the same region of the UV spectrum and thus interferes with data collection. High resolution data, however, show major differences for the other two analogs (Fig 5.4), with more extensive deviations observed for the tyrosine analog. In a molecule possessing a hydroxyl group, the direction of the spectral shifts induced by hydrogen bond formation depends upon whether the molecule acts as a proton donor or acceptor[29]. If the hydroxyl group acts as proton acceptor, then a blue shift is observed whereas if it acts as a proton donor, a red shift occurs as observed here. When phenol forms a hydrogen bond with a proton acceptor like DMSO, there occurs charge transfer from the nonbonding orbital of DMSO to the O-H antibonding δ^* orbital, which makes the δ bond of O-H a little weaker. In other words, the electron density on the oxygen atom of phenol becomes larger than in the free molecule so that the coulomb potential energy of the oxygen $2p\pi$ electrons decreases, resulting in a red shift[27]. In addition, electron transfer from oxygen $2p\pi$ electrons to the benzene ring π system is larger in the π-π^* excited state than in the ground state. Therefore, the charge separation becomes larger in the excited state so that the proton donating ability of phenol is increased much more in the excited state compared to ground state. This is responsible for the red shifts observed (Planck's law).

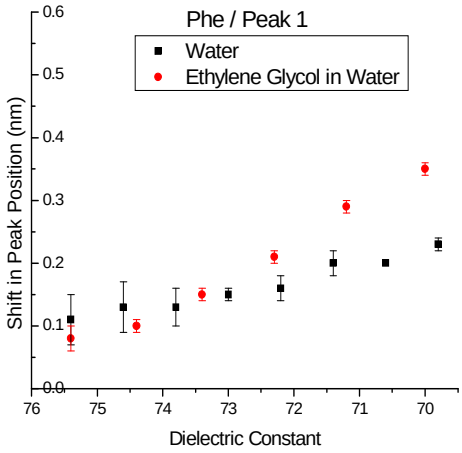
5.3.5. Effect of Microenvironment Viscosity

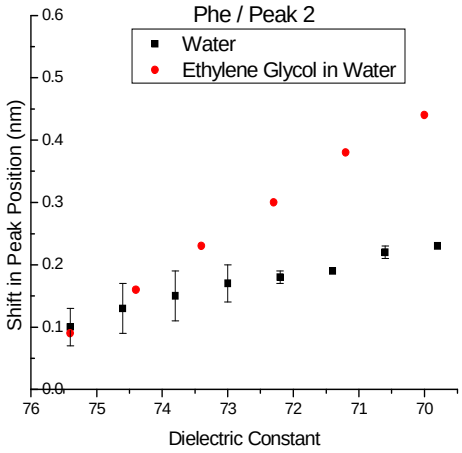
The extensive deviations observed for the mixtures containing DMSO could also potentially be due to its more viscous nature[20]. The side chains of the aromatic residues posses certain levels of rotational and vibrational motion in solution. Shapovalov *et al.*[30] have shown that direct correlation exist between the side chain capability to posses rotameric states and its electron density in which rotameric side chains posses larger electron densities than non-rotameric ones. Considering the nature of the molecular dipolar interactions herein as a major contributor to the spectral alterations observed, the reduction of the side chain electron densities should result in weaker solute-solvent dipolar interactions. Moreover, according to Coulomb's law, effective interaction of solvent-solute dipoles depends on the distance between the two dipoles (i.e., dipoles encountering each other in solution) and also on the proper alignment of the dipoles (i.e., the angle between the two dipoles). The probability of such effective interactions is expected to be statistically enhanced by increasing the mobility of the solute molecules in solution. We therefore speculate that the spatial restriction on the microenvironment of the aromatic residues due to higher viscosities may be partially responsible for the spectral alterations observed.

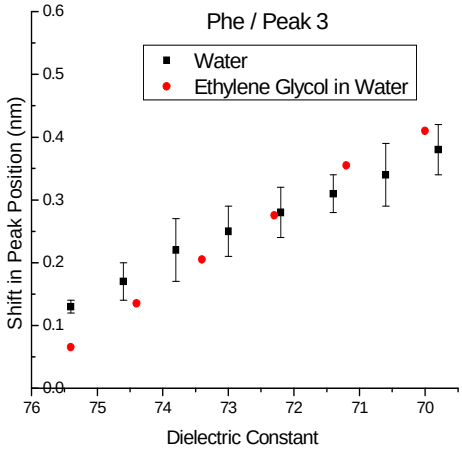
To further examine the effect of viscosity on the electronic spectra of the model aromatic analogs, mixtures of ethylene glycol in water were used to mimic changes in the dielectric constant of water induced thermally. Ethylene glycol is considerably more viscous than other solvents examined (~ 20 fold more viscous than

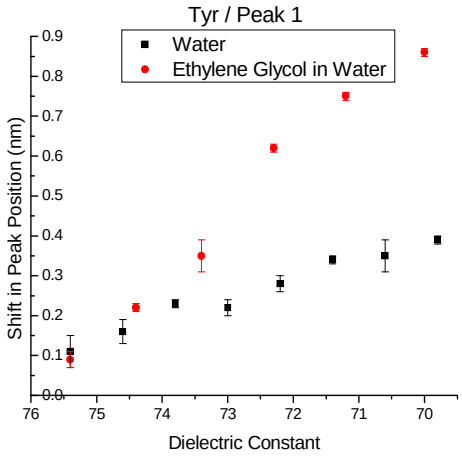
water)[20]. With the exception of phenylalanine peak 3, all data show enhanced red shifts with increasing ethylene glycol concentration in water confirming the restrictive effect of microenvironment viscosity on the electronic spectra of the aromatic residues (Fig 5.5).

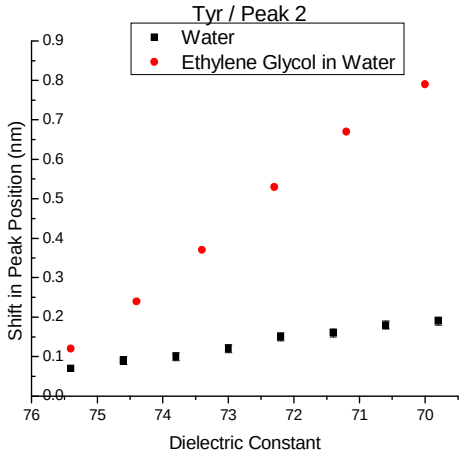
The restrictive effects of viscosity is manifested by red shifts in peak position similar to those observed due to the enhanced restrictive effects of hydrogen bonding and π-π interactions on the side chain mobility. The ethylene glycol has similar dielectric properties to methanol and both contain a functional hydroxyl group that can be involved in possible hydrogen bonding interactions with solute molecules. Increasing methanol concentrations in water produced similar red shifts to the ones induced thermally (Figure 4), although, ethylene glycol-water mixtures produced much larger red shifts, presumably due to the more viscous nature of ethylene glycol compared to methanol.

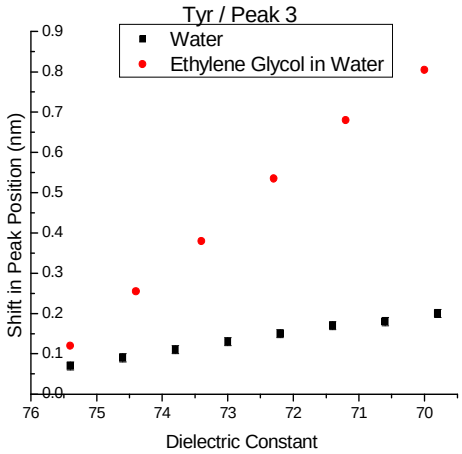

Phe / Peak 1
Water
Ethylene Glycol in Water
Shift in Peak Position (nm)
Dielectric Constant


Phe / Peak 2
Water
Ethylene Glycol in Water
Shift in Peak Position (nm)
Dielectric Constant


Phe / Peak 3
Water
Ethylene Glycol in Water
Shift in Peak Position (nm)
Dielectric Constant


Tyr / Peak 1
Water
Ethylene Glycol in Water
Shift in Peak Position (nm)
Dielectric Constant


Tyr / Peak 2
Water
Ethylene Glycol in Water
Shift in Peak Position (nm)
Dielectric Constant


Tyr / Peak 3
Water
Ethylene Glycol in Water
Shift in Peak Position (nm)
Dielectric Constant

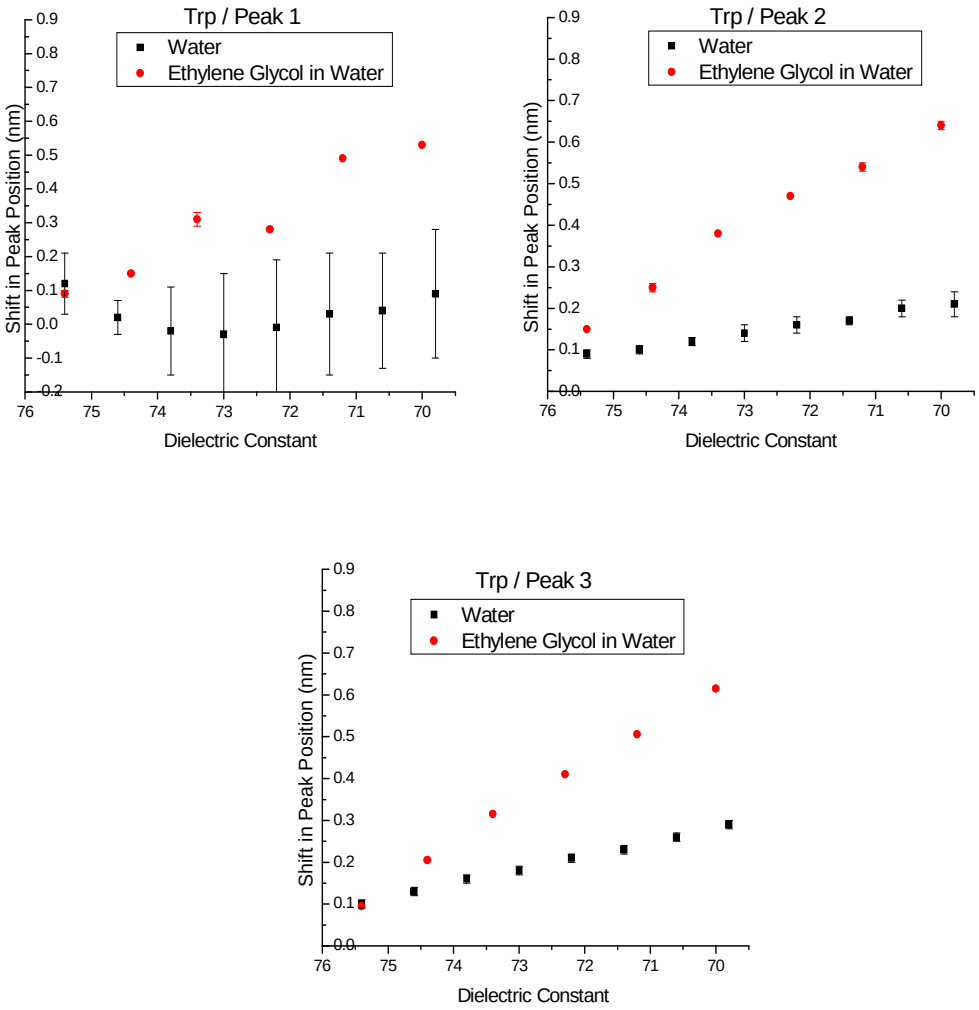


Figure 5.5.Shifts in the derivative absorbance peaks of model amino acids as a function of dielectric constant. Model compounds used were N-acetyl-L-phenylalanine ethyl ester, N-acetyl-L-tyrosine ethyl ester, and N-acetyl-L-tryptophan ethyl ester. (n=2)

5.3.6. Effect of Dispersive Interactions

The temperature dependent spectral alterations of the phenylalanine analog were monitored in hexane, a very non polar solvent with much lower dielectric constant values than water. Comparable peak shifts were observed in hexane to other organic solvents (Fig 5.6). Considering the highly non-polar nature of both solute (i.e., the benzene ring of phenylalanine) and the solvent, hexane, apolar interactions are expected to be the dominant contributor to the shifts observed. Current computational analyses are under way to examine this phenomenon in more detail.

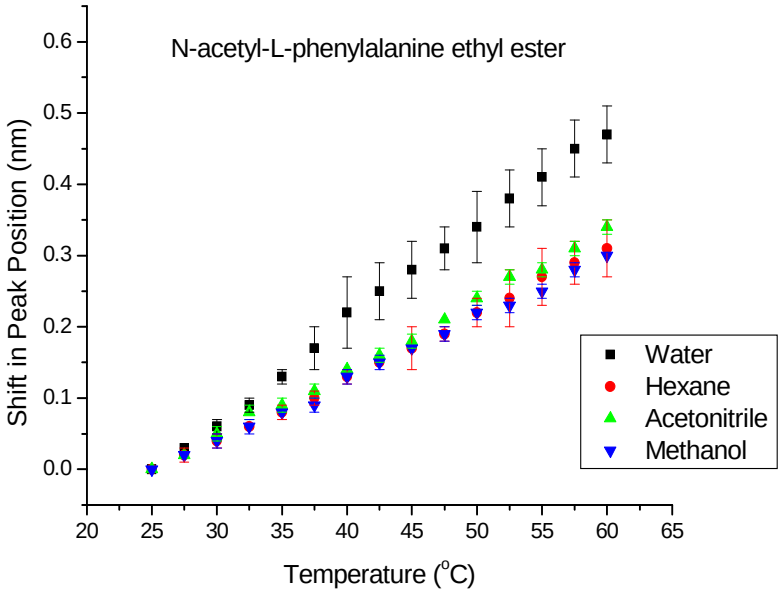


Figure 5.6.Temperature dependent derivative absorbance peak shifts of N-acetyl-L-phenylalanine ethyl ester in different organic solvents over the temperature range of 25-60 °C. (n=2)

5.3.7. Effect of Solute-Solvent Interactions on Temperature Dependent Peak Shifts

To examine the effect of the various solute-solvent interactions discussed here on the temperature dependent 2nd derivative UV peak shifts of the aromatic residues, the tryptophan and tyrosine analogs were studied in mixtures of DMSO and methanol in water.

The magnitude of the peak shifts of the tryptophan analog decreases in both DMSO and methanol mixtures in water with increasing organic solvent concentration (Fig 5.7). The data suggest that the slopes of the temperature dependent plots are sensitive to the dielectric properties of the solution environment as also supported by the computational analysis. Methanol exhibits enhanced effects on the slopes compared to DMSO due to its significantly lower dielectric constant values (Table 5.2).

The effect of other stabilizing interactions (i.e., hydrogen bonding) was also examined by the ability of the DMSO mixtures to alter the temperature dependent 2nd derivative peak shifts of the tyrosine and tryptophan analogs with different proton donating abilities to DMSO. (Phenylalanine was not included here due to both DMSO absorbance interference and also phenylalanine's lack of proton donating abilities compared to the other two amino acid analogs). The magnitude of the slopes as a function of increasing DMSO concentration in water decreases more significantly for the tyrosine analog due to more extensive hydrogen bonding between the tyrosine side chain and DMSO as discussed earlier (Fig 5.4).

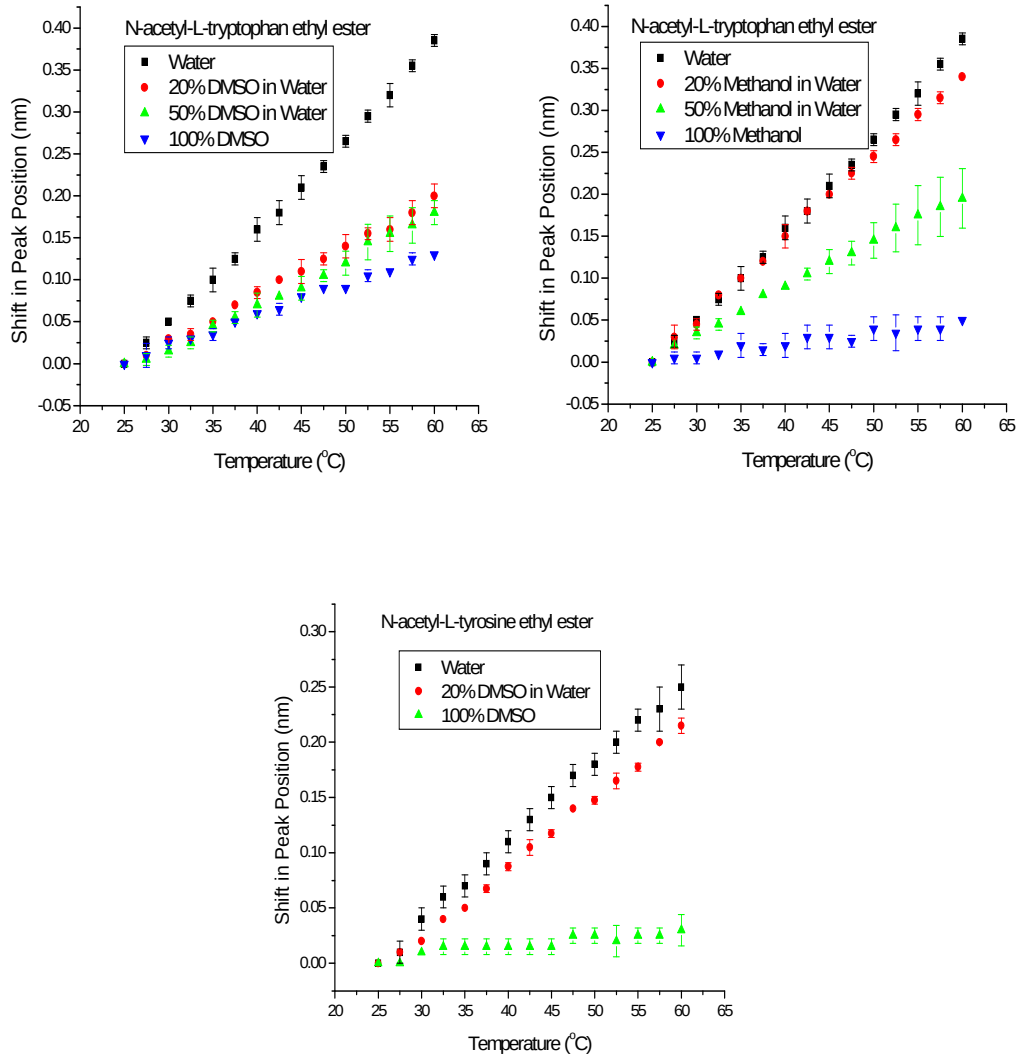


Figure 5.7.Temperature dependent derivative absorbance peak shifts of model aromatic amino acids in different organic solvent mixtures in water over the temperature range of 25-60 °C. Model compounds used were N-acetyl-L-tyrosine ethyl ester, and N-acetyl-L-tryptophan ethyl ester. (n=2)

5.3.8. Temperature Dependent 2nd Derivative Absorbance Spectroscopy of Aromatic Amino Acids in Proteins

Solvent penetration into protein apolar cores contributes significantly to the high polarizability of protein interiors where hydration of the interior results in dielectric constant values three to six fold larger than those observed for dry protein powders[31]. Focusing on protein pre-transition region, we have shown both experimentally and computationally the sensitivity of the 2nd derivative UV peak shifts of the three aromatic residues to thermally induced alterations of the solvent dielectric properties. We therefore hypothesize that the temperature dependent 2nd derivative UV peak shifts of the aromatic residues in proteins should correlate with the hydration of the interior aromatic residues and consequently solvent penetration into the protein core. One major contributor to such solvent penetration is protein intramolecular dynamic motions. Thus, we suggest that temperature dependent spectral alterations observed should be a qualitative measure of such dynamic behavior.

A direct correlation between the extent of peak shifts ($d\lambda/dT$) and the local mobility of the aromatic side chains was also investigated in which restrictive factors such as hydrogen bonding of the side chains and more viscous microenvironments lowered the magnitude of the shifts observed. We therefore further suggest that the temperature dependent peak shifts should also correlate with the local dynamic motions of the aromatic side chains and that the slopes observed are sensitive to both protein global and local dynamic motions.

The temperature dependent 2nd derivative UV analysis of the aromatic residues was therefore investigated in a variety of model proteins and peptides to test this hypothesis (Table 5.3). All proteins examined retained their native conformation in the temperature range of 10-35 °C examined as observed by CD, intrinsic fluorescence, and turbidity analysis (data not shown). Azurin, RNase-T1, and protein A were selected due to their highly buried tryptophan residues with limited solvent accessibility as corroborated by tryptophan peak positions of ~ 319, 323, and 326 nm respectively (data not shown).

Protein	No. Subunits	Residues per Subunit	Subunit MW (kDa)	No. Phe per Subunit	No. Tyr per Subunit	No. Trp per Subunit
Protein A	1	60	7.1	0	1	1
RNase-T1	1	130	14.0	4	10	1
Azurin	4	128	13.9	6	2	1
Melittin	n.a.	26	2.85	0	0	1
Leucine Enkephaline	n.a.	5	0.55	1	1	0
Substance P	n.a.	11	1.35	2	0	0
Angiotensin I	n.a.	10	1.3	1	1	0
HSA	2	584	66.2	31	18	1
L-Asparaginase	4	326	34.6	8	12	1

Table 5.3.Physical and chemical properties of proteins studied.

The 2nd derivative UV peak position of the tryptophan residue was monitored as a function of increasing temperature. The tryptophan analog exhibited the largest red shifts followed by melittin, a linear peptide, and then the model globular proteins containing highly buried tryptophan residues (Fig 5.8). The dramatic reduction in the temperature dependent slopes for the highly buried side chains supports the hypothesis that there is a direct correlation between the magnitude of the temperature dependent slopes and the extent of side chain hydration.

To evaluate the applicability of temperature dependent 2nd derivative UV spectroscopy of the aromatic residues to the study of protein dynamics, acrylamide quenching of the proteins intrinsic fluorescence was employed as a complimentary technique for comparison purposes[32]. Fluorescence acrylamide quenching (Fig 5.8, inset) exhibits a similar extent of quenching and therefore comparable dynamic motions for RNase-T1 and protein A. The 2nd derivative UV slopes are in agreement with the quenching studies in which similar slopes were obtained for the two model proteins over the temperature range examined.

The 2nd derivative absorbance slopes, however, showed larger slopes for azurin compared to RNase-T1 in disagreement with the fluorescence acrylamide quenching results reported by Eftink *et al*.[32] One possible explanation may be an increased sensitivity of the 2nd derivative technique to the local dynamic motions of the aromatic side chains, specifically to hydrogen bonding. The indole side chain of the tryptophan residue in Azurin is surrounded by a number of nonpolar side chains, establishing strong apolar interactions[33]. In RNase-T1, however, the tryptophan side

chain forms a hydrogen bond with a glutamic acid residue in its proximity[34]. Rueda *et al.*[35] show that the rigidifying force constants exerted by hydrogen bonds on protein dynamics are much larger (~ 4 fold) than those from apolar interactions. Considering the sensitivity of the 2nd derivative technique to the local flexibility of the side chains, the reversed order of slopes could be explained if the fluorescence technique (i.e., its excited state) is insensitive to such local dynamic properties.

One potential major advantage of the 2nd derivative UV absorption technique over fluorescence methods is the ability to also examine the phenylalanine and tyrosine side chains in addition to tryptophan residues. We therefore also examined the peak shifts of a variety of small linear peptides containing phenylalanine and tyrosine residues (Table 5.3).

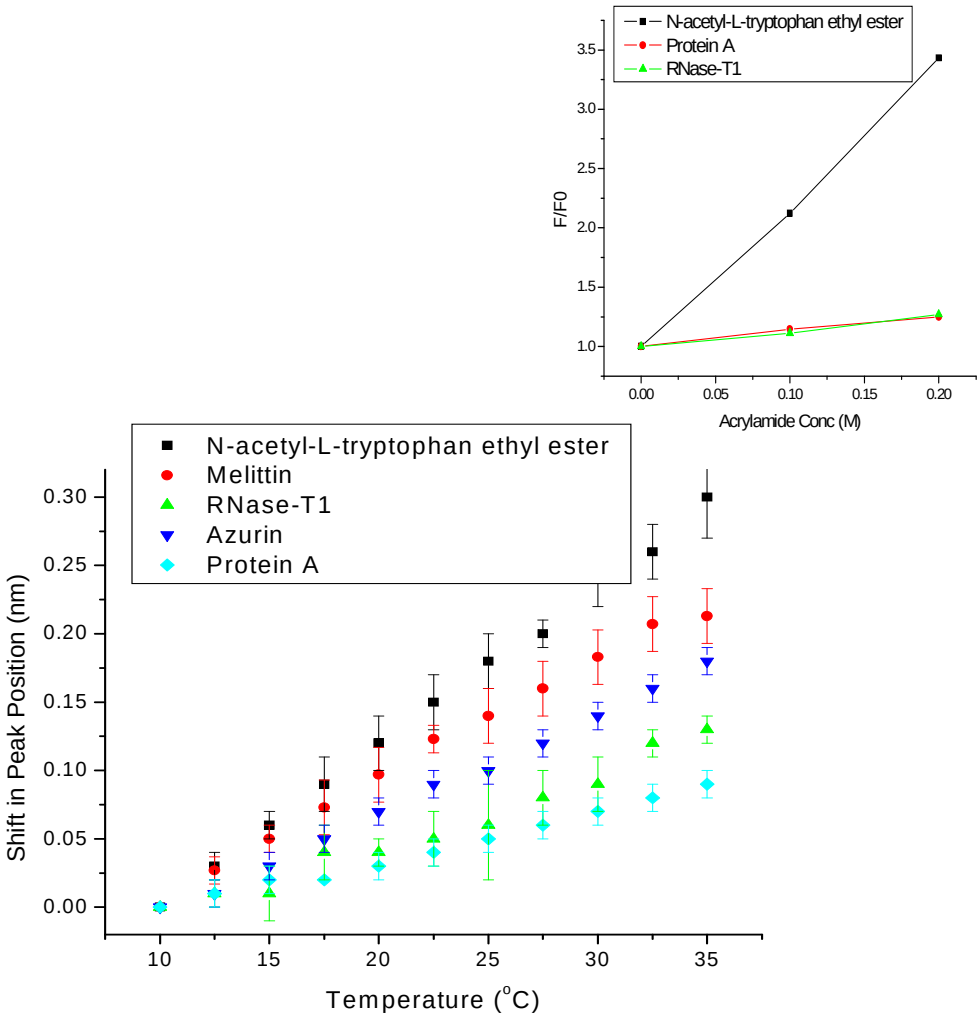


Figure 5.8.Temperature dependent derivative absorbance peak shifts of the tryptophan residue of model proteins over the temperature range of 10-35 °C. Inset shows acrylamide quenching of the fluorescence of proteins containing single tryptophan residues. (n=3)

Figure 5.9 demonstrates agreement between the data obtained for the linear peptides with those of the model aromatic analogs. Angiotensin I, however, exhibits decreased slopes when examining tyrosine side chain. This again could be explained by restricted local motions of the tyrosine side chain when it is involved in hydrogen bonding or cation-π interactions with the arginine side chain in its close proximity. Such interaction is absent in substance P due to presence of only apolar residues in the region surrounding tyrosine.

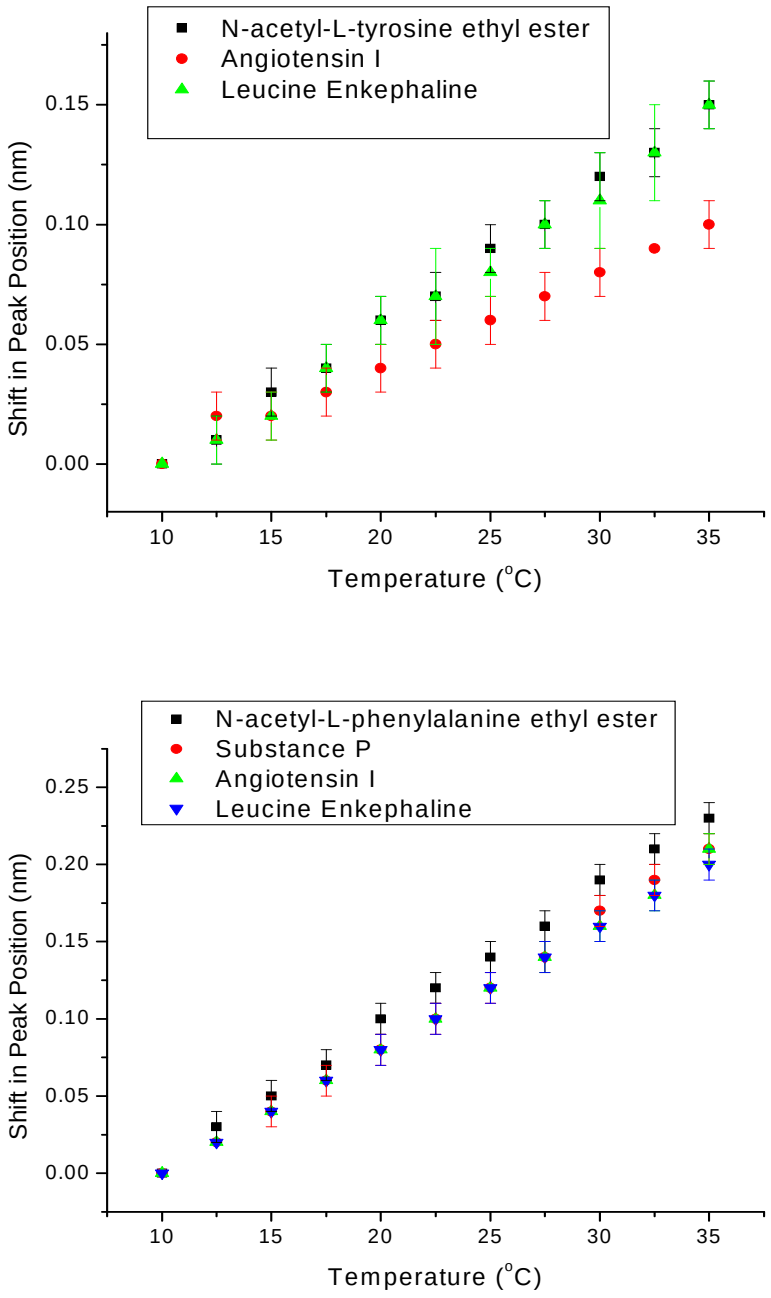
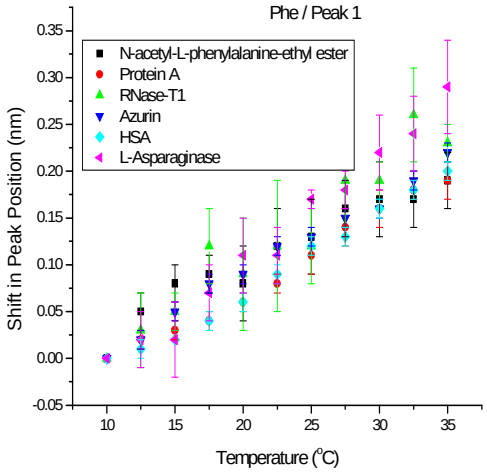
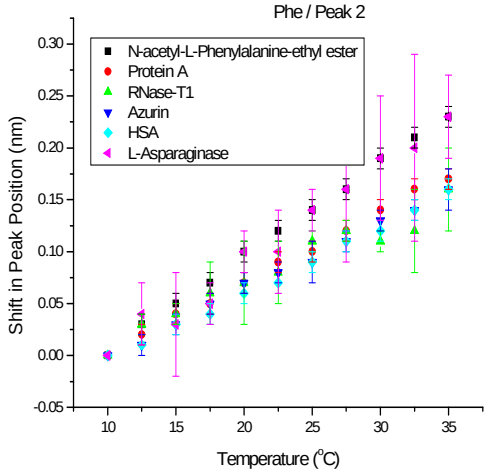


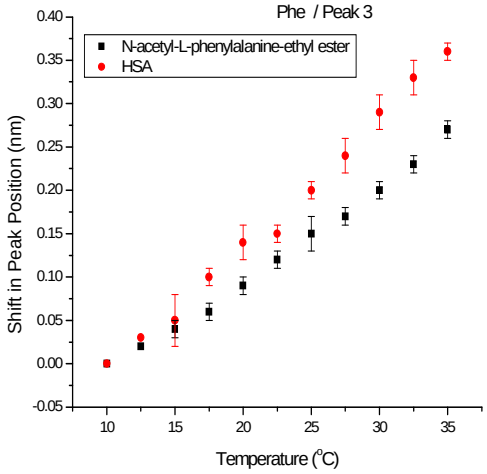
Figure 5.9.Temperature dependent derivative absorbance peak shifts of tyrosine and phenylalanine residues in model peptides over the temperature range of 10-35 °C. Model aromatic amino acids used were N-acetyl-L-tyrosine ethyl ester and N-acetyl-L-phenylalanine ethyl ester. (n=3)

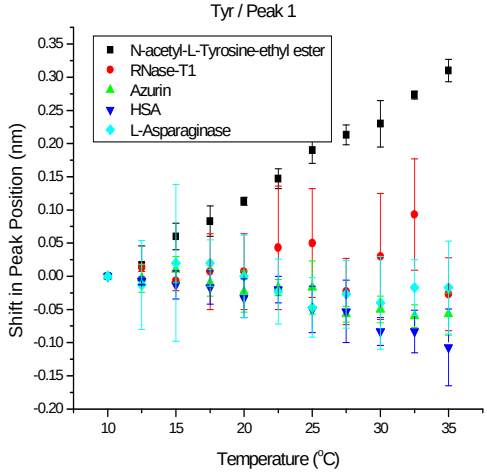
To explore the applicability of this approach to phenylalanine and tyrosine residues, we examined a variety of model proteins (Table 5.3). The temperature dependent shifts of each of the three derivative peaks of either phenylalanine or tyrosine vary when monitored for a given model protein (Fig 5.10). This can be explained based on the previously described quantum chemical excited state calculations from which the main $\pi \rightarrow \pi^*$ transition band was resolved into a number of discrete states, each of which exhibited a different dependence on solvation effects (Fig 5.2). The distinct temperature dependence of each of the three derivative peaks for a given aromatic residue was further supported through direct experimental observations as well (Fig 5.1 & Table 5.1).

Monitoring temperature dependent peak shifts of phenylalanine and tyrosine side chains in a variety of proteins, a general trend of reduced slopes (with a few exceptions) for globular proteins with partially to fully buried residues were observed (Fig 5.10). This again supports the proposed correlation between the magnitude of the observed temperature dependent peak shifts and extent of the hydration of buried aromatic side chains.


Phe / Peak 1
N-acetyl-L-phenylalanine-ethyl ester
Protein A
RNase-T1
Azurin
HSA
L-Asparaginase
Shift in Peak Position (nm)
Temperature (°C)


Phe / Peak 2
N-acetyl-L-Phenylalanine-ethyl ester
Protein A
RNase-T1
Azurin
HSA
L-Asparaginase
Shift in Peak Position (nm)
Temperature (°C)


Phe / Peak 3
N-acetyl-L-phenylalanine-ethyl ester
HSA
Shift in Peak Position (nm)
Temperature (°C)


Tyr / Peak 1
N-acetyl-L-Tyrosine-ethyl ester
RNase-T1
Azurin
HSA
L-Asparaginase
Shift in Peak Position (nm)
Temperature (°C)

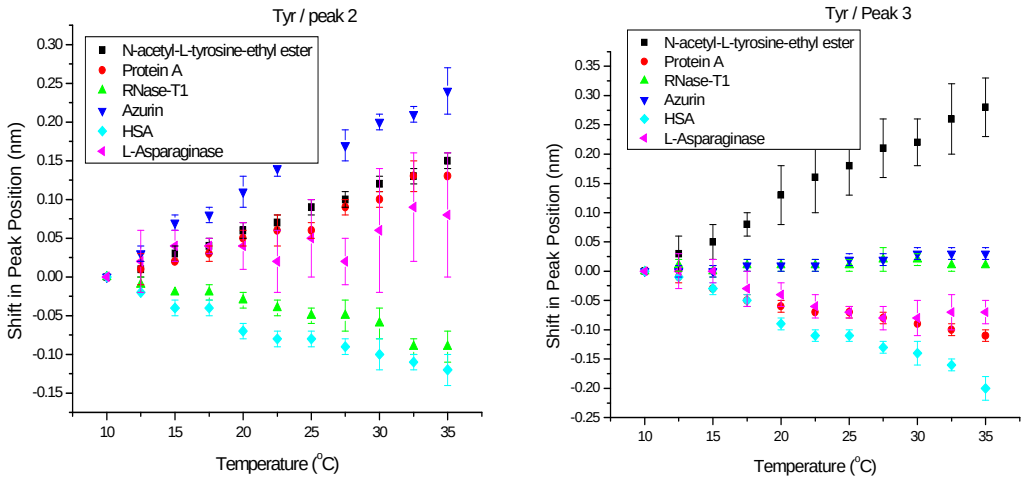


Figure 5.10.Temperature dependent derivative absorbance peak shifts of tyrosine and phenylalanine residues in model peptides and proteins over the temperature range of 10-35 °C. Model aromatic amino acids used were N-acetyl-L-tyrosine ethyl ester and N-acetyl-L-phenylalanine ethyl ester. (n=3)

The resolutions obtained from the 2nd derivative UV peak shifts of phenylalanine and tyrosine analogs, however, are significantly less than those of tryptophan. This is understandable considering the much larger extinction coefficient of tryptophan (ε_{280}=5540) compared to tyrosine (ε_{280}=1480) and phenylalanine (ε_{258}=197) which results in a better signal to noise ratio for the former[6]. Considering the local restrictive environmental effects (i.e., hydrogen bonding, π-π interaction, etc.) as discussed previously, the presence of a large number of phenylalanine and tyrosine residues in proteins (Table 5.3) could also contribute by producing a high degree of heterogeneity. The abundance of phenylalanine residues, however, is expected to be less problematic since they tend to be deeply buried in protein cores, maintaining dominantly apolar interactions with surrounding apolar side chains. Tyrosine residues, on the other hand, tend to be more dispersed throughout protein three dimensional structures, resulting in more heterogeneous microenvironments when present in large numbers. In addition, the ability of the phenol side chain of tyrosine residues to participate in a variety of different interactions with surrounding side chains adds to the complexity of their spectroscopic signals.

5.4. Discussion

Proteins are highly dynamic molecules residing in aqueous environments in which both their static and dynamic properties as well as the nature of the surrounding environment affects their structure and function. Central questions as to how protein dynamics bridge their structure and function, and how environmental

factors affect protein dynamical-structural-functional properties have engendered great interest. Various protein functions such as substrate binding, product release, regulation, allosteric behavior, as well as contractile and motor functions are heavily controlled by collective motions within proteins[36]. Fluctuations of the surrounding solvent molecules have been shown to control protein dynamic motions, engendering the use of the term "slaving" to describe the significance of the dominant role of solvent in controlling protein dynamics and function[2].

Here, we have investigated the applicability of the temperature dependent 2nd derivative ultraviolet absorption spectroscopy of the aromatic amino acids to study protein dynamics. This technique has been effectively employed for more than two decades to study protein conformational changes based on protein static properties (e.g., thermal unfolding). In previous applications of this technique, the origin of the temperature-dependent UV peak shifts in the protein pre-transition regions over which proteins retain their native structure has received little or no attention. Therefore, the scope of this work was to investigate the physical origin and application of such pre-transition alterations.

This work shows that plots of the temperature dependent 2nd derivative peak position of the aromatic residues have measurable quasi-linear slopes below the protein's melting temperature. Both computational and experimental analyses suggest that temperature dependent alterations of the solvent's physical properties and their effects on solvent-solute interactions are the source of the spectral changes observed. Extensive previous studies have explored the nature of the solvent-solute interactions

responsible for alterations of solute electronic spectra in which dipolar, van der Waals, and short range specific interactions (e.g., hydrogen bonding) have been shown to be major contributors[22]. Considering the temperature dependence of such interactions, they were investigated as potential sources of the spectral changes observed.

Our computational and experimental results suggest that the spectral alterations of the aromatic residues are sensitive to temperature dependent solvent dielectric changes. Both the absolute value and magnitude of solvent dielectric alterations appear to be important in producing the shifts observed. The largest temperature dependent alterations were observed in water which possesses the highest dielectric constant value and the greatest extent of temperature dependent dielectric alterations. The magnitude of the experimentally observed peak shifts was reduced significantly upon addition of organic solvents to water due to lowering of the solvent dielectric constant.

Based on the origin of spectral alterations observed, we proposed the applicability of such information to the study of protein dynamics. Previous studies have shown solvent penetration (due to protein global dynamic motions) to be responsible for the high polarizabilities and large dielectric constant values observed in protein interiors[31]. Considering the sensitivity of the observed temperature dependent peak shifts of the aromatic residues to the dielectric properties of the surrounding solvent, the utility of the 2nd derivative UV absorption method was examined as a qualitative tool to probe protein dynamical motions that at least partially control the hydration of buried aromatic residues in protein interiors. The

utility of this technique appears plausible when studying model proteins with extensive buried tryptophan residues. Model proteins showed considerably reduced temperature dependent slopes compared to the tryptophan analog and linear peptides with exposed indole moieties.

The sensitivity of the temperature dependent peak shifts to local dynamic motions of aromatic side chains was also investigated in which restrictive forces such as hydrogen bonding reduced the magnitude of the peak shifts. Based on this, the effect of protein local dynamic motions on temperature dependent peak shifts was also evaluated. Complimentary fluorescence acrylamide quenching studies showed agreement in dynamic behavior with the UV technique for RNase-T1 and protein A. Opposite trends, however, were obtained when comparing RNase-T1 and azurin which could be explained by the sensitivity of the UV technique to local rigidifying effects (i.e., hydrogen bonding) absent in the fluorescence technique.

Multiple peptides containing phenylalanine and tyrosine residues were also analyzed to demonstrate the advantages of the UV absorbance approach over fluorescence techniques in their ability to monitor spectral alterations of phenylalanine and tyrosine residues. This data exhibited slopes comparable to the model aromatic residues further supporting the applicability of this technique in the context of all three aromatic residues. Although the results with proteins supported the overall expected trends of reduced slopes compared to amino acid analogs, resolution of the data were not as good as those of tryptophan. This is presumably due to the much lower extinction coefficients of these residues resulting in poor signal to

noise ratios. Furthermore, the presence of large number of these residues that collectively contribute to the peak shifts observed provides another source of heterogeneity in the data. Future work is intended to investigate small synthetic proteins with single phenylalanine and/or tyrosine residues to shed more light on their contribution to the temperature dependent UV peak shift dynamics approach.

5.5. Conclusion

We have investigated the origin and application of the temperature-dependent peak shifts of the three aromatic residues in protein pre-transition regions. We conclude that such alterations are sensitive to temperature dependent changes of solvent physical properties and therefore can be used as a qualitative tool to study protein dynamic motions that produce alterations in the hydration of buried aromatic residues in protein interiors. We have shown that such alterations are sensitive to both protein global dynamic motions as well as side chain local mobility.

5.6. Bibliography

(1) Lindorff-Larsen K, Best RB, DePristo MA, Dobson CM, Vendruscolo M. Simultaneous determination of protein structure and dynamics. Nature (London, United Kingdom) (2005), 433, 128-132.

(2) Fenimore PW, Frauenfelder H, McMahon BH, Parak FG. Slaving: Solvent fluctuations dominate protein dynamics and functions. Proc. Nat. Acad. Sci. U.S.A. (2002), 99, 16047-51.

(3) Beaven GH, Holiday ER. Ultraviolet absorption spectra of proteins and amino acids. Adv. Protein. Chem. (Academic Press Inc., New York, N.Y.) (1952), 7, 319-86.

(4) Lucas LH, Ersoy BA, Kueltzo LA, Joshi SB, Brandau DT, Thyagarajapuram N, Peek LJ, Middaugh CR. Probing protein structure and dynamics by second-derivative ultraviolet absorption analysis of cation $-\pi$ interactions. Protein Sci. (2006), 15, 2228-2243.

(5) Mach H, Thomson JA, Middaugh CR, Lewis RV. Examination of phenylalanine microenvironments in proteins by second-derivative absorption spectroscopy. Arch. Biochem Biophys. (1991), 287, 33-40.

(6) Kueltzo LA, Middaugh CR. Ultraviolet absorption spectroscopy. Biotechnology: Pharmaceutical Aspects (2005), 3, 1-25.

(7) Kueltzo LA, Ersoy B, Ralston JP, Middaugh CR. Derivative absorbance spectroscopy and protein phase diagrams as tools for comprehensive protein characterization: A bGCSF case study. J. Pharm. Sci. (2003), 92, 1805-1820.

(8) Lange R, Balny C. UV-visible derivative spectroscopy under high pressure. Biophys. Acta, Protein Struct. Mol. Enzymol. (2002), 1595, 80-93.

(9) SYBYL 8.0, The Tripos Associates, St. Louis, MO, 2007.

(10) Clark M, Cramer RD, Van Opdenbosch N. Validation of the general purpose Tripos 5.2 force field. J. Comput. Chem. (1989), 10, 982-1012.

(11) Gasteiger J, Marsili M. Iterative partial equalization of orbital electronegativity: A rapid access to atomic charges. Tetrahedron (1980), 36, 3219-22.

(12) Gaussian 03, Revision C.02. Frisch GW, Trucks HB, Schlegel GE, Scuseria MA, Robb JR, Cheeseman JA, Montgomery T, Vreven KN, Kudin JC, Burant

JM, Millam SS, Iyengar J, Tomasi V, Barone B, Mennucci M, Cossi G, Scalmani N, Rega GA, Petersson H, Nakatsuji M, Hada M, Ehara,K, Toyota,R, Fukuda J, Hasegawa M, Ishida T, Nakajima Y, Honda O, Kitao H, Nakai M, Klene X, Li JE, Knox HP, Hratchian JB, Cross V, Bakken C, Adamo J, Jaramillo R, Gomperts RE, Stratmann O, Yazyev AJ, Austin R, Cammi C, Pomelli JW, Ochterski PY, Ayala K, Morokuma G A, Voth P, Salvador JJ, Dannenberg VG, Zakrzewski S, Dapprich AD, Daniels MC, Strain O, Farkas DK, Malick AD, Rabuck K, Raghavachari JB, Foresman JV, Ortiz Q, Cui AG, Baboul S, Clifford J, Cioslowski BB, Stefanov G, Liu A, Liashenko P, Piskorz I, Komaromi RL, Martin DJ, Fox T, Keith MA, Al-Laham CY, Peng A, Nanayakkara M, Challacombe PM, Gill B, Johnson W, Chen MW, Wong C, Pople JA; Gaussian, Inc; Wallingford CT; 2004.

(13) Becke AD. Density-functional thermochemistry III. The role of exact exchange. J. Chem. Phys. (1993), 98, 5648-52.

(14) Lee C, Yang W, Parr RG. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. Physical Review B: Condens. Matter (1988), 37, 785-9.

(15) Ditchfield R, Hehre WJ, Pople JA. Self-consistent molecular-orbital methods IX. Extended Gaussian-type basis for molecular-orbital studies of organic molecules. J. Chem. Phys. (1971), 54, 724-8.

(16) Frisch MJ; Pople JA, Binkley JS. Self-consistent molecular orbital methods 25. Supplementary functions for Gaussian basis sets. J. Chem. Phys. (1984), 80, 3265-9.

(17) Stratmann RE, Scuseria GE, Frisch MJ. An efficient implementation of time-dependent density-functional theory for the calculation of excitation energies of large molecules. J. Chem. Phys. (1998), 109, 8218-8224.

(18) Tomasi J, Mennucci B, Cances E. The IEF version of the PCM solvation method: An overview of a new method addressed to study molecular solutes at the QM ab initio level. THEOCHEM (1999), 464, 211-226.

(19) Fernandez DP, Goodwin AR. Lemmon EW, Sengers JM. Levelt WR. A formulation for the static permittivity of water and steam at temperatures from 238 K to 873 K at pressures up to 1200 MPa, including derivatives and Debye-Hueckel coefficients. J. Phys. Chem. Ref. Data (1997), 26, 1125-1166.

(20) Lide DR; Kehiaian HV. Handbook of Thermophysical and Thermochemical Data. CRC Press, Boca Raton, FL (1994), 518 pp.

(21) Kundt A. “Ann. Der physic” 1878, vol. 4, p. 34.

(22) Basu S. Theory of solvent effects on molecular electronic spectra. Adv. Quantum Chem. (Per Olov Lowdin, editor. Academic) (1964), 1, 145-69.

(23) Grimme S, Izgorodina EI. Calculation of 0–0 excitation energies of organic molecules by CIS(D) quantum chemical methods. Chem. Phys. (2004), 305, 223-230.

(24) Anslyn EV, Dougherty DA. Modern Physical Organic Chemistry. Wilsted&Taylor Publishing Services (2006), 1099 pp.

(25) Akerlof G. Dielectric constant of some organic solvent-water mixtures at various temperatures. J. Am. Chem. Soc. (1932), 54, 4125-39.

(26) Casteel JF, Sears PG. Dielectric constants, viscosities, and related physical properties of 10 liquid sulfoxides and sulfones at several temperatures. J. Chem. Eng. Data (1974), 19, 196-200.

(27) Usacheva TM, Lifanova NV, Zhuravlev VI, Novozhilov AA, Matveev VK. Calculations of the structure parameters of weakly associated liquids: acetonitrile. Zhurnal Fizicheskoi Khimii (2000), 74, 1962-1966.

(28) Mataga N, Kubota T. Molecular Interactions and Electronic Spectra, Marcel Dekker, Inc. (New York, NY) (1970), 504 pp.

(29) Ito M. Ultraviolet absorption study of the molecular association of phenols. J. Mol. Spectrosc. (1960), 4, 125-43.

(30) Shapovalov MV, Dunbrack RL. Statistical and conformational analysis of the electron density of protein side chains. Proteins: Structure, Function, and Bioinformatics (2007), 66, 279-303.

(31) Dwyer JJ, Gittis AG, Karp DA, Lattman EE, Spencer DS, Stites WE, Garcia-Moreno EB. High apparent dielectric constants in the interior of a protein reflect water penetration. Biophys. J. (2000), 79, 1610-1620.

(32) Eftink MR, Ghiron CA. Exposure of tryptophanyl residues in proteins. Quantitative determination by fluorescence quenching studies. Biochemistry (1976), 15, 672-80.

(33) Turoverov KK, Kuznetsova IM, Zaitsev VN. The environment of the tryptophan residue in Pseudomonas aeruginosa azurin and its fluorescence properties. Biophys. Chem. (1985), 23, 79-89.

(34) Irie M. Studies on the state of tryptophan residue in ribonuclease T1 and carboxymethyl ribonuclease T1. J. Biochem. (1970), 68, 31-7.

(35) Rueda M, Ferrer-Costa C, Meyer T, Perez A, Camps J, Hospital A, Gelpi JL; Orozco M. A consensus view of protein dynamics. Proc. Nat. Acad. Sci. U.S.A. (2007), 104, 796-801.

(36) Berendsen HJ, Hayward S. Collective protein dynamics in relation to function. Curr. Opin. Struct. Biol. (2000), 10, 165-169.

Chapter 6

Conclusions & Future Work

6.1. Overview

Protein instability is a particularly important issue in the pharmaceutical industry as the number of therapeutic protein products continue to increase[1]. The marginal stability associated with many macromolecular-based agents introduces significant difficulties in development of proteins and their complexes as safe and efficacious therapeutics. Macromolecules, due to their large size and complexity are prone to a variety of physical and chemical degradations during their manufacturing, storage, and shipment. This often results in loss of their biological activity and altered immunogenicity, jeopardizing their potential benefits to human health[2].

Given this high degree of structural complexity and associated degradation pathways, an optimized formulation that retains a protein's stability and activity over its entire shelf life is crucial. The FDA requires that the stability and activity of potential protein therapeutics be demonstrated in real time under the proposed labeled storage conditions. Considering the desired long shelf life of therapeutic proteins (18-24 months) [3], however, these types of studies demand that significant resources be dedicated to them for extended periods of time. One time and cost effective approach involves conducting accelerated stability studies to identify potential formulations with properties that lead to enhanced stability under more moderate storage conditions. Such accelerated studies are typically conducted under a variety of different stressed conditions (e.g., elevated temperatures, suboptimal pH, high or low ionic strengths, etc.) to induce protein degradation over a short period of time.

Herein, a rapid and systematic approach has been employed for biophysical characterization of complex macromolecular systems under accelerated degradation conditions. The "Empirical phase diagram" (EPD) approach[4] employed integrates a large library of data obtained from a variety of experimental techniques (e.g., circular dichroism, intrinsic and extrinsic fluorescence, static and dynamic light scattering, etc.) into an easy-to-visualize color coded map to provide a global picture of protein structural response to a wide range of stressed solution conditions.

6.2. Chapter Summaries & Future Works

6.2.1. Chapter 2

In chapter 2, the EPD approach was utilized for the characterization of recombinant vault particles[5-8], a large and complex ribonucleoprotein. With the current interest in employing recombinant vaults as nanocapsules for the delivery of biomolecules, characterization of the physical stability of vault particles under a variety of solution conditions has the potential to provide important information concerning vault's structural integrity and its potential use as drug delivery vehicles. Our studies identified ten different conformational states of the vaults over the pH and temperature range studied with the most stable region found at pH 6-8 below 40 °C and the least stable condition at pH 4-6 above 60 °C. A unique intermediate molten globule-like state was also identified at pH 6 and ~55 °C. EM imaging showed the opening of intact vaults into flower-like structures when transitioning from neutral to acidic pH supporting the potential utility of these particles as

nanocapsules for drug delivery since one mechanism by which therapeutic agents entrapped in vaults could be released is through an opening of the intact vault structure. *In vivo*, naturally occurring vault nanocapsules posses a dynamic structure and appear to be highly interactive with their surrounding environment[9]. Therefore, future work should focus on construction of an EPD based on techniques such as hydrogen/deuterium exchange, red edge shift spectroscopy, time correlated single photon anisotropy measurements, pressure perturbation calorimetry, and high resolution ultrasonic spectroscopy that are more sensitive to the dynamic behavior of vault internal structure.

6.2.2. Chapter 3

In chapter 3, the EPD method was utilized as a rapid and systematic approach to pre-formulation studies of a complex fusion cytotoxin composed of mutated versions of Interleukin-13 and Pseudomonas Exotoxin A[10-13] with the ultimate goal of developing a stable liquid formulation. The EPD approach identified apparent phase boundaries over which the protein's degradation pathways were accelerated. One such boundary (pH 5, 45 °C) over which the fusion protein showed extensive aggregation was used as a basis for the development of a high throughput excipient screening assay employing a kinetic-based turbidity assay. The ability of members of a large library of Generally Regarded as Safe (GRAS) excipients to inhibit aggregation of the fusion protein was examined. Polyanions (dextran sulfate), polysaccharides (sucrose, dextrose, lactose, and trehalose), polyols (sorbitol and

glycerol), salt (sodium citrate), and detergents (Tween 20 and 80) all inhibited protein aggregation by ~ 25 to 80%. Combinations of more promising stabilizers were tested to see if there was an additive or synergistic effect of these compounds on the inhibition of protein aggregation. The combination of sodium citrate (≥ 0.1M) with sugars or polyols (≥10% w/v) inhibited the aggregation of the fusion protein almost completely. To further explore whether the inhibitors of aggregation were also stabilizing the structure of the protein, the tertiary structural stability of the fusion protein was examined in the presence of the aforementioned combinations. Combinations of sodium citrate with dextrose, trehalose, or sorbitol all delayed the unfolding of the protein by shifting the T_m values to higher temperatures by ~ 3 to 6 °C. They were therefore identified as optimal formulation additives for this particular fusion cytotoxin.

In addition to the utility of the EPD approach in providing a thorough in vitro characterization of this fusion complex, comparison of the fusion protein's thermal stability to the work of others involving the stability of the individual domains suggested that they are strongly correlated and that the toxin domain comprising ~ 84% of the total mass contributes significantly to the fusion protein structural alterations. The fusion protein was shown to undergo an acid-triggered α-helix disruptive event similar to its toxin domain, further adapting an unfolded state with strong hydrophobicity at or below pH 4, essential for insertion of fusion cytotoxins into endosomal membranes[14]. These studies showed interaction of this protein with polyanions in agreement with the presence of such interactions between the toxin

domain and anionic lipids of endosomal membranes, as well as IL-13 with its cell surface receptor, IL13Rα2. Therefore, our studies show that the thermal stability of a complex fusion protein can at least in some cases be understood in terms of the behavior of its individual domains.

Future work should be directed towards further optimization of the final concentrations and combinations of more promising stabilizers identified herein to provide a suitable liquid formulation compatible with the protein's therapeutic application and its route of administration. Future studies should be focused on evaluating and optimizing parameters such as ionic strength, osmolarity, and viscosity of the protein formulation and its integrity over its shelf life upon addition of such excipients. In addition, a more thorough thermal characterization of the individual domains should help to better understand the effects and contributions of the individual domains to the overall thermal stability of the fusion complex.

6.2.3. Chapter 4

In chapter 4, the EPD approach was used for thermal characterization of three reassortant human-bovine rotavirus[15] strains with the ultimate goal of developing a thermostable vaccine formulation resistant to temperature excursions. The biophysical studies employed herein have the potential to provide insight into the mechanisms of inactivation as a function of a variety of environmental conditions. This provides an opportunity for a more rational and systematic formulation of such vaccines based on inhibition of such degradation processes. Thermostable vaccines

has major benefits including reducing vaccine wastage, increasing vaccine effectiveness, reducing dependency on cold chain supplies and equipment, and enabling vaccine delivery to remote populations[16].

Examination of the EPD of the three rotavirus strains revealed distinct conformational phases across the pH and temperature range examined. The spectroscopic data suggested that all three strains were stable across the pH range of 5 to 8 below 40 °C. This was also supported by EM image analysis of the virus strains in which the presence of intact, triple layer infectious particles were observed. All three strains showed a region of minimum stability with perturbed secondary and tertiary structures along with the presence of aggregates across the pH range of 5 to 8 at elevated temperatures above 55 to 60 °C. This is supported by EM analysis at 70 °C which exhibits extensive aggregation. The onset temperatures at which the structural alterations occurred was lowered in a stepwise manner as a function of increasing pH for all three strains suggesting reduced virus thermal stability as a function of increasing pH. All strains exhibited intermediate phases across the pH range of 5 to 8 in the region of ~ 40-60 °C with quite subtle secondary or tertiary structural perturbations and the presence of micro-aggregates as manifested by static light scattering data. EM analysis showed larger populations of virus particles with partially disrupted outer-layers over these intermediate regions. The empirical phase diagram approach did not identify any major thermal stability differences among the three strains encouraging the potential use of all three in a single rotavirus vaccine formulation.

We further employed complementary *in vitro* activity based studies of the distinct conformational states identified by the spectroscopic techniques to provide a more meaningful interpretation of such complex moieties and their structural stability as it is related to their biological activity. These studies showed that secondary structural alterations of the virus complex as well as their aggregation resulted in loss of virus activity. Virus activity was shown to be highly pH and temperature dependent and varied among the three strains.

Previous Studies of a variety of vaccine candidates have shown that identification of regions of marginal stability in the EPD serve as a basis for the development of high throughput screening assays and identification of solution stabilizers[2,17,18]. Therefore, future work should be directed towards developing screening assays and the consequent identification of optimized liquid formulations containing individual as well as combinations of all three strains in a single formulation.

Comparing to the other two macromolecular systems discussed in Chapters 2 and 3, the biophysical data obtained for the virus systems were found to be harder to interpret due to the more complex nature of this macromolecular system. The results, however, still can be used to detect stability changes due to the distinctive and highly reproducible signals produced. Rotavirus exhibited well defined sizes and shapes employing dynamic light scattering. In addition, biophysical techniques such as CD and fluorescence generated reproducible results across the pH and temperature range

examined. A further significance of such an approach was indicated by analysis of virus activity and its relationship to the structural alterations observed herein.

It should be emphasized that although such accelerated stability studies provide a rapid and systematic approach for a thorough biophysical characterization of macromolecular complexes under artificially induced degradation conditions, real time stability must be directly examined in actual long-term stability studies. This is due to the possible presence of different degradation mechanisms under accelerated versus more moderate storage conditions[19].

6.2.4. Chapter 5

In addition to the aforementioned structural alterations and aggregation (i.e. time averaged or static properties) affecting protein stability, protein dynamics have been shown to play a role in controlling their stability and function[20]. Studies involving protein dynamics suggest that analysis of protein flexibility along with their static properties is necessary for successful characterization of protein structural stability. It follows that the availability of techniques that detect protein structural fluctuations is crucial for optimal characterization of macromolecular-based therapeutics. In Chapters 2-4, we have presented the EPD approach based on a variety of biophysical techniques that are sensitive to conformational alterations and aggregation (i.e., static properties) of macromolecules. More recent efforts have employed data from spectroscopic techniques sensitive to protein internal motions for construction of empirical phase diagrams that should be more sensitive to transitions

involving alterations of protein motions[21]. Such a dynamic-based approach revealed that EPDs constructed from such methods provide information above and beyond that obtained by the static approach. As part of efforts to utilize available spectroscopic tools for a better understanding of protein dynamic properties, for the first time we here examined the utility of temperature dependent UV absorption spectroscopy of aromatic residues in protein pre-transition regions to show that information obtained from such analyses hold valuable information regarding protein dynamic fluctuations (Chapter 5). Our studies showed that plots of the temperature dependent 2nd derivative peak positions of the aromatic residues have measurable, quasi-linear slopes below the protein's melting temperature. Our computational and experimental results suggested that the spectral alterations of the aromatic residues are sensitive to temperature dependent solvent dielectric changes in which both the absolute value and the magnitude of solvent dielectric alterations appear to be important in producing the shifts observed. The sensitivity of the temperature dependent peak shifts to local dynamic motions of aromatic side chains was also investigated in which restrictive forces such as hydrogen bonding were shown to affect the magnitude of the peak shifts.

Previous studies have shown solvent penetration (due to protein global dynamic motions) to be responsible for the high polarizabilities and large dielectric constant values observed in protein interiors[22]. Considering the sensitivity of the observed temperature dependent peak shifts of the aromatic side chains to the dielectric properties of the surrounding solvent, the utility of the 2nd derivative UV

absorption method was examined as a qualitative tool to probe protein dynamic motions that at least partially control the hydration of buried aromatic residues in protein interiors. The utility of this technique appeared promising when studying model proteins with extensively buried tryptophan residues. Model proteins showed considerably reduced temperature dependent slopes compared to the tryptophan analog and linear peptides with exposed indole moieties.

Multiple peptides and proteins containing phenylalanine and tyrosine residues were also analyzed to demonstrate the advantages of the UV absorbance approach over fluorescence techniques in their ability to monitor spectral alterations of phenylalanine and tyrosine residues. Although the results with proteins supported the overall expected trends of reduced slopes compared to amino acid analogs, resolution of the data was not as good as those of tryptophan, presumably due to the much lower extinction coefficients of these residues which resulted in poor signal to noise ratios[23]. Furthermore, the presence of large numbers of these residues that collectively contribute to the peak shifts observed provided another source of heterogeneity in the data. Future work is intended to investigate small synthetic proteins with single phenylalanine and/or tyrosine residues to shed more light on their contribution to the temperature dependent UV peak shift dynamics approach.

In summary, we have employed a rapid and systematic approach employing accelerated degradation conditions referred to as the “Empirical phase diagram” (EPD) approach to characterize the structural stability of three complex macromolecular systems; a large recombinant ribonucleoprotein (chapter 2), a novel multi-domain

fusion cytotoxin (chapter 3), and a live-attenuated double stranded RNA virus (chapter 4). This approach utilizes advanced algebraic procedures to integrate a large library of data obtained from a variety of biophysical techniques into an easy-to-interpret color coded map to provide a global picture of protein structural behavior under a wide range of solution conditions. We demonstrate the ability of this approach in providing a thorough *in vitro* characterization of large and complex macromolecular systems. We further show the utility of the EPD approach in designing high throughput screening assays for identification of solution stabilizers and its use in interpretation of the biological activity and function of these systems *in vivo*. In contrast to simple macromolecules, the more complex systems such as those examined here, the measured stability is presumably the sum of all components. This can make interpretations at the molecular level somewhat difficult. We show that employing the EPD approach, the structural stability of the complex macromolecular systems can at least partially be explained in terms of the behavior of their individual domains and components.

Empirical phase diagrams constructed based on techniques sensitive to transitions due to alterations of protein motions (i.e., dynamic-based EPDs) have been shown to provide information above and beyond that obtained by the static approach. Therefore, integration of techniques that detect extensive and subtle conformational alterations of proteins, as well as structural fluctuations into an empirical phase diagram is beneficial. Herein, we have for the first time examined the temperature dependent UV absorption spectroscopy of aromatic residues in protein pre-transition

regions (chapter 5). We show that such analyses hold valuable information regarding protein dynamic fluctuations which can essentially be integrated into dynamic-based EPDs to study protein dynamics and its effect on stability.

6.3. Bibliography

(1) Chi EY, Krishnan S, Randolph TW, Carpenter JF. Physical Stability of Proteins in Aqueous Solution: Mechanism and driving forces in nonnative protein aggregation. Pharm. Res. (2003), 20, 1325-1336.

(2) Ausar SF, Espina M, Brock J, Thyagarayapuran N, Repetto R, Khandke L, Middaugh CR. High-throughput screening of stabilizers for respiratory syncytial virus: Identification of stabilizers and their effects on the conformational thermostability of viral particles. Human Vaccines (2007), 3, 94-103.

(3) Randolph TW, Carpenter JF. Engineering challenges of protein formulations. AIChE J. (2007), 53, 1902-1907.

(4) Kueltzo LA, Ersoy B, Ralston JP, Middaugh CR. Derivative absorbance spectroscopy and protein phase diagrams as tools for comprehensive protein characterization: A bGCSF case study. J. Pharm. Sci. (2003), 92, 1805-1820.

(5) Kedersha NL, Rome LH. Isolation and characterization of a novel ribonucleoprotein particle: Large structures contain a single species of small RNA. J. Cell. Biol. (1986), 103, 699-709.

(6) Suprenant KA. Vault ribonucleoprotein particles: Sarcophagi, gondolas, or safety deposit boxes? Biochemistry (2002), 41, 14447-14454.

(7) Kickhoefer VA, Garcia Y, Mikyas Y, Johansson E, Zhou JC, Raval-Fernandes S, Minoofar P, Zink JI, Dunn B, Stewart PL, Rome LH. Engineering of vault nanocapsules with enzymatic and fluorescent properties. Proc. Nat. Acad. Sci. U.S.A. (2005), 102, 4348-4352.

(8) Anderson DH, Kickhoefer VA, Sievers SA, Rome LH, Eisenberg D. Draft crystal structure of the vault shell at 9-A° resolution. PLoS biology (2007), 5, e318

(9) Poderycki MJ, Kickhoefer VA, Kaddis CS, Raval-Fernandes S, Johansson E, Zink JI, Loo JA, Rome LH. The vault exterior shell is a dynamic structure that allows incorporation of vault-associated proteins into its interior. Biochemistry (2006), 45, 12184-12193.

(10) Debinski W. Molecular targeting of brain tumors with cytotoxins: Novel bacterial toxin-containing anti-brain tumor therapeutics. Cellular and Molecular Mechanisms of Toxin Action (2002), 4, 222-246.

(11) Husain SR, Puri RK. Interleukin-13 receptor-directed cytotoxin for malignant glioma therapy: from bench to bedside. Journal of neuro-oncology (2003), 65, 37-48.

(12) Debinski W. Anti-brain tumor cytotoxins. Science & Medicine (Philadelphia) (1998), 5, 36-42.

(13) Debinski W. Recombinant cytotoxins specific for cancer cells. Annals of the New York Academy of Sciences (1999), 886, 297-299.

(14) Mere J, Morlon-Guyot J, Bonhoure A, Chiche L, Beaumelle B. Acid-triggered membrane insertion of Pseudomonas Exotoxin A involves an original mechanism based on pH-regulated tryptophan exposure. J. Biol. Chem. (2005), 280, 21194-21201.

(15) Fields BN; Knipe DM, Howley PM. Fields Virology: 4th ed, volume 2, Lippincott Williams&Wilkins (2001), 1445 pp.

(16) Chen D, Kristensen D. Opportunities and challenges of developing thermostable vaccines. Expert Review of Vaccines (2009), 8, 547-557.

(17) Peek LJ, Brandau DT, Jones LS, Joshi SB, Middaugh CR. A systematic approach to stabilizing EBA-175 RII-NG for use as a malaria vaccine. Vaccine (2006), 24, 5839-5851.

(18) Kissmann JM, Ausar SF, Rudolph A, Braun C, Cape SP, Sievers RE, Federspiel MJ, Joshi SB, Middaugh CR. Stabilization of measles virus for vaccine formulation. Human vaccines (2008), 4, 350-9.

(19) Yoshioka S, Stella VJ. Stability of drugs and dosage forms, Kluwer Academic/Plenum Publishers (New York, NY) (2000), 187-199.

(20) Kamerzell TJ, Middaugh CR. The complex inter-relationships between protein flexibility and stability. J. Pharm. Sci. (2008), 97, 3494-3517.

(21) Ramsey JD, Gill ML, Kamerzell TJ, Price ES, Joshi SB, Bishop SM, Oliver CN, Middaugh CR. Using empirical phase diagrams to understand the role of intramolecular dynamics in immunoglobulin G stability. J. Pharm. Sci. (2009), 98, 2432-2447.

(22) Dwyer JJ, Gittis AG, Karp DA, Lattman EE, Spencer DS, Stites WE, Garcia-Moreno EB. High apparent dielectric constants in the interior of a protein reflect water penetration. Biophys. J. (2000), 79, 1610-1620.

(23) Kueltzo LA, Middaugh CR. Ultraviolet absorption spectroscopy. Biotechnology: Pharmaceutical Aspects (2005), 3, 1-25.