1997

Monitoring protein refolding using diverse separation and spectroscopy techniques

Pamela Kay Jensen
Iowa State University

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

Monitoring protein refolding using diverse separation and spectroscopy techniques

by

Pamela Kay Jensen

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Analytical Chemistry

Major Professor: Cheng S. Lee

Iowa State University

Ames, Iowa

1997

Graduate College
Iowa State University

This is to certify that the Doctoral dissertation of

Pamela Kay Jensen

has met the dissertation requirements of Iowa State University

Signature was redacted for privacy.

Major Professor

Signature was redacted for privacy.

For the Major Program

Signature was redacted for privacy.

For the Graduate College

This thesis is dedicated to my husband, Brett Jensen. Without his love and encouragement, this thesis would not have been possible.

TABLE OF CONTENTS

ABSTRACT

Rapid growth in the biotechnology industry has led to a dramatic increase in attention to the protein folding problem. Understanding of protein folding is essential to the production of biopharmaceuticals since commercial production of recombinant proteins often requires a protein refolding process to recover high yields.

Capillary zone electrophoresis (CZE) equipped with laser induced fluorescence detection (LIFD) is developed as a tool for monitoring the refolding of a model protein, phage P22 tailspike endorhamnosidase. Intermediates on the folding pathway are separated in CZE and monitored by their intrinsic tryptophan fluorescence. Monitoring refolding of tailspike by traditional methods, sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and fluorescence spectrophotometry, are used to confirm the refolding kinetics and yield as measured by CZE. The agreement between these methods shows CZE to be an effective technique for simple, rapid detection of changes in protein conformation.

Effects of temperature on the refolding kinetics and yield of tailspike are investigated using CZE-LIFD. Refolding kinetics are extremely slowed at 0°C and results in a loss of the protrimer in the refolding pathway. Results obtained from CZE-LIFD and polyacrylamide gel electrophoresis indicate the loss of folding intermediates through the adsorption of polypeptide chains onto the wall of the refolding vial.

By increasing the reaction temperature at different stages of tailspike refolding, structured monomeric species are identified as the critical intermediates at the junction between productive and aggregation pathways of tailspike protein. Optimal refolding kinetics and yields of tailspike endorhamnosidase can be achieved by increasing the refolding temperature from 10 to 35°C only after passage of the stage for accumulation of the thermolabile intermediate.

Capillary isoelectric focusing-electrospray ionization mass spectrometry (CIEF-ESIMS) is developed as another technique for monitoring protein folding intermediates. After the initiation of the refolding process, the reduced and denatured disulfide bonded protein, ribonuclease A (RNase A), is blocked at different refolding stages by alkylation of free cysteines with iodoacetate. This alkylation reaction results in the introduction of charge (-1) and mass (58) differences for each alkylation site, providing the means for predictable separation and direct identification of intermediates by CIEF-ESIMS.

CHAPTER 1. GENERAL INTRODUCTION

Dissertation Organization

The first chapter of this dissertation is the general introduction, designed to offer a general background and a review of literature related to this research. The following three chapters each consist of an individual paper accepted or submitted for publication, complete with references formatted for the particular journal to which it was submitted. The final chapter offers a general conclusion of this work. References for the general introduction immediately follow the last chapter.

Refolding of Recombinant Proteins

The development of recombinant protein technology brought not only promises of unlimited supplies of therapeutic proteins, but also problems of recovering these proteins in their native and biologically active conformation. Frequently, newly synthesized polypeptide chains from the over-expression of cloned genes are unable to fold correctly in the host cell, resulting in the formation of insoluble aggregates or inclusion bodies[1]. The formation of aggregates and the recovery of active protein has become a major challenge for the biotechnology industry and in biomedical research. However, the use of *in vitro* refolding techniques can result in recovery of native active protein from these inclusion bodies. Considerable study is essential in improving refolding techniques for recovery of active protein in high yields. Studies into the protein folding problem

also help in elucidating the "second half of the genetic code" desired for the prediction of structural information from amino acid sequence. Experimental data on folding mechanisms is necessary to determine optimal conditions for protein recovery and to identify the critical intermediates at the junction between productive folding and misfolding or aggregation.

Analytical Techniques for Monitoring Protein Refolding

Techniques used for characterizing protein folding reactions can be grouped into two main classes, those based on spectroscopy or those involving separations. Spectroscopy techniques most commonly used to study proteins include absorbance, fluorescence, and circular dichroism[2-4]. However, these techniques depend on the measurement of some average physical properties that are sensitive to changes in the protein structure and only give information on gross conformational changes that occur. More sophisticated techniques have also been developed. Both nuclear magnetic resonance (NMR) and mass spectrometry (MS) have been used to examine the formation of hydrogen bonds during protein folding in several studies[5,6]. Mass spectrometry has also been a useful tool for accurate molecular mass determination of proteins and for structural characterization[7,8].

Common separation techniques used to study protein unfolding and refolding processes are polyacrylamide gel electrophoresis (PAGE), size exclusion chromatography (SEC), and ion-exchange chromatography (IEC)[9-14]. For example, native PAGE has long been a standard technique for resolving differences in

charge state and/or hydrodynamic volume of protein conformations. Charge differences resulting from the substitution of just one amino acid can be easily detected[15]. When used for studying protein folding reactions, native PAGE can resolve monomers, dimers, trimers, etc. that may accumulate on the folding and aggregation pathway[16].

Capillary Zone Electrophoresis-Laser Induced Tryptophan Fluorescence Detection for Monitoring Refolding of a Large Multimeric Protein

Capillary zone electrophoresis (CZE) in open-tubular glass capillaries was first developed by Jorgenson and Lukacs[17,18] which eliminated many of the problems associated with traditional electrophoretic methods and resulted in high efficiency separations of charged analytes. The typical CZE instrument consists of a fused silica capillary with internal diameters ranging from 25 to 200 μm and lengths from a few centimeters to a meter. This capillary is filled with buffer solution and extended between two buffer reservoirs connected with high voltage electrodes. A potential field of 100 to 1000 V/cm is then applied across the capillary creating migration of charged analytes due to both electrophoretic mobility and electroosmotic flow.

Electroosmotic flow arises from the surface charges on the capillary wall due to silanol groups of fused silica which are ionized at pH's above 3. Positive ions in the buffer are attracted to these anionic sites, forming an electric double layer. Under an electric field, these hydrated cations migrate towards the cathode causing movement of the bulk solution. In this manner, anions, cations and neutrals are all

swept past the detector which is positioned towards the cathodic end of the capillary. Separation of analytes is obtained from their differences in electrophoretic mobility, governed mainly by charge and to a lesser degree by size.

Unlike chromatography, the separation principle in CZE is based on intrinsic properties, mass and charge, and does not involve interaction and mass transfer between a mobile and stationary phase. Therefore, CZE offers a fast and efficient means of separating various molecular states present during protein unfolding and refolding reactions[19-22], thus allowing better dissection of protein folding and unfolding pathways. For example, the mobility data obtained from CZE were used to estimate the apparent thermodynamic parameters of enthalpy and entropy changes associated with the unfolding transition of lysozyme[19]. Furthermore, CZE was utilized for the analysis of bovine trypsinogen, which underwent oxidation from fully reduced protein through a distribution of intermediate species until reaching the disulfide bond conformation corresponding to the native structure[21].

When molecules absorb light, electrons are promoted to an excited state. Fluorescence emission occurs when these electrons return to the ground state. Since some energy is lost by non-radiative processes, the energy of the emitted light is always less than that of the absorbed light. Therefore, the wavelength of fluorescence emission is longer than the wavelength of absorption. In proteins, fluorescence originates from the aromatic amino acids, which includes phenylalanine, tyrosine and tryptophan. When present, tryptophan usually dominates the fluorescence spectrum due to its much higher absorbance and

quantum yield of emission. The wavelength of excitation for tryptophan is 280 nm, whereas its emission occurs at values around 350 nm.

The emission wavelength can vary depending on the environment surrounding the tryptophans. Usually an emission maximum around 350 nm corresponds to fluorescence of tryptophan in aqueous solutions, as when a protein is in an unfolded denatured state. However, this maximum is usually shifted to shorter wavelengths as tryptophan is incorporated into the interior of the protein as it refolds. This is because tryptophan becomes shielded from quenching by the solvent. Additionally, a change in fluorescence intensity is usually observed upon protein folding, although it may be either higher or lower than that of the denatured state. These changes in fluorescence (intensity and emission wavelength) can be significant and are highly sensitive to environment, making intrinsic tryptophan fluorescence a valuable probe of conformational changes during protein refolding.

Refolding intermediates separated in CZE can be monitored by their native tryptophan fluorescence at 340 nm. This is accomplished by using laser-induced fluorescence detection (LIFD), originally designed by Yeung and co-workers[23]. UV absorbance measurements not only suffer from the strong background absorbance of denaturant, but also exhibit higher concentration detection limits in comparison with the LIFD. Additionally, tryptophan fluorescence reflects the compactness of the molecule during the refolding reaction. The relatively long lived refolding intermediates in the well-characterized pathway of tailspike protein, as discussed in

the next session, makes tailspike ideal for monitoring refolding using CZE equipped with LIFD.

The *in vitro* refolding of phage P22 tailspike endorhamnosidase, a model system for studying refolding mechanisms, is the subject of Chapter 2. Its pathway has been well characterized both *in vivo*[24-27] and *in vitro*[16,28-30] using conventional techniques, including PAGE and fluorescence spectroscopy. In this study, rapid identification of transient partially folded intermediates of tailspike protein is demonstrated using CZE-LIFD.

The P22 tailspike protein is a large oligomeric protein consisting of three identical monomers containing 666 amino acids (M_r = 71,759). Tailspike structure consists mostly of β-sheet as determined by Raman spectroscopy[31] and x-ray crystallography[32]. Tailspike is thermostable, withstanding temperatures up to 88°C, and is resistant to protease and SDS. Upon initiation of folding, tailspike polypeptides rapidly fold into structured monomers indicated by significant secondary structure. These monomeric species associate to form the protrimer intermediate that is susceptible to SDS and protease. Conversion of the protrimer into the native, SDS resistant tailspike is the rate-limiting step in the folding pathway[16,28] (see Figure 1.1).

Although both the *in vivo* and *in vitro* pathways have been well studied, more thorough research on how changing temperature affects tailspike refolding kinetics and yields is needed. An early monomeric intermediate of tailspike has been

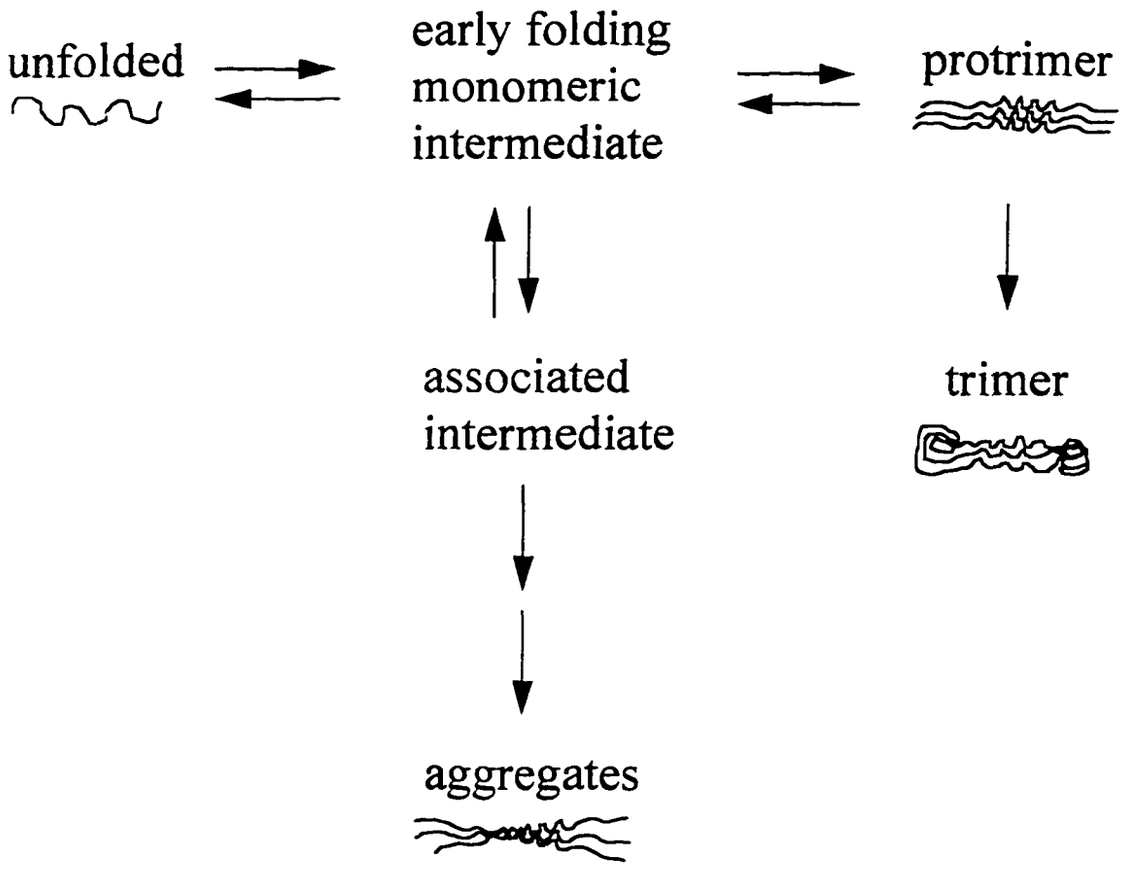

Figure 1.1. In vitro refolding pathway for P22 tailspike endorhamnosidase.

identified as being thermolabile with an increased tendency to form aggregates at higher temperatures[15,27,30]. Several temperature sensitive folding (tsf) mutations of tailspike have also been identified, indicating the significance of amino acid sequence in acquiring native structures[33-35]. At restrictive temperatures (>39°C), these tsf mutants act to further destabilize the thermolabile monomeric intermediate causing a shift from productive folding to the aggregation pathway. Even less is known about the effects of extremely low temperatures on tailspike refolding. Chapter 3 attempts to answer these questions by analyzing tailspike refolding at various temperatures and during temperature shifts using CZE-LIFD.

Capillary Isoelectric Focusing-Electrospray Ionization Mass Spectrometry for Monitoring Disulfide Bond Induced Protein Refolding

Capillary isoelectric focusing (CIEF), as first developed by Hjerten and Zhu[36], combines the resolving power of traditional gel isoelectric focusing with the high efficiency and speed of capillary electrophoresis. CIEF separates proteins based on their isoelectric points (pIs) and can resolve differences as small as 0.02 pI units. In CIEF the fused silica capillary is coated with linear polyacrylamide to eliminate electroosmotic flow[37]. The sample of proteins to be separated is mixed with a solution of ampholytes, which are a mixture of synthetic polyamino and polycarboxyl molecules.

To form the pH gradient, an acidic solution (pH~2) is placed at the anode and a basic solution (pH~12) is placed at the cathode. When an electric field is applied, the ampholytes containing a net charge will migrate towards the oppositely charged electrode. However, the ampholytes cannot migrate out of the capillary because they will encounter a pH greater or lower than their pI, resulting in loss of charge or charge reversal. Since the ampholytes have a broad range of pIs, a continuous pH gradient is formed in the capillary. The extent of the gradient is controlled by the range of pIs of the ampholytes.

At the same time the pH gradient is being formed, protein components also begin to migrate. Proteins, initially charged, will migrate until they reach the pH equal to their pI where they become neutral and produce a focused zone. After focusing, protein bands must be mobilized past an on-line detector (either UV or MS). The most effective method is a combination of cathodic and gravity mobilization. This entails replacing the basic catholyte with a solution containing a non-hydroxyl anion. The migration of this non-hydroxyl anion into the capillary results in a reduction of the concentration of hydroxyl anions causing a decrease in pH that eventually propagates down the capillary. This causes a progressive shift in pH which mobilizes the zones past the detector. Gravity is added to this process to reduce problems of salt mobilization for proteins at the anodic end of the capillary[37] which can destroy the linearity of the pH gradient during mobilization.

Mass spectrometry (MS) is an extremely valuable technique for detailed structural characterization and purity evaluation of biomolecules. MS has long been

valued as a tool for highly accurate mass determinations. With such high accuracy and the advent of new ionization methods, researchers are now able to detect and characterize post-translational modifications, chemical modifications, mutations and even protease degradations of proteins[38,39]. Electrospray Ionization (ESI) and Matrix Assisted Laser Desorption/Ionization (MALDI) are two of the most common techniques presently used for protein characterization by mass spectrometry.

ESI is a soft ionization technique, first developed by Yamashita and Fenn in 1984[40,41], allowing formation of macromolecular ions that are multiply charged. This multiple charging is ideally suited for analysis of large biomolecules since it brings the mass-to-charge ratios of the analytes within the limited range of typical mass analyzers. Electrospray ionization occurs as an aqueous solution containing analyte flows from a capillary in a high electric field at near atmospheric pressure, producing a fine spray of charged droplets. As the droplets undergo evaporation, they reach a point at which the surface charge density exceeds the surface tension of the droplet (called the Rayleigh limit), resulting in Coulombic explosion or droplet fission. These droplets may undergo solvent evaporation and droplet fission several times until very small highly charged droplets are formed and eventually gas phase ions are produced[42]. These gaseous ions are then swept into the high vacuum of the mass spectrometer for subsequent detection.

On-line integration of CIEF with ESI-MS is employed for monitoring protein refolding induced by disulfide bond formation (see Chapter 4). Some of the pioneering work of Anfinsen and co-workers[43], showing that proteins could

spontaneously refold in vitro and thus contained all the required information in their amino acid sequence, was done with ribonuclease A (RNase A). Since then, RNase A has been the subject of many thorough studies[44-46]. The small, globular structure (MW ~ 13,700 Da) of RNase A and the presence of four disulfide bonds makes it ideal for studying protein refolding.

Since it is believed that proteins fold by a non-random process passing through defined intermediates, then it should be possible to block folding by interfering with the pathway. In the case of RNase A, folding is coupled to the formation of disulfide bonds between pairs of cysteine residues. By covalently modifying the free cysteines during different stages in the folding pathway, intermediates can be trapped and folding is halted. The intermediates can then be studied directly and in detail. Another advantage of disulfide coupled refolding is that since disulfide formation between thiols (cysteines) is an oxidation/reduction process, the redox potential can be experimentally controlled through the use of thiol reagents, such as dithiothreitol (DTT) or glutathione (GSSG)[47]. This method has been employed for the study of various disulfide containing proteins[48,49].

Blocking of the cysteines can be accomplished by several methods. Since the reactive species for disulfide formation is S^-, lowering the pH will push the equilibrium toward the formation of SH. This method does not actually stop thiol disulfide exchange, but simply slows it down by reducing the amount of reactive S^-. Although it is a fast and reversible method, it severely limits the analytical methods which can be used to characterize the intermediates, since low pH must be

maintained always. Thiosulfonate reagents, such as 2-aminoethyl methanethiosulfonate (AEMTS), also react extremely quick and can be reversible[46]. AEMTS has the added advantage of increasing the charge of the protein by +1 for each thiol modified, providing basis for separation of intermediates in a predictable manner. More commonly used are the alkylating reagents of iodoacetate (IAA) or iodoacetamide (IAM)[50,51]. Although IAA is considered less efficient in some cases than IAM, it was chosen for this work for its ability to add -1 charge for each alkylation site, reducing the pI of basic RNase A. The increase in molecular mass and the reduction in pI of alkylated RNase A folding intermediates facilitates the study of refolding kinetics and yields using CIEF-ESIMS.

CHAPTER 2. MONITORING THE REFOLDING PATHWAY FOR A LARGE MULTIMERIC PROTEIN USING CAPILLARY ZONE ELECTROPHORESIS

A paper published in the Journal of Chromatography A

Pamela K. Jensen, Z. Hugh Fan, Jonathan King, and Cheng S. Lee

Abstract

Rapid identification of transient partially folded intermediates formed during protein refolding and aggregation has been difficult, particularly with separation methods relying on solid matrices. Capillary zone electrophoresis equipped with laser-induced fluorescence detection provides a fast sensitive means of identifying folding and aggregation intermediates using the intrinsic tryptophan fluorescence. The *in vitro* refolding of the trimeric P22 tailspike, a model system for the study of protein folding, misfolding, and aggregation, has been monitored after dilution out of denaturant. Both monomeric and trimeric folding intermediates were resolved.

The refolding kinetics and yields measured by capillary zone electrophoresis were in good agreement with those obtained via fluorescence spectrophotometry and polyacrylamide gel electrophoresis. In comparison with typical UV detection, laser-induced tryptophan fluorescence increased detection sensitivity. In addition,

the fluorescence signal carries information on the packing of the tryptophan residues in the folding intermediates.

For tailspike and many other proteins, the off pathway aggregation reactions proceed from a thermolabile intermediate at the junction with the productive pathway. By monitoring refolding intermediates after temperature shifts, the structured monomeric intermediate was identified as the thermolabile junctional intermediate between the productive and aggregation pathways.

Introduction

Newly synthesized polypeptide chains within cells pass through a series of partially folded intermediates in reaching their native state. Association of these intermediates into the aggregated inclusion body state often competes with productive folding into the native state in both prokaryotic and eukaryotic cells[1]. The generation of misfolded and generally aggregated chains has emerged as a practical problem both in the biotechnology industry and in biomedical research. It has been particularly, though not exclusively, associated with expression of cloned genes in heterologous hosts[2,3].

Many proteins can fold *in vitro* from a completely denatured state[4,5], providing evidence that a polypeptide chain may contain all the information required for folding. Various refolding studies performed to characterize the effects of denaturant, redox reagent, and protein concentration have improved the efficiency and yield of the refolding process. The study of protein refolding is also crucial to

understand the underlying mechanisms, as well as for the elucidation of "the second half of the genetic code" in the prediction of protein structure.

Most techniques for characterizing protein folding reactions usually depend on direct measurements of some average physical properties that are sensitive to the change in the protein structure. Examples include the changes in optical density and steady state fluorescence intensity[6-10]. Capillary zone electrophoresis (CZE) is an attractive alternative to experimental observation and provides the access to a cross section of the population of molecular states within an equilibrium system during a protein refolding reaction. Additionally, separation in CZE is a function of intrinsic molecular properties and does not involve the interaction and mass transfer between a mobile and a stationary phase such as in size-exclusion[11,12] and ion-exchange[11,13,14] HPLC.

CZE has been used to monitor the unfolding of human serum transferrin in urea[15] and the temperature-dependent unfolding transitions of lysozyme at low pH[16]. The unfolding of transferrin isoforms is dependent on the iron content, but not the carbohydrate content of isoforms. The mobility data obtained from CZE were used to estimate the apparent thermodynamic parameters of enthalpy and entropy changes associated with the unfolding transition of lysozyme. Furthermore, CZE was utilized for the analysis of bovine trypsinogen, which underwent oxidation from a fully reduced molecule through a distribution of intermediate species until it reached the disulfide bond conformation corresponding to the native structure[17]. In comparison to the results obtained from size-exclusion HPLC, slab gel

electrophoresis, and gel isoelectric focusing, CZE provides an effective technique for the analysis of protein refolding.

In addition to protein unfolding studies, CZE has been applied for the analysis of point mutations in DNA using temporal thermal gradients in the capillary filled with sieving liquid polymers[18]. The temperature gradient was used to exploit the differences in melting temperature (T_m) of DNA molecules differing by a single nucleotide[18]. When the temperature in the capillary equaled a DNA fragment's T_m, the double helix partially unwound and resulted in a marked retardation of its electrophoretic mobility. The method was demonstrated for the identification of three point mutations located in the cystic fibrosis transmembrane conductance regulator gene.

In the biotechnology industry, it is often important to be able to monitor species which associate off pathway into insoluble aggregates. The absence of a solid matrix, and the speed of separation, suggests that CZE may be a method of choice for such monitoring. In this study, CZE equipped with native protein fluorescence detection is employed for monitoring the refolding of phage P22 tailspike endorhamnosidase, a model system for the study of folding and aggregation processes. Refolding intermediates separated in CZE can be monitored by their tryptophan fluorescence at 340 nm. The current CZE-UV absorbance measurements not only suffer from the strong background absorbance of denaturant, but also exhibit the higher concentration detection limits in comparison

with the laser-induced fluorescence detection (LIFD). Additionally, tryptophan fluorescence reflects the compactness of the molecule during the refolding reaction.

The P22 tailspike endorhamnosidase is an oligomeric protein and consists of three identical monomers containing 666 amino acids. Six tailspikes assemble onto the phage capsid to form the cell recognition and attachment apparatus of the phage. The native protein is thermostable, with a T_m of 88°C, and is resistant to sodium dodecyl sulfate (SDS) and proteases[19,20]. Refolding experiments, starting from unfolded single-chain monomer in acid-urea, have shown that the *in vitro* refolding pathway closely resembles the folding pathway *in vivo*[21,22]. Upon initiation of refolding, tailspike polypeptides rapidly fold into structured monomer intermediates with a high content of secondary structure. These monomeric species associate to form the triple-chain defined folding intermediates, the protrimers, that are susceptible to SDS and proteases. Conversion of the protrimer into the native, SDS-resistant tailspike protein is the rate-limiting step in the folding pathway[21,22].

Experimental

Capillary Zone Electrophoresis-Laser Induced Fluorescence Detection

The CZE apparatus shown in Fig. 2.1 was constructed in-house using a CZE 1000R high-voltage (HV) power supply (Spellman High-Voltage Electronics, Plainview, NY). A negative electric potential was applied at the detector end of the capillary for electrokinetic injection and for electrophoretic separation while the

injection end was grounded. Fused silica capillaries of 50 μm i.d. x 150 μm o.d. were obtained from Polymicro Technologies (Phoenix, AZ) and were cut to a length of 40 cm. The separation distance was 27 cm between the injection point and the LIFD. Except for those specified, CZE separations were carried out in 25 mM sodium phosphate, 140 mM urea at pH 7.6. The CZE apparatus was enclosed in a plexiglas interlock box for operator safety.

A cooling jacket (see Fig. 2.1) was employed to maintain a constant temperature of 10°C at the inlet reservoir and around most of the separation capillary for refolding measurements. There was a 3 cm cooling gap between the end of the cooling jacket and the LIFD. The capillary was inserted inside a 2 mm i.d. x 5 mm o.d. C-Flex tubing (Cole-Parmer, Chicago, IL) where the coolant was circulated around the capillary. The coolant, a mixture of ethylene glycol and water in a volume ratio of 2 to 1, was thermostatted in a low temperature circulating bath (Fisher Scientific, Pittsburgh, PA). A peristaltic pump (Cole-Parmer) was used to circulate the coolant in the cooling jacket at a flow rate of 35 ml/min.

The LIFD originally designed by Yeung and his co-workers[23] was employed for measuring tryptophan fluorescence of refolding intermediates separated by CZE. Briefly, the 248 nm line of a KrF excimer laser (Potomac Photonics, Lanham, MD) was used for excitation. A mirror directed the laser beam through a 248 nm bandpass filter (Oriel, Stratford, CT) after which a lens (Oriel) focused the beam onto the capillary window. The detection window was made by removing the polyimide coating with heated sulfuric acid. A 10x microscope objective (Edmund

Scientific, Barrington, NJ) collected the fluorescence emission and a second mirror directed the fluorescence through a 340 nm filter (Melles Griot, Irvine, CA) and then onto a photomultiplier tube (Hamamatsu, Bridgewater, NJ). Data collection was performed by an HP 35900C analog to digital interface board with the HP G1250C General Purpose Chemstation Software (Hewlett Packard, Fullerton, CA).

Protein Denaturation and Refolding

The P22 tailspike endorhamnosidase was denatured in a solution containing 7 M urea and 25 mM sodium phosphate at pH 3 for one hour at room temperature[21,22,24]. Refolding was initiated by rapidly diluting the unfolded protein with the refolding buffer of 25 mM sodium phosphate and 1 mM ethylenediaminetetraacetic acid at pH 7. The refolding buffer was pre-equilibrated at 10°C in a low temperature circulating bath (Fisher Scientific). Based on a 50 fold dilution, the final tailspike and urea concentrations in the refolding reaction were 22 μg/ml and 140 mM, respectively.

The progress of the refolding reaction was monitored by subjecting an aliquot of refolding sample to CZE separation at various reaction times. Additionally, the refolding samples were analyzed by slab gel electrophoresis and the results were compared with the CZE separations. For slab gel electrophoresis, a 15 μl aliquot of refolding sample was placed on ice to quench refolding. Before it was placed on the gel, the refolding sample was mixed with 7.5 μl of SDS solution containing

0.28 g/ml glycerol, 0.16 M tris(hydroxymethyl)aminomethane at pH 6.8, 5% SDS, 0.25 mg/ml bromophenol blue, and 50 mM dithiothreitol.

Slab Gel Electrophoresis

A Bio-Rad Mini-Protean II dual slab cell (Hercules, CA) was used to perform all SDS polyacrylamide gel electrophoresis measurements. The separation and stacking gels were cast as described in the instruction manual. A protein mixture of rabbit muscle myosine, *E. Coli* b-galactosidase, rabbit muscle phosphorylase, bovine albumin, egg albumin, and bovine erythrocyte carbonic anhydrase was used as the molecular mass standard. The electrophoresis voltage was increased to 150 V after the protein bands had entered the separation gel. Electrophoresis was stopped when the bromophenol blue dye front reached about 1 cm from the gel bottom. Silver staining of protein bands was performed as described elsewhere[25].

Tryptophan Fluorescence Spectrophotometry

Tryptophan fluorescence of tailspike protein during the refolding reaction was monitored by a Perkin Elmer LS 50B spectrophotometer (Norwalk, CT). Fluorescence was excited at 280 nm with a slit width of 5 nm. Fluorescence emission spectrum was scanned between 300 nm and 400 nm with a slit width of 5 nm. Refolding was initiated by rapidly diluting the unfolded protein with the refolding buffer in a 200 μl quartz cell with a 3 mm path length (Starna

Spectrophotometer Cells, Atascadero, CA). The quartz cell was thermostatted at 10°C by a low temperature circulating bath (Fisher Scientific).

Chemicals

The P22 tailspike endorhamnosidase with a concentration of 8.9 mg/ml was prepared and purified by standard procedures in King and Yu[26]. The protein molecular mass standard used in slab gel electrophoresis was purchased from Sigma (St. Louis, Missouri).

L-tryptophan received from Sigma was used as the electroosmosis marker for the CZE separations. Monobasic and dibasic sodium phosphate, urea, and dithiothreitol were obtained from Aldrich (Milwaukee, WI). Tris(hydroxymethyl)-aminomethane and ethylenediaminetetraacetic acid were purchased from Bio-Rad and Fisher Scientific, respectively. All solutions were prepared using water purified by a NANOpure II system (Dubuque, IA) and further filtered with a 0.22 μm membrane (Millipore, Bedford, MA).

Results and Discussion

Monitoring Tailspike Refolding Using Capillary Zone Electrophoresis-Laser Induced Fluorescence Detection

The near-UV detection of tailspike refolding intermediates separated in CZE suffered from the strong background absorbance of typical denaturants including

guanidine hydrochloride and urea. For the far-UV detection of tailspike protein, each single-chain polypeptide of 666 amino acids contained only 7 tryptophan residues. To reduce single-chain polypeptide self-association and aggregation, a low protein concentration around 10-50 μg/ml has been required for achieving *in vitro* reconstitution of native tailspikes from completely unfolded polypeptides[21,22,24]. The small UV extinction coefficient of tailspike protein at 280 nm and the use of low protein concentration in productive refolding led to the selection of LIFD of refolding intermediates in CZE.

By measuring native tryptophan fluorescence emitted at 340 nm, a concentration detection limit of 1×10^{-8} M for conalbumin in CZE was demonstrated by using the excitation at 257 nm as the result of frequency doubling from an argon ion laser[27]. An even lower detection limit of 1×10^{-10} M for conalbumin was later achieved by Lee and Yeung[28] with the application of a 275 nm line of an argon ion laser. The good match of the laser line at 275 nm with tryptophan's maximum absorbance at 280 nm accounted for the improvement on the detection limit by two orders of magnitude. In this study, the 248 nm line of a KrF excimer laser was used for excitation in the measurement of native tryptophan fluorescence. A concentration detection limit of 1 μg/ml (or 1.4×10^{-8} M) for native tailspike protein was obtained in normal CZE separation (data not shown).

As shown in Fig. 2.2A, the CZE separation of tryptophan (the first peak) as the electroosmosis marker and native tailspike protein (the second peak) was obtained in the electrophoresis buffer of 25 mM sodium phosphate, 100 mM urea at pH 7.6.

At pH 7.6, the tailspike protein with a pI of 5.2 exhibited a net negative charge. The high thermostability of phage P22 tailspike protein was also reflected in its resistance to unfolding by chemical denaturants. At neutral pH, the tailspike protein trimer withstood urea up to the limit of solubility of the denaturant at room temperature. By using a combination of urea and acid pH (6-7 M urea at pH 3), tailspike trimers were dissociated and completely unfolded.

To maintain the denatured state of tailspike polypeptide chain and compare its electrophoresis behavior with that of native tailspike in CZE, the electrophoretic separations of native and denatured tailspike proteins were carried out in the electrophoresis buffer containing 25 mM sodium phosphate and 6 M urea at pH 7.6, as shown in Figs. 2.2B and 2.2C, respectively. The increase in the migration times of tryptophan and native tailspike in Figs. 2.2A and 2.2B presumably reflects the increase in the solution viscosity due to the addition of 6 M urea. In the presence of 6 M urea, the absolute value of negative electrophoresis mobility of tailspike endorhamnosidase was 0.69×10^{-4} cm^2/V-sec for native tailspike (see Fig. 2.2B) and decreased to 0.39×10^{-4} cm^2/V-sec for denatured single-chain polypeptide (see Fig. 2.2C). In comparison with the contracted chains of native tailspike, the random coil configuration of denatured tailspike resulted in a greater frictional drag and a smaller electrophoretic mobility under the influence of an electric field. Additionally, the peak broadening of denatured tailspike might reflect the slowness of the interchange reactions among various denatured forms and the greater protein-wall interactions due to the hydrophobic nature of unfolded protein.

The formation of native tailspikes during reconstitution from denatured polypeptides occurred very slowly at 10°C[21]. Thus, a constant temperature of 10°C was selected for tailspike refolding and for refolding monitoring by CZE equipped with LIFD. As shown in Fig. 2.1, nearly 90% of the separation capillary between the injection point and the LIFD was thermostatted at 10°C by using a cooling jacket.

The refolding reaction was performed at 10°C by rapidly diluting the unfolded tailspike protein with the refolding buffer of 25 mM sodium phosphate and 1 mM ethylenediaminetetraacetic acid at pH 7. Based on a 50 fold dilution, the final tailspike and urea concentrations in the refolding reaction were 22 μg/ml and 140 mM, respectively. The refolding conditions, including refolding temperature and final denaturant and tailspike concentrations, were selected for the enhancement of productive refolding over aggregation. As shown in Fig. 2.3, a series of tailspike refolding samples taken at various reaction times were analyzed by CZE-LIFD. Two partially resolved protein peaks were observed in the electropherograms and their tryptophan fluorescence intensities all increased during the first hour of tailspike refolding (see Figs. 2.3A-2.3C). The fluorescence intensity of the early eluting peak at 9.6 min reached its maximum during the first hour of refolding reaction and continuously decreased to its disappearance at 48 hours (see Figs. 2.3C-2.3G). The early eluting peak was assumed to be the formation of structured monomers which formed rapidly at the beginning of reconstitution[21,22,24].

In contrast, the fluorescence intensity of the late eluting peak at 10.0 min continued to increase and reached its maximum around the refolding time of 3 hours

(see Fig. 2.3D). The late eluting peak was formed by association from the monomeric polypeptides, and represented a precursor (the protrimer) of the native tailspike protein trimer[19,22]. Its late migration time, in comparison with that of the monomeric intermediate, was attributed to the relative compactness of protrimer structure. The observed refolding kinetics in the formation of monomeric and trimeric intermediates during the first three hours of refolding reaction were in good agreement with those reported in the literature[21,22,24].

The tryptophan fluorescence intensity of the late eluting peak decreased slowly after the first three hours of refolding reaction (see Figs. 2.3D-2.3G). The conversion of the protrimer into the native, SDS resistant tailspike protein is the rate-limiting step in the folding pathway[21,22] and occurs very slowly at temperatures below 20°C[29]. In this study, the protrimer and the formation of native tailspike coeluted during the CZE separations. The coelution of native tailspike with trimeric precursor was illustrated by spiking the refolding sample with native tailspike (see Fig. 2.3H). The decrease in fluorescence intensity of the late eluting peak was contributed by the maturation of tailspike trimers and the difference in tryptophan fluorescence between the protrimer and native tailspike. By assuming the presence of only native tailspike in the refolding sample at 48 hours (see Fig. 2.3G), the comparison of peak intensities between the refolding sample and the authentic tailspike (see Fig. 2.3I) indicated a refolding yield of ~ 50% of total denatured protein.

Monitoring Tailspike Refolding Using Tryptophan Fluorescence Spectrophotometry

Fluorescence emission spectra of tailspike refolding samples excited at 280 nm are summarized in Fig. 2.4. The denatured tailspike protein had an emission maximum of 355 nm. Tailspike protein refolded by rapid dilution had a higher fluorescence intensity and an emission maximum that was shifted to 342 nm, indicating that some of the tryptophans became inaccessible to the solvent during refolding. The fluorescence intensity continued to increase for the first three hours of tailspike refolding and remained relatively steady for the next three to six hours. The increase in tryptophan fluorescence during the early stage of tailspike refolding corresponded to the formation of structured monomeric intermediates and the association of monomeric intermediates into the protrimer. The maturation of native tailspike from the protrimer was extremely slow at 10°C and accounted for the decrease in the fluorescence intensity during the late stage of tailspike refolding. The decrease in the fluorescence intensity was most likely due to a rearrangement of the hydrophobic environment surrounding some of the tryptophans from the protrimer to the native, SDS-resistant tailspike. Similar changes in tryptophan fluorescence were also observed during *in vitro* folding of the phage P22 coat protein[30]. The spectroscopic results illustrated in Fig. 2.4 were in good agreement with the CZE electropherograms summarized in Fig. 2.3 for tailspike refolding monitoring.

Monitoring Tailspike Refolding Using Slab Gel Electrophoresis

The mature native tailspikes remained native in the presence of SDS and migrated much more slowly than the denatured extended SDS polypeptide chain complex. Those refolding species that have not yet matured into native trimers, including single-chain and protrimer intermediates, were denatured by the detergent and easily distinguished from the native tailspikes. As shown in Fig. 2.5, the presence of native tailspike appeared as early as two hours after the initiation of tailspike refolding. Still, this observation confirmed our previous assumption that the two species partially resolved by CZE during the first three hours of tailspike refolding were mainly contributed by the structured monomers and the protrimers (see Figs. 2.3A-2.3D). The increase in the amount of native tailspike was clearly accompanied with the decrease in the band intensity of denatured monomers contributed by both the single-chain and protrimer intermediates. By comparing the band intensity of refolding sample at 48 hours (lane 8) with that of native tailspike control (lane 2), a refolding yield of ~ 63% was estimated, indicating a reasonable agreement with the result of CZE-LIFD measurement.

Temperature Dependence of Productive vs. Aggregation Pathways

As shown in Figs. 2.3E-2.3G, a minor peak was observed in the electropherogram with the migration time very close to that of tryptophan as the electroosmotic flow marker. This minor component might be the early intermediate in the aggregation pathway. To test this, protein reconstitution could be performed

at protein concentrations above 50 μg/ml, or at elevated temperatures, which increases off pathway aggregation. To further illustrate the effect of temperature on tailspike refolding, the refolding was initially carried out at 10°C, followed by a temperature shift at different stages of tailspike refolding. As shown in Fig. 2.6A, an increase of refolding temperature from 10° to 30°C during the formation and accumulation of structured monomers significantly affected the distribution and competition among misfolding (aggregation) and productive folding. In contrast, the aggregation intermediate remained in a relatively small amount when the refolding temperature was increased during or after the accumulation of protrimer tailspike (Fig. 2.6B). These preliminary results seemed to indicate that the monomeric species were the critical intermediates at the junction between the productive and aggregation pathways. By comparing with the results shown in Fig. 2.3, the increase in the migration times of refolding and aggregation intermediates was caused by the slower electroosmotic flow in this particular capillary.

Refolding yield of tailspike endorhamnosidase was strongly dependent on the reaction temperature. High refolding temperatures promoted protein self-association and aggregation. By shifting from high temperature to permissive temperature, the formation of protrimer from structured monomer in the folding pathway appeared, and then the native protein. The observation was consistent with the refolding being blocked at the stage of the single-chain polypeptide. Thus, all experimental results based on temperature shifts supported the role of monomeric species as the critical intermediates at the junction between productive

folding and aggregation. Furthermore, high refolding yield of tailspike protein, around 60-80% of the total protein, has been reported[24,31] and demonstrated in this study by initiating protein refolding at 10°C.

Acknowledgments

Assistance and discussion by Dr. Edward S. Yeung (Iowa State University) in the design and development of laser-induced fluorescence detection is appreciated. Special thanks to Drs. Anne Robinson and Margaret Speed (Massachusetts Institute of Technology) for providing P22 tailspike endorhamnosidase and for their assistance in tailspike refolding and slab gel electrophoresis protocols.

This work was supported by the Biotechnology Process Engineering Center at Massachusetts Institute of Technology through the National Science Foundation's Engineering Research Center Initiative under the Cooperative Agreement CDR-8803014. P.K.J. is a recipient of the Graduate Assistance in Areas of National Need (GAANN) Fellowship from the US Department of Education. C.S.L. is a National Science Foundation Young Investigator (BCS-9258652).

References

1. J. King, C. Haase-Pettingell, C. Gordon, S. Sather, A. Mitraki, *Amino Acid Sequence Determinants of Polypeptide Chain Folding and Inclusion Body Formation* in *Protein Folding: in Vivo and in Vitro,* J. L. Cleland, ed., American Chemical Society, Washington, D. C., 1993, pp. 24.

2. F. A. O. Marston, *Biochem. J.* 240 (1986) 1.

3. A. Mitraki, J. King, *Bio/Technology* 7 (1989) 690.

4. T. E. Creighton, *Proteins: Structure and Molecular Properties,* W. H. Freeman, New York. 1983.

5. F. M. Richards, *Folded and Unfolded Proteins - An Introduction* in *Protein Folding,* T. E. Creighton, ed., W. H. Freeman, New York, 1992.

6. P. J. Hagerman, R. L. Baldwin, *Biochemistry* 15 (1976) 1462.

7. A. M. Labhart, *Proc. Natl. Acad. Sci.* 81 (1984) 7674.

8. A. Galat, C. C. Yang, E. R. Blout, *Biochemistry* 24 (1985) 5678.

9. E. H. Zuniga, B. T. Nall, *Biochemistry* 22 (1983) 1430.

10. R. F. Kelley, J. Wilson, C. Bryant, E. Stellwagen, *Biochemistry* 25 (1986) 728.

11. F. E. Regnier, K. M. Gooding, *Anal. Biochem.* 103 (1980) 1.

12. R. J. T. Corbett, R. S. Roche, *Biochemistry* 23 (1983) 79.

13. F. E. Regnier, *Anal. Biochem.* 126 (1982) 1.

14. R. W. Stout, S. I. Sivakoff, R. D. Ricker, L. R. Snyder, *J. Chromatogr.* 353 (1986) 439.

15. F. Kilar, S. Hjerten, *J. Chromatogr.* 638 (1993) 269.

16. V. J. Hilser, G. D. Worosila, E. Freire, *Anal. Biochem.* 208 (1993) 125.

17. M. A. Strege, A. L. Lagu, *J. Chromatogr. A* 652 (1993) 179.

18. C. Gelfi, P. G. Righetti, L. Cremonesi, M. Ferrari, *Electrophoresis* 15 (1994) 1506.

19. D. Goldenberg, J. King, *J. Mol. Biol.* 145 (1981) 633.

20. J. Sturtevant, M. H. Yu, C. Haase-Pettingell, J. King, *J. Biol. Chem.* 264 (1989) 10693.

21. R. Seckler, A. Fuchs, J. King, R. Jaenicke, *J. Biol. Chem.* 264 (1989) 11750.

22. A. Fuchs, C. Seiderer, R. Seckler, *Biochemistry* 30 (1991) 6598.

23. E. S. Yeung, P. Wang, W. Li, R. W. Giese, *J. Chromatogr.* 608 (1992) 73.

24. A. Mitraki, M. Danner, J. King, R. Seckler, *J. Biol. Chem.* 268 (1993) 20071.

25. S. K. Sather, J. King, *J. Biol. Chem.* 269 (1994) 25268.

26. J. King, M. H. Yu, *Methods Ensymol.* 131 (1986) 250.

27. D. F. Swaile, M. J. Sepaniak, *J. Liq. Chromatogr.* 14 (1991) 869.

28. H. T. Lee, E. S. Yeung, *J. Chromatogr.* 595 (1992) 319.

29. D. P. Goldenberg, J. King, *J. Mol. Biol.* 145 (1981) 633.

30. C. M. Teschke, J. King, *Biochemistry* 32 (1993) 10839.

31. M. Danner, A. Fuchs, S. Miller, R. Seckler, *Eur. J. Biochem.* 215 (1993) 653.

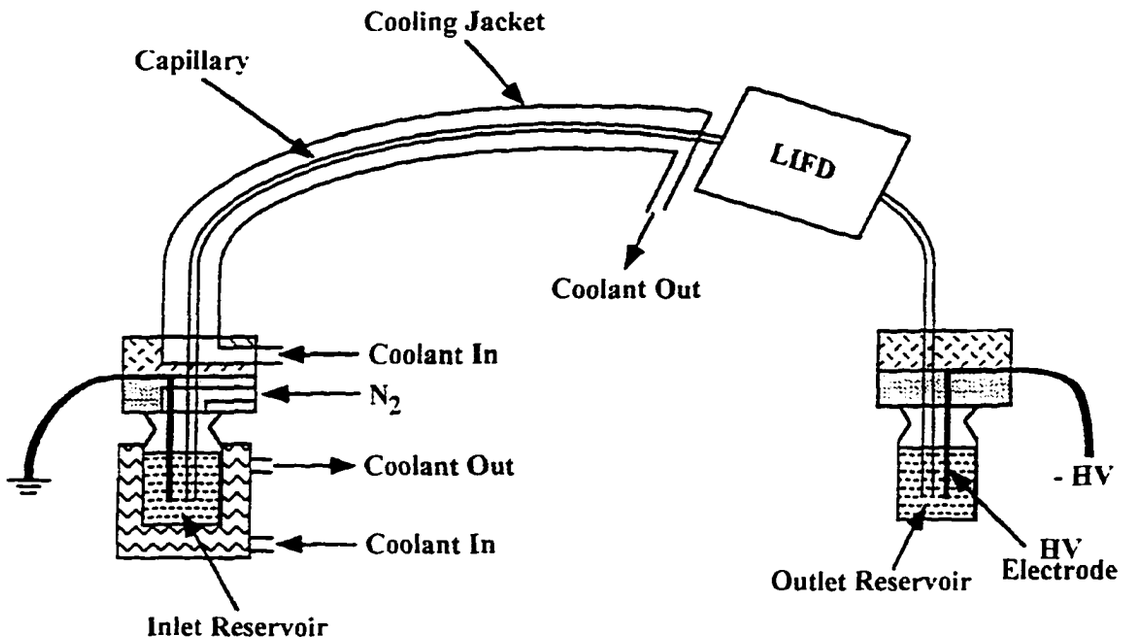

Figure 2.1. Schematic diagram of capillary zone electrophoresis equipped with the cooling jacket and the laser-induced fluorescence detection. Nitrogen gas was used for purging the capillary with washing and buffer solutions.

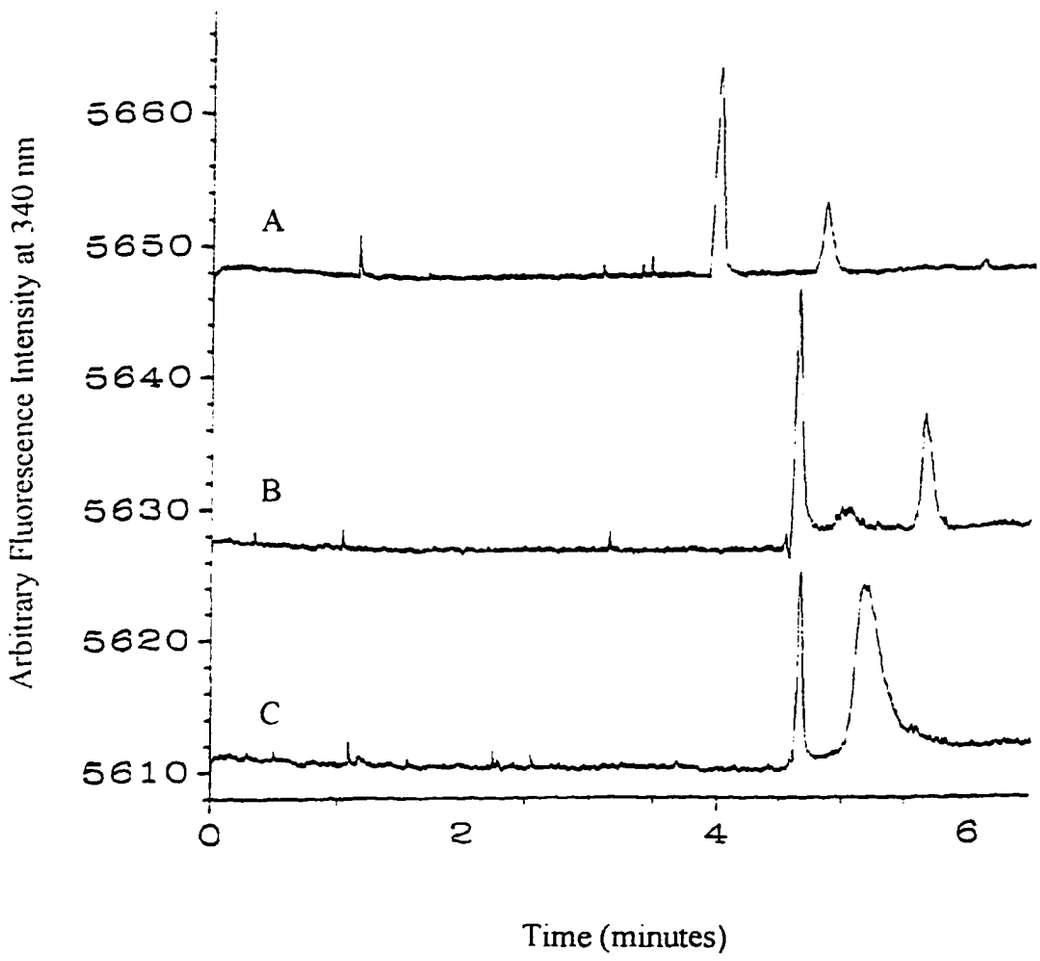

Figure 2.2. Capillary zone electrophoresis-laser induced fluorescence detection of native and denatured tailspikes at 10°C. Elution order was tryptophan, followed by tailspike protein. (A) 22 μg/ml native tailspike in the electrophoresis buffer of 25 mM sodium phosphate and 100 mM urea at pH 7.6; (B) 22 μg/ml native tailspike in the electrophoresis buffer of 25 mM sodium phosphate and 6 M urea at pH 7.6; (C) 85 μg/ml denatured tailspike in the electrophoresis buffer of 25 mM sodium phosphate and 6 M urea at pH 7.6. Capillary, 40 cm total length and 27 cm to detector, 50 μm i.d. and 150 μm o.d.; applied voltage, -10 kV and 2 sec for tailspike injection, -10 kV for electrophoresis; fluorescence detection at 340 nm.

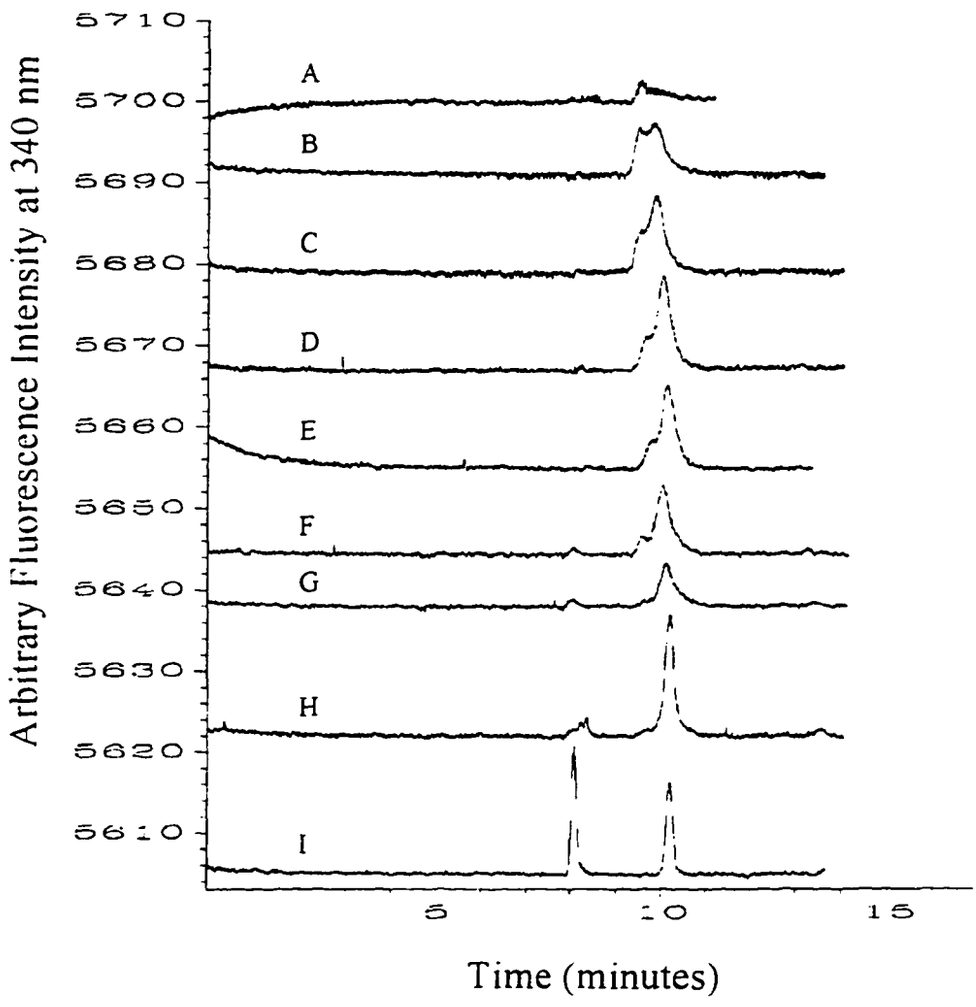

Figure 2.3. Monitoring refolding of tailspike protein by capillary zone electrophoresis-laser induced fluorescence detection at 10°C. Refolding samples were taken at (A) 1 minute, (B) 15 minutes, (C) 1 hour, (D) 3 hours, (E) 5 hours, (F) 22 hours, and (G) 48 hours after the initiation of tailspike refolding. The electropherograms marked as H and I contained the refolding sample at 48 hours spiked with native protein and the mixture of tryptophan and 22 μg/ml native tailspike, respectively. Electrophoresis buffer, 25 mM sodium phosphate and 140 mM urea at pH 7.6; capillary, 40 cm total length and 27 cm to detector, 50 μm i.d. and 150 μm o.d.; applied voltage, -6 kV and 10 sec for refolding sample injection, -6 kV for electrophoresis; total tailspike protein concentration, 22 μg/ml; fluorescence detection at 340 nm.

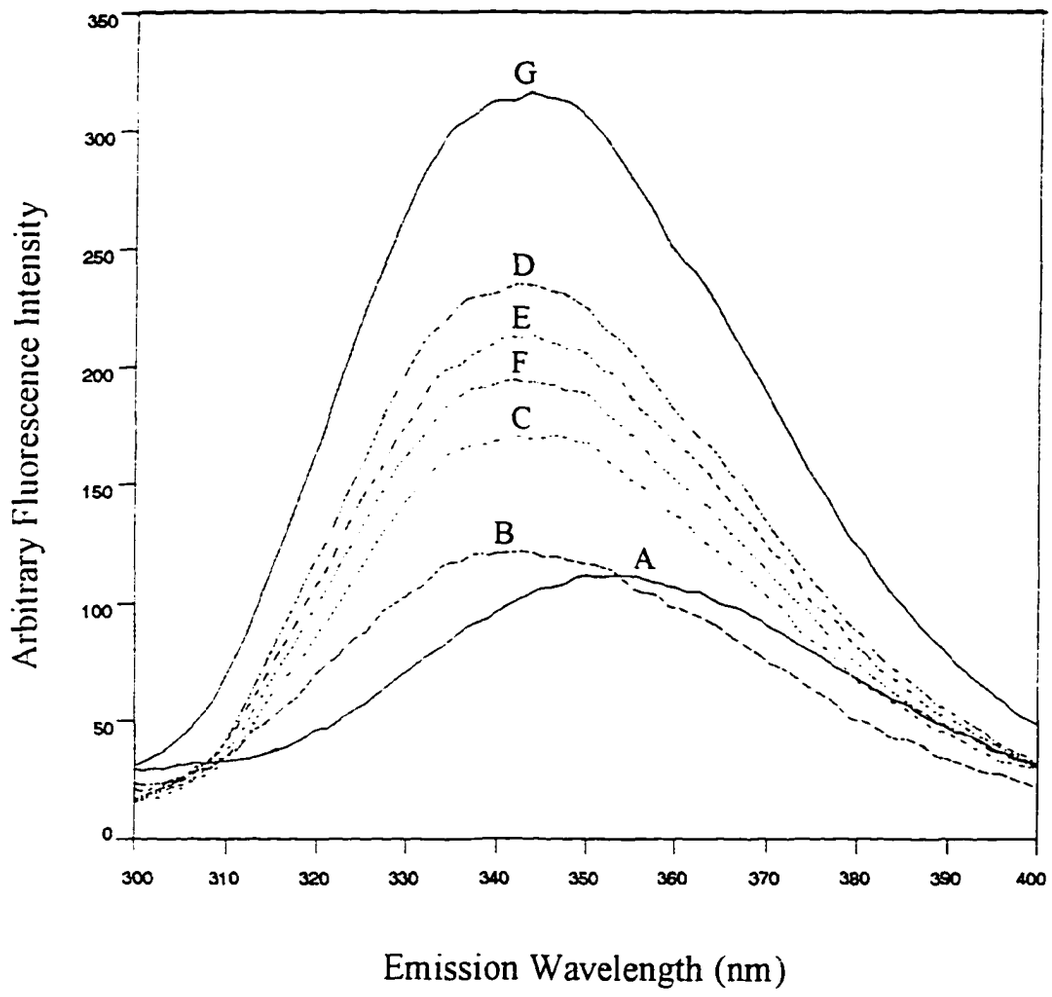

Figure 2.4. Monitoring refolding of tailspike protein by tryptophan fluorescence spectrophotometry at 10°C. All protein concentrations were the same at 22 μg/ml. Fluorescence emission spectra are shown for (A) denatured tailspike, (B) 1 minute, (C) 15 minutes, (D) 3 hours, (E) 24 hours, (F) 48 hours after the initiation of tailspike refolding, (G) native tailspike.

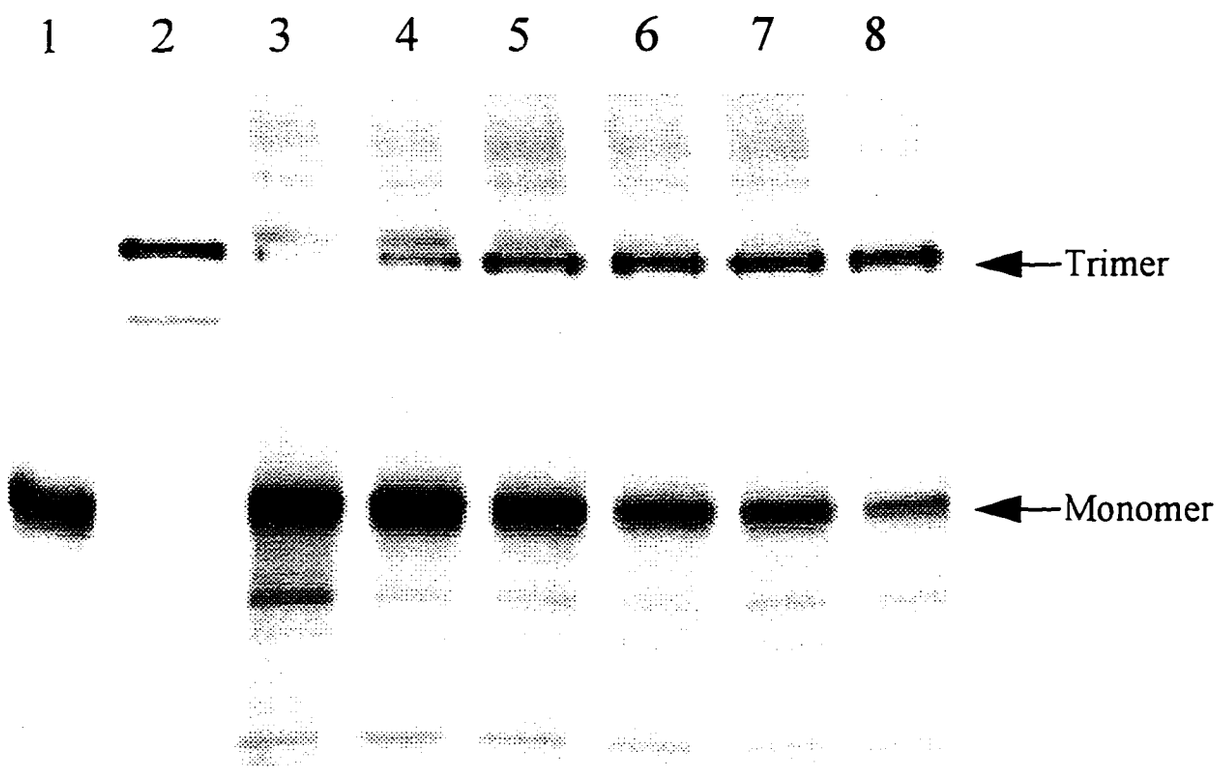

Figure 2.5. Monitoring refolding of tailspike protein by slab gel electrophoresis. Lane 1, denatured tailspike control; lane 2, native tailspike control at 22 μg/ml; lanes 3-8, refolding samples taken from 1 hour, 2 hours, 6 hours, 10 hours, 24 hours, and 48 hours after the initiation of tailspike refolding.

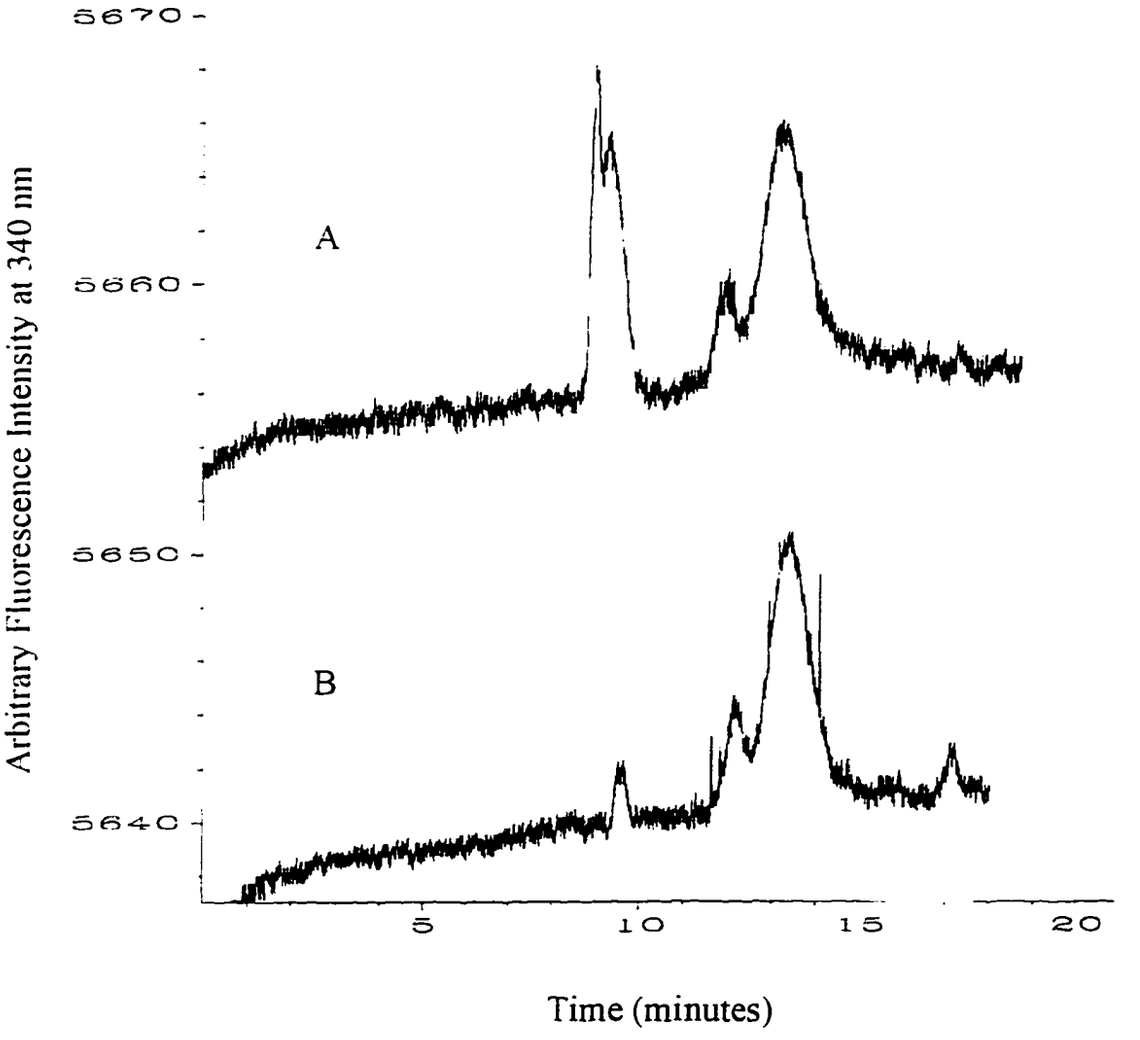

Figure 2.6. Effect of temperature shift on tailspike refolding. (A) refolding occurred at 10°C for 45 minutes, and then shifted to 30°C for 30 minutes. (B) refolding occurred at 10°C for 2 hours and then shifted to 30°C for 30 minutes. The separation conditions for performing capillary zone electrophoresis-laser induced fluorescence detection of samples (A) and (B) were the same as in Fig. 2.3.

CHAPTER 3. INVESTIGATING TEMPERATURE EFFECTS ON REFOLDING AND AGGREGATION OF A LARGE MULTIMERIC PROTEIN USING CAPILLARY ZONE ELECTROPHORESIS

A paper accepted to Analytical Chemistry

Pamela K. Jensen, Jonathan A. King, and Cheng S. Lee

Abstract

Capillary zone electrophoresis (CZE) equipped with laser-induced fluorescence detection (LIFD) allows rapid and sensitive identification of folding and aggregation intermediates of P22 tailspike endorhamnosidase using the intrinsic tryptophan fluorescence. Conversion of the protrimer into the native tailspike protein, which is the rate-limiting step, can be quenched at 0°C and results in a loss of the protrimer in the refolding pathway. Results obtained from CZE-LIFD and polyacrylamide gel electrophoresis indicate the loss of folding intermediates through the adsorption of polypeptide chains onto the wall of the refolding vial.

Early aggregation intermediates are studied using a temperature-sensitive folding tailspike mutant which shifts productive folding to the aggregation pathway at restrictive temperatures. By monitoring refolding intermediates during the temperature shift experiments, the early monomeric intermediate is identified as the

thermolabile junctional intermediate between the productive and aggregation pathways. After passing the stage for accumulation of thermolabile intermediate, an increase in refolding temperature from 10 to 35°C is carried out for achieving the optimal refolding kinetics and yields of tailspike endorhamnosidase.

Introduction

A critical problem both in the biotechnology industry and in biomedical research is the incorrect folding and association of newly synthesized polypeptide chains, resulting in formation of insoluble aggregates. It has been particularly, though not exclusively, associated with expression of cloned genes in heterologous hosts[1-3]. Studies of both refolding of denatured proteins *in vitro* and *in vivo* folding pathways, indicate that aggregates are formed from partially folded intermediates and not from mature native or fully unfolded proteins[4].

Many proteins can fold *in vitro* from a completely denatured state[5,6], providing evidence that the primary structure of a polypeptide chain may contain all the information required for folding. However, the rules governing the mechanisms by which the polypeptide sequence dictates the three dimensional conformation are not fully understood. The study of protein refolding is necessary to understand these underlying mechanisms, as well as for the prediction of protein structure.

Most techniques for characterizing protein folding reactions depend on direct measurement of some average physical properties that are sensitive to changes in protein structure. Examples include changes in optical density and steady state

fluorescence intensity[7-11]. Capillary zone electrophoresis (CZE) is an attractive alternative to experimental observation and provides access to a cross section of the population of molecular states within an equilibrium system during a protein refolding reaction. Additionally, separation in CZE is a function of intrinsic molecular properties including size and charge of macromolecules.

On the basis of CZE separation, the temperature-dependent unfolding transitions of lysozyme at low pH and the unfolding process of human serum transferrin in urea have been studied by Hilser et al.[12] and Kilar and Hjerten[13], respectively. CZE-UV detection was also utilized for the analysis of bovine trypsinogen, which underwent oxidation from a fully reduced molecule through a distribution of intermediate species until it reached the disulfide bond conformation corresponding to the native structure[14].

In our laboratory, CZE equipped with laser-induced fluorescence detection (LIFD) provided a fast sensitive means of analyzing folding and aggregation intermediates of phage P22 tailspike endorhamnosidase using the intrinsic tryptophan fluorescence[15]. The P22 tailspike endorhamnosidase is a homotrimer consisting of three 666 amino acid polypeptide chains. Each of the 72-kD monomeric subunits within the trimer is mainly comprised of a parallel b-helix structure[16]. The native trimer is thermostable, with a T_m of 88°C, and is resistant to sodium dodecyl sulfate (SDS) and protease[17,18]. Upon initiation of refolding, tailspike polypeptides rapidly fold into structured monomeric intermediates with a high content of secondary structure. These monomeric species associate to form

the triple-chain defined folding intermediates, the protrimers, that are susceptible to SDS and protease. Conversion of the protrimer into the native, SDS and protease-resistant tailspike protein is the rate-limiting step in the folding pathway[19,20].

The folding and aggregation pathways of P22 tailspike endorhamnosidase were well characterized both *in vivo*[17,21-23] and *in vitro*[19,20,24,25]. It has been stated that refolding of tailspike could essentially be quenched by lowering the temperature to 0°C[21,26], thus trapping the intermediates in a detergent-sensitive state. However, little is known about how such low temperatures may affect the distribution of folding intermediates and the competition between productive folding and aggregation reactions. Furthermore, an early monomeric intermediate of tailspike protein is thermolabile and has an increased tendency to form aggregates at higher temperatures[23,24,26]. This is due in part to the higher collisional frequencies and the changes in protein conformation[27].

In this study, CZE equipped with tryptophan fluorescence detection allows rapid identification of transient partially folded intermediates formed during tailspike refolding and aggregation at various folding temperatures. Additionally, a temperature-sensitive folding (tsf) tailspike mutant which acts to further destabilize an early folding intermediate at restrictive temperatures (at or above 39°C) shifts productive folding to the aggregation pathway before association into protrimer[22]. CZE monitoring of aggregation intermediates and pathways is demonstrated using a tsf mutant at restrictive temperatures. Optimization of refolding kinetics and yields measured by CZE is established on the basis of temperature shifts. In comparison

to the results obtained from SDS and native polyacrylamide gel electrophoresis (PAGE), CZE provides an effective technique for investigating the effects of folding temperatures on the distribution and competition among misfolding (aggregation) and productive folding.

Experimental

CZE-LIFD

The CZE apparatus was constructed in the laboratory using a CZE 1000R high voltage power supply (Spellman High-Voltage Electronics, Plainview, NY). Fused silica capillaries of 50 μm i.d. x 150 μm o.d. were obtained from Polymicro Technologies (Phoenix, AZ). The capillaries were cut to a length of 50 cm with a separation distance of 27 cm between the injection point and the LIFD system. A negative electric potential was applied at the detector end of the capillary for electrokinetic injection and for electrophoretic separation while the injection end was grounded. CZE separations were carried out in 25 mM sodium phosphate-140 mM urea at pH 7.6. The CZE apparatus was enclosed in an interlock box for operator safety.

A cooling jacket was utilized to maintain a constant temperature of 10°C at the inlet buffer reservoir and around most of the separation capillary during refolding measurements. The LIFD system originally designed by Yeung and co-workers[28] was employed for measuring tryptophan fluorescence of refolding intermediates

separated by CZE. The 248 nm line of a KrF excimer laser (Potomac Photonics, Lanham, MD) was used as the excitation source and fluorescence emission was collected at 340 nm using a photomultiplier tube (Hamamatsu, Bridgewater, NJ). CZE-LIFD equipped with cooling jacket was described in detail elsewhere[15].

Protein Denaturation and Refolding

The P22 tailspike endorhamnosidase was unfolded at room temperature in a solution of 6.4 M urea and 25 mM sodium phosphate at pH 3 for 1.5 hours. Refolding was initiated by rapidly diluting the denatured protein with 25 mM sodium phosphate buffer at pH 7, pre-equilibrated at the desired refolding temperature. Based on a 45-fold dilution, the final tailspike and urea concentrations in the refolding reaction were 33 μg/ml and 142 mM, respectively.

The progress of the refolding reaction was monitored by subjecting an aliquot of refolding sample to CZE separation at various reaction times. Refolding samples were also analyzed by both SDS and native PAGE. For SDS-PAGE analysis, 20 μl aliquots of refolding sample were placed in vials containing 10 μl of SDS buffer (0.16 M Tris-HCl at pH 6.8, 6% SDS, 6% dithiothreitol (DTT), and 18% glycerol) and were kept at 4°C until electrophoresed. For native PAGE analysis, 20 μl aliquots of refolding sample were placed in vials containing 10 μl of native buffer (15 mM Tris, 0.11 M glycine, and 30% glycerol) at 0°C and the mixtures were immediately electrophoresed.

Slab Gel Electrophoresis

A Bio-Rad Mini-Protean II dual slab cell (Hercules, CA) was used to perform all SDS and native PAGE experiments. The separation and stacking gels were cast as described in the instruction manual. Acrylamide concentrations for the separating gels were 7.5% for SDS and 9% for native, while the stacking gels were 3% for SDS and 4.3% for native. SDS gels were run at 150 V and electrophoresis stopped when the bromophenol blue dye front reached about 1 cm from the gel bottom. Native PAGE was carried out at 200 V for 2-3 hours at 4°C. Silver staining of protein bands was performed as described elsewhere[29]. Quantitation of the gel bands was achieved by scanning the dry gels and analyzing the intensities using the NIH Image software.

Chemicals

The P22 tailspike endorhamnosidase and one of its tsf mutants, H304, were prepared and purified by standard procedures in King and Yu[30]. Ultrapure urea (ICN, Aurora, OH) was used for all denaturation and refolding experiments. Monobasic and dibasic sodium phosphate were purchased from Fisher (Fair Lawn, NJ). DTT, glycine, glycerol, and SDS were obtained from Sigma (St. Louis, MO). All solutions were prepared using water purified by a NANOpure II system (Dubuque, IA) and further filtered with a 0.22 μm membrane (Costar, Cambridge, MA).

Results and Discussion

Productive Refolding of Tailspike Endorhamnosidase at 10°C

Tailspike refolding was initiated by rapidly diluting denatured protein with refolding buffer of 25 mM sodium phosphate (pH 7) at 10°C. The final tailspike and urea concentrations in the refolding reaction were 33 μg/ml and 142 mM, respectively. The refolding reaction was monitored using CZE-LIFD by electrokinetically injecting an aliquot of sample at various times after initiation of refolding. The early eluting peak corresponded to the formation of structured monomers which appeared rapidly at the beginning of reconstitution (Fig. 3.1A) and gradually decreased as the refolding reaction proceeded. A second, later eluting peak quickly emerged in the electropherogram (Fig. 3.1B) and the fluorescence intensity of the peak continued to increase until reaching a maximum around the refolding time of 1-2 h (Fig. 3.1C). The late eluting peak was formed by association from the monomeric polypeptides, and represented a precursor (the protrimer) of the native tailspike protein trimer[17,20].

The conversion of the protrimer into the native, SDS resistant tailspike protein is the rate-limiting step in the folding pathway[19,20] and occurs very slowly at temperatures below 20°C[17]. After the first 3-4 h of refolding, an additional peak arose at a migration time between the structured monomer and the protrimer (Fig. 3.1D) and the peak intensity continued to increase during the rest of the refolding reaction (Fig. 3.1E-G). This peak was rationalized to be the mature trimer and was

confirmed by spiking with a native tailspike standard (data not shown). Furthermore, the decrease in tryptophan fluorescence intensity of the protrimer was accompanied by the accumulation of native tailspike. The observed refolding kinetics in the formation of monomeric and trimeric intermediates, and native tailspike were in good agreement with those reported in the literature[19,20,24].

Quenching of Tailspike Refolding at 0°C

To better characterize the protrimer, the in vitro folding reaction was quenched by lowering the temperature rather than by denaturing all of the precursors with SDS. The refolding was carried out by diluting the denatured protein with refolding buffer equilibrated at 10°C and was allowed to refold for 1 h. At this time, the distribution of folding intermediates would largely be at the protrimer stage with a residual amount of structured monomer and very little mature trimer yet formed. The refolding temperature was then reduced to 0°C using a thermostatted water bath. Thus, the concentrations of folding intermediates present before the temperature shift should remain essentially constant and no further production of native tailspike should occur[21,25].

Tailspike endorhamnosidase was allowed to refold for 1 h at 10°C (Fig. 3.2A). Upon shifting the refolding temperature to 0°C and continuing the refolding reaction for another 1 h, the amount of protrimer, instead of remaining constant, was noticeably reduced (Fig. 3.2B). The tryptophan fluorescence intensity of the protrimer peak continued to decrease over a period of several hours at 0°C (Fig.

3.2C-D). At the same time, there was a slight increase in the concentration of structured monomer as analyzed by CZE-LIFD. Conversion of the protrimer into the native, SDS-resistant tailspike protein was quenched at 0°C together with a loss of the protrimer in the refolding pathway.

Both native and SDS-PAGE were also employed for analyzing the same refolding samples. The native PAGE (Fig. 3.3) displayed three major bands corresponding to the protrimer, trimer, and monomer components of tailspike endorhamnosidase[21,31]. Upon reducing the refolding temperature to 0°C, the staining intensities of the protrimer and trimer bands continuously decreased for a total of 6 h refolding at 0°C (Lanes 3-8). By comparing the band intensities shown in Lanes 2 and 8, the temperature shift from 10 to 0°C resulted in an approximate loss of 47% for protrimer and 52% for trimer. A minor protein band which migrated ahead of the trimer was attributed to a dimer species[25,32]. The concentrations of both the monomer and dimer bands increased during tailspike refolding at 0°C. By adding up the band intensities of the monomer, dimer, protrimer, and trimer in Lanes 2 and 8, an approximate 25% reduction in total protein concentration was measured during the temperature shift experiments.

A very faint protein band corresponding to the native, SDS-resistant tailspike was observed and remained relatively constant in SDS-PAGE (Fig. 3.4). A refolding yield of ~5% was estimated for the refolding sample taken from 6 h after the temperature shift from 10 to 0°C (Lane 7). It has been reported that an additional species whose mobility appeared to be that of native trimer in native PAGE was

possibly an intermediate between the protrimer and the native trimer[32]. This intermediate was still SDS-sensitive and would be denatured into monomer in SDS-PAGE. This would explain the higher trimer concentration measured in native PAGE (Fig. 3.3) than those by CZE-LIFD (Fig. 3.2) and SDS-PAGE (Fig. 3.4).

Misfolding aggregates as well as those refolding species that have not yet matured into native trimers, including single-chain and protrimer intermediates, were denatured by the detergent and easily distinguished from the native tailspike. The monomer band on the SDS gel (Fig. 3.4) appeared to decrease gradually (Lanes 3-7), ending with a concentration that was 23% less than that present before the temperature shift (Lane 3). This large decrease in polypeptide concentration analyzed by SDS-PAGE was in good agreement with the results of native PAGE and CZE-LIFD. The blockage in the conversion of the protrimer into the native tailspike at 0°C shifted the folding equilibrium toward the monomeric intermediates and promoted the loss of polypeptide chains through the adsorption of monomeric species onto the wall of the refolding vial. In fact, the SDS buffer was added to the vial at the end of a refolding reaction to recover any adsorbed protein chains. An intense monomer band did appear on the SDS-PAGE gel (data not shown), indicating the adsorption of tailspike protein onto the vial.

To rescue tailspike protrimer and shift the folding equilibrium toward the formation of native tailspike, the refolding temperature was increased from 0 to 10°C after the refolding reaction was quenched at 0°C for 2 h (Fig. 3.5A-C). As shown in Fig. 3.5D-G, the productive folding pathway was resumed and the conversion of the

protrimer into the native tailspike resulted in a refolding yield of ~40%. This refolding yield was about 10% less than that obtained from the refolding reaction at 10°C for 24 h (Fig. 3.1G), indicating an irreversible loss of structured monomer and protrimer during the temperature shift experiments. The refolding kinetics and yields measured by CZE-LIFD were in good agreement with those obtained via SDS and native PAGE (data not shown).

Competition between Productive and Aggregation Pathways at High Temperatures

Refolding yield of tailspike endorhamnosidase was strongly dependent on the reaction temperature. High refolding temperatures promoted protein self-association and aggregation. To monitor the early and rapid folding events at high temperatures, the refolding reaction was initiated by diluting denatured protein with refolding buffer pre-equilibrated at 35°C and the refolding sample was immediately injected into the capillary. The refolding sample was re-injected three times, each at intervals of 3 min, and continuously analyzed using CZE-LIFD (Fig. 3.6A). Besides the peaks corresponding to the folding intermediates of structured monomer and protrimer, there was another distinct species migrating after the protrimer.

The accumulation of native tailspike, together with the disappearance of protrimer, already occurred within 30 min after the initiation of refolding at 35°C (Fig. 3.6B), indicating faster reaction kinetics at high temperatures. The refolding yield obtained at 1 h after the initiation of refolding (Fig. 3.6C) was about 26% less than that of tailspike refolding at 10°C. The low refolding yield was attributed to the

fact that more polypeptide chains were shunted to the off-pathway aggregation versus the productive refolding pathway.

Two partially resolved components were observed in the electropherogram with migration times past those of native tailspike and protrimer. These components might be the early intermediates in the aggregation pathway. To further investigate the identity of these two additional peaks, a refolding reaction was carried out using a tsf tailspike mutant, designated H304. This tailspike mutant acts to further destabilize an early folding intermediate at restrictive temperatures (at or above 39°C) and shifts productive folding to the aggregation pathway before association into protrimer[22]. By refolding H304 for 2 h at 39°C, a very small amount of native tailspike was formed, barely distinguishable above the background (Fig. 3.7A). At the same time, similar aggregate peaks as the ones shown in Fig. 3.6 were observed in the electropherogram.

To optimize both refolding kinetics and yields of tailspike endorhamnosidase, the refolding reaction was initiated at 10°C for 12 min. The refolding temperature was then increased to 35°C for the rest of the refolding reaction. After the temperature shift, the refolding sample was injected four times, each at intervals of 3 min, and continuously monitored using CZE-LIFD (Fig. 3.8A). The distribution of folding components, including monomer, protrimer, and native tailspike changed rapidly during the first 30 min after the temperature shift (Fig. 3.8A-B). The refolding yield at 1 h after the temperature shift (Fig. 3.8C) was estimated around 75% and was the best yield over those obtained at constant 10°C (Fig. 3.1G) and

35°C (Fig. 3.6C).

It is known that an early monomeric intermediate of tailspike endorhamnosidase is thermolabile and has an increased tendency to form aggregates at higher temperatures[23,24,26]. Thus, the refolding reaction was initiated at 10°C during the formation and accumulation of early thermolabile monomer species. Once the folding reaction has reached the stage for the production of association competent monomer, the folding temperature was increased to 35°C for the enhancement of multimer formation including the protrimer and native tailspike. Rapid multimer formation in the productive folding pathway also reduced the amount of monomeric intermediates adsorbed onto the wall of the refolding vial and increased the overall refolding yield of native tailspike.

Conclusion

Rapid identification of transient partially folded intermediates formed during protein refolding and aggregation has been difficult, particularly with separation methods relying on solid matrices. CZE in combination with the intrinsic tryptophan fluorescence detection provides a fast sensitive means for identifying folding and aggregation intermediates, and for studying the effects of folding temperatures on the distribution and competition among misfolding (aggregation) and productive folding. In comparison to the results obtained from SDS and native PAGE, CZE is demonstrated as an effective technique for the analysis of tailspike refolding.

Tailspike refolding at low temperatures reduces the amount of aggregate

formation at the expense of slow refolding kinetics. The refolding yields at low temperatures suffer from the adsorption of monomeric species onto the wall of the refolding vial, instead of the generation of aggregated polypeptide chains. For tailspike and many other proteins, the off-pathway aggregation reaction proceeds from an early thermolabile intermediate at the junction with the productive pathway. By monitoring refolding intermediates during the temperature shift experiments, the early monomeric intermediate is identified as the thermolabile junctional intermediate between the productive and aggregation pathways for tailspike endorhamnosidase. Optimal refolding kinetics and yields of tailspike protein are obtained by initiating the refolding reaction at low temperature and increasing the refolding temperature after passing the stage for formation and accumulation of the thermolabile intermediate.

Acknowledgments

Special thanks to Dr. Anne Robinson (Massachusetts Institute of Technology) for providing P22 tailspike endorhamnosidase and for helpful discussions. This work was supported by the Biotechnology Process Engineering Center at Massachusetts Institute of Technology through the National Science Foundation's Engineering Research Center Initiative under the Cooperative Agreement CDR-8803014. P.K.J. is a recipient of the Graduate Assistance in Areas of National Need (GAANN) Fellowship from the U.S. Department of Education. C.S.L. is a National Science Foundation Young Investigator (BCS-9258652).

References

1. Marston, F. A. O. *Biochem. J.* **1986**, *240*, 1.

2. DeBernardez-Clark, E.; Georgiou, G. In *Protein Refolding;* Georgiou, G, DeBernardez-Clark, E., Eds.; American Chemical Society: Washington, D. C., **1991**; ACS Symposium Series 470, Chapter 1.

3. King, J.; Haase-Pettingell, C.; Gordon, C.; Sather, S.; Mitraki, A. In *Protein Folding: in vivo and in vitro*; Cleland, J. L., Ed.; American Chemical Society: Washington, D. C., **1993**; ACS Symposium Series 526, Chapter 2.

4. Mitraki, A.; King, J. *Bio/Technology* **1989**, *7*, 690.

5. Creighton, T. E. *Proteins, Structures and Molecular Properties;* W. H. Freeman and Company: New York, **1983**.

6. Richards, F. M. *Scientific American* **January 1991**, 54.

7. Hagerman, P. J.; Baldwin, R. L. *Biochemistry* **1976**, *15*, 1462.

8. Labhart, A. M. *Proc. Natl. Acad. Sci. USA* **1984**, *81*, 7674.

9. Galat, A.; Yang, C. C.; Blout, E. R. *Biochemistry* **1985**, *24*, 5678.

10. Zuniga, E. H.; Nall, B. T. *Biochemistry* **1983**, *22*, 1430.

11. Kelley, R. F.; Wilson, J.; Bryant, C.; Stellwagen, E. *Biochemistry* **1986**, *25*, 728.

12. Hilser, V. J.; Worosila, G. D.; Freire, E. *Anal. Biochem.* **1993**, *208*, 125.

13. Kilar, F.; Hjerten, S. *J. Chromatogr.* **1993**, *638*, 269.

14. Strege, M. A.; Lagu, A. L. *J. Chromatogr. A* **1993**, *652*, 179.

15. Jensen, P. K.; Fan, Z. H.; King, J.; Lee, C. S. *J. Chromatogr. A* **1997**, *769*, 315.

16. Steinbacher, S.; Seckler, R.; Miller, S.; Steipe, B.; Huber, R.; Reinemer, P. *Science* **1994**, *265*, 383.

17. Goldenberg, D.; King, J. *J. Mol. Biol.* **1981**, *145*, 633.

18. Sturtevant, J. M.; Yu, M.-H.; Haase-Pettingell, C.; King, J. *J. Biol. Chem.* **1989**. *264*, 10693.

19. Seckler, R.; Fuchs, A.; King, J.; Jaenicke, R. *J. Biol. Chem.* **1989**, *264*, 11750.

20. Fuchs, A.; Seiderer, C.; Seckler, R. *Biochemistry* **1991**, *30*, 6598.

21. Goldenberg, D.; King, J. *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 3403.

22. Goldenberg, D. P.; Smith, D. H.; King, J. *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 7060.

23. Haase-Pettingell, C. A.; King, J. *J. Biol. Chem.* **1988**, *263*, 4977.

24. Mitraki, A.; Danner, M.; King, J.; Seckler, R. *J. Biol. Chem.* **1993**, *268*, 20071.

25. Speed, M. A.; Wang, D. I. C.; King, J. *Protein Sci.* **1995**, *4*, 900.

26. Goldenberg, D. P.; Berget, P. B.; King, J. *J. Biol. Chem.* **1982**, *257*, 7864.

27. Zettlmeissl, G.; Rudolph, R.; Jaenicke, R. *Biochemistry* **1979**, *18*, 5567.

28. Yeung, E. S.; Wang, P.; Li, W.; Giese, R. W. *J. Chromatogr.* **1992**, *608*, 73.

29. Sather, S. K.; King, J. *J. Biol. Chem.* **1994**, *269*, 25268.

30. King, J.; Yu, M.-H. *Methods in Enzymology* **1986**, *131*, 250.

31. Speed, M. A.; Morshead, T.; Wang, D. I. C.; King, J. *Protein Sci.* **1997**, *6*, 99.

32. Robinson, A. S.; King, J. *Nature Structural Biology* **1997**, *4*, 450.

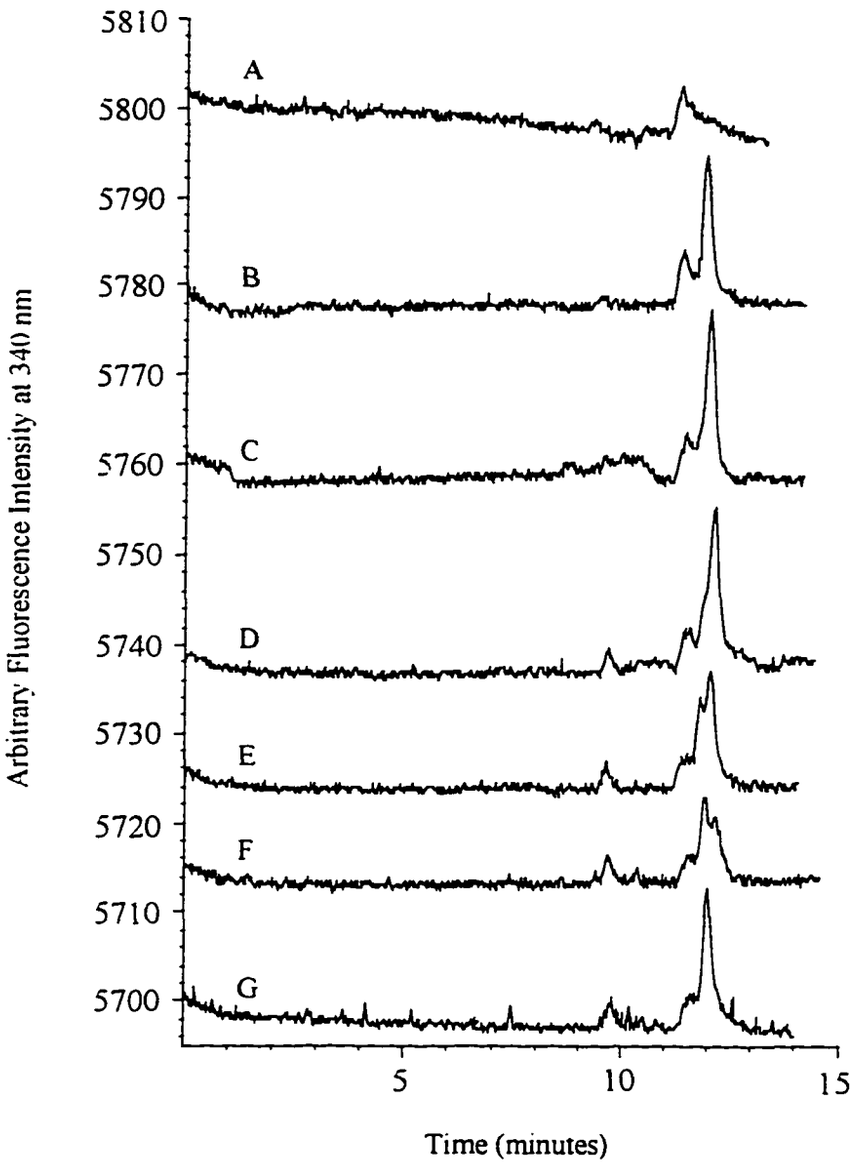

Figure 3.1. Monitoring refolding of tailspike protein at 10°C by CZE-LIFD. Refolding samples were taken at (A) 1 min, (B) 1 h, (C) 2 h, (D) 4 h, (E) 6 h, (F) 8 h and (G) 24 h after the initiation of refolding. Electrophoresis buffer, 25 mM sodium phosphate and 142 mM urea at pH 7.6; capillary, 50 cm (27 cm to detector) x 50 μm i.d. x 150 μm o.d.; applied voltage, -6.5 kV and 10 s for refolding sample injection, -6.5 kV for electrophoresis; total tailspike protein concentration, 33 μg/ml; fluorescence detection at 340 nm.

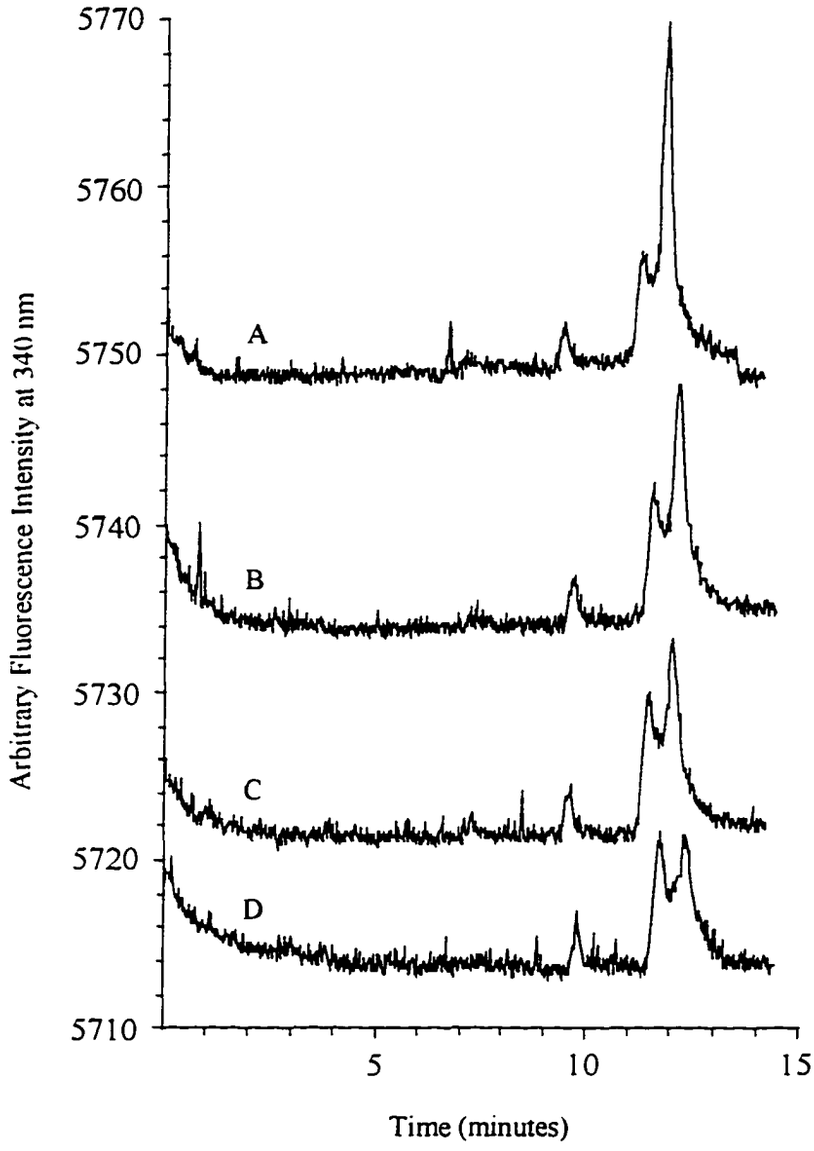

Figure 3.2. Quenching of tailspike folding reaction by reducing the refolding temperature from 10 to 0°C. Refolding reaction was monitored by CZE-LIFD. Refolding sample was taken at (A) 1 h after the initiation of refolding at 10°C. The temperature was then shifted to 0°C and the reaction continued for (B) 1 h, (C) 2 h and (D) 4 h after the shift. All other conditions were the same as in Fig. 3.1.

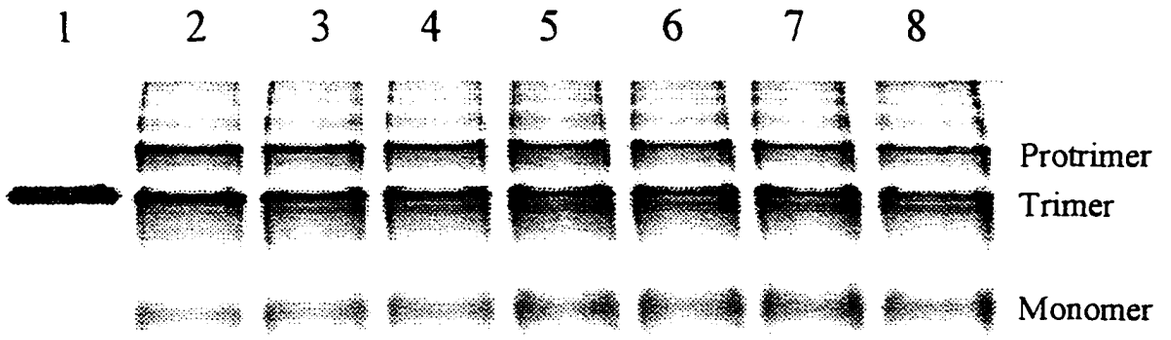

Figure 3.3. Monitoring refolding of tailspike protein during the temperature shift experiment using native PAGE. Lanes: 1 = native tailspike control; 2 = 1 h after the initiation of refolding at 10°C; 3-8 = refolding samples taken from 1 h, 2 h, 3 h, 4 h, 5 h and 6 h after the temperature shift from 10 to 0°C.

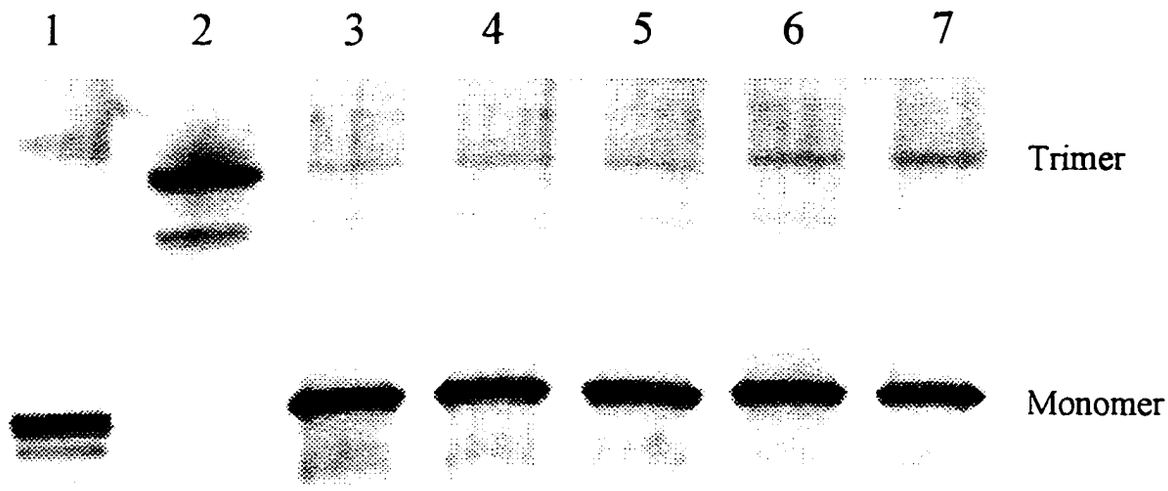

Figure 3.4. Monitoring refolding of tailspike protein during the temperature shift experiment using SDS-PAGE. Lanes: 1 = denatured tailspike control; 2 = native tailspike control; 3 = 1 h after the initiation of refolding at 10°C; 4-7 = refolding samples taken from 1 h, 2 h, 3 h and 6 h after the temperature shift from 10 to 0°C.

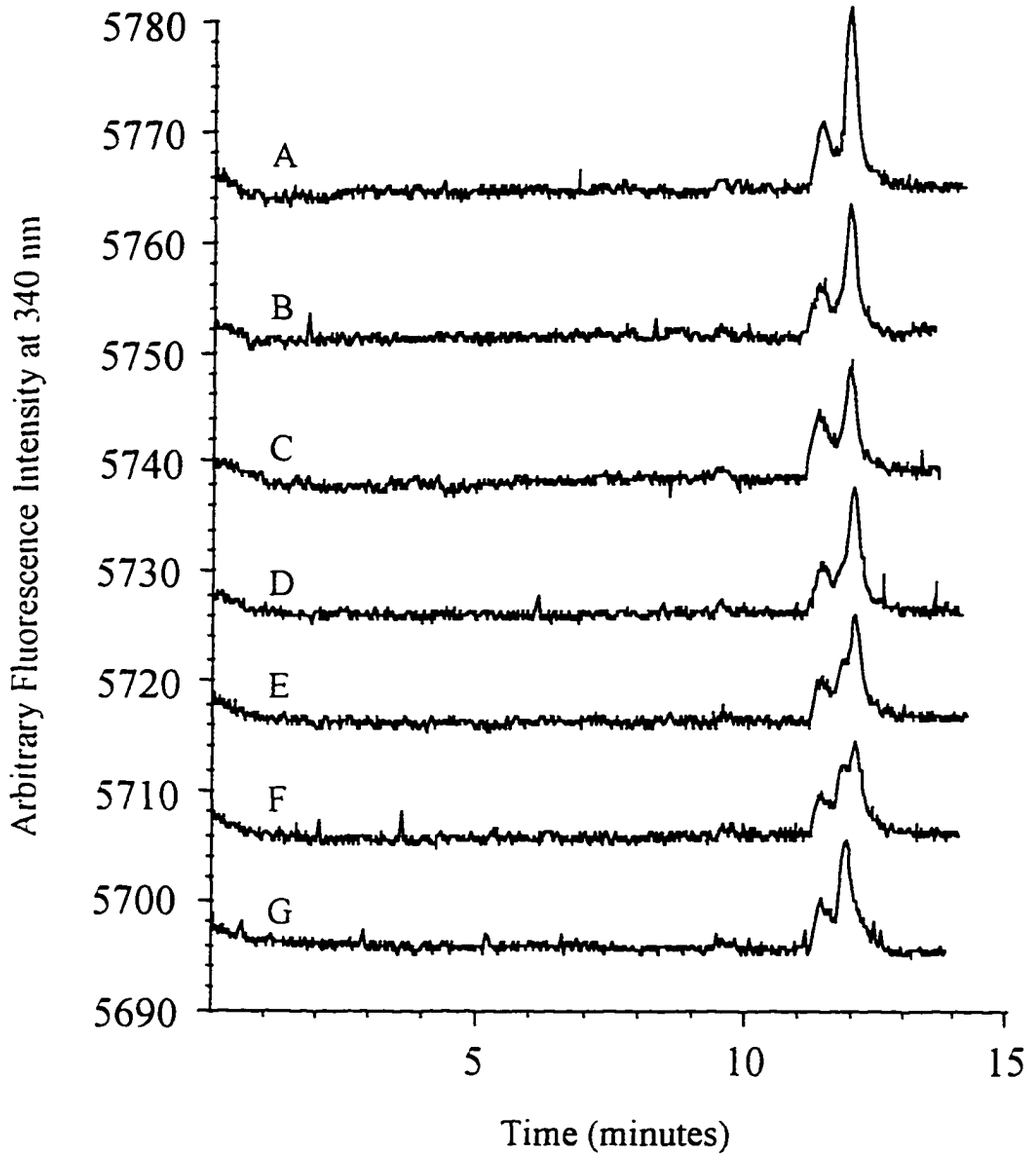

Figure 3.5. Rescue of tailspike protrimer by increasing the refolding temperature from 0 to 10°C. Refolding reaction was monitored by CZE-LIFD. Refolding sample was taken at (A) 1 h after the initiation of refolding at 10°C. The temperature was then shifted to 0°C and the reaction continued for (B) 1 h and (C) 2 h after the shift. The temperature was increased back to 10°C and the reaction proceeded for another (D) 1 h, (E) 3 h, (F) 5 h and (G) 24 h. All other conditions were the same as in Fig. 3.1.

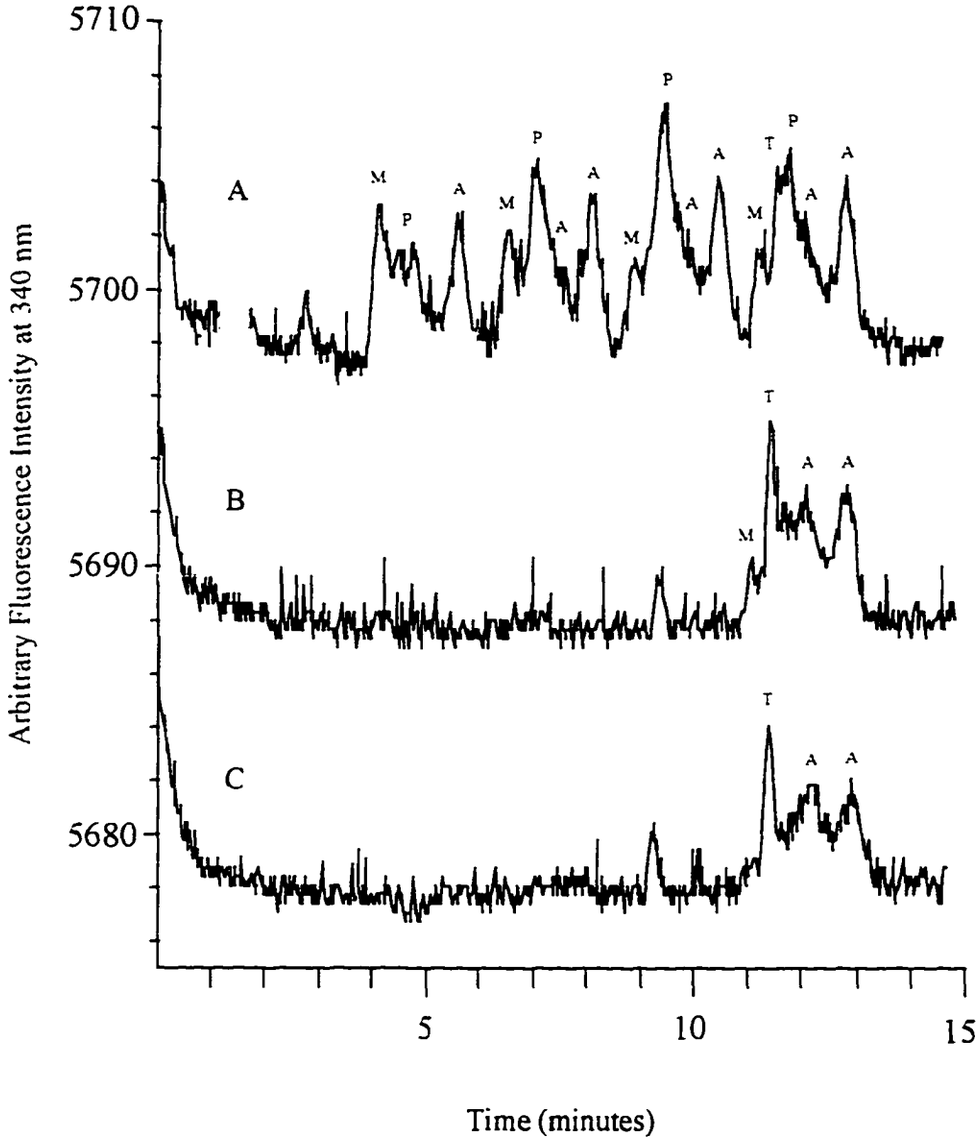

Figure 3.6. Monitoring refolding of tailspike protein at 35°C by CZE-LIFD. Refolding samples were taken at (A) 1 min, 3 min, 6 min and 9 min, (B) 30 min and (C) 1 h after the initiation of refolding. All other conditions were the same as in Fig. 3.1. The brackets in (A) represent the part of electropherogram removed for aligning the last sample injection in (A) with the electropherograms shown in (B) and (C). The migration order: M (monomer), T (native trimer), P (protrimer) and A (aggregate).

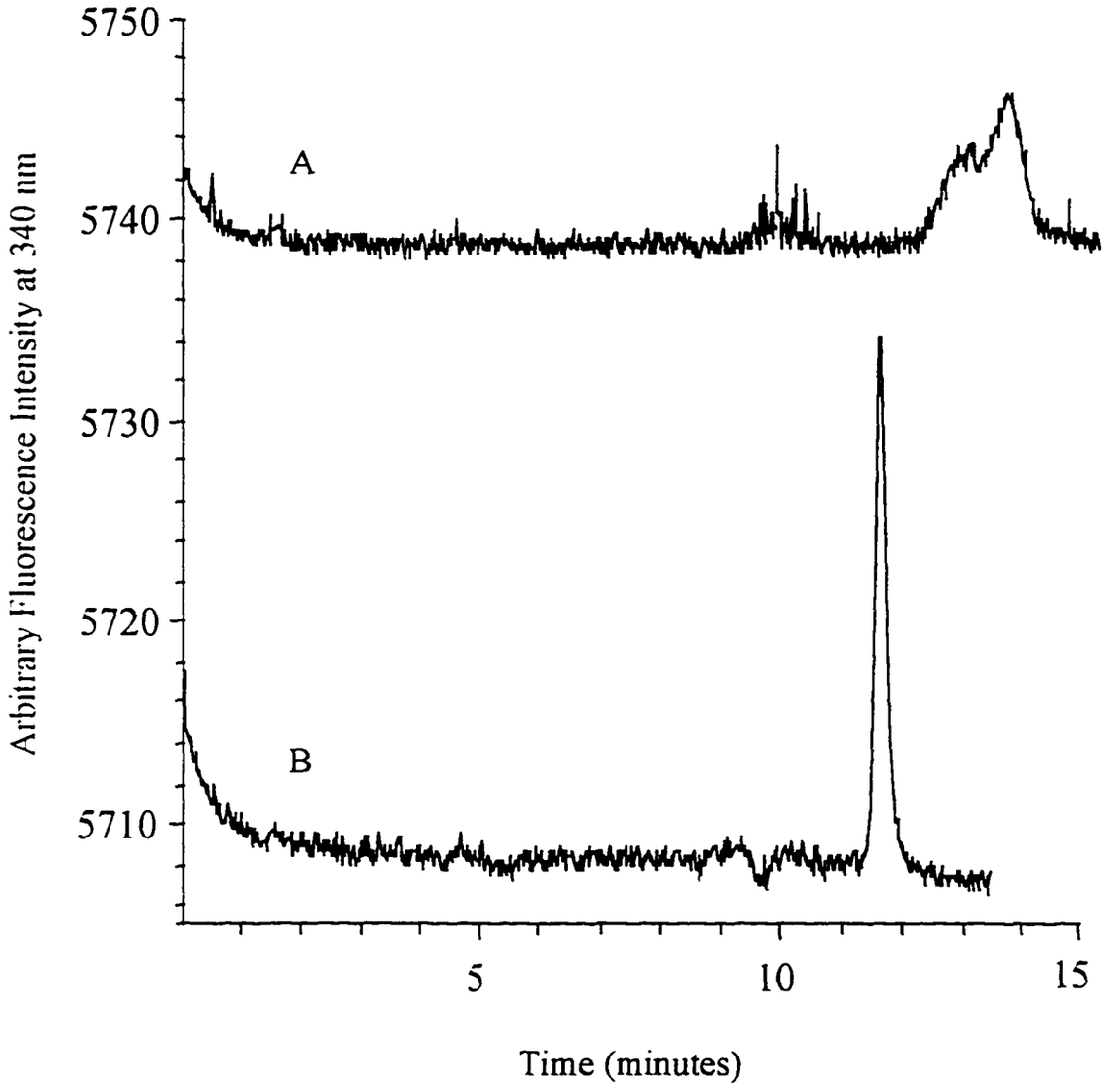

Figure 3.7. Monitoring refolding of tsf H304 tailspike mutant at 39°C by CZE-LIFD. (A) 2 h and (B) H304 native tailspike control. All other conditions were the same as in Fig. 3.1.

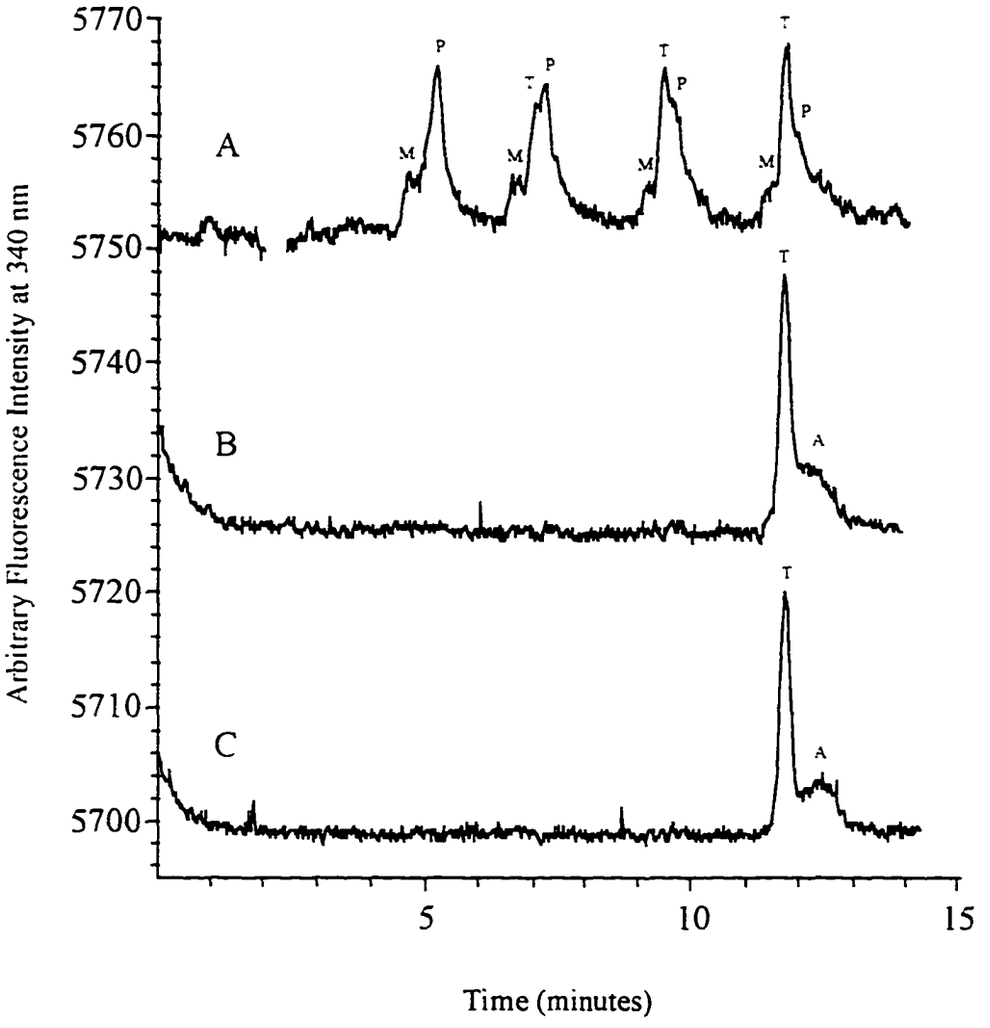

Figure 3.8. Optimization of refolding kinetics and yields by increasing the refolding temperature from 10 to 35°C after the passage of thermolabile monomeric intermediate. Refolding reaction was initiated at 10°C and monitored by CZE-LIFD. The temperature was then shifted to 35°C at 12 min after the initiation of refolding. Refolding samples were taken at (A) 3 min, 6 min, 9 min and 12 min, (B) 30 min and (C) 1 h after the temperature shift. All other conditions were the same as in Fig. 3.1. The brackets in (A) represent the part of electropherogram removed for aligning the last sample injection in (A) with the electropherograms shown in (B) and (C). The migration order: M (monomer), T (native trimer), P (protrimer) and A (aggregate).

CHAPTER 4. MONITORING PROTEIN REFOLDING INDUCED BY DISULFIDE FORMATION USING CAPILLARY ISOELECTRIC FOCUSING-ELECTROSPRAY IONIZATION MASS SPECTROMETRY

A paper submitted to Analytical Chemistry

Pamela K. Jensen, A. Kamel Harrata, and Cheng S. Lee

Abstract

Rapid growth in the biotechnology industry has led to a dramatic increase in attention to the protein folding problem. Understanding protein folding pathways is essential to the production of biopharmaceuticals since commercial production of recombinant proteins often requires a protein refolding process for recovery of high yields.

Protein folding coupled to the formation of disulfide bonds presents one of the simplest approaches to studying folding intermediates. On-line capillary isoelectric focusing-electrospray ionization mass spectrometry (CIEF-ESIMS) is demonstrated for kinetic studies of disulfide bond induced protein refolding. Refolding intermediates of bovine pancreatic ribonuclease A (RNase A), a model system for this study, are blocked at different stages by alkylating free thiols with iodoacetate. The alkylation reaction results in the introduction of charge (-1) and mass (58)

differences for each alkylation site, providing the means for predictable separation and direct identification of refolding intermediates using CIEF-ESIMS. Besides the observation of refolding intermediates containing different numbers of disulfide bonds and even mixed disulfides, the two-dimensional resolving power of CIEF-ESIMS allows the determination of conformational heterogeneity among groups of refolding intermediates.

Introduction

Since the days of Anfinsen's[1] pioneering work, there have been considerable efforts put forth to characterize protein folding, to identify intermediates responsible for rapid folding, and to define the folding pathways. Elucidation of protein folding pathways relies on measuring the structural properties of folding intermediates. More detailed information on protein folding intermediates can be obtained by measuring some physical properties that are sensitive to changes in protein structure. Examples include changes in optical density and steady state fluorescence intensity[2-6].

Protein folding coupled to the formation of disulfide bonds presents one of the simplest approaches for studying folding intermediates. Creighton and others[7-10] have employed this method for the study of various disulfide containing proteins, most notably bovine pancreatic trypsin inhibitor (BPTI) and bovine pancreatic ribonuclease A (RNase A). One advantage of disulfide induced protein folding is that disulfide formation between cysteines (thiols) is an oxidation/reduction process.

The redox potential in the reaction can be directly controlled through the use of thiol reagents such as dithiothreitol (DTT) or glutathione[11]. The most significant advantage of disulfide coupled protein folding comes from the ability to covalently block the free thiols at different stages in the folding pathway, kinetically trapping intermediates which can then be studied in more detail.

Blocking of the cysteines can be accomplished by several methods. The simplest approach is to lower the pH to 3-4 which suppresses disulfide rearrangements by protonation of the reactive thiolate anion[9]. Although it is a fast, unintrusive, and reversible method, the requirement in the maintenance of low solution pH, however, severely limits the selection of analytical methods for analyzing and characterizing the folding intermediates. Thiosulfonate reagents, such as 2-aminoethyl methanethiosulfonate (AEMTS), react extremely quickly and specifically to cysteines[10]. In addition to reversible reaction of AEMTS with cysteines, the advantage of increasing protein charge by +1 for each thiol modified provides a basis for separation of folding intermediates in a predictable manner. More commonly used are the alkylating reagents of iodoacetate (IAA) or iodoacetamide (IAM)[12,13].

Electrospray ionization mass spectrometry (ESIMS) has been employed in many studies for probing the changes in protein conformation. Chowdhury et al.[14] examined the acid unfolding of cytochrome c observed by changes in the charge distributions of the electrosprayed protein at different pHs. Mirza et al.[15] developed a special device to control the temperature just prior to electrospraying for

investigating the heat-induced denaturation of several proteins. The effect of disulfide bonds on protein conformation was examined by Loo and co-workers[16]. The reduction of disulfide bonds resulted in a more extended tertiary structure which allowed otherwise inaccessible sites to be protonated in ESIMS. Besides protein denaturation and unfolding, ESIMS provided a fast sensitive means of monitoring protein folding by detecting an increase in mass due to trapping of deuterons in hydrogen exchange experiments[17].

The identification of RNase A refolding intermediates containing different numbers of disulfide bonds and even mixed disulfides were demonstrated by Torella et al. using ESIMS[18]. However, ESIMS alone cannot resolve and determine heterogeneity among groups of folding intermediates. On the other hand, capillary zone electrophoresis (CZE)-UV detection was utilized for the analysis of bovine trypsinogen, which underwent oxidation from a fully reduced molecule through a distribution of intermediate species until it reached the disulfide bond conformation corresponding to the native structure[19]. Additionally, Rush et al.[20] observed a conformational change in α-lactalbumin upon raising the column temperature in CZE. The initially sharp peak of the native protein became broad and somewhat asymmetric when higher temperatures were applied, indicating the presence of both the folded and a conformationally altered species. The UV absorbance measurements suffer from the lack of structural information in comparison with ESIMS.

In this study, on-line capillary isoelectric focusing (CIEF)-ESIMS is introduced for monitoring protein refolding induced by disulfide bond formation. The *in vitro* folding of RNase A, a model system for this study, is blocked at different stages by alkylating free thiols with IAA. IAA is chosen for its ability to add a -1 charge for each alkylation site. The alkylated RNase A intermediates are focused and cathodically mobilized in a polyacrylamide-coated capillary. At the end of the CIEF capillary, the mobilized protein zones are analyzed by ESIMS using a coaxial sheath flow configuration. The increase in molecular mass and the reduction in pI of alkylated RNase A folding intermediates facilitate monitoring the refolding pathway and studying the refolding kinetics using CIEF-ESIMS.

Experimental

Materials and Chemicals

Bovine pancreatic RNase A, PD-10 desalting columns, and Pharmalyte 3-10 were obtained from Pharmacia (Uppsala, Sweden). Acetic acid, DTT, ethylene-diaminetetraacetic acid (EDTA), IAA, and oxidized and reduced forms of glutathione (GSSG and GSH) were acquired from Sigma (St. Louis, MO). Tris(hydroxymethyl)-aminomethane (Tris) and ultrapure urea were purchased from Bio-Rad (Hercules, CA) and ICN (Aurora, OH), respectively. N,N,N',N'-tetramethylethylenediamine (TEMED) was obtained from Aldrich (Milwaukee, WI). All solutions were prepared

using water purified by a NANOpure II system (Dubuque, IA) and further filtered with a 0.22 mm membrane (Costar, Cambridge, MA).

Protein Denaturation/Reduction, Refolding, and Alkylation

RNase A was completely denatured and reduced at a concentration of 25 mg/ml in a solution containing 8 M urea, 65 mM DTT, and 0.1 M Tris-HCl (pH 8.0) for approximately 2 hours at room temperature under a nitrogen atmosphere. The reduced RNase A was desalted using a PD-10 column equilibrated with 0.6% acetic acid. Eluted protein was vacuum dried and stored at -20°C.

The dried denatured RNase A was first dissolved in 0.6% acetic acid containing 0.1 mM EDTA. The solution pH was then adjusted to between 7.70 and 7.85 using 1.0 M Tris base. Oxidative refolding was initiated by adding a concentrated solution of equimolar GSH and GSSG which was degassed with nitrogen. Final concentrations of each reagent were 1-1.2 mg/ml of RNase A, 0.5 mM each of GSH and GSSG, 0.1 mM EDTA, and 0.1 M Tris-acetate (pH 7.70-7.85).

Reshuffling of disulfide bonds at room temperature was stopped at various refolding times by adding an aliquot of the refolding mixture into a concentrated IAA solution. The refolding sample was always diluted into the IAA solution to ensure that the folding intermediates were always surrounded with excess IAA reagent (0.1 M final concentration). The alkylation reaction was allowed to proceed for 3 minutes in the dark and the reaction mixture was immediately desalted using a PD-10

column equilibrated with 0.6% acetic acid. The alkylated (carboxymethylated) RNase A was then vacuum dried and stored at -20°C.

Mass Spectrometer and Electrospray Interface

The mass spectrometer was a Finnigan MAT TSQ 700 (San Jose, CA) triple quadrupole equipped with an electrospray ionization source. The Finnigan MAT electrospray adapter kit, containing both gas and liquid sheath tubes, was used to couple CIEF with ESIMS without any modifications. The first quadrupole was used for the mass scanning of protein ions, while the second and third quadrupoles were operated in the radio frequency-only mode. The electron multiplier was set at 1.4 kV, with the conversion dynode at -15 kV. Tuning and calibration of the mass spectrometer were established by using an acidic solution (methanol/water/acetic acid, 50:49:1 v/v/v) containing myoglobin and a small peptide of methionine-arginine-phenylalanine-alanine.

Capillary Isoelectric Focusing-Electrospray Ionization Mass Spectrometry

For on-line coupling of CIEF with ESIMS, a 30-cm-long (50 mm i.d. and 192 mm o.d.) fused silica capillary (Polymicro Technologies, Phoenix, AZ) coated with linear polyacrylamide[21] was mounted within the electrospray probe. The capillary was filled with a solution containing 0.5% Pharmalyte 3-10, 0.05% TEMED, and alkylated refolding sample. The outlet reservoir for the CIEF separation, containing 20 mM sodium hydroxide as the catholyte, was located inside the electrospray

housing during the focusing step. The inlet reservoir, containing 20 mM phosphoric acid as the anolyte, was kept outside at the same height as the outlet reservoir. Focusing was performed at a 10 kV constant voltage for approximately 10 minutes using a CZE 1000R high-voltage (HV) power supply (Spellman High-Voltage Electronics, Plainview, NY).

Once the focusing was completed, the electric potential was turned off, and the outlet reservoir was removed. The capillary tip was fixed about 1 mm outside the electrospray needle. A sheath liquid composed of 50% methanol, 49% water, and 1% acetic acid (v/v/v) at pH 2.6 was delivered at a flow rate of 3 ml/min using a Harvard Apparatus 22 syringe pump (South Natick, MA). During the mobilization step, two HV power supplies (Spellman) were used for delivering the electric potentials of 15 kV and 5 kV to the inlet electrode and electrospray needle, respectively. Detailed configuration of CIEF-ESIMS including sheath liquid and electrical connections was described elsewhere[22,23].

Gravity mobilization was combined with cathodic mobilization by raising the inlet reservoir 10 cm above the level of the electrospray needle. No sheath gas was employed during the CIEF-ESIMS measurements. The first quadrupole was scanned from m/z 1000 to m/z 2000 at a scan rate of 1.5 sec/scan. The deconvoluted mass spectra of protein analytes were obtained using the LC/MS BioToolBox analysis software from Perkin-Elmer (Foster City, CA).

Results and Discussion

Electrospray Ionization Mass Spectra of Native, Denatured and Reduced, and Alkylated RNase A

The mass spectra of native, denatured and reduced, and alkylated RNase A (Fig. 1) were obtained by electrospraying protein samples in an acetic acid solution (methanol/water/acetic acid, 50:49:1 v/v/v) at pH 2.6. The mass spectrum obtained from a native RNase A solution (Fig. 1A) exhibited five peaks, each one corresponding to a different protonation state of RNase A. These protonation states ranged from +7 to +11 with +9 being the most intense. The deconvoluted mass spectrum of native RNase A (Fig. 1A) displayed the molecular masses of 13,680 and 13,707 daltons. The molecular mass of 13,707 daltons was attributed to a sodium adduct of RNase A, due to the high salt content in the lyophilized RNase A sample.

The ability to produce a dramatic increase in the number of positive charges for disulfide-containing proteins such as lysozyme and bovine albumin upon the reduction of disulfide bonds with DTT was demonstrated by Loo et al.[16] using ESIMS. The enhanced protonation was presumably a result of allowing the protein molecule to attain a more extended tertiary structure, thus allowing otherwise inaccessible sites to be protonated. The addition of DTT and urea for the denaturation and reduction of RNase A allowed the molecule to acquire up to +19 charges (Fig. 1B). Upon denaturation and reduction, the protonation states

centered around +13 and the ion intensity of the distribution increased by approximately 10 fold relative to that of native RNase A electrospray mass spectrum. Reduction of disulfide bonds with subsequent carboxymethylation using IAA (Fig. 1C) yielded similar results, but with increase in m/z value due to alkylation. Since each carboxymethylation adds a mass of 58 daltons, the expected molecular mass of completely alkylated RNase A (four disulfide bonds and thus eight alkylation sites) would be 14,154 daltons. The deconvoluted masses of native, denatured and reduced, and alkylated RNase A measured from the electrospray mass spectra are summarized in Table I. A relative standard deviation of 0.01-0.04% between the expected and measured molecular masses of RNase A was demonstrated for the precise mass determination and quantitative alkylation of RNase A with IAA.

The observation of two discrete distributions of ions with no ions having intermediate charge states provided evidence for a highly cooperative transition between the native and denatured RNase A. These results are in concert with the earlier conclusion of Creighton[24]. Unfolding of native RNase A was an all-or-none transition to the fully reduced protein, with no accumulation of disulfide intermediates. On the other hand, the acid unfolding of cytochrome c was studied by Chowdhury et al.[14] using ESIMS. Three distinct conformational states of cytochrome c were observed for electrosprayed solutions in the pH range 2.6-5.2.

Refolding Monitoring of RNase A Using Capillary Isoelectric Focusing-Electrospray Ionization Mass Spectrometry

In Creighton's work, intermediates with one, two, three, or four disulfide bonds which accumulated during refolding of the reduced RNase A were trapped by rapid alkylation with IAA and separated by ion-exchange chromatography[24]. The intermediates in refolding were separated from the fully folded state by the highest energy barrier in the folding transition. Measured elements of the final conformation were not acquired during formation of the first three disulfide bonds, but appeared simultaneously with formation of the fourth native disulfide bond. Correctly refolded RNase A was indistinguishable from the original native protein.

In this study, oxidative refolding was initiated by rapidly adding a concentrated solution of equimolar GSH and GSSG into the unfolded protein solution. Immediately after the addition of GSH and GSSG, the refolding reaction was blocked by reacting the free thiols with IAA. As shown in Fig. 2A, three major RNase A refolding intermediates, corresponding to zero (0 S-S), one (1 S-S), and two (2 S-S) disulfide species, were resolved in CIEF based on their differences in pI. The reaction of free thiols with IAA added a -1 charge for each alkylation site, resulting in the decrease of pI and the increase of migration time of refolding intermediates. The numbers of carboxymethyl groups in the zero, one, and two disulfide species were eight, six, and four, respectively.

In CIEF-ESIMS, the refolding intermediates were directly identified on the basis of the mass spectra taken from the average scans under the peaks. The

mass spectra obtained from the average scans under the peaks of zero, one, and two disulfide species of RNase A are shown in Fig. 3 together with the deconvoluted masses. All three mass spectra exhibited similar charge state distribution of enhanced protonation, indicating an extended conformation of RNase A refolding intermediates.

Both the zero and one disulfide species disappeared at 5 minutes after the initiation of RNase A refolding (Fig. 2B). The refolding sample was dominated by the presence of the two disulfide species (2 S-S), together with the appearance of several new refolding intermediates. The third migrating peak shown in Fig. 2B was a mixture of the three disulfide species with various degrees of carboxymethylation and exogenous glutathione (3 S-S and mixed disulfide). The mass spectrum taken from the average scans under the peak (Fig. 4A) displayed two major components with molecular masses of 13,793 and 14,043 daltons. The mass of 13,793 daltons corresponded to a three disulfide species with two carboxymethyl groups. The second component with molecular mass of 14,043 daltons exhibited three intramolecular disulfide bonds, one intermolecular mixed disulfide with glutathione, and one carboxymethyl group.

Intramolecular protein disulfide bond formation using thiol-disulfide exchange with a disulfide reagent like glutathione is a two-step process. Initial reaction of reagent with a protein thiol, to form a mixed disulfide, is followed by intramolecular exchange with the thiol of a second cysteine residue, which reflects the conformational properties of the polypeptide chain. The mixed disulfide

accumulates only if there are conformational restrictions on the ability of a cysteine residue to form rapidly a disulfide with any of the other free cysteine residues[25]. In the absence of ESIMS, the identification of mixed disulfide RNase A refolding intermediates was achieved by Creighton using ion-exchange chromatography in combination with a radiolabeled disulfide reagent[24].

The first two peaks resolved by CIEF (designated as 4 S-S N' and 4 S-S N") contained four disulfide bonds with deconvoluted masses of 13,672 and 13,684 daltons, respectively (Fig. 5A-B). Even though all eight cysteines were involved in disulfide bond formation, these disulfide bonds might not be the correct pairings within the native RNase A. Thus, the tertiary structures of 4 S-S N' and 4 S-S N" remained in an extended conformation and allowed buried basic residues available for protonation to yield higher charged molecular ions than those of native RNase A in ESIMS (Fig. 5A-B).

The observation of four disulfide conformers, different from native RNase A, was entirely consistent with the earlier results of Creighton[12,24] achieved by ion-exchange chromatography and polyacrylamide gel electrophoresis. In fact, RNase A molecules with the shape of the native protein were formed only by slow rearrangement of at least some the disulfide bonds of the species containing four disulfide bonds. In contrast, the use of ESIMS alone by Torella et al.[18] was unable to resolve and determine heterogeneity among groups of refolding intermediates such as 4 S-S N', 4 S-S N", and native RNase A. Additionally, the peak migrating between 2 S-S and the mixture of 3 S-S and mixed disulfide was identified as

another conformer mixture of the three disulfide species (molecular mass of 13,807 daltons) and the component containing three intramolecular disulfide bonds, one intermolecular mixed disulfide, and one carboxymethyl group (molecular mass of 14,050 daltons). Finally, two minor components (designated as M) migrating within the refolding intermediates in Fig. 2A were identified as conformers of the 1 S-S and 2 S-S RNase A species based on the measured pIs and molecular masses. No positive identifications were able to be made for the components migrating after the zero disulfide species.

The peak intensities of four and three disulfide species continued to increase at 10 minutes after the initiation of protein refolding (Fig. 2C). The two disulfide species, however, was replaced with a mixture of refolding intermediates containing two intramolecular disulfide bonds and various degrees of intermolecular mixed disulfides (2 S-S mixed disulfides). The mass spectrum taken from the average scans under the peak (Fig. 4B) displayed two major components with molecular masses of 14,156 and 14,402 daltons. The mass of 14,156 daltons corresponded to a refolding intermediate containing two intramolecular disulfide bonds, one intermolecular mixed disulfide with glutathione, and three carboxymethyl groups. The second component with molecular mass of 14,402 daltons exhibited two intramolecular disulfide bonds, two intermolecular mixed disulfides, and two carboxymethyl groups. The deconvoluted masses of RNase A refolding intermediates are summarized in Table II with a relative standard deviation of 0.01-0.06% between the expected and measured molecular masses.

There was a very small amount of yet another four disulfide species (4 S-S N) which was barely distinguishable above the background at 10 minutes after the initiation of protein refolding (Fig. 2C). The peak intensity continued to increase at 30 minutes after the initiation of protein refolding (Fig. 2D). The mass spectrum taken from the average scans under the peak exhibited (Fig. 5C) a deconvoluted mass of 13,677 daltons. The electrospray ionization envelope shifted to high m/z region, indicating a compact structure of refolded RNase A. The accumulation of refolded RNase A was accompanied by the decrease in the peak intensities of the refolding intermediates. Although the gain in the intensity of refolded RNase A was unmatched by the loss of the refolding intermediates, the difference is attributed to the lower electrospray ionization efficiency of refolded RNase A due to the compactness of the molecule and the consequent shielding of ionizable residues. The refolding yield of RNase A at 30 minutes after the initiation of protein refolding was estimated around 85% on the basis of the CIEF-ESIMS measurements.

Acknowledgments

This work was supported by the Biotechnology Process Engineering Center at Massachusetts Institute of Technology through the National Science Foundation's Engineering Research Center Initiative under the Cooperative Agreement CDR-8803014. P.K.J. is a recipient of the Graduate Assistance in Areas of National Need (GAANN) Fellowship from the U.S. Department of Education. C.S.L. is a National Science Foundation Young Investigator (BCS-9258652).

References

1. Anfinsen, C. B.; Haber, E.; Sela, M.; White, F. H., Jr. *Proc. Natl. Acad. Sci. U.S.A.* **1961**, *47*, 1309.

2. Hagerman, P. J.; Baldwin, R. L. *Biochemistry* **1976**, *15*, 1462.

3. Labhart, A. M. *Proc. Natl. Acad. Sci. USA* **1984**, *81*, 7674.

4. Galat, A.; Yang, C. C.; Blout, E. R. *Biochemistry* **1985**, *24*, 5678.

5. Zuniga, E. H.; Nall, B. T. *Biochemistry* **1983**, *22*, 1430.

6. Kelley, R. F.; Wilson, J.; Bryant, C.; Stellwagen, E. *Biochemistry* **1986**, *25*, 728.

7. Creighton, T. E.; Darby, N. J.; Kemmink, J. *FASEB J.* **1996**, *10*, 110.

8. Creighton, T. E. *Proteins, Structures and Molecular Properties;* W. H. Freeman and Company: New York, **1983**.

9. Konishi, Y.; Scheraga, H. A. *Biochemistry* **1980**, *19*, 1308.

10. Rothwarf, D. M.; Scheraga, H. A. *Biochemistry* **1993**, *32*, 2671.

11. Rothwarf, D. M.; Scheraga, H. A. *Biochemistry* **1993**, *32*, 2690.

12. Creighton, T. E. *J. Mol. Biol.* **1977**, *113*, 329.

13. Gray, W. R. *Protein Sci.* **1993**, *2*, 1732.

14. Chowdhury, S. K.; Katta, V.; Chait, B. T. *J. Am. Chem. Soc.* **1990**, *112*, 9012.

15. Mirza, U. A.; Cohen, S. L.; Chait, B. T. *Anal. Chem.* **1993**, *65*, 1.

16. Loo, J. A.; Edmonds, C. G.; Udseth, H. R.; Smith, R. D. *Anal. Chem.* **1990**, *62*, 693.

17. Miranker, A.; Robinson, C. V.; Radford, S. E.; Dobson, C. M. *FASEB J.* **1996**, *10*, 93.

18. Torella, C.; Ruoppolo, M.; Marino, G.; Pucci, P. *FEBS Letters* **1994**, *352*, 301.

19. Strege, M. A.; Lagu, A. L. *J. Chromatogr. A* **1993**, *652*, 179.

20. Rush, R. S.; Cohen, A. S.; Karger, B. L. *Anal. Chem.* **1991**, *63*, 1346.

21. Kilar, F.; Hjerten, S. *Electrophoresis* **1989**, *10*, 23-29.

22. Tang, Q.; Harrata, A. K.; Lee, C. S. *Anal. Chem.* **1995**, *67*, 3515-3519.

23. Tang, Q.; Harrata, A. K.; Lee, C. S. *Anal. Chem.* **1996**, *68*, 2482-2487.

24. Creighton, T. E. *J. Mol. Biol.* **1979**, *129*, 411.

25. Creighton, T. E. *Progr. Biophys. Mol. Biol.* **1978**, *33*, 231.

Table 4.1. Molecular Weights of Native, Denatured and Reduced, and Alkylated RNase A

RNase A	Measured Molecular Weight (Dalton)	Expected Molecular Weight (Dalton)[a]	RSD (%)
Native	13,680	13,682	+0.01
Denatured and Reduced	13,696	13,690	+0.04
Alkylated	14,150	14,154	-0.03

a) Taken from reference (18).

Table 4.2. Molecular Weights of RNase A Refolding Intermediates

RNase A Refolding Intermediates	Measured Molecular Weight (Dalton)	Expected Molecular Weight (Dalton)[a]	RSD (%)
0 S-S	14,145	14,154	-0.06
1 S-S	14,026	14,031	-0.04
2 S-S	13,909	13,914	-0.04
2 S-S and 1 mixed S-S	14,156	14,161	-0.04
2 S-S and 2 mixed S-S	14,402	14,408	-0.04
3 S-S	13,793	13,798	-0.03
3 S-S and 1 mixed S-S	14,043	14,045	-0.01

a) Taken from reference (18).

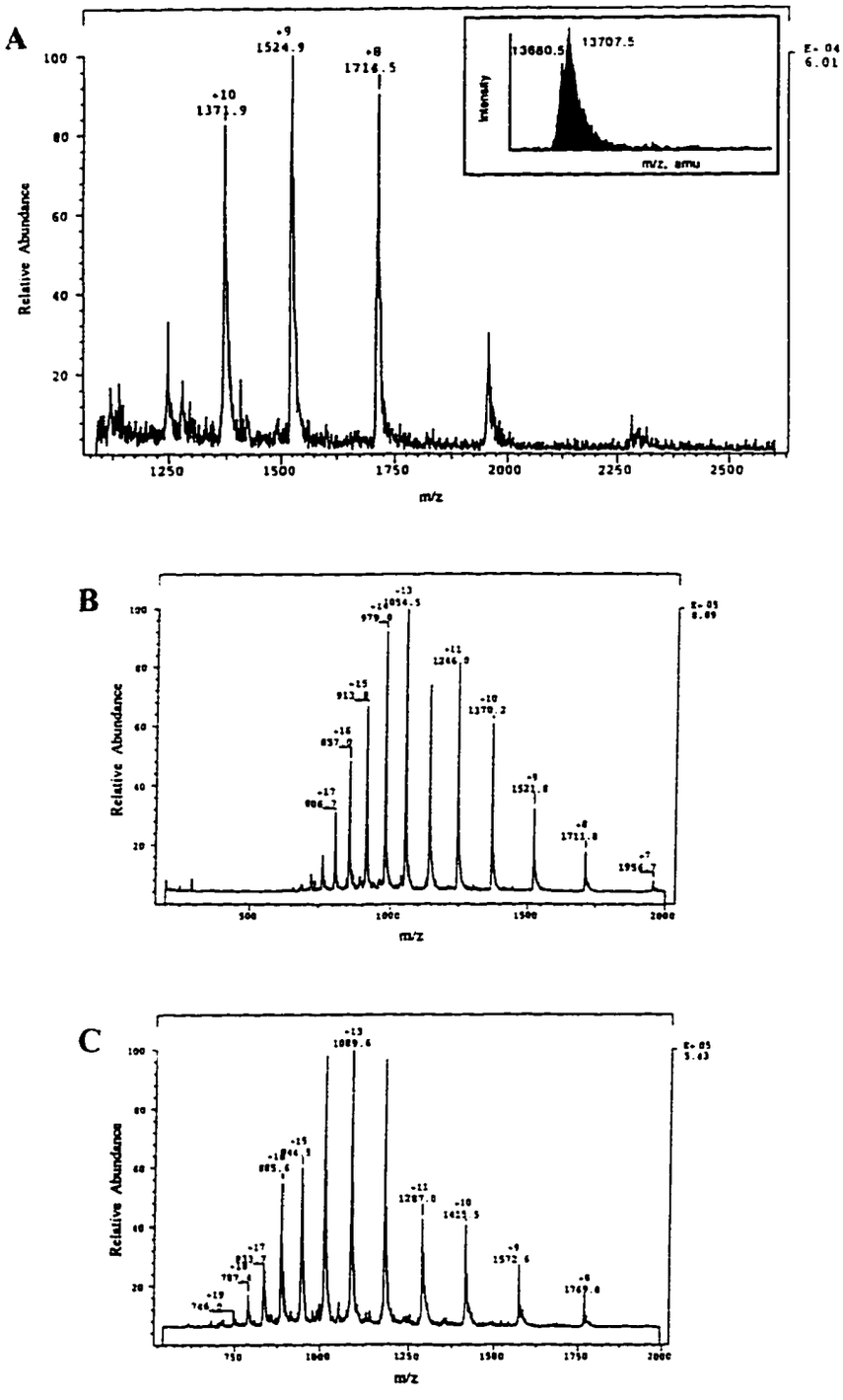

Figure 4.1. Positive electrospray ionization mass spectra of 1 mg/ml each of (A) native, (B) denatured and reduced, and (C) denatured and alkylated RNase A.

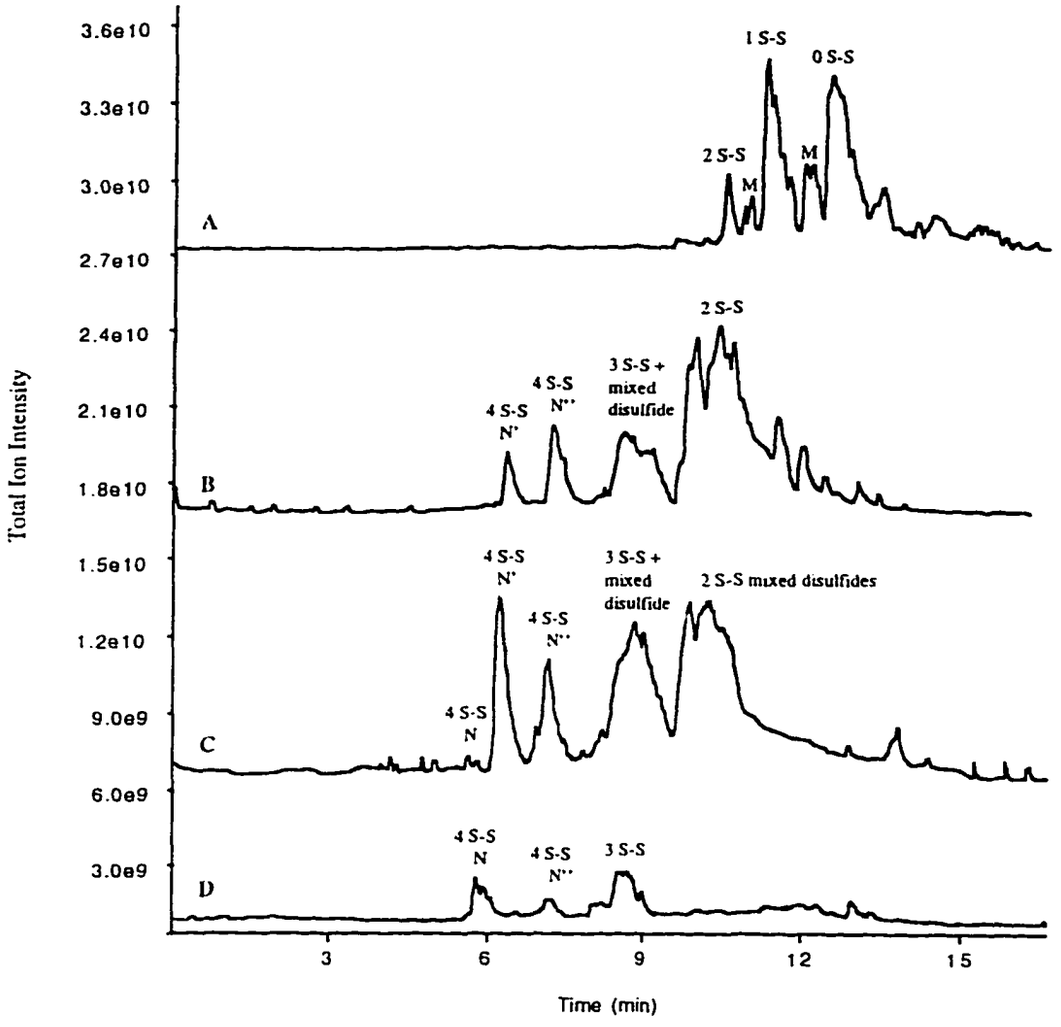

Figure 4.2. Monitoring refolding of RNase A using CIEF-ESIMS. Refolding samples were alkylated with IAA at (A) 0 minutes, (B) 5 minutes, (C) 10 minutes, and (D) 30 minutes after the initiation of RNase A refolding. Capillary, 30 cm total length, 50 mm i.d. and 192 mm o.d.; electrophoresis buffer, 2 mg/ml RNase A, 0.5% 3-10 Pharmalyte, and 0.05% TEMED; applied voltages, 15 kV for focusing and mobilization, 5 kV for electrospray; sheath liquid, methanol/water/acetic acid (50:49:1 v/v/v) at pH 2.6, 3 ml/min; mass scan, m/z 1000 to 2000 at 1.5 sec/scan.

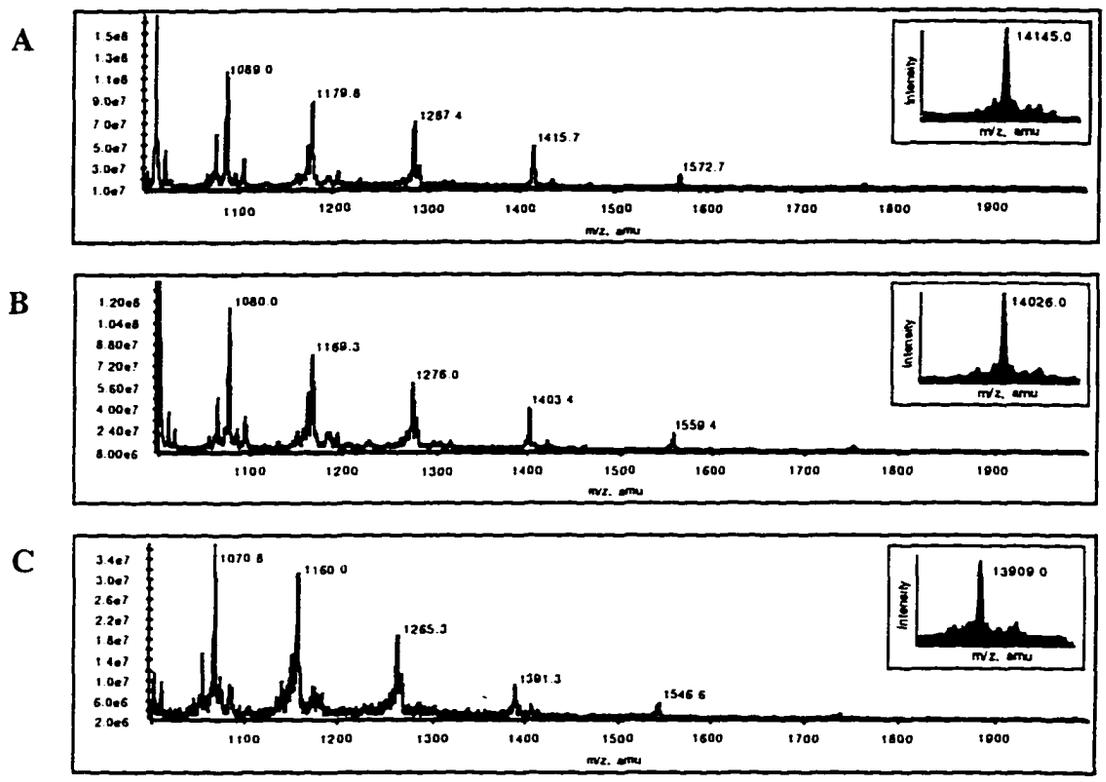

Figure 4.3. Positive electrospray ionization mass spectra taken from the average scans under the peaks of (A) zero, (B) one, and (C) two disulfide RNase A intermediates in Fig. 2A.

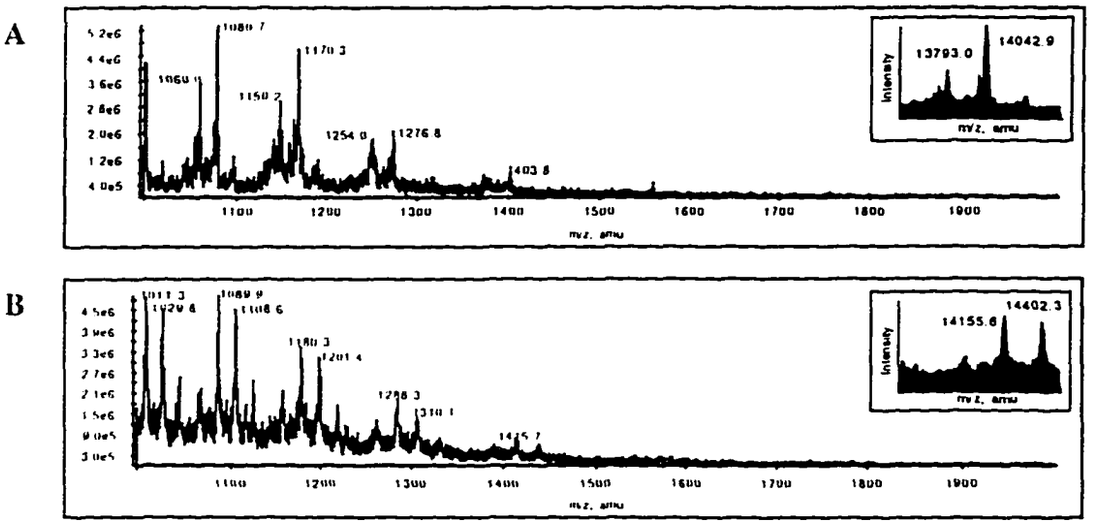

Figure 4.4. Positive electrospray ionization mass spectra of (A) three disulfide species with various degrees of exogenous glutathione (from Fig. 2B) and (B) two disulfide species with various degrees of exogenous glutathione (from Fig. 2C).

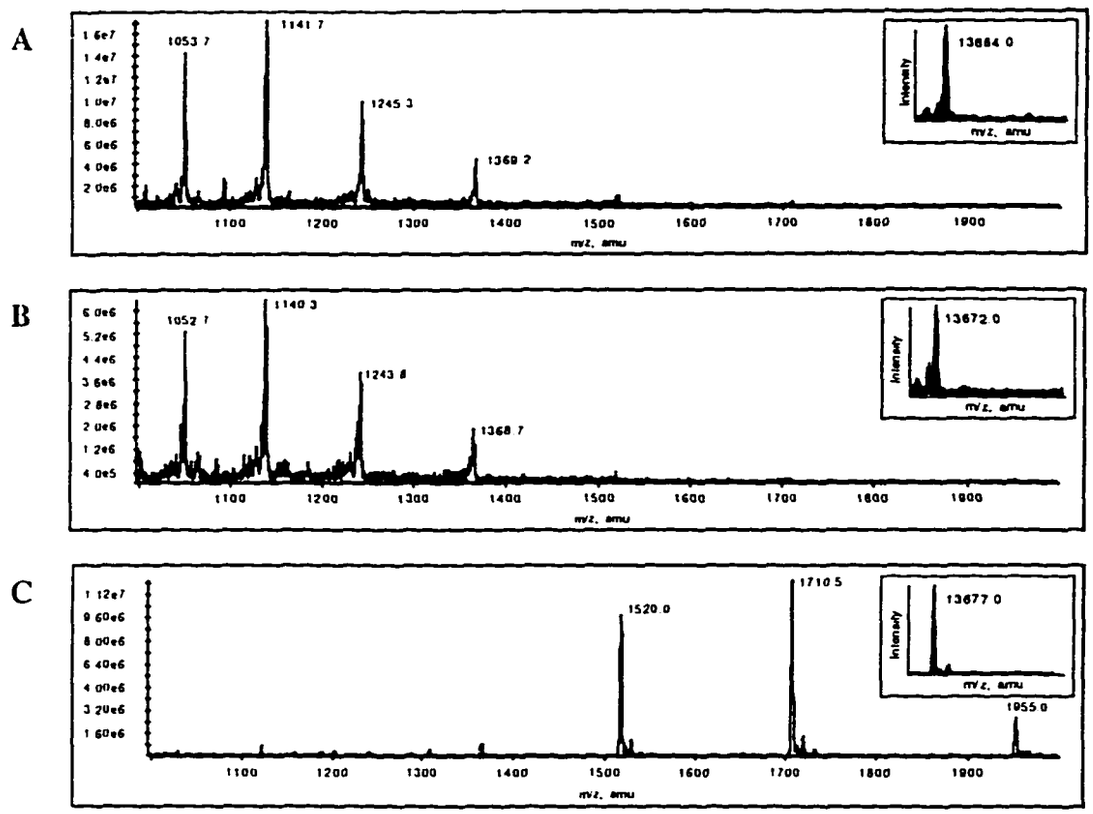

Figure 4.5. Positive electrospray ionization mass spectra of four disulfide conformers. (A) 4 S-S N" (from Fig. 2B), (B) 4 S-S N' (from Fig. 2B), and (C) 4 S-S N (from Fig. 2D).

CHAPTER 5. GENERAL CONCLUSIONS

Rapid identification of transient partially folded intermediates formed during protein refolding and aggregation has been difficult, particularly with separation methods relying on solid matrices. Capillary zone electrophoresis (CZE) in combination with the intrinsic tryptophan fluorescence detection provides a fast sensitive means for identifying folding and aggregation intermediates. This has proven particularly useful in studying the effects of folding temperatures on the distribution and competition among productive folding and aggregation pathways. In comparison to the results obtained from conventional methods (polyacrylamide gel electrophoresis and fluorescence), CZE is demonstrated as an effective technique for the analysis of tailspike refolding.

Additionally, capillary isoelectric focusing (CIEF) coupled with electrospray ionization mass spectrometry (ESIMS) is explored as a two-dimensional separation method with high mass accuracy for analysis of protein refolding pathways. Intermediates of the disulfide bond induced refolding of RNase A are trapped by alkylation with iodoacetate at various times during refolding, providing a means for separation in a predictable manner. Species containing varying degrees of both intramolecular and intermolecular disulfide bonds as well as those exhibiting conformational heterogeneity are resolved according to pI using CIEF and directly identified by ESIMS.

REFERENCES

1. DeBernardez-Clark, E.; Georgiou, G. In *Protein Refolding:* Georgiou, G., DeBernardez-Clark, E., Eds.; American Chemical Society: Washington, D. C., **1991**; ACS Symposium Series 470, Chapter 1.

2. Schmid, F. X. In *Protein Structure: A Practical Approach;* Creighton, T. E., Ed.; IRL Press: New York, **1990**, Chapter 11.

3. Semisotnov, G. V.; Uversky, V. N.; Sokolovsky, I. V.; Gutin, A. M.; Razgulyaev, O. I.; Rodionova, N. A. *J. Mol. Biol* **1990**, *213*, 561.

4. Egan, D. A.; Logan, T. M.; Liang, H.; Matayoshi, E.; Fesik, S. W.; Holzman, T. F. *Biochemistry* **1993**, *32*, 1920.

5. Englander, S. W. *Science* **1993**, *262*, 848.

6. Miranker, A.; Robinson, C. V.; Radford, S. E.; Dobson, C. M. *FASEB J.* **1996**, *10*, 93.

7. Perkins, J. R.; Smith, B.; Gallagher, R. T.; Jones, D. S.; Davis, S. C.; Hoffman, A. D.; Tomer, K. B. *J. Am. Soc. Mass Spectrom.* **1993**, *4*, 670.

8. Feng, R.; Konishi, Y.; Bell, A. W. *J. Am. Soc. Mass Spectrom.* **1991**, *2*, 387.

9. Goldenberg, D. P.; Creighton, T. E. *Anal. Biochem.* **1984**, *138*, 1.

10. Dunn, M. J. In *Gel Electrophoresis: proteins;* Dunn, M. J., Ed.; Bios Scientific Publishers: Oxford, England, **1993**.

11. Batas, B.; Jones, H. R.; Chaudhuri, J. B. *J. Chromatogr. A* **1997**, *766*, 109.

12. Light, A.; Higaki, J. N. *Biochemistry* **1987**, *26*, 5556.

13. Regnier, F. E. *Anal. Biochem.* **1982**, *126*, 1.

14. Stout, R. W.; Sivakoff, S. I.; Ricker, R. D.; Snyder, L. R. *J. Chromatogr.* **1986**, *353*, 439.

15. Goldenberg, D. P.; Berget, P. B.; King, J. *J. Biol. Chem.* **1982**, *257*, 7864.

16. Speed, M. A.; Wang, D. I. C.; King, J. *Protein Sci.* **1995**, *4*, 900.

17. Jorgenson, J. W.; Lukacs, K. D. *Anal. Chem.* **1981**, *53*, 1298.

18. Jorgenson, J. W.; Lukacs, K. D. *Science* **1983**, *222*, 266.

19. Hilser, V. J.; Worosila, G. D.; Freire, E. *Anal. Biochem.* **1993**, *208*, 125.

20. Kilar, F.; Hjerten, S. *J. Chromatogr.* **1993**, *638*, 269.

21. Strege, M. A.; Lagu, A. L. *J. Chromatogr. A* **1993**, *652*, 179.

22. Thannhauser, T. W.; Rothwarf, D. M.; Scheraga, H. A. *Biochemistry* **1997**, *36*, 2154.

23. Yeung, E. S.; Wang, P.; Li, W.; Giese, R. W. *J. Chromatogr.* **1992**, *608*, 73.

24. Goldenberg, D.; King, J. *J. Mol. Biol.* **1981**, *145*, 633.

25. Goldenberg, D.; King, J. *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 3403.

26. Goldenberg, D. P.; Smith, D. H.; King, J. *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 7060.

27. Haase-Pettingell, C. A.; King, J. *J. Biol. Chem.* **1988**, *263*, 4977.

28. Seckler, R.; Fuchs, A.; King, J.; Jaenicke, R. *J. Biol. Chem.* **1989**, *264*, 11750.

29. Fuchs, A.; Seiderer, C.; Seckler, R. *Biochemistry* **1991**, *30*, 6598.

30. Mitraki, A.; Danner, M.; King, J.; Seckler, R. *J. Biol. Chem.* **1993**, *268*, 20071.

31. Sargent, D.; Benevides, J. M.; Yu, M.-H.; King, J.; Thomas, G. J., Jr. *J. Mol. Biol.* **1988**, *199*, 491.

32. Steinbacher, S.; Seckler, R.; Miller, S.; Steipe, B.; Huber, R.; Reinemer, P. *Science* **1994**, *265*, 383.

33. Smith, D. H.; Goldenberg, D. P.; King, J. In *The Protein Folding Problem;* Wetlaufer, D. B., Ed.; Westview Press, Inc.: Colorado, **1984**; AAAS Selected Symposium 89, Chapter 6.

34. Sturtevant, J. M.; Yu, M.-H.; Haase-Pettingell, C.; King, J. *J. Biol. Chem.* **1989**, *264*, 10693.

35. King, J.; Fane, B.; Haase-Pettingell, C.; Mitraki, A.; Villafane, R.; Yu, M.-H. In *Protein Folding: Deciphering the Second Half of the Genetic Code;* Gierasch, L.; King, J., Eds.; AAAS: Washington, D. C., **1989**, Chapter 23.

36. Hjerten, S.; Zhu, M. *J. Chromatogr.* **1985**, *346*, 265.

37. Hjerten, S. *J. Chromatogr.* **1985**, *347*, 191.

38. Andersen, J. S.; Svensson, B.; Roepstorff, P. *Nature Biotechnology* **1996**, *14*, 449.

39. Fenselau, C. In *Advances in Biophysical Chemistry;* Bush, C. A., Ed.; JAI Press Inc: London, England, **1994**, 31.

40. Yamashita, M.; Fenn, J. B. *Phys. Chem.* **1984**, *88*, 4451.

41. Yamashita, M.; Fenn, J. B. *Phys. Chem.* **1984**, *88*, 4671.

42. Fenn, J. B. *J. Am. Soc. Mass Spectrom.* **1993**, *4*, 524.

43. Anfinsen, C. B.; Haber, E.; Sela, M.; White, F. H., Jr. *Proc. Natl. Acad. Sci. U.S.A.* **1961**, *47*, 1309.

44. Creighton, T. E. *J. Mol. Biol.* **1979**, *129*, 411.

45. Konishi, Y.; Scheraga, H. A. *Biochemistry* **1980**, *19*, 1308.

46. Rothwarf, D. M.; Scheraga, H. A. *Biochemistry* **1993**, *32*, 2671.

47. Rothwarf, D. M.; Scheraga, H. A. *Biochemistry* **1993**, *32*, 2690.

48. Creighton, T. E. In *Protein Folding;* Creighton, T. E., Ed.; W. H. Freeman and Company: New York, **1992**, Chapter 7.

49. Creighton, T. E.; Darby, N. J.; Kemmink, J. *FASEB J.* **1996**, *10*, 110.

50. Gray, W. R. *Protein Sci.* **1993**, 2, 1732.

51. Creighton, T. E. *J. Mol. Biol.* **1977**, *113*, 329.

ACKNOWLEDGMENTS

I would like to express my gratitude to several people, who contributed either personally or professionally in making it possible for me to achieve this goal.

First and foremost, I am forever indebted to my husband, Brett Jensen. His unconditional love and support has been the one thing I could always count on. His words of encouragement helped me through the tough times and kept me focused on accomplishing my goals. I appreciate all his sacrifice, patience and understanding over these past five years.

I am eternally grateful to my parents, Joe and JoAnn Cheney for always showing interest in my education and instilling in me from the beginning the desire to learn. The lessons they taught were the most valuable and gave me the foundation I needed on which to build. My siblings, Janet, Mike and Dave, played an unknowing role in my desire to learn and to succeed at whatever I do. I would also like to thank my mother-in-law, Betty Jensen, for her continued support and understanding.

I want to acknowledge Dr. Cheng Lee for his guidance and support throughout my graduate career. I would also like to thank Dr. Hugh Fan for his initial work on the tailspike system. Additionally, I would like to thank Dr. Kamel Harrata for his assistance in the electrospray ionization mass spectrometry work. I also appreciate the aid and advice of Dr. Jonathan King and members of his group at MIT, especially Drs. Margaret Speed and Anne Robinson, who also provided the tailspike protein.

To all Lee group members, past and present, you have made my time here most interesting and memorable. Special thanks to Wendy Nelson, who has been an important part of both my personal and professional development. I also appreciate Jing Wei for the valuable time and effort she has given to me these last few months. To other group members, Jean, Meng, Yan, Yun, Qing and Liyu, I appreciate your helpful discussions, your cooperation and your friendship throughout our time together. Best wishes to everyone.