ABSTRACT

Chen, Yun-Ru. Equilibrium and Kinetic Folding Properties of Alpha-helical Greek Key Protein Domains. (Under the direction of Dr. A. Clay Clark)

I have characterized the equilibrium and kinetic folding properties of members of α-helical Greek key protein domain, caspase recruitment domains of RICK (RICK-CARD) and procaspase-1 (Pro-1-CARD). At equilibrium, folding of both RICK-CARD and Pro-1-CARD is well described by a two-state mechanism representing the native and unfolded ensembles. The proteins are marginally stable, with a ΔG^{H_2O} of 3.0 and 1.1 kcal/mol and m-values of 1.27 and 0.68 kcal/mol/M for RICK-CARD and Pro-1-CARD, respectively (30 mM Tris-HCl, pH 8, 1 mM DTT, 25 ºC). The folding pathways of RICK-CARD are complex and contain at least two or three non-native conformations. The major folding event of RICK-CARD has a folding rate constant of 30 s^{-1}. Studies on seven mutants of RICK-CARD suggest that the three prolyl residues contribute to its structural stability and folding complexity. The folding of Pro-1-CARD is different than that of RICK-CARD. The folding and unfolding kinetics of Pro-1-CARD are fast but with kinetically trapped intermediates present, which appear to be similar to that of RICK-CARD. The work suggests that different α-helical Greek key protein domains fold differently but may share some similar folding pathways.

EQUILIBRIUM AND KINETIC FOLDING PROPERTIES OF ALPHA-HELICAL GREEK KEY PROTEIN DOMAINS

by

Yun-Ru Chen

A dissertation submitted to the Graduate Faculty of

North Carolina State University

in partial fulfillment of the

requirements for the Degree of Doctor of Philosophy

Molecular and Structural Biochemistry

Raleigh

2003

APPROVED BY:

_____ _____

(Chair of Advisory Committee)

_____ _____

DEDICATION

This dissertation is dedicated to my grandmothers,

Hsiu-Ching Chen Tseng (陳曾秀琴) and Sun-Ti Liu (劉順娣),

and my parents, Ding-Shinn Chen (陳定信) and Hsu-Mei Hsu (許須美).

BIOGRAPHY

Yun-Ru Chen (Ruby, 陳韻如) was born in Taipei, Taiwan. She grew up in a warm and successful family with her parents, brother, and grandmother. She was very close to her grandmother, who is a strong and smart woman that raised her family in the period of World War II. Her father is the most important person that influences her life and studies. He is a well-known and successful medical doctor and a scientist. He is now the dean of the medical college of National Taiwan University and the president-elect of International Association for the Study of the Liver. Her mother is a career woman and an idol for her to look up to. By growing up in the family, she was a good and honest child although she felt the pressure of the success from her parents. She was educated under the education system in Taiwan. She first passed the entrance exam to the Taipei First Girl High School for happy high school years. Then, she entered the Department of Agricultural Chemistry in National Taiwan University for four years of undergraduate study. At the time, she has made many good friends from the guitar club and has a wonderful friendship with them until now. She began to get interested in biochemistry in her senior year by doing research in the laboratory of Professor Jong-Ching Su, and further worked as a research assistant under Dr. Shang-Fa Yang in Institute of Botany in Academia Sinica, Taipei. In the summer of 1998, Yun-Ru came to the United States for her graduate studies. It was her first time leaving home. She entered the Biochemistry Department at North Carolina State University as a Ph.D. student and has been in love with the place. Yun-Ru is expecting to become an academic professor in Taiwan and is looking forward to be close to her family in the future.

ACKNOWLEDGMENTS

First, I would like to deeply appreciate my advisor Dr. Clay Clark for his guidance and solid training. I would not become an independent researcher and this project will not be completed if there is not Clay's significant input, great patient, and strong support for the research.

Second, I would like to thank my committee members, Drs. Dennis Brown, Carla Mattos, Harold Swaisgood, and Clay Clark for your advisory and care. You are there with me from the beginning to this fruitful end. I also appreciate the cooperation and tolerance from my labmates and lab stuff, Kakoli Bose, Cristina Pop, Brett Feeney, Justin Stern, Luming He, and Randy Durren.

Moreover and very importantly, I would like to thank Dr. Stefan Franzen (North Carolina State University) for assistance with the homology modeling and Dr. Gabriel Núñez (University of Michigan) for providing the plasmid pcDNA3-RICK. Specially thank our lab manager Brett Feeney for assisting and generate most of the proline mutants of RICK-CARD. Also, I appreciate Dr. Carla Mattos (North Carolina State University) and Nathan Nicely for helping and input on the structure analysis and preparation of crystal trials. I am thankful for the technical support from Dr. Ashutosh Tripathy (UNC, Chapel Hill) for performing analytical ultracentrifugation studies. Also, I thank Richele Thompson and Dr. John Cavanagh (North Carolina State University) on helping the preparation of NMR samples. I appreciate Dr. Yun-Chorng Chang (National Cheng Kung University) for generating the long-range order program for our calculation.

TABLE OF CONTENT

CHAPER TWO

MATERIALS AND METHODS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Chapter 3

Chapter 4

Chapter 5

LIST OF ABBREVIATIONS

CARD, caspase recruitment domain; RICK, RIP (Receptor interacting protein)-like interacting CLARP kinase; RICK-CARD, CARD of RICK; Pro-1-CARD, CARD of procaspase-1; DD, death domain; DED, death effector domain; APAF-1, apoptotic protease activating protein; IPTG, isopropyl β-D-1-thiogalactopyranoside. CO, relative contact order; LRO, long-range order; CD, circular dichroism. Standard abbreviations are used for the amino acids.

Chapter 1

INTRODUCTION

Chapter 1

INTRODUCTION

1.1 Importance of protein folding

The field of protein folding has not been more important than today in this post genomic era. Because of the resolution of human genome and the quick development of proteomic technology, the information of protein structures is much more accessible than the past decade [1]. The increasing discovery of diseases caused by protein misfolding shows the importance of protein folding in the cell and provides the practical aspects of the field [2; 3]. Although the rate of solving protein structures has been accelerating and the number of proteins that been studied has increased, the fundamental problem of protein folding still remains unsolved.

Protein is translated from the translation machinery in cells as a linear polypeptide chain that is composed by various amino acids. There are 20 naturally occurring amino acids in human cells. The amino acid sequence of a polypeptide chain that encodes the functional gene is referred to primary structure of protein. Different amino acid compositions result in secondary protein structures including alpha helix, beta-sheet, and beta-turns. Numerous arrangements of the secondary structures lead to different protein tertiary structures. Larger proteins further organize tertiary structures to quarternary structures such as subunits in the proteins. According to the central dogma of biology,

generation of protein initiates from transcription of DNA to RNA followed by translation from RNA to polypeptide chains (DNA-RNA-Protein); then the polypeptide chain folds to a protein with a unique structure. It is the final folded protein that is responsible for the function and specificity. Clearly, protein folding is the required step for primary sequence of protein to reach its tertiary structure. Solving protein folding problem would provide information on the missing link between the primary sequence and the three-dimensional structure.

1.2 Protein folding problem

Anfinsen has shown that protein folding is a spontaneous and reversible process [4] and suggests that the primary sequence of a protein provides information on both its structure and folding. In the view of thermodynamics, the conformational energy, Gibbs free energy (ΔG), of protein accounting for stability is contributed by both entropy and enthalpy ($\Delta G = \Delta H - T\Delta S$). There are large differences in both entropy (ΔS) and enthalpy (ΔH) between the native and unfolded ensembles. However, the difference of free energy between the two states is small, only a few kcal/mol, 5 –15 kcal/mol, because of the compensation of the two large values. The native protein conformation is believed to have the minimal conformational energy compared to all other conformations in the folding process so that an unfolded polypeptide chain would search for its most stable conformation, which is its native state, along the folding process. For further explanation, Levinthal has suggested the concept of energy landscape. Energy landscape is simply a function of free energy level with the coordinates of different conformations of a

polypeptide chain formed upon folding. The coordinates could be referred to degree of freedom for the chain. According to the thermodynamic concept that native proteins adopt minimum energy, a golf course folding landscape with smooth and flat energy surface and a needle shape energy minimum was suggested (Figure 1). The golf-course folding landscape indicates that there is no bias, but there are evenly distributed pathways for protein folding [5; 6].

However, the protein folding problem arises based on this simple assumption. The problem can be exemplified by the drastic difference between the folding rate assumed and the actual folding rate. For example, for folding of a polypeptide chain with 100 amino acids, each peptide bond can adopt three possible conformations. They are α–helix, β-sheet, and loop based on Ramachandran plot. Simple calculations indicate that 3^{100} conformations are available. It will take 10^{33} years for this peptide to fold to its native conformation if we assume 10^{-11} seconds for a peptide bond to find its correct conformation. However, protein folds much faster than the calculated value [4]. This problem is called Levinthal's paradox. The folding rates of proteins range from milliseconds to thousand of seconds. Therefore, there must be pathways involved in the folding event rather than a random searching process. Rather than golf-course energy landscape, the folding process could be presented by a rough folding funnel with the presence of local energy minima causing kinetic traps [6].

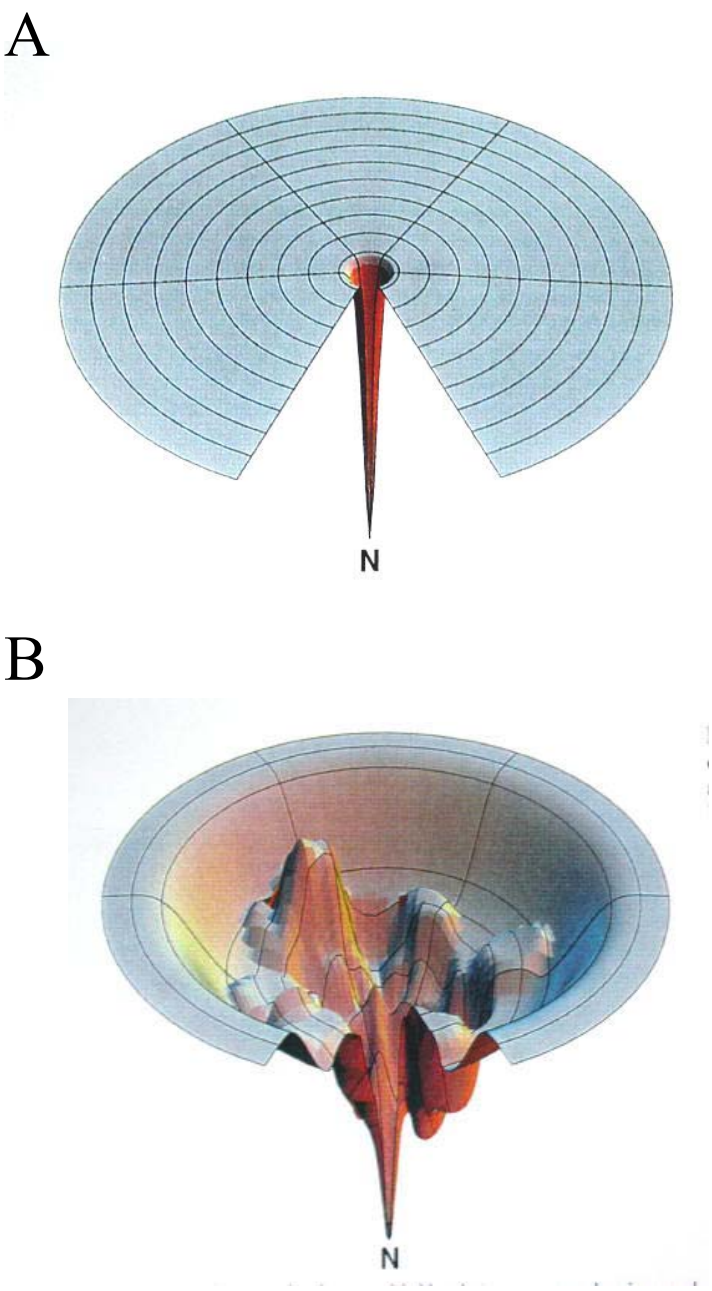


Figure 1. Conceptional framework of energy landscapes[6]. (A) Golf-course energy landscapes (B) Rough energy landscape. The surface of the landscape represents the energy level of each conformation. The native conformation containing the minimal energy is labeled as "N".

Today, after experimental and theoretical research, a new view of protein folding mechanism as well as several theories of protein folding mechanism have been established. For a small protein that folds in microsecond timescale, the compact conformation forms rapidly and serves as a folding nucleus on the folding process. The folding rates are limited by slow interconversion of the different unfolded species such as the unfolded proteins with different proline isomers. For the protein with complex fold, the probability of forming an appropriate nucleus is much smaller, this results in slow folding kinetics[7]. If the nucleus or the collapsed state is more native-state, the folding rate will be faster. If the collapsed state contains many non-native characters, the folding will be slower because of the process of complex rearrangement of the conformation[7; 8]. Recently, based on experiments and computational simulation, researchers find that it is necessary for proteins to go through a series of identifiable intermediates upon protein folding[5].

1.3 Determinants of protein folding kinetics

Protein structure and its folding kinetics are not likely determined only by the primary sequence, although protein originates from a polypeptide chain and primary sequences are the basic difference among proteins. Various other factors have been discussed [9] including topology, stability, and length of the protein. Baker *et al*. have analyzed the correlations of different factors and the folding kinetics based on proteins adopting two-state folding kinetics. Two-state folding kinetics represent the simplest folding mechanism that consists of native and unfolded ensembles in the folding process, without detectable intermediates (N ⇆ U). They concluded that the complexity of topology is a

dominant but not the only factor to determine the folding rate[9]. In contrast, sequences, stability, and length have little correlations with the folding rates. In summary, they suggest that there is fundamental similarity in the folding mechanism of single domain proteins despite the diversity of the structures.

1.4 Folding of homologous proteins

In order to determine the fundamental mechanism of protein folding, researchers have begun to study proteins that share similar structures and the same folding topology. By comparing the differences among them, the folding mechanism can be examined and the correlations will be observed. Several groups of proteins have been examined and some are listed here. They are immunity protein, acyl coenzyme A-binding protein, apomyoglobin, SH3 domain, cold shock protein, immunoglobulin, fatty acid binding proteins, protein G, acylphosphastase, and lysozyme [10]. By examining their folding properties and comparing the results among the homologous proteins, a few guidelines have emerged, as described below.

1.4.1 Homologous proteins in the same family

Radford's group has studied the bacterial immunity proteins as a model system for protein folding of small α-helical proteins [11; 12; 13]. The sequence identity among the immunity proteins is about 50 %, and they share the same topology as a four-helix bundle with sequential helices that pack adjacently in alternate directions (up-down pattern). At

equilibrium, all the immunity proteins adopt a two state folding mechanism with stabilities ranging from 3 to 5 kcal/mol (16 to 23 kJ/mol).

For folding kinetics, most of the immunity proteins follow a simple two-state mechanism except that immunity protein-7 follows a three-state mechanism describing an intermediate present in the folding pathway. By using point mutations and segment swapping, they found the unfolding rates, but not refolding rates, of the immunity proteins correlate well with their stabilities. However, they all have conserved and highly compacted transition states. Therefore, the residues that are important for stabilizing the native state are distinct from those for the transition state. These results are also observed for the cold stock protein family [14].

1.4.2 Homologous proteins in different families

Clarke *et al*. have examined six Immunoglobulin (Ig)-like β-sandwich proteins in different families[15]. The folding topology of Immunoglobulin is categorized as Greek key fold. Since homologous proteins in the same family often share the same function, their primary sequences are possibly responsible for the function but not for the folding. Clarke's group believes the folding mechanism should be less ambiguous by eliminating the factor of function shared by those homologous proteins in the same family. They found that the logarithm of folding rates, but not unfolding rates, of those Ig-like proteins have linear relationship with their stabilities. They conclude that those proteins share the same nucleation event in folding and their transition state is conserved. The results indicate that those Ig-like proteins with low sequence identity fold though common folding pathways.

Therefore, structure or topology may be the most important factor for folding of Ig-like proteins.

1.5 Folding of small helical proteins

The folding properties of several small helical bundles, such as cytochrome c [16; 17], acyl-CoA binding protein [18; 19], λ-repressor [20; 21], and immunity protein [11; 12; 13; 22], have been studied. In general, their folding rates are fast (from ~300 to ~5000 s^{-1}) [23; 24] and most of the folding kinetics follow a two-state mechanism. Their folding rates can be predicted by computational algorithm such as "contact order" and "long-range order".

Contact order, which describes the topology of proteins, is an algorithm that calculates the average sequence distance between all pairs of contacting residues in the native structure [9; 25]. The relative contact order (CO) is the contact order normalized by the total length of the sequence. CO is obtained by equation [1].

$$CO = \frac{1}{L * N} \sum_{1}^{N} \Delta S_{i,j} \qquad [1]$$

where CO is the relative contact order, N is the total number of contacts, ΔS_{ij} is the sequence separation between contacting residues i and j, and L is the total number of residues in the protein.

Since long-range interactions play a vital role in stability, the interactions caused by the residues that are far in the sequence should also be important for protein folding. Long-range order (LRO) is an algorithm that considers the long-range interactions. LRO also represents the topology of the protein. It is defined to consider the contacting residues

that are close in space in the native structures but far in distance in the primary sequence. LRO could be calculated as follows:

$$LRO = \sum n_{ij} / N \qquad n_{ij}=1, \text{ if } |i-j| > 12 \text{, otherwise } n_{ij}= 0 \qquad [2]$$

where *i* and *j* are two residues for which the distance between two alpha-carbons is no more than 8 Å, and N is the total number of residues in a protein.

1.6 Proline isomerization

The rate-limiting steps for protein folding are generally due to proline isomerization or disulfide bond formation because of their high activation energy barriers. The peptide bonds of proteins are usually in the *trans* configuration meaning the two alpha-carbons of the two residues point in opposite directions according to the plane formed by the peptide bond (Figure 2). Alternatively, a *cis*- configuration shows the two alpha-carbons of the two residues point in the same directions according to the plane formed by the peptide bond. The *cis* configuration is about 1000 times less stable. Therefore, *cis*-peptide bond is rarely found in native proteins. However, *cis*-proline residue is found in polypeptide chain and is sometimes essential for activity or for conformational flexibility [26; 27]. Since the ratio of *trans* proline to *cis* proline in the denatured solution is 4 to 1, there are more chances for prolyl peptide bonds (Xaa-Pro) to adopt the *cis* configuration in the native protein relative to the non-proyl peptide bonds. Since prolyl *cis*-*trans* isomerization is involved in rotation of the peptide bond, a partial double bond, the activation energy required is high and the reaction is slow. The half time of prolyl isomerization from *trans* to *cis* ranges from 10 to 100 sec. The activation energy is about 20 kcal/mol.

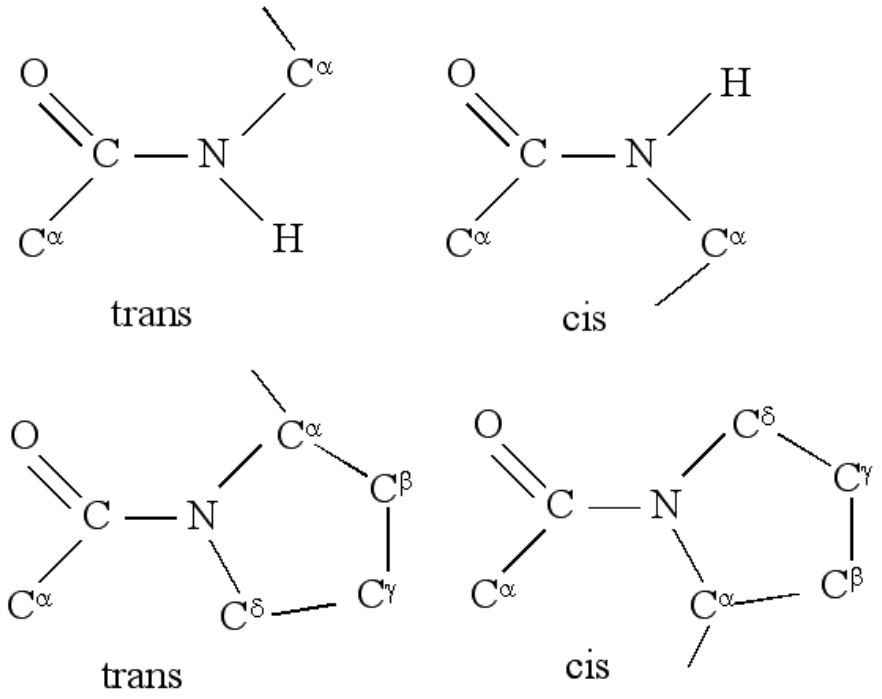


Figure 2. Non-prolyl (upper panel) and prolyl peptide bonds (lower panel) in *cis* or *trans* configurations.

1.7 Folding of β- Greek key proteins

There are a few β-Greek key proteins, comprised of β-strands in the Greek key topology, which have been examined, such as immunoglobulin-like proteins [28; 29; 30], βγ–crystallin [31; 32], fibronectin type III domain [15; 33] and azurin [34; 35; 36]. While the mechanisms of folding are not known for the complex Greek key topology, two nucleation models have been suggested. The β-hairpin model suggests that folding occurs initially through a long β-hairpin, and then the protein folds up two strands at a time [37; 38]. Therefore, the sequential β-sheet is not next to each other but in the Greek key topology. However, the hairpin model is a hypothesis without any experimental data. Alternatively, a proposed nucleation site called β-zipper model is suggested in folding. It suggests that there are well-conserved, large hydrophobic residues on two sequential β-strands joined by a short loop [39]. The two strands face each other from opposite sides of the barrel and interact through their hydrophobic side chains. β-zipper model is observed in many β-Greek key proteins, however, they are also present in some non-Greek key proteins. In addition, a highly conserved tyrosine in the "tyrosine corner" motif that is found exclusively in the β-Greek key proteins was shown to be important for the stability and the early structure formation of fibronectin type III domain [33] and for the equilibrium folding of pseudoazurin [33; 40]. However, the tyrosine corner does not form in the transition state meaning the motif may not be the nucleation site upon folding. Such a motif is not found in the α-Greek key proteins.

1.8 Caspase recruitment domain (CARD)

1.8.1 Alpha Greek key topology

A unique protein fold, the α-helical Greek key, is found only in the death domain (DD) superfamily. The death domain superfamily includes members containing the death domain (DD), the death effector domain (DED), the pyrin domain, and the caspase recruitment domain (CARD) [41; 42; 43]. These protein domains are small. They are generally composed of about 100 residues and consist of six α-helices folded into a Greek key topology (Figure 3). Many charge residues are found in their primary sequences. The family members share the same topology and have similar structures, but their sequence identities are low in general. The sequence alignments of homologous CARD domains are shown in Figure 4.

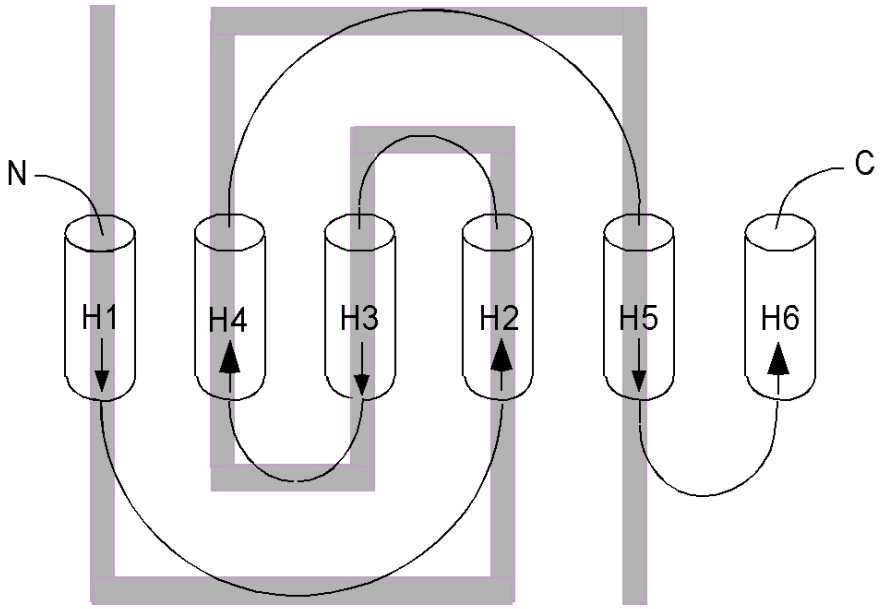


Figure 3. Folding topology of α-helical Greek key of the death domain superfamily. The helices are represented as cylinders with orientation indicated by the arrows. The N- and C-termini are indicated as well.

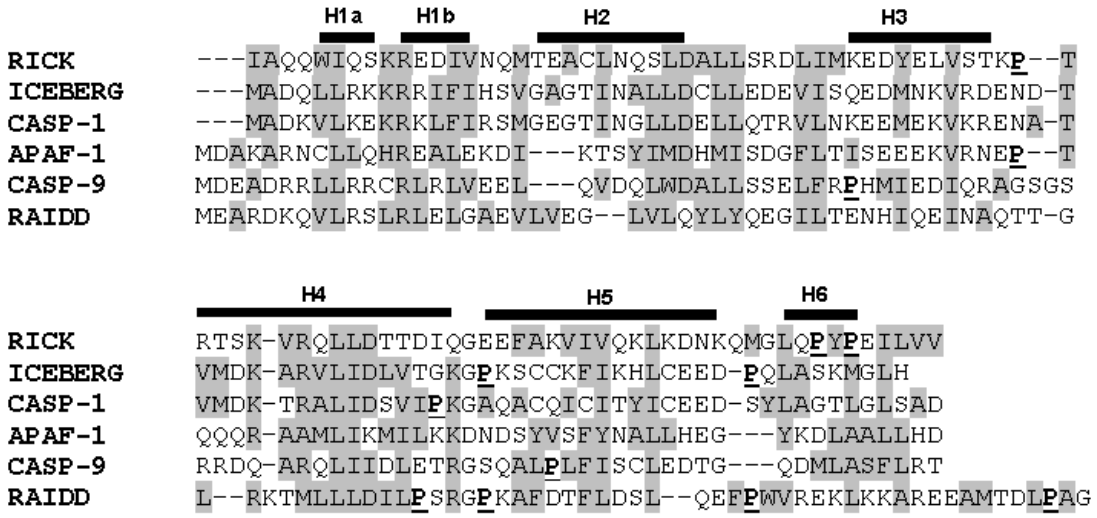


Figure 4. Sequence alignment of CARDs. The helical regions of ICEBERG are indicated above the alignment. The CARDs of RICK, ICEBERG, procaspase-1 (CASP-1), APAF-1, procaspase-9 (CASP-9), and RAIDD are listed. The homologous regions are highlighted in gray. Proline residues are underlined and in bold.

1.8.2 Death domain superfamily

The DD superfamily members include many biologically important proteins that play key roles in essential mechanisms for cell survival or cell death [44; 45] such as apoptosis, transcription activation, and inflammatory responses. The mechanisms will be further described in the following chapters.

A few examples of how the members of DD superfamily are involved in cellular mechanisms are described here. For example in apoptosis, cytochrome c is released from mitochondria at the beginning of apoptosis because of cellular stress. Cytochrome c then interacts with APAF-1 and procaspase-9. APAF-1 is a multi-domain protein consisting a N-terminal CARD domain, an ATPase domain, and a C-terminal region containing WD40 repeats. Procaspase-9 belongs to upstream caspases in the caspase family, and it is the zymogen form of caspase-9. APAF-1 interacts with procaspase-9 through CARD-CARD interaction. By forming apoptosome consisting cytochrome c, APAF-1, and procaspase-9, procaspase-9 is activated and further activates the downstream executioner caspase, caspase-3. Caspase-3 further proteolyses proteins that are required for cell survival. In an alternative pathway of apoptosis, DD in the death receptor interacts with an adaptor DD when apoptosis is initiated from receptor-ligand binding. The signal is further transferred to DED-DED interactions such as that in FADD and procaspase-8. Once procaspase-8 is activated, it further activates caspase-3.

For example in transcription activation, the same ligand-receptor binding for apoptosis could lead to transcription activation. RIP, a kinase containing a N-terminal DD domain, interacts with adaptor protein through DD-DD interactions and also recruits IKK.

IKK is a kinase that can phosphorylate and inactivate IκB. Since IκB is an inhibitor of the transcription factor NF-κB, the event results in transcription activation.

The biological function of the DD superfamily members is to convey cellular signals, and this is accomplished through hetero- or homo-oligomerizations that allow the domains to function as bridging proteins to larger protein complexes [41; 43; 46; 47]. A general schematic drawing of the cellular signal pathways of DD superfamily members is illustrated in Figure 5.

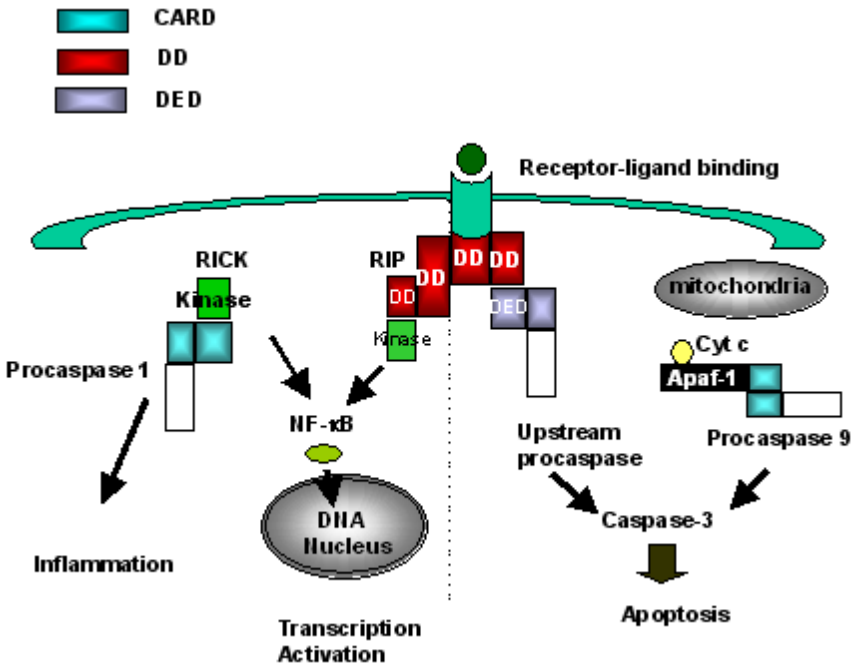


Figure 5. Schematic cellular pathways of the DD family members. Three major cellular mechanisms in which DD are involved are apoptosis, transcription activation, and inflammation. The color schemes of the DD members are indicated on the left upper corner. Each domain is labeled and the details of the mechanisms are described in text.

1.8.3 CARD-CARD interactions

The CARD-CARD interactions are highly specific [44]. While hydrophobic interactions are considered the primary driving force for the interactions, electrostatic contacts between surface patches of the CARDs determine the binding specificity [48]. There are many charged residues in the sequences of CARDs. The solution structure of the complex of APAF-1-CARD and procaspase-9-CARD shows details of the binding interface. The interface is composed of an acidic patch on the surface of APAF-1-CARD formed by helices-2 and -3 and the basic patch on the surface of procaspase-9-CARD formed by helices-1 and -4 [49] (Figure 6). The residues involved in the acidic patch of APAF-1-CARD are aspartate 27 and glutamates 40 and 41. The residues involved in the basic patch of procaspase-9-CARD are arginines 10, 11, 13, 52, and 56. In addition, the two patches in the interface are complementary in shape. The acidic patch of APAF-1 forms a convex surface, whereas, the basic patch of procaspase-9 forms a concave surface.

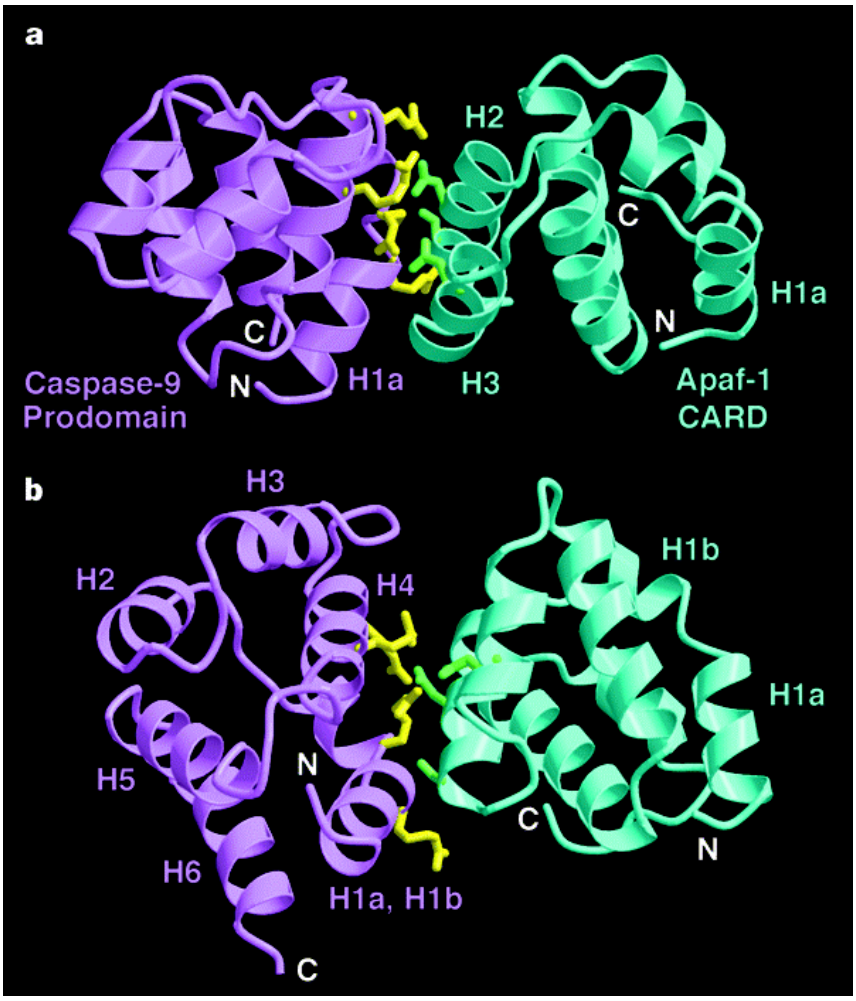


Figure 6. Representation of the structure of APAF-1-CARD [50] in complex with procaspase-9 prodomain (purple). The picture is obtained from the paper of Qin *et al.* [48].
A 90 degree rotation along the horizontal axis are shown in Panel (a) and (b). The interacting side chains are shown in stick figures. The N and C termini and helices are labeled.

1.8.4 CARD of RICK

A novel serine/threonine kinase, RICK (CARDIAK, RIP2) (RIP-like interacting CLARP kinase), is a CARD-containing protein found in the cytosol [51; 52]. RICK is expressed in multiple human tissues and regulates several cellular mechanisms: apoptosis initiated by CD95 (also called Fas) through enhancement of caspase-8 activation, activation of JNK (Jun N-terminal kinase) and NF-κB, and initiation of inflammatory responses by activating procaspase-1 [51; 52; 53] (Figure 5). It was shown that RICK mediates signal transduction responses for both innate and adaptive immune systems and is a key therapeutic target for bacterial-induced inflammation [54; 55]. RICK contains an N-terminal serine/threonine kinase catalytic domain and a C-terminal CARD. Both the catalytic domain and the CARD are required for CD-95 mediated apoptosis and NF-κB activation [51; 52], although only the CARD domain is required for the activation of procaspase-1 [53].

1.8.5 CARD of Procaspase-1

A group of the death domain superfamily members, CARD or DED, are present in the prodomain region of caspases. Caspases are cysteinyl proteases that are involved in apoptosis and inflammation. The CARD is located at the long prodomain region (~150 residues) in the N- terminus of upstream caspase including caspase-1,-2,-4, -5, -9, -11,-and -12 [46]. After specific CARD-CARD interactions [44] between the upstream caspase and its homologous binding partner located in signaling receptors or in adaptor proteins, the upstream caspase is activated by cleavage of the linker region of the subunits followed by releasing or inactivating the prodomain.

Procaspase-1, a zymogen form of caspase-1, is known as interleukin-1β-converting enzyme (ICE), a primary cytosolic enzyme involved in inflammation. The activation of the beta isoform of interleukin specifically by caspase-1 is related to inflammatory diseases [56]. Procaspase-1 is a protein with its mass of 45 kDa and is formed by a single polypeptide chain. The 15 kDa prodomain, including the CARD, is located at the N-terminus. The CARD of procaspase-1 (Pro-1-CARD) is located at the N-terminus of prodomain from methionine 1 to aspartate 92. It is responsible for oligomerization of procaspase-1 through CARD domains to enable the activation through close proximity. There are several binding partners of procaspase-1 employing CARD-CARD interactions, such as RICK (CARDIAK, RIP2), a pro-inflammatory serine/threonine kinase, ICEBERG, an inhibitor of procaspase-1 activation, and COP, an protein that has very high sequence identity with CARD of procaspase-1 and inhibits the pro-inflammation responses [57]. According to the crystal structures of the CARD-CARD complex of APAF-1 and procaspase-9, helices-1 and -4 of procaspase-1 are suggested to be the interacting helices to its partners [48].

1.9 Objective

My Ph.D. research is focused on this unique folding topology, α-helical Greek Key, of protein domains in the CARD family. The research details are described in the following chapters. Chapter two is the details of the experimental materials and methods. Chapter three describes the equilibrium and kinetic folding of RICK-CARD. These results have been published in *Biochemistry* in 2003 [58]. Chapter four describes the research on folding properties of Pro-1-CARD. Chapter five describes the research on site-specific

proline mutants of RICK-CARD. The project will have an impact on the field of protein folding since these are the first data on alpha-helical Greek Key domains and also a valuable report on new folding mechanisms of homologous proteins.

Chapter 2

MATERIALS AND METHODS

Chapter 2

Materials and Methods

2.1 Reagents

Ultra-pure urea was purchased from Nacalai Tesque Inc. (Kyoto, Japan). Ultra-pure Guanidium hydrochloride (GdnHCl) was obtained from ICN. Dithiothreitol (DTT) was either from Sigma or Denvil. Tris(2-carboxyethyl)-phosphine hydrochloride (TCEP) was from Pierce. Trizma-Base (Tris), sodium chloride (NaCl), potassium chloride (KCl), monobasic and dibasic potassium phosphate (KH_2PO_4, K_2HPO_4), sodium sulfate (Na_2SO_4), isopropyl β-D-1-thiogalactopyranoside (IPTG), ethylenediaminetetraacetic acid (EDTA), DEAE sepharose, Q-sepharose, ampicillin, antifoam C, DNase I, RNase A, and molecular weight markers for the size-exclusion column were purchased from Sigma. All buffers were filtered through either 0.45 or 0.22 μm filter membranes. The urea-containing and GdnHCl-containing buffers were prepared by weight measurement and refractive index as described previously [59]. Briefly, the molarity of urea in the buffer is calculated by weight based on the following descriptions. First, the ratio of the density of the solution to that of water, d/d_0, is calculated by equation [3]. ω is the weight fraction of urea to the total weight of the solution. The total weight of the solution was measured after pH adjustment.

$$d / d_0 = 1 + 0.2658\omega + 0.0330\omega^2 \qquad [3]$$

The volume of the solution is calculated based on equation [4], where *weight* $_{total}$ is the total weight of the solution.

$$volume = weight_{total} / (d / d_0) \quad [4]$$

Then, the molarity of urea-containing buffer is calculated based on equation [5], where *weight* $_{urea}$ is the weight of urea added. The *M.W.urea* is the molecular weight of urea, 60.056.

$$Molarity = weight_{urea} / (M.W.urea)(volume) \quad [5]$$

The molarity of urea was also calculated by refraction based on equation [6].

$$Molarity = 117.66(\Delta N) + 29.753(\Delta N)^2 + 185.56(\Delta N)^3 \quad [6]$$

where ΔN is the difference of refraction of the urea-containing buffer and the buffer without urea.

The molarity of GdnHCl was obtained in the same methods but using the following equations: The density ratio, d/d_0, is calculated by equation [7], where ω is the weight fraction of GdnHCl added to the total weight of solution measured after pH adjustment.

$$d / d_0 = 1 + 0.2710\omega + 0.0330\omega^2 \quad [7]$$

The volume of the solution was calculated by equation [8], where *weight* $_{total}$ is the total weight of the solution.

$$volume = weight_{total} / (d / d_0) \quad [8]$$

Then, the molarity of GdnHCl-containing buffer is calculated based on equation [9], where *weight* $_{GdnHCl}$ is the weight of GdnHCl added. The *M.W. GdnHCl* is the molecular weight of GdnHCl, 95.53.

$$Molarity = weight_{GDnHCl} / (M.W.GdnHCl)(volume) \quad [9]$$

The molarity of GdnHCl was calculated by refraction based on equation [10].

$$Molarity = 57.147(\Delta N) + 38.68(\Delta N)^2 - 91.60(\Delta N)^3 \quad [10]$$

where ΔN is the difference of refraction of the GdnHCl-containing buffer and the buffer without GdnHCl. The urea-containing or GdnHCl-containing buffers were used only if the calculated molarities from weight and refraction differ within 1 %.

2.2 Plasmid construction

2.2.1 pCARD-RICK

The cDNA encoding the CARD domain of RICK was obtained from PCR amplification using the template pcDNA3-RICK [52] (kindly provided by Dr. Gabriel Núñez, University of Michigan). The forward and reverse primers used for PCR amplification were 5'-GTCTGCAGCATATGATAGCCCAG-3' and 5'-GTAAATTTAACTCGAGTCATCTAGAAACC-3', respectively. The PCR product was subcloned into pET21a (Novagen) using the restriction enzymes *Nde*I and *Xho*I. The resulting plasmid was named pCARD-RICK. The resulting protein, CARD of RICK (RICK-CARD), is composed of 95 residues, from isoleucine 426 to arginine 519 of RICK. In this construct, an N-terminal methionine was generated by the cloning procedure. The calculated molecular weight of RICK-CARD is 10,917 Da.

2.2.2 pCARD-Procaspase 1

The cDNA encoding the CARD domain of procaspase-1 was obtained from PCR amplification using the template cDNA of full-length procaspase-1 (unpublished data). The forward and reverse primers used for PCR amplification were 5'-GTCGCGGATCCCATATGGCCGACAAGG-3' and 5'-GGTAATTTCCTCGAGTTTAATCTGCTG-3', respectively. The PCR product was subcloned into pET21a (Novagen) using the restriction enzymes *Nde*I and *Xho*I. The resulting plasmid was named pCARD-procasp1. The resulting protein, CARD of procaspase-1 (Pro-1-CARD), is composed of 92 amino acids from methionine 1 to aspartate 92 of Procaspase-1. The calculated molecular weight of Pro-1-CARD is 10,328 Da.

2.2.3 Plasmid constructions for proline mutants of RICK-CARD

The plasmids of single proline mutants of RICK-CARD including P47A, P85A, and P87A were done by site-directed mutagenesis (Stratagene) using pCARD-RICK as template in PCR amplification. The palindrome PCR primers used for mutagenesis are as follows

: The primer sequence for P47A is 5'-CTTGTTAGTAC**T**AAG*G*CTACAAGGACCTC-3'. A unique *Sca*I restriction site was created and was underlined in the above sequence. The resulting protein contains a silent mutation at 3rd base of threonine 45, which is shown in bold (see Figure 4 for amino acid sequence). The mutated base for proline 47 to alanine is in italic. The primer sequence for P85A is 5'-CAAATGGGTCT**G**CAG*G*CT TAC CCG-3'. A unique *Pst*I restriction site is created and underlined in the above sequence. A silent

mutation at the 3rd base of leucine 83 is generated and labeled in bold. The mutated base for proline 85 to alanine is in italic. The primer sequence for P87A is 5'-CAAATGGGTCTTCAG<u>CCA</u>TA**T***G*CGGAAATACTTGTG -3'. A unique *Nde*I restriction site is created and underlined in the above sequence. The resulting protein contains two silent mutations. One is at 3rd base of proline 85 and the other is at the 3rd base of tyrosine 86. They are shown in bold. The mutated base for proline 87 to alanine is in italic.

The plasmids of double proline mutations of RICK-CARD include pP47/85A, pP47/87A, and pP85/87A. The plasmid pP47/85A was done by using P85A primers and pP47A as template for site-directed mutagenesis. Two unique sites, *Sca*I and *Pst*I were generated. The plasmid pP47/87A was done by using P87A primers and pP47A as template. Two unique sites, *Sca*I and *Nde*I were generated. pP85/87A was done by using 5'-GGGTCTGCAG*G*CATA T*G*CGGAAATACTTGTGG -3' as mutagenesis primers and pP87A as template. Two unique sites, *Pst*I and *Nde*I were generated.

The plasmid of triple proline mutations of RICK-CARD, P47A/P85A/P87A or P_3A, was constructed by using pP47A/87A as template and P85/87A primers in site-directed mutagenesis. All constructed plasmids were confirmed by sequencing from both strands (University of Maine, Sequencing Facility). A summary of the plasmid information is listed in Table 1.

Table 1. Plasmid of the RICK-CARD proline mutants

Plasmid name	**Mutagenesis primer**	**template**	**Mutation** **Silent mutation (Bold)** ***Pro to Ala (Italic)***	**Unique restriction site**	**Protein name**
pP47A	P47A primers	PCARD-RICK	ACCAAGCCT (45-TKP-47) to AC**T**AAG*G*CT (45-TKA-47)	*Sca*I	P47A
pP85A	P85A primers	PCARD-RICK	CTTCAGCCT (83-LQP-85) to CT**G**CAG*G*CT (83-LQA-85)	*Pst*I	P85A
pP87A	P87A primers	PCARD-RICK	CCTTACCCG (85-PYP- 87) to CC**A**TA**T***G*CG (85-PYA-87)	*Nde*I	P87A
pP47/85A	P85A primers	PP47A	Same mutations described in pP47A and pP85A	*Sca*I and *Pst*I	P47A/P85A
pP47/87A	P87A primers	PP47A	Same mutations described in pP47A and pP87A	*Sca*I and *Nde*I	P47A/P87A
pP85/87A	P85/87A primers	PP87A	CTTCAG CCTTACCCG (83-LQPYP-87) to CT**G**CAG*G*C**A**TA**T***G*CG (83-LQAYA-87)	*Pst*I and *Nde*I	P85A/P87A
p P_3A	P85/87A primers	PP47/87A	The same mutation as described in pP47A and pP8587A	*Sca*I, *Nde*I, and *Pst*I	P_3A

2.3 Protein Purification

2.3.1 RICK-CARD

E. coli BL21 (DE3) cells were transformed with pCARD-RICK. Cells were grown in Fernbach flasks containing 1 liter of LB media, ampicillin (50 μg/ml) and antifoam-C (0.003%) at 37 °C and 300 rpm. When the A_{600} of the culture reached 1.2, IPTG was added to a final concentration of 0.4 mM to induce protein expression. The culture was harvested 4 hours after induction by centrifugation for 20 min at 5,000 rpm (GS3 rotor, Sorvall) and 4 °C. The cells were resuspended in about 15 ml lysis buffer (30 mM Tris-HCl, pH 8, 1 mM DTT, and 1 mM EDTA) per one liter of culture, and were broken by French Press at 16,000 psi followed by centrifugation for 30 min at 14,000 rpm (SA-600 rotor, Sorvall) to separate the soluble and insoluble portions. The pellet was resuspended and washed for 1 hr at room temperature by stirring the pellet in buffer A (30 mM Tris-HCl, pH 8, and 1 mM DTT) containing 50 mM $MgCl_2$ and 0.1 mg DNase I per 100 ml of lysate. After the wash, the solution was centrifuged at 14,000 rpm for 30 min at 4 °C. The pellet was dissolved in 5 ml buffer A containing 8 to 9 M urea per one liter culture, and was incubated on ice for one hour. The solution was then centrifuged for 30 min at 14,000 rpm and 4 °C to remove cell debris. The solubilized protein was rapidly diluted into ice-cold buffer A containing 0.1 M NaCl, and the solution was incubated for 15 min. The final protein concentration for refolding was approximately 0.05 mg/ml based on estimation by A_{280}. The refolded protein was concentrated to approximately 50 ml and centrifuged for 30 min at 14,000 rpm. The supernatant was dialyzed against buffer A (4 °C).

RICK-CARD eluted from a DEAE-sepharose column (2.6 x 25 cm, Sigma) equilibrated with buffer A at 4 °C using the following procedure. The protein was loaded onto the column, and the column was washed with 200 ml of buffer A. A two-step gradient procedure was used to elute the protein. First, a linear gradient of buffer A containing 0 to 300 mM NaCl, with a total volume of 500 ml, was used. This was followed by a linear gradient of buffer A containing 300 to 500 mM NaCl, with a total volume of 500 ml. The flow rate was 4 ml/min. RICK-CARD eluted between 200 and 250 mM NaCl. The fractions were analyzed by 4-25 % SDS-PAGE. The pure protein fractions were pooled, concentrated and dialyzed against buffer A. The extinction coefficients at 280 nm of RICK-CARD was calculated [60] to be 8,250 $M^{-1}cm^{-1}$.

2.3.2 Pro-1-CARD

Pro-1-CARD was purified using protocols described previously for RICK-CARD in the previous paragraph with the following modifications. Pro-1-CARD was loaded onto a DEAE-sepharose column (2.6 x 25 cm, Sigma) equilibrated with buffer A (30 mM Tris, pH8, 1 mM DTT) at 4°C. The column was then washed with 200 ml of buffer A, and a two-step gradient procedure was used to elute the protein. First, a linear gradient of buffer A containing 150 to 400 mM NaCl, with a total volume of 800 ml, was used. This was followed by a linear gradient of buffer A containing 400 to 500 mM NaCl, with a total volume of 400 ml. The flow rate was 4 ml/min. Pro-1-CARD eluted between 300 and 350 mM NaCl. The fractions were analyzed by 4-25% SDS-PAGE. The pure protein fractions

were pooled, concentrated and dialyzed against buffer A. The extinction coefficients at 280 nm of Pro-1-CARD was calculated [60] to be 2,650 $M^{-1}cm^{-1}$.

2.3.3 Proline mutants of RICK-CARD

E. coli BL21 (DE3) Lys S cells were transformed with the plasmids of the proline mutants individually and used for the protein purification. The proline mutants P47A, P85A, and P85A/P87A were purified using protocols described previously for RICK-CARD in section 2.3.1. The P47A/85A, P47A/P87A, and P_3A were purified using protocols described for RICK-CARD but with different conditions for growing the culture. Those proteins were produced in the bacteria at 37°C until the absorbance at 600 nm reached 0.6. The culture was then induced with 0.8 mM IPTG and grew at 20°C overnight. The proteins were present in inclusion bodies. P87A was purified using the protocol as described in section 2.3.1 except the following steps. After washing and dissolving the insoluble fraction in about 8 M urea-containing buffer A, the dissolved protein was centrifuged at 14,000 rpm, 4 °C for 30 min, and the supernatant was filtered through 0.45 μm filter membrane. The protein solution was loaded onto a Q-sepharose column (3.2 x 14) equilibrated with 4 M urea-containing buffer A. A linear gradient of 400 ml of 0-400 mM NaCl in 4 M urea-containing buffer A was used to elute the protein. The flow rate was about 4 ml/min. P87A eluted around 100 mM NaCl. The pure protein fractions were pool, concentrated and refolded against ice-cold buffer A containing 0.1 M NaCl at approximately 0.5 mg/ml by either the rapid dilution described for RICK-CARD or dialysis. The pure protein was obtained after refolding and was dialyzed against buffer A.

2.4 Molecular modeling

The homology models of RICK-CARD, Pro-1-CARD, and RICK-CARD mutants were generated using the homology module of Insight II (Molecular Simulation Inc.). The sequences of the CARDs were aligned based on the conserved residues, hydrophobic core residues, and the secondary structural elements, and the alignment used to generate the structural model is shown in Figure 7B, 18B, and 29. The NMR structure of ICEBERG (from Protein Data Bank, accession number 1DGN), another CARD containing protein [61], was used as the modeling template for RICK-CARD and Pro-1-CARD (Figure 7A and 18A). The RICK-CARD model was used as template for the models of the proline mutants of RICK-CARD. The sequence identity of ICEBERG and RICK-CARD is about 21% and for ICEBERG and Pro-1-CARD is 51%. The superimpositions of the models are generated by "lsqman", a part of the coordinate manipulation package of programs, which aligns the alpha-carbons in the structures.

2.5 Fluorescence spectroscopy

Fluorescence emission spectra were obtained with a PTI C-61 spectrofluorometer (Photon Technology International). Fluorescence spectra of RICK-CARD and the mutants were obtained by exciting the proteins (3 μM) at 280 or 295 nm, and the emission spectra were measured from 300 to 400 nm. Pro-1-CARD (6 μM) was excited at 280 nm, and the emission spectra were measured from 300 to 400 nm in buffer B (10 mM Tris, pH 8, 1 mM DTT). All spectra were corrected for buffer signals. The experiments were maintained at a constant temperature of 25 °C using a circulating water bath.

2.6 Circular dichroism (CD) spectroscopy

CD spectra were obtained with a J600A spectropolarimeter (Jasco Inc.) at a constant temperature of 25 °C using a circulating water bath. The path length of the CD cells used was 0.1 cm (far-UV) or 1 cm (near-UV). All spectra were corrected for buffer signals. The protein concentrations of RICK-CARD for both far- and near-UV CD were 57 μM, and the samples were in buffer B (10 mM Tris-HCl, pH 8, 1 mM DTT). The protein concentrations of Pro-1-CARD for far-UV CD in buffer B and 6 M urea-containing buffer B were 92 and 35 μM, respectively, whereas, the Pro-1-CARD concentrations for near-UV CD in buffer B and 6 M urea-containing buffer B were 225 μM and 117 μM, respectively. The Pro-1-CARD concentration for both near-UV and far-UV CD in 6 M GdnHCl-containing buffer B was 183 μM. For far and near-UV CD, the protein concentrations of P47A, P85A, P87A, P47A/P85A, P47A/P87A, P85A/P87A, and P_3A were 23.7, 17.3, 38.9, 19.2,45.4, 10.5, and 28 μM, respectively. All CD spectra of the proline mutants of RICK-CARD were done in buffer B.

2.7 Size-exclusion chromatography

A Superose 6HR 10/30 FPLC column (10x300 mm) (Amersham Pharmacia) is used to characterize the molecular weight of Pro-1-CARD. Molecular weight markers (100 μl) were loaded onto the column individually in buffer C (30 mM Tris, pH 8, and 250 mM NaCl). The concentrations of the markers including dextran blue (2000 kD), bovine serum albumin (66.2 kD), carbonic anhydrase (29 kD), cytochrome c (12.4 kD), and aprotinin (6.5

kD) were 2, 5, 2, 2, and 3 mg/ml, respectively. The flow rate was 0.1 ml/min, and the fractions were collected with 10 drops per fraction (~0.47 ml). The fractions were measured by absorbance at 280 nm. A volume of 150 μl of Pro-1-CARD, 85 μM, was loaded onto the column in buffer C containing 0.5 mM TCEP. The molecular mass of Pro-1-CARD obtained from the standard curve is 10,656 kD as shown in Figure 22.

2.8 Fluorescence Titration Studies for CARD-CARD Interaction.

The fluorescence titration studies were preformed by monitoring the tryptophanyl fluorescence emission of RICK-CARD (excitation at 295 nm) as described in the methods of fluorescence spectroscopy. Pro-1-CARD, from 0 to 12 μM, was titrated into a solution containing RICK-CARD or its proline mutants at 3 μM individually in buffer A. The final ratio of Pro-1-CARD to RICK-CARD was 0 to 4. The sample was excited at 295 nm. The signals were corrected for dilution and background signal. The fluorescence emission of RICK-CARD at 330 nm was plotted against the ratio of Pro-1-CARD/RICK-CARD. The result was fitted to equation [13] as derived in the following equations.

Assuming a simple protein to ligand binding reaction,

$$P + L \rightarrow PL \qquad [11]$$

$$\gamma = [PL]/[P_T] \qquad [12]$$

where P is the protein concentration, L is the ligand concentration, γ is the fraction of protein bounded with ligand (PL) and total protein concentration (P_T). Therefore, the dissociation constant, K_d, for the reaction will be

$$K_d = [P][L]/[PL] = (P_T - \gamma P_T)(L_T - \gamma P_T)/\gamma P_T \qquad [13]$$

where L_T is the total ligand concentration. The protein concentration is equal to the concentration of the free protein, which does not bind to ligand. In terms of fluorescence signal changes, γ can be described as fraction of total signal changes (equation [14]).

$$Fobs = F0 - \gamma(F0 - F1) \qquad [14]$$

where F_{obs} is the observed fluorescence signal, F0 is the initial fluorescence signal, F1 is the final fluorescence signal. By solving g in equation [13], we can put γ value to equation [14] and obtain equation [15].

$$Fobs = F0 + (F1 - F0)(2Pt)^{-1}[(Kd + Lt + Pt) - ((Kd + Lt + Pt)^2 - 4PtLt)] \qquad [15]$$

Here, in our experiments, P_T is RICK-CARD at 3 μM, L_T is Pro-1-CARD from 0 to 4 μM.

For the proline mutants of RICK-CARD, the same experiments were performed. The data were fit to equation [15]. In Figure 23, the changes of the fluorescence signal at 330 nm were normalized with the initial fluorescence of the RICK-CARD in the absence of Pro-1-CARD. The normalized fluorescence signals at 330 nm were plotted against the ratio of Pro-1-CARD and RICK-CARD.

2.9 Equilibrium Folding/Unfolding

The equilibrium folding/unfolding studies were done with either a titrator (Olis Inc.) or manual preparation. The program and operation of titrator were examined by reproducing the equilibrium folding/unfolding of RNase A. The details of the program and the results from RNase A were described in Appendix D.

2.9.1 RICK-CARD

Stock solutions of native RICK-CARD were prepared in buffer A for fluorescence measurements and in the buffer B for CD measurements. Stock solutions of unfolded RICK-CARD were prepared in the same buffers containing 6 M urea. Both stocks (native and unfolded protein) contained the same protein concentration and were incubated at room temperature for at least 30 min prior to the experiments. A titrator (Olis Inc.) was used for equilibrium folding studies by fluorescence emission, and an incubation time of 1,500 sec was used after each titration. It was determined that this time is sufficient to allow for equilibration. Samples were excited at 280 or 295 nm, and fluorescence emission was monitored from 300 to 400 nm.

The emission spectrum at each urea concentration was examined as described previously [62] to determine the average emission wavelength, <λ>, using equation [16],

$$< \lambda >= \sum_{i=1}^{N} (I_i \lambda_i) / \sum_{i=1}^{N} (I_i) \qquad [16]$$

where I_i is the fluorescence emission signal at wavelength λ_i. The normalized average emission wavelength was plotted versus urea concentration. The data were fit to a two-state equilibrium process (N ⇆ U) described by Santoro and Bolen [63]. For the CD experiments, titrations were performed manually by mixing unfolded protein with the native protein. A quartz cuvette with 1 cm path length was used in these experiments. The signal at 220 nm was measured for 120 sec, and the data were averaged and corrected for background signals.

2.9.2 Pro-1-CARD

Equilibrium folding/unfolding experiments of Pro-1-CARD were done by manual preparation or by titration operation. For the manual preparation, each sample was prepared in buffer B to contain the desired urea concentrations as indicated in the figures and a desired protein concentration. The experiments were repeated with different protein concentrations ranging from 3 to 40 μM. The samples were incubated overnight at 25ºC, then, the fluorescence and CD spectra were examined. For the titration operation, stock solutions of native Pro-1-CARD were prepared in buffer B. Stock solutions of unfolded Pro-1-CARD were prepared in buffer B containing 8 M urea. Both stocks (native and unfolded proteins) contained the same protein concentration as desired and were incubated at room temperature for at least 40 min prior to the experiments. The average emission wavelength was calculated based on the method described in section 2.9.1. The equilibrium folding/unfolding experiments of Pro-1-CARD in the absence of DTT were performed in buffer B without DTT. The protein stock was dialyzed against buffer B without DTT. The final DTT concentration left in the samples was estimated to be 0.05 μM (Figure 21C).

2.9.3 Proline mutants of RICK-CARD

Equilibrium folding/unfolding experiments of proline mutants were done by manual preparation described in section 2.9.2 or by titration operation described in section 2.9.1. The experiments were done with 3 μM protein in buffer B. The samples were incubated for more than 30 min at 25ºC, then fluorescence and CD properties were examined.

2.10 Stopped-Flow Fluorescence Studies

The kinetic folding/unfolding experiments were performed using a stopped-flow spectrofluorometer (SX.18MV, Applied Photophysics, UK). The temperature was controlled at 25°C using a circulating water bath. The samples were excited at 280 nm and the fluorescence emission was measured using a cutoff filter of 305 nm.

2.10.1 RICK-CARD

The refolding experiments were done by mixing stock protein solutions (33 μM) that had been prepared in buffer A-containing 4 M urea (unfolded protein) with buffer A. The final protein concentration was 3 μM, and a mixing ratio of 1:10 was used. Stock protein (33 μM) in 4 M urea-containing buffer was mixed with buffer A-containing urea between 0 and 4 M, as shown in the figures. For unfolding experiments, stock protein (33 μM) in buffer A was mixed with buffer A-containing urea between 0 and 4 M. The final protein concentration was 3 μM, and a mixing ratio of 1:10 was used. The final urea concentrations are shown in the figures.

2.10.2 Pro-1-CARD

The single mixing experiments were preformed as described in section 2.10.1 except the following: The refolding experiments were done by mixing stock protein solutions (33 μM) that had been prepared in buffer B-containing 6 M urea (unfolded

protein) with buffer B. For unfolding experiments, stock protein (33 μM) in buffer B was mixed with buffer B-containing urea between 0 and 6 M.

2.10.3 Proline mutants of RICK-CARD

The experiments were preformed as described in section 2.10.2 except the experiments for P87A. The experiments for P87A were done as described as for RICK-CARD in section 2.10.1 by using buffer B.

2.11 Temperature studies of RICK-CARD refolding

The refolding experiments of RICK-CARD were performed in stopped-flow fluorescence as described in section 2.10.1. The refolding experiments were done in different temperatures ranging from 10 to 30 °C. We have normalized the refolding traces based on the controls. The refolding traces were fit to a three exponential equation and the observed rate constants were obtained. The activation energy was calculated based on Arrhenius equation.

The Arrhenius equation is derived from an empirical equation as shown in equation [17],

$$d \ln k / dT = E_a / RT^2 \qquad [17]$$

where E_a is the activation energy, R is the gas constant, and T is the absolute temperature. Equation [15] can be integrated to give equation [18] or [19]. A is a constant.

$$\ln k = \ln A - E_a / RT \qquad [18]$$

$$k = A \exp(-E_a / RT) \qquad [19]$$

By plotting the natural logarithm of the rate constant versus the reciprocal value of the absolute temperature (Arrhenius plot), the slope of the linear fit to the data is equal to the value of activation energy divided to the gas constant.

2.12 Double jump stopped-flow experiments

The experiments were performed with the stopped-flow spectrofluorometer (SX.18MV, Applied Photophysics, UK) using the sequential mixing mode. The temperature was held constant at 25 °C. In general, the first jump was performed with a mixing ratio of 1:1 using two buffer solutions, described below. After a specified delay time, the solution from the first jump was rapidly mixed with a third solution using a mixing ratio of 1:5. Incubation (delay) times from 0.1 to 900 sec were used after the first jump.

2.12.1 Double jump of RICK-CARD

For double jump experiments [64], native RICK-CARD (36 μM) in buffer A was mixed with buffer A-containing 8 M urea for the first jump. After various delay times, as described in the text, the protein solution was mixed with buffer A for the second jump. The signal trace was monitored for 200 sec. The final protein concentration was 3 μM, and the final urea concentration was 0.67 M. The native protein signal was determined from the final signal of the experiment that used a 900 sec delay time. To obtain the signal of the unfolded protein, the third solution, described above, was changed to buffer A-containing 4 M urea.

2.12.2 Double jump of Pro-1-CARD

Double jump experiments of Pro-1-CARD were described in section 2.12.1 with exception as follows. Native Pro-1-CARD (36 μM) in buffer B was mixed with buffer B-containing 9.6 M urea for the first jump with a mixing ratio of 1:1. After a specified delay time, as indicated in the figures, the protein solution was mixed with buffer B for the second jump using a mixing ratio of 1:5. The signal trace was monitored for at least 10 sec. The final protein concentration was 3 μM, and the final urea concentration was 0.8 M. The native protein signal was determined from the protein signal in 0.8 M urea containing buffer B. The unfolded protein signal was obtained from the protein signal in 4.8 M urea containing buffer B.

2.12.3 Double jump of proline mutants of RICK-CARD

Double jump experiments of the proline mutants were described in section 2.12.2 with the following modifications. Native proline mutants (36 μM) in buffer B was mixed with buffer B-containing 9.6 M urea for the first jump. After various delay times, as described in the text, the protein solution was mixed with buffer B for the second jump. The signal trace of refolding from the second jump was monitored. The final protein concentration was 3 μM, and the final urea concentration was 0.8 M. The control signal of the native protein was determined from a solution containing Pro-1-CARD at 3 μM in 0.8 M urea-containing buffer B. The control signal of the unfolded protein was obtained by a solution containing Pro-1-CARD, 3 μM, in 4.8 M urea-containing buffer B.

2.13 Interrupted refolding

2.13.1 RICK-CARD and the proline mutants

For interrupted refolding experiments [65; 66], unfolded RICK-CARD (36 μM) in buffer A-containing 3 M urea was mixed with buffer A for the first jump. After various delay times, as shown in the figures, the protein solution was mixed with buffer A-containing 4 M urea. The final urea concentration was 3.58 M, and the final protein concentration was 3 μM. The signal for the unfolded protein was obtained from the final signal of the experiment that used a 900 sec delay time. To obtain the signal of the native protein, the third solution was changed to buffer A-containing 1.5 M urea. The same conditions were used for the interrupted refolding experiments of the proline mutant of RICK-CARD, P87A.

2.13.2 Pro-1-CARD

The interrupted refolding experiments for Pro-1-CARD were set up differently than that for RICK-CARD. The first jump was performed with a 1:10 mixing ratio and the second jump was using 1:5 mixing ratio. Unfolded Pro-1-CARD (198 μM) in buffer B-containing 5 M urea was mixed with buffer B for the first jump. After various delay times, as shown in the figures, the protein solution was mixed with buffer B-containing 6 M urea. After the first jump, the urea concentration is 0.454 M. After the second jump, the urea concentration was 5 M, and the final protein concentration was 3 μM. The signal for the unfolded protein was obtained from the final signal of the experiment that used a 100 sec

delay time. To obtain the signal of the native protein, the third solution was changed to buffer B.

Chapter 3

RESULT I

RICK-CARD

Chapter 3

Results: Part I

RICK-CARD

3.1 General properties of RICK-CARD.

The CARD of RICK (called RICK-CARD) used in these studies is a protein composed of 95 residues. The protein contains six α-helices arranged in a Greek key topology (Figure 3). With a few exceptions, the amino acid sequence identities of the CARD family members are less than 20% (see also Figure 4) [41; 57; 67], although structural data for several CARDs demonstrate that the Greek key topology is conserved [61; 68; 69; 70; 71; 72; 73; 74]. While the structure of RICK-CARD has not been determined, we generated a homology model using the solution structure of ICEBERG [61] as the template and the sequence alignment shown in Figure 4. The results of the model are shown in Figure 7, where the model of RICK-CARD is compared to the structure of ICEBERG. The model of RICK-CARD shows a well-packed core of hydrophobic residues with exposed charged residues. The least defined region of the RICK-CARD model is helix-6, but as described below, this is true of other CARD proteins as well. RICK-CARD contains one tryptophan, which is located in helix-1, and two tyrosines, which are located in helices-3 and -6. A single phenylalanine is located in helix-5. There are no disulfide bonds or prosthetic groups.

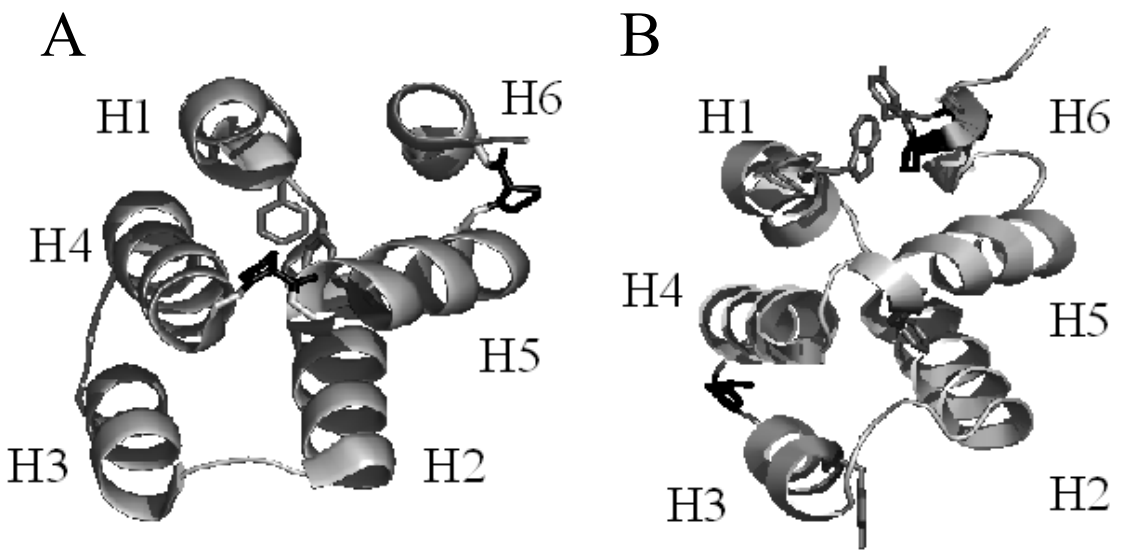


Figure 7. (A)The structure of ICEBERG and (B) the homology model of RICK-CARD. Structures were drawn with Pymol (DeLano scientific, CA). ICEBERG (PDB code 1DGN) is shown in ribbon diagram with the proline residues in black and the aromatic residues, two phenylalanines, in gray. The homology model of RICK-CARD in ribbon diagram was generated by Insight II using ICEBERG (PDB code 1DGN) as the template. The sequence alignment for modeling is shown in Figure 4. The three prolines are colored in black. The aromatic residues, one tryptophan in helix-1, two tyrosines in helices-3 and -6, and one phenylalanine in helix-5, are in gray.

3.2 Fluorescence and CD spectra of RICK-CARD

Fluorescence emission spectra of RICK-CARD, using intrinsic fluorescence probes, are shown in Figure 8A. When the native protein is excited at 295 nm, which excites only the tryptophanyl residue, the emission maximum is 330 nm, demonstrating that the tryptophanyl residue is relatively buried. The spectrum shifted to 350 nm when the protein was unfolded in urea-containing buffer, demonstrating that the tryptophanyl residue is exposed under these conditions. In addition to fluorescence emission, we examined the secondary and tertiary structures of RICK-CARD using near- and far-UV circular dichroism. The far-UV CD spectrum (Figure 8B) shows double minima at 206 nm and 218 nm, similar to that of other CARDs [42]. The near-UV CD spectrum (Figure 8C) shows a minimum at 278 nm. Both signals decreased when the protein was incubated in urea-containing buffer. Overall, the data show that the protein is folded with a well-packed tertiary structure. In 4 M urea-containing buffer, the spectra are consistent with those of unfolded proteins.

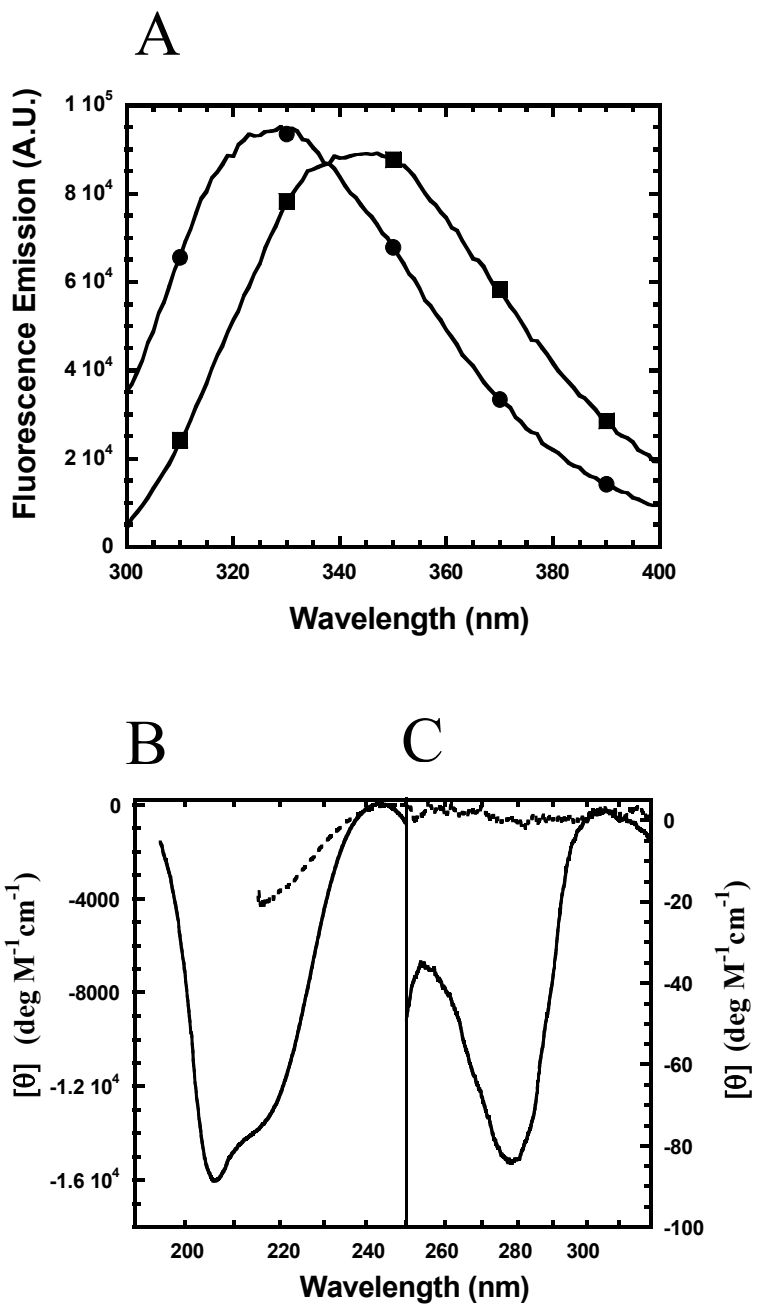


Figure 8. (A) Fluorescence emission spectra of RICK-CARD. The protein (3 μM) in buffer (●) or in 3 M urea-containing buffer (■) was excited at 295 nm. The fluorescence emission was collected from 300 to 400 nm. (B) Far-UV and (C) near-UV CD spectra of RICK-CARD. The spectra of native protein in buffer (solid line) and the denatured protein in 4M urea-containing buffer (dash line) are shown.

3.3 Equilibrium folding of RICK-CARD.

The equilibrium folding/unfolding of RICK-CARD was examined by monitoring the changes in fluorescence emission upon incubation in urea-containing buffer. The samples were excited either at 280 nm, to excite all aromatic residues, or 295 nm, to excite the single tryptophanyl residue. In addition, the changes in CD at 220 nm were determined. Representative data are shown in Figure 9A. The results show that the three spectroscopic probes overlay, suggesting that the loss of secondary structure is concomitant with the loss of tertiary structure. The experiments were repeated over a range of protein concentrations (1.5 to 12 μM) and did not vary from those shown in Figure 9B. While some CARDs have been shown to form homo-oligomers at higher protein concentrations, RICK-CARD is a monomer under the experimental conditions described here (30 mM Tris-HCl, pH 8, 1 mM DTT, 25 ºC, 12 μM or less protein).

The results shown in Figure 9 are well-described by a two-state equilibrium folding model (N ⇆ U), representing only the native and unfolded states of the protein. The conformational free energy, ΔG^{H_2O}, and the cooperativity index, m, determined from fits of the data to the two-state equilibrium model are 3.0 ± 0.15 kcal/mol and 1.27 ± 0.06 kcal/mol/M, respectively. The m-value is thought to represent the change in surface area of the protein exposed to solvent upon unfolding [75; 76]. Scholtz and coworkers [77] have described a qualitative correlation between the m-value and the change in the solvent exposed surface area (ΔASA). The equations are as follows:

$$m = 374 + 0.11(\Delta ASA) \qquad [20]$$

$$\Delta ASA = -907 + 93(\#residue) \quad [21]$$

where m is the m-value in cal/mol/M, #residue is the calculated number of residues that the protein contains. Using their equations, we calculate the ΔASA of RICK-CARD to be 8,200 $Å^2$, corresponding to 98 residues. This is in excellent agreement with the number of amino acids of RICK-CARD, 95.

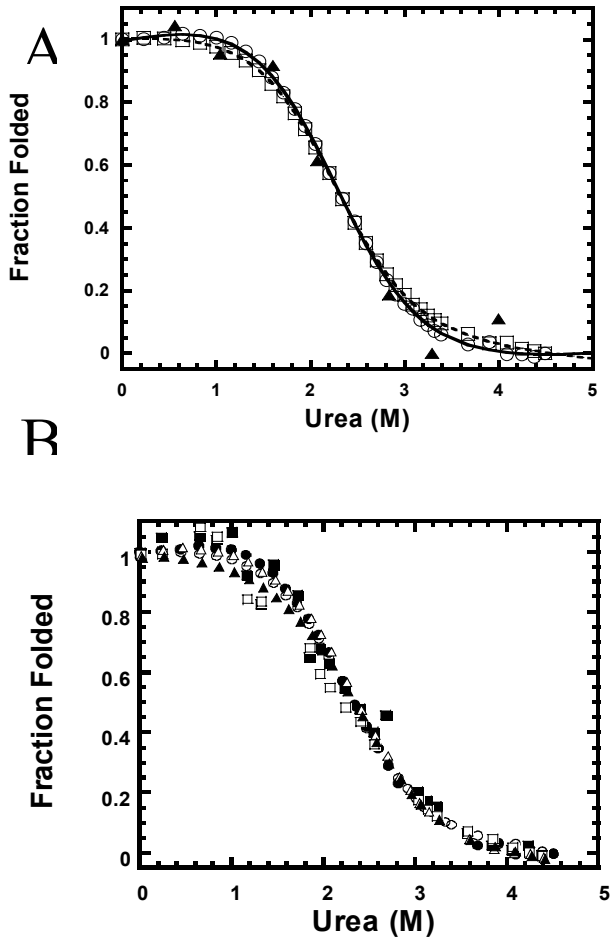


Figure 9. Equilibrium unfolding of RICK-CARD. The average emission wavelength was determined for fluorescence emission spectra as described in methods. The normalized average emission wavelength and CD data versus urea concentrations are plotted. (A) The average emission wavelength from excitation at 280 (○), or 295 nm (□) of RICK-CARD (3 μM) as well as the CD data at 220 nm (▲) of RICK-CARD (6 μM) are shown. The solid line and dashed line represent fits of the data from excitation at 280 and 295 nm, respectively, to a two-state equilibrium folding mechanism. The ΔG^{H_2O} and m-value obtained for excitation at 280 nm are 2.9 kcal/mol and 1.3 kcal/mol/M, respectively. The ΔG^{H_2O} and m-value for excitation at 295 nm are 3.1 kcal/mol and 1.4 kcal/mol/M, respectively, at 25°C. (B) The average emission wavelength from excitation at 280 (closed symbols), or 295 nm (open symbols) of RICK-CARD at 1.5 (square), 3 (circle), 6 (triangle) μM.

3.4 Influence of salts on protein stability.

Long-range electrostatic interactions on the surface of several small proteins [78; 79; 80] have been shown to affect the conformational free energy. In some cases, charge repulsions on the protein surface have been shown to destabilize the native conformation of the protein [78; 79; 80]. In addition, CARD-CARD interactions are known to be influenced by the charged groups on the protein surface [48]. The calculated pI of RICK-CARD is 5.3, although it contains 13 basic and 14 acidic residues, and the charge of the protein at pH 8, the conditions of the equilibrium unfolding experiments, is about -1.5. Because of these considerations, we examined the effects of various salts on the protein stability.

The urea unfolding studies were done in buffers containing NaCl, KCl, KH_2PO_4/K_2HPO_4, or Na_2SO_4. Representative data for the experiments in which NaCl was included in the buffers are shown in Figure 8A. Overall, the unfolding profiles were similar over the entire range of salt, although the amount of urea required to unfold the protein increased between 200 mM and 2 M NaCl. Similar results were obtained for experiments in which the other salts were examined (data not shown), with the exceptions noted below. The data were well-described by a two-state equilibrium model, and the results of the fits are summarized in Figure 8B. The m-value did not change significantly in the presence of salt, indicating that equilibrium folding intermediates are not stabilized under these conditions. At low concentrations of NaCl or KCl, between 0 and 200 mM, the mid-point, or $urea_{½}$, decreased from 2.4 M to ~1.8 M urea, and the ΔG^{H_2O} decreased from 3.0 kcal/mol to 2.5 kcal/mol. These results suggest that in RICK-CARD electrostatic interactions are stabilizing, and the favorable coulombic interactions are screened at low ionic strength.

Alternatively, Lee *et al*. [81] have shown that at ionic strengths near physiological values, pairwise coulombic interactions are weak at distances greater than 5 Å, but in aggregate, the long-range interactions can affect the properties of the protein. For example, long-range, repulsive, interactions can have a significant effect if not balanced by favorable interactions. Thus, it is possible that for RICK-CARD unfavorable coulombic interactions become more important at low ionic strength. Above 200 mM salt, the $urea_{½}$ increased so that at 2 M NaCl, ΔG^{H_2O} was 7.3 kcal/mol. These results are consistent with Hofmeister effects [82; 83]. Similar results have been described for the cold shock protein (Csp) from the thermophilic organism *Bacillus caldolyticus* [78; 84]. As expected for electrostatic interactions that are screened by counterions, the divalent ions were more effective than the monovalent ions, and the effects are shifted to lower salt concentrations. The data show that the response is not cation or anion dependent and probably precludes the binding of ions to specific sites on the protein.

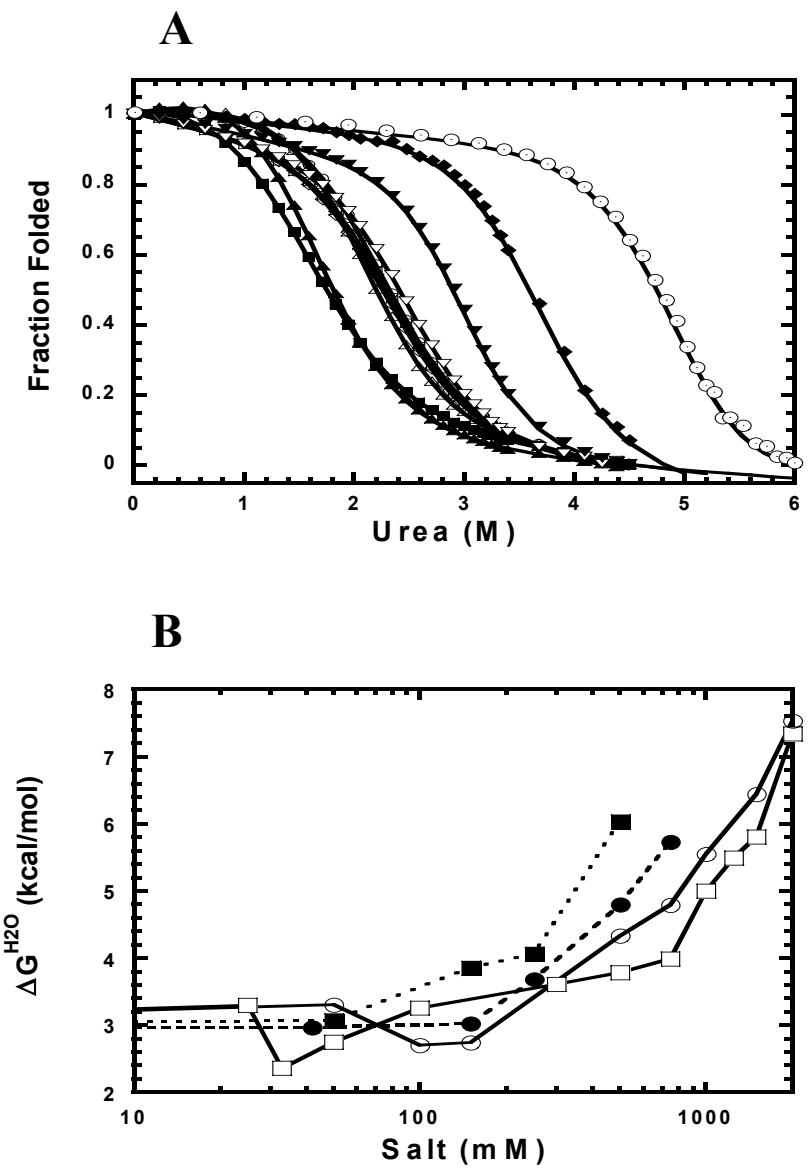


Figure 10. (A) Equilibrium unfolding of RICK-CARD in the presence of NaCl. NaCl concentration ranges from 0 to 2 M are indicated: 0 (○), 33 (■), 50 (▲), 100 (△), 300 (◇), 500 (□), 750 (▽), 1000 (▼), 1500 (◆), 2000 (⊙) mM. The solid lines are fits to a two-state equilibrium mechanism, and the results are shown in Panel B. (B) Conformational free energy versus salt concentration. The ΔG^{H_2O} was obtained from the fits of the equilibrium unfolding of RICK-CARD (3 μM) in the presence of four salts: NaCl (□), KCl (○), KH_2PO_4/K_2HPO_4 (●), and Na_2SO_4 (■). The solid lines are not fits but are present as a guide.

3.5 Folding kinetics of RICK-CARD

The time courses of refolding and of unfolding RICK-CARD were examined in single-mixing stopped-flow experiments by monitoring changes in fluorescence emission. The data are shown in Figure 11A, and several kinetic phases are observed in both experiments. First, a burst phase is present in both the unfolding and refolding profiles. In unfolding, the burst phase accounts for approximately 55% of the total signal change, whereas in refolding the burst phase accounts for approximately 65% of the total signal change. Second, following the burst phase the unfolding data are fit best to a three exponential process in which the three phases have similar amplitudes. The refolding data also are fit best to a three exponential process. Residuals for fits of the refolding data to two or three exponential processes are shown in Figure 11B. The results demonstrate that although the amplitude of the middle phase is small, the data are fit better to a three exponential process.

For refolding, the apparent rate constant of the fast phase is 21 s^{-1}, whereas those of the intermediate and slow phases are 1 s^{-1} and 0.012 s^{-1}, respectively. The intermediate phase accounts for only ~3% of the refolding signal change, whereas those of the fast and slow phases account for ~54% and ~43%, respectively. For unfolding in 4 M urea-containing buffer, the apparent rate constant of the fast phase is ~1 s^{-1}, whereas those of the intermediate and slow phases are 0.08 s^{-1} and 0.014 s^{-1}, respectively. Each phase accounts for about 15% of the total signal change. Overall, the results demonstrate that there are at least four phases in unfolding and in refolding: a burst phase within the instrumental dead time, a fast phase, with a half-time of ~50 msec (refolding) or ~0.5 sec (unfolding), an

intermediate phase with a half-time of ~1 sec (refolding) or ~10 sec (unfolding) and a slow phase, with a half-time of ~50 sec.

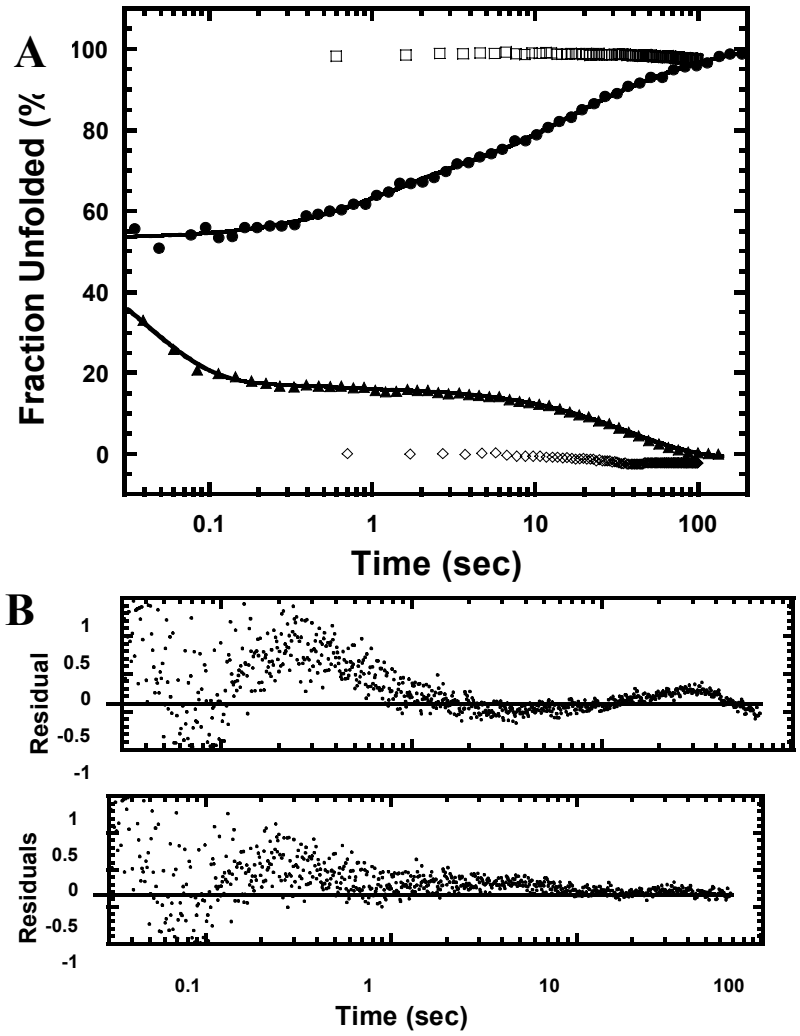


Figure 11. The unfolding and refolding kinetics of RICK-CARD. (A) The refolding and unfolding signals versus time. The native (◇) and unfolded protein signals (□) are indicated. Both unfolding (●) and refolding traces (▲) are shown. Solid lines represent fits of the unfolding and refolding traces to a three exponential equation. Parameters obtained from the fits are given in the text. (B) Residuals of fits of the refolding data (▲) shown in Figure 11A. The upper and lower panels are the residuals of two or three exponential fits, respectively.

To examine whether the burst phases represent the formation of transient intermediates, we performed the experiments in several final urea concentrations, from 0 to 4 M. The signal traces of refolding are shown in Figure 12A and of unfolding are shown in Figure 12B. The initial and final signals are normalized and plotted against urea concentration (Figure 12C). For refolding, the data show that there is a linear decrease of the initial signal from 4 to 0.5 M urea. This demonstrates that the burst phase in refolding does not correlate with the formation of a transiently populated partially folded, or intermediate, conformation. In contrast, the change in the burst phase amplitude is cooperative in the unfolding experiments. The data show that the burst phase in unfolding does correlate with formation of a transiently populated intermediate, and in 4 M urea-containing buffer, the burst phase amplitude accounts for approximately 55% of the total signal change. The burst phase amplitudes for unfolding were fit to a two-state equilibrium unfolding model, and the ΔG^{H_2O} and m-value determined from the fits are 2 kcal/mol and 0.93 kcal/mol/M, respectively. These values account for about 70% of the total ΔG^{H_2O} and m determined from equilibrium unfolding (Figure 9).

As shown in Figure 12C, the final signal (at ~100 sec) accounts for the total change in signal between native and unfolded protein. The final amplitudes were fit to a two-state equilibrium model, and the ΔG^{H_2O} and m-values determined from the fits are 2.6 kcal/mol and 1.3 kcal/mol/M, respectively. These values are in good agreement with those determined from the equilibrium unfolding experiments (Figure 9). As shown in Figure 13, the final signals of unfolding and refolding superimpose to the equilibrium unfolding data,

demonstrating that the final conditions of the kinetic experiments are equivalent to those at equilibrium.

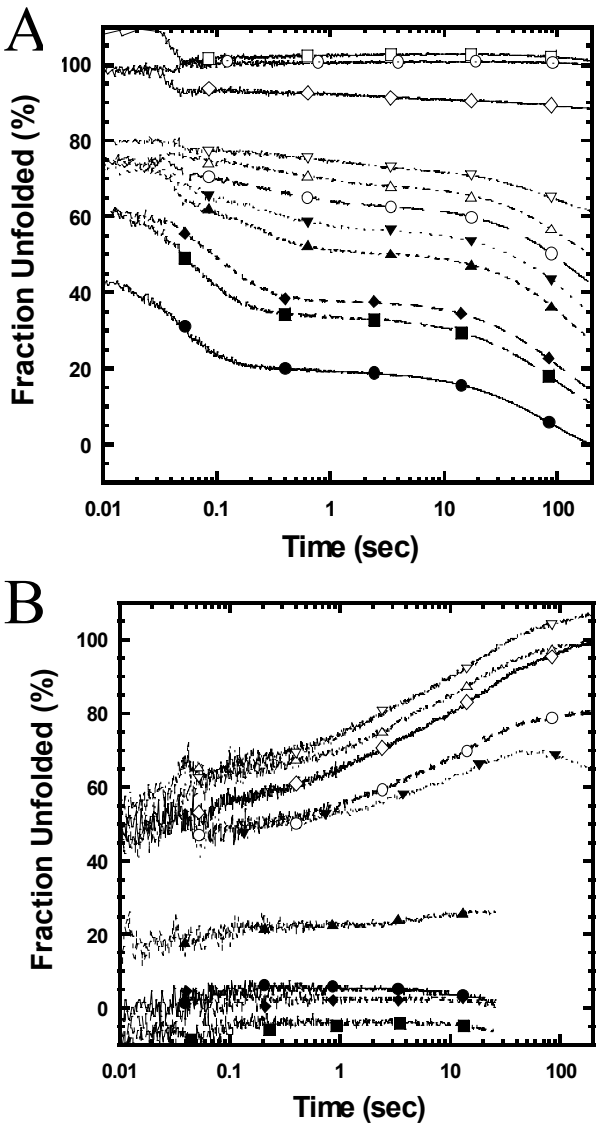


Figure 12. Refolding and unfolding traces of RICK-CARD. (A) The refolding experiments were performed by mixing RICK-CARD in 4 M urea containing buffer to different final urea concentrations. The final urea concentration (M) are as follows: 0 (●), 0.18 (■), 0.36 (◆), 0.54 (▲), 1.26 (▼), 1.44 (○), 1.62 (△), 1.8 (▽), 2.16 (◇), 2.88 (□), and, 6 (⦿). (B) The unfolding experiments were performed by mixing RICK-CARD in buffer to different final urea concentrations. The final urea concentration (M) are as follows: 0 (●), 0.5 (■), 1 (◆), 2 (▲), 2.8 (▼), 3.2 (○), 3.4 (△), 3.8 (▽), and 4 (◇). The signals more than 20 sec are not plotted due to diffusion problem for the unfolding traces in final urea from 2 to 4 M.

C

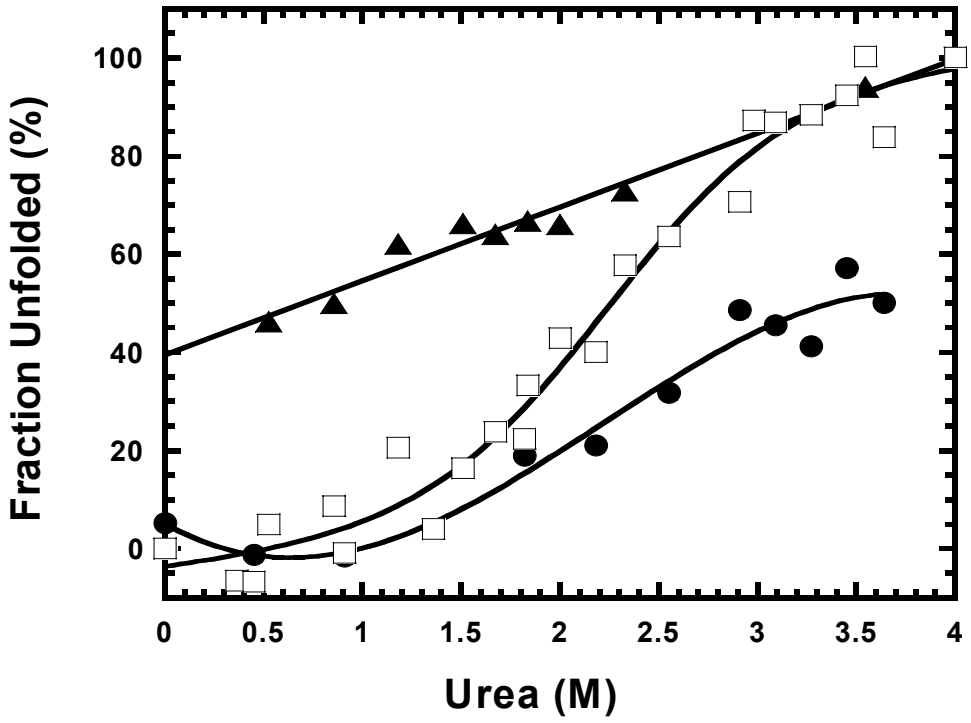


Figure12. (C) Burst phase signal of the refolding and unfolding reactions versus urea concentration. The initial signal of refolding (▲), the initial signal of unfolding (●), the final signal (□) from both studies, and the fits (solid lines) are plotted. The initial signals of refolding were fit to a linear equation. The initial signals of unfolding were fit to a two-state folding mechanism where the ΔG^{H_2O} is 2.0 kcal/mol and the m-value is 0.93 kcal/mol/M. The final signals from both studies were fit to a two-state folding mechanism where the ΔG^{H_2O} is 2.8 kcal/mol and the m-value is 1.26 kcal/mol/M.

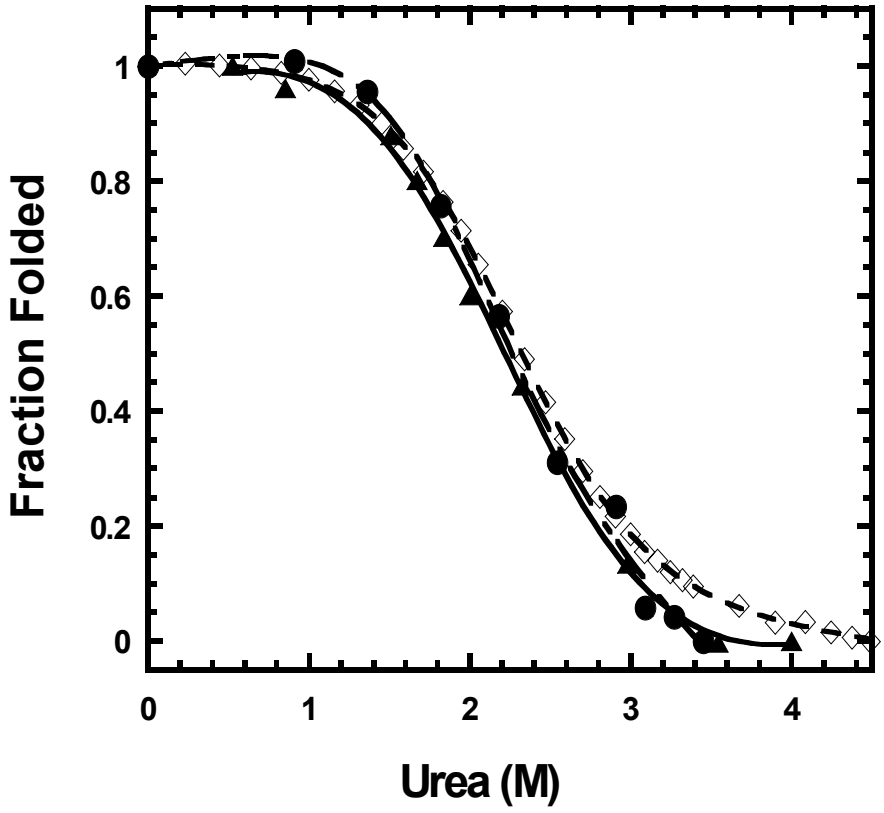


Figure 13 The final signals of refolding (▲) and unfolding (●) versus urea concentrations. Both final signals are superimposed to the equilibrium folding experiment (◇). The lines are fits to a two-state folding mechanism.

In addition to the burst phases, there are three other phases in the refolding and unfolding profiles. The apparent rate constants of the phases were determined in several final urea concentrations, from 0 to 4 M, and the results are shown in Figure 14. For the fast phases, the apparent rate constant, k_{obs}, generates a chevron plot, where the midpoint of the curve is approximately 2.7 M urea, similar to the mid-point determined from the equilibrium unfolding data (2.4 M urea). For the intermediate phase of refolding, k_{obs} decreased with urea such that at concentrations >1.5 M, the refolding data were equally well fit to a two exponential process. For the intermediate phase of unfolding as well as the slow phases, k_{obs} did not vary over two fold, suggesting that those rates were independent of the final urea concentration.

The data for the fast phases were fit to a two-state kinetic mechanism using equation 22 [12; 85],

$$k_{obs} = k_{UN} \exp(-m_{U-TS} * [urea] / RT) + k_{NU} \exp(m_{N-TS} * [urea] / RT) \qquad [22]$$

where k_{obs} is the observed rate constant of the fast phase, k_{UN} and k_{NU} are the rate constants of refolding and of unfolding, respectively, in the absence of denaturant, and $m_{U\text{-}TS}$ and $m_{N\text{-}TS}$ reflect the change in solvent accessible surface area between the unfolded and native states, respectively, relative to the transition state ensemble. Values of 32 and 0.13 s^{-1} were obtained for k_{UN} and k_{NU}, respectively, and values of 0.90 and 0.37 kcal/mol/M were obtained for $m_{U\text{-}TS}$ and $m_{N\text{-}TS}$, respectively.

The Gibbs free energy and m-value between the native and unfolded states of RICK-CARD can be calculated from the kinetic experiments using equations 23 and 24, respectively [12; 85].

$$\Delta G_{H_2O} = -RT \ln(k_{NU} / k_{UN}) \quad [23]$$

$$m_{UN} = m_{U-TS} + m_{N-TS} \quad [24]$$

The calculations show that the Gibbs free energy of RICK-CARD is 3.3 kcal/mol, and the m-value is 1.26 kcal/mol/M. These values are in good agreement with those described for the equilibrium unfolding studies (Figure 9), where the ΔG^{H_2O} is 3.0 ± 0.15 kcal/mol and the m-value is 1.27 ± 0.06 kcal/mol/M. Using equation 25,

$$\beta^{\ddagger} = m_{U-TS} / m_{UN} \quad [25]$$

We calculated the $\beta^{\ddagger}$ value of RICK-CARD to be 0.71. As described by Chen and coworkers [86], the $\beta^{\ddagger}$ value describes the compactness of the transition state ensemble relative to that of the native state ($\beta^{\ddagger}$=1) and of the unfolded state ($\beta^{\ddagger}$=0) ensembles and refers to the position of the transition state on a folding reaction coordinate. The data presented here demonstrate that the transition state ensemble of RICK-CARD is compact, which indicates that many native-state contacts have formed in this folding reaction.

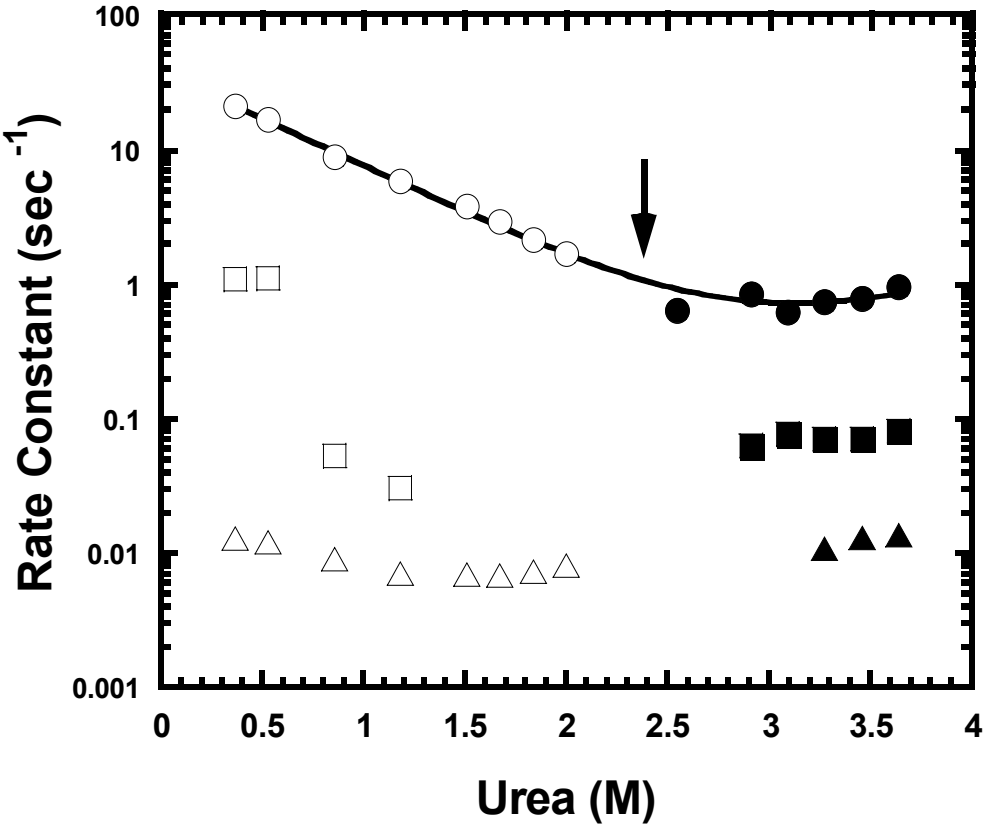


Figure 14. The observed rate constants (k_{obs}) for unfolding and for refolding versus urea concentration. The k_{obs} is determined from the three-exponential fit from the unfolding (closed symbols) and refolding kinetics (open symbols). The midpoint from the equilibrium unfolding studies (Figure 9) is indicated (2.4 M urea). The solid line represents a fit to equation [22]. Parameters obtained from the fits are given in text.

3.6 Double jump stopped-flow studies

Sequential mixing, or double jump, experiments were performed to investigate further the slow phases observed in the folding reactions of RICK-CARD. In the first jump, the native protein was mixed with urea-containing buffer so that the final urea concentration was 4 M. As shown in Figure 9, this urea concentration is sufficient to unfold the protein. During the first jump the time in which the protein was incubated in urea was varied from 0.1 to 900 sec. In the second jump, the protein was returned to native conditions, and refolding was monitored. This procedure allows one to examine slow equilibration processes in the unfolded protein and has been used to examine the isomerization of prolyl (Xaa-Pro) peptide bonds [87; 88]. The homology model of RICK-CARD (Figure 7) suggests that the three prolyl residues are found in the *trans* configuration in the native protein. However, due to the low sequence identity of CARD proteins (Figure 4), the isomeric states of the prolines are not well determined in our structural model. If one of the three prolines resides in the cis isomer, then the unfolded state ensemble would consist of $(1/6)x(5/6)^2$ or 11.6% of molecules in their native isomer at equilibrium, if one assumes a *cis:trans* ratio of 1:5 [34; 89]. If two prolyl residues reside in the *cis* isomer, then the unfolded state ensemble would consist of 2.3% of molecules in their native isomer at equilibrium. Thus, Xaa-Pro bonds in the *cis* configuration can have significant effects on the population of native isomers in the unfolded state ensemble. Assuming a three state model ($N \leftrightarrows U_f \leftrightarrows U_s$) holds for RICK-CARD, in which N refers to the native protein, and U_f and U_s refer to the unfolded protein in the native or non-native isomer, respectively, the double jump experiment should reflect the population of molecules with the native isomeric state prior to prolyl isomerization.

The signal traces and the results of the double jump experiments for RICK-CARD are shown in Figure 15. Surprisingly, the short delay times resulted in little or no native protein formation upon refolding. The data show a lag phase of approximately 20 seconds followed by a second phase in which the population of native protein increases exponentially. At a minimum, the results rule out the sequential model described above. In that mechanism, the non-native isomer develops slowly from the unfolded protein, and the population of native protein decreases with longer delay times. Qualitatively, the two phases observed in the double jump experiment are consistent with the apparent rate constants obtained from the two slow phases of the single mixing unfolding experiment (Figure 11A), 0.08 s^{-1} ($t_{1/2}$=8.6 s) and 0.014 s^{-1} ($t_{1/2}$=49.3 s), respectively. Overall, the results indicate that the intermediates that form during the unfolding of RICK-CARD represent misfolded, or kinetically trapped, species. That is, when RICK-CARD unfolds for a short time, it is trapped in an unfolded-like intermediate state that cannot refold to the native state unless the time for unfolding is prolonged. RICK-CARD is able to refold only after the unfolded species is formed, and this occurs with a half-time of ~50 seconds.

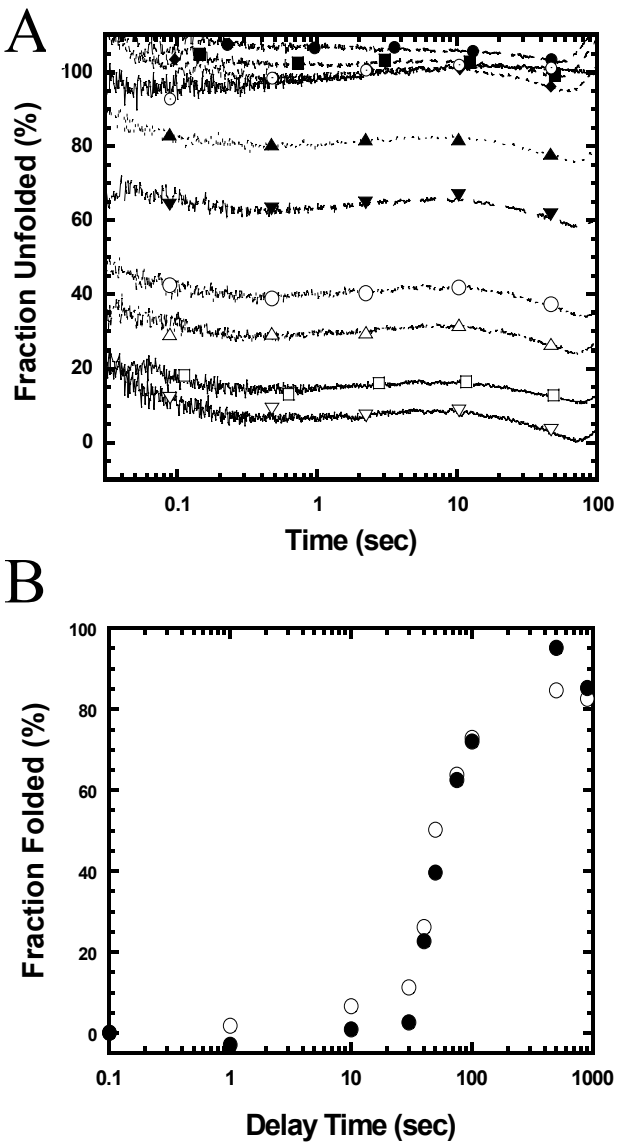


Figure 15. (A) The refolding signal traces from the second jump of double jump experiments. The delay time are indicated as follows: 1 (●), 10 (■), 30 (◆), 40 (▲), 50 (▼), 75 (○), 100 (△), 500 (▽), and 900 sec (□). The unfolded control is indicated (⦿). The signals are normalized to fraction unfolded. (B) Fraction of native protein versus delay time in double jump experiments. The initial (○) and final signals (●) of the signal traces were normalized to fraction folded and plotted versus delay time.

3.7 Interrupted refolding experiments.

The interrupted refolding experiment is similar to the double jump experiment except that it is employed to examine details of the refolding pathway. In the first jump, the unfolded protein is refolded for various amounts of time, and then it is returned to urea-containing buffer, where unfolding is monitored. The initial signal indicates the amount of native protein formed in various times of refolding, thus the interrupted refolding experiment gives a time course for the formation of native protein. The signal traces are shown in Figure 16A and the results for RICK-CARD are shown in Figure 16B. The data demonstrate a lag phase of ~30 seconds, in which there is little or no formation of native protein. Following the lag phase, a second phase occurs in which the population of native protein increases exponentially. These results provide evidence for the transient population of intermediate, or partially folded, conformations. Only the second phase produces native RICK-CARD. The half time (~100 sec) of the appearance of the native state correlates well with the apparent rate constant for the slow phase in refolding (0.008 s^{-1}, $t_{1/2}$= 86 s in 1.5 M urea, Figure 14).

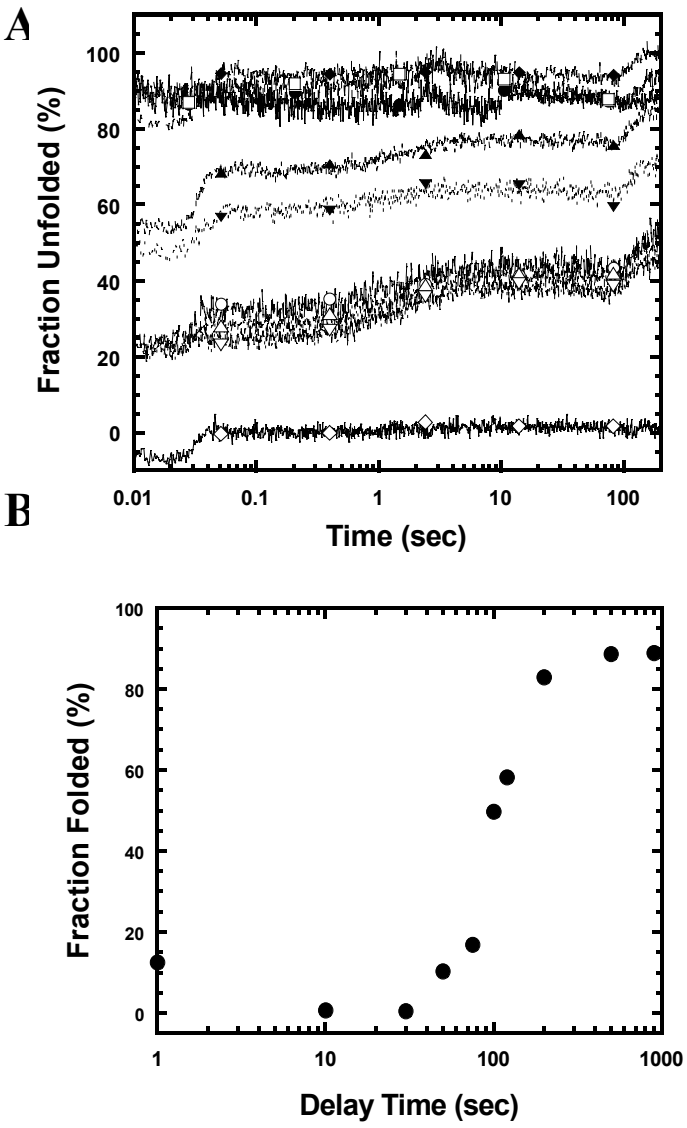


Figure 16. (A) The unfolding signal traces of the second jump in interrupted refolding experiments. The delay time is labeled as follows: 1 (●), 50 (■), 75 (◆), 100 (▲), 120 (▼), 200 (○), 500 (△), and 900 sec (▽). The native control (◇) is indicated. The signal traces were normalized to fraction unfolded and plotted versus time. (B) Fraction of native protein versus delay time in interrupted refolding experiments. The initial signal was normalized and plotted versus delay time.

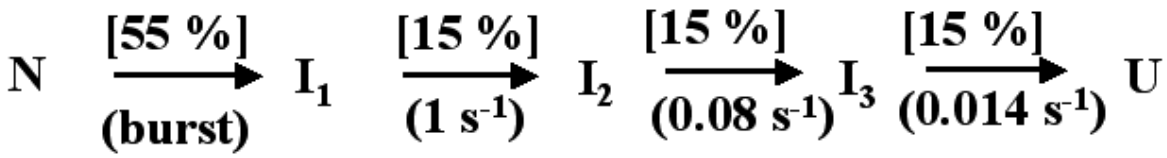
3.8 Discussion of Chapter 3

We have characterized the equilibrium and kinetic folding of a unique protein domain, CARD of RICK, which adopts a α-helical Greek key fold. We have shown that at equilibrium, the folding of RICK-CARD is well described by a two-state mechanism representing the native and unfolded ensembles. The protein is marginally stable, with a ΔG^{H_2O} of 3.0 ± 0.15 kcal/mol and an m-value of 1.27 ± 0.06 kcal/mol/M (30 mM Tris-HCl, pH 8, 1 mM DTT, 25 ºC). While the m-value is constant, the protein stability decreases in the presence of moderate salt concentrations (below 200 mM) and then increases at higher salt concentrations. The results are similar to those of the cold shock protein (Csp) from the thermophilic organism *Bacillus caldolyticus* [78; 84] in which it was shown that electrostatic interactions are stabilizing in the native protein. Above 200 mM salt, the results are consistent with Hofmeister effects [82; 83].

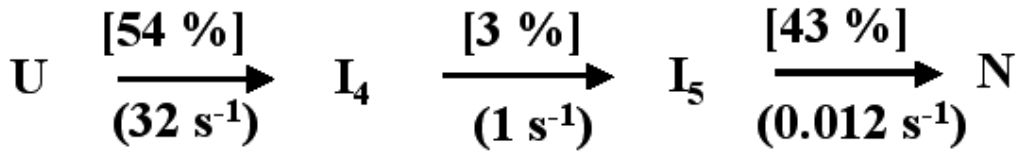
The unfolding pathway of RICK-CARD is complex and contains several non-native conformations. For unfolding, the data are consistent with the sequential unfolding model shown in Scheme 1 (Figure 17). In this model, the native protein unfolds via three intermediate conformations prior to reaching the unfolded state. The first step in unfolding, N→I_1, is rapid and occurs during the dead-time for mixing (~3 msec). The ΔG^{H_2O} and m-value, 2 kcal/mol and 0.93 kcal/mol/M, respectively, for the burst phase reaction account for about 70% of the total ΔG^{H_2O} and m determined from equilibrium unfolding. This suggests that most of the surface area is exposed during this phase of unfolding. Following the burst phase, at least three additional kinetic phases occur, with rate constants of ~1 s^{-1}, 0.08 s^{-1}, and 0.014 s^{-1}, respectively. The intermediate and slow phases are observed in double jump

experiments in which it was shown that the protein must completely unfold (to U in Scheme 1) prior to refolding.

The refolding pathway of RICK-CARD also is complex, and the data are consistent with the sequential folding model shown in Scheme 2 (Figure 17). In this model, the unfolded protein folds via (at least) two intermediate conformations prior to reaching the native state. The first step ($U \rightarrow I_4$) is rapid, with k_{obs} ~30 s^{-1} in the absence of denaturant, and accounts for about 54% of the total signal change. The second step ($I_4 \rightarrow I_5$) occurs with an apparent rate of 1 s^{-1} and accounts for ~3% of the total signal change. According to Scheme 2, this suggests that the fluorescence properties are similar for I_4 and I_5. This phase is apparent only at low urea concentrations (<1.5 M). The slow phase of refolding ($I_5 \rightarrow N$) occurs with an apparent rate of 0.012 s^{-1} and accounts for ~43% of the total signal change. The intermediate and slow phases are observed in the interrupted refolding experiments, where the time course is followed for the appearance of native protein. In that experiment, it was shown that formation of the native conformation occurs only from the slowest phase, consistent with the sequential model shown in Scheme 2.



Scheme 1



Scheme 2

Figure 17. Scheme 1. A proposed pathway for unfolding of RICK-CARD. The percent change in fluorescence emission associated with each reaction is shown above and the folding rates are shown below the reactions. N and U represent native and denatured states, respectively. I_1, I_2, and I_3 represent the non-native sequential intermediates on the pathway. Scheme 2. A proposed pathway for refolding of RICK-CARD. The percent change in fluorescence emission associated with each reaction is shown above and the folding rates are shown below the reaction. I_4, and I_5 represent the non-native sequential intermediates on the pathway.

The data suggest that for unfolding, most of the surface area is exposed in the first step ($N \rightarrow I_1$), where the m-value was determined to be 0.93 kcal/mol/M. Based on the correlation between the m-value and the change in the solvent exposed surface area (ΔASA) described by Scholtz and coworkers [77], this represents about 66 of the 95 residues in RICK-CARD. The remaining surface area appears to be exposed in the second phase ($I_1 \rightarrow I_2$). The data suggest that the intermediates in unfolding are more like the unfolded conformation in that most of the accessible surface area is solvent exposed. For refolding, the data suggest that the fast phase ($U \rightarrow I_4$) is the major folding reaction for the protein. This is reflected in the dependence of k_{obs} on urea, demonstrating that most of the surface area is buried in this phase. The data indicate that the intermediates in refolding are more like the native protein in terms of buried surface area rather than the unfolded protein.

Kinetic complexity in unfolding and refolding reactions has been observed for other small proteins [12; 34; 65; 87]. While the data presented here do not rule out other more complicated off-pathway mechanisms, results from the double jump and interrupted refolding experiments are difficult to reconcile if the intermediates are off-pathway. Overall, the data for RICK-CARD suggest the presence of kinetically trapped, or misfolded, species that are on-pathway both in refolding and in unfolding. At this point, the nature of the species is not clear, but the data suggest that the intermediates are important. Recently, an on-pathway, misfolded, kinetically trapped species has been described for apo-pseudoazurin [34; 36]. Interestingly, that protein also has a Greek key topology, although it is comprised of β-strands rather than the α-Greek key described for RICK-CARD. In the case of apo-pseudoazurin, folding was shown to be dependent on the isomerization of proline residues

from the intermediate conformation. Prolyl isomerization has also been observed in unfolding [65; 88] and may be the source of the slow unfolding reactions described here.

In the case of RICK-CARD, the isomeric states are not well determined for the prolyl residues. The three prolines (P47, in turn-3, and P85 and P87, in helix-6) are predicted to be in the trans configuration based on homology modeling, but the sequence identity is quite low in the CARD family, and this likely affects the reliability of the structural model. Indeed, none of the three prolines is conserved in general (see Figure 4), although P47 is found also in APAF-1. Prolines are, in some cases, observed in turn-4, between helices-4 and-5, and turn-5, between helices-5 and -6 [43]. Although some single prolines are predicted to be presented in helix-6, such as procaspase-4, -5,and -13, RICK-CARD and ARC are the only members of the CARD family with two prolines in helix 6 [43]. This helix is the least defined region of our structural model, but this is true also for other CARD proteins for which structures have been determined [48; 61]. With the exceptions of RAIDD and APAF-1, the amino acid sequences found in helix-6 within the CARD family have low helical propensities. In the case of RICK-CARD, P85 and P87 reside in the middle of the helix. Consequently, the propensity of the amino acid sequence to adopt a helical structure, as predicted by AGADIR [90], is lower than any other member of the CARD family. While both prolines are predicted to be in the trans configuration, the PxP motif distorts and disrupts the helix. In other CARDs [48; 61; 68; 69; 71], the docking of helices-1 and -6 is stabilized by hydrophobic contacts from each helix. In the case of RICK-CARD, disrupting helix-6 by insertion of the PxP motif may prevent tight packing of helix-6 with helix-1. Thus, in addition to the *cis/trans* isomeric states of the prolyl residues, these features may also contribute to the kinetic complexity observed in folding and unfolding.

Other sources of kinetic complexity, such as nonprolyl cis peptide bonds, cannot be ruled out. In this regard, at 25 ºC, the rate of isomerization, $k_{cis \rightarrow trans}$, for nonprolyl bonds is ~2.5 s^{-1} [66]. Although this has been described as an equilibrium process occurring in the unfolded protein, the rate associated with the reaction of $I_4 \rightarrow I_5$ in RICK-CARD is consistent with this reaction. However, if one assumes a value of 0.15% for the fraction of cis isomer of nonprolyl peptide bonds [66], then about 14% of the polypeptide chains would be predicted to have a cis peptide bond in RICK-CARD. The reaction of $I_4 \rightarrow I_5$ accounts for only ~3% of the total signal change. If this phase is due to the isomerization of *cis:trans* nonprolyl peptide bonds, then only a fraction of the peptide bonds are involved in the slow folding reaction.

The presence of intermediates in folding remains controversial [91; 92] as to whether they represent productive, on-pathway species or misfolded, non-productive off-pathway species. The results presented here suggest that the intermediates in the refolding of RICK-CARD are obligatory and on-pathway. A growing body of evidence for complex folding topologies [12; 34] or larger proteins (>100 amino acids) [93; 94] indicates that the intermediates may limit the search to the native state ensemble.

Chapter 4

RESULT II

PRO-1-CARD

Chapter 4

Results: Part II

Pro-1-CARD

4.1 General properties of Pro-1-CARD

The CARD of Procaspase-1 (called Pro-1-CARD), composed of 92 residues, is also a family member in the DD family. There are two tyrosines, a phenylalanine, and no tryptophan in the sequence. The protein contains six α-helices arranged in a Greek key topology [58; 69]. The amino acid sequence identities among Pro-1-CARD, RICK-CARD, and other family members are low with some exceptions such as Pro-1-CARD and ICEBERG [41; 57; 67]. The two tyrosines are located in helices-5 and -6, and the phenylalanine is located in helix-1. There are three cysteinyl residues in helix-4, but there are no disulfide bonds. There is one proline in Pro-1-CARD located at turn-4 connecting helices-3 and -4. The configuration of proline is predicted to be in *trans* based on the homology model. Since the structure of Pro-1-CARD has not been determined, we generated a homology model using the solution structure of ICEBERG [61] as the template because of their high sequence identity. The sequence alignment is shown in Figure 4, and the results of the model are shown in Figure 18, where the model of Pro-1-CARD is compared to the structure of ICEBERG.

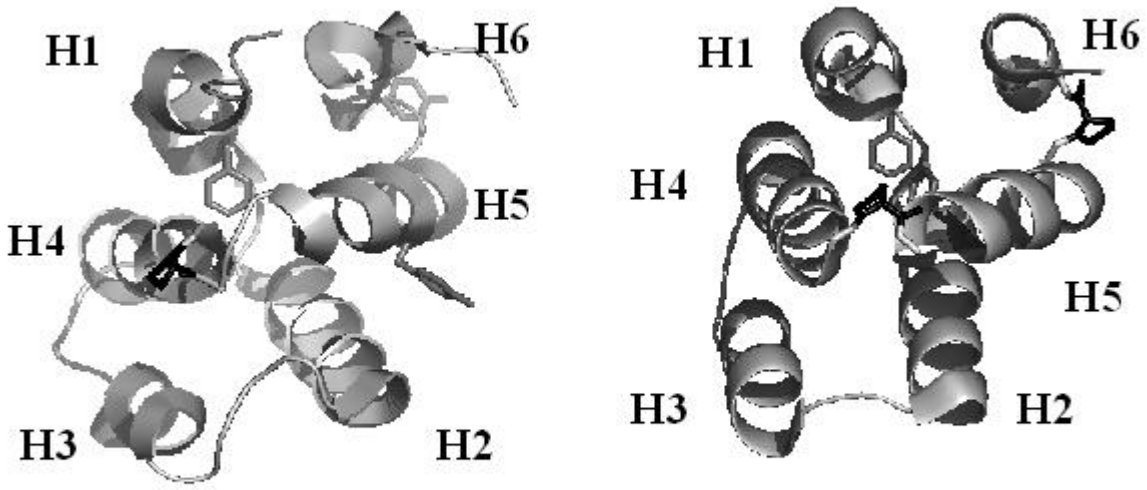


Figure 18. The structure of ICEBERG (right) and the homology model of Pro-1-CARD (left). Structures were drawn with Pymol (DeLano scientific, CA). ICEBERG (PDB code 1DGN) is shown in ribbon diagram with the proline residues in black and the aromatic residues, two phenylalanines, in gray. The homology model of Pro-1-CARD in ribbon diagram was generated by Insight II using ICEBERG **(**PDB code 1DGN**)** as the template. The sequence alignment for modeling is shown in Figure 4. The proline is colored in black. The aromatic residues, two tyrosines in helices-5 and -6, and one phenylalanine in helix-1, are in gray.

4.2 Fluorescence and CD spectroscopy of Pro-1-CARD

Fluorescence emission spectra of Pro-1-CARD, using intrinsic fluorescence probes, are shown in Figure 19. The emission maximum of Pro-1-CARD is at 330 nm in the native condition when the protein is excited at 280 nm. The emission maximum shifted to 340 nm when the protein was unfolded in 6 M urea-containing buffer, demonstrating that the aromatic residues are exposed in the unfolded condition. In general the signal is low because there are no tryptophanyl residues. We also examined the secondary and tertiary structures of Pro-1-CARD using near- and far-UV circular dichroism. The far-UV CD spectrum (Figure 20A) of Pro-1-CARD in native condition (buffer B, 25 °C) shows double minima similar to those of RICK-CARD and the full length Pro-domain of Pro-1-CARD [42; 58]. However, the signal is much lower for Pro-1-CARD when compared to that of RICK-CARD. The molar ellipticity of Pro-1-CARD at 203 nm is approximately 2,000 deg cm^{-1} M^{-1}, whereas, RICK –CARD is 16,000 deg cm^{-1} M^{-1} at 206 nm. This result is not expected for a well-packed six helical bundle. The spectra of the unfolded Pro-1-CARD in high concentrations of urea and guanidine hydrochloride show the protein signal decreases, demonstrating that the secondary structure is lost under these conditions. The near-UV CD spectrum (Figure 20B) shows a minimum at 277 nm. The signal decreases when the protein is unfolded in 6 M urea-containing buffer, but some residual tertiary structure remaining. We have also examined the near UV CD spectrum of Pro-1-CARD in 6 M guanidine hydrochloride (GdnHCl)-containing buffer. The results show that there is no remaining tertiary structure but similar secondary structure. Overall, the data show that the protein is folded with the presence of tertiary and secondary structures, and it is unfolded in 6 M urea-

containing buffer with some residual tertiary structures. The protein is completely unfolded in 6 M GdnHCl.

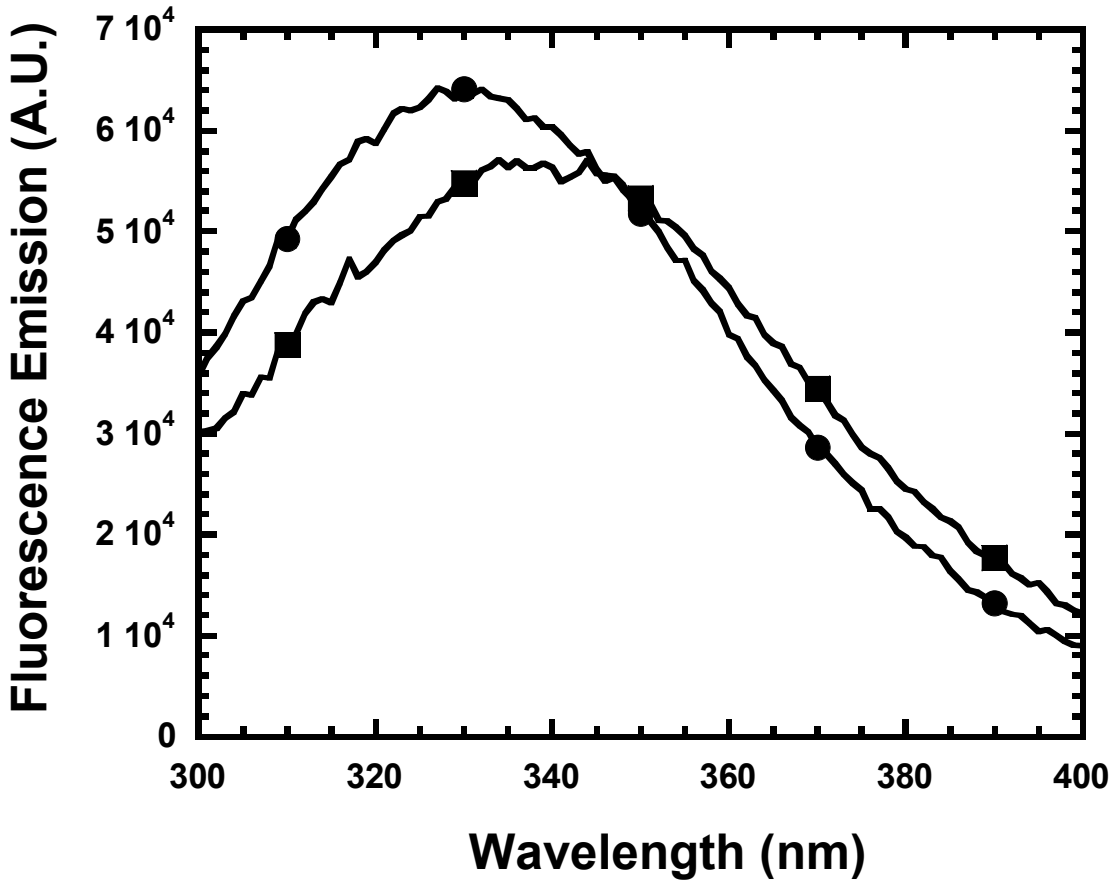


Figure 19. Fluorescence emission spectra of Pro-1-CARD. The protein (3 μM) in buffer (●) or in 5.8 M urea-containing buffer (▲) was excited at 280 nm. The fluorescence emission was collected from 300 to 400 nm.

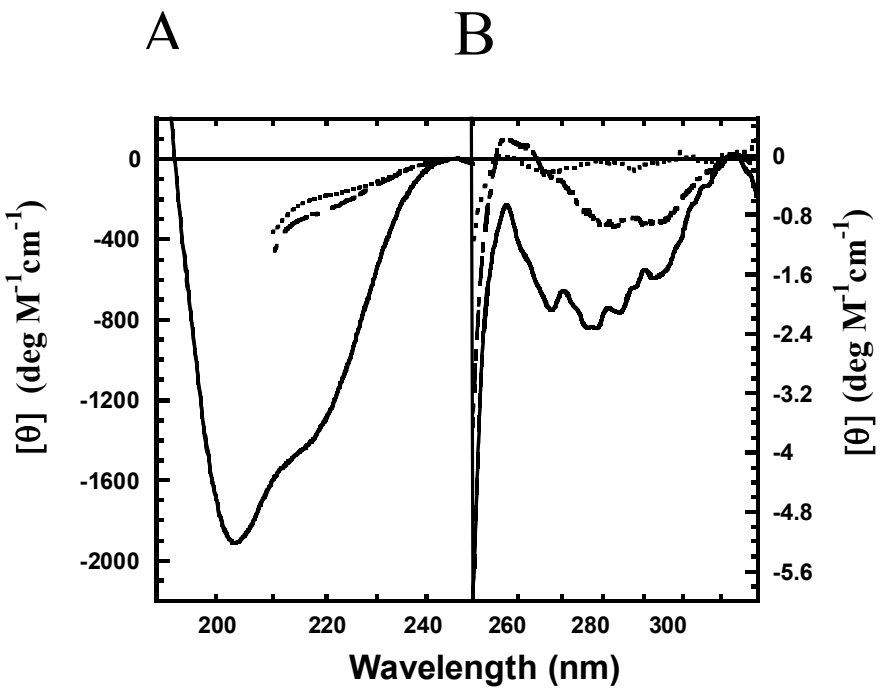


Figure 20. (A) Far-UV and (B) near-UV CD spectra of Pro-1-CARD. The spectra of native protein in buffer (solid line), the denatured protein in 6 M urea-containing buffer (dash line), and the denatured protein in 6 M GdnHCl -containing buffer (dot line) are shown.

4.3 Equilibrium folding of Pro-1-CARD

The equilibrium folding/unfolding of Pro-1-CARD was examined by monitoring the fluorescence emission and CD signals. In addition, the changes in CD at 220 nm representing the changes of the secondary structure were determined. The averaged fluorescence and CD data are shown in Figure 21A. The results show that the data from the two spectroscopic probes overlay, suggesting that the loss of secondary structure is concomitant with the loss of tertiary structure. The experiments were performed with different protein concentrations ranging from 3 to 40 μM. The data also overlay demonstrating the equilibrium folding/unfolding is not protein concentration dependent under the experimental conditions (Figure 21B) (buffer B, 25 °C). The results suggest that Pro-1-CARD is a monomeric species below 40 μM. Since there are three cysteinyl residues in helix-4, we would like to determine whether the cysteinyl residues are involved in disulfide bonds formation. We have examined the folding properties of Pro-1-CARD at equilibrium without the presence of the reducing agent DTT. The equilibrium folding properties of Pro-1-CARD with or without the presence of DTT are shown in Figure 21C. The results show the protein stability does not change in the presence or absence of DTT. The results indicate that there is no disulfide bond formation in the protein.

The results shown in Figure 21A are fitted by a two-state equilibrium folding model, representing only the native and unfolded ensembles of the protein. We determine the conformational free energy, ΔG^{H_2O} and the cooperativity index, m, from fits of the fluorescence data to be 1.1 ± 0.24 kcal/mol and 0.68 ± 0.05 kcal/mol/M and the fits of the CD data to be 1.27 ± 0.12 kcal/mol and 0.65 ± 0.06 kcal/mol/M. The fits are shown in

Figure 21A, and the residuals from both fits are shown in the lower panel of Figure 21A. The m-value correlates to the change in solvent exposed surface area of the protein upon unfolding [75; 76]. Scholtz and coworkers [77] have described a qualitative correlation between the m-value and the change in the solvent exposed surface area (ΔASA). We calculate the ΔASA of Pro-1-CARD to be around 3000 $Å^2$, corresponding to 42 residues using the methods described in section 3.3, Chapter 3 ($m = 374 + 0.11(\Delta ASA)$; $\Delta ASA = -907 + 93(\#residue)$). The number of amino acid calculated does not agree with the number of amino acids of Pro-1-CARD, 92.

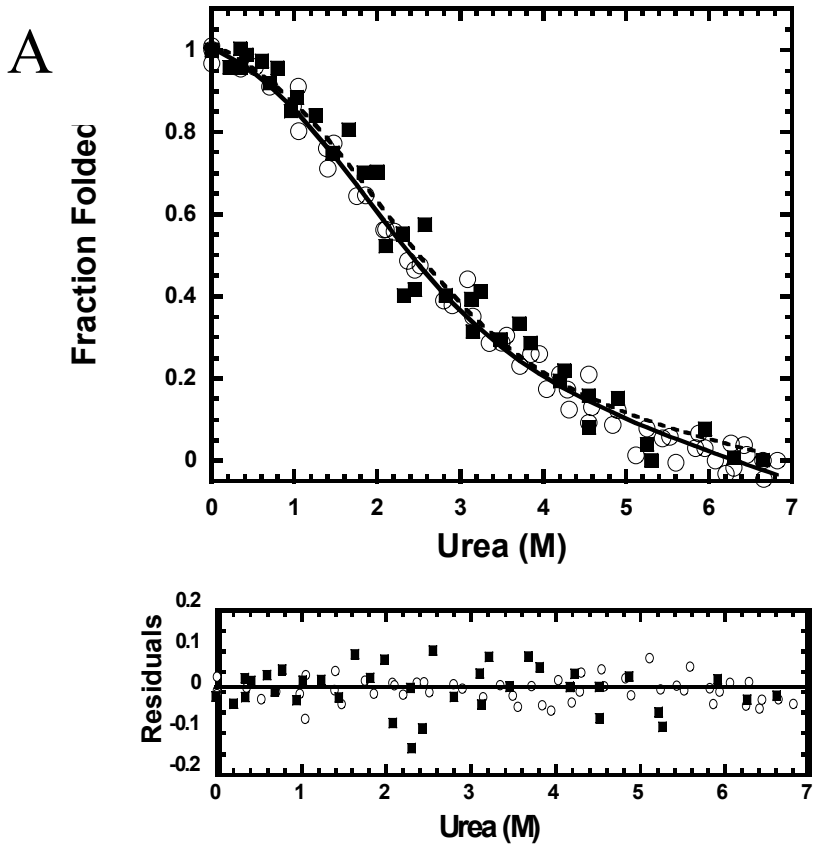


Figure 21. Equilibrium folding of Pro-1-CARD. The average emission wavelength was determined for fluorescence emission spectra as described in Methods. The normalized and averaged CD data and average emission wavelength versus urea concentration are plotted. (A) The fluorescence emission data from excitation at 280 nm (○) of Pro-1-CARD (6 μM) as well as the CD data at 220 nm (■) of Pro-1-CARD (6 μM) are shown. The solid line and dashed line represent fits of the fluorescence and the CD data, respectively, to a two-state equilibrium folding mechanism. The ΔG^{H_2O} and m-value obtained for fluorescence at 280 nm are 1.1 kcal/mol and 0.68 kcal/mol/M, respectively. The ΔG^{H_2O} and m-value for CD are 1.27 kcal/mol and 0.65 kcal/mol/M, respectively, at 25°C. The residuals of the fits of fluorescence (○) and CD (■) are shown in the lower panel.

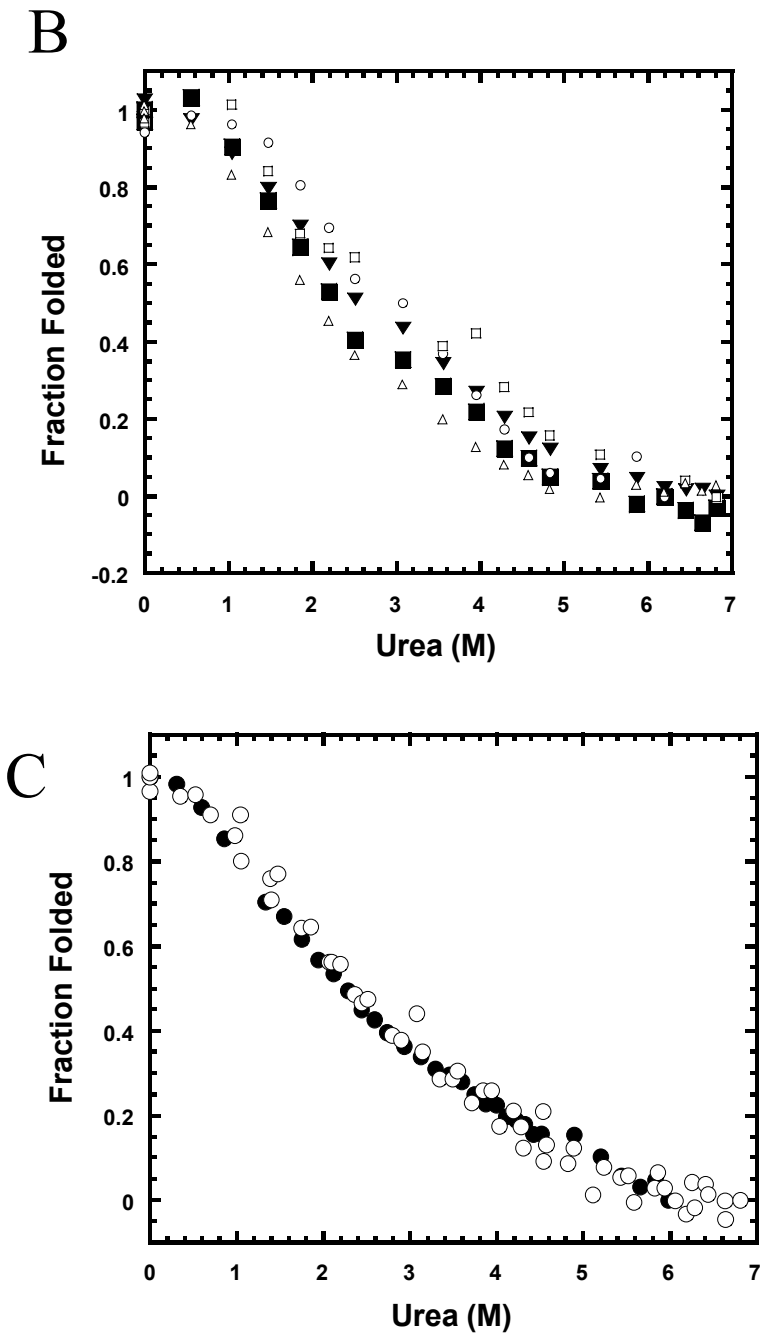


Figure 21. (B) Equilibrium folding of Pro-1-CARD by fluorescence emission as described in panel A. The concentrations of Pro-1-CARD are as follows: 3 (□), 6 (■), 9 (△), 10.5 (○), and 40 (▼) μM. (C) Equilibrium folding of Pro-1-CARD at 6 μM with (○) or without (●) the presence of DTT.

To further examine the oligomeric state of Pro-1-CARD, analytical size-exclusion chromatography was employed. Approximately 88 % of Pro-1-CARD at 85 μM elutes in the range of 10 kD based on the calculation of the standard curve (Figure 22). The protein elutes between the elution of cytochrome c, 12.4 kD, and aprotinin, 6.5 kD. There is a small population of protein (12%) that elutes in the range of a tetrameric species, however, it is insignificant compared to the monomeric species. Therefore, we suggest Pro-1-CARD is a monomer under the described experimental conditions (buffer C, pH 8, 4°C, 85 μM).

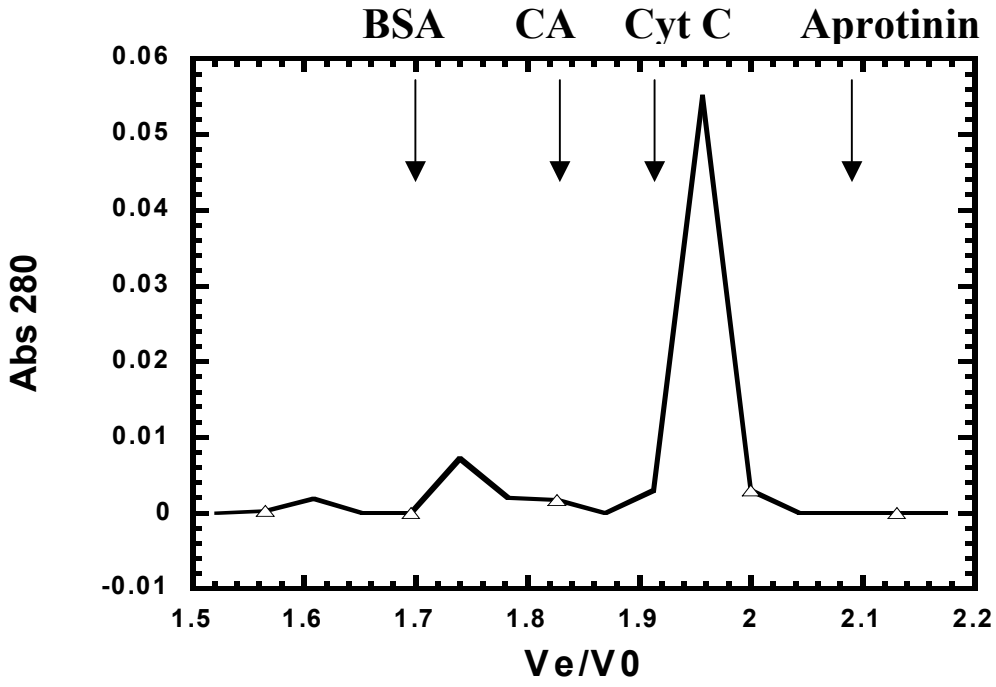


Figure 22. Analytical size exclusion chromatography of Pro-1-CARD. The elution volumes of the standards are indicated and of Pro-1-CARD (85 μM) (△) are plotted against the fraction numbers. The standards include bovine serum albumin (BSA, 66.2 kD), carbonic anhydrase (CA, 29 kD), cytochrome c (Cyt C, 12.4 kD), and aprotinin (6.5 kD).

4.4 Interaction of Pro-1-CARD and RICK-CARD

To further examine whether the structure of Pro-1-CARD is native and functional, we have examined the function of Pro-1-CARD through assay of protein-protein interactions in native condition. In these assays, fluorescence titration studies were performed to examine the interaction of Pro-1-CARD and RICK-CARD. The fluorescence emission of the tryptophanyl residue in RICK-CARD is monitored upon titration with Pro-1-CARD. The fluorescence emission is quenched upon binding of Pro-1-CARD (Figure 23A). We fit the fluorescence signal as a function of titrant (Pro-1-CARD) concentration with equation [15] as described in Methods.

The result of fluorescence titration correlates well with the assumption of simple protein and ligand binding in which the two proteins interact with a stoichiometry of 1:1 (Figure 23B). The dissociation constant obtained from the fit is 1.4 μM. This is in a good agreement with that obtained for the interaction of Pelle and Tube, two DD proteins, of 0.5 μM [95]. The results show that Pro-1-CARD in native condition is functional and suggest that the structure of Pro-1-CARD required for interactions with RICK-CARD are in their native conformation.

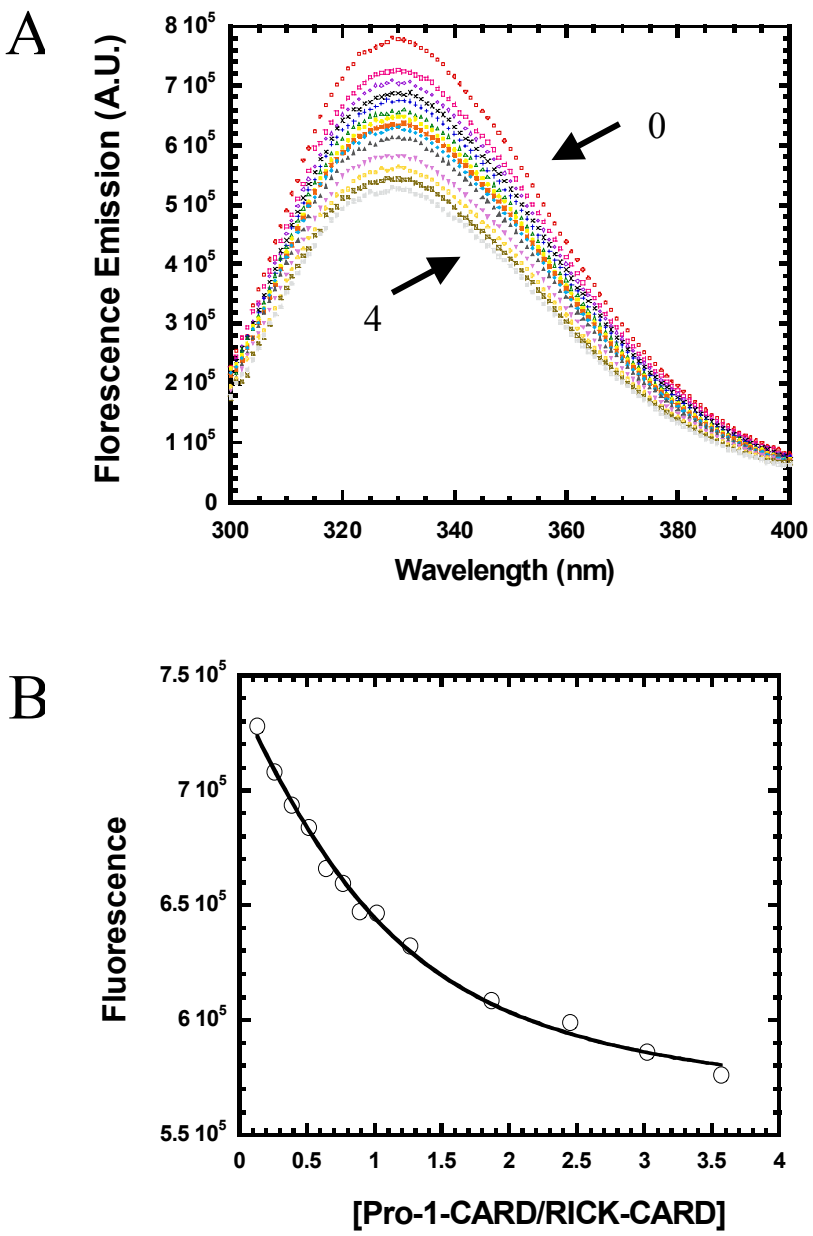


Figure 23. (A) Fluorescence emission spectra of RICK-CARD in the presence of Pro-1-CARD. The spectrum of RICK-CARD, 3 μM, is quenched in the presence of Pro-1-CARD from 0 to 12 μM. The ratio of Pro-1-CARD to RICK-CARD is indicated. (B) The fluorescence titration for binding of Pro-1-CARD and RICK-CARD. The fluorescence signals (○) at 330 nm are plotted against the ratio of Pro-1-CARD and RICK-CARD. The solid line is fit to the data describing simple protein-ligand binding.

4.5 Folding Kinetics of Pro-1-CARD

The refolding and unfolding kinetics of Pro-1-CARD were examined in both single-mixing and sequential-mixing stopped-flow experiments by monitoring changes in fluorescence emission. The single-mixing data are shown in Figure 24. The refolding traces overlay with the control of the native protein signal, and the unfolding traces overlay with the control of unfolded protein signal in the observable time. Both refolding and unfolding of Pro-1-CARD show no signal change from 10 ms to 500 s. Therefore, the results suggest that within the dead-time of mixing (~10 ms), the reactions are complete. There is no detectable kinetic phase in both unfolding and refolding experiments.

Furthermore, we monitored the refolding and unfolding traces (the refolding traces are shown in Figure 24A) in the final urea concentration from 0 to 6 M. The traces are linear without observable exponential phases. The initial signals (10 ms) and final signals (10 s) along with the equilibrium folding data are plotted in Figure 24B. The results show that the initial and final signals of unfolding and refolding overlay with the equilibrium folding data. The results are very different from the kinetics of RICK-CARD (Figure 11A, described in Chapter 3), where it contains a burst phase intermediate in unfolding and a fast refolding phase of 30 s^{-1} [58]. Moreover, the kinetics of Pro-1-CARD do not contain observable slow kinetic phases such as the ones observed for RICK-CARD. In that case we observed the slow phases with half-times of 46 s and 86 s for unfolding and refolding, respectively.

A

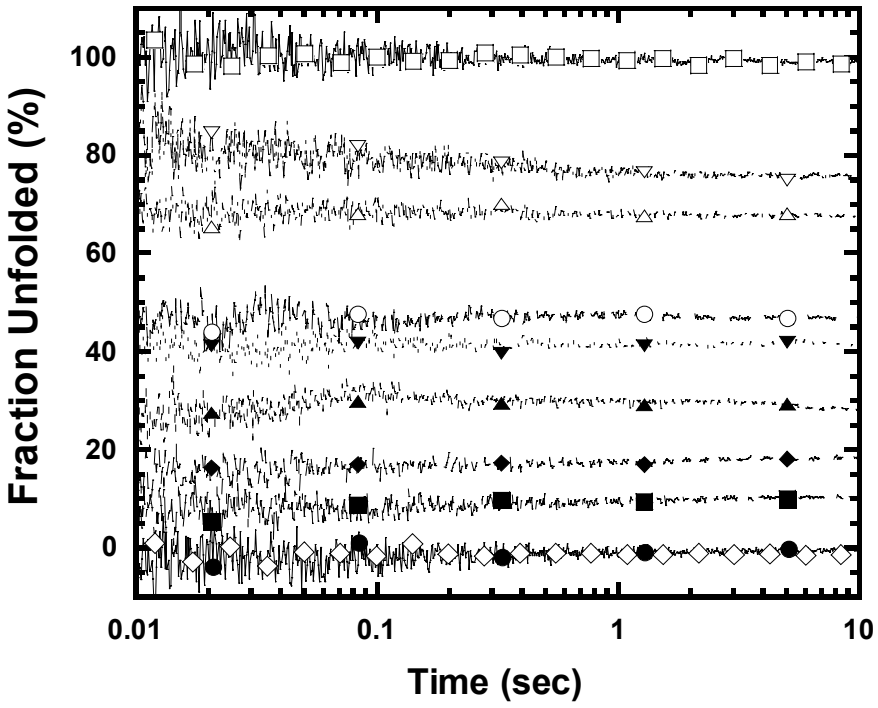


Figure 24. (A) The refolding traces of Pro-1-CARD. Each trace represents a refolding event from Pro-1-CARD in 6 M urea containing buffer to a buffer containing a desired final urea concentration. The final urea concentration (M) are as follows: 0 (●), 1.04 (■), 1.69 (◆), 2.02 (▲), 2.51 (▼), 3.73 (○), 4.36 (△), 4.9 (▽), and, 6 (□). The native control is shown in open diamond (◇).

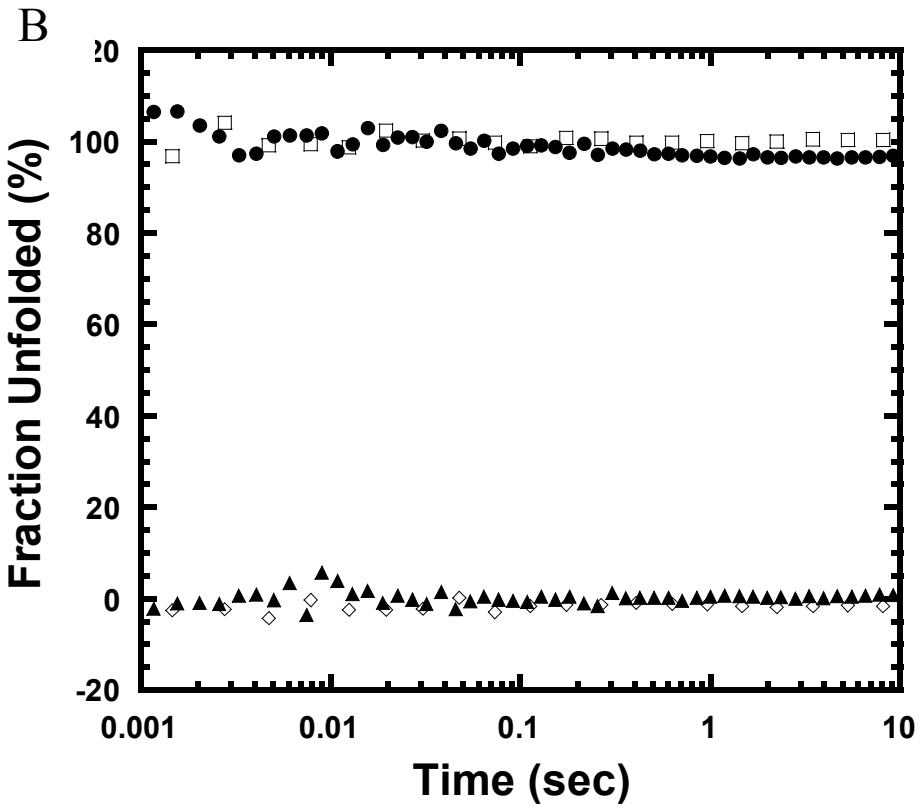


Figure 24. (B) The unfolding and refolding kinetics of Pro-1-CARD. The signals are normalized to fraction unfolded. The native (◇) and unfolded protein signals (□) are indicated. Both unfolding (●) and refolding traces (▲) are shown.

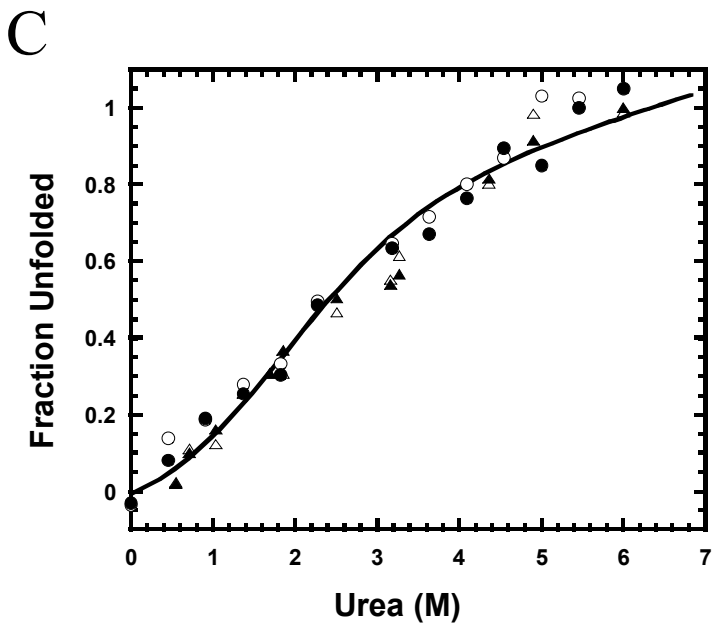


Figure 24. (C) The signals of unfolding and refolding versus urea concentrations. The initial (○) at 10 ms and final signals (●) at 10s of unfolding are shown. The initial (△) at 10 ms and final signals (▲) at 10s of refolding are shown as well. All the initial and final signals are superimposed to data from the equilibrium folding experiment (Figure 21). The solid line is fit of a two-state folding mechanism to the equilibrium data.

4.6 Double jump of Pro-1-CARD

Double jump experiments were performed to investigate further details of the unfolding pathway. We employed the double jump experiments as described in the Methods. Briefly, we mix a native Pro-1-CARD to a denatured condition (from 0 to 4.8 M urea) for various amounts of time (delay time), then further mix it to the native buffer (Figure 19). The final urea concentration is 0.8 M. Surprisingly, the results show a transition between 1 and 30 sec with a half-time of 10 sec. The data show that with a short delay time less than 1 sec, there is no formation of native protein. As the delay time is increased, the native protein is populated. There is 100 % of native protein populated when the protein unfolds for 30 sec or more in the first jump. A similar result is observed for the double jump experiments of RICK-CARD (Figure 15A, Chapter 3), although the half-time is about 50 sec. We fit the data to a single exponential equation and the observed rate constant is 0.07 sec^{-1}. This half-time of Pro-1-CARD at 10 sec is similar to the lag phase of RICK-CARD observed in double jump experiments (Figure 15A) and to the intermediate phases in unfolding of RICK-CARD. The results may suggest that the same intermediate forms in unfolding of RICK-CARD and Pro-1-CARD.

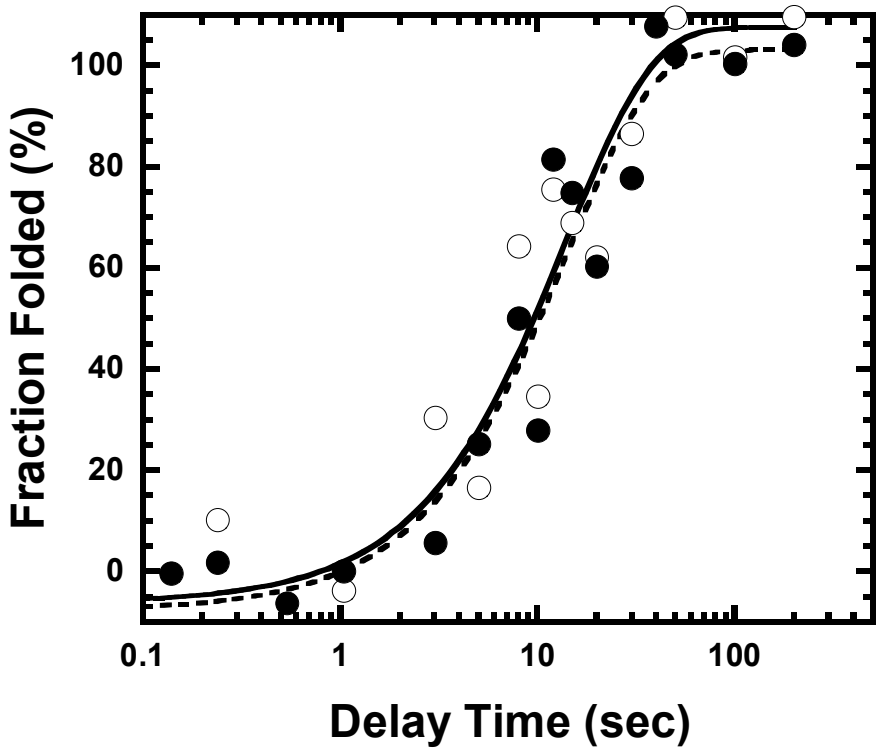


Figure 25. Fractions of native protein versus delay time in double jump experiments. The initial (○) and final signals (●) of the signal traces were normalized to fraction folded and plotted versus delay time.

4.7 Interrupted refolding of Pro-1-CARD

The refolding details of Pro-1-CARD are further examined by the interrupted refolding experiments. The set-up of the instrument of Pro-1-CARD is different than that for RICK-CARD. The details are described in Methods. Briefly, the denatured Pro-1-CARD in 5 M urea-containing buffer is mixed with buffer for various amounts of refolding time. The urea concentration is diluted 0.45 M. Then, the protein is unfolded by mixing with the urea-containing buffer. The final urea concentration is about 5 M.

The signal traces are linear as observed in the single mixing experiments. The final signals at 10 sec are normalized and plotted against the delay time. The results show that the protein is trapped in an intermediate state, which cannot unfold until the delay time of refolding is more than 50 sec. The half-time of the transition is about 63 sec. This half time observed in interrupted refolding is different than that in RICK-CARD, ~100 sec (Figure 15B, Chapter 3). The results suggest that an intermediate is also present on the refolding pathway of Pro-1-CARD.

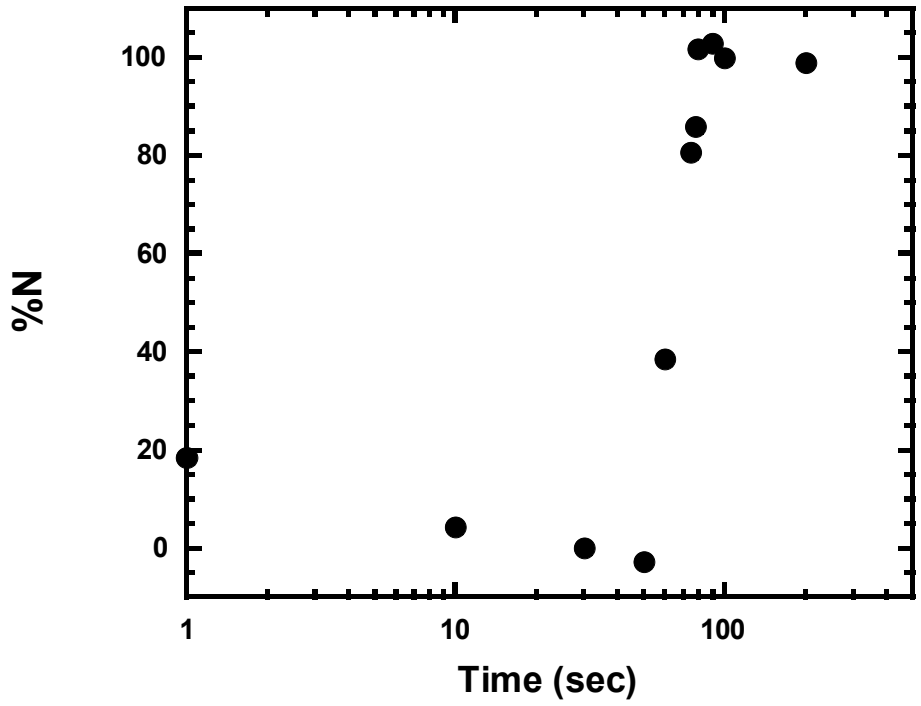


Figure 26. Fraction of native protein versus delay time in interrupted refolding experiments. The final signals (●) of the signal traces were normalized to fraction folded and plotted versus delay time.

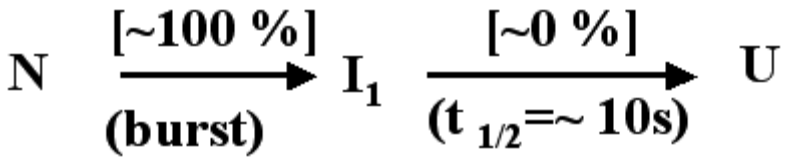
4.8 Discussion of Chapter 4

We have characterized the equilibrium and kinetic folding properties of Pro-1-CARD, a family member of α-helical Greek key proteins. We have shown the folding of Pro-1-CARD is described by a two-state mechanism representing the native and unfolded ensembles. The protein is marginally stable, with a ΔG^{H_2O} of ~1.1 ± 0.24 kcal/mol and an m-value of 0.65 ± 0.06 kcal/mol/M (10 mM Tris-HCl, pH 8, 1 mM DTT, 25 ºC). Compared to the homologous protein, RICK-CARD, the equilibrium folding of these two small protein domains are both described by two-state folding mechanisms, however, Pro-1-CARD is less stable than RICK-CARD. The native Pro-1-CARD is fully functional and has its native structure for binding of its binding partner, RICK-CARD, although the m-value is lower than the prediction. The amplitudes of the far-UV and near-UV CD spectra are quite low. The low m-values at equilibrium folding and low signals in CD spectroscopy could be due to the following reasons. First, Pro-1-CARD is partially folded in the native condition. However, the interaction studies show that it contains the native structure for the interaction. Alternatively, there might be a low energy barrier between the native and unfolded Pro-1-CARD causing fast exchange between the two states. Therefore, the unfolded protein may have some population in the native condition.

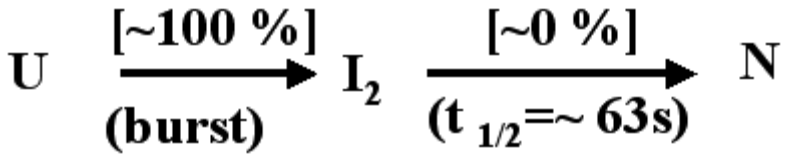
The results of the kinetic experiments show that the refolding and unfolding kinetics of Pro-1-CARD are faster than those of RICK-CARD. Since there is no observable slow kinetic phase, proline isomerization does not seem to occur in the folding of Pro-1-CARD. From the single mixing experiments, the unfolding seems to be complete within the burst phase. However, from the sequential mixing experiments, the results show that there is

an unfolding intermediate present, and the unfolding event is complete in about 30 sec. Therefore, the results indicate that there is an on pathway intermediate present as kinetic traps in the unfolding pathways, and the fluorescence signal of the intermediate is indistinguishable from the unfolded protein signal. Furthermore, we found that the half-time of the transition phase observed in double jump experiments of Pro-1-CARD (~10 sec) is close to the lag phase observed in the double jump experiments of RICK-CARD (~15 sec) (Figure 17, Chapter 3). It might suggest that the intermediates (I_1 of Pro-1-CARD and I_2 of RICK-CARD) are the same. Alternatively, from the results of single mixing experiments, the data suggest that the refolding of Pro-1-CARD is complete in the burst phase. However, we still observe an intermediate in the interrupted refolding experiments. Based on the results, the refolding of Pro-1-CARD is complete in ~80 sec. Therefore, the kinetics of Pro-1-CARD are different than those of RICK-CARD.

Here, we suggest a minimal kinetic model for Pro-1-CARD (Figure 27). Major folding or refolding phases are complete within the burst phase. Overall, the kinetics of Pro-1-CARD contain fast unfolding and refolding with the presence of misfolded, or kinetically trapped, intermediates.



Scheme 1



Scheme 2

Figure 27. Proposed unfolding (scheme 1) and refolding (scheme 2) pathways of Pro-1-CARD. The percent change in fluorescence emission associated with each reaction is shown above, and the half-time of the folding phases are shown below the reactions. N and U represent native and denatured states, respectively. I_1 and I_2 represent the non-native sequential intermediates on the pathways.

Chapter 5

RESULT III

PROLINE MUTANTS OF RICK-CARD

Chapter 5

Results: Part III

Proline Mutants of RICK-CARD

Foreword

We first studied the folding properties of RICK-CARD and found this protein has complex folding kinetics rather than simple two-state folding kinetics observed for many small helical proteins. We know that RICK-CARD has slow kinetic phases with the half-time of 50 and 86 s for unfolding and refolding, respectively. The slow folding phase may be due to proline-isomerization of the protein since the half-time of proline isomerization has been reported to be 10 to 100 s. The three prolyl residues in RICK-CARD are predicted to be *in trans* configuration in the homology model. However, since the homology model is based on a template that has low sequence identity with RICK-CARD, we cannot rule out the possibility of the presence of *cis*-proline. To examine further details of folding of RICK-CARD, we have generated seven proline mutants of RICK-CARD and have examined their properties. The seven mutants are all proline to alanine mutants including single mutants (P47A, P85A, and P87A), double mutants (P47A/P85A, P47A/P87A, and P85A/P87A), and a triple mutant (P47A/P85A/P87A; this mutant will be referred to P_3A in the text). The

figures in this chapter will have the following color scheme: black, wild type RICK-CARD; red, P47A; orange, P85A; yellow, P87A; green, P47A/P85A; blue, P47A/P87A; gray, P85A/P87A; purple, P3A.

5.1 Temperature effects on the refolding rate of RICK-CARD

A way to examine the proline isomerization is to calculate the activation energy of the slow folding steps. Since proline isomerization requires high activation energy due to the rotation of the peptide bonds from *trans* to *cis*, we can calculate the activation energy of the reaction based on Arrhenius equation as described in Methods.

We have performed the stopped-flow refolding experiments of RICK-CARD in different temperature ranging from 10 to 30 ºC. The refolding reactions were fit to a three exponential equation and the rate constants were obtained. The Arrhenius plot describing the slowest refolding rate is shown in Figure 28. The slope of the linear fit is equal to $-E_a/R$ based on Arrhenius equation (see Methods). Therefore, we obtained the activation energy of the slow refolding step to be 17 kcal/mol, which is similar to the 20 kcal/mol for proline isomerization described in the literature [96]. This is suggesting that the slow phase is due to proline isomerization.

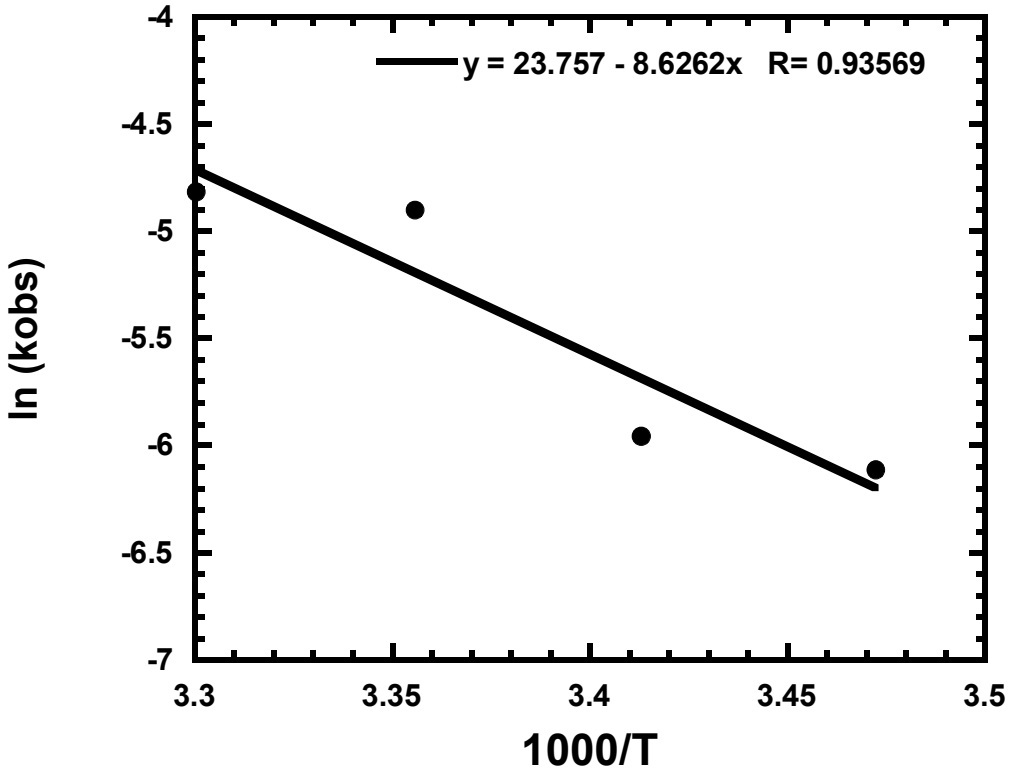


Figure 28. Arrhenius Plot of refolding rate of RICK-CARD versus temperature. The natural logarithm of the slowest rate of refolding of RICK-CARD is plotted against the reciprocal value of temperature. The solid line is a fit to the Arrhenius equation, E_a=-RTln(k_{obs}).

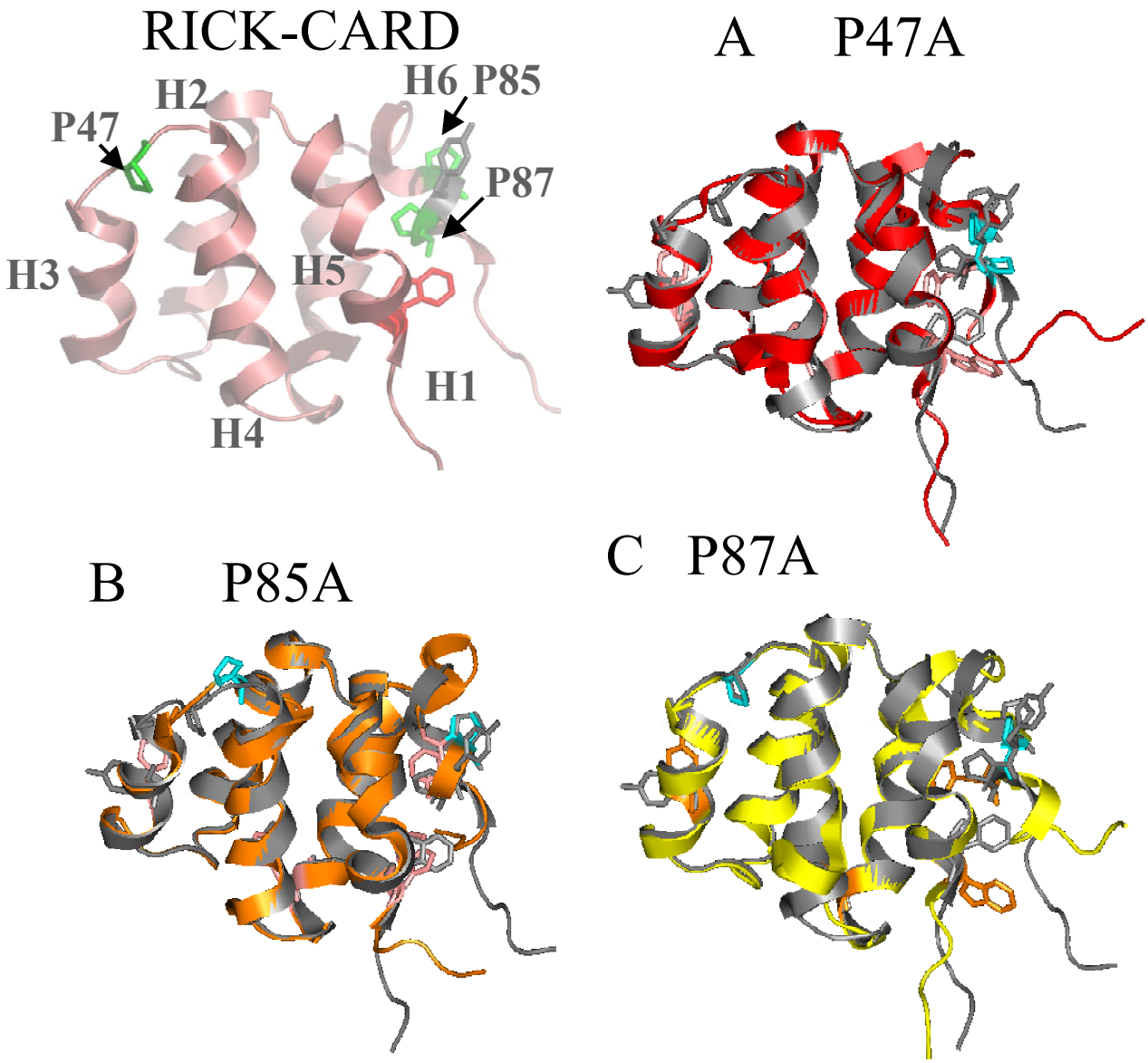
5.2 General properties of RICK-CARD proline mutants

The general properties of the RICK-CARD mutants are similar to the wild-type RICK-CARD (see chapter 3). They are proteins with molecular weight of 10 kD and are composed of 95 residues. The mutation we made are proline to alanine mutants in the following positions: Pro 47 in turn-3, Pro 85 in helix-6, and Pro 87 in helix-6 as shown in Figure 29 and 7. Therefore, three single mutants, three double mutants, and a triple mutant are generated.

The homology models are shown in Figure 29. They are generated as described in Methods. Since the homology models of the mutants are based on the homology model of RICK-CARD, the details of the structural arrangement may not be reliable. The models are generated for convenience of viewing the structures. A general observation of the models is described in next paragraph.

The root mean square deviations of the alignments of the peptide backbone between each mutant and the wild type RICK-CARD are as follows: P47A, 1.122 Å; P85A, 1.063 Å; P87A, 0.984 Å; P47A/P85A, 0.951 Å; P47A/P87A, 0.994 Å; P85A/P87A, 0.927 Å; P_3A, 0.947 Å. The helix-6 of the wild type RICK-CARD contains a full helical turn followed with a bending backbone. However, the helices-6 of P87A, P_3A and P85A/P87A appear to have two full helical turns. The helix-6 of P47A/P87A is similar to that of the wild type but seems to have 1.5 helical turns. The helices-6 of P47A and P47A/P85A are similar to that of the wild type. The helix-6 of P85A is twisting rather than forming helical turns. The tryptophanyl residues in helix-1 of the mutants are in different position than the wild type beside P85A and P85A/P87A. The tyrosines in helix-6 of the mutants are all flipped toward

the hydrophobic core rather than pointing toward the surface of the protein as observed in the wild type.



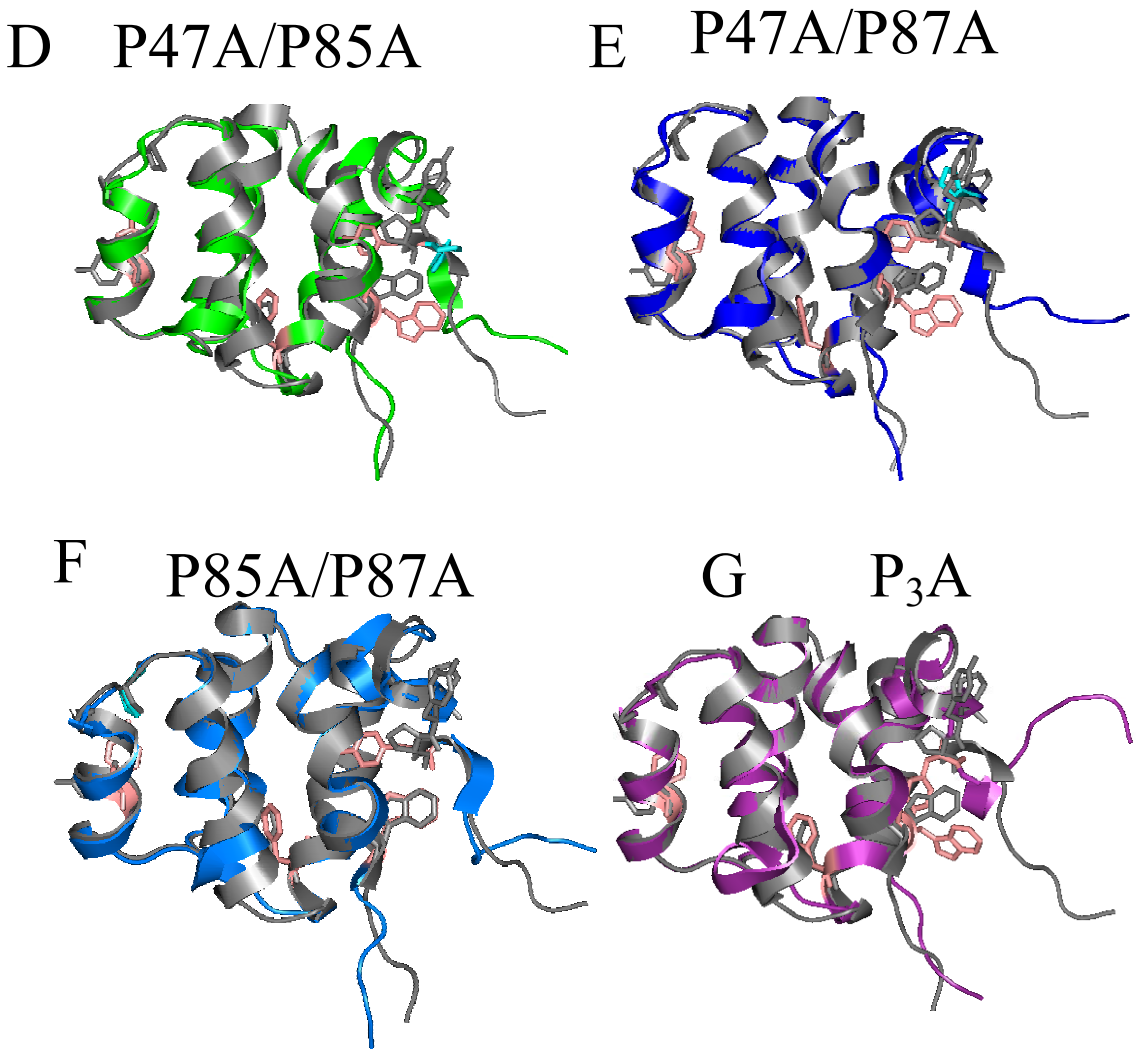


Figure 29. Homology models of RICK-CARD and its proline mutants. The model of the wild type RICK-CARD (the same model as Figure 7) is shown in the upper left panel in the top and the helices and prolyl residues are labeled. Panel (A) to (G) are the homology models of the proline mutants as indicated. The proline residues of the mutants are colored in cyan and the aromatic residues are colored in pink. The superimposed wild type RICK-CARD is colored in gray.

5.3 Fluorescence spectra of the proline mutants

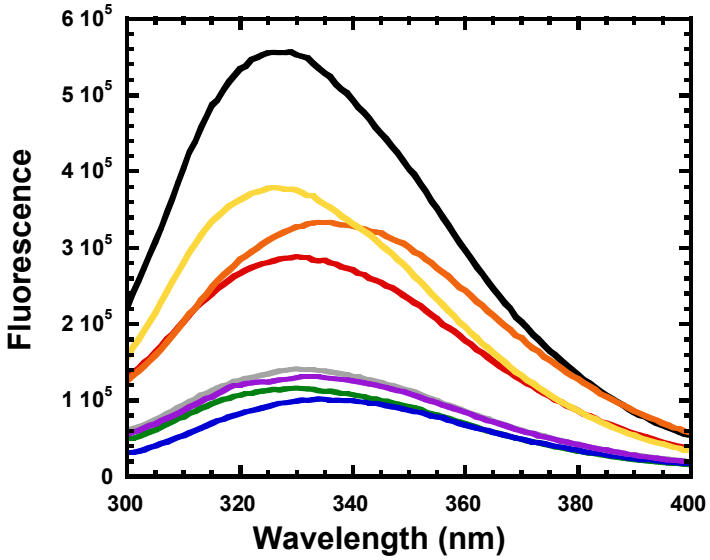
Fluorescence spectra are monitored as described in Methods. The fluorescence emission spectra of the proteins, 3 μM, in the native condition (buffer B, pH 8, 25 ºC) are shown in Figure 30A (excitation at 280 nm) and Figure 30B (excitation at 295 nm), whereas, the spectra of the proteins, 3 μM, in the denatured conditions (4 or 6 M urea containing buffer B, pH 8, 25 ºC) are shown in Figure 30C (excitation at 280 nm) and Figure 30D (excitation at 295 nm). A summary of the emission maxima in different conditions is listed in Table 2.

By comparing the protein spectra in native condition, most of the proteins have the same emission maximum as the wild type, ~ 328 nm for excitation at 280 nm, ~330 nm for excitation at 295 nm, although the mutants have lower amplitude. However, the emission maxima of P85A shows a red shift of 5 nm indicating the aromatic residues of P85A in the native condition are more exposed. The same situation is observed for P47A/P87A but with a smaller amount of red shift. In contrast, we observed slightly blue-shifted spectra for P87A. This indicates the aromatic residues of P87A are in a more hydrophobic environment than the wild type. In the denatured conditions, the emission maxima shift to ~ 345 nm for most of the proteins, but P47A/P87A, P85A/P87A, and P_3A have emission maxima at ~ 340 nm.

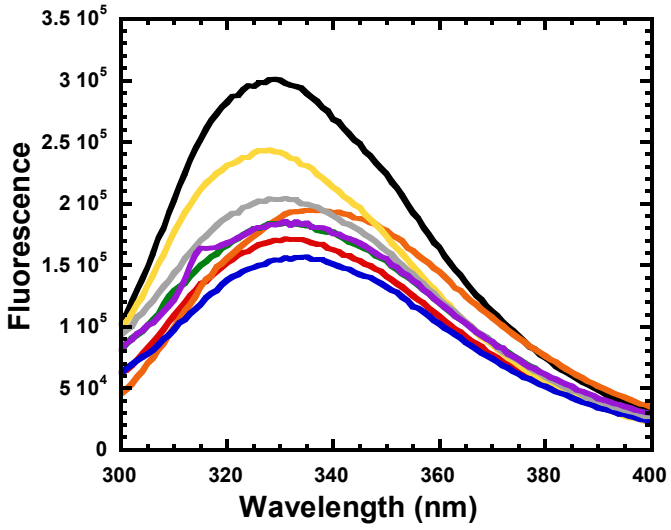
Table 2. The fluorescence emission maxima of the RICK-CARD, the proline mutants, and Pro-1-CARD

(Buffer A or B, pH 8, 25 ºC)	Native condition (in 0 M urea)		Denatured condition (in 4* or 6 M urea)	
	Excitation at 280 nm	Excitation at 295 nm	Excitation at 280 nm	Excitation at 295 nm
RICK-CARD*	328 nm	330 nm	345 nm	348 nm
P47A	330 nm	330 nm	345 nm	345 nm
P85A	336 nm	338 nm	345 nm	348 nm
P87A*	325 nm	325 nm	345 nm	345 nm
P47A/P85A	330 nm	330 nm	350 nm	353 nm
P47A/P87A	336 nm	335 nm	340 nm	340 nm
P85A/P87A	330 nm	330 nm	340 nm	345 nm
P_3A	330 nm	330 nm	340 nm	340 nm
Pro-1-CARD	330 nm	x	340 nm	x

A



B



RICK-CARD
P47A
P85A
P87A
P47A/P85A
P47A/P87A
P85A/P87A
P_3A

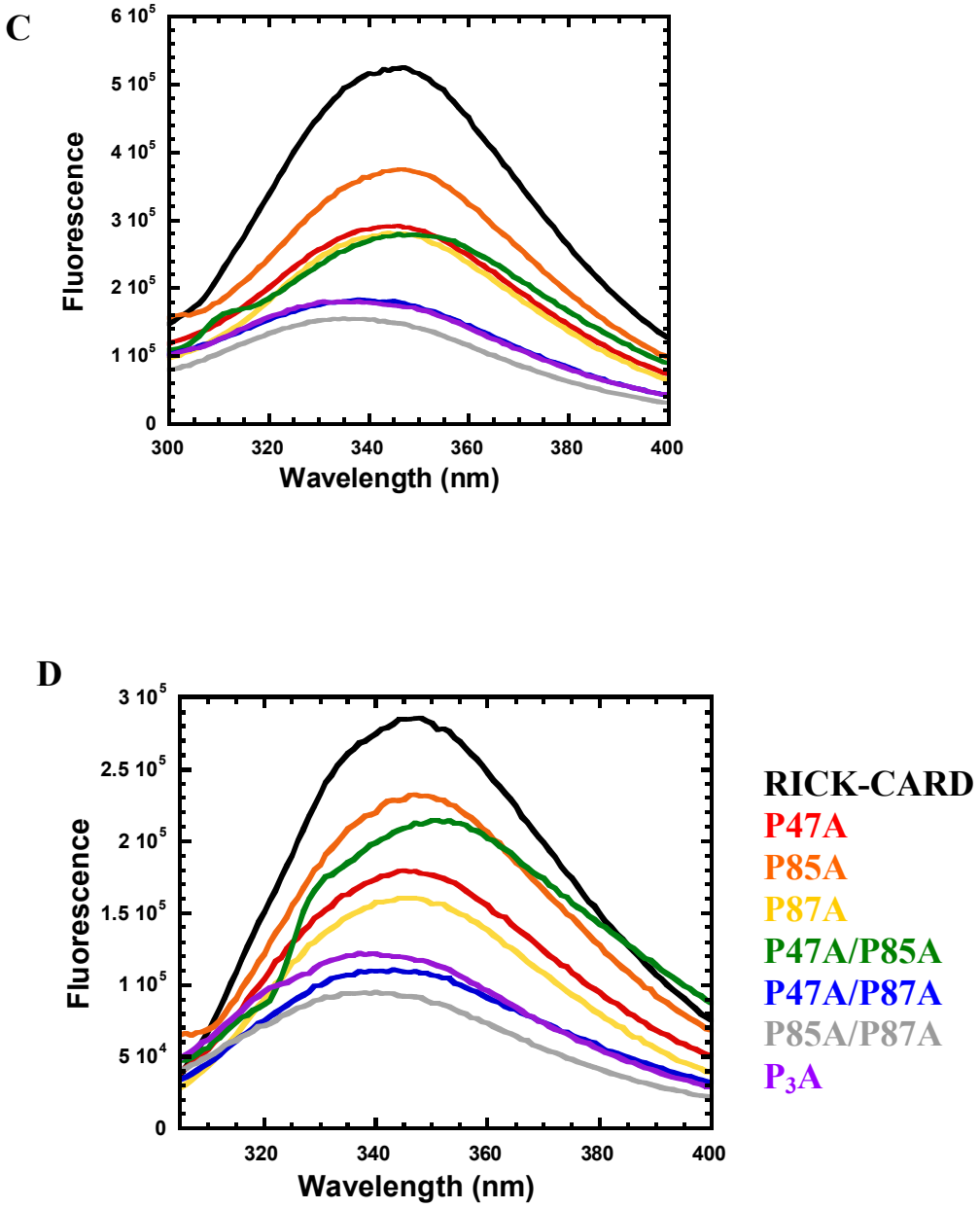


Figure 30. Fluorescence emission spectra of the proline mutants of RICK-CARD. The proteins (3 μM) in buffer B were excited at (A) 280 nm or (B) 295 nm. The proteins, 3 μM, in urea-containing buffer B were excited at (C) 280 nm or (D) 295 nm. All proteins were in 6 M urea-containing buffer except the wild type and P87A are in 4 M urea-containing buffer. The color scheme for each protein is indicated.

5.4 CD spectra of the proline mutants

The near-UV and far-UV CD spectra of the proline mutants are shown in Figure 31. All mutants have smaller amplitude than the wild type. In the near UV spectra (Figure 31B), we found that P87A has a higher signal than the other mutants. The order for the ones containing the most structure contents to the least is P87A > P47A/P85A > P47A > P_3A. There is little tertiary structure in P85A/P87A and no tertiary structure in P85A. Similar results are obtained in far-UV CD (Figure 31A). The order from the most structure content to the least is P87A > P47A/P85A ≈ P47A > P85A/P87A > P_3A, where the spectra of P47A/P85A and P47A are almost identical. The minima of the spectra of P47A/P87A and P85A are shifted from 208 nm to ~205 nm, especially P85A, indicating there are more random structures in the protein. Overall, the CD spectra show the mutation at position 85 disrupts the native protein structure dramatically, however, by compensating with another mutation at position 47, the structure is retained or recovered somewhat. In addition, P87A seems to remain a well-packed structure like the wild type protein.

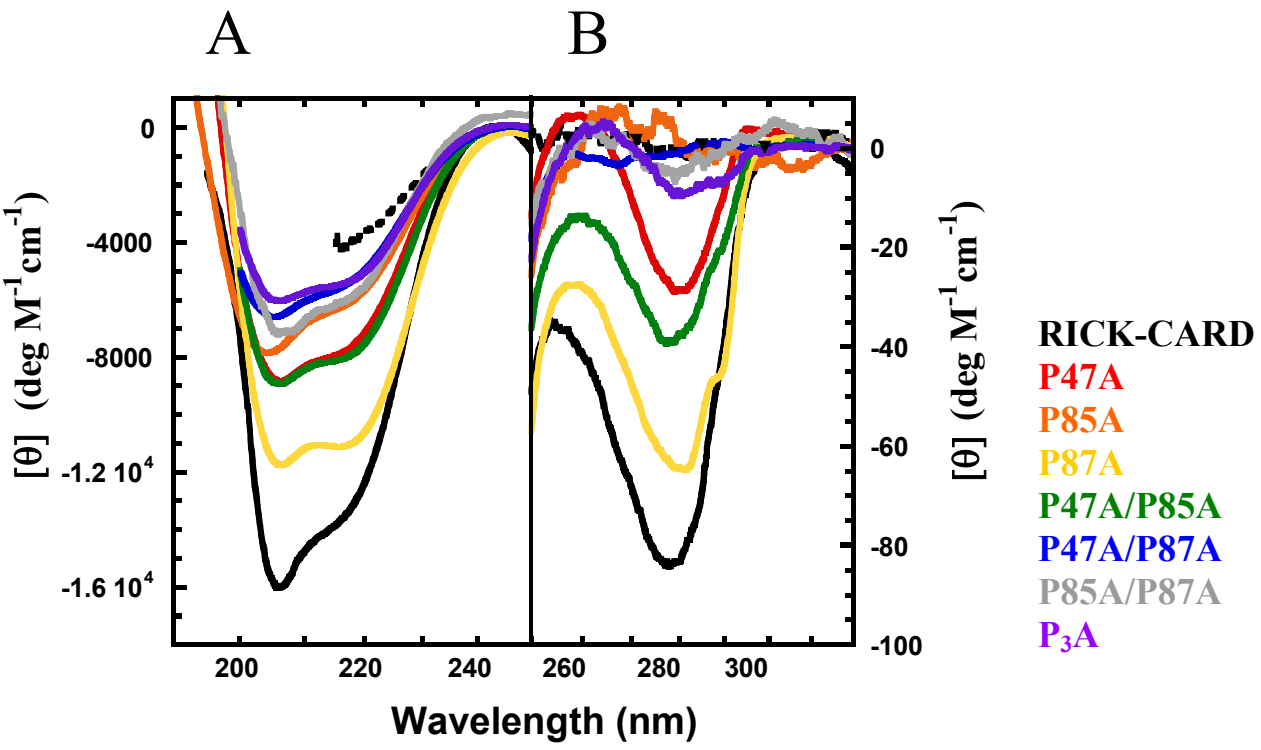


Figure 31. (A) Far-UV and (B) near-UV CD spectra of RICK-CARD and the proline mutants. The spectra of native RICK-CARD in buffer (solid line) and the denatured RICK-CARD in 4 M urea-containing buffer (dash line) are shown in black. The mutant spectra were in the native condition.

5.5 Interaction between Pro-1-CARD and the proline mutants

Previously, we characterized the interaction between Pro-1-CARD and RICK-CARD by fluorescence titration studies and obtained the dissociation constant as 1.4 μM (Figure 23, Chapter 4). Here, we would like to know whether the structural changes by the proline mutations will affect the binding interactions. The results are shown in Figure 32. P87A and P47A have similar patterns with the wild type, whereas P85A and P85A/P87A do not appear to bind Pro-1-CARD.

For the other mutants (P47A/P85A, P47A/P87A, and P_3A) have similar binding compared to the wild type, however, the quenching of the fluorescence signals are lower. It is reduced to 50% of the signal change observed for the wild type. The results might be due to the tryptophanyl residues of the mutants are exposed in different environment upon binding. Overall, beside P85A and P85A/P87A, the mutants seem to have similar binding affinity to Pro-1-CARD compared to that of the wild type. In addition, P87A containing the most similar structure of the wild type has nearly identical binding pattern to the wild type.

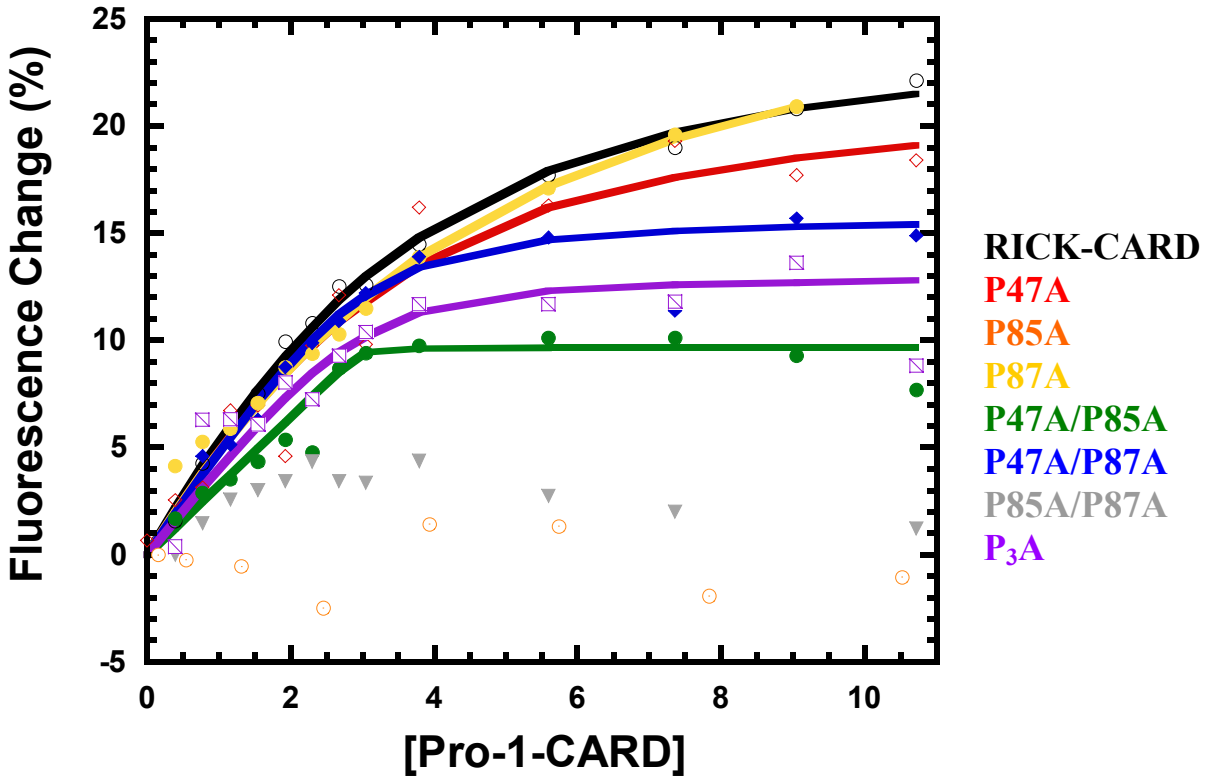


Figure 32. Interaction of proline mutants of RICK-CARD and Pro-1-CARD. Fluorescence titration studies were performed as described in the Methods. The symbols were the fluorescence signal changes at 330 nm normalized to the initial fluorescence signal. The solid lines are fits to a simple protein-ligand binding equation.

5.6 Equilibrium folding of proline mutants of RICK-CARD

We have examined the equilibrium folding properties of the proline mutants following the description in the Methods. The data of each mutant obtained from fluorescence and CD are plotted in Figure 33. We found that the equilibrium folding properties of the proline mutants have changed dramatically. Only P87A has a similar pattern with the wild type. Others look less cooperative than the wild type and P87A. The free energy and m-value from the fits are listed in Table 3.

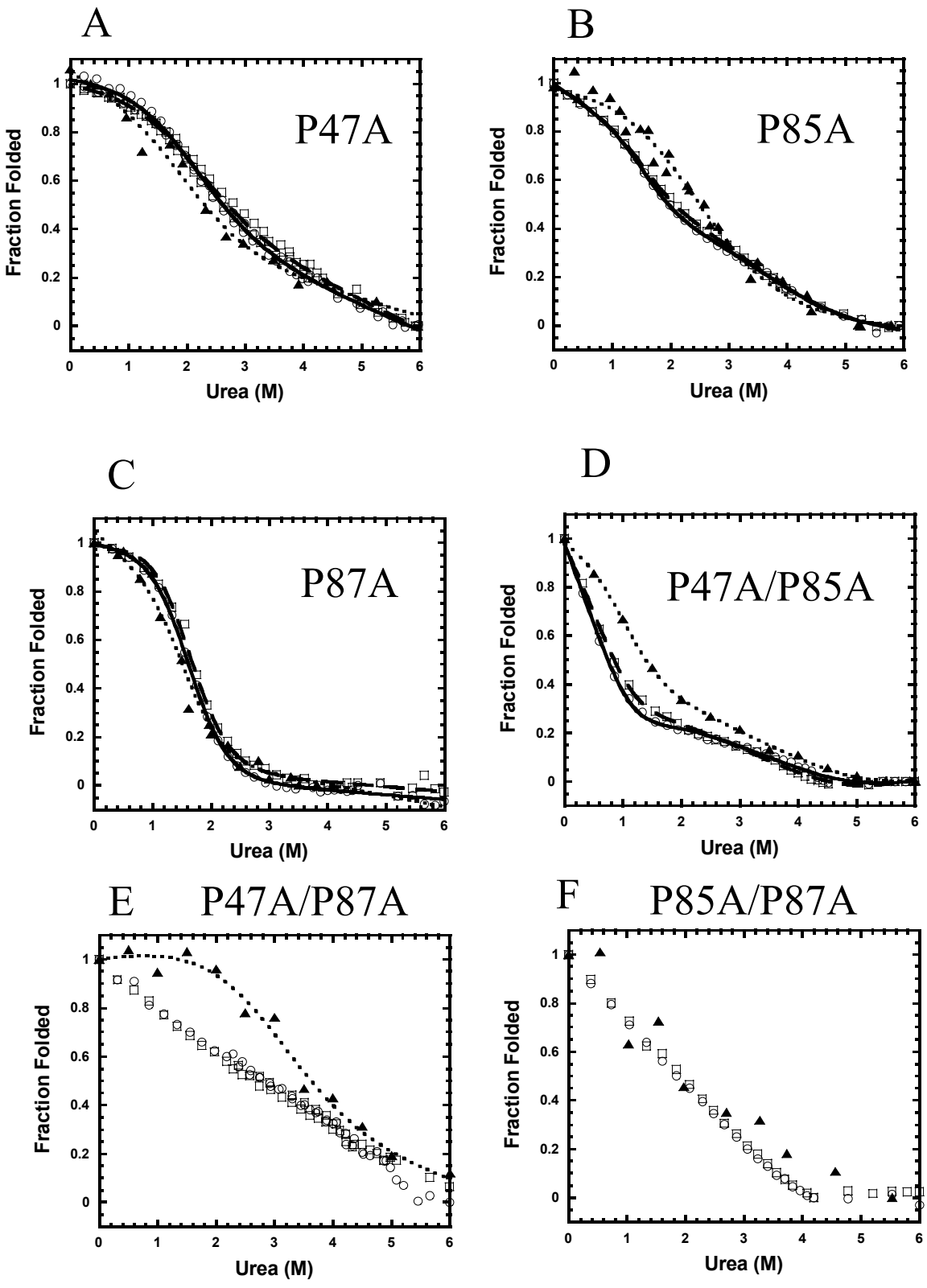
P87A appears to have similar stability with the wild type. Since its m-value increases (1.8 kcal/mol/M), it indicates that P87A has more surface area exposed upon unfolding. In the other words, P87A is more compact and has more hydrophobic residues buried in the native state. The results correlate well with the blue-shifted emission maxima of P87A that are observed in the fluorescence emission in the native conditions compared to that of the wild type (Figure 30A and B). From its m-value, 1.8 kcal/mol/M, we can obtain the solvent exposed surface area as 12963 $Å^2$ using the previously described methods. The solvent exposed surface area accounts for 149 residues in the protein. The predicted value is larger than the actual number of amino acids, 95.

For P47A, we fit the data to a two-state model and obtain the ΔG^{H_2O} as 2 kcal/mol and m-value of 0.99 kcal/mol/M. P47A appears to be less stable than the wild type and has less surface area buried in the native state. For P47A/P87A, the fluorescence data show non-cooperative changes on urea concentration. Therefore, we can only fit the CD data to a two-state model and obtain the ΔG^{H_2O} as 2.4 kcal/mol and m-value of 0.81 kcal/mol/M. This suggests that P47A/P87A is less stable than the wild type, and there is less surface area

exposed upon unfolding. The results correlate well with the CD spectra that show there is very little tertiary structure in P47A/P87A.

For P85A, we no longer can fit the data to a two-state model, but we can fit it to a three-state model (N ⇆ I ⇆ U). The details of the three-state model are described in Appendix F. The total ΔG^{H_2O} is 5 kcal/mol and m-value of 2.25 kcal/mol/M. This suggests that P85A is more stable than RICK-CARD. Furthermore, we also fit the data of P47A/P85A to a three-state model and obtain a total ΔG^{H_2O} of 5.4 kcal/mol and m-value of 2.5 kcal/mol/M. There is a relatively flat transition in 3 M urea for P47A/P85A as well as for P85A. It indicates the intermediates are populated in the equilibrium of P85A and P47A/P85A, and the population of the intermediates is dominated in 3 M urea.

The CD signals overlay with the fluorescence signals in the equilibrium folding studies of P47A, P87A, P85A/P87A, and P_3A. That indicates the loss of the tertiary structure is concomitant with the loss of the secondary structures. However, there are some differences observed in P85A, P47A/P85A, and P47A/P87A indicating for those mutants the loss of tertiary structure is not concomitant with the loss of the secondary structure.


A
P47A
Fraction Folded
Urea (M)
B
P85A
Fraction Folded
Urea (M)
C
P87A
Fraction Folded
Urea (M)
D
P47A/P85A
Fraction Folded
Urea (M)
E
P47A/P87A
Fraction Folded
Urea (M)
F
P85A/P87A
Fraction Folded
Urea (M)

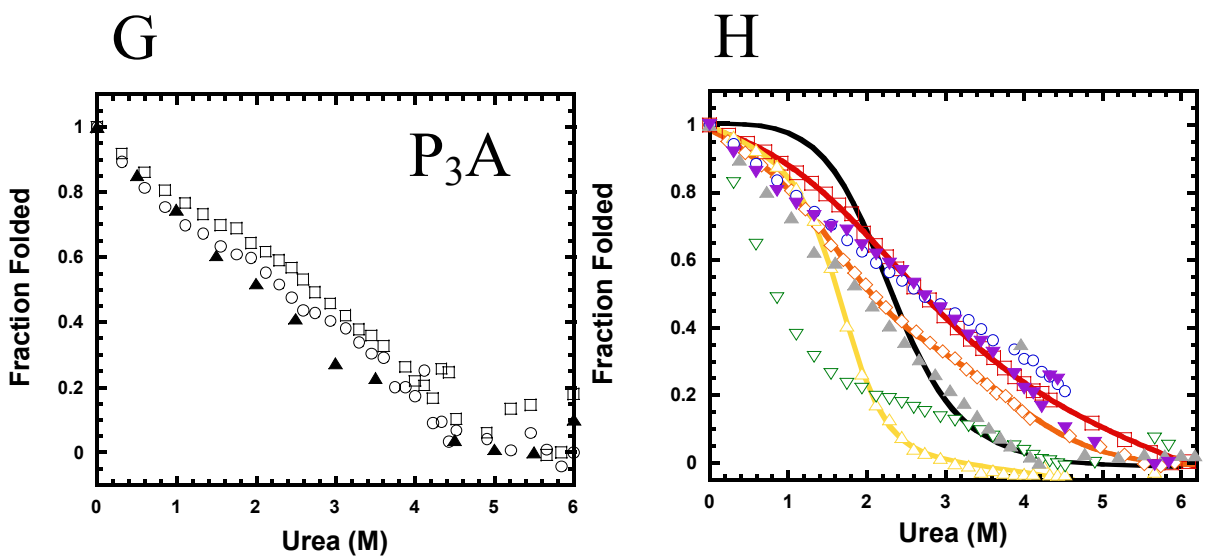


Figure 33. Equilibrium folding of the proline mutants of RICK-CARD. The fluorescence (excitation at 280 (○) and 295 nm (□)) and CD (▲) signals at 220 nm are plotted against urea concentration, and the global fits of the data are indicated by solid, dashed, dot lines, respectively. Besides the data of P85A and P47A/P85A were fit to a three-state mechanism, the others are fit to a two-state mechanism with the parameters listed in Table 3. (A) P47A, (B) P85A, (C) P87A, (D) P47A/P85A, (E) P47A/P87A, (F) P85A/P87A, and (G) P_3A are shown. (H) The fluorescence data (excitation at 295 nm) of the wild type and all the mutants are shown. The color scheme is the same as previous figures.

Table 3 Free energy and m-value of the proline mutants of RICK-CARD

	Folding mechanism	ΔG (kcal/mol)	m (kcal/mol/M)
Wild type	2-state	3.0 ± 0.15	1.27 ± 0.06
P47A	2-state	2.0 ± 0.36	0.99 ± 0.13
P85A	3-state	4.9 ± 0.5 (ΔG_1=3.0; ΔG_2=1.95)	2.25 ± 0.32 (m_1=1.8; m_2=0.45)
P87A	2-state	2.9 ± 0.32	1.78 ± 0.15
P47A/P85A	3-state	5.4 ± 0.5 (ΔG_1=1.91; ΔG_2=3.48)	2.5 ± 0.25 (m_1=1.8; m_2=0.7)
P47A/P87A	2-state (from CD data only)	2.4	0.81
P85A/P87A			
P_3A			

5.7 Kinetic properties of the proline mutants of RICK-CARD

Both refolding and unfolding are performed as described in the Methods. The events will be described individually in the following sections.

5.7.1 Unfolding and refolding of P47A

In the unfolding and refolding experiments of P47A, we observed no kinetic phase for unfolding but a kinetic phase within 200 msec in refolding (Figure 34A). The refolding phase is complete in the same time range as observed for the fast refolding phase of the wild type. The amplitude of the phase is about 20%. The initial and final signals are plotted along with the equilibrium data (Figure 35A). The initial signal of the refolding phases is different from the final signal in low urea concentration (< 2 M) but overlay with the final signal in higher urea concentration (> 2M).

By fitting the refolding traces in different final urea concentrations to a single exponential equation, we obtain rate constants in different final urea concentration. We fit the data to the refolding part of equation [22] as described for the wild type in Chapter 3. The refolding part of the equation is shown as equation [26].

$$k_{obs} = k_{UN} \exp(-m_{U-TS} * [urea] / RT) \qquad [26]$$

We obtain the rate constant of 71 s^{-1} in the absence of urea (Figure 37) and the value of m_{U-TS} is 0.94. The values in comparison with the wild type and other mutants are listed in Table 4.

5.7.2 Unfolding and refolding of P85A

Both unfolding and refolding of P85A are complete within the instrumental dead time (burst phase) based on the fluorescence (Figure 34B). There are no kinetic phases observed. The initial and final signals of the folding and unfolding overlay with the equilibrium data (Figure 35B).

5.7.3 Unfolding and refolding of P87A

The plot of unfolding and refolding traces of P87A are shown in Figure 34C. The unfolding initial (10 msec) and final (10sec) signals of P87A versus urea are shown in Figure 35C.

The unfolding of P87A is complete within the burst phase. The refolding of P87A shows that there is a fast refolding phase that finishes within 200 msec (Figure 34C). It has the same amplitude as the wild type, ~ 40%. In Figure 35C, we can see the initial signal of refolding of P87A has a linear dependence to urea indicating there is no intermediate in the burst phase. In addition, there is no slow phase observed in the refolding of P87A. The rate constants obtained from the single exponential fits to the refolding traces are plotted in Figure 37. The rates versus urea were fit to equation [26].

Since P87A is the only mutant that retains similar structure and equilibrium folding properties with the wild type, we compare the results of the kinetic experiments to that of the wild type. We found that the unfolding events are very different. The unfolding of P87A is complete in the burst phase without having three other phases observed in the wild type. The refolding of P87A does not form a burst phase intermediate neither does the wild type. The fast refolding phase of P87A has similar time scale with the wild type but the

dependence of the rate constant to urea is different (Figure 37). The logarithm of the rate constants in lower urea concentration (< 1 M) has a linear dependence on urea, but the rate constants remain unchanged in higher urea concentration (> 1 M). This result is different from the linear dependence on urea of the rate constants of the wild type. The slow refolding phase observed in the wild type is no longer present in the refolding of P87A.

5.7.4 Unfolding and refolding of P47A/P85A

The unfolding of P47A/P85A is complete within the burst phase (Figure 34D). The refolding of P47A/P85A contains a fast phase with 15% of the amplitude. The initial and final signals of the folding and unfolding are shown in Figure 35D. The rate constants of the refolding fast phase are obtained from fitting to equation [26]. We obtain the rate constant of 55 s^{-1} in the absence of urea (Figure 37) and the value of $m_{U\text{-}TS}$ is 0.84.

5.7.5 Unfolding and refolding of P47A/P87A

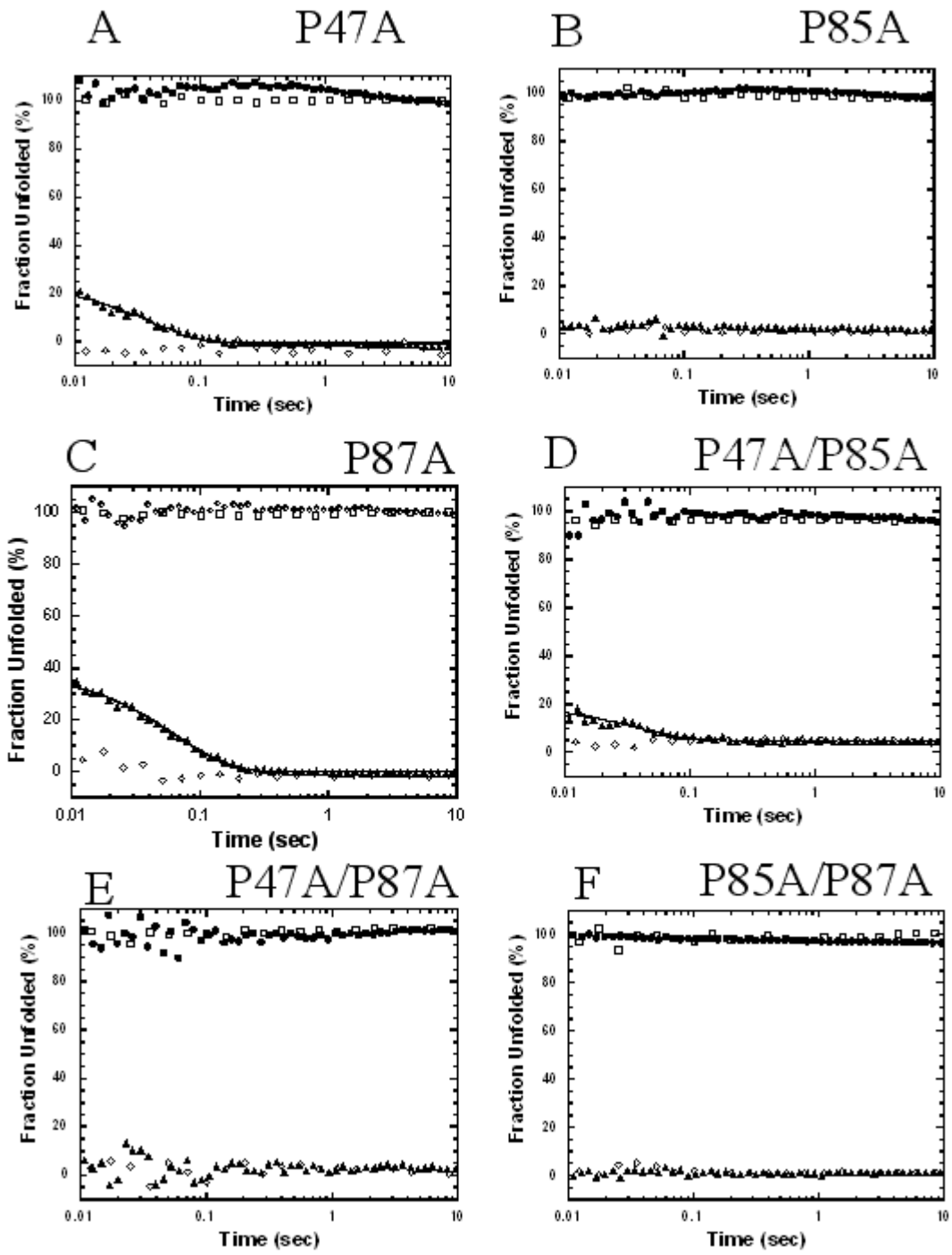
Both unfolding and refolding of P47A/P87A are complete within the burst phase (Figure 34E). The initial and final signals of the folding and unfolding overlay with the equilibrium data (Figure 35E).

5.7.6 Unfolding and refolding of P85A/P87A

Both unfolding and refolding of P85A/P87A are complete within the burst phase (Figure 34F). The initial and final signals of the folding and unfolding overlay with the equilibrium data (Figure 35F).

5.7.7 Unfolding and refolding of P_3A

The unfolding P_3A is complete within the burst phase (Figure 34G). The refolding of P_3A shows a small amplitude (< 10%) in the first 200 msec. The initial and final signals of the folding and unfolding are plotted (Figure 35B). The refolding fast phase was fit to a single exponential equation, and the rate constants versus urea were fit to equation [26]. We obtain the rate constant of 77 s^{-1} in the absence of urea (Figure 37) and the value of $m_{U\text{-}TS}$ is 1.0.


A
P47A
B
P85A
C
P87A
D
P47A/P85A
E
P47A/P87A
F
P85A/P87A
Fraction Unfolded (%)
Time (sec)
0
20
40
60
80
100
0.01
0.1
1
10

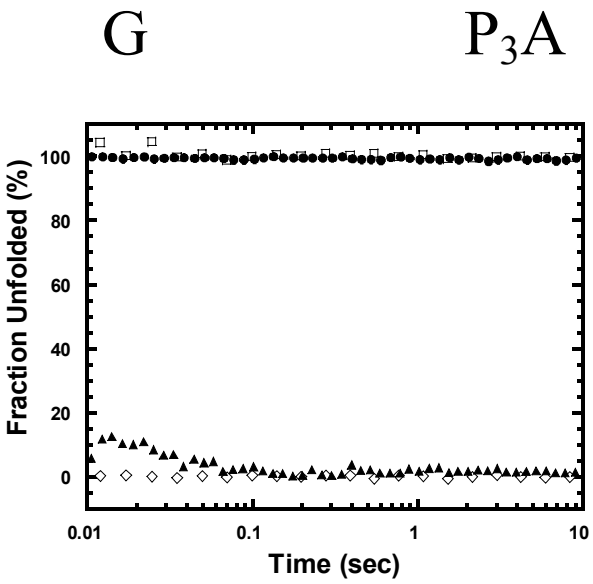
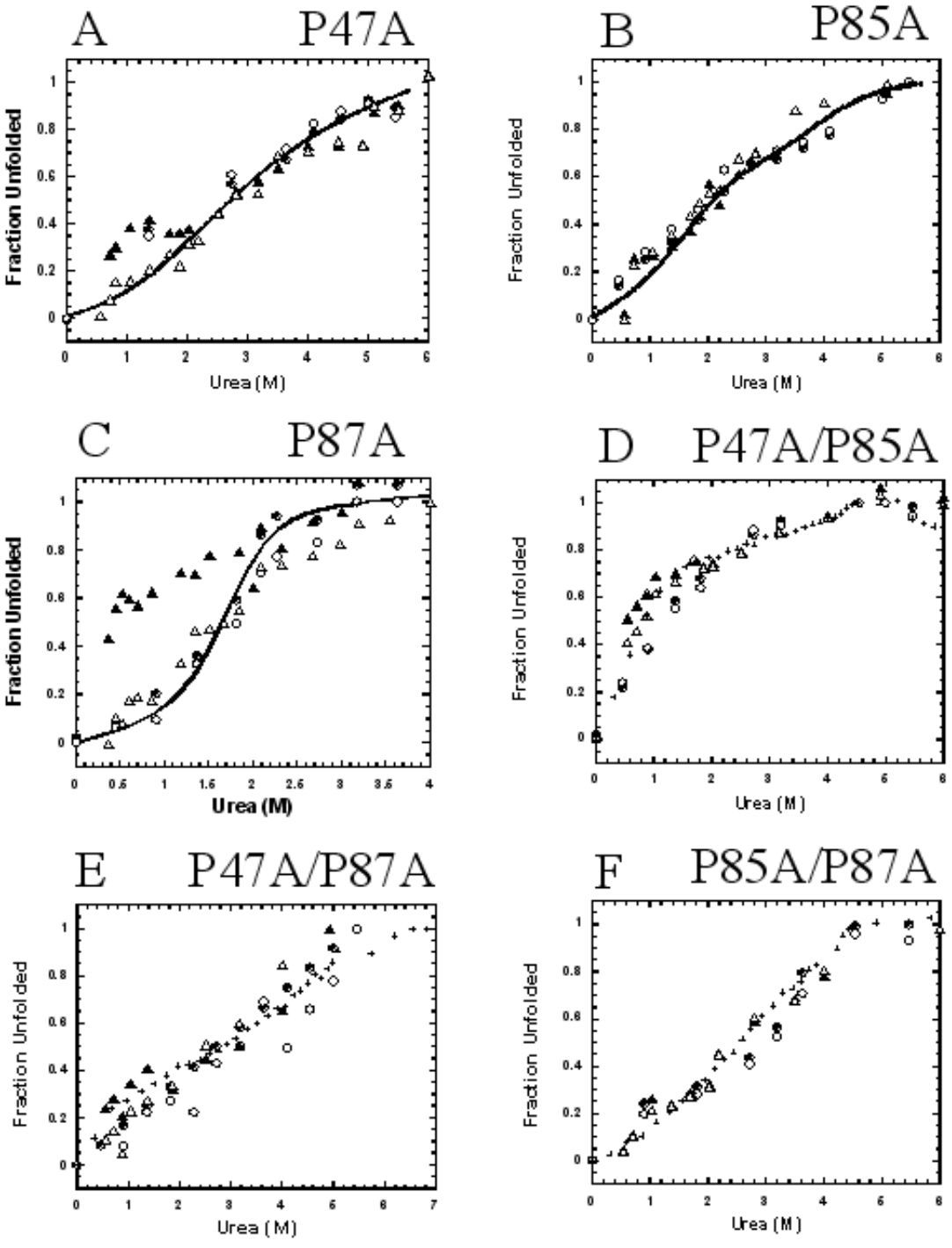


Figure 34. Refolding and unfolding traces of (A) P47A; (B) P85A; (C) P87A; (D) P47A/P85A; (E) P47A/P87A; (F) P85A/P87A; and (G) P_3A. The native control is labeled in diamond (◇), and the unfolded control is labeled in square (□). The unfolding data are labeled in solid circles (●) and the refolding data are labeled in solid triangles (▲). The solid lines shown in Panel (A), (C), and (D) are single exponential fits of the refolding data.


A
P47A
Fraction Unfolded
Urea (M)
B
P85A
Fraction Unfolded
Urea (M)
C
P87A
Fraction Unfolded
Urea (M)
D
P47A/P85A
Fraction Unfolded
Urea (M)
E
P47A/P87A
Fraction Unfolded
Urea (M)
F
P85A/P87A
Fraction Unfolded
Urea (M)

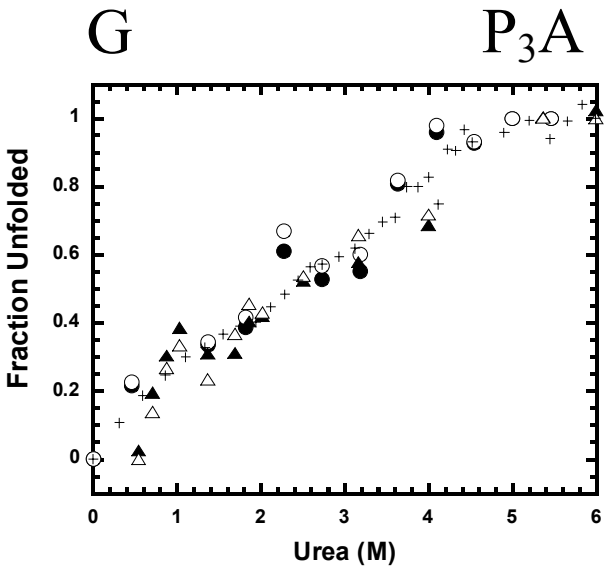


Figure 35. Signals of unfolding and refolding versus urea concentration of the proline mutants: (A) P47A; (B) P85A; (C) P87A; (D) P47A/P85A; (E) P47A/P87A; (F) P85A/P87A; and (G) P_3A. The initial signals (10 msec) of unfolding (●) and refolding (▲) and the final signals (10 sec) of unfolding (○) and refolding (△) are plotted.

5.7.8 Summary of kinetics of the mutants from single mixing experiments

Overall, for unfolding, all the mutants finish the unfolding within the burst phase (~5 msec). In the case of refolding, we observed a single refolding phase in refolding of P47A, P87A, P47A/P85A, and P_3A (Figure 34A, C, D, and G), whereas, refolding of the other mutants including P85A, P47A/P87A, and P85A/P87A are complete in the burst phase (Figure 34B, E, and F). The refolding phases of these mutants have a time scale similar to that of the refolding fast phase observed in the wild type (see Figure 11A). The amplitudes of the refolding phase for P87A, ~40%, is about the same as observed in the wild type. The amplitude of the refolding phase for P47A, P47A/P85A, and P_3A are smaller, less or equal to 20% of the total amplitude (summary in Table 4). All the refolding of P87A, P47A, and P47A/P85A are complete within 200 msec. There are no observable slow phases in the refolding of the mutants. By observation, the amplitude of the refolding phase seems to correlate with the tertiary structure observed in CD spectroscopy (Figure 31). The structure contents from high to low are P87A > P47A/P85A > P47A > P_3A, and the amplitudes of the refolding phases are P87A (40%) > P47A (20%) > P47A/P85A (15%) > P_3A (10%).

Table 4. Summary of the folding and unfolding rate constants and the amplitudes

	$k_{UN\ (s^{-1})}$ (fast)	$Amp_{\ ini}$ (%U)	$m_{\ U\text{-}TS}$	$k_{NU\ (s^{-1})}$
WT(RICK-CARD)	30; 1; 0.012	40	0.9	Burst, 1; 0.08; 0.014
P47A	71	20	0.94	Burst
P85A	Burst	100	X	Burst
P87A	25*	40	Unknown	Burst
P47A/P85A	55	15	0.84	Burst
P47A/P87A	Burst	100	X	Burst
P85A/P87A	Burst	100	X	Burst
P_3A	77	10	1.0	Burst
Pro-1-CARD	Burst	100	X	Burst

The rate constants and m-values were obtained from fits of the data shown in Figure 37 using equation [26]. The ones without refolding phase are marked as "X". "Burst" represents the kinetics are complete in the dead time. The subscript UN is for refolding (U ⇆N), and that NU is for unfolding (N⇆U).

* The refolding rate constant of P87A is obtained from the fit using equation [26] to the data in low urea concentration (<1 M) as shown in Figure 37.

Since the structure content and equilibrium folding of P87A are closer to the wild type, we compare the initial amplitude of P87A in different urea concentrations to examine details of the burst phase. As shown in Figure 36, the initial signals have a linear dependence on urea suggesting there is no burst phase intermediate present for P87A in the burst phase. The results are consistent with that observed from the wild type (see Figure 12C).

I plot the rate constants obtained from the refolding of the proline mutants including P47A, P87A, P47A/P85A, and P_3A as functions of urea concentration (Figure 37). The values are listed in Table 4. The rate constant of P47A has a similar pattern to the wild type but with a rate constant two times faster. Similar rate constants are obtained for P47A/P85A, and P_3A, although there are fewer data points for these two proteins. By comparing the m-value describing the urea dependence between the unfolded and the transition states ($m_{U\text{-}TS}$), we found that the $m_{U\text{-}TS}$ values do not have significant changes. The results indicate that the refolding process of those mutants have similar solvent exposed surface area to the wild type upon refolding from the unfolded to the transition states. Therefore, the transition states of the mutants may have similar structures with that of the wild type. Although the rates for those mutants are faster, the process may be quite similar.

The rate constant of P87A (25 s^{-1}) is very closed to the wild type (30 s^{-1}). However, the patterns of the rates depending on urea are very different (Figure 37). The observed rate constants in low urea concentrations (< 0.9 M) are close to those of the wild type. But the rate constants of P87A in relatively higher urea concentrations (> 1 M) are faster and remain constant.

The refolding kinetics of P87A show similar properties to the wild type with a few differences. First, the refolding slow phases of the wild type are not observed in P87A. Second, the mid point of the transition in kinetics is changed. Overall, P87A still retains the fast refolding phase with the same rate, and it is likely that the refolding of P87A is the same as the wild type beside the few difference mentioned above.

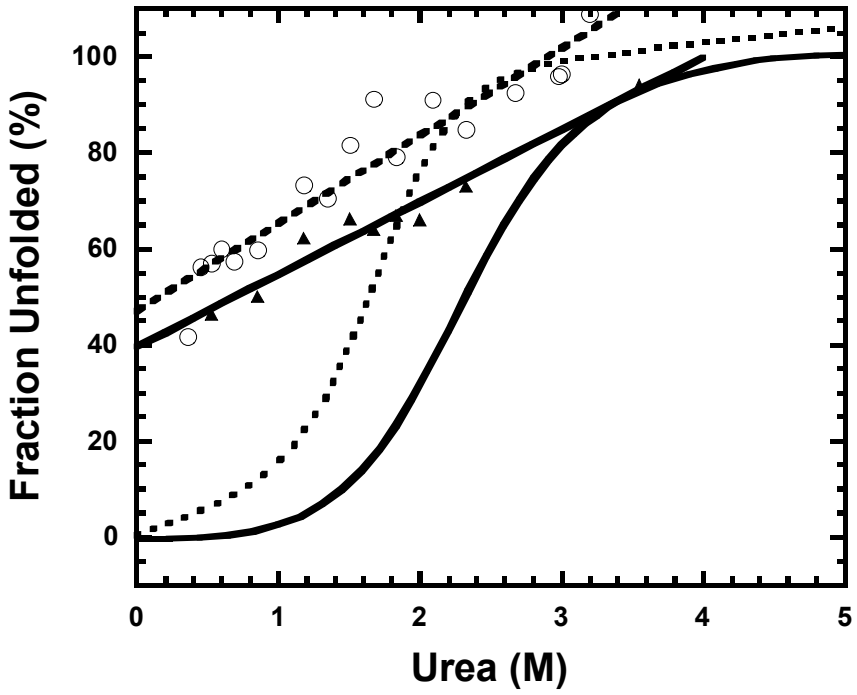


Figure 36. Comparison of the refolding burst phase signals between P87A (○) and the wild type (▲). The initial signal of P87A and the wild type are plotted. The dashed and solid lines are the linear fits of the signals of P87A and the wild type. The fits of equilibrium folding data where the samples were excited at 295 nm are plotted in dashed and solid curves for P87A and the wild type.

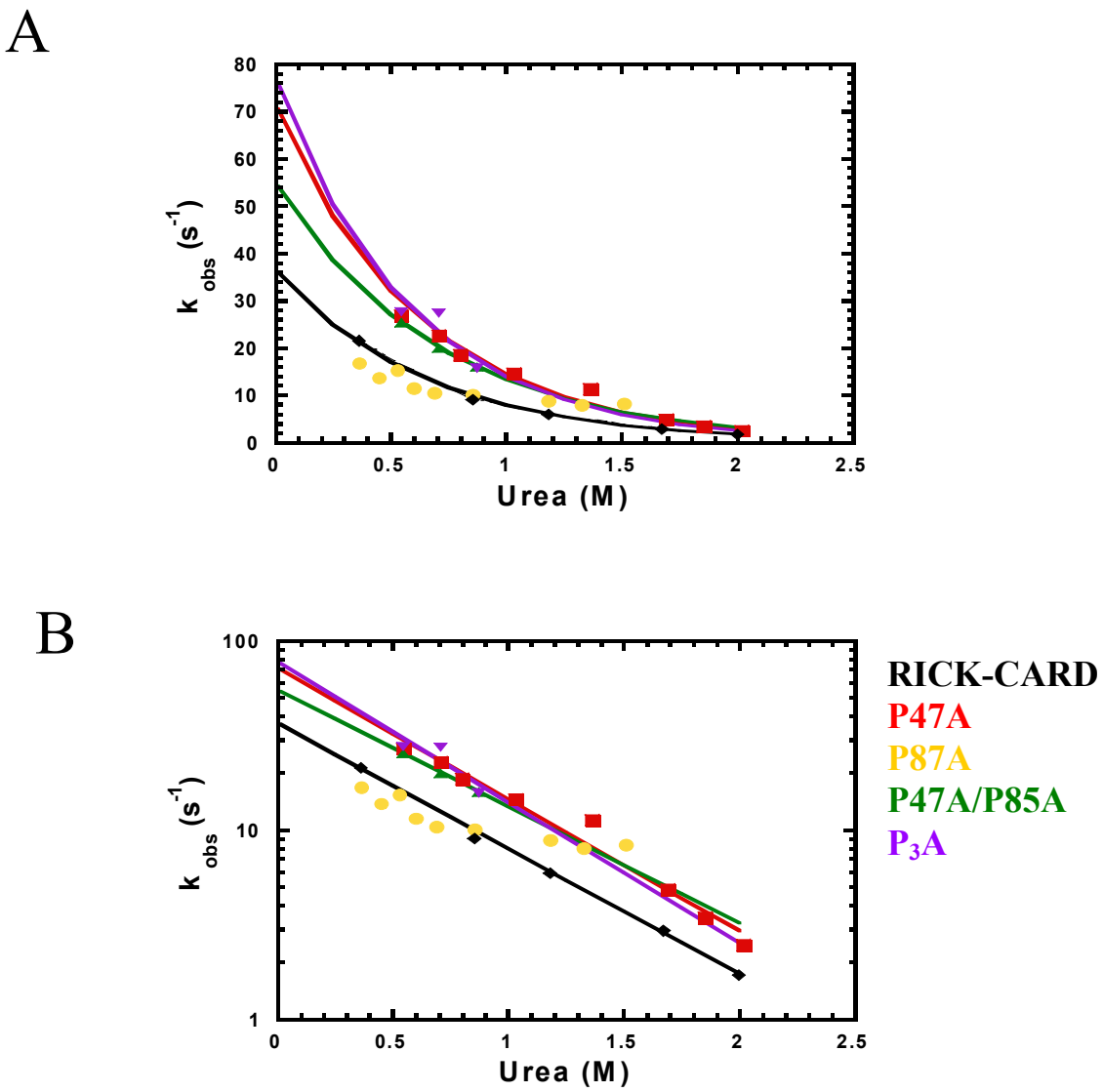


Figure 37. Comparison of the refolding rate constants of P87A (●), P47A (■), P47A/P85A(▲), P_3A (▼), and the wild type (◆). The rate constants were obtained from single exponential fit to the refolding data and plotted versus urea concentration in linear scale (A) or log scale (B). The rate constants of the wild type were obtained as described previously (Figure 14).

5.8 Double jump experiments

In order to examine the detail of unfolding process, we have employed double jump experiments as described in the Methods. We first unfold the protein for various amount of time then refolding it to the native state. The refolding traces from the second jump show linear traces and move from the unfolded to the native controls while the delay time changes (Figure 38). We observed signal changes with different delay time. The results show a similar pattern as we observed for the wild type protein but the transition occurs earlier ($t_{1/2}$= ~ 20 sec) (Figure 39).

Based on the results of the unfolding experiments, we know that for all the mutants unfolding is completed within 10 msec because the signal at 10 msec overlays with the equilibrium data meaning the event is completed (see Figure 35). However, based on the results of double jump, if we unfold a native protein for relatively short delay time (< 10 sec) then we refold the protein, we do not observe formation of the native species. When the delay time is prolonged (> 10 sec), the population of the native protein increased. The results show that the protein can unfold and refold to 100% of the native structure only if the time of unfolding is longer than 30 sec. This result indicates that there are still kinetically trapped intermediates on the unfolding pathway. Furthermore, the data show that the fluorescence signal of the intermediate is the same as that of the unfolded protein. Therefore, the results suggest that an intermediate containing the unfolded fluorescence signal is present, and the intermediate is not able to refold to its native structure. The native protein can be totally recovered only if the unfolding is complete at about 30 sec.

The half times of the transitions of the mutants in the double jump experiments are all very close. They range from 23 to 15 sec. However, the half time is much faster than that

of the wild type, ~50 sec. The results may indicate the proline mutants share a common unfolding step before they reach their unfolded states. If one suggests the lag phase and the transition of the wild type observed from the double jump experiments correlate with its slow unfolding phases with half-time 8.6 and 50 sec, respectively, the transition of the mutants could also be an unfolding phase with a half time of ~20 sec. The unfolding rates of the mutants are clearly faster than the wild type, however, they may unfold through similar kinetic traps but in different time scales.

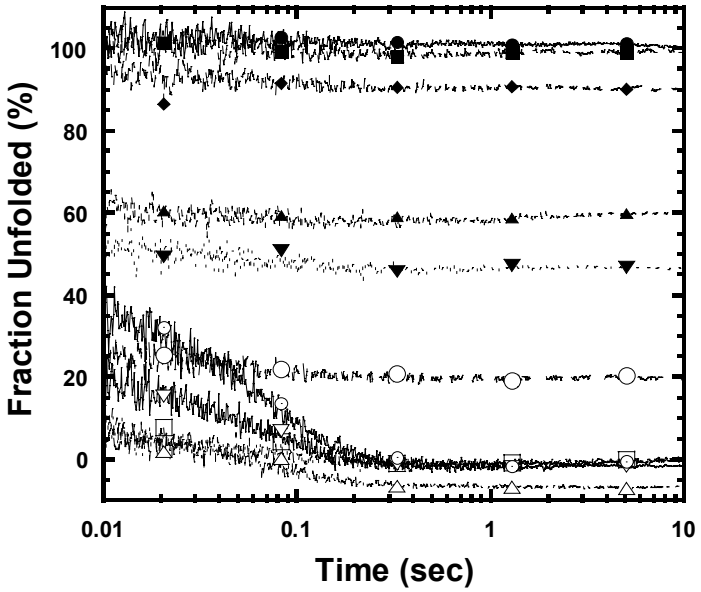


Figure 38. The refolding traces of P87A in double jump experiments. The refolding signal traces from the second jump of double jump experiments. The delay time is indicated as follows: 1 (●), 5 (■), 10 (◆), 15 (▲), 20 (▼), 30 (○), 50 (□), 100 (△), 200 (▽), and 500 sec (⦿). The signals are normalized to fraction unfolded.

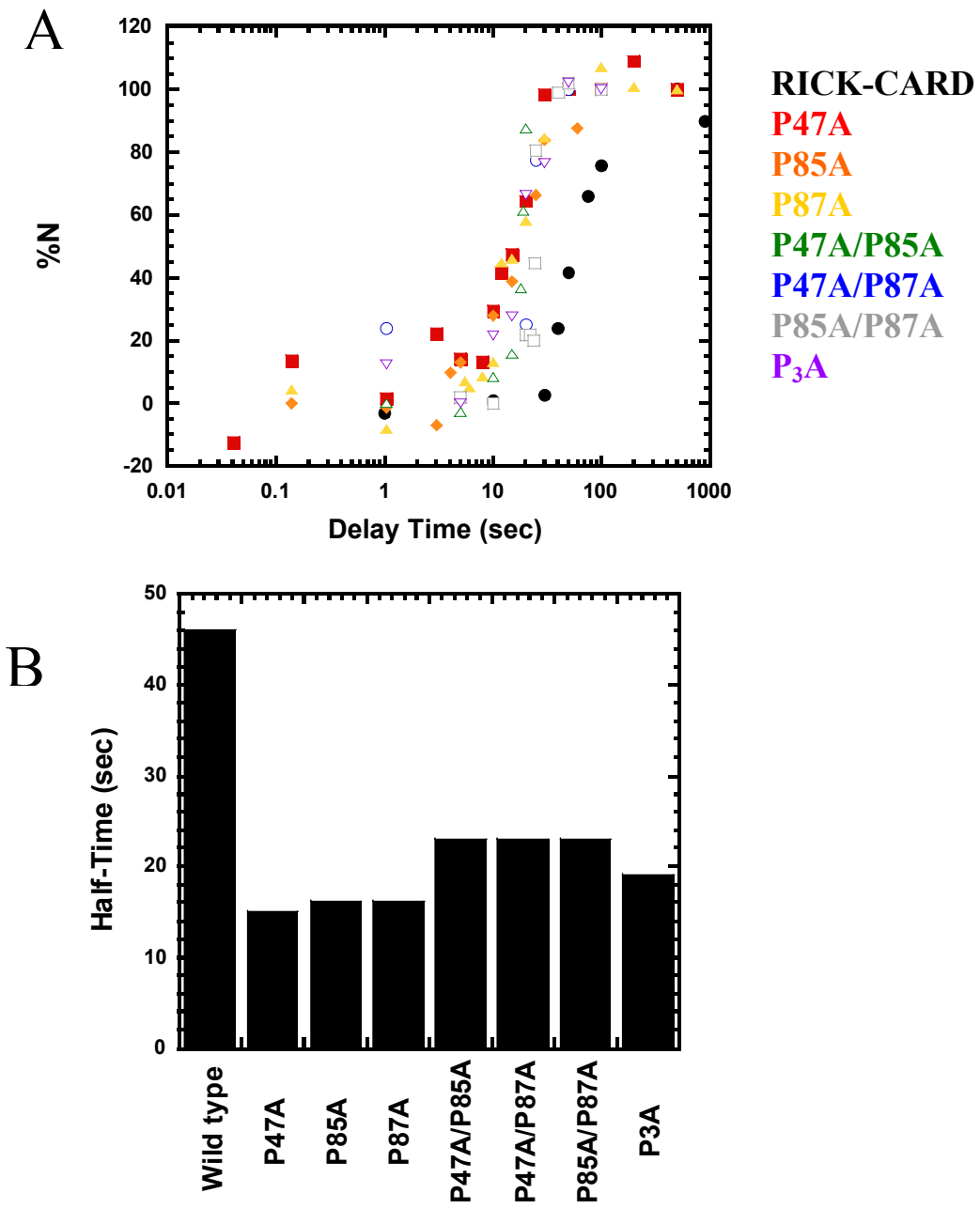


Figure 39. Double jump experiments of the proline mutants. (A) The percentage of native protein versus delay time of unfolding is plotted based on the final signal at 10 sec. The mutants are labeled as the color scheme indicated. (B) Half-time of the transitions for proline mutants in the double jump experiments. The half-times were estimated from the data shown in Panel (A).

5.9 Interrupted refolding of the proline mutants

We have preformed interrupted refolding experiments to examine the details of refolding process. We are only able to do the experiment with P87A because of the lower cooperativity of all the other mutants (Figure 33A, B, D, E, F, G). The proteins are unfolded in 6 M urea rather than 4 M urea as with wild type and P87A. With the standard set up of the sequential mixing, the final urea concentration would be 3 M if the proteins were in 6 M urea to begin with. This final urea concentration is not sufficient for the proteins to fully refold.

For P87A, we first refolded the protein for various amount of time then unfolded the protein. The signals of the unfolding in the second jump are monitored. The unfolding process shows linear traces moving from the native to unfold controls while the delay time of refolding increases. When the delay time of refolding is less than 10 sec, the signal is close to the native control (Figure 40).

One would predict that if the initial signal of unfolding in the second jump is close to the native control, there should be a high population of native species present after the first jump (U → N). However, when the protein unfolds after a short amount of refolding in the first jump, the observed signal does not change. The final signal of the trace remains constant and is close to that of the native signal even in the denatured condition. The results suggest that the protein first refolds to an intermediate, which contains similar fluorescence signal with the native species, but the intermediate is trapped and not able to unfold. When the delay time of refolding is prolonged, the native species starts to form so that the signal is close to the unfolded control. This means that there is a population of native species after

the first jump that is able to unfold to the unfolded state. This result is different from what we have observed from RICK-CARD, although the patterns after normalization are similar (Figure 40).

The half time of the transitions of P87A in the interrupted refolding experiments is about 23 sec, which is faster than that of the wild type, ~86 sec. The results indicate that there is kinetically trapped intermediate on the refolding pathway. This phase is not dependent on proline isomerization since we do not see a slow refolding phase in P87A. Moreover, the proline mutants may share a common folding intermediate with the wild type before they reach their native states. Overall, we find that P87A and the wild type protein have similar burst phase and fast phase kinetics. But formation of native protein from the intermediate is faster in P87A than in the wild type.

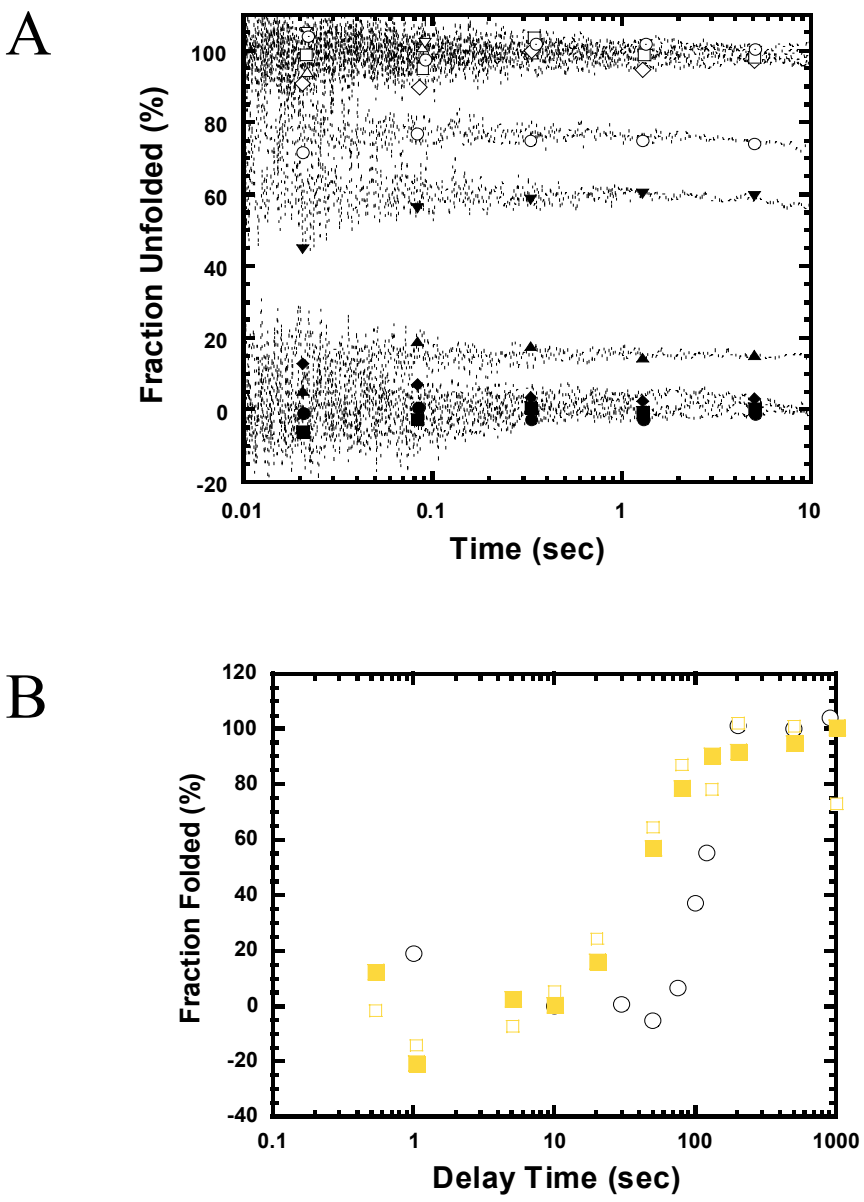


Figure 40. (A) Signal traces of P87A in the interrupted refolding experiments. The delay time is as follows: 0.5 (●), 5 (■), 10 (◆), 20 (▲), 50 (▼), 80 (○), 90 (△), 100 (▽), 130 (◇), 200 (□), and 500 sec (⦿). (B) Fraction folded versus delay time in interrupted refolding experiments of P87A (square) and the wild type (circle). The normalized data of the initial signal at 10 msec (open symbol) and final signal at 10 sec (closed symbol) are shown.

5.10 Discussion of Chapter 5

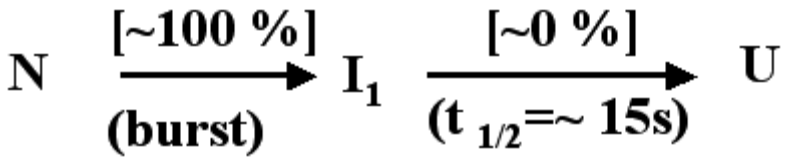
In this Chapter, I have shown the results from fluorescence emission spectra, CD spectra, interaction studies, equilibrium folding studies, and kinetic folding studies of all seven proline mutants generated from RICK-CARD.

In all the mutants, P87A appears to retain the wild type native structure and the function, although it is more stable and appears to be more compact than the wild type. P87A retains some of the kinetic properties of the wild type for example having the same fast folding rate and no burst phase intermediate. However, the slow refolding phase is gone. It is probably due to the mutation at position 87 either by eliminating the proline isomerization or slow folding events occurred in local conformation. Form the interrupted refolding experiments, we know that there is still an intermediate which acts as a kinetic trap on the refolding pathway with a half time of ~ 23 sec. The proposed kinetic pathways for P87A are shown in Figure 41. From the results, wee suggest that the intermediate (I_2 in Figure 41) of P87A might be the same intermediate (I_4) of RICK-CARD as shown in Figure 17.

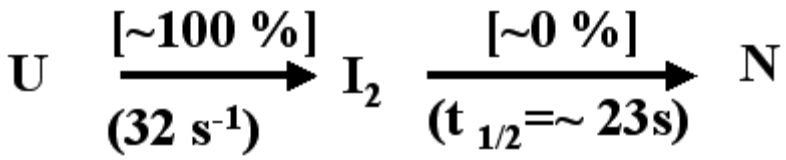
The slow kinetic phase observed in refolding of the wild type protein does not appear in that of P87A. Moreover, the activation energy of RICK-CARD is similar to that required for proline isomerization. These results suggested proline 87A is in *cis* configuration in the native condition. Therefore, the slow refolding phase is due to incorrect isomeric form, trans-proline, converting to the correct isomeric from, cis-proline. From our results, we cannot determine the isomeric states of P85A and P47A. However, their kinetics are fast and without slow phases. It is more likely that prolines 85 and 47 are in trans than they are in cis but do not go through isomerization. There are some proteins containing PXP

motifs and the second proline is in *cis*[96]. We cannot rule out the possibility of local structure arrangement that occurs in slow processes, such as slow docking of helix-6 with helix-1.

The unfolding of P87A is fast. The major unfolding phase occurs in the burst phase. However, we still observed a slow unfolding phase with a half-time of 16 sec from the double jump experiments. The results suggest that by mutating the proline at 87, the unfolding process is changed. The three phases observed in the unfolding of the wild type are no longer observed in that of P87A. However, the third unfolding phase with a half time of 9 sec in the wild type protein is close to the slow unfolding phases ($t_{½}$ = 16 sec) of P87A. Hence, the intermediate phase in unfolding of RICK-CARD (I_3 in Scheme 1, Figure 17) might be the same as the one in P87A (I_1 in Scheme1, Figure 41). However, we do not have direct evidence to prove these assumptions.



Scheme 1



Scheme 2

Figure 41. Proposed unfolding (scheme 1) and refolding (scheme 2) pathways of P87A. The percent change in fluorescence emission associated with each reaction is shown above and the half-time of the phases are shown below the reactions. N and U represent native and denatured states, respectively. I_1 and I_2 represent the non-native sequential intermediates on the pathways.

Mutation at position 85 (P85A) affects the function of the protein. P85A no longer binds to its binding partner, Pro-1-CARD. It loses the tertiary structure and appears to be less compact (based on fluorescence and CD spectroscopy) in the native condition. From the equilibrium folding study, we found this protein adopts a three state equilibrium folding mechanism with higher stability than the wild type. The equilibrium intermediate is dominated in ~3 M urea. Both folding and unfolding of the protein are fast and seem to complete in the burst phase. However, we found the kinetically trapped intermediate still is present in the unfolding pathway of P85A. Overall, P85A contains some structure but not the native wild type structure. It is more stable and with an intermediate populated at equilibrium. The kinetics are different than the wild type.

The combination of mutations at positions 85 and 87 (P85A/P87A) results in a protein with similar characteristics to those of P85A. There is not much tertiary structure and little function. The kinetics are fast but with an intermediate in the unfolding phase. The equilibrium folding is not cooperative. Overall, P85A/P87A does not contain the wild type native structure and loses both function and folding cooperativity. These results suggest that position P85A is important for maintaining native wild type structure, and it does not recover the wild type properties with another mutation at proline 87.

Interestingly, we thought by replacing the proline residues in a helix would increase the helical content of the protein and further make the protein more compact. In contrast, we find that mutating a single proline at position 85 disrupts the native wild type structure. It suggests the mutation may have altered the position of the side-chains of the nearby residues and somewhat disrupted the stabilizing interactions for RICK-CARD. Further discussion is in Chapter 6.

P47A located at turn 3 region contains some of the native wild type structure. It has the same affinity to Pro-1-CARD. It still adopts a two state equilibrium mechanism but with about 1/3 of the stability of the wild type. The unfolding is fast but contains an intermediate on the pathway. The amplitude of the refolding burst phase is about 20 to 40% in low urea concentration but the burst phase signals overlap with the equilibrium data in more than 2 M urea (Figure 35A). The results indicate that the burst phase kinetics of P47A is different than the wild type. There might be a burst phase intermediate populated in low urea concentration. There is still a fast refolding phase in P47A with a faster folding rate than the wild type. The transition state of P47A may be similar to the wild type due to similar $m_{U\text{-}TS}$ values (Table 4).

With another mutation at position 85 along with the mutation at position 47 (P47A/P85A), the double mutant retains some tertiary structure compared to P85A but still less than that of P47A (Figure 24). The equilibrium folding of P47A/P85A is fit to a three-state mechanism (N ⇆ I ⇆ U, more details in Appendix F). The pattern is very different than any of the other mutants. It shows two transitions from 1 to 2 M urea and from 3 to 4 M urea with a plateau at ~ 2 to 3 M urea. The intermediate is predominated in ~ 3 M urea indicating the intermediate could be similar to that of P85A at equilibrium. In kinetic experiments, the unfolding is also fast and contains an intermediate on pathway. There is a little amplitude of the refolding burst phase (~15 %). The phase is observed in less than 2 M urea concentration, but the phase is gone in more than 2 M urea concentration. The results indicate P47A/P85A may contain a similar burst phase intermediate with that of P47A. Since the refolding rate is similar to that of P47A, P47A/P85A may adopt similar refolding fast phase with P47A.

With the combination of mutation at position 47 and 87 (P47A/P87A), the protein also loses the tertiary structure of the native protein. In urea unfolding, the change of the fluorescence is noncooperative, although there is cooperativity based on the secondary structure.

Both unfolding and folding kinetics are fast but with an intermediate present on the unfolding pathway. By comparing the results of P47A/P87A to P47A or P87A separately, we found that P47A/P87A behaves neither like P47A nor P87A. It seems that by mutating the two positions together, we have further disrupted the native structure and the cooperativity.

Finally, if we mutate all three proline residues (P_3A), we do not obtain a well-packed protein. P_3A contains little tertiary structure and is not cooperative in equilibrium folding studies. The unfolding kinetics are fast but with an intermediate present. The refolding has a very small amplitude of the fast phase. The rate is similar to that of P47A, and P47A/P85A. Therefore, it seems P_3A has similar refolding fast phase to those of P47A, and P47A/P85A.

In addition, all the $m_{U\text{-}TS}$ values for the refolding phase of the mutants who contains a refolding fast phase (P47A, P87A, P47A/P85A, and P_3A) are similar to that of the wide type. This indicates they all have similar urea dependence and solvent exposed surface area upon refolding. It further suggests that the transition state in refolding of those mutants may be similar to that of the wild type.

Overall, our results suggest that proline 85 is a critical position to maintain the native structure and function. Proline 87 is not critical for the structure and function but contributes to refolding slow phase. It suggests proline 87 is in cis in the native structure. Proline 47 is also important for the structure, but it is not critical for the function. Most of

the double and triple mutations of RICK-CARD, except P47A/P85A, result in little tertiary structure. The tertiary structures of the proteins are less cooperative.

The kinetics of the mutants are quite different. In general, their unfolding are fast and the refolding are fast. The refolding fast phase in the wild type is present in P87A with the same amplitude. Some mutants also have the fast phase but with much smaller amplitude. They all have similar intermediates serve as kinetic traps on refolding and unfolding pathways. Here, we conclude that the mutations have changed the kinetic properties of RICK-CARD.

Chapter 6

DISCUSSION

Chapter 6

Discussion

Foreword

In this chapter, the general properties of RICK-CARD and Pro-1-CARD are discussed, and their folding properties are compared.

6.1 Difference between RICK-CARD and Pro-1-CARD

Anfinsen suggests that the information for protein structure and protein folding mechanism are defined in its primary sequence [4; 5]. Some research indicates the topology, or structure, is the major determinant of folding kinetics [9; 15]. Here, we have shown that the folding properties of Pro-1-CARD and RICK-CARD have distinctly different behaviors, although their structures are similar. I will discuss the difference based on their primary sequences in the following sections.

6.1.1The primary sequence

The sequence identity between RICK-CARD and Pro-1-CARD based on the alignment shown in Figure 4 is about 21% (20 identical residues in total of 95 residues) and the similarity is about 38% (36 residues in total of 95 residues). Both numbers are not high. RICK-CARD and Pro-1-CARD both contain many charged residues that are also seen in

other CARDs. There are 12 basic and 14 acidic residues in RICK-CARD and 13 basic and 15 acidic residues in Pro-1-CARD. Both the proteins do not contain disulfide bonds. The three prolyl residues in RICK-CARD are located in turn-3 and helix-6, whereas, the only proline in Pro-1-CARD is located in turn-4. In Pro-1-CARD, the corresponding positions for the prolines of RICK-CARD are Asn, Gly, and Leu. Those prolines are not conserved. However, there are proline residues at position 47 in APAF-1-CARD and ARC, another CARD containing protein. There are also two proline residues in helix-6 of ARC but it is a PDP motif rather than PYP[43]. There are some proline residues located in turn-4 in ICEBERG and RAIDD (Figure 4).

6.1.2 Helical propensity of RICK-CARD and Pro-1-CARD

Through the algorithm AGADIR [90; 97], which calculates the helical propensity of a peptide chain based on its sequence, we obtain the helical propensities of RICK-CARD, Pro-1-CARD, and ICEBERG. AGADIR has implemented the energy contributions for various interaction occurring in an α-helix including intrinsic helical propensities, side-chain -side chain interactions, main-chain-main-chain hydrogen bonds, and capping effects [90; 97] . The helical propensities are plotted in Figure 42. In general, their helical propensities are very different. However, the distribution of helical propensities through six helices is closer between Pro-1-CARD and ICEBERG compared to that of RICK-CARD and ICEBERG. This is due to higher sequence identity between Pro-1-CARD and ICEBERG.

Pro-1-CARD is not predicted to have much helical content in helices-5 and -6. The helices, helices-1 and –4, which are responsible for binding are predicted to helical. Helix-5

of RICK-CARD contains high helical propensity. In addition, there are some helical propensities in helices -2 and –3, which are responsible for the CARD-CARD interaction. Changing the primary sequences of RICK-CARD to the sequence of the proline mutants has little effect on the predicated helical propensity of RICK-CARD.

The results of helical propensities of CARDs are for reference since AGADIR only considers short-range interactions. Also, the algorithm does not predict well on some residues such as Pro, His, Cys and Trp due to insufficient data[90; 97]. The helical propensity could be influenced by the tertiary interactions in longer peptide chains. Therefore, the actual helical content could be different in RICK-CARD and Pro-1-CARD than the predicted values.

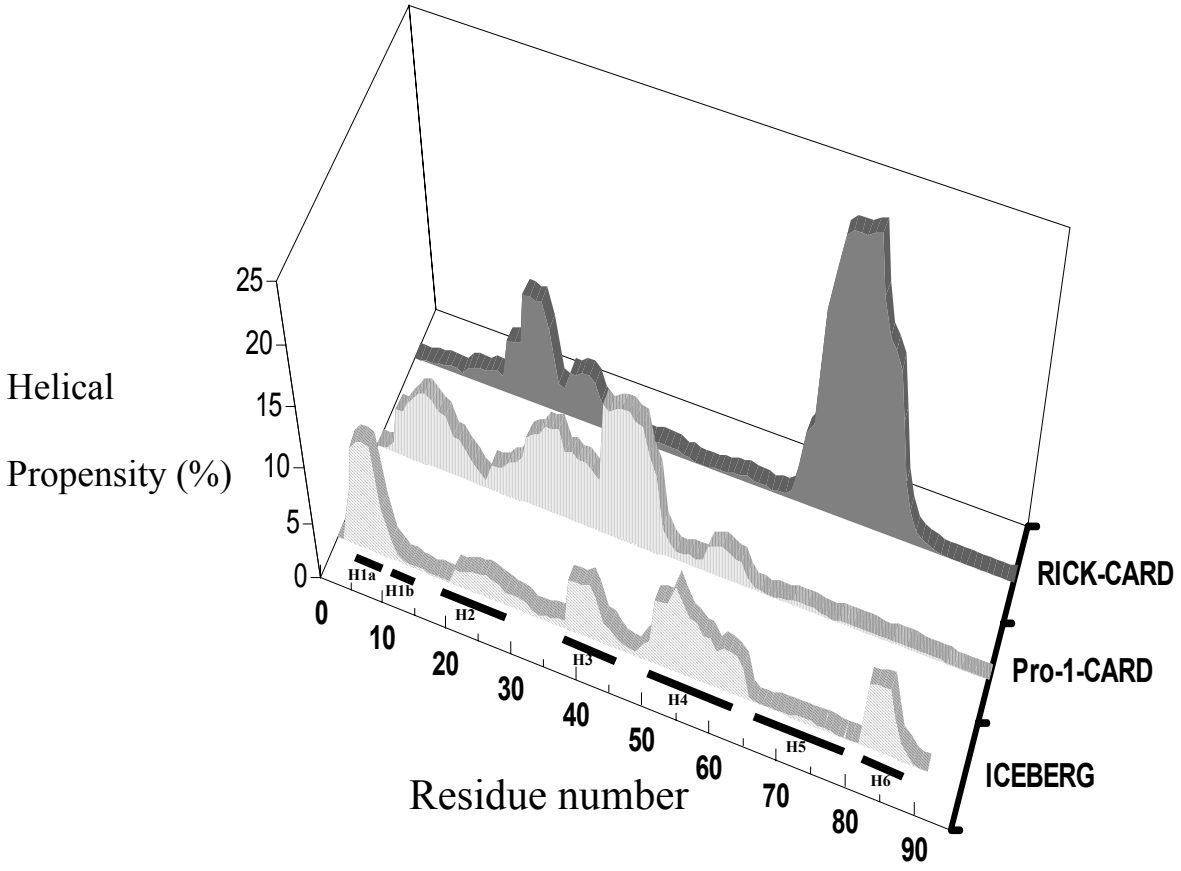


Figure 42. Helical propensities of CARDs. The values are predicated by AGADIR based on their primary sequences.

6.2 Predicated folding rates from contact order and long-range contact

While the determinants of two-state folding kinetics are still ambiguous, Baker et. al. have suggested the topology as one of the critical determinants for homologous two-state proteins[9; 25]. The contact order (CO means relative contact order) represents the average sequence separation between the pairs of contacting residues in the native structures, and is employed to describe the topology. According to the calculations, we obtained CO of Pro-1-CARD and RICK-CARD as 7% and 8%, respectively. Based on the prediction, the folding rate of RICK-CARD and Pro-1-CARD should be about $1*10^5$ and $2*10^5$ s^{-1}, respectively. Another algorithm, LRO, which takes into account of long-range interactions in the native protein [98] is defined as contacts between two residues that are close in space (≤8 Å) and far in the sequence (>12 residues). We calculate the LRO of Pro-1-CARD, RICK-CARD, ICEBERG, and APAF-1-CARD to be 1.0, 1.413, 1.056, and 1.674, respectively. The predicted folding rates of Pro-1-CARD, RICK-CARD, ICEBERG and APAF-1-CARD are 6900, 2700, 6070, and 750 s^{-1}, respectively (Figure 43). The predicted folding rates of CARDs are not far from the folding rates of small helical bundles ranging from ~300 to ~5000 s^{-1} [23; 24]. However, we obtained the fastest rate of RICK-CARD as 30 s^{-1}, which can be due to folding complexity of RICK-CARD since the folding kinetics is not a simple two-state mechanism. Moreover, the predictive rate of Pro-1-CARD falls into the range if the folding is complete before 10 msec. For example, if Pro-1-CARD folds with a half time of 1 msec, the folding rate would be 700 s^{-1}. However, from our interrupted refolding experiments, we found an intermediate in refolding with a half time of 10 sec. According to the results, we suggest that the major refolding event of Pro-1-CARD could be consistent with the predicted value but overall the kinetics are more complex.

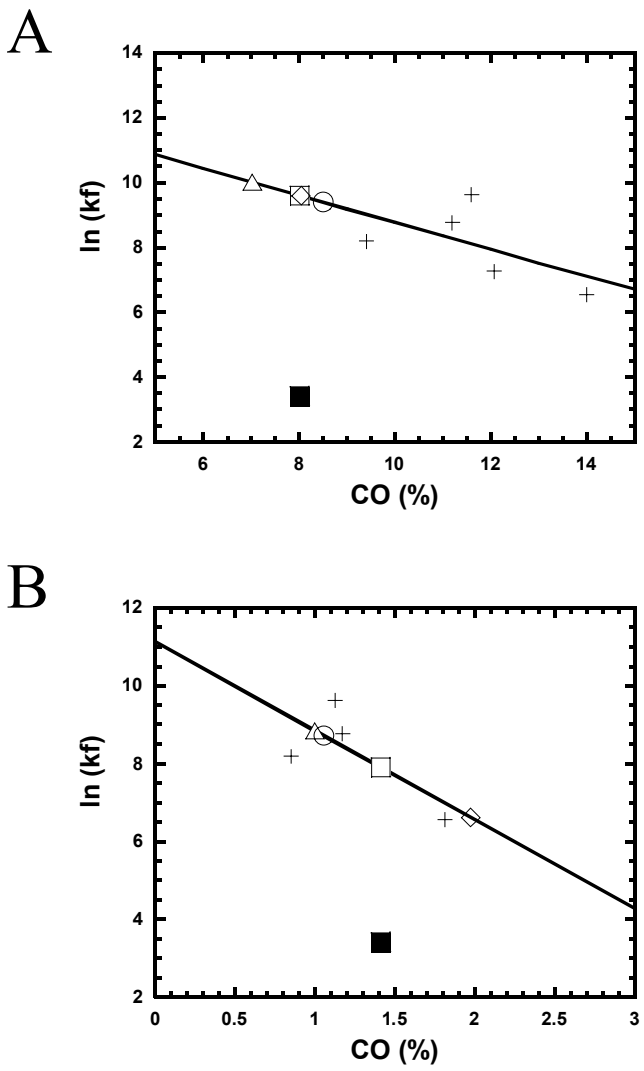


Figure 43. The folding rates of CARDs with other helical bundles predicted by (A) CO and (B) LRO. The plus symbols (+) are the helical proteins: cytochrome c (CO: 11.2%, LRO: 1.173), λ repressor (CO: 9.4%, LRO: 0.851), acyl-CoA binding protein (CO: 14%, LRO: 1.814), cytochrome c reduced form (CO: 11.6%, LRO: 1.126), and Immunity protein 9 (CO: 12.07%) [9; 98]. The solid lines are the fit to the published data. They are ln(k) = -0.4164 CO + 12.93 for Panel A and ln(k)=-2.29 LRO+11.13 for Panel B. The CO and LRO of RICK-CARD (□), Pro-1-CARD (△), ICEBERG (○), and APAF-1-CARD (◇) are plotted. The actual rate of the RICK-CARD is labeled as solid square (■).

6.3 The folding mechanism of α-helical Greek key

6.3.1 What is in Common?

From the studies of RICK-CARD and Pro-1-CARD, I found that they both adopt two state equilibrium folding mechanism. This is in agreement with most small proteins which has fewer than 100 residues. However, they do not adopt simple two-state folding kinetics as seen in many small proteins. Instead, several intermediates are present in both folding and unfolding. The only common feature in the folding kinetics is that they have kinetically trapped intermediates involved in unfolding and refolding. Although the time scales are different, the patterns are similar (Figure 44).

As shown in Figure 44, the half-time of the transition of Pro-1-CARD is close to that of the mutant of RICK-CARD. Since Pro-1-CARD does not contain a slow refolding phase that might be caused by proline isomerization, folding of Pro-1-CARD is likely to be independent of proline isomerization. It is possible that the folding kinetics of Pro-1-CARD contain some common folding steps with those of RICK-CARD such as a common kinetically trapped intermediate. The intermediate in refolding contains a half-time of ~30 to 50 sec in the absence of proline isomerization or slow steps of refolding. The intermediate in unfolding contains a half-time of ~10 to 20 sec when there are no other slow unfolding steps. The complex folding kinetics observed in the slow phases of RICK-CARD may be due to the proline residues or some local structural rearrangements that are only for RICK-CARD but not for all the CARD family members.

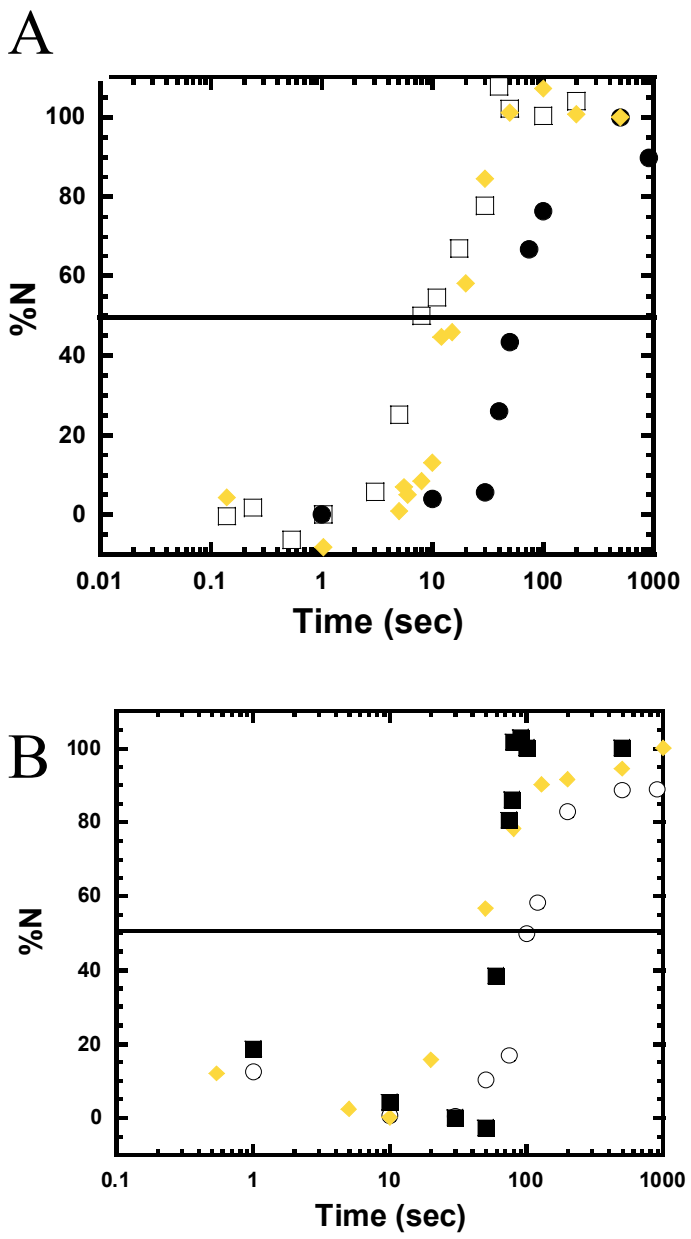


Figure 44. The comparison of native species formation in unfolding and refolding of Pro-1-CARD (square), RICK-CARD (circle), and RICK-CARD mutant P87A (diamond) from (A) double jump experiments and (B) interrupted refolding experiments.

6.3.2 What is different?

The stability of RICK-CARD and Pro-1-CARD are different. RICK-CARD is almost three times more stable than Pro-1-CARD. RICK-CARD is a well-packed six-helical bundle, however, Pro-1-CARD contains some unfolded protein in the native condition. Interestingly, by mutating some of the prolines in RICK-CARD, the equilibrium properties of the mutants of RICK-CARD, such as P47A, P47A/P87A, and P_3A, are similar to those of Pro-1-CARD. This might suggest that lacking proline residues in these positions could be a factor that contributes to the low stability of Pro-1-CARD. Further details on conformational effect by the proline mutants of RICK-CARD will be discussed in the following sections.

There are many differences in kinetics observed from Pro-1-CARD and RICK-CARD. There are no unfolding and refolding kinetic phases in Pro-1-CARD besides the kinetically trapped intermediates. Therefore, the major refolding phase of Pro-1-CARD should occur with a very fast rate, which is much faster than that, 30 s^{-1}, for RICK-CARD. The major unfolding phase of Pro-1-CARD could also be complete in the burst phase. Nevertheless, the unfolding of RICK-CARD contains additional three slower phases beside the major unfolding in the burst phase. Overall, Pro-1-CARD and RICK-CARD fold through different mechanisms but with some similar intermediates on the pathways.

The studies of Immunoglobulin-like fold proteins, which adopt β-Greek key fold [15], show that the unfolding rates are constant, and the protein stability correlates with the refolding rates. The β-Greek key proteins have common folding transition states that contain many native-contacts. However, it is not known whether the unfolding rates (k_{NU}) of

RICK-CARD and Pro-1-CARR are close. From the kinetic studies, we know that the major unfolding events of RICK-CARD and Pro-1-CARD both occur in burst phase, and the refolding rates are very different. Using the equation $\Delta G_{NU} = -RT \ln(k_{NU} / k_{UN})$ for RICK-CARD (ΔG_{NU}= 3 kcal/mol, k_{UN}=30 s^{-1}), we obtain an unfolding rate constant, k_{NU}, as ~0.19 s^{-1} and a half time as 4 sec. The number does not agree with the results that the major unfolding event of RICK-CARD occurs in the burst phase (< 5 msec). For Pro-1-CARD (ΔG_{NU}= 1.1 kcal/mol), neither refolding nor unfolding rates are known. Since the kinetics of CARDs are complicated, we cannot directly correlate the stability of the proteins to their kinetic properties.

6.4 The conformational stability of the CARD

From the proline mutants of RICK-CARD, we not only see the changes of the kinetic properties but also the changes of the structure and conformational stability. A surprising result was obtained by mutating the proline residue at position 85 (P85A). The mutation disrupts the native wild type structure and the protein loses the function of binding to Pro-1-CARD. However, P87A does not change much of the function and structure. P87A even contains more buried area meaning the protein is more compact. Mutations on both position 85 and 87 (P85A/P87A) do not change much the fact that the protein structure is disrupted and the protein loses the function. Here, I would like to discuss the possible factors on conformation and stability.

By labeling the hydrophobic residues of the structural models of RICK-CARD and Pro-1-CARD, we found a hydrophobic pocket in between helices-1 and –6 (Figure 45). Although there are several hydrophobic groups on the protein surface, this pocket seems to be the largest one. Also, In RICK-CARD, there is a bulky tryptophan in helix 1 pointing toward the pocket and there is a tyrosine sandwiched between proline 85 and 87. It is possible that the proline 85 of RICK-CARD orients the side-chains of the nearby hydrophobic residues, such as tyrosine 86, to a correct position so that the protein is stabilized by those contacts. There are methionine 81 and leucine 83 before proline 85. Position 87 does not change much of the protein conformation and does not affect the function. But P87A is more stable and more compact than the wild type. It is possible that by taking out the proline at position 87 increases the helical content of the protein and further enhances the stabilizing contacts. We also found that Pro-1-CARD contains a phenylalanine in helix-1 and a tyrosine in helix-6. But we do not know if they are involved

in stabilizing the protein. Maybe because there is no proline residue in Pro-1-CARD such as proline 85 of RICK-CARD to orient the side chains of the favorable contacts, the protein is not as stable as RICK-CARD.

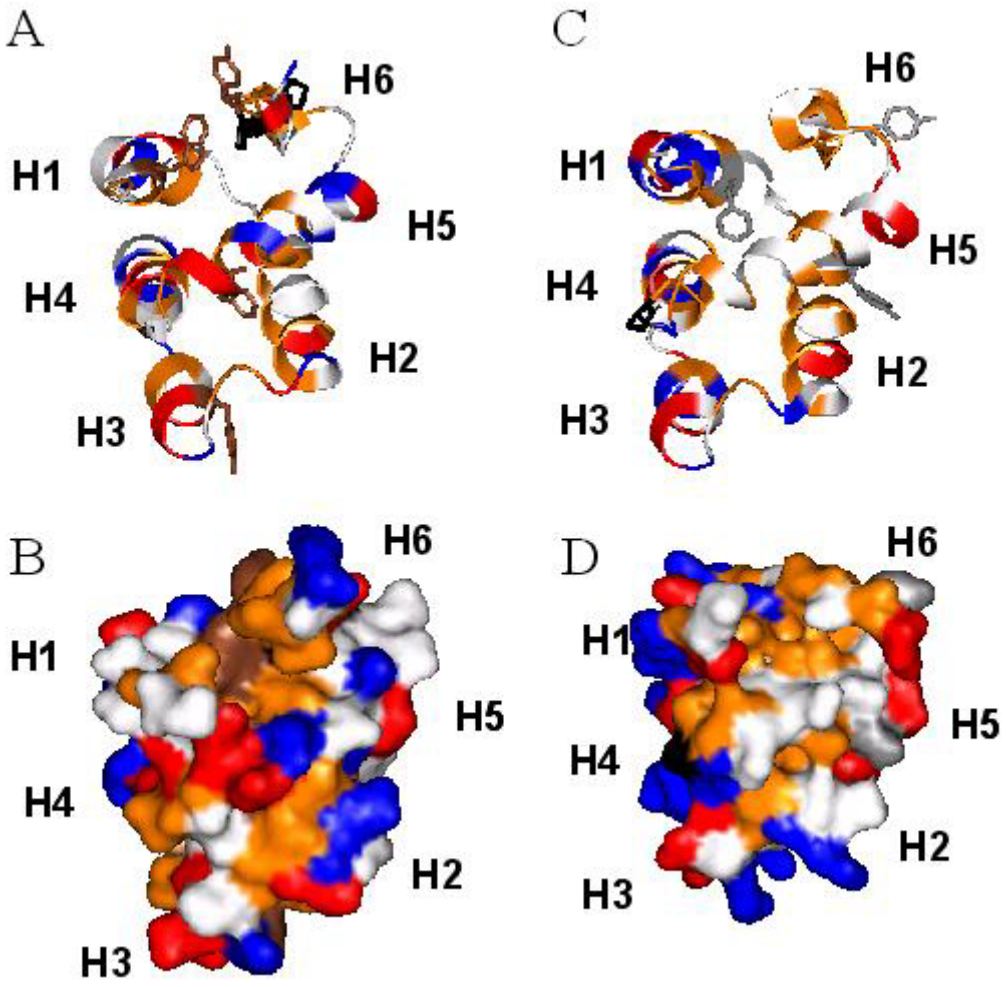


Figure 45. The hydrophobic areas of RICK-CARD (A, B) and Pro-1-CARD (C, D). Panel A and C are the homology models in cartoon and penal B and D show the protein surface. The hydrophobic residues are colored in orange. The basic and acid residues are colored in blue and red, respectively.

6.5 Conclusions

This is the first folding study on α-helical Greek key fold proteins. We contribute to the folding field with new mechanisms on death domain superfamily members. We have provided valuable information of folding of α-helical Greek key fold proteins, RICK-CARD and Pro-1-CARD. In addition, we also found critical positions to maintain the well-packed helical structure. According to the results and discussion of RICK-CARD, its proline mutants, and Pro-1-CARD, I summarize the conclusions as follows:

1. The equilibrium unfolding studies of RICK-CARD and Pro-1-CARD show a two-state mechanism (N ⇆ U).
2. Folding kinetics of RICK-CARD show that they are complex with several kinetically trapped, misfolded, intermediates present on both unfolding and refolding pathways.
3. Folding kinetics of Pro-1-CARD show that both unfolding and refolding are fast but with a kinetically trapped, misfolded, intermediate present. The folding kinetics of RICK-CARD and Pro-1-CARD are mainly different but may have similarity in the pathways.
4. Proline 85 of RICK-CARD is critical for maintaining the native conformation of the wild type protein.
5. Proline 87 of RICK-CARD contributes to the kinetic complexity observed in both refolding and unfolding. Proline 87 is also responsible for the slow refolding phase observed in the wild type RICK-CARD.

References

1. Radford, S. E. (2000). Protein folding: progress made and promises ahead. *Trends Biochem Sci* 25, 611-8.
2. Cohen, F. E. (1999). Protein misfolding and prion diseases. *J Mol Biol* 293, 313-20.
3. Bross, P., Corydon, T. J., Andresen, B. S., Jorgensen, M. M., Bolund, L. & Gregersen, N. (1999). Protein misfolding and degradation in genetic diseases. *Hum Mutat* 14, 186-98.
4. Anfinsen, C. B. (1973). Principles that govern the folding of protein chains. *Science* 181, 223-30.
5. Karplus, M. (1997). The Levinthal paradox: yesterday and today. *Fold Des* 2, S69-75.
6. Dill, K. A. & Chan, H. S. (1997). From Levinthal to pathways to funnels. *Nat Struct Biol* 4, 10-9.
7. Dobson, C. M. (2000). The nature and significance of protein folding. In *Mechanisms of protein folding* (Pain, R., ed.), pp. 1-33. Oxford.
8. Zhou, Y. & Karplus, M. (1999). Interpreting the folding kinetics of helical proteins. *Nature* 401, 400-3.
9. Plaxco, K. W., Simons, K. T., Ruczinski, I. & Baker, D. (2000). Topology, stability, sequence, and length: defining the determinants of two-state protein folding kinetics. *Biochemistry* 39, 11177-83.
10. Gunasekaran, K., Eyles, S. J., Hagler, A. T. & Gierasch, L. M. (2001). Keeping it in the family: folding studies of related proteins. *Curr Opin Struct Biol* 11, 83-93.
11. Ferguson, N., Li, W., Capaldi, A. P., Kleanthous, C. & Radford, S. E. (2001). Using chimeric immunity proteins to explore the energy landscape for alpha-helical protein folding. *J Mol Biol* 307, 393-405.

12. Ferguson, N., Capaldi, A. P., James, R., Kleanthous, C. & Radford, S. E. (1999). Rapid folding with and without populated intermediates in the homologous four-helix proteins Im7 and Im9. *J Mol Biol* 286, 1597-608.
13. Capaldi, A. P., Kleanthous, C. & Radford, S. E. (2002). Im7 folding mechanism: misfolding on a path to the native state. *Nat Struct Biol* 9, 209-16.
14. Perl, D., Welker, C., Schindler, T., Schroder, K., Marahiel, M. A., Jaenicke, R. & Schmid, F. X. (1998). Conservation of rapid two-state folding in mesophilic, thermophilic and hyperthermophilic cold shock proteins. *Nat Struct Biol* 5, 229-35.
15. Clarke, J., Cota, E., Fowler, S. B. & Hamill, S. J. (1999). Folding studies of immunoglobulin-like beta-sandwich proteins suggest that they share a common folding pathway. *Structure Fold Des* 7, 1145-53.
16. Chan, C. K., Hu, Y., Takahashi, S., Rousseau, D. L., Eaton, W. A. & Hofrichter, J. (1997). Submillisecond protein folding kinetics studied by ultrarapid mixing. *Proc Natl Acad Sci U S A* 94, 1779-84.
17. Mines, G. A., Pascher, T., Lee, S. C., Winkler, J. R. & Gray, H. B. (1996). Cytochrome c folding triggered by electron transfer. *Chem Biol* 3, 491-7.
18. Kragelund, B. B., Robinson, C. V., Knudsen, J., Dobson, C. M. & Poulsen, F. M. (1995). Folding of a four-helix bundle: studies of acyl-coenzyme A binding protein. *Biochemistry* 34, 7217-24.
19. Kragelund, B. B., Hojrup, P., Jensen, M. S., Schjerling, C. K., Juul, E., Knudsen, J. & Poulsen, F. M. (1996). Fast and one-step folding of closely and distantly related homologous proteins of a four-helix bundle family. *J Mol Biol* 256, 187-200.
20. Huang, G. S. & Oas, T. G. (1995). Structure and stability of monomeric lambda repressor: NMR evidence for two-state folding. *Biochemistry* 34, 3884-92.
21. Burton, R. E., Huang, G. S., Daugherty, M. A., Fullbright, P. W. & Oas, T. G. (1996). Microsecond protein folding through a compact transition state. *J Mol Biol* 263, 311-22.
22. Gorski, S. A., Capaldi, A. P., Kleanthous, C. & Radford, S. E. (2001). Acidic conditions stabilise intermediates populated during the folding of Im7 and Im9. *J Mol Biol* 312, 849-63.

23. Marianayagam, N. J., Khan, F., Male, L. & Jackson, S. E. (2002). Fast folding of a four-helical bundle protein. *J Am Chem Soc* 124, 9744-50.
24. Jackson, S. E. (1998). How do small single-domain proteins fold? *Fold Des* 3, R81-91.
25. Plaxco, K. W., Simons, K. T. & Baker, D. (1998). Contact order, transition state placement and the refolding rates of single domain proteins. *J Mol Biol* 277, 985-94.
26. Simeoni, F., Masotti, L. & Neyroz, P. (2001). Structural role of the proline residues of the beta-hinge region of p13suc1 as revealed by site-directed mutagenesis and fluorescence studies. *Biochemistry* 40, 8030-42.
27. del Camino, D., Holmgren, M., Liu, Y. & Yellen, G. (2000). Blocker protection in the pore of a voltage-gated K+ channel and its structural implications. *Nature* 403, 321-5.
28. McLachlan, A. D. (1980). Repeated folding pattern in copper-zinc superoxide dismutase. *Nature* 285, 267-8.
29. Boissinot, M., Karnas, S., Lepock, J. R., Cabelli, D. E., Tainer, J. A., Getzoff, E. D. & Hallewell, R. A. (1997). Function of the Greek key connection analysed using circular permutants of superoxide dismutase. *Embo J* 16, 2171-8.
30. Getzoff, E. D., Tainer, J. A., Stempien, M. M., Bell, G. I. & Hallewell, R. A. (1989). Evolution of CuZn superoxide dismutase and the Greek key beta-barrel structural motif. *Proteins* 5, 322-36.
31. Rajini, B., Shridas, P., Sundari, C. S., Muralidhar, D., Chandani, S., Thomas, F. & Sharma, Y. (2001). Calcium binding properties of gamma-crystallin: calcium ion binds at the Greek key beta gamma-crystallin fold. *J Biol Chem* 276, 38464-71.
32. Rudolph, R., Siebendritt, R., Nesslauer, G., Sharma, A. K. & Jaenicke, R. (1990). Folding of an all-beta protein: independent domain folding in gamma II-crystallin from calf eye lens. *Proc Natl Acad Sci U S A* 87, 4625-9.
33. Hamill, S. J., Cota, E., Chothia, C. & Clarke, J. (2000). Conservation of folding and stability within a protein family: the tyrosine corner as an evolutionary cul-de-sac. *J Mol Biol* 295, 641-9.

34. Capaldi, A. P., Ferguson, S. J. & Radford, S. E. (1999). The Greek key protein apo-pseudoazurin folds through an obligate on-pathway intermediate. *J Mol Biol* 286, 1621-32.
35. Jones, S., Reader, J. S., Healy, M., Capaldi, A. P., Ashcroft, A. E., Kalverda, A. P., Smith, D. A. & Radford, S. E. (2000). Partially unfolded species populated during equilibrium denaturation of the beta-sheet protein Y74W apo-pseudoazurin. *Biochemistry* 39, 5672-82.
36. Reader, J. S., Van Nuland, N. A., Thompson, G. S., Ferguson, S. J., Dobson, C. M. & Radford, S. E. (2001). A partially folded intermediate species of the beta-sheet protein apo-pseudoazurin is trapped during proline-limited folding. *Protein Sci* 10, 1216-24.
37. Richardson, J. S., Richardson, D. C., Tweedy, N. B., Gernert, K. M., Quinn, T. P., Hecht, M. H., Erickson, B. W., Yan, Y., McClain, R. D., Donlan, M. E. & et al. (1992). Looking at proteins: representations, folding, packing, and design. Biophysical Society National Lecture, 1992. *Biophys J* 63, 1185-209.
38. Zhang, C. & Kim, S. H. (2000). A comprehensive analysis of the Greek key motifs in protein beta-barrels and beta-sandwiches. *Proteins* 40, 409-419.
39. Hazes, B. & Hol, W. G. (1992). Comparison of the hemocyanin beta-barrel with other Greek key beta-barrels: possible importance of the "beta-zipper" in protein structure and folding. *Proteins* 12, 278-98.
40. Hemmingsen, J. M., Gernert, K. M., Richardson, J. S. & Richardson, D. C. (1994). The tyrosine corner: a feature of most Greek key beta-barrel proteins. *Protein Sci* 3, 1927-37.
41. Weber, C. H. & Vincenz, C. (2001). The death domain superfamily: a tale of two interfaces? *Trends Biochem Sci* 26, 475-81.
42. Fairbrother, W. J., Gordon, N. C., Humke, E. W., O'Rourke, K. M., Starovasnik, M. A., Yin, J. P. & Dixit, V. M. (2001). The PYRIN domain: a member of the death domain-fold superfamily. *Protein Sci* 10, 1911-8.
43. Hofmann, K. (1999). The modular nature of apoptotic signaling proteins. *Cell Mol Life Sci* 55, 1113-28.

44. Martin, S. J. (2001). Dealing the CARDs between life and death. *Trends Cell Biol* 11, 188-9.
45. Li, H. & Yuan, J. (1999). Deciphering the pathways of life and death. *Curr Opin Cell Biol* 11, 261-6.
46. Shi, Y. (2002). Mechanisms of caspase activation and inhibition during apoptosis. *Mol Cell* 9, 459-70.
47. Shiozaki, E. N., Chai, J. & Shi, Y. (2002). Oligomerization and activation of caspase-9, induced by Apaf-1 CARD. *Proc Natl Acad Sci U S A* 99, 4197-202.
48. Qin, H., Srinivasula, S. M., Wu, G., Fernandes-Alnemri, T., Alnemri, E. S. & Shi, Y. (1999). Structural basis of procaspase-9 recruitment by the apoptotic protease-activating factor 1. *Nature* 399, 549-57.
49. Lauber, K., Appel, H. A., Schlosser, S. F., Gregor, M., Schulze-Osthoff, K. & Wesselborg, S. (2001). The adapter protein apoptotic protease-activating factor-1 (Apaf-1) is proteolytically processed during apoptosis. *J Biol Chem* 276, 29772-81.
50. Beere, H. M., Wolf, B. B., Cain, K., Mosser, D. D., Mahboubi, A., Kuwana, T., Tailor, P., Morimoto, R. I., Cohen, G. M. & Green, D. R. (2000). Heat-shock protein 70 inhibits apoptosis by preventing recruitment of procaspase-9 to the Apaf-1 apoptosome. *Nat Cell Biol* 2, 469-75.
51. McCarthy, J. V., Ni, J. & Dixit, V. M. (1998). RIP2 is a novel NF-kappaB-activating and cell death-inducing kinase. *J Biol Chem* 273, 16968-75.
52. Inohara, N., del Peso, L., Koseki, T., Chen, S. & Nunez, G. (1998). RICK, a novel protein kinase containing a caspase recruitment domain, interacts with CLARP and regulates CD95-mediated apoptosis. *J Biol Chem* 273, 12296-300.
53. Thome, M., Hofmann, K., Burns, K., Martinon, F., Bodmer, J. L., Mattmann, C. & Tschopp, J. (1998). Identification of CARDIAK, a RIP-like kinase that associates with caspase-1. *Curr Biol* 8, 885-8.
54. Kobayashi, K., Inohara, N., Hernandez, L. D., Galan, J. E., Nunez, G., Janeway, C. A., Medzhitov, R. & Flavell, R. A. (2002). RICK/Rip2/CARDIAK mediates signalling for receptors of the innate and adaptive immune systems. *Nature* 416, 194-9.

55. Jeyaseelan, S. (2002). Rip2: a novel therapeutic target for bacteria-induced inflammation? *Trends Microbiol* 10, 356.

56. Zhu, S. G., Sheng, J. G., Jones, R. A., Brewer, M. M., Zhou, X. Q., Mrak, R. E. & Griffin, W. S. (1999). Increased interleukin-1beta converting enzyme expression and activity in Alzheimer disease. *J Neuropathol Exp Neurol* 58, 582-7.

57. Lee, S. H., Stehlik, C. & Reed, J. C. (2001). Cop, a caspase recruitment domain-containing protein and inhibitor of caspase-1 activation processing. *J Biol Chem* 276, 34495-500.

58. Chen, Y. R. & Clark, A. C. (2003). Equilibrium and Kinetic Folding of a alpha-Helical Greek Key Protein Domain: Caspase Recruitment Domain (CARD) of RICK. *Biochemistry* 42, 6310-20.

59. Pace, C. N., Shirley, B., and Thompson, J. (1989). In *Protein Structure, A Practical Approach* (Creighton, T. E., ed.), pp. 311-330. IRL Press, New York.

60. Edelhoch, H. (1967). Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry* 6, 1948-54.

61. Humke, E. W., Shriver, S. K., Starovasnik, M. A., Fairbrother, W. J. & Dixit, V. M. (2000). ICEBERG: a novel inhibitor of interleukin-1beta generation. *Cell* 103, 99-111.

62. Royer, C. A., Mann, C. J. & Matthews, C. R. (1993). Resolution of the fluorescence equilibrium unfolding profile of trp aporepressor using single tryptophan mutants. *Protein Sci* 2, 1844-52.

63. Santoro, M. M. & Bolen, D. W. (1988). Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants. *Biochemistry* 27, 8063-8.

64. Schmid, F. X. (1986). Fast-folding and slow-folding forms of unfolded proteins. *Methods Enzymol* 131, 70-82.

65. Schmid, F. X. (1983). Mechanism of folding of ribonuclease A. Slow refolding is a sequential reaction via structural intermediates. *Biochemistry* 22, 4690-6.

66. Pappenberger, G., Aygun, H., Engels, J. W., Reimer, U., Fischer, G. & Kiefhaber, T. (2001). Nonprolyl cis peptide bonds in unfolded proteins cause complex folding kinetics. *Nat Struct Biol* 8, 452-8.
67. Druilhe, A., Srinivasula, S. M., Razmara, M., Ahmad, M. & Alnemri, E. S. (2001). Regulation of IL-1beta generation by Pseudo-ICE and ICEBERG, two dominant negative caspase recruitment domain proteins. *Cell Death Differ* 8, 649-57.
68. Zhou, P., Chou, J., Olea, R. S., Yuan, J. & Wagner, G. (1999). Solution structure of Apaf-1 CARD and its interaction with caspase-9 CARD: a structural basis for specific adaptor/caspase interaction. *Proc Natl Acad Sci U S A* 96, 11265-70.
69. Vaughn, D. E., Rodriguez, J., Lazebnik, Y. & Joshua-Tor, L. (1999). Crystal structure of Apaf-1 caspase recruitment domain: an alpha-helical Greek key fold for apoptotic signaling. *J Mol Biol* 293, 439-47.
70. Day, C. L., Dupont, C., Lackmann, M., Vaux, D. L. & Hinds, M. G. (1999). Solution structure and mutagenesis of the caspase recruitment domain (CARD) from Apaf-1. *Cell Death Differ* 6, 1125-32.
71. Chou, J. J., Matsuo, H., Duan, H. & Wagner, G. (1998). Solution structure of the RAIDD CARD and model for CARD/CARD interaction in caspase-2 and caspase-9 recruitment. *Cell* 94, 171-80.
72. Jeong, E. J., Bang, S., Lee, T. H., Park, Y. I., Sim, W. S. & Kim, K. S. (1999). The solution structure of FADD death domain. Structural basis of death domain interactions of Fas and FADD. *J Biol Chem* 274, 16337-42.
73. Eberstadt, M., Huang, B., Chen, Z., Meadows, R. P., Ng, S. C., Zheng, L., Lenardo, M. J. & Fesik, S. W. (1998). NMR structure and mutagenesis of the FADD (Mort1) death-effector domain. *Nature* 392, 941-5.
74. Xiao, T., Towb, P., Wasserman, S. A. & Sprang, S. R. (1999). Three-dimensional structure of a complex between the death domains of Pelle and Tube. *Cell* 99, 545-55.
75. Alonso, D. O. & Dill, K. A. (1991). Solvent denaturation and stabilization of globular proteins. *Biochemistry* 30, 5974-85.
76. Schellman, J. A. (1978). Solvent Denaturation. *Biopolymers* 17, 1305-1322.

77. Myers, J. K., Pace, C. N. & Scholtz, J. M. (1995). Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci* 4, 2138-48.
78. Perl, D. & Schmid, F. X. (2001). Electrostatic stabilization of a thermophilic cold shock protein. *J Mol Biol* 313, 343-57.
79. Nishimura, C., Uversky, V. N. & Fink, A. L. (2001). Effect of salts on the stability and folding of staphylococcal nuclease. *Biochemistry* 40, 2113-28.
80. Shaw, K. L., Grimsley, G. R., Yakovlev, G. I., Makarov, A. A. & Pace, C. N. (2001). The effect of net charge on the solubility, activity, and stability of ribonuclease Sa. *Protein Sci* 10, 1206-15.
81. Lee, K. K., Fitch, C. A. & Garcia-Moreno, E. B. (2002). Distance dependence and salt sensitivity of pairwise, coulombic interactions in a protein. *Protein Sci* 11, 1004-16.
82. Baldwin, R. L. (1996). How Hofmeister ion interactions affect protein stability. *Biophys J* 71, 2056-63.
83. Record, M. T., Jr., Zhang, W. & Anderson, C. F. (1998). Analysis of effects of salts and uncharged solutes on protein and nucleic acid equilibria and processes: a practical guide to recognizing and interpreting polyelectrolyte effects, Hofmeister effects, and osmotic effects of salts. *Adv Protein Chem* 51, 281-353.
84. Dominy, B. N., Perl, D., Schmid, F. X. & Brooks, C. L., 3rd. (2002). The effects of ionic strength on protein stability: the cold shock protein family. *J Mol Biol* 319, 541-54.
85. Jackson, S. E. & Fersht, A. R. (1991). Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition. *Biochemistry* 30, 10428-35.
86. Chen, B. L., Baase, W. A., Nicholson, H. & Schellman, J. A. (1992). Folding kinetics of T4 lysozyme and nine mutants at 12 degrees C. *Biochemistry* 31, 1464-76.
87. Kiefhaber, T., Quaas, R., Hahn, U. & Schmid, F. X. (1990). Folding of ribonuclease T1. 1. Existence of multiple unfolded states created by proline isomerization. *Biochemistry* 29, 3053-61.

88. Brandts, J. F., Halvorson, H. R. & Brennan, M. (1975). Consideration of the Possibility that the slow step in protein denaturation reactions is due to cis-trans isomerism of proline residues. *Biochemistry* 14, 4953-63.
89. Kiefhaber, T. (1995). Protein folding kinetics. *Methods Mol Biol* 40, 313-41.
90. Munoz, V. & Serrano, L. (1994). Elucidating the folding problem of helical peptides using empirical parameters. *Nat Struct Biol* 1, 399-409.
91. Baldwin, R. L. (1996). On-pathway versus off-pathway folding intermediates. *Fold Des* 1, R1-8.
92. Creighton, T. E. (1994). The energetic ups and downs of protein folding. *Nat Struct Biol* 1, 135-8.
93. Heidary, D. K., Gross, L. A., Roy, M. & Jennings, P. A. (1997). Evidence for an obligatory intermediate in the folding of interleukin-1 beta. *Nat Struct Biol* 4, 725-31.
94. Hosszu, L. L., Craven, C. J., Parker, M. J., Lorch, M., Spencer, J., Clarke, A. R. & Waltho, J. P. (1997). Structure of a kinetic protein folding intermediate by equilibrium amide exchange. *Nat Struct Biol* 4, 801-4.
95. Schiffmann, D. A., White, J. H., Cooper, A., Nutley, M. A., Harding, S. E., Jumel, K., Solari, R., Ray, K. P. & Gay, N. J. (1999). Formation and biochemical characterization of tube/pelle death domain complexes: critical regulators of postreceptor signaling by the Drosophila toll receptor. *Biochemistry* 38, 11722-33.
96. Satumba, W. J. & Mossing, M. C. (2002). Folding and assembly of lambda Cro repressor dimers are kinetically limited by proline isomerization. *Biochemistry* 41, 14216-24.
97. Lacroix, E., Viguera, A. R. & Serrano, L. (1998). Elucidating the folding problem of alpha-helices: local motifs, long-range electrostatics, ionic-strength dependence and prediction of NMR parameters. *J Mol Biol* 284, 173-91.
98. Gromiha, M. M. & Selvaraj, S. (2001). Comparison between long-range interactions and contact order in determining the folding rate of two-state proteins: application of long-range order to folding rate prediction. *J Mol Biol* 310, 27-32.

APPENDICES

Appendix A

Enzyme activity of mature caspase –3

Foreword

In Appendix A are details of oligomerization, proten stability, and spectrascopic properties of procaspase-3 and the role of prodomain in the pro-enzyme. The results are published, and I am the second author of the paper. The paper is attached in this section.

For this paper, I have cloned the pro-less variant (P2901), which contains the full-length caspase-3 without the prodomain and have accomplished the purification of mature caspase-3 (see Methods in the paper). I have done enzyme assay of caspase-3 (Figure 7 in the paper) and helped with labeling the peptide. I obtained caspase-3 by purifying the individual subunits (the large and small subunits) and refolded them together. The catalytic ability of caspase-3 is comparable to the published value. I examined the inhibition of caspase-3 by a peptide inhibitor by fluorescence and further examine the inhibition of a wild type pro-peptide and a mutant pro-peptide.

Removal of the Pro-Domain Does Not Affect the Conformation of the Procaspase-3 Dimer†

Cristina Pop,‡ Yun-Ru Chen,‡ Brandye Smith,§ Kakoli Bose,‡ Benjamin Bobay,‡ Ashutosh Tripathy,|| Stefan Franzen,§ and A. Clay Clark*,‡

Department of Molecular and Structural Biochemistry and Department of Chemistry, North Carolina State University, Raleigh, North Carolina 27695, and Macromolecular Interactions Facility, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599

Received May 21, 2001; Revised Manuscript Received August 13, 2001

ABSTRACT: We have investigated the oligomeric properties of procaspase-3 and a mutant that lacks the pro-domain (called pro-less variant). In addition, we have examined the interactions of the 28 amino acid pro-peptide when added in trans to the pro-less variant. By sedimentation equilibrium studies, we have found that procapase-3 is a stable dimer in solution at 25 °C and pH 7.2, and we estimate an upper limit for the equilibrium dissociation constant of ∼50 nM. Considering the expression levels of caspase-3 in Jurkat cells, we predict that procaspase-3 exists as a dimer in vivo. The pro-less variant is also a dimer, with little apparent change in the equilibrium dissociation constant. Thus, in contrast with the long pro-domain caspases, the pro-peptide of caspase-3 does not appear to be involved in dimerization. Results from circular dichroism, fluorescence anisotropy, and FTIR studies demonstrate that the pro-domain interacts weakly with the pro-less variant. The data suggest that the pro-peptide adopts a β-structure when in contact with the protein, but it is a random coil when free in solution. In addition, when added in trans, the pro-peptide does not inhibit the activity of the mature caspase-3 heterotetramer. On the other hand, the active caspase-3 does not efficiently hydrolyze the pro-domain at the NSVD^9 sequence as occurs when the pro-peptide is in cis to the protease domain. Based on these results, we propose a model for maturation of the procaspase-3 dimer.

Caspase activation, more than any other event, defines a cellular response to apoptosis [see review by Earnshaw (*1*)]. Caspases are a family of cysteinyl proteases that have an unusual requirement for aspartate at the P1 position in the substrate. Although many proteins are cleaved during apoptosis, caspases cleave key structural components of the cytoskeleton and nucleus as well as numerous proteins in signaling pathways, DNA repair or degrading enzymes, and transcription factors (*1*). In addition, caspases degrade amyloid precursor protein, presenilins, tau, and huntington, suggesting that they may have a role in neurodegenerative diseases (*2*). In short, caspase activation results in an irreversible, organized dismantling of the cell.

Currently, 14 caspases have been identified (*3*), and it is generally accepted that caspase-3 (CPP32, apopain, Yama) is the primary executioner during apoptosis (*4*, *5*). The mature, enzymatically active caspase-3 is a tetramer with a M_r of approximately 60 000 and consists of two copies of one large (α) and one small (β) subunit, described as dimers of heterodimers, $(\alpha\beta)_2$ (*6*, *7*). Interestingly, the two subunits in the ($\alpha\beta$) heterodimer form a single domain, which consists of a six-stranded β-sheet core flanked by α-helices. Each ($\alpha\beta$) heterodimer contains an active site, although the heterodimer appears to be inactive (*8*). While the catalytic dyad (Cys163, His121) is contributed by the large subunit, both subunits contribute residues that form the S1−S4 binding pockets (*7*). All other caspases for which the crystal structure is known (caspases-1, -7, and -8) display a similar fold (*9*−*11*), demonstrating the structural homology within the family.

Like the other caspases, caspase-3 exists in normal cells as an inactive zymogen and must be activated by proteolytic processing. While the three-dimensional structures have not been determined, procaspases are organized with an amino-terminal pro-domain, which varies in size and function, followed by the large subunit, an optional intersubunit linker, and the small subunit, as shown in Figure 1. Although processing appears to be more complicated for long pro-domain caspases, in general, caspases are activated in a two-step sequential process (*12*). For procaspase-3, the first cleavage occurs at Asp175 (*13*), which separates the covalent linkage between the large and small subunits. The small subunit remains associated through noncovalent interactions. The cleavage at Asp175 generally occurs by an upstream activator caspase. Based on the increase in enzyme activity (*14*) as well as other biochemical data (*15*, *16*), significant conformational rearrangements must occur as a result of the first cleavage. In the second step, the pro-domain is auto-

† This work was supported by a grant from the American Diabetes Association.

* To whom correspondence should be addressed at the Department of Molecular and Structural Biochemistry, 128 Polk Hall, North Carolina State University, Raleigh, NC 27695-7622. Phone: (919) 515-5805. Fax: (919) 515-2047. E-mail: clay_clark@ncsu.edu.

‡ Department of Molecular and Structural Biochemistry, North Carolina State University.

§ Department of Chemistry, North Carolina State University.

|| Macromolecular Interactions Facility, University of North Carolina at Chapel Hill.

10.1021/bi011037e CCC: $20.00
Published on Web 10/30/2001

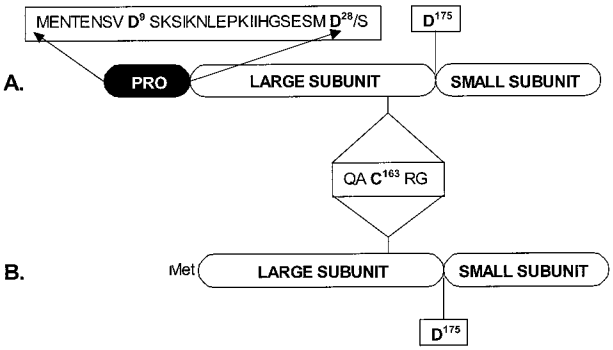


FIGURE 1: Procaspase-3 and pro-less variant. (A) Schematic representation of full-length procaspase-3, showing the pro-domain sequence, the conserved catalytic motif, and the cleavage sites, D9, D28, and D175. (B) Procaspase-3 lacking the 28 amino acid propeptide, referred to in the text as 'pro-less variant'.

catalytically removed from the large subunit (*13*) by a rapid proteolysis at Asp9 (see Figure 1) followed by a slow cleavage at Asp28 (*14*). The activity of the protein following cleavage at Asp9 ($NSVD^{9}/S$) appears indistinguishable from that of the mature heterotetramer (*14*), in which the N-terminus begins at Ser29 ($ESMD^{28}/S$). The significance of the two cleavage sites within the pro-domain is not yet understood, but many of the procaspases contain more than one processing site in the pro-domain (*17*).

In the case of procaspase-1 and other long pro-domain caspases, the pro-domain (119 amino acids) is involved in dimerization of the procaspase, leading to autoprocessing. This is due to the presence of a dimerization motif within the pro-domain. Procaspase-1, for example, contains a caspase recruitment domain (CARD) within the first 92 amino acids of the pro-domain. The CARD is a six-helix bundle folding motif and is responsible for highly specific interactions between the procaspase and homologous CARD motifs in adapter proteins (*18*).

Dimerization represents an early event in maturation of procaspase-1 (*19*). For downstream caspases, such as caspase-3, there is no evidence that dimerization is an early event. Indeed, procaspase-3 is frequently drawn as a monomer in the literature reviews (*20*), although the oligomeric properties have not been examined. Thus, the role of the pro-domain, if any, in dimerization and maturation of the short pro-domain procaspases is less clear. In addition, the pro-domain (28 amino acids) does not inhibit the protease activity when present in cis of the mature heterotetramer (*14*), suggesting the possibility that the short pro-domain of executioner caspases represents a vestigial linker between the CARD motif and the amino terminus of the large subunit. Indeed, procaspase-1 contains 27 amino acids between the CARD motif and the large subunit. However, recent data argue that the short pro-domain may also have a function in maturation. Recombinant procaspase-3 has been shown to autoprocess during overexpression in *E. coli*, yeast, and insect cells or during purification after overexpression in mammalian cells (*7*, *21*−*23*), suggesting that the pro-enzyme forms oligomers at high concentrations, facilitating autoprocessing. Recently, Meergans and co-workers (*24*) demonstrated that caspase-3 constructs lacking the pro-domain were able to autoprocess in HeLa cells, in contrast with the intact precursors. Thus, the pro-sequence may not inhibit the activity of the active protease, but rather may function somehow as an additional control to decrease the low levels of procaspase activity, preventing spontaneous activation.

Because of these considerations, we have investigated the oligomeric state of procaspase-3 and the role of the pro-domain on procaspase-3 stability and caspase-3 activity. We have used a catalytically inactive mutant (C163S) of the full-length human procaspase-3 as well as a variant of procaspase-3(C163S) in which the pro-domain is removed, referred to as the pro-less variant[1] (see Figure 1). We show here that procaspase-3(C163S) is a dimer in solution at 25 °C, and based on these data as well as estimations of the procaspase-3 concentration in vivo, we suggest that the protein is a dimer in vivo. Removal of the pro-domain has no effect on the oligomeric or conformational properties of procaspase-3. In addition, we demonstrate that the pro-domain interacts with the pro-less variant in trans, suggesting that the pro-peptide may bind weakly to the procaspase dimer at site(s) away from the active site region.

MATERIALS AND METHODS

Chemicals. Ampicillin, antifoam-C, bovine serum albumin, carbonic anhydrase, CHAPS, cytochrome *c*, DEAE-Sepharose, dansyl chloride, DMF, DMSO, DTT, EDTA, EGTA, glycerol, IPTG, kanamycin, nickel sulfate, PMSF, monobasic and dibasic potassium phosphate, Sephacryl-S15, Sephacryl-S100, sodium bicarbonate, TLCK, TPCK, and Trizma base were from Sigma. Ultrapure ammonium sulfate, guanidine hydrochloride, imidazole, and urea were from ICN. Sodium chloride was from Fisher. Tryptone and yeast extract were from Difco. His-bind resin was from Novagen. Potassium chloride and sucrose were from Mallinckrodt. Acetonitrile and HEPES were from Acros. α-Cyano-4-hydroxycinnamic acid was from Aldrich.

Plasmid Construction. The procaspase-3 gene was amplified by PCR from pET21b-CPP32 (*25*), kindly provided by Dr. Emad Alnemri, using the primers HC3P32F (GTCGCGGATCATATGGAGAACACTG) and HC3P12R (GTGGTGGTGGTGCTCGAGGTG). This introduced an *Nde*I site at the 5′ end of the gene and an *Xho*I site at the 3′ end of the gene. The amplified gene product was inserted into pET21b that had been digested with *Nde*I and *Xho*I. This strategy removed 14 amino acids at the amino terminus of procaspase-3 that arise from the vector in pET21b-CPP32. The resulting plasmid, pHC332, produces procaspase-3 with the correct amino terminus. The active site cysteine (Cys163) was mutated to serine using QuickChange site-directed mutagenesis kit (Stratagene), with the primers HCP3CS1 (5′-CATTATTCAGGCCT**C**CCGCGGTACAGAACTGGACTGTGG-3′) and HCP3CS2 (5′-CAGTTCTGTACCGCGG**G**AGGCCTGAATAATGAAAAGTTTGG-3′), and plasmid pET21b-CPP32. This strategy also introduced a unique *Sac*II

[1] Abbreviations: Pro-less variant, procaspase-3 (C163S) lacking the pro-domain; CHAPS, 3-[(3-chloroamidopropyl)dimethylammonio]-1-propanesulfonate; DMSO, dimethyl sufoxide; DMF, *N*,*N*-dimethylformamide; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; IPTG, isopropyl β-D-thiogalactopyranoside; PMSF, phenylmethylsulfonyl fluoride; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; TFA, trifluoroacetic acid; TLCK, *N*-α-*p*-tosyl-L-lysine chloromethyl ketone; TPCK, *N*-tosyl-L-phenylalanine chloromethyl ketone; Z-VAD-FMK, *N*-benzyloxycarbonyl-Val-Ala-Asp-fluoromethyl ketone; Z-DEVD-AFC, *N*-benzyloxycarbonyl-Asp-Glu-Val-Asp-7-amino-4-trifluoromethyl coumarin.

site (underlined) downstream of the C163S mutation (shown in boldface type). Plasmids were first screened by digestion with *Sac*II, and positive clones were sequenced to confirm the mutation. The mutated gene was cloned into pET21b, as described above, to produce the plasmid pHC33201.

Plasmid pHC32901 was constructed by subcloning the DNA for the procaspase-3 large and small subunits from pHC33201. The primers for PCR amplification were HC3P17F (5′-GCGAATCACATATGTCTGGAATATCCC-3′) and HC3P12R (5′-GTGGTGGTGGTGCTCGAGGTG-3′), generating *Nde*I and *Xho*I sites at the 5′ and 3′ ends, respectively. The ∼750 bp fragment was inserted into pET21b digested with *Nde*I and *Xho*I. The resulting proteins [pro-caspase-3(C163S) and pro-less variant] have carboxyl termini consisting of the sequence Leu-Glu-His_6 that arise from the vector.

Plasmids pHC317 and pHC312, which harbor the genes for the caspase-3 large and small subunits, respectively, were constructed by subcloning the DNA corresponding to each subunit from pHC332, described above, into pET21b. The PCR primers for the large subunit were HC3P17F and HC3P17R (5′-CATCATCAACCTCGAGGTCTGTCTC-3′), whereas those for the small subunit were HC3P12F (5′-GCATTGAGCATATGAGTGGTGTTGATG-3′) and HC3P12R. In both cases, *Nde*I and *Xho*I sites were introduced at the 5′ and 3′ ends, respectively. All constructs were sequenced (both strands) in order to confirm the correct sequence.

Protein Purification. All steps were performed at 4 °C unless otherwise noted. The purification protocols are modifications of those described previously (*26*). In separate experiments, human procaspase-3(C163S) and the pro-less variant were purified as C-terminal-$(His)_6$-tagged proteins from *E. coli* BL21(DE3) harboring the plasmid pHC33201 or pHC32901, respectively. The protocols for purification of both proteins were similar. Cells were grown in Fernbach flasks containing 1 L of LB media with 50 μg/mL ampicillin and 0.003% antifoam-C at 30 °C. When the cultures reached an A_{600} of ∼1.2, protein expression was induced by the addition of IPTG to a final concentration of 0.1 mM. The cells were harvested after ∼16 h by centrifugation at 10000*g* for 15 min (GS-3 rotor). The bacterial pellets were resuspended in buffer A (50 mM Tris-HCl, pH 7.9, 100 μg/mL PMSF, 50 μg/mL TLCK, 100 μg/mL TPCK) (∼10 mL/L of culture) and lysed on ice using a French pressure cell (16 000 psi). The supernatant was separated from cell debris by centrifugation at 28000*g* for 30 min (SA-600 rotor). The pellet from this step was washed once with buffer A (5 mL/L of culture) and centrifuged for 30 min. The resulting supernatant was combined with the first. The proteins were fractionated between 30% and 80% ammonium sulfate, centrifuged at 22000*g* for 15 min, resuspended in buffer B (20 mM Tris-HCl, pH 7.9, 500 mM NaCl, 100 μg/mL PMSF, 50 μg/mL TLCK, 100 μg/mL TPCK) containing 5 mM imidazole, and dialyzed against the same buffer (2 × 80 volumes). The samples were then batch-bound for 2 min to His-bind resin (5 mL) equilibrated in buffer B containing 5 mM imidazole, and the resin was loaded onto a column (2 cm diameter). The resin was washed and the protein was eluted using step gradients of imidazole (20, 40, 100, 150, 250, and 500 mM imidazole in buffer B, 50 mL per step). The fractions were analyzed by SDS−PAGE (4−20% gradient gels) (*27*). For procaspase-3(C163S), the fractions containing the protein (150−250 mM imidazole) were pooled, concentrated (YM10 membrane), and dialyzed first against buffer C (50 mM Tris-HCl, pH 7.9, 250 mM NaCl, 1 mM EDTA, 1 mM EGTA) and then against buffer D (50 mM KH_2PO_4/K_2HPO_4, pH 7.5, 1 mM DTT, 1 mM EDTA, 1 mM EGTA) (2 × 80 volumes for both buffers). For the pro-less variant, the fractions from the His-bind resin containing the protein (100−250 mM imidazole) were pooled, concentrated (YM10 membrane), and dialyzed first against buffer E (50 mM KH_2PO_4/K_2HPO_4, pH 7.9, 250 mM KCl, 1 mM EDTA, 1 mM EGTA) and then buffer F (50 mM KH_2PO_4/K_2HPO_4, pH 7.5, 25 mM KCl, 1 mM DTT, 1 mM EDTA, 1 mM EGTA), respectively (2 × 80 volumes for both buffers). In each case, the sample was applied to a DEAE-Sepharose column (4 cm × 18 cm) that had been equilibrated with buffers D [procaspase-3(C163S)] or F (pro-less variant), respectively. The proteins were eluted at a flow rate of 4 mL/min with a linear gradient of 0−400 mM KCl [procaspase-3(C163S)] or 25−250 mM KCl (pro-less variant). Each fraction was tested using a mini-Bradford assay (*28*), and the positive fractions were analyzed by SDS−PAGE (4−20% gradient gels). The fractions (75−125 mM KCl) containing procaspase-3(C163S) or the pro-less variant (100−150 mM KCl) were pooled, concentrated, and dialyzed overnight against buffer D or buffer F, respectively. The proteins were stored at −80 °C. The protein purity was greater than 95% as assessed by SDS−PAGE.

The concentrations of procaspase-3(C163S) and the pro-less variant were determined using $\epsilon_{280} = 26\,500\ M^{-1}\ cm^{-1}$ and $\epsilon_{280} = 25\,300\ M^{-1}\ cm^{-1}$, respectively. The extinction coefficients were determined by the method of Edelhoch (*29*) and are in good agreement with that determined previously for procaspase-3 (*30*). The concentrations shown here are those of the monomers.

To generate the active caspase-3 heterotetramer, the subunits were purified separately and refolded together. Briefly, *E. coli* BL21(DE3) containing either pHC317 or pHC312 were grown at 37 °C in LB media. When the cultures reached an A_{600} of ∼1.2, protein expression was induced by the addition of IPTG to a final concentration of 0.1 mM. After 4 h of induction, the cells were harvested, resuspended in lysis buffer (20 mM Tris, pH 7.9, 500 mM NaCl), and broken in a French pressure cell as described above. The samples were centrifuged at 14 000 rpm for 30 min, and the insoluble portion was washed extensively with lysis buffer and dissolved in lysis buffer containing 6 M guanidine hydrochloride. The proteins were filtered (0.22 μm membrane) and purified by HPLC using a preparative C8 guard column (9.4 mm × 15 mm) and a preparative C18 column (21.2 mm × 250 mm; 300SB-C18) (Agilent Technologies). The flow rate was 1 mL/min. The column was developed with a linear gradient of 5% acetonitrile, 0.1% TFA to 85% acetonitrile, 0.1% TFA. The proteins eluted at 48% acetonitrile. The acetonitrile was removed by vacuum (1 h in a speedvac), and the proteins were dissolved in 20 mM Tris-HCl, pH 7.0, containing 6 M guanidine hydrochloride. The two purified subunits were mixed and refolded by rapid dilution in refolding buffer [100 mM HEPES, pH 7.5, 10% (w/v) sucrose, 0.1% (w/v) CHAPS, 0.5 M NaCl, 10 mM DTT] at 25 °C, as described previously (*6*). The final protein concentration was 8 μg/mL. To purify the

heterotetramer, the refolded protein was eluted from a size exclusion column (1 cm × 14 cm; Sephacryl S-100) with 50 mM KH_2PO_4/K_2HPO_4, pH 7.2, containing 1 mM DTT.

The active site concentration of caspase-3 was determined by enzymatic assay using the tetrapeptide substrate Z-DEVD-AFC (nonmethylated form) (Alexis) and titration with the irreversible inhibitor Z-VAD-FMK (CalBiochem), as described previously (*26*, *31*).

Analytical Ultracentrifugation. Sedimentation equilibrium experiments were performed at 25 °C in a Beckman XL-A ultracentrifuge equipped with absorbance optics and a four-hole AnTi60 rotor. The proteins were dialyzed at 4 °C into 50 mM KH_2PO_4/K_2HPO_4, pH 7.2, with either 1 mM DTT (280 nm data) or 0.05 mM DTT (230 nm data). The samples were equilibrated at three rotor speeds (14 000, 18 000, and 24 000 rpm), and the absorbance was measured at 280 or 230 nm. For the 280 nm data, the protein concentrations were as follows: (a) procaspase-3, 3.5, 9.4, and 18.9 μM; (b) pro-less variant, 1.26, 4.9, and 12.9 μM. For 230 nm, the protein concentrations were the following: (a) procaspase-3, 0.45, 0.78, 1.1, 3, and 4.1 μM; (b) pro-less form, 0.75, 1.25, 1.85, 2.5, and 3.7 μM. The experimental data were fit using the ORIGIN (*32*) version of the NONLIN algorithm (*33*) supplied by Beckman.

Fluorescence and Circular Dichroism Spectroscopy. Fluorescence emission was measured using a PTI C-61 spectrofluorometer (Photon Technology International). The excitation wavelength was either 280 or 295 nm, and fluorescence emission was measured from 300 to 400 nm. Circular dichroism was measured using a Jasco J600A spectropolarimeter using either a 0.1 cm (far-UV) or a 1 cm (near-UV) cell. All measurements were corrected for background signal. Both instruments were equipped with thermostated cell holders, and the temperature was held constant at 25 °C (±0.1 °C) using a circulating water bath.

Fluorescence Anisotropy. Labeled pro-peptide (1 μM), either wild-type or D9A mutant, was incubated at 25 °C in a buffer of 50 mM potassium phosphate, pH 7.5, 1 mM DTT in a final volume of 2 mL. The peptide was titrated with pro-less variant between 0 and 10 μM, and the fluorescence anisotropy was measured as described (*34*). The PTI C-61 spectrofluorometer in the T-based format was used (excitation wavelength of 345 nm and emission wavelength of 450 nm). The data were fit to a simple binding model as described (*35*).

Infrared Spectroscopy. FTIR spectra were collected by attenuated total reflection (ATR) spectroscopy using an ATR objective in a Bio-Rad UMA 500 infrared microscope equipped with a liquid nitrogen-cooled MCT detector and attached to a Bio-Rad FTS 6000 FTIR spectrometer. The ATR objective consists of a Ge crystal as the internal reflection element at the focus of a Cassagranian objective in a single-pass configuration. The sample was placed beneath the ATR objective, and spectra were recorded continuously as the sample was concentrated slowly by dehydration. This method presents an alternative means of determining the amide I line shape in H_2O. Previous studies on small amounts of sample in H_2O required cells with path lengths <6 μM. The details of the method and comparison with bench FTIR spectra are described elsewhere (*61*). The method has a number of advantages. Measurements can be made on small quantities of sample obtained in H_2O. Solvent exchange into D_2O is not required. Solvent subtraction is not typically required unless the protein appears to denature under the dehydration conditions. The state of the sample can be monitored continuously in order to verify that the spectrum is not changing during the experiment. Finally, the sample is fully recoverable.

The procaspase proteins and pro-peptide were concentrated to approximately 100 μM (~3 mg/mL procaspase, ~0.3 mg/mL pro-peptide) and dialyzed against a buffer of 20 mM potassium phosphate, pH 7.2, 1 mM DTT. Using a Wheaton automatic pipet, 20 μL of sample was injected onto a cylindrical sample well that is milled in a Teflon block. All protein spectra were recorded at room temperature as averages of 64 scans of the interferometer with a resolution of 2 cm^{-1}. Spectra were recorded immediately after the sample was deposited onto the Teflon block and continuously, as the samples were gently dehydrated using a steady stream of N_2 gas. Spectra were acquired until a protein gel had formed onto the Ge crystal. Background spectra were obtained subsequently. The Ge crystal was rinsed with H_2O and allowed to dry prior to loading another protein sample. The spectral data were acquired using the software package Win-IR-Pro v2.97 manufactured by Bio-Rad. Data analysis was performed using the software package Igor-Pro v3.1. The spectral range of 600−4200 cm^{-1} was used for protein analysis.

Enzyme Activity and Inhibition Studies. Activity measurements were carried out as described (*26*). Caspase-3 was diluted into enzyme assay buffer (50 mM HEPES, pH 7.4, 100 mM NaCl, 1 mM EDTA, 10% glycerol, 0.1% CHAPS, 10 mM DTT) and incubated at 25 °C for 5 min. The protein concentration was 10 times that used in the experiment. The enzyme was then added to enzyme assay buffer that contained the caspase-3-specific tetrapeptide substrate Z-DEVD-AFC. The final enzyme concentrations were 0.1, 1, or 10 nM, and the substrate concentration was 10 μM. In separate assays, an inhibitor (Z-VAD-FMK) was added so that the final concentrations ranged from 10 to 800 nM, as indicated in the figures. For all assays, the final volume was 200 μL. The fluorescence emission of AFC generated from the catalytic reaction was monitored for 5000 s using an excitation wavelength of 400 nm and an emission wavelength of 505 nm. The resulting fluorescence intensity curves were fit to eq 1 (*36*) for an irreversible inhibitor, where v_0 is the initial velocity, k_{obs} is the observed rate constant, and A is a baseline offset.

$$Fl = A + v_0[1 - \exp(-k_{obs}t)]/k_{obs} \quad (1)$$

The second-order rate constant for binding of inhibitor to enzyme was obtained from a plot of k_{obs}, determined from fits of the enzyme assay data to eq 1, versus inhibitor concentration.

The wild-type and mutant (D9A) pro-domains were synthesized by the Peptide Facility at the University of North Carolina at Chapel Hill. The peptides were dissolved in DMSO at a concentration of 100 mM. In separate enzyme assays, as described above, the pro-domains were added in place of the inhibitor so that the final concentration of pro-domain was between 0 and 500 μM. The final DMSO concentration was less than 0.5%, and had no effect on enzyme activity as determined in control experiments of

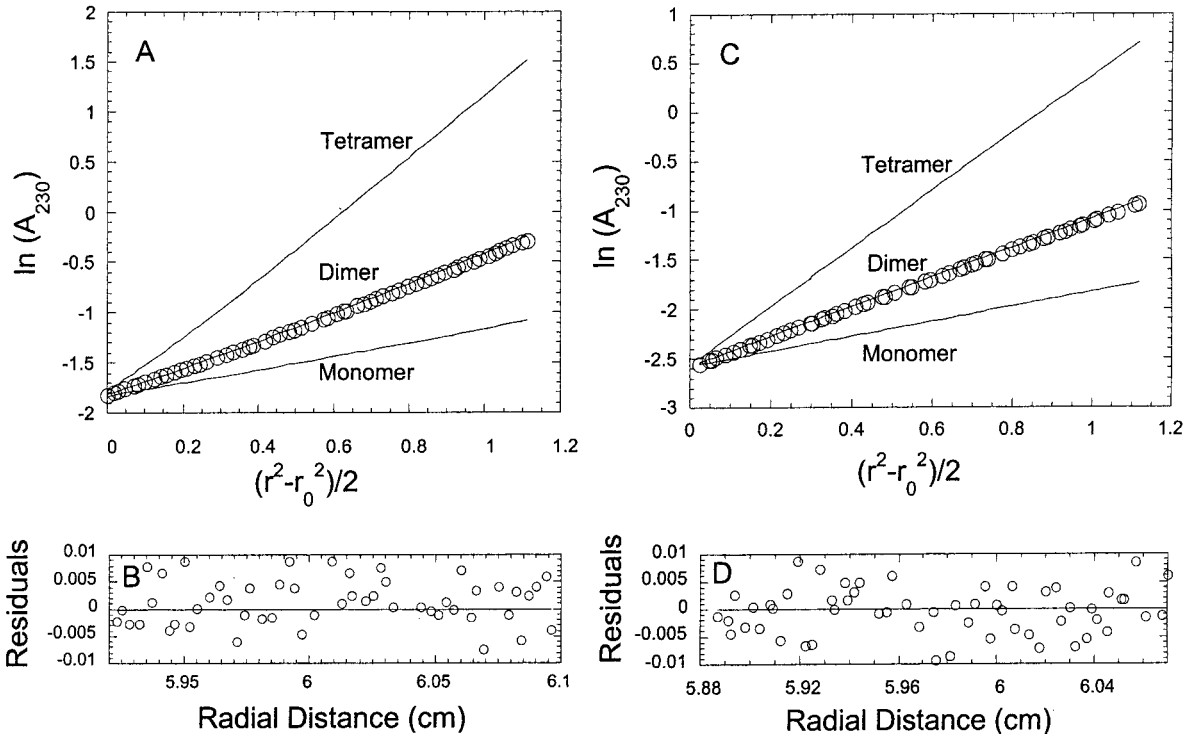


FIGURE 2: Sedimentation equilibrium studies of procaspase-3(C163S) and pro-less variant. Proteins (0.45−19.8 μM) were dialyzed against 50 mM KH_2PO_4/K_2HPO_4, pH 7.2, containing either 1 or 0.05 mM DTT. The spectra were recorded at 230 or 280 nm at 25 °C, using three rotor speeds (14 000, 18 000, and 24 000 rpm). The solid lines are the theoretical plots if the proteins were monomers, dimers, or tetramers, as indicated. The circles represent the experimental data for procaspase-3(C163S) (panel A) and pro-less variant (panel B). Residuals to the fits are shown in panels B and D.

enzyme in the presence of 0.5% DMSO. The enzymatic assays in the presence or in the absence of the pro-domain also were examined in a buffer containing 10% enzyme assay buffer, 10% water, and 80% 50 mM KH_2PO_4/K_2HPO_4, pH 7.2, 1 mM DTT, and were found to be nearly identical to those in enzyme assay buffer.

MALDI-TOF Mass Spectrometry. MALDI-TOF mass spectrometry was performed on a Bruker Proflex III equipped with a nitrogen laser (λ = 337 nm, 3 ns pulse width), deflection capabilities, delayed extraction, and an extended flight tube. The mass accuracy for analyte molecules is ±0.1% with external calibration. The matrix used was a saturated solution of α-cyano-4-hydroxycinnamic acid in acetonitrile, water, and TFA (50:50:0.1) (v/v/v). The instrument was calibrated against angiotensin I (1296.7 Da) and ubiquitin (8564.8 Da). Sample solutions (2 μL of 500 μM concentration) in 50 mM KH_2PO_4/K_2HPO_4, pH 7, with 1 mM DTT, containing 10% enzyme assay buffer, were desalted with Ziptip pipet tips (Millipore) and crystallized on a target plate with matrix solution (1:1). Experiments were performed at 25−30 attenuation and 50 shots. Each experiment was repeated at least 5 times, and the averaged data were collected.

Pro-Peptide Labeling with Dansyl Chloride. Labeling of the wild-type or mutant (D9A) pro-peptide was performed as described (*37*), with few modifications. Briefly, wild-type or mutant pro-peptide was dissolved in 0.1 M sodium bicarbonate, pH 8.6, to a final concentration of 10 mg/mL. The peptide solution was mixed while vortexing with 0.1 volume of 10 mg/mL dansyl chloride in DMF, followed by incubation overnight at 4 °C on a rotative wheel. The labeled peptide was separated from unreacted reagent using a Sephadex G-15 column (1 × 18 cm), equilibrated with 50 mM potassium phosphate, pH 7.5. The flow-through, containing the conjugated pro-peptide, was stored at −20 °C. The degree of labeling was determined by calculating the concentration of dansyl chloride conjugate (ϵ_{345} = 3400 M^{-1} cm^{-1}) and by comparison to the initial concentration of pro-peptide. The degrees of labeling of mutant pro-peptide and wild-type pro-peptide were 50% and 40%, respectively.

RESULTS

Comparison of Procaspase-3(C163S) and the Pro-less Variant. Procaspase-3 consists of 277 amino acids, with a M_r of 32 642 (including the LEH_6 sequence used for purification). Although under normal conditions in human cells the protein does not autoprocess, it has been shown to autoprocess under conditions in which the protein is over-expressed or the pro-domain is removed (*7*, *24*). To examine the oligomeric and spectroscopic properties of procaspase-3 without the complicating autoprocessing reactions, we mutated the active site cysteinyl residue to serine. Mutations of Cys163 have been shown to abrogate activity, but are not structurally perturbing (*19*, *38*−*40*). To examine the structure of the pro-domain within the procaspase, we constructed a mutant of procaspase-3(C163S), called the pro-less variant, in which the pro-peptide was removed by cloning (M_r = 29 658). In this protein, the amino terminus begins with a methionine residue, which is absent from processed protease purified from cells. The amino terminus of protease purified from human cells begins with Ser29. Both procaspase-3(C163S) and the pro-less variant were analyzed by MALDI-TOF mass spectrometry to confirm the correct mass (data not shown).

Procaspase-3 and the Pro-less Variant Are Dimers. To examine the oligomeric properties of procaspase-3 and the pro-less variant, we initially examined their elution profiles by size exclusion chromatography using Sephacryl S-100 HR resin at room temperature (data not shown). Results from these experiments showed that procaspase-3(C163S) eluted with an apparent molecular weight of 64 773, approximately equal to the mass of a dimer (for the monomer, $M_r = 32\,642$). Likewise, the pro-less variant eluted with an apparent molecular weight of 59 106 (for the monomer, $M_r = 29\,658$). There were no additional peaks corresponding to the monomer molecular weight, suggesting that at the protein concentrations of these experiments, 10 μM, both procaspase-3(C163S) and the pro-less variant are dimers.

To confirm that both proteins are dimers, we used sedimentation equilibrium to examine their oligomeric properties. The proteins were examined over a wide range of protein concentrations and at several rotor speeds (see Materials and Methods). Representative data are shown in Figure 2, panels A [procaspase-3(C163S)] and C (pro-less variant).

For both proteins, the data were best fit to a single, thermodynamically ideal species that corresponds to a dimer. The average molecular weights determined from the fits of all data sets were 65 573 for procaspase-3(C163S) and 58 342 for the pro-less variant. The residuals to the fits for both proteins (Figure 2, panels B and D) demonstrate that the data are well fit to the model.

In these experiments, there was no evidence for formation of a monomeric species, even at the lowest protein concentration (~450 nM) of the experiment. An accurate determination of the equilibrium dissociation constant is not possible under these conditions because we were unable to populate the monomeric species at the lowest protein concentration, which was at the detection limit of the instrument when the absorbance at 230 nm was monitored. However, if we assume a simple model for dissociation of a dimer to two monomers, and given the fact that the proteins are dimers at protein concentrations of 450 nM, we can estimate the upper limit for the equilibrium dissociation constant of the dimer to the monomer to be 50 nM, although it may be much lower.

Secondary and Tertiary Structural Properties of Procaspase-3 Do Not Change Significantly after Pro-Domain Removal. Procaspase-3 has 2 tryptophans in the C-terminal region (Trp206 and Trp214), as well as several tyrosines (10) and phenylalanines (15) that are well-dispersed throughout the sequence. Although the structure of the procaspase is not known, in the mature heterotetramer the two tryptophans are in close contact. Tryptophan 206 forms part of the S2 binding pocket (*7*), whereas tryptophan 214 is found in the S4 site and hydrogen bonds to the substrate. There are no aromatic amino acids in the pro-domain of procaspase-3. These features allow for the study of potential changes in the fluorescence emission of the procaspase-3 dimer upon removal of the pro-domain by monitoring fluorescence emission between 300 and 400 nm following excitation either at 280 nm, to excite all aromatic amino acids, or at 295 nm, to excite primarily the tryptophanyl residues.

As shown in Figure 3, panel A, the fluorescence emission spectra following excitation at 280 nm are superimposable for procaspase-3(C163S) and the pro-less variant. Both proteins demonstrate an emission maximum at 335 nm. The

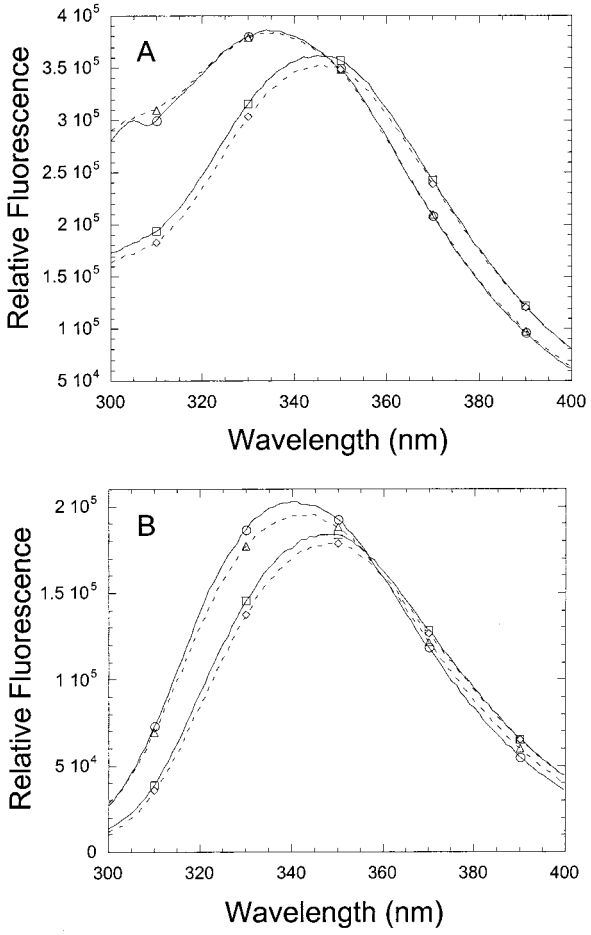


FIGURE 3: Fluorescence emission spectra of native and unfolded procaspase-3(C163S) and pro-less variant. Proteins (1 μM) were in 50 mM KH_2PO_4/K_2HPO_4, pH 7.2, 1 mM DTT, 25 °C, with or without 8 M urea. The samples were excited at 280 nm (panel A) or 295 nm (panel B), and the fluorescence emission was collected from 300 to 400 nm. For panels A and B: (○) native procaspase-3(C163S); (△) native pro-less variant; (□) unfolded procaspase-3(C163S); (◇) unfolded pro-less variant.

fluorescence emission following excitation at 295 nm shows similar results (Figure 3, panel B). In this case, the emission maximum for procaspase-3(C163S) is 340 nm, whereas that for the pro-less variant is slightly red-shifted to 341 nm. In addition, the fluorescence emission of the pro-less variant is quenched slightly (~2%). Following unfolding in 8 M urea, the emission maxima are red-shifted to 346 nm (Figure 3, panel A) or 348 nm (Figure 3, panel B), and the fluorescence emission decreases by ~10%. These results are consistent with our equilibrium unfolding data (see preceding paper, *62*), which show that procaspase-3(C163S) is largely unfolded under these conditions.

We compared the secondary and tertiary structures of procaspase-3(C163S) and the pro-less variant by monitoring near- and far-UV circular dichroism, and the data are shown in Figure 4. In the far-UV, procaspase-3(C163S) demonstrates two minima (208 and 215 nm) (Figure 4, spectrum 1), although the amplitudes of the minima are very similar. While it is not possible to deconvolute these data into their component secondary structural elements due to the limitations of our instrument, based on reference spectra described by Johnson (*41*), the data are consistent with a protein that contains both α-helix and β-sheet elements, with the α-heli-

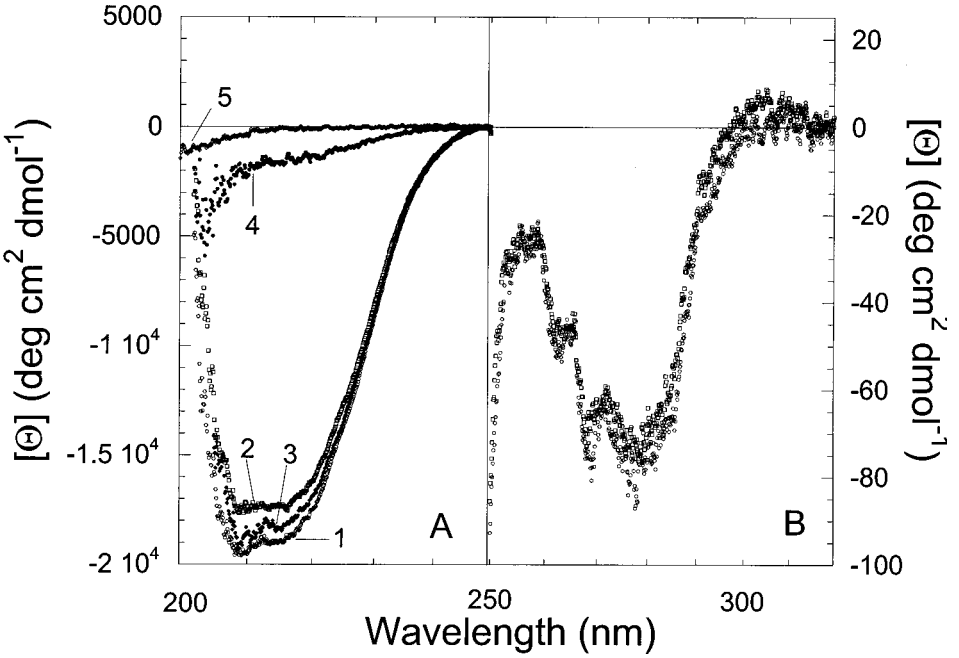


FIGURE 4: Near- and far-UV circular dichroism spectra of procaspase-3(C163S), pro-less variant, and pro-peptide. The experiments were performed at 25 °C in a buffer containing 20 mM KH_2PO_4/K_2HPO_4, pH 7.2, 0.1 mM DTT. Protein concentrations were 20 μM. (Panel A) Far-UV CD spectra: (1) procaspase-3(C163S); (2) pro-less variant; (3) pro-less variant with pro-peptide in trans; (4) procaspase-3(C163S) minus pro-less variant (spectrum 1 minus spectrum 2); (5) pro-peptide. (Panel B) Near-UV CD spectra: (○) procaspase-3(C163S); (□) pro-less variant.

cal structures contributing to the minimum at 208 nm. The far-UV CD spectrum for the pro-less variant (Figure 4, spectrum 2) is similar to that of procaspase-3 except that the minimum at 208 nm is less pronounced. In addition, there is a decrease in amplitude of approximately 10% between the pro-less variant and procaspase-3(C163S), consistent with the loss of about 10% of the total number of amino acids in the pro-less variant (28 amino acids from 285 amino acids). The spectra are superimposable when the data (spectra 1 and 2 in Figure 4) are normalized by the total number of amino acids in each protein, that is, the mean residue ellipticity (data not shown). This suggested that the secondary structure of the procaspase-3(C163S) dimer was unaffected by the loss of the pro-domain.

The near-UV CD spectra (Figure 4) for procaspase-3 and the pro-less variant are superimposable. The data demonstrate minima at 280 and 270 nm as well as a maximum at 255 nm, indicative of well-packed tertiary structures. The fluorescence emission data shown in Figure 3, panel B, suggested that there might be a conformational change in the protein due to loss of the pro-domain. However, results from circular dichroism (Figure 4) suggest that the structural changes may be localized to the environment of the two tryptophan residues in the pro-less variant. While it is not yet clear whether the localized structural changes affect protein stability, there is no effect on dimerization (Figure 2).

The Structure of the Pro-Domain Bound to the Procaspase Dimer Is Different than the Pro-Domain in Solution. A comparison of the far-UV CD spectra (Figure 4) for procaspase-3(C163S) and the pro-less variant suggested that the pro-domain may contain secondary structural elements that contribute to the spectrum of the full-length protein. We have subtracted the two spectra (Figure 4, spectra 1 and 2), and the results are shown in Figure 4 (spectrum 4). Because the mean residue ellipticity spectra are superimposable for the two proteins (data not shown), conformational changes, if any, in the protein as a result of removal of the pro-domain do not affect the far-UV spectral properties. Therefore, we suggest that spectrum 4 in Figure 4 represents the pro-domain of procaspase-3. We have examined the pro-domain (28 amino acids) under the same solution conditions as those described for procaspase-3 (Figure 4) in order to compare the far-UV CD properties of the pro-domain with those of the subtracted spectra. The far-UV CD spectrum for the pro-domain in solution is shown as spectrum 5 in Figure 4. In the absence of the protease domain, the pro-peptide appears to contain primarily random structure. We also examined the CD spectra at several concentrations of pro-peptide (10−100 μM) and found no change in the CD spectra with an increase in protein concentration (data not shown). These results demonstrated that the pro-domain does not self-associate under the conditions of these experiments. In support of this, we examined the peptide by native gel electrophoresis (data not shown) at concentrations of ~500 μM and observed only a monomeric peptide. Together, these results demonstrate that the pro-domain of procaspase-3 does not self-associate and suggest that it contains secondary structure when present in cis to the remaining protease domain.

FTIR Spectra Suggest the Pro-Domain Is a β-Strand. To examine the interactions of the pro-domain with the pro-less variant, especially the structure of the pro-domain, we examined procaspase-3(C163S), the pro-less variant, and the pro-peptide by FTIR, and the results are shown in Figure 5. While the details are described elsewhere (*42*), a method was developed for collecting FTIR spectra using attenuated total reflection spectroscopy in a Fourier transform infrared microscope that presents an extremely rapid and sensitive means of determining the amide I line shape in H_2O.

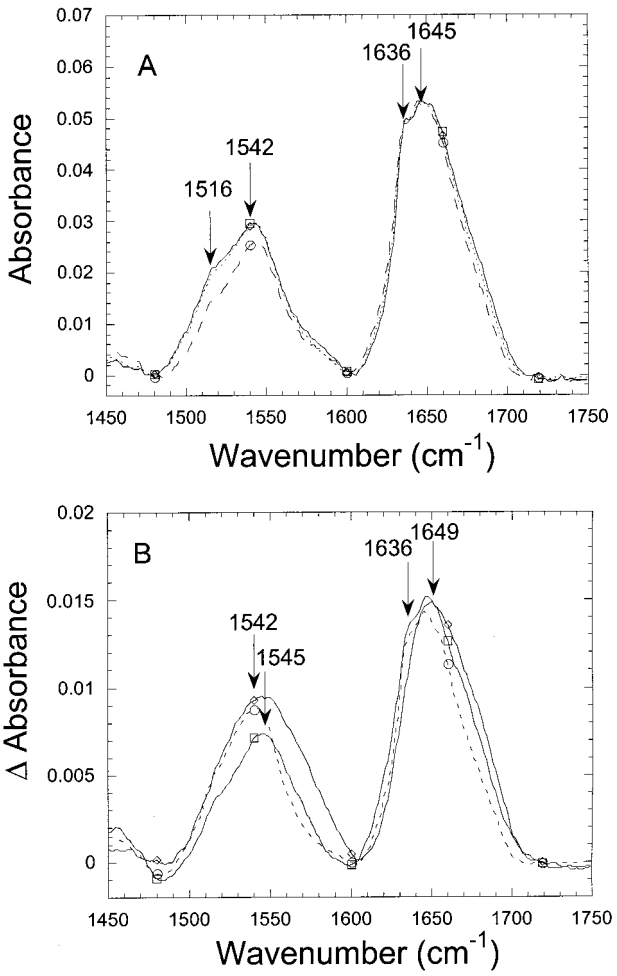


FIGURE 5: FTIR spectra of procaspase-3(C163S), pro-less variant, and pro-peptide. (Panel A) FTIR spectra for procaspase-3(C163S) (□), pro-less variant (◇), and a mixture of pro-less variant with pro-peptide (○). (Panel B) FTIR spectra of pro-peptide (◇), procaspase-3(C163S) minus pro-less variant (○), and pro-less variant in a mixture with pro-peptide minus the spectrum of the pro-less variant (□).

Representative spectra for procaspase-3(C163S) and the pro-less variant are shown in Figure 5, panel A. The data demonstrate similar spectral shapes for both proteins in the amide I region with peaks at 1636 and 1645 cm^{-1}. In addition, both proteins have peaks in the amide II region at 1542 and 1516 cm^{-1}. These spectra are consistent with an α/β protein (*42*). Consistent with the fluorescence and CD spectra (Figures 3 and 4), the FTIR spectra demonstrate that removal of the pro-peptide has little effect on the remainder of the procaspase-3 dimer.

As described for the circular dichroism spectra in Figure 4, the FTIR spectra shown in Figure 5A were subtracted in order to obtain the contribution of the pro-domain to the spectrum, and the data were compared to the pro-peptide free in solution. The results are shown in Figure 5, panel B. The spectrum of the pro-peptide in solution (Figure 5B, diamonds) shows a single peak at 1649 cm^{-1} in the amide I region as well as a peak at 1545 cm^{-1} in the amide II region. In comparison, the subtracted spectrum (procaspase-3 minus pro-less variant) (Figure 5B, circles) demonstrates peaks at 1645 and 1636 cm^{-1} in the amide I region as well as a peak at 1542 cm^{-1} in the amide II region. The correlations between protein structure and amide I bands are well-known (*42*) and indicate that the pro-peptide, in solution, is a random coil, whereas the pro-peptide in cis to the protease domain (that is, in procaspase-3) contains β-structure. Currently, we cannot interpret the shifts in the amide II region because there are no correlations as for the amide I band. When these spectra are interpreted with the results from circular dichroism (Figure 4), the data suggest that the pro-domain of procaspase-3 is not a random coil when in cis to the protease domain, but rather may be a β-strand that forms hydrogen bonds with the protease domain. The pro-peptide free in solution, however, appears to be a random coil.

The Pro-Domain Interacts with the Pro-less Variant in Trans. The pro-less variant was incubated with the pro-domain in trans in order to determine whether the pro-peptide interacts with the protease domain when not covalently attached to the protease domain. As shown in Figure 4 (spectrum 3), a mixture of pro-less variant with the pro-domain (20 μM of each) resulted in an increase in ellipticity at 208 nm and at 215 nm. Compared to the spectrum of procaspase-3 (Figure 4, spectrum 1), approximately 70% of the ellipticity was obtained. No further increase occurred when the pro-domain was in excess of the pro-less variant (data not shown). In addition, the near-UV CD spectra (data not shown) were identical to those shown in Figure 4. These results suggest that the pro-domain interacts with the pro-less variant and adopts a structure similar to the pro-domain in full-length procaspase-3. Formally, it remains possible that the differences in CD spectra are due to conformational changes in the pro-less variant. While we cannot rule out this possibility, the data demonstrate that the pro-peptide interacts with the protease domain in trans. It seems unlikely that the structure of the pro-peptide would remain a random coil when bound to the protease domain.

The mixture of pro-less variant plus pro-domain was also analyzed by FTIR. As shown in Figure 5, panel A, the spectrum for the mixture was superimposable with that of full-length procaspase-3(C163S) in the amide I region. This does not occur simply by adding the individual spectra (data not shown), indicating that structure is induced in the pro-domain while in the presence of the pro-less variant. The spectrum for the pro-less variant was subtracted from that for the mixture (pro-less variant plus pro-peptide), and the results are shown in Figure 5, panel B (squares). The resulting spectrum is intermediate with those for the peptide in solution (Figure 5B, diamonds) and the pro-peptide in cis to the protease domain (Figure 5B, circles). For example, the peptide in trans has two peaks in the amide I region at 1647 and 1636 cm^{-1}, similar to the pro-peptide in cis. However, the peak in the amide II region is similar to that for the pro-peptide in solution. These spectra are consistent with both β-sheet and random coil components to the pro-peptide when it is in the presence of the protease domain, which may be due to weak binding of the pro-peptide to the pro-less variant.

To confirm the interactions of the pro-peptide with the pro-less variant, we measured the changes in fluorescence anisotropy of a pro-peptide labeled with dansyl chloride as it was titrated with the pro-less variant. The data are shown in Figure 6. In these experiments, either the dansylated wild-type pro-peptide (Figure 6, squares) or the D9A mutant pro-peptide (Figure 6, circles) was titrated with the pro-less variant. The data show that both pro-peptides bind to the

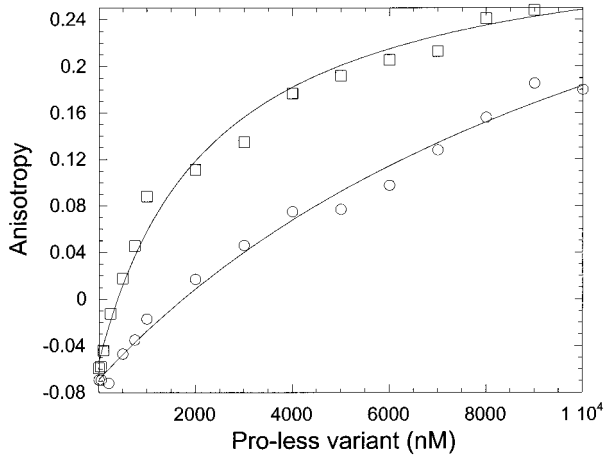


FIGURE 6: Fluorescence anisotropy of pro-peptide in the presence of pro-less variant. The dansyl-labeled wild-type pro-peptide (□) or the dansyl-labeled D9A mutant pro-peptide (○) (1 μM) was titrated with the pro-less variant. All assays were carried out at 25 °C in a buffer of 50 mM potassium phosphate, pH 7.5, 1 mM DTT. The solid lines represent fits of the data to a simple binding model, as described under Materials and Methods.

pro-less variant, although the wild-type pro-peptide appears to bind more tightly than the D9A mutant peptide. While the interactions may be more complex, we assumed a simple binding model to describe the interactions (solid lines in Figure 6). Overall, the data demonstrate the pro-peptides bind to the pro-less variant with equilibrium dissociation constants in the low micromolar range and confirm the interactions observed by circular dichroism (Figure 4) and FTIR (Figure 5).

Pro-domain Is neither a Good Substrate nor a Good Inhibitor of Caspase-3. Because of the spectroscopic data described above, we examined the activity of the mature caspase-3 protease in the presence of the pro-peptide in order to examine interactions with the protease domain in trans. The results of these studies are shown in Figure 7. The data in Figure 7, panel A, show the enzyme activity of caspase-3 in the presence of a fluorescence substrate (DEVD-AFC) and with increasing amounts of the irreversible caspase inhibitor Z-VAD-FMK. The data demonstrate a decrease in activity with an increase in inhibitor concentration. As shown previously (*26*), the data are well-described by eq 1, and the second-order rate constant for binding of inhibitor to caspase-3 (described under Materials and Methods) is in agreement with that determined previously (*6*), $\sim 3 \times 10^3$ M^{-1} s^{-1}. In contrast to the caspase inhibitor, there was no change in enzyme activity in the presence of the pro-peptide, as shown in Figure 7, panel B. In these studies, the concentration of the pro-peptide was varied (from 0 to 500 μM), and there was no change in activity. In agreement with previous reports (*14*), these results show that the pro-peptide does not inhibit the activity of the caspase-3 protease.

While the enzymatic data (Figure 7, panel A) indicated that the pro-domain does not inhibit the protease, it is possible that the pro-peptide may be a substrate of the caspase-3 protease. Simulations of the enzymatic reactions (data not shown) using the program Kinsim (*43*) demonstrate that production of AFC product from the DEVD-AFC substrate would be unaffected so long as the binding of the pro-peptide were weaker than that of the fluorescent substrate

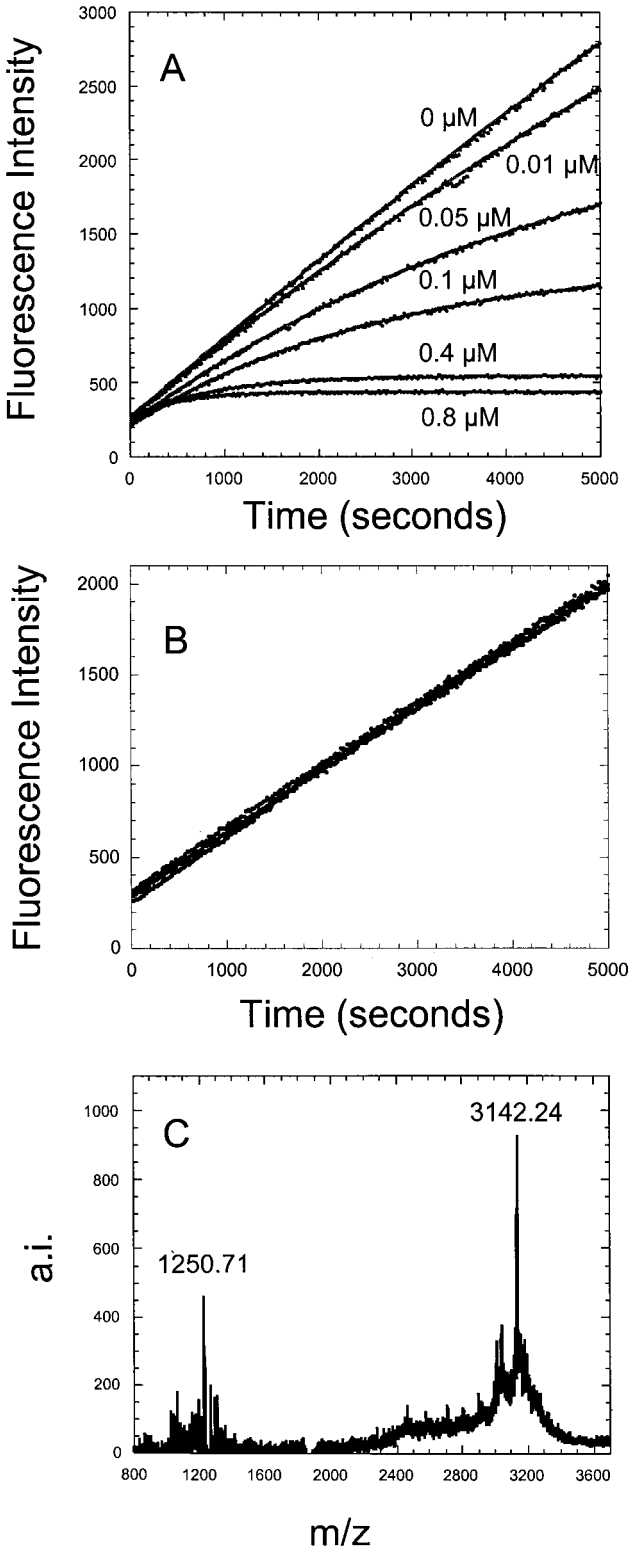


FIGURE 7: Pro-domain influence on caspase-3 activity. Caspase-3 (0.1 nM) and DEVD-AFC substrate (10 μM) were incubated in assay buffer in the presence of inhibitor or pro-peptide as described under Materials and Methods. Fluorescence emission of generated AFC was monitored at 505 nm after excitation at 400 nm. (Panel A) Caspase-3 assay in the presence of the inhibitor Z-VAD-FMK. The inhibitor concentrations are indicated. (Panel B) Caspase-3 assay in the presence of wild-type and mutant (D9A) pro-peptide (0−500 μM). (Panel C) MALDI-TOF analysis of caspase-3 reaction products after hydrolysis of wild-type or mutant (D9A) pro-domain by caspase-3. The peak at 3142 Da represents the full-length pro-peptide.

(DEVD-AFC) and the binding of the pro-peptide were reversible. Based on the structure described recently (*44*) of the procaspase-3 maturation intermediate bound to an inhibitor, XIAP, both of these criteria seem likely. In that structure, the pro-peptide is disordered in the crystal.

We examined whether the pro-peptide is a substrate for caspase-3 in two ways. First, we tested the enzyme activity of the caspase-3 protease in the presence of a mutant of the pro-peptide in which Asp9 was replaced with alanine ($NSVA^9$). This peptide would not be cleaved by the protease. As shown in Figure 7, panel B, there was no change in the enzyme activity in the presence of up to 500 μM pro-peptide, confirming that the pro-peptide does not inhibit activity. In separate experiments, the enzyme concentration was 0.1, 1, or 10 nM, and the results were the same as those shown in Figure 7B. Second, in separate reactions, we incubated the pro-peptide (wild-type or D9A mutant) with active caspase-3 enzyme under conditions similar to those in Figure 7 except that the substrate (DEVD-AFC) was not added. We then examined the products, if any, of the enzymatic reaction by SDS−PAGE and MALDI-TOF mass spectrometry. If cleavage occurred at Asp9 in the wild-type pro-peptide, then the molecular weight would decrease from one species of ~3130 to two species of ~1030 and ~2100. The SDS−PAGE analysis of the enzymatic reactions showed a single species for the wild-type pro-peptide as well as for the D9A mutant pro-peptide (data not shown). Results from the MALDI-TOF analysis are shown in Figure 7C. For both the wild-type and mutant pro-peptides, only a single species was observed with M_r ~3100. In Figure 7C, the peak at M_r 1250 arises from the buffer, as determined from control experiments of buffer alone (data not shown). These data suggest that while the pro-domain does not inhibit the activity of caspase-3, it is not a good substrate when in trans to the mature protease.

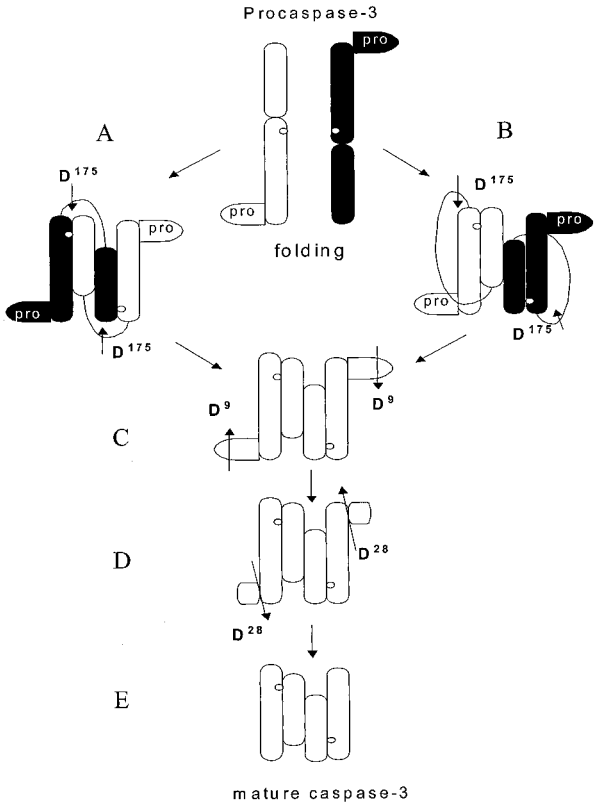


FIGURE 8: Model for procaspase-3 assembly and maturation. Before processing, two molecules of procaspase-3 form a dimer by interdigitation (structure A) or association (structure B), generating a structure similar to the one of caspase-3 heterotetramer. The dimeric procaspase is cleaved at Asp175 to form the maturation intermediate (structure C). The pro-domain is then cleaved rapidly at Asp9 (structure D), then more slowly at Asp28 to give the mature heterotetramer (structure E). The catalytic site is schematized as a white ball.

DISCUSSION

We have shown that human procaspase-3 is a dimer in solution, and we estimate the upper limit for the equilibrium dissociation constant to be ~50 nM. Removal of the pro-peptide has little effect on the equilibrium dissociation constant of the dimer or on the spectroscopic properties of the dimer, indicating that the dimeric structure is likely to be unaltered in the pro-less mutant. Results from circular dichroism, fluorescence anisotropy, and FTIR studies suggest that the pro-peptide interacts weakly with the dimer. In addition, the data suggest that the pro-peptide adopts a β-structure when in contact with the protein, either covalently attached (in cis), as for the procaspase-3, or when added in trans, as for the pro-less variant. Surprisingly, while the pro-peptide does not inhibit the activity of the mature caspase-3 heterotetramer, our data suggest that the pro-peptide is a poor substrate for the enzyme.

The question remains, however, whether the procaspase dimer observed in vitro is relevant to the oligomeric state of the protein in vivo. First, our folding data (*62*) demonstrate that dimerization of procaspase-3(C163S) occurs through a folding intermediate rather than through the native, monomeric procaspase-3, demonstrating that dimerization is a folding event. Second, based on the average volume of a Jurkat T cell, ~9.5×10^{-10} mL (*45*), and the quantity of caspase-3 heterotetramer in apoptotic Jurkat cells (6.6 ng/10^6 cells) (*46*), we estimate the concentration in vivo of the active caspase-3 to be ~100 nM. Salvesen and co-workers have quantified caspase-3 in 293 cells and found the same concentration (~100 nM) (*14*). In the case of caspase-1, the concentration of the procaspase has been shown to be >10-fold higher than the activated form (*47*, *48*). If this is true for caspase-3, then the concentrations of procaspase-3 may be much higher than 100 nM. Third, Cohen and co-workers (*49*, *50*) demonstrated that procaspase-3 from THP.1 cell lysates eluted as a dimer from a gel filtration column. This occurred in both control (nonapoptotic) and apoptotic cells (*49*). Together, the data strongly suggest that procaspase-3 is a dimer in vivo.

These results are important because they suggest that dimerization is an early event in procaspase-3 maturation, similar to that described for procaspase-1 and activator procaspases (*40*). Our results as well as previous studies (*51*) suggest the model shown in Figure 8 for maturation of procaspase-3. In this model, dimeric procaspase is the result of either interdigitation or association of two precursors. In the first case, large and small subunits from different molecules contribute to the catalytic center, whereas in the latter case this is preformed within the same monomer. In both cases, the interface is generated predominantly by interactions between the small subunits, as in mature caspase-3. The dimeric procaspase-3 is cleaved at Asp175 to separate the covalent linkage of the large and small subunits. The

pro-peptide is then cleaved rapidly at Asp9, then more slowly at Asp28, generating the mature heterotetramer. This model is consistent with the recent crystal structure of the procaspase-3 maturation intermediate with the inhibitor XIAP bound, which shows that the intermediate of maturation is a dimer with inhibitor bound to each active site (*44*). While it is not known whether the dimeric procaspase structure involves a domain-swap as shown in the Figure 8, this was suggested from the crystal structures of the mature heterotetramer (*6*, *7*) based on the positions of the C-terminal ends of the large subunits and the N-terminal ends of the small subunits. In the case of procaspase-1 (*40*), both intercalated and nonintercalated structures have been shown to form. Whether the procaspase-3 dimer results from domain-swapping remains to be determined, although one should note that following cleavage at Asp175, the maturation intermediate from either form of the procaspase would be indistinguishable, as shown in Figure 8.

The question remains why the procaspase-3 dimer is inactive in vivo such that autocatalytic processing does not occur. We suggest two reasons for this. First, rather than the canonical caspase-3 recognition sequence (DEVD), the tetrapeptide sequence at Asp175 (IETD^{175}) is optimal for caspases in subclass III, caspases-6, -8, -9, all of which activate procaspase-3 (*52*). Thornberry and co-workers (*53*) and Salvesen and co-workers (*54*) have shown that substitutions at the P4 position of the substrate have a dramatic effect on the ability of caspase-3 to cleave the substrate. In some cases, the caspase-3 activity is only a small fraction when compared to that for the optimal sequence. Second, the low activity is coupled to low protein concentrations (~100 nM). Increasing either of these two factors (activity or concentration) would lead to autoactivation. In the case of protein expressed from heterologous expression systems, the procaspase-3 dimer autoactivates because of the high protein concentrations. Although we have shown that the pro-peptide is not a good substrate for caspase-3 in trans, the locally high concentrations when it is present in cis to the protease domain allows for cleavage of the pro-peptide. Therefore, an increase in activity, as observed in the maturation intermediate, could lead to cleavage of the pro-peptide. The active intermediate or mature enzyme would feed back to activate procaspase-3.

Pro-domains from many protease zymogens have been shown to interact with the protease domain either to inhibit activity (*55*) or to influence folding as an intramolecular chaperone (*56*, *57*). In those systems, the pro-domains are generally much larger than that for procaspase-3. Recently, however, the pro-peptide of β-secretase, which consists of 24 amino acids, was shown to increase the yield of correctly folded protein when present in cis or in trans (*58*). Although the nature of the interactions is not yet known, we have shown that the caspase-3 pro-peptide interacts weakly with the protease domain of the zymogen. In contrast with long pro-domain caspases, procaspase-3 does not utilize its pro-domain in dimerization, because the pro-less variant described here is also a dimer. Most probably, procaspase-3 forms dimers using the same residues involved in the mature caspase-3 heterodimer−heterodimer interface. The nature of this interface, with Val266−Val266# at the 2-fold symmetry axis, suggests that the driving force in procaspase-3 dimerization is based on hydrophobic interactions. This is not the case for caspases-1, -4, and -5, where the interface between the small subunits contains two salt-bridges. In those caspases, the CARD−CARD interactions between the pro-domains and adapter proteins apparently precede other contacts, but are not sufficient to keep the dimer in a stable conformation (*19*, *59*).

The possibility remains that the pro-peptide of procaspase-3 may act to attenuate the low levels of activity of the procaspase dimer, which would further serve as a control over activation. Alternatively, in comparison with subtilisin (*60*), for example, the pro-peptide may function as an intramolecular chaperone to assist in procaspase-3 folding. Either of these possibilities would suggest that pro-domains from the activator caspases, which contain the CARD motifs, are bifunctional. The remaining possibility is that the pro-peptides of the executioner caspases are simply vestigial linkers that connect the CARD motif to the amino terminus of the large subunit. Because the executioner caspases are activated by other caspases in the cascade, the CARD motif may have been lost through evolution, whereas the linker was retained. Further experiments will determine the role of the procaspase pro-peptide among these intriguing possibilities.

ACKNOWLEDGMENT

We thank Luming He and Randy Durren for their technical expertise. We also thank Dr. Dennis Brown for use of the HPLC. We are grateful to Dr. Emad Alnemri for supplying procaspase-3 cDNA and to Dr. Jim Hu for helpful discussions and for use of the analytical ultracentrifuge, allowing us to perform initial experiments.

REFERENCES

1. Earnshaw, W. C., Martins, L. M., and Kaufmann, S. H. (1999) *Annu. Rev. Biochem. 68*, 383−424.
2. Marks, N., and Berg, M. J. (1999) *Neurochem. Int. 35*, 195−220.
3. Wolf, B. B., and Green, D. R. (1999) *J. Biol. Chem. 274*, 20049−20052.
4. Tewari, M., Quan, L. T., O'Rouke, K., Desnoyers, S., Zeng, Z., Beidler, D. R., Poirier, G. G., Salvesen, G. S., and Dixit, V. M. (1995) *Cell 81*, 801−809.
5. Kuida, K., Zheng, T. S., Na, S., Kuan, C.-Y., Yang, D., Karasuyama, H., Rakic, P., and Flavell, R. A. (1996) *Nature 384*, 368−372.
6. Rotonda, J., Nicholson, D. W., Fazil, K. M., Gallant, M., Gareau, Y., Labelle, M., Peterson, E. P., Rasper, D. M., Ruel, R., Vaillancourt, J. P., Thornberry, N. A., and Becker, J. W. (1996) *Nat. Struct. Biol. 3*, 619−625.
7. Mittl, P. R. E., DiMarco, S., Krebs, J. F., Bai, X., Karanewsky, D. S., Priestle, J. P., Tomaselli, K. J., and Grutter, M. G. (1997) *J. Biol. Chem. 272*, 6539−6547.
8. Thornberry, N. A., Bull, H. G., Calaycay, J. R., Chapman, K. T., Howard, A. D., Kostura, M. J., Miller, D. K., Molineaux, S. M., Weidner, J. R., Aunins, J., Elliston, K. O., Ayala, J. M., Casano, F. J., Chin, J., Ding, G. J.-F., Egger, L. A., Gaffney, E. P., Limjuco, G., Palyha, O. C., Raju, S. M., Rolando, A. M., Salley, J. P., Yamin, T.-T., Lee, T. D., Shively, J. E., MacCross, M., Mumford, R. A., Schmidt, J. A., and Tocci, M. J. (1992) *Nature 356*, 768−774.
9. Walker, N. P. C., Talanian, R. V., Brady, K. D., Dang, L. C., Bump, N. J., Ferenz, C. R., Franklin, S., Ghayur, T., Hackett, M. C., Hammill, L. D., Herzog, L., Hugunin, M., Houy, W., Mankovich, J. A., McGuiness, L., Orlewicz, E., Paskind, M., Pratt, C. A., Reis, P., Summani, A., Terranova, M., Welch, J. P., Xiong, L., Moller, A., Tracey, D. E., Kamen, R., and Wong, W. W. (1994) *Cell 78*, 343−352.

10. Watt, W., Koeplinger, K. A., Mildner, A. M., Heinrikson, R. L., Tomasselli, A. G., and Watenpaugh, K. D. (1999) *Structure 7*, 1135−1143.
11. Chai, J., Shiozaki, E., Srinivasula, S. M., Wu, Q., Dataa, P., Alnemri, E. S., and Shi, Y. (2001) *Cell 104*, 769−780.
12. Yamin, T.-T., Ayala, J. M., and Miller, D. K. (1996) *J. Biol. Chem. 271*, 13273−13282.
13. Han, Z., Hendrickson, E. A., Bremner, T. A., and Wyche, J. H. (1997) *J. Biol. Chem. 272*, 13432−13436.
14. Stennicke, H. R., Jurgensmeier, J. M., Shin, H., Deveraux, Q., Wolf, B. B., Yang, X., Zhou, Q., Ellerby, M., Ellerby, L. M., Bredesen, D., Green, D. R., Reed, J. C., Froelich, C. J., and Salvesen, G. S. (1998) *J. Biol. Chem. 273*, 27084−27090.
15. Mancini, M., Nicholson, D. W., Roy, S., Thornberry, N. A., Peterson, E. P., Casciola-Rosen, L. A., and Rosen, A. (1998) *J. Biol. Chem. 140*, 1485−1495.
16. Shin, S., Sung, B.-J., Cho, Y.-S., Kim, H.-J., Ha, N.-C., Hwang, J.-I., Chung, C.-W., Jung, Y.-K., and Oh, B.-H. (2001) *Biochemistry 40*, 1117−1123.
17. Cohen, G. (1997) *Biochem. J. 326*, 1−16.
18. Chou, J., Matsuo, H., Duan, H., and Wagner, G. (1998) *Cell 94*, 171−180.
19. Van Criekinge, W., Beyaert, R., Van de Craen, M., Vandenabeele, P., Schotte, P., De Valck, D., and Fiers, W. (1996) *J. Biol. Chem. 271*, 27245−27248.
20. Stennicke, H. R., and Salvesen, G. S. (1998) *Biochim. Biophys. Acta 1387*, 17−31.
21. Ryser, S., Vial, E., Magnenat, E., Schlegel, W., and Maundrell, K. (1999) *Curr. Genet. 36*, 21−28.
22. Fernandes-Alnemri, T., Litwack, G., and Alnemri, E. S. (1994) *J. Biol. Chem. 269*, 30761−30764.
23. Dorstyn, L., Kinoshita, M., and Kumar, S. (1998) *Results Probl. Cell Differ. 24*, 1−24.
24. Meergans, T., Hildebrandt, A.-K., Horak, D., Haenisch, C., and Wendel, A. (2000) *Biochem. J. 349*, 135−140.
25. Fernandes-Alnemri, T., Litwack, G., and Alnemri, E. S. (1995) *Cancer Res. 55*, 2737−2742.
26. Stennicke, H. R., and Salvesen, G. S. (1999) *Methods 17*, 313−319.
27. Laemmli, U. K. (1970) *Nature 227*, 680−685.
28. Clark, A. C., Ramanathan, R., and Frieden, C. (1998) *Methods Enzymol. 290*, 100−118.
29. Edelhoch, H. (1967) *Biochemistry 6*, 1948−1954.
30. Stennicke, H. R., and Salvesen, G. S. (1997) *J. Biol. Chem. 272*, 25719−25723.
31. Zhou, Q., Krebs, J. F., Snipas, S. J., Price, A., Alnemri, E. S., Tomaselli, K. J., and Salvesen, G. S. (1998) *Biochemistry 37*, 10757−10765.
32. Hedegs, J., Sarrafzadeh, S., Lear, J. D., and McRorie, D. K. (1984) *Modern Analytical Ultracentrifugation* (Schuster, T. M., and Laue, T. M., Eds.) Birkhäuser, Boston, MA.
33. Johnson, M., Correa, I., Yphantis, D., and Halvorson, H. (1981) *Biophys. J. 36*, 575−588.
34. Lakowicz, J. R. (1983) *Principles of fluorescence spectroscopy* (Orton, C. G., Ed.) Plenum Press, New York.
35. Connolly, K. M., Ilangovan, U., Wojciak, J. M., Iwahara, M., and Clubb, R. T. (2000) *J. Mol. Biol. 300*, 841−856.
36. Morrison, J. F., and Walsh, C. T. (1998) *Adv. Enzymol. Relat. Areas Mol. Biol. 59*, 201−301.
37. Haugland, R. (1995) *Methods Mol. Biol. 45*, 205−221.
38. Colussi, P. A., Harvey, N. L., Shearwin-Whyatt, L. M., and Kumar, S. (1998) *J. Biol. Chem. 1998*, 26566−26570.
39. Stennicke, H. R., Deveraux, Q. L., Humke, E. W., Reed, J. C., Dixit, V. M., and Salvesen, G. S. (1999) *J. Biol. Chem. 274*, 8359−8362.
40. Gu, Y., Wu, J., Faucheu, C., Lalanne, J.-L., Diu, A., Livingston, D. J., and Su, M. S.-S. (1995) *EMBO J. 14*, 1923−1931.
41. Johnson, W. C., Jr. (1990) *Proteins: Struct., Funct., Genet. 7*, 205−214.
42. Jackson, M., and Mantsch, H. H. (1995) *Crit. Rev. Biochem. Mol. Biol. 30*, 95−120.
43. Barshop, B. A., Wrenn, R. F., and Frieden, C. (1983) *Anal. Biochem. 130*, 134−145.
44. Riedl, S. J., Renatus, M., Schwarzenbacher, R., Zhou, Q., Sun, C., Fesik, S. W., Liddington, R. C., and Salvesen, G. S. (2001) *Cell 104*, 791−800.
45. Ivanov, I. B., Hadjiiski, A., Denkov, N. D., Gurkov, T. D., Kralchevsky, P. A., and Koyasu, S. (1998) *Biophys. J. 75*, 545−556.
46. Saunders, P. A., Cooper, J. A., Roodell, M. M., Schroeder, D. A., Borchert, C. J., Isaacson, A. L., Schendel, M. J., Godfrey, K. G., Cahill, D. R., Walz, A. M., Loegering, R. T., Gaylor, H., Woyno, I. J., Kaluyzhny, A. E., Krzyzek, R. A., Mortari, F., Tsang, M., and Roff, C. F. (2000) *Anal. Biochem. 284*, 114−124.
47. Ayala, J. M., Yamin, T.-T., Egger, L. A., Chin, J., Kostura, M., and Miller, D. K. (1994) *J. Immunol. 153*, 2592−2599.
48. Miossec, C., Decoen, M. C., Durand, L., Fassy, F., and Diu-Hercend, A. (1996) *Eur. J. Immunol. 26*, 1032−1042.
49. Cain, K., Brown, D. G., Langlais, C., and Cohen, G. M. (1999) *J. Biol. Chem. 274*, 22686−22692.
50. Bratton, S. B., Walker, G., Srinivasula, S. M., Sun, X.-M., Butterworth, M., Alnemri, E. S., and Cohen, G. M. (2001) *EMBO J. 20*, 998−1009.
51. Stennicke, H., and Salvesen, G. (2000) *Biochim. Biophys. Acta 1477*, 299−306.
52. Talanian, R. V., Quinlan, C., Trautz, S., Hackett, M. C., Mankovich, J. A., Banach, D., Ghayur, T., Brady, K. D., and Wong, W. W. (1997) *J. Biol. Chem. 272*, 9677−9682.
53. Thornberry, N. A., Rano, T. A., Peterson, E. P., Rasper, D. M., Timkey, T., Carcia-Calvo, M., Hotzager, V. M., Nordstrom, P. A., Roy, S., Vaillancourt, J. P., Chapman, K. T., and Nicholson, D. W. (1997) *J. Biol. Chem. 272*, 17907−17911.
54. Stennicke, H., Renatus, M., Meldal, M., and Salvesen, G. (2000) *Biochem. J. 350*, 563−568.
55. Cygler, M., and Mort, J. S. (1997) *Biochimie 79*, 645−652.
56. Wiederanders, B. (2000) *Adv. Exp. Med. Biol. 477*, 261−270.
57. Baker, D., Shiau, A., and Agard, D. (1993) *Curr. Opin. Cell Biol. 5*, 966−970.
58. Shi, X.-P., Chen, E., Yin, K.-C., Na, S., Garsky, V., Lai, M.-T., Li, Y.-M., Platchek, M., Register, R. B., Sardana, M. K., Tang, M.-J., Thiebeau, J., Wood, T., Shafer, J. A., and Gardell, S. J. (2001) *J. Biol. Chem. 276*, 10366−10373.
59. Butt, A. J., Harvey, N. L., Parasivam, G., and Kumar, S. (1998) *J. Biol. Chem. 273*, 6763−6768.
60. Li, Y., Hu, Z., Gordon, F., and Inouye, M. (1995) *J. Biol. Chem. 270*, 25127−25132.
61. Smith, B., and Franzen, S. (2001) *Anal. Chem.* (submitted for publication).
62. Bose, K., and Clark, A. C. (2001) *Biochemistry 40*, 14236−14242.

BI011037E

Appendix B:

Cloning of caspase-3 variants and subunits

The purpose of the cloning is to provide the laboratory different versions of caspase-3 to investigate its folding pathway. A total of 9 different clones were generated. The clones are named after their size. The protein with the mutation at the catalytic cyteine 163 to serine is named "01." The schematic diagram is shown in Figure 54.

In general, the variants of caspase-3 with the catalytic cysteine were obtained from PCR amplification by using the designed PCR primers and the pHC332 (full length caspase-3) as template. The 5' end of the gene contains an *Nde*I restriction site and the 3' end of the gene contains an *Xho*I restriction site. The PCR solution (50 μl) includes 1X PCR buffer, dNTP (25 mM), Taq polymerase (2.5 unit), template DNA (~25 ng), $MgCl_2$ (4 mM), and forward and reversed PCR primers (2 μM each). The PCR process was started from denaturing the DNAs at 95 °C for 5 min, repeated for 28 of the PCR cycles (95°C for 30 sec, 55 °C for 1 min, then 68 °C for 2 min), then, finished with an extension time of 5 min at 68°C. The amplified gene was then purified and cloned into pET21b (Novagen) by *Nde*I and *Xho*I sites. The resulting protein contains a C-terminal His-tag. The variants of caspase-3 with the catalytic cysteine mutation were also obtained from PCR following the description above expect using pHC3201 as template. All the clones are confirmed by sequencing both strands of the entire gene.

Table 5. The PCR primers and conditions of caspase-3 variants

Plasmid/Protein Name	forward primer	reverse primer	Size of the gene (base pair)	With or without C163S
P29	HCP17F	HCP12R	~747	C163
P2901	HCP17F	HCP12R	~747	C163S
P19	HCP32F	HCP19R	~525	C163
P1901	HCP32F	HCP19R	~525	C163S
P17	HCP17F	HCP17R	~441	C163
P1701	HCP17F	HCP17R	~441	C163S
P12	HCP12F	HCP12R	~306	C163

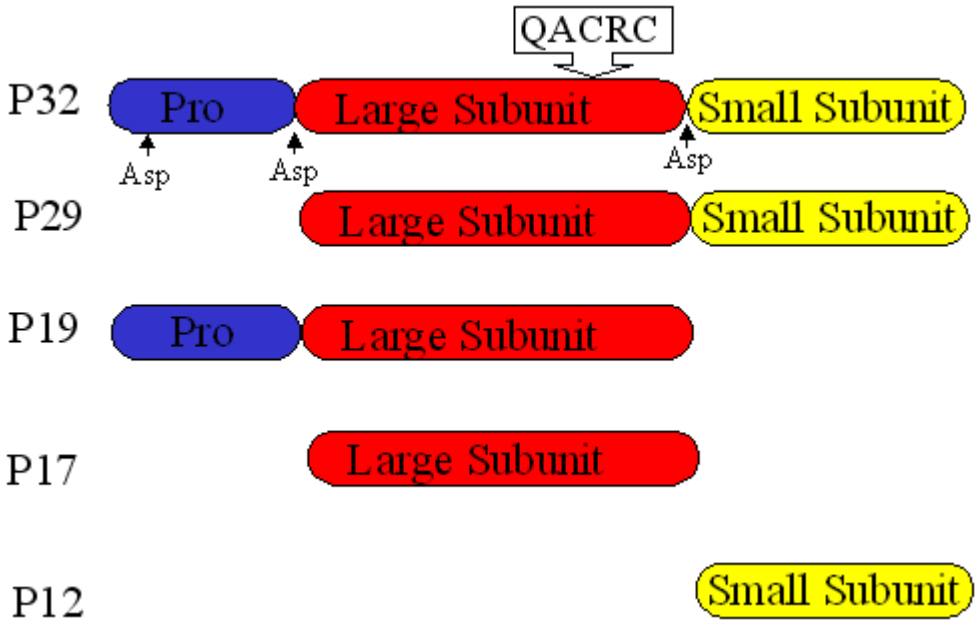


Figure 54. The different proteins of caspase-3 constructed by subcloning. The pro-domain (Pro), large subunit (P17), and small subunit (P12) are labeled. The specific aspartate cleavage sites are indicated. The proteins are indicated by their molecular weights (in kDa) on the left. The active site residues (QACRG) of caspase-3 are shown above P32.

Appendix C:

Folding and Assembly of the Caspase-3 Subunits

C.1 Overview

Caspase-3 (Yama, CPP32), a heterotetrameric enzyme, is the key executioner protease that serves to amplify apoptotic signals. The two subunits of caspase-3, P17 and P12, have molecular weights of 17,058 and 12,957 respectively. Caspase-3, a down-stream caspase, cleaves a number of structural proteins, kinases, DNA repair enzymes, and cell cycle control proteins. Caspase-3 is activated by the upstream caspases. Caspase-3 is produced initially as an inactive zymogen, procaspase-3, which contains a short pro-domain (28 amino acids) at its N-terminus and is followed by the P17 and P12 subunits.

Maturation of the procaspases may be viewed as two sequential pathways (1) folding of the procaspase to a conserved native conformation followed by (2) limited proteolysis in which the two subunits are excised while the active heterotetramer can be assembled *in vitro* from the two subunits. In common with other caspases, procaspase-3 is activated following limited proteolysis in an interdomain linker segment at the highly conserved Asp-175. Cleavage at Asp-28 removes the pro-domain to release the two subunits. It is suggested that a cleavage occurring at Asp-9 within the prodomain prior to the cleavage at Asp-28. It is not yet known at which point in maturation the subunits associate to give a mature tetrameric enzyme. Although in upstream caspases the long pro-domain region is involved in protein:protein interactions (such as CARD-CARD interaction); however, the role of the short pro-domain (28 amino acids) of caspase-3 is not clear.

The objective to study refolding kinetics of caspase-3 is to investigate the maturation pathways of procaspase-3. In order to begin to study the two pathways in maturation, we are examining the folding of the subunits as well as formation of the heterotetramer from the individual subunits.

C.2 Methods: Refolding experiments in the presence of substrate

The refolding buffer for the subunits contains 50 mM HEPES, pH 7.4, 100 mM NaCl, 10 mM DTT, 1 mM EDTA, 10% glycerol, 0.1% CHAPS. The fluorogenic substrate, Ac-DEVD-AFC (Acetyl-Asp-Glu-Val-Asp-7-amido-4-trifluoromethyl-coumarin) was used to monitor the activity of the active caspase-3. Active caspase-3 can cleave at the second aspartate and release the fluorogenic group, AFC. The concentrations of the subunits of caspase-3 are varied and indicated in Figure 55. Fluorescence emission of AFC of the substrate was measured at 505 nm, and the samples were excited at 400 nm.

In the first experiment (as shown in Figure 55A), we refold the two subunits, P17 and P12 in the presence of the substrate. P17 and P12 subunits were in 50 mM Tris-HCl, pH 7.5 containing 2 M GdnHCl, then, were refold by a 1:10 rapid dilution into refolding buffer that included 25 μM of the fluorogenic substrate, Ac-DEVD-AFC.

In the second experiments (as shown in Figure 55B), refolding was initiated as described in the first experiments but without substrate in the buffer. After refolding for 1 hour, substrate was added to a final concentration of 25 μM, and release of AFC was monitored.

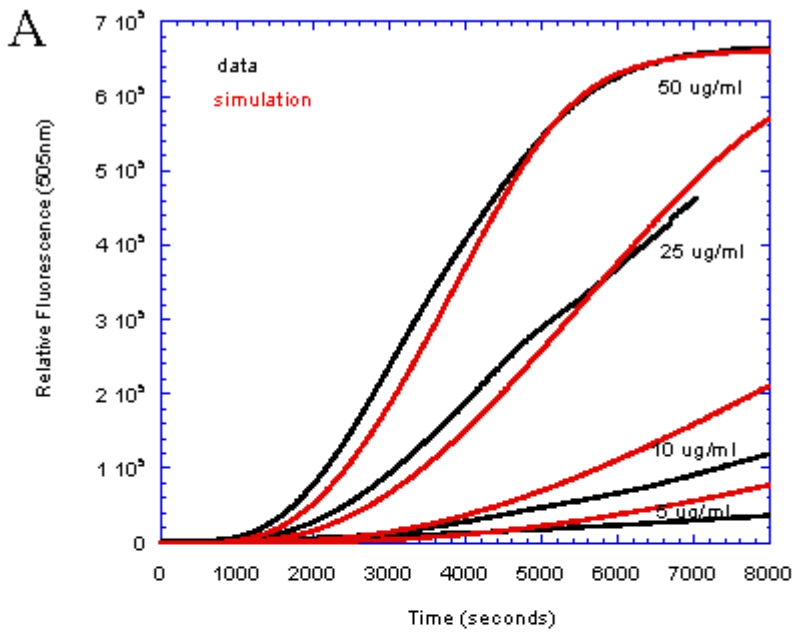
The data were simulated using the program Kinsim (Barshop et al., (1993) *Anal. Biochem*. 130, 134-145), the models were shown in Figure 56, and the rate constants were shown in Table 6. Results of the simulations are shown as the continuous red lines.

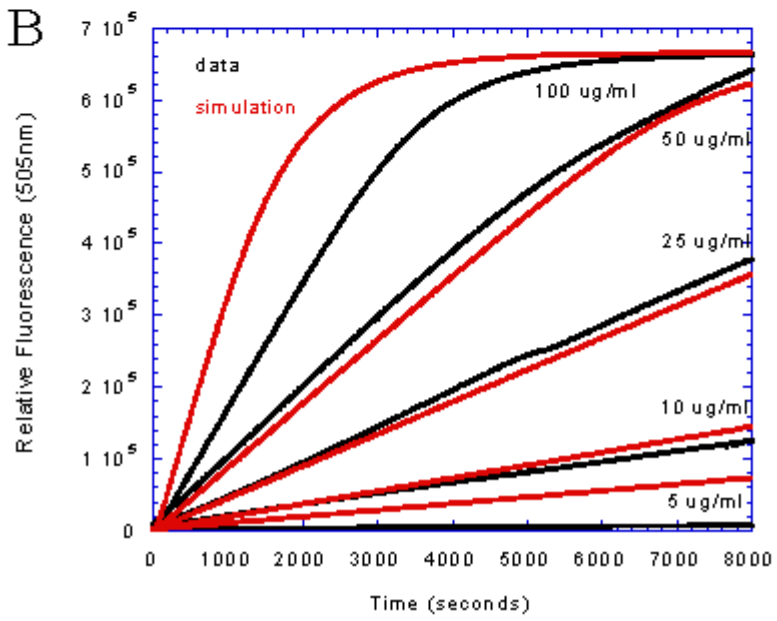
C.3 Results

Using heterologous expression systems, we have expressed and purified separately the two subunits as described in Appendix A. This allows us to examine subunit folding versus association as independent events. I have refolded the two subunits and monitored the formation of active enzyme at several protein concentrations with or without the presence of substrate. The experimental details are described in section C.2.

The data show that folding and assembly are slow processes. Also, the first phase was a long lag for approximately 1000 seconds; the second phase showed increase in activity and was strongly dependent on the protein concentration. The data were simulated by the kinetic simulation program, Kimsim. The diagram and rate constants are shown in Figure 56 and Table 6.

Overall, we are able to propose a kinetic mechanism of the folding and assembly of the two subunits based on the simulation. The simulation of the data indicates at least two first-order reactions in the folding of the subunits prior to assembly followed by two second-order assembly reactions to form the heterotetramer. In addition, the data suggest a transiently active form of caspase-3.


A
data
simulation
50 ug/ml
25 ug/ml
10 ug/ml
5 ug/ml
Relative Fluorescence (505nm)
Time (seconds)
0
1000
2000
3000
4000
5000
6000
7000
8000


B
data
simulation
100 ug/ml
50 ug/ml
25 ug/ml
10 ug/ml
5 ug/ml
Relative Fluorescence (505nm)
Time (seconds)
0
1000
2000
3000
4000
5000
6000
7000
8000

C

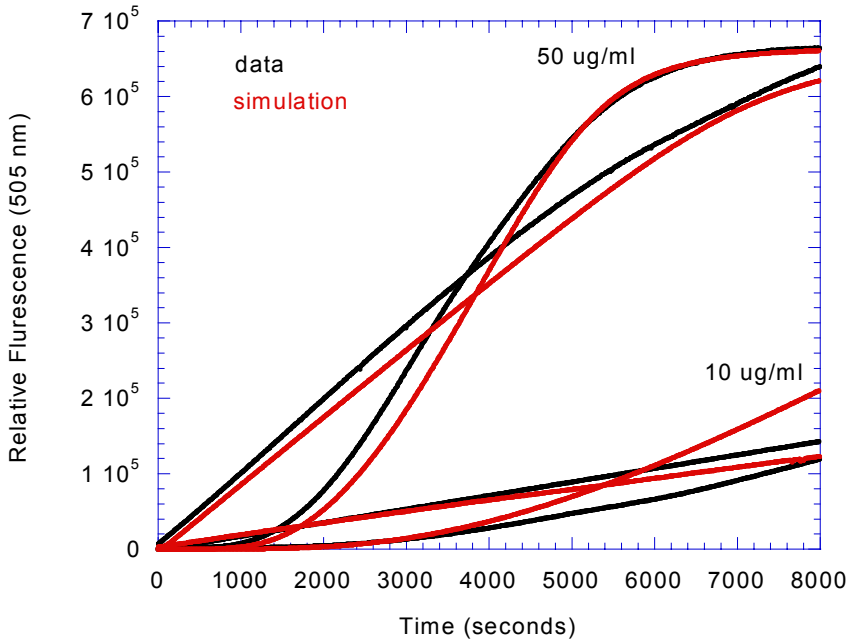


Figure 55. Refolding of caspase-3 and the activity. (A**)** The P17 and P12 subunits were rapidly diluted in the refolding buffer with the presence of the substrate, Ac-DEVD-AFC. The final concentrations of subunits are shown in the figure. Fluorescence emission was measured at 505 nm, and the samples were excited at 400 nm. (B**)** Refolding was initiated but without the presence of the substrate. After refolding for >1 hour, substrate was added, and release of AFC was monitored. (C) Comparison of the 10 and 50 μg/ml data from panels A and B. The black lines are experimental data. The data were simulated using the program Kinsim with the models shown in Figure 56. The rate constants were shown in Table 6. Results of the simulations are shown as the continuous red lines.

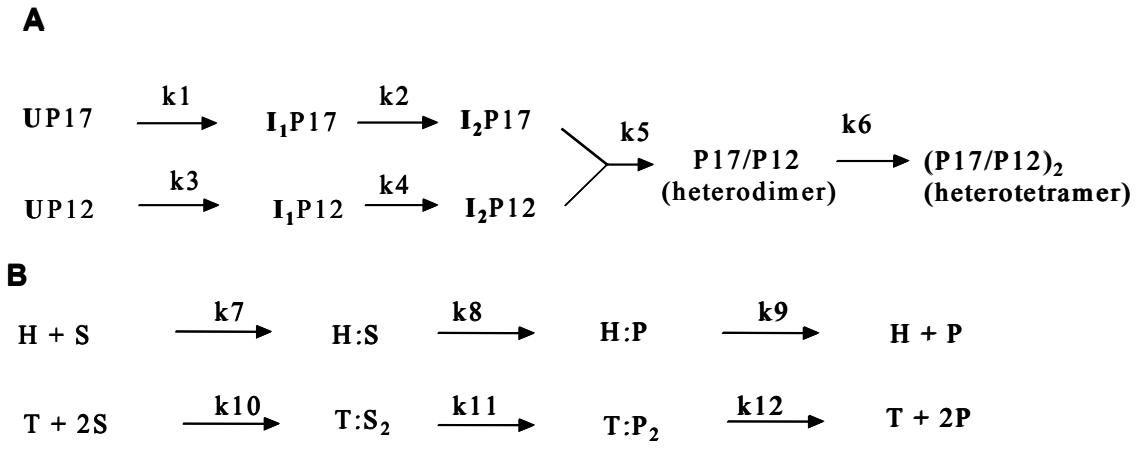


Figure 56. Kinetic model of caspase-3. (A) Model for folding and assembly of the caspase-3 subunits. (B) A continuation of panel A to include enzyme catalysis of the heterodimer (H) and tetramer (T). S refers to the substrate, and P refers to the product, AFC. The rate constants, determined from the data shown in Figure 54 and the model shown here, are given in Table 6.

Table 6. Rate Constants for Folding and Activity of Caspase-3

Rate Constant	(sec^{-1})
k_1	3×10^{-3}
k_2	4×10^{-4}
k_3	3×10^{-3}
k_4	4×10^{-4}
k_5	1×10^{3} $M^{-1}s^{-1}$
k_6	0.5×10^{3} $M^{-1}s^{-1}$
k_7	1.5×10^{3} $M^{-1}s^{-1}$
k_8	1×10^{-2}
k_9	2×10^{-2}
k_{10}	0.4×10^{3} $M^{-1}s^{-1}$
k_{11}	1.2×10^{-3}
k_{12}	2×10^{-2}

Appendix D:

Titrator Operation

D.1. Ribonuclease A

Equilibrium folding studies were done by tritrator operation (Olis Inc.). The titrator operation was examined by urea denaturation using RNase A XII (Sigma). The unfolding experiments were done several times, and the result is consistent with the published data (Pace, N. C. et. al. (1990). *Biochemistry* 29, 2564-2572). The experiments were performed at 25 °C. The concentration of RNase A is 0.65 mg/ml in a buffer containing 30 mM sodium formate (Sigma) at pH 3.55. Two protein samples were prepared: the denatured protein in sodium formate buffer containing 9 M urea and the native protein in sodium formate buffer. By setting the computer macro in software Felix, desired amount of the denatured protein would be titrated into the native protein and mix four times by trafficking the mixture back and forth in the tubing. The equilibration time after mixing was set to 180 sec. Then the sample was excited at 280 nm and the fluorescence emission scan from 300 to 400 nm was taken. On the other hand, refolding experiments were also performed. The equilibrium refolding experiment was done by mixing the protein in urea-containing buffer with the native buffer. Extra washing was required for refolding. The tubing should be washed with the buffers in the reservoirs at least for three times. The washing should include the connections between syringes and reservoirs and the connection between syringe 1 and cuvette. The cartoon of the titrator is shown in Figure 57.

After the emission scans were collected, the fluorescence signals at 350 nm are normalized by the native and unfolded protein signals. The resulting denaturation curves are plotted in Figure 58. The data were fit to a two state equilibrium mechanism. The ΔG and m-values are approximately 4.5 kcal/mol and 1.5 kcal/mol/M , respectively. The $[urea]_{1/2}$ is around 3M with is close to the published data, 3.2M (Pace, N. C. et. al. (1990). *Biochemistry* 29, 2564-2572).

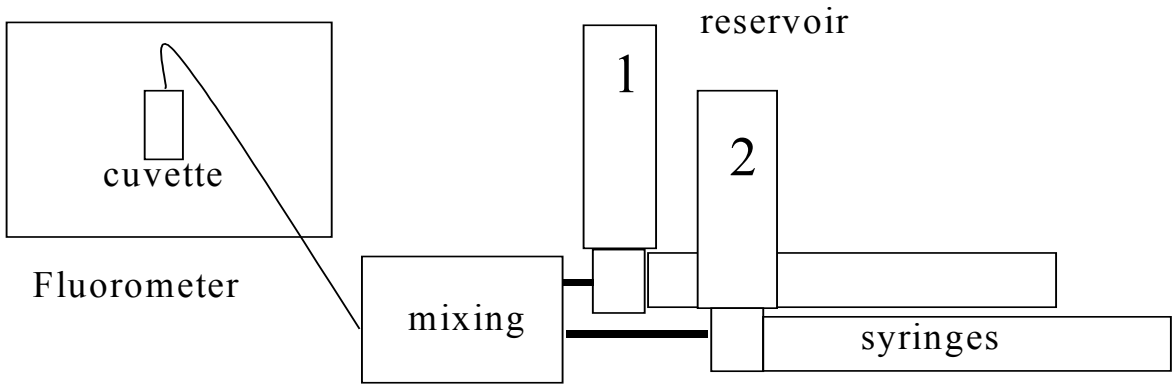


Figure 57. A cartoon of the titrator.

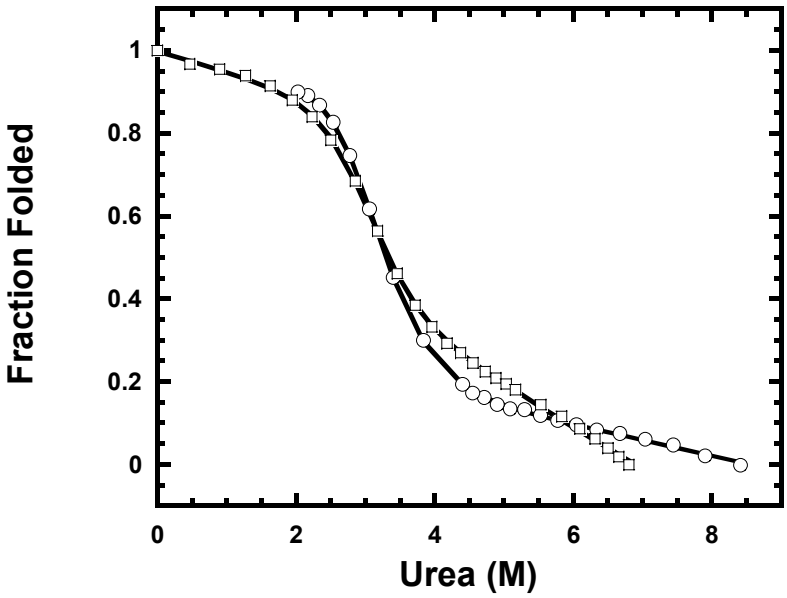


Figure 58. The equilibrium folding of RNase A. The data from the equilibrium unfolding are labeled in squares and the data from equilibrium refolding experiments are labeled in circles.

D.2. RICK-CARD

Protocol for urea denaturation studies of RICK-CARD

1. Make buffers precisely. Make 60 ml of Tris buffer (30 mM Tris, pH 8, 1 mM DTT with desired salts) and 20 or 10 ml of urea buffer (10 M urea in Tris buffer). The concentrations of salts in the buffers are calculated to reach the desire final concentration considering a dilution after mixing with the protein stock. Minimize addition of NaOH while adjusting pH.
2. Filtered the buffers through 0.22um filters. Make the two protein stocks, vortex immediately and let it sit for 30min. Also make the blank buffers.
3. Turn on the fluorometer. Lamp voltage at 75 volt and the photomultiplier voltage at 1000 volt. Temperature at 25 °C.
4. Clear the existed files.
5. Unclick the “always notice” option in Hardware\titrator\controller\set up. Also calibrate titator by Tris buffer. Calibrate the wavelengths of the fluorometer. And fill in the numbers shown in excitation and emission windows on the machine.
6. Run the following programs: “Clean 1” (check leftover in the cuvette) and “Clean 2” with Tris buffer.
7. Run program “wash 2” twice by filling syringe 2 with high molar urea blank buffer (ex: 6 M).
8. Fill syringe 2 with urea blank and run program “blkcard” (under folder Ruby)
9. After program “blkcard” is done, run clean 1, check leftover and run program “wash 2” again.

10. Put the protein in buffer in syringe 1 and the protein in urea buffer in syringe 2
11. Run programs “load 1” then “load 2”
12. Run program “cr-RICK” (make sure the total and individual volumes of titrant added are correct, as well as the number of loops)
13. After the experiment is complete, wash again with water with program “clean 1 and 2”
14. Fill the cuvette with .1 N HCl to allow cleaning.

D.3 Macro program of titrator operation of RICK-CARD

The computer macro program of RICK-CARD for the equilibrium unfolding experiments is listed as follows:

(file location: C:/Felix/Ruby/crrick.mac)

```
"Version1.42"
"Set Acquisition Mode   Emission Scan"
3
"User Name  Ruby"
"Ruby"
"Sample Name      pCardrick"
"pCardrick"
"Comments   Titration at 1.5uM"
"Titration at 1.5uM"
"File Root Name   C:\FELIX\RUBY\TITRATOR\CRD01.FLX"
"C:\FELIX\RUBY\TITRATOR\CRD01.FLX"
"Titrator Operation     Prepare for Titration 1000µL 3000µL Fill also
Tubes"
20 1000 3000 1
"Titrator Operation     Add to Cuvette from Sample Syringe (1) 900µL"
30 900 0 0
"Emission Scan Setup     280  300  300  100  1  .5  1"
280 300 300 100 1.5 1 0 25
"Prep/Start Acquisition"
"Increment EX Wavelength      Index:1 Value:15"
"1" "15"
"Prep/Start Acquisition"
"Titrator Operation     Mixing (For Titration) 0µL 5sec 4"
21 0 5 4
"Titrator Operation     Add to Cuvette from Sample Syringe (1) 900µL"
```

```
30 900 0 0
"Save Data File"
"Start Loop"
"Emission Scan Setup     280  300  300  100  1  .5  1"
280 300 300 100 1 .5 1 0 25
"Prep/Start Acquisition"
"Increment EX Wavelength      Index:1 Value:15"
"1" "15"
"Prep/Start Acquisition"
"Titrator Operation     Mixing (For Titration) 40µL 5sec 4"
21 40 5 4
"Time Out   Value: 1500"
1500
"Titrator Operation     Add to Cuvette from Sample Syringe (1) 900µL"
30 900 0 0
"Save Data File"
"Number of loops  Value: 13"
13
"Start Loop"
"Emission Scan Setup     280  300  300  100  1  .5  1"
280 300 300 100 1 .5 1 0 25
"Prep/Start Acquisition"
"Increment EX Wavelength      Index:1 Value:15"
"1" "15"
"Prep/Start Acquisition"
"Titrator Operation     Mixing (For Titration) 60µL 5sec 4"
21 60 5 4
"Time Out   Value: 1500"
1500
"Titrator Operation     Add to Cuvette from Sample Syringe (1) 900µL"
30 900 0 0
"Save Data File"
"Number of loops  Value: 13"
13
"Start Loop"
"Emission Scan Setup     280  300  300  100  1  .5  1"
280 300 300 100 1 .5 1 0 25
"Prep/Start Acquisition"
"Increment EX Wavelength      Index:1 Value:15"
"1" "15"
"Prep/Start Acquisition"
"Titrator Operation     Mixing (For Titration) 280µL 5sec 4"
21 280 5 4
"Time Out   Value: 1500"
1500
"Titrator Operation     Add to Cuvette from Sample Syringe (1) 900µL"
30 900 0 0
"Save Data File"
"Number of loops  Value: 6"
6
"Emission Scan Setup     280  300  300  100  1  .5  1"
280 300 300 100 1 .5 1 0 25
"Prep/Start Acquisition"
"Increment EX Wavelength      Index:1 Value:15"
```

```
"1" "15"
"Prep/Start Acquisition"
"Save Data File"
"Display Message  experiment is complete"
```
"experiment is complete"

Appendix E:

Calculation of "Long range order" (LRO)

The window based long-range order program is created by Dr. Yun-Chorng Chang according to the definition (see Chapter 1). The procedure of using the LRO program is as follows:

1. Download the PDB file
2. Open the PDB file by Microsoft Excel, choose delimited, then choose "space" to separate the data into columns
3. Sort the data by Excel/Data/Sort so that the alpha carbon for all residues will be in the same place. (Usually sort column 3)
4. Cut and paste four columns to a new excel file. The four columns include the followings in order: amino acid 1 to n, x coordinate, y coordinate, z coordinate.
5. Save the new excel file as a text file
6. Open the LRO program and open the new text file.
7. The values will show in the dialog box. The LRO value here is half of that in the LRO paper (Gromiha, M. M. & Selvaraj, S. (2001). *J Mol Biol* 310, 27-32).

For examples: Pro-1-CARD LRO=0.5*2=1, RICK-CARD LRO=0.7065*2=1.413

Appendix F:

Calculation of three-state mechanism (N ⇆ I ⇆ U)

To fit the equilibrium folding data to a three state mechanism, we can start with the basic concept that total property observed, for examples: fluorescence or CD properties, is the sum of the properties from each species at equilibrium. The equation [27] is shown. *P* is the properties and *f* is the fraction of the species. The subscripts are the species. *N* is the native state, *I* is the intermediate state, and *U* is the unfolded state.

$$P_{total} = f_N P_N + f_I P_I + f_U P_U \quad [27]$$

The equilibrium constant between *N* and *I* is K_1 and that between *I* and *U* is K_2.

$$K_1 = \frac{[I]}{[N]} \quad [28]$$

$$K_2 = \frac{[U]}{[I]} \quad [29]$$

Therefore, equation [27] can be rewrite as the follows equations:

$$P_{total} = \frac{[N]}{[N]+[I]+[U]} P_N + \frac{[I]}{[N]+[I]+[U]} P_I + \frac{[U]}{[N]+[I]+[U]} P_U \quad [30]$$

Put in the equilibrium constants, K_1 and K_2 based on equations [28] and [29].

$$P_{total} = \frac{[N]}{[N]+K_1[N]+K_1K_2[N]} P_N + \frac{K_1[N]}{[N]+K_1[N]+K_1K_2[N]} P_I + \frac{K_1K_2[N]}{[N]+K_1[N]+K_1K_2[N]} P_U$$

[31]

$$P_{total} = \frac{1}{1+K_1+K_1K_2}P_N + \frac{K_1}{1+K_1+K_1K_2}P_I + \frac{K_1K_2}{1+K_1+K_1K_2}P_U \quad [32]$$

Since the equilibrium constant can be represented as a function of the free energy,

The following equations will substitute K_1 and K_2 in equation [32].

$$K_1 = \exp(-\Delta G_1 / RT) \quad [33]$$

$$K_2 = \exp(-\Delta G_2 / RT) \quad [34]$$

where R is the gas constant, T is the absolute temperature, ΔG_1 is the free energy between N and I states, and ΔG_2 is the free energy between I and U states.

In addition, several parameters will be put into equation [32] for the fitting because of the dependence of the denaturant (For examples: urea). The subscript, H_2O, of the parameter indicates the condition is in the absence of urea. They are listed as follows:

$$\Delta G_1 = \Delta G_{1(H_2O)} - m_1(urea)$$

$$\Delta G_2 = \Delta G_{2(H_2O)} - m_2(urea)$$

$$P_N = P_{N(H_2O)} + m_3(urea)$$

$$P_U = P_{U(H_2O)} + m_4(urea)$$

$$P_I = m_5$$

ΔG, P, and m are the parameters that will be obtained from the fitting.