

ABSTRACT

MILAM, SARA LIS. Kinetic Folding Studies of Apaf-1 CARD and Procaspase-3. (Under the direction of A. Clay Clark.)

Apoptosis regulates the balance between cell growth and death. Apoptosis consists of two main signaling pathways, extrinsic and intrinsic. We have examined the kinetic folding mechanism of two proteins, Apaf-1 CARD and procaspase-3, which play a major part in these signaling cascades.

The caspase recruitment domain (CARD) of Apaf-1 (apoptotic protease activating factor) is the 97 amino acid N-terminal domain involved in protein-protein interactions, allowing for the activation of procaspase-9. Apaf-1 CARD consists of six antiparallel α -helices arranged in a Greek key topology. Single and sequential mixing stopped-flow studies showed that Apaf-1 CARD folds and unfolds rapidly and suggest a folding mechanism that contains parallel channels with two unfolded conformations folding to the native conformation. KINSIM simulations show that a slow folding phase is described by a third conformation in the unfolded ensemble that interconverts with one or both unfolded species. Overall, the native ensemble is formed rapidly upon refolding.

Procaspase-3, like all other caspases, exists in the cell as an inactive zymogen and is the final step in the apoptotic pathway. Once activated it cleaves numerous substrates, leading to the dismantling of the cell. The refolding pathway of homodimeric procaspase-3 is complex, consisting of multiple monomeric intermediates with a slow rate of dimerization. The refolding and unfolding burst phase revealed multiple species formed within milliseconds of folding. Kinetic data support the hypothesis of two native conformations, one

of which is enzymatically active. Collectively, these results demonstrate that dimerization is an important aspect in both folding and activation of procaspase-3.

Overall, the kinetic folding data for Apaf-1 CARD and procaspase-3 provide an improved picture of the function and regulation of apoptosis. These studies propose new targets for therapeutic design to combat diseases associated with apoptosis.

Kinetic Folding Studies of Apaf-1 CARD and Procaspase-3

by
Sara Lis Milam

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

Biochemistry

Raleigh, North Carolina

2008

APPROVED BY:

Michael B. Goshe

Carol K. Hall

William L. Miller

A. Clay Clark
Chair of Advisory Committee

DEDICATION

To my parents, Julianne and John, for their love and support throughout my life and my
brother, Mark, for his perception and humor.

BIOGRAPHY

Sara Milam grew up in Indianapolis, IN where she attended Cathedral High School. She graduated from Miami University in May 2003 with a Bachelor of Science degree in Biochemistry. In the summer of 2003, she moved to Raleigh, NC in pursuit of a Ph.D. in biochemistry at North Carolina State University under the direction of A. Clay Clark. In 2007 she won the Anne A. J. Work Award given in recognition of outstanding graduate achievement in research by women in Biochemistry at North Carolina State University. While in Raleigh, she was able to rekindle her love of soccer by playing in the Triangle Adult Soccer League. Following completion of her Ph.D., Sara hopes to have a successful career in biochemistry / biophysics.

ACKNOWLEDGMENTS

I would like to thank Dr. A. Clay Clark for his guidance and support through some frustrating times and his commitment to correcting my vocabulary. I would also like to thank my committee members, Dr. Michael Goshe, Dr. William Miller and Dr. Carol Hall, for their advice and helpful feedback.

I would like to thank my family for their unconditional love. My parents, John and Julianne, have always bent over backwards for my education and pushed me to broaden my horizons (even when they do not understand the science world). The lab will also miss the cookie donations. I would like to thank my brother, Mark, for teaching me patience. You could always make me smile with your quick wit and humor. I expect big things from you!! I would also like to thank my second family, the Dyers (Debsie, Tom, Emily and Claire), for all of their encouragement.

I would like to thank my friends (too many to mention here) for their support and entertainment throughout the years. I would also like to thank my fellow Clay-mates; Cristina Pop, Justin Stern, Ruby Chen, Brett Feeney, Denise Appel, Jad Walters and Sarah MacKenzie. I learned a lot about research from merely listening to the many “discussions” held between Brett and Cristina. Eventually I was the only student in the lab, and Brett helped me survive and flourish. I am the scientist that I am today because of him. I would like to thank Denise for all her companionship in- and outside the lab. I appreciate all the help in editing/writing e-mails, papers and even the dissertation. I would be incoherent

without her. Denise has been a great soundboard for project ideas and life in general. Thanks D! I would like to thank Jad for being my counterpart in science and novel lab games. Jadly - “Whatever tears at us, whatever holds us down, and if nothing can be done, we’ll make the best of what’s around” (Dave Matthews Band) and “People I respect, heroes of mine, would be Bob Hope... Abraham Lincoln, definitely. Bono. And probably God would be the fourth one. And I just think all those people really helped the world in so many ways that it's really beyond words. It's really incalculable.” (Michael Scott, The Office). Thank you for all the entertainment and support throughout the years. Your friendship is truly incalculable.

TABLE OF CONTENTS

LIST OF TABLES	xi
LIST OF FIGURES	xiii
CHAPTER I. Introduction	1
A. The importance of cell death.....	2
B. The apoptotic pathway	3
B.1. Extrinsic pathway.....	3
B.2. Intrinsic pathway.....	5
C. Apaf-1.....	5
C.1. Death domain superfamily	5
C.2. Apaf-1 domain.....	6
C.3. Apaf-1 CARD	6
D. Caspases in apoptosis.....	8
D.1. Inactive oligomeric state of the protein.....	9
D.2. Prodomain: length and function	10
D.3. Function in apoptosis	10
D.4. Maturation process	12
D.5. The importance of protein dimerization in cell death.....	12
E. Procaspase-3	14
F. Project aims	15

REFERENCES	17
CHAPTER II. Practical Approaches to Protein Folding and Assembly: Spectroscopic Strategies in Thermodynamics and Kinetics.....	19
ABSTRACT.....	20
INTRODUCTION	21
MEASURING FOLDING KINETICS	24
A. Experimental Protocol.....	24
A.1. General Considerations	24
A.2. Initial Parameterization	26
A.3. Sample Preparation for Measuring Refolding and Unfolding Kinetics	28
A.4. Instrument Procedure	29
B. Differential Quenching by Acrylamide.....	31
B.1. Sample Preparation	32
B.2. Experimental Procedure	33
C. Data Analysis	33
C.1. Burst Phase.....	34
C.2. Exponential Fits.....	35
C.3. Simulations.....	37
REFERENCES	42

CHAPTER III. Rapid Folding and Unfolding of Apaf-1 CARD	54
ABSTRACT.....	55
INTRODUCTION	56
RESULTS	59
A. Structural Studies	59
B. Equilibrium Unfolding	60
C. Single Mixing Refolding Studies	61
D. Single Mixing Unfolding Studies	64
E. Proposed Folding Models.....	67
F. Sequential Mixing Studies	69
G. Kinetic Simulations.....	72
DISCUSSION	77
METHODS	82
ACKNOWLEDGMENTS	87
REFERENCES	88
CHAPTER IV. Folding and Assembly Kinetics of Procaspase-3	109
ABSTRACT.....	110
INTRODUCTION	111
RESULTS	115
A. Multiple Monomeric Species are Formed in the Burst Phase.....	116

B. Procaspase-3 Folds through Multiple Monomeric Intermediates	118
C. Dimerization Occurs Slowly in the Folding of Procaspase-3	123
D. Proposed Folding Models for Procaspase-3	124
E. Unfolding Data Reveal Multiple Species	125
F. Procaspase-3 Dimer Dissociates Slowly	128
DISCUSSION	130
METHODS	136
REFERENCES	141
CHAPTER V. Conclusion	167
APPENDICES	170
APPENDIX A. Apaf-1 CARD Supplemental Data.....	171
I. Analytical ultracentrifugation.....	172
II. Circular dichroism spectroscopy	174
APPENDIX B. Procaspase-3 Supplemental Data	177
I. Chemical cross-linking	178
II. Native gels.....	188
III. Procaspase-3 labeling with dansyl chloride.....	190
IV. Anisotropy	192

V. Collaboration with David A. Johnson at University of California,	
Riverside	197
VI. Differential quenching by acrylamide	210
VII. Procaspase-3 D3A studies	213
A. Activity assays	213
B. Kinetic studies	214
APPENDIX C. Protocols	217
I. Preparation of 10 M urea stock.....	218
II. Preparation of 8 M guanidine hydrochloride stock.....	220
III. KINSIM	223
IV. Derivation of two-state equilibrium folding model	232
V. Derivation of three-state equilibrium folding model	235
VI. Anisotropy calibrations.....	239

LIST OF TABLES

CHAPTER I

Table I. Diseases with Dysregulated Apoptosis	2
---	---

CHAPTER II

Table I. Example Calculations for Urea and 0.77 M

Acrylamide/Urea Stock Solutions.....	46
--------------------------------------	----

Table II. Example Calculations for 55 μ M Procaspase-3 Protein Stock Solutions for

Refolding/Unfolding Kinetic Experiments	46
---	----

CHAPTER III

Table 1. Summary of statistics for data collection and

refinement for Apaf-1 CARD	93
----------------------------------	----

Table 2. Summary of thermodynamic parameters for Apaf-1 CARD	94
--	----

Table 3. Summary of calculated and simulated rates using Scheme 3	95
---	----

CHAPTER IV

Table I. Refolding burst phase signals fit to two- or three-state

equilibrium folding models.....	145
---------------------------------	-----

Table II. Unfolding phases and detection methods.....	145
---	-----

Table III. Comparison of unfolding burst phase species and proposed native species.....	146
--	-----

APPENDICES

APPENDEIX B

Table 1. Protein concentration dependence and plateau anisotropy for procaspase-3 mutants from Figure 15	200
---	-----

LIST OF FIGURES

CHAPTER I

Figure 1. Apoptosis.....	4
Figure 2. Apaf-1 structure.....	7
Figure 3. Caspase domains	9
Figure 4. Adaptor protein complexes are responsible for the activation of initiator caspases.....	11
Figure 5. Caspases: architecture and activation.....	13
Figure 6. Homology model of procaspase-3 (panel A) and crystal structure of caspase-3 (panel B).....	15

CHAPTER II

Figure 1. Data analysis.....	47
Figure 2. The 4.36 M refolding data from Fig. 1B were fit to a one (A), two (B) or three (C) exponential equation	48
Figure 3. Chevron and KINSIM analysis	50
Figure 4. One sequential (A) and two parallel folding pathways (B and C) used in KINSIM simulations	52

CHAPTER III

Figure 1. Structure of Apaf-1 CARD with GS sequence at the amino terminus	96
Figure 2. Equilibrium studies of Apaf-1 CARD	97
Figure 3. Single-mixing stopped-flow fluorescence emission experiments	98
Figure 4. Sequential-mixing stopped-flow fluorescence studies	100
Figure 5. Simulations of single-mixing stopped-flow fluorescence emission experiments.....	101
Supplemental Figure 1. Single-mixing stopped-flow circular dichroism experiments in MTA buffer pH 6.....	102
Supplemental Figure 2. Single-mixing stopped-flow circular dichroism experiments in MTA buffer pH 8.....	104
Supplemental Figure 3. -mixing stopped-flow fluorescence emission experiments in MTA buffer pH 8.....	106
Schemes	108
 CHAPTER IV	
Figure 1. Far-UV CD (228 nm) refolding data	146
Figure 2. Complex burst phase	147
Figure 3. Emission scanning of the burst phase species	149
Figure 4. CD lag.....	150
Figure 5. Rate versus the final urea concentration for fluorescence excitation at 280 nm (panel A) or 295 nm (panel B) and differential quenching by acrylamide (panel C)...	152

Figure 6. Intermediate phase of refolding.....	153
Figure 7. Slow phase of refolding.....	154
Figure 8. Activity studies with procaspase-3 D3A.....	155
Figure 9. Proposed sequential refolding mechanism.....	156
Figure 10. Unfolding kinetic data.....	158
Figure 11. Fluorescence (excitation at 280 nm) burst phase signals versus urea.....	159
Figure 12. Unfolding rates.....	159
Figure 13. Proposed sequential unfolding mechanism.....	160
Figure 14. Double jump anisotropy studies.....	162
Figure 15. Dimer interface sequences of caspase-3 and -8.....	163
Figure 16. Homology models of inactive (N ₂ , panel A) and active (N ₂ *, panel B) procaspase-3.....	164
Supplemental Figure 1. Slow phases in refolding.....	166
Supplemental Figure 2. Unfolding rates.....	166

APPENDICES

APPENDIX A

Figure 1. Sedimentation equilibrium studies of Apaf-1 CARD.....	173
Figure 2. Far- and near-UV CD spectra of Apaf-1 CARD.....	175
Figure 3. Far- and near-UV CD spectra of Apaf-1 CARD.....	176

APPENDIX B

Figure 1. Homology model of procaspase-3.....	179
Figure 2. A) DMS cross-linker; B) DMS cross-linking reaction.....	179
Figure 3. DMS initial parameterization	181
Figure 4. A) BS ³ cross-linker; B) BS ³ cross-linking reaction.....	181
Figure 5. BS ³ cross-linking control reactions	182
Figure 6. Refolding reactions.....	184
Figure 7. 4-25% SDS-PAGE	186
Figure 8. 8-15% SDS-PAGE	187
Figure 9. Native PAGE analysis	189
Figure 10. Equilibrium unfolding experiments of dansyl labeled procaspase-3 monitored by anisotropy.....	191
Figure 11. Anisotropy results.....	193
Figure 12. Stopped-flow anisotropy burst phase studies	195
Figure 13. Equilibrium unfolding detected by anisotropy with excitation at 280 nm (panel A) or 295 nm (panel B)	196
Figure 14. Homology model of procaspase-3.....	200
Figure 15. Equilibrium anisotropy of procaspase-3 C163S (panels A and B), W206Y C163S (panels C and D) and W214V C163S (panels E and F) versus urea concentration.....	201
Figure 16. Far-UV (panel A) and near-UV (panel B) CD spectra of procaspase-3 mutants.....	203

Figure 17. Time-resolved anisotropy data for procaspase-3 mutants under native conditions	205
Figure 18. Absorbance (panel A) and fluorescence emission with excitation at 295 nm (panel B) scans of procaspase-3 mutants	206
Figure 19. Time-resolved anisotropy data for procaspase-3 C163S (blue) and weighted average (red) of W206Y C163S and W214V C163S	208
Figure 20. Differential quenching by acrylamide	212
Figure 21. Activity studies with procaspase-3 D3A	214
Figure 22. Kinetic studies with procaspase-3 D3A	216

APPENDIX C

Figure 1. KINSIM main screen.....	224
Figure 2. KINSIM mechanism.....	225
Figure 3. KINSIM simulations of Apaf-1 CARD.....	231
Figure 4. Procedure for a two-state equilibrium folding model in Igor Pro	235
Figure 5. Procedure for a three-state equilibrium folding model in Igor Pro	238

CHAPTER I

Introduction

A. The importance of cell death

Apoptosis is one of the pathways by which cell death can occur (MacKenzie and Clark, 2008). Cell population is kept relatively constant through the opposing processes of cell death and division. It is, therefore, of crucial importance that the balance between cell death and proliferation be tightly regulated. If this balancing process is skewed in either direction, diseased state results (Table I). For example, excessive apoptosis can lead to neurodegenerative disorders (such as Alzheimer's Parkinson's and Huntington's), cardiovascular disease and autoimmune disorders, such as diabetes. On the other hand, deficient apoptosis can lead to cancer and other autoimmune disorders, such as rheumatoid arthritis (Chamond et al., 1999; Saikumar et al., 1999). Since apoptosis is such a critical pathway in the cell, further understanding of this process is essential for development of therapeutic strategies to combat these diseases.

Table I. Diseases with Dysregulated Apoptosis (Saikumar et al., 1999)

Excessive Apoptosis	Deficient Apoptosis
Degenerative neurological diseases (Alzheimer's, Huntington's, Parkinson's)	Autoimmune lymphoproliferative syndrome (Canale-Smith syndrome)
Aplastic anemia	Graves' disease
Acquired immunodeficiency syndrome	Hypereosinophilia syndrome
Hashimoto's thyroiditis	Hashimoto's thyroiditis
Lupus erythematosus	Lupus erythematosus
Liver failure	Lymphoma
Multiple sclerosis	Leukemia
Myelodysplastic syndrome	Solid tumors
Type I diabetes mellitus	Type I diabetes mellitus
Ulcerative colitis	Osteoporosis
Wilson's disease	Developmental defects
Chronic neutropenia	
Developmental defects	

B. The apoptotic pathway

The apoptotic pathway consists of two main signaling pathways; extrinsic and intrinsic (Figure 1). The extrinsic pathway is initiated by extracellular pro-apoptotic ligands, while the intrinsic pathway is initiated by cellular stress, such as DNA damage, chemotherapeutic agents, cytokines and reactive oxygen species (MacKenzie and Clark, 2008). Both pathways result in the activation of caspase-3, which is the commitment step in cell death (Taylor et al., 2008). Once activated, caspase-3 is able to cleave over 400 substrates involved in a variety of cellular functions ranging from structural proteins to regulators of transcription, ultimately leading to the dismantling of the cell (Luthi and Martin, 2007; Saikumar et al., 1999; Taylor et al., 2008).

B.1. Extrinsic pathway

The extrinsic pathway is initiated by pro-apoptotic ligands, such as FasL (also known as CD95L), which binds to a transmembrane death receptor, such as Fas (also known as APO-1/CD95) (Peter and Krammer, 2003; Riedl and Salvesen, 2007). The death receptors belong to the tumor necrosis factor (TNF) family, which contain a single DD (death effector) in the intracellular compartment. The binding of the death ligand to the receptor promotes oligomerization of the receptor and formation of the death-inducing signaling complex (DISC). The adaptor protein FADD, which contains DED (death effector domain) and DD domains, mediates the interaction between the death receptor and initiator caspase-8 or -10 (Peter and Krammer, 2003). The DISC complex activates caspase-8 and -10 through

dimerization (Bao and Shi, 2006). These proteins are then able to activate downstream executioner caspases.

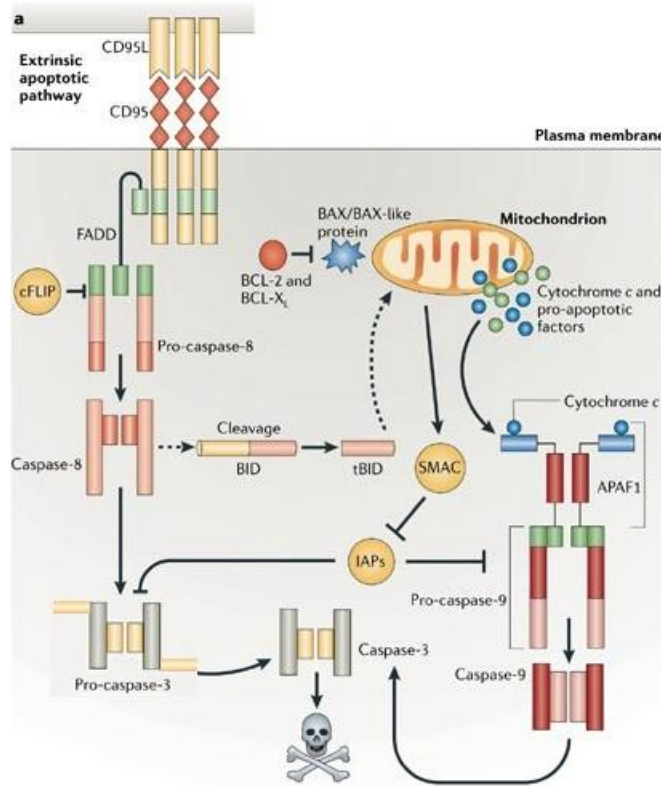


Figure 1. Apoptosis. There are two death pathways: an extrinsic (death-receptor) and intrinsic (mitochondrial) pathway. The extrinsic pathway is mediated by caspase-8 whereas the intrinsic pathway is mediated by caspase-9. FADD (FAS-associated via death domain) is an adapter protein that couples death receptors, such as CD95, to caspase-8. Cytochrome c is released from the mitochondria and together with Apaf-1 and procaspase-9 form the apoptosome. Both pathways activate the downstream caspase, caspase-3. (Hotchkiss and Nicholson, 2006) adapted

B.2. Intrinsic pathway

Initiation of the intrinsic pathway, results in the increase in permeability of the mitochondrial membrane and release of pro-apoptotic factors, such as cytochrome c. Cytochrome c binds to monomeric Apaf-1 and in the presence of dATP/ATP promotes the assembly of the apoptosome. The apoptosome is a ~1 MDa wheel-like particle with seven spokes that radiate from a central hub (Acehan et al., 2002; Riedl and Salvesen, 2007). Once formed, this complex binds procaspase-9, which increases the local concentration of monomeric procaspase-9, promoting dimerization and activation the protein (Renatus et al., 2001; Shiozaki et al., 2002). Caspase-9 is then able to activate procaspase-3.

C. Apaf-1

C.1. Death Domain superfamily

Apaf-1 CARD is a member of the death domain superfamily (Weber and Vincenz, 2001). This superfamily is subdivided into 4 families consisting of CARD (caspase recruitment domain), DD (death domain), DED (death effector domain) and PYD (pyrin domain) (Fairbrother et al., 2001; Weber and Vincenz, 2001). Collectively, these domains mediate homotypic protein-protein interactions in the apoptotic pathway. Within the death domain superfamily there is low sequence identity, but all share a common “death fold” composed of six antiparallel α -helices arranged in a Greek key folding topology (Vaughn et al., 1999; Weber and Vincenz, 2001). The Death Domain superfamily has a similar structural topology with different contact surfaces, which allow for specificity in binding. In addition,

these domains are composed of a hydrophobic core (Vaughn et al., 1999; Weber and Vincenz, 2001).

C.2. Apaf-1 domains

Apaf-1 (apoptotic protease activating factor 1) is a ~130 kDa protein that consists of an N-terminal CARD (caspase recruitment domain), a central ATPase domain and 13 WD40 repeats at the C-terminus (Acehan et al., 2002; Zou et al., 1997) (Figure 2A). Each domain has its own function in the apoptosome. The WD40 repeats bind cytochrome c that is released from the mitochondria upon induction of the intrinsic apoptosis pathway (Riedl and Salvesen, 2007; Zou et al., 1997). The ATPase domain, which consists of Walker's A- and B-box consensus sequences, can then bind dATP/ATP (Zou et al., 1997). Apaf-1 exists as an inactive monomer in the cell. The binding of cytochrome c and dATP/ATP promotes oligomerization of Apaf-1 (Acehan et al., 2002). Finally, the CARD domain of Apaf-1 binds the procaspase-9 CARD in a 1:1 stoichiometry, which allows for the activation of procaspase-9 (Shiozaki et al., 2002).

C.3. Apaf-1 CARD

The N-terminus of Apaf-1 consists of a small, 97 amino acid domain, which exists as a monomer in solution (Shiozaki et al., 2002). Apaf-1 CARD adopts a tightly packed six helix bundle, as shown in Figure 2B (Vaughn et al., 1999). The secondary structure is primarily α -helical and surrounds a hydrophobic core. The six helices are arranged in an antiparallel bundle with Greek key folding topology (Figure 2C) (Weber and Vincenz, 2001; Zhou et al., 1999). The Greek key structural motif was based on a design on Greek vases,

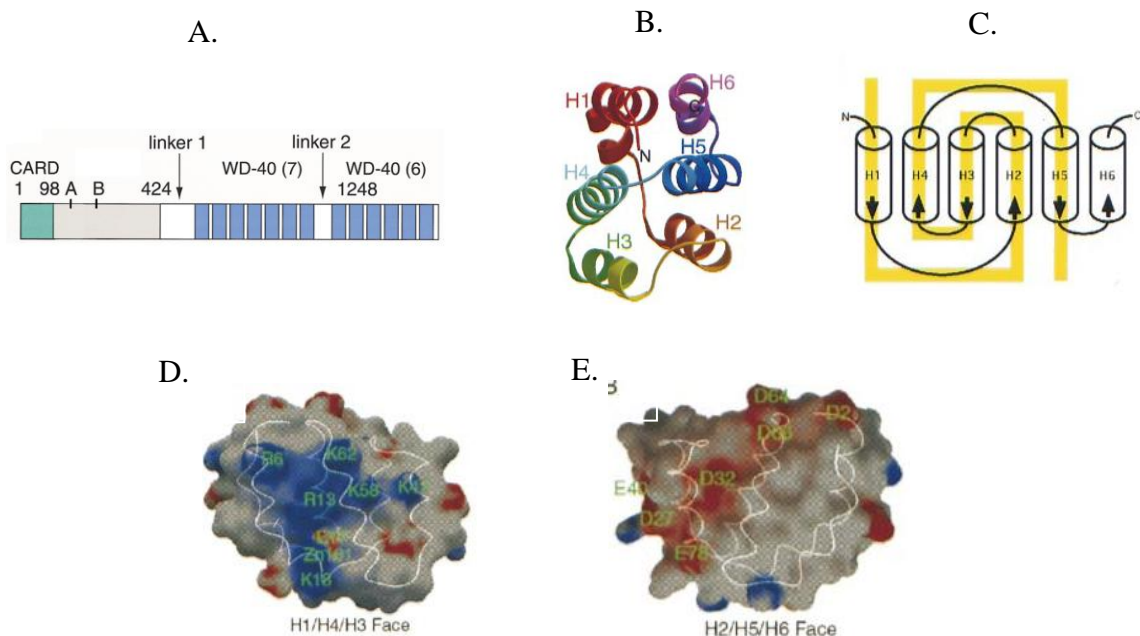


Figure 2: Apaf-1 structure. A) Domain layout of Apaf-1. The protein consists of a CARD (green), ATPase (purple, Walker A- and B-boxes are marked) and 13 WD40 repeats (blue) (Acehan et al., 2002). B) Apaf-1 CARD structure. The 6 α -helices are labeled. C) Greek key motif. Cylinders represent the helices are arrows represent the direction of the polypeptide chain. D and E) The surfaces are semitransparent with the C^α trace of the molecule drawn in a white line. The basic H1/H3/H4 face (panel D) and acidic H2/H5/H6 face (panel E) of Apaf-1 CARD. (Vaughn et al., 1999)

which illustrates the connection of the β -strands or α -helices, in this case (Richardson, 1977a). The surface of Apaf-1 CARD includes both electrostatic and hydrophobic surfaces. The H1/H4/H3 surface is primarily comprised of basic residues, while H2/H5/H6 (the opposing face) is primarily acidic, as shown in Figure 2D and E (Vaughn et al., 1999; Zhou et al., 1999).

Apaf-1 CARD and the CARD of procaspase-9 have been shown to interact in the apoptosome and in solution (Acehan et al., 2002; Shiozaki et al., 2002). The specificity of binding is due to the electrostatic interaction between the complementary charged surfaces on each CARD. Helices 2 and 3 are primarily acidic in Apaf-1 CARD, while helices 1 and 4 of the CARD of procaspase-9 are basic. These four helices form an antiparallel four helix bundle between Apaf-1 CARD and the CARD of procaspase-9 (Weber and Vincenz, 2001; Zhou et al., 1999).

Chapter III focuses of the kinetic pathway of Apaf-1 CARD. As mentioned above, this CARD is completely α -helical. Therefore, far-UV circular dichroism (CD) was utilized to examine the secondary structure of the protein. In addition, Apaf-1 CARD contains 4 tyrosine residues. Therefore, fluorescence emission (with excitation at 280 nm, which excites the tyrosine residues) is utilized to monitor the tertiary structure of the protein. The results of these studies were published in the *Journal of Molecular Biology* (Milam et al., 2007).

D. Caspases in apoptosis

Caspases (cysteine-dependent aspartate-specific proteases) are highly specific proteases that recognize their substrates after specific tetrapeptide motifs (P4-P3-P2-P1), where cleavage occurs after P1, which is an aspartate residue (Riedl and Salvesen, 2007; Taylor et al., 2008). Caspases involved in the apoptotic pathway are divided into two groups based on their function in the apoptotic pathway. These groups are termed initiators and executioners (Figure 3) (Bao and Shi, 2006). Both groups of caspases have similar domain structures; an N-terminal prodomain, large subunit, intersubunit linker and a small subunit.

The two groups differ in their inactive oligomeric state of the enzyme, the length and function of the prodomain, their function in apoptosis and the maturation processes of these enzymes.

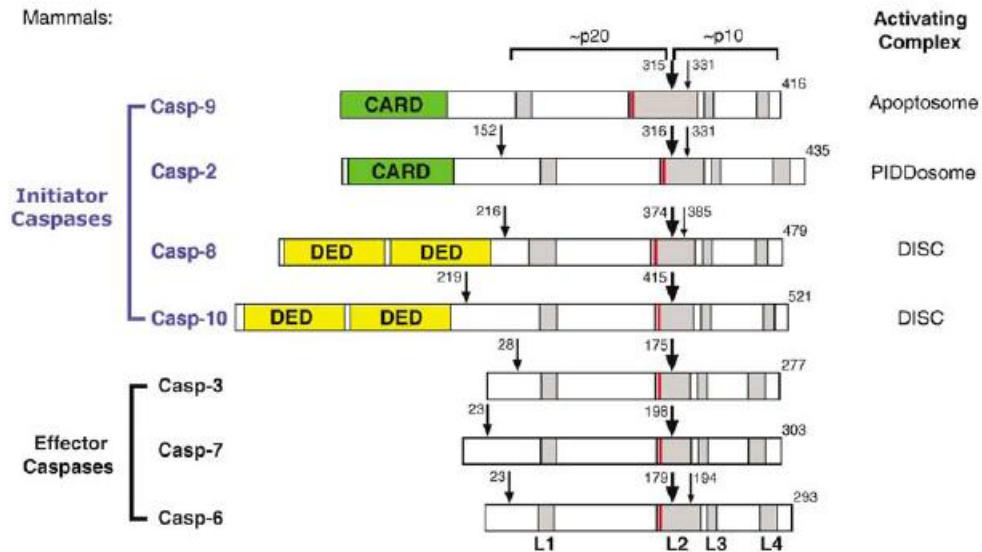


Figure 3. Caspase domains. Four initiator caspases and three effector (executioner) caspases are drawn to scale. The position of the first intrachain cleavage (between the large and small subunits) is highlighted by a large arrow whereas additional cleavages are represented by medium and small arrows. The prodomains in initiator caspases invariably contain homotypic interaction motifs, such as the CARD and the DED. The four surface loops, labeled L1 through L4, form the active site of a caspase. The catalytic cysteine residue is shown as a red line at the beginning of loop L2. The p20 and p10 subunits together form a caspase monomer. (Bao and Shi, 2006)

D.1. Inactive oligomeric state of the protein

Caspases exist in the cell as catalytically inactive zymogens, which allow the cell to

regulate their activity. In the inactive form of the enzyme, the active sites are not aligned properly for catalysis (Riedl and Salvesen, 2007). Initiator caspases, such as -2, -8, -9 and -10, exist in the cell in a monomeric state (Pop et al., 2007; Renatus et al., 2001). Executioner caspases, on the other hand, (such as -3, -6 and -7) exist in the cell in a dimeric state (Pop et al., 2001; Taylor et al., 2008). The maturation processes allow for full activation of the enzymes and are described below.

D.2. Prodomain: length and function

All caspases contain an N-terminal prodomain. For initiator caspases, the prodomain is long (152-219 amino acids) and contains death domains, such as CARD (caspase recruitment domain) or DED (death effector domain) (Bao and Shi, 2006; Taylor et al., 2008). These domains mediate protein-protein interactions that are necessary for dimerization and activation of initiator caspases. The prodomain of executioner caspases is shorter (23-28 amino acids) (Bao and Shi, 2006; Taylor et al., 2008). The function of these prodomains is unknown. However, the procaspase-3 prodomain has been shown to act as an intramolecular chaperone during assembly of the (pro)caspase subunits and increases the efficiency of formation of the native conformation (Feeney and Clark, 2005).

D.3. Function in apoptosis

Initiator caspases are activated through the extrinsic and intrinsic apoptotic pathways. The activation of initiator caspases depends on the formation of multiprotein complexes, such as the PIDDosome, the death-inducing signaling complex (DISC) and the apoptosome (Acehan et al., 2002; Peter and Krammer, 2003). These multiprotein complexes recruit the

initiator caspases through the prodomain, as shown in Figure 4. Briefly, caspase-2 is recruited to the PIDDosome by its CARD domain, though the mechanism by which caspase-2 is activated is unknown (Bao and Shi, 2006). As described above, caspase-8 and -10 and caspase-9 are recruited to the DISC through their DED and to the apoptosome through its CARD, respectively. Once recruited to the appropriate multiprotein complex, the dimerization/activation is induced via the induced proximity model (Bao and Shi, 2006). The initiator caspases then activate downstream executioner caspases by proteolytic cleavage in the intersubunit linker (Riedl and Salvesen, 2007; Taylor et al., 2008).

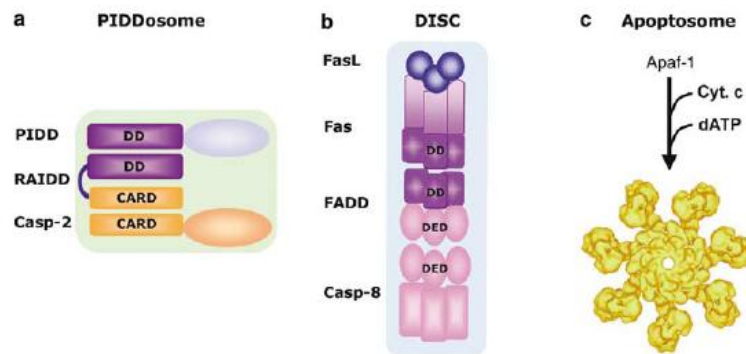


Figure 4. Adaptor protein complexes are responsible for the activation of initiator caspases. The activation of caspase-2, caspase-8, and caspase-9 in mammalian cells are mediated by the PIDDosome (panel A), the DISC (panel B), and the apoptosome (panel C), respectively. PIDDosome contains at least three components, PIDD, RAIDD, and caspase-2. DISC is assembled following binding of death ligand to its receptor and contains FADD and caspase-8 (or -10). Apoptosome is composed of seven molecules of Apaf-1 bound to cytochrome c in the presence of ATP/dATP. (Bao and Shi, 2006)

D.4. Maturation process

The maturation process for both groups of apoptotic caspases is shown in Figure 5. Initiator caspases, inactive monomers, require adaptor molecules in order to dimerize. Dimerization of initiator caspases results in activation of the enzyme (Bao and Shi, 2006). Cleavage of the protein in the intersubunit linker further stabilizes the protein, but is not required for activation (Renatus et al., 2001).

On the other hand, executioner caspases, inactive dimers, require cleavage in the intersubunit linker for full catalytic activity (Riedl and Salvesen, 2007; Taylor et al., 2008). The cleavage event allows for rearrangement of the active site loop, including the cleaved loop, L2', from one monomer helping to form the active site of the other monomer. The resulting active site is formed properly for catalysis in the mature form of the enzyme.

D.5. The importance of protein dimerization in cell death

Despite similar domain structure among initiator and executioner caspases, the functions and maturation processes for these proteins are different. Both apoptotic caspase groups exist as inactive zymogens. The collective differences in the oligomeric states of both groups the prodomain length and its function in activation stress the importance of maturation as a way to regulate apoptotic activity.

Once the initiator caspases are activated, the apoptotic signal is relayed to the executioner caspases. This activation signal needs to be controlled to prevent unintentional activation. Initiator caspases need to be sensitive to upstream activation signals as well as

deactivation (activated reversibly). Therefore, these proteins are activated by dimerization, which is induced by an upstream signal, and not by cleavage of the intersubunit linker.

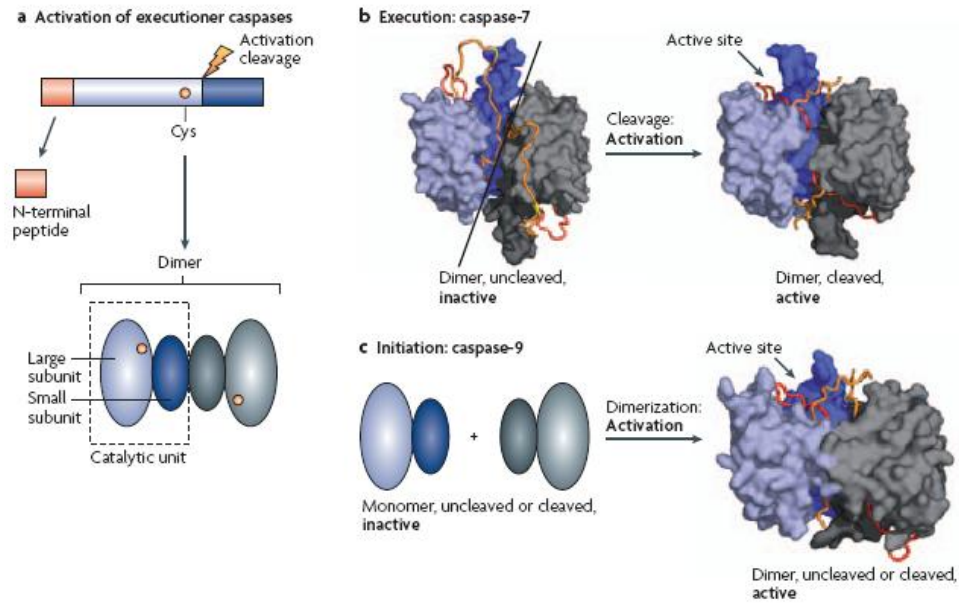


Figure 5. Caspases: architecture and activation. A) Activation of executioner caspases. An executioner caspase is typically cleaved twice, leading to the release of a short N-terminal peptide. The actual activation cleavage divides the catalytic unit into a large and small subunit. The position of the active-site Cys residue is indicated in orange. B) Surface rendering of an executioner caspase, caspase-7, before and after activation cleavage. The same color code as in panel A is used, and important loop regions are shown as ribbons. Cleavage releases strains on surface loops (red and orange) and the chains rearrange. The newly formed termini of the large and small subunits (orange) interact with each other across the other catalytic unit and with the red loops to align the substrate-binding pockets at the bottom of the active-site cleft. This results in a highly active caspase (right). C) Although they can be cleaved, initiator caspases, such as caspase-9, show full activity in their uncleaved forms. Their activity is regulated by dimerization instead of by cleavage. Initiator caspases exist as inactive monomers (left). Dimerization allows for the formation of a productive active site, shown here in the structure of cleaved, dimeric caspase-9 (right). (Riedl and Salvesen, 2007)

Executioner caspases, however, are the last step in the apoptotic pathway. These proteins must exist in the cell in a form that is easily activated. Cleavage of the intersubunit linker by initiator caspases, allows for irreversible activation of these proteins. The irreversibility of this step is important for the total execution of apoptosis. Therefore, these proteins are activated by cleavage and not dimerization.

E. Procaspase-3

As mentioned above, procaspase-3 exists in the cell as an inactive zymogen. This executioner caspase can be activated through the intrinsic and/or extrinsic apoptotic pathway. Chapter IV discusses on the kinetic pathway of procaspase-3. A homology model of procaspase-3 is shown in Figure 6. This homodimeric protein contains 10 tyrosines located throughout its sequence and 2 tryptophans located in the active site of each monomer. Therefore, we are able to utilize the intrinsic fluorescence of the protein to monitor the folding pathway of the protein. Monitoring fluorescence with excitation at 280 nm (excites the aromatic residues of the protein) provides information on the tertiary structure of the protein, while excitation at 295 nm (excites the tryptophan residues) provides information on the active site formation of the protein. We are also able to monitor the secondary structure by use of far-UV circular dichroism (CD).

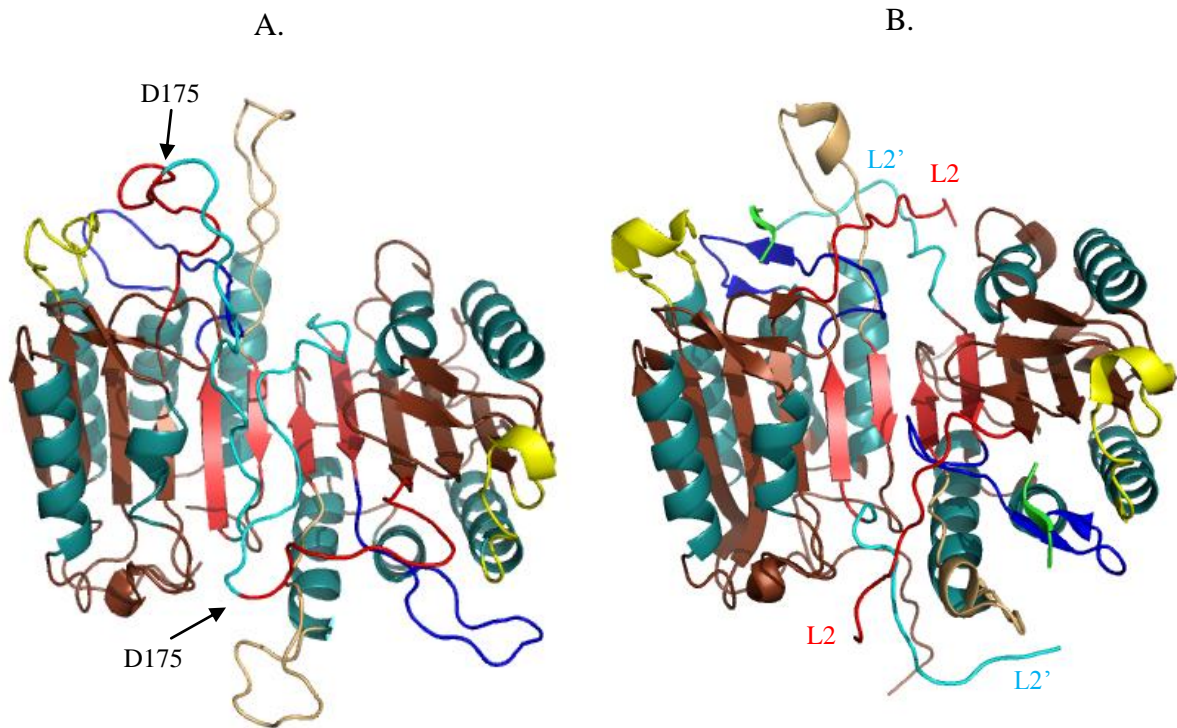


Figure 6: Homology model of procaspase-3 (panel A) and crystal structure of caspase-3 (panel B). Loops are colored; loop 1 is yellow, loop 2 is red and cyan, loop 3 is blue and loop 4 is brown. During maturation, D175 is cleaved in the intersubunit linker, loop 2 (panel A), creating loops L2 and L2' (panel B). This cleavage is followed by conformational changes in the protein, which allow for rearrangement of loops 1-4 and proper active site formation.

F. Project Aims

The following discussion focuses on elucidation of the kinetic folding pathway of two proteins involved in the apoptotic pathway; Apaf-1 CARD and procaspase-3. Chapter II discusses the sample preparation and data analysis procedures for kinetic folding experiments, while Chapters III and IV examine the kinetic folding mechanism of Apaf-1

CARD and procaspase-3, respectively. The following questions will be addressed in these chapters. Does the folding pathway of Apaf-1 CARD contain kinetic traps, as was detected for other CARDS (RICK and procaspase-1 CARD)? How quickly does the monomer of procaspase-3 form? Are there intermediates in the folding pathway of procaspase-3? If so, what are the roles of the intermediates? When does the procaspase-3 dimer form? The appendices contain additional information on the folding pathways of Apaf-1 CARD and procaspase-3 and protocols for buffers and KINSIM simulations.

REFERENCES

- Acehan, D., Jiang, X., Morgan, D. G., Heuser, J. E., Wang, X., and Akey, C. W. (2002). Three-dimensional structure of the apoptosome: implications for assembly, procaspase-9 binding, and activation. *Molecular Cell* **9**, 423-432.
- Bao, Q., and Shi, Y. (2006). Apoptosome: a platform for the activation of initiator caspases. *Cell* **14**, 56-65.
- Chamond, R. R., Anon, J. C., Aguilar, C. M., and Pasadas, F. G. (1999). Apoptosis and disease. *Allegol Immunol Clin* **14**, 367-374.
- Fairbrother, W. J., Gordon, N. C., Humke, E. W., O'Rourke, K. M., Starovasnik, M. A., Yin, J.-P., and Dixit, V. M. (2001). The PYRIN domain: a member of the death domain-fold superfamily. *Protein Science* **10**, 1911-1918.
- Feeney, B., and Clark, A. C. (2005). Reassembly of active caspase-3 is facilitated by the propeptide. *Journal of Biological Chemistry* **280**, 39772-39785.
- Hotchkiss, R. S., and Nicholson, D. W. (2006). Apoptosis and caspases regulate death and inflammation in sepsis. *Nature Reviews Immunology* **6**, 813-822.
- Luthi, A. U., and Martin, S. J. (2007). The CASBAH: a searchable database of caspase substrates. *Cell Death and Differentiation* **14**, 641-650.
- MacKenzie, S. H., and Clark, A. C. (2008). Targeting cell death in tumors by activating caspases. *Current Cancer Drug Targets* **8**, 98-109.
- Milam, S. L., Nicely, N. I., Feeney, B., Mattos, C., and Clark, A. C. (2007). Rapid folding and unfolding of Apaf-1 CARD. *Journal of Molecular Biology* **369**, 290-304.
- Peter, M. E., and Krammer, P. H. (2003). The CD95(APO-1/Fas) DISC and beyond. *Cell Death & Differentiation* **10**, 26.
- Pop, C., Chen, Y.-R., Smith, B., Bose, K., Bobay, B., Tripathy, A., Franzen, S., and Clark, A. C. (2001). Removal of the pro-domain does not affect the conformation of the procaspase-3 dimer. *Biochemistry* **40**, 14224-14235.
- Pop, C., Fitzgerald, P., Green, D. R., and Salvesen, G. S. (2007). Role of proteolysis in caspase-8 activation and stabilization. *Biochemistry* **36**, 4398-4407.

- Renatus, M., Stennicke, H. R., Scott, F. L., Liddington, R. C., and Salvesen, G. S. (2001). Dimer formation drives the activation of the cell death protease caspase 9. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 14250-14255.
- Richardson, J. S. (1977). [beta]-Sheet topology and the relatedness of proteins. **268**, 495-500.
- Riedl, S. J., and Salvesen, G. S. (2007). The apoptosome: signalling platform of cell death. **8**, 405-413.
- Saikumar, P., Dong, Z., Mikhailov, V., Denton, M., Weinberg, J. M., and Venkatachalam, M. A. (1999). Apoptosis: definition, mechanisms, and relevance to disease. *The American Journal of Medicine* **107**, 489-506.
- Shiozaki, E. N., Chai, J., and Shi, Y. (2002). Oligomerization and activation of caspase-9, induced by apaf-1 CARD. *Proceedings of the National Academy of Sciences* **99**, 4197-4202.
- Taylor, R. C., Cullen, S. P., and Martin, S. J. (2008). Apoptosis: controlled demolition at the cellular level. **9**, 231-241.
- Vaughn, D. E., Rodriguez, J., Lazebnik, Y., and Joshua-Tor, L. (1999). Crystal structure of apaf-1 caspase recruitment domain: an α -helical greek key fold for apoptotic signaling. *Journal of Molecular Biology* **293**, 439-447.
- Weber, C. H., and Vincenz, C. (2001). The death domain superfamily: a tale of two interfaces? *TRENDS in Biochemical Sciences* **26**, 475-481.
- Zhou, P., Chou, J., Olea, R. S., Yuan, J., and Wagner, G. (1999). Solution structure of apaf-1 CARD and its interaction with caspase-9 CARD: a structural basis for specific adaptor/caspase interaction. *Proceedings of the National Academy of Sciences* **96**, 11265-11270.
- Zou, H., Henzel, W. J., Liu, X., Lutschg, A., and Wang, X. (1997). Apaf-1, a human protein homologous to *C. elegans* CED-4, participates in cytochrome c-dependent activation of caspase-3. *Cell* **90**, 405-413.

CHAPTER II

Practical Approaches to Protein Folding and Assembly: Spectroscopic Strategies in Thermodynamics and Kinetics

Jad Walters, Sara L. Milam and A. Clay Clark

Department of Molecular and Structural Biochemistry, North Carolina State University,
Raleigh, North Carolina, 27695-7622

Published in Methods in Enzymology (modified, in press)

ABSTRACT

We describe here the use of several spectroscopies, such as fluorescence emission, circular dichroism and differential quenching by acrylamide, in examining the kinetic folding of proteins. This paper focuses on the experimental design and interpretation of kinetic data, such as burst phase analysis and exponential fits, used in elucidating kinetic folding pathways. In addition, simulation programs are used routinely to support folding models generated by kinetic experiments, and the fundamentals of simulations are covered.

INTRODUCTION

Protein folding is a central theme in structural biochemistry and in biotechnology. While the forces that stabilize protein structure have been examined for over one hundred years (Clark, 2008), protein chemists still are unable to predict the native structure of a protein from a known amino acid sequence. Under physiological conditions, proteins exist in equilibrium between ensembles of unfolded states (U) and native states (N), where each ensemble is characterized by a closely related set of structures that fluctuate around a local (or global) energy minimum. Protein function depends upon attaining the native conformation. While the forces that drive proteins to adopt their native conformations are, in general, defined as the difference between the unfavorable chain entropy and the favorable enthalpic interactions, the stability of protein native structures can vary drastically. Moreover, the kinetic pathway a protein utilizes to adopt its native conformation can vary from a relatively simple two-state process, where only the native and unfolded ensembles are populated significantly, to more complex reactions where the structure passes through one or more non-native, partially folded, intermediates before reaching the native conformation. It has been well documented that the pathways can be sequential, in which the intermediates are found between the unfolded and native ensembles, or parallel, where multiple intermediates form simultaneously, eventually coalescing to the native ensemble (Wallace, 2002). In addition, not all intermediates lead to the native conformation but rather can lead to misfolded, or off-pathway, structures (Ikai and Tanford, 1971). Finally,

proteins have been shown to fold over a wide range of time regimes, from microseconds to hours (Creighton, 1990). Consequently, characterizing the kinetic and thermodynamic folding of proteins can be a daunting task. However, the benefits to understanding the folding process can prove to be invaluable, for example, in revealing motifs or regions of the protein that are critical to function, as potential drug targets, or in determining the mechanisms for protein misfolding or aggregation (Cohen and Kelly, 2003; Soto, 2003)

Outlined in this chapter is practical information for characterizing the kinetic folding properties of a protein, exploiting intrinsic probes such as fluorescence emission and circular dichroism. One should note other excellent sources that describe the use of extrinsic probes in protein folding (Lakowicz, 2006; Weber, 1951; Waggoner, 1995). Fluorescence techniques are extremely useful for this application, and the advantages of fluorescence emission in comparison with other techniques make it an attractive method of choice to examine protein tertiary structure. These include high sensitivity, the use of low protein concentrations, the ability to selectively monitor regions or motifs within a protein, and the use of a multitude of solution conditions (Eftink, 2000). Circular dichroism also is employed to examine the protein secondary structure and/or tertiary structure during unfolding and refolding and to validate the findings from the fluorescence emission experiments. While proteins can be unfolded using a variety of agents, the focus here will be on a well identified chaotrope, urea. The kinetic techniques will aid in deciphering the folding pathway and in examining intermediates that may not be detectable in equilibrium experiments.

This chapter aims to provide a comprehensive protocol for examining the kinetic folding

properties of simple systems, so-called two-state, as well as more complex systems where multiple intermediates are present. More complex analyses for parsing the conformational free energy into component parts (entropy, enthalpy, heat capacity) as well as studies of the transition state can be found elsewhere (Dill, 1990; Privalov, 1989; Royer, 2008). Often, it is of interest not only to examine the conformational stability of wild-type proteins, but also of mutants or other proteins that differ slightly in structure. This type of analysis can be useful in comparing proteins within the same family or comparing structural motifs in general. Such studies have revealed critical residues and regions of proteins that make significant contributions to the overall stability (Wilson and Wittung-Stafshede, 2005). The protocols outlined in this chapter also are useful for comparing multiple proteins in a family.

MEASURING FOLDING KINETICS

In contrast to the equilibrium experiments, kinetic experiments examine refolding or unfolding processes by monitoring changes in a spectroscopic signal over time following the initiation of the reaction. A number of methods have been developed to examine protein folding kinetics, from continuous flow instruments, which measure reactions on the microsecond time scale (Shastry *et al.*, 1998), to stopped-flow instruments, which measure reactions on the millisecond to minute time scale. Stopped-flow instruments, the focus of this section, can employ absorbance, fluorescence emission, circular dichroism or other spectroscopies as a detection method. We describe here experimental protocols and data fitting procedures for kinetic folding experiments utilizing stopped-flow fluorescence emission and circular dichroism spectroscopies.

A. Experimental Protocol.

A.1. General Considerations.

Kinetic experiments with a stopped-flow instrument require rapid mixing of two solutions. For both refolding and unfolding studies, one routinely uses asymmetric mixing, 1:10 for example, either to dilute (refolding) or increase (unfolding) the amount of denaturant to span the unfolding/refolding transition region that was determined from equilibrium folding experiments. The 1:10 ratio is obtained typically by using one small drive syringe (250 μ L, for example) and one large drive syringe (2.5 mL, for example). One should consult the instrument manufacturer to determine the mixing dead-time, which is the shortest time at

which one can measure the kinetic signal due to the time required to mix the two solutions, typically 1-10 msec.

While the instrumental setup will vary depending on the manufacturer, there are several parameters common to all instruments, including signal detection, temperature, slit widths, wavelengths and time scale. Each parameter will be discussed briefly. As described above, it is recommended to use multiple spectroscopic probes to study protein folding because each detection method provides different information about the structures formed during refolding or unfolding. For fluorescence emission studies, there are two options available. The detection photomultiplier tube can be attached directly to the sample handling unit, or it can be attached to a monochromator. The advantage to the latter method is that it allows the user to select a particular detection wavelength. The disadvantage is that the overall signal will be decreased because other emission wavelengths are filtered out such that only a fraction of the total emission is detected. In contrast, a cut-off filter is used if the photomultiplier tube is attached to the sample handling unit. For this case, there are a variety of filters available, where the most common for intrinsic protein fluorescence emission use cut-off wavelengths of 305 nm or 320 nm.

It is important to maintain a constant temperature around the observation cell and drive syringes. The most common temperature for kinetic folding experiments reported in the literature is 25 °C (Maxwell *et al.*, 2005; Zarrine-Afsar and Davidson, 2004). This is due to the fact that 25 °C is slightly above room temperature, which allows easy control with heating, and it is an adequate endpoint for temperature-jump studies. If a reaction is too fast

to be detected by a stopped-flow instrument, it may be helpful to slow the chemistry by lowering the temperature. One should note that studies at low temperatures ($<10\text{ }^{\circ}\text{C}$) will require low temperature syringes to prevent sample leakage during injections.

Kinetic traces should be collected until the signal reaches that of the native or unfolded control in order to examine the full time course for the reaction. Depending on the protein, equilibration could occur in milliseconds or in hours. One should note that longer collection times can result in artifacts due to re-mixing of solutions of different densities in the observation cell and flow lines. One can test for mixing artifacts using various fluorescent compounds, such as tryptophan. Hand mixing experiments, which are performed on a steady-state instrument set for time-based data collection, also can verify slow reactions.

All experiments described below should be completed at multiple protein concentrations in (at least) duplicate experiments in order to verify the apparent kinetic rates and amplitude changes. One should use a 10-fold or greater variation in protein concentration, which will allow for the identification of protein concentration-dependent steps if the protein is an oligomer. If the protein is a monomer, then the apparent rate(s) should not depend on the protein concentration. If one observes this experimentally, then the data from the various protein concentrations can be averaged. If one observes a protein concentration-dependent rate during the folding or unfolding of a monomer, then the data indicate that the protein forms aggregates during the reaction.

A.2. Initial Parameterization.

Two samples are used to determine the initial parameters (final protein concentration and

slit widths) needed for a kinetic experiment: native protein in buffer and unfolded protein in urea-containing buffer. First, native protein stock is placed in the small syringe, mixed with buffer (from the large syringe), and data are collected for several seconds. Second, unfolded protein stock is placed in the small syringe, mixed with urea-containing buffer (from the large syringe), such that there is no dilution of the urea, and data are collected for several seconds. For each experiment, there should be no change in signal over time because the samples serve as controls for the native and unfolded signals. The unfolded protein stock should contain the same final urea concentration as the urea-containing buffer to which it is mixed. The general procedure for these initial experiments is to examine one protein concentration and slit width for various time scales, then to change the slit widths and repeat the experiment until the maximum difference is observed between the native and unfolded control samples. One may find that the optimized slit widths and protein concentrations are similar to those used in equilibrium unfolding experiments. In general, total volumes of 1 mL protein and 10 mL buffer solutions are sufficient for determining one set of parameters (slit width and voltage used for a particular final protein concentration, for example).

In the experiments described below, unfolded protein will be mixed with different urea-containing buffers in order to examine refolding at several final urea concentrations. The number of urea steps and volumes required will vary, depending on the protein and on the refolding conditions. The procedure outlined here is useful for examining burst-phase kinetics, apparent rates and amplitudes of observable phases, and for generating chevron plots of the apparent rate constants.

A.3. Sample Preparation for Measuring Refolding and Unfolding Kinetics.

First, one should determine the final urea and protein concentrations. Equation (1) describes the necessary calculation assuming a 1:10 mixing ratio.

$$\frac{([urea]_{proteinstock} \times 1) + ([urea]_{ureastock} \times 10)}{11} = \text{final urea concentration} \quad (1)$$

For example, if the protein stock solution is in 8 M urea-containing buffer (1 part = 250 μ L syringe) and is mixed with a 0 M urea-containing buffer (10 parts = 2.5 mL syringe), then the final urea concentration will be 0.73 M using Eq. (1). Similar calculations are used to determine the final protein concentration as well.

To measure refolding kinetics, one should prepare two stock protein samples (unfolded and native protein) and various urea-containing buffers that encompass the refolding transition. For example, in a typical experiment that measures the refolding of procaspase-3, we routinely use 17 urea concentrations between 0 M and 8 M at 0.5 M increments, and we average ~20 injections per urea step. This requires 10 mL urea-containing buffer for each final urea concentration and 15 mL protein stock. Table I shows the procedure for preparing urea-containing buffer solutions as well as the final urea concentration calculated using Eq. (1). If necessary, the urea concentrations should be adjusted in order to cover all regions of the equilibrium unfolding curve. Table II shows the calculations for making the unfolded and native protein stocks, assuming that the protein unfolds in 8 M urea-containing buffer and

that the final protein concentration in the observation cell is 5 μM .

A native protein stock is used as a control in order to determine the final signal in the refolding process (Table II). The native stock should contain the same buffer as the final refolding sample, 0.73 M urea-containing buffer in this example. Incubate the protein samples for the predetermined equilibration time, and then follow the instrument procedure below.

To measure unfolding kinetics, one will mix native protein with different urea-containing buffers. The sample preparation basically is the same as for the refolding experiment except for the initial protein stock. Prepare the urea-containing buffers as described above (Table I) and the unfolded protein and controls as shown in Table II. Note that the unfolded control should contain the same urea concentration as the final unfolding sample, 8 M in this example. Incubate the protein samples for the predetermined equilibration time, and then follow the instrument procedure below.

A.4. Instrument Procedure.

The general procedure for a kinetic refolding experiment on a stopped-flow instrument is outlined here. The experimental protocol for unfolding follows the same steps except that one starts with native rather than unfolded protein. The basic idea is to refold/unfold the protein at several final urea concentrations and to monitor the signal until no further change occurs. Final signals for the lowest and highest urea concentrations should match those of the controls (native and unfolded protein).

- 1) Flush the system with distilled, deionized water and then with buffer. Make sure

that the drive syringes contain no air bubbles as this will cause mixing artifacts in the data. For 1:10 asymmetric mixing, ~10 injections are required to move samples and/or buffers completely through the system, depending on the tubing length and stop syringe volume.

- 2) Place the unfolded protein stock in the small drive syringe and the 8 M urea-containing buffer in the large drive syringe.
- 3) Fill the drive syringes and flush the system.
- 4) Set the instrument parameters determined above (wavelength, temperature, slit width, signal detection of choice, voltage).
- 5) Obtain data by acquiring multiple injections at one set of conditions. The goal is to collect enough repetitions to average, the final number of which will depend on the signal-to-noise ratio. The averages also depend on the final protein concentration and detection method. Collect data for various time frames, such as 1, 10, 100 and 500 seconds, as needed.
- 6) Once the unfolded control sample is completed, remove the 8 M urea-containing buffer. For a refolding experiment, fill the large drive syringe with the next urea-containing buffer, in this case 7.5 M urea. One should note that whenever the solution conditions are drastically changed, one should rinse the syringe either with distilled, deionized water or with buffer to remove any remaining solution in the syringe.
- 7) Flush the system and acquire data as in step 5.

- 8) Continue to change the urea-containing buffer in the large drive syringe and acquire data by repeating steps 5-7 until the final buffer (0 M urea) is reached.
- 9) Remove unfolded protein and buffer solutions from the syringes.
- 10) For the native control, add the native protein (Table II) to the small syringe and 0.73 M urea-containing buffer to the large syringe.
- 11) Flush the system and acquire data as in step 5.
- 12) Once the experiment is complete, clean the instrument by removing the protein and urea-containing buffer solutions from the reservoir syringes. Flush the system with buffer and distilled, deionized water as before.

B. Differential Quenching by Acrylamide.

Differential quenching describes the relative solvent accessibility of aromatic residues during refolding/unfolding compared to the same residues in the native or unfolded states. The technique of differential quenching may allow one to identify and characterize folding intermediates not detected by other methods such as stopped-flow fluorescence emission or CD spectroscopies. For example, Vanhove *et al.* (1998) used this technique to investigate a non-native intermediate formed during the refolding of TEM-1 β -lactamase. Four tryptophan residues are located on the protein surface in the native state, and it is proposed that the tryptophans are less accessible to solvent during refolding than in the native state. This group verified, through quenching by acrylamide and other techniques, the presence of a hydrophobic collapse in the non-native intermediate.

One should first identify the quencher of choice. It is desirable to use a quencher that

reacts differentially to the native and unfolded states of the protein. Equilibrium quenching measurements should be conducted first to provide the optimal concentration of quencher for use in kinetic studies. In the case of procaspase-3, for example, it was shown that 0.7 M acrylamide provided the greatest difference between quenching of native and unfolded protein fluorescence emission (Bose *et al.*, 2003). In kinetic experiments, however, it is important to verify that the quencher does not affect the kinetic folding process. One should test a variety of quencher concentrations and, if necessary, different quenchers. Also, if an ionic quencher (such as iodide) is used, it is important to verify that the increased ionic strength has no effect on the folding process. One can do this simply by increasing the ionic strength in the folding experiments described above (in the absence of quencher).

B.1. Sample Preparation.

Sample preparation for differential quenching by acrylamide is performed as described above for the refolding and unfolding experiments except that acrylamide is included in the urea-containing buffers (Table I). Protein stock solutions should be prepared as described in Table II. Equation (1) is used to determine the initial concentration of acrylamide needed in the urea-containing buffers. Using procaspase-3 as a model, a final acrylamide concentration of 0.7 M requires the initial concentration to be 0.77 M, so the 0.77 M acrylamide urea-containing buffers are made as shown in Table I. The acrylamide stock solution (8 M) is made by weighing 11.37 g of acrylamide (wear a mask) and bringing the volume up to 20 mL with buffer.

B.2. Experimental Procedure.

The experimental procedure for stopped-flow differential quenching is the same as those for the refolding and unfolding experiments described above for stopped-flow fluorescence emission. Briefly, unfolded or native protein (Table II) is placed in the small syringe and is mixed 1:10 with acrylamide- and urea-containing buffers (Table I). Data also are collected for controls of native and unfolded protein (Table II). In the examples shown here, the fluorescence emission of the native protein control is examined in 0.7 M acrylamide, 0.73 M urea-containing buffer, and that of the unfolded protein control is examined in 0.7 M acrylamide, 8 M urea-containing buffer. The excitation wavelength should be set to 295 nm in order to avoid inner filter effects from acrylamide (Lakowicz, 2006).

The kinetic data in the presence of quencher and at various final urea concentrations are plotted with the native and unfolded controls, which also are in the presence of quencher (Fig. 1A). In this way, one observes the differential quenching of tryptophan fluorescence emission during refolding/unfolding as compared to the quenching of the native and unfolded proteins.

C. Data Analysis.

The first step in data analysis is to determine the number of kinetic phases in the experimental data. To achieve this, there are multiple data analysis programs available including KaleidaGraph (Synergy Software), Excel (Microsoft Corporation) and Origin (Origin Lab).

A phase is defined here as any change in signal over time. A phase can be described by a

single rate constant, or, in the case of a burst phase, a change in amplitude. In order to determine the number of kinetic phases, plot the signal versus time in seconds. A logarithmic scale can be used for short times (1 to 100 milliseconds) or for combining multiple time scales. An example of a refolding experiment is shown in Figure 1B, which shows the signal acquired versus time at six different final urea concentrations. One should note that the data will vary for different proteins. In this hypothetical refolding experiment, three phases are detected and are referred to here as the burst, fast and slow phases in order to illustrate the analysis below. The burst phase occurs within the mixing dead-time for the instrument and is observed at all final urea concentrations shown here. The fast phase, from ~2 milliseconds to ~100 milliseconds, occurs at 3.45, 4.36, 5.27 and 6.18 M urea. The slow phase, from ~200 milliseconds to ~10 seconds, occurs at 3.45 and 4.36 M urea.

C.1. Burst Phase.

The burst phase is defined as the change in signal that occurs during the mixing dead-time, and while illustrated here, not all kinetic experiments will display a burst phase. If the data display a change in the initial signal with different final urea concentrations, as shown in Figure 1B, then one should plot the burst phase signal versus final urea concentration (Fig. 1C). In order to determine the signal, extrapolate the data from ~1 to 5 milliseconds (the boxed area in Fig. 1B) to a time of zero seconds (arrows in Fig. 1B). These times are chosen because the signal typically is too noisy at shorter times.

A plot of the burst phase signals versus urea (Fig. 1C) may show linear or non-linear transitions. A linear transition indicates a non-cooperative process and simply may represent

a change in signal for the unfolded protein with a change in final urea concentration. In contrast, a non-linear transition (Fig. 1C) indicates a cooperative folding process. In this case, one can determine the free energy and m-value for the formation of the burst phase species by fitting the data to a relevant equilibrium folding model. Typically, one uses a two-state equilibrium model, although more complicated models also can be used. An examination of data determined by fluorescence emission, following excitation at 280 nm and 295 nm, CD, and/or differential quenching may show whether multiple species form in the burst phase (Zaidi *et al.*, 1997; Georgescu *et al.*, 1998). By comparing the free energy and m-values to those obtained from equilibrium unfolding experiments, one will obtain further information regarding the species formed during the burst phase.

C.2. Exponential Fits.

Most observable phases comprise exponential increases or decreases in signal. Therefore, this section will focus on the use of multiple exponential equations to fit the experimental data. Two valuable pieces of information will be obtained from these fits: the amplitude and the apparent rate of each phase. The amplitude provides information on the change in signal with each phase and is related to the population of species that occurs in a phase (Wallace and Matthews, 2002). The apparent rate constant provides the rate at which the transition occurs. One should note that since only the rate limiting step is detected, a single phase may consist of multiple transitions with multiple rate constants. In addition, the amplitude of a particular phase may differ based on the detection method used, while the rate will be consistent between all methods (so long as the phase is detected by each method). As stated

above, kinetic measurements are collected until the equilibrium signal is reached. So, in general, the total change in amplitudes of the kinetic phases should equal the difference in signals between native and unfolded control proteins.

The data are fit as a sum of exponentials, as shown in Eq. (2).

$$A(t) - A(\infty) = \sum_{i=1}^n A_i e^{-k_i t} \quad (2)$$

In this case, $A(t)$ is the amplitude at time t , $A(\infty)$ is the offset value, A_i is the change in signal for phase i , k is the apparent rate, and t is time (Bieri and Kiefhaber, 2000; Utiyama and Baldwin, 1986). To fit exponential phases, it is best to use a nonlinear least squares fitting program, such as those mentioned above. Two and three (or more) exponential equations have the same format as a single exponential but with additional terms to account for the total number of phases. That is, data that contain more than one phase are fit to a sum of exponentials.

As an example, for the refolding experiment shown in Figure 1B, each urea step is plotted and analyzed individually. Figure 2 shows a plot of data collected at a final urea concentration of 4.36 M fit to either a one (Fig. 2A), two (Fig. 2B) or three (Fig. 2C) exponential equation. An analysis of the residuals to the fits (below each plot) shows the difference between the fit and the experimental data, where the distribution of the points should be random around zero. A single exponential (Fig. 2A) does not adequately describe the data, and the residuals for the three exponential fit (Fig. 2C) are not significantly different from those of the two exponential fit (Fig. 2B). Thus, using the simplest model, one would

describe the experimental data as containing two phases, each defined by an apparent rate constant and amplitude.

Continue the fitting process until all of the urea steps are analyzed, and then plot the rates and amplitudes from the different phases and detection methods versus denaturant concentration. Using a monomeric protein as a simple example (all observed phases are first order reactions), we generated a plot of the log of apparent rate constant versus denaturant concentration. The data display a V-shape, and this plot is referred to as a chevron plot (Fig. 3A). The chevron analysis has been described extensively for a simple 2-state folding model (Wallace and Matthews, 2002; Ferguson *et al.*, 1999; Parker *et al.*, 1995; Zarrine-Asfar and Davidson, 2004), and the reader is referred to previous reviews on the subject. The information obtained from the chevron analysis includes the folding rates (both refolding and unfolding) in the absence of denaturant, the m-values for refolding and unfolding, and information on the transition state for folding.

For an oligomeric protein, the phase in which the protein forms the oligomer will display a protein concentration dependence to the apparent rate constant. For a dimeric protein, for example, the apparent rate will be second order with regard to the protein concentration. To determine the rate of dimerization, plot the apparent rate (sec^{-1}) versus the protein concentration (μM). The second order rate plot should show a linear dependence on protein concentration, with a slope equal to the rate of dimerization (Jaenicke and Rudolph, 1986).

C.3. Simulations.

Creating a kinetic folding model is outside of the scope of this article but is covered in

other reviews (Wallace and Matthews, 2002; Utiyama and Baldwin, 1986; Creighton, 1988; Bieri and Kiefhaber, 2000). Once a model is created, it can be tested using other stopped-flow techniques, such as sequential mixing (Wallace and Matthews, 2002; Schmid, 1986; Eftink and Shastry, 1997). The protein folding mechanism could be very complex, depending on the number of kinetic phases and the oligomeric properties of the protein, and analysis of the data may yield several possible folding mechanisms. The overall goal is to devise the simplest mechanism that adequately explains the experimental data. In this regard, simulation programs are useful for distinguishing between possible mechanisms. It is important to note that simulations cannot prove a mechanism, but rather they allow one to determine whether a mechanism is supported by the experimental data. This section will describe the use of the simulation program KINSIM (Barshop *et al.*, 1982) but other programs also are available, such as KinTekSim (KinTek Corp.), KINFITSIM (Svir *et al.*, 2002) and DynaFit (BioKin, Ltd.).

There are various examples in the literature of the use of KINSIM in the field of protein folding, ranging from determining if an intermediate is on- or off-pathway (Heidary *et al.*, 2000) to examining the complex parallel folding pathway of a homodimeric protein (Mallam and Jackson, 2006). KINSIM is a free program available at the following website: <http://www.biochem.wustl.edu/cflab/message.html>. The website also has a help manual with instructions for the program, but a brief overview will be described here.

KINSIM allows the user to define a mechanism in a text file. A simple example is shown in Figure 3B for the folding of a monomer with one on-pathway intermediate ($U \leftrightarrow I \leftrightarrow N$). In

this sequential mechanism, the unfolded protein (U) folds to an intermediate species (I) before forming the native structure (N). As shown in Figure 3B, a reversible reaction is described by two equal signs (= =). The concentration of each species is multiplied by an output factor (X1-X3 in Fig. 3B), and the signal observed is the sum of the contribution of each species.

Once the mechanism is loaded into KINSIM, the user can change protein concentration (μM), rate constants (sec^{-1}), output factors (signal/[protein]) and simulation time (sec). The protein concentration and simulation time are chosen based on the experimental conditions. The rate constants determined from fits of the experimental data (described above) are used as initial guesses for the simulated rate constants. For example, if the data analysis reveals a rate of 50 sec^{-1} for the transition of U to I, then one would use 50 for k_1 . Output factors are calculated by dividing the signal of the species, taken from the experimental data, by the protein concentration. Finally, one can output any combination of species. As an example, using the sequential mechanism in Figure 3, the simulation screen will show four lines based on the four specified outputs, U, I, N and U+I+N (Fig. 3C). Experimental data also can be uploaded in order to examine the agreement between simulated and real data, which allows the user to fine tune rate constants and output factors and to distinguish between possible mechanisms.

As an example, KINSIM simulations were used to investigate several potential kinetic folding pathways of the caspase recruitment domain (CARD) of apoptotic protease activating factor 1 (Apaf-1) (Milam *et al.*, 2007). Apaf-1 CARD is a small (<100 amino acids),

monomeric protein consisting of six α -helices arranged in an α -Greek key folding topology (Weber and Vincenz, 2001). One sequential (Fig. 4A) and two parallel (Fig. 4B-C) kinetic mechanisms were proposed based on three phases present in the experimental data (burst, fast and slow). In order to demonstrate the ability of KINSIM to discriminate between possible mechanisms, this brief discussion will focus on simulations of the slow phase in the Apaf-1 CARD kinetic folding pathway. In the sequential folding model (Fig. 4A), the unfolded protein folds through two intermediates before forming the native conformation. KINSIM simulations with this mechanism agree with the experimental data (Fig. 4D). However, more complicated mechanisms were proposed based on stopped-flow sequential mixing studies that showed the presence of multiple unfolded species (Milam *et al.*, 2007). The proposed parallel mechanisms, shown in Figure 4B and C, consist of multiple unfolded and/or native conformations. In the model shown in Figure 4B, two unfolded conformations, U_1 and U_2 , form two native conformations, N_1 and N_2 , and the unfolded and native species are able to interconvert. Simulations with this mechanism were not able to recapitulate the slow phase of refolding, assuming that the slow phase was due to the interconversion of U_1 to U_2 or of N_1 to N_2 (Fig. 4E). As a result, the folding model in Figure 4C was proposed, with the inclusion of a third unfolded species, U_3 . In this parallel pathway, both unfolded conformations, U_1 and U_2 , can form the native species (N). The conversion of U_3 to U_1 or to U_2 represents the slow phase, as shown in Figure 4F. Therefore the models in Figures 4A and 4C adequately explain the single mixing stopped-flow data (Fig. 4D and F, respectively), but only the model in Figure 4C agrees with the sequential mixing stopped-flow data, which

showed multiple unfolded species.

REFERENCES

- Barshop, B. A., Wrenn, R. F., and Frieden, C. (1983). Analysis of numerical methods for computer simulation of kinetic processes: Development of KINSIM - a flexible, portable system. *Anal. Biochem.* **130**, 134-145.
- Bieri, O., and Kiefhaber, T. (2000). Kinetic models in protein folding. In "Mechanisms of Protein Folding" (R. H. Pain, Ed.), pp. 34-64. Oxford University Press.
- Bose, K., Pop, C., Feeney, B., and Clark, A. C. (2003). An uncleavable procaspase-3 mutant has a lower catalytic efficiency but an active site similar to that of mature caspase-3. *Biochem.* **42**, 12298-12310.
- Clark, A. C. (2008). Protein folding: Are we there yet? *Arch. Biochem. Biophys.* **469**, 1-3.
- Cohen, F. E., and Kelley, J. W. (2003). Therapeutic approaches to protein-misfolding diseases. *Nature* **426**, 905-909.
- Creighton, T. E. (1988). Toward a better understanding of protein folding pathways. *Proc. Natl. Acad. Sci.* **85**, 5082-5086.
- Creighton, T. E. (1990). Protein folding. *Biochem. J.* **270**, 1-16.
- Dill, K. A. (1990). Dominant forces in protein folding. *Biochem.* **29**, 7133-7155.
- Eftink, M. R. (2000). Use of fluorescence spectroscopy as thermodynamics tool. *Meth. Enzymol.* **323**, 459-473.
- Eftink, M. R., and Shastry, M. C. R. (1997). Fluorescence methods for studying kinetics of protein-folding reactions. *Meth. Enzymol.* **278**, 258-286.
- Ferguson, N., Capaldi, A. P., James, R., Kleanthous C., and Radford, S. E. (1999). Rapid folding with and without populated intermediates in the homologous four-helix proteins Im7 and Im9. *J. Mol. Biol.* **286**, 1597-1608.

- Georgescu, R. E., Li, J. H., Goldberg, M. E., Tasayco, M. L., and Chaffotte, A. F. (1998). Proline isomerization-independent accumulation of an early intermediate and heterogeneity of the folding pathways of a mixed alpha/beta protein, *Escherichia coli* thioredoxin. *Biochem.* **37**, 10286-10297.
- Heidary, D. K., O'Neill, J. C., Roy, M., and Jennings, P. A. (2000). An essential intermediate in the folding of dihydrofolate reductase. *Proc. Natl. Acad. Sci.* **97**, 5866-5870.
- Ikai, A., and Tanford, C. (1971). Kinetic evidence for incorrectly folded intermediate states in the refolding of denatured proteins. *Nature* **230**, 100-102.
- Jaenicke, R., and Rudolph, R. (1986). Refolding and association of oligomeric proteins. *Meth. Enzymol.* **131**, 218-250.
- Lakowicz, J. R. (2006). "Principles of fluorescence spectroscopy." Springer, New York.
- Mallam, A. L., and Jackson, S. E. (2006). Probing nature's knots: The folding pathway of a knotted homodimeric protein. *J. Mol. Biol.* **359**, 1420-1436.
- Maxwell, K. L., Wildes, D., Zarrine-Afsar, A., De Los Rios, M. A., Brown, A. G., Friel, C. T., Hedberg, L., Horng, J. C., Bona, D., Miller, E. J., Vallee-Belisle, A., Main, E. R. G., Bemporad, F., Qiu, L., Teilum, K., Vu, N. D., Edwards, A. M., Ruczinski, I., Poulsen, F. M., Kragelund, B. B., Michnick, S. W., Chiti, F., Bai, Y., Hagen, S. J., Serrano, L., Oliveberg, M., Raleigh, D. P., Wittung-Stafshede, P., Radford, S. E., Jackson, S. E., Sosnick, T. R., Marqusee, S., Davidson, A. R., and Plaxco, K. W. (2005). Protein folding: Defining a "standard" set of experimental conditions and a preliminary kinetic data set of two-state proteins. *Protein Sci.* **14**, 602-616.
- Milam, S. L., Nicely, N. I., Feeney, B., Mattos, C., and Clark, A. C. (2007). Rapid folding and unfolding of Apaf-1 CARD. *J. Mol. Biol.* **369**, 290-304.
- Parker, M. J., Spencer, J., and Clarke, A. R. (1995). An integrated kinetic analysis of intermediates and transition states in protein folding reactions. *J. Mol. Biol.* **253**, 771-786.
- Privalov, P. L. (1989). Thermodynamic problems of protein structure. *Ann. Rev. Biophys. Biophys. Chem.* **18**, 47-69.

- Royer, C. A. (2008). The nature of the transition state ensemble and the mechanisms of protein folding. *Arch. Biochem. Biophys.* **469**, 34-45.
- Schmid, F. X. (1986). Fast-folding and slow-folding forms of unfolded proteins. *Meth. Enzymol.* **131**, 70-82.
- Shastry, M. C. R., Luck, S. D., and Roder, H. (1998). A continuous-flow capillary mixing method to monitor reactions on the microsecond time scale. *Biophys. J.* **74**, 2714-2721.
- Soto, C. (2003). Unfolding the role of protein misfolding in neurodegenerative diseases. *Nat. Rev. Neurosci.* **4**, 49-60.
- Svir, I. B., Klymenko, O. V., and Platz, M. S. (2002). 'KINFITSIM' - a software to fit kinetic data to a user selected mechanism. *Computers and Chemistry* **26**, 379-386.
- Utiyama, H., and Baldwin, R. L. (1986). Kinetic mechanisms of protein folding. *Meth. Enzymol.* **131**, 51-70.
- Vanhove, M., Lejeune, A., Guillaume, G., Virden, R., Pain, R. H., Schmid, F. X., and Frere, J. M. (1998). A collapsed intermediate with nonnative packing of hydrophobic residues in the folding of TEM-1 beta-lactamase. *Biochem.* **37**, 1941-1950.
- Waggoner, A. (1995). Covalent labeling of proteins and nucleic acids with fluorophores. *Meth. Enzymol.* **246**, 362-373.
- Wallace, L. A., and Matthews, C. R. (2002). Sequential vs. parallel protein-folding mechanisms: Experimental tests for complex folding reactions. *Biophys. Chem.* **101-102**, 113-131.
- Weber, C. H., and Vincenz, C. (2001). The death domain superfamily: A tale of two interfaces? *TRENDS Biochem. Sci.* **26**, 475-481.
- Weber, G. (1951). Polarization of the fluorescence of macromolecules. *Biochem. J.* **51**, 145-155.
- Wilson, C. J., and Wittung-Stafshede, P. (2005). Role of structural determinants in folding of the sandwich-like protein pseudomonas aeruginosa azurin. *Proc. Natl. Acad. Sci.* **102**, 3984-3987.

Zaidi, F. N., Nath, U., and Udgaonkar, J. B. (1997). Multiple intermediates and transition states during protein folding. *Nat. Struct. Biol.* **4**, 1016-1024.

Zarrine-Afsar, A., and Davidson, A. R. (2004). The analysis of protein folding kinetic data produced in protein engineering experiments. *Methods* **34**, 41-50.

TABLES AND FIGURES

Table I. Example Calculations for Urea and 0.77 M Acrylamide/Urea Stock Solutions

	10 M Urea Stock (mL)	Buffer (mL)	8 M Acrylamide Stock (mL)	Final Urea Concentration (M)
Urea Stocks (M)				
0	0	10		0.73
1	1	9		1.64
2	2	8		2.55
3	3	7		3.45
4	4	6		4.36
5	5	5		5.27
6	6	4		6.18
7	7	3		7.09
8	8	2		8
0.77 M Acrylamide/Urea Stocks (M)				
0	0	9.037	0.963	0.73
1	1	8.037	0.963	1.64
2	2	7.037	0.963	2.55
3	3	6.037	0.963	3.45
4	4	5.037	0.963	4.36
5	5	4.037	0.963	5.27
6	6	3.037	0.963	6.18
7	7	2.037	0.963	7.09
8	8	1.037	0.963	8

Table II. Example Calculations for 55 μ M Procaspase-3 Protein Stock Solutions for Refolding/Unfolding Kinetic Experiments

Protein Stock Solutions	10 M urea (mL)	1 M DTT (μ L)	825 μ M protein ^a (μ L)	Buffer (mL)
Unfolded Control (1 mL)	0.8	1	66.7	0.1323
Native (15 mL)	0	15	1000	13.985
Native Control (1 mL)	0.073	1	66.7	0.8593
Unfolded (15 mL)	12	15	1000	1.985

^a Initial protein concentration of 825 μ M is diluted to 55 μ M for kinetic folding experiments.

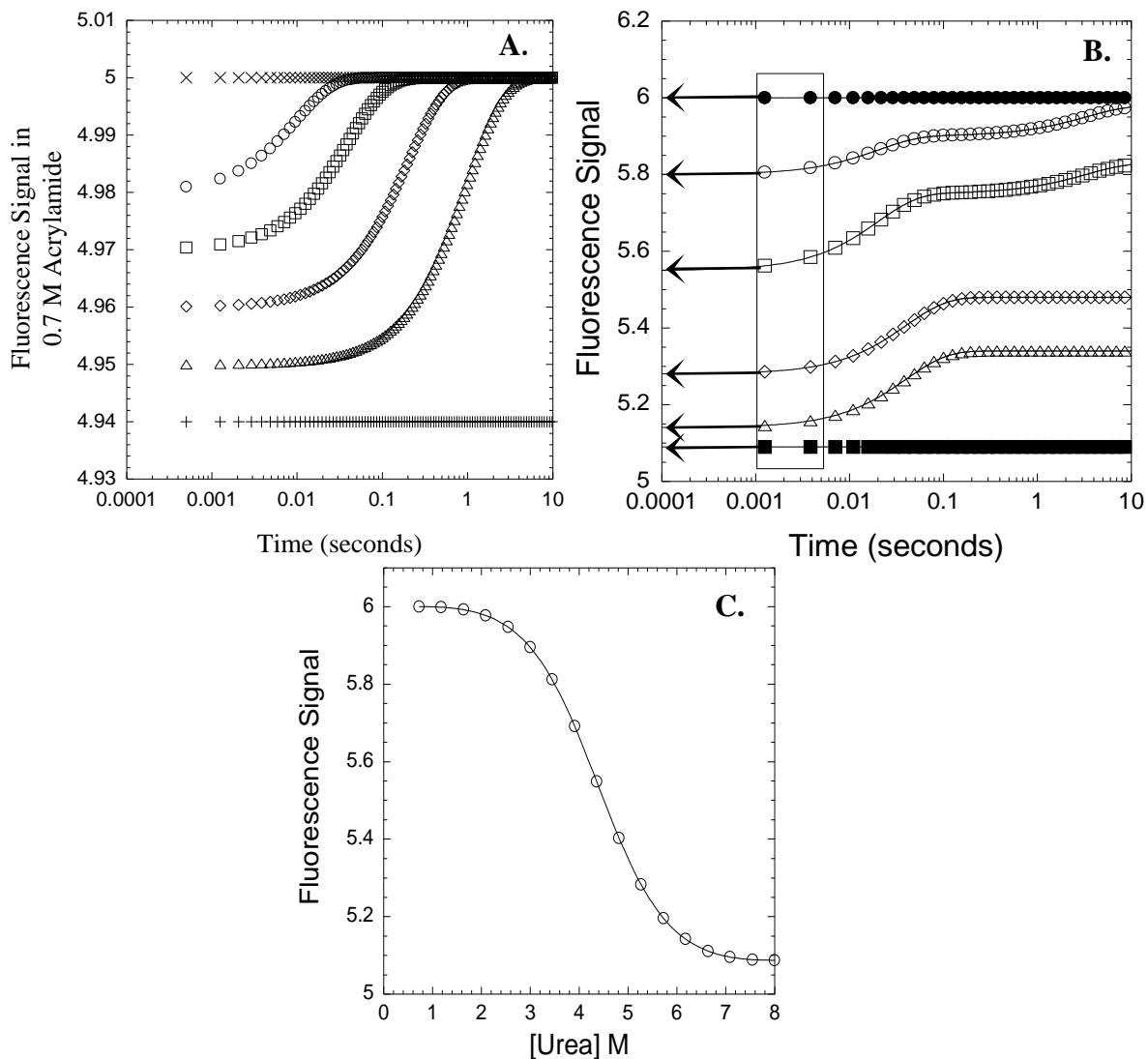


Figure 1. Data analysis. **A.** Hypothetical plot of fluorescence signal in the presence of 0.7 M acrylamide versus time for a refolding reaction. The unfolded (+) and native (x) protein signals in the presence of acrylamide are shown. Data at several final urea concentrations are shown (□,◇,□,○). **B.** Hypothetical plot of the fluorescence emission signal versus time for a refolding reaction. The unfolded (■) and native (□) protein signals are shown. Final urea concentrations of 3.45 M (○), 4.36 M (□), 5.27 M (◇), and 6.18 M (△) are shown. The boxed area marks the time from one to five milliseconds. The arrows show the extrapolation of the signal to time zero. **C.** Plot of burst phase signal versus final urea concentration. The continuous line represents a fit to a two-state equilibrium folding model as described in the text.

Figure 2. The 4.36 M refolding data from Fig. 1B were fit to a one (A), two (B) or three (C) exponential equation. Residuals to the fits are shown below each panel. **A.** A single exponential fit of the data provides an amplitude of 0.22 and rate of 33 sec^{-1} . **B.** A two exponential fit of the data provides amplitudes of 0.20 and 0.08 and rates of 50 and 0.3 sec^{-1} , respectively. **C.** A three exponential fit of the data provides amplitudes of 0.2, 0.08, and 0.05 and rates of 50, 0.3 and $3.3 \times 10^{-7} \text{ sec}^{-1}$, respectively.

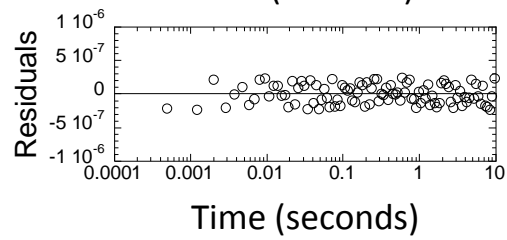
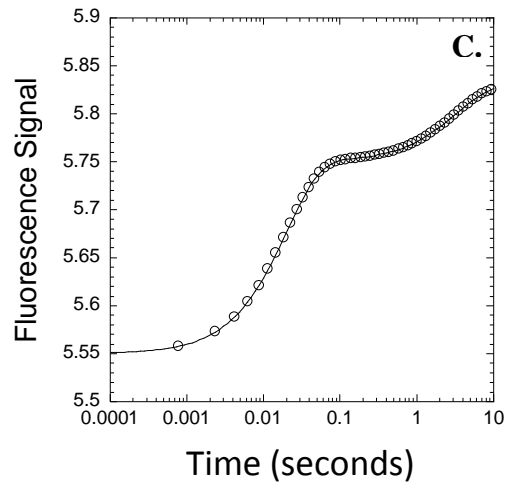
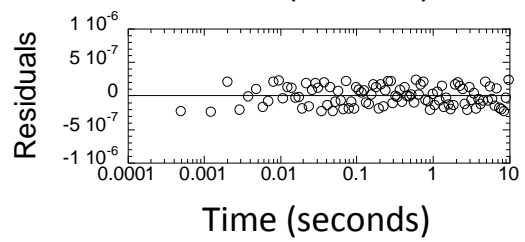
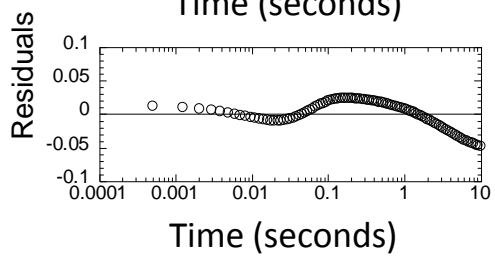
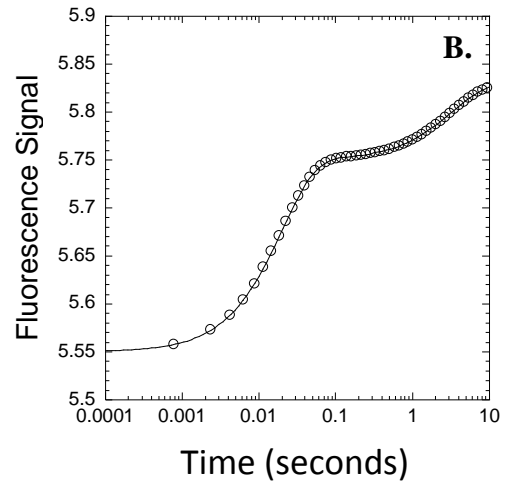
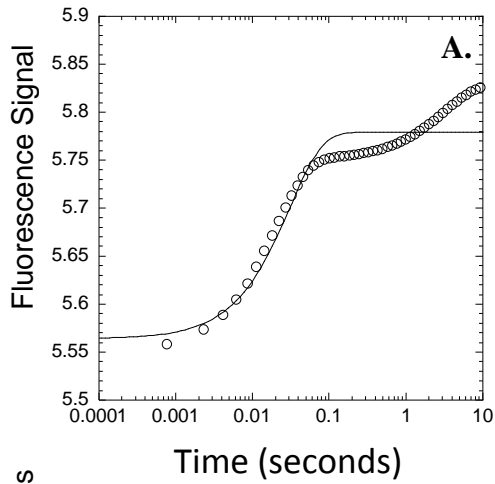
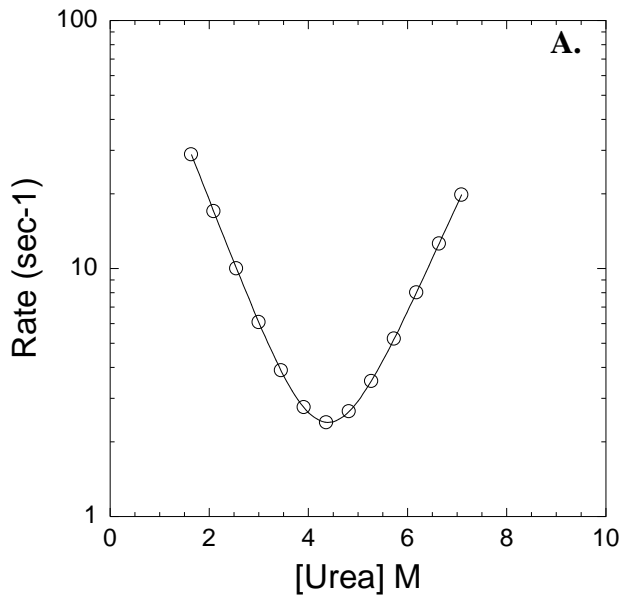


Figure 3. Chevron and KINSIM analysis. **A.** Hypothetical chevron plot of the apparent refolding and unfolding rates versus the final urea concentration. The continuous line represents a fit to a two-state kinetic folding model (Ferguson *et al.*, 1999) with the following parameters: $k_f^{H_2O} = 200 \text{ sec}^{-1}$, $k_U^{H_2O} = 0.015 \text{ sec}^{-1}$, $m_{N-TS} = 0.7$, $m_{U-TS} = 0.6$, $\Delta G^{H_2O} = 5.62$ kcal/mol, $m\text{-value} = 1.3 \text{ kcal/mol/M}$, $\text{urea}_{1/2} = 4.3 \text{ M}$. **B.** Example of a sequential pathway with one on-pathway intermediate written in the text format for KINSIM. X1, X2 and X3 are the extinction coefficients of U, I and N, respectively. **C.** Hypothetical example of a refolding reaction of 10 μM protein for 10 seconds. The populations of species, shown by the continuous lines, are labeled as U, I, N and U+I+N. The rates of the U to I and I to N transitions are 1 and 0.3 sec^{-1} , respectively. The values of X1, X2 and X3 are 0.01, 0.05 and 0.06, respectively.



B.

\$ sequential pathway
 !
 U == I
 I == N
 !
 *OUTPUT
 X1*U
 X2*I
 X3*N
 (X1*U)+(X2*I)+(X3*N)

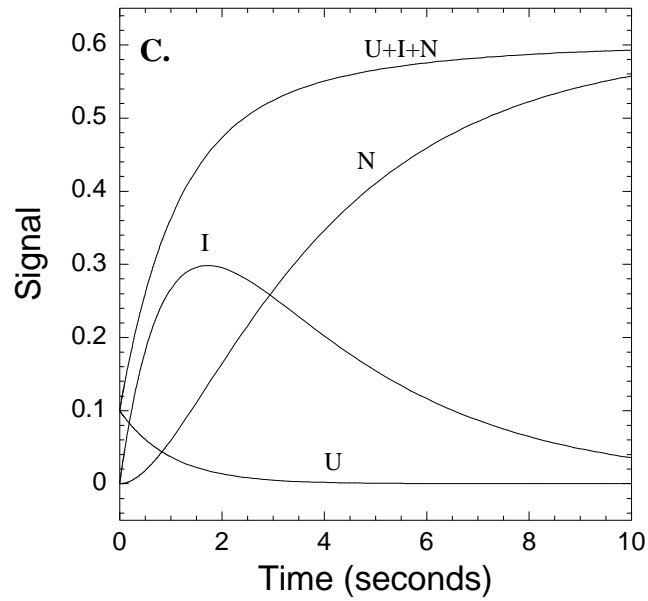
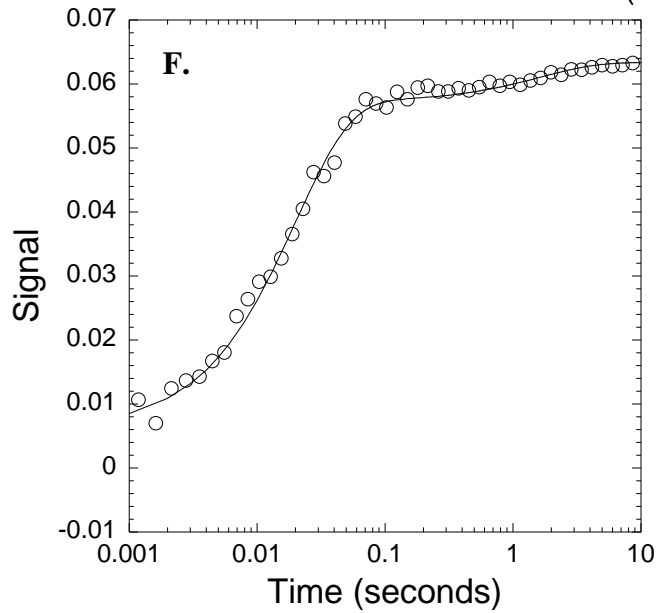
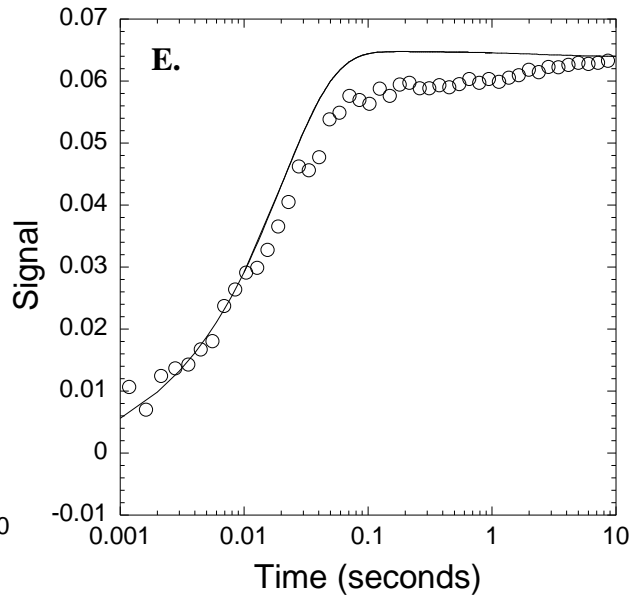
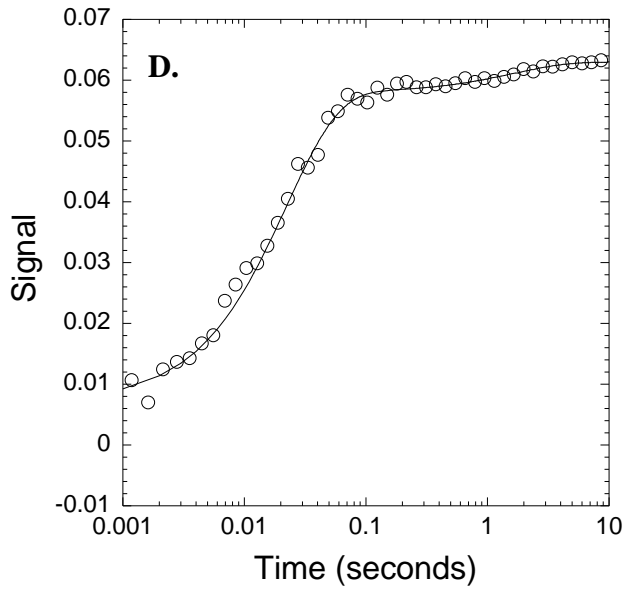
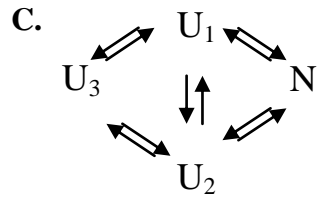
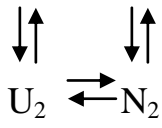
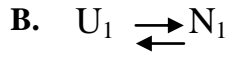
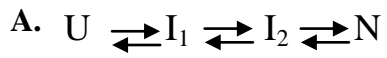


Figure 4. One sequential (**A**) and two parallel folding pathways (**B** and **C**) used in KINSIM simulations. **A**. The unfolded species folds through two intermediates before forming the native species. **B**. Two unfolded and native species exist and interconvert. **C**. Three unfolded species exist, where U_1 and U_2 fold to the native state and U_3 interconverts with U_1 or U_2 . **D-F**. Plots of an Apaf-1 CARD refolding reaction in 3.64 M urea. KINSIM simulations for each plot are shown as a continuous line. For pathway A, the concentration of I_1 was 10 μM . The extinction coefficient for U, I_1 , I_2 and N were 0, 0.0007, 0.0058 and 0.0063, respectively. The rates for I_1 to I_2 and I_2 to N were 45 and 0.6 sec^{-1} , respectively. For pathway B, the concentrations of U_1 and U_2 were 1 and 9 μM , respectively. The extinction coefficients for N_1 and N_2 were 0.0074 and 0.0064. The rates of U_1 to N_1 , U_1 to U_2 , U_2 to N_2 and N_1 to N_2 were 500, 100, 50 and 0.4 sec^{-1} , respectively. For pathway C, the concentrations of U_1 , U_2 and U_3 were 0.7, 8.3 and 1 μM . The extinction coefficient of N was 0.0086. The rates of U_1 to N, U_2 to N and U_3 to U_1 or U_2 were 5000, 37 (back rate of 13) and 0.6 (back rate of 0.01) per second.



CHAPTER III

Rapid Folding and Unfolding of Apaf-1 CARD

Sara L. Milam, Nathan I. Nicely, Brett Feeney, Carla Mattos and A. Clay Clark

Department of Molecular and Structural Biochemistry, North Carolina State University,
Raleigh, North Carolina, 27695-7622

Published in Journal of Molecular Biology, 2007, Vol. 369, p.290-304

ABSTRACT

Caspase recruitment domains (CARDs) are members of the death domain superfamily and contain six antiparallel helices in an α -helical Greek key topology. We have examined the equilibrium and kinetic folding of the CARD of Apaf-1 (apoptotic protease activating factor 1), which consists of 97 amino acid residues, at pH 6 and pH 8. The results showed that an apparent two state equilibrium mechanism is not adequate to describe the folding of Apaf-1 CARD at either pH, suggesting the presence of intermediates in equilibrium unfolding. Interestingly, the results showed that the secondary structure is less stable than the tertiary structure, based on the transition mid-points for unfolding. Single mixing and sequential mixing stopped-flow studies showed that Apaf-1 CARD folds and unfolds rapidly and suggest a folding mechanism that contains parallel channels with two unfolded conformations folding to the native conformation. Kinetic simulations show that a slow folding phase is described by a third conformation in the unfolded ensemble that interconverts with one or both unfolded species. Overall, the native ensemble is formed rapidly upon refolding. This is in contrast to other CARDs in which folding appears to be dominated by formation of kinetic traps.

INTRODUCTION

Studies with several small helical bundle proteins (Huang and Oas, 1995; Kragelund et al., 1996; Kubelka et al., 2004; Sosnick et al., 1996) have shown that folding can be very fast, on the order of 300 s^{-1} to $>5000 \text{ s}^{-1}$, and in many cases folding rates can be predicted by the contact order, an algorithm based on the average sequence distance between all contacting residues in the native conformation (Plaxco et al., 1998). The correlations suggest a relationship between the rate of folding and the topology of the native state (Plaxco et al., 1998). For the more complex topology found in the caspase recruitment domain (CARD) of RICK (RIP-like interacting CLARP kinase) or of procaspase-1 (CP1-CARD), which consists of six antiparallel α - helices in the Greek key topology, folding is too slow to be accounted for by the contact order (Chen and Clark, 2003; Chen and Clark, 2004). For those CARDS, the folding pathways appear to contain kinetically trapped species that significantly slow the folding rate. Kinetic traps in folding have been observed for many other proteins (Capaldi et al., 1999; Chang et al., 2001; Hoyer et al., 2002; Hua et al., 2002; Kiefhaber, 1995).

While relatively little is known about the folding of α -Greek key proteins, much more is known about β -Greek key proteins, as the folding of proteins in this class has been studied for many years. For proteins such as $\beta\gamma$ -crystallin, superoxide dismutase, fibronectin type III domain and azurin (Bosissinot et al., 1997; Capaldi et al., 1999; Hamill et al., 2000; Rajini et al., 2001), several nucleation models have been proposed to explain the folding properties. The models focus on early formation of a β -hairpin (Richardson et al., 1992), hydrophobic

zipper (Hazes and Hol, 1992), tyrosine corner motifs (Bosissinot et al., 1997; Hemmingsen et al., 1994), or hydrogen bonding (Richardson, 1977b). For example, a highly conserved “tyrosine corner” motif is found in several β -Greek key proteins and is important for early structure formation or stability (Bosissinot et al., 1997; Richardson et al., 1992). However, no such folding models have been proposed for the α -Greek key family, and the tyrosine corner motif is not found in this family.

At present, the nature of the intermediates found in the folding of the CARDS is not known, nor is it known why the folding process is so complex. In order to further examine this issue, we studied the folding of Apaf-1 CARD, a 97 amino acid residue domain found at the amino terminus of apoptotic protease-activating factor 1 (Apaf-1). Apaf-1 is a ~130 kDa multidomain protein that is involved in the intrinsic apoptotic pathway (Zou et al., 1997), which is activated when cytochrome c is released from the mitochondria into the cytoplasm (Shiozaki et al., 2002; Zou et al., 1997). The binding of cytochrome c and of ATP cause Apaf-1 to oligomerize into a complex called the apoptosome (Acehan et al., 2002; Hajra and Liu, 2004), a 1 MDa wheel-like particle with seven spokes that radiate from a central hub (Acehan et al., 2002). The Apaf-1 CARD binds the CARD of procaspase-9 (Acehan et al., 2002; Shiozaki et al., 2002; Zhou et al., 1999) and results in dimerization of the inactive procaspase-9 to form the active protease, which in turn activates executioner caspases such as caspases-3 and -7 (Zou et al., 1997). Apaf-1 CARD is a member of the death domain superfamily, which is subdivided into four families consisting of CARD, DD (death

domain), DED (death effector domain) and the PYD (pyrin domain) (Fairbrother et al., 2001; Weber and Vincenz, 2001). To date, the α -Greek key proteins have been found only in this superfamily. The function of these domains is to mediate homo- and heterotypic protein–protein interactions in apoptotic pathways. Within the death domain superfamily there is a low sequence identity, but all members share a common “death fold” consisting of six antiparallel α -helices arranged in a Greek key topology (Weber and Vincenz, 2001).

We show here that the equilibrium folding of Apaf-1 CARD at pH 6 and at pH 8 is not well described by a two-state process, in contrast to RICK-CARD and CP1-CARD, demonstrating the presence of non-native folding intermediate(s) during equilibrium folding. Single mixing kinetic studies indicate that both folding and unfolding occur rapidly, while sequential mixing kinetic studies show that both reactions apparently occur without the presence of folding intermediates. Contrary to RICK-CARD, kinetically trapped or misfolded species are not present in the folding of Apaf-1 CARD, so the protein both folds and unfolds rapidly.

RESULTS

A. Structural studies

The structure of Apaf-1 CARD was determined previously by crystallography (Vaughn et al., 1999) and NMR (Zhou et al., 1999). In addition, a structure of the complex of Apaf-1 CARD bound to caspase-9 (CP9-) CARD was determined (Qin et al., 1999; Zhou et al., 1999). Our clone of Apaf-1 CARD contains two extra residues at the amino terminus, a GS sequence that arises from the cloning vector (Feeney et al., 2006). These residue positions are labeled -2 and -1, respectively (Figure 1). Since the side-chain of M1 forms part of the hydrophobic core (Vaughn et al., 1999), it was important to determine whether the two residues affected the structure in any significant way. We were able to grow crystals of the new Apaf-1 CARD construct using the published crystallization conditions (Vaughn et al., 1999), and they were isomorphous with those previously reported, with symmetry of the space group $P2_12_12_1$ and one molecule in the asymmetric unit (Table 1). Diffraction data were collected at the SER-CAT beam line at APS (Argonne National Laboratory, Argonne, IL) and the structure was refined to 2.0 Å resolution using the published model for initial phases (PDB code 1CY5). While the N-terminal G and S residues were both clearly visible in the electron density maps, they did not make hydrogen bonds or van der Waals contacts with other protein atoms. The sidechain of M1 contributed to the hydrophobic core of the protein exactly as in the published structure where the GS segment is not present. The backbone root mean square deviation (RMSD) between our model and the previously published 1CY5

model was 0.36 Å, demonstrating that the GS sequence had no significant effect on the protein structure.

B. Equilibrium unfolding

We examined the secondary structure of Apaf-1 CARD by far-UV circular dichroism (CD) (data not shown), and the results demonstrated double minima at 206 nm and 218 nm, consistent with the α -helical CARD proteins (Chen and Clark, 2003; Chen and Clark, 2004). In the presence of 6 M urea-containing buffer, the signal decreased to that of a random coil, and the spectrum was consistent with that of an unfolded protein.

The equilibrium unfolding of Apaf-1 CARD was examined by monitoring changes in CD at 222 nm and fluorescence emission at 305 nm following incubation of the protein in urea-containing buffers. The experiments were performed using three protein concentrations and two different buffers as described in Materials and Methods. In addition, the experiments were performed at pH 6, which is near the isoelectric point of the protein, and pH 8. The data collected at pH 8 allow for a comparison with that of RICK-CARD and CP1-CARD (Chen and Clark, 2003; Chen and Clark, 2004; Chen and Clark, 2006), while the data collected at pH 6 should provide the stability of the protein where the net charge is approximately zero (Pace et al., 1990). Overall, the data demonstrate that Apaf-1 CARD does not form an oligomer under these conditions (25 °C, pH 6 or pH 8, <20 μ M protein) (data not shown).

Based on the midpoints of the transitions, the results showed that the protein was more stable at pH 6 than at pH 8 in both CD and fluorescence emission measurements (Figure 2(a) and (b)). While the individual data sets could be fit to a two-state equilibrium model (Table

2), an overlay of the CD and fluorescence emission data showed that the curves are not coincident (Figure 2(c) and (d)). This demonstrated that a two-state folding model was not adequate to explain the equilibrium folding of Apaf-1 CARD. Interestingly, the tertiary structure, as measured by fluorescence emission, appeared to be more stable than the secondary structure. At both pH values, the midpoint of the transition as monitored by fluorescence emission was greater than that monitored by CD, and the difference between the two curves was greater at pH 8 than at pH 6 (Figure 2(c) and (d)). In addition, the free energy values, determined from fits of the data to a two-state equilibrium model, were higher for the fluorescence emission data than for the CD data (Table 2). We note that the data were not better fit to a three-state model that included an equilibrium intermediate ($U \rightleftharpoons I \rightleftharpoons N$), so at present the conformational free energy of Apaf-1 CARD is not known. The thermodynamic parameters obtained from fits of the equilibrium data are listed in Table 2 as a comparison to the kinetic data described below.

C. Single mixing refolding studies

The protein was unfolded initially in 6 M urea-containing buffer, and it was refolded by rapid dilution to conditions that favor the native conformation. As with the equilibrium folding studies described above, the experiments were performed at pH 6 and 8, and both fluorescence emission (>305 nm) and far-UV CD (222 nm) were used to monitor the time-course for refolding. Unless otherwise noted, the results from fluorescence emission at pH 6 are presented in Figure 3, and those using CD at pH 6 as well as both probes at pH 8 are shown in the Supplementary Data Figures. As shown in Figure 3(a) (circles), refolding

occurred within the mixing dead-time as the signal at 1 millisecond was similar to that of the native protein control. Following this burst phase, a small increase in signal occurred in the first 10 seconds that encompassed less than 10% of the total signal change. This phase is described in more detail below. Stopped-flow CD studies at pH 6 showed similar results (Supplementary Data Figure 1A), and results of fluorescence emission and CD studies at pH 8 were similar to those at pH 6 (Supplementary Data Figures 2 and 3).

We monitored the amplitude of the refolding burst phase at several final urea concentrations, from 0.55 M to 6.9 M, and representative data are shown in Figure 3(b) for fluorescence emission at pH 6. Data collected by CD at pH 6 and by fluorescence emission and CD at pH 8 are presented in Supplementary Data Figures 1, 2, 3. The data show a cooperative decrease in signal so that the burst phase of refolding was no longer observed beyond 4 M urea (Figure 3(b)) at pH 6 or 3 M urea at pH 8 (Supplementary Data Figures), demonstrating that the protein folds cooperatively during the burst phase. The results were fit to a two-state equilibrium folding model (Santoro and Bolen, 1988), and the conformational free energies, $\Delta G^{\text{H}_2\text{O}}$, and m-values are presented in Table 2. At both pH values, the conformational free energy determined by fluorescence emission was greater than that determined by CD by >1 kcal/mol. This is consistent with the equilibrium unfolding data (Figure 2, Table 2) that showed the secondary structure unfolded prior to the tertiary structure. Overall, the mid-points for the transitions were lower at pH 8 compared to pH 6, again consistent with the equilibrium unfolding data.

In addition to the burst phase amplitudes, the final signals at 10 seconds were

determined at several final urea concentrations (Figure 3(b), squares) (Supplementary Data Figures 1, 2, 3). The data showed that the transition mid-point was significantly higher than that observed for the burst phase amplitude. The data also were fit to a two-state folding model, and the parameters obtained from the fits are presented in Table 2. In contrast to the burst phase amplitudes, the $\Delta G^{\text{H}_2\text{O}}$ and m -values determined from fluorescence emission or CD were similar at pH 6, 5 kcal mol⁻¹ and 1.3 kcal mol⁻¹ M⁻¹, respectively, although we note that the m -values were similar for the burst phase refolding and the final signals, within experimental error of 1.2 kcal mol⁻¹ M⁻¹. The same was true for data collected at pH 8 (Table 2).

The mid-points for the transitions of the burst phase and final amplitudes (Figure 3(b)) are not coincident because of a kinetic phase that occurs at intermediate urea concentrations. For example, an increase in fluorescence emission was observed over 10 seconds when the protein was refolded in 3.6 M urea-containing buffer (Figure 3(a), diamonds). Under these conditions, the amplitude of the burst phase was small (see also Figure 3(b)), although it still was present, and most of the signal change occurred from the observable kinetic phase. Similar results were obtained by CD at pH 6 (Supplementary Data Figure 1) and by both probes at pH 8 (Supplementary Data Figures 2 and 3). The data were best fit to a two exponential process (Figure 3(a), lower panels), where the slower phase was the same as that described above for refolding at low urea concentrations. Therefore, the data showed that the protein refolded rapidly (burst phase) when it was refolded in low urea concentrations, and refolded through a somewhat slower kinetic phase (half-time of

approximately 20 milliseconds) at higher urea concentrations. Under the latter conditions, the amplitude of the burst phase was low. In both cases, a slow kinetic phase was observed that encompassed less than 10% of the total signal change.

In order to further study refolding at higher urea concentrations, where the burst phase amplitude is low, the two apparent rate constants and their associated amplitudes were determined at several final urea concentrations. While the observed rate associated with the first phase is described in detail later (Figure 3(f)), the amplitude of this phase reached a maximum at 3.6 M urea (pH 6) or 2.6 M urea (pH 8) (Figure 3(c)) and decreased at lower and higher urea concentrations. Although the experiments in Figure 3(c) used fluorescence emission as a probe, the amplitude maxima determined from the data were similar to the midpoint of unfolding determined from equilibrium experiments when the secondary structure was monitored by CD (Figure 2 and Table 2). Overall, the data showed that during refolding under intermediate urea concentrations (~ 2 M to ~ 4 M urea), both the burst phase and the two kinetic phases were observed, and the amplitudes of the burst and fast phases depended on the final urea concentration. In contrast, only the burst and the slower kinetic phases were observed under conditions that strongly favor the native conformation (0.55 M urea). We note that neither the rate associated with the slowest phase (0.6 s^{-1}) (data not shown) nor the amplitude of this phase (Figure 3 (c), squares) varied with final urea concentration or with pH.

D. Single mixing unfolding studies

We monitored the time-course for unfolding Apaf-1 CARD by rapidly mixing

native protein with urea-containing buffer. Representative data are shown in Figure 3(a) for protein unfolded in 6 M urea at pH 6 (triangles). Unfolding monitored by CD at pH 6 or by both probes at pH 8 is presented in Supplementary Data Figures 1–3. The data showed that unfolding occurred in a single phase that encompassed the entire signal between native and unfolded protein. There was no further change in signal between 0.1 and 200 seconds (data not shown), suggesting that the protein unfolded rapidly under these conditions, with a half-time of about 10 milliseconds.

We determined both the amplitude and the rate of the unfolding transition from fits of the unfolding data to a single exponential equation, and this was done at several final urea concentrations. We also monitored the final signal at each final urea concentration. As shown in Figure 3(d), the amplitude of the unfolding phase increased cooperatively as the urea concentration increased. The data were fit to a two-state equilibrium process, and the $\Delta G^{\text{H}_2\text{O}}$ and *m*-value were determined to be $5.1(\pm 0.8)$ kcal mol⁻¹ and $1.2(\pm 0.4)$ kcal mol⁻¹ M⁻¹, respectively, at pH 6 (Table 2). The parameters were similar at pH 8 (Supplementary Data Figure 3 and Table 2). The final signals in unfolding (Figure 3(e) and Supplementary Data Figures) also were fit to a two-state process, and the $\Delta G^{\text{H}_2\text{O}}$ and *m*-value determined from the fits were 5 kcal mol⁻¹ and 1.20 kcal mol⁻¹ M⁻¹, respectively, at pH 6 (Table 2). The conformational free energy was somewhat lower at pH 8 (4.5 kcal mol⁻¹), but the *m*-value was similar to that at pH 6. Because the thermodynamic parameters were the same regardless of whether one monitored the amplitude of the unfolding phase or the final signal of

unfolding versus urea, the data showed that the protein unfolded in a single transition that reached equilibrium within 100 milliseconds.

The apparent rate constants, k_{obs} , of the fast phase of refolding and of unfolding, obtained from fits of the data as described above, generate a chevron plot (Figure 3(f)), demonstrating that these phases represent the folding and unfolding of the same species. Similar rates were obtained regardless of whether fluorescence emission or CD was used to monitor refolding or unfolding (Table 2). In addition, the minimum of the chevron was similar to that determined from equilibrium unfolding at pH 6 (4.3 M versus 4.4 M urea, respectively). At pH 8, however, this was true only for the fluorescence emission data (3.9 M versus 3.5 M urea, respectively). In contrast, the data collected by CD at pH 8 demonstrated a mid-point that was significantly higher than that determined by equilibrium unfolding (3.8 M versus 2.8 M, respectively). The data were fit to a two-state kinetic mechanism (Ferguson et al., 1999a; Jackson and Fersht, 1991) using equation (1):

$$k_{obs} = k_{UN} \exp(-m_{U-TS} \times [urea] / RT) + k_{NU} \exp(m_{N-TS} \times [urea] / RT) \quad (1)$$

where k_{obs} is the observed rate constant of the refolding or unfolding phase, and k_{UN} and k_{NU} are the rate constants for refolding and for unfolding, respectively, in the absence of urea.

The parameters m_{U-TS} and m_{N-TS} reflect the change in solvent exposed surface area between the unfolded or refolded states, respectively, and the transition state ensemble. At pH 6, values of $3537(\pm 640) \text{ s}^{-1}$ and $0.72(\pm 0.05) \text{ kcal mol}^{-1} \text{ M}^{-1}$ were obtained for k_{UN} and m_{U-TS} , respectively, while values of $0.4(\pm 0.3) \text{ s}^{-1}$ and $0.50(\pm 0.11) \text{ kcal mol}^{-1} \text{ M}^{-1}$ were obtained for k_{NU} and m_{N-TS} , respectively. Overall, the data show that the higher pH affects both the rate of

refolding and of unfolding. At pH 8, values of $1427(\pm 174) \text{ s}^{-1}$ and $0.68(\pm 0.04) \text{ kcal mol}^{-1} \text{ M}^{-1}$ were obtained for k_{UN} and $m_{\text{U-TS}}$, respectively, while values of $0.21(\pm 0.3) \text{ s}^{-1}$ and $0.64(\pm 0.19) \text{ kcal mol}^{-1} \text{ M}^{-1}$ were obtained for k_{NU} and $m_{\text{N-TS}}$, respectively. Although there are large errors in these values due to the long extrapolations to zero urea, the conformational free energy and total m-value were calculated from these parameters, where $\Delta G^{\text{H}_2\text{O}} = -RT \ln (k_{\text{UN}}/k_{\text{NU}}) = -5.3 \text{ kcal mol}^{-1}$ at pH 6 and $-5.1 \text{ kcal mol}^{-1}$ at pH 8. The m-values were calculated as $m_{\text{total}} = m_{\text{U-TS}} + m_{\text{N-TS}} = 1.22 \text{ kcal mol}^{-1} \text{ M}^{-1}$ at pH 6 and $1.32 \text{ kcal mol}^{-1} \text{ M}^{-1}$ at pH 8. These results suggest that the fast phase represents rapid folding to the native conformation, since the conformational free energy and m-value are similar to those obtained from equilibrium unfolding data (Table 2).

The compactness (β^\ddagger) of the transition state ensemble relative to the native ($\beta^\ddagger=1$) and unfolded ensembles ($\beta^\ddagger=0$) can be calculated from the ratio of $m_{\text{U-TS}}/m_{\text{total}}$ (Tanford, 1968). For Apaf-1 CARD, $\beta^\ddagger=0.59$ at pH 6, and $\beta^\ddagger=0.52$ at pH 8, demonstrating that the transition state for folding is somewhat more like the native ensemble than the unfolded ensemble.

E. Proposed folding models

The single mixing kinetic studies show that three phases are observed during refolding, two of which depend on the final urea concentration: a burst phase, a fast phase with a time constant of about 10 milliseconds, and a slow phase with a time constant of about 1 second. A minimal model that might explain these results is presented in Scheme 1, which shows that the unfolded protein folds sequentially through two non-native intermediate

structures prior to forming the native conformation. In this Scheme, the burst phase is represented by the transition of U to I₁, the fast phase is represented by the transition of I₁ to I₂, and the slow phase is represented by the transition of I₂ to N. There are features of the kinetic data, however, that are not compatible with this sequential folding scheme. Under conditions that strongly favor the native conformation, Apaf-1 CARD folded within the mixing dead-time to a conformation, presumably the native structure that had a $\Delta G^{\text{H}_2\text{O}}$ and m-value similar to those determined from equilibrium unfolding data. Thus, a significant amount of hydrophobic surface area was buried in the burst phase. As shown in Scheme 1, if the hydrophobic surface is buried in the burst phase, then there should be little buried in the fast and slow phases. The data show, however, that a significant amount of hydrophobic surface area was buried during the fast folding phase as well. In fact, the conformational free energy and m-value derived from the observed rate constants of the fast phase also were similar to those determined from equilibrium folding studies, suggesting that the native conformation formed during this phase. Consequently, there is seemingly a paradox in which two separate refolding transitions result in burial of surface area that is approximately equivalent to that of the native protein. We note that this is not due to formation of transient aggregates because the kinetic data were independent of protein concentration between 5 and 20 μM (data not shown).

A second model that might explain the single mixing kinetic data is the parallel folding model shown in Scheme 2. In this Scheme, the unfolded ensemble consists of two conformations, U₁ and U₂, each of which fold to native conformations, N₁ and N₂. This

Scheme allows refolding in the burst phase or in the fast phase to be separated into parallel folding channels. For example, folding in the burst phase yielded a species with a conformational free energy of 3–4 kcal mol⁻¹, depending on the probe used to monitor refolding, and an m-value of 1.2– 1.4 kcal mol⁻¹ M⁻¹. Regardless of the probe, the conformational free energy was about 1 kcal mol⁻¹ lower than that obtained from fitting the equilibrium unfolding data to a two-state model, although the m-values were similar. Likewise, folding in the fast phase yielded a species with a conformational free energy of 5.3 kcal mol⁻¹ and an m-value of 1.22 kcal mol⁻¹ M⁻¹, also similar to those determined from equilibrium unfolding. If the transition of U₁→N₁ represents folding that occurs in the burst phase, then the transition of U₂→N₂ represents folding in the fast phase. Under conditions that strongly favor the native conformation, that is, low urea concentrations, the protein folds from U₁→N₁ in the burst phase. At intermediate urea concentrations, refolding occurs from both channels, and at higher urea concentrations (>4 M), the protein refolds from U₂→N₂ in the fast phase. This simple scheme explains the burial of hydrophobic surface area from two major transitions and the urea dependence to each major folding phase. The slow phase, which is independent of urea concentration, could represent isomerizations in the unfolded ensemble or the native ensemble. We performed sequential mixing stopped-flow studies, described below, to further examine the proposed folding schemes.

F. Sequential mixing kinetic studies

Initially, we examined the transitions using double jump studies. In these experiments, the native protein was mixed rapidly with urea, and after a specified time of

unfolding (delay time), the protein was returned to native conditions (Brandts et al., 1975; Kiefhaber et al., 1990a). The fluorescence signal was monitored experimentally after the second mix. This experiment examined the formation of unfolded species that were able to fold to the native conformation upon dilution of the denaturant, and the refolding traces should resemble those shown in Figure 3(a) only when those species are formed. As described above for single mixing refolding studies, when the protein was refolded in buffer containing low urea concentrations, only the burst phase and slow refolding phases were observed. Thus, in the double jump experiment, one should monitor the formation of the unfolded conformation by monitoring these two phases as a function of unfolding delay time. A measure of the refolding burst phase signal versus unfolding delay time showed that the signal reached that of the native conformation within 100 milliseconds of unfolding (Figure 4, circles). That is, when the protein was unfolded for 100 milliseconds, it formed a species that could refold to the native conformation. This was true at pH 6 and pH 8 (Figure 4). The data are in agreement with the single mixing unfolding studies (Figure 3(a)), which showed that Apaf-1 CARD unfolded rapidly in a single transition. The double jump experiments showed that the signal of the refolding burst phase remained constant even after extended periods of unfolding (>200 seconds). In addition, the amplitude of the slow phase reached a maximum at 1 second and did not change with longer delay times (data not shown).

The double jump experiment was repeated to monitor refolding in 4 M urea. In this experiment, the protein was initially unfolded in 5 M urea-containing buffer for various delay

times, and then the urea was diluted to 4 M. As shown in Figure 3(b), the burst phase of refolding was not observed at urea concentrations of 4 M or greater; although the fast and slow phases were observed under these conditions (Figure 3(a) and (c)). The amplitude of the fast phase versus delay time is shown in Figure 4 (triangles) for data collected at pH 6 (Figure 4(a)) or pH 8 (Figure 4(b)). The results showed that the fast phase of refolding appeared at short unfolding delay times (100 milliseconds), and the amplitude of the phase did not change with longer times of unfolding. Therefore, the unfolded species that refolds via this phase formed fully within 100 milliseconds of unfolding. As described above, the amplitude of the slow phase reached a maximum at 1 second and did not vary with longer delay times (data not shown).

In addition to double jump experiments, we examined the refolding pathways in interrupted refolding sequential mixing studies (Pappenberger et al., 2001; Schmid, 1983). In these experiments, the unfolded protein was rapidly mixed with buffer, conditions that favor the native conformation, and the protein was returned to denaturing conditions after a specified time of refolding (delay time). The fluorescence signal was monitored experimentally following the second mix. This experiment monitored formation of the native conformation as a function of refolding time, and the unfolding traces should resemble those shown in Figure 3(a) when the native species is formed. As shown in Figure 3(a), unfolding occurred from a single transition with a half-time of 10 milliseconds, so in the interrupted refolding experiment one should observe the fast unfolding transition only when the native conformation is formed. The results from this experiment are summarized in Figure 4

(squares) for data collected at pH 6 (Figure 4(a)) or pH 8 (Figure 4(b)). The amplitude of the unfolding transition and the apparent rate constant, k_{obs} , were obtained by fitting the raw data to a single exponential process. We note that k_{obs} did not vary with refolding delay time and was similar to that determined from the single mixing unfolding studies (Figure 3(a)). In addition, the maximum amplitude of the phase was obtained within 1 millisecond of refolding, and the amplitude did not change with longer refolding delay times (Figure 4). Attaining the maximum amplitude within 100 milliseconds of refolding showed that the native conformation was fully populated within this time and further demonstrated the absence of slow folding species on the folding pathway.

G. Kinetic simulations

The sequential mixing kinetic studies described above ruled out the sequential folding model shown in Scheme 1 because the native ensemble formed rapidly upon refolding, either in low urea concentrations or in 4 M urea. As described for the single mixing studies, these refolding phases were observed as burst phase or as fast phase kinetics, respectively. Furthermore, the unfolded ensemble formed within 100 milliseconds of unfolding. Together, the results showed that the native ensemble was directly accessible from the unfolded ensemble. Formally, we have presented evidence for three kinetic phases during refolding: a burst phase, in which the native ensemble formed at low concentrations of urea, a fast phase, in which the native ensemble also formed at higher urea concentrations, and a slow phase, which encompassed less than 10% of the total signal change but was observed at all urea concentrations examined. In intermediate concentrations of urea, the native conformation

formed simultaneously from the burst phase and the fast phase of refolding. While the data are consistent with parallel folding channels, there is no direct evidence for four species as presented in Scheme 2. We further examined this folding model, as well as variations of the model, using kinetic simulations (Barshop et al., 1983).

We used the following reasoning in the simulations. The rate of refolding in the burst phase was assumed to be 5000 s^{-1} , which is sufficiently fast to allow for a burst phase in the simulations. In simulating refolding in 0.55 M urea, the final conditions in low urea concentrations, the fast phase was calculated to be 1800 s^{-1} (Figure 3(f)). The rate of refolding in 3.6 M urea was calculated to be 41 s^{-1} (Figure 3(f)). Those values were used to simulate refolding in low urea or in higher urea concentrations, respectively. In addition, simulations with a single unfolded species demonstrated that refolding occurred primarily in the burst phase, since the rate was significantly faster than that of the fast phase in 3.6 M urea (5000 s^{-1} versus 41 s^{-1}). Thus, in order to observe refolding in the burst phase and the fast phase simultaneously, which is detected at intermediate urea concentrations, two unfolded species were populated.

We used a single native conformation in the simulations. This assumption was supported by single mixing unfolding experiments that showed that the protein unfolded in a single transition (Figure 3(a)). Furthermore, unfolding occurred from the native species that formed during the fast phase of refolding (Figure 3(f)). Therefore, regardless of whether the protein was refolded in low urea concentrations or in 4 M urea, unfolding occurred from a single species. If two native species were present at equilibrium, as suggested in Scheme 2,

one might expect two transitions during unfolding. Otherwise, one must assume either that the two native species share the same rate-limiting step in unfolding or that one native species rapidly relaxes to the second native species prior to unfolding. Without further evidence to support these assumptions, we elected to use a single native conformation in the simulations. We note that sequential mixing unfolding experiments (that is, interrupted refolding studies) (Figure 4) demonstrated that the burst phase species was not positioned on-pathway between the unfolded and native ensembles. Indeed, those results showed that the native conformation formed within 100 milliseconds of refolding. Likewise, sequential mixing refolding studies (that is, double jump experiments) (Figure 4) demonstrated that the unfolded ensemble formed rapidly from the native species. This was true also when the experiments were carried out in 4 M urea. In addition, structural studies of Apaf-1 CARD using NMR (Zhou et al., 1999) did not support the presence of two native conformations.

The slow phase in refolding was somewhat more problematic to simulate. For example, the slow phase was observed under conditions where the burst phase was not observed (Figure 3(b)). The slow phase also was observed when refolding occurred under conditions that favored the burst phase (Figure 3(a) and (c)). Neither the amplitude nor the rate of the slow phase was affected by urea concentration (Figure 3(c)). The simplest explanation for the slow refolding phase is the presence of slowly interconverting species in the unfolded ensemble. The simulations suggested, however, that the slow phase was not due to interconversion of the two unfolded species that yield the native conformation. Because the refolding rates were significantly faster than that of the slow phase, all of the unfolded

species rapidly converted to the native conformation upon simulated refolding, leaving no species in the unfolded conformation to interconvert. In order to simulate the slow phase, we assumed that the unfolded ensemble contained a third unfolded conformation, U_3 , that slowly interconverted with U_1 and U_2 .

Based on those rationales, we propose the model shown in Scheme 3. In this model, the native conformation is accessible either from U_1 or U_2 . Refolding from U_1 to N occurs in the burst phase, and refolding from U_2 to N occurs in the fast phase. The slow phase is represented by the transition of U_3 either to U_1 or to U_2 . Furthermore, in this model, N unfolds to U_2 regardless of whether the protein was refolded in 0.55 M urea ($U_1 \rightleftharpoons N$) or in 3.6 M urea ($U_2 \rightleftharpoons N$). Thus, under the conditions used here to unfold the protein (6 M urea-containing buffer at pH 6), U_2 is the preferred conformation of the unfolded ensemble, with $\sim 10\%$ found as U_3 . In the simulations, we assumed that the protein unfolded through a single channel ($N \rightleftharpoons U_2$), so we have no information on the rate constant k_{-1} , that is, unfolding from N to U_1 . Upon refolding in 3.6 M urea, approximately 80% of the protein refolds from U_2 , and about 10% refolds from U_1 . Thus, by changing the distribution of U_1 and U_2 at intermediate urea concentrations, the burst phase and the fast phase are observed simultaneously during refolding. At low urea concentrations, the protein refolds primarily from U_1 ($\sim 80\%$). We note that the rates of interconversion between U_1 and U_2 are not known at present (k_2 and k_{-2} in Table 3), although the formation of U_1 from U_2 upon dilution from urea must be rapid in order to observe the burst phase. Under all conditions, about 10% of the protein refolds from U_3 . The rate constants determined from simulations of the data at pH 6

and at pH 8 are presented in Table 3 and are compared to those calculated from the experimental data. In addition, simulations of the refolding and unfolding data are shown in Figure 5(a) and (b) for data at pH 6 and 8, respectively, and are compared to the experimental data. Overall, the results show good agreement between the simulated and experimental data at both pH values.

DISCUSSION

Since our Apaf-1 CARD construct contained two extra residues at the amino terminus, we crystallized the protein and determined the structure to 2 Å resolution in order to examine the effect of the residues on the structure. The results showed that the GS sequence does not hydrogen bond, nor does it form van der Waals contacts, with other atoms of the protein and thus does not significantly affect the structure of the protein.

We showed that refolding and unfolding of Apaf-1 CARD occurred rapidly. For example, at low urea concentrations, refolding occurred within the dead-time for mixing. An examination of the burst phase signals versus urea concentration showed that the mid-point of the transition was significantly lower than that of equilibrium unfolding regardless of whether fluorescence emission or CD kinetic studies were performed. At higher urea concentrations, the protein folded to the native conformation simultaneously through two distinct phases. Together the kinetic data strongly suggest that the protein folds through parallel pathways, and the flux through each pathway depends on the final urea concentration. Simulations of the data suggested the model shown in Scheme 3 in which the unfolded ensemble consists of three conformations. The native conformation is accessible from U_1 and/or U_2 , depending on the final urea concentration, and isomerizations from U_3 account for the slow phase observed in the refolding experiments. At high urea concentrations (>5 M), the majority of the protein favored one unfolded conformation, which we designated U_2 .

We also have shown that equilibrium unfolding of Apaf-1 CARD is not consistent with a two-state process, since unfolding of the secondary and tertiary structures were not coincident. Interestingly, the tertiary structure appeared to be more stable than the secondary structure, since the midpoint of the unfolding transition ($urea_{1/2}$) was higher when fluorescence emission was monitored compared to far-UV CD. These results are contrary to those for RICK-CARD (Chen and Clark, 2003) and CP1-CARD (Chen and Clark, 2004), where there was no evidence for partially folded intermediates at equilibrium.

The equilibrium and kinetic studies of Apaf-1 CARD were carried out at pH 6, in order to examine folding near the isoelectric point, and at pH 8, in order to compare to the folding of other CARD proteins. The equilibrium data suggested that Apaf-1 CARD is less stable at pH 8 than at pH 6, since the mid-points of the unfolding transitions were lower compared to those at pH 6. In addition, there were larger differences between the two structural probes, CD and fluorescence emission, at pH 8 compared to pH 6. Because of the differences over this pH range, we suggest that charged histidine residues favor stability at pH 6. An examination of the structure shows that one histidine resides on each of helices 1, 2, 5, and 6 (Figure 1(c)). The tyrosine residues reside on helix 2 (Y24), helix 5 (Y68 and Y72), and turn 5 (Y80), which is located between helices 5 and 6. Based on the positioning of the tyrosine residues on one half of the protein, as shown in Figure 1, our equilibrium unfolding data suggest that helices 2 and 5, and possibly helix 6, are more stable than helices 1, 3 and 4, since the fluorescence emission unfolding profiles demonstrated a higher $urea_{1/2}$ than did the CD unfolding profiles, which monitor changes in secondary structure. The

intermediates observed at equilibrium (Figure 2) likely represent the heterogeneity observed in the unfolded ensemble (Scheme 3). Changes in the final urea concentration affect the relative populations of U_1 and of U_2 , and we speculate that this also affects the rates at which helices 1, 3, and 4 fold and dock against the core. This interpretation is consistent with structural studies of the pyrin domain of Nalp1, which showed that helix 3 is disordered (Hiller et al., 2003). For Apaf-1 CARD at pH 8, where the histidine residues are uncharged, the secondary structure appeared much less stable than the tertiary structure. This may be due to loss of interactions between histidine 88, on helix 6, and tyrosine 72, on helix 5, which are 4.6 Å apart (Figure 1(c)). Although this requires further study, the loss of these putative interactions may destabilize helices 5 and 6.

When compared to previous results for RICK-CARD (Chen and Clark, 2003) or CP1-CARD (Chen and Clark, 2004), the results described here show that the folding of Apaf-1 CARD is complex, as it is for the other CARDS. Although the rates differed, the folding pathways were similar at pH 6 and pH 8 for Apaf-1 CARD, and the data collected at pH 8 allow a direct comparison of the three proteins. As noted above, there is no evidence to date for equilibrium intermediates in the folding of RICK or CP1 CARDS, and both proteins are adequately described by an apparent two-state equilibrium folding model. In addition, the apparent m -value for Apaf-1 CARD, as shown in Table 2, is similar to that of RICK-CARD ($\sim 1.3 \text{ kcal mol}^{-1} \text{ M}^{-1}$) and is much larger than that of CP1-CARD ($0.65 \text{ kcal mol}^{-1} \text{ M}^{-1}$). As noted previously (Chen and Clark, 2003), the m -values of RICK-CARD and of Apaf-1 CARD are consistent with proteins of 97 amino acid residues, based on the work of Scholtz

and co-workers (Myers et al., 1995b). It was suggested, however, that CP1-CARD contains non-native structure in the native ensemble (Chen and Clark, 2004), explaining the low m -value. The folding pathways of CP1-CARD and of RICK-CARD are complex and contain misfolded or kinetically trapped species. While CP1-CARD appeared to fold rapidly in single mixing kinetic studies, sequential mixing studies showed that the native conformation formed fully only after about 100 seconds of refolding (Chen and Clark, 2004). This was true also for RICK-CARD, except that several phases were observed during folding (Chen and Clark, 2003). In the case of RICK-CARD, an analysis of several proline-to-alanine replacement mutants demonstrated that the kinetic complexity was not due to the three proline residues present in that protein (Chen and Clark, 2006). That is, the kinetically trapped species was present in a triple mutant that lacked proline residues. As noted, the data for RICK-CARD are difficult to reconcile with a sequential folding pathway, and it was suggested that the protein folds via parallel channels (Chen and Clark, 2006). Like RICK-CARD, Apaf-1 CARD also appears to fold via parallel channels. Apaf-1 CARD contains two proline residues, P47, in turn three, and P92 near the C terminus of the protein (Figure 1), although at present it is not clear whether either proline is responsible for the slow isomerization observed in the refolding of Apaf-1 CARD. P47 is conserved between Apaf-1 CARD and RICK-CARD but is not conserved in other CARDS. While removal of P47 destabilized RICK-CARD, the folding kinetics were similar to those of the wild-type protein. The results described here show that folding of Apaf-1 CARD is complex, as observed with other CARDS, although the complexity does not result from formation of kinetically trapped

conformations. In contrast to Apaf-1 CARD, we showed previously that the fast phase in refolding of RICK-CARD has an apparent rate of $\sim 32 \text{ s}^{-1}$, and this transition represents the primary folding phase in terms of burial of hydrophobic surface area (Chen and Clark, 2003). Interestingly, the transition state placements of Apaf-1 CARD ($\beta_{\ddagger}^{\ddagger} = m_{\text{U-TS}}/m_{\text{UN}} = 0.52$ at pH 8) and of RICKCARD ($\beta_{\ddagger}^{\ddagger} = 0.71$) (Chen and Clark, 2003) correlate with their refolding rates. As described previously (Plaxco et al., 1998), this suggests that the more poorly packed transition state for Apaf-1CARD results in a lower loss in chain entropy while overcoming the rate-limiting step to folding. In contrast, the more highly organized folding transition state for RICK-CARD may result in higher chain entropy losses and hence a decrease in the folding rate. To date, this is the best correlation that explains the differences in folding rates between the CARDS, although it does not explain the formation of kinetically trapped species in the folding of RICK-CARD or CP1-CARD.

METHODS

Materials

Ultrapure urea was purchased from ICN. Dithiothreitol (DTT) was from Acros. 2-(4-morpholino)-ethane sulfonic acid (Mes), sodium chloride (NaCl), 2-mercaptoethanol, trizma base (Tris), and reduced glutathione were from Fisher. Acetic acid was from VWR. DEAE Sepharose, monobasic and dibasic potassium phosphate (KH_2PO_4 , K_2HPO_4), ampicillin, and molecular weight markers were from Sigma. Isopropyl β -D-1-thiogalactopyranoside (IPTG) was from Anatrace. All buffers were filtered through either 0.45 or 0.22 μm filter membranes. The urea-containing buffers were prepared as described (Pace et al., 1989). Apaf-1 CARD was expressed and purified from *Escherichia coli* BL21(DE3) pLysS cells harboring plasmid pApaf-1CARD+GS as described (Feeney et al., 2006). An extinction coefficient (ϵ_{280}) of $5120 \text{ M}^{-1} \text{ cm}^{-1}$ was used to determine protein concentration. Caspase-3 was purified as described (Feeney et al., 2004; Pop et al., 2001).

Crystallography

Apaf-1 CARD was dialyzed at a concentration of 5 mg/ml in a buffer of 10 mM Tris-HCl (pH 7.5), 10 mM NaCl. The best crystals were formed using the hanging-drop method by mixing 2 μl of protein with 2 μl of reservoir buffer containing 200 mM Tris-HCl (pH 6.1), 200 mM zinc acetate, 17% (w/v) PEG 6000. Crystals were grown for approximately one week at 18 °C and flash frozen using glycerol (10%, v/v) as a cryoprotectant. Data sets were collected at 100 K at the SER-CAT synchrotron beamline (Advance Photon Source, Argonne

National Laboratory, Argonne, IL) and processed with HKL2000 (Otwinowski and Minor, 1997). Ten percent of the unique reflections were randomly selected and reserved for the calculation of R_{free} (Brünger, 1997). The crystal structure was solved by the molecular replacement method using the CNS software package (Brunger et al., 1998) with the previously published coordinates (PDB code 1CY5 (Vaughn et al., 1999)) as the search model. The model was refined with alternating rounds of manual rebuilding in O (Jones et al., 1991) with visualization of F_o-F_c and 2F_o-F_c electron density maps and crystallographic refinement in CNS (Brunger et al., 1998). The final structure was checked with the online resource MolProbity (Lovell et al., 2003) and deposited in the Protein Data Bank.

Equilibrium unfolding

Equilibrium unfolding experiments were performed as described (Chen and Clark, 2003; Chen and Clark, 2004) by mixing native protein with urea-containing MTA (10 mM Mes, 20 mM Tris, 10 mM acetate (pH 6.0 or pH 8.0), 1 mM DTT) or potassium phosphate (50 mM KH₂PO₄/K₂HPO₄ (pH 6.0), 1 mM DTT) buffers. The final protein concentrations were 5, 10 or 15 μ M, and a quartz cuvette with a 1 cm path length was used for data collection. The CD signal at 222 nm was measured for 20 seconds, and the data were averaged. In addition, fluorescence emission was measured at 305 nm following excitation at 280 nm. The data versus urea concentration were fit globally to a two-state equilibrium process (Santoro and Bolen, 1988) using the program IgorPro (Wavemetrics Inc.). All samples were maintained at a constant temperature of 25 °C using a circulating water bath. In control experiments, stock protein was initially unfolded in 6 M urea-containing

MTA or phosphate buffer, the samples were diluted to several final urea concentrations that encompass the folding transition, and the samples were incubated overnight at 25 °C. The signals were identical to those shown in Figure 2, in which the protein stock was not initially unfolded (data not shown).

Single mixing stopped-flow fluorescence

Single mixing unfolding and refolding fluorescence (SX18, Applied Photophysics, UK) and circular dichroism (PiStar, Applied Photophysics, UK) experiments were performed as described (Chen and Clark, 2003) with the following modifications. For refolding experiments, protein stock solutions (55, 110 or 220 μM) that had been prepared in MTA buffer containing 6 M urea were rapidly diluted into MTA buffer containing urea between 0 and 7 M. For unfolding experiments, protein stock solutions (55, 110 or 220 μM) in MTA buffer were mixed with MTA buffer containing urea between 0 and 7M. For both refolding and unfolding, a mixing ratio of 1:10 was used, the final protein concentrations were 5, 10, or 20 μM , and the final urea concentrations are shown in the Figures.

Sequential mixing stopped-flow experiments

Double jump stopped-flow experiments were performed as described (Chen and Clark, 2003) with the following modifications. For experiments at pH 6 or at pH 8, native protein (120 μM) in MTA buffer was mixed 1:1 with 10 M urea-containing MTA buffer for the first jump. After various delay times, as described in the text, the protein solution was mixed 1:5 with MTA buffer, and the signal trace was monitored for 10 seconds. The final protein concentration was 10 μM , and the final urea concentration was 0.83 M. For the native

protein signal, a protein stock solution (120 μM) in 0.83 M urea-containing MTA buffer was mixed with 0.83 M urea-containing MTA buffer. For the unfolded protein signal, a protein stock solution (120 μM) in 5 M urea-containing buffer was mixed with 5 M urea-containing MTA buffer.

For sequential mixing double jump studies in 4 M urea (pH 6) or 3.18 M urea (pH 8) (final concentrations), native protein (120 μM) in MTA buffer (pH 6 or pH 8) was mixed 1:1 with 10 M urea-containing buffer for the first jump. After various delay times, as described in the text, the protein solution was mixed 1:5 with 3.8 M (pH 6.0) or 2.82 M (pH 8.0) urea-containing MTA buffer, and the signal trace was monitored for 10 s. The final protein concentration was 10 μM , and the final urea concentration was 4 M (pH 6) or 3.18 M (pH 8). For the native protein signal at pH 6, a protein stock solution (120 μM) in 4 M urea-containing MTA buffer was mixed with 4 M urea-containing MTA buffer. For the native protein signal at pH 8, a protein stock solution (120 μM) in 3.18 M urea-containing MTA buffer was mixed with 3.18 M urea-containing MTA buffer. For the unfolded protein signal at pH 6 and pH 8, a protein stock solution (120 μM) in 5 M urea-containing buffer was mixed with 5 M urea-containing MTA buffer.

Interrupted refolding stopped-flow experiments were performed as described (Chen and Clark, 2003) with the following modifications. Unfolded protein (360 μM) in 5 M urea-containing MTA buffer was mixed 1:5 with MTA buffer (pH 6 or pH 8) for the first jump. After various delay times, as described in the text, the protein solution was mixed 1:5 with 5.9 M urea-containing MTA buffer (pH 6 or pH 8). The final protein concentration was 10

μM , and the final urea concentration was 5.1 M. For the native protein signal, a protein stock solution (120 μM) in 0.83 M urea-containing MTA buffer (pH 6 or pH 8) was mixed with 0.83 M urea-containing MTA buffer. For the unfolded protein signal, a protein stock solution (120 μM) in 5 M urea-containing buffer was mixed with 5 M urea-containing MTA buffer.

Protein Data Bank accession code

The atomic coordinates have been deposited in the RCSB Protein Data Bank and are available under accession code 2P1H.

ACKNOWLEDGMENTS

This work was supported by a grant from the National Institutes of Health (GM065970 to A.C.C.). The authors thank Dr Paul Swartz for assisting in the final stages of structure refinement and the research agencies of North Carolina State University and the North Carolina Agricultural Research Service for continued technical support. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. W-31-109-ENG-38.

REFERENCES

- Acehan, D., Jiang, X., Morgan, D. G., Heuser, J. E., Wang, X., and Akey, C. W. (2002). Three-dimensional structure of the apoptosome: implications for assembly, procaspase-9 binding, and activation. *Molecular Cell* **9**, 423-432.
- Barshop, B. A., Wrenn, R. F., and Frieden, C. (1983). Analysis of numerical methods for computer simulation of kinetic processes: development of KINSIM--a flexible, portable system. *Analytical Biochemistry* **130**, 134-145.
- Bosissinot, M., Karnas, S., Lepock, J. R., Cabelli, D. E., Tainer, J. A., Getzoff, E. D., and Hallewell, R. A. (1997). Function of the Greek key connection analysed using circular permutants of superoxide dismutase. *European Molecular Biology Organization Journal* **16**, 2171-2178.
- Brandts, J. F., Halvorson, H. R., and 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**, 4953-4963.
- Brünger, A. T. (1997). Free R value: Cross-validation in crystallography. *Macromolecular Crystallography Part B*, 366-396.
- Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallographica Section D* **54**, 905-921.
- Capaldi, A. P., Ferguson, S. J., and Radford, S. E. (1999). The greek key protein apopseudoazurin folds through an obligate on-pathway intermediate. *Journal of Molecular Biology* **286**, 1621-1632.
- Chang, J.-Y., Li, L., and Lai, P.-H. (2001). A major kinetic trap for the oxidative folding of human epidermal growth factor. *Journal of Biological Chemistry* **276**, 4845-4852.
- Chen, Y.-R., and Clark, A. C. (2003). Equilibrium and kinetic folding of a α -helical greek key protein domain: caspase recruitment domain (CARD) of RICK. *Biochemistry* **42**, 6310-6320.

- Chen, Y.-R., and Clark, A. C. (2004). Kinetic traps in the folding/unfolding of procaspase-1 CARD domain. *Protein Science* **13**, 2196-2206.
- Chen, Y.-R., and Clark, A. C. (2006). Substitutions of prolines examine their role in kinetic trap formation of the caspase recruitment domain (CARD) of RICK. *Protein Science* **15**, 395-409.
- Fairbrother, W. J., Gordon, N. C., Humke, E. W., O'Rourke, K. M., Starovasnik, M. A., Yin, J.-P., and Dixit, V. M. (2001). The PYRIN domain: a member of the death domain-fold superfamily. *Protein Science* **10**, 1911-1918.
- Feeney, B., Pop, C., Tripathy, A., and Clark, A. C. (2004). Ionic interactions near the loop L4 are important for maintaining the active-site environment and the dimer stability of (pro)caspase 3. *Biochemical Journal* **384**, 515-525.
- Feeney, B., Soderblom, E. J., Goshe, M. B., and Clark, A. C. (2006). Novel protein purification system utilizing an N-terminal fusion protein and a caspase-3 cleavable linker. *Protein Expression and Purification* **47**, 311-318.
- Ferguson, N., Capaldi, A. P., James, R., Kleanthous, C., and Radford, S. E. (1999). Rapid folding with and without populated intermediates in the homologous four-helix proteins Im7 and Im9. *Journal of Molecular Biology* **286**, 1597-1608.
- Hajra, K. M., and Liu, J. R. (2004). Apoptosome dysfunction in human cancer. *Apoptosis* **9**, 691-704.
- Hamill, S. J., Cota, E., Chothia, C., and Clarke, J. (2000). Conservation of folding and stability within a protein family: the tyrosine corner as an evolutionary cul-de-sac. *Journal of Molecular Biology* **295**, 641-649.
- Hazes, B., and Hol, W. G. J. (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: Structure, Function, and Genetics* **12**, 278-298.
- Hemmingsen, J. M., Gernert, K. M., Richardson, J. S., and Ricardson, D. C. (1994). The tyrosine corner: A feature of most Greek key beta-barrel proteins. *Protein Science* **3**, 1927-1937.

- Hiller, S., Kohl, A., Fiorito, F., Herrmann, T., Wider, G., Tschopp, J., Grütter, M. G., and Wüthrich, K. (2003). NMR structure of the apoptosis- and inflammation-related NALP1 pyrin domain. *Structure* **11**, 1199-1205.
- Hoyer, W., Ramm, K., and Plückthun, A. (2002). A kinetic trap is an intrinsic feature in the folding pathway of single-chain Fv fragments. *Biophysical Chemistry* **96**, 273-284.
- Hua, Q.-X., Jia, W., Frank, B. H., Phillips, N. F. B., and Weiss, M. A. (2002). A protein caught in a kinetic trap: structures and stabilities of insulin disulfide isomers. *Biochemistry* **41**, 14700-14715.
- Huang, G. S., and Oas, T. G. (1995). Submillisecond folding of monomeric lambda repressor. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 6878-6882.
- Jackson, S. E., and Fersht, A. R. (1991). Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition. *Biochemistry* **30**, 10428-10435.
- Jones, T. A., Zou, J.-Y., Cowan, S. W., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallographica Section A* **47**, 110-119.
- Kiefhaber, T. (1995). Kinetic traps in lysozyme folding. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 9029-9033.
- Kiefhaber, T., Quaas, R., Hahn, U., and Schmid, F. X. (1990). Folding of ribonuclease T1. 1. Existence of multiple unfolded states created by proline isomerization. *Biochemistry* **29**, 3053-3061.
- Kragelund, B. B., Højrup, P., Jensen, M. S., Schjerling, C. K., Juul, E., Knudsen, J., and Poulsen, M. (1996). Fast and one-step folding of closely and distantly related homologous proteins of a four-helix bundle family. *Journal of Molecular Biology* **256**, 187-200.
- Kubelka, J., Hofrichter, J., and Eaton, W. A. (2004). The protein folding "speed limit". *Current Opinion in Structural Biology* **14**, 76-88.
- Lovell, S. C., Davis, I. W., Arendall, W. B. r., de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S., and Richardson, D. C. (2003). Structure validation by Calpha geometry: phi, psi and Cbeta deviation. *Proteins: Structure, Function, and Genetics* **50**, 437-450.

- Myers, J. K., Pace, C. N., and Scholtz, J. M. (1995). Denaturant m values and heat capacity changes: Relation to changes in accessible surface areas of protein unfolding. *Protein Science* **4**, 2138-2148.
- Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology*, 307-326.
- Pace, C. N., Laurents, D. V., and Thomson, J. A. (1990). pH dependence of the urea and guanidine hydrochloride denaturation of ribonuclease A and ribonuclease T1. *Biochemistry* **29**, 2564-2572.
- Pace, C. N., Shirley, B. A., and Thomson, J. A. (1989). Measuring the conformational stability of a protein. In "Protein Structure and Function, A Practical Approach" (T. Creighton, Ed.), pp. 311-329. IRL Press, New York.
- Pappenberger, G., Aygun, H., Engels, J. W., Reimer, U., Fischer, G., and Kiefhaber, T. (2001). Nonprolyl cis peptide bonds in unfolded proteins cause complex folding kinetics. *Nature Structural Biology* **8**, 452.
- Plaxco, K. W., Simons, K. T., and Baker, D. (1998). Contact order, transition state placement and the refolding rates of single domain proteins. *Journal of Molecular Biology* **277**, 985-994.
- Pop, C., Chen, Y.-R., Smith, B., Bose, K., Bobay, B., Tripathy, A., Franzen, S., and Clark, A. C. (2001). Removal of the pro-domain does not affect the conformation of the procaspase-3 dimer. *Biochemistry* **40**, 14224-14235.
- Qin, H., Srinivasula, S. M., Wu, G., Fernandes-Alnemri, T., Alnemri, E. S., and Shi, Y. (1999). Structural basis of procaspase-9 recruitment by the apoptotic protease-activating factor 1. *Nature* **399**, 549-557.
- Rajini, B., Shridas, P., Sundari, C. S., Muralidhar, D., Chandani, S., Thomas, F., and Sharma, Y. (2001). Calcium binding properties of gamma-crystallin. *Journal of Biological Chemistry* **276**, 38464-38471.
- Richardson, J. S. (1977). beta-sheet topology and the relatedness of proteins. *Nature* **268**, 495-500.

- 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., and Surlus, M. C. (1992). Looking at proteins: representations, folding, packing, and design. *Biophysical Journal* **63**, 1185-1209.
- Santoro, M. M., and 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**, 8063-8068.
- Schmid, F. X. (1983). Mechanism of folding of ribonuclease A. Slow refolding is a sequential reaction via structural intermediates. *Biochemistry* **22**, 4690-4696.
- Shiozaki, E. N., Chai, J., and Shi, Y. (2002). Oligomerization and activation of caspase-9, induced by apaf-1 CARD. *Proceedings of the National Academy of Sciences* **99**, 4197-4202.
- Sosnick, T. R., Mayne, L., and Englander, S. W. (1996). Molecular collapse: the rate-limiting step in two-state cytochrome c folding. *Proteins: Structure, Function, and Genetics* **24**, 413-426.
- Tanford, C. (1968). Protein denaturation. *Advances in Protein Chemistry* **23**, 121-282.
- Vaughn, D. E., Rodriguez, J., Lazebnik, Y., and Joshua-Tor, L. (1999). Crystal structure of apaf-1 caspase recruitment domain: an a-helical greek key fold for apoptotic signaling. *Journal of Molecular Biology* **293**, 439-447.
- Weber, C. H., and Vincenz, C. (2001). The death domain superfamily: a tale of two interfaces? *TRENDS in Biochemical Sciences* **26**, 475-481.
- Zhou, P., Chou, J., Olea, R. S., Yuan, J., and Wagner, G. (1999). Solution structure of apaf-1 CARD and its interaction with caspase-9 CARD: a structural basis for specific adaptor/caspase interaction. *Proceedings of the National Academy of Sciences* **96**, 11265-11270.
- Zou, H., Henzel, W. J., Liu, X., Lutschg, A., and Wang, X. (1997). Apaf-1, a human protein homologous to *C. elegans* CED-4, participates in cytochrome c-dependent activation of caspase-3. *Cell* **90**, 405-413.

TABLES AND FIGURES

Table 1. Summary of statistics for data collection and refinement for Apaf-1 CARD

Temperature (K)	100
Resolution (Å)	35–2.0 (2.00–2.07)
Space group	P2 ₁ 2 ₁ 2 ₁
Unit cell (Å, deg.)	
<i>a</i> (Å)	33.47
<i>b</i> (Å)	47.52
<i>c</i> (Å)	54.75
α, β, γ (deg.)	90
No. reflections	5981 (458)
Completeness (%)	96 (75)
I/σ	24.4 (4.8)
R _{merge} (%) ^a	10.6 (23.1)
Redundancy	5.6 (1.4)
R _{work} (%) ^b	23.9 (31.7)
R _{free} (%) ^c	29.7 (32.0)
Stereochemistry (r.m.s. deviation)	
Bond lengths (Å)	0.006
Bond angles (deg.)	1.14
Dihedral angles (deg.)	17.4
Improper (Å)	0.6
Average <i>B</i> factor (Å ²)	42.4
No. protein atoms	753
No. Zn ²⁺	6
No. water molecules	55

^a $R_{\text{merge}} = \frac{\sum_h \sum_i |I(h,i) - I(h)|}{\sum_h \sum_i I(h,i)}$, where $I(h,i)$ are symmetry-related intensities and $I(h)$ is the mean intensity of the reflection with unique index h .

^b $R_{\text{work}} = \frac{\sum ||F_{\text{obs}}| - |F_{\text{calc}}||}{\sum |F_{\text{obs}}|}$, where F_{obs} and F_{calc} are observed and calculated structure factors.

^c $R_{\text{free}} = \frac{\sum_T ||F_{\text{obs}}| - |F_{\text{calc}}||}{\sum_T |F_{\text{obs}}|}$, where T is a test data set of 10% of the total reflections randomly chosen and set aside prior to refinement.

Table 2. Summary of thermodynamic parameters for Apaf-1 CARD

Experiment	ΔG^{H_2O} (kcal mol ⁻¹) FL	ΔG^{H_2O} (kcal mol ⁻¹) CD	m (kcal mol ⁻¹ M ⁻¹) FL	m (kcal mol ⁻¹ M ⁻¹) CD	Urea _{1/2} (M) ^a FL	Urea _{1/2} (M) ^a CD
pH 6						
Equilibrium unfolding ^b	5.2±0.9	4.2±0.6	1.17±0.20	1.12±0.17	4.4	3.8
Refolding burst phase ^c	4.2±0.6	2.9±0.8	1.40±0.43	1.23±0.30	3.0	2.4
Refolding final signal ^c	5.1±1.4	4.9±0.4	1.10±0.28	1.18±0.11	4.6	4.2
Unfolding fast phase amplitude ^d	5.1±0.8		1.20±0.40		4.3	
Unfolding final signal ^e	4.9±0.9	5.1±0.2	1.20±0.21	1.31±0.54	4.1	3.9
Fast phase rates ^f	5.3		1.22		4.3	
pH 8						
Equilibrium unfolding ^b	4.8±0.8	3.1±0.5	1.38±0.23	1.11±0.16	3.5	2.8
Refolding burst phase ^c	3.6±0.3	2.0±0.5	1.40±0.12	1.00±0.30	2.3	2.0
Refolding final signal ^c	5.0±0.5	4.4±0.6	1.60±0.18	1.32±0.16	3.1	3.3
Unfolding fast phase amplitude ^d	4.7±0.01		1.35±0.02		3.5	
Unfolding final signal ^e	4.4±1.8	4.6±0.3	1.26±0.54	1.34±0.41	3.5	3.4
Fast phase rates ^f	5.1	5.1	1.32	1.33	3.9	3.8

^aCalculated as Urea_{1/2} = $\Delta G^{H_2O} / m$
^bFrom fits of data in Figure 2
^cFrom fits of data in Figure 3B or Supplemental Figures 1-3
^dFrom fits of data in Figure 3D or Supplemental Figure 3C
^eFrom fits of data in Figure 3E or Supplemental Figures 1-3
^fFrom fits of data in Figure 3F

Table 3. Summary of calculated and simulated rates using Scheme 3

	k_1^a (s ⁻¹)	k_3 (s ⁻¹)	k_{-3} (s ⁻¹)	k_4 (s ⁻¹)	k_{-4} (s ⁻¹)	k_5 (s ⁻¹)	k_{-5} (s ⁻¹)
<i>Refolding pH 6</i>							
(3.6 M) calculated rate ^b	Burst	0.6		0.6		41	8.4
(3.6 M) simulated rate	5000	0.6	0.01	0.6	0.01	37	13
(0.55 M) calculated rate	Burst	0.6		0.6		1800	0.63
(0.55 M) simulated rate	5000	0.6	0.01	0.6	0.01	1800	0.6
<i>Refolding pH 8</i>							
(3.2 M) calculated rate	Burst	0.6		0.6		37	6.6
(3.2 M) simulated rate	5000	0.6	0.01	0.6	0.01	35	8
(0.45 M) calculated rate	Burst	0.6		0.6		850	0.34
(0.45 M) simulated rate	5000	0.6	0.01	0.6	0.01	850	0.3
<i>Unfolding pH 6</i>							
(6 M) calculated rate						61	2.3
(6 M) simulated rate						70	
<i>Unfolding pH 8</i>							
(5 M) calculated rate						48	4.6
(5 M) simulated rate						55	

^a Rates correspond to those shown in Scheme 3. k_{-1} , k_2 , and k_{-2} are not shown since no information was obtained in the simulations for these rates.

^b Final urea concentration is indicated in parentheses for each experiment.

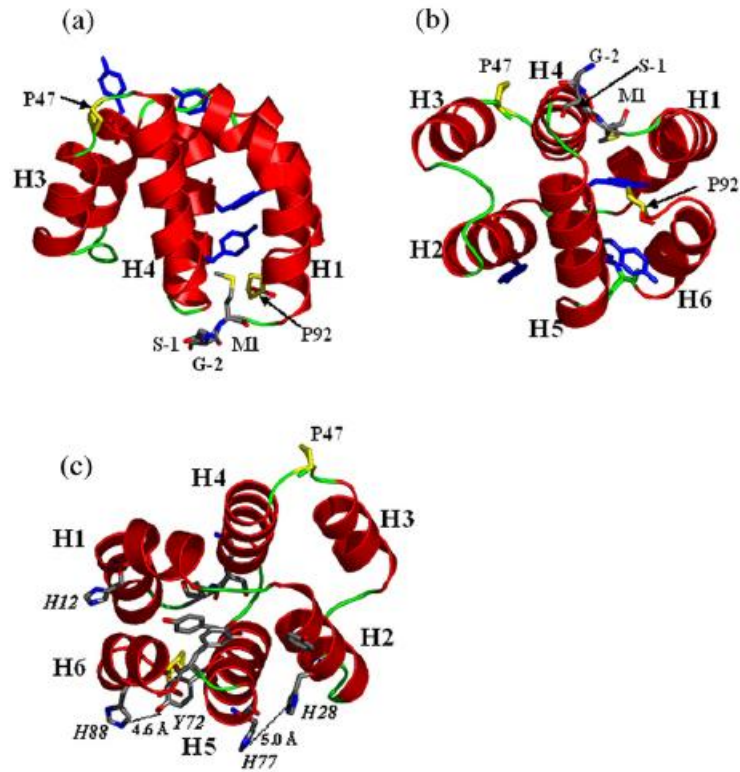


Figure 1. Structure of Apaf-1 CARD with GS sequence at the amino terminus. Side view (a) and top view (b) of Apaf-1 CARD. Helices 1 through 6 are labeled H1–H6. The four tyrosine residues are colored blue, and the two proline residues are colored yellow and are labeled. The first three residues (G-2, S-1, M1) are colored cpk and are labeled. (c) The four histidine residues of Apaf-1 CARD (H12, H28, H77, H88) are highlighted relative to their position in helices -1, -2, -5, and -6, respectively. The distances between H28 and H77 and between H88 and Y72 are shown. This Figure was generated with PyMol (Delano Scientific, San Carlos, CA).

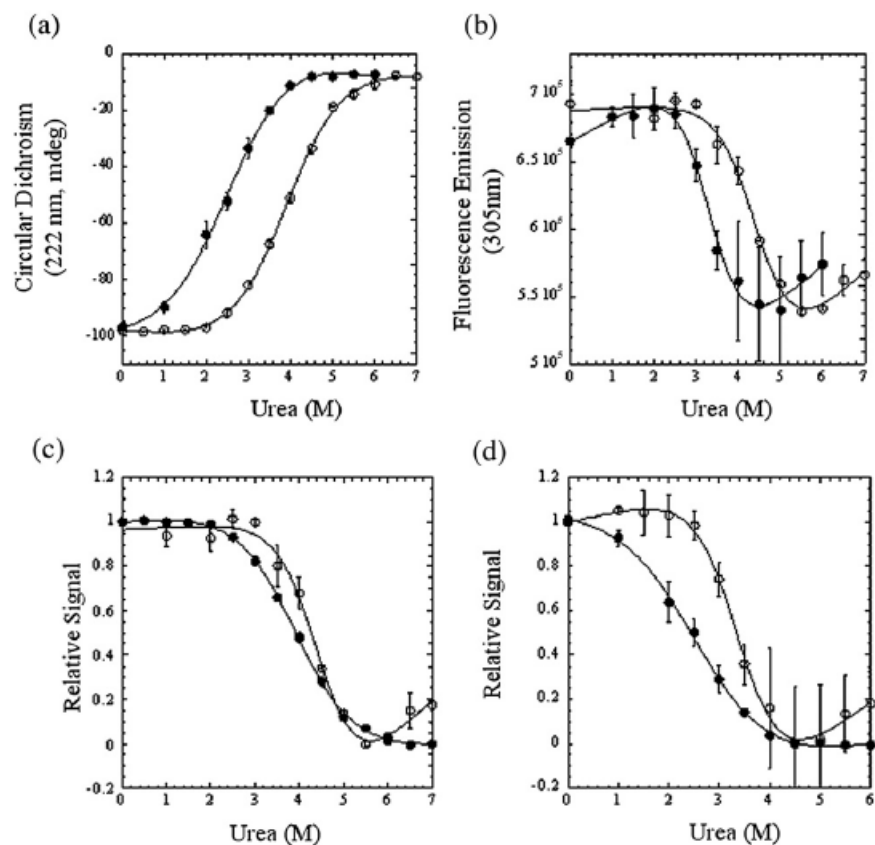
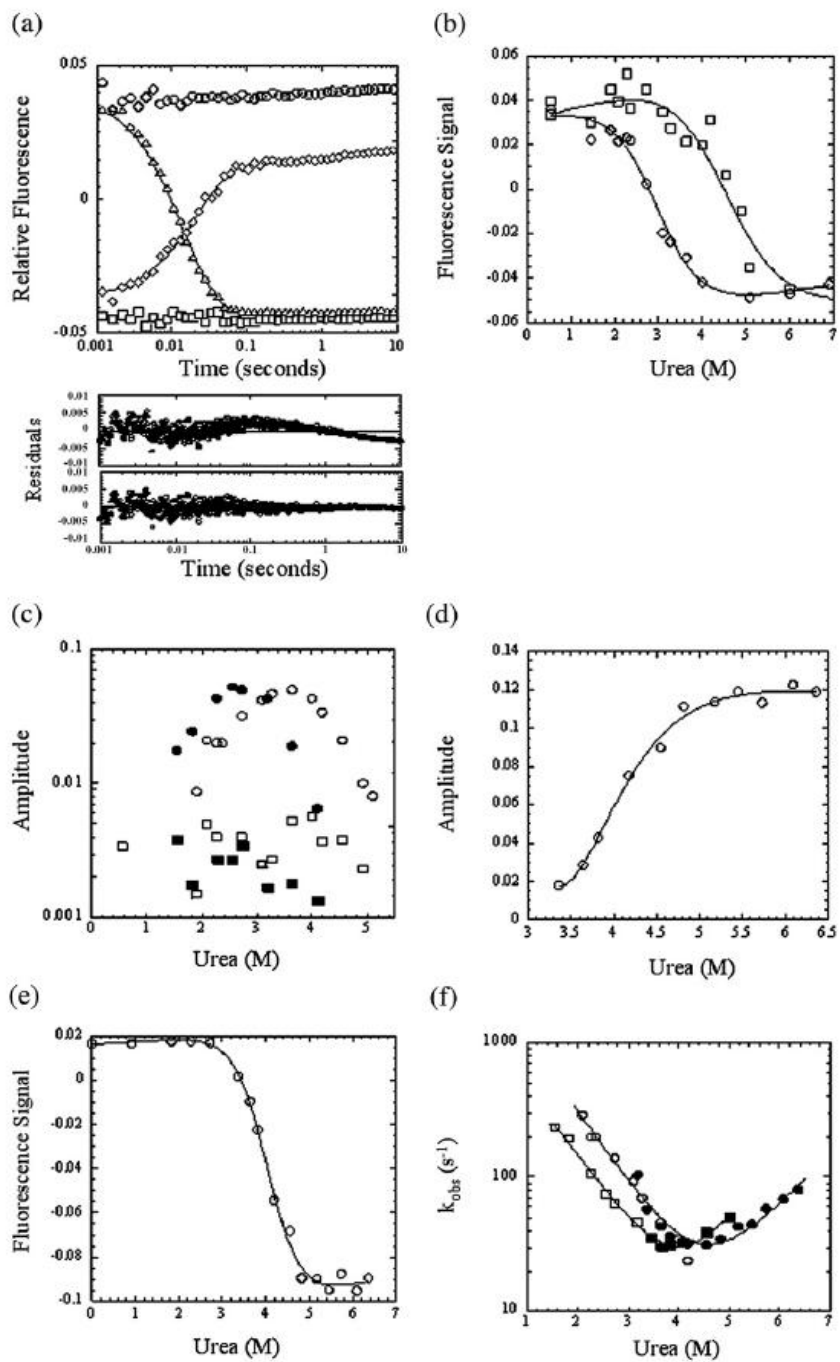


Figure 2. Equilibrium studies of Apaf-1 CARD. (a) Equilibrium unfolding of Apaf-1 CARD was measured by CD at 222 nm in urea-containing MTA buffer at pH 6 (\circ) or pH 8 (\bullet). Error bars show the standard deviation from three independent experiments. (b) Equilibrium unfolding of Apaf-1 CARD was measured by fluorescence emission at 305 nm with excitation at 280 nm in MTA buffer at pH 6 (\circ) or pH 8 (\bullet). Error bars show the standard deviation from three independent experiments. (c) and (d) Normalized equilibrium unfolding of Apaf-1 CARD was measured by CD at 222 nm (\bullet) or by fluorescence emission at 305 nm with excitation at 280 nm (\circ) in urea-containing MTA buffer at pH 6 (c) or pH 8 (d). Error bars show the standard deviation from three independent experiments. For (a)–(d), continuous lines represent fits of the data as described in the text.

Figure 3. Single-mixing stopped-flow fluorescence emission experiments. (a) Refolding of Apaf-1 CARD in 0.55 M(\circ) or 3.6 M (\diamond) urea-containing MTA buffer at pH 6. Unfolding of Apaf-1 CARD in 6 M urea-containing MTA buffer (\triangle). A control of Apaf-1 CARD in 6 M urea is shown (\square). Continuous lines represent fits to single (unfolding) or double (refolding) exponential equations. Parameters obtained from the fits are described in the text. The upper and lower panels are residuals of one or two exponential fits of 3.6 M refolding data (\diamond), respectively. (b) Burst phase (\circ) or final (\square) fluorescence emission signal versus urea concentration obtained from single-mixing stopped-flow refolding experiments. The final signal was measured at 10 s. (c) Amplitude of the fast (pH 6 (\circ) or pH 8 (\bullet)) and slow (pH 6 (\square) or pH 8 (\blacktriangleright)) phases versus urea concentration from single-mixing stopped-flow refolding experiments. Refolding data such as those in (a) were fit to a single or double exponential equation, as described in the text, and the amplitudes obtained from the fits are plotted versus urea concentration. (d) Single-mixing stopped-flow unfolding experiments were performed at the urea concentrations shown in the Figure, and the data were fit to a single exponential equation. The amplitude associated with the unfolding transition at each urea concentration is shown. (e) Final fluorescence emission signal of stopped-flow unfolding experiments. The final signals for the unfolding experiments shown in (d) were obtained after 10 seconds of unfolding. In (b), (d) and (e) the continuous lines represent fits of the data to a two-state equilibrium folding model. Parameters obtained from the fits are described in the text and are shown in Table 2. (f) Apparent rates of refolding (pH 6 (\circ) and pH 8 (\square)) and of unfolding (pH 6 (\bullet) and pH 8 (\blacktriangleright)) versus urea concentration. The apparent rates are associated with the refolding and unfolding data described in (c) and (d), respectively. The continuous lines represent fits to a two-state kinetic folding model. Parameters obtained from the fits are described in the text.



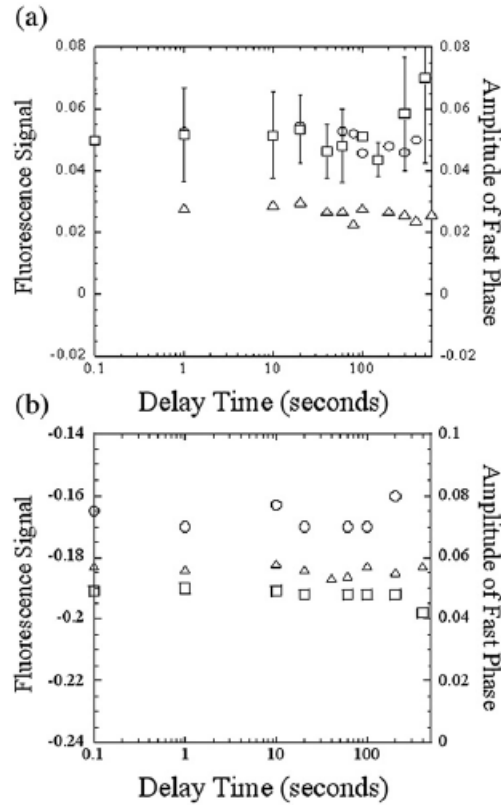


Figure 4. Sequential-mixing stopped-flow fluorescence studies. (a) Double-jump experiments were performed at final urea concentrations of 0.83 M (\circ) or 4 M (Δ) in MTA buffer at pH 6. The fluorescence emission signal of the burst phase (\circ) and the amplitude of the fast refolding phase (Δ) were determined versus unfolding delay time. Interrupted refolding experiments (\square) were performed at a final urea concentration of 5 M in MTA buffer at pH 6. The amplitude versus refolding delay time was determined from fits of the unfolding transition to a single exponential equation. The signals of the native protein in 0.83 M or in 4 M urea are 0.05 and 0.028, respectively. (b) Double-jump experiments were performed at final urea concentrations of 0.83 M (\circ) or 3.18 M (Δ) in MTA buffer at pH 8. The fluorescence emission signal of the burst phase (\circ) and the amplitude of the fast refolding phase (Δ) were determined versus unfolding delay time. Interrupted refolding experiments (\square) were performed at a final urea concentration of 5 M in MTA buffer at pH 8. The amplitude versus refolding delay time was determined from fits of the unfolding transition to a single exponential equation. The signals of the native protein in 0.83 M or in 3.18 M urea are -0.17 and -0.185 , respectively.

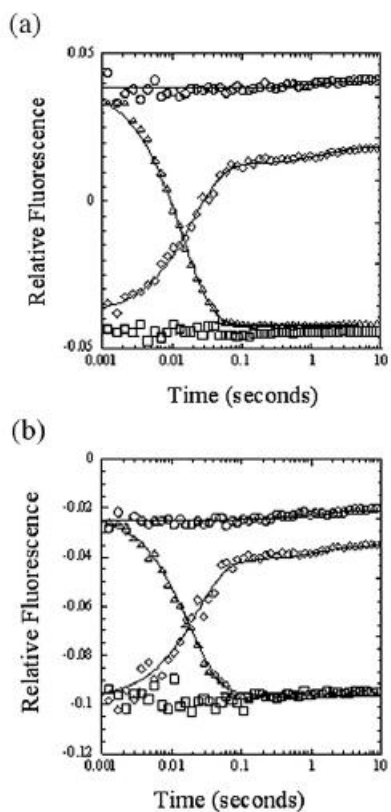
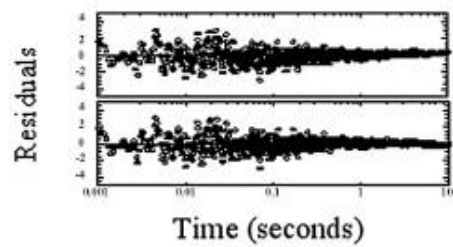
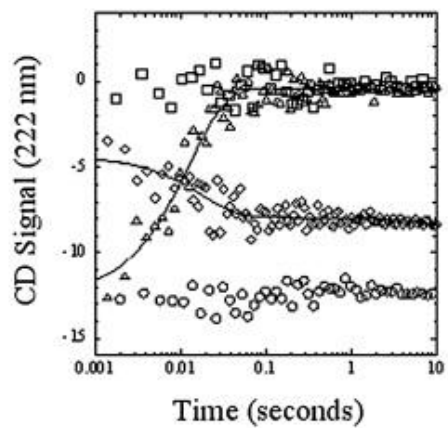


Figure 5. Simulations of single-mixing stopped-flow fluorescence emission experiments. (a) Refolding and unfolding of Apaf-1 CARD are as shown in Figure 3(a) (pH 6). (b) Refolding of Apaf-1 CARD in 0.45 M (\circ) or 3.2 M (\diamond) urea-containing MTA buffer at pH 8. Unfolding of Apaf-1 CARD in 5 M urea-containing MTA buffer (\triangle). A control of Apaf-1 CARD in 5 M urea is shown (\square). For (a) and (b), the data were simulated as described in the text using the model shown in Scheme 3. The rate constants used in the simulations are shown in Table 3, and the results are represented by the continuous lines.

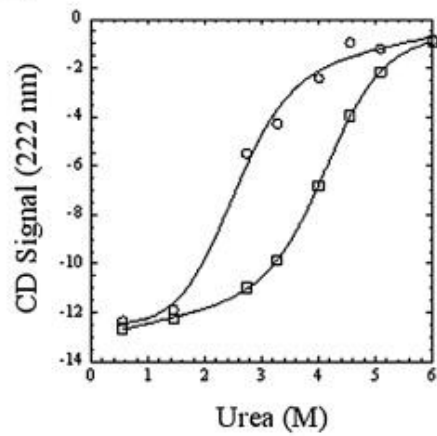
SUPPLEMENTAL FIGURES

Figure 1. Single-mixing stopped-flow circular dichroism experiments in MTA buffer pH 6. (A) Refolding of Apaf-1 CARD in 0.55 M (\circ) or 3.27 M (\diamond) urea-containing MTA buffer. Unfolding of Apaf-1 CARD in 6 M urea-containing MTA buffer (Δ). A control of Apaf-1 CARD in 6 M urea is shown (\square). Solid lines represent fits to single (unfolding) or double (refolding) exponential equations. Parameters obtained from the fits are described in the text. The upper and lower panels are residuals of one or two exponential fits of the 3.27 M refolding data (\diamond), respectively. (B) Burst phase (\circ) or final (\square) CD signal *versus* urea concentration obtained from single-mixing stopped-flow refolding experiments. The final signal was measured at 10 seconds. (C) Final CD signal of stopped-flow unfolding experiments. The final signals for the unfolding experiments shown in Panel C were obtained after 10 seconds of unfolding. In Panels B and C, the solid lines represent fits of the data to a two-state equilibrium folding model. Parameters obtained from the fits are described in the text and are shown in Table 2.

(a)



(b)



(c)

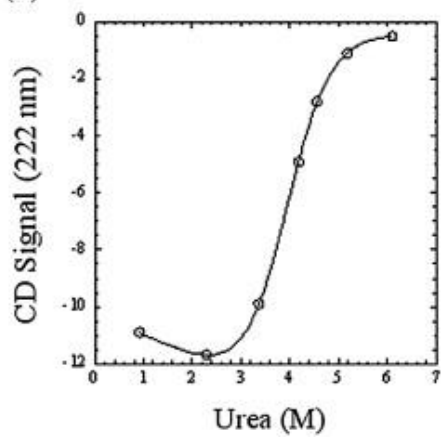


Figure 2. Single-mixing stopped-flow circular dichroism experiments in MTA buffer pH 8. (A) Refolding of Apaf-1 CARD in 0.45 M (\circ) or 2.73 M (\diamond) urea-containing MTA buffer. Unfolding of Apaf-1 CARD in 5 M urea-containing MTA buffer (Δ). A control of Apaf-1 CARD in 5 M urea is shown (\square). Solid lines represent fits to single (unfolding) or double (refolding) exponential equations. Parameters obtained from the fits are described in the text. The upper and lower panels are residuals of one or two exponential fits of 2.73 M refolding (\diamond), respectively. (B) Burst phase (\circ) or final (\square) CD signal *versus* urea concentration obtained from single-mixing stopped-flow refolding experiments. The final signal was measured at 10 seconds. (C) Final CD signal of stopped-flow unfolding experiments. The final signals for the unfolding experiments shown in Panel C were obtained after 10 seconds of unfolding. In Panels B and C, the solid lines represent fits of the data to a two-state equilibrium folding model. Parameters obtained from the fits are described in the text and are shown in Table 2.

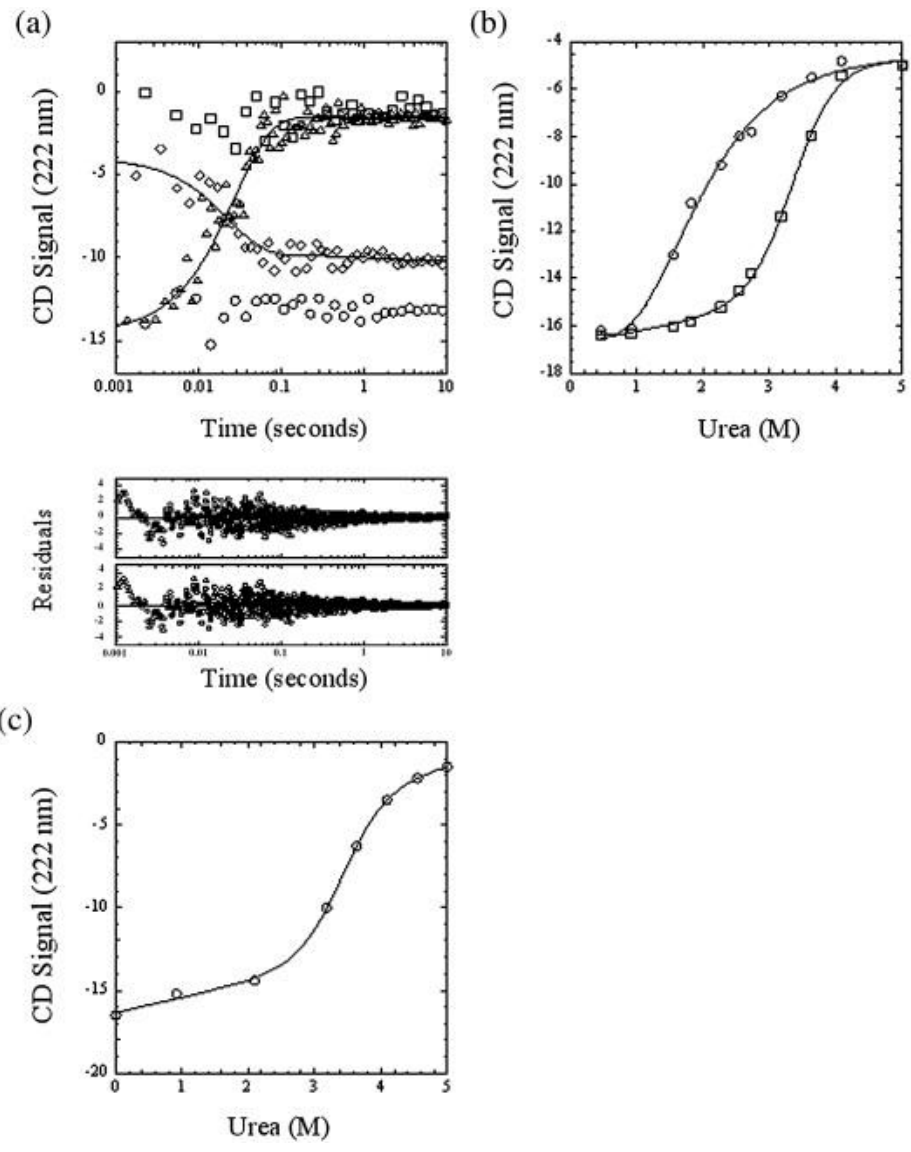
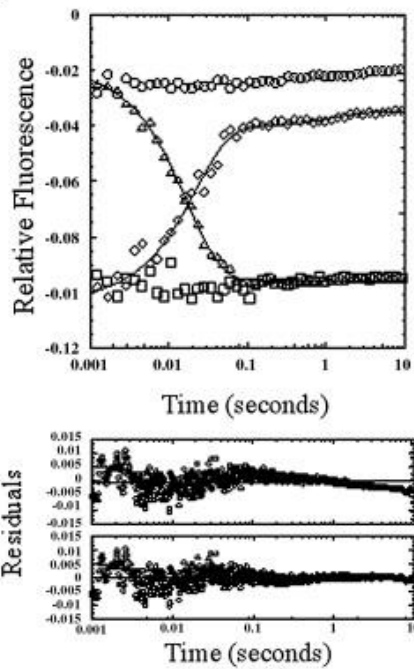
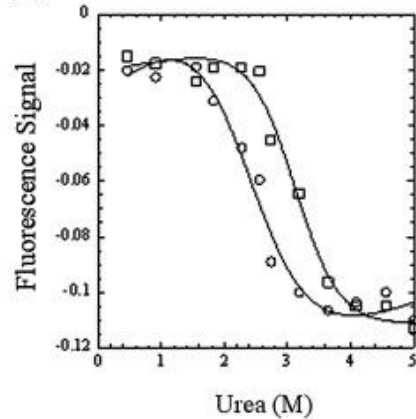


Figure 3. Single-mixing stopped-flow fluorescence emission experiments in MTA buffer pH 8. (A) Refolding of Apaf-1 CARD in 0.45 M (\circ) or 3.18 M (\diamond) urea-containing MTA buffer. Unfolding of Apaf-1 CARD in 5 M urea-containing MTA buffer (Δ). A control of Apaf-1 CARD in 5 M urea is shown (\square). Solid lines represent fits to single (unfolding) or double (refolding) exponential equations. Parameters obtained from the fits are described in the text. The upper and lower panels are residuals of one or two exponential fits of 3.18 M refolding data (\diamond), respectively. (B) Burst phase (\circ) or final (\square) fluorescence emission signal *versus* urea concentration obtained from single-mixing stopped-flow refolding experiments. The final signal was measured at 10 seconds. (C) Single-mixing stopped-flow unfolding experiments were performed at the urea concentrations shown in the figure, and the data were fit to a single exponential equation. The amplitude associated with the unfolding transition at each urea concentration is shown. (D) Final fluorescence emission signal of stopped-flow unfolding experiments. The final signals for the unfolding experiments shown in Panel C were obtained after 10 seconds of unfolding. In Panels B, C and D, the solid lines represent fits of the data to a two-state equilibrium folding model. Parameters obtained from the fits are described in the text and are shown in Table 2.

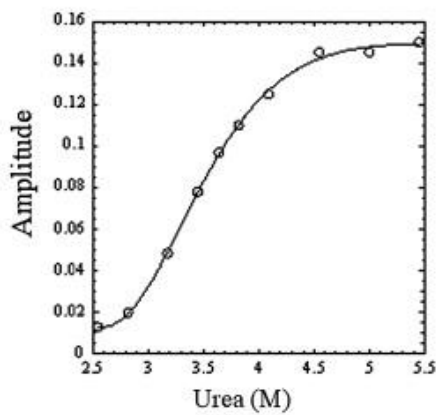
(a)



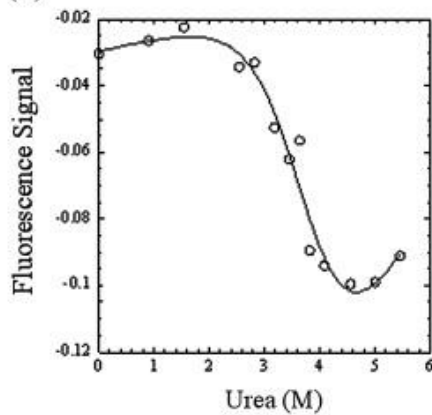
(b)



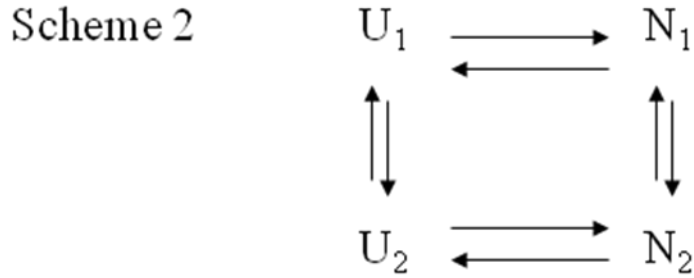
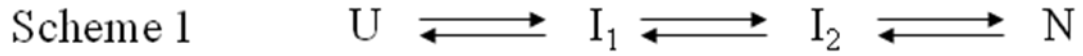
(c)



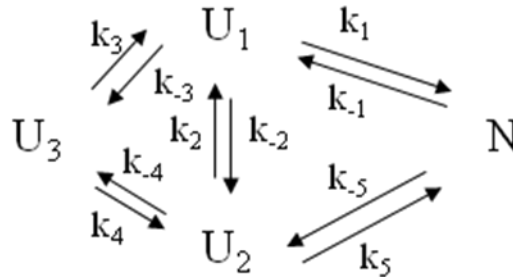
(d)



SCHEMES



Scheme 3



Scheme 1. Sequential folding model in which the unfolded protein, U, folds through two non-native intermediates, I₁ and I₂, prior to forming the native conformation, N.

Scheme 2. Parallel folding model in which the unfolded ensemble contains at least two conformations, U₁ and U₂. Both unfolded conformations fold to native conformations, N₁ and N₂, respectively. In this scheme, both the unfolded and the native conformations interconvert.

Scheme 3. Parallel folding model in which the unfolded ensemble contains at least three conformations, U₁, U₂, and U₃. Unfolded conformations U₁ and U₂ fold to a single native conformation, N. Conversion of U₃ either to U₁ or U₂, or both, represent slow isomerizations observed during refolding. Rate constants determined from kinetic experiments as well as those estimated from simulations of the data are shown in Table 3.

CHAPTER IV

Folding and Assembly Kinetics of Procaspase-3

ABSTRACT

Caspases are vital to the apoptotic pathway and exist in the cell as inactive zymogens (termed procaspase). We have examined the kinetic folding of the homodimeric procaspase-3 with fluorescence, circular dichroism, differential quenching by acrylamide, anisotropy and activity assays. Single-mixing refolding stopped-flow refolding studies showed a complex burst phase, in which the protein forms multiple monomeric species within milliseconds. The data also revealed several other monomeric intermediates, some of which appear to be off-pathway or misfolded, that eventually form a dimerization-competent species. Return of activity studies demonstrated slow rate of dimerization ($\sim 70 \text{ M}^{-1}\text{s}^{-1}$). In addition, single-mixing unfolding stopped-flow studies reveal a complex pathway with a slow rate of dimer dissociation. Interestingly, multiple dimeric species were observed in the unfolding burst phase, which support the theory of two native conformations of the protein. Collectively, these results demonstrate that dimerization is an important aspect in both folding and activation of procaspase-3.

INTRODUCTION

The “protein folding problem” is to understand how the polypeptide chain encodes the native structure of the protein. The majority of kinetic folding studies have focused on small (<100 amino acid) monomeric proteins. These studies are advantageous as all observed phases are first order processes. However, monomeric proteins are the minority, representing ~19% of the proteins in *E. coli*, while ~38% of the proteins are dimeric (Goodsell and Olson, 2000). A lot is unknown about the kinetic folding and assembly of dimeric proteins. Data analysis of dimeric proteins is more complicated due to the fact that second order processes are involved. Similar to monomeric proteins, dimeric proteins fold through a diverse array of folding models ranging from simple two-state transitions to complex parallel pathways with monomeric and/or dimeric intermediates (Rumfeldt et al., 2008). For example, the folding of a small dimeric P22 arc repressor is best described by a two-state reaction with a dimerization rate of $\sim 10^7 \text{ M}^{-1}\text{s}^{-1}$ (Milla and Sauer, 1994). The large dimeric SecA has also been shown to dimerize near the diffusion limit ($\sim 10^9 \text{ M}^{-1}\text{s}^{-1}$), as well as fold through parallel channels with sequential intermediates (Doyle et al., 2004).

To complicate these pathways further, monomeric or dimeric intermediates may exist in the folding pathway. Intermediates have been shown to either slow (off-pathway) or accelerate (on-pathway) the rate of forming the native ensemble (Baldwin, 1996). Chevron analysis allows for the detection of intermediates in the folding pathway. A rollover (non-linearity) in a chevron plot points to the presence of an intermediate and is interpreted as a

change in the position of the rate-limiting transition state (Brockwell and Radford, 2007). Intermediates may also be detected in the burst phase.

Cooperative burst phase transitions are associated with structural changes in the protein during the first few milliseconds of the folding reaction. These data are typically best described by a two-state equilibrium folding model, which reveal the thermodynamic parameters associated with the transition. By comparing these values to those determined from equilibrium experiments, additional information about the burst phase species is acquired. For example, folding studies of monomeric Apaf-1 CARD revealed a cooperative burst phase that was best described to a two-state equilibrium model. The free energy and m -values obtained were similar to equilibrium values, indicating that the native species was formed in the burst phase (Milam et al., 2007). Two-state transitions do not always indicate the formation of a monomeric species. Topping *et al.* has observed the formation of a dimeric intermediate in the burst phase for the folding of FIS (Topping et al., 2004). In addition, it is best to use multiple probes to study protein folding because each detection method provides different information about the structures formed during the folding process. Mann and Matthews have shown that dimeric trp aporepressor folds through a monomeric intermediate. The coincidence of the burst phase transitions monitored by far-UV CD and ANS spectroscopy revealed that this intermediate consists of a significant amount of secondary structure and hydrophobic surface (Mann and Matthews, 1993). If the burst phase data monitored by multiple probes reveal non-coincident transitions, it suggests that at least two intermediates exist during folding (Zaida et al., 1997). Bhuyan and Udgaonkar have shown

that the unfolding burst phase of horse cytochrome c in the oxidative state monitored by far-UV CD and fluorescence quenching reveal non-coincident transitions, indicating the accumulation of two kinetic intermediates during unfolding (Bhuyan and Udgaonkar, 1998).

Here we discuss the kinetic folding pathway of homodimeric procaspase-3. Equilibrium unfolding studies reveal a four-state equilibrium folding process (Bose and Clark, 2001). To further investigate the folding pathway of this protein, kinetic folding studies were performed. Procaspase-3 exists in the cell as an inactive zymogen, which can be activated through the intrinsic or extrinsic apoptotic pathway (Pop et al., 2001; Taylor et al., 2008). Once activated, (pro)caspase-3 is able to cleave a variety of substrates, leading to the dismantling of the cell (Luthi and Martin, 2007; Saikumar et al., 1999; Taylor et al., 2008). Homodimeric procaspase-3 consists of an N-terminal prodomain, large subunit, intersubunit linker and a small subunit with ten tyrosine and two tryptophan residues per monomer. The tryptophans are both located in the active site, while the tyrosines are distributed throughout the protein.

For these kinetic folding studies fluorescence and circular dichroism spectroscopies were used to determine the folding mechanism of procaspase-3. These methods were unable to detect the dimerization step in the folding pathway. For this reason other techniques were employed, such as anisotropy, differential quenching by acrylamide and activity assays. Refolding kinetic studies with procaspase-3 reveal a complex burst phase, multiple monomeric intermediates and slow dimerization. Unfolding kinetic studies with procaspase-3 reveal multiple native species and slow dimer dissociation. In addition, the unfolding results

support the hypothesis of two native conformations of the protein.

RESULTS

Single mixing refolding and unfolding studies were performed using a variety of detection methods in order to elucidate the folding mechanism of procaspase-3. Fluorescence and circular dichroism (CD) studies were first performed to examine the refolding of procaspase-3. Representative CD data are shown in Figure 1. The CD data displayed 4 phases: burst, lag, intermediate and slow. The fluorescence data (excitation at 280 nm) displayed only 2 phases: burst and fast. All of these phases monitored the folding of monomeric species of the protein (see below). Therefore, protein dimerization might only slightly change the secondary or tertiary characteristics of the protein, which would account for a small change in CD or fluorescence signal, respectively. As a result, other detection methods, such as anisotropy, differential quenching by acrylamide and return of activity studies, were used to examine the rate of dimerization of procaspase-3.

The rest of this chapter will be presented based on different time frames in the refolding and unfolding reactions. Therefore, each detection method is briefly described here.

- Fluorescence (excitation at 280 nm): examines the global or tertiary structure of the protein due to excitation of tryptophans and tyrosines
- Fluorescence (excitation at 295 nm): examines the formation of the active site due to selective excitation of tryptophans
- CD (far-UV CD; 228 nm): examines the secondary structure of the protein

- Anisotropy (excitation at 280 or 295 nm): examines the rotation or tumbling of the protein
- Differential quenching by acrylamide (excitation at 295 nm): provides more specific information about the orientation of tryptophan residues in the protein
- Return of activity studies: provides information on when the dimer is formed since only the dimer is enzymatically active

The kinetic studies presented here were performed with a catalytically inactive mutant of procaspase-3 in which the active site cysteine has been mutated to a serine (C163S). In the refolding studies, unfolded protein was mixed with various urea-containing buffer solutions, which allows for the characterization of intermediates in the folding pathway. The data for each detection method revealed multiple phases, which indicated the possibility of multiple species formed during the folding pathway of procaspase-3. Further analysis of each phase revealed a complex burst phase, multiple monomeric intermediates and a slow dimerization step, as described below.

A. Multiple monomeric species are formed in the burst phase

The protein was unfolded in 8 M urea-containing buffer, followed by rapid dilution of the urea, which allowed the protein to refold. The burst phase signal was collected at multiple final urea concentrations using multiple detection methods, such as fluorescence emission (excitation at 280 or 295 nm), far-UV CD (228 nm), anisotropy (excitation at 280 or 295 nm) and differential quenching by acrylamide (excitation at 295 nm). The burst phase amplitudes

were plotted versus the final urea concentration, shown in Figure 2A.

Each detection method was plotted separately for clarity in Figures 2B-E. The data displayed a cooperative folding transition, indicating that the species formed in the burst phase is associated with structural changes in the protein. The error bars represent the standard deviation from at least two independent experiments and multiple protein concentrations (ranging from 2-15 μM). These data were not protein concentration dependent, demonstrating that the species formed in the burst phase is monomeric. Furthermore, the burst phase studies with different detection methods reveal that the data were non-coincident indicating the presence of multiple species. CD, anisotropy and differential quenching by acrylamide burst phase signals were best described by a three-state equilibrium folding model, whereas the fluorescence emission burst phase signal was best described by a two-state equilibrium folding model. The conformational free energies, ΔG^{H_2O} , and m-values are presented in Table I. The ΔG^{H_2O} values for the first folding transition from CD, anisotropy and differential quenching by acrylamide were similar ~ 1.3 , 0.8 and 1.4 kcal mol⁻¹, respectively. The ΔG^{H_2O} values for the second folding transition from CD, anisotropy, differential quenching by acrylamide and fluorescence emission were also similar ~ 3.1 , 3.1 , 3 and 3.2 kcal mol⁻¹, respectively. The stability of the monomer, as determined from equilibrium folding experiments is ~ 8 kcal mol⁻¹. Therefore, the species formed in the burst phase has half the stability of the monomer ($\Delta G^{H_2O} \sim 4.3$ kcal mol⁻¹).

In order to investigate the species formed in the burst phase, emission scanning

experiments were performed. The refolding burst phase signal detected by fluorescence emission was monitored at multiple emission wavelengths with excitation at either 280 (Fig 3A) or 295 nm (Fig 3B). The emission profiles of the native protein (open circles) had maxima at 337 nm and 341 nm with excitation at 280 or 295 nm, respectively. The emission profiles of the unfolded protein (open squares) had maxima at 342 nm (280 nm) and 345 nm (295 nm). These data were similar to measurements obtained on a steady-state instrument (Bose and Clark, 2001). The emission profile of the refolding burst phase (closed circles) was different from the native and unfolded protein signals with a maxima at 335 nm (280 nm) or 338 nm (295 nm). Thus, within 1-2 milliseconds, the refolding protein signal was blue-shifted compared to native and unfolded protein signals, indicating that the aromatic residues are buried in this species.

B. Procaspase-3 folds through multiple monomeric intermediates

As mentioned above, refolding studies reveal multiple kinetic phases from milliseconds to 100 seconds. The CD data (228 nm) revealed four phases; burst, lag, intermediate and slow (Figure 1). The burst phase data were described above. The lag, intermediate and slow phases for CD will be described in this section. The length of the lag time at different final urea concentration and protein concentrations was determined (Figure 4A). The lag time increased with increasing urea concentration for all protein concentrations investigated (ranging from 2-12 μ M), indicating that this lag was not due to dimerization. In order to investigate the lag in the CD data, other techniques were employed, such as fluorescence emission and differential quenching by acrylamide.

Shown in Figures 4B-D are examples of raw refolding data for fluorescence emission with excitation at 280 (B) or 295 (C) nm and differential quenching by acrylamide (D). These results show that multiple species formed during the lag phase observed by far-UV CD, and the data were best fit to three exponentials as shown by the residuals below each plot. The apparent rates of the three phases are shown in Figure 5A-C. For fluorescence emission with excitation at 280 or 295 nm and differential quenching by acrylamide, two of the apparent rates are independent of the final urea concentration (~ 40 and 0.2 sec^{-1}). In addition, none of the three rates were dependent upon the protein concentration, which indicates that each species is a monomer.

The second phase shown in Figure 5A-C is dependent on the final urea concentration, and the apparent rates generated a chevron plot (Figure 5A-C). The data were fit to a two-state kinetic mechanism (Ferguson et al., 1999; Jackson and Fersht, 1991) using Equation 1:

$$k_{obs} = k_{UN} \exp(-m_{U-TS} \times [urea] / RT) + k_{NU} \exp(m_{N-TS} \times [urea] / RT) \quad (1)$$

where k_{obs} is the observed rate constant of the refolding phase, and k_{UN} and k_{NU} are the rate constants for refolding and unfolding, respectively, in the absence of urea. The parameters m_{U-TS} and m_{N-TS} reflect the change in solvent exposed surface area between the unfolded or refolded states, respectively, and the transition state ensemble. For differential quenching by acrylamide (Figure 5C), values of $4.4 \pm 0.90 \text{ sec}^{-1}$ and $0.50 \pm 0.24 \text{ kcal mol}^{-1} \text{ M}^{-1}$ were obtained for k_{NU} and m_{N-TS} , respectively, while values of $0.78 \pm 0.43 \text{ sec}^{-1}$ and $0.23 \pm 0.074 \text{ kcal mol}^{-1} \text{ M}^{-1}$ were obtained for k_{UN} and m_{U-TS} , respectively. The conformational free energy

and total m-value were calculated from these parameters, where $\Delta G^{H_2O} = -RT \ln (k_{UN}/k_{NU}) = 1.02 \text{ kcal mol}^{-1}$. The m-value was calculated as $m_{\text{total}} = m_{U-TS} + m_{N-TS} = 0.73 \text{ kcal mol}^{-1} \text{ M}^{-1}$. Overall, this low conformational energy correlates with the lack of change in secondary structure of the protein detected by CD.

On the other hand, the apparent rates for fluorescence emission with excitation at 280 or 295 nm (Figure 5A and B) displayed a rollover on the unfolding limb of the chevron plot. The data were fit to a three-state kinetic mechanism (Ferguson et al., 1999; Parker et al., 1995) using Equation 2:

$$k_{obs} = k_{UI} \exp(-m_{U-TS} \times [urea] / RT) + \frac{k_{IU} \exp(m_{I-TS} \times [urea] / RT)}{1 + \frac{1}{K_{IN} \exp(-m_{IN} \times [urea] / RT)}} \quad (2)$$

where k_{obs} is the observed rate constant of the refolding phase, k_{UI} and k_{IU} are the rate constants for refolding and unfolding between the unfolded and intermediate states, respectively, in the absence of urea. The parameters m_{U-TS} and m_{I-TS} reflect the change in solvent exposed surface area between the unfolded and intermediate states, respectively, and the transition state ensemble. In this analysis, it is assumed that the intermediate is a prerequisite to achieving the native state and that a pre-equilibrium exists for formation of the intermediate (defined as $K_{IN} = k_{IN}/k_{NI}$), which occurs much faster than the rate-limiting step to native state formation. Using fluorescence emission with excitation at 295 nm as an example (Figure 5B), values of $6.7 \pm 0.68 \text{ sec}^{-1}$ and $0.61 \pm 0.082 \text{ kcal mol}^{-1} \text{ M}^{-1}$ were obtained for k_{UI} and m_{U-TS} , respectively, while values of $0.016 \pm 0.013 \text{ sec}^{-1}$ and 1.3 ± 0.18

kcal mol⁻¹ M⁻¹ were obtained for k_{IU} and m_{I-TS} , respectively. Values of 208 ± 128 and 1.2 ± 0.16 kcal mol⁻¹ M⁻¹ were obtained for K_{IN} and m_{IN} , respectively. The conformational free energy and total m-value were calculated from these parameters, where $\Delta G^{H_2O} = 1.6$ kcal mol⁻¹ and 3.2 kcal mol⁻¹ and $m_{total} = 1.9$ kcal mol⁻¹ M⁻¹ and 1.2 kcal mol⁻¹ M⁻¹ for each transition. The values from the first transition were similar to those calculated from the data obtained by differential quenching by acrylamide, indicating that this transition monitors formation of the active site since the tryptophan residues are located in the active site. The rollover, which was not detected in the differential quenching by acrylamide studies, represents a change in the rate-limiting step at high urea concentrations.

The next phase (termed the intermediate phase) observed in the CD data had a rate of 0.01 sec⁻¹ at low urea concentrations. This rate decreased with increasing urea as shown by the open triangles in Figure 6C. There was no change in the fluorescence emission data with excitation at 280 nm after 10 seconds (data not shown). However, both fluorescence emission with excitation at 295 nm (Figure 6A) and differential quenching by acrylamide (Figure 6B) revealed another, slower, phase. These data were best fit to a three exponential equation as shown by the residuals below each plot. The first two phases were the same as those described by the 10 second data. The rates of the intermediate phase versus the final urea concentration are shown in Figure 6C. The rate of this phase in low urea as detected by fluorescence emission (295 nm) and differential quenching with acrylamide was 0.015 sec⁻¹. None of these phases were dependent upon the protein concentration.

The last refolding phase detected by CD is the slow phase (Figure 7A). For these experiments, unfolded protein was hand-mixed with buffer, and the data were collected on a steady-state instrument. The CD data revealed a halftime of about 25 minutes, which was independent of the protein concentration, and the data were fit to a single exponential. The rate was also independent of the final urea concentration up to 4 M urea (Supplementary Figure 1). Hand-mixing differential quenching by acrylamide studies were also performed (Figure 7B). These studies revealed two additional phases, which were not protein concentration dependent. In the first phase, the fluorescence emission signal increased, and the halftime for the phase was about 8 minutes. These data indicate that the active site tryptophans are in a more hydrophobic environment. The fluorescence emission signal decreased during the second phase, and the halftime for the phase was about 25 minutes, indicating exposure of the active site tryptophans to solvent.

It is important to note that all of the phases described in this section were not dependent upon the protein concentration, indicating that all of these species are monomeric. Five out of the six of these phases detected here are independent of the final urea concentration, indicating that there are small changes in the accessible surface area in these transitions. The phase that was dependent on the final urea concentration (Figure 5) revealed a low free energy associated with the conformational transition. This phase was detected by fluorescence emission and not CD, indicating that there are small structural changes in the protein which only affect the tertiary structure and not the secondary structure of the protein.

C. Dimerization occurs slowly in the folding of procaspase-3

In order to verify that the native protein was formed, activity studies were performed. The aforementioned refolding studies were carried out with a catalytically inactive procaspase-3 (C163S). In order to test for activity but still examine procaspase-3, a catalytically active mutant was used in which the three processing sites were mutated to an alanine (D9A, D28A, D175A). The above kinetic studies were also performed with this mutant to verify that these mutations had no effect on the folding pathway (see Appendix). Due to improper formation of the active sites, the monomeric form of procaspase-3 is not active, while the dimeric form is active. Therefore, one will detect activity only once the native dimer is formed. In this assay, unfolded protein in 8 M urea-containing buffer was manually mixed with buffer in a 1:10 ratio to allow the protein to refold. Then activity was measured in aliquots of the refolding protein using a fluorogenic substrate. These experiments were performed at a 10-fold range of final protein concentration (0.5 – 5 μM procaspase-3 during refolding).

The return of activity data displayed two phases (Figure 8A). The first phase was a lag of about 10 minutes in which there is no detectable activity, which was observable at all protein concentrations examined. The lag phase is consistent with folding of the monomer described above. The second phase was best fit to a single exponential. This phase also was dependent upon the protein concentration. The apparent rates of this phase were plotted against the protein concentration to create a second order rate plot (Figure 8B). A linear fit to the data gave a second order rate of $\sim 70 \text{ M}^{-1}\text{s}^{-1}$.

KINSIM simulations were also performed to support this conclusion (Barshop et al., 1983). The mechanism used in the simulations consisted of two steps; a protein conformational change followed by protein dimerization. The conformational change represented the transition of the monomeric intermediate to the dimerization-competent monomeric intermediate, both of which are enzymatically inactive. The rate of this phase as determined from hand-mixing CD and differential quenching by acrylamide studies was 0.0005 and 0.002 sec⁻¹, respectively. The value determined from KINSIM simulations for this lag / conformational change was in good agreement with the experimental values, 0.00071 sec⁻¹. In addition, the rate of dimerization determined from KINSIM simulations, 115 M⁻¹s⁻¹, was similar to the rate obtained from activity studies. These results support the conclusion that the rate of procaspase-3 dimerization is slow.

D. Proposed folding models for procaspase-3

The simplest kinetic folding model to describe the single-mixing kinetic refolding data is shown in Figure 9. In this sequential mechanism, the unfolded protein forms multiple monomeric species in the burst phase, followed by a series of monomeric intermediates. At longer times, the protein forms a dimerization competent species, which then slowly dimerizes. The detection method(s) used to determine each step in the folding reaction are listed. This mechanism is most likely not the accurate; however it does describe the refolding kinetic data. Additional data, described below, suggest the presence of misfolded intermediates and parallel pathways.

E. Unfolding data reveal multiple species

Single mixing unfolding studies were performed using fluorescence emission (excitation at 280 or 295 nm), CD and anisotropy (excitation at 280 and 295 nm) as detection methods. Representative data are shown in Figure 10. In these studies, native protein was mixed 1:10 with various urea-containing buffer solutions. The observed phases and the corresponding detection method are listed in Table II and briefly described here.

Fluorescence emission data (excitation at 280 or 295 nm) displayed five phases (Figure 10A and B): burst, fast, lag and two slow phases. Anisotropy data (excitation at 280 or 295 nm) did not display a burst phase but did display a lag phase, followed by a decrease in anisotropy toward the unfolded protein signal (Figure 10C). The second phase was fit to two exponentials. Overall, the unfolding reaction was slow (green, ●, Figure 10C), reaching the unfolded protein signal by ~500 seconds, suggesting that the protein is monomeric by 8 minutes. Far-UV CD (228 nm) data also did not contain a burst phase but rather displayed a lag phase followed by an increase in signal toward the unfolded protein signal (Figure 10D). The latter phase was best described by a single exponential. The CD signal during unfolding (green, ●, Figure 10D) reached that of the unfolded protein by ~200 seconds, suggesting that the protein loses its secondary structure by ~3 minutes.

As mentioned above, the anisotropy and CD data did not display a burst phase (Figure 10). On the other hand both fluorescence emission data sets displayed an increase in fluorescence emission at higher urea concentrations. A plot of the burst phase amplitude versus final urea concentration demonstrated a cooperative burst phase transition in which a

hyperfluorescent species formed within the mixing dead-time. The data were best described by a three-state equilibrium folding model (Figure 11), indicating that at least two species formed in the burst phase for unfolding that, as described below, represents the formation of multiple dimeric forms of the protein. The ΔG^{H_2O} and m-value for the first transition are 2.45 ± 0.55 kcal/mol and 0.70 ± 0.21 kcal/mol/M, respectively. The ΔG^{H_2O} and m-value for the second transition are 11.7 ± 3.33 kcal/mol and 2.00 ± 0.54 kcal/mol/M, respectively.

The fluorescence emission data also displayed a fast phase with an apparent rate of $\sim 50 \text{ sec}^{-1}$ (Figure 12A). When examined as a function of final urea concentration, the data show that this phase aligns with the fastest phase detected in the single-mixing refolding data (Figure 5A and B), indicating that these phases describe the interconversion of the same species. This fast phase did not exist in the anisotropy and CD data where there was no change in signal for the first 5 seconds of unfolding. We believe that procaspase-3 is still a dimer at this stage because both the anisotropy and CD unfolding signals are close to the native dimer protein signal (Figure 10C and D).

After the lag phase, the far-UV CD data revealed an increase in signal, indicating the loss of secondary structure, that was best fit to a single exponential. Anisotropy (excitation at 280 or 295 nm) and fluorescence emission (excitation at 280 or 295 nm) data were best fit to a two exponential equation. A representative plot of these apparent rates versus the final urea concentration is shown in Figure 12B (Supplementary Figure 2). The first exponential observed in the fluorescence emission and anisotropy data is associated with that observed by CD, indicating a concomitant loss of tertiary and secondary structure. The slow phase

observed by fluorescence emission and anisotropy was dependent on the final urea concentration and displayed a chevron shape. As described above, these data were fit to a two-state kinetic mechanism. For fluorescence emission, values of $0.0050 \pm 0.0020 \text{ sec}^{-1}$ and $0.11 \pm 0.075 \text{ kcal mol}^{-1} \text{ M}^{-1}$ were obtained for k_{NU} and $m_{\text{N-TS}}$, respectively, while values of $6.7 \times 10^{-6} \pm 1.3 \times 10^{-5} \text{ sec}^{-1}$ and $0.50 \pm 0.13 \text{ kcal mol}^{-1} \text{ M}^{-1}$ were obtained for k_{UN} and $m_{\text{U-TS}}$, respectively. For anisotropy, values of $0.014 \pm 0.0020 \text{ sec}^{-1}$ and $0.19 \pm 0.033 \text{ kcal mol}^{-1} \text{ M}^{-1}$ were obtained for k_{NU} and $m_{\text{N-TS}}$, respectively, while values of $3.1 \times 10^{-5} \pm 3.8 \times 10^{-5} \text{ sec}^{-1}$ and $0.37 \pm 0.09 \text{ kcal mol}^{-1} \text{ M}^{-1}$ were obtained for k_{UN} and $m_{\text{U-TS}}$, respectively. The conformational free energy and total m-value were calculated from these parameters, where $\Delta G^{H_2O} = -RT \ln(k_{\text{UN}}/k_{\text{NU}}) = 3.9 \text{ kcal mol}^{-1}$ for fluorescence emission and $3.6 \text{ kcal mol}^{-1}$ for anisotropy. The m-values were calculated as $m_{\text{total}} = m_{\text{U-TS}} + m_{\text{N-TS}} = 0.6 \text{ kcal mol}^{-1} \text{ M}^{-1}$ for fluorescence and $0.56 \text{ kcal mol}^{-1} \text{ M}^{-1}$ for anisotropy. These results suggested that this phase represents partial unfolding of the monomer, since the conformational free energy and m-value are about half of those determined previously for the monomer ($7.0 \pm 0.5 \text{ kcal mol}^{-1}$ and $1.2 \text{ kcal mol}^{-1} \text{ M}^{-1}$) (Bose and Clark, 2001).

The proposed mechanism based on the single-mixing unfolding data presented here is shown in Figure 13. In this sequential mechanism, the native protein forms multiple dimeric species in the burst phase, followed by a series of dimeric intermediates, which then slowly dissociate and unfold. The detection method(s) used to determine each step in the folding reaction are listed.

F. Procaspase-3 dimer dissociates slowly

Sequential-mixing studies were performed to further investigate the folding pathway of procaspase-3. Double-jump studies consist of an unfolding and refolding jump. In this experiment, native protein is mixed with a high concentration of urea, which allows the protein to unfold. The protein is unfolded for various amounts of time (delay time), followed by dilution of the urea to allow the protein to refold. The refolding signal is monitored over time. The purpose of this experiment is to examine intermediates in the refolding pathway or those intermediates that equilibrate with the unfolded protein ensemble due to cis-trans proline isomerization (Brandts et al., 1975; Kiefhaber et al., 1990). In the experiments presented here, the signal was monitored by fluorescence emission, CD and anisotropy.

Double-jump anisotropy data revealed a burst phase (Figure 14A), where a large increase in signal was observed with delay times of less than ~0.1 min. This increase in amplitude is similar to that observed with the burst and lag phases observed in the single mixing unfolding data monitored by anisotropy. As described above, the protein is still in its dimeric state during this phase. Therefore, the intermediate that forms during the burst phase of unfolding persists to at least 0.1 minutes in 8 M urea and represents a conformational transition of the dimer.

As the delay time (unfolding time) was increased, the burst phase amplitude decreased. This decrease is associated with dimer dissociation. As described above, single mixing anisotropy unfolding studies (Figure 10C) suggested that the protein is completely unfolded within 8 minutes. As shown in Figure 14A, at delay times greater than 8 minutes,

the double jump anisotropy burst phase signal is close to that of the native signal, consistent with the results obtained in the single-mixing unfolding data. In addition, the anisotropy signal from single-mixing anisotropy data revealed two phases (data not shown). Two phases also were observed in the double-jump data (Figure 14B). These data were fit to one or two exponentials and the amplitudes are shown in Figure 14C. The amplitude of the slow phase increased over time, the half time of which is associated with the dimer dissociation. Overall, these data suggest that the protein unfolds slowly in 8 M urea.

DISCUSSION

Stopped-flow fluorescence emission and CD refolding studies were employed in order to elucidate the kinetic folding mechanism of procaspase-3. These experiments revealed the formation of multiple monomeric species, but were unable to detect the rate of dimerization of procaspase-3. Therefore, other detection methods (such as anisotropy, differential quenching by acrylamide and return of activity studies) were used. All of these methods revealed a complex burst phase. During these first few milliseconds of the folding reaction, multiple monomeric species were formed that are associated with the burial of aromatic residues, as indicated by a blue shift in fluorescence emission. Emission scans of an equilibrium intermediate in 4 M urea also revealed a blue shift in fluorescence emission, which shows that the burst phase species is not the native dimer (Bose and Clark, 2001).

The burst phase was followed by multiple phases, none of which are protein concentration dependent. Therefore, these phases represent conformational changes in the monomeric form of procaspase-3. For example, the far-UV CD data reveal a lag of about 100 seconds in which there was no change signal, indicating no change in secondary structure in this phase. During this same time frame, three phases were observed in the fluorescence emission data, indicating that there are structural changes in the protein that affect the tertiary structure. Further analysis of these data revealed that these phases have low free energy values, which correlates to the fact that there are no apparent secondary structural changes in the protein as indicated by CD studies. Therefore, the monomeric intermediates that are

formed during these phases consist of structural rearrangements to the protein that affect the aromatic residues, but not the secondary structure of the protein. These monomeric intermediate species are most likely off-pathway or misfolded forms of the protein.

As mentioned above, the observed phases were not dependent upon the protein concentration, indicating the formation of monomeric species. These phases were also associated with low free energy and m-values, indicating small structural changes for each transition. Together, these results point toward the misfolding of these monomeric intermediates before the dimerization-competent species is formed. Additional sequential mixing studies are necessary to determine if these intermediate are on- or off-pathway.

Eventually procaspase-3 was able to form a dimerization-competent species, which dimerized at a rate of $\sim 70 \text{ M}^{-1} \text{ s}^{-1}$ as determined by return of activity studies. This dimerization rate is slow compared to the rate of dimerization of other homodimers, such as SecA ($\sim 10^9 \text{ M}^{-1} \text{ s}^{-1}$) (Doyle et al., 2004), P22 arc repressor ($\sim 10^7 \text{ M}^{-1} \text{ s}^{-1}$) (Milla and Sauer, 1994) and triosephosphate isomerase from *Trypanosoma cruzi* ($\sim 2.6 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$) and *Trypanosoma brucei* ($\sim 2 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$) (Zomosa-Signoret et al., 2003). These proteins are involved in a variety of functions, from transcriptional regulation to the breakdown of sugar in glycolysis. The rate of dimerization has only been measured for only one other caspase. The initiator procaspase-8 is mostly monomeric at physiological conditions, but forms dimers in the presence of activating scaffolds *in vivo* (Kischkel et al., 1995; Medema et al., 1997) or in the presence of kosmotropes *in vitro* (Boatright et al., 2003). Consequently, the rate of dimerization of procaspase-8 was measured in high salt concentrations (Pop et al., 2007). In

1 M sodium citrate, for example, the second order rate of dimerization was determined to be $5.03 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$. Extrapolation of the rates determined at several salt concentrations to the absence of salt indicates a second order rate of dimerization of approximately zero for procaspase-8. Therefore, even though the rate of dimerization of procaspase-3 is slow, that of procaspase-8 is even slower. This conclusion supports the fact that procaspase-3 exists as a stable dimer, whereas procaspase-8 exists as a stable monomer. Consistent with these results is the fact that the dimer interfaces of these two proteins are extremely different (Figure 15). The dimer interface of procaspase-3 is hydrophobic, while that of procaspase-8 contains a proline in the dimer interface, which presumably hinders dimerization as a negative design element. In addition, caspase-8 lacks charge-charge interactions in helices 5 and 5' that are thought to stabilize the dimer (MacKenzie and Clark, in press). Together, these two differences may decrease the rate of dimerization of caspase-8.

The kinetic unfolding pathway was also complex, as was shown by the stopped-flow fluorescence emission, CD and anisotropy data. Unfolding studies suggested that the dimer dissociates slowly in high concentrations of urea. These results were confirmed by sequential mixing studies. Equilibrium studies have shown that the dimerization and subsequent conformational changes contribute about 70% of the total conformational stability of the protein (Bose and Clark, 2001). Therefore, slow dissociation is expected. In addition, the unfolding burst phase was complex consisting of multiple dimeric species. We proposed that these species formed in the first few milliseconds of unfolding are N_2 , N_2^* and I_2 . The N_2 and N_2^* species represent two conformations of the native protein: inactive (N_2) and active (N_2^*).

Evidence for the two native conformations of procaspase-3 will be discussed briefly, beginning with the inactive form (N_2). First, it had been shown that procaspase-3 exists in the cell as an inactive dimer (Pop et al., 2001). Currently, there is no structure of procaspase-3. However, there is a crystal structure of another executioner caspase, procaspase-7 (Chai et al., 2001; Riedl et al., 2001). These two proteins have high sequence identity (55%), and their mature enzyme structures are similar. Therefore, procaspase-7 has been used for a structural model for procaspase-3. This structure is assumed to be in the inactive conformation due to the fact that the active sites are misaligned for substrate binding and catalysis (Riedl and Salvesen, 2007). Finally, there is a tripeptide sequence in the intersubunit linker of the protein termed the “safety catch”, D179-D181. Mutations in this sequence lead to activation of the enzyme (Roy et al., 2001).

The active form (N_2^*) of procaspase-3 was originally found by mutational studies in the dimer interface. The procaspase-3 mutant V266E has a 60-fold increased level of activity, which is 2-fold less than the activity of the fully mature wild-type caspase-3 (Pop et al., 2003). These studies demonstrated that it is not necessary to cleave the intersubunit linker in order to achieve maximum activity. Based on all of this evidence, homology models of the active and inactive native forms of the protein were created (Figure 16) (Walters, et al., unpublished). In the inactive conformation the intersubunit linker is bound, as it is in the procaspase-7 structure, in the dimer interface, preventing insertion of the active site loops into the active site and formation of the substrate binding pocket. In the active conformation, structural changes in the protein move the intersubunit linker out of the dimer interface,

which allows for proper active site formation.

The kinetic unfolding burst phase data further support the theory of two native conformations of procaspase-3. From the previous enzymatic assays (Pop et al., 2003), we calculated the $\Delta\Delta G$ between the inactive (N_2) and active (N_2^*) forms of procaspase-3 to be 2.23 kcal/mol at 25 °C. This value was similar to that obtained from the first transition in the unfolding kinetic burst phase experiments, 2.45 ± 0.55 kcal/mol. In addition, we calculated the change in solvent accessible surface area between the two conformations by comparing the two homology models (Novotny et al., 2007). The ΔASA was then correlated to the number of residues effected by the transition and to the m-value for the transition using the analysis described by Myers *et al.* (Myers et al., 1995a). Using their correlations, the m-value and number of residues were determined for the homology models, and the results are presented in Table III. The experimental data for the unfolding kinetic burst phase provided an m-value for the transition, and this also was correlated to the ΔASA and number of effected residues (Myers et al., 1995a). The results also are presented in Table III. A comparison of the values determined from the homology models and the kinetic burst phase experiments were similar. The analysis shows that the number of residues affected by the conformational transition is between 40 and 60, which correlates to the changes in the intersubunit linker and active site upon the transition of N_2 to N_2^* . Together, the data strongly suggest that the unfolding kinetic burst phase experiments monitor the interconversion of two native conformations during unfolding, and based on our previous data (Pop et al., 2003), we suggest that one conformation is enzymatically active. The

unfolding burst phase results are important because, thus far, this is the only detection method we have found that is able to monitor the two native conformations. Therefore, *in vivo* the cell most likely maintains an apoptotic balance by keeping procaspase-3 locked in the inactive conformation.

METHODS

Materials

Ultrapure urea was purchased from ICN. Dithiothreitol (DTT), acrylamide and imidazole were from Acros. Sodium chloride (NaCl), tryptone, yeast extract and trizma base (Tris) were from Fisher. Q-Sepharose was from Amersham Biosciences. Monobasic and dibasic potassium phosphate (KH_2PO_4 , K_2HPO_4), ampicillin, and molecular weight markers were from Sigma. IPTG (isopropyl β -D-1-thiogalactopyranoside) was from Anatrace. YM 10 membranes were from Millipore. His bind was from Novagen. Ac-DEVD-AFC (acetyl-DEVD-7-amino-4-trifluoromethyl) was from Calbiochem. All buffers were filtered through either 0.45 or 0.22 μm filter membranes. Urea stock solutions were prepared as described previously (see Appendix protocols).

Protein Purification

All steps were performed at 4 °C unless otherwise noted. Procaspase-3 (C163S and D9A,D28A,D175A also termed D3A) were purified as C-terminal-(His)₆-tagged proteins from *E. coli* BL21(DE3) lys S cells. The purification protocol of both proteins was similar. Cells were grown in Fernbach flasks containing 1 L of LB media with 50 $\mu\text{g}/\text{mL}$ ampicillin at 37 °C. When the cultures reached an absorbance (at 600 nm) of 1.2, protein expression was induced with a final IPTG concentration of 0.6 mM (C163S) or 0.8 mM (D3A). The temperature was reduced to 20 °C for 20 hours (C163S) or 25 °C for 6 hours (D3A). The cells were harvested by centrifugation at 5000 rpm for 16 minutes. The cells were

resuspended in lysis buffer (50 mM imidazole, 50 mM Tris, 50 mM NaCl pH 7.9) and lysed with a French press. The supernatant was separated from cell debris by centrifugation at 20,000 rpm for 20 minutes. The protein was initially purified with a histidine affinity column and eluted with 500 mM imidazole, 50 mM Tris, 50 mM NaCl pH 7.9. The protein was dialyzed against 50 mM Tris, 50 mM NaCl pH 7.9 with 1 mM DTT overnight. The protein was further purified with a Q-Sepharose column with a 50-350 mM NaCl salt gradient (50 mM Tris pH 7.9 plus 1 mM DTT). Fractions were collected and tested with a Bradford assay. The positive fractions were analyzed by SDS-PAGE. The procaspase-3 fractions were pooled, concentrated (YM 10 membrane) and dialyzed overnight against 50 mM potassium phosphate pH 7.5. The protein was stored at -20 °C. The concentrations of procaspase-3 (C163S and D3A) were determined using ϵ_{280} of 26,500 M⁻¹ cm⁻¹.

Single mixing stopped-flow fluorescence and CD

Single mixing kinetic folding/unfolding experiments were performed using a stopped-flow spectrofluorometer (SX18) or spectrophotometer (PiStar) from Applied Photophysics. The temperature was controlled at 25 °C using a circulating water bath. For fluorescence and anisotropy studies, the samples were excited at 280 or 295 nm and fluorescence emission was measured using a 305 nm cut-off filter. The CD signal was monitored at 228 nm. For the differential quenching by acrylamide studies, samples were excited at 295 nm and fluorescence emission was monitored using a 305 nm cut-off filter. The slits for fluorescence and differential quenching by acrylamide, anisotropy and CD were 0.76 mm, 2 mm and 2 nm, respectively. The final concentration of DTT in the urea-containing buffers was 1 mM.

The refolding experiments were performed by mixing unfolded protein (22, 44, 88 or 132 μM for fluorescence with excitation at 280 nm; 22, 55 or 165 μM for fluorescence with excitation at 295 nm and differential quenching by acrylamide; 44, 88 or 132 μM for CD; 35.75, 55 or 165 μM for anisotropy with excitation at 280 or 295 nm) in 8 M urea-containing buffer with 50 mM potassium phosphate buffer-containing urea between 0 and 8 M, as shown in the figures. The final protein concentration was 2, 4, 8 or 12 μM for fluorescence with excitation at 280 nm, 2, 5 or 15 μM for fluorescence and differential quenching by acrylamide with excitation at 295 nm, 4, 8 or 12 μM for CD and 3.25, 5 or 15 μM for anisotropy with excitation at 280 or 295 nm. A mixing ratio of 1:10 was used.

The unfolding experiments were performed by mixing native protein (44 or 165 μM for CD; 38.5 or 143 μM for anisotropy and fluorescence with excitation at 280 or 295 nm) in 50 mM potassium phosphate buffer with urea-containing buffers, as shown in the figures. The final protein concentration was 4 or 15 μM for CD and 3.5 or 13 μM for anisotropy and fluorescence with excitation at 280 or 295 nm. A mixing ration of 1:10 was used.

Emission Scanning

Emission scanning experiments were performed using a stopped-flow (SX18 Applied Photophysics). The emission photomultiplier tube was connected to a monochromator and the signal was measured at multiple emission wavelengths. The excitation and emission slits were 2 and 5 mm, respectively. 55 μM unfolded procaspase-3 C163S in 8 M urea-containing buffer was mixed 1:10 with 50 mM potassium phosphate buffer pH 7.5. The fluorescence burst phase signal was collected (excitation at 280 or 295 nm) at several emission

wavelengths.

Hand-mixing fluorescence and CD spectroscopy

Hand-mixing refolding studies were performed with a C-61 spectrofluorometer (differential quenching by acrylamide fluorescence, PTI) and J600A spectropolarimeter (CD, Jasco Inc.) Unfolded protein (22, 55 or 110 μM for differential quenching by acrylamide and 20, 40 or 80 μM for CD) in 8 M urea-containing buffer was mixed with 50 mM potassium phosphate buffer pH 7.5 (CD) or with 0.7 M acrylamide-containing phosphate buffer (differential quenching by acrylamide). The dead time was ~ 10 seconds. The signal at 228 nm (CD) or excitation at 295 nm with emission at 335 nm (differential quenching by acrylamide) was monitored continuously and discontinuously over time.

Return of activity assays

The activity of procaspase-3 D3A was examined using a C-61 spectrofluorometer (PTI). Unfolded protein (88, 55, 44, 27.5, 22, 16.5 or 5.5 μM) in 8 M urea-containing buffer was hand-mixed with 50 mM potassium phosphate buffer pH 7.5. Aliquots of the refolding reaction were tested for activity. The activity reaction volume was 200 μL , with final concentrations of 400 nM protein, 1 mM DTT, 100 μM substrate (Ac-DEVD-AFC) and 0.73 M urea-containing buffer. The fluorescence emission was measured at 505 nm after excitation at 400 nm.

Sequential mixing stopped-flow experiments

Sequential mixing experiments were performed with the stopped-flow spectrofluorometer (SX18, Applied Photophysics). The temperature was held constant at 25

°C. In general for double jump experiments, the first jump was performed with a mixing ratio of 1:5 (native protein : urea solution). 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 (unfolding protein : buffer). Delay times range from 0.01 to 900 seconds.

Native protein (126 or 396 μM) was mixed 1:5 with 10 M urea-containing buffer for the first jump. After various delay times, the unfolding protein solution was mixed 1:5 with 50 mM potassium phosphate buffer for the second jump. The anisotropy signal (with excitation at 280 or 295 nm) was monitored for 500 seconds. The final protein concentration was 3.5 or 11 μM , and the final urea concentration was 1.38 M. For the native protein signal, a protein stock solution (126 or 396 μM) in 1.38 M urea-containing buffer was mixed with 1.38 M urea-containing buffer. For the unfolded protein signal, a protein stock solution (126 or 396 μM) in 8.3 M urea-containing buffer was mixed with 8.3 M urea-containing buffer.

REFERENCES

- Baldwin, R. L. (1996). On-pathway versus off-pathway folding intermediates. *Folding and Design* **1**, R1-R8.
- Bao, Q., and Shi, Y. (2006). Apoptosome: a platform for the activation of initiator caspases. *Cell* **14**, 56-65.
- Barshop, B. A., Wrenn, R. F., and Frieden, C. (1983). Analysis of numerical methods for computer simulation of kinetic processes: development of KINSIM--a flexible, portable system. *Analytical Biochemistry* **130**, 134-145.
- Bhuyan, A. K., and Udgaonkar, J. B. (1998). Multiple kinetic intermediates accumulate during the unfolding of horse cytochrome c in the oxidized state. *Biochemistry* **37**, 9147-9155.
- Bose, K., and Clark, A. C. (2001). Dimeric procaspase-3 unfolds via a four-state equilibrium process. *Biochemistry* **40**, 14236-14242.
- Brandts, J. F., Halvorson, H. R., and 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**, 4953-4963.
- Brockwell, D. J., and Radford, S. E. (2007). Intermediates: ubiquitous species on folding energy landscapes? *Current Opinion in Structural Biology* **17**, 30-37.
- Boatright, K. M., Renatus, M., Scott, F. L., Sperandio, S., Shin, H., Pedersen, I. M., Ricci, J.-E., Edris, W. A., Sutherlin, D. P., Green, D. R., and Salvesen, G. S. (2003). A unified model for apical caspase activation. *Molecular Cell* **11**, 529-541.
- Chai, J., Shiozaki, E. N., Srinivasula, S. M., Alnemri, E. S., and Shi, Y. (2001). Crystal structure of a procaspase-7 zymogen mechanisms of activation and substrate binding. *Cell* **107**, 399-407.
- Chamond, R. R., Anon, J. C., Aguilar, C. M., and Pasadas, F. G. (1999). Apoptosis and disease. *Alergol Inmunol Clin* **14**, 367-374.

- Doyle, S. M., Bilsel, O., and Teschke, C. M. (2004). SecA folding kinetics: a large dimeric protein rapidly forms multiple native states. *Journal of Molecular Biology* **341**, 199-214.
- Feeney, B., and Clark, A. C. (2005). Reassembly of active caspase-3 is facilitated by the propeptide. *Journal of Biological Chemistry* **280**, 39772-39785.
- Ferguson, N., Capaldi, A. P., James, R., Kleanthous, C., and Radford, S. E. (1999). Rapid folding with and without populated intermediates in the homologous four-helix proteins Im7 and Im9. *Journal of Molecular Biology* **286**, 1597-1608.
- Goodsell, D. S., and Olson, A. J. (2000). Structural symmetry and protein function. *Annual Review of Biophysics & Biomolecular Structure* **29**, 105-154.
- Hotchkiss, R. S., and Nicholson, D. W. (2006). Apoptosis and caspases regulate death and inflammation in sepsis. *Nature Reviews Immunology* **6**, 813-822.
- Jackson, S. E., and Fersht, A. R. (1991). Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition. *Biochemistry* **30**, 10428-10435.
- Kiefhaber, T., Quaas, R., Hahn, U., and Schmid, F. X. (1990). Folding of ribonuclease T1. 1. Existence of multiple unfolded states created by proline isomerization. *Biochemistry* **29**, 3053-3061.
- Kischkel, F. C., Hellbardt, S., Behrmann, I., Germer, M., Pawlita, M., Krammer, P. H., and Peter, M. E. (1995). Cytotoxicity-dependent APO-1 (Fas/CD95)-associated proteins form a death-inducing signaling complex (DISC) with the receptor. *The European Molecular Biology Organization Journal* **14**, 5579-5588.
- Luthi, A. U., and Martin, S. J. (2007). The CASBAH: a searchable database of caspase substrates. *Cell Death and Differentiation* **14**, 641-650.
- Mann, C. J., and Matthews, C. R. (1993). Structure and stability of an early folding intermediate of Escherichia coli trp aporepressor measured by far-UV stopped-flow circular dichroism and 8-anilino-1-naphthalene sulfonate binding. *Biochemistry* **32**, 5282-5290.

- Medema, J. P., Scaffidi, C., Kischkel, F. C., Shevchenko, A., Mann, M., Krammer, P. H., and Peter, M. E. (1997). FLICE is activated by association with the CD95 death-inducing signaling complex (DISC). *The European Molecular Biology Organization Journal* **16**, 2794-2804.
- Milam, S. L., Nicely, N. I., Feeney, B., Mattos, C., and Clark, A. C. (2007). Rapid folding and unfolding of Apaf-1 CARD. *Journal of Molecular Biology* **369**, 290-304.
- Milla, M. E., and Sauer, R. T. (1994). P22 Arc repressor: folding kinetics of a single-domain, dimeric protein. *Biochemistry* **33**, 1125-1133.
- Myers, J. K., Pace, C. N., and Scholtz, J. M. (1995). Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Science* **4**, 2138-2148.
- Novotny, M., Seibert, M., and Kleywegt, G. J. (2007). On the precision of calculated solvent-accessible surfaces areas. *Acta Crystallographica Section D* **D63**, 270-274.
- Parker, M. J., Spencer, J., and Clarke, A. R. (1995). An integrated kinetic analysis of intermediates and transition states in protein folding reactions. *Journal of Molecular Biology* **253**, 771-786.
- Peter, M. E., and Krammer, P. H. (2003). The CD95(APO-1/Fas) DISC and beyond. *Cell Death & Differentiation* **10**, 26.
- Pop, C., Chen, Y.-R., Smith, B., Bose, K., Bobay, B., Tripathy, A., Franzen, S., and Clark, A. C. (2001). Removal of the pro-domain does not affect the conformation of the procaspase-3 dimer. *Biochemistry* **40**, 14224-14235.
- Pop, C., Feeney, B., Tripathy, A., and Clark, A. C. (2003). Mutations in the procaspase-3 dimer interface affect the activity of the zymogen. *Biochemistry* **42**, 12311-12320.
- Pop, C., Fitzgerald, P., Green, D. R., and Salvesen, G. S. (2007). Role of proteolysis in caspase-8 activation and stabilization. *Biochemistry* **36**, 4398-4407.
- Renatus, M., Stennicke, H. R., Scott, F. L., Liddington, R. C., and Salvesen, G. S. (2001). Dimer formation drives the activation of the cell death protease caspase 9. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 14250-14255.

- Riedl, S. J., Fuentes-Prior, P., Renatus, M., Kairies, N., Krapp, S., Huber, R., Salvesen, G. S., and Bode, W. (2001). Structural basis for the activation of human procaspase-7. *Proceedings of the National Academy of Sciences* **98**, 14790-14795.
- Riedl, S. J., and Salvesen, G. S. (2007). The apoptosome: signalling platform of cell death. **8**, 405-413.
- Roy, S., Bayly, C. I., Gareau, Y., Houtzager, V. M., Keargman, S., Keen, S. L. C., Rowland, K., Seiden, I. M., Thornberry, N. A., and Nicholson, D. W. (2001). Maintenance of caspase-3 proenzyme dormancy by an intrinsic "safety catch" regulatory peptide. *Proceedings of the National Academy of Sciences* **98**, 6132-6137.
- Rumfeldt, J. A. O., Galvagnion, C., Vassall, K. A., and Meiering, E. M. (2008). Conformational stability and folding mechanisms of dimeric proteins. *Progress in Biophysics and Molecular Biology* **98**, 61-84.
- Saikumar, P., Dong, Z., Mikhailov, V., Denton, M., Weinberg, J. M., and Venkatachalam, M. A. (1999). Apoptosis: definition, mechanisms, and relevance to disease. *The American Journal of Medicine* **107**, 489-506.
- Shiozaki, E. N., Chai, J., and Shi, Y. (2002). Oligomerization and activation of caspase-9, induced by apaf-1 CARD. *Proceedings of the National Academy of Sciences* **99**, 4197-4202.
- Taylor, R. C., Cullen, S. P., and Martin, S. J. (2008). Apoptosis: controlled demolition at the cellular level. **9**, 231-241.
- Topping, T. B., Hoch, D. A., and Gloss, L. M. (2004). Folding mechanism of FIS, the intertwined, dimeric factor for inversion stimulation. *Journal of Molecular Biology* **335**, 1065-1081.
- Zaida, F. N., Nath, I., and Udgaonkar, J. B. (1997). Multiple intermediates and transition states during protein unfolding. *Nature Structural Biology* **4**, 1016-1024.
- Zomosa-Signoret, V., Hernandez-Alcantara, G., Reyes-Vivas, H., Martinez-Martinez, E., Garza-Ramos, G., Perez-Montfort, R., Tuena de Gomez-Puyou, M., and Gomez-Puyou, A. (2003). Control of the reactivation kinetics of homodimeric triosephosphate isomerase from unfolded monomers. *Biochemistry* **42**, 3311-3318.

FIGURES AND TABLES

Table I: Refolding burst phase signals fit to two- or three-state equilibrium folding models.

	FL	CD	Anisotropy	Quench
Two-State				
ΔG (kcal mol ⁻¹)	3.17±0.37	3.82±0.44	2.80±0.21	3.88±0.99
m (kcal mol ⁻¹ M ⁻¹)	0.57±0.057	0.66±0.064	0.59±0.039	0.91±0.22
urea _{1/2} (M)	5.6	5.8	4.7	4.3
Three-State				
ΔG_1 (kcal mol ⁻¹)		1.34±0.0048	0.78±0.048	1.4±0.04
m_1 (kcal mol ⁻¹ M ⁻¹)		0.52±0.013	0.31±0.16	0.50±0.10
urea _{1/2} (M)		2.6	2.5	2.8
ΔG_2 (kcal mol ⁻¹)		3.11±0.031	3.14±0.066	3.00±0.11
m_2 (kcal mol ⁻¹ M ⁻¹)		0.60±0.17	0.66±0.34	0.80±0.42
urea _{1/2} (M)		5.2	4.8	3.7

Table II. Unfolding phases and detection methods.

	Burst	Fast (s ⁻¹)	Lag	1 st exponential (s ⁻¹)	2 nd exponential (s ⁻¹)
Detection Method					
Fluorescence emission	Yes	~40	Yes	~0.04	~0.01
Anisotropy			Yes	~0.03	~0.005
Far-UV CD			Yes	~0.018	

Table III. Comparison of unfolding burst phase species and proposed native species.

	homology models	unfolding kinetic data
change in solvent accessible surface area	4587 Å ²	3018 Å ²
m-value	0.87	0.7 ± 0.21
number of residues	59	42

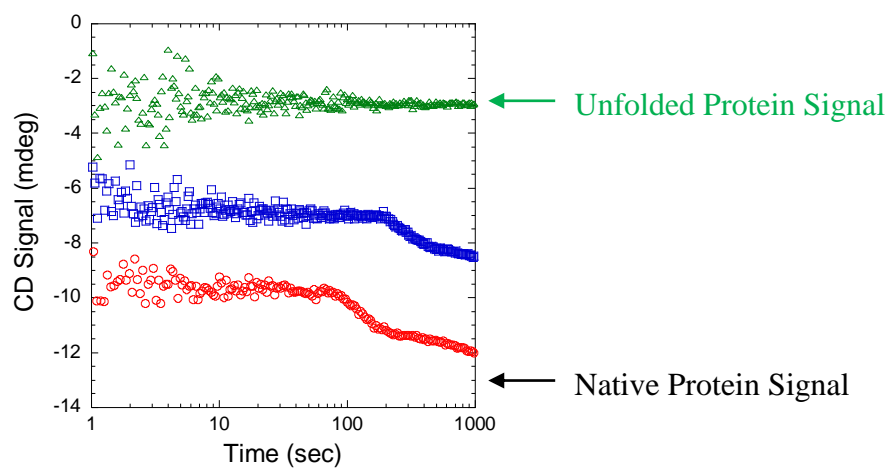
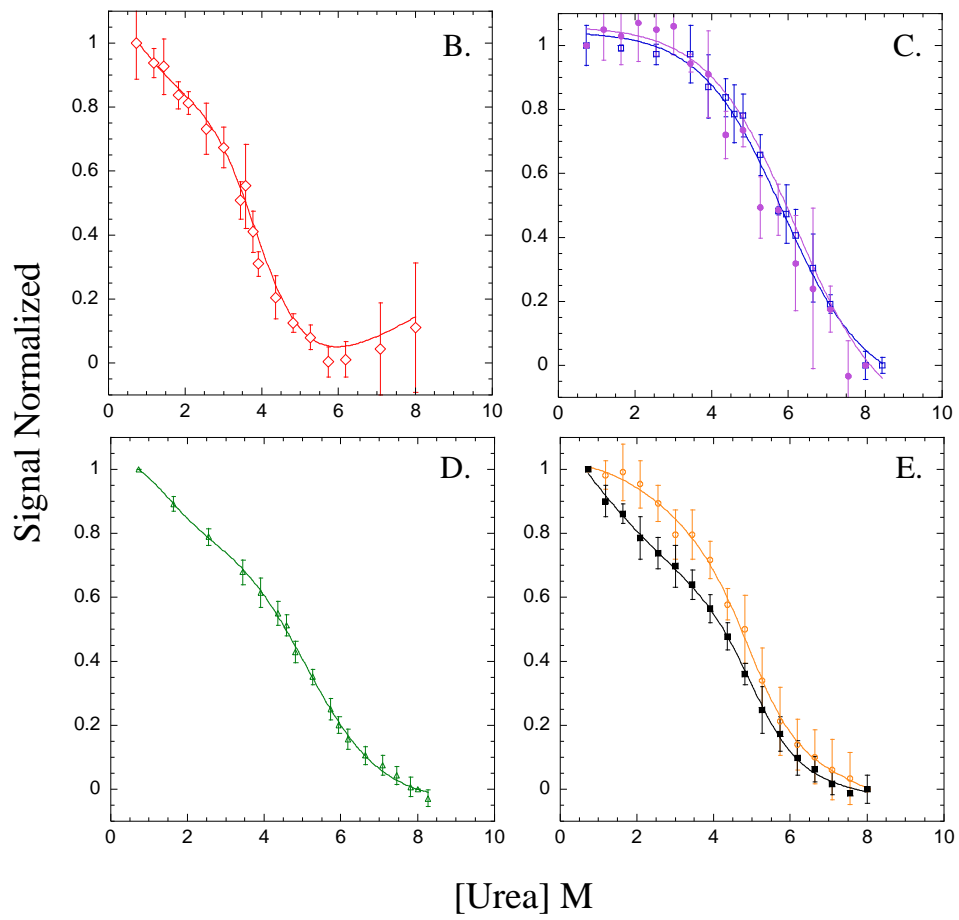
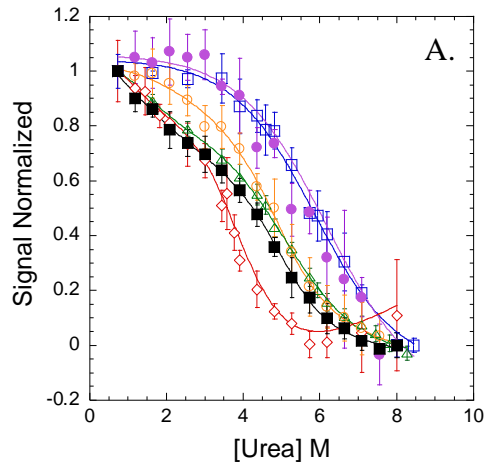


Figure 1: Far-UV CD (228 nm) refolding data. The unfolded (green, Δ) and native protein signals are marked. Unfolded protein is mixed 1:10 with 4.8 M urea-containing buffer, for a final urea concentration of 4.36 M (blue, \square). Unfolded protein is mixed 1:10 with buffer, for a final urea concentration of 0.73 M (red, \circ).

Figure 2: Complex burst phase. A) The burst phase signal for each detection method is plotted versus the final urea concentration. Error bars represent the standard deviation of two independent experiments at multiple protein concentrations. The differential quenching by acrylamide burst phase signal (red, \diamond) is shown in panel B. The fluorescence burst phase signals (excitation at 280 nm (blue, \square) and 295 nm (purple, \bullet)) are shown in panel C. The far-UV CD burst phase signal (228 nm, green, Δ) is shown in panel D. The anisotropy burst phase signals (excitation at 280 nm (orange, \circ) and 295 nm (black, \blacksquare)) are shown in panel E. The error bars represent the standard deviation from multiple protein concentrations in at least 2 independent experiments.



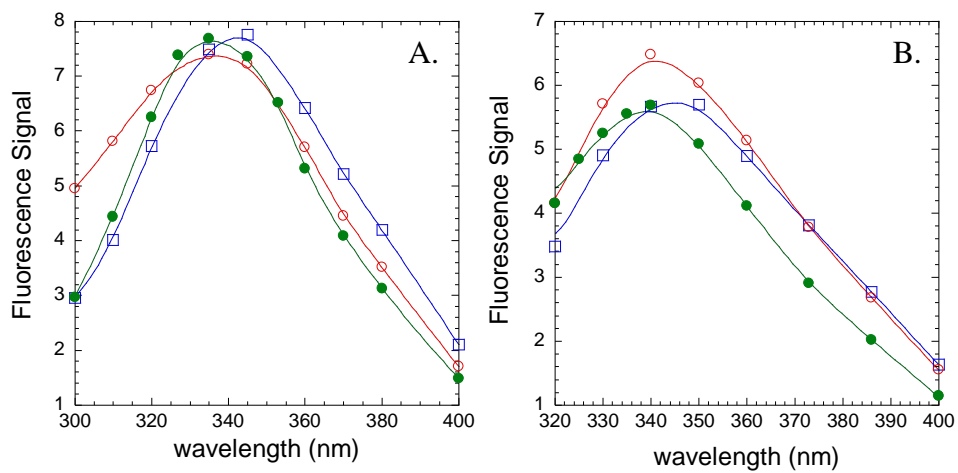
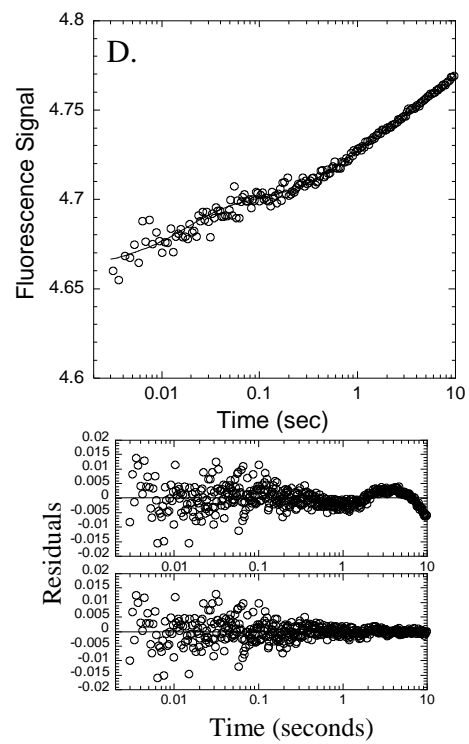
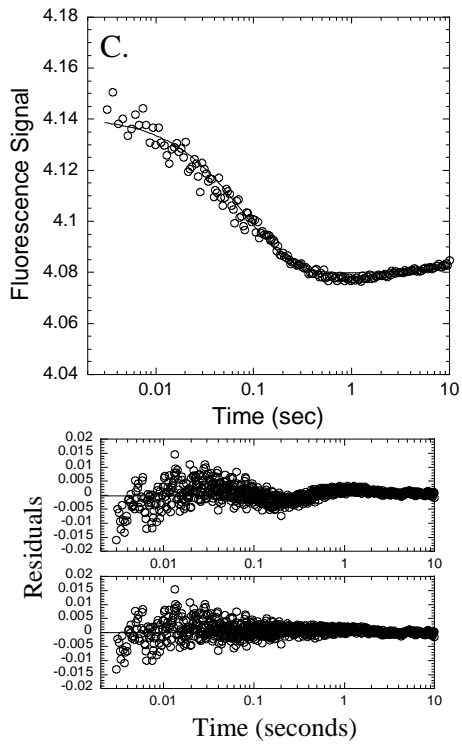
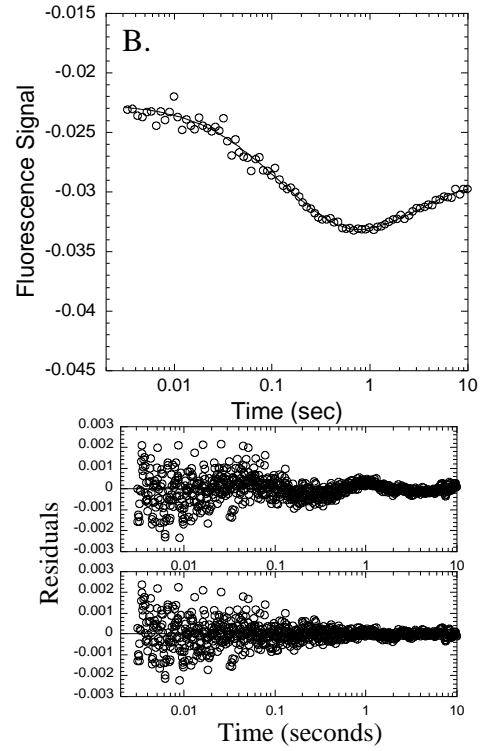
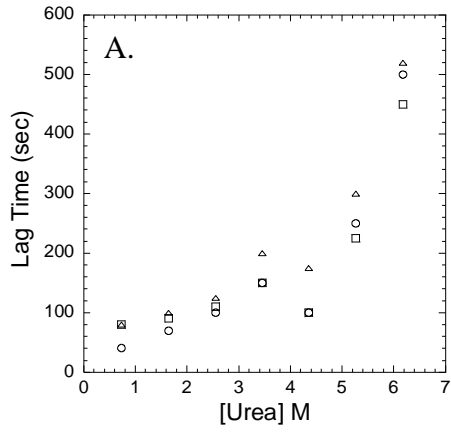


Figure 3: Emission scanning of the burst phase species. 5 μ M unfolded (blue, \square) and native (red, \circ) procaspase-3. Unfolded protein was mixed 1:10 with buffer. The refolding burst phase signal (green, \bullet) was collected at various emission wavelengths. Excitation at 280 nm (panel A) and 295 nm (panel B).

Figure 4: CD lag. A) CD (228 nm) lag phase versus the final urea concentration for 4 (\circ), 8 (\square) and 12 (Δ) μ M procaspase-3. B and C) Refolding procaspase-3 in 3.91 M urea-containing buffer detected by fluorescence with excitation at 280 nm (panel B) or 295 nm (panel C). D) Refolding procaspase-3 in 2.09 M urea-containing buffer detected by differential quenching by acrylamide (excitation 295 nm). The lines through the data represent a three exponential fit. The residuals for 2 and 3 exponential fits are shown below each plot, respectively.



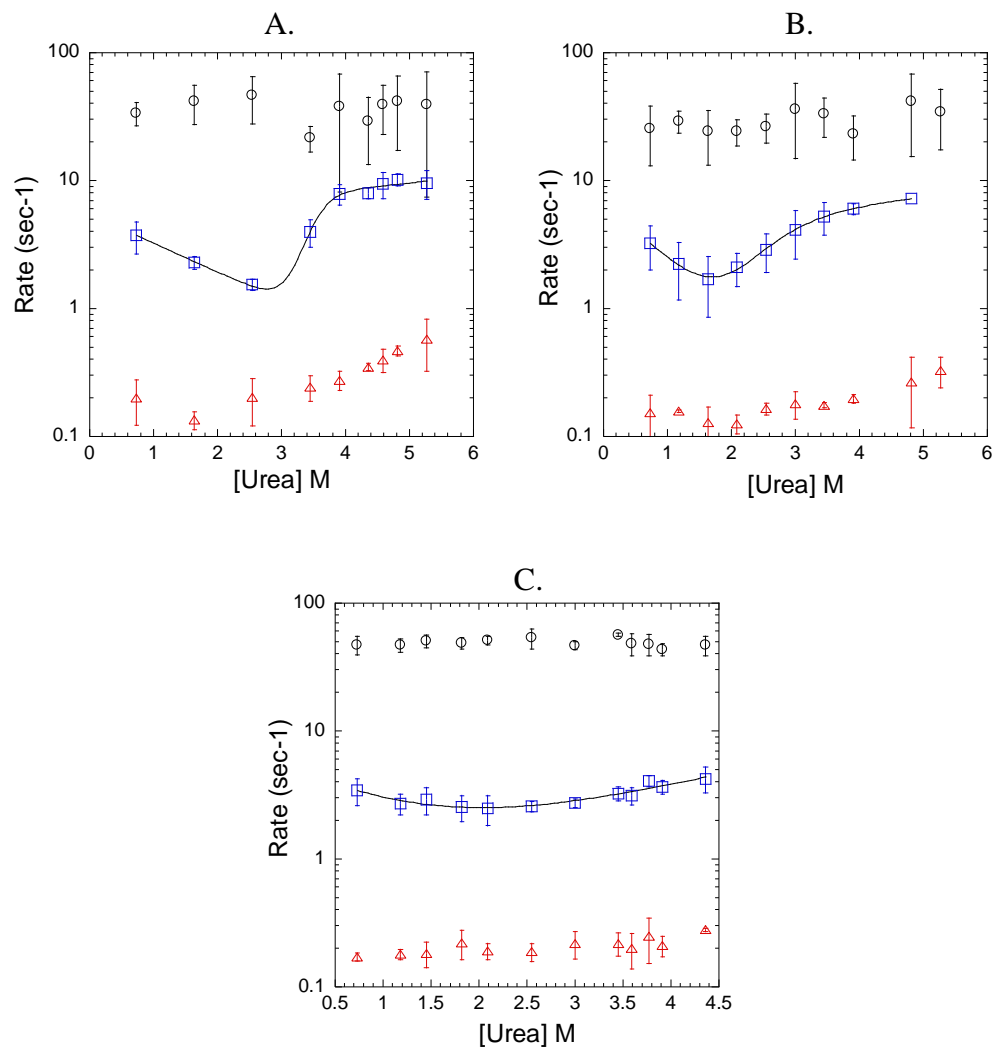


Figure 5: Rate versus the final urea concentration for fluorescence emission following excitation at 280 nm (panel A) or 295 nm (panel B) and differential quenching by acrylamide (panel C). Refolding rates 1 (black, \circ), 2 (blue, \square) and 3 (red, Δ) for the first 10 seconds of the folding reaction. Error bars represent the standard deviation from 2-3 independent experiments and various protein concentrations (ranging from 2-15 μ M procaspase-3). Lines through rate 2 are fit to a two-state kinetic mechanism for panel C and with a three-state kinetic mechanism for panels A and B.

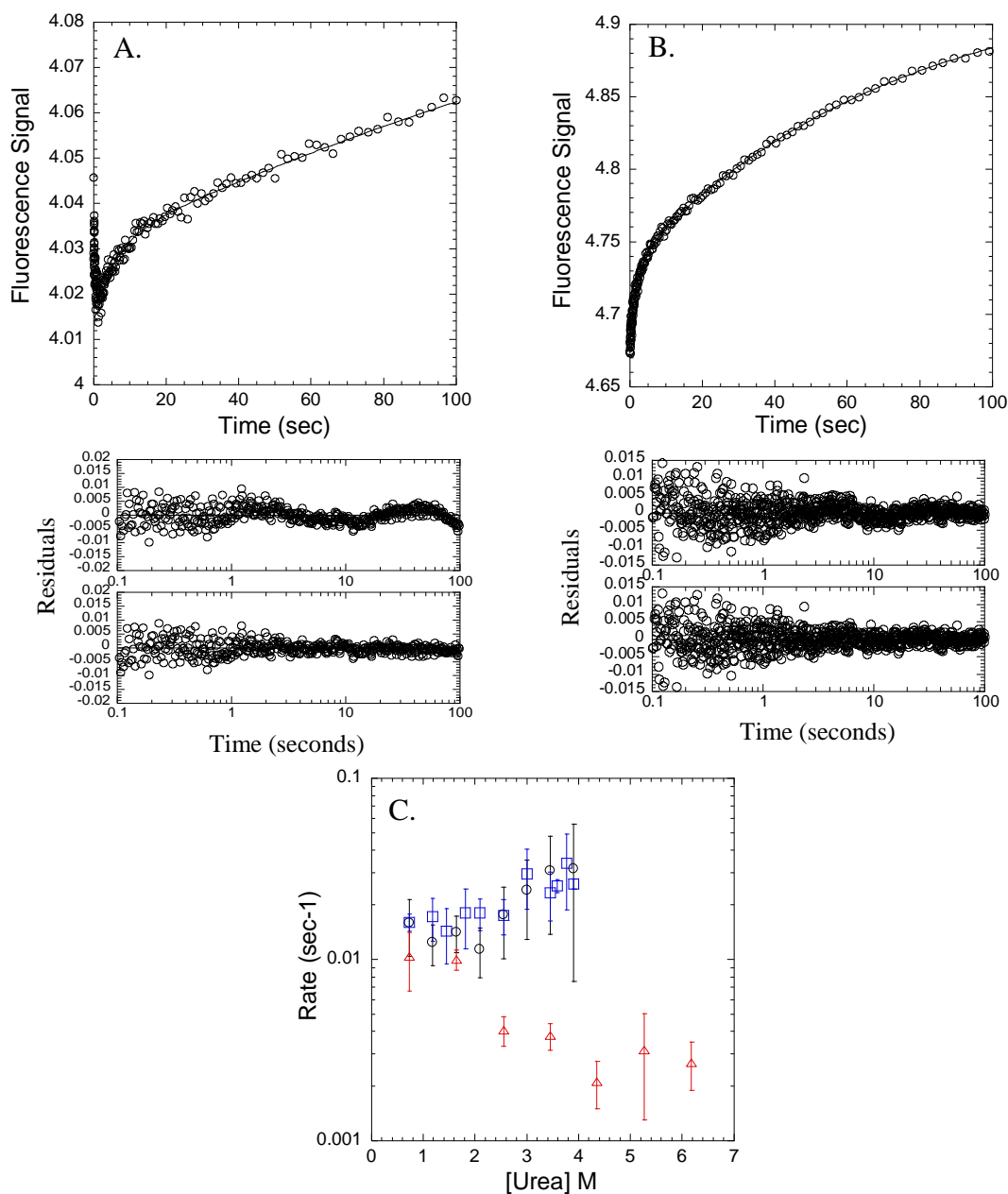


Figure 6: Intermediate phase of refolding. A and B) 100 seconds of refolding in 2.55 M (fluorescence excitation at 295 nm, panel A) or 1.82 M (differential quenching by acrylamide, panel B) urea-containing buffer. The lines through the data represent a three exponential fit. The residuals for 2 and 3 exponential fits are shown below each plot, respectively. C) Rate 4 for refolding as detected by fluorescence (black, \circ), differential quenching by acrylamide (blue, \square) and CD (red, Δ). The rates for the CD data are from single exponential fits of the intermediate phase shown in Figure 1. Error bars represent the standard deviation from two independent experiments at multiple protein concentrations.

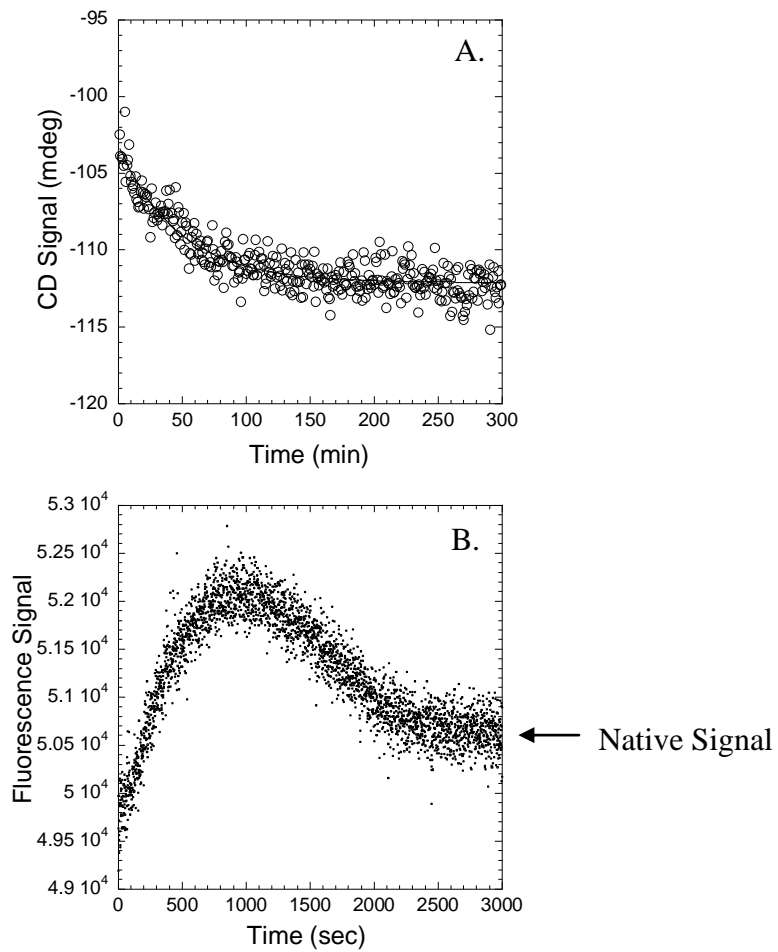


Figure 7: Slow phase of refolding. Unfolded protein was manually mixed with buffer. The CD (panel A) and fluorescence (panel B) signals were collected on a steady-state instrument. The native protein signal is marked.

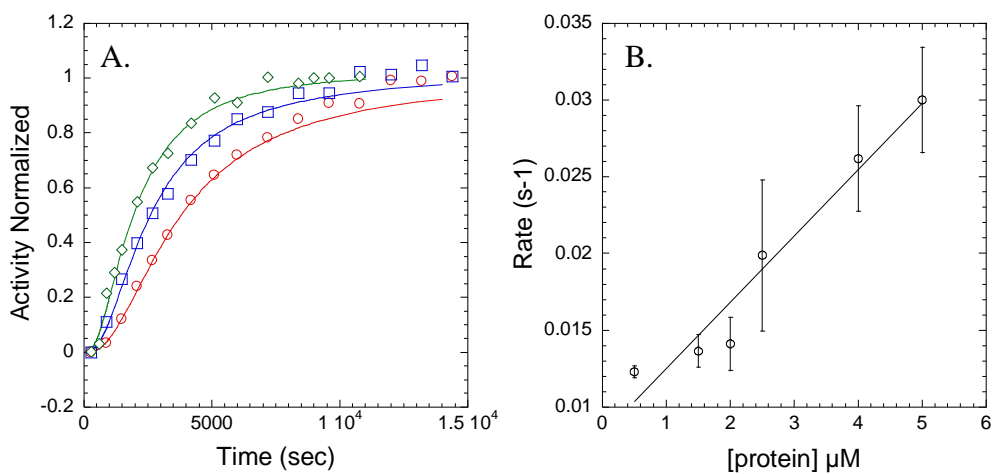
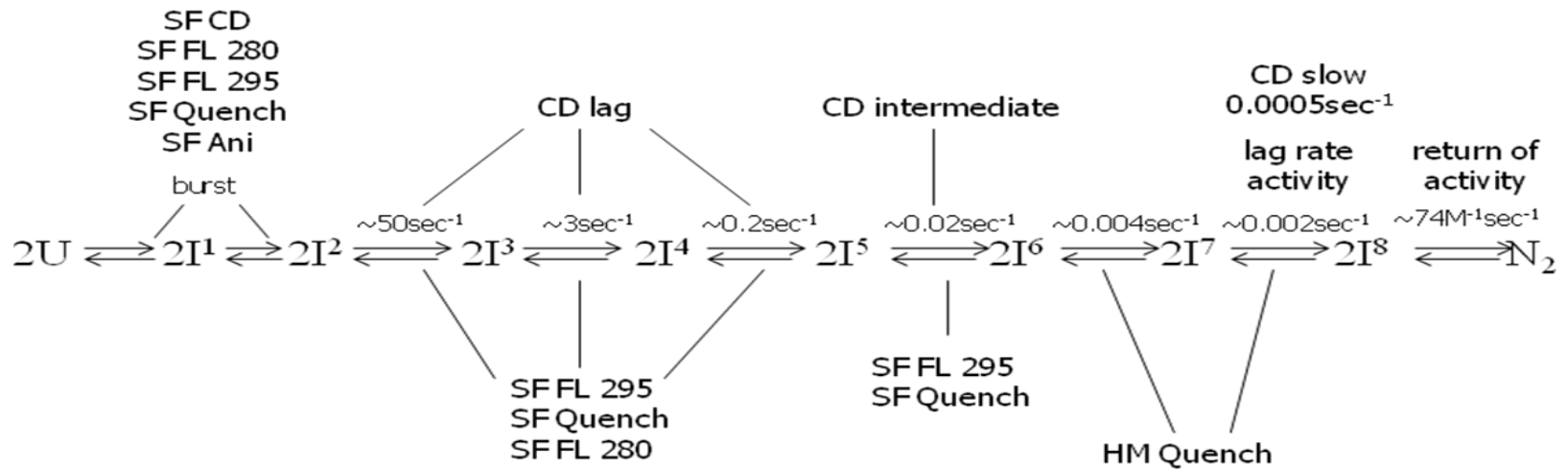


Figure 8: Activity studies with procaspase-3 D3A. A) Unfolded protein was mixed with buffer. Aliquots of the refolding reaction (1.5 (red, \circ), 2.5 (blue, \square) and 5 (green, \diamond) μM procaspase-3 D3A) were tested for activity. The lines through the data are KINSIM simulations ($k_1 = 0.00071 \text{ s}^{-1}$, $k_2 = 115 \text{ M}^{-1}\text{s}^{-1}$). B) Second order rate plot. Rate of second exponential from activity studies plotted versus the protein concentrations. Error bars represent standard deviation from two independent experiments. The data are fit to a line with a slope of $70 \text{ M}^{-1}\text{s}^{-1}$.

Figure 9: Proposed sequential refolding mechanism. The unfolded protein folds through a series of monomeric intermediates (I^1 to I^8). The dimerization competent monomeric intermediate (I^8) dimerizes slowly. Listed above each arrow is the rate of formation of each species. The detection methods are listed; SF = stopped-flow, CD = circular dichroism (228 nm), FL 280/295 (fluorescence with excitation at 280/295 nm), Quench = differential quenching by acrylamide, Ani = anisotropy (excitation at 280 or 295 nm), HM = hand-mixing.



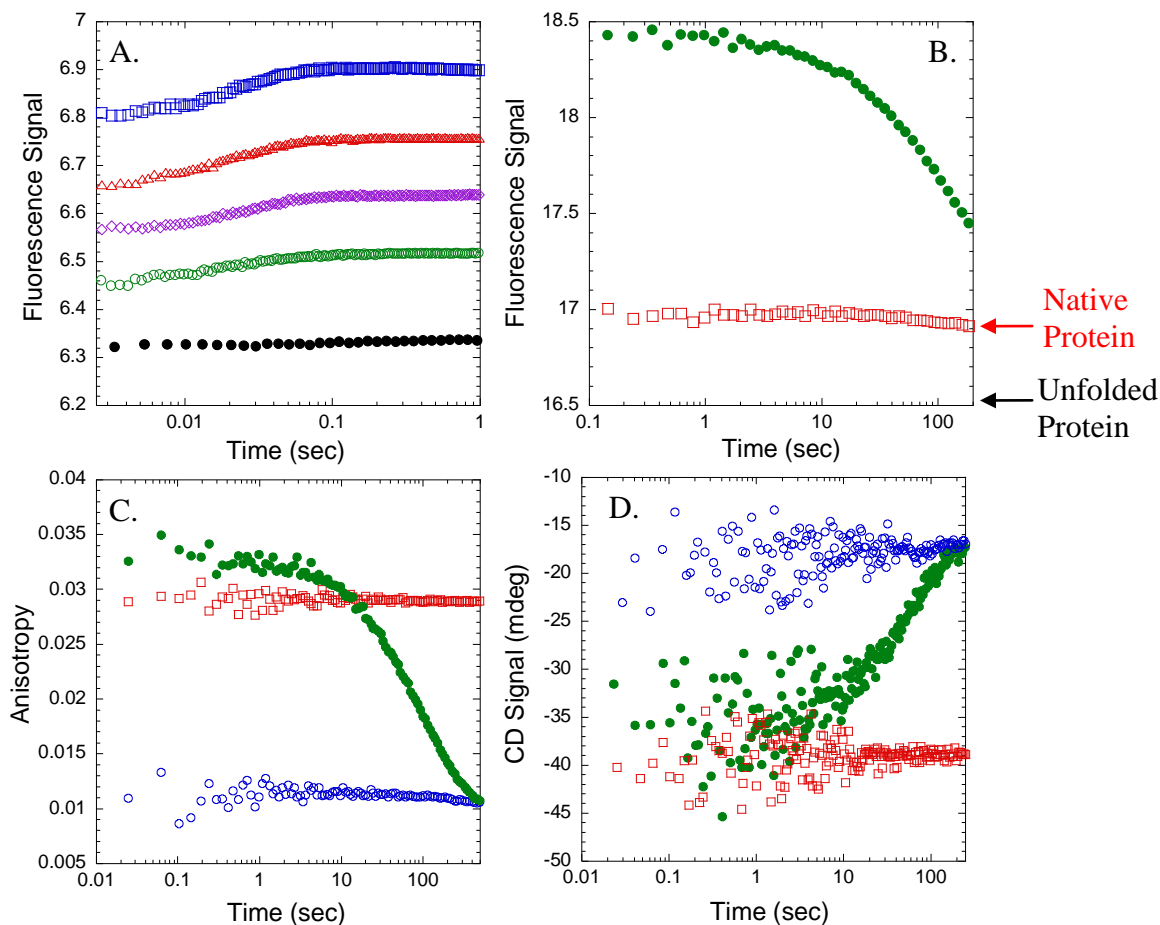


Figure 10: Unfolding kinetic data. A) 165 μM native procaspase-3 was mixed with various urea concentrations for a final urea concentration of 0 M (black, \bullet), 4.55 M (green, \circ), 5.45 M (purple, \diamond), 5.91 M (red, Δ) and 6.82 M (blue, \square). Fluorescence emission (excitation at 280 nm) was collected for 1 sec. B) 38.5 μM native procaspase-3 was mixed with 8.5 M urea-containing buffer for a final urea concentration of 7.73 M (green, \bullet) and the fluorescence emission (excitation at 280 nm) was collected for 200 sec. The native and unfolded protein signals are marked. C and D) 55 μM (anisotropy, excitation 280 nm, panel C) and 165 μM (far-UV CD, 228 nm, panel D) native procaspase-3 were mixed with urea-containing buffer for a final urea concentration of 8 M (green, \bullet). The native and unfolded protein signals are shown in red squares and blue circles, respectively.

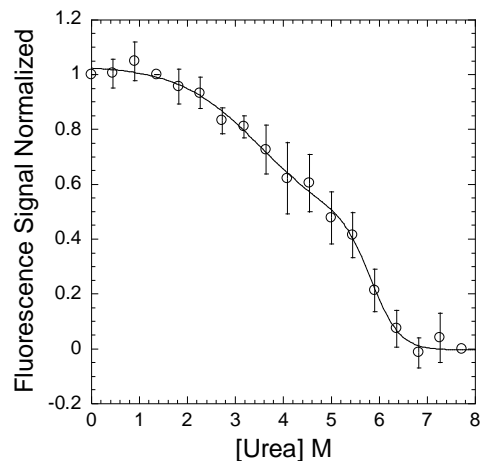


Figure 11: Fluorescence (excitation at 280 nm) burst phase signals versus urea. Error bars represent the standard deviation from two independent experiments at multiple (2, 5 and 15 μM procaspase-3) protein concentrations. The line through the data is a fit to a three-state equilibrium folding model.

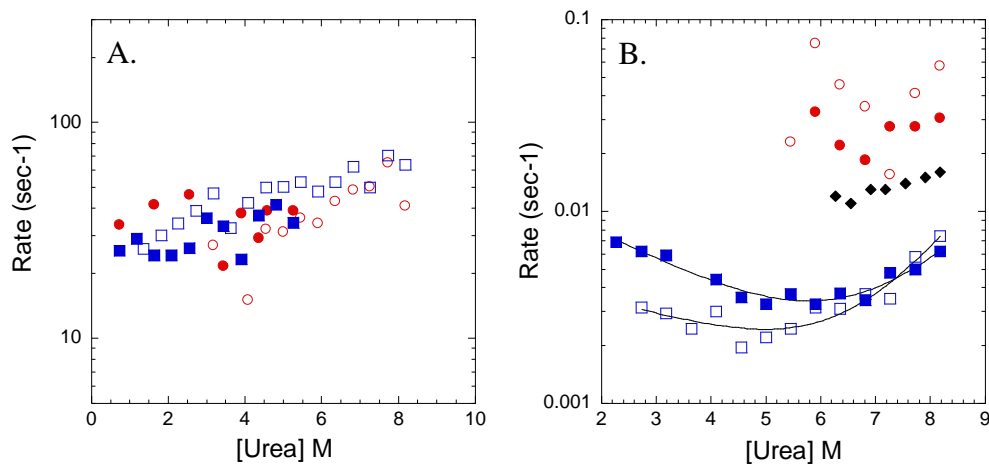
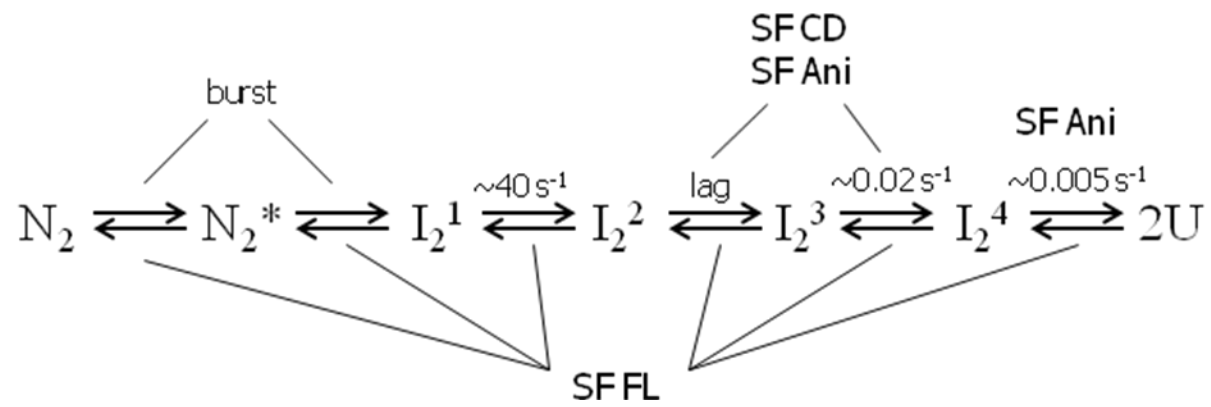


Figure 12: Unfolding rates. A) Fast phase of unfolding detected by fluorescence excitation at 280 (red, circles) and 295 (blue, squares) nm. Refolding data are represented by the closed symbols. Unfolding data are represented by the open symbols. B) Unfolding phases detected by fluorescence emission (excitation at 280 nm) (\circ , \square), anisotropy (excitation at 280 nm) (\bullet , \blacksquare) and CD 228 nm (\blacklozenge). The lines through the data represent fits to a two-state kinetic folding analysis. The values are listed in the text.

Figure 13: Proposed sequential unfolding mechanism. The native dimer (N_2 and N_2^*) unfolds through a series of dimeric intermediates (I_2^1 through I_2^2) before dimer dissociation. Listed above each arrow is the rate of formation of each species. The detection methods are listed; SF = stopped-flow, FL = fluorescence with excitation at 280 or 295 nm, CD = circular dichroism (228 nm), Ani = anisotropy with excitation at 280 or 295 nm.



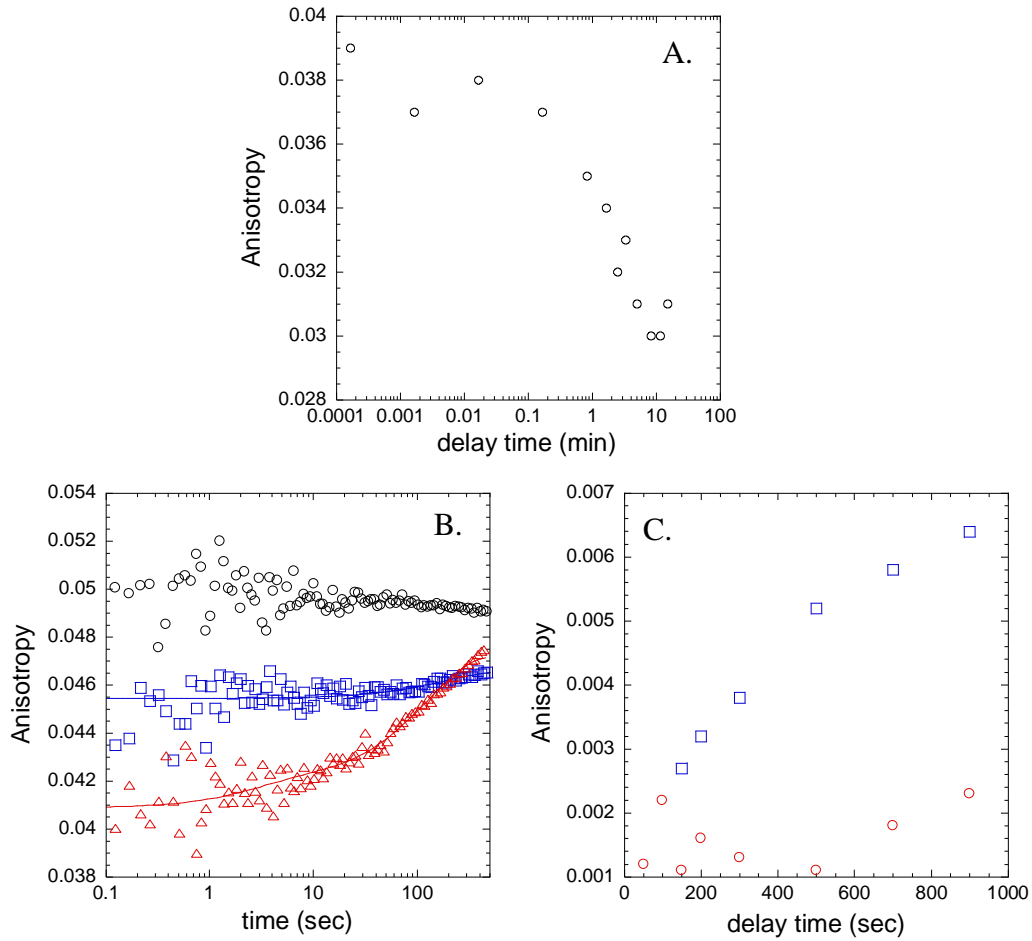


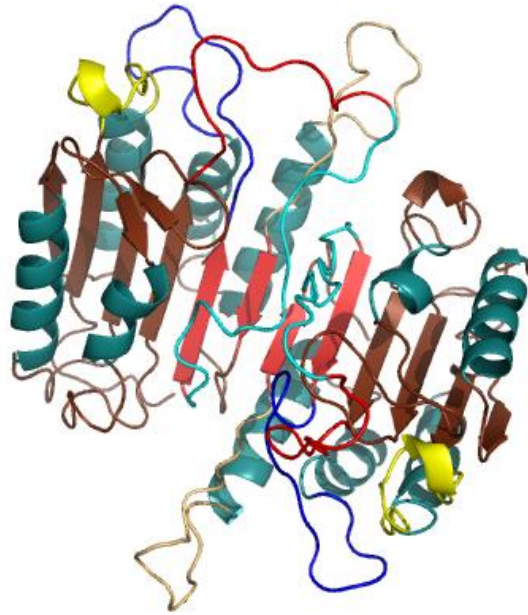
Figure 14: Double jump anisotropy studies. A) Burst phase signal versus the delay time. B) Raw data for 0.1 (black, \circ), 50 (blue, \square) and 500 (red, Δ) second delay times. The lines represent one (50 sec) and two (500 sec) exponential fits. C) Amplitudes of one and two exponential fits of double jump anisotropy data (as in panel B). The slow phase is represented by blue squares. The fast phase is represented by red circles.

```
      dimer interface
Caspase-3  --KQIPCIVSM--
Caspase-8  --KQMPQPAFT--
```

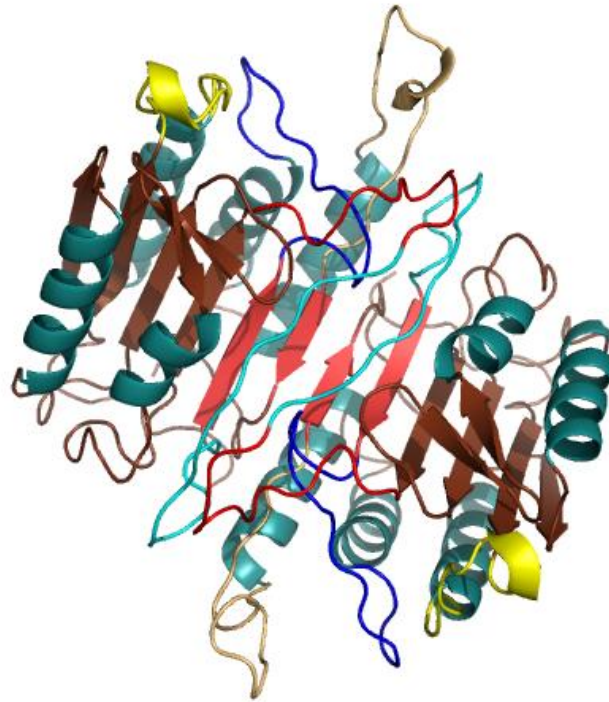
Figure 15: Dimer interface sequences of caspase-3 and -8.

Figure 16: Homology models of inactive (N_2 , panel A) and active (N_2^* , panel B) procaspase-3.

A.



B.



SUPPLEMENTAL FIGURES

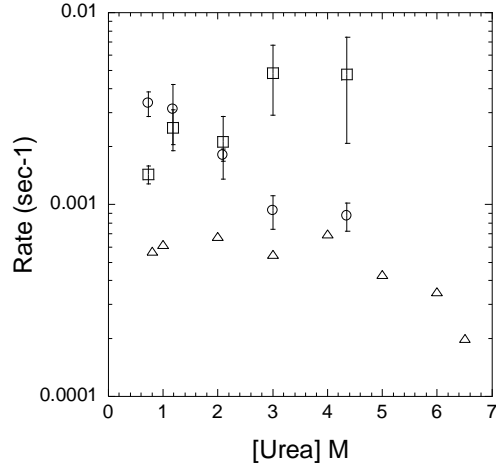


Figure 1. Slow phases in refolding. Hand-mixing far-UV CD slow rate (Figure 7A) versus urea (Δ). First (\circ) and second (\square) phase observed by hand-mixing differential quenching by acrylamide (Figure 7B).

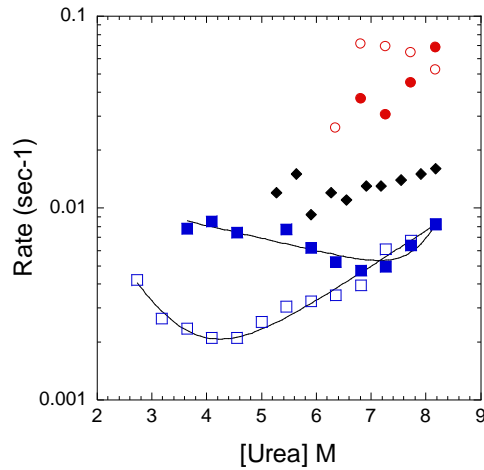


Figure 2. Unfolding rates. Unfolding phases detected by fluorescence emission (excitation at 295 nm) (\circ , \square), anisotropy (excitation at 295 nm) (\bullet , \blacksquare) and CD 228 nm (\blacklozenge). The lines through the data represent two-state kinetic folding analysis.

CHAPTER V

CONCLUSION

Investigating apoptosis is essential to understanding the pathology of a multitude of diseases, such as neurodegenerative disorders and cancer. The proteins examined here, Apaf-1 CARD and procaspase-3, are involved in both the intrinsic and extrinsic apoptotic pathways. In the intrinsic apoptotic pathway, Apaf-1 CARD mediates the interaction between the apoptosome and procaspase-9, leading to the activation of caspase-9. The activation of procaspase-3 is the final step in both the intrinsic and extrinsic apoptotic pathways. As these two proteins are key regulators of apoptosis, kinetic studies were performed to investigate the folding pathway of these proteins.

We have shown that Apaf-1 CARD folded and unfolded rapidly. Kinetic folding studies, as well as KINSIM simulations, suggested a folding mechanism that contains parallel folding channels with three unfolded conformations, two of which are able to fold into the native conformation. In contrast to other CARDS, such as RICK (RIP-like interacting CLARP kinase) and procaspase-1, the folding of Apaf-1 CARD does not reveal kinetic traps or misfolded species. These results suggest that there is not one unifying mechanism to folding for these α -Greek key proteins, which will allow for the design of specific targets for each protein.

We have shown that procaspase-3 folded through a complex pathway involving multiple monomeric species and slow dimerization. The data suggested that these monomeric intermediates were misfolded, which is the cause of slow dimer formation. *In vivo* these misfolded species can be corrected by chaperone proteins, allowing for procaspase-3 to dimerize quickly into its native dimeric state inside the cell. Future studies will focus on

comparison of the kinetic folding pathway of executioner caspases (ie procaspase-3) to initiator caspases (ie procaspase-8) in order to elucidate the differences between these two groups of caspases (ie the inactive zymogen's oligomeric state, prodomain function and maturation process). These studies will help uncover how the cell regulates apoptosis, leading to the discovery of new therapeutic targets to combat diseases.

The unfolding kinetic data of procaspase-3 revealed a slow rate of dimer dissociation. These data supported the equilibrium folding studies which reveal that the stability of procaspase-3 is largely due to dimerization. We have shown that multiple dimeric species were formed in the unfolding burst phase. These data support the hypothesis of two native conformations of the protein, one of which is enzymatically active. The current maturation model states that the intersubunit linker of executioner caspases (ie caspase-3) must be cleaved for full activation of the enzyme. These studies oppose this current dogma and suggest that procaspase-3 can be activated without cleavage of the intersubunit linker. It is possible that there are multiple mechanisms for regulation of procaspase-3 activation, which remain to be discovered. The proposal of two native conformations, active and inactive, of procaspase-3 implies a new target for therapeutic strategies to enhance apoptosis. Future studies will focus on the interconversion of the two species.

Overall, the kinetic folding data for Apaf-1 CARD and procaspase-3 provided an improved picture of the function and regulation of these proteins in apoptosis. These studies proposed new targets for the regulation of apoptosis.

APPENDICES

APPENDIX A

Apaf-1 CARD Supplemental Data

I. Analytical ultracentrifugation

Analytical ultracentrifugation experiments were conducted in order to determine the oligomeric state of Apaf-1 CARD. These experiments were performed as described previously (Pop et al., 2001) with the following exceptions. Apaf-1 CARD was dialyzed at 4 °C into 50 mM potassium phosphate and 50 mM MES, 100 mM Tris, 50 mM acetate pH 7.5 with 0.5 mM DTT. The absorbance was measured at 280 nm and samples with an absorbance of 0.1, 0.25 and 0.5 were prepared. The experimental data were collected with a Beckman XL-A ultracentrifuge at the University of North Carolina (Chapel Hill) Macromolecular Interactions Facility and fit using ORIGIN software.

The representative data are shown in Figure 1A and B. The molecular weights determined from the data fits were 13,1001 and 14, 231 for Apaf-1 CARD in 50 mM potassium phosphate and 50 mM MES, 100 mM Tris, 50 mM acetate buffer pH 7.5, respectively. The residuals signify that the data are best fit to the model. The molecular weight of Apaf-1 CARD is 11, 280, which is in good agreement with the AUC data. Therefore, Apaf-1 CARD is a monomer under the conditions used for the equilibrium and kinetic folding experiments.

REFERENCES

- Pop, C., Chen, Y.-R., Smith, B., Bose, K., Bobay, B., Tripathy, A., Franzen, S., and Clark, A. C. (2001). Removal of the pro-domain does not affect the conformation of the procaspase-3 dimer. *Biochemistry* **40**, 14224-14235.

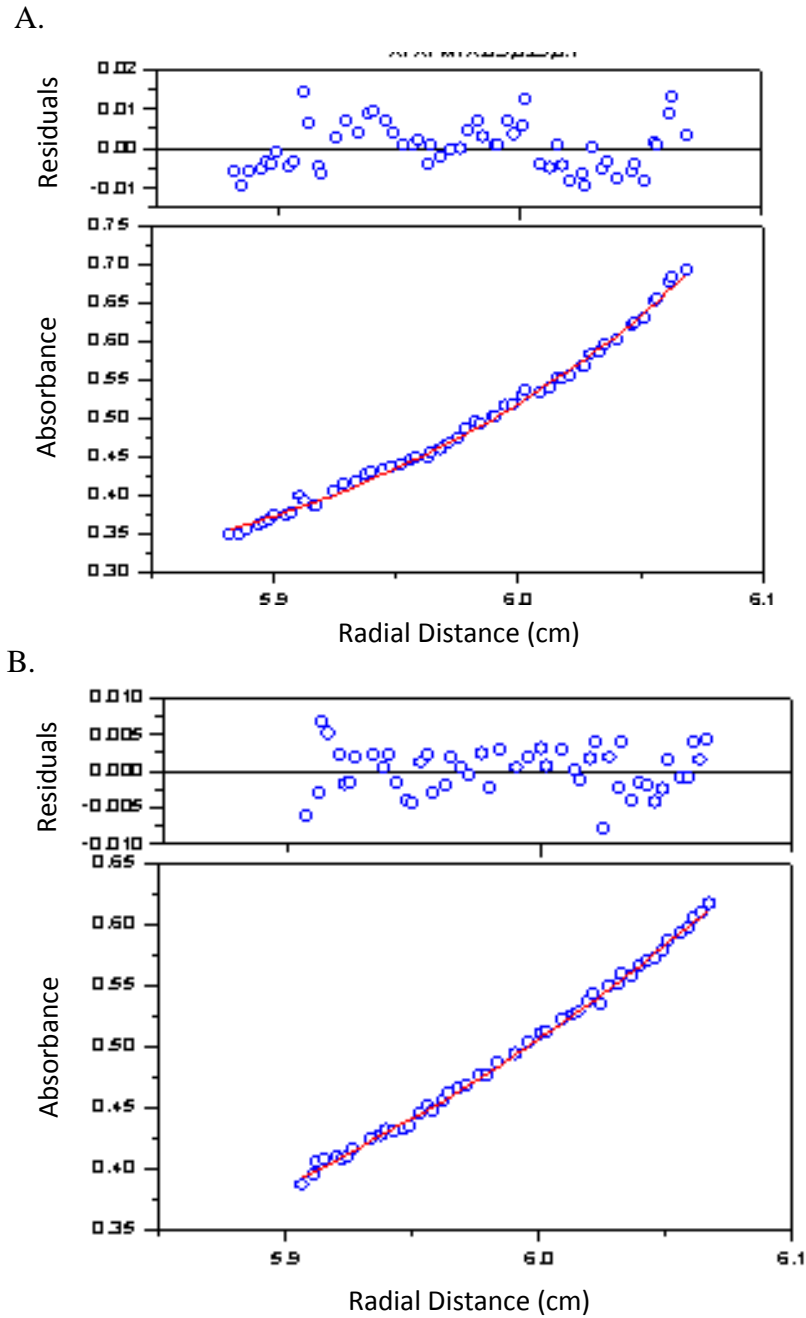


Figure 1: Sedimentation equilibrium studies of Apaf-1 CARD. The circles represent the experimental data for 50 mM MES, 100 mM Tris, 50 mM acetate pH 7.5 (panel A) or 50 mM potassium phosphate pH 7.5 (panel B). The solid lines are the theoretical plots assuming the protein is a monomer. Residuals to the fits are shown above each plot.

II. Circular dichroism spectroscopy

Far-UV (200-250 nm) and near-UV (250-300 nm) CD spectra were obtained for Apaf-1 CARD in order to examine the structural changes with different buffers and pH's. Circular dichroism spectra were measured using a J600A spectropolarimeter (Jasco Inc.) with either a 0.1 cm (far-UV) or 1 cm (near-UV) cell. The protein concentration was 97.7 μ M. The samples were incubated in 0 or 8 M urea-containing buffer (PB or MTA pH 6 or 7.5). All spectra were corrected for background signals. The temperature was kept constant at 25 °C using a circulating water bath.

The far-UV and near-UV spectra for Apaf-1 CARD are shown in Figure 2. The far-UV results show double minima at 206 nm and 218 nm, consistent with α -helical proteins. In the presence of 8 M urea-containing buffer, the signal decreased to that of a random coil, and the spectrum is consistent with that of an unfolded protein. In addition, the far- and near-UV spectra under native conditions reveal that the secondary and tertiary structures of Apaf-1 CARD are similar in both buffers at both pH's.

In order to verify that the protein was completely unfolded in 8 M urea-containing buffer, far- and near-UV spectra were collected using guanidinium hydrochloride (GdmHCl) as a denaturant (Figure 3). The far- and near-UV spectra of Apaf-1 CARD in 6 M GdmHCl-containing potassium phosphate buffer pH 7.5 reveal that the signal is decreased in GdmHCl, similar to the results obtained for the urea studies. Therefore Apaf-1 CARD is completely unfolded in 8 M urea-containing buffer.

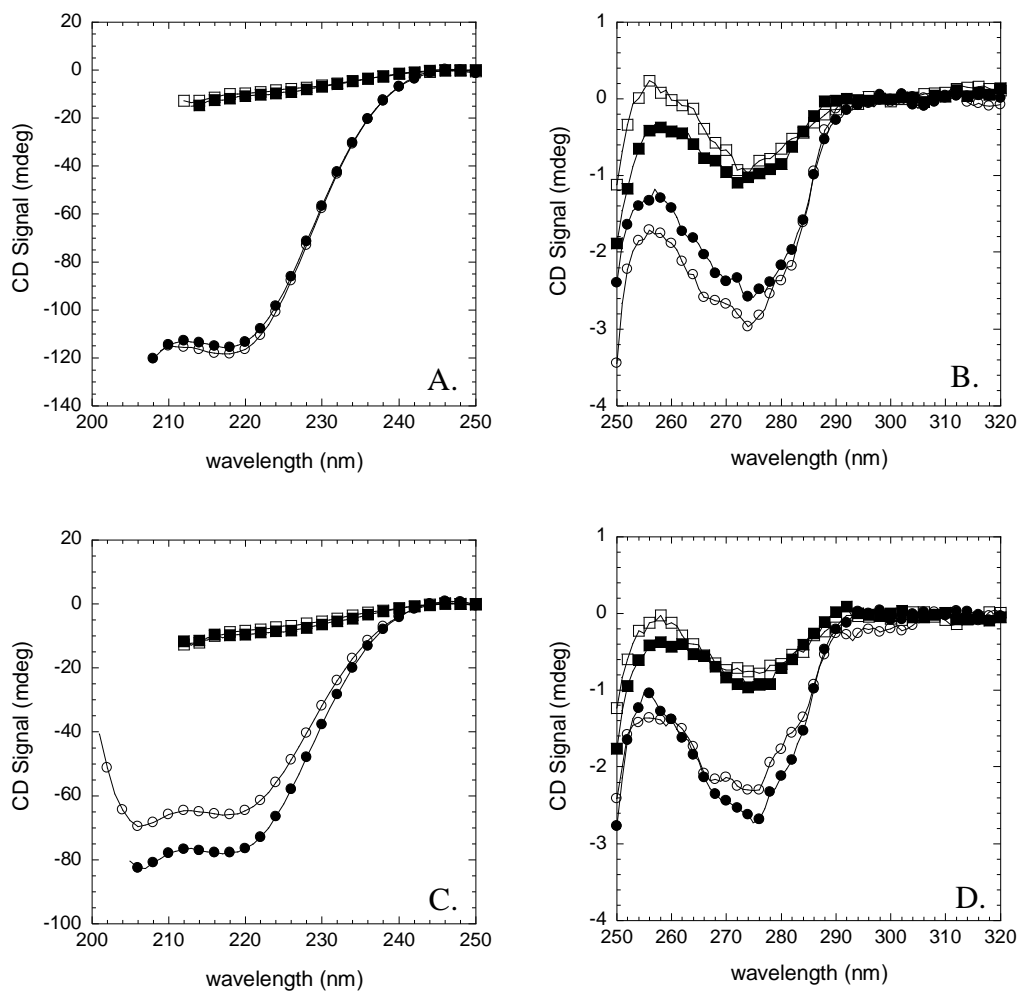


Figure 2: Far- and near-UV CD spectra of Apaf-1 CARD. A and C) Far-UV CD spectra in 0 M urea-containing buffer (PB = \circ or MTA = \bullet) and 8 M urea-containing buffer (PB = \square or MTA = \blacksquare). B and D) Near-UV CD spectra in 0 M urea-containing buffer (PB = \circ or MTA = \bullet) and 8 M urea-containing buffer (PB = \square or MTA = \blacksquare). Panels A and B are at pH 6. Panels C and D are at pH 7.5.

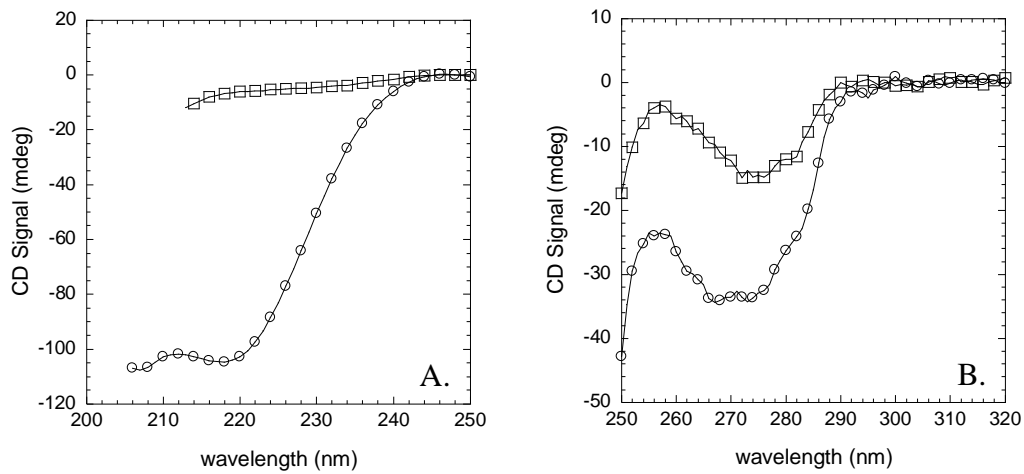


Figure 3: Far- and near-UV CD spectra of Apaf-1 CARD. A) Far-UV CD spectra in 0 M (\circ) and 6 M GdmHCl-containing potassium phosphate buffer (\square). B) Near-UV CD spectra in 0 M (\circ) and 6 M GdmHCl-containing potassium phosphate buffer (\square).

APPENDIX B

Procaspase-3 Supplemental Data

I. Chemical cross-linking

Cross-linking experiments were performed in order to confirm the dimerization step in the folding reaction. The rationale behind these experiments was that intermolecular cross-links would be more likely to form in the dimer rather than the monomer. The specific cross-linkers chosen for these experiments, DMS and BS³ (see below), react with primary amines (such as lysine residues and the N-terminus of proteins). Dimeric procaspase-3 contains 44 lysines residues, 4 of which are on the surface of the protein and within 11 Å to cross-link across the dimer interface, as shown in Figure 1.

Kinetic studies rely on initial experiments to determine the correct amount of cross-linker and the length of time to cross-link the protein. In order to avoid false positives (such as cross-linked dimers when the dimer is not yet formed), the cross-linking reactions must be quenched quickly, within seconds, as the protein is refolding during the cross-linking reaction. Therefore the basic experimental design for the initial parameterization is as follows. First the protein and cross-linker were mixed together for various amounts of time, followed by the addition of a quencher (1M Tris pH 7). These reactions were then placed on ice to further slow the cross-linking reaction. SDS-PAGE gels were then used to identify the cross-linked products.

Initial experiments were employed with the cross-linking reagent, DMS (dimethyl suberimidate), following previously determined conditions for procaspase-3 by a former graduate student. DMS (Figure 2A) is a water soluble, homobifunctional imidoester cross-linker. It reacts with primary amines at alkaline pH values (pH 8-10) (Figure 2B) to produce

an amidine linkage.

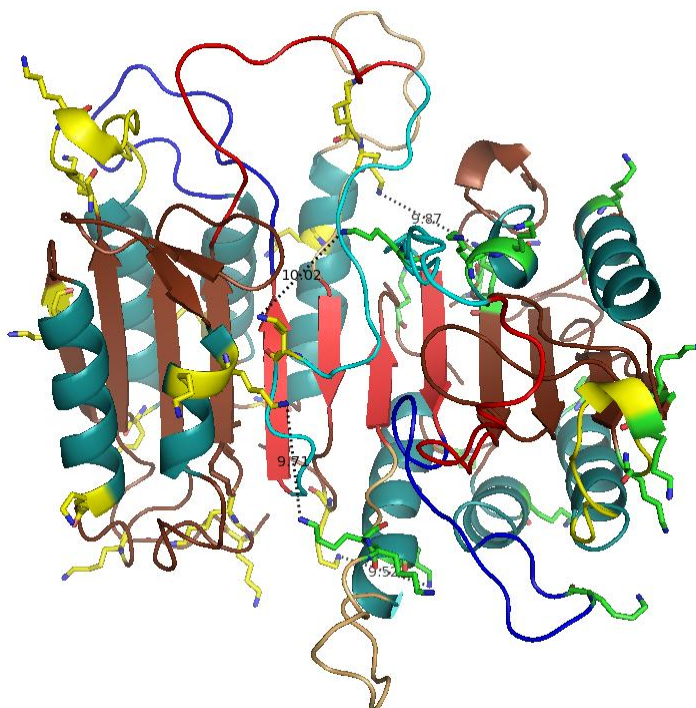


Figure 1: Homology model of procaspase-3. The lysine residues on each monomer are colored green or yellow. Based on this model, there are 4 lysines within 11 Å across the dimer interface (K186-K186, 10.02 Å; K260-K137, 9.71 Å; K242-K271, 9.52 Å; K156-K260, 9.87 Å).

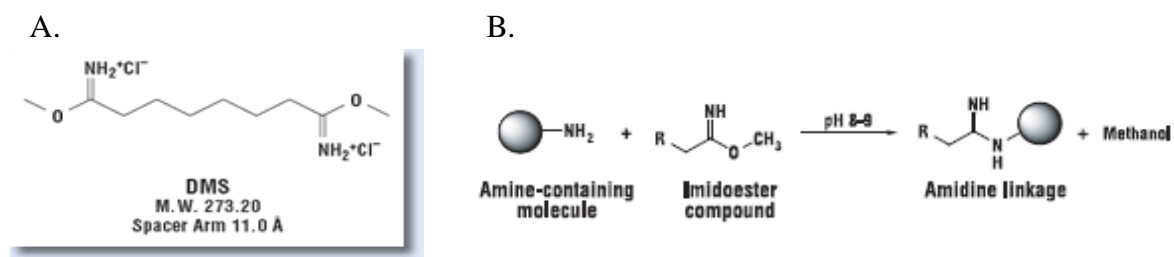


Figure 2: A) DMS cross-linker; B) DMS cross-linking reaction

Shown in Figure 3 are representative gels for the initial parameterization for DMS cross-linking reactions with procaspase-3 in 0.73 M urea-containing buffer pH 7.5. The protein was equilibrated in 0.73 M urea as a refolding / native control, previously determined as the endpoint for the folding reaction. These experiments were performed at cross-linker concentrations ranging from 5 to 1830 molar excess DMS to protein. Representative gels are shown here. Procaspase-3 was mixed with either 50 (panel A) or 1830 (panel B) molar excess of DMS for the times indicated in the figure. The reaction was then quenched with 1M Tris pH 7, followed by SDS-PAGE analysis.

The gels in Figure 3 display both monomer and dimer bands for both concentrations of cross-linker (50 and 1830 molar excess DMS to protein). As expected, the amount of dimer cross-linked with 50 molar excess of DMS (Figure 3A) is less than that cross-linked with 1830 molar excess DMS (Figure 3B). In addition, the intensity of the bands does not change over time indicating that the amount of protein cross-linked in short times (5 seconds) is the same as at longer times (2 hours). Theoretically, the amount of protein cross-linked should increase with longer cross-linking incubation times until the protein is 100% cross-linked, which is not the case for DMS. This could be due to the fact that DMS works best at high pH values (pH 8-10), while the protein in these experiments is in 0.73 M urea-containing buffer at pH 7.5.

As an alternative to DMS, BS³ is a cross-linker whose optimal activity fits within the experimental conditions. BS³ (Bis(Sulfosuccinimidyl) suberate) (Figure 4A) is a water soluble, homobifunctional NHS-ester (N-hydroxysuccinimide) cross-linker. It reacts rapidly

with any primary amine-containing molecule in pH 7-9 to form covalent amide bonds (Figure 4B).

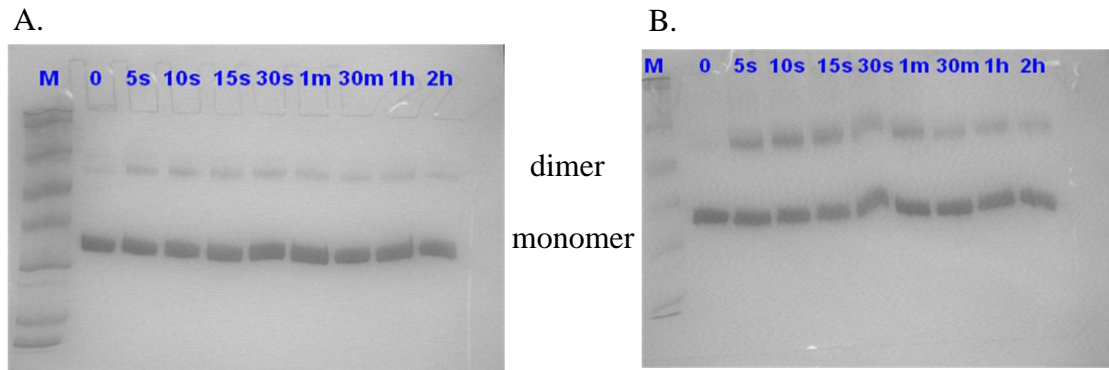


Figure 3: DMS initial parameterization. 8-15% SDS-PGE gel. Lanes: marker followed by incubation time with cross-linker (no cross-linker, 5 sec, 10 sec, 15 sec, 30 sec, 1 min, 30 min, 1 hr, 2 hr). Monomer and dimer bands are marked. A) 50 molar excess DMS; B) 1830 molar excess DMS

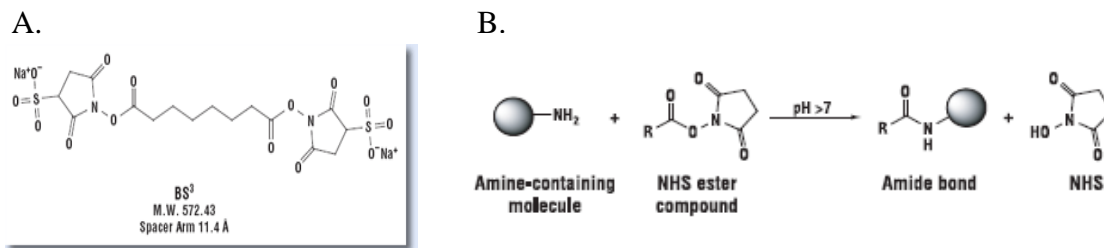


Figure 4: A) BS³ cross-linker; B) BS³ cross-linking reaction

Shown in Figure 5 are representative gels for the initial parameterization for BS³ cross-linking reactions with 20 μ M procaspase-3 in 0.73 and 8 M urea-containing buffer pH

7.5 corresponding to the native and unfolded protein, respectively. These experiments were performed at cross-linker concentrations ranging from 10 to 150 molar excess BS^3 to protein. Representative gels are shown below. Procaspase-3 was mixed with either 10 (panel A) or 100 (panel B and C) molar excess of BS^3 for the times indicated in the figure. The reaction was then quenched with 1M Tris pH 7, followed by SDS-PAGE analysis.

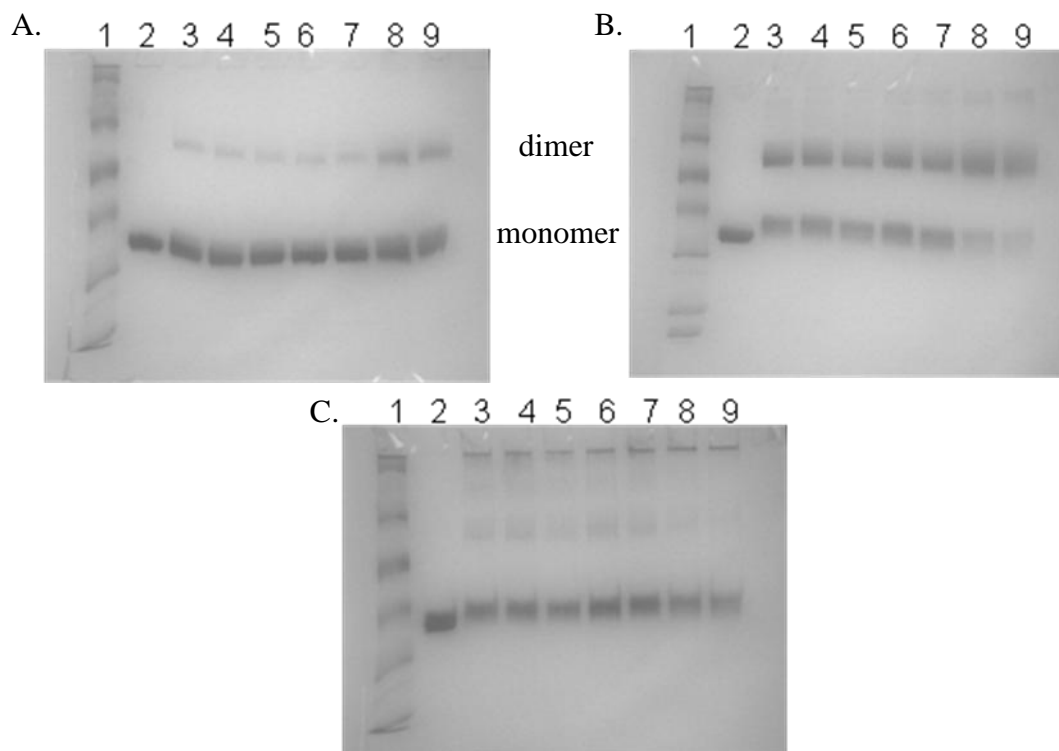


Figure 5: BS^3 cross-linking control reactions. Times listed are incubation times. 8-15% SDS-PAGE. Lanes: 1 - marker, 2 - no BS^3 , 3 - 5 sec, 4 - 10 sec, 5 - 30 sec, 6 - 45 sec, 7 - 1 min, 8 - 30 min, 9 - 1 hr. Monomer and dimer bands are marked. A) 10 fold molar excess of BS^3 to protein in 0.73 M urea-containing buffer; B) 100 fold molar excess BS^3 to protein in 0.73 M urea-containing buffer; C) 100 fold molar excess BS^3 to protein in 8 M urea-containing buffer

Figure 5A and B, SDS-PAGE gels for the native protein, display the migration of both the monomer and dimer. As expected, the intensity of the dimer band increases with increasing cross-linker incubation time, while the intensity of the monomer band decreases. These results confirm that the dimer of procaspase-3 can be detected by this technique. The SDS-PAGE gel for unfolded protein reveals that there is a background level of cross-linking in 8 M urea-containing buffer. From these data, the cross-linker conditions were chosen. Even though 100 fold cross-linker molar excess (Figure 5B) displays a greater shift in the intensity of the monomer and dimer bands at short incubation times, the data also display higher order species. Therefore lower concentrations of cross-linker were used in the refolding experiments.

For the refolding reaction, 126.3 μM unfolded protein in 8 M urea-containing buffer was mixed 1:11 with buffer, which diluted the final urea concentration to 0.73 M and allowed the protein (11.5 μM final) to refold. Aliquots of the refolding reaction were mixed with BS^3 for 5 seconds. Then the cross-linking reaction was terminated with 1 M Tris pH 7 and analyzed by SDS-PAGE. Figure 6 shows representative gels for the refolding reaction with 25 (panels A and B) or 10 (panels C and D) molar excess of BS^3 .

The native protein control and BS^3 (lanes 3, 8, 2 and 9 for Figure 3A, B, C and D, respectively) confirm that the procaspase-3 dimer is able to be cross-linked with an incubation time of 5 seconds. As was seen with the cross-linking controls (Figure 5), oligomeric states of the protein are detected with the higher concentration of cross-linker. All of the lanes show both monomer and dimer bands, which do not change in intensity over

time (from refolding for 5 seconds to refolding overnight), indicating that the protein dimerizes very quickly.

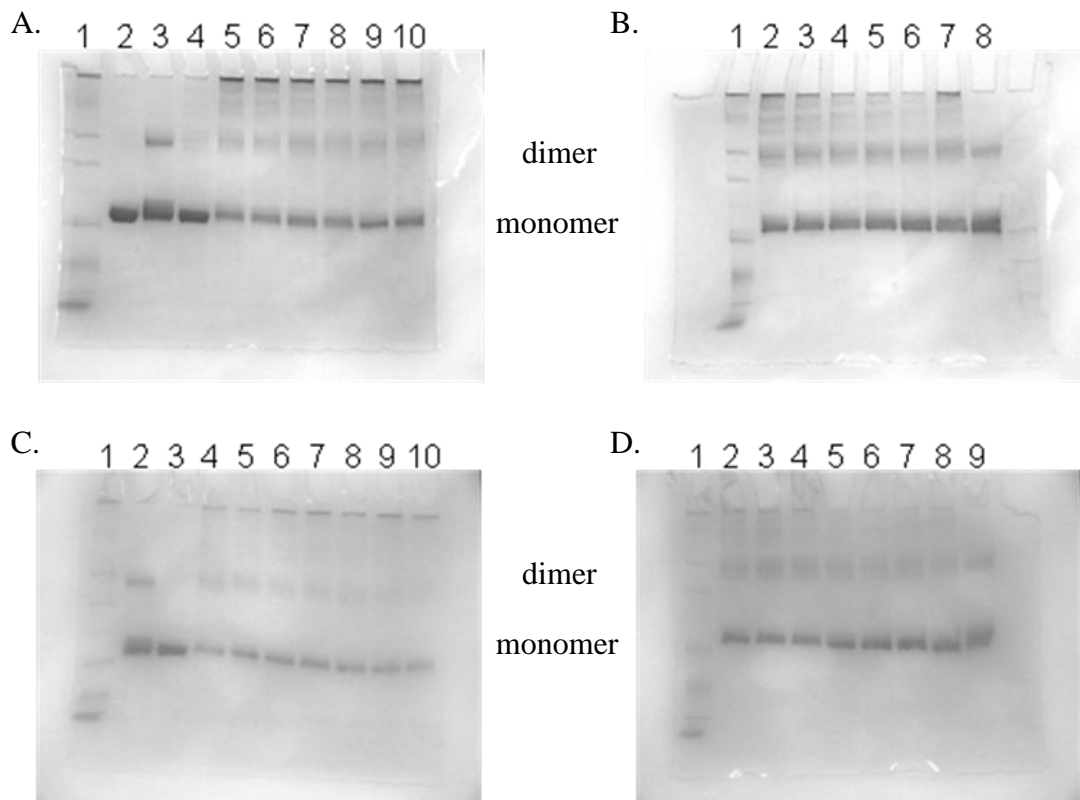


Figure 6: Refolding reactions. Monomer and dimer bands are marked. 8-15% SDS-PAGE. Panels A and B: 25 molar excess of BS³. Panels C and D: 10 molar excess of BS³. Times listed are refolding times. A) Lanes: 1 – marker, 2 – native protein, 3 – native protein control and BS³, 4 – 5 sec, 5 – 15 sec, 6 – 3- sec, 7 – 45 sec, 8 – 1 min, 9 – 1.5 min, 10 – 2 min. B) Lanes: 1 – maker, 2 – 3.3 hr, 3 – 3.5 hr, 4 – 3.7 hr, 5 – 3.8 hr, 6 – 4 hr, 7 – overnight, 8 – native protein and BS³. C) Lanes: 1 – marker, 2 – native protein and BS³, 3 – native protein, 4 – 5 sec, 5 – 10 sec, 6 – 15 sec, 7 – 20 sec, 8 – 25 sec, 9 – 30 sec, 10 – 40 sec. D) Lanes: 1 – marker, 2 – 2.2 hr, 3 – 2.3 hr, 4 – 2.5 hr, 5 – 2.7 hr, 6 – 2.8 hr, 7 – 3 hr, 8 – overnight, 9 – native protein plus BS³.

In order to confirm these results, a control experiment was conducted. Multiple quenchers (1 M Tris pH 7.5 and 8.5 and SDS loading dye) were tested to verify the complete termination of the cross-linking reaction. The controls for this experiment are native protein alone (which should display a monomer band) and native protein incubated with BS³ for 1 hr, the recommended incubation time by the manufacturer (which should display both monomer and dimer bands). 100 μM BS³ (10 fold molar excess) was mixed with quencher alone. After 5 seconds, 10 μM procaspase-3 was added. SDS-PAGE was used to verify the cross-linked products (Figure 7). Assuming the quencher binds to all available cross-linker, only a monomer band for the protein should be detected. However, if the quencher does not sequester all the cross-linker from the solution, BS³ will still be able to bind to the protein, resulting in dimer bands on the gel.

As expected, native protein (#1) only displays a monomer band and the 1 hr incubation time (#3) displays both monomer and dimer bands for all three quenchers. The results for the Tris quencher (#2, first and second column) at pH 7.5 or 8.5 display both monomer and dimer bands, which indicate that these quenchers do not fully sequester the cross-linker from the protein. Therefore all the results from the previous refolding cross-linking experiments, in which Tris was used as the quencher, are most likely due to background levels of cross-linking. However, SDS-loading dye was able to terminate the cross-linking reaction (#2, third column) since only the monomer band is shown for this reaction. Thus, SDS-loading dye was used to determine the conditions in which refolding cross-linking experiments could be performed (Figure 8).

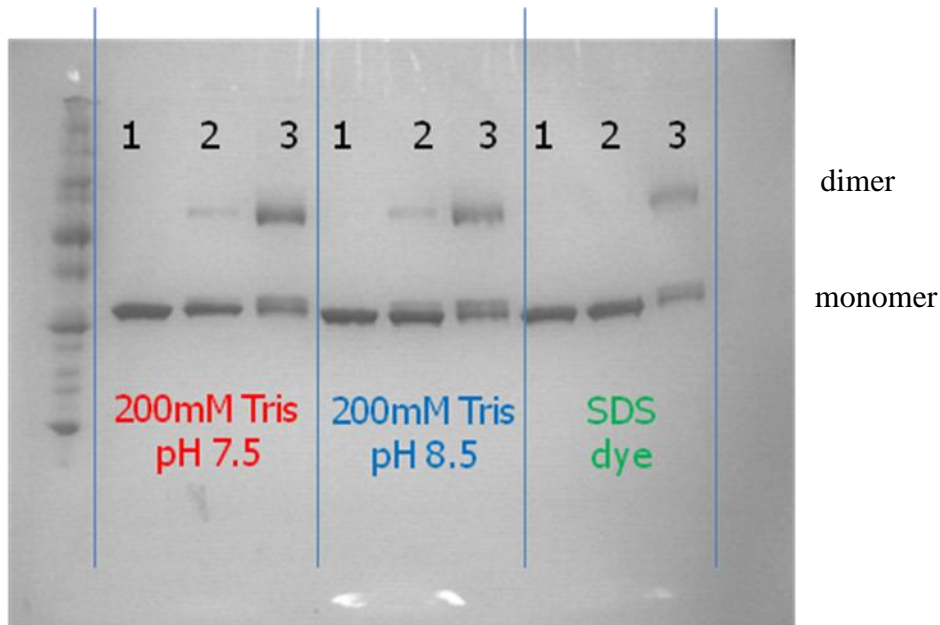


Figure 7: 4-25% SDS-PAGE. Lanes: 1 – native protein; 2 – incubate BS³ and quencher for 5 seconds, then add protein; 3- native protein and BS³ incubated for 1 hr

Initial parameterization was conducted in the same manner as described above, except using SDS to quench the reaction. Briefly, 10 μM native procaspase-3 in phosphate buffer pH 7.5 was mixed with BS³ concentrations ranging from 10 to 500 molar excess to protein for the incubation times listed in Figure 8, followed by the addition of SDS-loading dye to quench the cross-linking reaction.

As predicted, the SDS-loading dye quencher was able to eliminate the false positives from the cross-linking reaction. The low BS³ concentration, 100 μM , only cross-linked the procaspase-3 dimer at longer incubation times (>15 min, lane 9 Figure 8A). 5000 μM BS³ cross-linked the procaspase-3 dimer at shorter incubation times (~1 min, lane 8 Figure 8B),

completely cross-linking all the protein by 30 min (as shown by no monomer band in lane 10, Figure 8B). However, in order to use this technique to examine dimer formation in the folding process of procaspase-3, BS^3 must be able to cross-link the protein in a short time frame. Otherwise, the data will not accurately correspond to dimerization. Therefore, these concentrations of cross-linker will not be able to detect dimer formation in the time frame necessary for this experiment. Another option would be to increase the BS^3 concentration. However, this option would use an abundant amount of cross-linker and is therefore not a feasible experiment. Therefore, cross-linking experiments are not a viable technique to study the rate of procaspase-3 dimerization.

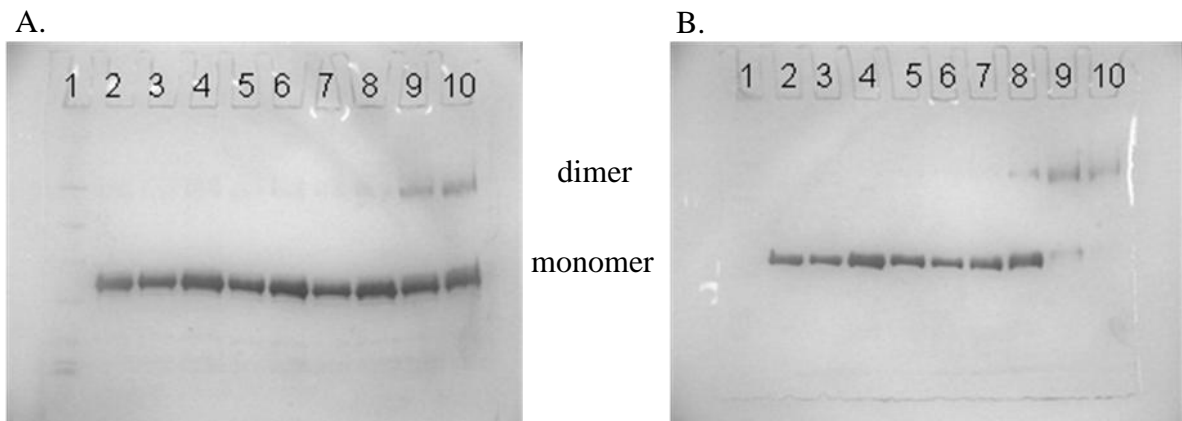


Figure 8: 8-15% SDS-PAGE. Monomer and dimer bands are marked. Lanes: 1 – marker, 2- native protein, 3 – incubate BS^3 and quencher for 5 seconds, then protein was added, 4 – 5 sec, 5 – 10 sec, 6 – 20 sec, 7 – 45 sec, 8 – 1.5 min, 9 – 15 min, 10 – 30 min. A) 10 fold molar excess of BS^3 . B) 500 fold molar excess of BS^3

II. Native gels

The rate of dimerization was accessed by examining a refolding reaction on a native gel. First different protein concentrations (5, 10 and 20 μM protein) were tested in order to determine the optimal amount of protein needed to visualize the results. As shown in Figure 9A, 20 μM procaspase-3 displayed the most visual bands and was therefore used in future experiments.

Next an equilibrium urea titration was performed in order to examine the shift in the monomer and dimer bands on the gel. As shown in Figure 9B, 0, 1 and 2 M urea-containing buffer displayed 2 bands, whereas the higher urea concentrations (>3 M) displayed a very high molecular weight band as well as smears. There were also very faint bands in the high urea concentrations around the same location as the 0-2 M samples. However, the protein species that exist in higher urea concentrations might not be populated in the refolding experiment since it is performed in 0.8 M urea-containing buffer. Therefore, a refolding experiment was conducted and the results examined on a native gel.

For the refolding reaction, 180 μM unfolded protein in 8 M urea-containing buffer was mixed in a 1:10 ratio with buffer. Aliquots of the refolding reaction were added to SDS-loading dye and placed on ice in order to stop the reaction. The samples were then analyzed on a native gel. As shown in Figure 9C, the native protein control was the only visible band on the gel. All other lanes displayed a very high molecular weight band (even after refolding overnight, in which it is known that the protein is completely folded by then). Therefore,

these experiments are not a viable technique to study the rate of dimerization of procaspase-3.

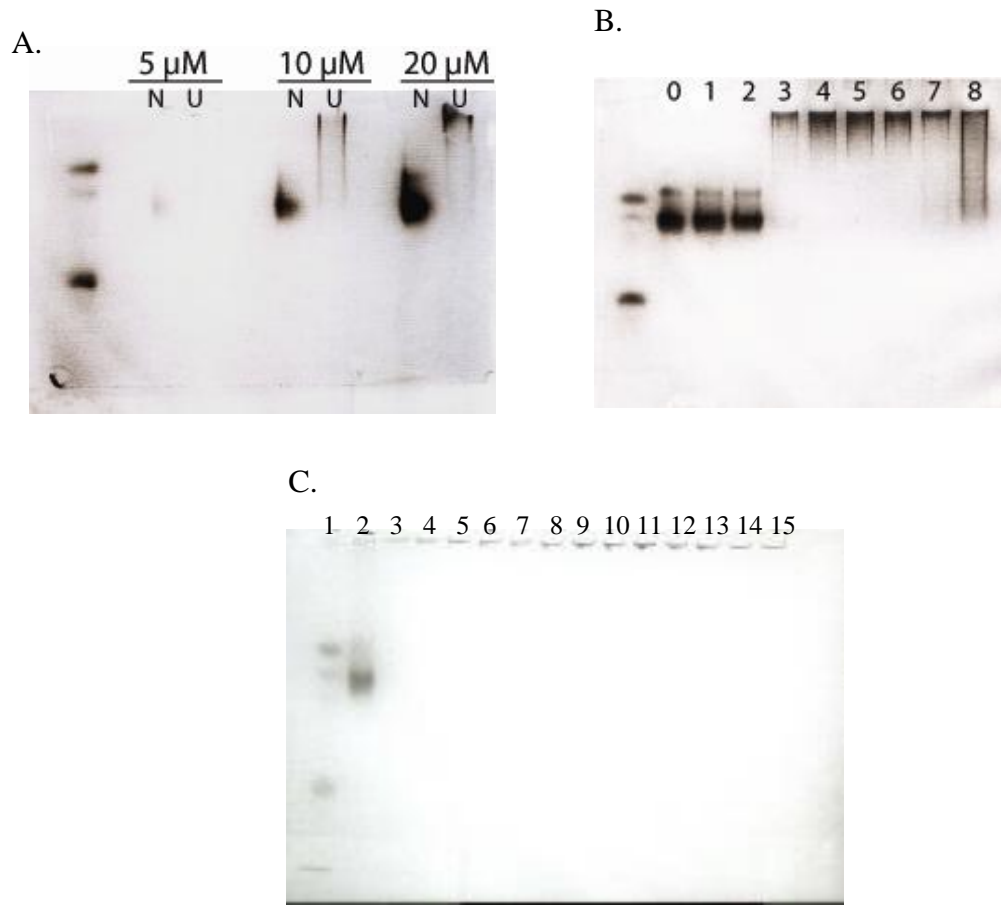


Figure 9: Native PAGE analysis A) Initial parameterization N = native protein in 0.73 M urea-containing buffer, U = unfolded protein in 8 M urea-containing buffer. Three protein concentrations of 5, 10 and 20 μ M are marked. B) Equilibrium urea titration experiment. The numbers (0-8) listed above the lanes indicate the concentration of urea. The protein concentration was 20 μ M. The samples were incubated over night. C) Refolding experiment. Times listed are refolding times. Lanes: 1 – marker, 2 – native protein in 0.8 M urea-containing buffer, 3 – 30 sec, 4 – 2.5 min, 5 – 5 min, 6 – 15 min, 7 – 30 min, 8 – 45 min, 9 – 1 hr, 10 – 1.5 hr, 11 – 2 hr, 12 – 3 hr, 13 – 4 hr, 14 – 5 hr, 15 – overnight refolding.

III. Procaspase-3 labeling with dansyl chloride

The purpose of these experiments was to determine if anisotropy could be used to monitor the formation of monomeric and dimeric procaspase-3. A previous graduate student had successfully labeled the pro-peptide with dansyl chloride and used our steady-state fluorometer to detect binding to procaspase-3 (Pop et al., 2001). Therefore, these studies focus on labeling procaspase-3 with dansyl chloride and measuring the anisotropy of the protein versus urea concentration.

The labeling reaction was performed as described previously (Pop et al., 2001), with some exceptions. 10 mg/mL procaspase-3 was incubated with 1 mg/mL dansyl chloride for 1 hr at 4 °C. Absorbance measurements at 340 nm (for dansyl chloride) and a Bradford assay (for procaspase-3) were completed to ascertain the number of dansyl labeled amino acids per procaspase-3. From multiple labeling experiments, it was determined that procaspase-3 was labeled once to twice per monomer. These results were verified by mass spectrometry analysis (Dr. Goshe lab, North Carolina State University).

Equilibrium unfolding experiments were performed with dansyl labeled procaspase-3 and the data were monitored by anisotropy. The results are shown in Figure 10. The data from the urea denaturation experiment (Figure 10A) revealed protein concentration dependence between 4 and 8 M urea. However, due to the absence of a post-transition, equilibrium unfolding experiments were performed with GdmHCl since it is a stronger denaturant. These studies (Figure 10B) demonstrate protein concentration dependence between 1 and 2.5 M GdmHCl. Equilibrium unfolding experiments (monitored by CD and

fluorescence) reveal that the protein is unfolded by 7.5 M urea (Bose and Clark, 2001). These results indicate the dansyl labeled procaspase-3 is more stable than unlabeled protein.

Therefore, the label must be interacting with the protein somehow to increase its stability.

Future anisotropy studies were not performed with dansyl labeled procaspase-3.

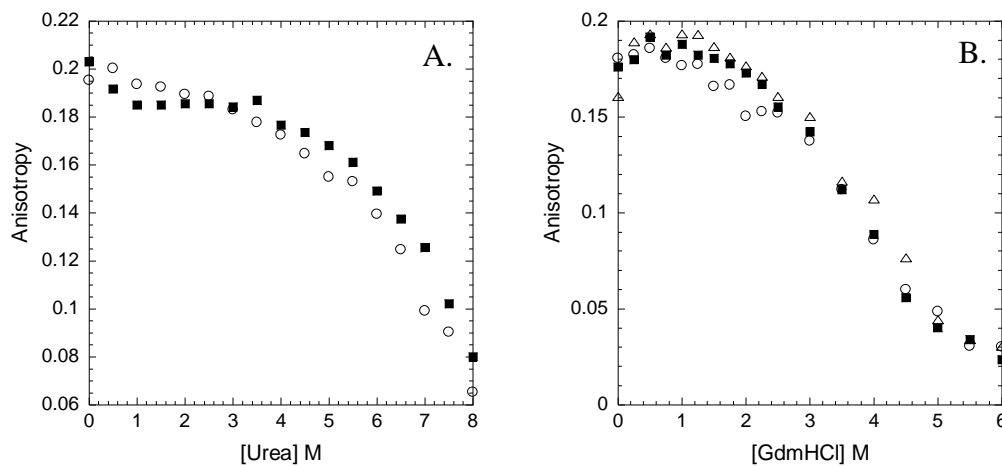


Figure 10: Equilibrium unfolding experiments of dansyl labeled procaspase-3 monitored by anisotropy. A) Equilibrium unfolding in urea of 2 μM (\circ) and 4 μM (\blacksquare) procaspase-3 was measured by anisotropy (excitation 345 nm, emission 450 nm). B) Equilibrium unfolding in GdmHCl of 2 μM (\circ), 4 μM (\blacksquare) and 6 μM (\triangle) procaspase-3 was measured by anisotropy (excitation 345 nm, emission 450 nm).

REFERENCES

Bose, K., and Clark, A. C. (2001). Dimeric procaspase-3 unfolds via a four-state equilibrium process. *Biochemistry* **40**, 14236-14242.

IV. Anisotropy

Anisotropy studies were performed to determine the rate of procaspase-3 dimerization. Initial studies were conducted with a SPEX Fluorolog-3 Research T-format Spectrofluorometer at the Macromolecular Interactions Facility at the University of North Carolina (Chapel Hill). The purpose of these experiments was to determine if there was a difference in anisotropy between the native dimer and unfolded monomer forms of procaspase-3 using the intrinsic fluorescence of the protein. Equilibrium samples at two protein concentrations (5 and 15 μM) were prepared for the native and unfolded protein in 0.72 and 8 M urea-containing buffer, respectively. Excitations at 280 and 295 nm (Figure 11A) revealed a change in anisotropy between the native dimer (0.72 M urea) and unfolded protein (8 M urea). This graph also displays the anisotropy of the dansyl labeled procaspase-3 (see Section III). These data support the conclusion that the label affected the protein, as the native sample (Figure 11A) had a much higher anisotropy than the anisotropy determined from the intrinsic fluorescence of the protein experiments.

Since the equilibrium data revealed a change in anisotropy between the native and unfolded forms of the protein, continuous and discontinuous kinetic refolding assays were performed in order to determine the rate of protein dimerization. These experiments were conducted at multiple protein concentrations (5 and 15 μM), as well as excitation at both 280 and 295 nm. Unfolded protein was manually mixed with buffer, and the anisotropy was collected continuously or at various time intervals. The dead time of these measurements was about 20 – 30 seconds. The data, from both methods, overlay with the anisotropy value of the

native protein as determined from equilibrium experiments (Figure 11B). These data indicate that the protein reaches the anisotropy of the native dimer within the first 20 – 30 seconds of the refolding reaction. Anisotropy measurements were then performed on a stopped-flow in order to examine the short times in the refolding reaction.

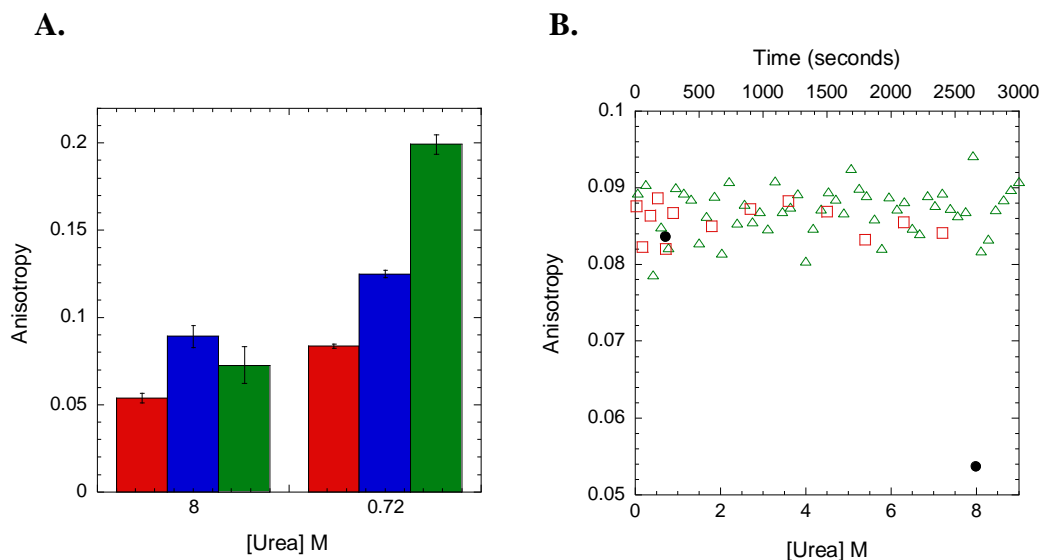


Figure 11: Anisotropy results. A) Equilibrium anisotropy of procaspase-3 in 0.72 and 8 M urea. Intrinsic protein fluorescence was used to determine anisotropy with excitation at 280 nm (red) or 295 nm (blue) with emission at 335 nm. Dansyl-labeled procaspase-3 (green) was excited at 345 nm with emission at 450 nm (see Section III). The error bars represent the standard deviation from two protein concentrations (5 and 15 μ M for intrinsic fluorescence and 2 and 4 μ M for dansyl-labeled protein). B) Continuous and discontinuous kinetic anisotropy studies. Equilibrium anisotropy for native and unfolded procaspase-3 with excitation at 280nm, emission 335 nm (●). Anisotropy of the refolding protein (excitation 280 nm, emission 335 nm) was determined either by a continuous (5 μ M protein, green Δ) or discontinuous (15 μ M protein, red \square) assay.

As stated in Chapter III, stopped-flow anisotropy refolding studies revealed a

cooperative burst phase versus urea. Unfolded protein in 8 M urea-containing buffer was mixed with various urea concentrations and the burst phase signal plotted against the final urea concentration. Multiple independent stopped-flow anisotropy experiments were performed at various protein concentrations (3.25, 5 and 15 μ M) in order to determine the protein concentration dependent step in the folding reaction. The anisotropy burst phase signal is not protein concentration dependent (Figure 12A), suggesting that the species formed in the burst phase was monomeric. On the other hand, the burst phase signals revealed a cooperative transition that is different from the anisotropy of the unfolded monomer. In addition, the burst phase signal for the refolding protein at low urea concentrations was similar to the anisotropy of the native dimer (Figure 12B and C). These data suggest that the dimer was formed in the burst phase.

To confirm these results, equilibrium unfolding anisotropy studies were performed on the SX18 stopped-flow. The dansyl label affects the stability of procaspase-3. Therefore, the intrinsic fluorescence of the protein was used to monitor anisotropy. The SX18 has a polarization accessory, which accommodates monitoring intrinsic fluorescence in house. Equilibrium unfolding samples (5 mL) were prepared at various final urea concentrations. A 1:1 mixing ratio was used on the stopped-flow and data were collected with excitation at 280 or 295 nm, emission with a cut-off filter at 305 nm.

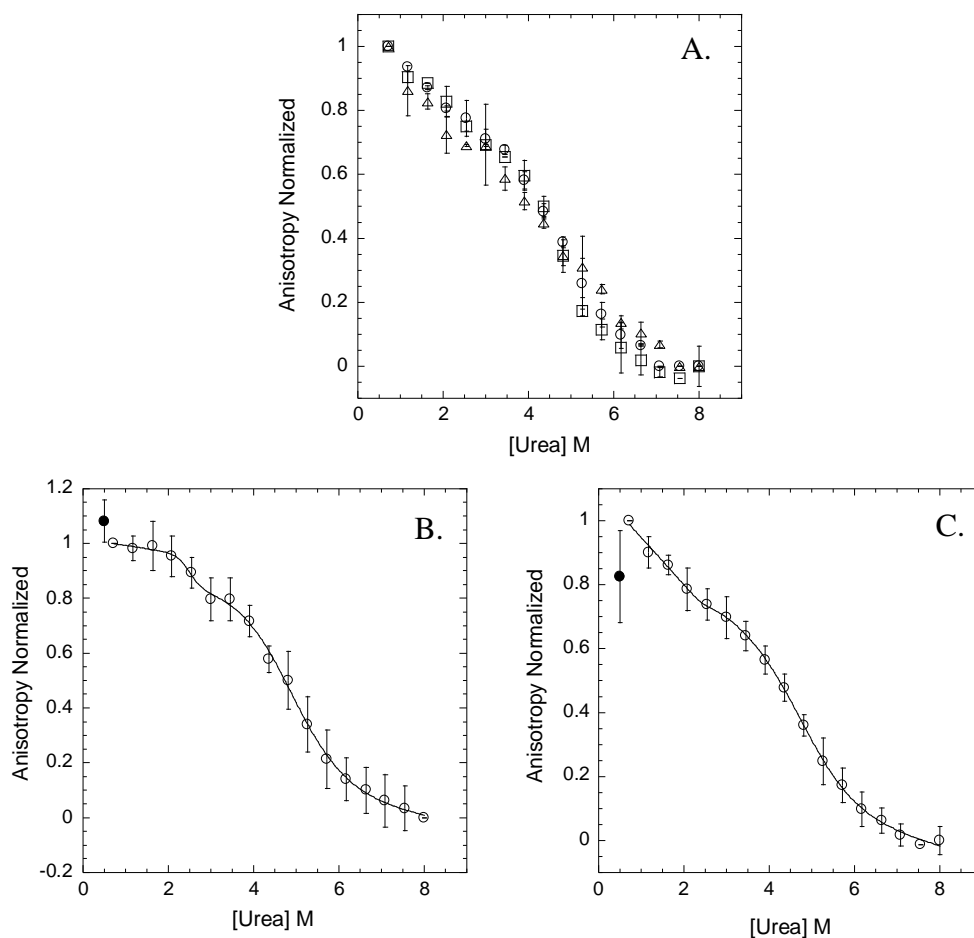


Figure 12: Stopped-flow anisotropy burst phase studies. A) Refolding burst phase with excitation at 295 nm versus several final urea concentrations for 3.25 (\square), 5 (\circ) and 15 (Δ) μ M procaspase-3. Average refolding burst phase (\circ) with excitation 280 nm (panel B) or 295 nm (panel C). The anisotropy of the native dimer is shown (\bullet). The error bars represent 3 protein concentrations and 2 independent experiments.

The equilibrium unfolding anisotropy results are shown in Figure 13. The data demonstrate protein concentration dependence between 5 and 8 M urea with excitation at both 280 and 295 nm. These results are in agreement with dimer dissociation found by

equilibrium unfolding studies with fluorescence and CD studies (Bose and Clark, 2001). In addition, the results indicate that if dimerization was occurring in the burst phase, anisotropy studies would display protein concentration dependence. Therefore, steady-state anisotropy results support the hypothesis that the burst phase species in anisotropy are monomeric.

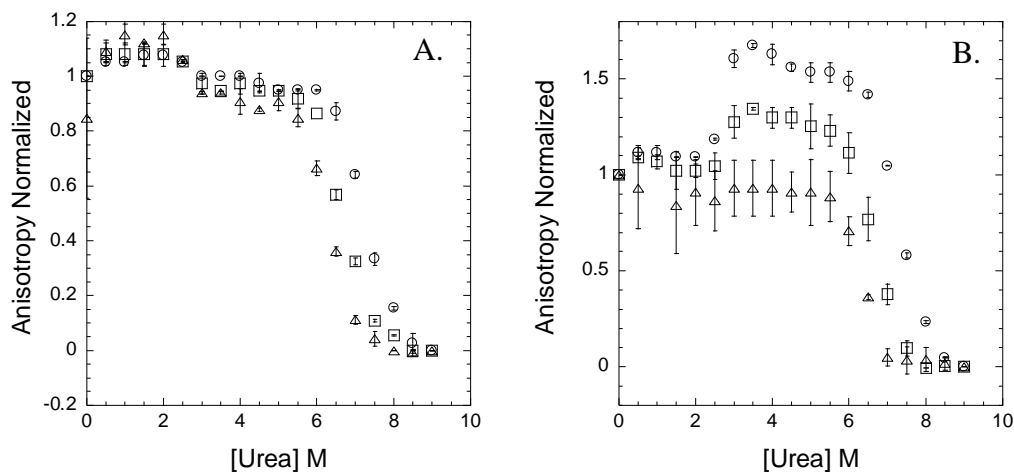


Figure 13: Equilibrium unfolding detected by anisotropy with excitation at 280 nm (panel A) or 295 nm (panel B). 2 μ M (Δ), 5 μ M (\square) and 15 μ M (\circ) procaspase-3. Error bars represent the standard deviation for two independent experiments.

REFERENCES

Bose, K., and Clark, A. C. (2001). Dimeric procaspase-3 unfolds via a four-state equilibrium process. *Biochemistry* **40**, 14236-14242.

V. Collaboration with David A. Johnson at University of California, Riverside

The main purpose of this collaboration was to determine the lifetime of the monomeric and dimeric forms of procaspase-3. Similar lifetimes of monomeric and dimeric procaspase-3 would indicate that the differences observed in the steady-state anisotropy are significant. However, if the lifetimes of these two forms were different, the change in steady-state anisotropy would be an artifact. Time-resolved anisotropy studies were performed on the native dimer and unfolded monomer of procaspase-3. This technique measures the fluorescence intensity / anisotropy decay associated with one light pulse, as opposed to continuous illumination with steady-state anisotropy.

The time dependence of fluorescence anisotropy was determined by the single-photon-counting technique (Hanson et al., 1981) with an EEY scientific nanosecond spectrofluorometer (La Jolla, CA). The vertically and horizontally polarized fluorescence decays were collected by exciting samples with vertically polarized light (295 nm) while orienting the emission polarizer in a vertical (I_{VV}) or horizontal (I_{VH}) direction. Excitation and emission bands were selected with a Glan Thompson polarizer and filter (#59460), respectively.

Native and unfolded procaspase-3 samples were prepared. The fluorescence intensity decay data, $I(t)$, were best fit by a sum of two exponentials:

$$I(t) = \alpha_1 \exp\left(\frac{-t}{\tau_1}\right) + \alpha_2 \exp\left(\frac{-t}{\tau_2}\right) \quad (1)$$

described as the intensity (I), the pre-exponential factors (α_1 and α_2) and fluorescence lifetime

(τ_1 and τ_2). Time-resolved anisotropy analysis of the native dimer and unfolded monomer of procaspase-3 revealed similar amplitude weighted-averaged lifetime values ($\langle\tau\rangle$) of 3.26 ± 1.5 nsec and 3.41 ± 1.5 nsec, respectively. Therefore, the steady-state anisotropy results cannot be explained by a change in the lifetime of the protein. In other words, the difference in anisotropy between the native dimer and unfolded monomer is authentic.

Further collaboration with David A. Johnson focused on identifying the location of the two tryptophan residues in the active site (*ie* buried or solvent exposed). The homology model of procaspase-3 (Figure 14) suggests that W214 is buried in the active site, whereas W206 is solvent exposed in each monomer. Based on the homology model, W214 is stabilized by its location at the end of a helix, implying this residue is less flexible than the other active site residue W206, which is found on one of the active site loops. However, previous biochemical studies indicated that the tryptophan residues of procaspase-3 were in similar orientations as in caspase-3 (Bose et al., 2003). To validate these results, steady-state and time-resolved anisotropy studies were performed on wild-type procaspase-3 (C163S) and two single tryptophan procaspase-3 mutants (W206Y C163S and W214V C163S). These proteins were purified as described previously for procaspase-3 (Chapter IV).

Equilibrium samples (1.5, 5 and 15 μ M procaspase-3) from 0 to 9 M urea-containing buffer pH 7.5 at half molar increments were prepared in duplicate. Experiments were conducted on a FluorMax spectrofluorometer (Horiba Jobin Yvon). Fluorescence intensity data (I_{VV} , I_{VH} , I_{HH} and I_{HV}) were collected with excitation at 280 or 295 nm (emission at 335

nm) and Glan Thompson polarizers. The correction factor, G, was calculated using Equation 2. The anisotropy was then calculated using Equation 3.

$$G = \frac{I_{HV}}{I_{HH}} \quad (2)$$

$$r = \frac{I_{VV} - 2 \times G \times I_{VH}}{I_{VV} + 2 \times G \times I_{VH}} \quad (3)$$

These results were then plotted versus the final urea concentration (Figure 15).

Equilibrium anisotropy results (Figure 15 and Table 1) revealed protein concentration dependence between ~5 and 7 M urea for all three proteins. This transition is associated with the dissociation of the dimer, as was detected by equilibrium fluorescence and CD studies (Bose and Clark, 2001). These data were similar to the results obtained on the SX18 stopped-flow instrument for procaspase-3 C163S, confirming the steady-state anisotropy data collected on the stopped-flow. In addition, the plateau values were similar for all three proteins, indicating that the form of the protein at low concentrations of urea has a similar rigidity in all three proteins. The change in anisotropy at low urea concentrations is scatter caused by precipitation in the samples.

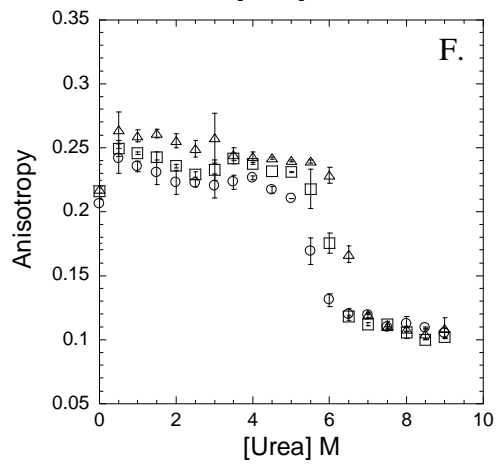
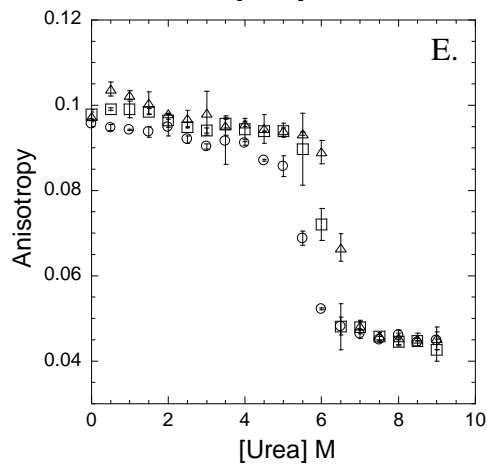
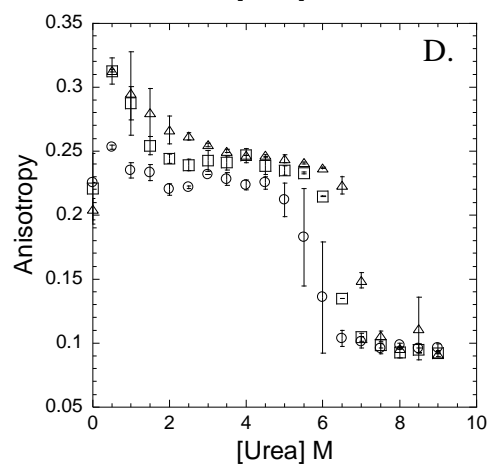
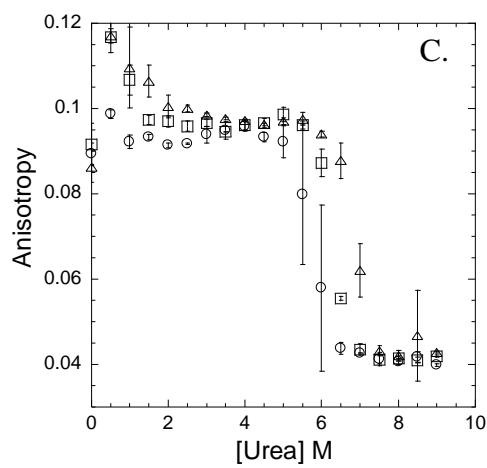
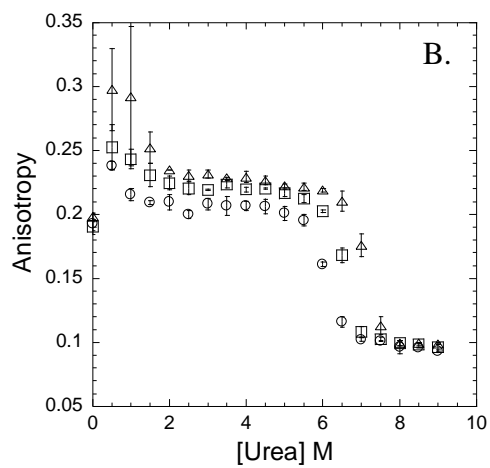
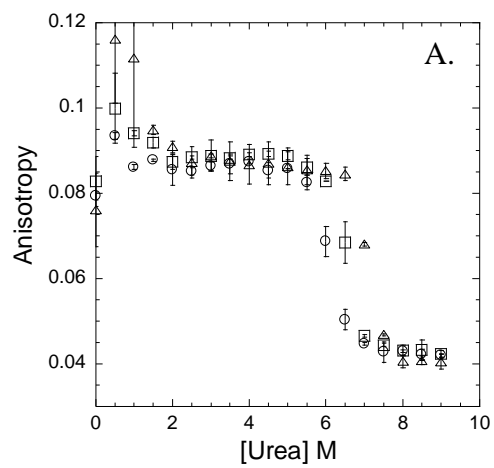


Figure 14: Homology model of procaspase-3. W206 is yellow and W214 is green.

Table 1: Protein concentration dependence and plateau anisotropy for procaspase-3 mutants from Figure 15.

procaspase-3	ex 280 nm [protein] dep.	ex 280 nm plateau	ex 295 nm [protein] dep.	ex 295 nm plateau
C163S	6-7 M urea	0.088	6-7.5 M urea	0.22
W206Y C163S	5.5-7 M urea	0.096	5.5-7 M urea	0.24
W214V C163S	5-6.5 M urea	0.095	5-6.5 M urea	0.23

Figure 15: Equilibrium anisotropy of procaspase-3 C163S (panels A and B), W206Y C163S (panels C and D) and W214V C163S (panels E and F) versus urea concentration. Excitation at 280 nm (panels A, C and E) and 295 nm (panels B, D and F) and emission at 335 nm. Protein concentrations: 1.5 μ M (\circ), 5 μ M (\square) and 15 μ M (Δ). Error bars represent the standard deviation from two independent experiments.



The equilibrium anisotropy data can also provide information on the stability of a protein. These data indicated that W214V C163S is less stable than C163S and W206Y C163S since the protein concentration dependent transition occurred at lower urea concentrations. The results suggested that W206 assists in stabilizing the protein. Since equilibrium anisotropy studies only provided comparison of the proteins (such as stability and rigidity), time-resolved anisotropy studies were performed in order to examine the orientation of the tryptophan residues in the active site.

In order to verify that the mutations do not greatly affect the secondary or tertiary structure of the protein, near- and far-UV studies were performed. As shown in Figure 16, the tryptophan mutants had similar secondary and tertiary structures to wild-type procaspase-3.

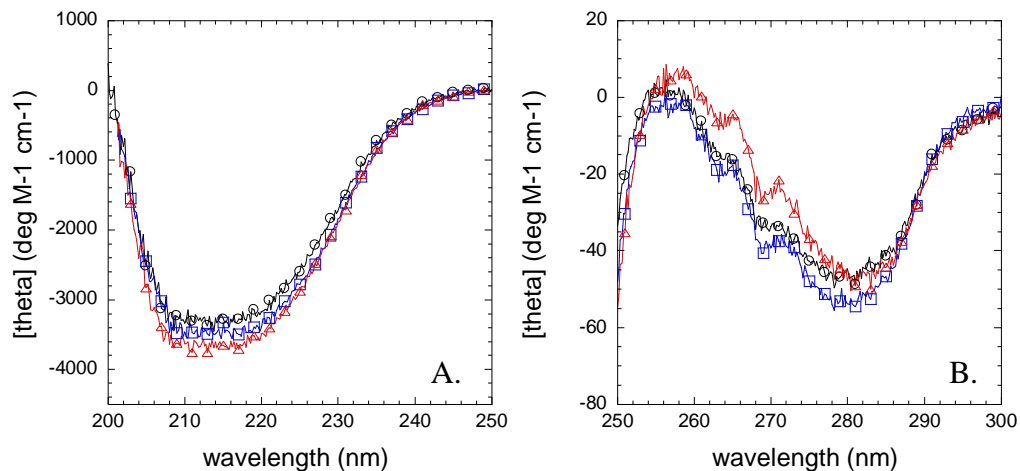


Figure 16: Far-UV (panel A) and near-UV (panel B) CD spectra of procaspase-3 mutants: C163S (black \circ), W206Y C163S (blue \square) and W214V C163S (red Δ).

Time-resolved anisotropy studies were performed on native procaspase-3 C163S, W206Y C163S and W214V C163S (Figure 17). The fluorescence anisotropy data, $r(t)$, were best fit by a sum of two exponentials (Gangal et al., 1999):

$$r(t) = r_0[f_S \exp(-t/\Phi_S) + f_F \exp(-t/\Phi_F)] \quad (4)$$

$$f_S + f_F = 1 \quad (5)$$

where r_0 is the limiting anisotropy, f_S and f_F are the pre-exponential weighting factors, and Φ_S and Φ_F are the slow and fast rotational correlation times, respectively. The fast exponential (f_F or Φ_F) represents the backbone motion of the tryptophan. The slow exponential (f_S or Φ_S) represents whole-body rotations of the protein. The time-resolved anisotropy data (Figure 17A) were fit with Equation 4 in order to determine the rotational correlation times (Φ_S and Φ_F) of each exponential. The first rotational correlation time (ϕ_1) for W206 Y (W214 reporting) was lower (slower tryptophan movement) than the other two mutants. This observation suggested that W214 was more immobile than W206. The second rotational correlation time (ϕ_2) was similar for all three proteins. For a ~60 kDa protein, the whole-body rotational correlation time is ~25 nsec. Therefore, these results confirmed that the procaspase-3 mutants were dimers under native conditions.

To further examine the orientation of the tryptophan residues in the active site of procaspase-3, quantum yield measurements were performed. Fluorescence quantum yield is defined as the ratio between emitted and absorbed photons (Lakowicz, 2006). Absorbance and fluorescence emission spectra of the procaspase-3 mutants were collected and compared

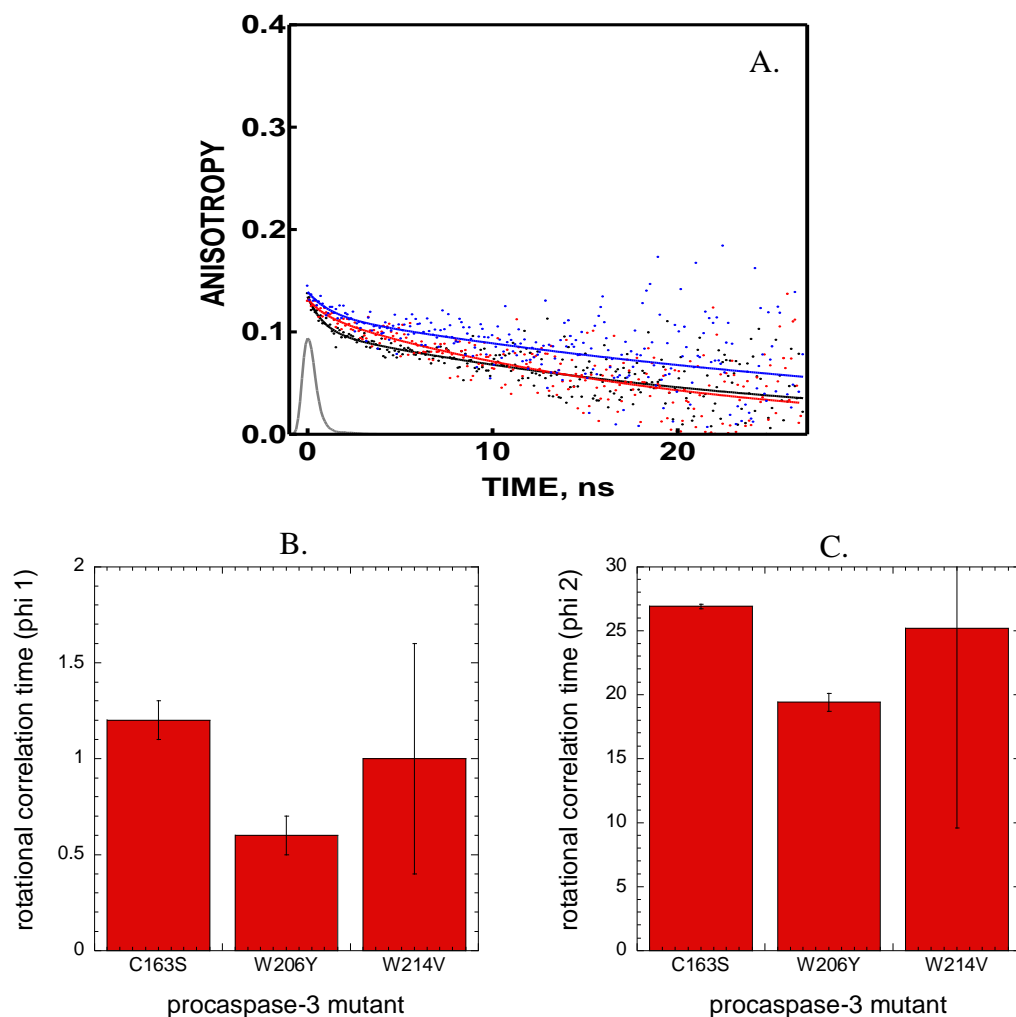


Figure 17: Time-resolved anisotropy data for procaspase-3 mutants under native conditions. A) Anisotropy decay data for procaspase-3 C163S (black), W206Y C163S (red) and W214V C163S (blue) versus time. IRF (instrument response function) is grey. Rotational correlation time 1 (panel B) and 2 (panel C) for procaspase-3 mutants determined from anisotropy data analysis.

to a tryptophan control. The quantum yield was calculated using Equation 6.

$$Q = Q_R \times \left(\frac{I}{I_R}\right) \times \left(\frac{OD_R}{OD}\right) \quad (6)$$

where Q is the quantum yield, I is integrated intensity, OD is optical density and subscript R is reference (tryptophan).

The absorbance scan (250-350 nm) for each protein and the tryptophan control is shown in Figure 18A. All of these scans displayed the characteristic peak for tryptophan residues, as shown by the tryptophan control. This provides evidence that the two tryptophans in the active site were not directly interacting.

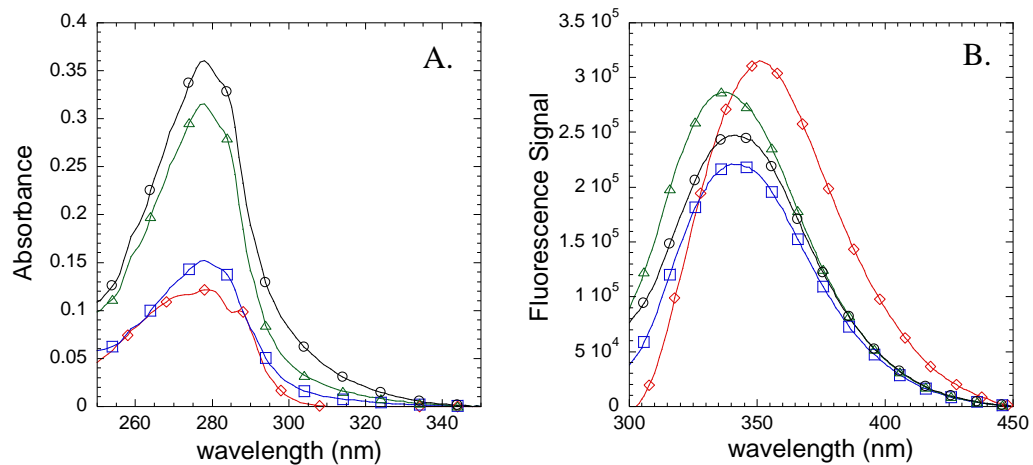


Figure 18: Absorbance (panel A) and fluorescence emission with excitation at 295 nm (panel B) scans of procaspase-3 mutants: 10 μM C163S (blue □), 20 μM W206Y C163S (green Δ), 20 μM W214V C163S (black ○). 20 μM tryptophan (red ◇).

The fluorescence emission spectrum for each procaspase-3 mutant and tryptophan control is shown in Figure 18B. As expected, the emission maximum of the tryptophan control was 350 nm. The emission maximum for the W206Y C163S (W214 reporting) was blue

shifted compared to wild-type C163S, indicating that W214 was buried in the active site. This was supported by the procaspase-3 homology model.

Using Equation 6, the quantum yields for procaspase-3 C163S, W206Y C163S and W214V C163S were 0.074, 0.047 and 0.036, respectively. These values suggested that the tryptophans could be interacting due to the fact that addition of the quantum yields from the two single tryptophan mutants (0.083) was larger than wild-type C163S.

In order to determine if there is an interaction between the tryptophan residues in the active site, the weighted average of W206Y and W214 was calculated using Equation 7.

$$r(t) = \frac{a(t) \times Q_a + b(t) \times Q_b}{Q_a + Q_b} \quad (7)$$

where $r(t)$ is anisotropy, $a(t)$ is anisotropy decay of W206Y, $b(t)$ is anisotropy decay of W214V and Q_a and Q_b are the quantum yields of W206Y and W214V, respectively. The result is shown in Figure 19. Wild-type procaspase-3 C163S decayed slightly faster than the weighted average. This result is consistent with some Förster Resonance Energy Transfer (FRET) between the two tryptophans, as indicated by an increase in the anisotropy decay.

The collaboration with David A. Johnson at the University of California, Riverside provided the following conclusions. First, the lifetime of the native dimer and unfolded monomer were similar, indicating that the changes in anisotropy observed in the steady-state results are authentic. Second, time-resolved anisotropy studies revealed that W214 is immobile, supporting the fact that it is buried in the active site. Third, the fluorescence emission spectrum of W214 revealed a blue shift compared to wild-type procaspase-3,

indicating that it is buried in the active site. Fourth, quantum yield values suggested that the tryptophans may be interacting. The microenvironments of the two tryptophan residues in each monomer are not in the same orientation (based on the procaspase-3 homology model). Therefore, an average of four signals can be detected in wild-type procaspase-3, which results in complicated data analysis. This complexity still exists in the single tryptophan mutants, since the two tryptophans could be in different configurations in each active site. Thus, time-resolved anisotropy data is extremely complicated and unresolved.

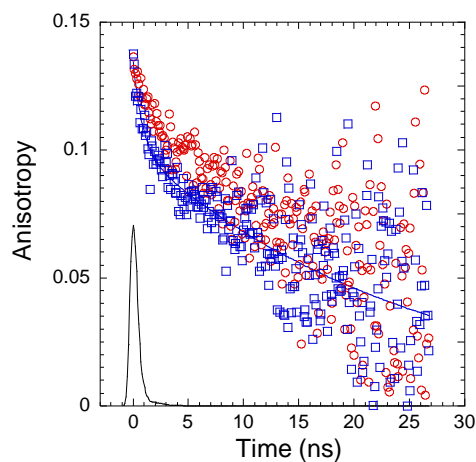


Figure 19: Time-resolved anisotropy data for procaspase-3 C163S (blue) and weighted average (red) of W206Y C163S and W214V C163S. IRF is grey.

REFERENCES

Bose, K., and Clark, A. C. (2001). Dimeric procaspase-3 unfolds via a four-state equilibrium process. *Biochemistry* **40**, 14236-14242.

Bose, K., Pop, C., Feeney, B., and Clark, A. C. (2003). An uncleavable procaspase-3 mutant has a lower catalytic efficiency but an active site similar to that of mature caspase-3. *Biochemistry* **42**, 12298-12310.

Gangal, M., Clifford, T., Deich, J., Cheng, X., Taylor, S. S., and Johnson, D. A. (1999). Mobilization of the A-kinase N-myristate through an isoform-specific intermolecular switch. *Proceedings of the National Academy of Sciences* **96**, 12394-12399.

Hanson, D. C., Yguerabide, J., and Schumaker, V. N. (1981). Segmental flexibility of immunoglobulin G antibody molecules in solution: a new interpretation. *Biochemistry* **20**, 6842-6852.

Lakowicz, J. R. (2006). "Principles of fluorescence spectroscopy." Springer, Baltimore.

VI. Differential quenching by acrylamide

Stopped-flow fluorescence and CD studies were unable to detect the dimerization step in the kinetic folding of procaspase-3. Therefore, differential quenching by acrylamide was adapted for kinetic folding experiments. These experiments are typically used in equilibrium experiments to investigate the solvent accessibility of tyrosine or tryptophan residues in a protein. Based on examples of differential quenching used in kinetic folding studies (Schulenburg et al., 2007; Vanhove et al., 1998), initial experiments focused on the practical aspects of the experiment (concentration of quencher, where and when the quencher should be added to the folding experiment, etc).

Previous studies showed that 0.7 M acrylamide provided the greatest difference in signal between native and unfolded protein (Bose et al., 2003). Therefore, this concentration of quencher was used in the initial studies. Acrylamide was tested in both protein and buffer solutions in order to determine if the quencher affects the folding process. Initial parameterization results suggested that the acrylamide interacts with the unfolded form of the protein when added to the protein solution, but not with the refolding reaction when it is included in the buffer solution. Other groups have used quencher in the buffer solutions (Schulenburg et al., 2007; Vanhove et al., 1998), which supports these results. Therefore acrylamide was used in the urea solutions for these experiments.

Chapter IV describes the results obtained for procaspase-3 kinetic refolding studies with 0.7 M acrylamide. In order to examine if acrylamide affects the folding process of the protein, multiple quencher concentrations (0.1 and 0.4 M acrylamide) were tested. These

experiments were performed with the same procedure as was used with 0.7 M acrylamide. Multiple protein concentrations (2 and 5 μ M procaspase-3) were also examined. Representative results are shown in Figure 20. The burst phase (Figure 20A) displayed a cooperative transition, which was similar for all three acrylamide concentrations examined. These data verified that the acrylamide did not affect the structure or formation of the burst phase species. The rates of procaspase-3 refolding were also examined versus acrylamide concentration. As shown in Panels B and C of Figure 20, the concentration of acrylamide did not affect the rate of refolding of the protein. In addition, the data did not display protein concentration dependence, as was observed in the 0.7 M acrylamide reactions. Combined, these results showed that acrylamide did not interfere with the folding process of procaspase-3. Therefore, the results determined from 0.7 M acrylamide were valid.

REFERENCES

- Bose, K., Pop, C., Feeney, B., and Clark, A. C. (2003). An uncleavable procaspase-3 mutant has a lower catalytic efficiency but an active site similar to that of mature caspase-3. *Biochemistry* **42**, 12298-12310.
- Schulenburg, C., Martinez-Senac, M. M., Low, C., Golbik, R., Ulbrich-Hofmann, R., and Arnold, U. (2007). Identification of three phases in Onconase refolding. *Federation of European Biochemical Societies Journal* **274**, 5826-5833.
- Vanhove, M., Lejeune, A., Guillaume, G., Virden, R., Pain, R. H., Schmid, F. X., and Frere, J.-M. (1998). A collapsed intermediate with nonnative packing of hydrophobic residues in the folding of TEM-1 β -lactamase. *Biochemistry* **37**, 1941-1950.

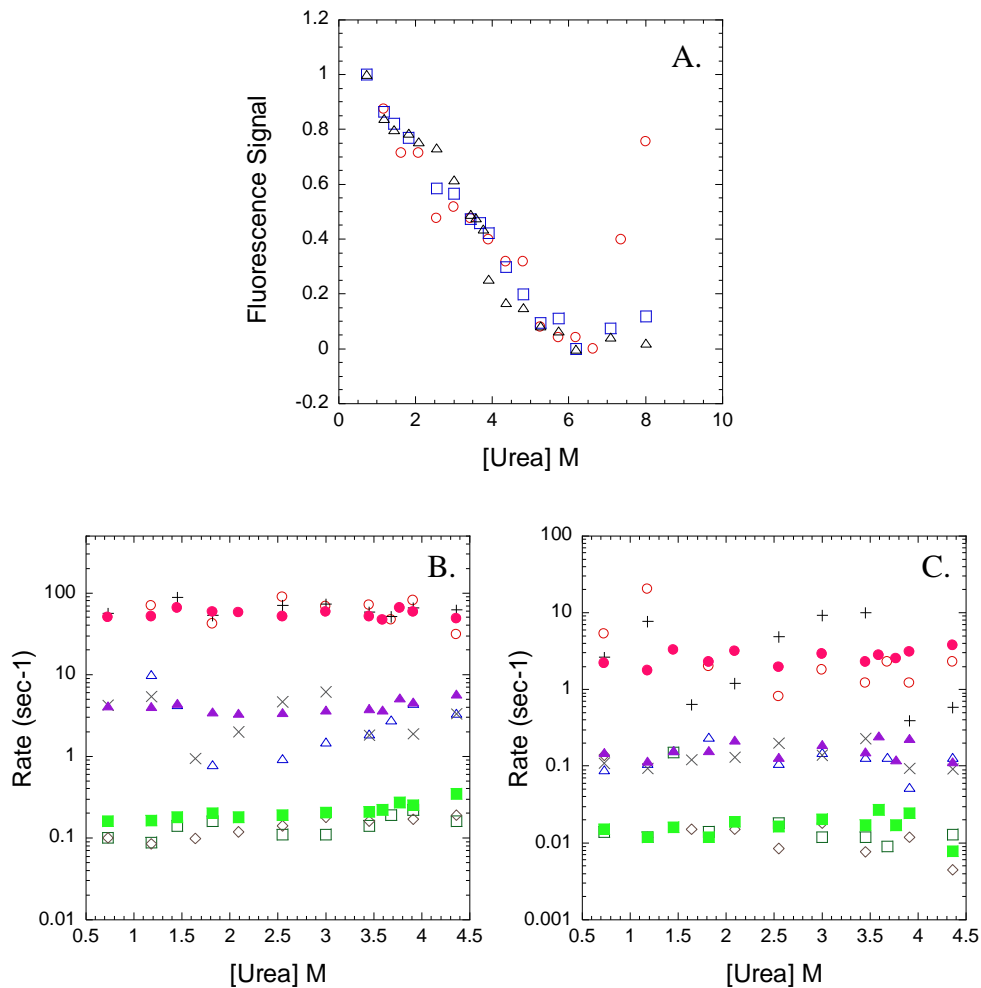


Figure 20: Differential quenching by acrylamide. A) Burst phase versus urea concentration for 5 μ M procaspase-3 with 0.1 (\circ), 0.4 (\square) or 0.7 (\triangle) M acrylamide. B and C) Rates (sec^{-1}) of procaspase-3 refolding in the presence of 0.1 M (+, \times , \diamond), 0.4 M (\circ , \triangle , \square) or 0.7 M (\bullet , \blacktriangle , \blacksquare) acrylamide. Shown in panels B and C are procaspase-3 concentrations of 5 μ M for 0.7 and 0.1 M acrylamide and 2 μ M for 0.4 M acrylamide.

VII. Procaspase-3 D3A studies

A. Activity assays

Activity assays were performed with a catalytically active mutant of procaspase-3, in which the three processing sites have been mutated to alanines (D3A). Various concentrations of the protein (ranging from 0.5 to 7 μM) were incubated in 0.73 M urea-containing buffer overnight. The activity was tested with a fluorogenic substrate (100 μM) and 400 nM protein (red columns, Figure 21A). To verify that the volume of protein added to the cuvette was not effecting the activity reading, protein samples were diluted to 0.5 μM , incubated for 5 minutes, then tested for activity (blue bars, Figure 21A). The results of these two experiments were similar. The activity decreased as the protein concentration is increased. These results were most likely due to aggregation of the protein at higher protein concentrations.

Similar results were observed in the return of activity studies with procaspase-3 (D3A). Unfolded protein was mixed 1:10 with buffer and aliquots of the refolding reaction were tested for activity. As shown in Figure 21B, the activity increased with increasing refolding time. The higher protein concentrations reached maximal activity faster than the lower protein concentrations, indicating that this folding process was the dimerization of procaspase-3. In addition, the maximum activity obtained was higher with lower protein concentrations, as was seen with the native controls.

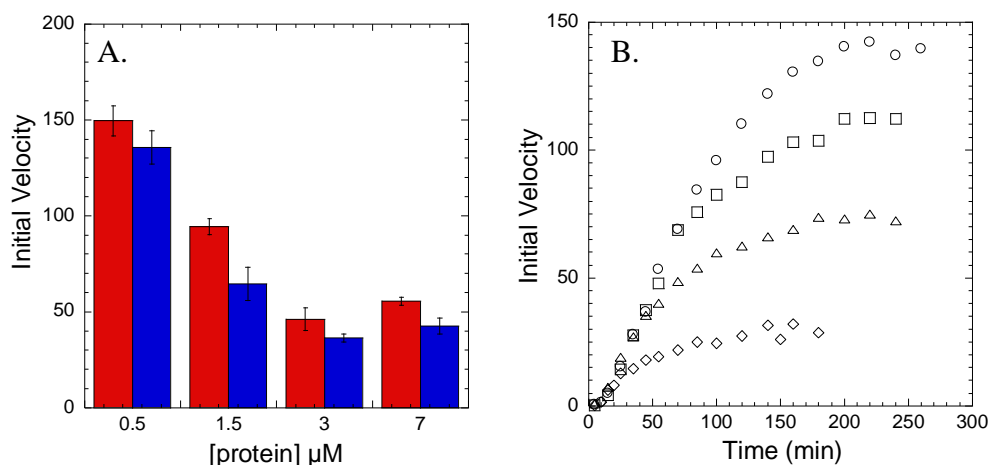


Figure 21: Activity studies with procaspase-3 D3A. A) Initial velocity of native 0.5, 1.5, 3 and 7 μM procaspase-3 D3A was determined under standard (red) or diluted (blue) conditions. Standard conditions are 400 nM procaspase-3, 1 mM DTT, 100 μM substrate and 0.73 M urea-containing potassium phosphate buffer pH 7.5. For the diluted reactions, the protein was diluted to 0.5 μM , incubated for 5 minutes at 25 $^{\circ}\text{C}$, and then tested for activity. Error bars represent the standard deviation from 2 independent experiments. B) Initial velocity measurements observed at various refolding times in minutes (as indicated on the x-axis) of 0.5 μM (\circ), 1.5 μM (\square), 2.5 μM (\triangle) and 5 μM (\diamond) procaspase-3 D3A.

B. Kinetic studies

Activity studies were used to determine the rate of procaspase-3 dimerization.

However, these studies were performed with active procaspase-3 (D3A), while the kinetic data was collected with the inactive procaspase-3 (C163S). Therefore, stopped-flow and hand-mixing experiments were performed with the active procaspase-3 mutant to verify that there was no difference in the kinetic folding pathway of these two proteins.

Stopped-flow fluorescence (excitation at 280 nm) and far-UV CD (228 nm) experiments were performed with procaspase-3 D3A. These data revealed the same number

of phases and rates of these phases, representative plots of which are shown in Figure 22. The burst phase signals (Figure 22A) were similar for both the fluorescence and CD measurements. Stopped-flow CD studies with procaspase-3 D3A also revealed a lag phase, as was shown with procaspase-3 C163S. The lag phase for both mutants is shown in Figure 22B versus the final urea concentration. These data were similar and were not protein concentration dependent. The slow phase in the refolding reaction of procaspase-3 was also investigated. The results for the CD hand-mixing experiments are shown in Figure 22C. In these experiments, unfolded protein was mixed with buffer, and the far-UV CD (228 nm) signal was monitored over time. Procaspase-3 C163S data displayed a single exponential phase with a rate of 0.027 min^{-1} and were not protein concentration dependent (see Chapter IV). Procaspase-3 D3A data also displayed a single exponential phase with a rate of 0.030 min^{-1} and were not protein concentration dependent. The kinetic studies with the D3A mutant were similar to the C163S mutant. Therefore the activity studies provide a valid method to determine the rate of dimerization of procaspase-3.

To verify that the substrate does not increase the rate of refolding of procaspase-3, hand-mixing CD studies were performed with an inhibitor, DEVD-CHO. The inhibitor did not affect the native or unfolded protein signal (data not shown). In addition, the inhibitor did not increase the rate of procaspase-3 D3A refolding (Figure 20C). The data were fit to a single exponential with a rate of 0.033 min^{-1} , which was similar to the rate in the absence of inhibitor. Therefore, the inhibitor had no effect on the rate of refolding of procaspase-3.

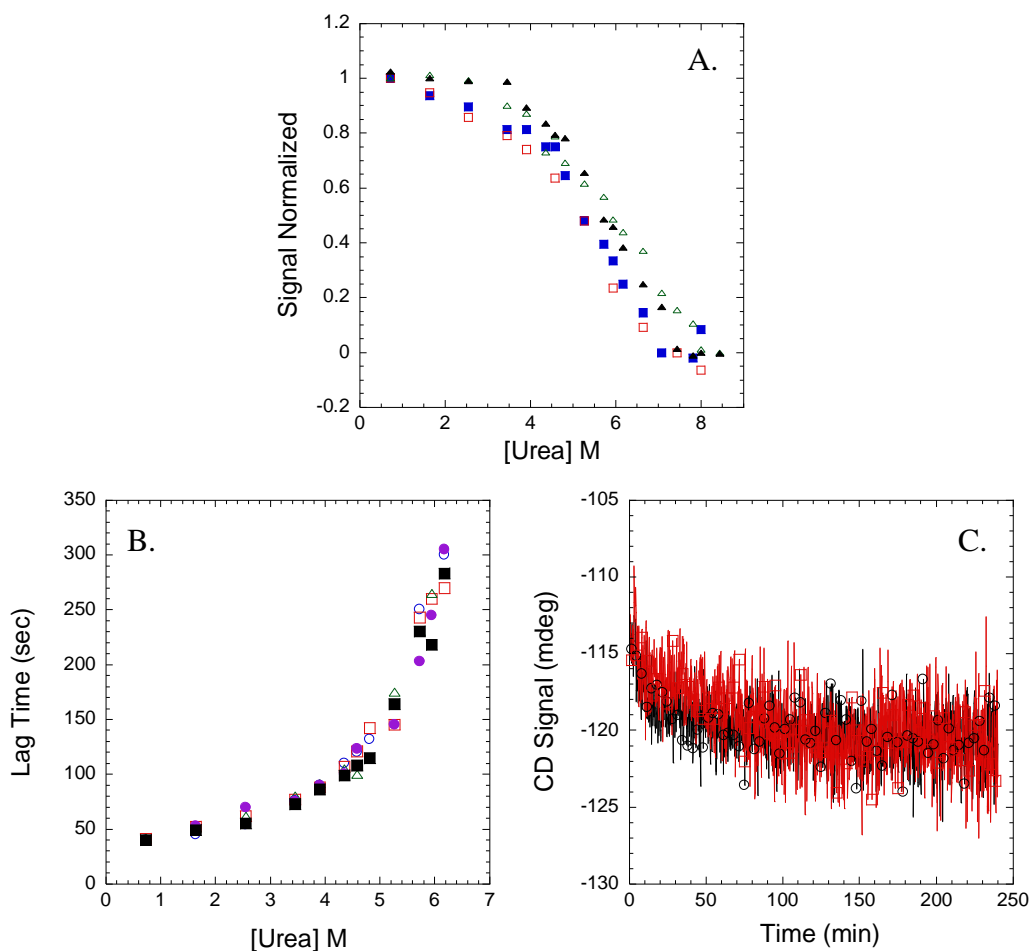


Figure 22: Kinetic studies with procaspase-3 D3A. A) Burst phase signal. Unfolded protein was mixed with various urea-containing buffers for the final urea concentrations shown in the figure. 8 μM stopped-flow CD, procaspase-3 C163S (black, \blacktriangle) and D3A (green, \triangle). 1 μM stopped-flow fluorescence, excitation at 280 nm, procaspase-3 C163S (red, \square) and D3A (blue, \blacksquare). B) Lag time versus the final urea concentration for stopped-flow CD data. 4 (\circ), 8 (\square) and 12 (\triangle) μM procaspase-3 C163S. 8 (\bullet) and 10 (\blacksquare) μM procaspase-3 D3A. C) Slow phase in refolding monitored by far-UV CD (228 nm). 80 μM unfolded procaspase-3 was hand-mixed with buffer, and the far-UV CD signal was monitored over the time (red, \square). For a control, the refolding reaction also included 20 μM inhibitor (DEVD-CHO) (black, \circ).

APPENDIX C

Protocols

I. Preparation of 10 M urea stock

This protocol, adapted from the method described by Pace (1986), describes the preparation of 100 mL of a 10 M urea stock containing 50 mM potassium phosphate buffer, pH 7.5. One should use an analytical balance that measures accurately to ± 0.1 mg. Potassium phosphate buffer is prepared separately in a volumetric flask (100 mL) using the chemicals in step 1 below and distilled, deionized water.

1. Separately weigh 191 mg of potassium phosphate monobasic, KH_2PO_4 , and 650 mg of dibasic, anhydrous dipotassium hydrogen phosphate, K_2HPO_4 on weigh paper.
2. Weigh 60.0 g of ultra pure urea (purity > 99%) in a weigh boat.
3. Combine both phosphates in a beaker with a minimum capacity of 200 mL, add a clean, dry stir bar, place on a scale and tare. Add the urea into the beaker and record the weight.
4. Add distilled, deionized water into the beaker to a weight of 114.6 g. Record the weight.
5. Cover the beaker and stir until the urea dissolves completely. One should expect this to take three to four hours.
6. When dissolved, check the pH using a recently calibrated pH meter. If the pH needs to be adjusted, then correct to pH 7.5, accordingly.
7. Filter the urea stock with 0.45 μm filter.
8. Once the urea stock is prepared, the molarity is determined based on the recorded weights from steps 3 and 4 and from the refractive index. First, calculate the molarity

based on the recorded weights using Equations 1-4.

$$d/d_0 = 1 + 0.2658w + 0.0330w^2 \quad (1)$$

where d/d_0 is the specific gravity of the solution and w equals the weight of urea (g) divided by the total weight, as shown in Equation 2.

$$w = 60.0 \text{ g} \div 114.6 \text{ g} = 0.52356$$

$$d/d_0 = 1 + 0.2658(0.52356) + 0.0330(0.52356)^2 \quad (2)$$

$$d/d_0 = 1.1482$$

The volume of the solution is then calculated as shown in Equation 3.

$$V = (\text{total weight}) \div (d/d_0) = 114.6 \text{ g} \div 1.1482 \text{ g/mL} = 99.8084 \text{ mL} \quad (3)$$

Finally, the molarity of the solution based on the recorded weight is calculated as shown in Equation 4.

$$M = (\text{weight of urea}) \div (MW_{\text{urea}} \times \text{Vol (L)}) \quad (4)$$

$$M = (60.0 \text{ g}) \div (60.06 \text{ g/mol} \times 0.0998084 \text{ L}) = 10.00 \text{ M}$$

Next, a refractive index is obtained for two solutions, a buffer blank, which in this case is 50 mM potassium phosphate, and the prepared urea stock. Add one drop of the phosphate buffer to the refractometer, and record the reading. Clean the lens with ethanol, and repeat this step using the 10 M urea stock.

Once the refractive indices have been recorded, calculate ΔN , the difference in the refractive index between the two solutions, as shown in Equation 5.

$$\Delta N = RI_{\text{urea}} - RI_{\text{buffer}} = 1.4208 - 1.3379 = 0.0829 \quad (5)$$

This value is used in the following empirical formula (Equation 6) to calculate the molarity

of the solution based on the refractive index:

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

$$M = 117.66(0.0829) + 29.753(0.0829)^2 + 185.56(0.0829)^3 = 10.06 \text{ M}$$

If the difference in the calculated molarity from each method is less than 10%, then the urea may be used. If it is greater than 10%, then the urea must be prepared again. Upon completion, the urea may be stored at -80 °C until used.

REFERENCES

Pace, C. N. (1986). Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods in Enzymology* **131**, 266-28

II. Preparation of 8 M guanidine hydrochloride stock

This protocol, adapted from the method described by Pace (1986), describes the preparation of 100 mL of a 8 M guanidine hydrochloride (GdmHCl) stock containing 50 mM potassium phosphate buffer, pH 7.5. One should use an analytical balance that measures accurately to ± 0.1 mg. Potassium phosphate buffer is prepared separately in a volumetric flask (100 mL) using the chemicals in step 1 below and distilled, deionized water.

9. Separately weigh 191 mg of potassium phosphate monobasic, KH_2PO_4 , and 650 mg of dibasic, anhydrous dipotassium hydrogen phosphate, K_2HPO_4 on weigh paper.

10. Weigh 76.4 g of ultra pure GdmHCl in a weigh boat.
11. Combine both phosphates in a beaker with a minimum capacity of 200 mL, add a clean, dry stir bar, place on a scale and tare. Add the GdmHCl into the beaker and record the weight.
12. Add distilled, deionized water into the beaker to a weight of 118.7 g. Record the weight.
13. Cover the beaker and stir until the GdmHCl dissolves completely.
14. When dissolved, check the pH using a recently calibrated pH meter. If the pH needs to be adjusted, then correct to pH 7.5, accordingly.
15. Filter the urea stock with 0.45 μm filter.
16. Once the GdmHCl stock is prepared, the molarity is determined based on the recorded weights from steps 3 and 4 and from the refractive index. First, calculate the molarity based on the recorded weights using Equations 7-9.

$$d/d_0 = 1 + 0.2710w + 0.0330w^2 \quad (7)$$

where d/d_0 is the specific gravity of the solution and w equals the weight of GdmHCl (g) divided by the total weight, as shown in Equation 8.

$$w = 76.4 \text{ g} \div 118.7 \text{ g} = 0.64364$$

$$d/d_0 = 1 + 0.2710(0.64364) + 0.0330(0.64364)^2 \quad (8)$$

$$d/d_0 = 1.1881$$

The volume of the solution is then calculated as shown in Equation 9.

$$V = (\text{total weight}) \div (d/d_0) = 118.7 \text{ g} \div 1.1881 \text{ g/mL} = 99.9074 \text{ mL} \quad (9)$$

Finally, the molarity of the solution based on the recorded weight is calculated as shown in Equation 10.

$$M = (\text{weight of urea}) \div (\text{MW}_{\text{urea}} \times \text{Vol (L)}) \quad (10)$$

$$M = (76.4 \text{ g}) \div (95.53 \text{ g/mol} \times 0.0999074 \text{ L}) = 8.00 \text{ M}$$

Next, a refractive index is obtained for two solutions, a buffer blank, which in this case is 50 mM potassium phosphate, and the prepared GdmHCl stock. Add one drop of the phosphate buffer to the refractometer, and record the reading. Clean the lens with ethanol, and repeat this step using the 8 M GdmHCl stock.

Once the refractive indices have been recorded, calculate ΔN , the difference in the refractive index between the two solutions, as shown in Equation 11.

$$\Delta N = \text{RI}_{\text{GdmHCl}} - \text{RI}_{\text{buffer}} = 1.4655 - 1.3345 = 0.131 \quad (11)$$

This value is used in the following empirical formula (Equation 12) to calculate the molarity of the solution based on the refractive index:

$$M = 57.147(\Delta N) + 38.68(\Delta N)^2 + 91.60(\Delta N)^3 \quad (12)$$

$$M = 57.147(0.131) + 38.68(0.131)^2 + 91.60(0.131)^3 = 7.94 \text{ M}$$

If the difference in the calculated molarity from each method is less than 10%, then the GdmHCl may be used. If it is greater than 10%, then the GdmHCl must be prepared again. Upon completion, the GdmHCl may be stored at $-80 \text{ }^\circ\text{C}$ until used.

REFERENCES

Pace, C. N. (1986). Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods in Enzymology* **131**, 266-28

III. KINSIM

KINSIM allows the user to simulate user-defined mechanisms. A brief description of the program and its advantages is given in Chapter II. A more detailed procedure will be described here.

- 1) KINSIM can be downloaded from Carl Frieden's website (<http://www.biochem.wustl.edu/cflab/message.html>). There is also a help manual and basic instructions listed on the website.
- 2) Write the mechanism. An example mechanism is shown in Figure 2. Save the mechanism as a text file in the KINSIM folder.
- 3) To open, double click "KINSIM40". The program will look like Figure 1. From this screen, the user will be able to upload a mechanism and change parameters, such as time, protein concentration and the rate constants.
- 4) The mechanism text file needs to be converted to a "sim" file. Type "O" for "compile mechanism to .sim file" on the KINSIM main screen. Type the mechanism file name (ie apaf.txt). Enter. Type the output file name (ie apaf.sim). Enter. The screen should then read "successful completion".

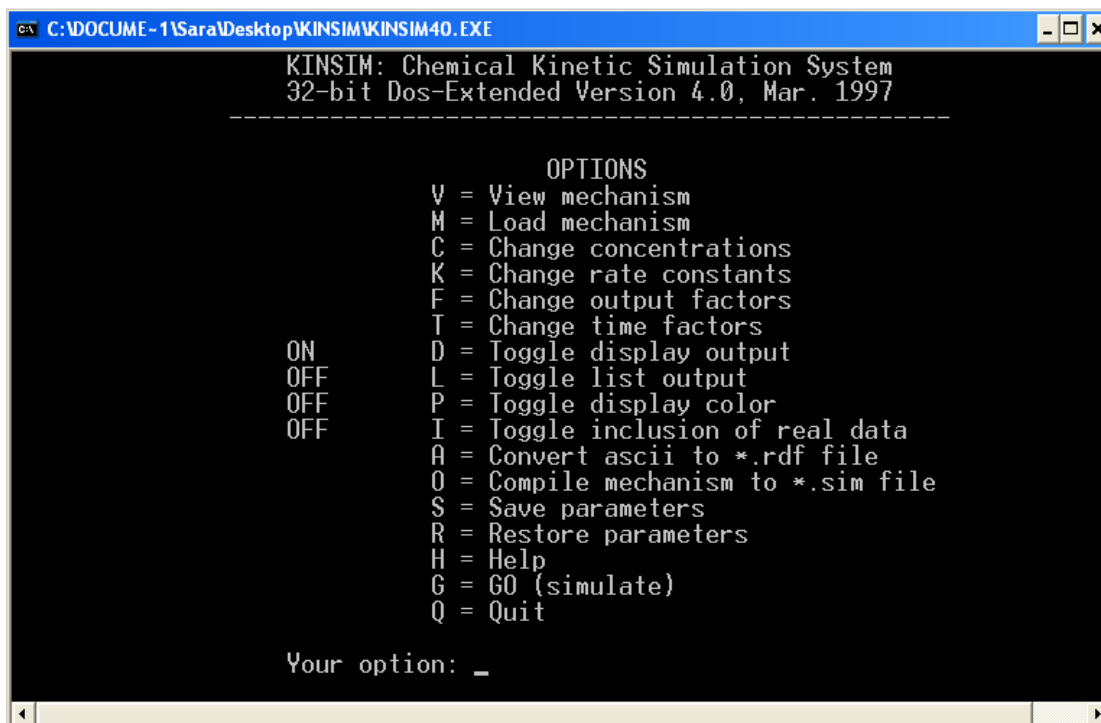


Figure 1: KINSIM main screen.

Figure 2: KINSIM mechanism. A) Mechanism from panel B written in text format. The first line is the title of the mechanism. A reversible reaction is denoted by 2 equal signs (\rightleftharpoons). The reaction can be written on one line or on separate lines as shown here. A The output list ($X1*B$, $X2*A$, etc) will be what is displayed on the screen. B and C) Parallel mechanism for Apaf-1 CARD. In panel C, there are 3 unfolded states of the protein, U_1 , U_2 and U_3 . U_1 and U_2 are able to fold directly into the native state, N. U_3 is able to form U_1 or U_2 . The unfolded states of the protein are able to interconvert. Panel B is the same as panel C except that the letters are different due to the fact that the program would read U_2 and U_3 as a dimer and trimer of U, respectively. Therefore, these letters were changed to A (U_1), C (U_2) and D (U_3) for simplicity. The red numbers written above each step of the reaction indicate the rate. For example, k_1 is the rate from U_1 to N or A to B and k_{-1} is the rate from N to U_1 or B to A. The rates are assigned in the order that they are written in the text file. D) View mechanism from the main screen. The title of the mechanism is listed at the top of the screen. The mechanism will be broken into steps. This allows the user to verify that the mechanism is correct.

A. \$Apaf folding

!

A == B

A == C

E == A

E == C

C == B

!

*OUTPUT

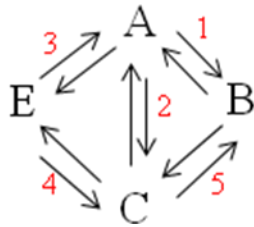
X1*B

X2*A

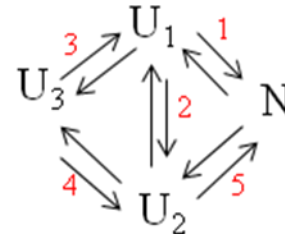
X3*C

X4*E

B.



C.



D.

```
C:\DOCUME-1\Sara\Desktop\KINSIM\KINSIM40.EXE
Apaf folding
      K+1      K+2      K+3      K+4      K+5
A == B ; A == C ; E == A ; E == C ; C == B
      K-1      K-2      K-3      K-4      K-5

      OUTPUT 1 = X1*B
      OUTPUT 2 = X2*A
      OUTPUT 3 = X3*C
      OUTPUT 4 = X4*E

Return to continue:
```

- 5) Load the mechanism. Type “M” for “load mechanism” on the KINSIM main screen. Type the .sim file name (ie apaf.sim). Enter.
- 6) View mechanism. Type “V” for “view mechanism on the KINSIM main screen. An example is shown in Figure 2D.
- 7) Parameters. Listed are the parameters and descriptions of each parameter. Also listed are two examples (i and ii under each parameter) for the Apaf-1 CARD simulations with the mechanism in Figure 2. After entering the values for each parameter, “control E” enter to return to the main menu.
 - a. “C” for change concentrations. The protein concentration is in μM . For the Apaf-1 CARD example, the total protein concentration is $10 \mu\text{M}$. These values are similar to percentage.
 - i. $A = 0.7, C = 8.3, E = 1$ (7% A, 83% B, 1% E)
 - ii. $A = 9, E = 1$ (90% A, 10% E)
 - b. “K” for rate constants (sec^{-1}). Try to set these values as close to the experimental values as possible. It is a good idea to start with one branch/step of a pathway, instead of filling out all the values at once. Burst phases are difficult to simulate. For a true burst phase, the rate needs to be about 5000 sec^{-1} . However, occasionally this value causes the program to crash. If this occurs, change the protein concentrations. For example, if A goes to B in the burst phase, instead of $10 A \rightarrow B$ with a rate of 5000 sec^{-1} , use $A \rightarrow 10 B$ with a rate of 0 sec^{-1} .

i) $k_1 = 5000$, $k_3 = 0.6$, $k_{-3} = 0.01$, $k_4 = 0.6$, $k_{-4} = 0.01$, $k_5 = 37$, $k_{-5} = 13$ for both Apaf-1 CARD examples

c. “F” for output factors. Output factors are defined as $\text{signal}/[\text{protein}]$. Output factors are calculated by dividing the signal of the species, taken from the experimental data, by the protein concentration. For the Apaf-1 CARD example, the protein concentration is $10 \mu\text{M}$. Therefore, the signal of the species is divided by 10 in order to obtain the output factor.

i. $X1 = 0.0086$ for both Apaf-1 CARD examples

d. “T” for time factors.

i. delta time: is the time step used in solution of the mechanism (0.001 seconds). If the program crashes, try changing this value.

ii. log scale = 1 yes

iii. iterations/pt = 1

iv. run time: length of simulation (10 seconds)

v. Ymax: signal maximum on the screen (obtained from the experimental data) – 0.1 for Apaf-1 CARD examples

vi. Flux tolerance and integral tolerance: control the numerical method used for solution of the mechanism (if the program crashes, try lowering the flux tolerance)

e. “G” for go (simulate). Simulates the mechanism with the parameters entered.

8) Experimental data can also be uploaded into KINSIM. First, the file must be

converted to an .rdf file. In order to do this, open the file in KaleidaGraph. Each experimental run needs to be saved individually in KaleidaGraph and then exported to a text file. In KaleidaGraph, “file”, “export”, “tab-delimited text”. In KINSIM, type “A” (convert asii to .rdf file). Enter.

- a. Use an old batch conversion file (Y/N) [N]: enter
- b. Name for the conversion file is DEFAULT.CVT (Y/N)[Y]: enter
- c. Does the ascii file have headers or comments (Y/N)[N]: Y enter
- d. Do you want the ascii files inverted (Y/N)[N]: enter
- e. Do you want the y-values shifted (Y/N)[N]: enter
- f. How many files to process: insert the number of files to process
- g. Input the name of the ascii data file to convert: insert file name .txt Enter
- h. Enter a name for the RDF file: insert file name .rdf Enter. Now KINSIM will display the file. Be sure to take note where the first and last lines of the data are. The first line will most likely be line 2. Keep hitting Enter until you reach the next question. Type “2” to continue, unless you need to see the lines again.
- i. Enter 1st line of data: example “2”. Enter
- j. Enter last line of data: example “1001”. Enter.
- k. The X,Y data is in 2 columns, so hit Enter.
- l. Do you want to modify x and y-values (Y/N)[N]: Enter
- m. You need to enter a zero time point, 0.0001 Enter
- n. Enter name of creator: Enter

The results for the Apaf-1 CARD example simulations are shown in Figure 3. The simulation parameters in Figure 3A agree with the experimental data. KINSIM simulations can help the user investigate multiple mechanisms and variations of these mechanisms. KINSIM results can support a mechanism (not prove).

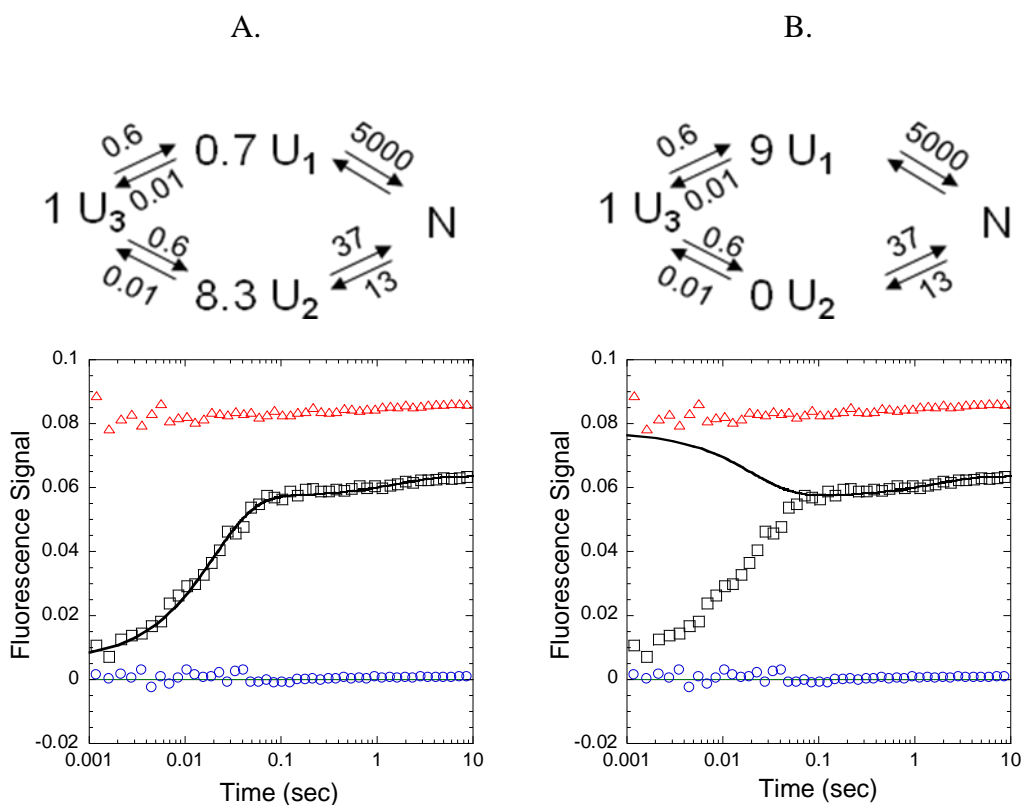


Figure 3: KINSIM simulations of Apaf-1 CARD. The difference between panel A and B is the initial concentration of the protein. The simulation results are shown below each mechanism. The native (red, Δ) and unfolded (blue, \circ) protein signals are shown. Unfolded Apaf-1 CARD is mixed with urea-containing buffer for a final urea concentration of 3.64 M (black, \square). The black line represents the simulation with the parameters listed above. The orange, purple and green lines represent the unfolded protein species (U_1 , U_2 and U_3).

REFERENCES

Barshop, B. A., Wrenn, R. F., and Frieden, C. (1983). Analysis of numerical methods for computer simulation of kinetic processes: Development of KINSIM - a flexible, portable system. *Anal. Biochem.* **130**, 134-145.

IV. Derivation of two-state equilibrium folding model



The two-state mechanism is shown in Equation 13, in which the native monomer (N) cooperatively unfolds to the unfolded protein (U). K_{eq} is the equilibrium constant for the reaction. The total molar concentration of the protein (P_T) is shown in Equation 14.

$$[P_T] = [N] + [U] \quad (14)$$

The molar fraction of each species is shown in Equations 15 and 16.

$$f_N = \frac{N}{P_T} \quad (15)$$

$$f_U = \frac{U}{P_T} \quad (16)$$

The sum of these fractions is unity, as shown in Equation 17.

$$1 = f_N + f_U \quad (17)$$

The equilibrium constant, K_{eq} , is defined as:

$$K = \frac{f_U}{f_N} \quad (18)$$

From the above equations, the fraction of the unfolded species (f_U) is shown by Equation 19.

$$f_U = \frac{K}{K + 1} \quad (19)$$

The equilibrium constant for the transition is related to the free energy (ΔG) as shown in Equation 20.

$$\Delta G = -RT \ln(K_{eq}) \quad (20)$$

where R is the gas constant ($\text{cal K}^{-1} \text{mol}^{-1}$) and T is temperature (K). If one assumes that the free energy change for each step of the reaction is linearly dependent on the denaturant concentration (Pace, 1986), then the free energy change in the absence of denaturant is shown in Equation 21.

$$\Delta G = \Delta G^{H_2O} - m_1[\text{denaturant}] \quad (21)$$

where ΔG^{H_2O} is the free energy change in the absence of denaturant and m_1 represents the cooperativity index associated with the reaction. The amplitude of the spectroscopic signal determined at each urea concentration was assumed to be a linear combination of the fractional contribution from each species (Equation 22).

$$Y = Y_N f_N + Y_U f_U \quad (22)$$

where Y_N and Y_U are the amplitudes of the signals for the native and unfolded species, respectively. Finally, the amplitudes associated with the native and unfolded forms of the protein were assumed to be linearly dependent on the urea concentration, as shown in Equations 23 and 24.

$$Y_N = Y_{N'} + m_2[\text{urea}] \quad (23)$$

$$Y_U = Y_{U'} + m_3[\text{urea}] \quad (24)$$

where $Y_{N'}$ and $Y_{U'}$ are the amplitudes of the signals in the absence of urea for the native and unfolded species, respectively.

Data were plotted in KaleidaGraph. The equation for a two-state equilibrium folding model in KaleidaGraph format is shown in Equation 25.

$$\frac{((m1 + m2 * m0) + (m3 + m4 * m0) * \exp(-(m5 + m6 * m0)/(0.5921)))}{(1 + \exp(-m5 + m6 * m0)/0.5921))} \quad (25)$$

where m1 and m3 are the y-intercepts of the pre- and post-transitions, respectively; m2 and m4 are the slopes of the pre- and post-transition, respectively; m0 is the urea concentration; m5 is the ΔG^{H_2O} ; m6 is the m-value. The fitting procedure for Igor Pro is shown in Figure 4.

REFERENCES

Pace, C. N. (1986). Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods in Enzymology* **131**, 266-28

```

#include <Global Fit>
Function TwoState (w, x): FitFunc

    Wave w

    Variable x

    Variable K=exp(-(w[0]-(w[1]*x))/(0.001987*298))
    Variable Fu=K/(1+K)
    Variable Fn=1-Fu
    Variable Yf=w[2]+w[3]*x
    Variable Yu=w[4]+w[5]*x
    Variable y=Yf*Fn+Yu*Fu

    Return y

End

```

Figure 4: Procedure for a two-state equilibrium folding model in Igor Pro. $w[2]$ is Y_N , $w[3]$ is m_2 , $w[4]$ is Y_U , $w[5]$ is m_3 . The other parameters have been defined above.

V. Derivation of three-state equilibrium folding model



The three-state mechanism is shown in Equation 26, in which the native monomer (N) cooperatively unfolds to an intermediate (I), which then unfolds to the unfolded protein (U).

K_1 and K_2 are the equilibrium constant for first and second transitions in the reaction, respectively. The total molar concentration of the protein (P_T) is shown in Equation 27.

$$[P_T] = [N] + [I] + [U] \quad (27)$$

The molar fraction of each species is shown in Equations 28, 29 and 30.

$$f_N = \frac{N}{P_T} \quad (28)$$

$$f_I = \frac{I}{P_T} \quad (29)$$

$$f_U = \frac{U}{P_T} \quad (30)$$

The sum of these fractions is unity, as shown in Equation 31.

$$1 = f_N + f_I + f_U \quad (31)$$

The equilibrium constants, K_1 and K_2 , are defined as:

$$K_1 = \frac{f_I}{f_N} \text{ and } K_2 = \frac{f_U}{f_I} \quad (32)$$

From the above equations, the fraction of each species is shown by Equations 33, 34 and 35.

$$f_U = \frac{K_1 K_2}{1 + K_1 + K_1 K_2} \quad (33)$$

$$f_I = \frac{f_U}{K_2} \quad (34)$$

$$f_N = \frac{f_U}{K_1 K_2} \quad (35)$$

The equilibrium constant for the transition is related to the free energy (ΔG) as shown in Equation 36.

$$\Delta G = -RT \ln(K_{eq}) \quad (36)$$

where R is the gas constant ($\text{cal K}^{-1} \text{mol}^{-1}$) and T is temperature (K). If one assumes that the free energy change for each step of the reaction is linearly dependent on the denaturant concentration (Pace, 1986), then the free energy change in the absence of denaturant is shown in Equations 37 and 38.

$$\Delta G_1 = \Delta G_1^{H_2O} - m_1[\text{denaturant}] \quad (37)$$

$$\Delta G_2 = \Delta G_2^{H_2O} - m_2[\text{denaturant}] \quad (38)$$

where $\Delta G_1^{H_2O}$ and $\Delta G_2^{H_2O}$ are the free energy change in the absence of denaturant corresponding to K_1 and K_2 , respectively, and m_1 and m_2 represent the cooperativity indices associated with each step of the reaction. The amplitude of the spectroscopic signal determined at each urea concentration was assumed to be a linear combination of the fractional contribution from each species (Equation 39).

$$Y = Y_N f_N + Y_I f_I + Y_U f_U \quad (39)$$

where Y_N , Y_I and Y_U are the amplitudes of the signals for the native, intermediate and unfolded species, respectively. Finally, the amplitudes associated with the native, intermediate and unfolded forms of the protein were assumed to be linearly dependent on the urea concentration, as shown in Equations 40, 41 and 42.

$$Y_N = Y_{N'} + m_3[\text{urea}] \quad (40)$$

$$Y_I = Y_{I'} + m_4[\text{urea}] \quad (41)$$

$$Y_U = Y_{U'} + m_5[\text{urea}] \quad (42)$$

where $Y_{N'}$, $Y_{I'}$ and $Y_{U'}$ are the amplitudes of the signals in the absence of urea for the native, intermediate and unfolded species, respectively.

The fitting procedure for Igor Pro is shown in Figure 5.

```
#include <Global Fit>
Function ThreeState (w, x): FitFunc

    Wave w

    Variable x

    Variable K1=exp(-((w[0]-(w[1]*x)/(0.001987*298)))
    Variable K2=exp(-((w[2]+w[3]*x)/(0.001987*298)))
    Variable Ff=1/(1+K1+K1*K2)
    Variable Fi=K1/(1+K1+K1*K2)
    Variable Fu=(K1*K2)/(1+K1+K1*K2)
    Variable Yf=w[4]+w[5]*x
    Variable Yu=w[6]+w[7]*x
    Variable Yi=w[8]+w[9]*x
    Variable y=Yf*Ff+Yu*Fu+Yi*Fi

    Return y

End
```

Figure 5: Procedure for a three-state equilibrium folding model in Igor Pro. $w[4]$ is $Y_{N'}$, $w[5]$ is m_3 , $w[6]$ is $Y_{U'}$, $w[7]$ is m_4 , $w[8]$ is $Y_{I'}$, $w[9]$ is m_5 . The other parameters have been defined above.

REFERENCES

Pace, C. N. (1986). Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods in Enzymology* **131**, 266-28

VI. Anisotropy calibration

- 1) Anisotropy accessory set-up is described in the user manual.
- 2) Parameters: excitation at 280 or 295 nm; 305 nm cut-off filters for emission; slits of 2 mm; signal: fluorescence polarization; check “total fluorescence”
- 3) Flush the system with deionized water and buffer as usual.
- 4) Flush the system with refolded control protein and 0.73 M urea-containing buffer.
(This procedure can also be performed with the unfolded protein and 8 M urea-containing buffer. The signal and G-factors will be the same using either control to calibrate the instrument.)
- 5) Click on the “set-up” button on the control panel. Follow the directions on the screen, which are described below.
- 6) Move the polarizer to the H (set-up) position. Set the target signal voltage level to 30% and click on the autoPM button. Record the voltages.
- 7) Set the acquisition channel offsets to compensate for scattered light. Ensure that the cell is filled with fluorophore absent chemistry. Remove the protein and buffer from the syringes. Add deionized water to the syringes and flush the system. The instrument will not allow the user to drive the syringes, so the user has to manually

- flush the deionized water through the system. Turn the stop-syringe knob clockwise and manually empty the stop-syringe and then turn the knob back. Manually drive deionized water through the system three times (refill the syringes each time). Then move the polarizer to the V (measure) position. Click on “measure offsets” to measure the zero signal offsets for the V position. Record the offset values.
- 8) Set acquisition channel offsets to compensate for scattered light. Move the input polarizer to the H position. Click on “measure offsets” to measure the zero signal offsets for the H position. Record the offset values.
 - 9) Ensure the input polarizer is in the H position. Flush the system manually with protein and buffer as in step 7. Press the “measure” button to measure the G-factor with the required precision of 15%. The resultant G-factor is displayed below along with the number of iterations. Record the G-factor.
 - 10) You have completed the fluorescence polarization set-up. Move the input polarizer to the V position. Start experiment with (refolded) control first (flush not necessary as the tubing already contains refolded protein and buffer).
 - 11) Save data for both anisotropy and total fluorescence separately (right click on the y-axis label to change the detection method).