

PROTEIN FOLDING IN LIVING CELLS AND UNDER PRESSURE

BY

ANNA JEAN WIRTH

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Chemistry
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2015

Urbana, Illinois

Doctoral Committee:

Professor Martin Gruebele, Chair
Professor Yann Chemla
Professor Prasanth Kumar Kannanganattu
Professor Deborah Leckband

ABSTRACT

Protein folding, the process through which proteins gain their functional structure, can be approached from the perspective of many disciplines. Starting with biology, we can probe how protein structure relates to function and consider how the fold of a protein interacts with the biological environment. From the chemical perspective, we can treat protein folding as a chemical reaction and study the thermodynamics and kinetics of the structural transition. With physics, we can understand the underlying forces that give rise to protein folding and use theory and simulation to describe the protein folding process on an atomic level. This thesis studies protein folding through the lens of all three of these fields with two interdisciplinary methodological themes: one at the interface of chemistry and physics and the other at the interface of biology and chemistry.

In section 1, we study in detail the kinetics of fast-folding reactions following pressure-jump perturbation and pair experiment with molecular dynamics simulations. The first chapter is a review of the effects of pressure on the structure of biomolecules as well as a brief literature review of pressure-probed protein folding kinetics. We see that the methodology to study pressure-jumps is generally limited by time-scale—very fast folding is hard to study by pressure—and chapter two presents an overview of a fast pressure-jump instrument that meets this challenge. Although this instrument was developed by the previous generation of graduate students, several significant improvements are summarized in the chapter with a detailed user manual for the instrumentation. Closing up the section, we use the fast pressure-jump instrumentation as well as temperature-jump instrumentation to study the microsecond pressure and temperature-jump refolding kinetics of the engineered WW domain FiP35, a model system for beta sheet folding. With a full complement of molecular dynamics experiments mimicking experimental conditions, we show that simulation and experiment are consistent with a four-state kinetic mechanism and highlight FiP35's position at the boundary where activated intermediates and downhill folding meet.

Section 2 focuses on the interface of biology and chemistry, where we study how the protein folding reaction is impacted by immersion in the crowded intracellular environment and explore whether perturbations to the intracellular folding landscape can be linked to protein function. A review of the forces at play in the intracellular environment and the role that ultra-weak “quinary” interactions play inside living cells is presented in chapter 4, which also includes a review of the most recent literature studying biomolecular dynamics in their native environments. In chapter 5 we study the time-dependence of protein folding inside living cells as probed by live-cell fluorescent microscopy. We find that both the rate of folding and the thermodynamic stability of yeast phosphoglycerate kinase (PGK-FRET) are cell cycle-dependent, a process strictly regulated in time, suggesting that the interplay between the intracellular environment and proteins may impact their function. In chapter 6, a new probe to study protein folding in the cell is explored, namely the GFP/ReAsH Forster resonance energy transfer (FRET) pair. We show that this FRET pair suffers from bleaching artifacts but that directly excited ReAsH is an appealing prospect for studying protein folding in living cells on fast and slow time-scales. Finally, chapter 7 builds on the work presented in chapter 5 and chapter 6 by seeking a protein candidate whose function and in cell folding dynamics are linked. Several constructs of p53, a transcription factor, are explored as potential candidates for answering the question of whether protein activity level indeed can correlate with stability in living cells.

ACKNOWLEDGEMENTS

I must thank first and foremost my advisor, Prof. Martin Gruebele, for his guidance over the past five years. I could not possibly list all the ways in which Martin has positively impacted my graduate career, but I especially want to express my gratitude for his receptiveness to new ideas and his scientific guidance. As I moved through my PhD it has become increasingly apparent how lucky I was to have an advisor like Martin—I cannot emphasize enough how grateful I am to have been a member of his research group.

Second, I owe much gratitude to those colleagues who trained me in the lab. Max Prigozhin taught me many of the techniques and instrumentation that have been the foundation of my PhD research. I'd especially like to thank Hannah Gelman for training me (and providing copious amounts of instrument support) on the live-cell instrumentation. Minghao Guo, as well, was particularly helpful with all sorts of instrumentation and analysis problems I experienced over the years. Irisbel Guzman Sanchez taught me how to do temperature-jump experiments as well as many biochemical techniques. I am also grateful to Max Platkov for his assistance with the PID.

I would also like to thank the members of the Gruebele group for their support and friendship: Krishnarjun Sarkar, Max Platkov, Eduardo Berrios, Kiran Girdhar, Ruopei Feng, Tatyana Perlova, Elaine Christman, Kapil Dave, Drishti Guin, Hannah Gelman, Maxim Prigozhin, Irisbel Guzman Sanchez, Minghao Guo, Jay Goodman, Shahar Suenik, Timothy Chen, Yangfan Xu, Lea Nienhaus, Duc Nuygen, and Sumit Ashketar. I am honored to have worked with such excellent colleagues. In particular, I am very thankful to have been office-mates with Hannah Gelman and Irisbel Guzman Sanchez who have helped me so much over the past five years.

I also must thank several collaborators. On the pressure-jump project, I was fortunate to collaborate with Yanxin Liu and Prof. Klaus Schulten. I am grateful for Yanxin's excellent work and for the opportunity to access the Schulten group's computational prowess and resources. On my ReAsH work, I had the pleasure of working with my colleague Hannah Gelman.

I am greatly indebted to many staff in the School of Chemical Sciences. I would like to give special thanks to the entire administrative staff in the physical chemistry office who have been incredibly helpful throughout the years: Karen Watson, Beth Myler, Connie Knight, Theresa Struss, and Stacy Dudzinski. The SCS machine shop has been an invaluable resource that has solved an uncountable number of problems and crises. Tom and Mike "Hodge" Harland, especially, have made a significant impact in the success of the pressure-jump project. In the High Throughput Screening Facility, Cheng Zhang was a valuable resource for many years on our mammalian cell culture experiments.

My first research experiences were at the College of William and Mary under the direction of Prof. Lisa Landino, and I am grateful to have had the opportunity to get this early start in research. Prof. Landino's mentorship has been very valuable to me. I am also grateful for the guidance of Prof. Joel Hockensmith at the University of Virginia, whose mentorship during the SRIP program helped me form the foundation of my understanding of molecular biology. I'd also like to thank Prof. Carey Bagdassarian at the College of William and Mary, who was the first to teach me the joys of using computational tools and modeling to better understand nature.

Last, I must thank my family for all their support and love. Most of anything I have achieved springs from my luck of being born into such a wonderful family. My husband's family has also been supportive and helpful far beyond the call of duty. Finally, my husband, Daniel, has been a calming, supportive presence in my life who has made the stresses of graduate school bearable and life in Champaign-Urbana fun—it's hard to imagine how I would have managed to finish my degree without him in my life.

TABLE OF CONTENTS

SECTION I: PROTEIN FOLDING UNDER PRESSURE	1
<i>CHAPTER 1</i>	1
PRESSURE PERTURBATION OF PROTEIN STRUCTURE	
<i>CHAPTER 2</i>	8
PRESSURE-JUMP INSTRUMENTATION TO STUDY PRESSURE-PROBED MICROSECOND FOLDING KINETICS	
<i>CHAPTER 3</i>	26
COMPARING FAST PRESSURE-JUMP AND TEMPERATURE-JUMP PROTEIN FOLDING EXPERIMENTS AND SIMULATIONS	
SECTION II: PROTEIN FOLDING IN LIVING CELLS	46
<i>CHAPTER 4</i>	46
QUINARY PROTEIN STRUCTURE AND THE CONSEQUENCES OF CROWDING IN LIVING CELLS	
<i>CHAPTER 5</i>	62
TEMPORAL VARIATION OF A PROTEIN FOLDING ENERGY LANDSCAPE IN THE CELL	
<i>CHAPTER 6</i>	83
DEVELOPMENT OF THE GFP-REASH FRET PAIR TO STUDY PROTEIN FOLDING	
<i>CHAPTER 7</i>	100
DEVELOPING CONSTRUCTS TO MEASURE P53 FOLDING INSIDE LIVING CELLS	
APPENDIXES AND REFERENCES	113
<i>APPENDIX A</i>	113
SUPPLEMENTARY INFORMATION FOR “COMPARING FAST PRESSURE-JUMP AND TEMPERATURE-JUMP PROTEIN FOLDING EXPERIMENTS AND SIMULATIONS”	
<i>APPENDIX B</i>	122
DONOR TO ACCEPTOR RATIO AS A REACTION COORDINATE FOR PROTEIN FOLDING IN LIVING CELLS	
<i>APPENDIX C</i>	126
SUPPLEMENTARY INFORMATION FOR “DEVELOPMENT OF THE GFP-REASH FRET PAIR TO STUDY PROTEIN FOLDING”	
<i>APPENDIX D</i>	128
PRESSURE-JUMP TROUBLE SHOOTING GUIDE	
<i>APPENDIX E</i>	131
LIVE CELL METHODS AND DATA ANALYSIS	
<i>REFERENCES</i>	134

SECTION I: PROTEIN FOLDING UNDER PRESSURE

CHAPTER 1 PRESSURE PERTURBATION OF PROTEIN STRUCTURE

1.1 MECHANISM OF PROTEIN PRESSURE DENATURATION AND THE PRESSURE UNFOLDED STATE

1.1.1 Pressure-probed folding thermodynamics For over 100 years, it has been known that application of high pressures has an effect on protein structure. The discovery was first made by observing the coagulation of albumin, from an egg white, under pressure¹. Since then, pressure has been used, although less popularly than temperature or denaturants, to probe structural transitions in proteins and other biomolecules, particularly oligomeric or aggregated proteins.

The effect of pressure on a protein's fold can be understood entirely through Le Chatelier's principle²: at high pressures, the energetically favored structure has the minimum volume. Generally speaking, the unfolded state of globular proteins is lower in volume than the folded state which gives rise to denaturation at high pressures.

The pressure dependence of the free energy of folding, ΔG , is given by:

$$\Delta G(P) = \Delta G_0 + \Delta V_0(P - P_m) - \frac{V\Delta\beta_T}{2}(P - P_m)^2$$

where ΔV_0 is the volume change upon unfolding, P_m is the pressure unfolding midpoint, and $\Delta\beta_T$ is the change in isothermal compressibility upon unfolding. Thus, while temperature denaturation accesses the enthalpy and heat capacity of a protein, a pressure denaturation thermodynamics experiment provides information on the volume change upon unfolding as well as the isothermal compressibility of the protein. Figure 1 shows a generic phase diagram for temperature and pressure denaturation of a protein and the trend in volume change exhibited for both denaturation methods.

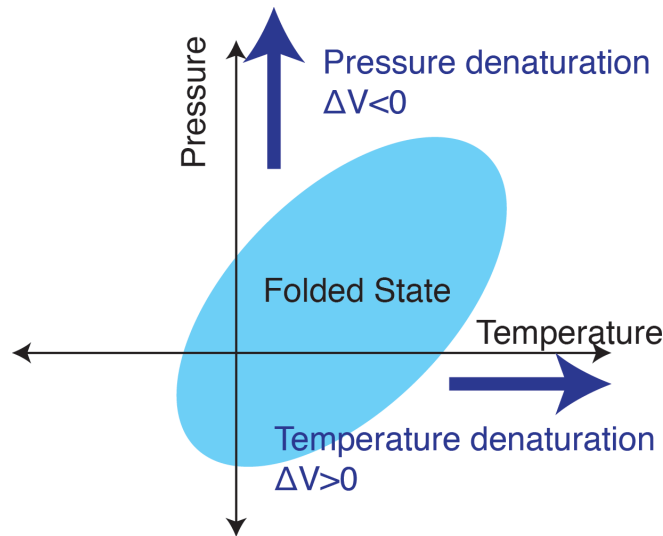


Figure 1 Temperature pressure phase diagram. Area in light blue shows the region of the phase diagram where the protein is folded. The border of the ellipse shows the T_m or P_m of the protein at any pressure or temperature, respectively.

1.1.2 The pressure unfolded state The contrast between the dependence of the free energy on temperature and the dependence of the free energy on pressure provides some insight into the often nuanced effect pressure can have on protein structure. While temperature denatures proteins by perturbing both volume and energy, pressure perturbs only volume, making it a somewhat more gentle denaturation method that can leave secondary structure behind in the unfolded state. Often, the radius of gyration of the pressure-unfolded ensemble is lower than what would be expected for a true random coil³.

In an extreme example of residual secondary structure in the pressure-denatured state, a transcription factor in its pressure denatured state retained its globular structure—thus appearing fully folded by NMR measurements—but had its core completely penetrated by water—thus appearing completely unfolded by tryptophan fluorescence measurements⁴. Numerous examples of the pressure denatured state showing residual secondary structure, particularly alpha-helical structure, have been reported across the literature^{3,5,6}. Indeed, it was shown that the volume change of unfolding for an isolated helical peptide is positive, suggesting that in some cases helices are stabilized at high pressures⁷.

1.1.3 Origins of pressure-induced unfolding Until recently⁸, the origin for the difference in volume between the folded and unfolded states of globular proteins (and, thus, the variation in the sensitivity of various protein folds to pressure denaturation) was poorly understood. Because protein

solutions are not ideal solutions, the change in molar volume upon unfolding is dependent on hydration effects as well as changes in volume due to the disruption of internal cavities that exist in the folded state. Solvent effects alone, including unfavorable solvation of non-polar groups at high pressures driving their packing in the core of the protein⁹ and changes in the bulk structure of water at high pressures¹⁰, cannot alone explain the difference in the molar volume between the folded and unfolded state². A recent exhaustive high pressure NMR study of cavity forming mutants of staphylococcal nuclease showed that the susceptibility of proteins to high pressure derives mostly from the presence of internal cavities in the folded state⁸. Thus, the general consensus in the literature is that the higher the volume of internal cavities in a protein or the higher degree of packing defects in the fold, the more susceptible a protein is to pressure denaturation.

1.2 PRESSURE-PROBED PROTEIN FOLDING KINETICS

Pressure-probed kinetics can provide useful information about a protein's transition state through measurement of the activation volume, ΔV_f^\ddagger , which is related to the measured pressure-probed folding rate, k_f , by:

$$k_f = k_m e^{-p\Delta V_f^\ddagger / RT}$$

where k_m is the folding prefactor and p is the pressure. The ratio of the folding activation volume and the overall volume change of folding, called the “V-value”¹¹, provides an estimate in the change in hydration between the unfolded state and the transition state. This measure is essentially the pressure analog to the Φ -value and provides a complimentary method to traditional T-jump and stopped-flow experiments to understand the nature of the transition-state.

Because the molar volume of the transition-state ensemble is higher than the unfolded state (or, in other words, the activation volume is very positive), proteins unfold very slowly at high pressures and upwards pressure-jumps greatly slow reaction kinetics¹². The slowing of kinetics enables use of instrumentation that otherwise would have lacked the time resolution to observe the kinetics. For example, the folding of staphylococcal nuclease was measured with site-specific precision via upwards pressure-jumps monitored by multi-dimensional NMR¹³. Here, the rate of folding for each individual residue was resolved and used to show that pressure effects on the fold are observed in different regions of the protein structure depending on mutant and denaturant concentrations—remarkably specific mechanistic insight.

1.3 PRESSURE AND PROTEIN AGGREGATION

Aside from its obvious utility in studying protein folding from a mechanistic perspective, pressure has useful applications in understanding other biomolecules, including nucleic acids¹⁴ and oligomeric proteins¹⁵. One area of active research is pressure's ability to disaggregate non-productive protein aggregates including amyloid fibrils and bacterial inclusion bodies.

1.3.1 Probing amyloid fibrils with pressure Amyloid fibrils are a type of protein aggregate that is common across many diseases, particularly neurodegenerative disorders, including Parkinson's Disease, Alzheimer's Disease, and Huntington's Disease. Amyloid fibrils are characterized by a beta-sheet topology that varies depending on the protein sequence that gives rise to the fibril. Until the past decade with the expansion of solid-state NMR techniques, amyloid fibrils were not understood structurally at the molecular level due to their poor crystallization and incompatibility with solution NMR¹⁶.

Pressure, however, offers an interesting, and relatively straightforward, way to probe the general attributes of fibril structure. Because the structural stability of biomolecules to pressure denaturation depends on internal cavities, the resistance to pressure denaturation of fibrils can be used to probe the packing efficiency of the fibril structure¹⁷. Indeed, the stability of amyloid fibrils to application of high pressure is not uniform, reflecting the known diversity of fibril structure. In some cases, amyloid fibrils were found most sensitive to pressure denaturation early in the fibril formation process before eventually growing over more time resistant to even very high pressures, presumably due to internal reorganization that occurs over time and reduces packing deficiencies in the fibril structure^{16,18}. In most other cases¹⁴, however, amyloid fibrils were found to be very sensitive to pressure denaturation, the behavior of hen lysozyme fibrils being typical which shows accelerated dissociation under pressures as low as 50 MPa¹⁹.

Even slight differences in fibril structure, as might be expected between amyloid fibrils formed from two mutants of the same protein, give rise to varying resistivity to applied pressure. A study on Alpha-Synuclein found that fibrils made from wild-type Alpha-Synuclein were more likely to resist denaturation than the mutant varieties that lead to inherited forms of Parkinson's Disease due to structural differences between the two types of fibrils²⁰. This implied that the in cell protease machinery was more likely to be able to break up the mutant amyloid fibrils into smaller aggregates (which are thought to be more neurotoxic than the full-length fibrils), offering a possible

explanation for why those with inherited mutations have a more severe form of Parkinson's Disease than those with sporadic cases.

High pressure is also useful to study aggregation kinetics because it can often stabilize partially folded intermediates that otherwise would be too transient to observe. Aggregation or amyloid formation is often thought to originate from partially folded, off pathway intermediates, and high hydrostatic pressure has been proposed as a way to access these intermediates in order to better understand the early steps of the amyloid formation²¹. For example, in the case of transthyretin, high pressure disaggregates fibrils into a partially folded intermediate that reaggregates upon return to atmospheric pressure more rapidly than the monomeric protein under aggregation-promoting conditions²⁰.

1.3.2 High pressure to purify proteins from inclusion bodies A common problem faced in the field of biotechnology are proteins that aggregate in bacterial expression systems. Rather than remain soluble in the bacterial cytoplasm, some proteins aggregate and amass into highly insoluble inclusion bodies from which purification of the target protein of interest is extremely difficult. The ability to purify difficult to express proteins at high yield is of great importance in the pharmaceutical industry, particularly as the development of biologics-based drugs becomes more prevalent.

Pressure's ability to disaggregate amyloid fibrils translates to the solubilization of inclusion bodies. Some of the first proteins to be folded from inclusion bodies by use of high pressure were human growth hormone, lysozyme, and beta-lactamase²². In many cases, very low concentrations of GuHCl and hydrostatic pressures of 20 MPa were sufficient to complete break up insoluble aggregates, and a return to GuHCl free buffer and atmospheric conditions resulted in complete refolding of the protein.

1.4 BRIDGING PRESSURE-JUMP EXPERIMENTS AND SIMULATIONS THROUGH LOW DEAD-TIME PRESSURE-JUMP INSTRUMENTATION

While simulated temperature-jumps²³ and long MD simulations at high temperatures where multiple folding and unfolding transitions can be observed^{24,25} are relatively common place, simulations at high pressures^{26,27} or of pressure-jumps^{28,29} are rarer, despite the utility of comparing pressure-perturbation results with experiment to improve the performance of force fields. The reason for the paucity of pressure-jump simulated data is that, until recently, there were no experimental reports of pressure-probed kinetics with rates below a millisecond and, thus, no

experimental benchmark available for comparison to the simulated data. Although the pressure-probed kinetics of several protein systems have been thoroughly studied, namely the ankyrin repeat protein^{30,31}, staphylococcal nuclease^{11,13,32,33}, and lambda repressor^{28,34,35}, only the lambda repressor has a kinetic phase less than a millisecond.

The biggest challenge to measure the pressure-induced kinetics of fast folders is instrumentation. Figure 2 summarizes the major types of pressure-jump instrumentation and highlights the differences in the size of the pressure-jumps the instrumentation is capable of administering and the fastest kinetics that can be resolved. The most common strategy to administer pressure changes is through the use of valve-based instrumentation. While this enables large jumps in pressure, the dead-time of most electrically-controlled valves or “fast-valves” in figure 2 is on the order of 5 milliseconds³⁶. For a downward pressure-jump, this limits the studied proteins to those that fold on the order of 10’s of milliseconds.

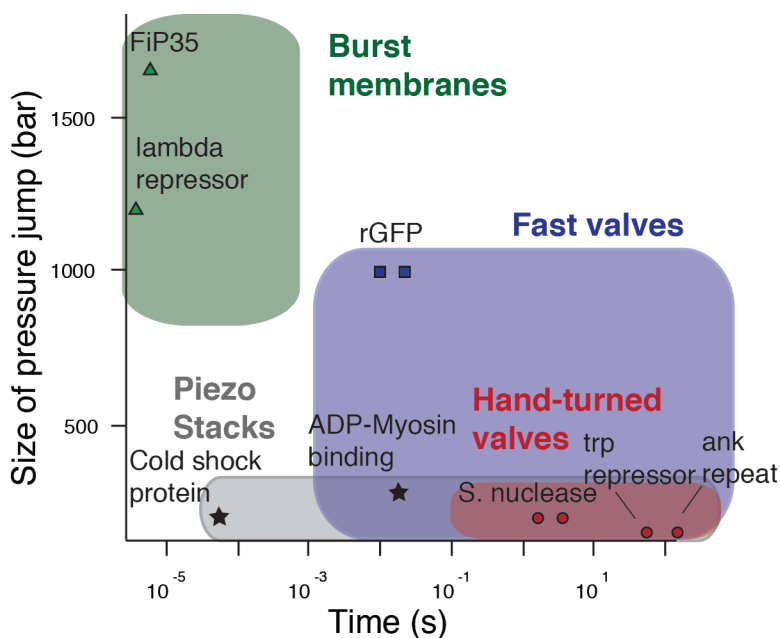


Figure 2 Summary of the three major types of pressure-jump instrumentation. For each method, the ranges of the pressure-jump size administered and kinetic rate resolution possible are indicated by the borders of the colored regions. Examples of kinetic rates observed for each method type are shown. The references for the various proteins are: Staphylococcal nuclease^{11,32}, ankyrin repeat protein³⁰, trp repressor⁶, lambda repressor^{28,35}, Myosin-ADP binding³⁷, and cold shock protein³⁸. Fip35 is the subject of chapter 3 of this thesis.

One useful strategy to circumvent the time-scale limitations of valve-based instrumentation is to employ upwards pressure-jumps. At high pressures, protein folding reactions proceed at a much slower rate, so instrumentation dead-time becomes less critical for resolving reaction kinetics.

In this case, hand-controlled valves can be used to administer jumps of up to a few hundred bar and monitor reaction kinetics with rates as fast as a few seconds (red highlighted area in figure 2). Upward pressure-jumps are particularly useful because reactions can proceed so slowly that kinetics can be monitored by multi-dimensional NMR, which allows residue by residue monitoring of protein folding transitions¹³. However, while slowing down the kinetics makes them easier to measure, this strategy is incompatible with directly comparing experiment to molecular dynamics simulations.

Chapter 2 of this thesis describes in detail a burst-membrane based approach to measure fast folding kinetics following pressure-drop. As is highlighted in figure 2, burst-membrane based instrumentation can access protein folding rates on the order of microseconds. In Chapter 3, this instrumentation is used to study the fast folding of FiP35 following pressure-jump and the resulting kinetics are directly compared to molecular dynamics simulations of pressure-jumps.

CHAPTER 2

PRESSURE-JUMP INSTRUMENTATION TO STUDY PRESSURE-PROBED MICROSECOND FOLDING KINETICS

2.1 INTRODUCTION

Previously, the Gruebele group developed a burst-membrane based pressure-jump instrument that enables very large pressure drops—exceeding 2000 times atmospheric pressure—with an instrument dead-time of only 1 μ s. With such resolution and jump size, the instrument has been used to measure the pressure-probed folding kinetics of fast folders including lambda repressor^{28,34,35} and the WW-domain mutant FiP35 (Chapter 3). In the case of Lambda repressor^{28,35} and FiP35 (Chapter 3), the resolved pressure-jump kinetics were fast enough to be compared directly to molecular dynamics simulations, which is only possible due to the microsecond dead-time of the pressure-jump instrumentation. From the original prototype reported in Nature Methods³⁴, a number of improvements to the instrumentation were made that enabled greater ease of use and, accordingly, higher throughput. Building upon these improvements, the work presented in chapter 3 represents an approximately 5-fold increase in the number of measurements obtained in a typical experimental run. The first section of this chapter summarizes the major instrument improvements and provides rationale for why they were needed. The remaining sections provide a detailed overview of assembly and operation of the pressure-jump instrumentation. Appendix D includes a troubleshooting guide for users which covers all major (and minor) problems encountered over the past 5 years.

2.2 IMPROVEMENTS MADE TO THE PRESSURE-JUMP INSTRUMENTATION

2.2.1 Laser stability Originally, the pressure-jump apparatus was mounted onto the optical table that is shared by the Ti:Saph laser, which provides (see section 2.4) the excitation source for fluorescence detection of protein folding. In order for the laser signal to be frequency tripled to tryptophan's 280 nm excitation maximum and to effectively trigger signal collection on the scope, the laser is mode-locked with a pulse train at 80 MHz. The stability and reliability of this mode-lock can be perturbed by shocks to the optical table.

The pressure drop instrumentation for jumps greater than 1000 bar delivers a significant shock to the table that often perturbs the mode-lock enough to interfere with signal collection. Several strategies were employed to avoid this problem including mounting the pressure-jump

apparatus onto a piece of shock absorbing honeycomb steel and attaching this breadboard to the optical table with shock absorbing rubber feet. While this was effective, the limited lifetime of the rubber feet (<1 year) made this approach non-sustainable.

The solution that has proved most effective in the long term is mounting the pressure-jump instrument to its own pedestal separate from the optical table. This pedestal is firmly mounted to the concrete floor via screws and the only connections between the shock-producing pressure-jump instrument and the optical table are flexible wires and a pmt light-guide. By mechanically isolating the pressure-jump assembly from the sensitive optical equipment, large jumps can now be accessed (as high as 3000 times atmospheric pressure), without laser disruption.

2.2.2 Signal corruption One persistent problem with the pressure-jump instrument has been random electrical noise evident in the fluorescent signal. This typically manifests itself as large baseline rolls in the signal that appear right around the jump. Many different hypothesis for the origin of this signal corruption were tested and ruled out including: 1) Contamination of the fluorescence signal by stray light from sparks (which occur when the membrane breaks), 2) Mechanical oscillation of the pressure-jump assembly causing the laser path length through the sample to change, 3) Mechanical disruption of the laser signal, and 4) Electrical artifacts from the capacitor bank discharge travelling to electrical equipment through the metal optical table. Ultimately, it was determined that the problem arises from electrical interference originating from the extremely large capacitor bank discharge partially through the air and partially through the electrical connections. The PMT appears to be exquisitely sensitive to these random electrical signals. The best strategies to combat this kind of electrical interference are to a) use the Faraday cage around the capacitor bank and P-jump assembly, b) ensure that the fluorescence signal is extremely robust (>100 mV on the scope) so that interference doesn't overwhelm the signal, and c) use signal processing to remove some electrical noise that breaks through the signal. These strategies are discussed in detail in Appendix D.

2.2.3 Equipment changes A number of changes were made to the equipment used throughout the instrument including the materials of the pressure fitting, the method the sample cavity is machined into the sapphire cube and the design of the non-conductive cylinder. These improvements are outlined in section 2.3.

2.3 PRESSURE-JUMP INSTRUMENT COMPONENTS

2.3.1 Pressure fitting

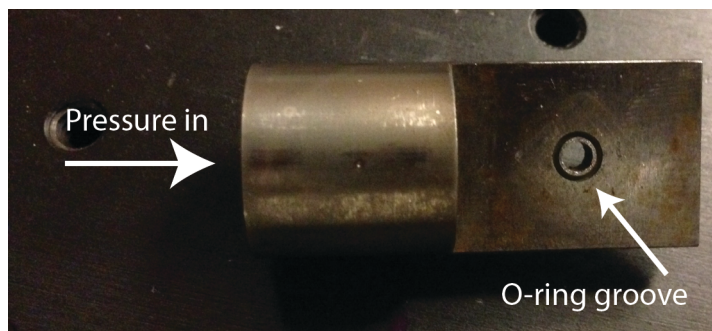


Figure 3 Pressure fitting showing groove where the O-ring fits and pressurization fluid inlet. The liquid emerges through the hole surrounded by the O-ring groove, which is placed above the sample.

What it does: Guides pressurization fluid to sample

Replacement guidelines: The area around the O-rings becomes corroded due to the electrical discharge (which can essentially weld the steel burst membrane to the pressure fitting) and this can lead to O-ring damage. When O-rings start breaking frequently, it's time to replace the pressure fitting. These shouldn't need to be replaced more than once a year with typical instrument usage. The SCS machine shop makes this part and has a diagram on file. It should be made of steel and then hardened. Titanium is also acceptable.

2.3.2 Sapphire cube

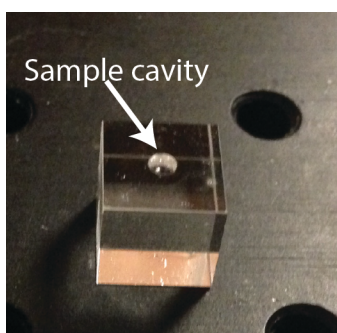


Figure 4 Sapphire cube with machined sample cavity. The sample cavity typically holds 5-10 μL of sample. The diameter of the sample cavity is slightly smaller than the diameter of the o-rings, enabling the bottom o-ring to form a seal around the edge of the sample cavity.

What it does: Holds the sample. Sapphire is used because it is strong and optically transparent.

Replacement guidelines: Sapphire cubes can be used until they break. Even cracked cubes can sometimes continue to be used for a long time provided the crack does not cross through the sample cavity. We order the sapphire cubes from Esco products and they have the measurements and specification on file. The sample hole is machined into the cube by the machine shop. The hole should be smaller in diameter than the O-ring and on the strong face of the sapphire. Finding the strong face:

- 1) Put a polarizer on a flashlight and turn on the flashlight
- 2) Set a cube on the polarizer
- 3) Hold the other polarizer on top and rotate it.
- 4) Watch the top face of the cube. If you have found a strong face, there will be a spectrum of many colors visible as you rotate the polarizer. The opposite face will have the same spectrum.
- 5) A weak face shows no color spectrum.

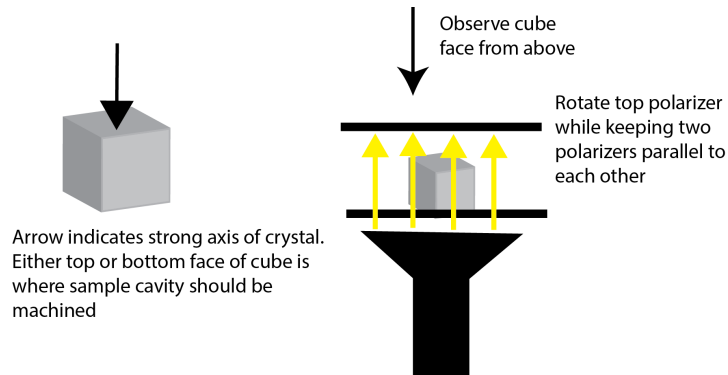


Figure 5 Finding the strong axis of the sapphire cubes. As shown in the above diagram, two polarizers and a flashlight can be used to identify the directionality of the crystal so that the strongest axis is used for machining of the sample cavity.

2.3.3 Burst membrane

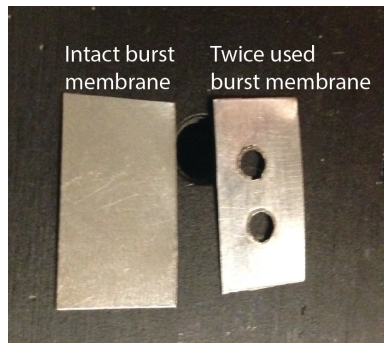


Figure 6 Burst membrane. Left membrane has not been used while the right membrane has been used twice, as evidenced by the two burst holes. The clean edges of the burst holes indicate that the jump was likely with a low dead-time.

What it does: Breaks when capacitor charge is delivered and allows the system pressure to drop very rapidly.

Replacement guidelines: These can be used twice (put the membrane into the assembly another way so that un-burst steel is over the sample cavity). We buy stainless steel shim stock to cut into burst membranes from McMaster Carr. One order from McMaster Carr will supply enough membranes

for several hundred or more jumps. In a pinch, the SCS machine shop usually has some shim stock. 5-7 thousandths of an inch thick stock can be cut with scissors into membrane sized pieces. These thicknesses access the range of jumps from 1000 to 2000 bar. Thicker shim stock can be cut in the machine shop with shears. The membrane should be about the width of a cube and 2/3 the length of the tongue of the pressure fitting.

2.3.4 Foil



Figure 7 Foil used to protect the sample from contamination by pressurization fluid. Visible here is the folded piece of aluminum coated Mylar that wraps around a piece of foil.

What it does: Protects the sample from the pressurization fluid and blocks sparks that form when the membrane bursts from contaminating the fluorescence signal.

Replacement: These need to be replaced every other jump or so. To make, wrap a piece of aluminum coated Mylar around a piece of foil. It helps to wet the Mylar so it sticks to the foil. To restock foil, Sigma sells many kinds of foil that can be used for the inner piece. The aluminum wrapped Mylar is an emergency blanket (the kind that marathon runners use after races).

2.3.5 Electrode



Figure 8 Copper electrode. Tips on either end are sharpened in order to cleanly break the burst-membrane. The double ended design allows a single electrode to be used twice before needing to be re-machined.

What it does: Copper electrode with sharpened tips on both sides. Delivers capacitor discharge to burst membrane.

Replacement guidelines: These can only be used once per side before they need to be re-machined. The end that comes into contact with the burst membrane must be sharp, or the break will not be clean.

The machine shop makes this part and can re-furbish old electrodes until they are too short to reach from the top of the pressure assembly to the burst membrane. A part diagram is on file at the machine shop and the material used is copper.

2.3.6 Non-conductive cylinder



Figure 9 Non conductive cylinder. This cylinder is seen from the top, where the electrode enters the cylinder before making contact with the burst membrane. The groove is meant to enable liquid that escapes during the pressure drop to travel away from the electrode.

What it does: Specially treated, heat and pressure resistant piece of ceramic that guides the electrode to the burst membrane. As the burst membrane is weakened by the pressure, it pushes up into the bottom of the cylinder so the cylinder also determines the size of the hole that is burst in the burst membrane.

Replacement guidelines: These can be used for approximately 50 jumps. When the edges of the hole on the bottom are no longer sharp, the jump resolution will suffer and its time for the cylinder to be replaced. These are ordered from AstroMet and all the information should be on file there. It is essential that the bottom of the through hole of the cylinder (where it contacts the burst membrane) is sharp and does not have a chamfer. Otherwise, sapphire cubes will break and jump resolution will be poor.

2.3.7 O-ring

What it does: Provides a seal between the pressure fitting and the foil and the pressure fitting and the burst membrane.

Replacement guidelines: These last a variable amount of time. Only replace if they are obviously broken or pressure is not holding well in the assembly. They are .158 x .02090 o-rings made of fluorocarbon from Apple Rubber Products (part number R00158-020-90VTB).

2.3.8 High pressure tubing, pressure fitting connector, and nipple

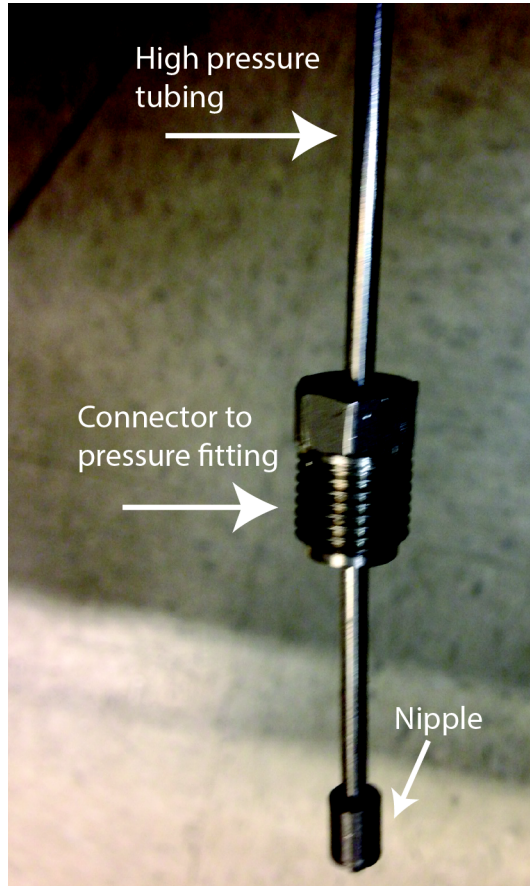


Figure 10 High pressure tubing diagram showing the end of the tubing. The nipple is attached via threading on the end of the high pressure tubing. The connector to the pressure fitting slides over the nipple, enabling a tight seal to form between the pressure-jump instrument and the high pressure tubing.

What it does: Connects pressure generator to the pressure-jump assembly and delivers pressurization fluid.

Replacement guidelines: Replace the tubing when it starts functioning poorly. Common problems are leaks, usually where the tubing has been bent, and clogs. The tubing can be purchased from High Pressure Equipment Company (HiP). Make sure to buy tubing that is rated for at least 3000 bar. The channel is machined here to add threads to either end so that the nipple end piece can be attached. Nipples and pressure fitting connectors rarely need replacing.

2.4 SAFETY

2.4.1 Capacitor bank

The capacitor bank is an electrocution risk. Always make sure to discharge and turn off the capacitor before touching anything connected to it. Even if it discharged during the experiment, discharge

again (by switching the discharge button) and turn it off before disconnecting the assembly. It is a good practice to attach or detach the electrical contacts to the P-jump assembly with one hand. Never try to increase the pressure or adjust any aspect of the P-jump assembly when the capacitor bank is charged and connected to the assembly!

2.4.2 Laser

The laser beam moves upwards through the pressure assembly and towards your face, so take care when aligning the beam. Always block the beam path when you are working on the pressure assembly in case the shutter opens. Always keep the shutter closed (except during alignment), but bear in mind that static electricity can (and does) occasionally cause the shutter to spontaneously open and close.

2.4.3 Pressure

A pressure release not initiated by a capacitor discharge as might occur during pressurization is not violent but does release a lot of pressurization fluid like a little geyser. When using ethanol as a pressurization fluid and working close to the pressure assembly, it's a good idea to wear glasses or goggles. The pressure release with the capacitor discharge generates significant sparks, so no one should be near the pressure assembly during an experiment. Keep flammable materials (for example, ethanol soaked tissue) away from the pressure assembly.

2.5 OPERATION INSTRUCTIONS

2.5.1 System overview

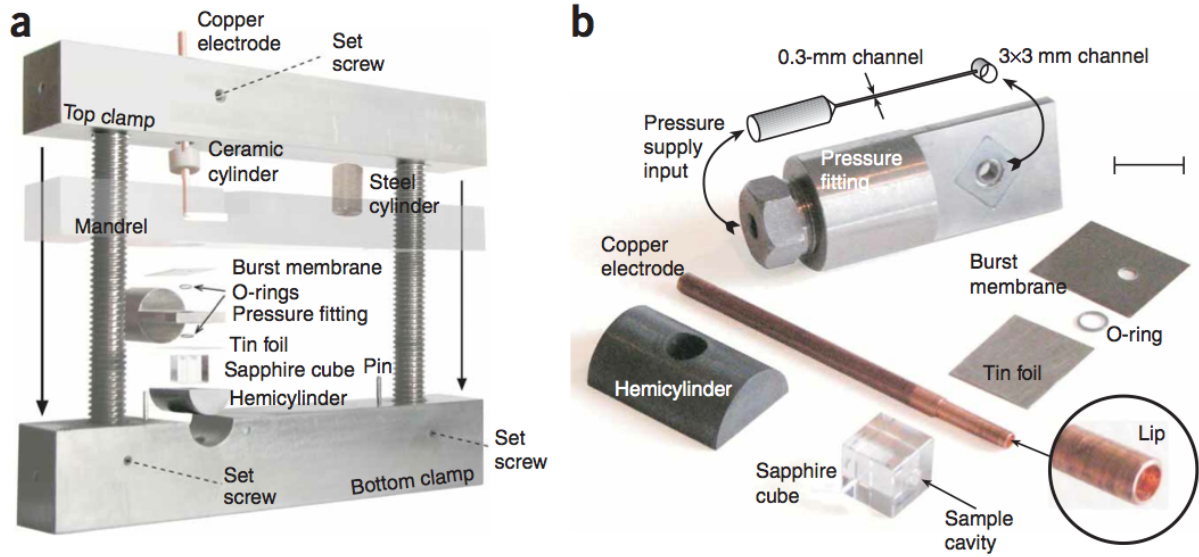


Figure 11 Pressure-jump assembly overview showing parts and their assembly. Figure from ref.³⁴

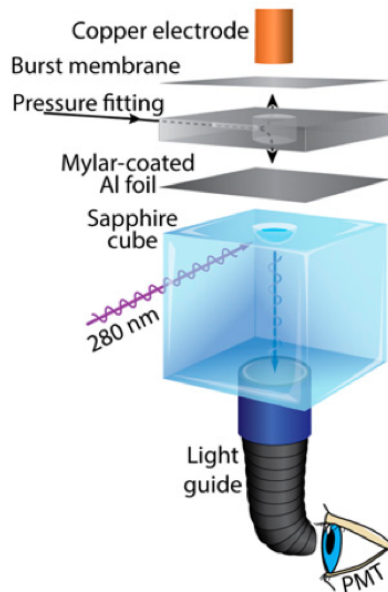


Figure 12 Schematic showing excitation and emission collection geometry relative to P-jump components. Figure from ref.³⁵

2.5.2 Alignment of the laser in the pressure assembly

- 1) Get laser aligned, mode locked, and stable. Power from the tripler should be at least 10 mW.
- 2) Send the laser through the optics between the shutter and the pressure assembly. Make sure that the beam is centered in every mirror and the lens.
- 3) Place a card in the empty pressure-jump assembly. Roughly align the laser to be centered where the top of the cube will be. Adjust the placement of the lens in the beam path, if necessary, so that the beam is focused in the center of the sample cavity.
- 4) Put a concentrated tryptophan solution into the sample cavity of the cube and put the cube into the mantle and into the assembly. Adjust the alignment using the final mirror so that the beam moving through the sample. The beam should be clearly visible in the tryptophan solution.
- 5) Check the PMT alignment by ensuring that there is plenty of signal with the tryptophan only solution. If the signal is low, check the placement of the pmt light guides both in the assembly and at the PMT itself.

2.5.3 Using the pressure generator

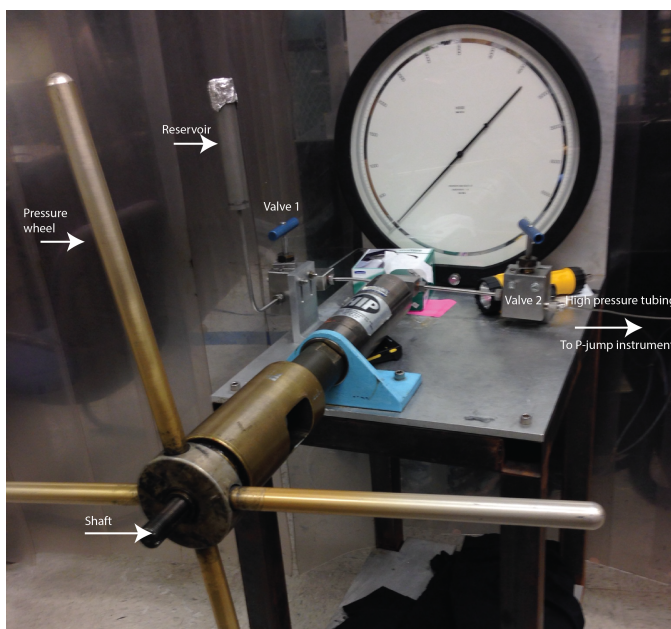


Figure 13 Hydrostatic pressure generator diagram with part labels. Pressure is read from gauge in center.

- 1) Make sure reservoir is filled with water or ethanol (either works fine, but if you switch between the two the system must be primed completely with the new pressurization fluid).

- 2) If the shaft is not visible, the system needs to be primed. If the shaft is fully exposed, the system is ready to use (in the diagram above, the system is ready to use).
- 3) To prime the system: Turn the reservoir valve (valve one) to the fully open position (turn counter clockwise until it stops). Turn valve two to the fully closed position (turn clockwise until it stops). Turn the pressure wheel counter clockwise until it stops. This fills the pressure generator tubing with pressurization fluid from the reservoir. Close valve one and open valve two. The system is primed and ready to go.
- 4) To pressurize: Make sure valve one is closed and valve two is open. Turn the pressure wheel clockwise and watch the outlet to make sure pressurization fluid comes out. Connect the pressure assembly and start pressurizing by continuing to turn the pressure wheel clockwise. Watch the gauge to monitor pressure.
- 5) To pressurize again, repeat steps 3 and 4.
- 6) NOTE: watch the shaft carefully. If it is not-visible, it's time to prime the system. If you turn too far, the wheel can come off.

2.5.4 Putting together the pressure assembly and pressurizing

- 1) Insert the cube into the mantle (“mantle” is called “mandrel” in the Nature Methods paper) and the mantle into the assembly.

Mantle in lowered position with cube inserted:

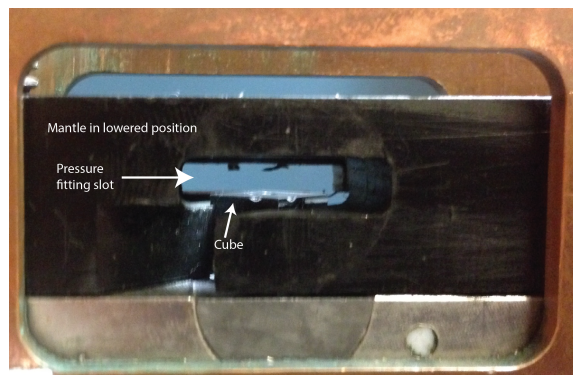


Figure 14 Assembly step 1. Cube is inserted with the slot for the pressure fitting clearly visible.

- 2) Carefully add just enough sample to fill the sample cavity with a pipette (access through the top).

3) Raise the mantle about .5 cm, leaving the cube in place.

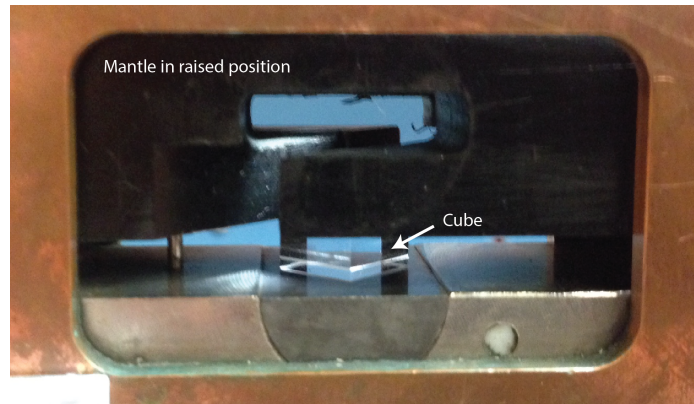


Figure 15 Assembly step 3. Mantel in raised position to facilitate easy insertion of the pressure fitting.

4) Place the foil carefully into the pressure fitting slot with tweezers. The foil should be slightly narrower than the pressure fitting and about 2 or 3 times as long as the width of the cube.

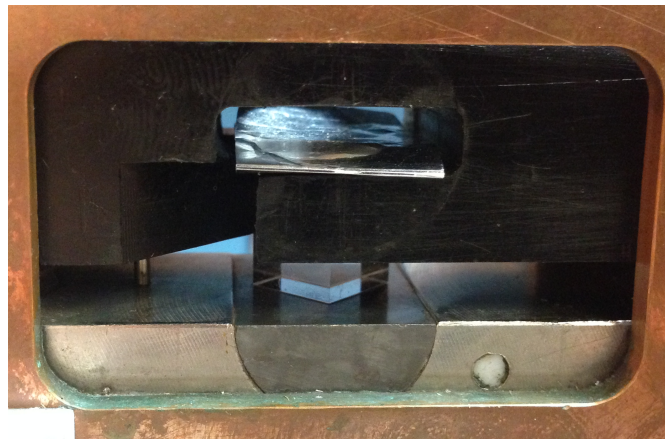


Figure 16 Assembly step 4. Mantel in raised position with foil inserted.

5) Make sure the o-rings are in place on the pressure-fitting. Place the burst membrane on top of the o-ring on the top of the pressure fitting.



Figure 17 Assembly step 5. Pressure fitting with burst membrane placed on top.

6) Carefully insert the pressure fitting with the burst membrane into the assembly on top of the foil. The burst membrane is on top of the tongue of the pressure fitting, as in figure 17. Try not to disturb the foil as you insert the pressure fitting. The pressure fitting should go all the way in until the round edge of the fitting is flush with the pressure-jump assembly.

7) Push the mantle all the way down so that the foil comes into contact with the cube. At this point, do not shift the pressure fitting or it will dislodge the sample from the cube.



Figure 18 Assembly step 7. Mantle pushed down with pressure fitting inserted.

8) Insert the non-conductive cylinder and the metal cylinder. The non-conductive cylinder should touch the top of the burst membrane.

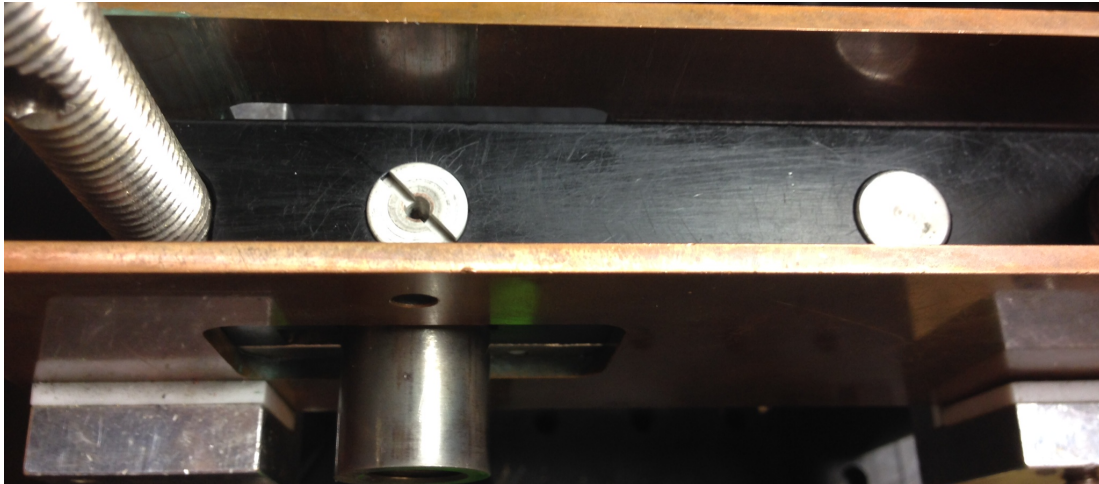


Figure 19 Assembly step 8. View from the top of the assembly with non-conductive cylinder and metal cylinder inserted

9) Put the top clamp on top of the mantle. Make sure that the hole through the top aligns with the hole through the non-conductive cylinder and that the set screw on the side of the top clamp aligns with the hole in the copper side of the assembly directly above the pressure fitting. Slide the washers and bolts into place and tighten with a wrench. Try to tighten both sides evenly.

10) Insert the pressurization tubing into the pressure fitting. Tighten using a wrench. System is now ready for pressurization.

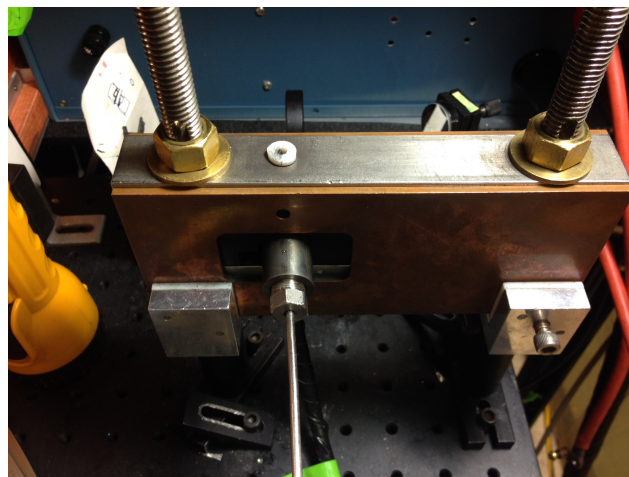


Figure 20 Assembly step 10. Assembly with top clamp tightened and pressure tubing attached.

2.5.5 Calibrating burst membranes

- 1) When to calibrate: Every time a new non-conductive cylinder is used, when there are significant changes in the thickness of the protective foil, or when a different size jump is desired, the system needs to be re-calibrated. Occasionally, changing the pressure fitting, pressurization fluid, or cube is enough of a perturbation that the system should be recalibrated.
- 2) Assemble the pressure-jump assembly and increase the pressure until the membrane breaks. There is a stainless steel “cube” which can be used to ensure that no cube breakage occurs during calibration.
- 3) Note the pressure at which the membrane broke. Repeat this three or more times.
- 4) The pressure to pressurize to for a pressure-jump experiment should be 50 bar less than the average pressure at which the membrane broke.

2.5.6 Doing a pressure-jump experiment

- 1) Assemble the pressure-jump assembly and pressurize to 50 bar less than the average pressure at which the membrane broke.
- 2) Put the electrode in through the top of the assembly. Make sure it makes contact with the burst membrane.

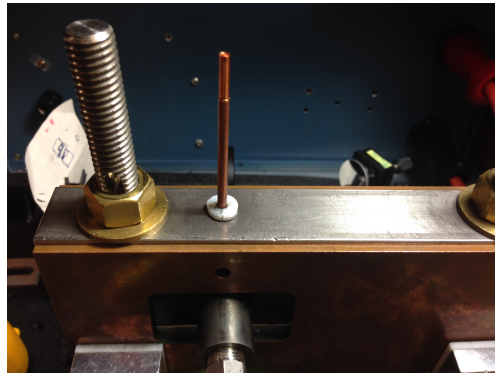


Figure 21 Inserting the electrode.

3) Secure the electrode into place by tightening the set-screw.

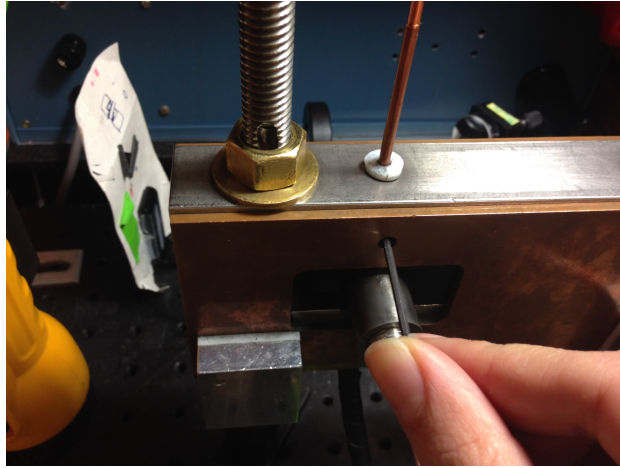


Figure 22 Fixing the electrode into place by tightening the set screw.

4) Attach the capacitor outlet to the electrode.

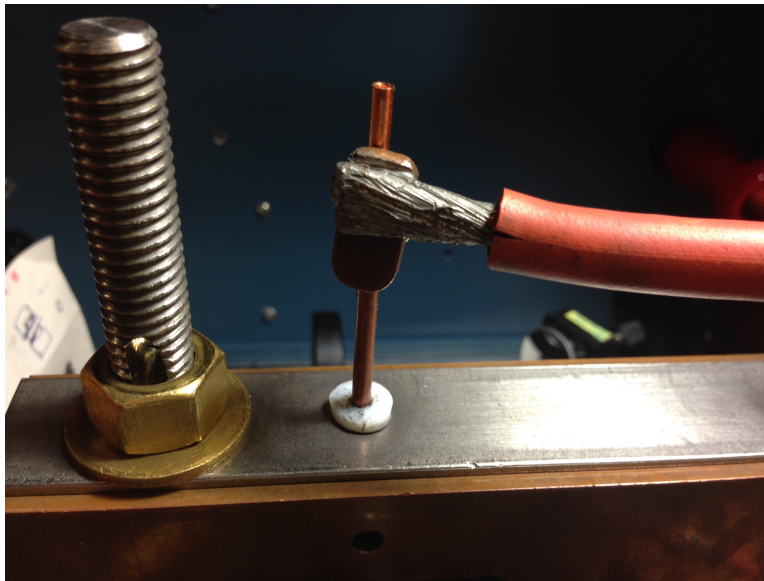


Figure 23 Capacitor outlet attached to the electrode.

5) Attach the grounding wire hub to one of the posts. Attach the alligator clips as shown in diagram. Pressure fitting, both screws on the bases, and the tall screw posts should be grounded.

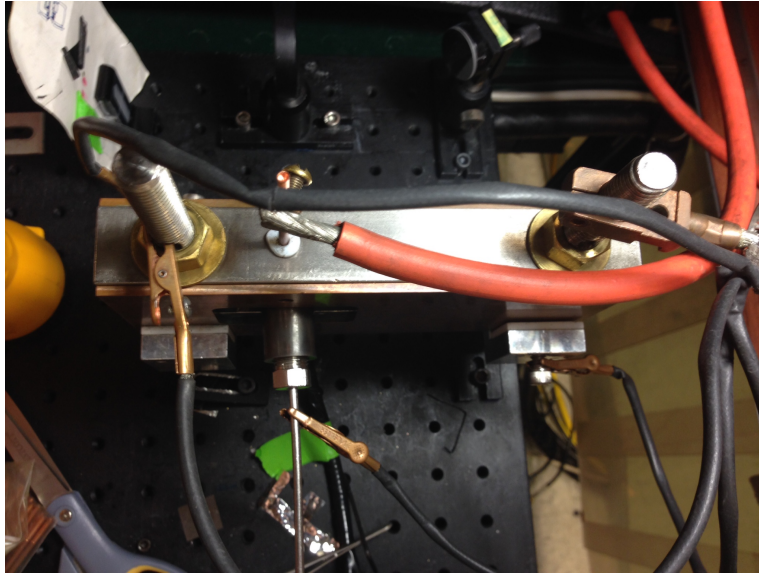


Figure 24 Assembly with all four grounding wires attached.

5) Turn on capacitor bank. Switch to “charge” and adjust voltage dial to maximum. Wait for voltage to read 100 V (or desired voltage) and turn voltage dial down to 0. **SAFETY NOTE: DO NOT TOUCH PRESSURE-JUMP ASSEMBLY AT THIS POINT BEFORE DISCHARGING THE CAPACITOR BANK.**



Figure 25 Capacitor bank control panel with key controls labeled.

- 6) Turn PMT on and turn voltage to 640 mV.
- 7) Set scope to collection mode to “single” and set trigger to “aux”. Turn scope to PMT channel and change time scale to allow for a full 5 ms of data collection.
- 8) Press control button to initiate jump. If you hear a snap and a hiss, the capacitor discharged and the membrane broke.

2.6 BEFORE BEGINNING EXPERIMENTS ON A NEW PROTEIN

- 1) Thoroughly understand the pressure unfolding thermodynamics of the system. If the protein is <30 or 40% unfolded at 2000 bar, guanadinium or urea will need to be added to destabilize the protein enough to see a significant enough population change over a jump.

- 2) Make sure there is a robust lifetime change upon unfolding. This can be accomplished by doing lifetime probed thermodynamics in the ISS pressure cell (pressure denaturation) or in the T-jump assembly (temperature denaturation). Any protein that has previously been studied by T-jump is likely to be a good candidate for P-jump, from a spectroscopic perspective, but success is not guaranteed as the pressure denatured state is often different from the temperature denatured state.

- 3) Consider the concentrations and conditions under which the protein of interest aggregates. Typically, concentrations of at least 200 μM are needed for a robust signal in the P-jump instrument.

- 4) Try to choose proteins whose kinetics are likely to be within a good timescale for the P-jump apparatus: τ_{obs} greater than 1 μs but less 2 ms.

CHAPTER 3

COMPARING FAST PRESSURE-JUMP AND TEMPERATURE-JUMP PROTEIN FOLDING EXPERIMENTS AND SIMULATIONS*

3.1 INTRODUCTION

The detailed mechanisms of protein folding reactions are becoming amenable to a direct comparison between theory, simulation and experiment.²⁹ FiP35, a 35 residue hybrid of the FBP and Pin1 WW domains,³⁹ is a triple stranded beta sheet model protein with a small but well-defined hydrophobic core. FiP35 refolds very rapidly, making this WW domain a popular target for folding experiments and molecular dynamics (MD) simulation.⁴⁰

Experiments indicate that at its melting temperature FiP35 is an apparent two-state folder, while well below its melting temperature it approaches downhill folding with a ‘molecular rate’ $k_m = (t_m)^{-1} = (0.7-2 \text{ ms})^{-1}$.^{41,42} (t_m is approximately the time required to cross the transition state region on a coarse-grained one-dimensional free energy landscape.⁴²⁻⁴⁴) Indeed, a variant of FiP35 (“GTT”), whose relaxation time of ≈ 3.7 ms nearly reaches the molecular time t_m , has been experimentally characterized and simulated.⁴⁵ MD trajectories well below T_m are not yet available for direct comparison with experiment. Long MD trajectories at T_m reveal a low barrier, consistent with experiments at T_m . The computed barrier lies between 1 and 5 kT on a one-dimensional free energy surface, depending on the analysis.^{25,41,42,46} A one-dimensional description is not complete. Markov analysis of multiple trajectories near T_m ,⁴⁷ and implicit solvent simulations of the Beta3s peptide⁴⁸ (structurally similar to the WW-domain), reveal complex kinetics among multiple short-lived intermediates prior to the main activated folding step. Thus the experimentally measured molecular time t_m probably lumps together into one number a complex network of fast processes whose 0.7-2 ms time scale collectively accounts for the time required to cross the transition state region on a rough free energy landscape.

The power of all-atom molecular dynamics (MD) simulations to illuminate protein folding experiments remains limited by time scale: It is computationally expensive to integrate the equations of motion for tens of thousands of atoms over the complete duration of protein refolding.²⁵ Until recently, only temperature-jump-induced kinetics (T-jumps) were quick enough to compare with

* This chapter is adapted with permission from Anna J. Wirth, Yanxin Liu, Maxim Prigozhin, Klaus Schulten, and Martin Gruebele, *Journal of the American Chemical Society*, 137 (22), 7152-7159, 2015. Copyright 2015 American Chemical Society. YL performed all molecular dynamics simulations referenced in this chapter and generated the data shown in figures 30 and 31. MP provided instrumentation support to this publication.

single-trajectory simulations.^{24,49} Now instrumentation improvements have enabled microsecond resolution pressure-jump experiments (P-jumps).^{28,34,35}

Here we present a unified study of microsecond P-jump and T-jump experiments of FiP35. Experiments reveal two distinct microsecond kinetic phases in both types of jumps, but the faster phase behaves differently in the P-jump than in the T-jump. We complement the experimental data with all-atom MD simulations that fairly closely mimic the actual P- and T-jump experiments: the protein is unfolded either by pressure or heat, allowed to equilibrate, and then jumped back to the original temperature or pressure, whereupon relaxation is observed. Several P- and T-jump trajectories lead to complete refolding of FiP35, allowing us to gather some statistics about mechanism. Our modeling using Bayesian analysis of the MD trajectories⁵⁰ and kinetic master equations shows that simulation is in good agreement with the experimentally observed kinetic curves. Although some reasonable assumptions have to be made about the fluorescence signal of the intermediates, and the agreement thus does not prove the mechanism extracted from the MD simulations, our results show that MD can reproduce complex experimental kinetics with multiple time scales quantitatively. This is an important step forward in using the combination of experiment and simulation to understand protein folding mechanisms in detail.

3.2 METHODS

3.2.1 Sample preparation FiP35 was expressed and purified as described previously,⁴² although with a slightly different plasmid construction. Briefly, a construct encoding a fusion protein consisting of Glutathione-S-transferase (GST), a thrombin cleavage site, and FiP35 was cloned into pDream (GenScript). The fusion protein was expressed in BL21(DE3)-Ripl (Agilent) *E. coli* and captured and purified from the cell extract on an immobilized glutathione column according to manufacturer's guidelines (GenScript). The protein was eluted by 10 mM glutathione in 50 mM Tris-HCl pH 8.0 and subsequently dialyzed against 10 mM sodium phosphate buffer.

FiP35 was cleaved from the purification tag by overnight incubation with biotinylated thrombin (EMD Millipore). Thrombin was removed by incubation with streptavidin-agarose resin (EMD Millipore) according to manufacturer's protocol. FiP35 was purified from cleaved GST via an ultra filtration cell with 10 kDa cutoff membrane (Millipore). The purified FiP35 was lyophilized and resuspended to a final concentration of 200-400 μ M FiP35 with 100 mM sodium phosphate buffer, pH 7.0.

Buffer conditions were identical throughout all experiments while concentrations of FiP35 and GuHCl varied. FiP35's single tryptophan on the first beta strand (loop or hairpin 1) enabled monitoring of folding via fluorescence, in particular the formation of strands 1 and 2.

3.2.2 Pressure thermodynamics The thermodynamics of FiP35 unfolding under pressure were measured via fluorimeter (Cary Eclipse, Varian) in a pressure cell (ISS) as described previously.³⁵ The sample was placed in a quartz cuvette with a flexible cap to enable pressure transduction and inserted into the pressure cell. With a hydrostatic pressure generator (High Pressure Equipment Company), the cell was filled with spectroscopic grade ethanol (95.0 %, Acros organics) and pressurized up to 250 MPa with 4 minutes equilibration every 10 MPa increase or decrease. A custom built stage was used to mount the pressure cell into the fluorimeter.

Solutions of 50 or 100 μ M FiP35 in 100 mM sodium phosphate buffer pH 7.0 and various concentrations of guanadinium hydrochloride (GuHCl) (Sigma) were assayed at room temperature. FiP35's single tryptophan was excited at 280 nm and emission was collected from 290 nm to 500 nm at each pressure.

Pressure denaturation traces at varying concentrations of GuHCl were analyzed using singular value decomposition (SVD). Because intensity is arbitrary across different pressure titrations, the 2nd SVD vector (SVD2), which contains contributions from intensity and spectral shape change, was normalized by the 1st SVD vector (SVD1), which contains only intensity contributions. Plotting SVD2/SVD1 allows all of the pressure titrations to be analyzed together.

The unfolding traces were fit globally to a two-state unfolding model:

$$S(P) = \frac{S_u(P) + S_f(P)e^{-\Delta G(P,M)/RT}}{1 + e^{-\Delta G(P,M)/RT}}$$

where $S_u(P)$ and $S_f(P)$ are the linear folded and unfolded baselines, and the free energy of folding is given by:

$$\Delta G(P,M)_{u \rightarrow f} = \Delta V_{u \rightarrow f}(P - P_m) + g_m M$$

Here, $\Delta V_{u \rightarrow f}$ is the volume change of folding, P is pressure, P_m is the midpoint of the denaturation transition at 0 M GuHCl, g_m is a constant that describes the linear dependence of the free energy on GuHCl concentration, and M is the concentration of GuHCl.

3.2.3 Temperature thermodynamics The thermodynamics of FiP35 unfolding under temperature were measured with a fluorimeter (Cary Eclipse, Varian) in an Agilent temperature controller. Tryptophan was excited at 280 nm and emission collected from 290 nm to 500 nm. Solutions of 10 μM FiP35 in 100 mM sodium phosphate buffer pH 7 and various concentrations of GuHCl were assayed and spectra were analyzed by SVD as described in the previous section.

The temperature-probed unfolding traces were fit globally to a two-state model analogously to the pressure-probed thermodynamics:

$$S(T) = \frac{S_u(T) + S_f(T)e^{-\Delta G(T,M)/RT}}{1 + e^{-\Delta G(T,M)/RT}}$$

where $S_u(T)$ and $S_f(T)$ are the linear folded and unfolded baselines, and the free energy of folding is given by:

$$\Delta G(T,M)_{u \rightarrow f} = g_l(T - T_m) + g_m M$$

where g_l is a constant describing the linear temperature dependence of the free energy, T is the temperature, T_m is the unfolding midpoint at 0 M GuHCl, g_m is a constant describing the linear dependence of the free energy on GuHCl concentration, and M is the GuHCl concentration.

3.2.4 P-jump and T-jump kinetics Pressure-jump kinetics were measured as described previously^{34,35} with a home-built fast pressure drop apparatus capable of pressure drops up to 250 MPa and a dead time of about a microsecond. A solution of 300 μM FiP35 in 100 mM phosphate buffer pH 7.0 and various concentrations of GuHCl was placed into a hemispheric sample cavity machined on the face of a sapphire cube (Esco Products). Hydrostatic pressure was applied by a pressure generator (High Pressure Equipment Company) through a pressure fitting to pressurize the sample, isolated from the pressurization fluid by a piece of Mylar coated aluminum foil. A capacitor bank discharge delivered by a sharpened copper electrode burst a 0.007" stainless steel membrane, weakened by the applied pressure, dropping the sample from 165 MPa to atmospheric pressure in about a microsecond.

Tryptophan fluorescence lifetime decays were probed every 12.5 ns before, during, and after the jump by a mode-locked Ti:Sapphire laser (KMLabs) that was frequency-tripled to 280 nm with a commercial tripler (CSK Optronics). Tryptophan fluorescence was passed through a B370 band-pass filter, collected via a PMT (R7400U-03, Hamamatsu), and digitized by an oscilloscope (DPO7254) at a sampling period of 100 ps per point for 5 ms total data collection time. Two to

three jumps were collected for each GuHCl concentration and, to improve signal to noise, 50 decays were binned in each jump before analysis. Plots shown throughout the text are further smoothed after analysis to 5 μ s per point.

FiP35 was jumped to the same final thermodynamic state in T-jumps as in P-jumps (figure 26a) enabling direct comparison of the measured kinetics. Temperature-jumps from 18°C to 23°C were carried out by a Surelite Q-switched Nd:YAG laser (Continuum) Raman-shifted to 1.9 μ m and sent through a 50/50 beam splitter such that the sample is heated from two sides. The starting temperature was held constant by a temperature controller (Lake Shore 330, Lake Shore Cryotronics) and the size of the jump was measured by comparing the tryptophan decay at the end of the jump to tryptophan decays measured at equilibrium at various temperatures.

300 μ M FiP35 in 100 mM sodium phosphate buffer pH 7.0 and varying concentrations of GuHCl samples or 200 μ M tryptophan for jump calibration were held in fused silica cuvettes for the experiments. Tryptophan fluorescence was excited and collected as described in the previous section. 50 decays were binned in each jump before analysis, as in pressure-jump experiments.

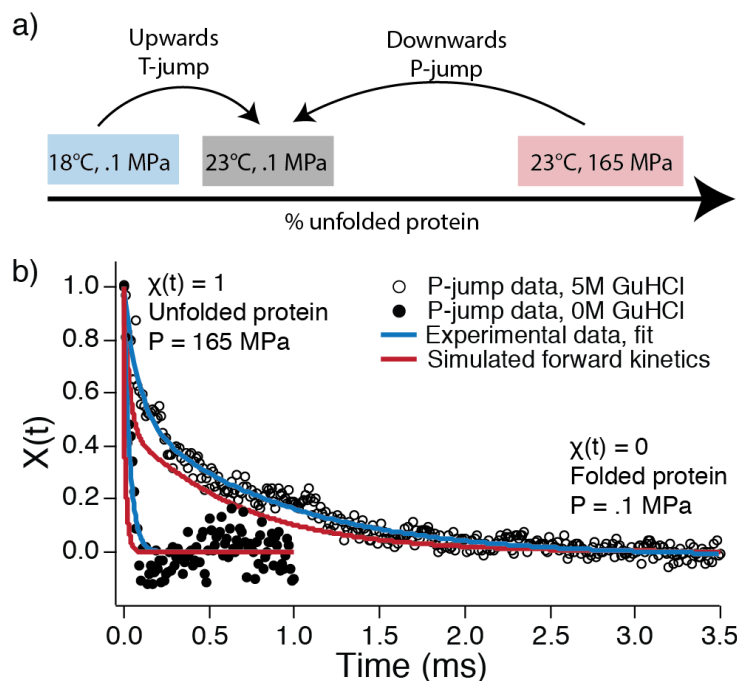


Figure 26 Experimental scheme and kinetic traces. a) Temperature and pressure were jumped to the same final condition, thus allowing their rates to be compared. At low GuHCl concentration, both observed kinetic rates monitor predominantly refolding kinetics. b) Two representative P-jump traces at 0 M or 5 M GuHCl. Both traces have been binned to 5 μ s intervals. Blue curves show double exponential fits to the experimental data. Red curves are the kinetics from a ‘two parallel intermediates’ master equation model (model IV in figure 5a) populated with rates from MD trajectories. Simulated kinetics are scaled to account for the viscosity difference between simulated and experimental water, and the 5 M GuHCl is also scaled by the experimental k_{0M}/k_{5M} ratio of rate coefficients (see SI methods).

3.2.5 Kinetics data analysis The change in tryptophan lifetime reflects changes in the probed tryptophan's surrounding environment as the protein unfolds and is a more reliable observable than overall fluorescence intensity as it is immune to slight fluctuations in laser intensity. Chi analysis⁵¹ is applied to quantify the change in lifetime over the course of the jump. The lifetime as a function of time, $S(t)$, is decomposed into a linear combination of the lifetime before the jump, f_1 , and the lifetime at the end of the jump, f_2 , such that:

$$S(t) = c_1(t)f_1 + c_2(t)f_2$$

$c(t)$ is defined as the normalized $c_1(t)$:

$$\chi(t) = \frac{c_1(t)}{c_1(t) + c_2(t)}$$

$c(t)$ thus ranges from 1 at the beginning of the jump to 0 at the end of the jump (figure 26b).

In T-jump experiments, a slow linear baseline due to photo-bleaching were fit and subtracted off. For P-jump experiments, portions of the kinetic traces with large amplitude random noise were removed and excluded entirely from analysis (See figure 61). For both P-jump and T-jump experiments, the kinetic signal expressed as $c(t)$ was fit to a double-exponential :

$$\chi(t) = y_0 + A_1 e^{-t/\tau_1} + (1 - A_1) e^{-t/\tau_2}$$

For P-jump experiments, all replicates (2-3 jumps per GuHCl concentration) were fit to a double-exponential individually and the fitted parameters averaged together to get single values for A_1 , τ_1 , and τ_2 for each GuHCl concentration. For T-jump experiments, two traces were binned together before each averaged pair was individually fit to double-exponentials.

The folding prefactor in the Arrhenius equation depends inversely on solvent viscosity and, consequently, the observed folding rate also shows viscosity dependence. High concentrations of GuHCl impart higher solvent viscosity, so observed folding rates from both pressure- and temperature-jump experiments were viscosity-corrected. Relative viscosities of GuHCl solutions to water,⁵² ν_r , were used to correct for the folding rate decrease:⁵³

$$k_{obs,corr} = \nu_r k_{obs}$$

Thermodynamic parameters from the global fits were used to calculate the free energy at the final equilibrium state after both temperature-jumps and pressure-jumps: 23°C, 0.1 MPa, and various GuHCl concentrations. The temperature-jump final state was chosen to match the final

thermodynamic state after pressure-jump. With the fitted observed folding rate, k_{obs} , from kinetics experiments, the approximate folding and unfolding rates (approximate because FiP35 is not a two-state folder) were calculated for each GuHCl concentration:

$$k_f = \frac{k_{obs}K_{eq}}{(1 + K_{eq})}$$

$$k_u = k_{obs} - k_f$$

where K_{eq} is given as:

$$K_{eq} = e^{-\Delta G / RT}$$

The thermodynamic parameters from the pressure-probed unfolding were used to calculate folding and unfolding rates for pressure-jumps and those from the temperature-probed unfolding were used for temperature-jump calculations. However, at the equilibrium conditions (room temperature and atmospheric pressure), the free energies from the temperature and pressure-probed thermodynamics were very close: -17 kJ/mol (temperature) *vs.* -16 kJ/mol (pressure).

3.2.6 Molecular dynamics simulations Molecular dynamics simulations, described in detail in Appendix A, were performed in explicit solvent using the TIP3P water model⁵⁴ and the CHARMM22* force field for the protein.⁵⁵ The structure of the WW domain (FiP35 mutant, residues 4 to 38) from the protein data bank (PDB code 2F21)³⁹ was placed in a cubic box of 10,232 water molecules and neutralized with 6 sodium ions and 9 chloride ions employing VMD.⁵⁶ All simulations were carried out with periodic boundary conditions in a constant particle number, temperature, and pressure ensemble (NPT). Starting from the native state of the protein, two types of simulations were performed: P-jump and T-jump. Unfolding simulations were performed using NAMD⁵⁷, and refolding simulations were performed on the Anton Platform.^{58,59} Data analysis and figure rendering were done using VMD.⁵⁶

In a P-jump simulation, pressure was increased from 0.1 MPa to 900 MPa in 0.3 μ s at a rate of 0.9 MPa/300 ps, followed by a 1- μ s high-pressure equilibrium simulation ($P = 900$ MPa) and a pressure-drop simulation in which pressure was jumped downward from 900 MPa to 0.1 MPa in 0.3 μ s at a rate of -0.9 MPa/300 ps. The temperature was maintained at $T = 325$ K through the P-jump simulation. In a T-jump simulation, the pressure was maintained at $P = 0.1$ MPa throughout the simulation and the system was heated up to 400 K for 1 μ s between the initial 100-ns equilibrium simulation and the final refolding simulation, both at $T = 325$ K. To generate multiple refolding

trajectories, the pressure or temperature unfolded state was equilibrated for an additional 200 ns respectively, during which the structures were taken at $t = 0$ ns, 100 ns, and 200 ns to continue the P-jump and T-jump simulations at ambient conditions. The simulated and experimental P-jumps are in the same direction, although the denaturation pressure is higher in simulations to facilitate rapid unfolding. Note that both P- and T-jump simulations leave the WW domain under strong refolding conditions, as is also the case in the experiments ($T=23$ °C, $P=0.1$ MPa), although the simulated and experimental T-jumps are in opposite directions.

3.2.7 Bayesian rates analysis of MD data and kinetic mechanism simulation Our main focus here is on comparing the novel P-jump experiments and simulations. We started with a Bayesian analysis⁵⁰ to estimate from the P-jump MD simulations the rate coefficients connecting the unfolded state to intermediate states and intermediate states to the native state. This strategy is particularly useful when MD trajectories that do not show a transition are present in the data set and enables inclusion of this data in the estimation of transition rates.

Briefly, Bayes rule states that the probability of a model corresponding to a given data set, or the posterior probability density, can be derived from the probability of the model to generate the given data. The posterior probability density depends on the prior probability density or the state of knowledge about the model given previous data. Taking a uniform distribution as the prior probability density (or assuming no prior knowledge of the model) and specifically treating a two-state transition model, the probability that the two state kinetic model with rate k corresponds to the molecular dynamics data set is given by:

$$p_U(k | D, X, I) = \frac{\Theta^{n+1}}{n!} k^n \exp[-k\Theta]$$

where Θ is the total simulation time, n is the number of simulations where a folding transition occurred, and k is the rate. The total simulation time is calculated by summing the time to the folding transition in each simulation. For simulations where no transition occurred, the folding time is the length of the simulation. The expectation value and variance of the distribution are given by:

$$\langle k \rangle_U = \frac{n+1}{\Theta}, \text{var}(k)_U = \frac{n+1}{\Theta^2}$$

The total folding times, values of n , and calculated rates and variance are summarized in table 11. To calculate the overall folding time for the pressure and temperature-jump simulations,

FiP35's folding to the native state is approximated as a two-state transition. For the pressure-jump simulations, the forward folding rates to the two intermediates and from the intermediates to the native state were calculated by approximating each of the intermediate forming transitions and the intermediate to native state transitions as separate two-state reactions. Because the P-jump simulations (and the experiments) were carried out under conditions strongly favoring the native state, we assume that reverse rate coefficients are negligible.

To assess whether the rates and models observed in the pressure-jump simulations were consistent with experimental data, we performed kinetic simulations. Four different kinetic models (figure 27) that are minimally consistent with experiment and simulation (i.e. they can yield a double exponential decay) were solved populated with transition rates from the MD simulations by numerically solving the differential rate equations. The time depended of the concentration of each species was simulated for 100 μ s with a time resolution of 0.01 μ s.

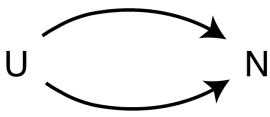
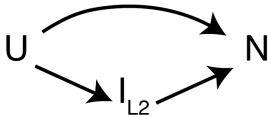
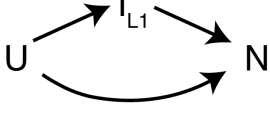
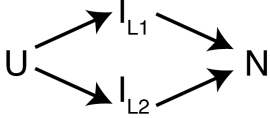
Model	Model score	Assessment
	29	Simulation poorly matches experimental data.
	29	Simulation poorly matches experimental data.
	.14	Simulation matches experimental data, but a long lived loop 2 intermediate was observed in MD data.
	.15	Simulation matches experimental data and model matches intermediates observed in MD data.

Figure 27 Summary of tested kinetic models and outcomes.

Various signal functions were also tested. Because the tryptophan residue of FiP35 is located between strand 1 and strand 2, it is reasonable to assume that all measureable signal in our experiment arises from the formation of the loop 1 intermediate or the native state. Thus, signal functions of the following form were evaluated:

$$S(t) = c_{I_{L1}}[I_{L1}](t) + c_N[N](t)$$

where $[I_{L1}]$ and $[N]$ are the time dependent concentrations of the loop 1 intermediate and the native state, respectively. Their coefficients, C_{IL1} and C_N , were varied from 0 to 1 in intervals of .01 such that:

$$C_{IL1} + C_N = 1$$

To test model and signal function combinations, each signal function was compared to experimentally generated data. To account for the disparity between simulated folding rates and experimental folding rates, the simulated data was scaled. To determine the scaling factor, the previously determined *in silico* folding rate of FiP35 on the Anton computer²⁵ was compared to the experimentally determined FiP35 folding rate at the same temperature.⁴² As our simulations were performed with the same force field on the same super-computer as the previous *in silico* studies, this calculation should provide a robust link between simulation and experimental time-scales. In previous simulations,²⁵ FiP35 gave a folding time of 10 μ s at the *in silico* melting temperature. At the experimentally determined melting temperature, ~ 44 K below the *in silico* melting temperature, FiP35 folds in 27 μ s.⁴² This gives a scaling factor of 3.7, in relatively close agreement with the 26 scaling factor necessary to account for the lessened viscosity of the TIP3 water model.⁶⁰

Since the P-jump experiment is carried under conditions strongly favoring the native state, our kinetic simulation assumes reverse rates are negligible. To account for the decrease in rate due to perturbation of the energy landscape by GuHCl as well as the change in solution viscosity at high concentrations of GuHCl, the simulation was further scaled for comparison to experiment. In this case, the scaling factor was obtained by taking the ratio of the rate from the 0 M GuHCl pressure-jump experiment and the viscosity uncorrected rate from the 5 M GuHCl jumps. This yielded a scaling factor of 38.7.

To assess each model and the various signal functions for each model, each simulated kinetic trace was fit to a double exponential. The fitted fast phase amplitude, fast phase rate, and slow phase rate were compared to those obtained from experiment via a scoring function:

$$Score = \sqrt{\left(\frac{k_{fast,obs} - k_{fast,sim}}{k_{fast,obs}}\right)^2 + \left(\frac{k_{slow,obs} - k_{slow,sim}}{k_{slow,obs}}\right)^2 + \left(\frac{A_{fast,obs} - A_{fast,sim}}{A_{fast,obs}}\right)^2}$$

The lowest scoring signal functions for each model were used to report the model's overall closeness to experimental data.

3.3 RESULTS

3.3.1 FiP35 is thermodynamically stable to temperature and pressure denaturation Unfolding of FiP35 with pressure, temperature and denaturant (GuHCl) was measured by monitoring fluorescence from its single tryptophan residue. FiP35 unfolds reversibly up to 90 °C at concentrations $\leq 50 \mu\text{M}$ ⁴² and up to 210 MPa at $\leq 100 \mu\text{M}$ (figure 62). At the highest pressures and temperatures used in kinetics experiments, no concentration dependence of the thermodynamics were observed and FiP35 was fully reversible at a concentration of 300 μM (figure 63). When FiP35 is destabilized, its activation barrier increases, and its unfolding can be fitted approximately to a two-state model, yielding its melting temperature and pressure (figure 28). FiP35 has an extremely stable fold: with no added GuHCl, the pressure unfolding midpoint is $\approx 680 \text{ MPa}$ (table 26).

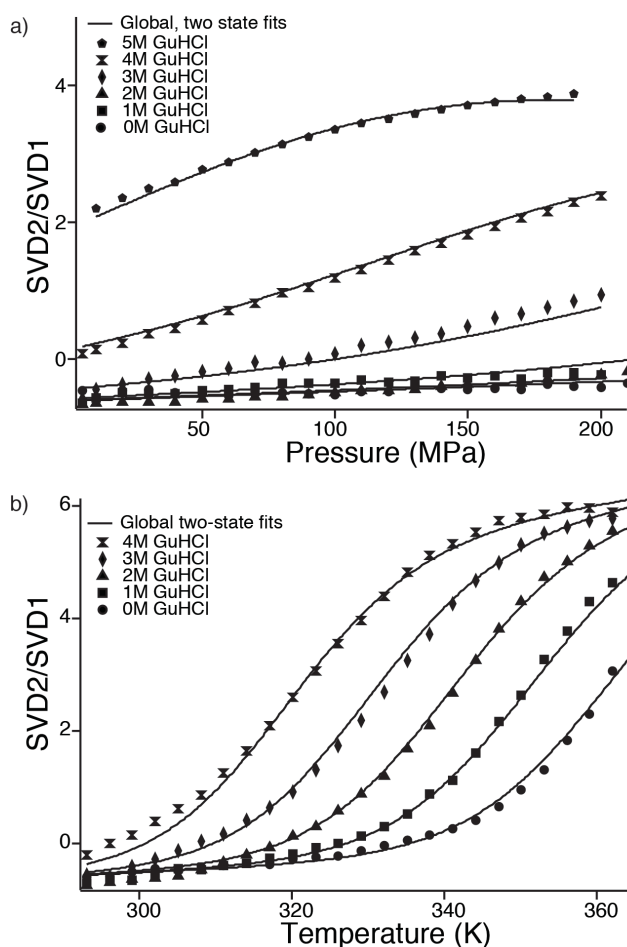


Figure 28 Pressure and temperature thermodynamics with global fits. Y-axis is the ratio of component 2 and 1 from the singular value decomposition (SVD), which corresponds to wavelength shift.

	g_1 or $\Delta V_{u \rightarrow f}$	T_m or P_m	g_m (kJ mol ⁻¹ M ⁻¹)
Pressure	3.00 ± 0.25 mL mol ⁻¹	677.0 ± 34 MPa	3.783 ± 0.16
Temperature	0.3039 ± 0.0068 kJ mol ⁻¹ K ⁻¹	362.7 ± 0.44 K	3.265 ± 0.11

Table 1 Thermodynamics. Values are derived from global fits and errors are fit uncertainties.

Due to FiP35's deviations from two-state behavior, fitted melting temperatures vary depending on how the fluorimeter data is analyzed. Here, singular value decomposed (SVD⁶¹) fluorescence spectra plotted versus temperature give a melting temperature close to 90 °C. When the same data is expressed as an integrated fluorescence intensity versus temperature, the fitted melting temperature is 75 °C, consistent with previous measurements.^{41,42} Here we use SVD to analyze the spectra in order to avoid intensity contributions to the signal, which in our pressure thermodynamics instrumentation can contain artifacts (see figure 62).

3.3.2 P- and T-jump kinetics are bi-exponential Pressure-drop-induced refolding kinetics were measured at various GuHCl concentrations using a home-built apparatus that delivered a 165 → 0.1 MPa downward P-jump at 23 °C with an instrument dead time of about 1 μs.³⁴ To compare P- and T-jump folding kinetics, 5 °C T-jumps were performed from 18 to 23 °C, the same final state as the P-jumps (figure 26a). In both P- and T-jumps, the reaction was monitored by tryptophan fluorescence lifetime.⁴²

Both T- and P-jump kinetics reveal a faster kinetic phase (≈ 6 μs for T-jump, ≈ 30 ms for P-jump; tables 2 and 3, figures 61 and 64) and a slower kinetic phase (200-500 μs; tables 2 and 3, figures 61 and 64). The 'slow' P-jump phase is much faster than the previously observed millisecond P-jump refolding of I₆₋₈₅, which was attributed to pressure-induced trapping in a highly helical non-native state.³⁵ There is no evidence of such millisecond traps in pressure-induced refolding of FiP35, making it attractive for full-atom simulations of complete refolding.

[GuHCl] (M)	%A ₁	τ_1 (μs)	τ_1 (μs), viscosity corrected	τ_2 (μs)	τ_2 (μs), viscosity corrected
3.5	34.1 ± 0.9	12.7 ± 7.7	11.3 ± 6.8	218 ± 54	193 ± 48
3	50.8 ± 5.2	13.6 ± 3.0	11.6 ± 2.6	398 ± 42	340 ± 36
4	32.6 ± 3.6	22.2 ± 6.0	17.4 ± 4.7	452 ± 24	355 ± 19
5	47.3 ± 3.3	30.0 ± 4.4	21.2 ± 3.1	727 ± 69	513 ± 48

Table 2 Pressure-jump kinetics. Errors are standard error (standard deviation of the mean).

[GuHCl] (M)	%A ₁	τ_1 (μ s)	τ_1 (μ s), viscosity corrected	τ_2 (μ s)	τ_2 (μ s), viscosity corrected
3	77.7 \pm 1.0	7.52 \pm .78	6.42 \pm .64	276 \pm 21	235 \pm 18
3.5	68.6 \pm 1.1	7.48 \pm .36	6.14 \pm .29	269 \pm 11	221 \pm 9.0
4	59.2 \pm 1.3	7.07 \pm .69	5.55 \pm .54	310 \pm 18	244 \pm 14
4.5	50.7 \pm 1.1	8.70 \pm .51	6.50 \pm .38	368 \pm 12	275 \pm 8.9
5	49.8 \pm 1.0	8.68 \pm .43	6.13 \pm .30	427 \pm 7.5	302 \pm 5.3

Table 3 Temperature-jump kinetics. Errors are standard error (standard deviation of the mean).

3.3.3 The slow kinetic phase is similar between perturbation methods The slow kinetic phase in both T- and P-jump is consistent with being FiP35's activated refolding phase. Both experiments in low or 0 M GuHCl and 23 °C are under conditions strongly favoring the native state. Kinetic and thermodynamic data were combined in a two-state analysis to extract folding and unfolding rates for the slow phase, an approximation given the observation of an additional fast phase (discussed below). The chevron plot in figure 29a shows that the denaturant dependence of both folding and unfolding rates is similar for T- and P-jump kinetics, although T-jump-probed folding rates are somewhat faster. All assayed GuHCl concentrations are below the GuHCl titration midpoint of FiP35,⁶² resulting in the expected decrease in the folding rate and increase in the unfolding rate with increasing GuHCl concentration.

Linear extrapolation of the activated folding rates versus GuHCl concentration yields activated folding rate coefficients k_f of (56 μ s)⁻¹ for P-jump and (52 μ s)⁻¹ for T-jump at 23 °C in the absence of denaturant. A P-jump of FiP35 without denaturant (figure 26b and 65) yielded a fitted folding rate coefficient of (29.5 μ s)⁻¹, within measurement error of the extrapolation (figure 29a). The T-jump activated rate was previously measured only above 60 °C,⁴² where $k_f \approx (20 \mu\text{s})^{-1}$ is faster than our value of (52 μ s)⁻¹ at 23 °C. The previous measurements cover a wide temperature range above 60 °C,⁴² and their extrapolation to room temperature in absence of GuHCl agrees well with our current measurements at room temperature extrapolated to 0 M GuHCl.

3.3.4 The fast kinetic phase differs between perturbation methods Unlike the slow phase, the fast kinetic phase of FiP35 behaves differently after T- and P-jump (figure 29b, 29c). Like the slow phase, the P-jump fast rate decreases as GuHCl concentration increases, even after viscosity correction (figure 2b); its amplitude shows no obvious trend when the GuHCl concentration is changed (figure 29c). In contrast, the fast T-jump rate has no GuHCl concentration dependence, even after viscosity correction (figure 29b), and its amplitude clearly increases when FiP35 is stabilized by removing denaturant (figure 29c). In addition, the T-jump fast phase also has a rate (6

$\mu\text{s})^{-1}$ at 23 °C that is about 2-3 times faster than the P-jump fast phase and approaches the folding “speed limit”⁶³ previously measured for FiP35 in the $(0.7\text{-}2 \mu\text{s})^{-1}$ range between 60-85 °C.^{41,42}

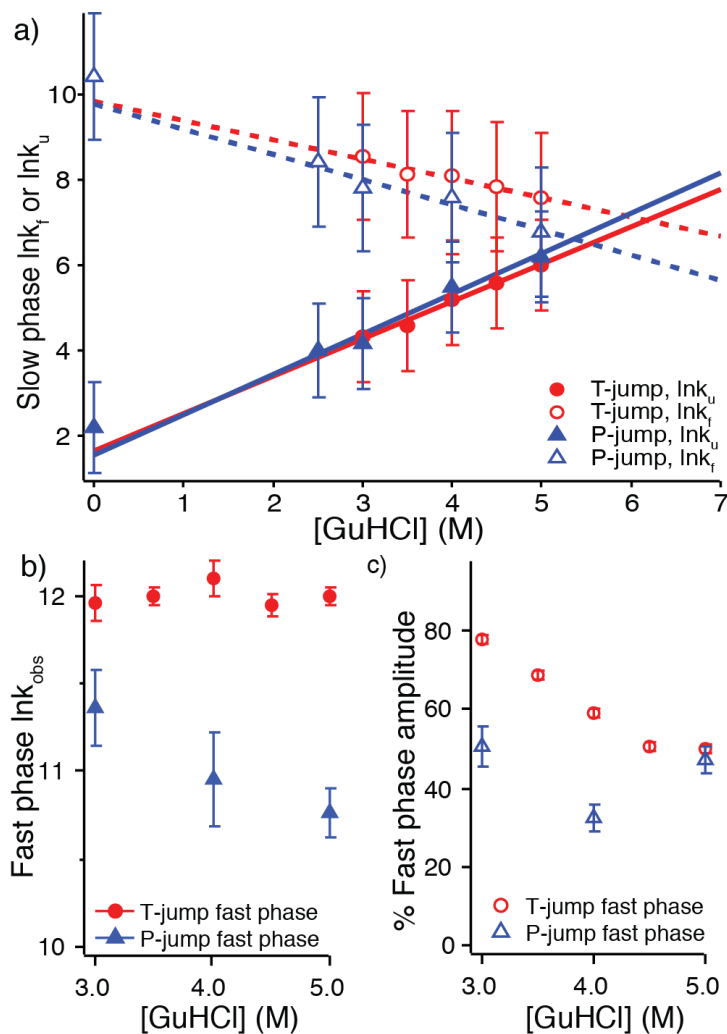


Figure 29 Experimental kinetics summary. a) Chevron plot illustrating the GuHCl dependence of the slow phase folding and unfolding rates. Rates are linearly extrapolated without viscosity correction.⁶⁴ Error bars include fit error from thermodynamic values and standard error from averaging kinetic fit parameters. b) Observed, solvent viscosity scaled relaxation rate of the fast phase for T- and P-jump (k in units of s^{-1} , \ln is the natural logarithm, error is standard error). c) The percentage of the amplitude corresponding to the fast phase. For both b) and c), error is standard error from averaging fitted values across multiple experimental traces.

3.3.5 FiP35 completely refolds in several P- and T-jump simulations We ran six long single-trajectory all-atom MD simulations for FiP35 to provide mechanistic information to complement experiment. Two types of simulations of unfolding were performed, high temperature and high pressure,

followed by a jump to ambient conditions where refolding ensued. Three refolding trajectories for each denaturation type were obtained yielding over 70 microseconds of refolding trajectories in total.

The P-jump simulations mimic the experimental protocol. A pressure of 900 MPa, which exceeds the experimentally determined unfolding midpoint, was used to unfold the protein. During the unfolding process, hairpin 2 lost its native structure first, followed by the denaturation of hairpin 1, as shown in figures 66a and 67a. Across the three refolding simulations (figures 30 and 68), the protein refolded into the native state in 4 μ s, 12.5 μ s, and 17.5 μ s.

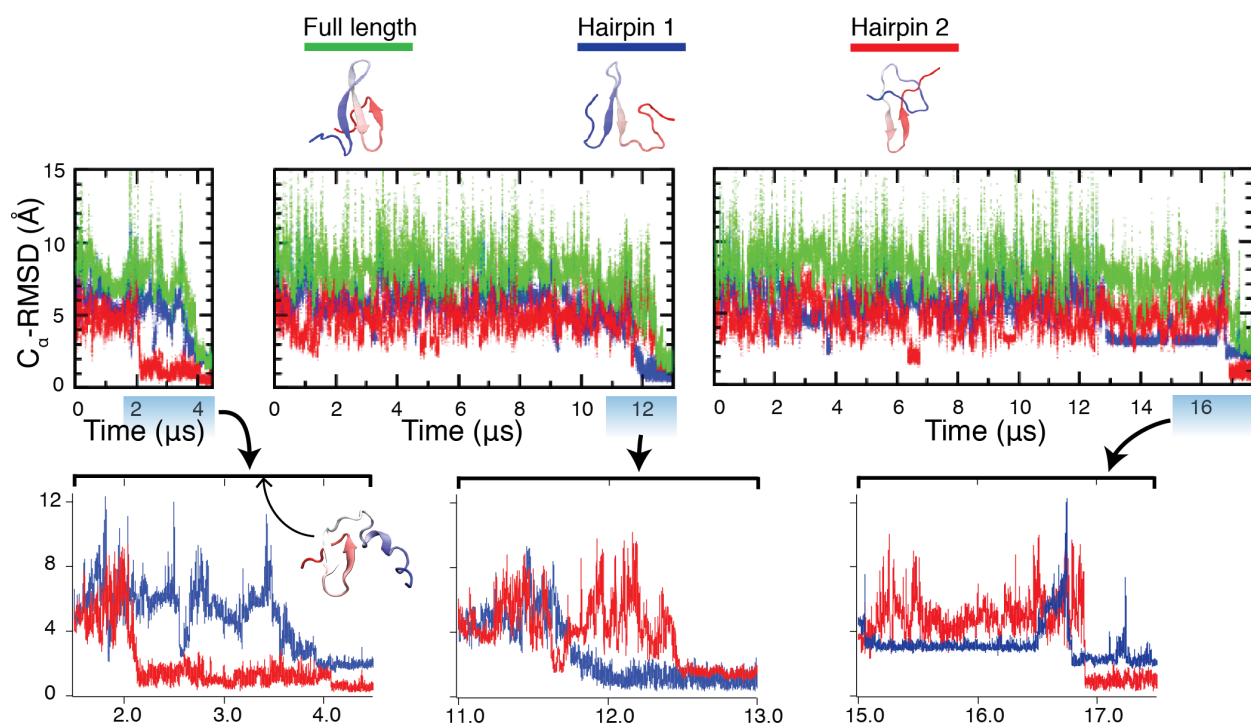


Figure 30 Structural characterization of the refolding trajectories following P-jump, at $T=325$ K and $P= 0.1$ MPa, from the pressure-denatured state. C_{α} -RMSD values have been calculated relative to the crystal structure (PDB:2F21).³⁹ Hairpin 1 contains residues 11 to 25 and hairpin 2 contains residues 22 to 33. Structures under the legend show the native structure, a hairpin 1 formed (or loop 1) intermediate, and a hairpin 2 formed (or loop 2) intermediate. The full length C_{α} -RMSD values are calculated using residues 7 to 35. **Bottom:** Highlights of the folding transitions. Inset structure in the first folding transition is a misfolded mirror image structure (see main text).

Two different folding mechanisms were observed in the simulations, with one simulation folding through hairpin 2 first, and the other two through hairpin 1 first. The persistence of the hairpin 2-containing structure (representative structures shown in figure 30 under the legend) for several microseconds before refolding to the native state may arise from the formation of a mirror image conformation where hairpin 1 begins to form on the wrong side of hairpin 2 (inset structure

in bottom panel of figure 30)—a kinetic trap arising from the energetic similarity between the native state and the mirror image conformation that has been observed theoretically and experimentally in symmetrical proteins.⁶⁵

T-jump simulations were also performed. A high temperature of 400 K was used to simulate temperature denaturation of FiP35 (figures 66b and 67b). The protein unfolded in 50 ns by losing both hairpins almost simultaneously. Upon return to room temperature, the protein folded to its native state in 0.5 and 6 μ s, hairpin 2 first. In one simulation, FiP35 formed hairpin 2 in 19 μ s but failed to fold in 27 μ s of simulation (figures 31 and 69). Note that while the simulation is a downward T-jump and the experiment upward, the experimental final condition, as in the P-jump, is under conditions very strongly favoring the native state (23 $^{\circ}$ C, far below previous experimental studies carried out at 50+ $^{\circ}$ C).

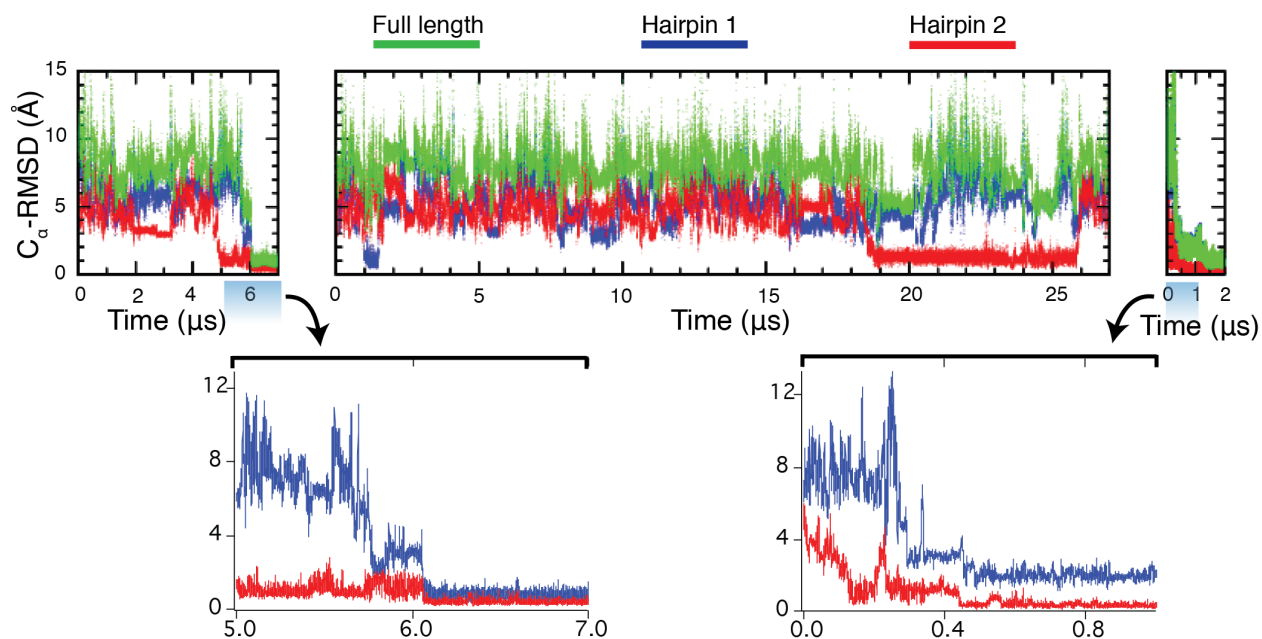


Figure 31 Structural characterization of the refolding trajectories following T-jump, at $T=325$ K and $P= 0.1$ MPa, from the temperature-denatured state. C_{α} -RMSD values have been calculated relative to the crystal structure (PDB:2F21).³⁹ Hairpin 1 contains residues 11 to 25 and hairpin 2 contains residues 22 to 33. The full length C_{α} -RMSD values are calculated using residues 7 to 35. **Bottom:** Highlights of the folding transitions.

3.3.6 Kinetic modeling suggests a parallel four-state folding mechanism following P-jump The P-jump simulations revealed two paths to the native state following pressure-jump: folding through hairpin 1 first or hairpin 2 first. We used Bayesian analysis of the molecular dynamics trajectories

and kinetic master equations to determine the kinetic mechanism that best matches both MD simulation and experiment. Four kinetic models (summarized in figure 27 and 32) were tested.

I: Heterogeneous transition state model

Phi-value analysis has shown that WW-domains can fold through two transition states with either loop 1 or loop 2 formed.⁶⁶ In this kinetic model, FiP35 would fold through either of these two transition states through parallel paths. Our molecular dynamics simulations imply that this model is unlikely because, particularly in the case of the loop 2 formed structure, we observe persistence of the partially folded structures.

Because this model has no intermediates, the change in concentration of the native state was used as the signal function:

$$S(t) = [N](t)$$

As expected, the model did not closely match experiment giving a score of 29. The fitted values for the simulated data are summarized in table 12.

II: Folding through the loop 2 intermediate

In this model, FiP35 folds through a stable loop 2 intermediate and through a parallel path to the native state with no intermediate. As in the two transition-state model, the signal function was given by the change in concentration of the native state. Because of the similarity of the signal functions, this gives very similar results to the two transition state model. The double exponential kinetics yielded by the simulation were a poor fit to the experimental data giving a score of 29. The fitted values for the simulated data are summarized in table 12.

III: Folding through the loop 1 intermediate

In this model, FiP35 folds through a loop 1 intermediate and through an alternate path to the native state with no intermediate. In this case, the signal function was given as a linear combination of the concentration of the native state and the loop 1 intermediate:

$$S(t) = c_{I_{L1}}[I_{L1}](t) + c_N[N](t)$$

The scores for each tested signal function are shown graphically in figure 70. The minimum score occurred for a signal function with $c_{I_{L1}}=0.94$ giving a score of 0.14. The fitted parameters for the kinetics were thus a close match for the experimental kinetics and are shown in table 12.

While this model provides an excellent fit to experimentally observed kinetics, it deviates substantially from the observed mechanism in MD simulations. In the simulations, one trajectory shows a stable loop 2 intermediate, which persists for several microseconds while the loop 1 intermediate is far more transient in the other two trajectories. This model is not consistent with this

data because it designates the most transient of observed structures an intermediate while the more stable observed structure a transition state.

IV: Folding through both the loop 1 and loop 2 intermediate

In this model, FiP35 folds through a loop 1 intermediate and through a parallel path to the native state through a loop 2 intermediate. As in case III, the signal function was comprised of contributions of both the native state and the loop 1 intermediate. The scores for each tested signal function are shown graphically in figure 70. The minimum score occurred for a signal function with $c_{IL1}=0.95$ giving a score of 0.15. This model is thus consistent with experimental data and also reflects the observed folding mechanism in MD simulation. figure 26b shows the very close agreement between this model (red traces) and the experimental data (black circles).

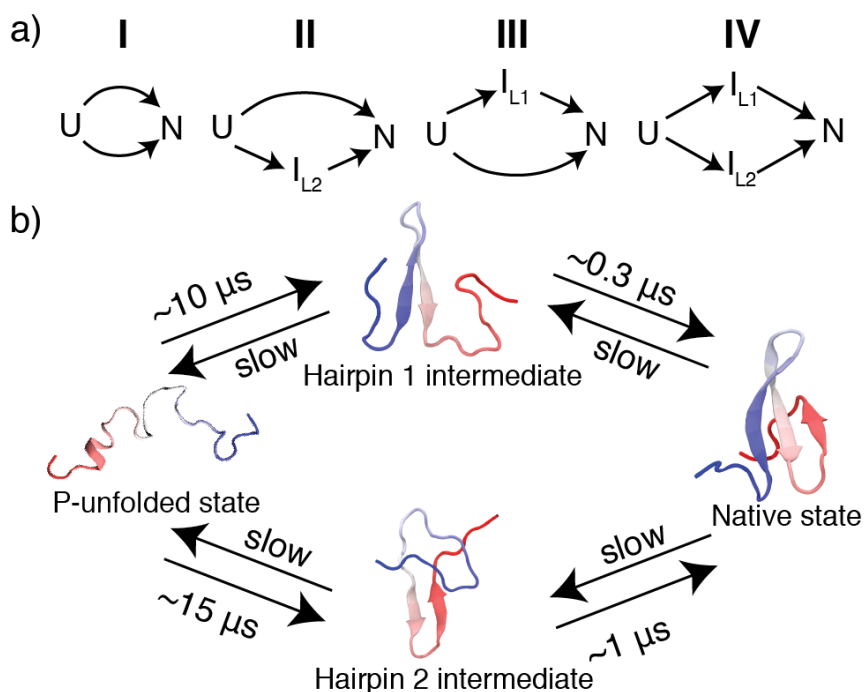


Figure 32 Kinetic models. (a) Mechanisms tested by master equation kinetics and compared with experiment. (b) Mechanism describing folding behavior of FiP35 following P-jump that is consistent with both experiment and simulation. Forward rates were estimated via Bayesian analysis from MD simulations.⁵⁰

3.4 DISCUSSION

We present here an experimental comparison of T- and P-jump refolding on the microsecond time scale and simulate T- and P-jump protocols in all-atom molecular dynamics simulations. In particular the P-jump protocol is similar in experiment and simulation, although a

higher denaturation pressure was used in the simulation to accelerate unfolding. FiP35 has been shown previously to fold following T-jump on a molecular dynamics-friendly timescale.^{41,42} Unlike the only other microsecond folder previously analyzed by P-jump,^{28,35} lambda repressor, which exhibited a slow millisecond phase in addition to a fast phase, we see that P-jump folds FiP35 on a microsecond time scale. For the most direct comparison with simulation, we were also able to obtain refolding data in the absence of any denaturant where folding occurred with a rate below $(30 \mu\text{s})^{-1}$. This fast folding behavior makes FiP35, and other fast WW-domain mutants such as GTT,⁴⁵ excellent candidates for future P-jump experimental and computational studies.

Experimentally, FiP35's refolding kinetics are bi-exponential when probed by both temperature and pressure. The slow phase, on the scale of hundreds of microseconds depending on GuHCl concentration, is very similar between T- and P-jump, showing comparable rate dependence on denaturant concentration and similar folding and unfolding rates. The fast phase, on the other hand, differs between the two perturbation methods.

The T-jump fast phase rate at 23 °C (over 50 °C below T_m) approaches the folding speed limit, shows no dependence on GuHCl concentration when adjusted for viscosity, and exhibits an increase in amplitude as solvent conditions become more stabilizing—the combination of these behaviors has been interpreted as a signature of downhill folding.^{29,42,67} Considering this data and previous work demonstrating FiP35's downhill folding behavior under strongly native state-promoting conditions,^{41,42} we assign the fast phase of FiP35 following T-jump as the “molecular” or “downhill” phase. No previous simulations are available under these conditions. Downhill kinetics occur when the folding barrier along a 1-D free energy profile is very low, so some population exists at the barrier top. This barrier-top population reacts promptly after T-jump and is followed sequentially by activated folding of population in the unfolded free energy well. Of course, on a multi-dimensional free energy surface, such prompt kinetics may correspond to numerous transiently populated states, as seen in Markov modeling and other analyses of complex beta sheet folding.^{47,48}

In contrast, (1) the P-jump fast phase has a slower rate of $(11\text{-}21 \mu\text{s})^{-1}$; (2) the rate decreases more upon addition of GuHCl than expected from increased solvent viscosity, indicating a barrier whose height increases with denaturant concentration; and (3) there is no discernible trend in the fast phase amplitude. These three behaviors are not consistent with downhill folding. Short-lived intermediate(s) on the unfolded side of the main barrier are the simplest explanation, although other kinetic scenarios are also possible.

The difference in folding mechanism between pressure and temperature probed folding observed experimentally is reflected in the MD simulations, where P-jump showed a more heterogeneous folding mechanism than T-jump. Our P-jump folding mechanism—folding through either hairpin 1 or hairpin 2—is remarkably consistent with experiment, as demonstrated through direct comparison of the kinetic model with the experimental data (Fig. 1b). The mechanism allowing both hairpin 1 or hairpin 2 to fold first is generally in agreement with scenarios reported in the literature when FiP35 is started in the unfolded state under native conditions.^{25,46,47,68,69} Markov state models proposed a heterogeneous folding mechanism, with one group finding four folding paths⁶⁸, including folding through both hairpin 1 and hairpin 2, and another group finding that folding proceeds 70% of the time through hairpin 1 and 30% of the time through hairpin 2.⁶⁹

Although the P-jump refolding trajectories partitioned between two parallel paths, the three T-jump trajectories refold exclusively through a mechanism where hairpin 2 was the first one to fold. Even the one trajectory that did not fold (Figures 31 and 69) formed hairpin 2 for an extended time, but not hairpin 1. Interestingly, the only all-atom MD simulation where FiP35 ultimately folded (prior to the work presented here) was interpreted by a folding mechanism where hairpin 1 exclusively formed first.²⁵ Re-analysis of that data found that 80% of the time hairpin 1 folded first and 20 % of the time hairpin 2 was the one to fold first.⁴⁶ Our work and previous simulation results suggest that folding *via* both paths is very likely following both P-jump and T-jump.

Although folding times from T-jump experimental data have been compared to folding times from MD simulations,²⁹ pressure perturbation offers an additional constraint through which force-fields can be tested. Our combination of computational and experimental results highlight three interesting aspects of FiP35's folding behavior: (1) P-jump refolding is complete in 10s of ms, an MD-friendly time scale. (2) When perturbed by temperature and pressure, FiP35 lies in the gray zone of the transition between downhill folding and multi-state folding through short-lived intermediates. That is, by T-jump the fast phase shows characteristics explained by the breakdown of transition state (Kramers) models discussed in detail previously.⁶³ By P-jump, the fast phase has characteristics that strongly support a short-lived intermediate—behavior that is reflected in some MD simulations.⁷⁰ (3) MD simulations of the P-jump induced refolding of FiP35 very closely match our experiment when the appropriate kinetic model is chosen. This remarkably close agreement allows us to assign a parallel kinetic mechanism for FiP35's folding following pressure drop (Figure 32b).

SECTION II: PROTEIN FOLDING IN LIVING CELLS

CHAPTER 4 QUINARY PROTEIN STRUCTURE AND THE CONSEQUENCES OF CROWDING IN LIVING CELLS*

4.1 INTRODUCTION

The complexity of the cell—jam-packed with macromolecules, inherently heterogeneous, and always dynamic—has led to widespread recognition that *in vitro* studies of proteins cannot always provide physiologically relevant structural and functional information⁷¹. The momentum of protein science towards probing proteins in their native environments is leading to a departure from the canonical protein purification-to-characterization workflow towards more *in vivo* studies. But which aspects of protein structure, folding, and function are most likely to be modulated by the cell, by how much, and to what adaptive advantage, if any?

With the caveat that living cells are never in equilibrium, thermodynamics tells us "how much" modulation is significant: interaction of a protein with its environment must alter occupancy of different protein populations significantly. A disordered protein chain poised to fold from a state "U" (unfolded) to a state "F" (folded) upon binding a partner will switch from 80% U to 80% F if binding favors "F" over "U" by just 7 kilojoules per mole (kJ/mole). By comparison, a single carbon-carbon bond in one amino acid of that protein requires fifty times more bond enthalpy to break. Teasing apart all the interactions between a protein of interest and the cellular water, ions, macromolecules and other co-solutes is difficult. Indeed, studying protein dynamics *in vivo* at all remains challenging.

The cell has evolved mechanisms to modulate proteins from the energy scale of primary structure (~370 kJ/mole amide bond enthalpy) to quaternary structure (20 kJ/mole hemoglobin-hemoglobin enthalpy increase in normal erythrocytes *vs.* sickle cells⁷²). Biomolecule populations depend not just on enthalpy, but also on entropy: reactions producing greater disorder are favored. Thus many biomolecular reactions have free energy changes (combining enthalpy and entropy effects) of just 10s of kJ/mole. For example, as strong as amide bonds are, they actually break spontaneously in acidic water, releasing about 10 kJ/mole of free energy⁷³: the broken C-N amide

* This chapter is adapted with permission from Anna J Wirth and Martin Gruebele, *BioEssays*, 35(11), 984-993, 2013. Copyright 2013 John Wiley and Sons.

bond is partly made up for by forming N-H and C-OH bonds, and on top of that, the two liberated peptides get to move about more randomly.

What about even weaker interactions - weaker than two protein surfaces binding or forming quaternary structure? Could such extremely weak interactions play a meaningful role in the cell? These transient protein-protein interactions are collectively known as “quinary structure”. Since the 1980s, quinary structure has been implicated in many cellular processes ranging from cellular metabolism⁷⁴ to protein translation⁷⁵. In a few cases, illustrated in figure 33, the strength of quinary interactions has been measured quantitatively *in vitro*⁷⁶. Their strength is difficult to quantify *in vitro* in ways that realistically mimic the interior of cells. Consequently, quinary structure is an area of research where *in vivo* study is warranted. Only recently have techniques become available to probe these weak transient interactions directly and quantitatively in the cell. With whole cell computational modeling now feasible, the potential even exists to bridge computation and experiment.

The aim of this Problems and Paradigms essay is to provide a perspective on when and how quinary structure in the cell affects folding and function of biomolecules, and especially proteins. First we will review forces at work in the cell. Then we discuss quinary structure and examples where it may play a role. Finally, we will discuss the newest advances in *in vivo* experimental methodology and whole cell modeling and speculate on the future of the field.

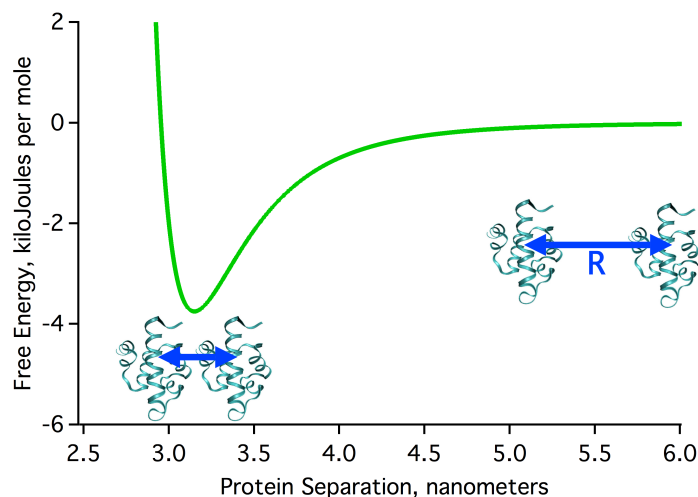


Figure 33 The interaction free energy for two lambda repressor fragment proteins, derived by us from small angle X-ray scattering measurements that look at how much proteins tend to cluster together as their concentration is increased. These proteins do not form quaternary structure or tightly bound dimers. Nonetheless, at a center-to-center separation R of 3.2 nanometers, the proteins sticking to each other with a small free energy release of 4 kJ/mole.

4.2 FORCES IN THE INTRACELLULAR ENVIRONMENT

The intracellular environment is highly crowded—macromolecular concentrations reach up to 400 grams per liter (or mg/mL)⁷¹. Such crowding by complex mixtures of biological molecules contributes only modestly to stabilization or destabilization of proteins, but when proteins interact very weakly to begin with, the consequences of crowding likely play an important role. We briefly review here how crowding, 'stickiness' of the cytoplasm, hydrodynamic forces, and water dynamics manifest themselves from the macromolecular to the cellular level.

4.2.1 How does macromolecular crowding affect protein-protein interactions and protein folding through the excluded volume effect? Rather than arising from attractive intermolecular interactions, the excluded volume effect depends on the size and concentration of molecules crowding the solution. In figure 33, at about 4.2 nanometers, the two proteins 'touch': a third protein simply could not squeeze through between them. If they separate by less than the diameter of the third protein, it will still not fit through, and the volume between the pair is excluded. Only when they separate by enough (pair on the right) is the volume between the two proteins no longer excluded.

The main consequence of excluded volume is that the larger and less ordered form of a protein or protein complex—such as the unfolded state or a non-associated dimer—is entropically disfavored due to the lack of available space⁷⁷. Figure 34a demonstrates how the more compact state of a protein is favored under crowded conditions. The excluded volume for a given protein in solution is defined as any region that protein's center of mass cannot occupy⁷⁸. As the radius of gyration of a protein increases (such as would occur from a transition from the folded to unfolded state) the excluded volume increases, and the apparent protein concentration in the left-over accessible volume is greater than the actual concentration.

The energetic consequences for protein folding are summarized in figure 34b. As more volume is excluded, the unfolded state has no place to fit and becomes less stable, and the unfolded protein crosses the barrier more easily and thus folds more rapidly. This behavior is entirely driven by the entropy term in the free energy relation. Likewise, a compact protein complex is favored over the separated protein monomers - unless the leftover volumes between crowding molecules are so small that the complex just can't fit.

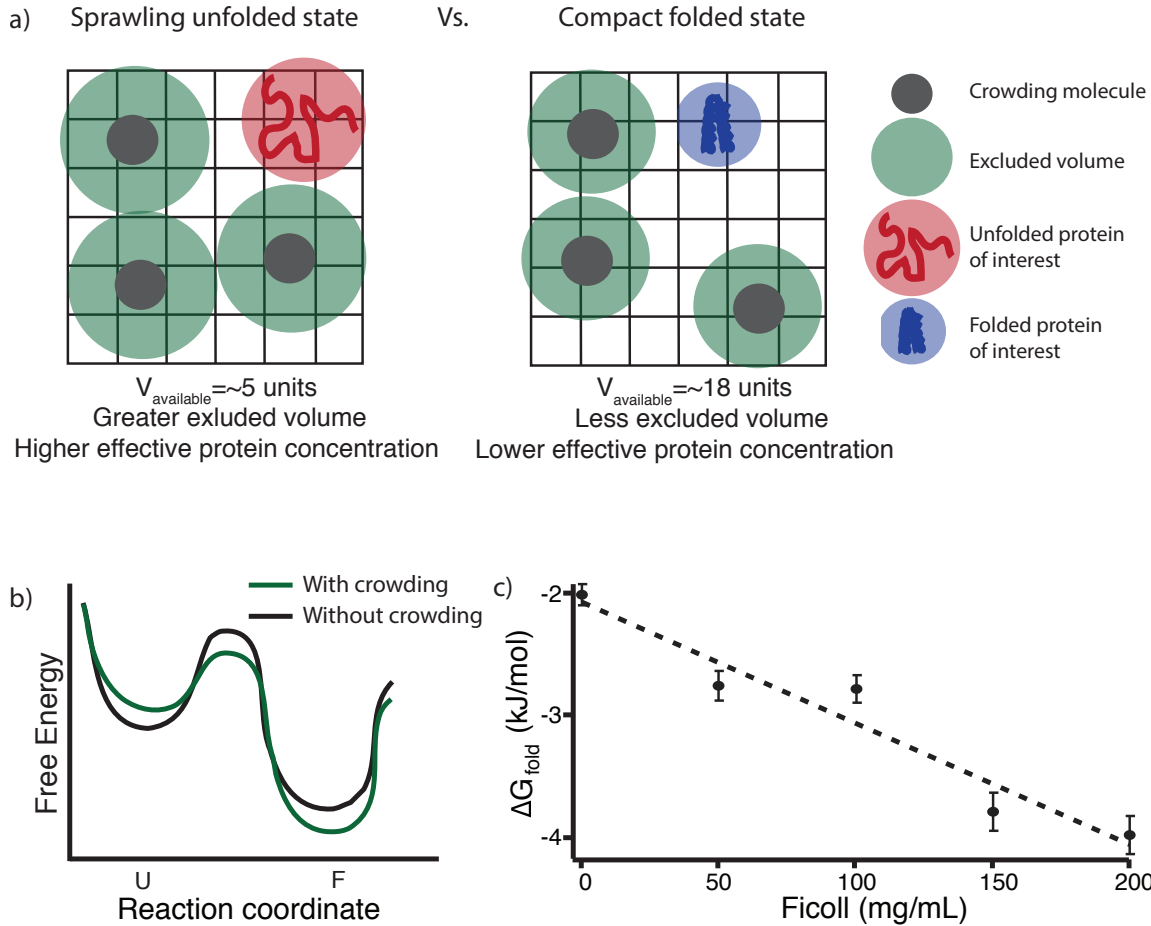


Figure 34 The excluded volume and its energetic consequences. A) The compact, folded state of a protein has less excluded volume than the less-ordered unfolded state even though the size and concentration of crowders is the same. The excluded volume (green) is defined as any region of space that the protein's center of mass cannot occupy. The apparent protein concentration is greater for the unfolded state. B) Increasing crowding decreases the stability of the unfolded state, and decreases the activation barrier for folding. C) Experimentally measured free energy of folding ΔG_{fold} for FRET-labeled yeast phosphoglycerate kinase (PGK) as macromolecular crowding is increased. The protein is stabilized as crowding increases, illustrating the net stabilization of the folded state due to the excluded volume effect⁷⁹. ΔG_{fold} is calculated at a reference temperature of 37 °C.

The role of crowding on protein structure and stability has been studied extensively *in vitro* and *in silico*, and experiment verifies the predictions of theory⁷⁹⁻⁸³. The role that the excluded volume effect plays in biological reactions in general was recently reviewed⁸⁴. In the heterogeneous environment of the cytoplasm, a given protein will be subject to a wide range of excluded volumes of different shapes.

An approximate free energy range for the effect can be assigned based on experimental and theoretical results. A computational study found that at high crowder concentration, the excluded volume effect increased the stability of the folded state by about 1-6 kJ/mole and protein-protein

interactions by 2.5-8.5 kJ/mole⁸⁵. Similarly, in a model of a bacterial cell, it was found that the stabilization of the folded state in the crowded cytoplasm due to the excluded volume effect was about 4-8 kJ/mole and for protein-protein interactions 4-5 kJ/mole⁸⁶. Experimentally, the folded state of FRET-labeled yeast phosphoglycerate kinase was stabilized by 2 kJ/mole when 200 mg/mL Ficoll were added to the solution (figure 34c)⁷⁹.

4.2.2 Can the 'stickiness' of the cytoplasm modulate protein-protein interaction and protein folding in a sequence dependent manner? The excluded volume effect predicts size-specific effects, but there are also sequence-specific effects. Non-specific interactions between proteins in the crowded cellular environment, either through electrostatics (due to charges) or hydrophobicity (mainly the protein's effect on solvent water disorder), make proteins stick to each other or other macromolecules. This makes the apparent viscosity of the cytoplasm larger than bulk water's viscosity⁸⁷.

The diffusion coefficient depends inversely on solution viscosity. Consequently, most experimental efforts to measure cytoplasmic viscosity have relied on diffusion measurements⁸⁸. This effort is hampered, however, by anomalous diffusion—basically a time-dependent diffusion coefficient—which abounds in the cell due to non-uniformity of the environment⁸⁹. In one striking example that highlights the challenges of using translational diffusion to directly measure cytoplasmic viscosity, the translational diffusion of chromosomes has been shown to be sensitive to the metabolic activity level of the cell⁹⁰.

Rotational diffusion, on the other hand, has been shown to be almost exclusively sensitive to localized intermolecular interactions⁹¹. Nuclear Magnetic Resonance (NMR) spectroscopy is highly sensitive to the tumbling of molecules (and thus can extract the rotational diffusion coefficient). Wang et al. found that rotational diffusion coefficients for a test protein varied widely between solutions containing inert crowders, which only increase the excluded volume, and those containing protein crowders, which can stick non-specifically to the protein of interest⁹¹. Another group used a whole cell NMR approach to measure the rotational diffusion of three globular proteins in *E. coli* cells and found that although the proteins studied had similar size and fold, they showed significantly different rotational diffusion⁸⁷. This result suggests that stickiness of the cytoplasm is sequence-dependent.

Computation is not far behind. McGuffee and Elcock used 'Brownian dynamics' to show that specific protein-protein binding interactions were destabilized consistently by cytoplasmic stickiness,

while its effect on protein folding varied significantly from about 10 kJ/mole of destabilization to the same amount of stabilization⁸⁶. MD simulations of proteins in crowded, “sticky”, solutions have shown that the enthalpic consequences of many protein-protein interactions may destabilize the fold of a protein and cause accumulation of non-native and partially denatured states⁹². The effect of specific interactions, on the other hand, are thought to impart general stabilization to a protein fold⁹³.

4.2.3 What happens when water and other molecules get dragged around by proteins inside the cell? Hydrodynamic interactions arise from collisions between macromolecules and solvent molecules and affect the kinetics of protein-protein binding. At physiological temperature, collisions with water molecules impart energy fluctuations of ± 0.04 kJ/mole to a average-size protein. This may seem like a small effect, and when these collisions between solute and solvent are completely random, they give rise to the familiar Brownian motion.

In a crowded environment, however, motions of particles are correlated, and collisions between protein and solute can result in non-random motion. Figure 35 illustrates how hydrodynamic interactions influence the binding of two proteins. As a protein moves towards its partner, the water molecules between the two proteins must be pushed out of the interfacial region, carrying away energy. Some of those displaced water molecules will also bump into the partner protein, pushing it away. Together, these two hydrodynamic effects slow down the kinetics of protein-protein binding⁹⁴.

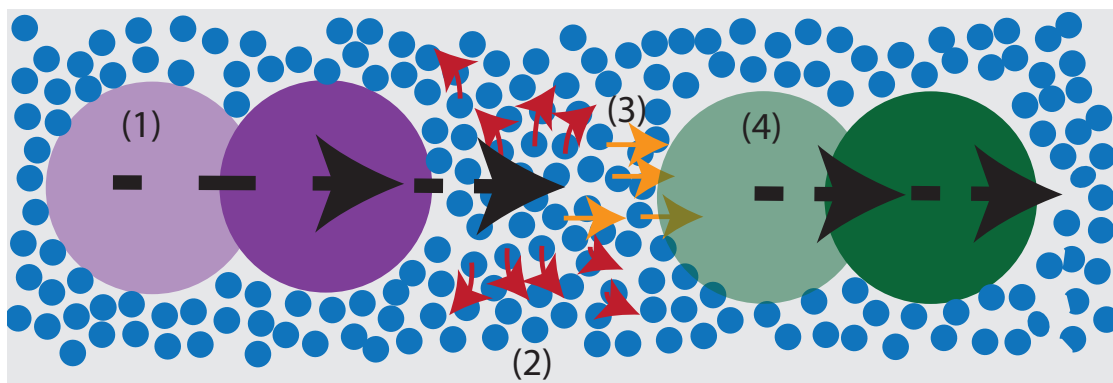


Figure 35 Cartoon depiction of the effect of hydrodynamic interactions on the kinetics of bimolecular association. (1) As one protein (purple) moves towards its binding partner (green), it displaces the water molecules between the two proteins. (2) The water molecules must move from the region between the proteins and reorganize (red arrows). (3) Some water molecules are displaced towards the binding partner (orange arrows) and push the binding partner away. (4) This hydrodynamic repulsion further slows the kinetics of association.

The energetic magnitude of hydrodynamic interactions on proteins—i.e. the reduction or increase of the transition state barrier due to hydrodynamic interactions—is again relatively small. Theory predicts that rates of association should decrease by less than 30% depending on the surface charge of the molecules in question ⁹⁵. Simulations also have explored the role that hydrodynamic interactions can play in protein folding and protein-protein association kinetics. Frembgen-Kesner and Elcock showed that the association rate of proteins decreased by 30-80% in the presence of hydrodynamic interactions ⁹⁶. The same researchers also found that inclusion of hydrodynamics increases the folding rate of proteins 1.5-3 fold, corresponding to at most 3 kJ/mole ⁹⁷. Moving towards simulating hydrodynamic effects in cytoplasmic models is extremely challenging. However, Ando et al. found that inclusion of hydrodynamic interactions in a model *E. coli* cytoplasm along with the excluded volume effect replicated the experimentally measured translational and rotational diffusion of GFP ⁹⁸.

4.2.4 Is there any bulk water left inside the cell? Water often mediates protein-protein interactions, or can be thought of as part of the protein structure. Water molecules inside protein cavities form on average 3 hydrogen bonds, each providing a stabilization of 2.5 kJ/mole ⁹⁹. Of more interest for water-mediated protein interactions is the intermediate regime between bulk solvent water and the “frozen”, long-resident time structural water molecules. This intermediate regime is often called the solvent-shell of the protein and consists of water molecules that are transiently bound to the protein surface ¹⁰⁰, as well as water molecules up to several nanometers away whose hydrogen bonds rearrange more slowly than in the bulk due to the influence of the protein ¹⁰¹. Although NMR experiments indicate some presence of bulk-like water ¹⁰², THz experiments that probe further from protein surfaces indicate that the hydrogen bond rearrangement time is retarded up to 2 nm away from protein surfaces ¹⁰³. At typical packing densities in the cell, the separation between macromolecules is only a few nanometers, leaving little room for any bulk water inside the cell.

Figure 36 shows an example of a protein-protein interaction that is mediated by water. Transiently bound water molecules have been proposed to mediate recognition during protein-protein binding and to contribute to the stability of the interaction ¹⁰⁴. Experiments have used double mutant cycles, which compare the overall loss of stability from binding site mutations that remove hydrogen bond donors or acceptors, to elucidate the energetic contribution of water-

mediated protein interactions to the overall stability of protein complexes. Stabilization, if any, ranges from 0 to about 4 kJ/mole¹⁰⁵⁻¹⁰⁷.

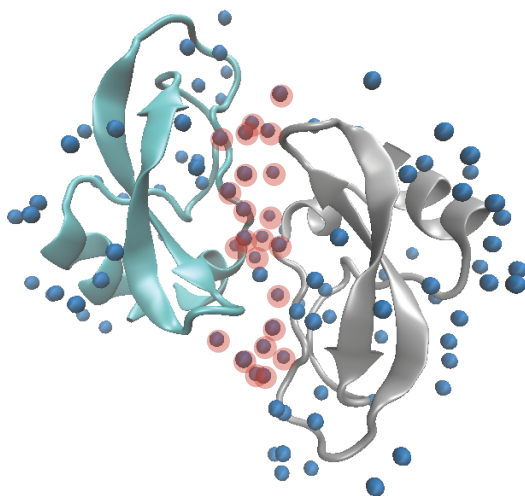


Figure 36 Structure of protease inhibitor domain of the amyloid beta-protein precursor with associated waters (pdb 1AA¹⁰⁸). The density of water molecules at the dimer interface (highlighted in orange) illustrates a water-mediated protein-protein interaction.

4.3 QUINARY STRUCTURE

4.3.1 Quinary structure is the fifth level of protein structural organization. The hierarchy of protein structure and its relation to energy and free energy is shown in figure 37. Quinary structure is the most fleeting, and thus the most likely to be affected by the weak forces acting at the macromolecule to cell level.

The term quinary structure was introduced by McConkey to define a fifth level of protein structural organization: protein-protein interactions that he described as “inherently transient”⁷⁵. Quinary structure, he argued, could be disrupted by the isolation methods of the time, and consequently had not received the recognition they deserved. We discuss quinary structure in terms of three criteria: functionality, thermodynamics, and kinetics.

The quinary structure terminology is intermittently used in the literature, but in our - and others⁷¹ - opinion it deserves to be explicitly recognized. Any individual quinary interaction may not confer a large functional advantage to the cell. Yet even small differences can serve as the basis for evolutionary selection, and a large number of them could add up to a significant advantage. Indeed, many interactions that eventually evolve to tight binding may start out highly transient, while others remain weak to be optimal for the cell.

Very low thermodynamic stability has been used as a criterion for quinary structure (i.e. K_d values $> 1 \mu\text{M}$ for complex formation⁸⁷). However, some interactions have variable affinity. They can be extremely weak unless the proper ligand assists the interaction¹⁰⁹. Such 'three body' effects abound in biology. One particularly important example is the binding of the molecular chaperone DnaK to an unfolded protein at low affinity when ATP is bound, but with high affinity following hydrolysis of ATP to ADP¹¹⁰. McConkey does not exclude such variable protein-protein interactions from his original definition. One of the examples he gives of quinary structure--calmodulin binding to actin fibers in interphase cells¹¹¹--is a variable affinity protein-protein interaction modulated by calcium binding. Thus, in our definition here, quinary structure includes constant low affinity protein-protein interactions, as well as the low affinity state of variable affinity protein-protein interactions.

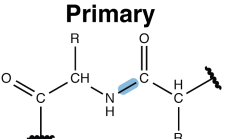
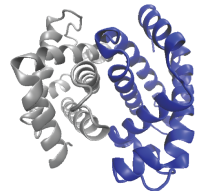
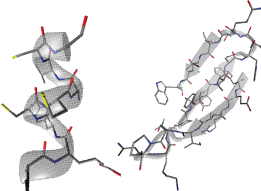
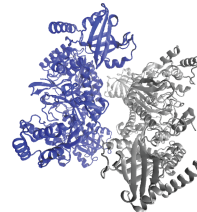
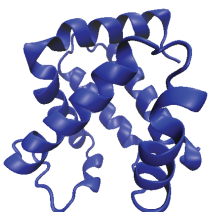
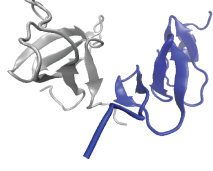
<p>Primary</p> 	<p>Single amide bond: $-\Delta G=8$ to 16 kJ/mol $-\Delta H=350$ to 400 kJ/mol</p>	<p>Quaternary</p> 	<p>H-bond in protein-interphase $-\Delta G=2$ to 7 kJ/mol $-\Delta H=14$ to 20 kJ/mol</p>
<p>Secondary</p> 	<p>H-bond in α-helix or β-sheet: $-\Delta G < 1$ kJ/mol $-\Delta H = 6$ kJ/mol</p>	<p>Quinary</p> 	<p>Quaternary structure: $-\Delta G=40$ to 60 kJ/mol $-\Delta H=60$ to 250 kJ/mol</p>
<p>Tertiary</p> 	<p>Disulfide bond: $-\Delta G=10$ to 13 kJ/mol $-\Delta H=225$ to 275 kJ/mol</p> <p>Salt bridge: $-\Delta G=3-5$ kJ/mol $-\Delta H=13-15$ kJ/mol</p> <p>Overall tertiary structure: $-\Delta G=0$ to 0.1 kJ/mol/res $-\Delta H=0.5$ to 3 kJ/mol/res</p>		<p>Lys1-tRNA synthetase tetramer $\Delta G=-39$ kJ/mol $k_d=3 \times 10^{-7}$ M</p> <p>Nck-2 SH3 domain and PINCH-1 LIM4 domain ultraweak complex: $\Delta G=-15$ kJ/mol $k_d=3 \times 10^{-9}$ M</p>

Figure 37 Structural and energy scales (absolute values around 37 °C) in the cell in terms of enthalpy change ΔH and free energy change ΔG . Enthalpy describes the strength of interactions, while free energy describes which interactions are most likely by also including the effect of disorder. For example, amide bonds are very strong, but nonetheless hydrolyze in water because new NH and OH bonds are formed and because the two peptide fragments have higher disorder. Primary structure^{73,112}, secondary structure^{113,114}, tertiary structure^{79,115-118}, quaternary structure^{119,120}, and two examples of quinary structure^{121,122} are shown. PDB structures (in order of appearance): 1EG3¹²³, 1HHO¹²⁴, 1U55¹²⁵, 3BJU¹²¹.

McConkey's definition of "inherently transient" implies rapid kinetics in addition to low stability. Only when both criteria are fulfilled is an interaction "inherently transient." Consider two states: protein dimer, *vs* two monomers. If these two states have similar free energy, so neither is very stable with respect to the other, their connecting barrier could still be high, trapping the system for a long time in one state or the other. If the two states are connected by a small barrier, but one state is much lower in free energy than the other, the system will simply end up in the lower state, and no function associated with switching states can occur. Only when stability and barriers are both low, is the system inherently transient.

We can use the thermodynamic, kinetic, and functional criteria to define quinary structure as inherently transient (none of the states is thermodynamically favored, nor are any two states separated by a large barrier that traps population), yet having some functional specificity subject to evolutionary selection. Thus quinary interactions between proteins or other biomolecules must be on the order of a few kJ/mole on the free energy scale: larger free energy differences would lock in the lowest energy population, and larger barriers would make each state, once formed, rather long lived. This low free energy scale makes quinary interaction subject to modulation by the forces in the cell acting on the macromolecular to whole-cell scale, as discussed in the previous section.

4.3.2 What is not a quinary structure? Interactions between water molecules are fleeting; they may even be functionally important. However, there is no functional specificity that can be selected by evolution. Water is just water and cannot evolve like a protein. Obligatory protein-protein interactions are also not quinary structure as they form strongly bound thermodynamically stable structures. An example is the arc repressor dimer, which as a monomer is unstable, but as a dimer very stable ^{109,126}. Another example of an obligatory protein-protein interaction is the intrinsically disordered protein P27 that gains structure when it tightly binds and inhibits various Cyclin/Cdk complexes throughout the cell cycle ¹²⁷. These types of interactions will not be classified as quinary because they result in tight binding and/or can be assigned to lower levels of protein structure: quaternary for the dimerization of the arc repressor and tertiary for the folding of p27.

4.3.3 What is quinary structure and what does it do? Quinary structure has been implicated in a number of cellular processes from cell signaling to metabolomics. Its inherent transience allows quinary structure to facilitate dynamic spatial organization of macromolecules in the cytoplasm. It can facilitate loose groupings of proteins when they are working together, but not otherwise. In

other examples, quinary structure can enable a single protein to serve multiple functions depending on its binding partner or serve as a dose-dependent molecular switches. Very recent work has presented empirical and theoretical evidence that a protein in a specific interaction stabilizes its partner's fold and hypothesizes that the protein interaction network is of critical importance to the overall thermodynamic stability of the proteome⁹³.

The most famous and well studied example of quinary structure is the metabolon, a multi-enzyme complex associated with multi-step metabolic pathways. These pathways produce numerous intermediates, ~80% of which have no use in the cell other than to be fed to the next processing enzyme. It would be entropically very unfavorable for a cell to support a random distribution of these molecules and their processing enzyme throughout the entire cytoplasm⁷⁴. Metabolons alleviate this problem by “substrate channeling”: intermediates along a metabolic pathway are kept from dispersing throughout the cytoplasm because they are passed between enzymes associated into quinary structures (thus effectively localizing reactions)¹²⁸.

The protein biosynthesis pathway is one metabolic process that features such quinary structure and is an example of quinary structure providing both spatial organization and enabling protein multifunctionality. Aminoacyl-tRNA synthetases, the family of enzymes which aminacylate tRNA molecules, were first observed in the 1980's to form high molecular weight complexes of then-unknown function¹²⁹. The multisynthetase complexes appear to play multiple roles depending on cellular conditions and since their original discovery have been associated with several processes from substrate channeling in protein biosynthesis to cytokine release in the inflammation response¹³⁰. The exact mode of assembly of these complexes is debated, but many of the protein-protein interactions in the complex are thought to be quite transient¹³¹.

One component of the complex that is involved in quinary structure is the lysl-tRNA synthetase (LysRS). Guo et al. showed that the normally dimeric LysRS forms a weakly associated ($k_d \sim 1 \mu\text{M}$) homotetramer that competes with the formation of the larger multi-synthetase complex¹²¹. Because release of LysRS from the complex is thought to be a trigger for the proinflammatory response, they hypothesize that the competition between these two equilibria allows for LysRS to shuttle between two functions. Thus, in this case, quinary structure not only provides a level of spatial organization for a metabolic pathway as part of a metabolon but also allows for protein multi-functionality.

One of the exciting new frontiers of quinary structure is the intracellular formation of membrane-free macromolecular structures that arise because of phase separation¹³². Weak

association between proteins at sufficiently high concentration has been shown to drive phase separation of proteins into a protein-rich and protein-poor liquid phase *in vitro*¹³³. Such protein-rich droplets behave as a separate liquid phase - that is, they are embedded in but distinct from the surrounding cytoplasm, and they can merge with each other to form larger droplets. These phase-separated regions without membranes have been observed recently *in vivo*, first as the P-granules in germ-line cells, and most recently in an actin regulatory pathway^{134,135}. P-granules are an example of quinary structure where interactions driving the phase transition derive from transient RNA-protein binding¹³⁴.

The recent observation of phase transitions in actin-signaling associated proteins, is another example that relies on transient protein-protein interactions only. Li et al. demonstrated that the SH3 domain and its ligand, the PRM domain, form liquid droplets *in vitro* and found that the ability to form multivalent associations is necessary for phase separation¹³⁵. Moving to the nephrin-NCK-N-WASP system, a signaling pathway associated with actin stimulation, they showed that this complex forms liquid droplets *in vivo* when nephrin is phosphorylated and that the downstream signaling activity sharply increases as phase separation occurs. This example of quinary structure is remarkable because it provides another degree of cytoplasmic structure—one that is far more spatially and temporally dynamic than membrane bound organelles.

When the phase transition is tied to a downstream response, quinary structure enables a dose-dependent, nearly binary switch for the signaling pathway. Only when a sufficient concentration of nephrin is phosphorylated will the phase transition occur, but when the threshold is reached the transition occurs rapidly thus “turning on” the downstream signaling pathway. Such a dose dependent response mediated by quinary structure has also been seen in extracellular signaling pathways associated with the immune system¹³⁶.

In addition to metabolism and signaling, quinary structure can also affect the folding stability and kinetics of proteins. For example, Denos et al. investigated the folding of low concentrations of lambda repressor embedded in a crowding matrix of SubL, a thermophilic protein that remains stable while lambda repressor is thermally unfolded⁵³. Protein-protein interaction raises the melting temperature of lambda repressor by 13 °C. This change is large enough that lambda repressor could switch from 80% unfolded to 80% folded simply by adding the crowder. The folding time of lambda repressor is less than 100 ms, so this is a highly transient equilibrium. Some proteins that are intrinsically disordered *in vitro* (IDPs) may undergo a similar equilibrium shift inside cells, promoted by only transient interactions¹³⁷, rather than the tight binding interactions that we exclude (like p27).

The key take-home message is that quinary structure abounds in cells, and produces far more fine-grained structure inside cells than was originally suspected. Cells are very far from random mixtures of macromolecules. Even in regions that are not bounded by membranes, compartmentalization exists.

4.4 LEAVING THE TEST-TUBE BEHIND

4.4.1 Experimental advances for detecting quinary structure *in vivo* are beginning to yield quantitative information on transient protein-protein interaction in its native environment.

Numerous advances in the past decade have allowed for measurement of protein-protein interactions inside cells; these are thoroughly reviewed elsewhere¹³⁸. A few standouts have great potential to yield useful data on the weakest and more transient protein-protein interactions.

For example, a recent solid-state NMR study of frozen cells produced atomic resolution protein structural information, and as the technique further develops, protein-protein interactions and structural changes of proteins *in vivo* may be characterized in great detail¹³⁹. With improving isotopic labeling schemes and pulse sequences, whole cell NMR of living cells also promises to soon be applicable to eukaryotic cell types¹⁴⁰. Similarly, mass spectrometry has been used to probe how the interior of bacterial cells changes protein stability¹⁴¹ and could be expanded to eukaryotic cells.

Super resolution imaging techniques, which provide imaging capabilities below the diffraction limit, have already provided rich information on the co-localization of proteins¹⁴². One of the more recent and spectacular improvements to this family of techniques include improved temporal and spatial resolution in living cells using STORM in both 2 and 3 dimensions¹⁴³. As time resolution improves even further, such imaging capabilities will undoubtedly allow direct detection of even transiently bound protein complexes.

Spatial and temporal patterns of protein stability and folding kinetics modulated by quinary structure can now be imaged inside living cells using Fast Relaxation Imaging (FRiI). Ebbinghaus et al. review micro-patterned stability of the enzyme phosphoglycerate kinase inside mammalian cells¹⁴⁴. Similarly to *in vitro* observations¹⁴⁵, crowding in the cell increases the average stability of the enzyme, but in addition, microenvironments that persist for at least minutes differ in protein stability and folding rate, as shown in Fig. 38. The folding free energy landscape of the protein is modulated by several kJ/mole inside the cell. It remains to be seen whether such modulation has functional significance.

NMR, mass spectrometry and imaging techniques are not the only areas in which progress has been made. Biochemistry has also offered new solutions for detecting quinary structure *in vivo*. M-track uses a “bait” and “prey” methylation based approach to detect very short-lived interactions between proteins in yeast ¹⁴⁶. Another technique, PICT, uses a “bait” and “prey” approach paired with fluorescence live cell spectroscopy to detect protein-protein interactions in living cells ¹⁴⁷.

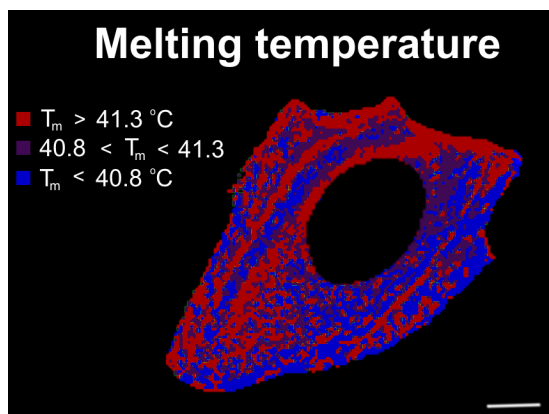


Figure 38 Microenvironments with different protein stabilities inside a U2OS cell (thermodynamic data from ref. ¹⁴⁴). Stability is characterized by the melting temperature T_m , at which 50% of the protein is folded, 50% unfolded. Stability patterns extend from the resolution limit (2 nm) to 10s of μ m in size. The scale bar is 10 μ m in length.

4.4.2 How can computation and experiment work together to study quinary structure? Only now are simulations of large protein complexes and even whole cells becoming feasible, allowing comparison with the new array of experiments probing transient structures inside the cell ^{86,87,148}. By combining computation, theory, and experiment we can attempt to dissect how each of the various forces at play in the intracellular environment affects a specific protein of interest. For the case of excluded volume and cytoplasmic stickiness, theory and experiment are beginning to paint a clear picture of how these two effects compete against one another in the cell. Several theoretical studies have predicted that attractive interactions between crowders and proteins forming a complex (an enthalpic effect) can counteract entropic stabilization of the complex by the excluded volume effect ^{149,150}. Jiao et al. demonstrated this principle for protein association in mixtures of attractive polymers and found that they could tune the stability of a protein complex by using polymers with stronger and weaker attractions to the protein of interest ¹⁵¹. Simulation studies have shown that moderate attraction between proteins and crowding molecules, on the order of 1 kJ/mole, can counteract the excluded volume effect ¹⁵².

The competition of stabilizing and destabilizing effects can also be seen in protein folding inside cells. Dhar et al. directly measured both protein folding thermodynamics and kinetics in U2OS cells¹⁵³. Using a FRET, they found that the folded state of yeast phosphoglycerate kinase was stabilized by 6.3 kJ/mole *in vivo* compared to *in vitro*. On the other hand, Ignatova et al. monitored the unfolding of CRABP I in living bacterial cells and found a net destabilization of 0.8 kJ/mole compared to *in vitro*¹⁵⁴. In an update to their original work, the same group found that although thermodynamic stability between *in vitro* and *in vivo* is not significant, the unfolding rate of CRABP I was significantly faster *in vivo*¹⁵⁵. Philip et al. used a similar FRET construct to directly measure the binding dynamics between β -lactamase to its inhibitor in HeLa cells and found that association was 0.8 to 1.5 kJ/mole destabilized *in vivo* compared to *in vitro*¹⁵⁶. These studies highlight a wide range of behaviors. The cell can both stabilize and destabilize proteins in a sequence-specific manner. Because of this sequence dependence, proteins that are energetically similar *in vitro* might be quite different *in vivo*. The holy grail for simulation will be to predict from protein sequence the effect that the intracellular environment will exert on specific protein-protein interactions and protein folding reactions. In particular, as the field of whole cell modeling continues to improve, it may be possible to predict the effect that a changing intracellular environment has on complex protein-protein interaction networks.

4.5 CONCLUSIONS AND PROSPECTS

The energy and distance scales of the forces that arise in the crowded intracellular environment are such that we can expect them to influence modestly protein tertiary and quaternary structure and to affect significantly quinary structure. The role that quinary structure plays in cell signaling, subcellular localization of proteins, metabolomics, and other systems is clearly significant, and studying these weak protein-interactions in their native environment will become increasingly important if we are to fully understand their role in biology. Experimental and computational advances will likely soon make these types of studies more routine and allow for careful consideration of how the interplay between intracellular forces and protein structure fits into the biology of the cell.

One intriguing possibility is to consider that structure could be dynamically tuned by the changing microenvironments inside the cell, in a yet unexplored mode of protein regulation. Is it conceivable that the physical characteristics of the cytoplasm could change substantially enough make quinary structure an important regulatory mechanism inside the cell?

For one cellular process, the cell cycle, it certainly is. During the procession of events leading to division into two genetically identical daughter cells, the nucleus dissolves during mitosis and dumps its highly charged chromosomes into the cytoplasm. Furthermore, the dry mass of the cell throughout the cell cycle under paces cell volume growth as the cell moves towards division, so the intracellular protein density may also fluctuate in a cell-cycle dependent fashion^{157,158}. Finally, Hinde et al. measured diffusion across the nucleus and showed that diffusion is limited during certain stages of the cell cycle by changing degrees of chromatin compaction¹⁵⁹. This result implies that the cell-cycle environment is extremely dynamic.

Such a strikingly changing environment during the cell cycle invites speculation that the interplay between the physical forces in the cell and quinary structure could be actively exploited by the cell for the purposes of process regulation. Indeed, the tightly regulated series of events that comprise the cell cycle are controlled by an incredibly complex, and still not fully elucidated, network of protein-protein interactions—some of which may be sensitive to environmental fluctuation. As more biophysical studies emerge that examine protein structure and protein-protein interactions across dynamic cellular processes, the possibility of new types of protein regulation can be explored and the *in vivo* structure-function relationship completely understood.

In the following section, protein folding *in vivo* will be studied across a dynamic process, the mammalian cell cycle (chapter 5). Chapter 6 discusses the development of a GFP-ReAsH FRET pair for studying protein folding, a potentially less disruptive dye to study intracellular folding reactions. Finally, chapter 7 discusses early work in characterizing the folding of p53, a cell-cycle related protein, inside living cells.

CHAPTER 5

TEMPORAL VARIATION OF A PROTEIN FOLDING ENERGY LANDSCAPE IN THE CELL*

5.1 INTRODUCTION

Like all chemical reactions, protein folding is dependent on its environment. Solvation, pH, crowding and ionic strength are just a few of the well-documented factors that affect the stability and folding rates of proteins in the test tube.^{84,160,161} Nonetheless, the energy landscape traversed by proteins during folding is usually considered to be constant in time, unless subject to artificial modulation.¹⁶² Yet inside the cell the environment is constantly changing in time. The prime example is the cell cycle itself¹⁶³, when the cell undergoes drastic changes in internal morphology between periods of quiescence and cell division. Here, we use protein stability and folding kinetics to probe this time-dependent cytoplasmic environment.

Here we use the folding reaction of the ubiquitous ATP-producing enzyme phosphoglycerate kinase (PGK) to probe the cell cycle of U2OS bone cancer cells in interphase (G1/S) and mitosis (prophase/metaphase). PGK is a 415 residue enzyme with two domains in the crystal structure,¹⁶⁴ although measurements in cells and in Ficoll crowder indicate that it is even more compact when folded in the crowded milieu of the cell⁷⁹. We label the enzyme with two fluorophores at the N and C termini, so the thermal stability and folding rate of the resulting PGK-FRET construct can be detected inside live cells by change in Förster Resonance Energy Transfer (FRET) upon thermal unfolding.¹⁴⁵ We then combine cell cycle arrest with transient transfection as a potentially higher throughput method than cell line generation to look at time-dependent protein stability and folding kinetics. By measuring changes of the protein melting temperature and folding rate at different times during the cell cycle, we can quantify the temporal variation of the PGK-FRET energy landscape as the cytoplasmic environment changes.

We find that mitotic cells increase the stability of PGK-FRET by about 5 kJ/mole relative to interphase cells. We also conclude that the lower viscosity of mitotic cells overcomes a small increase of the activation barrier during mitosis, leading to faster folding kinetics in mitotic cells relative to interphase cells.

* This chapter is adapted with permission from Anna J Wirth, Max Platkov, and Martin Gruebele, *Journal of the American Chemical Society*, 135(51), 19215-19221, 2013. Copyright 2013 American Chemical Society. MP contributed instrumentation support to this publication.

The same trends were observed previously when PGK-FRET was localized to the nucleus by a nuclear localization tag.¹⁶⁵ It was hypothesized that DNA-protein interaction rather than simple crowding (excluded volume) differences might underlie the greater stability of PGK in the nucleus. Our data is consistent with this earlier proposal, but we show that this scenario cannot by itself explain the mitotic-interphase folding rate difference: imaging of PGK-FRET relaxation kinetics in DNA-rich and DNA-poor regions of the cell reveals at best a small variation of rate coefficient.

5.2 MATERIALS AND METHODS

5.2.1 Cell cycle arrest U2OS cells were cultured using standard practices as described in Appendix E. Cells were arrested at the G1 and S border of interphase with a double thymidine block:¹⁶⁶ 12 hours incubation in 2 mM thymidine (Sigma) supplemented DMEM with 1% PS and 10% FBS, was followed by 12 hours in thymidine free media, and 12 hours of incubation again with thymidine. Cells were arrested in mitosis, between prometaphase and metaphase, by addition of nocodazole¹⁶⁶ (Sigma): 12 hours in DMEM with 1% PS, 10% FBS, and 100 ng/mL nocodazole.

To assess the degree of cell cycle arrest attained, DNA content in cells, which varies predictably in the cell cycle, was quantified by flow cytometry. Cell cultures were fixed in ice cold ethanol, treated with RNAase (Invitrogen), and stained with the intercalating dye, propidium iodide (Sigma). Propidium iodide fluorescence was quantified in each cell via flow cytometry on a BD LSR II with 488 nm excitation and a 695/40 emission filter. Resulting cell count versus fluorescence histograms were fit using standard cell cycle analysis algorithms (FCS Express) to derive percentages of cells in G1, S, and G2/M phases of the cell cycle. Results for mitotic, interphase, and un-arrested cell cultures are shown in figure 39.

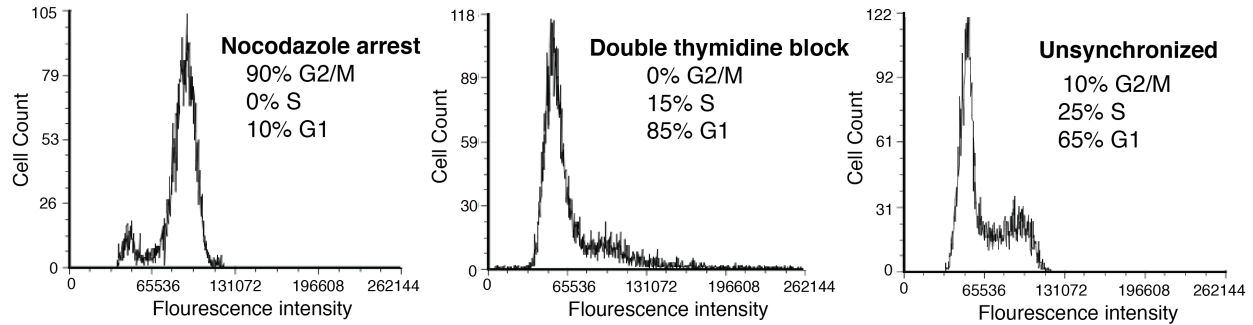


Figure 39 Flow cytometry verification of cell cycle arrest. Fluorescence intensity of propidium iodide stained DNA in each cell is histogrammed for mitotic and interphase arrested cultures and an un-arrested culture. In un-arrested cells (bottom panel) first peak corresponds to cells in G1, middle peak corresponds to cells in S, and third peak corresponds to cells in G2. Peaks centered at the same intensities in the arrested cultures correspond to the same cell cycle phases. To derive percentages of cells in each phase of the cell cycle, the FCS Express fitting cell cycle analysis fitting algorithm was used.

5.2.2 Transfection and plasmid constructs. The plasmid constructed for transfection encoded a FRET-labeled, destabilized, enzymatically active yeast phosphoglycerate kinase (PGK) mutant described previously.¹⁴⁵ A destabilized probe protein allows us to assay the folding within a temperature range acceptable for cell survival, and well below the melting temperature of the FRET labels. The fusion protein consisted of an N-terminal 6 histidine tag, AcGFP1, the destabilized Y122W W308F W333F PGK mutant, and mCherry at the C terminus. The fusion protein was cloned into the pDream 2.1 vector (Genscript) which contains both a T7 and CMV promoter for dual expression in bacterial and mammalian cells. For temperature calibrations, an mCherry-alone construct, also in pDream, was used and consisted of an N-terminal 6-histidine tag and mCherry.

Cells were transfected with 0.6 mg of plasmid per coverslip via Lipofectamine (Invitrogen) according to the manufacturer's protocol. For interphase arrest, cells were transfected during the 12 hour thymidine-free incubation period. Cells were imaged immediately following the second 12 hour thymidine incubation period. For mitotic arrest, cells were transfected 6 hours before addition of nocodazole and imaged immediately after 12 hours of incubation with the arresting agent. In both cases, the arresting agent was not present during imaging to maintain cell viability during experiments.

5.2.3 Measuring protein folding in cells via fast relaxation imaging Fast relaxation imaging has been described in detail elsewhere, and is illustrated in figure 40.¹⁴⁵ Briefly, the model protein PGK-FRET consists of PGK sandwiched between a FRET pair: green fluorescent AcGFP1 as a donor at the N-terminus, and red fluorescent mCherry as an acceptor at the C-terminus, both with melting

points > 70 °C. PGK-FRET has been studied extensively *in vitro* and in cells, making it a suitable model system for cell-cycle-dependent protein folding.^{145,165} It can change in size from about $r=2$ nm C to N terminal distance in the folded state, to as much as $r=17$ nm C to N distance upon unfolding to a random coil.¹⁶⁷ (The actual unfolded state inside cells will be more compact.) These values of r are a good match for the Förster distance R_0 (~ 5 nm) of the two chromophores.

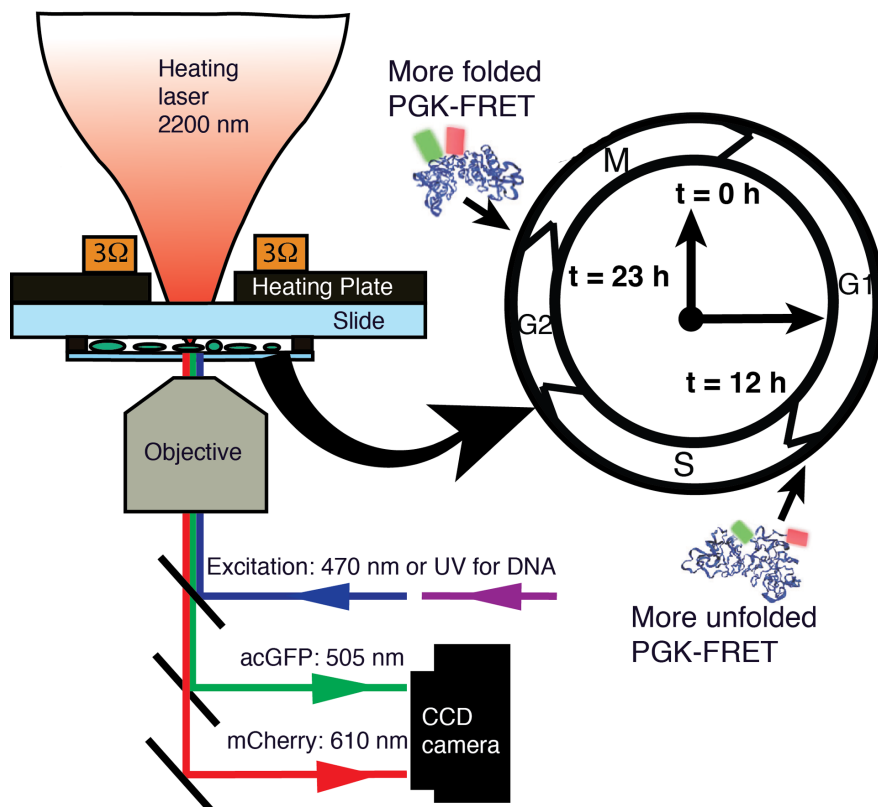


Figure 40 Schematic of the Fast Relaxation Imaging (FReI) experiment coupled to the cell cycle as a clock. U2OS cells divide every 24 hours. At the beginning of gap 1 (G1) or $t=0$, cells grow. In synthesis (S) or $t=12$ hours, the genome of the cell is duplicated. In gap 2 (G2), the cells continue to grow. In mitosis (M) or $t=23$ hours, the cell condenses its chromosomes, and finally divides into two daughter cells. Cells arrested at the end of G1 and at the beginning of M are temporally separated by about 11 hours. Such cells, transfected with PGK-FRET as a probe protein, are imaged to look at changes in stability and folding kinetics during the cell cycle (schematic proteins with arrows pointing at G1 and M). To probe protein stability and kinetics, a green fluorescent donor label is excited by a 470 nm LED, and green donor as well as red acceptor emission are imaged into channels with millisecond time resolution: acGFP (donor) at 505 nm and mCherry (acceptor) at 610 nm. The resulting movies of protein thermodynamics or kinetics are collected during different parts of the cell cycle. IR heating for temperature-jump kinetics and thermal unfolding titrations is provided by a 2200 nm laser.

The donor fluorescence intensity D increases, and the acceptor fluorescence intensity A decreases when the protein unfolds because the two fluorescent tags are further apart in the unfolded state. Thus the ratio D/A increases, the FRET efficiency $E = [1+(D/A)]^{-1}$ decreases, and

the difference D-A increases upon unfolding. PGK-FRET, as a consequence of its large fluorescent tags, shows slightly different folding stability and kinetics than the unlabeled PGK construct.¹⁶⁵ Therefore all comparisons among different environments are made with the same FRET construct, never with wild type PGK.

U2OS cells expressing our probe protein were imaged on a modified epifluorescent microscope, a schematic of which is shown in figure 40. After excitation at 470 nm, emission was split into red and green channels and imaged side by side on a CCD camera at 24 frames per second (fps) for thermodynamics and 60 fps for kinetics. For Hoechst labeling experiments, the dye was excited at 365 nm and resulting 465 nm emission collected in the green channel.

5.2.4 Thermal denaturation experiments Thermal unfolding was used to gauge PGK-FRET stability in G1 and M. The temperature of a cell was raised by laser heating as described previously¹⁶⁸ and in Appendix E. Briefly, a series of shaped pulses from a 2200 nm laser was coupled to a computer-controlled power supply (Thor Labs, LDC340) and a resistive heating stage, set to maintain a specific temperature through a PID (Thor Labs, ITC 510). The laser was mounted above the microscope stage and focused to a spot about the size of the field of view. At the end of each laser pulse, when the temperature had equilibrated, a 2 s video at 24 fps was taken.

To measure the temperature reached in cells at the end of each step, the temperature dependence of the mCherry quantum yield was used. The temperature dependence of the quantum yield of mCherry was measured *in vitro* using recombinantly expressed mCherry (figure 41a). Then, cells transfected with mCherry were heated and the intensity loss at each step used to calibrate the in cell temperature. 9 interphase and 18 mitotic cells were assayed, and their temperatures at each step were averaged (figure 41b). Mitotic cells experienced a greater degree of heating than interphase cells, and thus each cell population has its own temperature calibration. This difference in the heating rate is likely due to morphological differences between mitotic and interphase cells. Interphase cells are fully adhered to the coverslip and in greater contact with the glass, which serves as a heat sink. Mitotic cells, on the other hand, are almost completely detached from the glass and completely surrounded by the medium, which absorbs 2200 nm radiation.

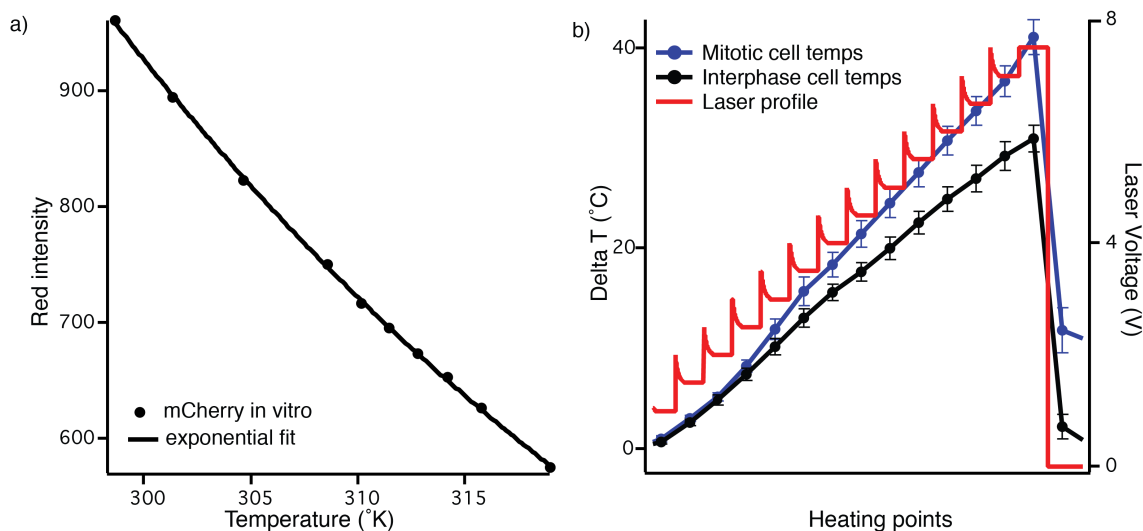


Figure 41 Fast heating thermodynamics temperature calibration. a) mCherry intensity versus temperature measured *in vitro* and corresponding fit used to extract temperature dependence of the mCherry quantum yield. b) Averaged mitotic and interphase in cell temperature calibration from mCherry expressing cells. Error bars show two standard deviations of the mean. Mitotic and interphase cells experience differential heating and require unique temperature calibration. The laser voltage profile is shown in red; as described in the Guo *et al.* reference in the main text, the spike at the beginning of each heating step convolutes with the instrument response to produce a square temperature step.

5.2.5 Thermodynamic data analysis The fluorescence was averaged over the cell area for each channel to generate an overall green and red cell intensity for each temperature. The raw donor to acceptor ratio for each frame is affected by bleaching. To correct, the initial fast exponential bleaching phase, which persists over the first 12 frames, was excluded from analysis. The remaining 36 frames have a bleaching baseline that is linear and straightforward to correct. The corrected intensity trace was then averaged to yield a single red and green intensity for each temperature.

The thermodynamic signal $S(T)$ was expressed as the ratio of donor and acceptor fluorescence intensities $D(T)$ and $A(T)$. A discussion of donor to acceptor ratio as a reaction coordinate is included in Appendix B and the thermodynamic analysis is addressed in detail in Appendix E. The temperature dependent signal was fit to a sigmoidal signal function after normalization¹⁶⁵ and the thermodynamic parameters extracted via a two-state model (Appendix E).

The major approximation made in this analysis is to treat PGK-FRET as a two state folder (two state model). Our criteria were to use the simplest (fewest parameter) model to fit all thermodynamic data within measurement uncertainty to yield the most robust relative changes in T_m across the cell cycle. Although PGK is a two domain multi-state folder,^{169,170} in thermal melts PGK-FRET behaves like an apparent two-state folder.¹⁴⁵ Thus, the thermodynamic traces were fitted to a simple two-state model assuming linear baselines for the folded and unfolded states to extract the most robust thermodynamic parameters, in particular the effective melting temperature (T_m) of the

protein. As seen in ref. ¹⁶⁵, the baselines arise mostly because of the nearly linear temperature-dependent chromophore quantum yields, and when they are eliminated by correcting for temperature-dependent quantum yield, very similar thermodynamic parameters are obtained.

5.2.6 Kinetic experiments and data analysis In a kinetics experiment (described in detail in Appendix E), the cells were heated to 1°C below the average melting temperature of their population with a resistive heating element on the stage and the temperature maintained within 0.1 °C with a PID. Thus, jumps are carried out at constant free energy between the populations rather than constant temperature. Five degree temperature-jumps administered by an IR laser enabled measurement of relaxation into a new thermodynamic state.

Kinetic data analysis is described in detail in Appendix E. Briefly, the fluorescence signal was averaged over the entire cell in each frame or over a region of the cell (for DNA-Hoechst experiments, see next section). The signal immediately following temperature-jump was expressed as the difference between donor and acceptor intensities.

Traces consist of a baseline before the jump, an instrument response lag (< 50 ms in figure 43) that includes the temperature-dependent change of fluorophore quantum yield, and folding kinetics. The magnitude of the instrument response was normalized for all traces and baseline adjusted to 0 (i.e. $S(t=0 \text{ s}) = 0$, where $t=0 \text{ s}$ is the beginning of the jump), the data truncated to the end of the instrument response, and the traces normalized to the end of the jump. All raw data traces as well as all scaled folding traces are summarized in figure 42.

The scaled folding data was fit to a double exponential where the first exponential is a single exponential with a positive amplitude and the second is a stretched exponential with a negative amplitude:

$$S(t) = S_0 + A_1 e^{-k_1 t} + A_2 e^{-(k_2 t)^\beta}$$

The first exponential is an instrument response that arises from differential rates of temperature response between GFP and mCherry and limits the kinetic time resolution. The second stretched exponential describes the actual folding kinetics where k_{obs} is the temperature-jump relaxation rate and β describes the deviation from two-state folding behavior. For fitting folding kinetics, the value of k_{inst} was fixed at the value of the average for each temperature: 7.3 s⁻¹ for jumps from 40.3°C (mitotic cells) and 10.8 s⁻¹ for jumps from 37.5°C (interphase cells).

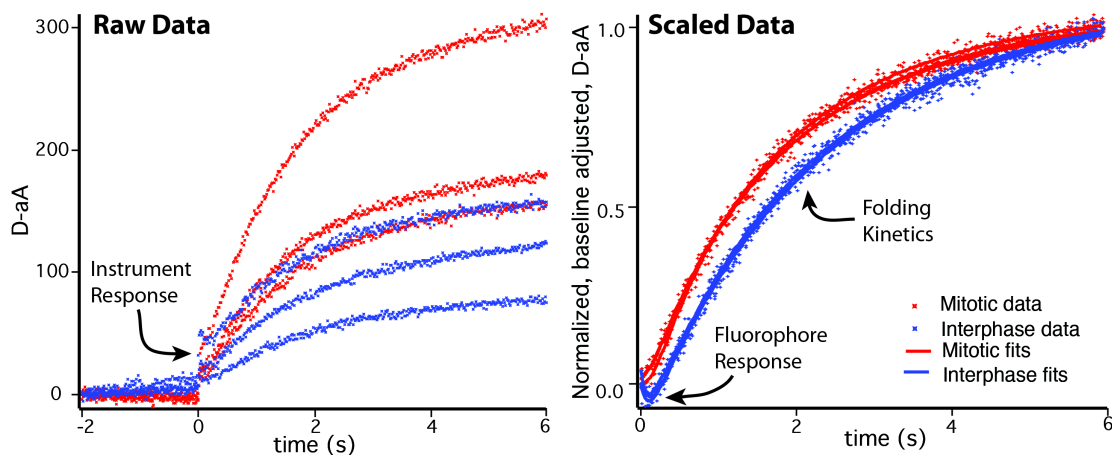


Figure 42 Kinetic data before and after normalization. Left panel shows raw data of 3 mitotic and 3 interphase cells with sharp, un-resolvable, negative phase exponential instrument response labeled. The temperature-jump is initiated at $t=0$. The right panel shows manipulated data with fluorophore and response and folding trace labeled. Data was normalized to top of instrument response, truncated after the instrument response and adjusted so the starting D-aA value is 0, and finally normalized to the maximum amplitude of the folding kinetics.

For figure 45, all mitotic cell traces and all interphase cell traces were averaged and the average traces fit as described above. The fluorophore temperature response exponential was subtracted from the average data and from the fit to yield single stretched exponential traces with fits corresponding to folding data only. The data was then baseline adjusted so that $S(t=0) = 0$ and normalized to the final amplitude so that $S(t=6s) = 1$.

5.2.7 Assessment of measurement heterogeneity To compare the standard deviations of two populations of different sizes, a one-tailed F-test was used to test the null hypothesis that the standard deviation of the mitotic T_m was equal to the standard deviation of the interphase T_m . The ratio of mitotic T_m variance and the interphase T_m variance was 3.53. This is greater than the upper bound F-value of 3.19 for a 95% confidence interval and the null hypothesis was rejected. Thus, despite sample size differences, the mitotic and interphase standard deviations of the T_m are significantly different.

5.2.8 Estimation of free energy contributions to folding rate differences The relative contributions of modulations in the free energy landscape (i.e. activation barrier) and the diffusion coefficient of the protein chain (i.e. the folding prefactor) were estimated as described previously (Dhar *et al.* 2010 in the main paper). Briefly, the observed folding rate is modeled by the equation, valid near T_m :

$$K_{obs}(T \approx T_m) = k_f + k_u = 2k_m e^{-\Delta G^\ddagger / RT}$$

where k_f and k_u are the folding and unfolding rates, k_m is the folding pre-factor (which is proportional to the intramolecular diffusion coefficient of the protein chain), and ΔG^\ddagger is the activation energy. Making the assumptions that

- (1) the phi value for the effects of the intracellular environment is about the same as the average for a single mutation, 0.3, and that energy landscape modulation can be described by the phi value alone,
- (2) the folding can be described by a two-state model and that the free energy of folding is approximately linear with respect to temperature around T_m , and
- (3) the temperature dependence of the folding pre-factor is insignificant over the ΔT_m between interphase and mitotic cells,

we can estimate the activation barrier for folding with thermodynamic data as:

$$\Delta G^\ddagger = \Delta G_0^\ddagger + \phi \delta g_1 (T - T_m)$$

The observed folding rate can then be expressed as:

$$k_{obs}(T, T_m) = k_0 (e^{-\phi \delta g_1 (T - T_m) / RT} + e^{-(\phi - 1) \delta g_1 (T - T_m) / RT})$$

where k_0 is given as:

$$k_0 = 2k_m e^{-\Delta G_0^\ddagger / RT}$$

We can use this expression of the observed folding rate to evaluate the relative contribution of activation energy and prefactor to the observed kinetics. We calculate this in two extremes: 1) assuming all energy landscape variation arises from the stability of the unfolded state, $\Phi = .3$ and 2) assuming all energy landscape variation arises from the stability of the folded state, $\Phi = .7$. For scenario (1):

$$\text{Mitotic cells: } k_{obs}(T=44.3 \text{ }^\circ\text{C}, T_m=41.1 \text{ }^\circ\text{C})=4.38 k_0$$

$$\text{Interphase cells: } k_{obs}(T=41.5 \text{ }^\circ\text{C}, T_m=38.7 \text{ }^\circ\text{C})=5.32 k_0$$

This result would indicate that interphase cells should show 22% faster folding due to lowering of the activation barrier compared to mitotic cells. For scenario (2):

$$\text{Mitotic cells: } k_{obs}(T=44.3 \text{ }^\circ\text{C}, T_m=41.1 \text{ }^\circ\text{C})=2.04 k_0$$

$$\text{Interphase cells: } k_{obs}(T=41.5 \text{ }^\circ\text{C}, T_m=38.7 \text{ }^\circ\text{C})=2.17 k_0$$

This result would indicate that interphase cells should show 7% faster folding than mitotic cells due to lowering of the activation barrier. For both cases, we used the measured k_{obs} to calculate k_0 values, the ratios of which describe the relative magnitude of the diffusion coefficient between mitotic and interphase cells.

5.2.9 Identification of DNA rich regions To identify regions of mitotic cells that were rich in DNA, cells were imaged under UV excitation after Hoechst staining, a UV excitable intercalating dye. Mitotic arrested cells on coverslips were incubated with 3 $\mu\text{g}/\text{mL}$ Hoechst 33258 (Invitrogen) in OptiMEM minimal media (Gibco) supplemented with 10% FBS and 100 ng/mL nocodazole for 30 minutes at 37°C and 5% CO₂. Cells were rinsed gently two times with PBS (Lonza) before imaging. Staining was carried out during the last 30 minutes of nocodazole incubation.

After a DNA-rich region was identified, the kinetics experiment described in the previous subsection was carried out under 470 nm excitation. UV and blue images were aligned by cross-correlation (figure 43). The overlay was used to identify DNA-rich and DNA-poor regions. Pixels corresponding to these two regions were averaged to yield spatially resolved kinetic traces.

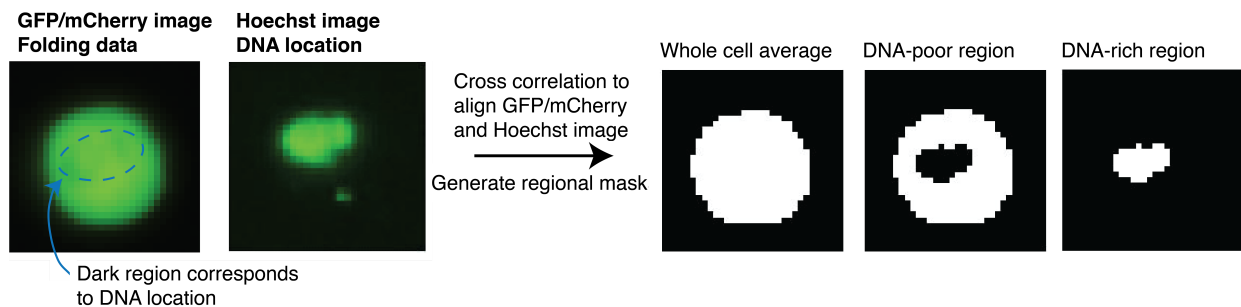


Figure 43 Schematic of the three-color, regional folding experiment. Red and green folding trace, top left panel, is acquired under blue excitation (470 nm), before a Hoechst image showing the position of chromosomes is taken under UV excitation. To align the images, the dark region in the GFP/mCherry region corresponding to the position of the chromosomes is aligned via cross correlation in the Matlab image toolbox to the image of the chromosomes. Masks corresponding to the entire cell, DNA-poor region, and DNA-rich region are generated and the pixels under each mask averaged to produce the spatially resolved kinetic traces.

5.2.10 Analysis of spatially resolved kinetic data FReI images an entire column through the cells. Due to the thickness of mitotic cells, the signal from the DNA-rich region includes contributions from PGK-FRET close to the chromosomes and PGK-FRET in a chromosome-free environment. To address this issue we used the known volume of human mitotic chromosomes¹⁷¹ to estimate the depth of chromosomes in each mitotic cells and assumed a spherical mitotic cell to estimate the

depth of the mitotic cell at each pixel location. The ratio of DNA depth to cell depth, D/C, provided a signal to noise threshold for our measurement. Cells with a D/C ratio below 1:6 were excluded from analysis.

The observed DNA-rich signal is given approximately by the following equation:

$$(C / D + 1)S_{DNA\text{-rich,obs}}(t) = (C / D)S_{DNA\text{-free,obs}}(t) + S_{DNA,actual}(t),$$

where $S_{DNA\text{-rich,obs}}(t)$ is the normalized observed signal from the DNA-rich region with fluorophore response subtracted out, $S_{DNA\text{-free,obs}}(t)$ is the normalized observed signal from the DNA-free region with fluorophore response subtracted out, and $S_{DNA,actual}(t)$ is the extrapolated signal from the chromosome-containing region only. $S_{DNA,actual}(t)$ was adjusted until the above equation agreed with the measured values in the DNA-rich and DNA-free regions. The rate of $S_{DNA,actual}(t)$ that produced the closest fit to the experimental data was reported as the folding rate in the DNA rich region. To generate the main text figure, the amplitude of each trace was weighted by its signal to noise before averaging. The average D/C ratio was then used to simulate $S_{DNA,actual}(t)$ as described above.

5.3 RESULTS

5.3.1 Harnessing the cell cycle “clock” to measure the time-dependence of the intracellular protein folding landscape. Immortalized cell lines, such as the U2OS line used here, constantly move through the cell cycle in a strictly time-regulated fashion completing a full cycle every 24 hours. Figure 39 illustrates the time dependence of this process as studied here. Arresting cells at a stage of the cell cycle allows for characterization of the intracellular environment across an entire population of cells at a single point in time.

We used this concept to develop a technique to study the time dependence of an intracellular protein folding landscape. Using common molecular biology techniques,¹⁶⁶ we arrested cells just before DNA synthesis begins, in interphase, which corresponds to $t \sim 12$ hours relative to the beginning of the cell cycle. To access a later time point, $t \sim 23$ hours, we arrested cells in the early stages of mitosis. Degree of arrest attained for both time points was verified by flow cytometry (figure 71). We combined this cell-cycle arrest approach with transient transfection to enable relatively rapid study of different proteins without the need to generate unique cell lines expressing a labeled protein of interest.

5.3.2 The thermodynamic stability of PGK-FRET is time-dependent in living cells The thermodynamics of PGK-FRET unfolding was measured in the cytoplasm of 18 mitotic cells and 9 interphase cells (figure 44). Fits of the two data sets yielded 41.3 ± 0.8 °C for mitotic cells and 38.8 ± 0.6 for interphase cells. PGK-FRET is more stable in mitotic cells, as evidenced by their 2.5 ± 0.6 °C higher mean effective melting temperature (T_m) compared to interphase cells (table 39).

The square root of the variance of T_m , which describes the degree of cell-to-cell variation rather than relative measurement uncertainty, was 1.5 °C in mitotic cells and 0.9 °C in interphase cells, indicating greater cell to cell variation in mitotic cells. The greater heterogeneity of mitotic cells (figure 44c) may arise from cell-to-cell differences in cellular structure only present during mitosis. Mitosis is a far more morphologically diverse phase than the G1/S border, with varying degrees of chromatin compaction or extent of mitotic spindle formation. Heterogeneity differences may also rise from the faster time-scale of mitosis. The many morphologically diverse events of mitosis occur over the course of an hour, so slight differences in cell cycle timing could produce greater heterogeneity. Interphase cells, on the other hand, experience no dramatic morphological changes for many hours, so the exact timing of cell-cycle arrest has less of an effect.

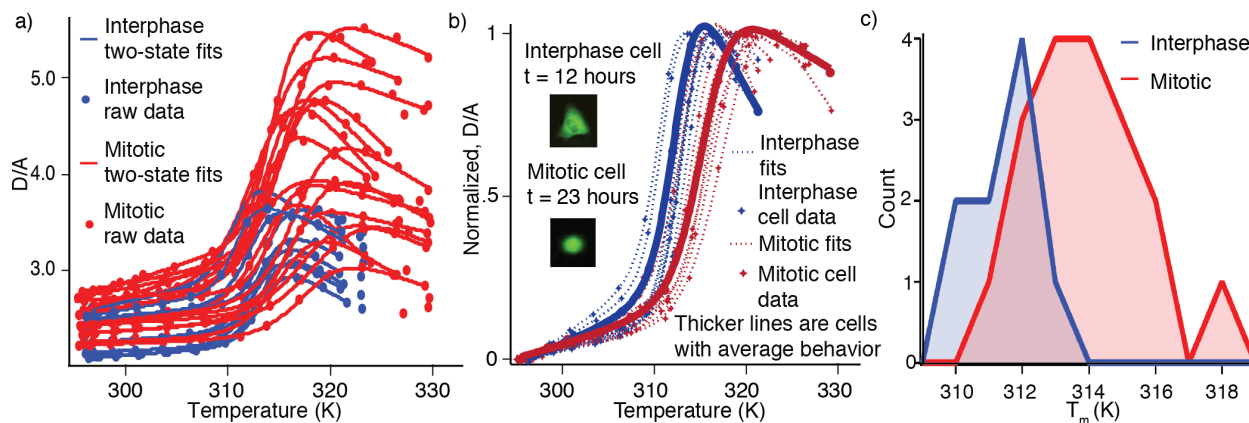


FIGURE 44 PGK-FRET thermal melts as a function of the cell cycle in mitotic- and interphase-arrested U2OS cells. Larger donor/acceptor (D/A) ratios correspond to unfolded protein. (A) Raw thermodynamic data. Un-normalized, baseline un-adjusted mitotic (18 cells) and interphase (9 cells) data is plotted with corresponding two-state fits. (B) For clarity, 9 randomly selected mitotic cells and all 9 interphase cells are displayed and the data is scaled from 0 to 1. The highlighted traces are for two cells that exemplify the average behavior of each cell cycle population. Insets show morphology of an interphase and a mitotic cell. Data is not corrected for temperature dependence of the Förster constant (see SI for corrected data, and a plot in terms of end-to-end distance R instead of D/A). (C) Histogrammed fitted T_m values for all interphase and mitotic cells showing overlapping populations and greater heterogeneity in mitotic cells. Note that some of this heterogeneity could come from variations in the timing of the cell cycle arrest.

	ΔT_m (°C)	$\Delta H_{U \rightarrow F}$ (kJ mol ⁻¹)	$\Delta S_{U \rightarrow F}$ (kJ mol ⁻¹ K ⁻¹)
Mitotic	+2.5 ± 0.6	-485 ± 54	-1.54 ± 0.17
Interphase	0	-665 ± 37	-2.13 ± 0.12
In vitro	-2.5 ± 0.6	-459 ± 201	-1.48 ± 0.65

Table 4 Effective thermodynamic parameters. All values based on two-state fits. Uncertainty reported is two standard deviations of the mean. See Methods for references to the in vitro data. Thermodynamic fits also extracted the effective two-state difference in enthalpy and entropy of PGK-FRET unfolding between interphase and mitotic cells. Both the Van't Hoff enthalpy ($\Delta H_{U \rightarrow F}$) and entropy ($\Delta S_{U \rightarrow F}$) of folding are greater for PGK-FRET in interphase cells than in mitotic cells (table 1).

5.3.3 Compactness of the unfolded state of PGK-FRET is time-dependent. Protein compactness over the course of a thermodynamics experiment can be assessed by the D/A ratio since $r = (D/A)1/6 R_0$. Below the unfolding temperature, interphase and mitotic cells show a similar value of D/A 2.33 ± 0.13 for interphase cells and 2.51 ± 0.17 for mitotic cells (figure 44a). However, upon unfolding the maximum D/A of PGK-FRET ranges from 3.1 to 5.5 in mitotic cells, but only from 2.9 to 3.6 for interphase cells. This implies that the unfolded state of PGK-FRET is more expanded in mitotic cells than in interphase cells, or at least that the FRET labels are less crowded.

5.3.4 PGK-FRET folding kinetics varies with cell cycle. Folding/unfolding relaxation kinetics traces of PGK-FRET were collected on 3 mitotic cells and 3 interphase cells via 5 °C temperature-jumps from 1 °C below their respective T_m (figure 45). The resulting kinetic traces of PGK-FRET relaxation were fit to stretched exponentials $\text{Signal} \sim \exp[-(kt)^\beta]$ as described in Methods. The observed relaxation rate k equals about twice the folding rate for jumps close to the melting temperature. The stretching factor β describes the deviation from two-state behavior, or multiple populations with different relaxation rates. Mitotic cells showed faster relaxation and a greater deviation from two-state folding with a mean k of 0.62 ± 0.04 s⁻¹ and β of 0.84 ± 0.01 compared to interphase cells with k of 0.44 ± 0.01 s⁻¹ and a β of 0.95 ± 0.03 .

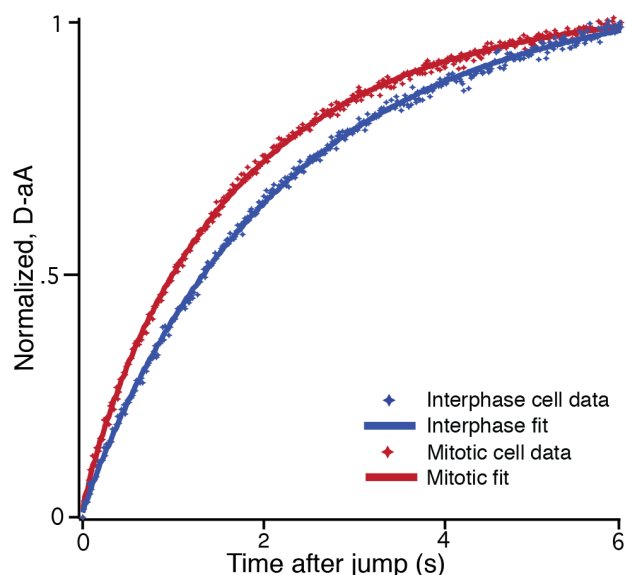


Figure 45 Averaged mitotic and interphase kinetic traces. As shown here, the fluorophore temperature response has been subtracted from the data and the data normalized. Figure 42 shows raw kinetic data traces for all cells, with the fluorophore temperature response included. The T-jump relaxation was fit to yield PGK-FRET relaxation rates k and an exponent b that describes deviations from two-state kinetics (see SI). T-jump begins at $t=0$ s, and higher D-aA means more unfolded protein. Mitotic cells (red) show a faster relaxation rate than interphase cells (blue).

To investigate whether the 'kinetic stretching' in mitotic cells is just due to averaging over different rates in different parts of the cell, or reflects a change in folding mechanism throughout the cell, we examined folding kinetics separately in two different regions of mitotic cells: a DNA-rich area whose location was identified by staining with Hoechst 33258 labeling, and the remaining DNA-poor region (figures 43 and 46). A difference between DNA-rich and poor regions is also of interest because faster folding of PGK-FRET has been observed in the nucleus of interphase cells than in the cytoplasm,¹⁶⁵ so protein-DNA interaction in mitotic cells, where the nuclear envelope has dissolved, could be the reason for the different rates observed in figure 45. The average difference between the DNA-rich kinetics (orange points in figure 46, table 5) and DNA-poor kinetics (green points in figure 46, table 5) is far too small to account for the faster mitotic kinetics in figure 45. When we take into account that the DNA does not occupy the whole column probed by our diffraction-limited imaging, the extrapolated DNA-only rate (orange dashed lines in figure 46) is 22% faster on average in mitotic cells, with a $\pm 24\%$ root-mean-square variation among 5 cells measured. Thus our results are consistent with the earlier proposal that DNA-PGK-FRET interaction could speed up folding relaxation kinetics,¹⁶⁵ but they cannot account for the mitotic-interphase difference observed here in figure 45.

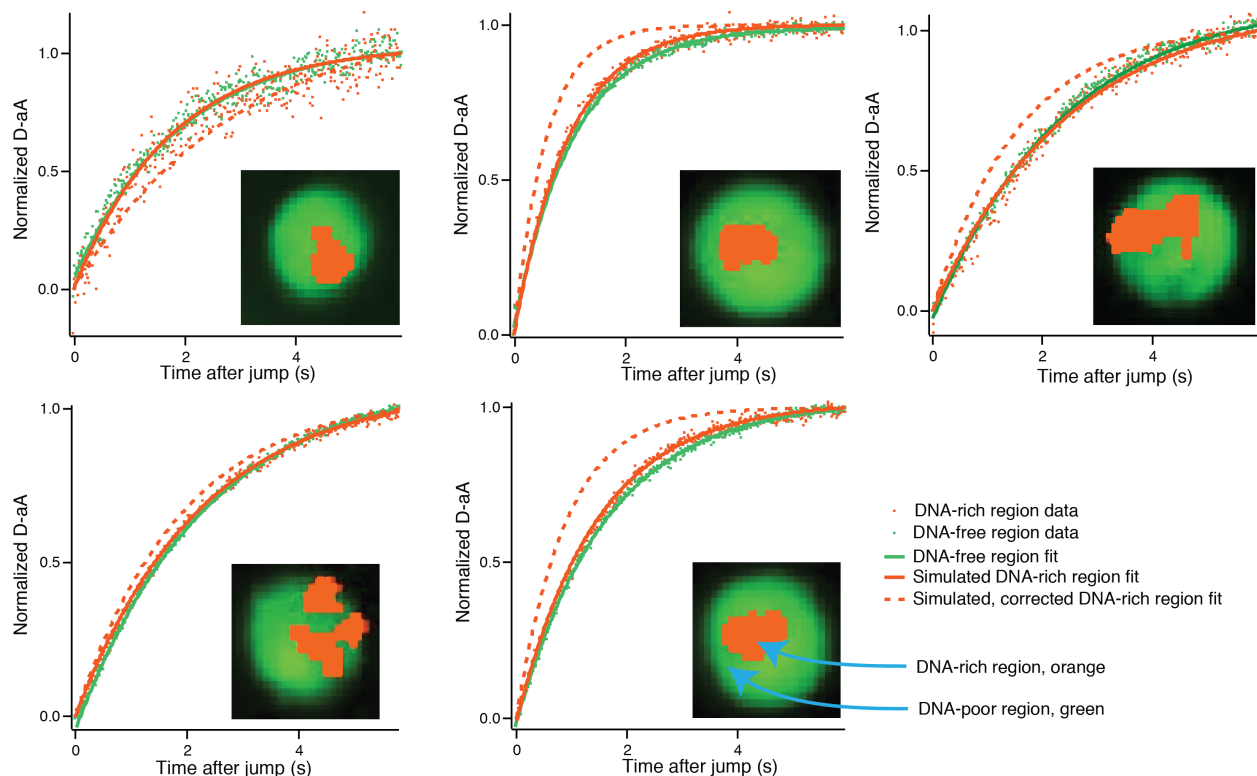


Figure 46: Spatially resolved kinetics. Normalized kinetics with the temperature response of the fluorophore removed for the DNA-rich and DNA-free region data are shown in orange and green. Bold orange traces show the simulated fit for the observed DNA-signal, $S_{DNA,obs}(t)$, which contains contributions from PGK-FRET close to and far away from DNA in the Z-dimension. Dashed orange lines show the corrected DNA-rich signal, $S_{DNA,act}(t)$. Bold green trace is the single exponential fit to the DNA-free region data. Inset shows each cell with Hoechst-stained DNA colored orange.

Cell #	% difference, observed	% difference, extrapolated to chromosomes only
Cell 1	-3.64	-19.3
Cell 2	9.96	21.8
Cell 3	8.85	45.6
Cell 4	3.21	17.9
Cell 5	9.62	44.5

Table 5 Percent difference between DNA and DNA-free region both observed and corrected. Percent difference is calculated as $100 \times (k_{DNA} - k_{no\ DNA}) / k_{DNA}$.

5.3.5 The diffusion coefficient for folding increases during mitosis. Like any chemical reaction in a solvent environment, protein folding rates depend on the prefactor and the activation barrier. The temperature dependence of the folding rate can be described by the Arrhenius equation:

$$k(T \approx T_m) = 2k_m \exp(-\Delta G^\ddagger / RT) \quad [1]$$

where ΔG^\ddagger is the free energy barrier for folding and k_m is the prefactor, which depends on internal friction of the polypeptide chain and on the material properties of the solvent, i.e. viscosity.^{172,173} The factor of 2 in equation [1] accounts for the fact that near the T_m , where our experiments are carried

out, the observed relaxation rate coefficient is the sum of equal folding and unfolding rate coefficients, or $k=k_f+k_u\approx 2k_f$.

A time-dependent energy landscape (barrier) and prefactor (diffusion) can both contribute to the observed time-dependent rate of reaction as the cell cycle progresses, and their relative contributions can provide insight into the mechanism underlying the rate differences between the two populations studied here.

Using Phi-value analysis,¹⁷⁴ we calculated the relative contributions of landscape vs. diffusion to the observed folding rates in interphase and mitotic cells. Phi values describe how much the activation free energy of a reaction changes when the free energy of the reactant (here: folded protein) changes because of a mutation or, in our case, a change in the cellular environment. For example, $\Phi=1$ implies that stabilization of the native state in mitotic cells lowers the activation energy by an equal amount whereas $\Phi=0$ implies that stabilization of the native state has no effect. The average Phi value for a single point mutation in a protein is 0.3,^{175,176} and we assume that the effects of a changing cellular environment can be approximated by this same Phi value.¹⁴⁵

To assign an upper and lower limit to calculated changes of the activation barrier and diffusion coefficient between interphase and mitotic cells, two extreme Phi value scenarios were considered. In the first, we assumed that most energy landscape differences between mitotic and interphase cells arise from free energy differences in the folded state ($\Phi=0.7$) and in the second, we assumed the complement: most changes result from free energy variations of the unfolded state ($\Phi=0.3$). Given the structural similarity of the folded states and the structural difference in the unfolded states in figure 44b, the latter scenario is more likely.

In both scenarios, the activation barrier for folding is higher in mitotic cells than in interphase cells. In other words, if the activation barrier alone were to control folding rates, PGK-FRET would fold faster in interphase cells than mitotic cells. Yet, the opposite is observed in figure 45. This indicates that the prefactor more than makes up for the slightly higher barrier in mitotic cells or, in other words, the difference in prefactors indicates that diffusion for the folding reaction is 1.5-1.7 times faster in mitotic cells than in interphase cells.

5.4 DISCUSSION

The lives of cells are dynamic by nature—biomolecule and ion concentrations fluctuate and cellular structure is reorganized across many cellular processes—and one must consider whether such changes influence the biophysical and chemical properties of the intracellular environment

enough to affect the energetics of protein function or folding with respect to time. Ideally, one would study such effects *in vitro* where experiments have improved signal and allow for greater manipulation of variables. Unfortunately, the cellular milieu is simply too complex to reconstruct in the test tube, at least at this point in time, particularly when one is interested in dynamics.

The folding free energy landscape of proteins *in vivo* is influenced by the intracellular environment.^{71,165} To measure the time dependence of a folding landscape, we used living cells as our test tube and a biological process, the cell cycle, as our clock. We studied a model protein, FRET-labeled yeast phosphoglycerate kinase, because its folding both *in vitro* and in the cell is well understood. We found that PGK-FRET is stabilized at a later time in the life of a cell, mitosis, compared to an earlier time, interphase. This is the first direct evidence that a protein energy landscape can fluctuate over time inside cells. We also found that our probe protein folds faster during mitosis due to an increased prefactor (faster diffusion) overcoming an increased barrier (energy landscape). Finally, we found that the rate effects are not associated with protein-DNA proximity as previously postulated for PGK localized to the nucleus,¹⁶⁵ so the exact mechanism behind these changes in stability and folding kinetics remains to be unraveled.

The time-dependent changes in the folding landscape can be rationalized by three factors: (1) Cell cycle-dependent excluded volume, an entropic effect; (2) variation in the interaction between PGK-FRET and the intracellular environment, which can make both enthalpic and entropic contributions; and (3) changes in the degree of spatial heterogeneity in folding environments. To apply these three factors to our data, we will make the assumption that the folded states of PGK-FRET are similar in mitotic and interphase cells (model in figure 47). This assumption is justified by the similar range of D/A values in these two environments (figure 44a). In contrast, the denatured state has a larger average D/A value in mitotic cells than in interphase cells, consistent with a less compact denatured state in mitosis than during interphase.

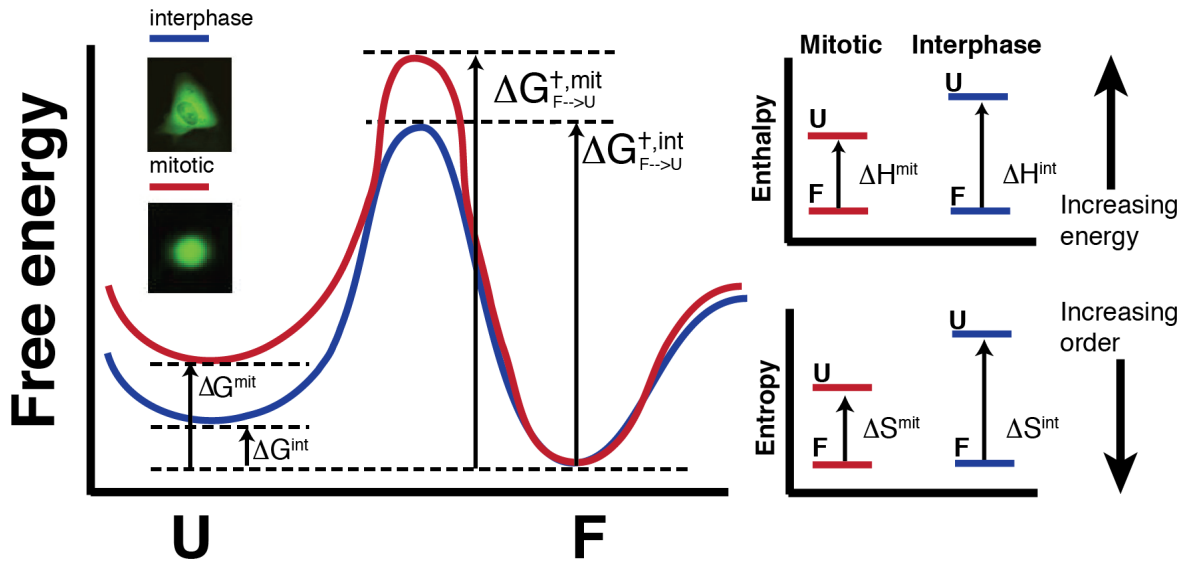


Figure 47 Schematic of proposed time-dependent energy landscape changes. Left panel: Free energy of PGK-FRET's folded state is similar between interphase and mitotic cells. The difference in thermodynamic stability arises from the greater instability of the unfolded state in mitotic cells which pushes the equilibrium towards the folded state. The activation barrier for folding is higher in mitotic cells. Right panels: Mitotic cells show enthalpic stabilization of the unfolded state, which destabilizes the folded state. However, the unfolded state of mitotic cells is more ordered than in interphase cells, a stabilizing effect for the folded state that overcomes the destabilizing enthalpic effects.

5.4.1 The cellular environment is more crowded than *in vitro*, but how do two in-cell environments separated by time differ from one another? Three lines of evidence indicate that excluded volume (macromolecular crowding) decreases as the cell cycle progresses from interphase to mitosis. First of all, the denatured state becomes less compact as the cell progresses from interphase to mitosis, as expected for less crowding. Secondly, the enthalpy and entropy of folding during mitosis more closely resemble the *in vitro* data measured in buffer (table 4). Most compelling is the kinetic data. Our Phi value analysis indicates that mitotic cells have slightly higher activation barriers for folding and ~ 1.6 times faster intramolecular diffusion than interphase cells, both expected for reduced crowding.^{80,177} Thus, we conclude that mitotic cells are less crowded than interphase cells. The altered cytoskeletal structure, evident in the round shape of mitotic cells could be a reason for decreased crowding in mitotic cells.

5.4.2 Usually, increased melting temperatures of proteins are explained by increased crowding. If crowding decreases during mitosis but T_m increases (table 4), another factor must be responsible. We believe that increased interaction of the unfolded state with the intracellular environment is a good candidate. First of all, the denatured state is less compact during

mitosis, giving the unfolded polypeptide chain more opportunity to interact with its environment. In table 4, the effective enthalpy of the denatured state is significantly lower during mitosis than interphase, consistent with increased binding of the denatured state to macromolecular surfaces, solvated ions, or other microstructure in the cell. (The same is incidentally true *in vitro*, indicating some interaction between FRET labels and PGK.)

A lower enthalpy of the denatured state alone would tend to decrease T_m during mitosis, so it must be offset by entropy. Indeed, in table 4 the effective entropy of the unfolded state is significantly lower during mitosis. This is also consistent with increased interactions of the unfolded polypeptide chain 'sticking' to other macromolecules in its environment, including the GFP and mCherry labels attached to the PGK. In this scenario, crowding is reduced in the mitotic state, and the resulting more extended unfolded state of PGK sticks to other cellular components, decreasing its enthalpy, but even more its entropy, thus destabilizing the unfolded state and increasing T_m .

Favorable enthalpic interactions between a protein's unfolded state and biomolecules or ions in the cell are not without precedent.^{149,151} For example, a recent molecular dynamics study showed that non-specific protein-protein interactions can lead to the accumulation of non-native or partially unfolded intermediates, suggesting that the unfolded state is enthalpically stabilized by such interactions (the simulations did not address the entropic effect of multiple crowders).⁹² If the unfolded state is less compact in mitotic cells, it may make more such stabilizing interactions (lower enthalpy, as observed), while being pinned down more due to such interactions (lower entropy, as observed). As another example, a statistical mechanical model¹⁵⁰ has shown that attractive interactions between protein and crowder can counteract excluded volume effects. And of course, many *in vitro* studies have shown that protein folding enthalpies depend on ionic strength.¹⁶⁰ Finally, previous *in cell* studies have shown that the stabilization imparted from moving a protein from *in vitro* to the cytoplasm can vary widely from a net stabilizing effect,¹⁵³ virtually no effect on stability,^{154,155} and destabilization.^{156,178} Such protein-specific effects implicate the importance of ionic or hydrophobic interactions between a protein and its surrounding environment, which evidence has shown are sequence dependent in cells.⁸⁷ Nonetheless, it is worth pointing out again that the overall effect of in-cell crowding on PGK is stabilizing in both mitotic and interphase cells relative to *in vitro*, and that the balance of enthalpic and entropic contributions plays a role in fine-tuning stability as a function of the cell cycle.

5.4.3 Some of the effects we observe here are consistent with both energy landscape and spatial variations of PGK folding in the cell. For example, consider our measurement of β averaged over the whole cell. Kinetics with $\beta < 1$ might simply reflect spatial averaging over a distribution of two-state rate coefficients inside the cell, indicating spatial heterogeneity. In that scenario, mitotic cells with smaller β are more spatially heterogeneous than interphase cells. On the other hand, the variation of β could also be caused by enthalpic stabilization of unfolded states in mitotic cells, as discussed above: Stabilization of folding intermediates would lead to a departure from two-state folding signaled by $\beta < 1$. In that scenario, mitotic cells could be spatially homogeneous, but the energy landscape of the protein has changed towards multi-state folding throughout the cell. The two scenarios are not inconsistent with one another, and both spatial and energy landscape variations are likely to exist within the cell.

To see if we could detect spatial variations of folding during the cell cycle, we measured folding kinetics in both DNA-rich and DNA-poor columns within mitotic cells, which have dissolved the nuclear envelope, but show condensed nuclear material (figure 46). Previous work has shown that PGK-FRET localized to the nucleus of interphase cells is more stable and folds faster than in the cytoplasm.¹⁶⁵ Therefore it was postulated that nuclear DNA-protein interaction could be responsible for faster folding and greater stability. We found that although there is a very weak trend towards faster folding close to the mitotic chromosomes, this difference cannot account for the faster folding observed in mitotic cells, nor is it as large as that observed between nucleus and cytoplasm of interphase cells. Thus, the loss of nuclear compartmentalization results in greater spatial homogeneity of the folding environment in mitotic cells than in interphase cells. We therefore conclude that the smaller β in mitotic cells is either due to cell cycle variation of the folding energy landscape, consistent with our scenario of less crowding and enthalpic stabilization of unfolded states in mitotic cells, or if it is due to spatial heterogeneity, the latter must be at the level of cellular microstructure below a few μm .

5.5 CONCLUSIONS AND OUTLOOK

The methodology developed here enables future studies of the time-dependence of the in-cell folding landscapes of other proteins. Of particular interest in the field of molecular biology is the study of cell-cycle related proteins. An intriguing possibility raised by our results is that of protein regulation by functional energy landscape or folding energy landscape modulation: could a

protein that is active in interphase be thermodynamically stabilized in that phase of the cell cycle and destabilized in others, or vice versa? Our results here are of particular importance to proteins with transient structural elements such as intrinsically disordered proteins. In these cases, modulations of the energy landscape of just ~ 5 kJ/mol observed here (based on ΔT_m in figure 44 and δg_i in table 1, $\Delta\Delta G = \Delta T_m \delta g_i$) could have dramatic effects on a protein's structure, stability or function, thanks to the amplifying effect of the exponential Boltzmann factor $\exp[-\Delta\Delta G/RT]$. Considering our results, these types of proteins may be far more structurally dynamic with respect to time in their native cellular environment than expected, especially if the cell poises them near structural transitions.

CHAPTER 6

DEVELOPMENT OF THE GFP-REASH FRET PAIR TO STUDY PROTEIN FOLDING*

6.1 INTRODUCTION

Since the development of the green fluorescent protein first made it possible image proteins inside living cells¹⁷⁹⁻¹⁸¹, an area of active research has been the development of new fluorophores for live-cell microscopy. A huge library of fluorescent proteins that can be appended to a protein of interest in a fusion construct have been developed covering the spectral range from 424 nm¹⁸² to 720 nm¹⁸³. Their versatility has enabled their use to study intracellular protein localization and function in a wide variety of optically demanding applications, even live-cell super resolution microscopy¹⁸⁴.

Fluorescent proteins as fluorophores and protein localization tags, however, have their drawbacks¹⁸⁵. For one, quantum yield and photostability of fluorescent proteins cannot match the best small-molecule dyes, which poses imaging limitations. Fluorescent proteins are also not necessarily inert when used as tags. Of concern to those that study protein structure or structure-sensitive processes, fluorescent proteins are large enough to substantially disrupt structure and function, as is the case for HSP70 which cannot tolerate N-terminal fluorescent protein fusions¹⁸⁶. Fluorescent proteins are also known to cause clustering or localization artifacts in living cells, which has lead to observation of artificial protein foci formation in bacteria¹⁸⁷.

A number of strategies to generate fluorescent protein alternatives for live-cell imaging have emerged over the years including quantum dot-labeled proteins¹⁸⁸, hijacking of native proteins to label a protein of interest with dyes¹⁸⁹, and using non-native amino-acids as bio-orthogonal handles for cell-membrane permeable dyes to bind¹⁹⁰. Another somewhat less technically demanding approach is to use labeling chemistry that links a cell-permeable dye to a specific, short peptide sequence appended to the termini of the protein of interest—essentially taking advantage of the ease-of-use of a genetically encoded tag while using a far smaller and less obtrusive fluorophore. The first example of this technique was FAsH¹⁹¹, which was developed in Roger Tsien's lab.

FAsH is a system that consists of a fluorescein-based dye conjugated to a bi-arsenical motif that covalently binds a cysteine rich sequence, CCPGCC (shown in figure 48). Before binding, the dye is minimally fluorescent. However, when it binds to the recognition sequence, the structure

* This work was done in equal collaboration with Hannah Gelman, and it is discussed half in this thesis and half in H.G.'s thesis. This work is partially reproduced from Gelman, Wirth, and Gruebele (in preparation).

rigidifies and becomes brightly fluorescent. Following the development of FAsH, a red fluorescent version of the system, ReAsH was developed¹⁹².

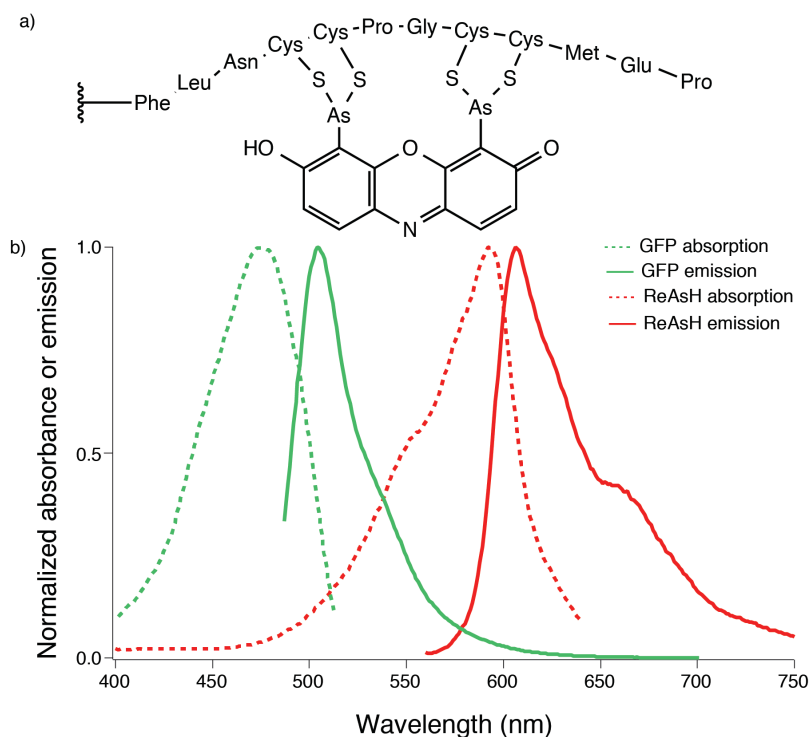


Figure 48 GFP/ReAsH FRET pair. A) ReAsH dye bound to the tetracysteine recognition sequence. In this configuration, the dye is rigid and fluorescent. B) Absorbance and fluorescence spectra for GFP and ReAsH. Spectral overlap between GFP emission and ReAsH absorption is significant, making the two fluorophores a good FRET pair. Spectra were acquired from Invitrogen Life Technologies (ReAsH) and Beckman Dickenson (acGFP1).

FAsH and ReAsH have been used successfully in many biological imaging applications including ReAsH enabled photo-activation of DAB (an electron microscopy probe) in the Golgi apparatus¹⁹³, genetically encoded calcium indicators¹⁹⁴, and FRET-based measurement of the activation of a G-protein coupled receptor¹⁹⁵. Researchers have also taken advantage of the sensitivity of FAsH and ReAsH labeling reaction efficiency to protein conformation. In one example, the tetracysteine recognition sequence was split across two protein domains or between two binding partners so that the amount of FAsH labeling correlated to the amount of folded protein or bound complex¹⁹⁶. Accessibility of FAsH or ReAsH binding sites has also been used to determine the solvent exposure of specific protein regions in amyloid fibrils¹⁹⁷. In some cases, FAsH fluorescence appears to be sensitive to its surrounding environment. Ignatova and coworkers

used the change in FAsH fluorescence intensity to monitor the unfolding of a model protein in living bacteria¹⁹⁸.

ReAsH and FAsH have also been used in Forster Resonance Energy Transfer (FRET) based applications including CFP-FAsH FRET to monitor a receptor's activation¹⁹⁵, GFP-ReAsH FRET to quantify ReAsH labeling in cells¹⁹⁹, and FAsH-ReAsH FRET to monitor alpha-synuclein fibril formation²⁰⁰. Figure 48b shows that ReAsH and acGFP1 make an excellent FRET pair with robust spectral overlap. Their estimated R_0 , based on quantum yields and spectral overlap is on the order of 6 nm.

In this chapter, we discuss the efforts to develop this FRET pair into a useful probe for protein folding both *in vitro* and in living cells. Here, the focus is placed on expression of this construct in mammalian U2OS cells while the partnering work in Hannah Gelman's thesis discusses use of this construct in bacterial cells. The main contribution of this work is development of ReAsH as a quantitative probe for protein folding both as a FRET acceptor and as a directly excited single fluorophore that reports on protein structure. While quantitative interpretation of ReAsH data must be approached with caution due to its strong tendency to bleach, our exhaustive characterization of the fluorophore pair across a variety of constructs provides a framework to interpret GFP/ReAsH data quantitatively both *in vitro* and in cells for protein folding thermodynamics and kinetics.

6.2 METHODS

6.2.1 Construct design The three constructs used throughout the experiment are shown in figure 49. In the dually labeled constructs, FRET serves to monitor the separation between termini, an established reaction coordinate that tracks folding for yeast phosphoglycerate kinase (PGK)⁷⁹. Various combinations of acGFP1, PGK, and the tetracysteine binding domain were generated as shown in figure 49 to provide full spectral characterization of the acGFP1-ReAsH FRET pair.

PGK-tc is a fusion of acGFP1, the destabilized Y122W W308 W333F PGK^{79,153}, and the ReAsH tetra-cysteine binding sequence: FLNCCPGCCMEP. The ReAsH binding sequence with flanking amino acids is the sequence optimized for binding in mammalian cells¹⁹⁹. GFP-tc is an identical construct to PGK-tc but without the internal PGK. Both protein constructs were cloned into the pDream 2.1 vector (Genscript) which contains both a T7 and CMV promoter for dual expression in bacterial and mammalian cells. For purification, each construct has an N-terminal 6-histidine tag. The tetracysteine peptide (tc-peptide) consisted only of the ReAsH binding site (FLNCCPGCCMEP).



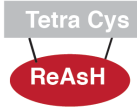
Construct	Purpose
	PGK folding probed by FRET
	Temperature dependence of acGFP and ReAsH bleaching
	GFP independent bleaching behavior of ReAsH

Figure 49 Schematic of the primary constructs used throughout the chapter to characterize the GFP/ReAsH FRET-pair: PGK-tc, GFP-tc, and tc-peptide (from top to bottom). The major utility of each construct is summarized in the figure.

6.2.2 Peptide synthesis and recombinant expression of PGK-tc and GFP-tc The tc-peptide was synthesized by GenScript. The peptide was soluble in water and stored at $-20\text{ }^{\circ}\text{C}$ at a concentration of 1.37 mg/mL per manufacturer's instructions.

GFP-tc and PGK-tc were both expressed recombinantly in bacteria. BL21-DE3-RIPL (Agilent) bacteria cells were transformed with the plasmid construct and selected overnight on ampicillin plates. Cells were grown under 1 mg/mL ampicillin selection to an OD_{600} of .6-1 before induction with 1 mM IPTG at room temperature. After 12-16 hours induction, cells were spun down and resuspended in a lysis buffer consisting of 300 mM NaCl, 50 mM Na_2HPO_4 and 5 mM BME at pH 8.0. Cells were broken via sonication and the cell extract containing soluble protein purified. The extract was loaded via FPLC onto a His-trap column (GE Healthcare) and the his-tagged constructs eluted by a stepped gradient of imidazole.

6.2.3 Cell culture and transfection U2OS cells were cultured as described in Appendix E. Cells were transfected via Lipofectamine (Invitrogen) 24 hours prior to ReAsH labeling and imaging with 2 μg of PGK-tc DNA per 75 cm^2 flask. After 4-6 hours of incubation with lipofectamine and DNA, cells were trypsinized and split onto coverslips in DMEM supplemented with 10% FBS. At the time of ReAsH labeling, cells were 80-90% confluent on the coverslips.

6.2.4 ReAsH labeling In cell ReAsH labeling was carried out according to manufacturer's (Invitrogen) protocols, with some modifications. Cells expressing PGK-tc and adhered to coverslips

were rinsed once with Optimem minimal media (Invitrogen) to remove all traces of phenol-red containing media. Cells were stained with 1 mL of Optimem minimal media supplemented with ReAsH dye (Invitrogen) to a final concentration of 1.75 μM . The labeling reaction was allowed to proceed for 1 hour. Cells were washed with 1 mL of Optimem with 250 μM BAL (Invitrogen) twice. Following labeling, cells prepared on coverslips as described in Appendix E were imaged immediately.

In vitro ReAsH labeling is discussed in detail in H. Gelman's thesis and is a modified version of the method described in Ignatova et al.¹⁵⁴ Briefly, recombinantly expressed PGK-tc or GFP-tc was reduced with BAL and TCEP or BME in degassed buffer for 2 hours. ReAsH dye in 2-10 fold excess was added and allowed to incubate for 2 hours while the reaction was monitored via fluorimeter. Centrifugal filter units (Amicon) were used to buffer exchange to remove unbound ReAsH after the reaction had completed.

6.2.5 Thermodynamics and kinetic measurements For both thermodynamic and kinetic measurements, the construct in living cells or *in vitro* was excited at 470 nm ("blue") for FRET-probed folding or at 590 nm ("amber") for direct excitation of the ReAsH. Imaging was carried out for both *in vitro* and in cell samples as described in Appendix E. For thermodynamic measurements, the slide was resistively heated by two resistors and the temperature measured on the slide surface by a thermocouple as previously described¹⁵³. Kinetics measurements were carried out as described in Appendix E: at each temperature, a 5 degree jump was initiated by an IR laser and the resulting relaxation measured for 6 seconds.

In *in vitro* experiments, kinetics were measured at each temperature for every sample. For *in vivo* measurements, this approach was only tenable in especially bright cells with very robust ReAsH labeling due to the tendency of ReAsH to bleach significantly over the course of an experiment with no recovery of signal, unlike *in vitro* experiments where diffusion from outside the field of view reduces overall intensity loss significantly. For less bright cells, kinetics were taken only at room temperature, two temperatures close to the melting temperature, and one temperature far above the melting temperature. This enabled measurement of kinetics without excessive bleaching of the cell.

6.2.6 Fluorimeter thermodynamic melts and analysis Thermodynamics of PGK-tc unfolding were measured via fluorimeter (Cary) equipped with a Peltier temperature controller (Agilent), where photo-bleaching is significantly less than on the live-cell microscope system. Samples of 2 μM

protein or peptide were assayed under blue (480 nm) or amber (590 nm) excitation. To provide a close comparison to thermodynamic melts on the live-cell microscope, which collect fluorescence in channels, green intensity was taken as the integrated intensity of the spectra from 490 to 550 nm and red intensity was taken as the integrated intensity of the spectra from 550 to 750 nm. Signal was expressed as the donor to acceptor ratio and fit to a two state thermodynamic model (described in Appendix E).

6.2.7 Microscope bleaching correction Baselines arising from irreversible photo-bleaching were a significant problem throughout the experiments. For temperature-jump measurements, the last second of the jump in each channel (after the kinetics were complete but where a bleaching baseline remained) was fit to a line, which was subtracted from the data.

Thermodynamics measurements were slightly more complex to correct due to the variation of bleaching behaviors at short time-scales (see figure 73) and between different constructs. Bleaching behavior could be characterized as either an initial fast exponential decay followed by a linear baseline or a 2nd order polynomial:

$$\text{Exponential: } S(t) = A_o + Ae^{-t/\tau_{bleach}} + m_{bleach}t$$

$$\text{Polynomial: } S(t) = A_0 + m_{bleach}t + c_{bleach}t^2$$

The data were fit to the function that best described the bleaching and the data corrected so that the average value of the trace matched the fitted A_o value.

6.2.8 Kinetics analysis Kinetics were fit to a double exponential as described in Appendix E:

$$S(t) = y_o + A_1e^{-t/\tau_1} + A_2e^{-(t/\tau_2)}$$

Before fitting, the kinetics were corrected for bleaching as described in the previous section. For FRET-probed kinetics, the kinetic signal was expressed as D-aA where D and A are the time dependent donor and acceptor fluorescence, respectively, and a is the initial donor to acceptor ratio at the start of the kinetics. Directly excited kinetics, which consist of only red channel (i.e. acceptor channel) fluorescence, were expressed as a normalized red intensity where the intensity before the jump was used to normalize.

6.2.9 Thermodynamics fitting In most cases, the average donor (acGFP) to acceptor (ReAsH) ratio yielded poor thermodynamics either due to bleaching artifacts or due to strong folded and unfolded baselines that obscured the folding transition. Using the method outlined in Girdhar *et al*²⁰¹ (referred throughout the thesis as the “thermodynamics from kinetics fit”) the kinetic measurements were used to extract folding thermodynamics. The fluorescence signal (either D-aA or normalized red intensity, depending on excitation method) in the last second of the jump, after the kinetics ended, was averaged. The resulting end of jump signal was plotted versus temperature and fit to the following equation to yield the melting temperature, T_m , and the free energy:

$$A(T) = \frac{-g_1 \Delta T T_m}{R(T - \Delta T / 2)^2} [A_o + m_o (T - T_m)] \\ \times \frac{e^{-g^1(T - \Delta T / 2 - T_m) / R(T - \Delta T / 2)}}{(1 + e^{-g^1(T - \Delta T / 2 - T_m) / R(T - \Delta T / 2)})^2}$$

where ΔT is the size of the temperature-jump and R is the gas constant. A_o and m_o are the slope and intercept of $B(T)$, which is the difference of the folded and unfolded baselines (as defined for a traditional, 2-state thermodynamic fit as discussed in Appendix E). The free energy is given by:

$$\Delta G(T) = g_1 (T - T_m)$$

The advantages of using this fitting equation is that it uses fewer parameters than the traditional thermodynamics fitting function (discussed in Appendix E) and can extract thermodynamics from data whose baselines are ill-defined or obscure the folding transition.

6.3 RESULTS AND DISCUSSION

6.3.1 Optimization of *in vivo* ReAsH labeling conditions Although the ReAsH dye is non-fluorescent when unbound from the tetra-cysteine sequence, inside living cells it does show some background fluorescence. This background fluorescence is thought to arise from transient binding to macromolecules in the cell or through partial covalent linkage to cysteine-rich proteins²⁰². The inset of figure 50 shows the background evident in cells that are not expressing the tetracysteine tagged construct. While the background itself would not be problematic, it does show a temperature dependence (figure 50), which can interfere with interpretation of the fluorescence from the labeled construct.

One method²⁰² to reduce background fluorescence is to treat cells with excess BAL, a reducing agent that competes with ReAsH to bind to the tetra-cysteine sequence, to eliminate non-

specific covalent linkage to proteins other than the tetracysteine-tagged construct. We found that treatment with BAL did not significantly reduce background fluorescence, suggesting that background we see in our experiments arises from transient non-covalent binding of ReAsH to other macromolecules in the cell. To reduce this background, we added Disperse Blue (Sigma), a dye that is spectrally distinct from ReAsH but competes with free ReAsH for transient binding to macromolecules. While this approach reduced background fluorescence in the red channel, it introduced new fluorescence background into the green-channel under blue excitation, making it incompatible with the FRET construct.

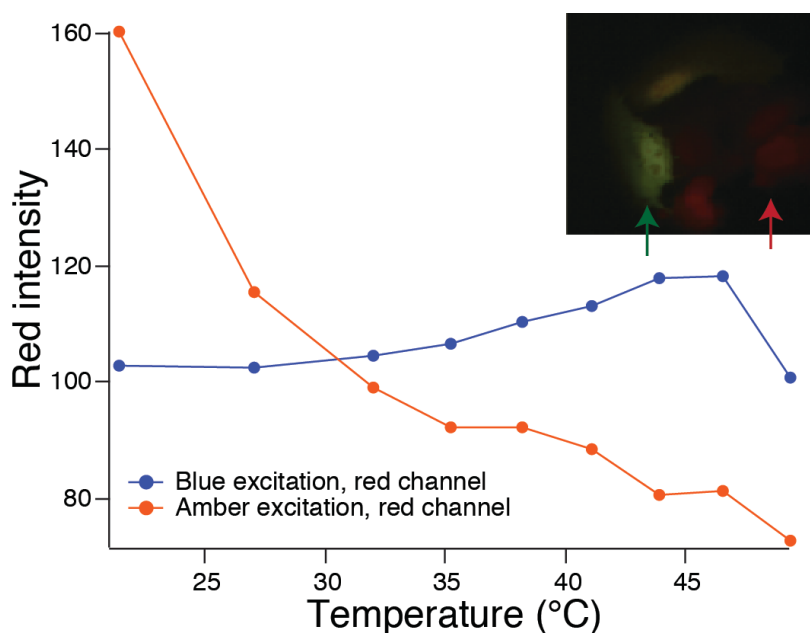


Figure 50 Unbound ReAsH in living cells shows significant fluorescence background under both blue (indirect) and amber (direct) excitation. Plot shows the temperature dependence of the red intensity of cells expressing no tetracysteine tag labeled proteins but treated with ReAsH under blue (blue line) or amber (red line) excitation. Inset shows a microscope field of view showing a cell transfected with a GFP and ReAsH labeled protein (green arrow) and a cell with no transfected protein (red arrow). The non-transfected cell shows significant red background fluorescence due to unbound ReAsH, even under blue excitation.

Thus, the most expedient way to eliminate background fluorescence was to strike a balance between adding enough ReAsH dye to ensure robust labeling but not so much dye that there was significant amounts of unbound ReAsH in the cell. The amount of ReAsH necessary to achieve this balance depends on the number of cells expressing the construct, the expression level of the construct in the cells, and the density of cells on the coverslip. Because these three characteristics vary slightly from experiment to experiment, there is a range of ReAsH concentrations for PGK-tc

in U2OS cells at a cell density (confluence) of 70-90% that produce appropriate labeling levels: 1.25 μM to 1.75 μM (these are 50-70% the manufacturer's recommendations for labeling concentrations). The best practice to achieve background free labeling is to label coverslips at 3-4 concentrations within this range and to choose the lowest concentration that produces an acceptable level of red fluorescence. With this approach, we were able to achieve labeling that had a very low or unobservable background at the blue excitation intensities used for imaging the FRET-labeled construct.

6.3.2 ReAsH shows significant bleaching when excited via FRET from acGFP ReAsH containing fusion constructs show significant photo-bleaching under blue excitation, a known limitation of FAsH and ReAsH²⁰³. Figure 73 (appendix C) shows the red (acceptor) channel fluorescence over time under continuous blue excitation for the PGK-tc and GFP-tc constructs. As might be expected, there is a resonance energy transfer efficiency dependence of the ReAsH bleaching under blue excitation. Table 6 shows the correlation between D/A ratio, which is inversely proportional to FRET-efficiency, and percent intensity loss over a video.

Construct	D/A	% intensity loss
PGK-tc	.9	27
GFP-tc	.62	38

Table 6 Correlation between FRET efficiency and bleaching. FRET efficiency was measured via fluorimeter at 20 °C. Percent intensity loss is in the red (donor) channel and is measured on an *in vitro* sample on the live-cell microscope at room temperature. Bound ReAsH to PGK-tc and GFP-tc were excited indirectly via FRET from GFP under continuous blue LED illumination (both samples were under identical illumination conditions). Over the 6 second video, the percent intensity loss is greater for the GFP-tc construct, which has a lower D/A or higher FRET efficiency.

Because the donor and acceptor fluorescence are linked through resonance energy transfer, the intensity loss in the acceptor channel is coupled to an intensity increase in the donor channel (figure 73 in appendix C). As more and more acceptor molecules bleach, there are fewer acceptor molecules available for energy transfer from the acGFP donor and, accordingly, the donor fluorescence increases creating a positively sloped donor fluorescence over time.

The strong photo-sensitivity of the ReAsH dye poses significant imaging challenges. First, because the bleaching rate is fast, bleaching during the times when the sample is illuminated but data is not recorded, such as during focusing of the microscope, is significant and may need to be considered when attempting to make quantitative intensity measurements. Second, tracking the change in FRET when such significant bleaching is present and when channel bleaching baselines go in opposite directions can be especially difficult. For example, in an intracellular measurement of

protein folding using a GFP-ReAsH FRET construct, the depletion of all ReAsH fluorophores in the cell due to bleaching results in a lower average red intensity and higher average green intensity. This fluorescence signature is difficult to distinguish from the loss of FRET due to expansion of the protein structure, as occurs during unfolding. This problem also persists in *in vitro* samples, although to a lesser degree as diffusion within the sample can replenish bleached molecules in the field of view. The temperature dependent diffusion rate, which is also protein construct dependent, provides additional difficulties in interpreting *in vitro* data.

There are a number of approaches that can be used ameliorate some of these problems. If possible, measurements should be made using ReAsH under direct excitation where bleaching is more straightforward to interpret. In the case of PGK-tc, the ReAsH dye alone appears to be sensitive to the protein folding structural transition (see following sections), which makes this approach feasible. Another alternative is to chemically modify ReAsH to reduce bleaching: one group has reported that fluorinated ReAsH and FAsH derivatives can improve photostability 70-fold²⁰³.

When FRET-type measurements are necessary, fewer measurements on a single cell should be made to avoid bleaching artifacts. For example, rather than performing temperature-jump kinetics at a range of temperatures on a single cell, it would be better to perform one or two kinetic measurements per cell and to use population level data to generate a full temperature dependence of the folding kinetics.

6.3.3 PGK-tc folding thermodynamics and kinetics *in vitro* and in living cells measured via FRET *In vitro* measurements of PGK-tc folding thermodynamics on the live-cell microscope were challenged by ReAsH bleaching, which made average intensities difficult to interpret. Bleaching was temperature-dependent and the temperature-dependent recovery from bleaching, due to diffusion from areas of the coverslip not within the field of view, made accurate bleaching corrections of average intensities non-trivial. Figure 51a shows that the thermodynamics measured on the microscope by taking the average intensity at each temperature before a temperature-jump experiment—corrected for bleaching as described in the methods section—are difficult to interpret, while in the fluorimeter, where little bleaching occurs, thermodynamics are well-defined giving a melting temperature of 315 K.

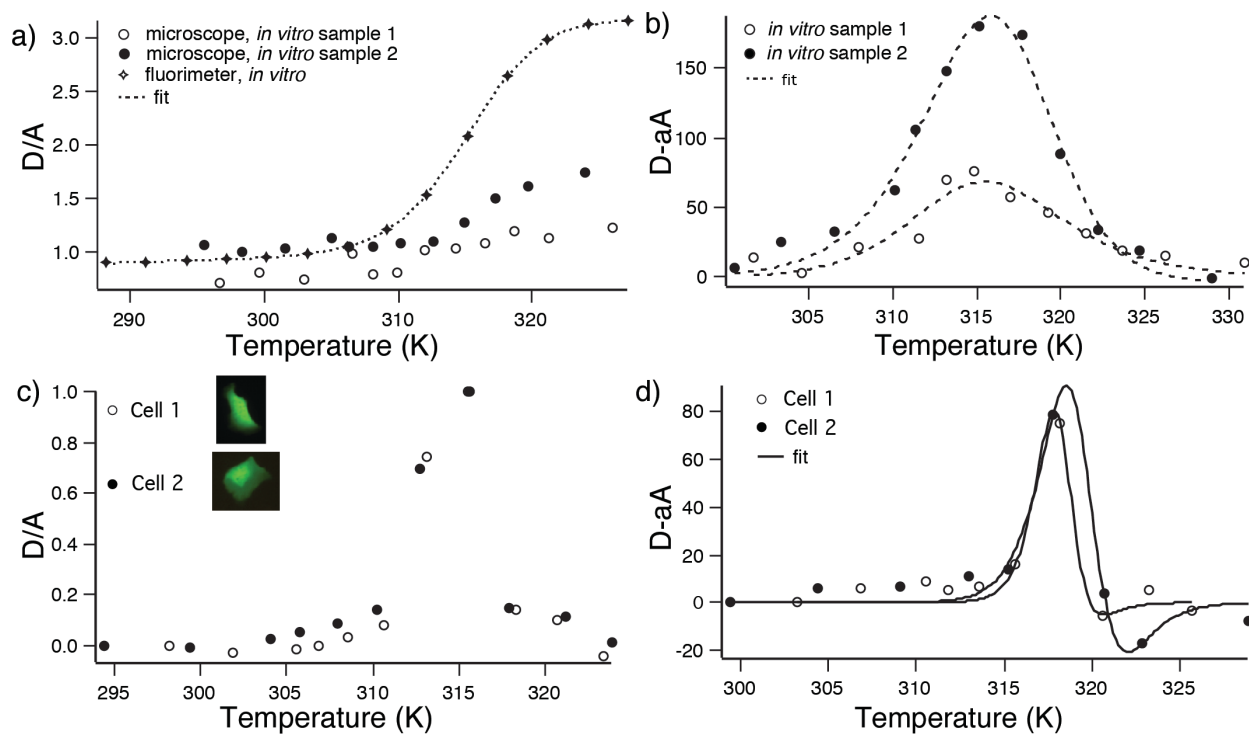


Figure 51 *In vitro* and in cell PGK-tc folding thermodynamics as measured via changes in FRET. a) Donor to acceptor ratios from thermodynamic titrations in the fluorimeter and on the live-cell microscope of ReAsH labeled PGK-tc *in vitro*. The average intensity measurement did not produce fittable thermodynamics on the live-cell microscope while on the fluorimeter, where bleaching is reduced, the observed thermodynamics fit to a melting temperature of 315 K. b) Temperature dependence of the signal at the end of T-jump experiments of *in vitro* PGK-tc on the live-cell microscope with a thermodynamics from kinetics fit. c) Donor to acceptor ratios from thermodynamic titrations on the live cell microscope of two cells expressing ReAsH labeled PGK-tc. Although a sigmoidal transition is visible, the lack of an unfolded baseline at high temperatures makes fitting thermodynamics difficult. d) Temperature dependence of the signal at the end of T-jump experiments of two cells expressing ReAsH labeled PGK-tc with thermodynamics from kinetics fit.

To circumvent this problem, temperature-jump probed kinetics, where the bleaching baseline is clearly distinguishable from the kinetic signal, were used to extract both folding kinetics and thermodynamics. Robust temperature-dependent kinetics were observed for PGK-tc *in vitro* (figure 74, in appendix C). Kinetics were bi-phasic with a fast (~ 2 -10 ms) phase, which is discussed in the subsequent section, and a second slower phase ($\tau_{\text{obs}} \sim 1$ -3 s), the latter of which is close to the folding time observed for PGK labeled with acGFP and mCherry. GFP-tc also shows a small, negative amplitude phase of comparable time-scale to the slow PGK-tc phase (figure 74 in appendix C) that is significantly lower in amplitude to PGK-tc. This may indicate that a small part of the observed PGK-tc folding phase arises from a bleaching or other spectral artifact. However, despite this minor artifact, plotting the temperature-dependent end of jump signal (figure 51b, table 7) to extract the thermodynamics with a thermodynamics from kinetics fit gave well defined thermodynamics far more reliably than fitting to the average intensity before the jump. The fits

yielded an average melting temperature of 314 ± 1 K, which is a little lower than the fluorimeter measurement but consistent with previous measurements for acGFP and mCherry labeled PGK^{79,153,165,168}.

As with *in vitro* PGK-tc, the thermodynamics of PGK-tc in living cells expressed as the average D/A ratio before the kinetic temperature-jump experiment were not fittable (figure 51c). In this case, a sigmoidal transition was visible, but the lack of an unfolded baseline at high temperatures made accurate fitting impossible. The same thermodynamics from kinetics approach that was applied to *in vitro* samples was thus used for PGK-tc expressed in living cells and labeled *in vivo* with ReAsH. The observed folding kinetics were bi-phasic, as *in vitro*. The thermodynamics from kinetics fits (figure 51d, table 7) yielded an average melting temperature of 316 ± 0.3 K which, as expected from previous studies of acGFP and mCherry labeled PGK, is higher than the *in vitro* value.

From the kinetics measurements, the folding rate for the slow phase was also extracted for ReAsH labeled PGK-tc both *in vitro* and in cell, although the in cell measurements were significantly noisier than *in vitro* due to low ReAsH intensity. The observed folding time for PGK-tc *in vitro* (figure 52a, table 7) at approximately the average melting temperature was 1.2 ± 0.1 s, which is slightly faster than what was observed previously for PGK labeled with acGFP and mCherry^{79,153,165}. The in cell folding was not significantly faster or slower than *in vitro* (within error), although the temperature dependence of the rates was not as reproducible as observed *in vitro* (figure 52b, table 7).

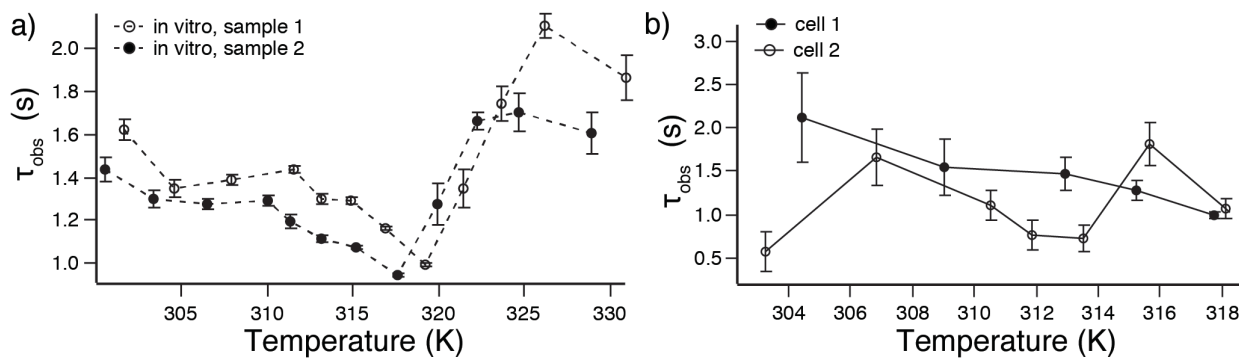


Figure 52 Temperature dependence of the observed folding time for PGK-tc *in vitro* (a) and in cells (b). Error bars shown are fit uncertainties.

	<i>In vitro</i>	in cell
T_m (K)	314 ± 1	316.5 ± 0.6
τ_{obs} (s)	1.2 ± 0.1	1.5 ± 0.3

Table 7 FRET-probed PGK-tc thermodynamics and kinetics as measured on the live-cell microscope. Values are averages of 2 measurements. Reported error is two standard deviation of the mean. Observed folding time is that observed at the average T_m .

The fast phase, which had the opposite amplitude of the slow-phase, showed a rate for both PGK-tc and GFP-tc that ranged from about .1 to .4 s—substantially faster than the expected folding rate for PGK-tc. Figure 53 shows the temperature dependence of this rate for both PGK-tc and GFP-tc, the latter of which should show no folding kinetics under the temperature ranges studied here. The fast phase rate and trend with temperature is nearly identical, within error, between PGK-tc and GFP-tc, suggesting that this phase does not correspond to a structural transition. Rather, it is more likely that it may arise from differential instantaneous temperature responses of the GFP and ReAsH fluorophores—such phases have been observed for GFP and mCherry (see chapter 5, methods).

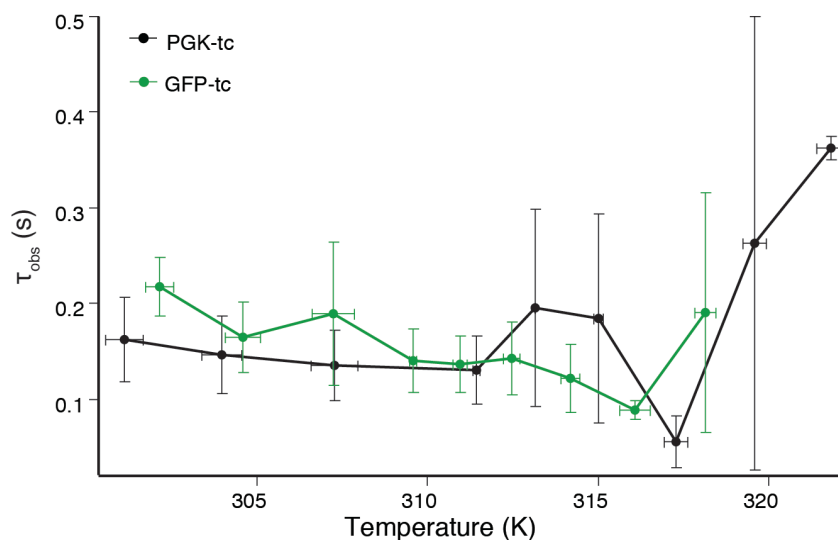


Figure 53 Temperature dependence of the observed fast phase folding time for PGK-tc and GFP-tc *in vitro*. For each trace, 2 (for PGK-tc) or 3 (for GFP-tc) individual experiments of the temperature dependence of the thermodynamics were averaged together. The uncertainties in both temperature and observed folding rate are the standard error from this average.

Interpretation of the kinetic and thermodynamic data for the FRET-probed transition remains challenging due to the bleaching of ReAsH, particularly in the case for cells where at higher temperatures the intensity of the ReAsH is extremely low due to a combination of temperature-dependent intensity loss and photo-bleaching. The small kinetic phases observed in the GFP-tc construct also suggest that bleaching artifacts may be convoluted with the kinetic signal of PGK-tc.

As will be discussed in section 6.3.4, the sensitivity of the ReAsH dye alone to the overall fold of PGK-tc likely also provides an additional convoluting factor into the FRET-probed kinetic and thermodynamic signal. Ignoring these confounding factors, however, PGK-tc shows similar thermodynamic characteristics both *in vitro* and in cells to the GFP/mCherry labeled PGK, providing evidence that the FRET-probed transition is indeed capturing folding behavior and not bleaching artifacts. Kinetically, folding of the PGK-tc construct is faster than the GFP/mCherry labeled PGK, which is not surprising given the small size of ReAsH compared to mCherry.

6.3.4 Direct excitation of ReAsH reports on PGK-tc folding ReAsH can also report on the folding of PGK-tc through an alternative, non-FRET based mechanism. Under direct excitation ReAsH fluorescence emission changes as a function of temperature, showing a trend that cannot be explained by simple temperature dependence of the fluorophore.

Figure 54a shows the thermodynamics as measured by the average red intensity on the fluorimeter. Steep folded and unfolded baselines are observed with a flattened region around PGK-tc's melting temperature, which fits to a two-state model with a melting temperature of 314 K. Melts of GFP-tc under amber excitation, while exhibiting a decreasing intensity, do not show the same thermodynamic features as PGK-tc (figure 54a). The average intensity before the jump at each temperature of *in vitro* PGK-tc on the live cell microscope (figure 54b), shows a similar shape to the fluorimeter measurements. Although GFP-tc and the tc-peptide red intensities do not show the same nearly linear temperature dependence on the live cell microscope as observed in the fluorimeter, their overall intensity loss is significantly less than PGK-tc.

Kinetic measurements on the live-cell microscope under amber excitation were used to extract the thermodynamic fitting parameters (figure 54d), as in the previous section for the FRET-probed structural transition, and the kinetics of the folding transition. Temperature-jump experiments yielded negative-phase (or, an increase in intensity) kinetics, which were fit to a single exponential. A small, fast positive phase was also observed in the kinetics (figure 75) that was too small and fast to fit reliably over all temperatures. The melting temperature from the thermodynamics from kinetics fit was 313 ± 2 K, slightly lower than the fluorimeter measurement and the FRET-probed thermodynamics. The temperature dependence of the observed folding rate was relatively flat (figure 54c), consistent with measurements for the GFP/mCherry labeled PGK¹⁶⁸, although the observed folding rate of $1.3 \pm .1$ s was somewhat faster than expected for PGK⁷⁹.

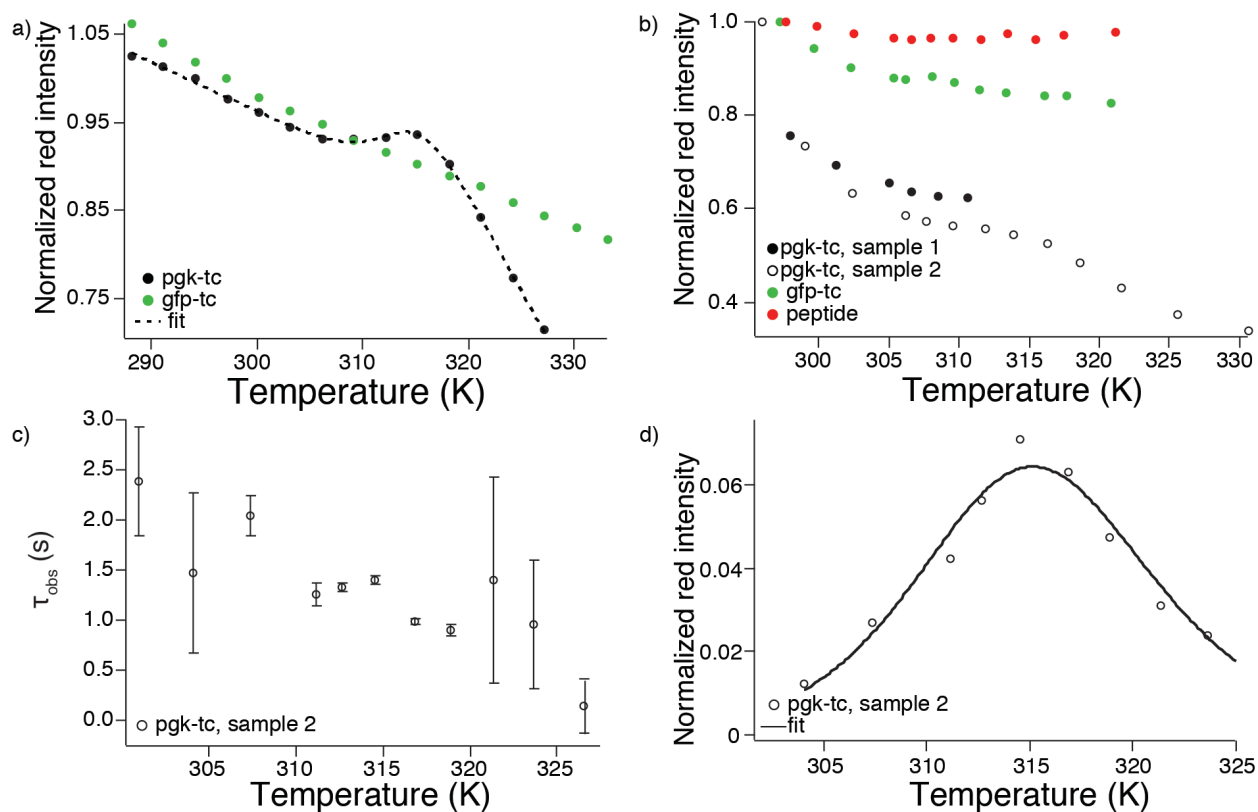


Figure 54 Thermodynamics and kinetics of PGK-tc unfolding monitored on the fluorimeter and live-cell microscope by direct excitation of the ReAsH label. a) Average red intensity at each temperature on the fluorimeter. Flat region from ~310 K to 315 K is the folding transition which fits to a melting temperature of 314K. GFP-tc shows no evidence of unfolding over the measured temperature range. b) Average, normalized red intensity before the temperature-jump experiment (analogous measurement to the fluorimeter experiments) on the live-cell microscope. Slight curvature of the folded and unfolded baselines makes fitting thermodynamics challenging. GFP-tc and the tc-peptide show significantly less intensity change over the melt than PGK-tc. c) Fitted folding times for PGK-tc on the live-cell microscope. d) Average, normalized, red intensity after the temperature-jump experiment with thermodynamics from kinetics fit. In this case, the intensity is normalized by the intensity before the jump.

Because the folding kinetics are relatively low amplitude, GFP-tc was used as a control to ensure that the observed kinetics were not an artifact. Figure 75 shows T-jump experiments of the two proteins at room temperature, where no kinetics should be observed, and 40 °C, where the maximum amplitude kinetics occur for PGK-tc. GFP-tc shows no sign of any kinetics (on the fast or slow time-scale), suggesting that the observed kinetics, both the fast and slow phase, for PGK-tc reflect PGK's unfolding structural transition.

The sensitivity of the directly excited ReAsH dye to the overall fold of PGK is not wholly unexpected, considering previous work that used an internally placed FAsH dye to monitor the folding of a protein¹⁵⁴. In that work, FAsH gained fluorescence through a mechanism that was not completely understood when the protein unfolded, and the dye appeared to report on both local and global unfolding. A comparable sensitivity of FAsH to the protein fold was also noted for a repeat

protein²⁰⁴. It is conceivable that FAsH or ReAsH fluorescence is affected by loss of proximity to quenching amino acid residues or by changes in the stiffness of the protein backbone when proteins unfold, although this effect is likely to be dependent on the protein studied and the placement of the tetra-cysteine recognition sequence. In the case of PGK-tc, it is not clear exactly which kind of structural transition the ReAsH probe is reporting on or through which mechanism the ReAsH fluorescence changes upon unfolding. The T_m is somewhat lower than the FRET-probed transition.

6.4 CONCLUSIONS

Although ReAsH's small size and cell-permeability make it an appealing fluorophore for intracellular folding measurements and other quantitative microscopy applications, the work presented here emphasizes that care must be taken to ensure that the ReAsH's poor photostability does not produce misleading artifacts in the data. Of particular challenge are FRET-based measurements where the change in energy transfer efficiency is correlated to the change in distance between the donor and acceptor. When ReAsH is the acceptor, the cross talk between the donor and acceptor as the ReAsH bleaches can lead to misleading signals that mimic the expected signal for increased distance between fluorophores. Within the cell, where ReAsH has its greatest utility, the situation becomes even more challenging as there is a much smaller reservoir of ReAsH labeled proteins to replace those that are inactivated by photobleaching. However, in this chapter, we used three constructs—PGK-tc, GFP-tc, and the tc-peptide—to characterize the GFP/ReAsH FRET pair and found that, with proper control constructs and data analysis, a GFP/ReAsH construct can reliably report on protein folding thermodynamics and kinetics both *in vitro* and in cells.

We also found that direct excitation of ReAsH yields information on protein structure through changes in emitted red fluorescence intensity as the surrounding protein environment fluctuates in its flexibility. This avenue of direct excitation is particularly appealing because it can evade some of the potential pitfalls of phototoxicity observed in the FRET-based system. In order to develop directly excited tetra-arsenic dyes as a method to detect protein folding two future experiments are proposed:

(1) *Development and characterization of FAsH labeled PGK* While PGK-tc labeled with ReAsH showed promise *in vitro*, in cell measurements were very challenging due to the poor photophysics of ReAsH. Switching to FAsH will likely ameliorate many of the intensity issues and enable measurement of PGK folding in cells with just a single, small reporter tag.

(2) Modulation of the placement of the tetracysteine tag Placement of the ReAsH at an alternate termini or internally may probe local structural transitions of PGK. The ability to probe how the intracellular environment affects different regions of a protein of interest, both thermodynamically and kinetically, is unprecedented and would be a valuable addition to the in-cell fluorescence tool-box.

With the possibility of use as both a FRET-based reporter or a single, directly excited fluorophore FAsH and ReAsH have great utility as quantitative fluorescence reporters in living cells. Although photo-bleaching must be carefully considered, this chapter presents a general strategy to avoid introduction of bleaching artifacts into quantitative data and gives specific insight in how GFP/ReAsH can be used to effectively quantify protein structural changes in living cells. As advances in fluorescence microscopy bring even more attention to studying the details of protein-protein interactions and protein folding in the intracellular environment, FAsH and ReAsH are excellent candidates for quantitative measurements.

CHAPTER 7

DEVELOPING CONSTRUCTS TO MEASURE p53 FOLDING INSIDE LIVING CELLS

7.1 INTRODUCTION

7.1.1 p53, “Guardian of the Genome” p53 is a tetrameric transcription factor essential for regulation of the anti-tumor response of cells²⁰⁵. Underscoring its importance, p53 has two paralogs in vertebrates—p64 and p73—and p53-like proteins are found across many invertebrates²⁰⁶. Often called the “Guardian of the Genome” for its role as a tumor suppressor, p53 is mutated in 50% of cancerous tumors and is the most frequent genetic anomaly associated with cancer²⁰⁵.

Figure 55 shows a very simplified schematic of p53 function and regulation. In response to stress such as hypoxia or signs of DNA damage, p53 is activated by a complex pathway of regulatory enzymes through phosphorylation or other post-translational modifications of its largely unstructured C-terminal domain and through regulatory protein binding to the N-terminal domain. Upon activation, p53 translocates into the nucleus where it acts as a transcription factor to upregulate the genes for apoptosis, cell cycle arrest, anti-angiogenesis, DNA repair, and others related to tumor-suppression. It also upregulates MDM4/MDM2, its primary negative regulators, constituting a feedback loop that eventually leads to reduction in its activity. Although the core function of p53 is understood, its regulation and ability to facilitate a wide and extremely sophisticated range of tumor-suppressor activities, depending on the exact stress and cellular conditions, remains an area of active research²⁰⁷.

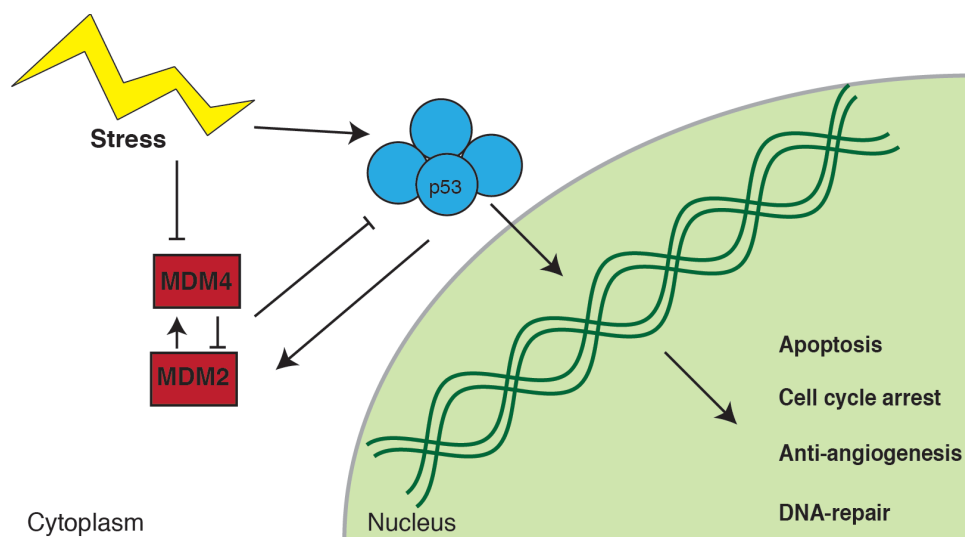


Figure 55 Basic biology of the p53 stress response showing p53 regulation and target gene categories.

p53 also has a role outside of the tumor response that is not necessarily that of a transcription factor. The spindle-check point, or the point at mitosis where cells decide whether proper chromosome segregation has occurred before cell-division, is p53 dependent, which suggests that p53 plays a role in ensuring that DNA is faithfully divided between two daughter cells²⁰⁸. In more recent work, p53 has been shown to localize near the Chromosome Passenger Complex (CPC), a protein complex that corrects faulty chromosome-mitotic spindle interactions²⁰⁹. Together, this evidence suggests that p53 may exhibit non-transcription factor activity during cell division.

7.1.2 p53 intracellular localization The mechanisms controlling the localization of p53 are complex but are thought to rely on a combination of post-translational modification and p53's nuclear localization and export signals, which lie in the C-terminal domain²¹⁰. For example, in one localization mechanism, p53 is ubiquitinated by its negative regulator, MDM2, which leads to unmasking of the nuclear export signal and active transport out of the nucleus²¹¹.

Generally speaking, p53's localization closely mimics its function. In cells with low stress, p53 is located relatively evenly between the cytoplasm and nucleus²¹². However, in response to stress signals, p53 localization skews towards the nucleus where it can act as a transcription factor. In some cell-lines, such as fibroblasts, p53 localization is strictly regulated by the cell cycle, and p53 is located in the nucleus (or nuclear region) from S phase, or DNA-synthesis, through mitosis²¹³ where it serves a non-transcription-factor role²⁰⁹.

7.1.3 Structural biology of p53 p53's structure can be roughly divided into 3 domains: an N-terminal "transactivation domain", the core DNA-binding domain, and the C-terminal domain. The transactivation domain is largely unstructured, gaining helical structure only when bound to regulatory proteins²⁰⁵, and is largely responsible for p53's response to regulatory signals. The C-terminal domain is also mostly unstructured and contains nuclear export and import signals as well as the tetramerization domain.

The core of p53 is folded and consists of residues 94 to 312. It was first crystallized in 1994²¹⁴ and its DNA binding is mediated by its immunoglobulin-like beta-sandwich fold, stabilized by a coordinated zinc ion²⁰⁵. The vast majority of cancer mutations occur in this domain either leading to thermodynamic destabilization of the fold or disruption of the DNA-binding surface²⁰⁵.

The first observation that fold stability underlies some cancer-causing p53 mutations was the discovery of the “temperature-sensitive” p53 mutants, which regain their function and native localization only when cells are incubated at lower temperatures²¹⁵⁻²¹⁷. Since then, the folding thermodynamics of the wild-type p53 as well as many thermodynamically destabilized cancer causing mutations have been assayed. Generally speaking, p53 unfolds irreversibly upon temperature denaturation but reversibly with urea denaturation²¹⁸, although the temperature-probed thermodynamics closely match those measured via urea²¹⁹. As might be expected, the temperature-dependent aggregation of p53 is more severe for many of the highly destabilized mutants²¹⁹. Interestingly, the folding thermodynamics of the core domain consisting of residues 94 to 312 are nearly identical to that of the full length protein²¹⁸, making the core a useful probe for p53’s folding trends as whole.

A few common cancer-causing mutations and an engineered super-stable p53 variant are categorized in table 8 with their thermodynamic melting temperatures. Generally speaking, mutations in the DNA binding domain—such as R249S, G245S, and R273H —have a relatively low impact on the overall stability of p53 and inhibit p53 activity by reducing the protein’s ability to bind DNA as a transcription factor. Those mutations that occur in the beta sandwich region (I195T and V143A), however, are generally highly destabilizing and abolish p53 activity by reducing the amount of folded protein available in the cell at physiological temperature.

Mutation	Type of mutant	T _m (°C)
p53C ^{218,219}	Wild type core domain	44
M133L/V203A/N239Y/N268D ^{220,221}	Engineered for stability	49
I195T ²¹⁹	Beta sandwich, highly destabilizing	~38*
V143A ^{219,220}	Beta Sandwich, highly destabilizing	~36*
R249S ^{219,220}	DNA binding domain, mildly destabilizing	40*
G245S ²²⁰	DNA binding domain, mildly destabilizing	42*
R273H ²²⁰	DNA contact mutation, no destabilization	44*

Table 8 Thermodynamic stability of various p53 mutants. p53C is the wild type core domain. All mutants listed are core-domain mutants. In some cases (indicated with a *), melting temperature was extrapolated based on the measured free energy difference between p53C WT and the mutant. Citations to thermodynamic data are in the “mutation” column.

7.1.4 p53 is an excellent candidate to test the relationship between fold stability and function

Considering previous work that has shown that protein folding stability can vary across a cellular process such as the cell-cycle²²², it is interesting to ask the question of whether protein fold stability

may correlate with functional activity. For example, if a protein is highly active in one phase of the cell cycle but inhibited in another, could its thermodynamic stability echo that activity? Such correlation of thermodynamic stability with function could constitute another level of protein regulation where the copy number of viable protein is not just controlled by transcriptional and post translational regulation but also by the number of folded, and thus potentially active, protein molecules present within the cell.

p53 has a number of characteristics that make it a promising candidate for testing this hypothesis. First, p53 is known to be thermodynamically unstable and its melting temperature across several organisms tracks with physiological temperature²⁰⁶. This suggests that p53 may have evolved to have a certain degree of thermodynamic stability that is essential to its function and regulation. With a melting temperature close to physiological temperature, p53 is poised at the point where slight perturbations in the cellular environment could have a significant impact on the amount of folded protein in the cell. Second, p53's activity is coupled to the cell-cycle—a process controlled in time—and to its subcellular localization. Based on its expected biological activity, table 9 shows a set of hypotheses for how thermodynamic stability would vary across time and cellular location if fold stability were positively correlated to activity level.

	Nucleus	Cytoplasm
Mitosis	n/a	More stable
Interphase	More stable	Less stable

Table 9 Hypothesis for thermodynamic stability by location and cell cycle phase correlating with p53's intracellular activity level. In this hypothesis, p53 would be more thermodynamically stable at the times and intracellular locations where it is biologically active.

Although there have been a few efforts to measure p53 thermodynamic stability in cells using the amount of expressed protein as a proxy for thermodynamic stability^{223,224}, there have been no direct measurements of folding thermodynamics of the protein in living cells. This chapter summarizes efforts to develop p53 FRET constructs that can be used to measure p53 thermodynamic stability in living cells, ultimately with the goal of testing the hypothesis shown in table 9. Two constructs that track folding via FRET are explored: one using ReAsH as the acceptor and GFP as the donor and another using mCherry as the acceptor and GFP as the donor. We find that the best fluorophore pair to pursue in future experiments is GFP/mCherry, which avoids excessive aggregation of the protein in living cells.

7.2 METHODS

7.2.1 Construct design p53 constructs were designed with a variety of fluorophores and expression tags and are summarized in table 10. The core domain of human p53 (not including residues 305-312, which overlap with the nuclear localization signal), residues 94-304, was sandwiched between acGFP1 on the N-terminus and either the tetracysteine recognition sequence¹⁹⁹ or mCherry on the C-terminus. The core domain was chosen because its folding thermodynamics are nearly identical to the folding thermodynamics of the full length protein²¹⁸, and removal of the nuclear localization domains, regulatory protein binding domains, and tetramerization domain ensures that the construct will probe only folding in living cells rather than higher order processes. GFP was chosen as the N-terminal fluorescent protein because GFP-p53 fusion constructs retain their function and native localization²²⁵. A construct with a Glutathione S Transferase (GST) protein before the GFP to aid in purification was also synthesized. All constructs were cloned into pDream (GenScript) for dual expression in bacterial and mammalian cells.

Construct	Fluorescent label	Purpose
GFP-p53-tc	GFP/ReAsH	Wild-type p53
GFP-p53*-tc	GFP/ReAsH	Destabilized p53 V143A mutant
GST-GFP-p53-tc	GFP/ReAsH	GST tag for improved expression
GFP-p53-mCherry	GFP/mCherry	Wild-type p53, ReAsH alternative

Table 10 Table of p53 constructs showing their nomenclature, FRET pairs, and experimental purpose.

In the construct design, we considered whether the termini of the p53 core domain were close enough that energy transfer could occur between acGFP1 and ReAsH (figure 56). The end to end distance between the N and C termini of p53 was estimated by the crystal structure of the p53 core domain²²⁶. The distance between residues 96 and 291 is 3.3 nm, leaving 15 residues in our construct that did not crystallize in the structure. Assuming these residues are entirely random coil, the end to end distance between fluorophores of the folded state would be about 6.5 nm. We calculated the Forster distance for acGFP1 and ReAsH as ~6-7 nm, so in the worst case scenario—assuming random coil structure in the uncrystallized amino acids—the folded state should show 50% FRET efficiency.

With these constructs, we use the donor to acceptor ratio as a proxy for end-to-end distance between the termini of the protein. Unfolded proteins have a high donor to acceptor ratio (D/A), as the tags are further apart, while folded proteins have a lower D/A reflecting the proximity of the two termini. Aggregated proteins are expected to have an even lower D/A as intermolecular FRET

between aggregated protein molecules bring fluorophores closer together than the intramolecular interaction in the folded state.

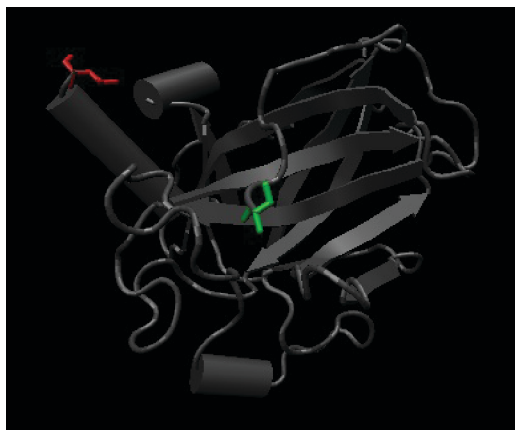


Figure 56 p53 structure showing dye placement. Red indicates C-terminus and green indicates N-terminus. Distance between termini in crystal structure is 3.3 nm. Crystal structure is from ref²²⁶ and figure was prepared using VMD.

7.2.2 Protein purification We attempted to express and purify p53 (see section 7.3), using the protocol of Mayer et al²²³ with slight pH and purification strategy modifications due to the altered isoelectric point of the fusion construct compared to the core domain of p53. Briefly, cells transformed with GFP-p53-tc were grown in BL21 DE3 RIPL cells to an OD₆₀₀ of .6 in 2xTY before induction with 1M IPTG in the presence of 100 μ M zinc sulfate (a folding co-factor). Cells were spun down and lysed via sonication in a buffer of 50 mM Tris HCl (pH 7.64) and 5 mM DTT. Purification was carried out via FPLC over an anion exchange column (HiTrap Q FF, GE), and the protein eluted via a gradient of NaCl. Further purification was attempted via affinity chromatography using a heparin column (GE).

We also attempted to express GST-GFP-p53-tc, although also unsuccessfully. Cells were grown in the same way as GFP-p53-tc, spun down, and lysed via sonication in a buffer of 10 mM sodium phosphate pH 7.64 and 5 mM DTT. Purification was carried out via FPLC over a reduced glutathione column (GSTrap, GE) and the protein eluted by a single step gradient of reduced glutathione.

7.2.3 Live cell imaging and data analysis Live cell imaging thermodynamics experiments were carried out by slow resistive heating, as described in Appendix E. Briefly, U2OS cells were transfected with GFP-p53-tc, GFP-p53*-tc, or GFP-p53-mCherry DNA at a concentration of 2 μ g

/10 mL culture. Incubation times were 24 hours for the tetracysteine constructs and 16-26 hours for the GFP/mCherry construct. For the tetra-cysteine constructs, cells were stained with ReAsH dye as described in Chapter 6.2. Slides were prepared as described in Appendix E, and cells were imaged under blue excitation with green and red emission collected simultaneously.

At each temperature, a 2 second video at 24 fps was taken. The green and red intensity vs. time was corrected for bleaching by subtraction of the linear baseline, and the average intensity used to calculate the ratio of donor (green fluorescence) to acceptor (red fluorescence) at each temperature. The thermodynamics were fit to a two-state thermodynamic model as described in Appendix E.

7.3 RESULTS AND DISCUSSION

7.3.1 GFP-p53-tc does not fold in bacteria Despite many different expression strategies attempted, we were unable to obtain folded GFP-p53-tc from bacteria. GFP-p53-tc was first expressed using the protocols of Mayer et al²²³, which yielded expressed protein (figure 57) that purified with two cleavage products. The small amount of full length protein obtained was not folded as evidenced by the lack of GFP absorbance/fluorescence, which is known to decrease in fusion proteins when the protein of interest is unfolded²²⁷. Despite varying induction conditions including the concentration of IPTG, the induction temperature, the concentration of ZnSO₄, and induction times, we obtained no folded full-length protein. Multiple *E. coli* cell lines including BL21 DE3 PlyS (Agilent), BL21 DE3 RIPL (Agilent), and C41 (Lucigern) were also tested with no obvious improvement. A GST-tagged construct also did not produce folded protein.

Considering that GFP-p53 expression has been reported in the literature²²³, it is likely that the addition of the tetra-cysteine tag caused aggregation. The core domain of p53 has 10 cysteines, so it is possible that the tetracysteine tag interfered with folding by making non-productive disulfide bonds with p53. The core domain of p53 is also a relatively unstable protein with a melting temperature of around 44 °C. We have observed that the addition of large fluorophore tags can often destabilize the overall protein fold and lead to a few degrees lower melting temperature (unpublished data). It is possible that the addition of the GFP tag made the construct too destabilized to fold effectively in bacteria, even under reduced temperature induction conditions.

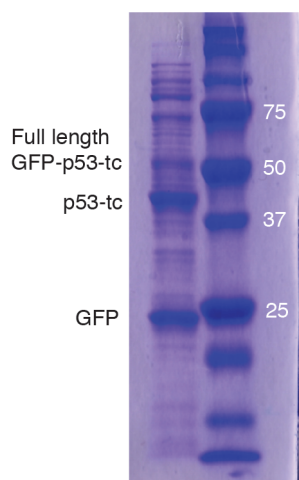


Figure 57 SDS-page gel showing a fraction from the anion exchange column containing recombinantly expressed GFP-p53-tc. In addition to the full length protein, two cleavage products are evident: GFP and p53-tc. Markers in rightmost lane are molecular weight standards with labels in kDa. Gel is 10% acrylamide run under reducing conditions and stained with Coomassie Brilliant Blue.

7.3.2 Optimization of GFP-p53-tc expression in mammalian cells Despite the challenges of expressing GFP-p53-tc in bacteria, the construct expressed in mammalian U2OS cells. A variety of DNA transfection amounts and overall DNA-incubation times were tested to find the best expression conditions (figure 58). With too much DNA, large inclusion bodies formed in the nuclei of the cell suggesting that when protein concentration in the cell is very high, aggregation occurs. With too short of a DNA incubation time, fluorescence was generally dim with smaller inclusion bodies of aggregated protein spread throughout the cytoplasm. After 24 hours incubation, uniform fluorescence was observed for most cells, although it was noted that ~10% of cells showed obvious signs of aggregation (which is higher than what is typically observed for other proteins such as PGK-tc). GFP-p53*-tc also expressed in mammalian cells, although with significantly more aggregation present—approximately 50% of cells showed signs of aggregation under the least aggregation-promoting conditions tested.

It is not clear why shorter incubation times with DNA yielded poor fluorescence and inclusion bodies. As a comparison, GFP-p53-mCherry showed uniform cell fluorescence at just 16 hours of incubation with DNA, suggesting that the issues with GFP-p53-tc aggregation arise from the tetracycline tag rather than from intrinsic properties of p53 itself. It is conceivable that upon initial expression GFP-p53-tc aggregates, but the mammalian chaperones are able to rectify most of the aggregation after sufficient incubation times. This could also explain why GFP-p53-tc does not express in bacterial cells, which have fewer folding chaperones than mammalian cells.

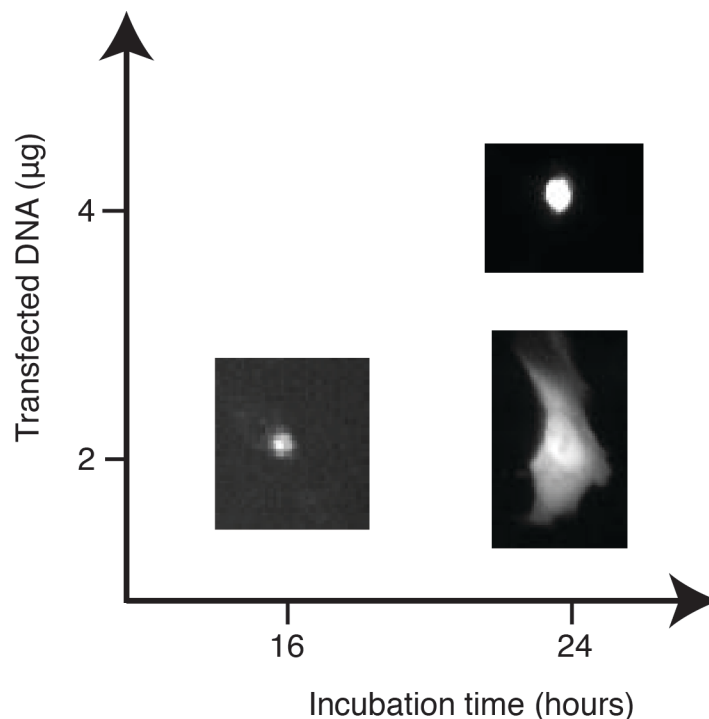


Figure 58 Green-channel images of U2OS cells expressing GFP-p53-tc with varying DNA incubation times and amounts of transfected DNA. At relatively short incubation times, cells are generally dim and show many inclusion bodies. With greater than 2 μg DNA transfected and long incubation times, large and very bright inclusion bodies are evident in many cells. At 24 hours incubation with 2 μg DNA, the most uniformly fluorescent cells are observed.

7.3.3 Folding monitored by ReAsH p53 folding thermodynamics were measured in living U2OS cells for the wild-type, GFP-p53-tc, and the V143A destabilized mutant, GFP-p53*-tc. Of many cells measured, only two wild-type and one mutant showed sufficient intensity and lack of visible signs of aggregation to assay folding thermodynamics (figure 59a). From this limited sample size, the wild type gives an average melting temperature of 39 $^{\circ}\text{C}$, which is 5 degrees lower than *in vitro* measurements of the core-domain alone²¹⁹. By the end of the melt, cells showed visible signs of aggregation, as is expected for p53 which aggregates at high temperatures²¹⁸. However, *in vitro* experiments have shown that even though temperature-probed thermodynamics of p53 are not reversible, their results are comparable to reversible thermodynamic measurements²¹⁹.

The V143A mutant, which *in vitro* melts at around 38 $^{\circ}\text{C}$ ²¹⁹, shows a higher melting temperature than the WT and the *in vitro* construct—an unexpected result. This is likely due to the tendency of GFP-p53*-tc to aggregate in cells. Even at low temperatures, many cells showed visible signs of aggregation and, even the few cells that lacked visible inclusion bodies had extremely low initial D/A ratios, a signature of aggregation. Of the cells tested, only one (figure 59a) showed a

thermodynamic transition, which was far more stable than expected, a result that is consistent with the thermodynamics monitoring disaggregation of inclusion bodies rather than a folding transition.

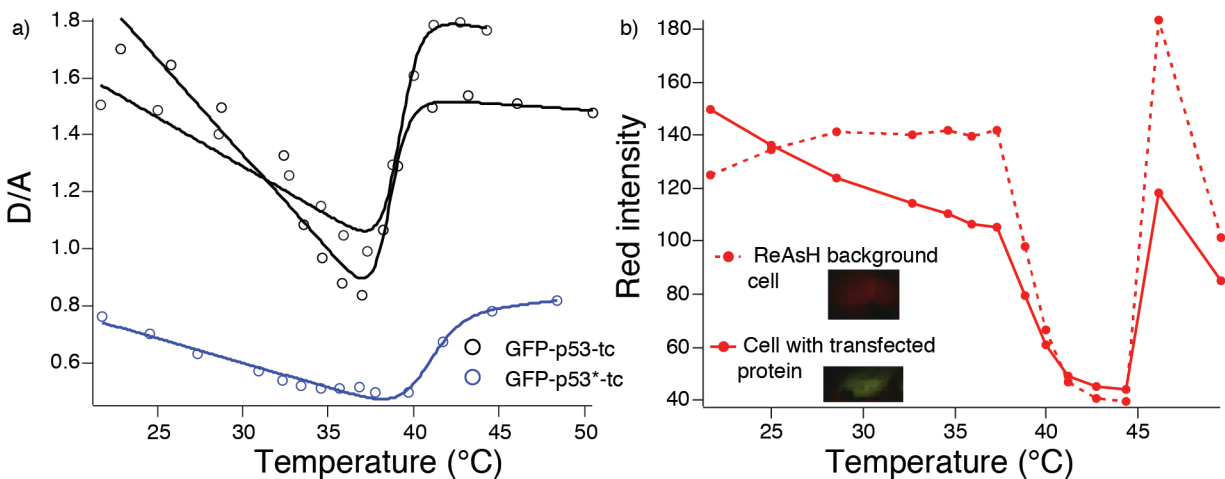


Figure 59 ReAsH probed thermodynamics of GFP-p53-tc and GFP-p53*-tc in living cells a) Temperature dependence of the D/A ratio for GFP-p53-tc (black) and GFP-p53*-tc (blue) with two-state thermodynamic fits. D/A is proportional to the FRET ratio: high D/A means the fluorophores are far apart and low D/A means the fluorophores are close together. GFP-p53*-tc's high melting temperature and low initial D/A ratio suggest that aggregation may be occurring. b) Temperature dependence of red fluorescence intensity of a cell expressing GFP-p53-tc and an adjacent cell showing only ReAsH background. The cell with only ReAsH background shows a similar temperature dependence to the cell with p53 suggesting that cells showing background may be expressing p53 with a non-fluorescent GFP fluorophore.

Considering the visible signs of aggregation of the V143A mutant and even the WT p53, it appears that the tetracysteine tag may be interfering with the measurement of folding thermodynamics. Compounding this problem is the behavior of the background fluorescence of cells. Figure 59b displays the red intensity of a cell that shows both red and green fluorescence under blue excitation—i.e. one that has been transfected with GFP-p53-tc and has had the GFP fluorophore form—and another that shows only red fluorescence. The red fluorescence only cell should show only a slight monotonic increase in intensity, which is the behavior of the ReAsH background under blue excitation (figure 50). However, it shows a nearly identical profile to the folding thermodynamics with a sharp drop in intensity at about the melting temperature of the GFP-p53-tc in cells. It is unclear exactly what causes this unusual background fluorescence, but it is possible that only a minority of cells are able to express a fully folded p53 with a fluorescent GFP and that many cells contain GFP-p53-tc in a partially folded, partially aggregated, or unfolded state. This problematic result lead us to see if an alternative, non-tetracysteine construct may provide more easily interpreted fluorescence results.

7.3.4 Folding of GFP-p53-mCherry in mammalian cells With the issues of aggregation noted for GFP-p53-tc, a non-tetracysteine construct was also tested where the tetracysteine binding site was replaced with mCherry. The construct expressed more readily in mammalian cells with fewer visible signs of aggregation than the GFP-p53-tc. After 16-24 hours incubation with DNA, cells showed uniform fluorescence throughout the cytoplasm—comparable to other proteins whose folding has been assayed in living cells like GFP/mCherry labeled PGK²²². However, by the end of the thermodynamic titrations all cells showed visible signs of aggregation, consistent with the known unfolding behavior of p53 *in vitro*.

GFP-p53-mCherry did not yield reproducible thermodynamics that could be fit to an unfolding model. The temperature dependence of the donor to acceptor ratio for the 19 cells that were assayed showed four general trends: 1) obvious aggregation at the approximate melting temperature of the construct characterized by a sharp decrease in the D/A ratio, 2) clear unfolding at the approximate melting temperature of the construct characterized by a sharp increase in the D/A ratio, 3) a combination of unfolding and aggregation, or 4) no trend (essentially a flat donor to acceptor ratio over the temperatures measured). Figure 60a shows the thermodynamics of unfolding for three cells, each a representative example of the first three trends observed in experiments. The vast majority of cells showed only very weak thermodynamics with some signs of aggregation (10 of 14 cells measured). Whether thermodynamics, aggregation, or a combination of the two behaviors was observed in any given cell did not correlate with cell brightness (a rough proxy for the amount of protein expressed in the cell), starting donor to acceptor ratio, the amount of time between DNA transfection and imaging, or the amount of DNA added. Figure 60b summarizes the measured thermodynamics for all assayed cells.

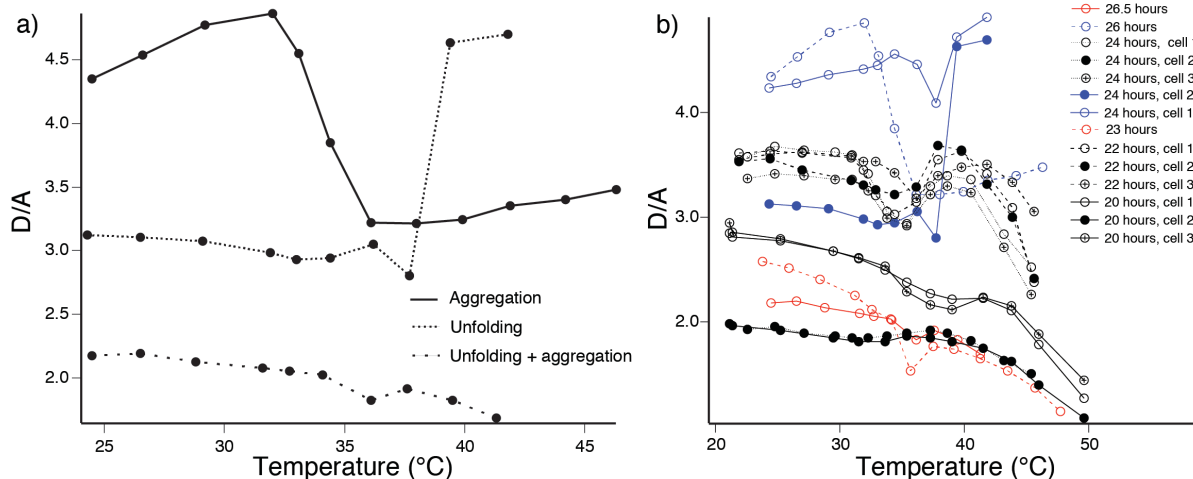


Figure 60 GFP-p53-mCherry in living cells. a) Representative cells for the three types of unfolding thermodynamics of GFP-p53-mCherry observed in living cells. Increasing D/A ratio indicates unfolding while decreasing D/A ratio indicates aggregation. b) Temperature dependence of the donor to acceptor ratio of all measured cells. Traces are colored by the day the data was taken (same color = same day). Line character and marker indicates cells that were from the same coverslip (same line character = same coverslip, different marker = different cells from same coverslip). Legend indicates time between DNA transfection and imaging for each cell.

7.4 CONCLUSIONS AND PROSPECTS

It is clear that the biggest hurdle to measurement of p53 folding in living cells, and comparing those results to comparable *in vitro* experiments, is p53's tendency to aggregate. Although a ReAsH labeled construct is appealing due to its small size compared to GFP/mCherry constructs and shows some promise in its ability to measure folding thermodynamics, the four cysteines within the tag appear to lead to an even stronger tendency to aggregate than p53 already possesses intrinsically—generally speaking, it may be that the tetracysteine label may be incompatible with aggregation prone or cysteine-rich proteins. It is unlikely that a tetracysteine labeled p53 construct can be expressed in bacteria, and without an *in vitro* construct to provide thermodynamic and kinetic behavior baselines, it is difficult to tease out exactly what kinds of structural transitions the in cell experiments are measuring. Furthermore, although aggregation is slightly better in U2OS cells than in bacteria, the tetracysteine labeled construct shows too many visible signs of inclusion body formation even at room temperature to be a reliable indicator in cells. Finally, when the difficult to interpret in cell thermodynamics are paired with ReAsH's tendency to give misleading results due to photobleaching (see Chapter 6), it is difficult to imagine a scenario where in cell results with this construct could be trusted enough to draw strong conclusions.

The GFP/mCherry labeled p53 shows somewhat more promise, at least in terms of aggregation tendencies, than the ReAsH labeled version and the well-characterized FRET pair

makes result interpretation somewhat more straightforward. However, the construct did not yield reproducible or fittable unfolding thermodynamics. The inability of this construct to report thermodynamics likely arises from two factors. First, because wild-type p53 folds on a minutes time-scale²²⁸ at physiological temperatures and even slower at room temperature, the amount of time of equilibration necessary at each temperature to reach thermodynamic equilibrium may exceed what is experimentally feasible. The lack of or very low amplitude thermodynamics of folding observed are consistent with a scenario where the system had not fully equilibrated at each temperature. Second, the p53 core domain has a handful of amino acids at either termini that are unstructured in the crystal structure²²⁶ which increases the separation between the two fluorophores. Although the termini should be close enough together to give 50% FRET efficiency in the folded state even with the unstructured residues (and, the low initial D/A ratios suggest that at least this FRET efficiency is achieved), it is possible that with the addition of the fluorescent protein tags the termini are too far apart to give a robust change in FRET signal upon unfolding. If this were the case, any unfolding thermodynamics would be low amplitude and difficult to measure.

A more successful approach may be to use a faster folding p53C mutant and a truncated version of the core domain that has less separation between the termini. One promising mutant is the p53C*, which was studied in cells with the GFP/ReAsH pair. The p53C* (V143A mutant) destabilized p53 mutant folds faster than the wildtype²²⁸ and, although it heavily aggregated in cells with the tetracysteine label, may perform well with the less aggregation prone GFP/mCherry pair. Truncating the core domain to comprise residues 96 to 289, rather than 94 to 312, vastly decreases the separation between N and C termini and would give a >90% FRET efficiency of the folded state.

In summary, this chapter offers a hypothesis for p53's folding trends in different locations of the cell as well as in different stages of the cell-cycle that mirror its biological functionality. From the characterization of GFP/ReAsH and GFP/mCherry labeled p53 constructs, we demonstrate that the ReAsH tags cause severe aggregation in cell and that the GFP/mCherry fluorophore pair shows much more promise. Furthermore, we offer a specific suggestion of a p53 mutant and core domain truncation that is a good candidate for a p53 construct that reliably reports on its fold inside living cells. With this construct, p53 may yet prove to be a useful protein to study the relationship between intracellular protein folding and biological activity level.

APPENDIXES AND REFERENCES

APPENDIX A SUPPLEMENTARY INFORMATION FOR “COMPARING FAST PRESSURE-JUMP AND TEMPERATURE-JUMP PROTEIN FOLDING EXPERIMENTS AND SIMULATIONS”*

A.1 SUPPLEMENTARY METHODS

A.1.1 Molecular Dynamics Simulations Molecular dynamics simulations were performed in explicit solvent using the TIP3P water model⁵⁴ and the CHARMM22* force field for protein.⁵⁵ The CHARMM22* force field is a modified version of the widely used CHARMM22 force field with CMAP corrections.²²⁹⁻²³¹ It has been employed successfully to simulate folding of a wide range of fast folding proteins with both α -helical and β -sheet native topology.²⁴ The structure of the WW domain (Fip35 mutant, residues 4 to 38) from the protein data bank (PDB code 2F21)³⁹ was placed in a cubic box of 10,232 water molecules and neutralized with 6 sodium ions and 9 chloride ions employing VMD.⁵⁶ The simulated systems, including protein, water molecules, and ions, measured 68.4 Å in each dimension at $T = 325$ K and $P = 0.1$ MPa and contained 31,273 atoms. All simulations were carried out with periodic boundary conditions in a constant particle number, temperature, and pressure ensemble (NPT). Starting from the native state of the protein, two types of simulations were performed: pressure-jump and temperature-jump. In both cases, the simulations started with 100 ns equilibrium simulation of the native state and ended with multi-microseconds refolding simulation, both at $T = 325$ K and $P = 0.1$ MPa. The temperature for ambient conditions, 325 K, was chosen because it is well below Fip35's melting temperature and directly comparable to previous P-jump simulations.^{28,35} The difference lies in the denaturing procedure which followed the 100-ns equilibrium simulation. In a pressure-jump simulation, pressure was increased from 0.1 MPa to 900 MPa in 0.3 μ s at a rate of 0.9 MPa/300 ps, followed by a 1- μ s high-pressure equilibrium simulation ($P = 900$ MPa) and a pressure-drop simulation in which pressure was jumped downward from 900 MPa to 0.1 MPa in 0.3 μ s at a rate of -0.9 MPa/300 ps. The temperature was maintained at $T = 325$ K through the pressure-jump simulation. In a temperature-jump simulation, the pressure was maintained at $P = 0.1$ MPa throughout the simulation and the system was heated up to 400 K for 1 μ s between the initial 100-ns equilibrium simulation and the final refolding simulation, both at $T = 325$ K. To generate multiple refolding trajectories, the pressure or temperature unfolded state was equilibrated for an additional 200 ns respectively, during which the structures were taken at $t = 0$ ns, 100 ns, and 200 ns to continue the P-jump and T-jump simulations at ambient conditions.

A.1.2 Molecular dynamics simulations using NAMD The simulation algorithm and features of the NAMD program are described previously.⁵⁷ The systems to be simulated were first subjected to 3000 steps of conjugate gradient minimization and equilibrated for 300 ps with harmonic restraints

* This appendix is adapted with permission from the supporting information for Anna J. Wirth, Yanxin Liu, Maxim Prigozhin, Klaus Schulten, and Martin Gruebele, *Journal of the American Chemical Society*, 137 (22), 7152-7159, 2015. Copyright 2015 American Chemical Society. YL performed all molecular dynamics simulations referenced in this appendix and generated the data shown in figures 66, 67, 68 and 69. MP provided instrumentation support to this publication.

applied to all the heavy atoms of the protein. The simulation was then continued for 3 ns without restraints at a constant pressure of 0.1 MPa using Nosé -Hoover Langevin piston barostat and at a constant temperature of 325 K with Langevin damping constant of 5.0 ps^{-1} . In the subsequent simulations, constant temperature was maintained using Langevin dynamics with a damping constant of 1.0 ps^{-1} and multiple time stepping employed with an integration time step of 2.0 fs, short-range forces being evaluated every time step and long-range electrostatics evaluated every three time steps. All bonds involving hydrogen in the protein were constrained using RATTLE,²³² while the geometries of water molecules were maintained using SETTLE.²³³

Cutoff for short-range nonbonded interactions was 8.0 Å with shifting beginning at 7.0 Å; long-range electrostatics was calculated using the particle-mesh Ewald method.²³⁴ The cutoff was chosen to accelerate the simulation on general purpose supercomputers while maintaining a sufficient level of accuracy. Indeed, it is accurate enough to fold two fast folding proteins: villin headpiece²³⁵ and lambda repressor.²³⁶

A.1.3 Molecular dynamics simulations on Anton The refolding simulations were carried out on the Anton platform.^{58,59} Multiple time stepping was employed, with an integration time step of 2.0 fs. The Multigrator integration procedure was employed.²³⁷ Short-range forces were evaluated every time step and long-range electrostatics every three time steps. Cutoff for the short-range nonbonded interactions was 14.18 Å, as recommended by the developers of the Anton machine at DE Shaw Research. Long-range electrostatics was calculated using the k-Gaussian Split Ewald method²³⁸ with a $32 \times 32 \times 32$ grid to better suit the Anton machine. All bonds involving hydrogen atoms were constrained using SHAKE.²³⁹

A.2 SUPPLEMENTARY TABLES

Transition	n	Θ (μs)	$\langle k \rangle_U$ ($\mu\text{s})^{-1}$
U \rightarrow F, temperature	2	33.8	0.0895 ± 0.052
U \rightarrow F, pressure	3	33.5	0.119 ± 0.059
U \rightarrow I _{L1} , pressure	2	30.7	0.0978 ± 0.056
U \rightarrow I _{L2} , pressure	1	30.7	0.0652 ± 0.046
I _{L1} \rightarrow F, pressure	2	0.940	3.19 ± 1.8
I _{L2} \rightarrow F, pressure	1	1.88	1.06 ± 0.75

Table 11 Rates from Bayesian analysis.

Model	τ_{fast} (μs)	τ_{slow} (μs)	%A _{fast}	%I _{loop 1}
II: parallel transition states	1	671	2.4×10^{-10}	0
III: loop 2 intermediate	1	671	2.4×10^{-10}	0
IV: loop 1 intermediate	32.7	666	47	94
V: two parallel intermediates	29.3	654	52	95

Table 12 Simulation summary. Rates shown are from the double exponential fits to the signal functions that most closely match experiment for each model. Amplitude shown is percent of the fitted amplitude that corresponds to the fast phase. The percentage of the loop 1 intermediate refers to the composition of the signal function that best matched the experimental data where the percentage of the signal function corresponding to the native state is given by $100 - \%I_{\text{loop 1}}$.

A.3 SUPPLEMENTARY FIGURES

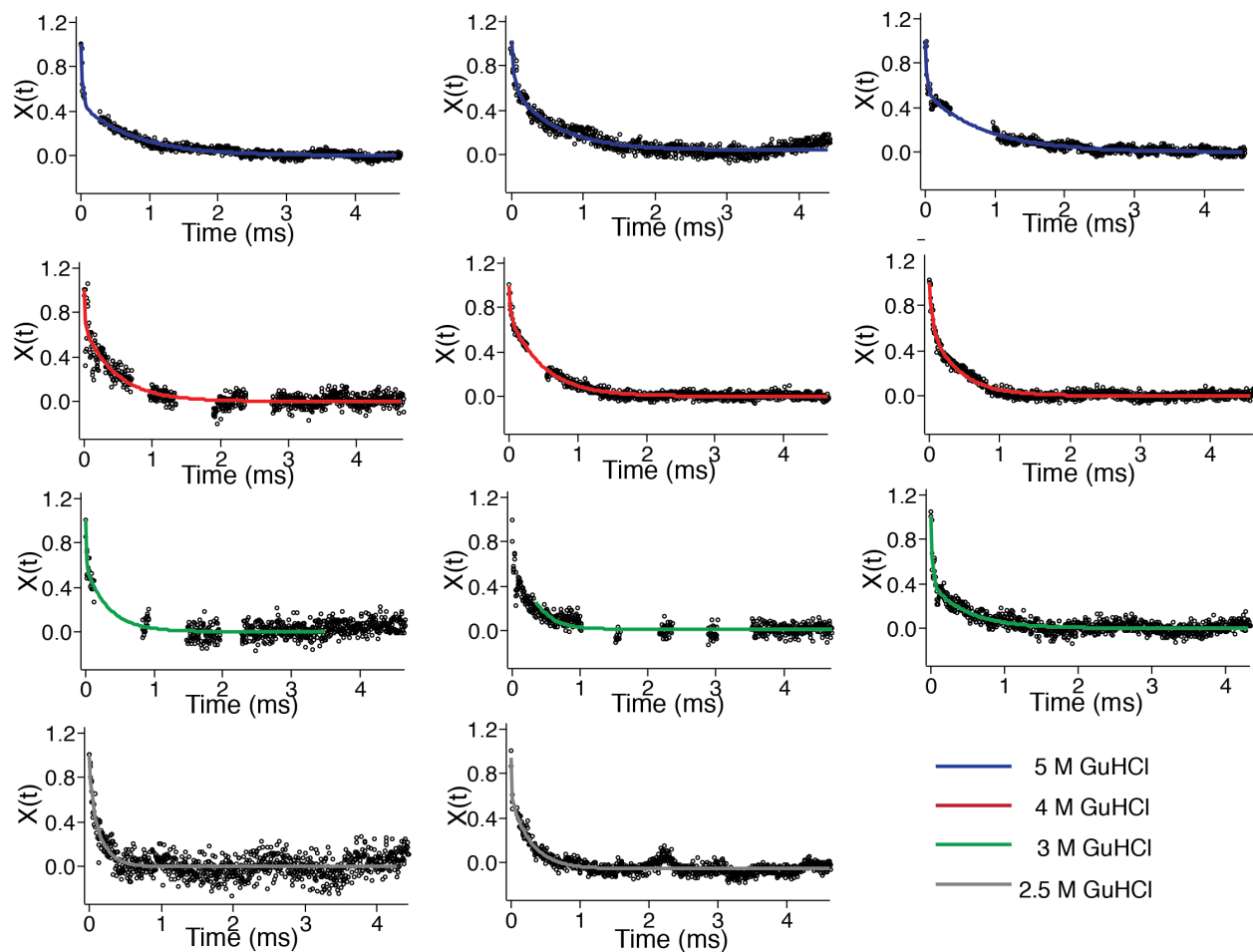


Figure 61 Pressure-jump kinetics. Data is smoothed and each data point corresponds to $5 \mu\text{s}$. Fits are double-exponential. Gaps in the kinetic traces correspond to regions of high noise that were excluded from analysis, but each GuHCl concentration has one trace with no gaps.

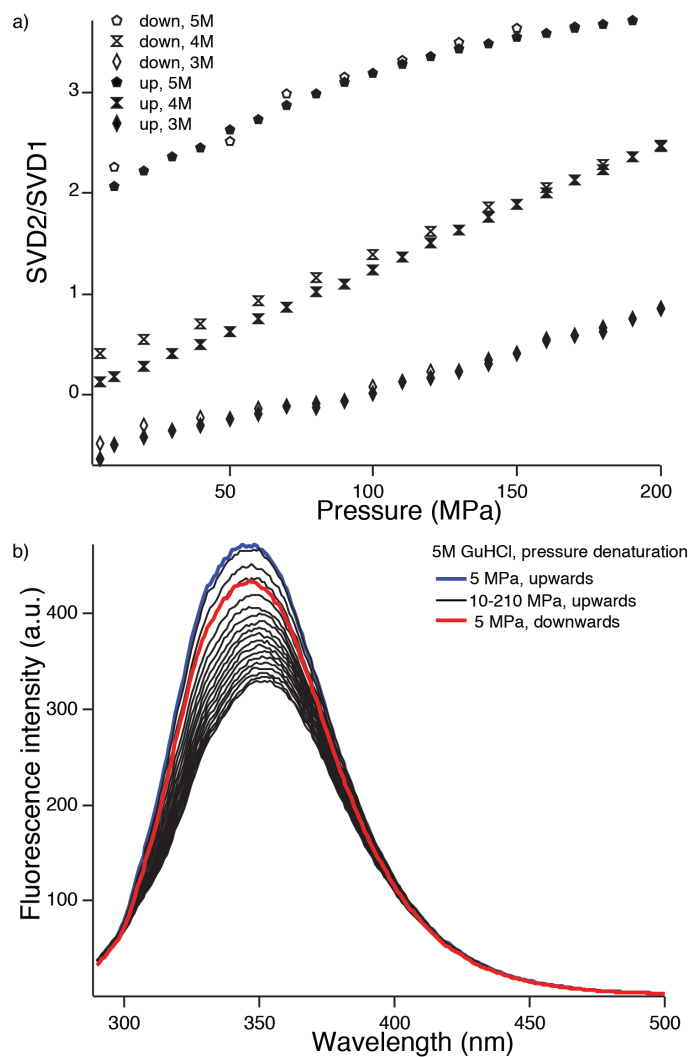


Figure 62 Pressure thermodynamics reversibility. a) High GuHCl concentration unfolding traces (closed markers) expressed as the ratio of SVD2 to SVD1 with refolding traces (open markers). Only high concentration GuHCl traces are shown because significant unfolding occurred at these concentrations. b) Fluorescence spectra of the 5M GuHCl unfolding trace. First unfolding (dark blue) and last refolding (red) trace are highlighted for comparison. The small difference in intensity between red and blue traces is due to intensity fluctuations that occur from pressurization and depressurization of the cell. SVD2/SVD1 shows nearly complete reversibility because it reflects wavelength shift (intensity component is normalized out).

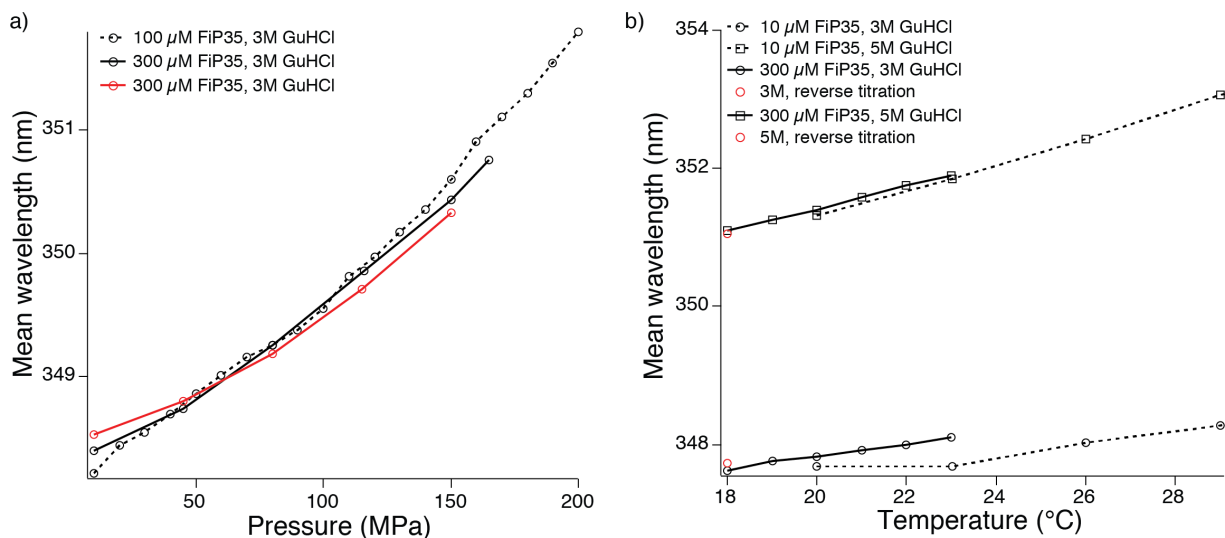


Figure 63 FiP35 reversibility at 300 μM a) Upwards (solid black line) and reverse (red line) pressure titration of 300 μM FiP35 in 3M GuHCl plotted with a pressure titration of 100 μM FiP35 in 3M GuHCl (dashed black line). No concentration dependence of the thermodynamics are observed and the high concentration titration shows complete reversibility. b) Upwards (solid black lines) and reverse (red points) temperature titration of 300 μM FiP35 in 3M (circles) or 5M (squares) GuHCl plotted with corresponding temperature titrations at 10 μM FiP35 (dashed lines). No concentration dependence of the thermodynamics are observed at the low temperatures assayed in kinetics experiments, 18 to 23 °C, and the high concentration titrations show complete reversibility.

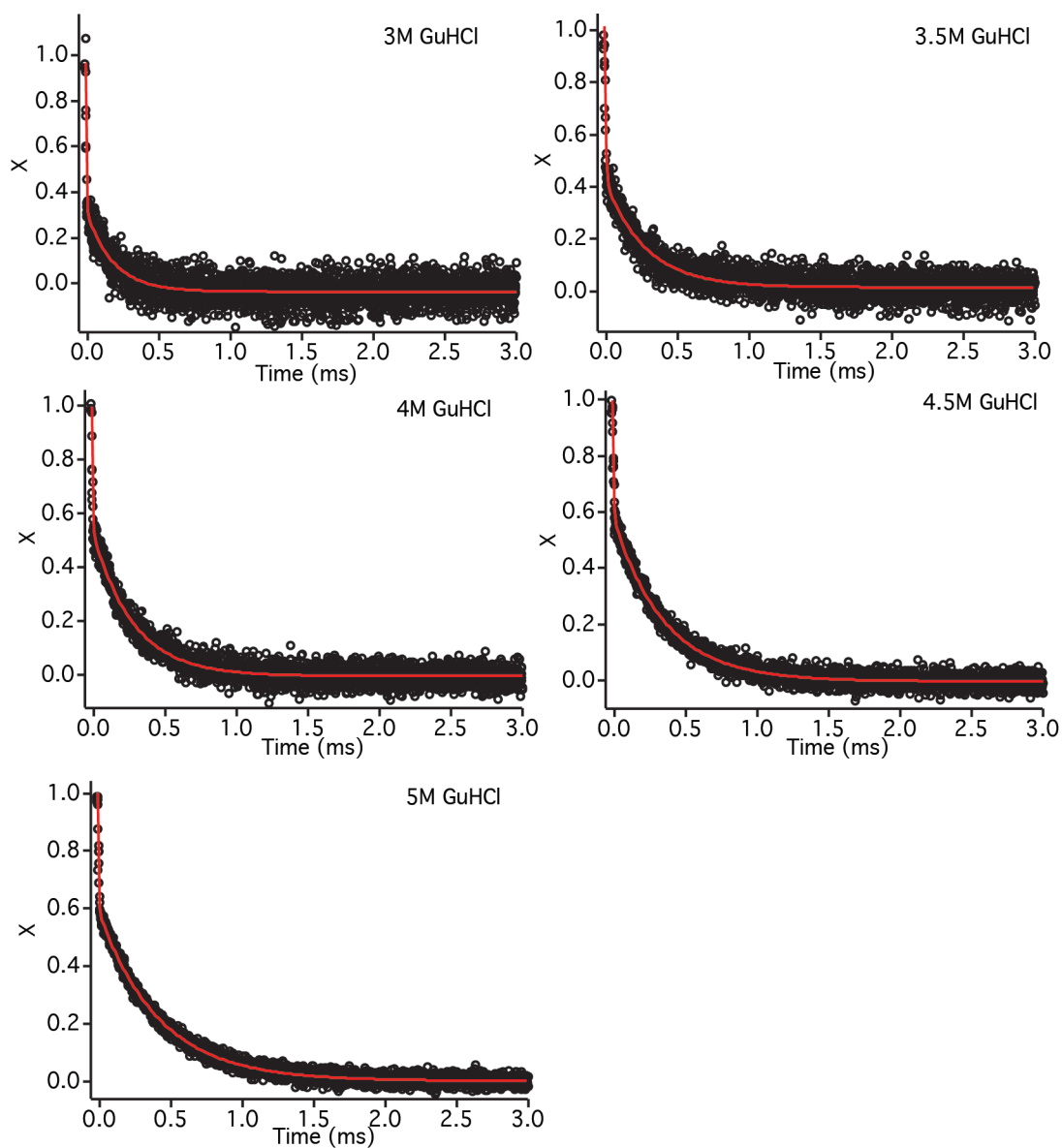


Figure 64 Temperature-jump kinetics. Double-exponential fits are shown in red. Traces and fits shown are averages over all collected traces (50). Each data point corresponds to $1.25 \mu\text{s}$. Only 3 ms of data were analyzed to avoid cooling effects at long timescales. A slow linear baseline corresponding to irreversible photobleaching has been removed from all traces.

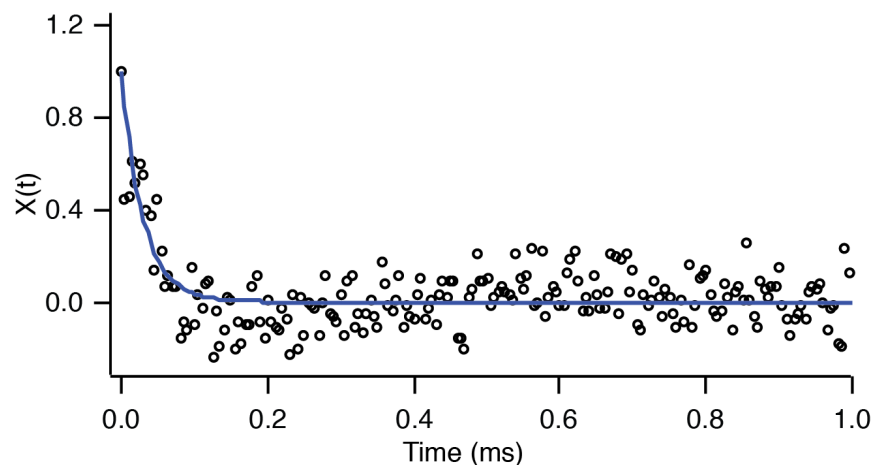


Figure 65 Pressure-jump kinetics at 0 M GuHCl. Data is smoothed and each data point corresponds to 5 μ s. The extremely small population of proteins undergoing transition makes the signal-to-noise ratio very low and only a single exponential phase can be fit with reasonable certainty. The red line represents a single-exponential fit with observed time constant of $29.5 \pm 1 \mu$ s (error is fit uncertainty).

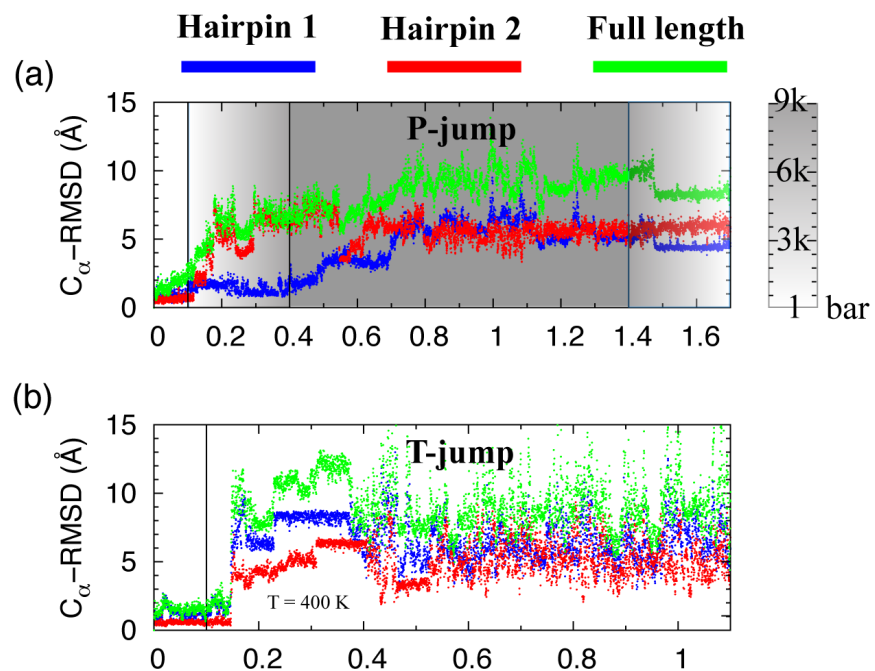


Figure 66 Structural characterization of the FiP35 unfolding trajectories. Time is in μ s. (a) Pressure-jump unfolding simulation. (b) Temperature-jump unfolding simulation. C_{α} -RMSD values have been calculated relative to the crystal structure 2F21.³⁹ The pressure is shown as the color background, which varies from 1 bar (white) to 9 kbar (dark gray). Hairpin 1 contains residues 11 to 25 and hairpin 2 contains residues 22 to 33. The full length C_{α} -RMSD are calculated using residues 7 to 35.

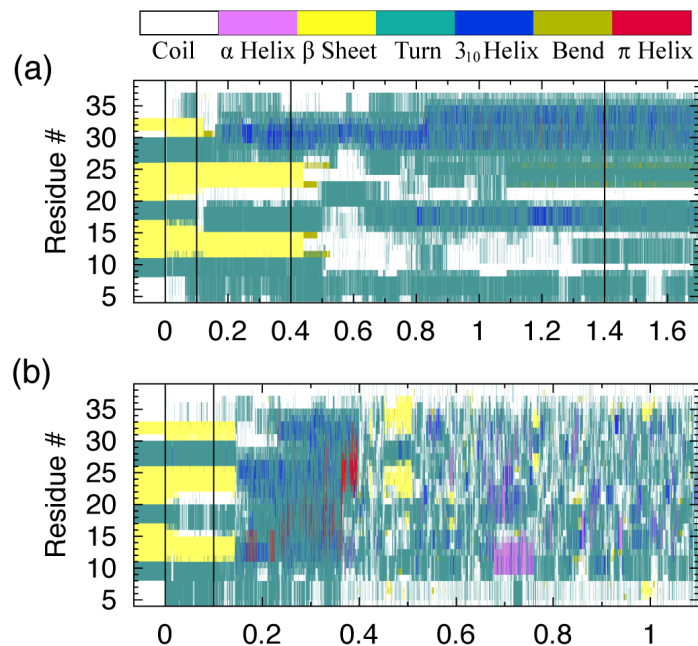


Figure 67 Time evolution of the secondary structure throughout the unfolding trajectories. Time is in μs . (a) Pressure-jump unfolding simulation. (b) Temperature-jump unfolding simulation. The secondary structure of the crystal structure is shown on the left.

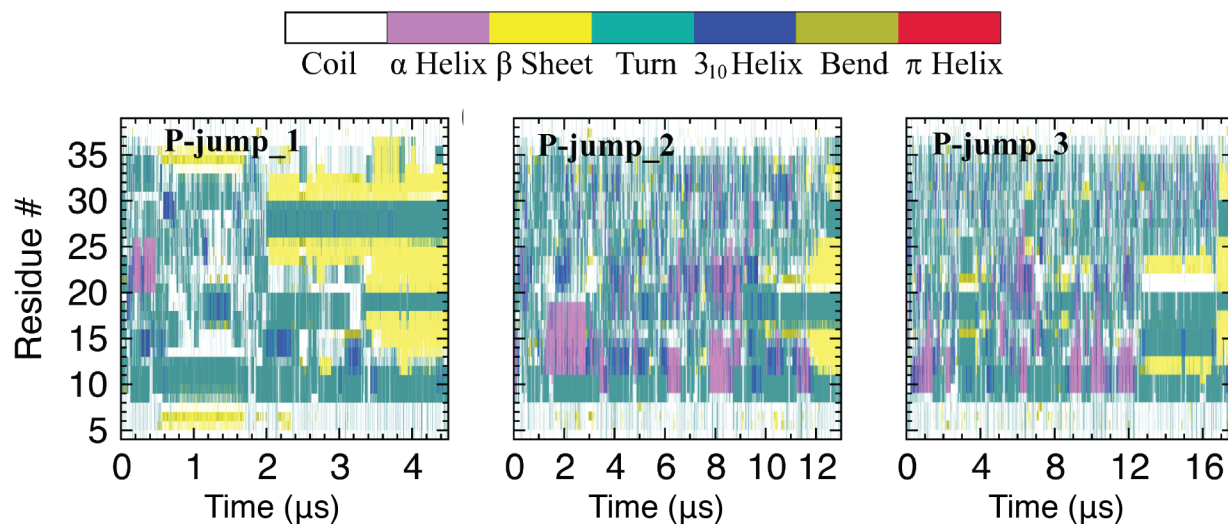


Figure 68 Time evolution of the secondary structure for the pressure-jump refolding.

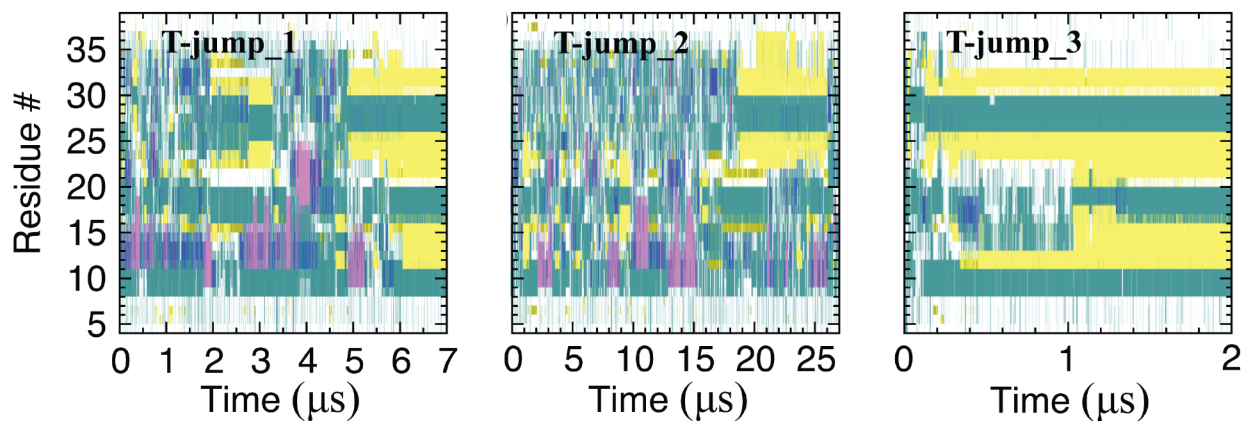


Figure 69 Time evolution of the secondary structure for the temperature-jump refolding trajectories. Secondary structure is as defined in SI figure 9.

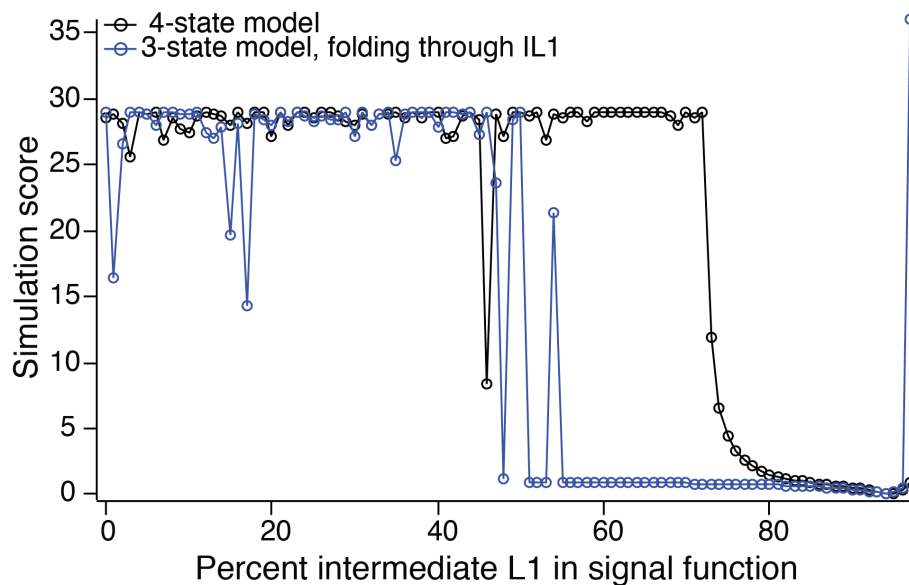


Figure 70 Score (describing distance between simulated kinetics and experimental kinetics) of all tested signal functions for the two best kinetic models. A low score indicates a closer match to experimental data. Signal function is represented on x-axis by percentage of the signal function corresponding to the loop 1 intermediate.

APPENDIX B

DONOR TO ACCEPTOR RATIO AS A REACTION COORDINATE FOR PROTEIN FOLDING IN LIVING CELLS*

B.1 FRET EFFICIENCY AND PROTEIN FOLDING

The Förster constant R_0 describes the relationship between the FRET efficiency E and the end-to-end distance between the donor and acceptor fluorophores, r :

$$E = \frac{1}{1 + \left(r/R_0\right)^6} = \frac{1}{1 + (D/A)}$$

To the best of our knowledge, no Förster constant for a fluorophore pair has been measured in the cell due to the experimental difficulties of determining exactly a distance between a FRET pair in the crowded and complex cell environment. Indeed, it is likely that the measured *in vitro* R_0 deviates somewhat from the R_0 in cell. Consequently, an absolute end-to-end distance between PGK-FRET in the cell, folded or unfolded, is not determined from our data.

To avoid using a distance measure that may be precise but not accurate, the FRET efficiency, or the donor to acceptor ratio, is used as a reaction coordinate, a detailed derivation of which for two state folding is treated in detail in earlier work¹⁶⁵. Briefly, the FRET efficiency of an ensemble of entirely folded protein (such as one would expect at room temperature), E_{fold} is related to the end to end distance of the two FRET labels in the folded state. At high temperatures, when the protein is completely unfolded, the lower FRET efficiency, $E_{denatured}$ reflects the greater distance between the two FRET labels. The temperature dependent signal observed during a thermodynamics experiment can thus be described, assuming the two-state scenario, as:

$$S(T) = f_f(T) \cdot E_{fold} + f_d(T) \cdot E_{denatured}$$

where $f_f(T)$ and $f_d(T)$ are the fraction of folded and unfolded protein, respectively. Although during a thermodynamics experiments an intermediate FRET efficiency between that of the folded and denatured state is observed at moderate temperatures, this does not correspond to an intermediate distance between the two fluorophores. Rather, it reflects the shifting fractions of folded and unfolded protein. Provided the two-state scenario remains true, only two populations of proteins exist throughout the entire experiment: one with an efficiency of E_{fold} and the other with an efficiency of $E_{denatured}$.

B.2 TEMPERATURE DEPENDENCE OF R_0

Although an accurate absolute r cannot be determined from our data, it is very instructive to see how R_0 is affected by temperature and whether the relative changes in r come out to be reasonable based on our experimental data. Such an analysis can show that our main result in chapter 5 (relative changes in T_m between interphase and mitosis) is robust irrespective of whether a D/A or a r analysis is used. Therefore we examined whether the temperature dependence of R_0

* This appendix is adapted with permission from the supplemental information for Anna J Wirth, Max Platkov, and Martin Gruebele, *Journal of the American Chemical Society*, 135(51), 19215-19221, 2013. Copyright 2013 American Chemical Society. MP contributed instrumentation support to this publication.

could significantly influence our results, which in the main text are not corrected for temperature dependence. The Förster constant is given by:

$$R_o^6 = \frac{Q(T)}{(n(T))^4} \frac{9 \ln(10) \kappa^2(T) J(T)}{128 \pi^5 N_A}$$

where $Q(T)$ is the donor quantum yield, $n(T)$ is the refractive index of the medium, $\kappa(T)$ is the orientation factor, and $J(T)$ is the overlap of the donor emission and acceptor excitation spectra.

J is given by:

$$J = \frac{\int F(\nu) \epsilon(\nu) \nu^{-4} d\nu}{\int F(\nu) d\nu}$$

Where $F(\nu)$ is the donor emission spectrum, $\epsilon(\nu)$ is the acceptor excitation spectrum (expressed as the frequency-dependent extinction coefficient) and ν is the frequency. We measured the acGFP1 emission spectrum and the mCherry excitation spectrum at two temperatures using recombinantly expressed acGFP1 and mCherry. Figure 71 shows the spectra, normalized for comparison. Over 10 degrees, the overlap does not change significantly. Assuming the overlap changes linearly with respect to temperature, the overlap changes by only 0.046% per degree and will thus will not be considered in the Forster constant temperature correction.

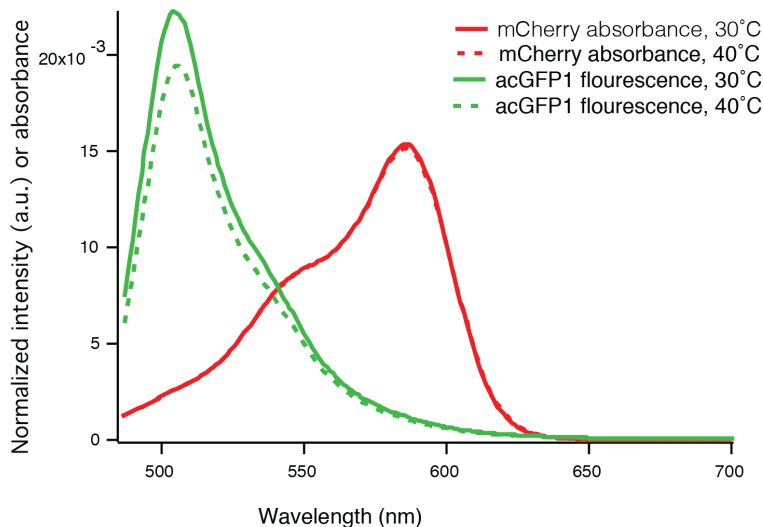


Figure 71 Temperature dependence of acceptor excitation and donor emission spectra. Spectra are normalized for comparison.

To correct for temperature effects in the donor quantum yield we used the following relationship:

$$Q(T, ^\circ C) = Q_0(1 - X)^{T-25}$$

where Q_0 is the initial donor quantum yield (0.82, from ClonTech) and X is the temperature dependence of the quantum yield ($X=0.0095$, fits our data such as in figure 2 of ref.¹⁵³). The

temperature dependence of the refractive index of a cell was approximated by the temperature dependence of the refractive index of water:²⁴⁰

$$\frac{\eta(T, ^\circ\text{C})}{\eta(20^\circ\text{C})} = \frac{1.333 + 10^{-5}[-8.376(T - 20) - .02644(T - 20)^2 + 0.00479(T - 20)^3]}{1.333}$$

The value of the non-temperature dependent portion of the R_o expression was obtained by using the published *in vitro* R_o value for EGFP and mCherry at 25 °C (5.1 nm), yielding:²⁴¹

$$R_o(T) = \sqrt[6]{\frac{Q(T)}{(n(T))^4} \cdot 67678}$$

To include the temperature dependence of the R_o in the thermodynamics calculations, the end to end distance between probes was expressed as a function of the donor to acceptor ratio D/A:

$$r(T) = \sqrt[6]{\frac{D(T)}{A(T)}} \cdot R_o(T)$$

$r(T)$ allows us to assess the effect of including temperature dependence in the Förster constant on our thermodynamics measurements but does not provide an accurate measure of the PGK-FRET end-to-end distance because it relies on the *in vitro* R_o value, rather than the unknown in-cell R_o value.

B.3 REACTION COORDINATE CHOICE FOR IN CELL DATA

Expressing the thermodynamic data from Chapter 5 as $r(T)$ instead of $D/A(T)$ does not significantly change the overall results. Thermodynamic signal traces, $S(T)$, for all interphase cells and all mitotic cells were averaged at each temperature and expressed as $r(T)$, which includes corrections for temperature dependent effects in R_o , or donor to acceptor ratio, which includes no temperature dependent corrections and is the form of data used in the main paper. Figure 72 shows the plotted data and table 13 shows a comparison of the major thermodynamics parameters calculated for each of the two scenarios. Although the shape of the thermodynamic trace and the absolute value of the thermodynamic parameters change slightly, the relative difference between interphase and mitotic cells for all parameters remains essentially the same. The T_m difference between interphase and mitotic cells is -2.12 ± 0.34 °C in the temperature corrected data versus -2.10 ± 0.52 °C in uncorrected data. Similarly, the difference between interphase and mitotic dg_1 is 479.1 ± 206.8 J mol⁻¹ K⁻¹ in the corrected data versus 522 ± 268 J mol⁻¹ K⁻¹. Because the true temperature dependence of R_o in the cell has not been determined and the above correction involves several approximations, the data is better reported without temperature correction as it is in Chapter 5. All conclusions about the system are drawn from relative differences in thermodynamic parameters between interphase and mitotic cells rather than absolute values, so changes in these values that may arise from temperature-dependent effects will not change the overall conclusions of the paper.

Although the orientation factor is also likely temperature dependent, we assume an isotropic value for $k^2(t)$ of 2/3 at all temperatures. At lower temperatures increased cellular viscosity may lead to deviations from isotropic behavior, but such a temperature dependence has yet to be measured even for *in vitro* FRET experiments. This temperature dependence should be similar between mitotic

and interphase cells, so any error introduced into our calculations from this effect will, like the other temperature effects, only affect absolute values and not relative values of main interest here.

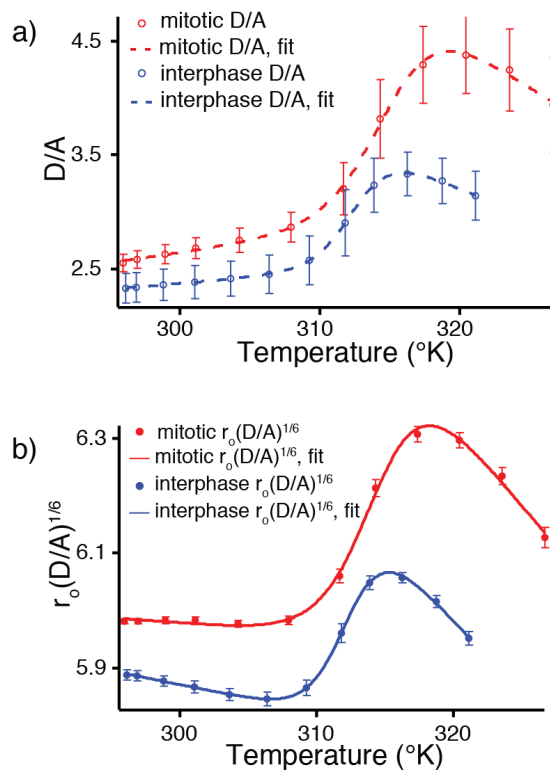


Figure 72 Correcting thermodynamic data by the estimated temperature dependence of the Förster constant. a) Raw, uncorrected D/A ratio with corresponding two-state fits. b) Un-normalized temperature-dependence corrected data. The signal function is the product of the temperature dependent Förster constant and the 6th root of the D/A ratio and corresponds to the estimated inter probe distance. For both graphs, error bars are two standard deviations of the mean for the averaging of all cell fluorescence data.

	Corrected, interphase	Uncorrected, Interphase	Corrected, Mitotic	Uncorrected, Mitotic
T_m (°K)	$312.49 \pm .07$	$311.84 \pm .12$	$314.61 \pm .27$	$313.94 \pm .40$
Δg_1 (J mol ⁻¹ K ⁻¹)	1839.6 ± 62.8	1715.7 ± 104	1360.5 ± 144	1193.5 ± 164

Table 13 Temperature corrected (r_0) vs. uncorrected (D/A) thermodynamic parameters. Although fitted thermodynamic quantities are slightly higher in the temperature corrected data, the difference between the mitotic and interphase populations is almost identical. The values for the uncorrected data reported here deviate slightly from those in text because the in text values report the average of parameters from individual fits, while these values report the parameters of a single fit to averaged raw data. The three methods of analysis show that the differences between mitotic and interphase cells are robust to the detailed method of analysis used.

APPENDIX C
SUPPLEMENTARY INFORMATION FOR “DEVELOPMENT OF THE
GFP-REASH FRET PAIR TO STUDY PROTEIN FOLDING”*

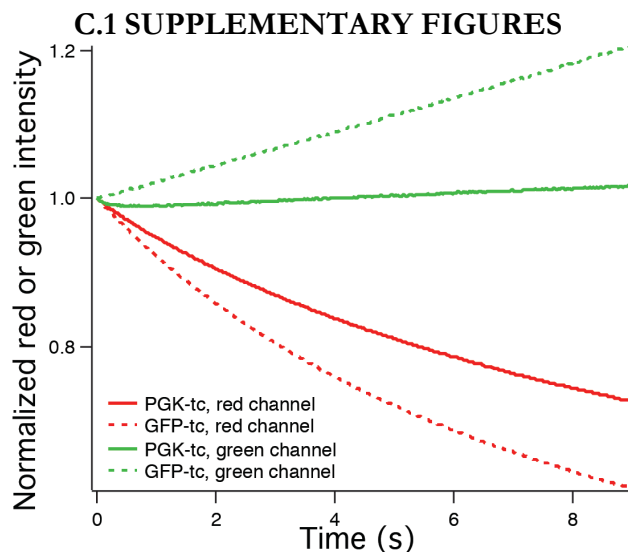


Figure 73 Red or green intensity vs. time under blue of *in vitro* protein samples on the live-cell microscope showing the effect of donor bleaching on acceptor fluorescence. Green channel intensity increases as a result of the significant bleaching of the ReAsH fluorophore, which reduces energy transfer. The slope of the green channel intensity is more positive for GFP-tc, which has a higher FRET efficiency than PGK-tc.

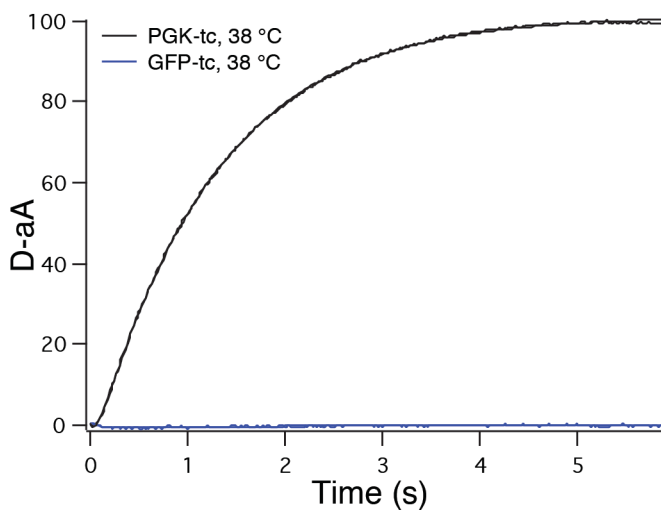


Figure 74 Kinetic traces of PGK-tc (black) and GFP-tc (blue) under blue excitation on the live-cell microscope at temperatures close to the melting temperature of PGK-tc. GFP-tc shows no sign of kinetics indicating that the observed kinetics for PGK-tc do not arise from bleaching of the ReAsH fluorophore or some other artifact. In this case, the folding is probed via FRET and the increase in D-aA indicates that the two fluorophores have moved further apart (i.e. the protein is unfolding).

* This work was done in equal collaboration with Hannah Gelman, and it is discussed half in this thesis and half in H.G.’s thesis. This work is partially reproduced from Gelman, Wirth, and Gruebele (in preparation).

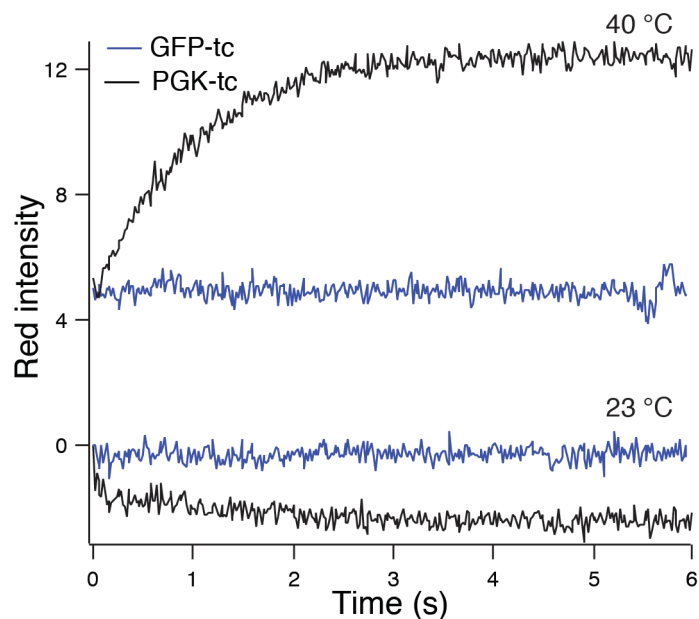


Figure 75 Selected kinetics traces of GFP-tc (blue) and PGK-tc (black) under amber (direct) excitation at room temperature and 40 °C where, for PGK-tc, the maximum folding amplitude occurs. Red intensity has been normalized to the starting point and adjusted to facilitate comparison of temperatures. GFP-tc shows no kinetics at either temperature while PGK-tc shows a robust folding amplitude close to its melting temperature. The increase in red intensity corresponds to PGK-tc unfolding and may arise from dequenching of the ReAsH fluorophore.

APPENDIX D PRESSURE-JUMP TROUBLE SHOOTING GUIDE

D.1 PROBLEMS WITH ASSEMBLY AND PRESSURIZATION

- 1) Slow pressure leak
 - Easy fixes that don't require taking apart everything:*
 - a) Assembly isn't tightened enough. Try tightening everything: pressure generator to pressure fitting connection and the bolts on top of assembly.
 - b) Valve that connects reservoir on pressure generator is not tightened enough.
 - Fixes that require taking apart everything:*
 - a) O-ring on pressure fitting is corroded. Try replacing it.
 - b) Sapphire cube is cracked.
 - c) Assembly is not put together correctly.

- 2) Sapphire cube is breaking very often
 - a) Make sure sample hole size on cube is less than the O-ring diameter
 - b) Make sure that the non-conductive cylinder's hole on the bottom does not have a chamfer
 - c) Ensure that the assembly is properly lined up during assembly. Pressure fitting hole should be directly over the sample hole. If this is slightly off center there will be too much mechanical stress on the weak edges of the sample cavity.
 - d) Make sure that the metal cylinder is in place if you are using the plastic mantle piece.
 - e) Don't tighten the assembly too much or more on one side than the other.

- 3) Assembly won't hold pressure at all
 - a) Problems with pressure generator
 - Is valve next to reservoir tightened?
 - Continue trying to increase the pressure. Are there any puddles of liquid forming anywhere around the pressure generator? If so, this is where the leak is. Sometimes the tubing that connects the pressure generator to the pressure assembly gets leaky and needs to be replaced. Sometimes valves go bad or connections can be loose. Try tightening the connections or replacing a valve.
 - Is the reservoir empty? Last time the system was filled with pressurization fluid, air could have gotten in. Refill the reservoir. Prime the system, expel everything (verify that liquid comes out), and then prime again to remove the air bubble.
 - Check for leaks within the pressure generator or very large air-bubbles. Close valve one and valve two and turn the pressurization wheel clockwise. The pressure gauge should go up and hold pressure. If it doesn't, there is a leak somewhere in the system or an air bubble.

 - b) Problems with the assembly
 - Is everything assembled and tightened?
 - Is the cube broken?
 - Are there puddles of liquid forming near the assembly, indicating a leak?

D.2 PROBLEMS WITH THE JUMP

- 1) Burst membrane doesn't break with capacitor discharge
 - a) Make sure voltage on capacitor is at least 50 V
 - b) Make sure the electrode is in contact with the burst membrane
 - c) Make sure that the assembly is grounded properly and that the electrode is connected to the capacitor.
 - d) Re-calibrate the burst membrane to make sure the starting pressure is not too low.

- 2) Pressure-jumps have a large dead-time (i.e. a tryptophan jump takes $>1-3 \mu\text{s}$ to drop from the high pressure state to the low pressure state. This can be determined by counting the number of points in the jump.)
 - a) Pressure before jump might be too low. Check again where the burst membrane breaks.
 - b) Pressure-jump might be too small. Jumps less 1500 bar are difficult to get to good resolution
 - c) Non-conductive cylinder might be getting old. Try replacing it (but make sure to re-calibrate the burst membrane! Often the breaking pressure changes when the non-conductive cylinder is switched out)
 - d) Are the membranes breaking cleanly (clean = hole completely burned through)? If not, this is the likely the origin of the low jump resolution. Try increasing the capacitor voltage or re-calibrating the burst membrane breaking pressure.

D.3 PROBLEMS WITH SIGNAL

- 1) There is not enough (or no) fluorescence signal
 - a) Find out where the problem is occurring:
 - Put a 200 μM tryptophan sample in the cube and put cube in pressure-jump apparatus but do not assemble any further. Is there signal? If not, you have either a laser problem, a pmt problem, or an alignment problem.
 - Put your protein sample in the cube and put cube in pressure-jump apparatus but do not assemble any further. Is there signal? If not, you have a protein sample problem.
 - With your sample or tryptophan still in the cube, put in the foil, pressure fitting, and burst membrane. Is there signal? If not, you are not doing a good job putting the assembly together. A likely (and very common) problem is that you are pushing the sample out of the cube when you insert the pressure fitting.
 - Fully assemble the apparatus. Pressurize up to 50 bar. Is there signal? If not, you are probably not doing a good job of putting the assembly together. A likely problem is that the foil is not adequately protecting the sample and pressurization fluid is leaking into the sample cavity.

 - b) Laser problems:
 - Is the laser lasing?
 - Is the laser mode-locked?
 - Is tripler power $>10 \text{ mW}$?
 - Is the laser's wavelength correct (should be 840 nm out of the laser cavity)?
 - Is the mode-lock stable and is the scope triggering properly?
 - Is the shutter on? (it shouldn't be open, but needs to be turned on)

c) PMT problems

- Is the correct filter on the PMT?
- Is the light guide correctly mounted in the pressure-jump assembly?
- Is the light guide on the PMT side mounted to the center of the PMT and is the light guide clear of obstruction?

d) Alignment problems

- Is the laser path blocked at any point before the pressure-jump assembly?
- Put a cube with concentrated tryptophan into the assembly but do not assemble any further. Is the laser focused and travelling through the tryptophan solution in the cube? You should be able to clearly see the beam path through the tryptophan.
- Is the laser hitting any of the mirrors or lenses improperly?

2) Signal looks terrible or weird with completely random seeming oscillations and baseline variations.

- Is the laser mode-lock stable and the scope triggering properly? Do a jump and collect the laser diode signal. A chi analysis should show no variations over time. If there are any, try tweaking the laser to improve stability or adjusting placement of the photodiode. Occasionally, the force of the jump can destabilize a laser if it is poorly aligned.
- Replace the foil. Old foil with thinned out areas can let the light from sparks come through.
- Make sure that the signal on the scope is at least 100 mV. The more signal there is, the less likely it is that electrical interference will make it through.
- Is the faraday cage up and closed?
- If noise is localized to just a few short periods of time in a data trace, the noise can be removed in data analysis. This produces data with time gaps and can be a work around, provided multiple traces under the same conditions are taken to supplement the "gapped" data.

APPENDIX E LIVE CELL METHODS AND DATA ANALYSIS

E.1 INTRODUCTION

This appendix summarizes the methods used throughout the chapters that describe imaging of protein folding thermodynamics and kinetics in living cells and in *in vitro* samples on the live-cell microscope. The Fast Relaxation Imaging (FreI) microscope¹⁵³ combines an inverted, epifluorescent microscope with an IR laser and a resistive heating element to measure protein folding. Two LED light sources provide blue or amber excitation and dual channel optics enable simultaneous measurement of emitted green and red fluorescence.

Typically, a protein of interest is expressed as a fusion sandwiched between a donor fluorescent protein (acGFP1) and an acceptor protein (mCherry), although as described in Chapter 6, alternative fluorophores can be used. U2OS cells are imaged due to their hardness and size. The following describes the sample preparation, thermodynamics and kinetics measurements, and data analysis used throughout section II of this thesis. A detailed description of the instrument and its optics is reported elsewhere¹⁵³.

E.2 SAMPLE PREPARATION

E.2.1 Cell culture Osteosarcoma cells (U2OS, ATCC) were cultured at 37°C and 5% CO₂ in Dulbecco's modified Eagle medium (DMEM, Corning Cellgro) supplemented with 10% fetal bovine serum (Fisher) and 1% penicillin-streptomycin (PS, Fisher Mediatech). Cells were split at approximately 80% confluence either into 75 cm³ tissue culture flasks (BD Falcon) for continued growth or onto 1.5 coverslips (Fisher) for imaging.

Cells were transfected with a plasmid encoding the fluorescently labeled protein of interest via Lipofectamine (Invitrogen). Proteins of interest were cloned into pDream (Genscript), a shuttle vector that contains both T7 and CMV promoters for dual expression in bacterial and mammalian cells. Depending on the protein, type of experiment, and cell culture conditions, imaging was carried out 12-48 hours following transfection.

E.2.2 Preparation of samples for imaging Cells expressing the fluorescently labeled protein of interest growing on coverslips were attached to a glass slide with a 120 μm spacer (Grace Bio Labs). Between the slide and coverslip was Optimem minimal media (Invitrogen) for dye labeling experiments to avoid excessive background fluorescence or Liebowitz medium (Invitrogen) supplemented with 30% FBS for other experiments to encourage cell viability.

For *in vitro* experiments, the protein sample at a concentration of 2-4 μM was pipetted into the well of a 120 μm spacer (Grace Bio Labs) attached to a glass slide. To aid in focusing the microscope, the protein solution is supplemented with fluorescent beads that adhere to the coverslip surface. The entire imaging chamber is sealed with a coverslip, as with the cell experiments.

E.3 PROTEIN FOLDING MEASUREMENTS

E.3.1 Thermodynamics Thermodynamics can be measured either via fast laser assisted heating or slow resistive heating. In laser assisted heating¹⁶⁸, a series of shaped IR laser pulses heat the cell in ~2 °C increments and, after a brief equilibration period, a video is taken to measure thermodynamics. This method is only suitable for relatively fast folding proteins ($\tau_{\text{obs}} < 5$ s) due to the short equilibration time between temperature steps. Because a full thermodynamics experiment is

complete in about 5 minutes, this approach is particularly useful for measurements where cell-viability is a concern.

A resistive heater mounted to the stage enables an alternative mechanism to measure folding thermodynamics. In this case, the temperature is raised through resistive heating and equilibrates ~ 5 minutes per ~ 3 °C temperature increase. In this case, a full thermodynamics experiment is complete in about an hour.

E.3.2 Temperature-jump kinetics measurements A shaped laser heating pulse (figure 76) heated the cells 4-6 °C with a resolution of ~ 30 ms. A 60 frames per second (fps) video was taken of the jump, capturing the kinetics. Temperature before the jump was held constant through resistive heating as in slow heat thermodynamics.

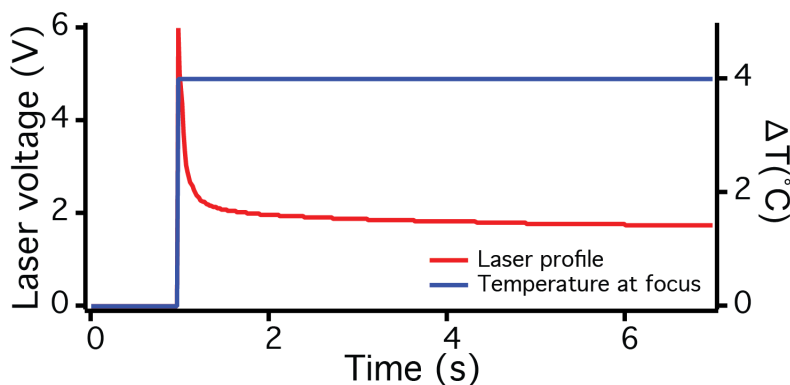


Figure 76 Temperature-jump profile. Temperature response as measured by mCherry quantum yield *in vitro* is shown in blue. The red line is the IR laser pulse shape.

E.4 PROTEIN FOLDING MEASUREMENTS

E.4.1 Thermodynamic data analysis At each temperature, the full video was averaged for each channel to produce an average green, $D(T)$, and red, $A(T)$, intensity. The temperature dependent signal was fit to a sigmoidal signal function after normalization¹⁶⁵:

$$\frac{D(T)}{A(T)} = S(T) = \frac{1 - B_F(T) + f_U(B_F(T) - B_U(T))}{B_F(T) + f_U(B_U(T) - B_F(T))}$$

Here, $B_F(T)$ and $B_U(T)$ are the linear folded and unfolded baselines and f_U describes the fraction of unfolded protein assuming a two state model:

$$f_U = \frac{1}{1 + e^{-\Delta G_{U \rightarrow F}(T)/RT}}$$

The free energy of folding is expanded to the first order such that:

$$\Delta G_{U \rightarrow F}(T) \approx \delta g_1(T - T_m)$$

where δg_1 is a Taylor series expansion coefficients and T_m is the melting midpoint. To the first order, the Van't Hoff Enthalpy at the T_m is thus given as:

$$\Delta H_{U \rightarrow F} = -\delta g_1 T_m$$

and the entropy of folding:

$$\Delta S_{U \rightarrow F} = -\delta g_1$$

E.4.2 Kinetic data analysis The fluorescence signal was averaged over the entire cell in each frame or over a region of the cell. The signal immediately following temperature-jump is expressed as:

$$S(t) = D(t) - \frac{D(t_0)}{A(t_0)} A(t)$$

where $A(t_0)$ and $D(t_0)$ are initial acceptor and donor intensities, respectively, and $A(t)$ and $D(t)$ are time-dependent donor and acceptor intensities. The raw data can be corrected for bleaching by subtracting off a linear baseline in the last second of the kinetic signal (when folding kinetics have ended)

The folding signal is generally fit to a double exponential where the first exponential is a single exponential with a positive amplitude and the second is a stretched exponential with a negative amplitude:

$$S(t) = S_0 + A_1 e^{-k_1 t} + A_2 e^{-(k_2 t)^\beta}$$

The first exponential is often an instrument response that arises from differential rates of temperature response between donor and acceptor and limits the kinetic time resolution. The second stretched exponential describes the actual folding kinetics where k_{obs} is the temperature-jump relaxation rate and β describes the deviation from two-state folding behavior.

REFERENCES

1. Bridgman, P.W. The coagulation of albumen by pressure. *Journal of Biological Chemistry* **19**, 511-512 (1914).
2. Meersman, F., Dobson, C.M. & Heremans, K. Protein unfolding, amyloid fibril formation and configurational energy landscapes under high pressure conditions. *Chemical Society Reviews* **35**, 908-917 (2006).
3. Schroer, M.A. et al. High-Pressure SAXS Study of Folded and Unfolded Ensembles of Proteins. *Biophysical Journal* **99**, 3430-3437 (2010).
4. Narayanan, S.P. et al. Extensively Hydrated but Folded: A Novel State of Globular Proteins Stabilized at High Pressure and Low Temperature. *Biophysical Journal* **102**, L8-L10 (2012).
5. Peng, X.D., Jonas, J. & Silva, J.L. Molten-Globule Conformation of Arc Repressor Monomers Determined by High-Pressure H-1-Nmr Spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 1776-1780 (1993).
6. Desai, G., Panick, G., Zein, M., Winter, R. & Royer, C.A. Pressure-jump studies of the folding unfolding of trp repressor. *Journal of Molecular Biology* **288**, 461-475 (1999).
7. Neumaier, S., Buttner, M., Bachmann, A. & Kiefhaber, T. Transition state and ground state properties of the helix-coil transition in peptides deduced from high-pressure studies. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 20988-20993 (2013).
8. Roche, J. et al. Cavities determine the pressure unfolding of proteins. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 6945-6950 (2012).
9. Chalikian, T.V. & Macgregor, R.B. Origins of Pressure-Induced Protein Transitions. *Journal of Molecular Biology* **394**, 834-842 (2009).
10. Grigera, J.R. & McCarthy, A.N. The Behavior of the Hydrophobic Effect under Pressure and Protein Denaturation. *Biophysical Journal* **98**, 1626-1631 (2010).
11. Mitra, L. et al. V-i-Value analysis: A pressure-based method for mapping the folding transition state ensemble of proteins. *Journal of the American Chemical Society* **129**, 14108-+ (2007).
12. Royer, C.A. The nature of the transition state ensemble and the mechanisms of protein folding: A review. *Archives of Biochemistry and Biophysics* **469**, 34-45 (2008).
13. Roche, J. et al. Effect of Internal Cavities on Folding Rates and Routes Revealed by Real-Time Pressure-Jump NMR Spectroscopy. *Journal of the American Chemical Society* **135**, 14610-14618 (2013).
14. Silva, J.L. et al. High-Pressure Chemical Biology and Biotechnology. *Chemical Reviews* **114**, 7239-7267 (2014).
15. Balny, C. Pressure effects on weak interactions in biological systems. *Journal of Physics-Condensed Matter* **16**, S1245-S1253 (2004).
16. Chiti, F. & Dobson, C.M. Protein misfolding, functional amyloid, and human disease. *Annual Review of Biochemistry* **75**, 333-366 (2006).
17. Meersman, F. & Dobson, C.M. Probing the pressure-temperature stability of amyloid fibrils provides new insights into their molecular properties. *Biochimica Et Biophysica Acta-Proteins and Proteomics* **1764**, 452-460 (2006).
18. Dirix, C., Meersman, F., MacPhee, C.E., Dobson, C.M. & Heremans, K. High hydrostatic pressure dissociates early aggregates of TTR105-115, but not the mature amyloid fibrils. *Journal of Molecular Biology* **347**, 903-909 (2005).

19. Shah, B.R., Maeno, A., Matsuo, H., Tachibana, H. & Akasaka, K. Pressure-Accelerated Dissociation of Amyloid Fibrils in Wild-Type Hen Lysozyme. *Biophysical Journal* **102**, 121-126 (2012).
20. Foguel, D. et al. Dissociation of amyloid fibrils of alpha-synuclein and transthyretin by pressure reveals their reversible nature and the formation of water-excluded cavities. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 9831-9836 (2003).
21. Foguel, B. & Silva, J.L. New insights into the mechanisms of protein misfolding and aggregation in amyloidogenic diseases derived from pressure studies. *Biochemistry* **43**, 11361-11370 (2004).
22. St John, R.J., Carpenter, J.F. & Randolph, T.W. High pressure fosters protein refolding from aggregates at high concentrations. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 13029-13033 (1999).
23. Snow, C.D. et al. Trp zipper folding kinetics by molecular dynamics and temperature-jump spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 4077-4082 (2004).
24. Lindorff-Larsen, K., Piana, S., Dror, R.O. & Shaw, D.E. How Fast-Folding Proteins Fold. *Science* **334**, 517-520 (2011).
25. Shaw, D.E. et al. Atomic-Level Characterization of the Structural Dynamics of Proteins. *Science* **330**, 341-346 (2010).
26. Paschek, D., Hempel, S. & Garcia, A.E. Computing the stability diagram Trp-cage miniprotein of the. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 17754-17759 (2008).
27. Garcia, A.E. & Paschek, D. Simulation of the pressure and temperature folding/unfolding equilibrium of a small RNA hairpin. *Journal of the American Chemical Society* **130**, 815-+ (2008).
28. Liu, Y.X., Prigozhin, M.B., Schulten, K. & Gruebele, M. Observation of Complete Pressure-Jump Protein Refolding in Molecular Dynamics Simulation and Experiment. *J. Am. Chem. Soc.* **136**, 4265-4272 (2014).
29. Prigozhin, M.B. & Gruebele, M. Microsecond folding experiments and simulations: a match is made. *Physical Chemistry Chemical Physics* **15**, 3372-3388 (2013).
30. Rouget, J.B. et al. Unique Features of the Folding Landscape of a Repeat Protein Revealed by Pressure Perturbation. *Biophysical Journal* **98**, 2712-2721 (2010).
31. Rouget, J.B. et al. Size and Sequence and the Volume Change of Protein Folding. *Journal of the American Chemical Society* **133**, 6020-6027 (2011).
32. Brun, L., Isom, D.G., Velu, P., Garcia-Moreno, B. & Royer, C.A. Hydration of the folding transition state ensemble of a protein. *Biochemistry* **45**, 3473-3480 (2006).
33. Vidugiris, G.J.A., Markley, J.L. & Royer, C.A. Evidence for a Molten Globule-Like Transition-State in Protein-Folding from Determination of Activation Volumes. *Biochemistry* **34**, 4909-4912 (1995).
34. Dumont, C., Emilsson, T. & Gruebele, M. Reaching the protein folding speed limit with large, sub-microsecond pressure jumps. *Nat. Methods* **6**, 515-519 (2009).
35. Prigozhin, M.B. et al. Misplaced helix slows down ultrafast pressure-jump protein folding. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 8087-8092 (2013).
36. Torrent, J. et al. The use of pressure-jump relaxation kinetics to study protein folding landscapes. *Biochimica Et Biophysica Acta-Proteins and Proteomics* **1764**, 489-496 (2006).
37. Pearson, D.S., Holtermann, G., Ellison, P., Cremo, C. & Geeves, M.A. A novel pressure-jump apparatus for the microvolume analysis of protein-ligand and protein-protein interactions: its application to nucleotide binding to skeletal-muscle and smooth-muscle myosin subfragment-1. *Biochemical Journal* **366**, 643-651 (2002).

38. Jacob, M. et al. Microsecond folding of the cold shock protein measured by a pressure-jump technique. *Biochemistry* **38**, 2882-2891 (1999).
39. Jager, M. et al. Structure-function-folding relationship in a WW domain. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 10648-10653 (2006).
40. Gelman, H. & Gruebele, M. Fast protein folding kinetics. *Q. Rev. Biophys.* **47**, 95-142 (2013).
41. Liu, F. et al. An experimental survey of the transition between two-state and downhill protein folding scenarios. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 2369-2374 (2008).
42. Liu, F., Nakaema, M. & Gruebele, M. The transition state transit time of WW domain folding is controlled by energy landscape roughness. *J. Chem. Phys.* **131**, 9 (2009).
43. Chandler, D. *Introduction to Modern Statistical Mechanics*, (New York, New York, 1987).
44. Chung, H.S., McHale, K., Louis, J.M. & Eaton, W.A. Single-Molecule Fluorescence Experiments Determine Protein Folding Transition Path Times. *Science* **335**, 981-984 (2012).
45. Piana, S. et al. Computational Design and Experimental Testing of the Fastest-Folding beta-Sheet Protein. *J. Mol. Biol.* **405**, 43-48 (2011).
46. Krivov, S.V. The Free Energy Landscape Analysis of Protein (FIP35) Folding Dynamics. *J. Phys. Chem. B* **115**, 12315-12324 (2011).
47. Noe, F., Schutte, C., Vanden-Eijnden, E., Reich, L. & Weikl, T.R. Constructing the equilibrium ensemble of folding pathways from short off-equilibrium simulations. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 19011-19016 (2009).
48. Settanni, G., Rao, F. & Caflisch, A. phi-Value analysis by molecular dynamics simulations of reversible folding. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 628-633 (2005).
49. Freddolino, P.L., Liu, F., Gruebele, M. & Schulten, K. Ten-microsecond molecular dynamics simulation of a fast-folding WW domain. *Biophysical Journal* **94**, L75-L77 (2008).
50. Ensign, D.L. & Pande, V.S. Bayesian Single-Exponential Kinetics in Single-Molecule Experiments and Simulations. *J. Phys. Chem. B* **113**, 12410-12423 (2009).
51. Ervin, J., Sabelko, J. & Gruebele, M. Submicrosecond real-time fluorescence sampling: application to protein folding. *J. Photochem. Photobiol. B* **54**, 1-15 (2000).
52. Kawahara, K. & Tanford, C. VISCOSITY AND DENSITY OF AQUEOUS SOLUTIONS OF UREA AND GUANIDINE HYDROCHLORIDE. *J. Biol. Chem.* **241**, 3228-& (1966).
53. Denos, S., Dhar, A. & Gruebele, M. Crowding effects on the small, fast-folding protein lambda(6-85). *Faraday Discuss.* **157**, 451-462 (2012).
54. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W. & Klein, M.L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **79**, 926-935 (1983).
55. Piana, S., Lindorff-Larsen, K. & Shaw, D.E. How Robust Are Protein Folding Simulations with Respect to Force Field Parameterization? *Biophys. J.* **100**, L47-L49 (2011).
56. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph. Mod.* **14**, 33-38 (1996).
57. Phillips, J.C. et al. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781-1802 (2005).
58. Shaw, D.E. et al. Anton, a special-purpose machine for molecular dynamics simulation. *Comm. ACM* **51**, 91-97 (2008).
59. Shaw, D.E. et al. Millisecond-Scale Molecular Dynamics Simulations on Anton. in *Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis (SC10)*, New Orleans, Louisiana (2010).
60. Mark, P. & Nilsson, L. Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *J. Phys. Chem. B* **105**, 24A-24A (2001).
61. Henry, E.R. & Hofrichter, J. Singular value decomposition: application to analysis of experimental data. *Methods Enzymol.* **210**, 129-192 (1992).

62. Gelman, H., Perlova, T. & Gruebele, M. Dodine as a Protein Denaturant: The Best of Two Worlds? *J. Phys. Chem. B* **117**, 13090-13097 (2013).
63. Kubelka, J., Hofrichter, J. & Eaton, W.A. The protein folding 'speed limit'. *Curr. Op. Struct. Biol.* **14**, 76-88 (2004).
64. Möglich, A., Krieger, F. & Kiefhaber, T. Molecular basis for the effect of urea and guanidinium chloride on the dynamics of unfolded polypeptide chains. *J. Mol. Biol.* **345**, 153-162 (2005).
65. Kachlishvili, K. et al. Accounting for a mirror-image conformation as a subtle effect in protein folding. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 8458-8463 (2014).
66. Jager, M., Nguyen, H., Crane, J.C., Kelly, J.W. & Gruebele, M. The folding mechanism of a beta-sheet: The WW domain. *Journal of Molecular Biology* **311**, 373-393 (2001).
67. Liu, F. & Gruebele, M. Tuning lambda(6-85) towards downhill folding at its melting temperature. *J. Mol. Biol.* **370**, 574-584 (2007).
68. Ensign, D.L. & Pande, V.S. The Fip35 WW Domain Folds with Structural and Mechanistic Heterogeneity in Molecular Dynamics Simulations. *Biophys. J.* **96**, L53-L55 (2009).
69. Beccara, S.A., Skrbic, T., Covino, R. & Faccioli, P. Dominant folding pathways of a WW domain. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 2330-2335 (2012).
70. Wagner, C. & Kiefhaber, T. Intermediates can accelerate protein folding. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 6716-6721 (1999).
71. Gierasch, L.M. & Gershenson, A. Post-reductionist protein science, or putting Humpty Dumpty back together again. *Nat. Chem. Biol.* **5**, 774-777 (2009).
72. Kuczera, K., Gao, J., Tidor, B. & Karplus, M. Free energy of sickling. A simulation analysis. *Proc. Nat. Acad. Sci. USA* **87**, 8481-8485 (1990).
73. Martin, R.B. Free energies and equilibria of peptide bond hydrolysis and formation. *Biopolymers* **45**, 351-353 (1998).
74. Srere, P.A. Complexes of sequential metabolic enzymes. *Annu. Rev. Biochem.* **56**, 89-124 (1987).
75. McConkey, E.H. Molecular evolution, intracellular organization, and the quinary structure of proteins. *Proc. Natl. Acad. Sci. U.S.A.* **79**, 3236-3240 (1982).
76. Kim, S.J., Dumont, C. & Gruebele, M. Simulation-based fitting of protein-protein interaction potentials to SAXS experiments. *Biophys. J.* **94**, 4924-4931 (2008).
77. Minton, A.P. Excluded volume as a determinant of macromolecular structure and reactivity. *Biopolymers* **20**, 2093-2120 (1981).
78. Minton, A.P. The influence of macromolecular crowding and macromolecular confinement on biochemical reactions in physiological media. *J. Biol. Chem.* **276**, 10577-10580 (2001).
79. Dhar, A. et al. Structure, function, and folding of phosphoglycerate kinase are strongly perturbed by macromolecular crowding. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 17586-17591 (2010).
80. Cheung, M.S., Klimov, D. & Thirumalai, D. Molecular crowding enhances native state stability and refolding rates of globular proteins. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 4753-4758 (2005).
81. Perham, M., Stagg, L. & Wittung-Stafshede, P. Macromolecular crowding increases structural content of folded proteins. *Febs Letters* **581**, 5065-5069 (2007).
82. Minton, A.P. Models for excluded volume interaction between an unfolded protein and rigid macromolecular cosolutes: Macromolecular crowding and protein stability revisited. *Biophys. J.* **88**, 971-985 (2005).

83. Hong, J.A. & Gierasch, L.M. Macromolecular Crowding Remodels the Energy Landscape of a Protein by Favoring a More Compact Unfolded State. *Journal of the American Chemical Society* **132**, 10445-10452 (2010).
84. Zhou, H.-X., Rivas, G. & Minton, A.P. Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences. *Annu Rev Biophys* **37**, 375-397 (2008).
85. Qin, S. & Zhou, H.-X. Atomistic Modeling of Macromolecular Crowding Predicts Modest Increases in Protein Folding and Binding Stability. *Biophys. J.* **97**, 12-19 (2009).
86. McGuffee, S.R. & Elcock, A.H. Diffusion, Crowding & Protein Stability in a Dynamic Molecular Model of the Bacterial Cytoplasm. *PLoS Comput. Biol.* **6**, e1000694 (2010).
87. Wang, Q., Zhuravleva, A. & Gierasch, L.M. Exploring Weak, Transient Protein-Protein Interactions in Crowded In Vivo Environments by In-Cell Nuclear Magnetic Resonance Spectroscopy. *Biochemistry* **50**, 9225-9236 (2011).
88. Verkman, A.S. Solute and macromolecule diffusion in cellular aqueous compartments. *Trends Biochem Sci* **27**, 27-33 (2002).
89. Dix, J.A. & Verkman, A.S. Crowding Effects on Diffusion in Solutions and Cells. *Annu Rev Biophys* **37**, 247-263 (2008).
90. Weber, S.C., Spakowitz, A.J. & Theriot, J.A. Nonthermal ATP-dependent fluctuations contribute to the in vivo motion of chromosomal loci. *Proc. Nat. Acad. Sci. USA* **109**, 7338-7343 (2012).
91. Wang, Y., Li, C. & Pielak, G.J. Effects of Proteins on Protein Diffusion. *J. Am. Chem. Soc.* **132**, 9392-9397 (2010).
92. Harada, R., Tochio, N., Kigawa, T., Sugita, Y. & Feig, M. Reduced Native State Stability in Crowded Cellular Environment Due to Protein-Protein Interactions. *J. Am. Chem. Soc.* **135**, 3696-3701 (2013).
93. Dixit, P.D. & Maslov, S. Evolutionary Capacitance and Control of Protein Stability in Protein-Protein Interaction Networks. *Plos Computational Biology* **9**(2013).
94. Bamford, C., Tipper, C. & Compton, R. *Diffusion-Limited Reactions*, 213-253 (Elsevier, 1985).
95. Wolynes, P.G. Slip boundary conditions and the hydrodynamic effect on diffusion controlled reactions. *J Chem Phys* **65**, 450 (1976).
96. Frembgen-Kesner, T. & Elcock, A.H. Absolute Protein-Protein Association Rate Constants from Flexible, Coarse-Grained Brownian Dynamics Simulations: The Role of Intermolecular Hydrodynamic Interactions in Barnase-Barstar Association. *Biophys. J.* **99**, L75-L77 (2010).
97. Frembgen-Kesner, T. & Elcock, A.H. Striking Effects of Hydrodynamic Interactions on the Simulated Diffusion and Folding of Proteins. *J. Chem. Theory Comput.* **5**, 242-256 (2009).
98. Ando, T. & Skolnick, J. On the Importance of Hydrodynamic Interactions in Lipid Membrane Formation. *Biophys. J.* **104**, 96-105 (2013).
99. Williams, M.A., Goodfellow, J.M. & Thornton, J.M. Buried waters and internal cavities in monomeric proteins. *Protein Sci.* **3**, 1224-1235 (1994).
100. Timasheff, S.N. Protein Hydration, Thermodynamic Binding, and Preferential Hydration. *Biochemistry* **41**, 13473-13482 (2002).
101. Ebbinghaus, S. et al. An extended dynamical hydration shell around proteins. *Proc. Nat. Acad. Sci. USA* **104**, 20749-20752 (2007).
102. Persson, E. & Halle, B. Cell water dynamics on multiple time scales. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 6266-6271 (2008).
103. Meister, K. et al. Long-range protein-water dynamics in hyperactive insect antifreeze proteins. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 1617-1622 (2013).

104. Papoian, G.A., Ulander, J. & Wolynes, P.G. Role of Water Mediated Interactions in Protein-Protein Recognition Landscapes. *J. Am. Chem. Soc.* **125**, 9170-9178 (2003).
105. Langhorst, U., Backmann, J., Loris, R. & Steyaert, J. Analysis of a Water Mediated Protein-Protein Interactions within RNase T1 †,‡. *Biochemistry* **39**, 6586-6593 (2000).
106. Reichmann, D., Phillip, Y., Carmi, A. & Schreiber, G. On the contribution of water-mediated interactions to protein-complex stability. *Biochemistry* **47**, 1051-1060 (2008).
107. Ikura, T., Urakubo, Y. & Ito, N. Water-mediated interaction at a protein-protein interface. *Chemical Physics* **307** 111-119 (2004).
108. Hynes, T.R., Randal, M., Kennedy, L.A., Eigenbrot, C. & Kossiakoff, A.A. X-ray crystal-structure of the protease inhibitor domain of alzheimers amyloid beta-protein precursor. *Biochemistry* **29**, 10018-10022 (1990).
109. Nooren, I. & Thornton, J.M. Diversity of protein-protein interactions. *EMBO J.* **22**, 3486-3492 (2003).
110. Sekhar, A., Lam, H.N. & Cavagnero, S. Protein folding rates and thermodynamic stability are key determinants for interaction with the Hsp70 chaperone system. *Protein Science* **21**, 1489-1502 (2012).
111. Welsh, M.J., Dedman, J.R., Brinkley, B.R. & Means, A.R. Calcium-dependent regulator protein: Localization in mitotic apparatus of eukaryotic cells. *Proc. Nat. Acad. Sci. USA* **75**, 1867-1871 (1978).
112. Buczek, O., Krowarsch, D. & Otlewski, J. Thermodynamics of single peptide bond cleavage in bovine pancreatic trypsin inhibitor (BPTI). *Protein Sci.* **11**, 924-932 (2002).
113. Scholtz, J.M. & Baldwin, R.L. The Mechanism of alpha-Helix Formation by Peptides. *Annu. Rev. Biophys. Biomol. Struct.* **21**, 95-118 (1992).
114. Wieczorek, R. & Dannenberg, J.J. Enthalpies of Hydrogen-Bonds in α -Helical Peptides. An ONIOM DFT/AM1 Study. *J. Am. Chem. Soc.* **127**, 14534-14535 (2005).
115. Doig, A.J. & Williams, D.H. Is the hydrophobic effect stabilizing or destabilizing in proteins? *Experimental Cell Research* **217**, 389-398 (1991).
116. Jones, K. & Wittung-Stafshede, P. The Largest Protein Observed To Fold by Two-State Kinetic Mechanism Does Not Obey Contact-Order Correlation. *J. Am. Chem. Soc.* **125**, 9606-9607 (2003).
117. Nguyen, H., Jäger, M., Kelly, J. & Gruebele, M. Engineering beta-sheet protein toward the folding speed limit. *J Phys Chem B* **109**, 15182-15186 (2005).
118. Shiroishi, M. et al. Structural Evidence for Entropic Contribution of Salt Bridge Formation to a Protein Antigen-Antibody Interaction. The case of hen lysozyme-HyHEL-10 Fv complex. *Journal of Biological Chemistry* **276**, 23042-23050 (2001).
119. Fersht, A.R. The hydrogen bond in molecular recognition. *Trends Biochem Sci* **12**, 301-304 (1987).
120. Klotz, I.M., Langebman, N.R. & Dahnall, D.W. Quaternary Structure of Proteins. *Annu. Rev. Biochem.* **39**, 25-62 (1970).
121. Guo, M., Ignatov, M., Musier-Forsyth, K., Schimmel, P. & Yang, X.-L. Crystal structure of tetrameric form of human lysyl-tRNA synthetase: Implications for multisynthetase complex formation. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 2331-2336 (2008).
122. Vaynberg, J. et al. Structure of an ultraweak protein-protein complex and its crucial role in regulation of cell morphology and motility. *Mol. Cell* **17**, 513-523 (2005).
123. Huang, X. et al. Structure of a WW domain containing fragment of dystrophin in complex with beta-dystroglycan. *Nature Structural Biology* **7**, 634-638 (2000).

124. Shaanan, B. Structure of human oxyhemoglobin at 2.1 Å resolution. *J. Mol. Biol* **171**, 31-59 (1983).
125. Pellicena, P., Karow, D.S., Boon, E.M., Marletta, M.A. & Kuriyan, J. Crystal structure of an oxygen-binding heme domain related to soluble guanylate cyclases. *Proc. Nat. Acad. Sci. USA* **101**, 12854-12859 (2004).
126. Bowie, J.U. & Sauer, R.T. Equilibrium dissociation and unfolding of the arc repressor dimer. *Biochemistry* **28**, 7139-7143 (1989).
127. Ou, L. et al. Incomplete Folding upon Binding Mediates Cdk4/Cyclin D Complex Activation by Tyrosine Phosphorylation of Inhibitor p27 Protein. *J. Biol. Chem.* **286**, 30142-30151 (2011).
128. Ovádi, J. & Srere, P.A. Macromolecular Compartmentation and Channeling. *sciencedirect.com* **192**, 255-280 (1999).
129. Dang, C.V., Ferguson, B., Burke, D.J., Garcia, V. & Yang, D.C.H. Interactions of aminoacyl-tRNA synthetases in high-molecular-weight multienzyme complexes from rat liver. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* **829**, 319-326 (1985).
130. Lee, S.W., Cho, B.H., Park, S.G. & Kim, S. Aminoacyl-tRNA synthetase complexes: beyond translation. *J. Cell. Sci.* **117**, 3725-3734 (2004).
131. Hausmann, C.D. & Ibba, M. Aminoacyl-tRNA synthetase complexes: molecular multitasking revealed. *FEMS Microbiology Reviews* **32**, 705-721 (2008).
132. Weber, S.C. & Brangwynne, C.P. Getting RNA and Protein in Phase. *Cell* **149**, 1188-1191 (2012).
133. Muschol, M. & Rosenberger, F. Liquid-liquid phase separation in supersaturated lysozyme solutions and associated precipitate formation/crystallization. *J Chem Phys* **107**, 1953 (1997).
134. Brangwynne, C.P. et al. Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science* **324**, 1729-1732 (2009).
135. Li, P. et al. Phase transitions in the assembly of multivalent signalling proteins. *Nature* **483**, 336-340 (2012).
136. Wu, H. Higher-order assemblies in a new paradigm of signal transduction. *Cell* **153**, 287-292 (2013).
137. Cino, E.A., Karttunen, M. & Choy, W.Y. Effects of Molecular Crowding on the Dynamics of Intrinsically Disordered Proteins. *Plos One* **7**(2012).
138. Hong, H., Goel, S. & Cai, W. Protein-Protein Interactions - Computational and Experimental Tools. *intechopen.com* (2012).
139. Renault, M. et al. Cellular solid-state nuclear magnetic resonance spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 4863-4868 (2012).
140. Selenko, P. & Wagner, G. Looking into live cells with in-cell NMR spectroscopy. *J. Struct. Biol.* **158**, 244-253 (2007).
141. Ghaemmaghani, S. & Oas, T.G. Quantitative protein stability measurement in vivo. *Nature Struct. Biol.* **8**, 879-882 (2001).
142. Xu, K., Zhong, G. & Zhuang, X. Actin, Spectrin, and Associated Proteins Form a Periodic Cytoskeletal Structure in Axons. *Science* **339**, 452-456 (2013).
143. Jones, S.A., Shim, S.-H., He, J. & Zhuang, X. Fast, three-dimensional super-resolution imaging of live cells. *Nat. Methods* **8**, 499-505 (2011).
144. Ebbinghaus, S. & Gruebele, M. Protein Folding Landscapes in the Living Cell. *J. Phys. Chem. Lett.* **2**, 314-319 (2011).
145. Dhar, A., Ebbinghaus, S., Shen, Z., Mishra, T. & Gruebele, M. The Diffusion Coefficient for PGK Folding in Eukaryotic Cells. *Biophys. J.* **99**, L69-L71 (2010).

146. Zuzuarregui, A. et al. M-Track: detecting short-lived protein-protein interactions in vivo. *Nat. Methods* **9**, 594-596 (2012).
147. Gallego, O. et al. Detection and Characterization of Protein Interactions In Vivo by a Simple Live-Cell Imaging Method. *PLoS ONE* **8**, e62195 (2013).
148. Roberts, E., Stone, J.E. & Luthey-Schulten, Z. Lattice Microbes: High-Performance Stochastic Simulation Method for the Reaction-Diffusion Master Equation. *J. Comp. Chem.* **34**, 245-255 (2013).
149. Douglas, J., Dudowicz, J. & Freed, K. Crowding Induced Self-Assembly and Enthalpy-Entropy Compensation. *Phys. Rev. Lett.* **103**, 135701 (2009).
150. Kim, Y.C. & Mittal, J. Crowding induced entropy-enthalpy compensation in protein association equilibria. *arXiv* (2012).
151. Jiao, M., Li, H.-T., Chen, J., Minton, A.P. & Liang, Y. Attractive Protein-Polymer Interactions Markedly Alter the Effect of Macromolecular Crowding on Protein Association Equilibria. *Biophys. J.* **99**, 914-923 (2010).
152. Rosen, J., Kim, Y.C. & Mittal, J. Modest Protein-Crowder Attractive Interactions Can Counteract Enhancement of Protein Association by Intermolecular Excluded Volume Interactions. *J Phys Chem B* **115**, 2683-2689 (2011).
153. Ebbinghaus, S., Dhar, A., McDonald, D. & Gruebele, M. Protein folding stability and dynamics imaged in a living cell. *Nat. Methods* **7**, 319-323 (2010).
154. Ignatova, Z. & Gierasch, L.M. Monitoring protein stability and aggregation in vivo by real-time fluorescent labeling. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 523-528 (2004).
155. Ignatova, Z. et al. From the test tube to the cell: Exploring the folding and aggregation of a beta-clam protein. *Biopolymers* **88**, 157-163 (2007).
156. Phillip, Y., Kiss, V. & Schreiber, G. Protein-binding dynamics imaged in a living cell. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 1461-1466 (2012).
157. Mir, M. et al. Optical measurement of cycle-dependent cell growth. *Proc. Nat. Acad. Sci. USA* **108**, 13124-13129 (2011).
158. Tzur, A., Kafri, R., LeBleu, V.S., Lahav, G. & Kirschner, M.W. Cell Growth and Size Homeostasis in Proliferating Animal Cells. *Science* **325**, 167 (2009).
159. Hinde, E., Cardarelli, F., Digman, M.A. & Gratton, E. Changes in chromatin compaction during the cell cycle revealed by micrometer-scale measurement of molecular flow in the nucleus. *Biophys. J.* **102**, 691-697 (2012).
160. Record Jr, M.T., Zhang, W. & Anderson, C.F. Analysis of Effects of Salts and Uncharged Solutes on Protein and Nucleic Acid Equilibria and Processes: A Practical Guide to Recognizing and Interpreting Polyelectrolyte Effects, Hofmeister Effects, and Osmotic Effects of Salts. in *Adv. Prot. Chem.*, Vol. Volume 51 (eds. Frederic M. Richards, D.S.E. & Peter, S.K.) 281-353 (Academic Press, Waltham, 1998).
161. Rosgen, J., Pettitt, B.M. & Bolen, D.W. Protein folding, stability, and solvation structure in osmolyte solutions. *Biophys. J.* **89**, 2988-2997 (2005).
162. Gelman, H., Platkov, M. & Gruebele, M. Rapid Perturbation of Free-Energy Landscapes: From In Vitro to In Vivo. *Chem. Eur. J.* **18**, 6420-6427 (2012).
163. Pardee, A.B., Dubrow, R., Hamlin, J.L. & Kletzien, R.F. ANIMAL-CELL CYCLE. *Annu. Rev. Biochem.* **47**, 715-750 (1978).
164. Watson, H.C. et al. Sequence and structure of yeast phosphoglycerate kinase. *Embo J.* **1**, 1635-1640 (1982).
165. Dhar, A. et al. Protein Stability and Folding Kinetics in the Nucleus and Endoplasmic Reticulum of Eucaryotic Cells. *Biophys. J.* **101**, 421-430 (2011).

166. Jackman, J. & Connor, P.M. Methods for synchronizing cells at specific stages of the cell cycle. *Curr. Prot. Cell Biol.* **8**, 1-20 (1998).
167. Fitzkee, N. & Rose, G. Reassessing random-coil statistics in unfolded proteins. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 12497-12502 (2004).
168. Guo, M.H., Xu, Y.F. & Gruebele, M. Temperature dependence of protein folding kinetics in living cells. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 17863-17867 (2012).
169. Osvath, S., Herenyi, L., Zavodszky, P., Fidy, J. & Kohler, G. Hierarchic finite level energy landscape model - To describe the refolding kinetics of phosphoglycerate kinase. *J. Biol. Chem.* **281**, 24375-24380 (2006).
170. Osvath, S., Sabelko, J.J. & Gruebele, M. Tuning the heterogeneous early folding dynamics of phosphoglycerate kinase. *J. Mol. Biol.* **333**, 187-199 (2003).
171. Heslopharrison, J.S. et al. The volumes and morphology of human-chromosomes in mitotic reconstructions *Hum. Gen.* **84**, 27-34 (1989).
172. Best, R.B. & Hummer, G. Coordinate-dependent diffusion in protein folding. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 1088-1093 (2010).
173. Hagen, S.J. Solvent Viscosity and Friction in Protein Folding Dynamics. *Current Protein & Peptide Science* **11**, 385-395 (2010).
174. Matouschek, A., Kellis, J.T., Serrano, L. & Fersht, A.R. MAPPING THE TRANSITION-STATE AND PATHWAY OF PROTEIN FOLDING BY PROTEIN ENGINEERING. *Nature* **340**, 122-126 (1989).
175. Naganathan, A.N. & Munoz, V. Insights into protein folding mechanisms from large scale analysis of mutational effects. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 8611-8616 (2010).
176. Weikl, T.R. Loop-closure events during protein folding: Rationalizing the shape of Phi-value distributions. *Proteins-Structure Function and Bioinformatics* **60**, 701-711 (2005).
177. Muramatsu, N. & Minton, A.P. TRACER DIFFUSION OF GLOBULAR-PROTEINS IN CONCENTRATED PROTEIN SOLUTIONS. *Proc. Natl. Acad. Sci. U.S.A.* **85**, 2984-2988 (1988).
178. Guzman, I., Gelman, H., Tai, J. & Gruebele, M. The extracellular protein VlsE is destabilized inside cells. *Journal of Molecular Biology* **426**, 11-20 (2013).
179. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W. & Prasher, D.C. Green Fluorescent Protein as a Marker for Gene-Expression. *Science* **263**, 802-805 (1994).
180. Ormo, M. et al. Crystal structure of the *Aequorea victoria* green fluorescent protein. *Science* **273**, 1392-1395 (1996).
181. Morise, H., Shimomura, O., Johnson, F.H. & Winant, J. Intermolecular Energy-Transfer in Bioluminescent System of *Aequorea*. *Biochemistry* **13**, 2656-2662 (1974).
182. Tomosugi, W. et al. An ultramarine fluorescent protein with increased photostability and pH insensitivity. *Nature Methods* **6**, 351-353 (2009).
183. Shcherbakova, D.M. & Verkhusha, V.V. Near-infrared fluorescent proteins for multicolor in vivo imaging. *Nature Methods* **10**, 751-+ (2013).
184. Shroff, H., Galbraith, C.G., Galbraith, J.A. & Betzig, E. Live-cell photoactivated localization microscopy of nanoscale adhesion dynamics. *Nature Methods* **5**, 417-423 (2008).
185. Shaner, N.C., Steinbach, P.A. & Tsien, R.Y. A guide to choosing fluorescent proteins. *Nature Methods* **2**, 905-909 (2005).
186. Hageman, J. & Kampinga, H.H. Computational analysis of the human HSPH/HSPA/DNAJ family and cloning of a human HSPH/HSPA/DNAJ expression library. *Cell Stress & Chaperones* **14**, 1-21 (2009).
187. Landgraf, D., Okumus, B., Chien, P., Baker, T.A. & Paulsson, J. Segregation of molecules at cell division reveals native protein localization. *Nature Methods* **9**, 480-U98 (2012).

188. Medintz, I.L., Uyeda, H.T., Goldman, E.R. & Mattoussi, H. Quantum dot bioconjugates for imaging, labelling and sensing. *Nature Materials* **4**, 435-446 (2005).
189. Chen, I., Howarth, M., Lin, W.Y. & Ting, A.Y. Site-specific labeling of cell surface proteins with biophysical probes using biotin ligase. *Nature Methods* **2**, 99-104 (2005).
190. Zhang, Z.W. et al. A new strategy for the site-specific modification of proteins in vivo. *Biochemistry* **42**, 6735-6746 (2003).
191. Griffin, B.A., Adams, S.R. & Tsien, R.Y. Specific covalent labeling of recombinant protein molecules inside live cells. *Science* **281**, 269-272 (1998).
192. Adams, S.R. et al. New biarsenical Ligands and tetracysteine motifs for protein labeling in vitro and in vivo: Synthesis and biological applications. *Journal of the American Chemical Society* **124**, 6063-6076 (2002).
193. Gaietta, G.M. et al. Golgi twins in late mitosis revealed by genetically encoded tags for live cell imaging and correlated electron microscopy. *Proc Natl Acad Sci U S A* **103**, 17777-82 (2006).
194. Tour, O. et al. Calcium Green FAsH as a genetically targeted small-molecule calcium indicator. *Nature Chemical Biology* **3**, 423-431 (2007).
195. Hoffmann, C. et al. A FAsH-based FRET approach to determine G protein-coupled receptor activation in living cells. *Nat Methods* **2**, 171-6 (2005).
196. Luedtke, N.W., Dexter, R.J., Fried, D.B. & Schepartz, A. Surveying polypeptide and protein domain conformation and association with FAsH and ReAsH. *Nature Chemical Biology* **3**, 779-784 (2007).
197. Coleman, B.M. et al. Conformational detection of prion protein with biarsenical labeling and FAsH fluorescence. *Biochemical and Biophysical Research Communications* **380**, 564-568 (2009).
198. Ignatova, Z. & Gierasch, L.M. Aggregation of a slow-folding mutant of a beta-clam protein proceeds through a monomeric nucleus. *Biochemistry* **44**, 7266-7274 (2005).
199. Martin, B.R., Giepmans, B.N.G., Adams, S.R. & Tsien, R.Y. Mammalian cell-based optimization of the biarsenical-binding tetracysteine motif for improved fluorescence and affinity. *Nature Biotechnology* **23**, 1308-1314 (2005).
200. Roberti, M.J., Bertocini, C.W., Klement, R., Jares-Erijman, E.A. & Jovin, T.M. Fluorescence imaging of amyloid formation in living cells by a functional, tetracysteine-tagged alpha-synuclein. *Nat Methods* **4**, 345-51 (2007).
201. Girdhar, K., Scott, G., Chemla, Y.R. & Gruebele, M. Better biomolecule thermodynamics from kinetics. *Journal of Chemical Physics* **135**(2011).
202. Crivat, G., Tokumasu, F., Sa, J.M., Hwang, J. & Wellems, T.E. Tetracysteine-Based Fluorescent Tags to Study Protein Localization and Trafficking in Plasmodium falciparum-Infected Erythrocytes. *Plos One* **6**(2011).
203. Spagnuolo, C.C., Vermeij, R.J. & Jares-Erijman, E.A. Improved photostable FRET-competent biarsenical-tetracysteine probes based on fluorinated fluoresceins. *Journal of the American Chemical Society* **128**, 12040-12041 (2006).
204. Tsytlonok, M. & Itzhaki, L.S. Using FAsH To Probe Conformational Changes in a Large HEAT Repeat Protein. *ChemBiochem* **13**, 1199-1205 (2012).
205. Joerger, A.C. & Fersht, A.R. Structural biology of the tumor suppressor p53. *Annual Review of Biochemistry* **77**, 557-582 (2008).
206. Khoo, K.H., Andreeva, A. & Fersht, A.R. Adaptive Evolution of p53 Thermodynamic Stability. *Journal of Molecular Biology* **393**, 161-175 (2009).
207. Vousden, K.H. & Prives, C. Blinded by the Light: The Growing Complexity of p53. *Cell* **137**, 413-431 (2009).

208. Cross, S.M. et al. A P53-Dependent Mouse Spindle Checkpoint. *Science* **267**, 1353-1356 (1995).
209. Gully, C.P. et al. Aurora B kinase phosphorylates and instigates degradation of p53. *Proceedings of the National Academy of Sciences of the United States of America* **109**, E1513-E1522 (2012).
210. O'brate, A. & Giannakakou, P. The importance of p53 location: nuclear or cytoplasmic zip code? *Drug Resistance Updates* **6**, 313-322 (2003).
211. Liang, S.H. & Clarke, M.F. Regulation of p53 localization. *European Journal of Biochemistry* **268**, 2779-2783 (2001).
212. Hayon, I.L. & Haupt, Y. p53: an internal investigation. *Cell Cycle* **1**, 111-6 (2002).
213. Shaulsky, G., Benzeev, A. & Rotter, V. Subcellular-Distribution of the P53 Protein during the Cell-Cycle of Balb/C 3t3 Cells. *Oncogene* **5**, 1707-1711 (1990).
214. Cho, Y.J., Gorina, S., Jeffrey, P.D. & Pavletich, N.P. Crystal-Structure of a P53 Tumor-Suppressor DNA Complex - Understanding Tumorigenic Mutations. *Science* **265**, 346-355 (1994).
215. Ginsberg, D., Michaelmichalovitz, D., Ginsberg, D. & Oren, M. Induction of Growth Arrest by a Temperature-Sensitive P53 Mutant Is Correlated with Increased Nuclear-Localization and Decreased Stability of the Protein. *Molecular and Cellular Biology* **11**, 582-585 (1991).
216. Zhang, W. et al. A Temperature-Sensitive Mutant of Human P53. *Embo Journal* **13**, 2535-2544 (1994).
217. Martinez, J., Georgoff, I., Martinez, J. & Levine, A.J. Cellular-Localization and Cell-Cycle Regulation by a Temperature-Sensitive P53-Protein. *Genes & Development* **5**, 151-159 (1991).
218. Bullock, A.N. et al. Thermodynamic stability of wild-type and mutant p53 core domain. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 14338-14342 (1997).
219. Bullock, A.N., Henckel, J. & Fersht, A.R. Quantitative analysis of residual folding and DNA binding in mutant p53 core domain: definition of mutant states for rescue in cancer therapy. *Oncogene* **19**, 1245-1256 (2000).
220. Ang, H.C., Joerger, A.C., Mayer, S. & Fersht, A.R. Effects of common cancer mutations on stability and DNA binding of full-length p53 compared with isolated core domains. *Journal of Biological Chemistry* **281**, 21934-21941 (2006).
221. Nikolova, P.V., Henckel, J., Lane, D.P. & Fersht, A.R. Semirational design of active tumor suppressor p53 DNA binding domain with enhanced stability. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 14675-14680 (1998).
222. Wirth, A.J., Platkov, M. & Gruebele, M. Temporal Variation of a Protein Folding Energy Landscape in the Cell. *Journal of the American Chemical Society* **135**, 19215-19221 (2013).
223. Mayer, S., Rudiger, S., Ang, H.C., Joerger, A.C. & Fersht, A.R. Correlation of levels of folded recombinant p53 in Escherichia coli with thermodynamic stability in vitro. *Journal of Molecular Biology* **372**, 268-276 (2007).
224. Khoo, K.H., Mayer, S. & Fersht, A.R. Effects of Stability on the Biological Function of p53. *Journal of Biological Chemistry* **284**, 30974-30980 (2009).
225. Norris, P.S. & Haas, M. A fluorescent p53GFP fusion protein facilitates its detection in mammalian cells while retaining the properties of wild-type p53. *Oncogene* **15**, 2241-2247 (1997).
226. Wang, Y., Rosengarth, A. & Luecke, H. Structure of the human p53 core domain in the absence of DNA. *Acta Crystallographica Section D-Biological Crystallography* **63**, 276-281 (2007).
227. Waldo, G.S., Standish, B.M., Berendzen, J. & Terwilliger, T.C. Rapid protein-folding assay using green fluorescent protein. *Nature Biotechnology* **17**, 691-695 (1999).

228. Friedler, A., Veprintsev, D.B., Hansson, L.O. & Fersht, A.R. Kinetic instability of p53 core domain mutants - Implications for rescue by small molecules. *Journal of Biological Chemistry* **278**, 24108-24112 (2003).
229. MacKerell, A.D. et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *Journal of Physical Chemistry B* **102**, 3586-3616 (1998).
230. Mackerell, A.D., Feig, M. & Brooks, C.L. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *Journal of Computational Chemistry* **25**, 1400-1415 (2004).
231. MacKerell, A.D., Feig, M. & Brooks, C.L. Improved treatment of the protein backbone in empirical force fields. *Journal of the American Chemical Society* **126**, 698-699 (2004).
232. Andersen, H.C. Rattle - a Velocity Version of the Shake Algorithm for Molecular-Dynamics Calculations. *Journal of Computational Physics* **52**, 24-34 (1983).
233. Miyamoto, S. & Kollman, P.A. Settle - an Analytical Version of the Shake and Rattle Algorithm for Rigid Water Models. *Journal of Computational Chemistry* **13**, 952-962 (1992).
234. Darden, T., York, D. & Pedersen, L. Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems. *Journal of Chemical Physics* **98**, 10089-10092 (1993).
235. Freddolino, P.L. & Schulten, K. Common Structural Transitions in Explicit-Solvent Simulations of Villin Headpiece Folding. *Biophysical Journal* **97**, 2338-2347 (2009).
236. Liu, Y.X., Strumpfer, J., Freddolino, P.L., Gruebele, M. & Schulten, K. Structural Characterization of lambda-Repressor Folding from All-Atom Molecular Dynamics Simulations. *Journal of Physical Chemistry Letters* **3**, 1117-1123 (2012).
237. Lippert, R.A. et al. Accurate and efficient integration for molecular dynamics simulations at constant temperature and pressure. *Journal of Chemical Physics* **139**(2013).
238. Shan, Y.B., Klepeis, J.L., Eastwood, M.P., Dror, R.O. & Shaw, D.E. Gaussian split Ewald: A fast Ewald mesh method for molecular simulation. *Journal of Chemical Physics* **122**(2005).
239. Ryckaert, J.P., Ciccotti, G. & Berendsen, H.J.C. Numerical-Integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of N-Alkanes. *Journal of Computational Physics* **23**, 327-341 (1977).
240. Dobbins, H.M. & Peck, E.R. Change of refractive-index of water as a function of temperature *J. of the Opt. Soc. of Am.* **63**, 318-320 (1973).
241. Albertazzi, L., Arosio, D., Marchetti, L., Ricci, F. & Beltram, F. Quantitative FRET Analysis With the E(0)GFP-mCherry Fluorescent Protein Pair. *Photochem. and Photobiol.* **85**, 287-297 (2009).