THERMODYNAMIC STUDIES OF THE *ESCHERICHIA COLI* FACTOR FOR INVERSION STIMULATION AND THE EUKARYOTIC NUCLEOSOME CORE PARTICLE

By

DUANE A. HOCH

A dissertation submitted in partial fulfillment of
the requirements for the degree of

DOCTOR OF PHILOSOPHY

WASHINGTON STATE UNIVERSITY
School of Molecular Biosciences

December 2006

To the Faculty of Washington State University

The members of the Committee appointed to examine the thesis of DUANE A. HOCH find it satisfactory and recommend that it be accepted.

_____
Chair

_____

_____

_____

_____

ACKNOWLEDGMENT

First and foremost, I would like to acknowledge Dr. Lisa M. Gloss for taking me in and mentoring me in many aspects of protein biochemistry during the course of my education at Washington State University. In addition I would like to acknowledge my family of the past 5 years, the Gloss lab, both past and present members who have helped me so much in becoming a better scientist. I am grateful to my committee members: Drs. Raymond Reeves, William Davis, and ChulHee Kang for serving on my committee and giving guidance throughout my graduate career. Lastly I want to acknowledge my family and friends, in particular my wife, for their love, kindness and support during both the good and the bad.

THERMODYNAMIC STUDIES OF THE *ESCHERICHIA COLI* FACTOR FOR INVERSION STIMULATION AND THE EUKARYOTIC NUCLEOSOME CORE PARTICLE

Abstract

By Duane A. Hoch, Ph.D.
Washington State University
December 2006

Chair: Lisa M. Gloss

The stability and kinetics of the Factor for Inversion Stimulation (FIS) were determined and compared with other intertwined segment swapped alpha-helical DNA binding dimers to aid in understanding oligomeric protein folding. The equilibrium stability of recombinant FIS had been determined by both urea and guanidinium-induced denaturation, using circular dichroism (CD). Stopped-flow CD was used to determine the folding mechanism of the FIS using both urea and guanidinium as the denaturants. Similar to the folding mechanisms of the eukaryotic histones determined previously, the kinetic mechanism of FIS is a sequential process: 1) unfolded monomers associate in a burst phase reaction to form a dimeric intermediate; 2) this intermediate further folds in a first-order reaction to yield the native dimer in the rate-limiting step of the folding reaction. The sequential formation of an on-pathway dimeric intermediate was confirmed using urea double jump experiments.

To investigate the role of protein stability in higher order protein-DNA complexes, a Förster resonance energy transfer (FRET) system was created to monitor the salt-induced unfolding of nucleosome core particles (NCPs) containing H2A.1 and

H2A.Z. The donor (D)-acceptor (A) pair chosen was tryptophan and Cys-AEDANS, respectively, with a Förster distance of 20 Å. The FRET system was able to monitor the dissociation of the H2A-H2B dimers from nucleosome core particles reconstituted with the 601 artificial positioning sequence. The free energy of dissociation, ΔG, for the dimers was determined from FRET equilibrium studies in which two transitions were observed. These transitions correlate to the dissociation of each dimer with positive cooperativity. Although the stability of the FRET NCPs incorporated with H2A.Z were not significantly altered, an increase in the cooperativity of the H2A.Z dimers was observed.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMTS iii

ABSTRACT iv

LIST OF TABLES viii

LIST OF FIGURES ix

CHAPTER

I. INTRODUCTION 1

1. Protein Folding and Stability 2

2. Energetics of Protein Folding 4

3. Thermodynamic Studies 6

4. Mechanisms of Protein Folding (Kinetics) 10

5. Models of Protein Folding 11

6. Thermodynamic Views of Protein Folding 12

5. ISSAD Oligomers 14

4. Nucleosome Core Particle 15

6. Histone Structure 17

5. Chromatin Structure and Regulation 17

7. Figures 23

8. References 30

II. EQUILIBRIUM AND KINETIC STUDIES OF FIS, FACTOR FOR INVERSION STIMULATION 38

1. Summary 39

2. Abbreviations..... 40

3. Introduction 41

4. Materials and Methods..... 42

5. Results 46

6. Discussion..... 57

7. Tables and Figures 67

8. References 77

III. MULTI-PROBE PROTEIN FRET ANALYSIS OF WT AND H2A.Z-INCORPORATED NUCLEOSOME CORE PARTICLE DYNAMICS 83

1. Summary 84

2. Abbreviations..... 85

3. Introduction 86

4. Material and Methods 93

5. Results 100

6. Discussion..... 112

7. Tables and Figures 120

8. References 136

IV. CONCLUSIONS AND FUTURE DIRECTIONS 142

1. Figures..... 151

2. References 155

LIST OF TABLES

CHAPTER II.

2.1 Parameters describing the equilibrium folding/unfolding reactions of FIS 67

2.2 Parameters describing the kinetic folding/unfolding reactions of FIS 68

CHAPTER III.

3.1 Distances between the Cα atoms of the residues mutated for incorporation of the FRET donors and acceptors 120

3.2 Fitted parameters describing the equilibrium stability of the histones modified with FRET donors and acceptors 121

LIST OF FIGURES

CHAPTER I: Introduction

1.1. Representative reaction coordinate diagram showing the free energy difference between native and unfolding species 23

1.2. Levinthal's paradox as a free energy landscape of protein folding 24

1.3. Three-dimensional representations of the new view of energy landscapes 25

1.4. Cartoon depicting domain swapping between two "closed" monomers 26

1.5. Ribbon diagram of the eukaryotic H2A/H2B histone dimer 27

1.6. Schematic of the assembly/disassembly pathway of the NCP 28

1.7. Models for chromatin remodeling 29

CHAPTER II

2.1 Ribbon diagram of the intertwined FIS homodimer 69

2.2 Far-UV CD spectra of FIS 70

2.3 Representative equilibrium GdmCl-induced unfolding transitions for FIS monitored by far-UV CD spectra using singular value decomposition 71

2.4 SF-CD refolding kinetics for GdmCl-denatured FIS 72

2.5 GdmCl dependence of SF-CD folding and unfolding kinetic data 73

2.6 Folding mechanisms for a dimeric protein with a dimeric intermediate 74

2.7 Generic diagram of a stopped-flow apparatus and double jump experiment 75

2.8 Urea double-jump assay for the formation of native FIS dimer 76

CHAPTER III

3.1 A. Scheme of the nucleosome unfolding/folding pathway

B. Crystal structure of NCP with mutated sites highlighted in various colors .. 123

3.2 F_{app} plots for the urea equilibrium denaturation transitions of the Cys-AEDANS modified H2A-H2B dimers..... 124

3.3 Apparent fraction of unfolded monomers (F_{app}) as a function of [Urea] for the unfolding transitions of WT and modified H2A•H2B dimers..... 125

3.4 F_{app} plots for the GdmCl equilibrium denaturation transitions of the Trp engineered H3-H4 tetramers..... 126

3.5 Apparent fraction of unfolded monomers (F_{app}) as a function of [GdmCl] for the unfolding transitions of the WT and Trp-containing H3-H4 oligomers at 4 □M monomer..... 127

3.6 Normalized Tyr FL monitoring the salt-induced dissociation of NCPs reconstituted with the 601 DNA sequence and recombinant WT histones 128

3.7 5% acrylamide native gels stained with EtBr of NCPs reconstituted with the 149 bp 601 DNA fragment..... 129

3.8 Representative FL data for the salt-induced dissociation of the FRET NCPs ... 130

3.9 Normalized donor FL data for the salt-induced dissociation of multiple FRET NCPs..... 131

3.10 Schematic representation of NCP dissociation considering relative D-A and D-A' distances..... 132

3.11 The salt dependence of the free energy of dissociation, □G°(H_2O), of the dimer from the H3-W78/H2A-108Cys-AEDANS 601-NCP 133

3.12 Equilibrium dissociation transitions for H2B-109Cys-AEDANS FRET NCPs with H2A and H2A.Z 134

3.13 The salt dependence of the free energy of dimer dissociation for H3-W78/H2B-109Cys-AEDANS NCPs with H2A and H2A.Z..... 135

CHAPTER IV

4.1 Ribbon model of a mutant FIS (K36E) in which the position of the N-terminal β-hairpins are resolved, depicting the locations of the Trp mutations..... 151

4.2 FL emission scans of folded and unfolded FIS S30W 152

4.3 Equilibrium urea-induced unfolding transitions for FIS M65W monitored by far-UV CD and Trp FL..... 153

4.3 Urea dependence of SF-FL refolding kinetics for FIS M65W 154

CHAPTER I

INTRODUCTION

The research described in this dissertation is divided into two focuses. The first, smaller body of work describes the equilibrium folding and kinetics of the *E. coli* DNA-binding transcription factor, Factor for Inversion Stimulation (FIS). The goal of this segment is to provide a deeper understanding of the folding reactions of oligomeric proteins, in particular, a group of proteins classified as intertwined segment-swapped □-helical DNA-binding dimers (ISSADDs). The second, larger body of work describes the design of a Förster resonance energy transfer (FRET) system to study the dynamics of the nucleosome core particle (NCP). This work establishes a system that will permit more detailed studies of the equilibrium and kinetic disassembly of the nucleosome core particle. The dissertation will then conclude with a discussion of this research and future directions for both of these areas of focus.

PROTEIN FOLDING AND STABILITY

In the wake of the human genome project, we are left with a huge quantity of data to analyze with our limited understanding of how to predict protein structure, folding and function from an amino acid sequence. Understanding how proteins fold is very important to many aspects of research. In biotechnology and medicine, it is of great interest to produce enzymes with novel, specific catalytic and/or therapeutic properties. The challenges in this field include: (1) the ability to design a novel fold, (2) the ability to design a more stable and/or more soluble fold, and (3) the ability to design precise features for specific binding or better active sites for new or more efficient catalytic activity in a known fold. There are many examples of these challenges just listed and how we have advanced in this field. A few of these examples will be discussed below.

Current focus of biotechnology in the U.S. is on renewable alternatives to fossil fuels. In particular; companies, such as Genencor Inc., are trying to engineer more thermostable and active cellulases to convert cellular biomass to ethanol for fuel and chemicals (*1*). Through the use of structural and biochemical studies of a family of cellulases, a more directed protein-engineering approach can be utilized to improve the cellulases for industrial purposes (*1*). Another common technique is the use of directed protein evolution to insert random mutations and then select those mutations that give the properties desired whether it be improved solubility, stability, or catalytic efficiency (*2*). One example of directed evolution in the industrial setting is the improvement of the catalytic efficiency of an esterase toward a specific antibiotic intermediate in an organic solvent (*3;4*). After multiple generations of random mutagenesis and gene recombination, a series of mutations were found to be able to improve the enzyme's catalytic efficiency over 100-fold. These mutations were modeled and found to be in areas away from the binding site for the antibiotic that would not have typically been found by directed enzyme design (*4*).

A better knowledge in areas of protein folding and design has many uses in medicine that range from designing therapeutic proteins to treat disease, to understanding various disease processes caused by protein misfolding. In treatment of diabetes, multiple insulin injections need to be given throughout the day in order to maintain appropriate insulin levels. Currently, insulin analogs are being designed to improve their pharmacokinetic and pharmacodynamic properties (for review, (*5*)). For example, insulin analogs that better mimic meal-induced insulin levels are being created through specific amino acid substitutions that have been found to disrupt higher-order insulin assemblies

that slow down the breakdown and absorption. Other analogs are being designed that exploit other physicochemical properties that promote normal basal-activity. Another application of protein folding in the treatment of diseases is in the creation of drugs that prevent oligomerization of proteins critical to disease processes. One class of HIV protease inhibitors are used to prevent a dimerization event important in HIV infection (*6*). The HIV protease is a 99-amino acid monomeric protein that is enzymatically inactive but gains aspartyl protease activity upon dimerization. The functional enzyme cleaves the gag and gag-pol proteins that are needed in HIV virion maturation and therefore produces noninfectious virus particles (*7*). A greater understanding of protein structure and folding is also critical in the understanding of various protein misfolding diseases, such Cystic fibrosis, amyloid-fibril forming diseases (Alzheimer's, Huntington's, Parkinson's) and prion-related diseases (transmissible spongiform encephalopathies).

ENERGETICS OF PROTEIN FOLDING

The laws of thermodynamics that guide all life processes both big and small also apply to protein stability and folding. The first law of thermodynamics states simply that energy is neither created nor destroyed, and therefore the total energy of the system and its surrounding is constant (*8*). The energy components that make up the system are enthalpy (the heat of the system), and entropy (the randomness of the system). The total entropy of the system and its surroundings according to the second law of thermodynamics always increases for a spontaneous process (*8*). The overall free energy change can be described by the classical thermodynamic equation:

$$\Delta G = \Delta H - T\Delta S \qquad \textbf{(1)}$$

where ΔG is Gibbs free energy of unfolding, ΔH is enthalpy change, T is temperature, and ΔS is entropy change.

In a simple scheme, N→U, where a native protein (N) unfolds into its unfolded species (U), the ΔG of this reaction is positive meaning that unfolding is not favored. Typically, the ΔG for most proteins is a small positive number, such that proteins are only marginally stable in the cell (*9*). This is of great importance to biological systems because it allows for flexibility of protein structure to aid in function as well as for protein turnover in the cell. The free energy surface of a protein folding reaction encompasses a vast number of interactions that involve both the making and breaking of weak chemical bonds and it is the balance of these many small interactions, both enthalpically and entropically, that dictate how a protein ultimately folds. To fully understand how a protein folds, one must consider all the possible protein-protein and protein-solvent interactions in the denatured species as well as in the folded species.

Enthalpy is defined as the heat content of the system, which results from the formation and breaking of all the chemical bonds, both weak and strong (*8*). The majority of enthalpic interactions that stabilize protein structure include noncovalent electrostatic interactions such as salt bridges, hydrogen bonds, and weak dipole-dipole interactions known as van der Waals interactions (*8*). In the unfolded/denatured form of a protein, the majority of the enthalpic interactions occur between the protein and the solvent. Even though many protein-solvent interactions in the unfolded state are replaced by protein-protein interactions in the native state, there is an overall favorable enthalpy change.

The randomness or disorder of a system and its surroundings is known as entropy. In the unfolded state of a protein, the entropy of the polypeptide chain is very high because of the astronomical number of possible conformations the chain can sample. This large number of conformations is highlighted in the seminal paper of Cyrus Levinthal (*10*). He made the simple calculation that an unfolded protein of 100 residues, with a minimum of three different conformations for each residue (an underestimate as some amino acids have up to seven possible rotamers), could sample 3^{100}, or 5×10^{47} different conformations. The implications of this number, in the context of folding kinetics, will be discussed in the next section. As a protein folds, nonpolar sidechains interact with each other and typically are packed into the core of the protein and excluded from the aqueous solvent. Despite a decrease in entropy of the polypeptide chain(s), there is a subsequent large increase in entropy of the solvent. This occurs because in the unfolded state, water molecules form well-ordered "cages" around nonpolar amino acid side chains to help solubilize them. Upon folding, the water molecules are released from this caged structure causing an increase in solvent entropy (*9*). This phenomenon of excluding nonpolar residues to interact and exclude solvent is known as the hydrophobic effect (*8;9*). The hydrophobic effect is thought to be the main driving force in protein folding but must be considered along with the enthalpy produced from the change in electrostatic interactions, in particular hydrogen bonding (*11*).

THERMODYNAMIC STUDIES

The goal of protein folding equilibrium studies is to determine the free energy difference, $\Delta G°$, between the native and unfolded species as indicated in Figure 1.1.

There are two critical requirements for thermodynamic information to be obtained: 1) the reaction must be reversible, so that N $\rightleftharpoons$ U; and 2) even if the reaction appears reversible, there must be no hysteresis, in other words, the pathway from N → U is the same as U → N. The stability of a protein can be defined by the equation $\Delta G^o = \Delta RT \ln K_{eq}$, where ΔG° is the difference in stability between the unfolded and folded states, K_{eq} is the equilibrium constant ([U]/[N]), R is the ideal gas constant (0.001987 kcal mol^{-1}), and T is the temperature in Kelvin (generally 298 °K). If you consider a protein of marginal stability, 5 kcal mol^{-1} and use the equation defined above, the ratio between the unfolded to native species is ~1/20,000.

Currently, there are no methods to detect such a small amount of unfolded species. Therefore we need to perturb the system such that the amount of U and N are simultaneously detectable and more similar in concentration. Various perturbants used in folding studies include temperature, pH, pressure and chemical denaturants. There are various advantages and disadvantages of using the different denaturation techniques. Some advantages of pressure denaturation include: 1) pressure-jumps do not significantly change solvent properties; 2) perturbations propagate rapidly so that sample homogeneity is not a problem; and 3) it permits measurements on the microsecond timescale for complementary kinetic studies. In addition, pressure denaturation can be applied in both positive and negative pressure directions (*12;13*). Pressure denaturation is thought to occur by destabilizing native structure by forcing water molecules into the protein (*14*). A major disadvantage is the requirement of specialized equipment. Temperature-denaturation can be performed rapidly and is commonly employed because of its ease of use, but proteins denatured by heat are less likely to be reversible. Although

temperature-jumps are rapid, an additional disadvantage is that they are only effective in the unfolding direction. Chemical denaturation is also easy to use but may not completely unfold a protein, with retention of some residual structure. In contrast, pressure is able to completely unfold most proteins but has been found to induce formation of partially structured intermediates in a significant number of proteins (*15*).

Equilibrium experiments performed in our lab utilize the use of chaotropic chemical denaturants such as urea and guanidinium chloride. Urea is a nonionic chemical denaturant that unfolds proteins by increasing the favorability of solvent interactions with the unfolded state by the increased hydrogen bonding of the solvent with the polypeptide backbone and weakening the hydrophobic effect. Guanidinium chloride is an ionic denaturant that also binds to the peptide backbone but is stronger than urea because of its ability to mask the electrostatic interactions important to the native structure (*16*).

The most common method utilized to study equilibrium folding is spectroscopy. Common spectroscopy methods include circular dichroism (CD), fluorescence (FL), absorbance (Tyr and Trp), nuclear magnetic resonance (NMR), mass spectrometry, hydrogen/deuterium exchange (H/D-ex), and Fourier Transform infrared spectroscopy (FTIR). NMR is a very powerful technique typically used to determine the three-dimensional structure of macromolecules in solution but can help provide residue specific information of proteins on time scales ranging from sub-nanoseconds to days. Unfortunately, a major disadvantage of NMR is that it can only be used with small proteins (< 25 kDa) and it requires a large amount of protein. Hydrogen-deuterium exchange monitored by mass spectrometry has become increasingly popular in protein

dynamics and folding because of its high sensitivity and its ability to provide high resolution backbone dynamics (for review, (*17;18*)). Besides being more sensitive (sub-femtomolar range) than other optical techniques (millimolar – sub-micromolar), a unique advantage of mass spectrometry is the ability to detect and characterize individual conformational states that may co-exist in solution at equilibrium (*17*). The other methods mentioned can only provide population-averaged conformational information. Fourier Transform infared spectroscopy is another technique that is rapidly emerging as a common tool in protein folding, especially for the folding of proteins with high beta-sheet content (for review, (*19;20*)). In proteins, FTIR monitors the vibrational stretching of the C=O amide bond that is very sensitive to secondary structure content because of the changes in hydrogen bonding, dipole-dipole interactions, and the geometry of the polypeptide backbone (*21*). This technique is particularly advantageous in monitoring beta-sheet structures because they give very distinct diagnostic IR bands, where α-helical structures give rise to close or overlapping bands that are more difficult to analyze. The two major spectroscopic techniques utilized in this thesis are circular dichroism (CD) and fluorescence (FL) spectroscopy. CD measures the difference in the UV absorption between left and right circularly polarized light. Far-UV CD (190-250 nm) gives information on the 2° structural properties of a protein and near-UV CD (260-300 nm) provides information on 3° and 4° structure (*22*). Fluorescence is a very sensitive method that probes the local 3° and 4° structure around the fluorophore. Many proteins contain the intrinsic amino acid fluorophores, tyrosine and tryptophan. Proteins that do not contain these amino acids can be engineered to contain unique tryptophan or tyrosine residues to be used in such studies.

In additional to more traditional FL techniques, Förster resonance energy transfer (FRET) is powerful tool that has become widely used (for review, (*23;24*)). FRET is a process that occurs when the emission spectrum of a fluorophore (donor) overlaps with the excitation of another fluorophore (acceptor) in close proximity. Resonance energy transfer from donor to acceptor is then able to occur rather than the fluorescent decay back to the donor ground state. Therefore, when the donor is excited in the presence of an acceptor, the resultant emission spectrum is that of the acceptor, and not of the donor, unless the two fluorophores are too far apart. FRET is strongly dependent on the donor-acceptor distance, thus it can serve as a molecular ruler to determine the distances between residues within a protein or macromolecular complex, as well as determine when two interacting protein partners are in close contact.

MECHANISMS OF PROTEIN FOLDING (KINETICS)

The information for the three-dimensional structure of a protein is contained in the primary structure of amino acids. This was first demonstrated in the work of C.B. Afinsen and colleagues with ribonuclease A (*25*). They observed that upon removal of chemical denaturants, ribonuclease A would spontaneously refold to its native structure. A large number of proteins both large and small have been shown to spontaneously refold upon reversal of a wide range of denaturation techniques (*26;27*).

Experimental work done in the late 1960's has influenced many of the concepts of modern protein folding. Cyrus Levinthal reasoned that a polypeptide chain of 100 amino acids would require a finite but very large amount of time to explore all possible conformations ($>3^{100}$), in order to reach its native state (*10*). In a random search of all

possible conformations, (see page 6 for Levinthal's calculation), with sampling of each state taking one pico-second, this 100 residue protein would take longer the lifespan of the universe to find the native state. This search is inconsistent with the observations that proteins fold rapidly, on the time scale of milli-seconds to minutes. Levinthal's paradox can be viewed as a flat free energy landscape (Figure 1.2) where a protein proceeds to its native state through an unguided search with low probability of reaching the its lowest energy state (*28*). Therefore proteins must fold through some type of defined pathway instead of a random search. Levinthal's work along with early studies on Ribonuclease A formed the "classical" view of protein folding in which proteins followed a specific pathway with distinct intermediate states (*29*).

MODELS OF PROTEIN FOLDING

Three classical models explaining how the linear amino acid sequence can coordinate and form its native structure came out of early studies done primarily on large monomeric proteins. The first model, nucleation-growth, describes the formation of tertiary structure through initial formation of secondary structure from a few residues. The local formation of a pocket, or nucleus, of secondary structure then quickly propagates to the final structure in downhill manner, with no defined intermediate species (*30;31*). This model was quickly dismissed because numerous studies indicated that intermediates were present in the folding of proteins (*32;33*). Two additional models that better explain the presence of intermediates in protein folding pathways are the framework or hierarchical model, and the hydrophobic collapse model. The framework model states that proteins fold by a stepwise formation of local secondary structure that

acts as a scaffold for further formation of secondary and tertiary structure (*32;34;35*). The stepwise formation of structure in this model occurs through diffusion-collision, where the secondary elements formed in early events diffuse and collide to promote formation of tertiary structure (*36*). The second model, the hydrophobic collapse model, starts with a initial rapid burial of hydrophobic sidechains followed by formation of secondary structure and packing to form a protein's tertiary structure (*37*).

It became more apparent, as experimental methods for detecting structure formation on shorter time-scales improved, that proteins fold through a combination of the two models, framework and hydrophobic collapse. A new view of protein folding was proposed with the discovery that intermediates were not required in the folding mechanisms of proteins such as chymotrypsin inhibitor 2 (CI2) (*38*). The most current view in protein folding is known as the nucleation-condensation model. This model incorporates both the framework and hydrophobic collapse models, in which secondary and tertiary interactions cause an initial collapse around a diffuse nucleus and then secondary and tertiary structure begin to form in a more stepwise manner as in the framework model (*39-41*).

THERMODYNAMIC VIEWS OF PROTEIN FOLDING

The "old" view of protein folding resulted from the limited availability of purified, well-behaved proteins. Recombinant DNA technology allowed for a massive increase in the number of proteins available for analysis and generation of smaller monomeric proteins and isolated domains to study. It was found that many small monomers (<100 residues) fold rapidly in a two-state kinetic manner with an unfolded

monomer (U) going to a folded monomer (N), U $\rightleftharpoons$ N, with no transient intermediates. Examples include: λ-repressor, cytochrome C, cold-shock protein B, and the Src SH3 domain. These monomeric, single domain proteins fold over a smooth energy surface (Figure 1.3A), which demonstrates that rapid, productive folding does not require kinetic intermediates (for review, (*42-44*)). In the classical view, protein folding is depicted by the reaction coordinate diagram shown in Figure 1.1, where $\Delta G^{\ddagger}$ is the free energy of activation or the amount of energy needed for the reaction to proceed. In contrast to these simple cases, larger multi-domain monomeric proteins have been shown to fold through complex pathways, with intermediates, multiple channels and rough energy surfaces in their folding reactions (*27;45*). From these studies and increasingly sophisticated theoretical/computational analyses (*46;47*), a "new" view of protein folding has emerged. Instead of a two-dimensional reaction coordinate diagram, the folding reaction can be viewed as free-energy landscape, described by a three-dimensional funnel depicting the large number of pathways leading to the native structure (*28;48*). The folding funnel shows the interplay between entropy, enthalpy and development of native contacts, and clearly depicts the roughness of the folding landscape and potential kinetic traps (Figure 1.3B). A key principle of the funnel view is that for a folding pathway to be kinetically favorable the overall slope of the energy landscape must be of sufficient steepness that the valleys flow toward the native structure. Another important aspect of the funnel is that as proteins proceed down the rough energy landscape, they will come to a thermodynamic bottleneck, the transition state, which consists of a large number of preferred pathways. The native-like states at this point are sensitive to sequence

mutations and proceed to discrete kinetic intermediates. The size of the bottleneck will dictate whether the search through kinetic intermediates becomes rate-limiting (*28*).

Although protein folding studies on monomers have provided many insights, there is still much unknown about how 2° and 3° structure formation is coordinated between multiple polypeptide chains in the formation of quaternary structure. This is of particular importance as many proteins rely on the formation of oligomeric structures for their function, and approximately one third of known cellular proteins are oligomers (*49*). Oligomerization is advantageous in a cell because it allows for a potential gain in function, an increase in regulation of protein function, increase in the local concentration of active sites, formation of new active sites, and larger binding surfaces to allow for increased protein interactions and networking. Oligomerization may also be advantageous through the enhancement of protein stability, as in the case of *T. maritima* DHFR (TmDHFR). This enzyme is the most thermostable DHFR identified thus far and is one of the rare DHFRs that forms a stable homodimer (*50*). Studying oligomeric systems may provide insights into protein evolution and is critical to fully understand the folding code.

ISSADD OLIGOMERS

Domain swapping describes one way in which protein oligomerization may have evolved and was first characterized from the crystal structure for diphtheria toxin (*51*). Domain (or segment) swapping was first defined in proteins that could populate both stable monomeric and dimeric forms where the domain of one monomer is exchanged with the identical domain of the other monomer (Figure 1.4) (*52*). Typically these

domain swaps occur with either the N-terminus or the C-terminus of a protein connected to the swapped domain by a hinge loop segment that varies in size, flexibility, and structure. A wider view of the swapped domain has been adopted as more structures have been determined (for review, (*53*)). Studies show that a swapped domain can be as small as one element of secondary structure, made up of several residues, or a large tertiary domain of hundreds of residues, and more than one domain can be swapped (for example, (*54-56*)).

Not only can domain swapping be advantageous to a cell as a mode for oligomerization, it can be potentially dangerous if not controlled, as in amyloid forming diseases (for review, (*57*)). The proteins studied in our lab belong to a class of domain swapped proteins referred to as the Intertwined Segment Swapped □-helical DNA-binding Dimers (both homo- and hetero-), or ISSADDs (for example, Figure 1.5). Examples important to this thesis include the prokaryotic transcription factor, FIS, and the eukaryotic core histones. The purpose of studying these proteins in our lab is to gain insight into the role of intermediates in protein folding, i.e. are they productive species or kinetic traps, as well as the relationships between folding landscapes and helical topology.

NUCLEOSOME CORE PARTICLE

The fundamental repeating unit of chromatin is the nucleosome core particle (NCP). 80% of the eukaryotic genome (*58*) is packaged by the NCP which compacts ~2 meters of DNA into a nucleus with a 6 □m diameter (*59*). The structure of the NCP is

critical for DNA compaction, as well as functioning as a key regulator of DNA-templated chemistries such as transcription, replication, and repair.

The nucleosome core particle consists of a ~150 bp piece of DNA wrapped around an octameric protein core consisting of 2 copies each of histones H2A, H2B, H3, and H4 (Figure 1.6). Histones H3 and H4 form a flat disc-shaped tetramer (two heterodimers) that binds to the central region of the 150 bp DNA to form a pseudo two-fold axis known as the dyad axis. Two H2A/H2B heterodimers then bind to the tetrameric core forming two ramps that complete the wrapping of the DNA around the octamer to form 1.65 superhelical turns. The large histone N-terminal tails extend out of the nucleosome, both over and between the coiled DNA.

The first low resolution (7 Å) crystal structure confirming the general structural characteristics of the NCP was published in 1984 (*60*). It was not until 1997 that the first high resolution structure (2.8 Å) was determined using recombinant histones, containing no post-translational modifications, and a defined homogeneous DNA sequence (*61*). Since then, many high resolution NCP crystal structures have been determined with alternative DNA sequences and different histone components—either from multiple organisms or mutant and natural histone variants (*61-68*).

The NCP assembly/disassembly pathway occurs in a step-wise manner both in the cell and in cell-free conditions (Figure 1.6). The interactions involved in nucleosome formation appear to be largely electrostatic and can be perturbed by the ionic strength of the solution. Reconstitution of the NCP from purified histones and DNA utilize this property by first combining the histones and DNA at a high salt concentration before

slowly reducing the ionic strength to first allow the association of the tetramer to the DNA followed by the binding of the H2A/H2B dimers to form the fully assembled NCP.

HISTONE STRUCTURE

The histones themselves are relatively small (~11-15 kDa), highly basic, DNA-binding proteins. Each histone contains a very basic N-terminal tail that extends from a globular domain containing three α-helices connected by two loops. The globular domain of each monomer dimerizes in an intertwined head-to-tail manner to form the "hand-shake motif" (*69*). The domain swapping that occurs in the eukaryotic core histones is generally referred to as quasi-domain swapped (*53*) since there are no stable monomeric forms. The hinge region common to domain swapped structures is not present in the histones and therefore no native-like monomeric structure can exist. This histone fold is generally referred to as a "ubiquitous architectural motif" that is commonly found in proteins involved in macromolecular DNA-protein assemblies (*70-73*). H2A and H2B both contain C-terminal tails; the H2A tail adopts an extended random structure, while that of H2B is helical. H3 and H4 fold to a similar heterodimeric structure, which oligomerizes to a tetramer through a four-helix bundle comprised of the C termini of each H3 protein.

CHROMATIN STRUCTURE AND REGULATION

The first level of chromatin structure is generally referred to as beads on a string where nucleosomes are lined up one after another with no interactions between them. In the presence of another histone commonly found associated with chromatin, histone H1,

the beads on a string can be further compacted into a fiber approximately 30 nm in diameter (*74*). How nucleosome arrays are arranged in the 30 nm fiber has been under debate. The two models that have been proposed are the one-start solenoid (linear coil of nucleosome arrays) and an irregular two-start helix. Recently the crystal structure of a tetranucleosome array had been determined and showed the linker DNA zig-zagging between two stacks of nucleosomes in a left-handed twisting manner (*75*) suggesting that the 30 nm fiber is a two-start helix.

In order for various processes involving DNA, such as transcription, replication and repair, to occur chromatin structure has evolved to be highly dynamic in nature. There are three main mechanisms for regulation of chromatin structure and dynamics that allow access of cellular machinery to compacted DNA. They are: 1) ATP-dependent chromatin remodeling (for review, (*76-79*)); 2) post-translational modifications (*80;81*); and 3) incorporation of histone variants (*82-86*).

As the name indicates, ATP-dependent chromatin remodeling couples ATP hydrolysis to the removal and relocation of histones from the DNA. The first chromatin remodeling complex was identified while characterizing the swi (switch) and snf (sucrose nonfermenting) mutations in yeast (*87*). The complex associated with these mutations was subsequently referred to as SWI/SNF. Other ATP-dependent chromatin remodelers include the SWR1 (SWI/SNF Related) and ISWI (Imitation SWI) complexes. Currently there are two proposed models for ATP catalyzed nucleosome mobility (*76*). The first model, the twisting model, occurs by disruption of the histone DNA-contacts at the entry sites of the linker DNA into the nucleosome (Figure 1.7B). This catalyzed distortion, or twisting, is transferred along the DNA wrapped around the nucleosome causing the

displacement of the histone octamer. The second model is the loop-recapture model, where the linker DNA is actually pulled into the nucleosome and causes a bulge or looped-out segment of DNA that can then be acted upon by various transcription factors (Figure 1.7C).

Another mechanism which cells use to alter chromatin states is post-translational modifications (PTMs). A myriad of histone modifying enzymes are known to covalently modify the histones proteins in complex patterns. Known modifications of histones include methylation, acetylation, phosphorylation, ADP ribosylation and ubiquination. The majority of the known PTMs, particularly methylation, acetylation, and phosphorylation, occur on the side-chains of the histone N-terminal tails (for review (*81;88;89*)). Although the tails are the most heavily studied sites for modifications, the globular histone cores have also been found to be modified in both the histone/DNA interface on the NCP lateral surface as well in locations not involved in protein-DNA contacts (*80*). These modifications, collectively referred as the histone code, can alter the charge state of the histones, which can alter the propensity of the nucleosome to translocate along the DNA, influence nucleosomal DNA accessibility and can recruit various nuclear proteins to the modified nucleosomes.

Changing the histone sequence by incorporation of natural variants can also alter nucleosome mobility and chromatin dynamics. There is a subset of mutations known as the SIN (Swi Independent) mutations that are located on the H3/H4 tetramer both near the dyad axis and the dimer-tetramer interface. These mutations were named for their ability to alleviate the need for SWI/SNF chromatin remodeling complexes in yeast (*90*). Although the SIN mutations were initially thought to create an altered NCP structural

state, it was found that there is only a slight increase in DNA accessibility at the entry/exits points of the NCP and no significant or detectable changes in NCP structure (*64*). It has been shown that these mutations actually lower the temperature required for thermal NCP redistribution. This may help alleviate the need for Swi/Snf by causing the NCPs to be more inherently mobile (*91*).

Most of the core histones have nonallelic variants that differ significantly in the amino acid sequence and are known to have distinct structural and functional roles when incorporated into NCPs. Many histone variants have now been identified, with H2A having highest number followed by histone H3. Histone variants for H2A include H2A.Z, H2A.X, H2AvD, H2A.Bbd, and macroH2A. H3 variants include H3.3, CenpA (also known as Cen3A). Little is known about H2B variants and no H4 variants have been identified. Variants may be used to alter the biophysical properties of the nucleosome or can mark specific regions of the genome. These variants can be found during various times in the cell cycle, as well as in different species (for review, (*82-86;92;93*). Post-translational modifications (PTMs), altering histone sequences and incorporations of histone variants are all ways to alter nucleosome dynamics and may overlap in function with ATP-dependent chromatin remodeling.

Of particular interest to this study is the variant H2A.Z, which has been extensively studied and has been found to be involved in repressive transcriptional control. Despite extensive study, the role of H2A.Z in chromatin dynamics is still under debate. In protists such as *Tetrahymena*, H2A.Z (H2Ahv1) is preferentially located in the transcriptionally active macronucleus and not in the silent micronucleus (*94*). H2A.Z (Htz1) in budding yeast (*S. cerevisiae*) was first implicated in the activation of Pho5 and

Gal1 promoters (*95*) and the silencing of the HMR locus and telomeres (*96*). Later, H2A.Z was found to mark euchromatin boundaries and prevent the migration of neighboring heterochromatin (*97*). H2A.Z enrichment in these intergenic regions influenced histone modifications and chromatin remodeling (*98*). The *Drosophila* homolog (H2Av) is found at both transcriptionally active regions and transcriptionally inactive loci within heterochromatin (*99*). When H2A.Z is present during early mammalian development, it is targeted to pericentric heterochromatin; H2A.Z appears to provide a signal to differentiate between constitutive and facultative heterochromatin (*100*). H2A.Z has also been implicated in chromosome segregation, given its co-localization with the silencing protein HP1alpha in mammalian cell lines, and H2A.Z depletion leads to high levels of genome instability (*101*). Genomic instability was evident by multiple, irregularly shaped nuclei in the same cell, as well as chromatin bridges and lagging chromosomes between separating nuclei. The *Drosophila* homolog of H2A.Z (H2Av) has also been shown to be critical in the recruitment of HP1alpha homolog (HP1) in the establishment of centromeric heterochromatin (*102*). HP1alpha interacts with the altered acidic patch in H2A.Z-containing NCPs and generates highly compacted constitutive heterochromatin (*103*) which is important in mammalian chromosome segregation.

Several biophysical studies show that H2A.Z may alter nucleosome stability (*104-106*), although whether it stabilizes, destabilizes, or causes no change at all, is yet to be resolved. Clearly, further studies are needed to decipher the functional role(s) and biophysical effects of H2A.Z on chromatin. Consideration of the context, such as types of histones used (i.e. chicken, mouse, recombinant, etc.), DNA sequence, post-

translational modifications and chromosomal location may be extremely important in elucidating the multiple facets of H2A.Z function. Multiple and more specific probes are required to monitor protein-protein interactions in the NCP than previously used. These probes need to be able to monitor the dissociation of the H2A/H2B dimers but not detect any non-specific interactions that may occur between the dissociated dimers and the DNA (*107*). Furthermore, the effects of incorporating the probes on the stability, and therefore structure, of the histones need to be carefully determined to aid in the interpretation of the system. Chapter three of this thesis will describe a multiple protein-protein FRET system that provides further insights into the equilibrium dissociation of histone dimers (WT and H2A.Z) from the NCP.

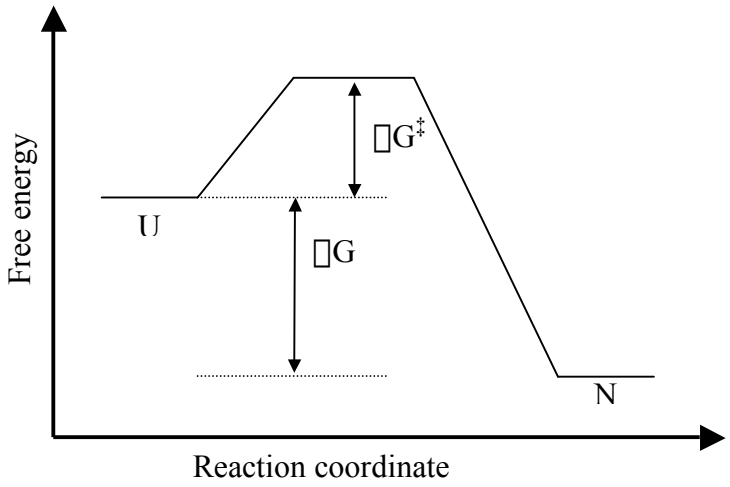


Figure 1.1. Representative reaction coordinate diagram showing the free energy differences between native (N) and unfolded (U) species. The overall free energy change, ΔG, is the difference between the unfolded and folded species. The difference in energy between the unfolded and transition state is known as the free energy of activation, $\Delta G^{\ddagger}$.

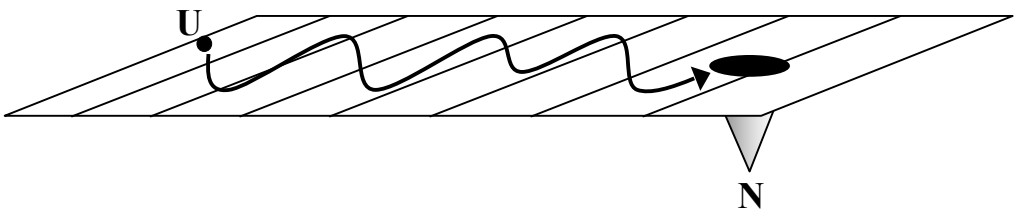


Figure 1.2. Levinthal's paradox as a free energy landscape of protein folding. The unfolded protein (U) proceeds unguided through a flat energy landscape with a single energy minima representing the native state (N). The search is random with a very low probability of reaching the native state.

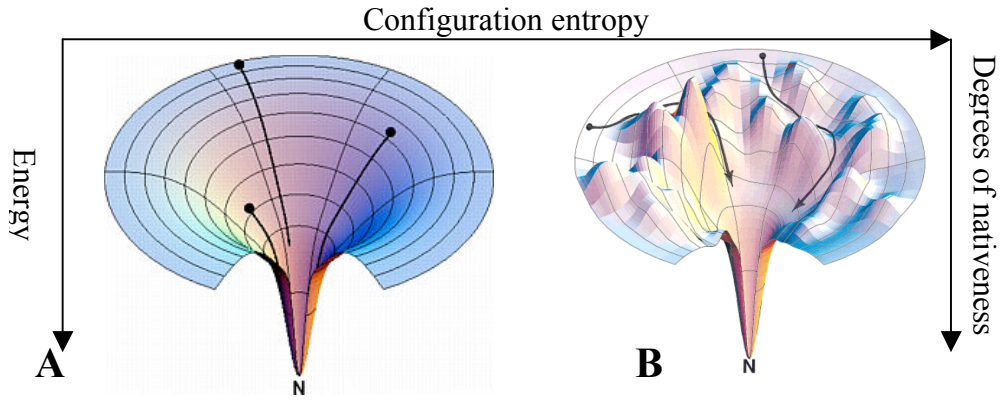


Figure 1.3. Three-dimensional representations of the new view of free energy landscapes. **A.** Single, well-defined funnel representative of fast folding single domain proteins. All unfolded protein conformations proceed through a downhill organization of partially folded ensembles that lead to the native folded structure. **B.** Rugged free energy landscape reflecting the transient trapping of configurations through local energy minima (depicted by the arrows). Pictures from H.S. Chan and K.A. Dill Nature Structural Biology, 1997 (*108*).

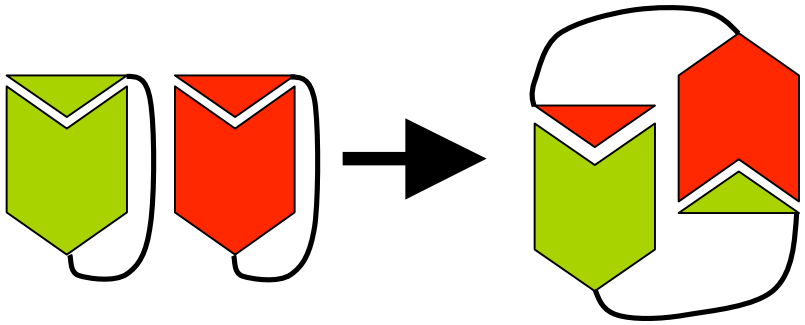


Figure 1.4. Cartoon depicting domain swapping between two "closed" monomers. The domain of one monomer makes the same contacts with the other monomer, and vice versa, forming the domain swapped dimer. The swapped domain can be as small as several amino acids or as large as entire tertiary globular domain.

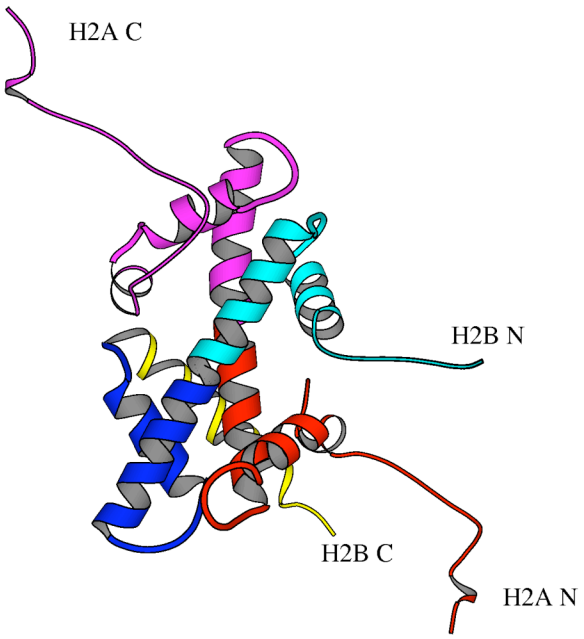


Figure 1.5. Ribbon diagram of the eukaryotic H2A/H2B histone dimer from the X-ray crystal structure of the nucleosome (*61*). H2A is red to magenta, and H2B is cyan to blue with the C-terminal helix in yellow. The domain swapped structure is depicted by the cyan N-terminal "domain" of H2B making contacts with the magenta C-terminal "domain" of H2A. These are identical to those made by the red to the blue interactions of the other half of the dimer. The diagram was generated with Molscript v2.0 (*109*).

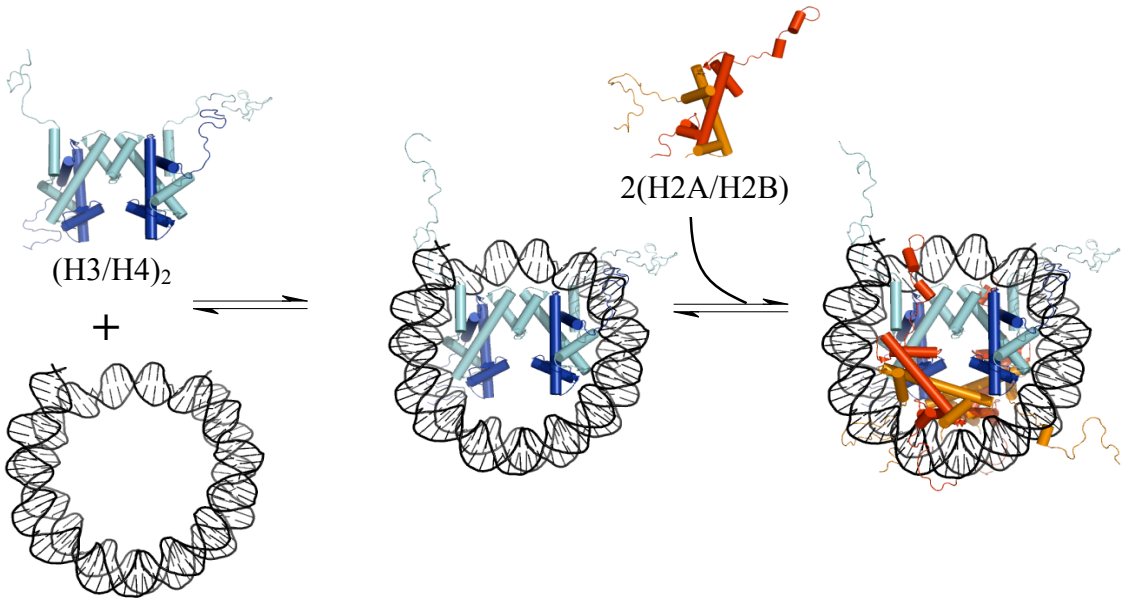


Figure 1.6. Schematic of the assembly/disassembly pathway of the nucleosome core particle. Although the full pathway is not reversible in the conditions used, the association of the H2A-H2B dimer is reversible. Structures were rendered in Pymol (*110*).

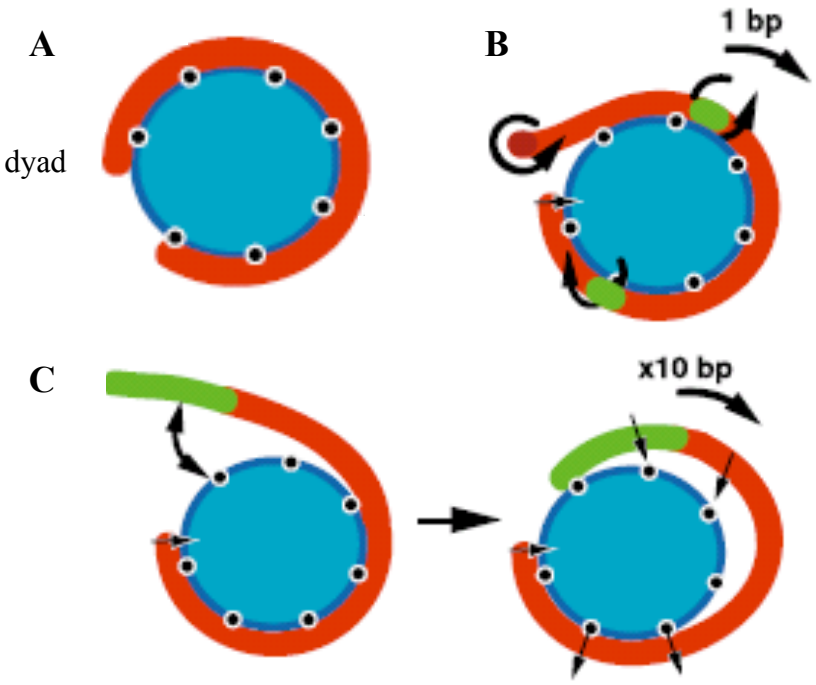


Figure 1.7. Models for chromatin remodeling. **A.** Cartoon of nucleosome core particle showing the locations of the protein-DNA contacts (black circles). **B.** The twisting model; a twist in the linker DNA gets propagated throughout the NCP. **C.** Loop recapture model; a segment of DNA becomes exposed at the exit of the NCP and make new histone-DNA contacts that create a "loop". This loop may be bound by various DNA binding proteins (transcription factors or chromatin remodeling complexes) or continue to make new histone-DNA contacts that cause the octamer to slide further upstream or downstream. Figure from Langst and Becker 2004 (*111*).

.

Literature Cited

1. Sandgren, M., Stahlberg, J., and Mitchinson, C. (2005) Structural and biochemical studies of GH family 12 cellulases: improved thermal stability, and ligand complexes, *Prog. Biophys. Mol. Biol. 89*, 246-291.

2. Kaur, J. and Sharma, R. (2006) Directed evolution: an approach to engineer enzymes, *Crit Rev. Biotechnol. 26*, 165-199.

3. Arnold, F. H. and Moore, J. C. (1997) Optimizing industrial enzymes by directed evolution, *Adv. Biochem. Eng Biotechnol. 58*, 1-14.

4. Moore, J. C. and Arnold, F. H. (1996) Directed evolution of a para-nitrobenzyl esterase for aqueous-organic solvents, *Nat. Biotechnol. 14*, 458-467.

5. DeFelippis, M. R., Chance, R. E., and Frank, B. H. (2001) Insulin self-association and the relationship to pharmacokinetics and pharmacodynamics, *Crit Rev. Ther. Drug Carrier Syst. 18*, 201-264.

6. Boggetto, N. and Reboud-Ravaux, M. (2002) Dimerization inhibitors of HIV-1 protease, *Biol. Chem. 383*, 1321-1324.

7. Kohl, N. E., Emini, E. A., Schleif, W. A., Davis, L. J., Heimbach, J. C., Dixon, R. A., Scolnick, E. M., and Sigal, I. S. (1988) Active human immunodeficiency virus protease is required for viral infectivity, *Proc. Natl. Acad. Sci. U. S. A 85*, 4686-4690.

8. Voet, D. and Voet, J. G. (1995) *Biochemistry* John Wiley & Sons, Inc., New York.

9. Fersht, A. R. (1999) *Structure and mechanism in protein science: A guide to enzyme catalysis and protein folding.* W.H. Freeman & Co., New York.

10. Levinthal, C. (1968) Are there pathways for protein folding?, *J. Chim. Phys. 65*, 44-45.

11. Dill, K. A. (1990) Dominant forces in protein folding, *Biochemistry 29*, 7133-7155.

12. Font, J., Benito, A., Lange, R., Ribo, M., and Vilanova, M. (2006) The contribution of the residues from the main hydrophobic core of ribonuclease A to its pressure-folding transition state, *Protein Sci. 15*, 1000-1009.

13. Ribo, M., Font, J., Benito, A., Torrent, J., Lange, R., and Vilanova, M. (2006) Pressure as a tool to study protein-unfolding/refolding processes: the case of ribonuclease A, *Biochim. Biophys. Acta 1764*, 461-469.

14. Weber, G. and Drickamer, H. G. (1983) The effect of high pressure upon proteins and other biomolecules, *Q. Rev. Biophys. 16*, 89-112.

15. Perrett, S. and Zhou, J. M. (2002) Expanding the pressure technique: insights into protein folding from combined use of pressure and chemical denaturants, *Biochim. Biophys. Acta 1595*, 210-223.

16. Monera, O. D., Kay, C. M., and Hodges, R. S. (1994) Protein denaturation with guanidine hydrochloride or urea provides a different estimate of stability depending on the contributions of electrostatic interactions, *Protein Sci. 3*, 1984-1991.

17. Eyles, S. J. and Kaltashov, I. A. (2004) Methods to study protein dynamics and folding by mass spectrometry, *Methods 34*, 88-99.

18. Last, A. M. and Robinson, C. V. (1999) Protein folding and interactions revealed by mass spectrometry, *Curr. Opin. Chem. Biol. 3*, 564-570.

19. Fabian, H. and Naumann, D. (2004) Methods to study protein folding by stopped-flow FT-IR, *Methods 34*, 28-40.

20. Jackson, M. and Mantsch, H. H. (1995) The use and misuse of FTIR spectroscopy in the determination of protein structure, *Crit Rev. Biochem. Mol. Biol. 30*, 95-120.

21. Krimm, S. and Bandekar, J. (1986) Vibrational spectroscopy and conformation of peptides, polypeptides, and proteins, *Adv. Protein Chem. 38*, 181-364.

22. *Circular Dichroism: Principles and Applications* (2000) Wiley-VCH, Inc., New York.

23. Lakowicz, J. R. (1999) Principles of Fluorescence Spectroscopy, 2nd ed., pp 367-394, Kluwer Academic/Plenum Publishers, New York.

24. Wu, P. and Brand, L. (1994) Resonance energy transfer: methods and applications, *Anal. Biochem. 218*, 1-13.

25. Haber, E. and Anfinsen, C. B. (1962) Side-chain interactions governing the pairing of half-cystine residues in ribonuclease, *J. Biol. Chem. 237*, 1839-1844.

26. Jaenicke, R. and Lilie, H. (2000) Folding and association of oligomeric and multimeric proteins, *Adv Protein Chem 53*, 329-401.

27. Matthews, C. R. (1993) Pathways of protein folding, *Annual Review of Biochemistry 62*, 653-683.

28. Wolynes, P. G., Onuchic, J. N., and Thirumalai, D. (1995) Navigating the folding routes, *Science 267*, 1619-20.

29. Tsong, T. Y., Baldwin, R. L., and Elson, E. L. (1972) Properties of the refolding and unfolding reactions of ribonuclease A, *Proc. Natl. Acad. Sci. U. S. A 69*, 1809-1812.

30. Wetlaufer, D. B. (1973) Nucleation, rapid folding, and globular intrachain regions in proteins, *Proc. Natl. Acad. Sci. U. S. A 70*, 697-701.

31. Wetlaufer, D. B. (1990) Nucleation in protein folding--confusion of structure and process, *Trends Biochem. Sci. 15*, 414-415.

32. Kim, P. S. and Baldwin, R. L. (1990) Intermediates in the folding reactions of small proteins, *Annu. Rev. Biochem. 59*, 631-660.

33. Ptitsyn, O. B. (1994) Kinetic and equilibrium intermediates in protein folding, *Protein Eng 7*, 593-6.

34. Baldwin, R. L. and Rose, G. D. (1999) Is protein folding hierarchic? I. Local structure and peptide folding, *Trends Biochem Sci 24*, 26-33.

35. Baldwin, R. L. and Rose, G. D. (1999) Is protein folding hierarchic? II. Folding intermediates and transition states, *Trends Biochem Sci 24*, 77-83.

36. Karplus, M. and Weaver, D. L. (1994) Protein folding dynamics: the diffusion-collision model and experimental data, *Protein Sci 3*, 650-68.

37. Ptitsyn, O. (1996) How molten is the molten globule?, *Nature Structural Biology 3*, 488-490.

38. Jackson, S. E. and Fersht, A. R. (1991) Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition, *Biochemistry 30*, 10428-35.

39. Daggett, V. and Fersht, A. (2003) The present view of the mechanism of protein folding, *Nat Rev Mol Cell Biol 4*, 497-502.

40. Fersht, A. R. (1997) Nucleation mechanisms in protein folding, *Curr. Opin. Struct. Biol. 7*, 3-9.

41. Fersht, A. R. (2000) Transition-state structure as a unifying basis in protein-folding mechanisms: contact order, chain topology, stability, and the extended nucleus mechanism, *Proc Natl Acad Sci U S A 97*, 1525-9.

42. Jackson, S. E. (1998) How do small single-domain proteins fold?, *Fold Des 3*, 81-91.

43. Plaxco, K. W., Simons, K. T., Ruczinski, I., and Baker, D. (2000) Topology, stability, sequence, and length: defining the determinants of two-state protein folding kinetics, *Biochemistry 39*, 11177-83.

44. Sosnick, T. R., Mayne, L., Hiller, R., and Englander, S. W. (1994) The barriers in protein folding, *Nat. Struct. Biol. 1*, 149-156.

45. Bilsel, O. and Matthews, C. R. (2000) Barriers in protein folding reactions, *Adv Protein Chem 53*, 153-207.

46. Clementi, C., Nymeyer, H., and Onuchic, J. N. (2000) Topological and energetic factors: what determines the structural details of the transition state ensemble and "en-route" intermediates for protein folding? An investigation for small globular proteins, *J Mol Biol 298*, 937-53.

47. Onuchic, J. N. and Wolynes, P. G. (2004) Theory of protein folding, *Curr Opin Struct Biol 14*, 70-75.

48. Lazaridis, T. and Karplus, M. (1997) "New view" of protein folding reconciled with the old through multiple unfolding simulations, *Science 278*, 1928-31.

49. Goodsell, D. S. and Olson, A. J. (2000) Structural symmetry and protein function, *Annu. Rev. Biophys. Biomol. Struct. 29*, 105-153.

50. Dams, T., Bohm, G., Auerbach, G., Bader, G., Schurig, H., and Jaenicke, R. (1998) Homo-dimeric recombinant dihydrofolate reductase from Thermotoga maritima shows extreme intrinsic stability, *Biol. Chem. 379*, 367-371.

51. Bennett, M. J., Choe, S., and Eisenberg, D. (1994) Refined structure of dimeric diphtheria toxin at 2.0 A resolution, *Protein Sci. 3*, 1444-1463.

52. Bennett, M. J., Schlunegger, M. P., and Eisenberg, D. (1995) 3D domain swapping: a mechanism for oligomer assembly, *Protein Sci 4*, 2455-68.

53. Liu, Y. and Eisenberg, D. (2002) 3D domain swapping: As domains continue to swap, *Protein Sci 11*, 1285-99.

54. Gotte, G., Bertoldi, M., and Libonati, M. (1999) Structural versatility of bovine ribonuclease A. Distinct conformers of trimeric and tetrameric aggregates of the enzyme, *Eur. J. Biochem. 265*, 680-687.

55. Liu, Y., Gotte, G., Libonati, M., and Eisenberg, D. (2001) A domain-swapped RNase A dimer with implications for amyloid formation, *Nat. Struct. Biol. 8*, 211-214.

56. Liu, Y., Gotte, G., Libonati, M., and Eisenberg, D. (2002) Structures of the two 3D domain-swapped RNase A trimers, *Protein Sci. 11*, 371-380.

57. Bennett, M. J., Sawaya, M. R., and Eisenberg, D. (2006) Deposition diseases and 3D domain swapping, *Structure. 14*, 811-824.

58. Noll, M. (1974) Subunit structure of chromatin, *Nature 251*, 249-251.

59. Alberts, B., Johnson, A., Lewis J., Raff, M., Roberts, K., and Walter, P. (2002) Molecular Biology of the Cell, 4th ed., Garland Science.

60. Richmond, T. J., Finch, J. T., Rushton, B., Rhodes, D., and Klug, A. (1984) Structure of the nucleosome core particle at 7 Å resolution, *Nature 311*, 532-537.

61. Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F., and Richmond, T. J. (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution, *Nature 389*, 251-260.

62. Chakravarthy, S. and Luger, K. (2006) The histone variant macro-H2A preferentially forms "hybrid nucleosomes", *J. Biol. Chem. 281*, 25522-25531.

63. Davey, C. A., Sargent, D. F., Luger, K., Maeder, A. W., and Richmond, T. J. (2002) Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution, *J Mol Biol 319*, 1097-1113.

64. Muthurajan, U. M., Bao, Y., Forsberg, L. J., Edayathumangalam, R. S., Dyer, P. N., White, C. L., and Luger, K. (2004) Crystal structures of histone Sin mutant nucleosomes reveal altered protein-DNA interactions, *EMBO J 23*, 260-271.

65. Richmond, T. J. and Davey, C. A. (2003) The structure of DNA in the nucleosome core, *Nature 423*, 145-150.

66. Schalch, T., Duda, S., Sargent, D. F., and Richmond, T. J. (2005) X-ray structure of a tetranucleosome and its implications for the chromatin fibre, *Nature 436*, 138-141.

67. Suto, R. K., Clarkson, M. J., Tremethick, D. J., and Luger, K. (2000) Crystal structure of a nucleosome core particle containing the variant histone H2A.Z, *Nat. Struct. Biol. 7*, 1121-1124.

68. White, C. L., Suto, R. K., and Luger, K. (2001) Structure of the yeast nucleosome core particle reveals fundamental changes in internucleosome interactions, *EMBO J 20, 5207-5218*.

69. Arents, G., Burlingame, R. W., Wang, B. C., Love, W. E., and Moudrianakis, E. N. (1991) The nucleosomal core histone octamer at 3.1 Å resolution: a tripartite protein assembly and a left-handed superhelix., *Proc. Natl. Acad. Sci. U. S. A. 88*, 10148-52.

70. Arents, G. and Moudrianakis, E. N. (1995) The histone fold: A ubiquitous architectural motif utilized in DNA compaction and protein dimerization, *Proc. Natl. Acad. Sci. USA 92*, 11170-11174.

71. Baxevanis, A. D., Arents, G., Moudrianakis, E. N., and Landsman, D. (1995) A variety of DNA-binding and multimeric proteins contain the histone fold motif, *Nucl. Acids Res. 23*, 2685-2691.

72. Birck, C., Poch, O., Romier, C., Ruff, M., Mengus, G., Lavigne, A. C., Davidson, I., and Moras, D. (1998) Human $TAF_{II}28$ and $TAF_{II}18$ interact through a histone fold encoded by atypical evolutionary conserved motifs also found in the SPT3 family., *Cell 94*, 239-249.

73. Xie, X., Kokubo, T., Cohen, S. L., Mirza, U., Hoffmann, A., Chait, B. T., Roeder, R. G., Nakatani, Y., and Burley, S. K. (1996) Structural similarity between TAFs and the heterotetrameric core of the histone octamer, *Nature 380*, 316-322.

74. Hansen, J. C. and Wolffe, A. P. (1994) A role for histones H2A/H2B in chromatin folding and transcriptional repression., *Proc. Natl. Acad. Sci. U. S. A. 91*, 2339-2343.

75. Dorigo, B., Schalch, T., Kulangara, A., Duda, S., Schroeder, R. R., and Richmond, T. J. (2004) Nucleosome arrays reveal the two-start organization of the chromatin fiber, *Science 306*, 1571-1573.

76. Becker, P. B. and Horz, W. (2002) ATP-dependent nucleosome remodeling, *Annu. Rev. Biochem. 71*, 247-273.

77. Flaus, A. and Owen-Hughes, T. (2004) Mechanisms for ATP-dependent chromatin remodelling: farewell to the tuna-can octamer?, *Curr. Opin. Genet. Dev. 14*, 165-173.

78. Korber, P. and Horz, W. (2004) SWRred not shaken; mixing the histones, *Cell 117*, 5-7.

79. Smith, C. L. and Peterson, C. L. (2005) ATP-dependent chromatin remodeling, *Curr. Top. Dev. Biol. 65*, 115-148.

80. Cosgrove, M. S., Boeke, J. D., and Wolberger, C. (2004) Regulated nucleosome mobility and the histone code, *Nat. Struct. Mol. Biol. 11*, 1037-1043.

81. Jenuwein, T. and Allis, C. D. (2001) Translating the histone code, *Science 293*, 1074-1080.

82. Ahmad, K. and Henikoff, S. (2002) Histone H3 variants specify modes of chromatin assembly, *Proc Natl Acad Sci U S A 99 Suppl 4*, 16477-16484.

83. Ausio, J., Abbott, D. W., Wang, X., and Moore, S. C. (2001) Histone variants and histone modifications: a structural perspective, *Biochem. Cell Biol. 79*, 693-708.

84. Ausio, J. and Abbott, D. W. (2002) The many tales of a tail: carboxyl-terminal tail heterogeneity specializes histone H2A variants for defined chromatin function, *Biochemistry 41*, 5945-5949.

85. Henikoff, S., Furuyama, T., and Ahmad, K. (2004) Histone variants, nucleosome assembly and epigenetic inheritance, *Trends Genet 20*, 320-326.

86. Kamakaka, R. T. and Biggins, S. (2005) Histone variants: deviants?, *Genes Dev 19*, 295-310.

87. Vignali, M., Hassan, A. H., Neely, K. E., and Workman, J. L. (2000) ATP-dependent chromatin-remodeling complexes, *Mol. Cell Biol. 20*, 1899-1910.

88. Strahl, B. D. and Allis, C. D. (2000) The language of covalent histone modifications, *Nature 403*, 41-45.

89. Verger, A. and Crossley, M. (2004) Chromatin modifiers in transcription and DNA repair, *Cell Mol Life Sci 61*, 2154-2162.

90. Kruger, W., Peterson, C. L., Sil, A., Coburn, C., Arents, G., Moudrianakis, E. N., and Herskowitz, I. (1995) Amino acid substitutions in the structured domains of histones H3 and H4 partially relieve the requirement of the yeast SWI/SNF complex for transcription., *Genes Dev. 9*, 2770-9.

91. Flaus, A., Rencurel, C., Ferreira, H., Wiechens, N., and Owen-Hughes, T. (2004) Sin mutations alter inherent nucleosome mobility, *Embo J 23*, 343-353.

92. Redon, C., Pilch, D., Rogakou, E., Sedelnikova, O., Newrock, K., and Bonner, W. (2002) Histone H2A variants H2AX and H2AZ, *Curr Opin Genet Dev 12*, 162-169.

93. Gloss, L. M. and Placek, B. J. (2002) The effect of salts on the stability of the H2A-H2B histone dimer, *Biochemistry 41*, 14960-14968.

94. Stargell, L. A., Bowen, J., Dadd, C. A., Dedon, P. C., Davis, M., Cook, R. G., Allis, C. D., and Gorovsky, M. A. (1993) Temporal and spatial association of histone H2A variant hv1 with transcriptionally competent chromatin during nuclear development in Tetrahymena thermophila, *Genes Dev. 7*, 2641-2651.

95. Santisteban, M. S., Kalashnikova, T., and Smith, M. M. (2000) Histone H2A.Z regulats transcription and is partially redundant with nucleosome remodeling complexes, *Cell 103*, 411-422.

96. Dhillon, N. and Kamakaka, R. T. (2000) A histone variant, Htz1p, and a Sir1p-like protein, Esc2p, mediate silencing at HMR, *Mol. Cell 6*, 769-780.

97. Meneghini, M. D., Wu, M., and Madhani, H. D. (2003) Conserved histone variant H2A.Z protects euchromatin from the ectopic spread of silent heterochromatin, *Cell 112*, 725-736.

98. Li, B., Pattenden, S. G., Lee, D., Gutierrez, J., Chen, J., Seidel, C., Gerton, J., and Workman, J. L. (2005) Preferential occupancy of histone variant H2AZ at inactive promoters influences local histone modifications and chromatin remodeling, *Proc. Natl. Acad. Sci. U. S. A 102*, 18385-18390.

99. Leach, T. J., Mazzeo, M., Chotkowski, H. L., Madigan, J. P., Wotring, M. G., and Glaser, R. L. (2000) Histone H2A.Z is widely but nonrandomly distributed in chromosomes of Drosophila melanogaster, *J. Biol. Chem. 275*, 23267-23272.

100. Rangasamy, D., Berven, L., Ridgway, P., and Tremethick, D. J. (2003) Pericentric heterochromatin becomes enriched with H2A.Z during early mammalian development, *EMBO J. 22*, 1599-1607.

101. Rangasamy, D., Greaves, I., and Tremethick, D. J. (2004) RNA interference demonstrates a novel role for H2A.Z in chromosome segregation, *Nat. Struct. Mol. Biol. 11*, 650-655.

102. Swaminathan, J., Baxter, E. M., and Corces, V. G. (2005) The role of histone H2Av variant replacement and histone H4 acetylation in the establishment of Drosophila heterochromatin, *Genes Dev. 19*, 65-76.

103. Fan, J. Y., Rangasamy, D., Luger, K., and Tremethick, D. J. (2004) H2A.Z alters the nucleosome surface to promote HP1alpha-mediated chromatin fiber folding, *Mol. Cell 16*, 655-661.

104. Abbott, D. W., Ivanova, V. S., Wang, X., Bonner, W. M., and Ausio, J. (2001) Characterization of the stability and folding of H2A.Z chromatin particles: implications for transcriptional activation, *J. Biol. Chem. 276*, 41945-41949.

105. Park, Y. J., Dyer, P. N., Tremethick, D. J., and Luger, K. (2004) A new fluorescence resonance energy transfer approach demonstrates that the histone variant H2AZ stabilizes the histone octamer within the nucleosome, *J Biol Chem 279*, 24274-24282.

106. Thambirajah, A. A., Dryhurst, D., Ishibashi, T., Li, A., Maffey, A. H., and Ausio, J. (2006) H2A.Z stabilizes chromatin in a way that is dependent on core histone acetylation, *J. Biol. Chem. 281*, 20036-20044.

107. Aragay, A. M., Diaz, P., and Daban, J.-R. (1988) Association of nucleosome core particle DNA with different histone oligomers: Transfer of histones between DNA-(H2A,H2B) and DNA-(H3,H4) complexes., *J. Mol. Biol. 204*, 141-154.

108. Dill, K. A. and Chan, H. S. (1997) From Levinthal to pathways to funnels, *Nat. Struct. Biol. 4,* 10-19.

109. Kraulis, P. J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures., *J. Applied Crystallography 24*, 946-950.

110. Delano, W. L. (2002) The PYMOL Molecular Graphics System, Delano Scientific, San Carlos, CA USA. http://www.pymol.org.

111. Langst, G. and Becker, P. B. (2004) Nucleosome remodeling: one mechanism, many phenomena?, *Biochim. Biophys. Acta 1677*, 58-63.

CHAPTER II

EQUILIBRIUM AND KINETIC STUDIES OF FIS, FACTOR FOR INVERSION STIMULATION

The work in this chapter was published in the Topping, T.B., Hoch, D.A. and Gloss, L.M. *Journal of Molecular Biology* (2004). **335**: 1065-81 (Cited throughout as reference 17 in bibliography). I am co-first author of this paper with Traci B. Topping who originally started this work and performed many of the thermodynamic and kinetic studies in urea. To aid in further understanding this protein and to help compare it to other systems, I performed the thermodynamic and kinetic experiments in guanidinium chloride at 10°C and the double jump kinetic experiments. This chapter is modified from the original paper to focus on the data I collected.

SUMMARY

FIS, the Factor for Inversion Stimulation, from *Escherichia coli* and other enteric bacteria, is an intertwined α-helical homodimer. The folding and unfolding of FIS were studied with both equilibrium and kinetic methods by circular dichroism using urea and guanidinium chloride (GdmCl) as the perturbants. The equilibrium folding is reversible and well-described by a two-state folding model, with stabilities at 10 °C of 15.2 kcal mol^{-1} in urea and 13.5 kcal mol^{-1} in GdmCl. The kinetic data are consistent with a two-step folding reaction where the two unfolded monomers associate to a dimeric intermediate within the mixing time for the stopped-flow instrument (<5 ms), and a slower, subsequent folding of the dimeric intermediate to the native dimer. Fits of the burst phase amplitudes as a function of denaturant showed that the free energy for the formation of the dimeric intermediate constitutes the majority of the stability of the folding (9.6 kcal mol^{-1} in urea and 10.5 kcal mol^{-1} in GdmCl). Folding-to-unfolding double jump kinetic experiments were also performed to monitor the formation of native dimer as a function of folding delay times. The data here demonstrate that the dimeric intermediate is obligatory and on-pathway. The folding mechanism of FIS, when compared to other intertwined, α-helical, homodimers, suggests that a transient kinetic dimeric intermediate may be a common feature of the folding of intertwined, segment-swapped, α-helical dimers.

ABBREVIATIONS

BCA, bicinchoninic acid; CD, circular dichroism; ΔASA, change in solvent-accessible surface area between two species; $\Delta G^{o}_{(H2O)}$, the free energy of unfolding in the absence of denaturant; Fapp; apparent fraction of unfolded monomer; FIS, E. coli factor for inversion stimulation; FL, fluorescence; GdmCl, guanidinium chloride; I_2, dimeric folding intermediate; KPi, potassium phosphate, pH 7.2; *m*-value, parameter describing the sensitivity of the unfolding transition to the [denaturant]; N_2, native dimer; SF, stopped-flow; 2U, two unfolded, disassociated monomers; WT, wild-type.

INTRODUCTION

The goal of protein folding is to understand the second half of the genetic code: how a linear sequence of amino acids encodes the three-dimensional structure of a protein as well as its equilibrium stability and the kinetic mechanism for the rapid, efficient transformation from a random coil, unfolded state into the native, functional form. Folding studies on monomeric proteins have provided many essential insights into the features of the collapse and assembly of a single polypeptide chain during the folding process (*1-3*). The assembly and function of oligomeric proteins is an important extension of efforts to understand protein folding. The folding of oligomers requires the coordination of secondary and tertiary structures with the docking and assembly of quaternary structures involving multiple polypeptide chains. Therefore, the folding code of a dimeric protein is contained in the sequence of not one, but two polypeptide chains. Currently, the understanding is incomplete of the relationships between folding mechanism and features of the oligomerization interface, such as secondary structure content and interface topology. Here we begin to address these relationships for the folding of dimers with intertwined, segment swapped, α-helical interfaces, using the *Escherichia coli* Factor for Inversion Stimulation (FIS) as a model system.

FIS is a small (98 residues per monomer), homodimeric, nucleoid protein from enteric bacteria. The protein was first identified because of its role in regulation of DNA inversion by the Hin-family of invertases (*4-6*). FIS is now known to function in the regulation of many DNA-based processes, including transcriptional activation and repression of a variety of genes. The three-dimensional structures of WT FIS and several mutants have been determined by X-ray crystallography (*7*). The most N-terminal

regions observed in the crystal structure form mobile β-hairpin arms (Figure 2.1), which interact with the DNA recombinase. The polypeptide fold then proceeds into a domain comprised of two relatively long α-helices, which dock in an anti-parallel orientation with the corresponding helices from the other monomer. The majority of the dimerization interface, which is largely hydrophobic in nature, is comprised of these four helices. The two C-terminal helices of each monomer form the helix-turn-helix DNA-binding domains. The intertwined subunit structure of FIS has been described as segment swapped (*8*), a domain swapped structure (*9*) in which segments (not necessarily compact, independent folding domains) are exchanged between monomers. Comparison of the folding mechanism of FIS with that of other dimers with predominantly α-helical interfaces may begin to elucidate some generalities in the relationships between folding landscapes and helical topology.

MATERIALS AND METHODS

Materials.

Ultra-pure urea and guanidinium chloride were purchased from ICN Biomedicals (Costa Mesa, CA). All other chemicals were of reagent or molecular biology grade.

Methods.

Protein overexpression. FIS was over-expressed, using the plasmid pRJ1077 (*10*), a pET11a derivative containing the gene for WT FIS, in the *E. coli* strain BL21(DE3) pLysS. FIS was purified by methods described elsewhere (*10*), with a slight modification of the brand of column resin, and using micro-crystallization as the final purification step. The homogeneity of the purified protein was assessed by SDS-PAGE,

and judged to be ≥ 98% pure. Protein concentration was estimated using an extinction coefficient at 280 nm of 4,790 $M^{-1}cm^{-1}$. This value was determined from the absorbance of samples whose protein concentrations had been determined by BCA assay kits (Sigma, St. Louis, MO).

Data collection. Equilibrium and stopped-flow circular dichroism data were collected on an AVIV 202SF spectrophotometer. Equilibrium data collection employed a Hamilton Model 500 automated titrator interfaced to the CD. SF experiments used an AVIV stopped-flow tower.

Buffer conditions for the folding experiments were 200 mM KCl, 20 mM KPi, pH 7.2; 1 mM EDTA was included in most experiments. EDTA was not included in the buffer for the CD spectra shown in Figure 2.2, to prevent interference below ~210 nm. Control equilibrium titrations demonstrated that there was no difference in stability in the presence or absence of EDTA. For the equilibrium data at 10 °C, wavelength spectra from 260 to 210 nm were collected at each denaturant concentration. The spectra were analyzed using singular value decomposition (*11*) performed with a subroutine of the fitting program Savuka 5.1. The application of SVD to equilibrium protein folding data has been described elsewhere (for example, (*12*)). SVD analyses yield the denaturant dependence of a series of basis vectors that, in total, describe the concerted response of all wavelengths to changing denaturant concentrations. The principal advantages of SVD data are: 1) better signal-to-noise relative to data at a single wavelength; and 2) ability to analyze data at multiple wavelengths to enhance either the detection of or confirm the absence of equilibrium intermediates. In general, only one or two basis vectors were strongly weighted in the SVD analyses, and used in the global fitting to determine the

free energy of unfolding. The equilibration time for the automated titrations was 3 min, which is more than two orders of magnitude longer than the slowest observed folding or unfolding kinetics.

For SF-CD data, ~40 shots were averaged for each kinetic trace to enhance the signal-to-noise ratio of the observed kinetics at 224 nm. The dead-time of the stopped flow instrument was ~5 ms, with a push velocity of 2 ml/s.

Analysis of folding data. The global fitting program, Savuka 5.1, described elsewhere (for example, (*13;14*)), was used to analyze the data. Data sets of five to seven equilibrium unfolding/refolding titrations were globally fit to determine the free energy of unfolding. The data were fit to a two-state model for the unfolding of a dimer to two unfolded monomers, with no populated intermediates. F_{app}, the apparent fraction of unfolded monomers is related to the equilibrium constant, K_{eq} and total monomer concentration, P_T, as well as the observed spectral properties, by the following equation:

$$F_{app} = \frac{\sqrt{K_{Eq}^2 + 8K_{Eq}P_T} - K_{Eq}}{4P_T} = \frac{Y_i - Y_N}{Y_U - Y_N} \qquad (2)$$

where Y_i is the CD signal measured at $[\text{denaturant}]_i$, and Y_N and Y_U are the spectral properties of the folded and unfolded species, *i.e.* the folded and unfolded linear baselines. A linear extrapolation between the free energy of unfolding at a given [denaturant], ΔG°, and the denaturant concentration was used (*15*): $\Delta G° = \Delta G° (H_2O) + m[\text{denaturant}]$, where ΔG ($H_2O$) is the free energy of unfolding in the absence of denaturant at a standard state of 1 M dimer, and the *m* value reflects the sensitivity of the

transition to the denaturant concentration. In the global fitting, the $\Delta G°$ (H_2O) and *m* values were treated as global parameters; the values of Y_N and Y_U were treated as local parameters and not linked between different unfolding transitions.

Kinetic folding and unfolding data were individually fit to a single, first-order exponential: $Y(t) = Y(t=\infty) + \Delta Y * \exp(-t/\tau)$, where τ is the relaxation time for the kinetic phase, ΔY is the signal change associated with the kinetic phase and $Y(t=\infty)$ is the final, equilibrium signal for the reaction. The denaturant dependence of the relaxation times were also globally fitted to the equation:

$$1/\tau_i = k^{H_2O} \exp\left(\frac{m^{\ddagger}[Denaturant]_i}{RT}\right) \qquad (3)$$

where k^{H_2O} is the folding or unfolding rate constant in the absence of denaturant and $m^{\ddagger}$ reflects the sensitivity of the reaction to denaturant. In global fits of data sets at multiple monomer concentrations, the value of $m^{\ddagger}$ was linked over all kinetic traces, and treated as independent of protein concentration; the value of k^{H_2O} was only linked between kinetic traces collected at the same monomer concentrations. The values of ΔY and $Y(t=\infty)$ were treated as local parameters and not linked between data sets. The burst-phase CD signals determined from the global kinetic fits were then fitted to the two-state equilibrium model described above.

The SF double jump experiments were performed using three syringe mixers provided by Aviv Instruments. A delay line was placed between the lower and upper mixer. For the shortest delay times (3 to 10 s), a mixer with a delay volume of 11 µl was

used. For longer folding delay times, a mixer with a delay volume of 198 μl was employed. Unfolded protein and folding buffer were mixed through the lower mixer, allowed to age in the delay lines, and the upper mixer was used to add unfolding buffer with urea to assay the unfolding reaction of the sample. With the small delay line, different delays were achieved by varying the push velocity.

RESULTS

Structural properties of FIS in the native state. The circular dichroism spectra of FIS are dominated by a helical signature, with double minima at 222 and 208 nm (Figure 2.2). This finding is expected based on the three-dimensional structure determined by X-ray crystallography (Figure 2.1). In the conditions used in this report (200 mM KCl, 20 mM KPi, pH 7.2, 10 °C), the mean residue ellipticity determined from the CD spectra was independent of protein concentration (Figure 2.2). These spectra suggest that in the range of 1 to 10 μM monomer, there is no protein concentration dependent change in the structure of FIS, such as dissociation of the dimer into partially folded monomers. Identical results were observed for CD spectra collected at 25 °C (data not shown).

This finding is in contrast to a previous report (*16*), in which a protein concentration dependence of the far-UV CD signal was observed, and were fitted to estimate a K_D value of ~0.4 μM. Under the same conditions as the previous report, we were unable to detect a significant protein-concentration dependence of the far-UV CD spectra (Figure 2.2, inset). There are only minor differences between the conditions of this report and the previous report (*16*) (see legend of Figure 2.2). We have no explanation for the difference between our results and those published previously (*16*),

except that the previous report used serial dilutions to generate the samples and the current study used individually prepared dilutions from a common stock.

A change in the elution of FIS from a size exclusion HPLC column was also reported previously with the protein eluting at later times at lower concentrations (*16*). These data were interpreted as a dissociation of the FIS dimer to monomeric intermediates below ~9 □M monomer. Using the same buffer conditions, a similar protein concentration shift in the elution time (*17*) was observed as reported previously (*16*). The column effluent was monitored by UV absorbance (222 and 280 nm) and by a static light-scattering detector and a refractive index detector. The latter data were used to determine the molecular mass of the eluting species. At 1-16 □M monomer concentrations, FIS eluted as a single, nearly symmetric peak, with coincident profiles for the three spectral probes. The molecular mass of the species was uniform across the peak and did not change with monomer concentration. At monomer concentrations of >25 □M, the apparent molecular mass at the leading edge of the elution peak increased with monomer concentration, indicating the presence of a small amount of aggregate.

The concentration dependence of the elution position observed at 1 to 16 μM monomer can be explained by a rapid exchange phenomenon on the time scale of the HPLC experiment. Cross-linking studies have shown that FIS subunits can exchange with a half-life of 4 to 10 min under similar buffer conditions (*18*). To assess the degree of unfolding/folding during this exchange, folded FIS dimer was subjected to a 5 minute pulse of D_2O (pD 7.4); exchange was then quenched by the addition of acid and the mass of the sample determined by MALDI-TOF mass spectrometry. The mass of FIS unfolded in D_2O (complete proton-deuteron exchange) was 11,378 Daltons. By comparison, the

mass after incubating folded dimer for 5 min in D_2O buffer was 11,362 Daltons, suggesting that at most only 16 protons per monomer (presumably those from hydrogen bonded backbone amides) are protected from exchange on this time scale. This number of protected protons is substantially lower than the ~45 hydrogen bonds in the α-helical regions observed in the X-ray crystal structure (*7*). No amides are protected after a 2 hr incubation in D_2O. This high degree of H/D exchange suggests that the subunit exchange characterized by cross-linking studies must transiently populate a poorly structured monomeric species in which most protons are accessible for facile exchange with solvent.

The subunit exchange of FIS occurs more rapidly than the HPLC-SEC elution time (18 to 19 min). Therefore, as FIS molecules are moving along the HPLC column, they are undergoing subunit exchange. The native dimer is the predominant equilibrium species, as indicated by the molecular weight measurements; however, the protein ensemble is dynamic, with transient population of the monomeric intermediate. The monomeric species is not separated from the dimer during chromatography, as its reassembly to dimer is thermodynamically favored and occurs faster (folding kinetics described below suggest faster than 1 sec) than transport/separation through the column. However, the observed elution peak position represents the population-weighted average of the predominate, folded, dimeric species and the transiently populated, monomeric intermediate. The population-weighted average of the elution profile will be protein concentration dependent, just as the rate of association and stability of the dimer are protein concentration dependent. The effect of [monomer] will be apparent even at concentrations well above the dissociation constant, for example as seen in the equilibrium unfolding transitions as a function of monomer concentration (see next

section). The effects of kinetic processes, such as protein folding and unfolding, on size-exclusion chromatography have been described in detail elsewhere (*19;20*).

The data in this report demonstrate that the native state of FIS is a dimer in the concentration range of 1 to 10 μM monomer; this is the protein concentration range employed in the following experiments. Our CD data suggests that the K_D value for the FIS dimer must be substantially below 1 μM. However, caution may be necessary in interpreting protein folding results at protein concentrations ≥25 μM, given the presence of higher order oligomers.

Equilibrium stability of FIS to chemical denaturants. The equilibrium stability of FIS was determined by chemically-induced denaturation, using both the non-ionic denaturant, urea (*17*), and the salt, GdmCl. Equilibrium data were collected at 10 °C, for comparison to the kinetic data described in this report. Unfolding transitions were monitored by far-UV CD. Data from intrinsic Tyr fluorescence were also collected, but the folded and unfolded baselines were too steep, relative to the transition region, to allow meaningful fitting of the data (data not shown). The reversibility of the far-UV CD equilibrium response was assessed by monitoring titrations in both the unfolding and refolding directions. Folded protein was titrated with a stock of unfolded protein, and conversely, unfolded protein was titrated with a stock of folded protein. The folding was shown to be highly reversible, with no evidence of hysteresis, as evidenced by superimposable transitions (Figure 2.3).

At 10°C, far-UV CD unfolding transitions were collected at several monomer concentrations (1, 2, 5 and 10 μM monomer). The F_{app} curves (Figure 2.3), describing the apparent fraction of unfolded monomers as a function of [denaturant], shift to higher

midpoints with increasing protein concentration; this trend is expected for the unfolding of a dimeric system. The data at multiple monomer concentrations were globally fit to determine the free energy of unfolding in the absence of denaturant, $\Delta G°$ (H_2O), and the *m* value, which is a measure of the sensitivity of the transition to the denaturant concentration. The fitted parameters show little temperature dependence between 25° C and 10° C (*17*).

The $\Delta G°$ (H_2O) and *m* values determined by urea denaturation are similar, but slightly larger than those reported previously (*16*). The $\Delta G°$ (H_2O) values determined from GdmCl denaturation are 10 to 15% less than those measured by urea denaturation (Table 2.1). These small differences between the two denaturants may be caused by the non-ionic nature of urea compared to the salt, GdmCl. Alternatively, the stability from the GdmCl experiments may be slightly underestimated because of the relatively short native baselines compared to the urea data (Figure 2.3).

Kinetic folding and unfolding of FIS. The kinetics of the folding and unfolding reactions of FIS were monitored by stopped-flow CD. At 25 °C, the reactions were too rapid to be detected by stopped-flow methods; therefore to slow the kinetic reactions, the data were collected at 10 °C. A typical refolding SF-CD kinetic trace is shown in Figure 2.4A. The observed kinetic phase is well described by the fit to a single exponential. The expected CD signal for the unfolded protein under folding conditions can be estimated from the unfolded baseline determined from equilibrium unfolding data. The initial CD signal observed after the ~5 ms stopped-flow dead time was substantially greater than that expected for unfolded protein. In fact, the burst-phase refolding CD signal is 65 to 70% that of the folded protein (Figure 2.4A). A similar quality of fit to a

single exponential and magnitude of burst-phase CD amplitude was observed for refolding experiments using urea (data not shown).

Unfolding kinetic traces, using GdmCl, were well-described by a single exponential fit (Figure 2.4B). The amplitudes associated with the single kinetic phase were similar to that expected from the equilibrium data for the conversion of N_2 to 2U. These data demonstrate that the initial step in unfolding is the rate-determining step, and that there are no detectable kinetic intermediates in the unfolding reaction.

Folding kinetics were determined as a function of final monomer concentration. The observed refolding kinetic phase exhibited a relaxation time that is largely independent of protein concentration, especially under strongly refolding conditions. A comparison between the observed folding relaxation times and those expected for a second order, association/folding reaction show a small increase in relaxation time at lower monomer concentrations that is not consistent with a second order reaction (*17*). The small protein concentration dependence (seen also in Figure 2.5A) is consistent with a shift in the equilibrium stability at the different monomer concentrations (compare the F_{app} curves in Figure 2.3). Similar results were obtained for the refolding relaxation times from urea-induced denaturation (*17*). The non-second order protein-concentration dependence suggests that the observed folding reaction is a first order process, and that dimerization may have occurred in the dead time of the SF-CD measurements.

Denaturant dependence of the folding and unfolding kinetics. GdmCl-dependent unfolding kinetics were collected at a final concentration of 5 μM monomer. Data sets of GdmCl refolding kinetics, at 5 and 10 μM monomer, were also collected, yielding the relaxation time of the observable folding reaction and the burst-phase amplitude as a

function of [GdmCl]. At all GdmCl concentrations, the folding and unfolding kinetic traces were well-described by a single exponential. The relaxation times of the GdmCl kinetic traces are presented in Figure 2.5A, in a typical "chevron plot." The chevron analysis of kinetic data describes the folding landscape of a protein (for review, (*21*)). The chevron plot examines the denaturant dependence of both the folding and unfolding relaxations times (or rates, k = 1/tau), typically at multiple protein concentrations for an oligomeric reaction. The log of the locally fitted relaxation times (folding and unfolding) are linearly dependent on [GdmCl] (Equation 3), and the plot takes on the shape of a Chevron (an inverted "V"). Three insights typically gained from chevron plots are: 1) the microscopic reversibility of the folding and unfolding reactions; 2) the $m^{\ddagger}$ value which reflects the change in solvent Accessible Surface Area (ΔASA) between the ground and transition states; and 3) the relaxation times in the absence of denaturant. The microscopic reversibility is shown by convergence of both the unfolding and refolding arms of the chevron plot. This can be seen near the mid-point of the equilibrium transition, ~1.5 M GdmCl, where the refolding and unfolding reactions exhibit similar relaxation times (Figure 2.5A).

Folding and unfolding kinetic traces were globally fitted to Equation 3. As shown in Figure 2.5A, there is good agreement between the results of the local fits and global fits. The parameters determined from the global fits are given in Table 2.2. There is significant denaturant dependence of folding and unfolding relaxation times. The relaxation times extrapolated to the absence of denaturant for 5 □M monomer concentrations are within 3-fold of those determined from the urea data (τ_{fold}, 0.071±0.05 s; τ_{unfold}, 146±6 s). The reduced Chi-squared values for the local and global fits of the

unfolding data (between 1.6 and 2.9 M GdmCl) were 1.097 and 1.076, respectively (72 and 44 parameters for the local and global fits, respectively). For the folding data fits, the local and global reduced Chi-squared values were 1.130 for 76 parameters and 1.152 for 45 parameters, respectively. The lack of change in the reduced Chi-squared between local and global fits demonstrates the adequacy of Equation 3 to describe the kinetic data.

The unfolding in the absence of denaturant is relatively rapid, with a relaxation time less than 50 s (Table 2.2). Therefore, FIS can access the unfolded state, and refold, on a time-scale that is consistent with the explanation of rapid-exchange phenomenon for the protein concentration dependence of the HPLC elution peak position described above.

The slopes of the chevron arms, the $m^{\ddagger}$ values, are given in Table 2.2. The sign of the slope, whether positive or negative, tells if there is an increase (negative) in or burial (positive) of surface assessable surface area associated with the kinetic phase. For folding reactions, this a good diagnostic in differentiating between on-pathway and off-pathway intermediates. For example, in a sequential mechanism with an off-pathway intermediate (Figure 2.6B), the refolding kinetic phase would actually correspond to the rate-limiting unfolding of the off-pathway intermediate, followed by its rapid conversion (burst-phase reaction) to the native dimer, which would yield a refolding chevron arm with a negative slope. As seen in Figure 2.5A, the $m^{\ddagger}_{fold}$ is positive and therefore the mechanism is either sequential with an on-pathway intermediate (Figure 2.6A), or parallel (Figure 2.6C) where there is a fast interconversion between the unfolded and native state in addition to the rapid formation of the intermediate. In order to fully elucidate between a sequential and parallel pathway, double jump kinetic experiments (for review, (*21*)), were performed (described below).

The results of the global fit of the folding data were used in the analysis of the burst-phase and final CD ellipticities. The GdmCl dependence of the SF-CD signals, for data at 5 μM monomer, are shown in Figure 2.5B. SF-CD data were also collected at 10 μM monomer, but aren't shown for clarity. The burst-phase signal decreases with increasing [GdmCl] in a manner suggesting the unfolding/destabilization of a kinetic intermediate. Similar results were obtained with urea (*17*). Given the lack of protein concentration dependence of the observed kinetic phase for both GdmCl (Figure 2.5A) and urea, it is hypothesized that the burst-phase intermediate is a dimeric ensemble, formed from a rapid association reaction that is complete within the 5 ms dead time of the stopped-flow instrument. This hypothesis would predict a protein-concentration dependence of the burst-phase signal. This dependence is observed, particularly with urea denaturation, as indicated by the shift to higher denaturant concentrations of the burst-phase amplitude pattern with higher monomer concentrations (*17*).

The burst-phase and final SF-CD signals were normalized for differences in monomer concentrations as well as cuvette and background buffer contributions so that comparisons could be made with far-UV CD equilibrium unfolding transitions. The final CD ellipticity (open symbols, Figure 2.5B) should reflect the GdmCl-dependence of the equilibrium response if the entire folding reaction has been completed during the SF experiment. If a slow, far-UV-CD-spectroscopically silent reaction occurred outside of the SF time scale of data collection (3 to 10 s) and led to a significantly more stable FIS species, the transition monitored by the final SF ellipticity and the equilibrium data should not coincide. As shown in Figure 2.5B, the final SF-CD folding and unfolding signals agree relatively well with the representative equilibrium data (dotted-dashed

lines). This agreement argues that the final species observed in the SF-CD folding kinetics of FIS is that of the native dimeric species.

To determine the stability of the apparent dimeric intermediate ensemble formed within ~5 ms, the burst-phase ellipticities at 5 and 10 μM monomer were globally fitted to a two-state unfolding equilibrium (Equation 2, Methods) between 2U and I_2. In the fitting, the unfolded baseline, determined from equilibrium experiments and the final CD signal of the unfolding SF-CD experiments (open diamonds, Figure 2.5B), was fixed. The slope of the folded baseline was fixed at that observed for the N_2 to 2U transition, as there was insufficient data to determine this parameter independently. The value of the intercept of the folded baseline for the burst phase data was linked between the normalized data sets at the two monomer concentrations. The solid line in Figure 2.5B is the result of the global fit of the burst phase data. The $\Delta G°$ (H_2O) and *m* values are given in Table 2.2. Given that the dimerization reaction seems to have been completed in the 5 ms SF dead time at monomer concentrations as low as 2 μM, the second order rate constant for the formation of the putative I_2 intermediate must be $\geq 10^7$ $M^{-1}s^{-1}$, approaching the diffusion limit.

Double jump experiments to assay when N_2 is formed. The kinetic data described above suggest that an on-pathway dimeric intermediate is transiently populated in the burst-phase folding of FIS. However, it is necessary to ascertain if this burst phase intermediate is obligatory, *i.e.* does the vast majority of the folding protein transiently populate this I_2 species in the most direct, accessible route from the unfolded monomers to native dimer OR does some fraction of protein fold to native dimer during the 5 ms SF dead time? Double jump kinetic experiments are a powerful method by which to address

this question. The general protocol is to dilute unfolded protein to folding conditions for varying lengths of time, and then jump to denaturant concentrations that promote unfolding (Figure 2.7). The goal is to detect the accumulation of the native state as a function of folding time; the native state is detected by virtue of its relatively slow unfolding reaction. Less stable, transient kinetic intermediates should unfold more rapidly than the fully folded native species (for example, (*22*)).

Double jump experiments were carried out as follows: 25 μM FIS monomer, unfolded in 5 M urea, was refolded by SF dilution to 2 M urea (10 μM monomer) (Step 1 in Figure 2.7A and B). The extent of native FIS dimer was assayed by a second jump to unfolding conditions, 4 M urea (5 μM FIS), monitored by SF-CD (Step 2 in Figure 2.7A and B). The kinetic traces of the unfolding phase were well-described by a single exponential and were globally fitted, with a linked relaxation time for unfolding. The relaxation time of the unfolding reaction detected by the double jump assay, 948 ± 14 ms, was similar to the relaxation times measured for direct unfolding jumps to 4 M urea (960 and 920 ms, for local and global fits, respectively). This similarity shows that the double-jump unfolding amplitudes are an assay for the presence of the N_2 state. The unfolding amplitudes as a function of folding delay time are shown in Figure 2.8. The increase in unfolding amplitude, as a function of folding time, was fit to a single exponential (line in Figure 2.8), yielding a relaxation time of 276 ± 28 ms. This value is in good agreement with the observed SF-CD relaxation time for direct folding to 2 M urea (10 μM monomer), 286 ± 40 ms. Unfolding after the shortest possible folding delay, 3.3 ms, yielded an unfolding amplitude of 0.02 mdeg, compared to an amplitude change of 4.7 mdeg after refolding for 2 to 5 s. These values suggest that less than 1% of the FIS

monomers have reached the native state in a folding time of 3 ms. The extent of N_2 achieved faster than the first order folding reaction can also be estimated by comparison of the fitted final amplitude for an infinite folding delay (data in Figure 2.8, -4.7 mdeg) and the fitted amplitude change as a function of delay time (4.4 mdeg); this comparison suggests that 93% of the FIS monomers reach the native state with the relaxation time of that observed by SF-CD refolding kinetics. The double-jump assay demonstrates that: 1) the burst-phase ensemble contains little or no fully folded N_2 dimer; and 2) folding to native dimer is constrained by the first-order folding reaction observed by SF-CD for the vast majority of unfolded FIS monomers.

DISCUSSION

Folding mechanism of FIS. The folding of FIS to the native dimer from denaturant unfolded monomers requires at least two steps: a rapid development of secondary structure in 5 ms and a slower folding reaction that leads to the native dimer. The double jump experiment (Figure 2.8) demonstrates that virtually no native dimer is formed in the reaction that occurs in SF dead time. The structure detected in the SF-CD burst phase represents an obligatory transient kinetic intermediate that is competent for conversion to the native dimer by the slower, subsequent folding reaction.

A small protein concentration dependence was detected for the observed refolding kinetic response. At GdmCl concentrations $\geq$ 1.4 M, there is an obvious difference between the locally fitted 5 and 10 μM data (■ and ○, Figure 2.5A), where the 10 μM data exhibits relaxation times 1.5 to 2.7-fold faster than those at 5 μM monomer. These differences, while significant, are substantially less than expected for a second order

association reaction. In effect, it appears that the 10 μM folding chevron arm has been shifted to higher urea concentrations, by ~ 0.2 M relative to the 5 μM data. This "GdmCl shift" and protein concentration dependence can be explained by the dependence of the unfolding transitions of the I_2 burst-phase species on monomer concentration, and is expected for a first order reaction coupled to a preceding association reaction. The 5 μM and 10 μM I_2 denaturation curves are offset by ~0.15 M urea (data not shown). Similar results are observed in the urea refolding data (*17*). The chevrons at 5 and 15 μM monomer are offset by ~0.4 M urea, and the unfolding transitions of the I_2 burst-phase intermediate are offset by ~0.5 M. Thus, the small effect of [monomer] on the observed folding relaxation times reflects the protein concentration dependence of the stability (equilibrium position or F_{app} value) of the folding species, in this case, I_2.

In combination, three results demonstrate that the burst phase kinetic intermediate is dimeric, a species in which the two monomer chains have associated and partially folded. Firstly, the relative insensitivity of the observed relaxation time for the conversion to the native dimer is consistent with a first-order reaction between a dimeric intermediate and the native dimer (Figure 2.5A). Secondly, the convergence of the folding and unfolding relaxation times near the midpoint of the equilibrium transition suggests that the first-order folding phase leads to the native dimeric species, and that the initial, rate-determining step in the unfolding of N_2 is the reverse of this first-order folding reaction. Therefore, dimerization must have preceded the observed first-order folding reaction. Thirdly, the stability of the kinetic intermediate, as measured by the dependence of the burst-phase ellipticity on denaturant concentration is protein concentration dependent in a manner expected for a dimeric species.

The most appropriate mechanism to describe the kinetic folding and unfolding data presented here for FIS is shown in Figure 2.6A. This scheme represents a minimal mechanism, and it is possible that other steps may exist that are not detected by SF-CD. Firstly, no information is available about the monomeric species that associate to form I_2. The association occurs faster than the time resolution of stopped-flow methods, so it is not possible for SF-CD methods to address if the monomers are partially folded in isolation before association to form I_2. Secondly, SF-CD assesses the formation of secondary structure. The development of tertiary structure has not been monitored directly by a method such as SF-FL, using the intrinsic Tyr FL. FIS contains four Tyr residues per monomer, and no Trp residues. However, the FL change associated with folding is very small, presumably because of the large solvent accessibility of the Tyr side chains in the folded state.

Stability of I_2 relative to N_2. In the context of the mechanism shown in Scheme 1, the results of the equilibrium and kinetic studies can be compared. The free energy, $\Delta G°$ (H_2O), and denaturant dependence, *m* value, for the conversion of 2U to I_2 was determined from the denaturant dependence of burst phase signals (Table 2.2). The magnitude of these parameters is less than those for the equilibrium unfolding of N_2 to 2U (Table 2.1), but demonstrates that the transient kinetic I_2 species exhibits substantial stability. The $\Delta G°$ (H_2O) values suggest that the K_D values for I_2 and N_2 are on the order of ~20 nM and ~20 pM, respectively. The *m* value is generally proportional to the ΔASA between the initial and final species (*23*). The *m* values determined for the formation of the I_2 species suggest that 40 to 60% of the total ΔASA between 2U and N_2 is buried in the formation of I_2.

The stability associated with the observed folding and unfolding reactions (between I_2 and N_2 in Figure 2.6A) was estimated from the following relationship: $\Delta G°$ (H_2O) = -RT ln($k_{unfold}^{H_2O} / k_{fold}^{H_2O}$) (Table 2.2). The corresponding *m* value was calculated by summing the absolute values of the $m^{\ddagger}$ values from folding and unfolding arms of the chevron plots (Figures 2.5A; Table 2.2). The *m* values indicate that this reaction proceeds with a significant burial of surface area, as expected for a folding reaction, and is not an isomerization reaction (such as proline isomerization), with little to no burial of surface area.

The stability and burial of surface area for neither the burst-phase or the slower reactions (Table 2.2) are as large as those that describe the equilibrium process, 2U to N_2, of FIS (Table 2.2). Clearly, the parameters associated with the observed first order folding/unfolding reactions, while leading to and from native dimer (as discussed above), do not describe, thermodynamically, the complete folding of FIS. Therefore, a second step, namely the burst-phase reaction, must also contribute to the overall stability and burial of surface area necessary for the complete folding of FIS. If the scheme in Figure 2.5A is the correct kinetic mechanism to describe the folding/unfolding of FIS, the summation of the $\Delta G°$ (H_2O) and *m* values determined for the two kinetic events should equal those determined by equilibrium experiments. For the GdmCl kinetic data, the $\Delta G°$ (H_2O) and *m* values for 2U to N_2 are 13.5 kcal mol^{-1} and 4.8 kcal $mol^{-1}M^{-1}$, respectively. The similarity to the equilibrium values (Table 2.1) supports the appropriateness of the scheme in Figure 2.6A. Similar agreement is observed for the urea kinetic and equilibrium data: $\Delta G°$ (H_2O) values of 13.9 and 15.2 kcal mol^{-1} and *m* values of 2.6 and 2.9 kcal $mol^{-1}M^{-1}$ for the kinetic and equilibrium data, respectively.

Differences between the urea and GdmCl energy landscapes. Both GdmCl and urea were used to help verify the apparent folding mechanism. However, the use of two denaturants may provide insights into the features that determine the details of the folding landscape of FIS. Refolding from urea and GdmCl can be described by the mechanism shown in panel A of Figure 2.6A. Relative to the urea I_2 species, the GdmCl I_2 species is somewhat more native-like, in terms of stability and burial of surface area.

The similarity of the transition state between I_2 and N_2 to the final N_2 state can be estimated from the Φ value:

$$\Phi = \frac{m^{\ddagger}_{fold}}{m^{\ddagger}_{fold} - m^{\ddagger}_{unfold}} \qquad (1)$$

The magnitude of Φ generally ranges from 0 (I_2-like) to 1 (N_2-like). For the transition states traversed in urea and GdmCl conditions, the Φ values are 0.46 and 0.26, respectively. Thus, the transition state traversed in the presence of GdmCl salt is much less native like than that traversed in the presence of urea.

Comparison to an equilibrium dimeric intermediate. The long, second α-helix of FIS contains a 20° kink centered at residue Pro61 (Figure 1). Mutation of this residue to Ala only slightly decreases the kink and does not significantly alter FIS function (*24*). However, protein folding studies have shown that the P61A mutation has a pronounced effect on the equilibrium folding and stability of FIS (*25*). P61A is significantly more stable than WT and populates a dimeric intermediate. There are interesting similarities between the I_2 species observed at equilibrium in the urea-induced unfolding of FIS P61A and the urea transient kinetic intermediate observed in this report. The stability of both dimeric intermediates are 63 to 68% of the native dimer. The relative ΔASA, from

comparison of *m* values relative to complete unfolding (2U to N_2), is also similar, ~52% for the P61A equilibrium intermediate and 45% for the kinetic I_2. Global analyses of the equilibrium data indicate that the I_2 species of P61A retains ~70% of the spectral signal associated with the native dimer. This magnitude of signal is similar to the burst-phase amplitude of I_2, relative to N_2, under strongly refolding conditions, 65 to 70%. On the basis of these comparisons, at least at the level of global thermodynamics and structure, the I_2 kinetic intermediate is strongly mimicked by the equilibrium intermediate observed for the P61A mutation.

FIS as a model folding system. A growing number of small (<100 residues), monomeric single domain protein folding model systems that can fold by a two-state kinetic mechanism have been identified (for example, cold-shock protein B (*26;27*), SH3 domains (*28-31*), and chymotrypsin inhibitor 2 (*32-34*). These model systems and insights from theoretical, computational studies (for example, (*35-39*)) have called into question the role of transient kinetic intermediates in protein folding. Do kinetic folding intermediates accelerate or improve the efficiency of protein folding OR are they kinetic traps that frustrate the folding protein and hinder rapid, productive conversion to the native species? There are certainly examples of intermediates as off-pathway kinetic traps that decrease the flux to the native state; a classic example is non-native heme ligation in the folding of cytochrome c (*40*). The signature of these kinetic traps is that structure formed in the intermediate must be *unfolded* before the native state can be achieved. Such an unfolding event of the burst phase intermediate populated in the folding of FIS is inconsistent with the positive $m_{fold}^{\ddagger}$ values seen for the denaturant dependence of the subsequent folding reaction (Figures2.5A and Table 2.2). Therefore, it

is reasonable to conclude that the structure and interactions formed in the FIS folding intermediate are productive in the formation of the native state.

Why is an intermediate necessary; why can't unfolded monomers fold directly to the native species? In larger proteins, with multiple domains and/or subunits and large contact orders (as defined by Plaxco *et al* (*41*)), kinetic intermediates may be required to lower the free energy barriers of folding, and increase the probability of folding. For a large protein to fold in a two-state manner, many more interactions—between greater numbers of side chains and backbone moieties separated by greater sequence distances—need to be formed in a concerted manner. The simultaneous coordination of the formation of these interactions could result in a relatively high energetic barrier, with a low probability of being traversed. This folding problem could be circumvented if the concerted nature of the folding reaction is broken into the formation of smaller, sub-global units of structure—potential intermediate ensembles—which coalesce into the native state as folding progresses. Therefore, it is conceivable that intermediates are required to accelerate the folding of larger proteins. Such proteins, with multiple domains and/or multiple subunits, are highly represented in the proteomes of most organisms. Therefore, it is important to understand the folding of these more complex structures and any intermediates that they transiently populate.

The folding of several oligomeric, particularly dimeric, systems have been studied. The dimeric model systems range from relatively small, coiled-coil peptides (for example, (*14;42-44*)) to moderate sized enzymes (for example, ketosteroid isomerases (*45;46*) and glutathione S-transferases (*47;48*)) to large homodimeric systems, such as malate dehydrogenase (*49;50*) and Sec A (*51*) and heterodimeric systems, such as

bacterial luciferase (*52;53*). Varying complexities of folding mechanisms have been observed, from apparent two-state folding of the leucine zipper peptides, to multiple phases for the larger, multi-domain systems.

It is of greatest interest to compare the folding of FIS to other small DNA-binding dimers, such as *Arc* repressor (*54*) (53 residues per monomer) and *Trp* repressor (*55*) (107 residues per monomer). Both proteins, like FIS, have monomer-monomer interfaces that are interdigitated or intertwined. Computational analyses of dimer interfaces and subunit structure were reported by Nussinov and colleagues (*8*), employing compactness profiles to assess the extent of domain swapping (*56*) and intertwining of various dimers; the study included *Arc* repressor, *Trp* repressor and FIS. These profiles provided insights into the classification of the dimer interfaces and possible folding mechanisms. The profile of *Arc* repressor was relatively flat, with the short chains being compact and much of the contacts made by each monomer being buried in the dimer interface. In contrast, FIS and *Trp* repressor exhibited profiles that could be described as domain or segment-swapped; the term segment swapped was used to specify that the swapped portions of the monomers were extended and did not fit the usual definition of a protein domain (*8*). For FIS and *Trp* repressor, it appeared that most of the stability of the dimer was contributed by the segment-swapping. Are there similarities and differences in the kinetic folding mechanisms of these proteins that relate to dimer interface structure?

Arc repressor monomers make contact through helices and two interdigitated □–strands, the latter providing the major contacts in the DNA binding interface. *Arc* repressor folds by a two-state kinetic mechanism, with no detected kinetic intermediates (*57;58*). *Arc* repressor as a model system demonstrates that the folding of quaternary,

tertiary and secondary structures can occur while traversing a single barrier in the folding reaction, and thus folding of small, dimeric structures does not require the formation of transient kinetic intermediates.

In FIS and *Trp* repressor, the secondary structure of the intertwining segments of the dimerization interface are exclusively α–helical. Both *Trp* repressor and FIS interact with DNA through helix-turn-helix motifs at or near their C-termini. The folding landscape of *Trp* repressor is complicated and rough (*59-62*), with three parallel channels, each populating monomeric intermediates and dimeric intermediates. The three channels arise from different modes of folding involving the three C-terminal helices. The $[2\text{-}66]_2$ core dimerization domain (without these C-terminal helices) folds more simply, through a single channel (*63;64*). The mechanism shown in Scheme 1 for FIS also describes the folding of the segment-swapped dimerization domain of *Trp* repressor.

These two examples suggest that dimeric intermediates may be a common feature in the folding landscape of segment-swapped, α-helical interfaces. It is unclear if this common folding mechanism is related to their similar function of binding DNA. To extend this potential link between topology, interfacial secondary structure content and segment-swapping (or compactness profile (*8*)), further model systems have been characterized. The α-helical histone fold, first characterized in the histones of the eukaryotic core nucleosome (*65*), is also an intertwined segment-swapped, α-helical DNA-binding Dimer (ISSADD). Data published for the H2A-H2B histone dimer and H3-H4 tetramer show that they similarly fold by the mechanism described by Figure 2.6A (*66;67*). The kinetic folding of the homodimeric archael histones, hPyA1, hMfB and hFoB, has also been examined by Topping and Gloss (*68*). Despite similar structures to

the eukaryotic histones, these histones do not populate a dimeric intermediate, although hMfB does fold through a sequential pathway with an on-pathway, monomeric intermediate (*68*). Interestingly, there appears to be a correlation between rate of folding and the presence of an on-pathway intermediate; the presence of intermediates appears to accelerate protein folding. This phenomenon is being further explored in comparative studies of H2A-H2B (diffusion limited folding with dimeric and presumed monomeric intermediates), hMfB (moderate folding rate a with monomeric intermediate) and hPyA1 (slowest folding by a two-state mechanism). Mutations are being introduced into these systems that will either stabilize or destabilize the kinetic intermediate(s), and monitor their effect on the folding rates. The hypothesis is that destabilization of a kinetic intermediate will decrease the folding rates.

ACKNOWLEDGEMENTS

The expression plasmid, pRJ1077, was generously provided by Dr. Reid Johnson (University of California, Los Angeles). This work was supported by grants to L.M.G. from the National Science Foundation (MCB-9983831) and the American Cancer Society (RPG-00-085-01-GMC). D.A.H. was partially supported by an N.I.H. Biotechnology training grant (GM08336-13).

Table 2.1. Parameters describing the equilibrium folding/unfolding reactions of FIS at 10ºC.[a]

Parameters	Urea data	GdmCl data
$\Delta G°$ (H_2O) (kcal mol^{-1})	15.2 (0.8)	13.5 (0.5)
m value (kcal $mol^{-1}M^{-1}$)	2.9 (0.3)	5.0 (0.3)

[a]Results of global fitting of the data. Conditions: 200 mM KCl, 1 mM EDTA, 20 mM KPi, pH 7.2. The values in parentheses are the error of one standard deviation determined from rigorous error analysis of global fits of the data.

Table 2.2. Parameters describing the kinetic folding/unfolding reactions of FIS.[a]

Parameters	GdmCl
τ_{fold} (s)	0.26 (0.03) 5 μM
	0.20 (0.02) 10 μM
$m^{\ddagger}_{fold}$ (kcal mol^{-1}M^{-1})	0.40 (0.06)
τ_{unfold} (s)[b]	49 (6)
$m^{\ddagger}_{unfold}$ (kcal mol^{-1}M^{-1})	-1.38 (0.04)
I_2 to N_2 values[b]	
$\Delta G°$ (H_2O) (kcal mol^{-1})	3.0
m value (kcal mol^{-1}M^{-1})	1.8
2U to I_2 values	
$\Delta G°$ (H_2O) (kcal mol^{-1})	10.5 (1.1)
m value (kcal mol^{-1}M^{-1})	3.0 (0.8)

[a]Results of global fitting of the data. Conditions: 200 mM KCl, 1 mM EDTA, 20 mM KPi, pH 7.2, 10 °C. The associated error of the values are shown in parentheses.

[b]For the data at 5 μM monomer.

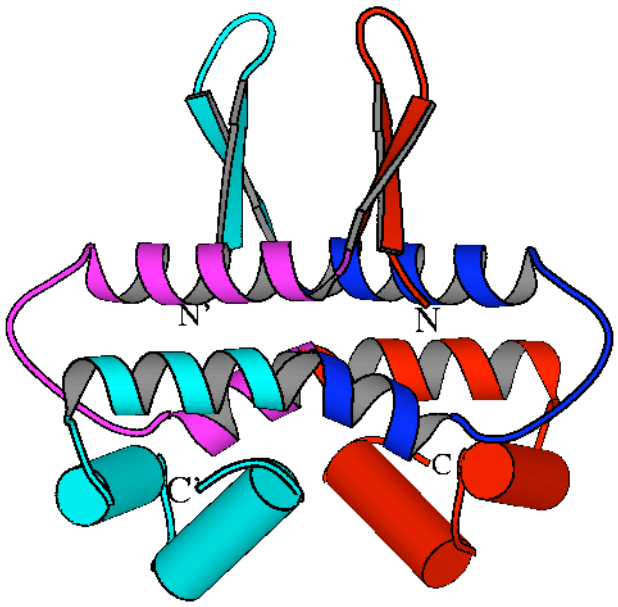


Figure 2.1. Ribbon diagram of the intertwined FIS homodimer. The helix-turn-helix DNA binding motifs are shown as cylinders. Monomer1 is colored cyan (N terminal β–strands, 10 to 25 residues, and C-terminal helices) and blue (central residues 26 to 57), and monomer2 is colored magenta and red with the same boundaries. The cyan portion of monomer1 makes contacts with the magenta portion of monomer2 that are nearly identical to those made by the red portion of monomer2 to the blue portion of monomer1. The figure is derived from the structure of a FIS mutant K36E (1f36.pdb (*7*)); in which the position of the N-terminal β-hairpins are resolved. The figure was rendered using Molscript v2.1 (*69*).

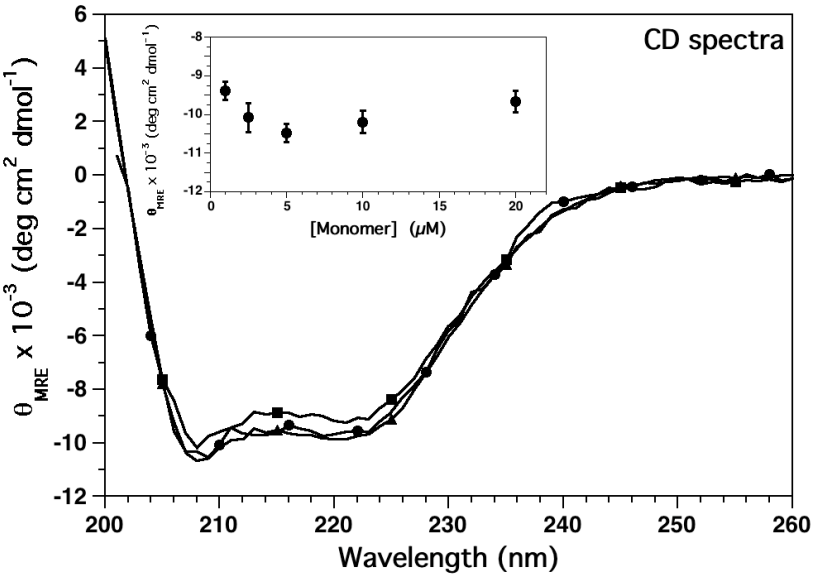


Figure 2.2. Far-UV CD spectra of FIS. Far-UV CD spectra at monomer concentrations of 1 μM, ●; 5 μM, ■; and 10 μM, ▲. The spectra have been converted to mean residue ellipticity. Conditions: 200 mM KCl, 20 mM KPi, pH 7.2, 10 °C. Inset of A) Far-UV CD signal at 222 nm as a function of [monomer] under the buffer conditions of Hobart *et al*.(*16*) Conditions: 170 mM NaCl, 10 mM KPi, pH 7.2

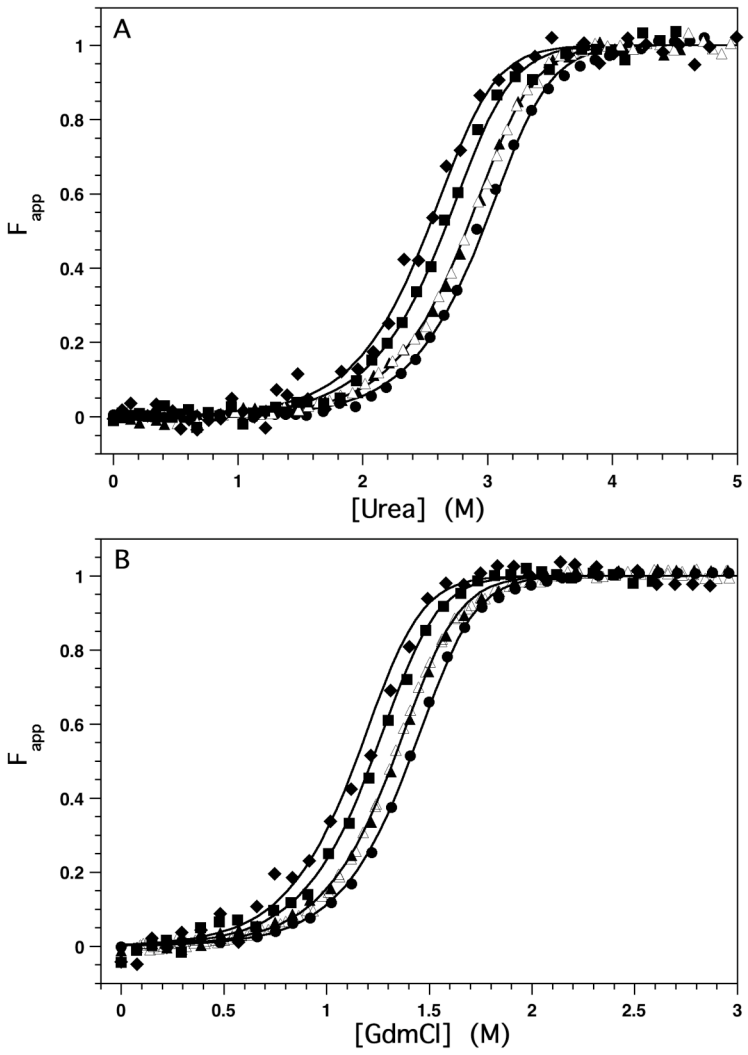


Figure 2.3. Representative equilibrium unfolding transitions for FIS monitored by far-UV CD spectra using singular value decomposition. The solid lines represent global fits of data at multiple protein concentrations to a two-state equilibrium model between native dimer and two unfolded monomers. Monomer concentrations: 1 μM, diamonds; 2 μM, squares; 5 μM, triangles; 10 μM, circles. Unfolding and refolding transitions are shown as solid and open symbols, respectively. Conditions: 200 mM KCl, 20 mM KPi, pH 7.2, 1 mM EDTA, 10 °C. **A.** Urea-induced transitions. **B.** GdmCl-induced transitions. Conditions: 200 mM KCl, 20 mM KPi, pH 7.2, 1 mM EDTA, 10 °C.

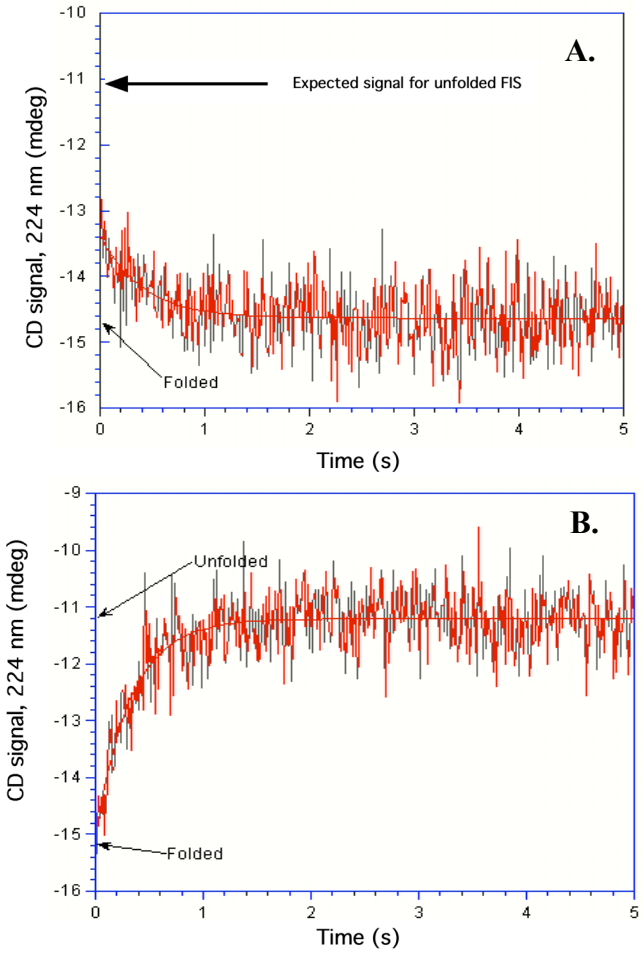


Figure 2.4. SF-CD refolding kinetics for GdmCl-denatured FIS. **A.** Representative SF-CD refolding kinetic trace at 5 µM FIS monomer, final [GdmCl] = 1.0 M. **B.** Representative SF-CD unfolding kinetic trace at 5 µM FIS monomer, final [GdmCl] = 2.0 M. The red line represents a global fit of the data to the model shown in Equation 3. The arrows indicate the expected ellipticity for the folded and unfolded protein under these conditions, extrapolated from the linear unfolded baseline observed in equilibrium experiments. Conditions: 200 mM KCl, 20 mM KPi, pH 7.2, 1 mM EDTA, 10 °C.

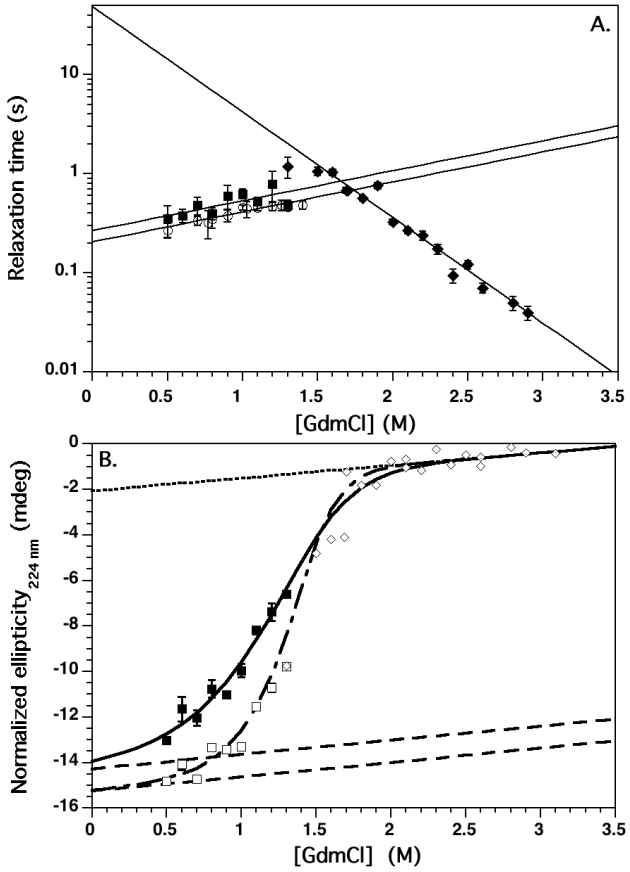


Figure 2.5. GdmCl dependence of SF-CD folding and unfolding kinetic data. **A.** Relaxation times. Local fits of the kinetic data are shown as data points. The solid lines represent global fits of the data to Equation 3; fitted parameters are given in Table 2.2. 5 μM folding, ■; 5 μM unfolding, ◆; 10 μM folding, ○. **B.** Burst-phase (closed symbols) and final (open symbols) CD ellipticity as a function of [GdmCl]. The data were normalized for different cuvette sizes and buffer contributions. 5 μM burst-phase folding signal, ●; 5 μM final folding signal, □; 5 μM final unfolding signal, ◇; unfolded baseline, dotted line; folded baselines, dashed lines; representative equilibrium transitions, dotted-dashed lines; fits of burst-phase data to a two-state model (Eq. 2), solid lines. The 10 μM folding kinetic data were used in the fit, but are not shown in the figure for clarity, because of the small difference between GdmCl dependence of the 5 and 10 μM data. Error bars are shown or are equal to or less than the size of the symbols. Conditions are given in the legend of Figure 2.4.

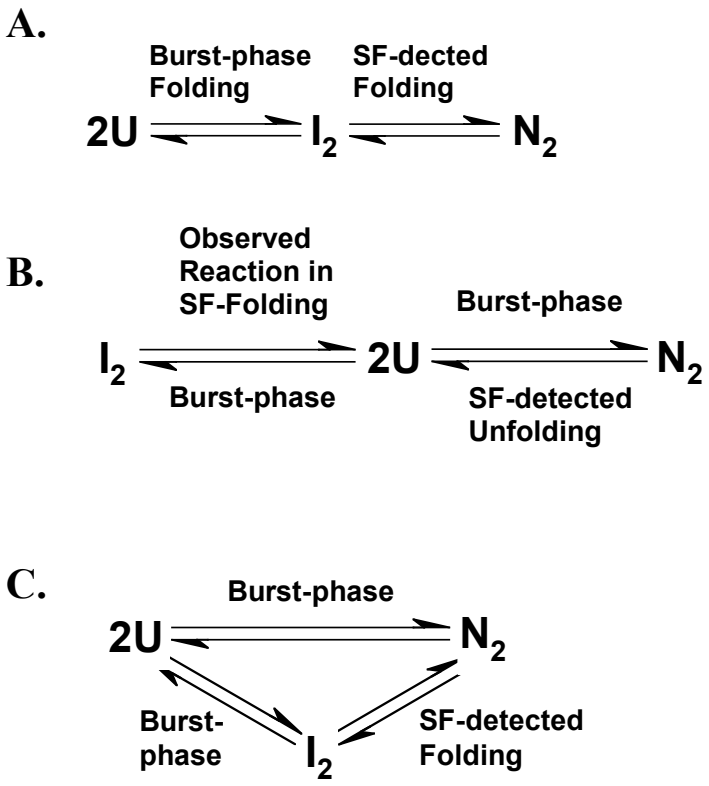


Figure 2.6. Potential folding mechanisms for a dimeric protein with a dimeric intermediate. **A.** Sequential on-pathway; **B.** Sequential off-pathway; and **C.** Parallel pathway.

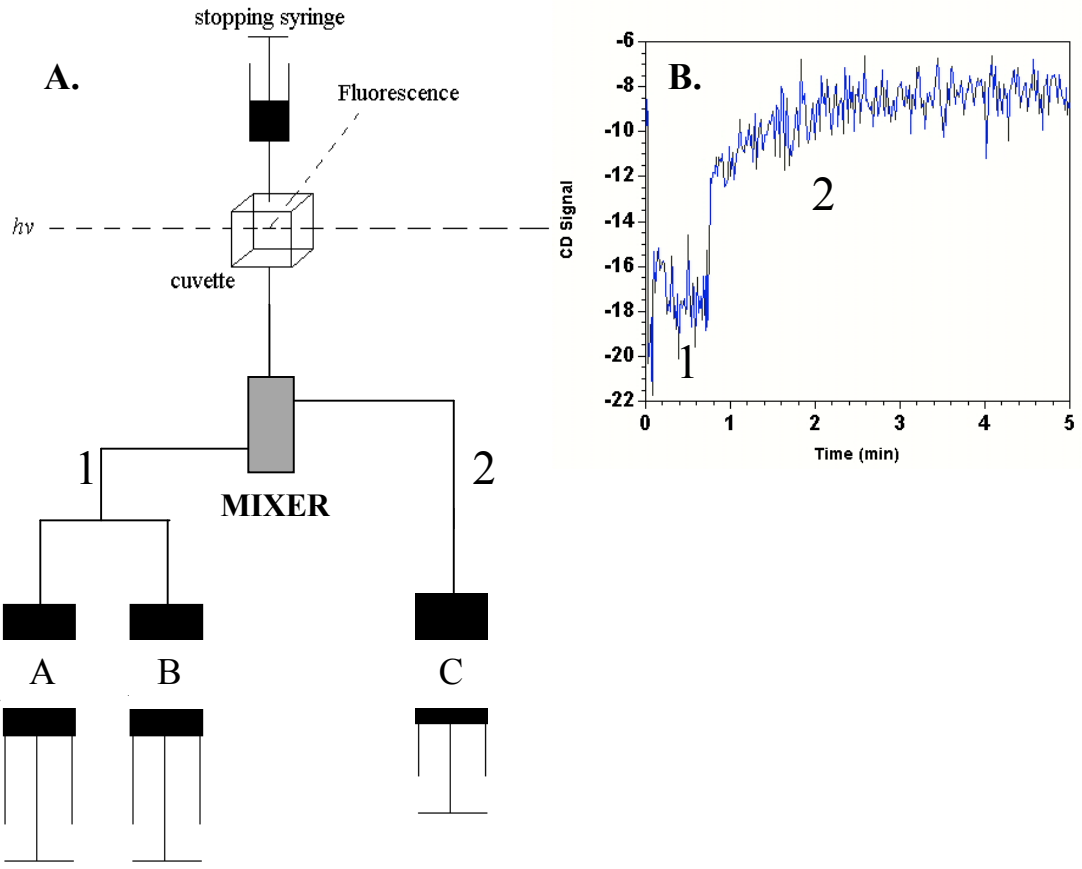


Figure 2.7. **A.** Generic diagram of a stopped-flow apparatus adapted from the thesis of Dr. Douglas Banks (2003, School of Molecular Bioscience, Washington State University). In double jump experiments, two driving syringes A and B are compressed in step 1, mixing unfolded protein and buffer and allowed to "age" for various refolding periods. After the desired delay time, syringe C is then compressed, mixing the aged sample with denaturant thus diluting the sample into unfolding conditions. During the time frame of the experiments, the reactions are monitored by CD. **B.** An example of a double jump kinetic trace fit with a delay time of 0.6 min.

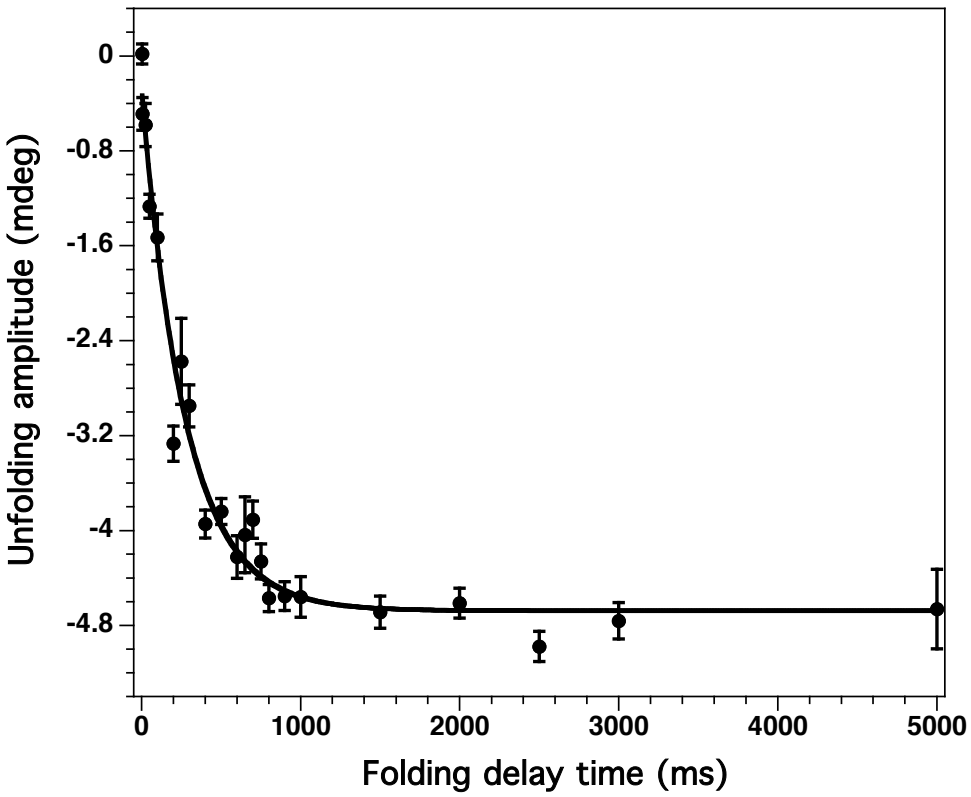


Figure 2.8. Urea double-jump assay for the formation of native FIS dimer. SF-CD unfolding amplitude data are shown as a function of refolding delay time. The solid line represents the fit of the data to a single exponential. The fitted relaxation time is 276 ± 33 ms, similar to the observed relaxation time for direct refolding of FIS at 10 μM monomer and 2 M urea, 286 ± 30 ms. The fitted amplitude change is 93% of the final, total signal change.

Literature Cited

1. Fersht, A. R. (2000) Transition-state structure as a unifying basis in protein-folding mechanisms: contact order, chain topology, stability, and the extended nucleus mechanism, *Proc Natl Acad Sci U S A 97*, 1525-1529.

2. Jackson, S. E. (1998) How do small single-domain proteins fold?, *Fold Des 3*, 81-91.

3. Matthews, C. R. (1993) Pathways of protein folding, *Annual Review of Biochemistry 62*, 653-683.

4. Finkel, S. E. and Johnson, R. C. (1992) The Fis protein: it's not just for DNA inversion anymore, *Mol Microbiol 6*, 3257-3265.

5. Travers, A., Schneider, R., and Muskhelishvili, G. (2001) DNA supercoiling and transcription in Escherichia coli: The FIS connection, *Biochimie 83*, 213-217.

6. Wolynes, P. G., Onuchic, J. N., and Thirumalai, D. (1995) Navigating the folding routes, *Science 267*, 1619-1620.

7. Cheng, Y. S., Yang, W. Z., Johnson, R. C., and Yuan, H. S. (2000) Structural analysis of the transcriptional activation on Fis: crystal structures of six Fis mutants with different activation properties, *J Mol Biol 302*, 1139-1151.

8. Xu, D., Tsai, C. J., and Nussinov, R. (1998) Mechanism and evolution of protein dimerization, *Protein Sci 7*, 533-544.

9. Liu, Y. and Eisenberg, D. (2002) 3D domain swapping: As domains continue to swap, *Protein Sci 11*, 1285-1299.

10. Pan, C. Q., Finkel, S. E., Cramton, S. E., Feng, J. A., Sigman, D. S., and Johnson, R. C. (1996) Variable structures of Fis-DNA complexes determined by flanking DNA-protein contacts, *J Mol Biol 264*, 675-695.

11. Henry, E. R. and Hofrichter, J. (1992) Singular value decomposition: Application of analysis of experimental data., *Methods Enzymol. 210*, 129-192.

12. Ionescu, R. M., Smith, V. F., O'Neill, J. C., Jr., and Matthews, C. R. (2000) Multistate equilibrium unfolding of Escherichia coli dihydrofolate reductase: thermodynamic and spectroscopic description of the native, intermediate, and unfolded ensembles, *Biochemistry 39*, 9540-50.

13. Bilsel, O., Zitzewitz, J. A., Bowers, K. E., and Matthews, C. R. (1999) Folding mechanism of the alpha-subunit of tryptophan synthase, an alpha/beta barrel protein: global analysis highlights the interconversion of multiple native, intermediate, and unfolded forms through parallel channels, *Biochemistry 38*, 1018-1029.

14. Zitzewitz, J. A., Ibarra-Molero, B., Fishel, D. R., Terry, K. L., and Matthews, C. R. (2000) Preformed secondary structure drives the association reaction of GCN4-p1, a model coiled-coil system, *J Mol Biol 296*, 1105-1116.

15. Pace, C. N. (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves, *Methods Enzymol. 131*, 266-280.

16. Hobart, S. A., Ilin, S., Moriarty, D. F., Osuna, R., and Colon, W. (2002) Equilibrium denaturation studies of the Escherichia coli factor for inversion stimulation: Implications for in vivo function, *Protein Sci 11*, 1671-1680.

17. Topping, T. B., Hoch, D. A., and Gloss, L. M. (2004) Folding mechanism of FIS, the intertwined, dimeric factor for inversion stimulation, *J Mol Biol 335*, 1065-1081.

18. Merickel, S. K., Sanders, E. R., Vazquez-Ibar, J. L., and Johnson, R. C. (2002) Subunit exchange and the role of dimer flexibility in DNA binding by the Fis protein, *Biochemistry 41*, 5788-5798.

19. Shalongo, W., Heid, P., and Stellwagen, E. (1993) Kinetic analysis of the hydrodynamic transition accompanying protein folding using size exclusion chromatography. 1. Denaturant dependent baseline changes, *Biopolymers 33*, 127-134.

20. Shalongo, W., Jagannadham, M., and Stellwagen, E. (1993) Kinetic analysis of the hydrodynamic transition accompanying protein folding using size exclusion chromatography. 2. Comparison of spectral and chromatographic kinetic analyses, *Biopolymers 33*, 135-145.

21. Wallace, L. A. and Matthews, C. R. (2002) Sequential vs. parallel protein-folding mechanisms: experimental tests for complex folding reactions, *Biophys Chem 101-102*, 113-131.

22. Martin, A. and Schmid, F. X. (2003) The folding mechanism of a two-domain protein: folding kinetics and domain docking of the gene-3 protein of phage fd, *J Mol Biol 329*, 599-610.

23. Myers, J. K., Pace, C. N., and Scholtz, J. M. (1995) Denaturant *m* values and heat capacity changes: Relation to changes in accessble surface areas of protein folding., *Protein Science 4*, 2138-2148.

24. Yuan, H. S., Wang, S. S., Yang, W. Z., Finkel, S. E., and Johnson, R. C. (1994) The structure of Fis mutant Pro61Ala illustrates that the kink within the long alpha-helix is not due to the presence of the proline residue, *J Biol Chem 269*, 28947-28954.

25. Hobart, S. A., Meinhold, D. W., Osuna, R., and Colon, W. (2002) From two-state to three-state: The effect of the P61A mutation on the dynamics and stability of

the Factor for Inversion Stimulation in an altered equilibrium denaturation mechanism., *Biochemistry 41*, 13744-13754.

26. Jacob, M., Geeves, M., Holtermann, G., and Schmid, F. X. (1999) Diffusional barrier crossing in a two-state protein folding reaction, *Nat Struct Biol 6*, 923-926.

27. Schindler, T., Herrler, M., Marahiel, M. A., and Schmid, F. X. (1995) Extremely rapid protein folding in the absence of intermediates., *Nature Struct. Biol. 2*, 663-673.

28. Grantcharova, V. P. and Baker, D. (1997) Folding dynamics of the src SH3 domain, *Biochemistry 36*, 15685-15692.

29. Grantcharova, V. P., Riddle, D. S., and Baker, D. (2000) Long-range order in the src SH3 folding transition state, *Proc Natl Acad Sci U S A 97*, 7084-7089.

30. Guerois, R. and Serrano, L. (2000) The SH3-fold family: experimental evidence and prediction of variations in the folding pathways, *J Mol Biol 304*, 967-982.

31. Viguera, A. R., Martinez, J. C., Filimonov, V. V., Mateo, P. L., and Serrano, L. (1994) Thermodynamic and kinetic analysis of the SH3 domain of spectrin shows a two-state folding transition, *Biochemistry 33*, 2142-2150.

32. Itzhaki, L. S., Otzen, D. E., and Fersht, A. R. (1995) The structure of the transition state for folding of chymotrypsin inhibitor 2 analysed by protein engineering methods: evidence for a nucleation-condensation mechanism for protein folding, *J Mol Biol 254*, 260-288.

33. Jackson, S. E. and Fersht, A. R. (1991) Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition, *Biochemistry 30*, 10428-10435.

34. Ladurner, A. G. and Fersht, A. R. (1999) Upper limit of the time scale for diffusion and chain collapse in chymotrypsin inhibitor 2, *Nat Struct Biol 6*, 28-31.

35. Bryngelson, J. D., Onuchic, J. N., Socci, N. D., and Wolynes, P. G. (1995) Funnels, pathways, and the energy landscape of protein folding: a synthesis, *Proteins 21*, 167-195.

36. Clementi, C., Nymeyer, H., and Onuchic, J. N. (2000) Topological and energetic factors: what determines the structural details of the transition state ensemble and "en-route" intermediates for protein folding? An investigation for small globular proteins, *J Mol Biol 298*, 937-953.

37. Plotkin, S. S. and Onuchic, J. N. (2002) Understanding protein folding with energy landscape theory. Part I: Basic concepts, *Q Rev Biophys 35*, 111-167.

38. Plotkin, S. S. and Onuchic, J. N. (2002) Understanding protein folding with energy landscape theory. Part II: Quantitative aspects, *Q Rev Biophys 35*, 205-286.

39. Plotkin, S. S. and Wolynes, P. G. (2003) Buffed energy landscapes: another solution to the kinetic paradoxes of protein folding, *Proc Natl Acad Sci U S A 100*, 4417-4422.

40. Sosnick, T. R., Mayne, L., and Englander, S. W. (1996) Molecular collapse: The rate-limiting step in two-state cytochrome c folding, *Proteins 24*, 413-426.

41. Plaxco, K. W., Simons, K. T., and Baker, D. (1998) Contact order, transition state placement and the refolding rates of single domain proteins, *J Mol Biol 277*, 985-994.

42. Bosshard, H. R., Durr, E., Hitz, T., and Jelesarov, I. (2001) Energetics of coiled coil folding: the nature of the transition states, *Biochemistry 40*, 3544-3552.

43. Durr, E., Jelesarov, I., and Bosshard, H. R. (1999) Extremely fast folding of a very stable leucine zipper with a strengthened hydrophobic core and lacking electrostatic interactions between helices, *Biochemistry 38*, 870-880.

44. Zitzewitz, J. A., Bilsel, O., Luo, J., Jones, B. E., and Matthews, C. R. (1995) Probing the folding mechanism of a leucine zipper peptide by stopped-flow circular dichroism spectroscopy, *Biochemistry 34*, 12812-12819.

45. Kim, D. H., Nam, G. H., Jang, D. S., Yun, S., Choi, G., Lee, H. C., and Choi, K. Y. (2001) Roles of dimerization in folding and stability of ketosteroid isomerase from Pseudomonas putida biotype B, *Protein Sci 10*, 741-752.

46. Kim, D. H., Jang, D. S., Nam, G. H., and Choi, K. Y. (2001) Folding mechanism of ketosteroid isomerase from Comamonas testosteroni, *Biochemistry 40*, 5011-5017.

47. Wallace, L. A., Sluis-Cremer, N., and Dirr, H. W. (1998) Equilibrium and kinetic unfolding properties of dimeric human glutathione transferase A1-1, *Biochemistry 37*, 5320-5328.

48. Wallace, L. A. and Dirr, H. W. (1999) Folding and assembly of dimeric human glutathione transferase A1-1, *Biochemistry 38*, 16686-16694.

49. Garel, J. R. (1992) Folding of large proteins: Multidomain and multisubunit proteins., (Creighton, T. E., Ed.) Freeman & Co., London, UK.

50. Jaenicke, R. and Lilie, H. (2000) Folding and association of oligomeric and multimeric proteins, *Adv Protein Chem 53*, 329-401.

51. Doyle, S. M., Braswell, E. H., and Teschke, C. M. (2000) SecA folds via a dimeric intermediate, *Biochemistry 39*, 11667-11676.

52. Apuy, J. L., Park, Z. Y., Swartz, P. D., Dangott, L. J., Russell, D. H., and Baldwin, T. O. (2001) Pulsed-alkylation mass spectrometry for the study of protein folding and dynamics: development and application to the study of a folding/unfolding intermediate of bacterial luciferase, *Biochemistry 40*, 15153-15163.

53. Clark, P. L., Weston, B. F., and Gierasch, L. M. (1998) Probing the folding pathway of a beta-clam protein with single-tryptophan constructs, *Fold Des 3*, 401-412.

54. Raumann, B. E., Rould, M. A., Pabo, C. O., and Sauer, R. T. (1994) DNA recognition by beta-sheets in the Arc repressor-operator crystal structure, *Nature 367*, 754-757.

55. Zhang, R.-G., Joachimiak, A., Lawson, C. L., Schevitz, R. W., Otwinowski, Z., and Sigler, P. B. (1987) The crystal structure of *trp*-aporepressor at 1.8 Å resolution shows how binding tryptophan enhances DNA affinity, *Nature 327*, 591-597.

56. Liu, Y. and Eisenberg, D. (2002) 3D domain swapping: As domains continue to swap, *Protein Sci 11*, 1285-1299.

57. Milla, M. E. and Sauer, R. T. (1994) P22 Arc repressor: folding kinetics of a single-domain, dimeric protein, *Biochemistry 33*, 1125-1133.

58. Srivastava, A. K. and Sauer, R. T. (2000) Evidence for partial secondary structure formation in the transition state for arc repressor refolding and dimerization, *Biochemistry 39*, 8308-8314.

59. Gittelman, M. S. and Matthews, C. R. (1990) Folding and stability of trp aporepressor from Escherichia coli, *Biochemistry 29*, 7011-7020.

60. Gloss, L. M., Simler, B. R., and Matthews, C. R. (2001) Rough energy landscapes in protein folding: dimeric E. coli Trp repressor folds through three parallel channels, *J. Mol. Biol. 312*, 1121-1134.

61. Mann, C. J., Xiao, S., and Matthews, C. R. (1995) Characterization of the slow folding reactions of *trp* aporepressor from *Escherichia coli* by mutational analysis of prolines and catalysis by a peptidyl-prolyl isomerase, *Biochemistry 34*, 14573-14580.

62. Shao, X., Hensley, P., and Matthews, C. R. (1997) Construction and characterization of monomeric tryptophan repressor: a model for an early intermediate in the folding of a dimeric protein, *Biochemistry 36*, 9941-9949.

63. Gloss, L. M. and Matthews, C. R. (1998) The barriers in the bimolecular and unimolecular folding reactions of the dimeric core domain of Escherichia coli Trp repressor are dominated by enthalpic contributions, *Biochemistry 37*, 16000-16010.

64. Gloss, L. M. and Matthews, C. R. (1998) Mechanism of folding of the dimeric core domain of Escherichia coli trp repressor: a nearly diffusion-limited reaction leads to the formation of an on-pathway dimeric intermediate, *Biochemistry 37*, 15990-15999.

65. Arents, G. and Moudrianakis, E. N. (1995) The histone fold: A ubiquitous architectural motif utilized in DNA compaction and protein dimerization, *Proc. Natl. Acad. Sci. U.S.A. 92*, 11170-11174.

66. Banks, D. D. and Gloss, L. M. (2004) Folding mechanism of the (H3-H4)2 histone tetramer of the core nucleosome, *Protein Sci. 13*, 1304-1316.

67. Placek, B. J. and Gloss, L. M. (2005) Three-state kinetic folding mechanism of the H2A/H2B histone heterodimer: the N-terminal tails affect the transition state between a dimeric intermediate and the native dimer, *J. Mol. Biol. 345*, 827-836.

68. Topping, T. B. and Gloss, L. M. (2004) Stability and folding mechanism of mesophilic, thermophilic and hyperthermophilic archael histones: the importance of folding intermediates, *J. Mol. Biol. 342*, 247-260.

69. Kraulis, P. J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures., *J. Applied Crystallography 24*, 946-950.

CHAPTER III

MULTI-PROBE PROTEIN-PROTEIN FÖRSTER RESONANCE ENERGY TRANSFER (FRET) ANALYSIS OF NUCLEOSOME CORE PARTICLES CONTAINING H2A AND H2A.Z

The work in this chapter will be prepared for submission to the *Journal of Molecular Biology.* I am the first author of this body of work and performed the majority of the experiments contained in this chapter. Jessica M Stratton, as an undergraduate researcher under my supervision, provided technical assistance in the construction and purification of several mutants. Ms. Stratton also determined the stability of the H2A-H2B-109Cys-AEDANS dimer. Ms. Stratton will be an author on the submitted manuscript.

SUMMARY

A protein-protein Förster resonance energy transfer (FRET) system, employing probes at multiple positions, was designed to investigate the thermodynamic properties of the nucleosome core particle (NCP), with a focus on the dissociation of the H2A-H2B dimer. Tryptophan and Cys-AEDANS were chosen as the donor and acceptor, respectively, because when compared to the fluorophores used in previous NCP FRET studies, they: 1) are smaller and less hydrophobic which should minimize perturbations of histone and NCP structure; and 2) have an R_0 of 20 Å, which is much less than the dimensions of the NCP (~50 Å width and ~100 Å diameter). CD and FL equilibrium unfolding titrations indicate that the donor and acceptor moieties do not alter the stability of the H2A-H2B dimer and H3-H4 tetramer. NCPs containing the various FRET pairs were reconstituted with the 601 artificial positioning sequence. The equilibrium NaCl-induced dissociation of the modified NCPs showed an increase in stability of the H2A-H2B dimer as compared to previous studies using weaker positioning sequences. The free energy of dissociation, ΔG, determined from reversible and well-defined transitions, revealed two distinct transitions depicting the dissociation of each dimer. There is an apparent cooperativity in the dissociation of the dimers. The incorporation of the histone variant H2A.Z into nucleosomes has been intensely studied both *in vivo* and *in vitro* with conflicting results. Biophysical studies on the structural consequences of H2A.Z incorporation have suggested both destabilizing and stabilizing effects. Comparison of the H2A and H2A.Z FRET NCP dissociation curves indicate no significant overall change in stabilities, but the dissociation of the H2A.Z-H2B dimers appears to be more cooperative.

ABBREVIATIONS

β-ME, beta-mercaptoethanol; CD, circular dichroism; Cm, midpoint of equilibrium titration curve; Fapp, apparent fraction dissociated; FL, fluorescence; FRET, Förster resonance energy transfer; ΔG^{o} (H_2O), the free energy of unfolding in the absence of denaturant; GdmCl, guanidinium chloride; KCl, potassium chloride; KPi, potassium phosphate, pH 7.2; *m* value, parameter describing the sensitivity of the unfolding transition to the [denaturant]; NCP, nucleosome core particle; PEG, polyethylene glycol; TMAO, trimethylamine-N-oxide; WT, wild-type**.**

INTRODUCTION

The fundamental repeating unit for DNA packaging into chromatin is the nucleosome core particle (NCP). Approximately 150 bp of DNA are wrapped around an octamer of histones, consisting of two copies each of H2A, H2B, H3 and H4. Although it was initially thought to be a static, highly regular packaging system, the nucleosome is now known to be a dynamic structure (*1*). Changes in the properties of the NCP are critical in the regulation of DNA-dependent processes, such as transcription, replication, recombination and repair. The appropriate regulation of the NCP and its packaging of DNA is essential for proper cellular function (for review, (*2-4*)).

The four histones of the core nucleosome contain a common three helix, heterodimerization motif, and the N-terminal tails as well as C-terminal extensions in the case of H2A and H2B (*5*). The H3-H4 dimer further oligomerizes to a tetramer through a four helix bundle comprised of the C-terminal helical regions of H3 and H3'. In the histone octamer, C-terminal regions of H2B and H4 form two additional four helix bundles in the octamer (*6*). Salt bridges and hydrogen bonds between buried charged side chains stabilize the four-helix bundles, as well as burial of hydrophobic surface area (*6*). H2A also participates in important interfaces in the octamer. The L1 loops of H2A and H2A' interact through multiple hydrogen bonds; it has been hypothesized that stabilization arising from these interactions may contribute to the cooperative association of the H2A-H2B dimers in formation of the NCP (*6-8*). The "docking domain" of the C-terminal tail of H2A also makes extensive contacts with the H3-H4 dimer and contributes to the helical ramp for the DNA (*6*).

The NCP assembly/disassembly process is similar in the cell and with purified proteins (*9;10*). (Figure 3.1A) $(H3\text{-}H4)_2$ tetramer binds first and positions the central portion of the ~150

bp piece of DNA wrapped in the NCP. The H2A-H2B dimers bind to surfaces on each side of the tetramer, completing the two helical ramps that position the 1.65 superhelical turns of DNA. Overall ~28 bp of DNA are organized by each H2A-H2B dimer with 4 bp stretches in between (for review, (*11*)). In the cell, histone deposition is controlled by histone chaperones such as CAF-1, RCAF and NAP-1 (for review, (*12*)). Cell-free reconstitutions of the NCP generally rely on the differential sensitivities to ionic strength of the association steps. As NaCl concentrations approach physiological ionic strength, the mobility of the DNA ends increase. The dissociation of the H2A-H2B dimers begins at ~0.4 M NaCl, and the $(H3\text{-}H4)_2$ tetramer dissociates above ~1 M NaCl (*9;13*).

The intrinsic dynamics of the NCP include the movement of either the DNA or the histone components themselves in the absence of any assistance from remodeling complexes. Widom and colleagues showed that target sites wrapped inside the NCP are transiently exposed to DNA-binding proteins of various sizes (*14;15*). The extent of transient exposure is DNA sequence dependent and decreases for target sites closer to the dyad axis of the NCP. The passive access of DNA-binding proteins, through transient exposure, is highly cooperative; binding of one protein greatly increases the likelihood of binding at a nearby site (*15*). Two mechanisms have been proposed for how a target site within the NCP is exposed (in the absence of additional factors): 1) translocation (or sliding) of the histone octamer along the DNA (*16-18*); or 2) spontaneous site-exposure of the DNA through dynamic breathing (*14*). The former mechanism was discounted by studies using DNA molecules with multiple positioning sequences of different strength. Competition assays showed that the histone octamer was unable to translocate to stronger positioning sequences on the DNA, nor was there a time-dependent shift from a "disfavored" nucleosome position to a more favorable one (*19*). Equilibrium and

kinetic studies of FRET (Förster resonance energy transfer) systems have been used to monitor the dynamic mobility of the DNA ends in the NCP (*20;21*). Spontaneous wrapping and unwrapping occurs on the millisecond timescale. The lifetime of the fully wrapped state was determined to be approximately 250 ms and the lifetime of unwrapped state was 10 to 50 ms (*21*). Therefore, a DNA-binding protein has a short but frequent opportunity to bind its target sequence through transient site exposure. Studies utilizing a FRET pair closer to the dyad access showed that an even larger scale, long-distance nucleosome unwrapping occurs rapidly (*22*). The long range DNA "on" and "off" timescales were determined to be 2.1 – 4.6 ms (depending on the [salt]) and 100 – 200 ms, respectively. The NCPs were found to be in the closed state over 95% of the time but once open, there was sufficient time to allow binding of additional proteins (*22*). Clearly, the nucleosome is in a state of constant flux between wrapped and unwrapped species that allows access to various DNA sequences. The magnitude and rate of this flux may be influenced by the incorporation of histone variants, post-translational modification and the actions of various chaperone and remodeling complexes. FRET experiments have been already proven to be useful in characterizing NCP dynamics and will continue to be a powerful tool to dissect this process.

Histone variants constitute a small, but essential fraction of a cell's histone complement and play distinct functional, and presumably structural, roles in chromatin dynamics (*23-28*). The greatest number of variants have been identified for H2A including the major H2A (H2A.1), H2A.Z, H2A.X, H2A-Bbd, and MacroH2A (*29*). The variants constitute ≤ 20% of the total H2A population and vary from 60% to 90% in sequence identity with the major H2A. They are distinguished from H2A.1 mainly in the sequence and length of the C-terminal α-helix and extended tail although changes also commonly occur in the L1 loop. Of these variants, H2A.Z is

the highest abundance in the genome after H2A.1, and has been highly studied. While yeast H2A.Z deficient strains are viable, this variant is essential in all higher eukaryotes (for review, (*28*)). Various studies have implicated H2A.Z in transcriptional activation and repression (*30-38*). An emerging view is that H2A.Z plays a role in defining and maintaining the boundaries of heterochromatin (*31;34;36*).

The X-ray crystal structure of H2A.Z NCP was very similar to that of NCPs containing the major H2A (*6;39;40*) with minimal changes in protein-DNA contacts. Three regions exhibited subtle structural differences that may influence biological function: 1) the interface between H2A and the H3-H4 dimer; 2) a novel metal binding site in H2A.Z; and 3) the H2A-H2A' L1 loop interactions. 1) Sequence differences in the C-terminal docking domains caused small localized changes in the interactions of H2A with H3 and H4. The substitution of major H2A Gln-104 with Gly in H2A.Z results in the loss of three hydrogen bonds—which may be potentially destabilizing. 2) Several changes were observed on the large protein surface of the NCP. His-112 and His-114 of H2A.Z formed a novel metal binding site, occupied by Mn^{2+} from the crystallization buffer. This site is adjacent to an acidic patch, which is slightly larger in the H2A.Z NCP. Asn-94 and Lys-95 in the major H2A are replaced by Asp and Ser, respectively, in H2A.Z. (*39*). Other sequence changes generate a pore in the H2A.Z NCP that would be occupied by K95 and R99 in major H2A NCP. These surface alterations may recruit a different cadre of histone chaperones or chromatin remodeling factors (*39*). 3) There are small changes in the interaction of the L1 loops of H2A and H2A'. The relative orientation of the L1 loops in the nucleosome is slightly altered as well as significant changes in the intermolecular hydrogen bonding between the two dimers—two and six hydrogen bonds in the H2A and H2A.Z NCPs, respectively. It was proposed that these changes would increase in the binding cooperativity for

H2A.Z-H2B dimers and disfavor "mixed" NCPs containing both major H2A and H2A.Z histones.

In addition to the debate on the biological function of H2A.Z, it is also unclear whether or not H2A.Z forms more stable NCPs. Early studies showed that incorporation of H2A.Z increased inherent nucleosome mobility (*41*). Other biophysical studies provided conflicting reports as to whether H2A.Z NCPs were destabilized (*42*) or stabilized (*13;42*). It has recently reported been reported that H2A.Z-NCPs were slightly stabilized in the absence of acetylation (*43*). Additional biophysical studies are required to confirm the relative stabilities of the H2A and H2A.Z containing NCPs.

The dynamics and assembly of the NCP are well suited for Förster resonance energy transfer (FRET) measurements. FRET has been used to monitor protein-protein, protein-nucleic acid and nucleic acid-nucleic acid interactions in diverse research areas from protein folding to macromolecular assembly and protein networking—both in the cell and in purified systems. FRET is a fluorescent-based method that can assess the distance between two interacting fluorophores, which may reside on a single macromolecule or on different partners in a macromolecular complex (*44;45*). Efficient FRET requires appropriately matched donor and acceptor fluorophores, dictated by the following criteria: 1) the extent of spectral overlap between the emission of the donor and the absorption of acceptor; 2) the quantum yield of the donor; 3) the relative orientation of the transition dipoles of the donor and acceptor fluorophores; and 4) the distance between the donor and acceptor moieties. The rate of energy transfer, k_T, between a donor and an acceptor is strongly dependent on the distance between the donor and acceptor (D-A) pair and is described by Equation 1:

$$k_T = \frac{1}{\tau_D}\left(\frac{R_O}{r}\right)^6 \qquad (1)$$

where τ_D is the decay relaxation time of the donor in the absence of acceptor, R_O is the Förster distance, and r is the distance between the donor and acceptor. The Förster distance is determined by the spectral properties of the components of the FRET pair. R_O is proportional to the overlap integral (J), the donor quantum yield (Φ_D) in the absence of acceptor and the orientation factor (κ^2). Transfer efficiency, E, is the ratio of the transfer rate, k_T, to the sum of the rates for the two decay modes of the donor (fluorescence and transfer):

$$E = \frac{k_T}{\tau_D^{-1} + k_T} \qquad (2)$$

Combining Equations 1 and 2 describes the transfer efficiency in terms of the Förster distance and D-A distance:

$$E = \frac{R_O^6}{R_O^6 + r^6} \qquad (3)$$

with limits of 1 (maximal FRET) and 0, when no FRET occurs because of large donor-acceptor separation. Equation 3 defines the relationship between the sensitivity (efficiency) of FRET and the distance between the donor-acceptor pair. When the Förster distance is equal to the D-A distance, the transfer efficiency is 50%. As r increases, FRET efficiency decreases precipitously; when the D-A distance is twice the R_O, the efficiency is ~1.5%.

In the design of a FRET system to study dynamics or relative distances, it is important to position the donor-acceptor pair at or near the Förster distance. Extrinsic donor and acceptor fluorophores, introduced through chemical modification, are commonly used in FRET

experiments because they afford a range of available probes and with varied R_0 values (10-80 Å). Extrinsic FRET functionalities can be introduced by modification of Cys residues, which can be complicated by the presence of multiple naturally occurring cysteines. Except for a non-conserved, unessential Cys in the C-terminal helix of H3, histones are devoid of Cys residues. In protein systems, intrinsic tryptophan fluorophores can be used as a donor chromophore. A potential disadvantage is the presence of multiple Trp donors in a protein; however, histones have no naturally occurring Trp residues. The disadvantage of using a Trp donor is the limited number of appropriate acceptor chromophores and their relatively short Förster distances (~15-30 Å). An advantage of a Trp-acceptor pair is the relatively small size of the chromophores compared to other D-A pairs, which means their introduction is less likely to perturb the structure, function and stability of a protein.

The FRET system used in this study employs tryptophan residues as donors and Cys modified with IAEDANS (5-(((2-iodoacetyl)amino)ethyl)amino)naphthalene-1-sulfonic acid) as acceptors. Unique Trp donors were engineered into the H3 and H4 histones. Single Cys residues, targets for IAEDANS modification, were engineered into H2A and H2B. Various combinations of the modified histones were reconstituted into NCPs using the 601 artificial positioning sequence. Five protein-protein FRET pairs were designed to monitor the salt-induced dissociation of the NCP. This system is sensitive to the dissociation of the dimer and relatively insensitive to other transitions such as DNA breathing or tetramer dissociation.

MATERIALS AND METHODS

Materials.

Ultrapure urea was purchased from MP Biomedicals (Solon, OH). Guanidinium chloride was purchased from VWR International (West Chester, PA). Trimethylamine *N*-oxide (TMAO) was purchased from Sigma (St. Louis, MO) and prepared as described previously (*46*). 1,5-IAEDANS was purchase from Invitrogen (Carlsbad, CA). The oligonucleotides used for mutagenesis and the production of pLMG601-23, were purchased from MWG-BIOTECH Inc. (High Point, NC).

Methods.

Sample preparation.

Histone Mutagenesis, Expression and Purification. The T_7 pET expression plasmids containing the *Xenopus laevis* histone genes (*47*) were used as templates for site-directed mutagenesis. The QuikChange® II Mutagenesis Kit from Stratagene (La Jolla, CA) was used to generate the following histone mutations: H2A-L108C, H2B-S109C, H3-F78W, H4-L49W and H4-V60W. The presence of the desired mutation, and absence of other mutations, was confirmed by DNA sequencing of the entire histone gene. The recombinant histones were overexpressed in the BL21(DE3)pLys *E. coli* strain and purified as described previously (*46;48*). H2A.Z was overexpressed in a pET vector containing the mouse gene (*39*) and purified by methods described previously (*49*). As analyzed by SDS-PAGE, the purity of the histone preparations was greater than 90%. Purified monomers were dialyzed against 10 mM HCl, 3 mM β-mercaptoethanol and in the case of H3 and H4, were lyophilized.

Histone Labeling and Refolding. Cys-containing H2A and H2B monomers were modified with the thiol reactive reagent 1,5-IAEDANS, 5-((((2-iodoacetyl)amino)ethyl)-amino)-naphthalene-1-sulfonic acid, from Invitrogen (Carlsbad, CA). The modification buffer was 6 M urea, 50 mM KCl and 20 mM Tris-Cl, pH 7.5 without any reducing agent. 4 mg/ml IAEDANS, in 20 mM Tris-Cl pH 7.5 was added slowly, with stirring, to urea unfolded monomers, to a final molar ratio of 20:1 IAEDANS to cysteine. After a three hour room temperature incubation (under vacuum), the labeling reaction was quenched with a 10 fold excess of β-ME followed by exhaustive dialysis into a buffer of 6 M urea, 20 mM KPi pH 7.2, 0.2 M KCl, 0.1 mM K_2EDTA, 3 mM β- ME for quantitation and reconstitution to the heterodimer.

The Cys-AEDANS content of the monomers was calculated using an extinction coefficient of 5,700 $M^{-1}cm^{-1}$ at 336 nm, and compared to the protein concentration determined by BCA assay (reagents from Sigma, St. Louis, MO). The extent of labeling was found to be ≥90%. This was confirmed by triton/acid/urea gels (TAU), which are able to resolve histone post-translational modifications (*50*).

Purified recombinant histones were refolded to their appropriate hetero-oligomeric form by methods described previously (*46;48*). Briefly, urea-unfolded H2A and H2B monomers were combined in a 1:1 ratio and refolded by 5-fold dilution into buffer without urea. Additional dialysis was performed to remove residual urea. The refolded H2A-H2B dimers were further purified with a Heparin column as described previously (*48*). Lyophilized H3 and H4 monomers were dissolved in a buffer of 20 mM KPi, pH 7.4, 5 mM K_2EDTA, 5 mM β-ME and 8 M GdmCl, and then diluted 10-fold into buffer without denaturant. The reconstituted tetramer preparation was dialyzed against buffer to remove residual GdmCl. The sample was concentrated in the dialysis tubing by placing the tubing in solid PEG 10,000, followed by

further dialysis. The concentrations of the folded WT oligomers and those containing Trp mutations were determined in 6 M GdmCl buffer using extinction coefficients calculated by the method of Gill and von Hippel (*51*). The concentrations of H2A-H2B dimers modified by IAEDANS were determined using the BCA assay (*52*) because of the overlapping absorbance of Cys-AEDANS at 280 nm.

601 DNA Production. The NCPs reconstituted used a 149 bp segment of the unnatural 601 sequence (*18;53*). The central 149 bp of the 601 sequence was amplified by PCR from the pGEM3Z plasmid, kindly provided by the lab of Jonathan Widom (Northwestern University). To enhance production of this DNA fragment, a pUC119-based plasmid, pLMG601-23 was constructed which contained 23 tandem repeats of the 149 bp of 601 sequence flanked by EcoRV restriction sites. The pLMG601-23 plasmid was produced and extracted from *E. coli* DH5 using a modified alkali lysis protocol described previously (*54*). 149 bp fragments of the 601 sequence were released by digestion with EcoRV and isolated with differential precipitation using PEG 6000 (*54*). The extent of digestion and isolation was monitored by PAGE on a 5% native gel. The concentration of the 601 fragments was determined by absorbance at 260 nm.

Octamer and Nucleosome Core Particle Formation. Folded H2A-H2B dimers and H3-H4 tetramers were combined at equimolar ratios in a buffer of 20 mM KPi, pH 7.45, 5 mM Na_2EDTA, 5 mM -ME and 2 M KCl, and dialyzed against this buffer to allow formation of histone octamers. After concentration with a Centricon YM-10 from Millipore Inc. (Billerica, MA), the octamer preparation was purified on a BIOSEP SEC-3000 HPLC column, equilibrated with the 2 M KCl buffer described above, to remove any improperly assembled species.

Fractions were analyzed by 15 % SDS-PAGE, and those with the appropriate histone ratio of 1:1:2 H3:H4: H2A-H2B (which co-migrate) were pooled.

NCP reconstitution was achieved by combining DNA and histone octamers in 2 M KCl buffer and lowering the ionic strength by step-wise dialysis against buffers with 0.85, 0.6, 0.2 and 0 M KCl. The concentrations of the histone octamers were estimated from the Cys-AEDANS absorbance at 336 nm. Since the labeling efficiency was not 100%, the octamer concentration was slightly inaccurate. Therefore, the optimal DNA:octamer ratio for NCP reconstitution was determined by setting up 50 µl reactions with DNA and octamer ratios varied from 0.4:1 to 1.2:1, with dialysis in mini-buttons (*55*). The quality of the reconstitutions were analyzed by 5% native PAGE to identify the DNA:octamer ratio which yielded the greatest amount of NCP and minimal free DNA, typically a ratio of 0.8:1. This ratio was then used in the large-scale reconstitutions (~0.6 mg/ml NCP) of volumes of 4 to 6 ml by conventional dialysis methods (6 – 8K MWCO). After dialysis, the newly reconstituted NCPs were centrifuged at 16.1 X g at 25 °C for 5 min to remove any aggregated species.

Data collection and analyses.

Instrumentation. Spectroscopy data were collected on an AVIV 202SF circular dichroism spectrophotometer and an AVIV ATF-105/305 differential/ratio spectrofluorometer. Both instruments were interfaced with Hamilton Model 500 automated titrators, used for collecting equilibrium stability data. The spectrometers were equipped with peltier-controlled sample holders; all spectroscopic data were collected at 25 °C.

H2A-H2B Dimer and H3-H4 Tetramer Equilibrium Stability Studies. Protein unfolding was monitored by the CD signal at 222 nm, with a 2 min equilibration time at each denaturant concentration. The equilibration times were significantly longer than the slowest observed folding and unfolding kinetic phases for the oligomers (*46*). The Cys-AEDANS FL of the modified H2A-H2B dimers was largely insensitive to protein unfolding, as expected for a highly solvent exposed fluorophore. Therefore, intrinsic Tyr FL emission at 305 nm was used to monitor quaternary and tertiary unfolding, with excitation at 280 nm. Trp FL emission at 340 nm was monitored for H3-H4 unfolding, with excitation at 290 nm.

The buffer conditions (20 mM KPi, pH 7.2; 0.2 M KCl, 0.1 mM EDTA) were the same used in previous studies (*46;48*) to allow comparison to results obtained with WT recombinant histones. Urea was used as the denaturant in the H2A-H2B equilibrium titrations. H3-H4 titrations employed GdmCl as the denaturant to improve the reversibility of tetramer unfolding (*46*); the buffers also contained 1 M TMAO to stabilize the tetramer and yield a well-defined native baseline which allows quantitative fitting of the thermodynamic data.

Equilibrium titrations at different protein concentrations or monitored by different spectroscopic properties can be normalized to F_{app}, the apparent fraction of unfolded monomers by Equation 4:

$$F_{app} = \frac{(Y_i - Y_N)}{(Y_U - Y_N)} \qquad (4)$$

where Y_i is the FL at a given [denaturant]; Y_N and Y_U describe the folded and unfolded baselines, respectively. Equation 5 relates the equilibrium constant, K_{Eq}, for a dimeric system to the F_{app} values:

$$F_{app} = \frac{\square K_{Eq} + \sqrt{K_{Eq}^2 + 8K_{Eq}P_T}}{4[P_T]} \qquad (5)$$

from which free energy of unfolding values, ΔG°, can be calculated with Equation 6:

$$\Delta G° = -RT \bullet \ln(K_{Eq}) \qquad (6)$$

The free energy of unfolding in the absence of denaturant at a standard state of 1 M dimer, $\Delta G°(H_2O)$ was determined using a linear extrapolation between ΔG° values and the concentration of denaturant (*46;48*):

$$\Delta G° = \Delta G°(H_2O) - m[\text{denaturant}] \qquad (7)$$

where the *m* value is the sensitivity of the transition to denaturant and correlates to the change in solvent accessible surface area upon unfolding (*56*). The data-fitting program Savuka 5.1 (*57;58*) was used to globally fit data from different spectral probes and protein concentrations. The error at one standard deviation of the fitted parameters were estimated using a rigorous analyses of the error surface as described elsewhere (*48;59*).

FL anisotropy of modified histones oligomers. The FL anisotropy of the Cys-AEDANS and Trp residues were measured for the folded and denaturant-unfolded histone oligomers and the folded and salt-unfolded NCPs. The G-factor (G), the ratio of the sensitivity of the instrument monochromaters for vertically and horizontally polarized light (*44*), was determined for each emission wavelength. Anisotropy (r) is defined by Equation 8:

$$r = \frac{I_{VV} \square GI_{VH}}{I_{VV} + 2GI_{VH}} \qquad (8)$$

The incoming excitation light is linearly polarized with an orientation along the vertical axis. The FL emission intensity is characterized by two components: 1) I_{VV}, the intensity that is

parallel to the electric vector of the polarized excitation light; and 2) I_{VH}, the intensity which is perpendicular.

Steady-state FL and FRET in the NCPs. A buffer of 20 mM Tris-Cl pH 7.6, 1 mM EDTA and 1 mM β-ME was used in all NCP studies. The NCP concentration in the salt-induced dissociation transitions was 250 nM, comparable to a previous NCP FRET study (*13*). Excitation was at 290 nm, and emission was monitored at both 350 and 490 nm for Trp and Cys-AEDANS, respectively. A relatively narrow excitation bandwidth of 1 nm was employed to minimize photobleaching, with an emission bandwidth of 4 nm. Salt-dependent FL intensities were determined by titrating NCP samples with buffer containing 5 M NaCl, resulting in less than 10% change in NCP concentration. Samples were equilibrated for 5 minutes after each incremental salt increase of 0.025 to 0.05 M NaCl. Data points represent the average of two or three measurements from transitions initiated at different salt concentrations.

The FL of the Trp donor (350 nm) and Cys-AEDANS acceptor (490 nm) were normalized to F_{app}, apparent fraction of NCP dissociated to H2A-H2B dimers and tetramer-DNA complex, as defined by Equation 9:

$$F_{app} = \frac{(Y_i - Y_{NCP})}{(Y_{T,D} - Y_{NCP})} \qquad (9)$$

where Y_i is the FL at a given [NaCl]; Y_{NCP} describes the linear response of the FL signal of the intact NCP (*i.e.* the folded baseline) and $Y_{T,D}$ is the linear response observed for the H3-H4/DNA complex after dissociation of H2A-H2B dimers. From the F_{app} values in the transition region (0.5 to 1.3 M NaCl), the concentrations of the NCP, H2A-H2B dimer and tetramer-DNA complex were calculated to determine an equilibrium constant (Equation 10):

$$K_{Eq} = \frac{[\mathrm{D}]^2[T \bullet DNA]}{[\mathrm{NCP}]} \qquad (10)$$

and the related ΔG° values by Equation 6.

RESULTS

Design of NCP components for FRET studies.

The histone octamer packages ~150 bp of DNA, which for many DNA sequences can populate multiple translational and rotational positions. To ensure a NCP population with homogeneous positioning, the current studies employed one of the strongest nucleosome positioning sequences characterized to date, the 601 sequence. This sequence was isolated by the Widom lab from a random library of synthesized DNA sequences in SELEX experiments based on its high affinity for nucleosome formation and accompanying strong positioning ability (*53*). The affinity of the 601 sequence for the histone octamer is ~150 times greater than that of the commonly used sea urchin 5S positioning sequence (*18*). This 5S sequence was used in a previous FRET study on NCPs reconstituted with recombinant histones (*13*).

Previous NCP FRET systems have used derivatives of coumarin, fluorescein, Cy3 and Cy5 (*13;20;22*). The Trp to Cys-AEDANS pair was chosen because: 1) the relatively small size and hydrophobic surface area of the fluorophores will minimize any tendency to interact with hydrophobic patches on or near the surface of the proteins; and 2) the R_0, 20 Å, is substantially less than the dimensions of the NCP (~50 Å width and ~100 Å diameter).

The choice of sites for introduction of the FRET donors (Trp on H3 and H4) and acceptors (Cys-AEDANS on H2A and H2B) was based on inspection of the crystal structure of the NCP (*6*). Four factors were considered in the choice of sites (Table 3.1 and Figure 3.1B): 1) the absence of Trp residues in the four histones and Cys residues in H2A and H2B; 2) high

solvent accessibility to permit rotational freedom of the fluorophore; 3) hydrophobicity of the WT residue to minimize structural perturbation; 4) proximity of the C_{α} atoms of the donor and acceptor sites. The H2A-L108C mutation is at the beginning of the C-terminal tail, which has close protein-protein interactions with the histone octamer. The H2B-S109C mutation is on the surface of the C-terminal helix, beyond the canonical histone fold. The H3-F78W mutation is the C-terminal residue of the α1 helix of the canonical histone fold. Because of distance requirements, Trp residues were introduced on the surface of the central helix, α2, of the histone fold of H4 by mutation of Leu-49 and Val-60. Five protein-protein FRET NCPs (Table 3.1) were designed to monitor the dissociation of H2A•H2B dimer. The distances from one donor to each acceptor (D-A and D-A') are similar for FRET NCPs where the acceptor is located on histone H2A, but dissimilar for H2B-acceptor NCPs.

Fluorescence emission and anisotropy of engineered Cys-AEDANS and Trp residues

Fluorescence emission spectra of folded and unfolded dimers and tetramers were compared to assess the solvent accessibility of the FRET donors and acceptors (Table 3.2). The emission wavelength maximum should red shift very little upon the addition of denaturants if the fluorophore is highly exposed to the solvent in the native oligomer. The folded H2A-H2B dimers containing the Cys-AEDANS acceptor had emission maxima of 485 to 490 nm; upon unfolding, there was a minimal red shift of 5 nm for the emission maxima. The Trp FL of the folded H3-H4 tetramers exhibited emission maxima of 345 to 350 nm, with 5 nm red shifts upon GdmCl-induced unfolding. These data demonstrate that the FRET donors and acceptors have high solvent accessibility in the folded histone oligomers.

If the fluorophores are not restricted by partial burial or interactions with hydrophobic surfaces on the histones or the NCP, they should also exhibit a high degree of rotational freedom, which is important for FRET efficiency. FL anisotropy is a measure of a fluorophore's rotational freedom through the detection of the relative orientations of the specific absorption and emission transition moments dictated by the structure of the fluorophore. The anisotropy parameter, r, is defined by Equation 8. A rigid fluorophore that doesn't tumble or reorient on the timescale of the fluorescence decay will emit photons with the same polarization as the incoming excitation beam. If the fluorophore is reorienting rapidly relative to the rate of fluorescence decay, the polarization of the excitation light will become scrambled, so that the emission has both vertical and horizontal components. Given the random orientation of molecules in solution, the theoretical maximum r value for a rigid, immobile fluorophore is 0.4, while the theoretical minimum is -0.2 (*44*). The relative magnitude of the r value depends on the environment of the fluorophore and the size of the protein to which it is attached; *i.e.* whether it is: 1) buried and rotation is limited by the tumbling of the macromolecular complex; 2) surface exposed and tumbles somewhat faster than the macromolecular complex; or 3) freely rotating in solvent with a rate that is independent of the tumbling time of the complex. Typical anisotropy values reported for immobile Trp range from 0.2 to 0.3 (*44*).

The anisotropy of the engineered Cys-AEDANS and Trp residues were determined under native conditions, as well as unfolded conditions, which should reflect the maximal anisotropy of the system (Table 3.2). The r values for the Cys-AEDANS fluorophores are ~0.06 in the native H2A-H2B dimers and decrease to ~0.02 upon unfolding. Consistent with a larger molecular weight oligomer and a shorter linker length between the fluorophore and the protein, the Trp residues in the H3-H4 tetramer exhibit r values between 0.1 and 0.08 in the native state and 0.06

to 0.04 in the unfolded species. In sum, the fluorescence anisotropies (Table 3.2) and emission maxima are consistent with FRET fluorophores that are solvent-exposed with substantial rotational freedom.

The stabilities of the FRET-engineered histone oligomers are not significantly altered

It is important to assess the effect of incorporation of the FRET fluorophores on histone stability; changes in protein stability may indicate significant structural alterations, which could alter NCP stability and dynamics. The stability to denaturant-induced unfolding of the engineered histone oligomers was determined by monitoring far-UV circular dichroism (CD) and tyrosine or tryptophan FL. Like the wild-type histone oligomers, the unfolding of the FRET-engineered histones was highly reversible, two-state and protein concentration dependent (Figures 3.2 to 3.5). The reversibility and absence of hysteresis was indicated by the coincidence of transitions of folded protein titrated with denaturant-unfolded protein and unfolded protein titrated with folded protein (Figure 3.2 and 3.4). Two-state unfolding was indicated by the similarity of far-UV CD and FL transitions, which report on secondary and tertiary/quaternary structure, respectively. The agreement of the protein-concentration dependence of equilibrium data with global fits to oligomeric two-state models provides further support for the absence of equilibrium intermediates.

The stabilities of the Cys-AEDANS H2A-H2B variants were determined from urea-induced unfolding transitions (Figures 3.2 and 3.3; Table 3.2). Titrations were performed at monomer concentrations of 5 and 10 μM, monitored by both far-UV CD and Tyr FL. The fluorescence of the solvent exposed Cys-AEDANS moiety was not used to monitor unfolding because there was no discernable difference between the folded and unfolded baselines. For

each H2A-H2B variant, a data set of six titrations was analyzed, including varied monomer concentrations, two spectroscopic probes and both folding and unfolding titrations. The fitted parameters from the global analyses are given in Table 3.2, and representative F_{app} transitions are shown in Figure 3.2. F_{app} curves monitored by CD and FL, at a given protein concentration, were superimposable as expected for a two-state equilibrium process, with no detectable intermediates. The stabilities of the modified H2A-H2B dimers are very similar to the WT dimer, exhibiting similar transition midpoints, C_M (Figure 3.3). The $\Delta G°(H_2O)$ and the *m* values for the variants are within the error of one standard deviation and indicate that there are no significant changes in structure.

The tryptophan H3-H4 variants were denatured with guanidinium chloride, GdmCl, as in previous wild-type studies (*46*). The instability of the H3-H4 oligomers yielded poorly defined native baselines, which complicated quantitative analyses. To stabilize the protein, 1.0 M trimethylamine-*N*-oxide (TMAO) was used, a naturally occurring osmolyte found in coelocanths, sharks, and rays which stabilizes the organisms' proteins against high urea concentrations in their serum. TMAO stabilizes the native state through preferential hydration (*46;60-62*) and extends the native, pre-transition baseline region in equilibrium denaturation studies (*46;63*).

Denaturation of the variant H3-H4 oligomers was monitored by far-UV CD and tryptophan FL at monomer concentrations of 2 and 4 μM. Data sets of 5 to 6 titrations were collected for each H3-H4 tryptophan variant. Despite the solvent-exposure of the Trp residues, their FL was still sensitive to local tertiary structural changes, unlike the Cys-AEDANS residues. As suggested by FL anisotropy, this difference may reflect that Trp is more closely tethered to the central histone fold compared to the structure and positions of the Cys-AEDANS residues. The far-UV CD signal of the WT $(H3\text{-}H4)_2$ tetramer was insensitive to tetramer dissociation to

native heterodimers, and the data were well described by a two-state dimeric mechanism (*46*). FL of the multiple Tyr residues of WT $(H3\text{-}H4)_2$ reported on both tetramer dissociation and dimer unfolding, requiring fits to a three-state mechanism. The FL of the single Trp residues are insensitive to tetramer dissociation, which is consistent with their lack of proximity to the H3-H3' dimerization interface. The transitions monitored by Trp FL and far-UV CD were coincident for the variant H3-H4 oligomers. The data sets containing FL and CD titrations at multiple protein concentrations were well-described by global fits to a two-state dimer unfolding model (as described above for H2A-H2B variants). The reduced Chi-squared values were not improved significantly in fits to a three-state model with tetramer dissociation and dimer unfolding.

The fitted parameters for the two-state global fit and representative F_{app} transitions are presented in Table 3.2 and Figure 3.4, respectively. The $\Delta G°(H_2O)$ values for the heterodimers containing H3-78W and H4-60W were identical to that of the WT H3-H4 dimer; the *m* values are slightly increased, resulting in a minor change in the C_M values (Figure 3.5). Trp-49 in the $\alpha 2$ helix of H4 slightly stabilizes the heterodimer as assessed by the $\Delta G°(H_2O)$ and *m* values, although the C_M value is identical to WT H3-H4. In summary, the introduction of Trp residues into H3 and H4 has minimal effects on the stability, and presumably the structure, of the H3-H4 heterodimer.

Salt-induced dissociation of the WT NCP

As in protein folding studies, perturbation methods are used to characterize NCP stability; the most commonly employed dissociation perturbant is salt, specifically NaCl. The stability of unmodified NCPs have been previously monitored by the intrinsic FL of the 30 Tyr

residues distributed throughout the eight core histones (*9;13;47*). At low ionic strength, the Tyr FL is quenched through direct interactions with the DNA in the fully assembled NCP. Salt-induced dissociation of the histone oligomers results in an increase in the Tyr FL intensity.

For comparison to published studies with other DNA sequences, the unfolding of 601-containing NCPs reconstituted with WT histones were monitored by the Tyr FL. Normalized Tyr FL data are shown in Figure 3.6. The description of multiple phases (Figure 3.1A) has been previously reported (*9;13*). The dissociation of the H2A-H2B dimers begins at ~0.6 M NaCl, compared to 0.3 to 0.4 M in NCPs reconstituted with the commonly used sea urchin 5S sequence (*13*). This suggests that the H2A-H2B dimer forms a more stable protein-DNA complex in the 601 NCPs. For both the 601 and 5S DNAs, complete histone dissociation is observed by 2 M NaCl.

Characterization of FRET nucleosome core particles.

Histones and 601 DNA were mixed in the appropriate molar ratios in 2 M KCl and reconstituted into NCPs by dialysis (*54*). NCPs were reconstituted with the appropriate FRET donor and acceptor pairs (Table 3.1) as well as with only donor- or acceptor-containing histones. Proper nucleosome formation was examined by native polyacrylamide gel electrophoresis with ethidium bromide staining (Figure 3.7A). The native gels were also stained with Coomassie Blue to visualize the protein component, which co-migrated with the DNA (data not shown). Similar gel mobility and homogeneity was observed for unmodified (WT) NCPs and those containing only donors or acceptors (data not shown).

The salt-dependence of the fluorescence of NCPs containing only donor or acceptor fluorophores was examined. There was no significant FL change for any of the Trp donors

between 0 and 2.4 M NaCl. H2A-108Cys-AEDANS NCPs exhibited a small linear increase in FL with salt, resulting in a 10% change between 0 and 2 M NaCl. This slight salt dependence is consistent with a solvent effect, rather than sensitivity to any structural transitions. The FL of the H2B-109Cys-AEDANS decreased ~20% between 0.5 and 1.5 M NaCl, suggesting a minor FL response to dimer dissociation; however, this FL change is much less than that observed in the FRET experiments (see below). To avoid potential complications, analysis of FRET data focused on Trp donor FL in the salt-induced dissociation of NCPs with the H2B acceptor.

To assess the FRET probes' rotational freedom, the anisotropies were determined for NCPs with only donor or acceptor fluorophores (Table 3.2). In the context of the NCP, H2A and H2B Cys-AEDANS residues exhibit r values that are intermediate between those of the folded dimer and unfolded monomer. The Trp r values decrease slightly (H4-49W) or exhibit slight to moderate increases (H4-60W and H3-78W, respectively).

In summary, native gel electrophoresis and fluorescence anisotropy demonstrate that the FRET probes are solvent-accessible, have rotational freedom and do not perturb NCP assembly. Two essential features required for efficient FRET have been demonstrated: 1) The FL intensity of the Trp donors is independent of the state of NCP assembly; and 2) at least one of the fluorophores in each FRET pair exhibits extensive rotational freedom.

H2A-H2B dimer dissociation monitored by FRET

The observed FRET signal between the H3-H4 donors and the H2A-H2B acceptors exhibits a salt-dependence that is consistent with dimer dissociation. Representative emission scans at low and high salt concentrations are shown in Figure 3.7A. In the intact NCP, Trp FL is quenched and there is substantial Cys-AEDANS FL at 490 nm. Upon dissociation of the dimer,

the Cys-AEDANS FL decreases, with a resulting increase in Trp FL at 350 nm. For the five NCPs, the Cys-AEDANS FL intensities decrease by 1.4 to 2.7-fold while the Trp FL increases 1.7 to 3.8-fold. The range of FL sensitivities correlates with the calculated Cα donor-acceptor distances and whether the D-A and D-A' distances are similar.

The donor and acceptor FL intensities exhibit a sigmoidal dependence on the concentration of NaCl. Representative transitions are shown in Figure 3.7B and C. Similar data were obtained for the four other NCP FRET pairs listed in Table 3.2. The relatively steep, sigmoidal changes in FRET-related fluorescence are coincident with the first major transition observed by intrinsic Tyr FL (Figure 3.6), between 0.6 and 1.5 M NaCl. The second transition observed in Tyr FL, above 1.5 M NaCl, reflects tetramer dissociation from DNA; as expected, the FRET signals are not sensitive to this unfolding step. Representative linear pre- and post-transition baselines are indicated in Figure 3.7B and C. The longer folded, pre-transition baselines were generated from at least ten NaCl concentrations. The shorter unfolded, post-transition baselines were determined from at least five NaCl concentrations, and usually ten concentrations. Between 0 and 0.2 M NaCl, the 490 nm FL of NCPs with the H2B-109Cys-AEDANS acceptor exhibited a small non-linear change of $\leq$ 6% (Figure 3.7C). This curvature was not included in determination of the folded baselines; the Trp FL baselines didn't exhibit curvature. The H2B-acceptor FL may reflect local structural rearrangements or DNA breathing as observed in previous NCP FRET studies (*13;20*).

Using the folded and unfolded baselines, the FL data were converted to F_{app} for comparison of different FRET NCPs (Equation 5). Donor FL intensities were generally used for these comparisons because the linearity of this signal from 0 to 0.5 M NaCl provided a less ambiguous determination of the folded baselines. Figure 3.9 shows the F_{app} values as a function

of NaCl concentration for the five FRET NCPs. The three NCPs with the H2A-108Cys-AEDANS acceptor, but different Trp donors, are nearly coincident, with midpoints of ~0.95 M NaCl.

The unfolding transitions of the two NCPs with the H2B-109Cys-AEDANS acceptor are broader than those with the H2A acceptor, despite common Trp donors. The difference in the steepness of the unfolding transitions between NCPs containing FRET acceptors on H2A and H2B may result from the differences in the distances of the D-A and D-A' pairs in the two sets of FRET NCPs. Figure 3.10 shows a schematic depiction of the location of the D and A fluorophores relative to each other in the each set of FRET NCPs. The D-A and D-A' distances are similar for NCPs containing the H2A 108Cys-AEDANS acceptor. In contrast, $\geq$ 95% of the FRET signal for the H2B-109Cys-AEDANS NCPs arise from the D-A pair. For the H2A-acceptor NCPs, both Trp donors are still within FRET distance of a potential acceptor (A or A') in the hexameric species formed by dissociation of the first H2A-H2B dimer. Thus, the donors in the hexameric population should be nearly homogeneous with respect to FRET signal. In contrast, 50% of the Trp donors are no longer quenched in the H2B-acceptor hexameric population, while 50% are still near their maximal FRET distance. Two spectroscopically distinct donors in the hexameric species may appear as an unfolding intermediate, giving rise to a broader transition. This may also be complicated by the small change in the FL of the H2B-109Cys-AEDANS between 0.5 and 1.5 M NaCl.

Estimates of the free energy of dimer dissociation from the NCP.

The dissociation of the H2A-H2B dimers appears to be reversible. Samples were prepared at 1.0 to 1.2 M NaCl, where the dimer is 50% to 90% dissociated. Under these

concentrations, Tyr FL indicates that the tetramer-DNA complex is still highly populated (Figure 3.6). These partially dissociated samples were diluted to 0.2 M NaCl by manual mixing and allowed to re-equilibrate for 10 min. The FL signals observed after refolding were within 5% of those of undissociated samples. This reversibility suggested that a thermodynamic free energy could be determined for the dissociation of the H2A-H2B dimers from the NCP.

Thermodynamic analyses of transitions monitored by intrinsic Tyr FL and DNA-protein FRET (*13*) are complicated by broad transitions that reflect contributions from multiple unfolding steps, including DNA breathing, dimer and tetramer dissociation. Thus, the relative concentrations of the various species can't be accurately determined. An advantage of this protein-protein FRET system is that a single sigmoidal transition is observed which reports specifically on H2A-H2B dissociation. This allows determination of the relative populations of folded and partially unfolded NCPs to estimate equilibrium constants at various NaCl concentrations (Equation 6). Figure 3.11 shows a representative plot of the salt-dependence of the ΔG values (Equation 7) in the transition region (0.6 to 1.3 M NaCl). The salt-dependence of the ΔG values wasn't linearly dependent on the NaCl concentration. The free energies for electrostatic interactions often vary linearly with the square root of the ionic strength (for review, (*64*)); however, the salt-dependence of the ΔG values was also non-linear in this analysis. The data for all NCPs were well described by two lines that intersect near the midpoint of the dissociation transition, 0.9 to 1 M NaCl. The two different linear phases presumably reflect the dissociation of the first and then second H2A-H2B dimer. The fitted lines yield slopes which are analogous to the *m* values from protein folding studies, as well as estimates of the free energy of dissociation in the absence of NaCl, $\Delta G°(H_2O)$ values. For the H2A-108Cys-AEDANS NCPs, which exhibit nearly coincident F_{app} curves, the average slopes for the first and second transition

regions were -15.8 ± 1.5 and -9.0 ± 0.8 kcal mol^{-1} M^{-1}. The corresponding $\Delta G°(H_2O)$ values were 32.5 ± 1.2 and 26.5 ± 0.8 kcal mol^{-1}. The slopes and $\Delta G°(H_2O)$ values for NCPs with the H2B-109Cys-AEDANS acceptor were ~60% and ~10% lower, respectively.

H2A.Z-incorporated FRET NCP dynamics

The multiple protein-protein FRET system is well-suited to characterize alterations of NCP equilibrium properties which result from incorporation of histone variants. We chose to apply this system to H2A.Z-containing NCPs because contradictions in the literature highlight the ambiguities of H2A.Z's biological function and effects on the thermodynamic properties of the nucleosome. Dimers of H2A.Z and H2B-109Cys-AEDANS were reconstituted into 601-NCPs with either H3-78W or H4-60W as FL donors. As analyzed by native PAGE, the quality and homogeneity of the reconstitutions were equivalent to those containing the major H2A (Figure 3.7B). A comparison of the salt-induced dissociation transitions of the H2A.Z and H2A NCPs are shown in Figure 3.12 A and C. Only small differences are observed for the salt-induced dimer dissociations; the largest apparent difference is a slightly steeper slope in the F_{app} curves of the H2A.Z NCPs.

To further characterize relative NCP stability, the concentration dependence of the FRET signals were monitored at 1.0 M NaCl, approximately the transition midpoint. The ratios of donor to acceptor FL as a function of NCP concentration dependence are shown in Panels B and D of Figure 3.12. The FL ratio was used to compensate for the protein concentration dependence of the FL intensities. The ratios were normalized using Equation 9 where Y_{NCP} is the ratio under fully assembled conditions (0.2 M) and $Y_{T,D}$ is the ratio observed at 2.0 M NaCl, where the dimer is fully dissociated from the NCP. Thus, a normalized ratio value of 0 indicates fully

intact NCPs. As expected, the NCPs have more folded-like FRET ratios at higher NCP concentrations. Like the salt-induced perturbations, the protein concentration dependence of the FL ratios revealed little difference in the stability of H2A and H2A.Z NCPs.

The ΔG values for dimer dissociation from the H2A and H2A.Z NCPs were also compared; representative data are shown in Figure 3.13. As observed for the H2A-acceptor NCPs, the ΔG values were best described by two linear terms, which intersect near the transition midpoint. At all NaCl concentrations in the transition region, the ΔG values for H2A and H2A.Z NCPs are within 10% of each other. However, the H2A.Z NCPs, with either the H3-78W and H4-60W donor, exhibit consistently lower ΔG values for the second linear phase. This difference indicates that the dissociation of the second dimer is more favorable in the H2A.Z NCP, suggesting higher cooperativity in the dimer dissociation. The H2A.Z NCPs exhibit steeper slopes for the first and second linear phases, 1.1 to 1.3-fold and 1.2 to 1.6-fold greater, respectively, which is also consistent with greater cooperativity in the H2A.Z NCP disassembly. In summary, although the differences in stability are small, there is an increase in the cooperativity of dimer dissociation from the H2A.Z NCPs.

DISCUSSION

Design of a histone-histone FRET system

Previous FRET systems engineered to study NCP dynamics used coumarin (CPM)/fluorescein (FM) derivatives (*13*) or Cy3/Cy5 (*20*) as the donor/acceptor pairs, usually with one of the FRET probes on the DNA termini. Multiple protein-protein FRET pairs were engineered in this study to specifically monitor the interactions of the H2A-H2B dimers with the $(H3\text{-}H4)_2$ tetramer. Tryptophan and Cys-AEDANS were chosen as the donor and acceptor,

repectively, because of their relatively small size and decreased hydophobicity in comparison to previously employed fluorophores.

A well-designed FRET system should employ fluorophores which are solvent accessible with substantial rotational freedom and do not significantly alter histone or NCP stability. FL emission spectra of individual D or A labeled histone oligomers in folded and unfolded conditions show a small red-shift in the emission maximum upon unfolding, demonstrating high solvent accessibility. In addition, the anisotropy r values in the dimers, tetramers and NCPs are consistent with substantial rotational freedom (Table 3.2). Equilibrium studies of H2A•H2B dimers with Cys-AEDANS and H3•H4 tetramers with tryptophans demonstrated that the stabilities, and therefore the structures, of the histone oligomers were not significantly altered (Table 3.2). The denaturation curves of all the histone oligomers were well described as highly reversible, two-state equilibrium processes (Figures 3.2 to 3.5). The NCPs reconstituted with the FRET probes were not significantly altered as observed by the similar gel mobility of the WT and variant NCPs. Furthermore, the salt-induced dissociation transitions monitored by FRET agreed well with the transition observed by Tyr FL in the WT NCP. Overall, the incorporation of the FRET probes into the various locations on the histones minimally perturbed their thermodynamic properties as individual olimomers as well as when reconstituted into a NCP.

In order to calculate a free energy, ΔG, for the dissociation of the H2A-H2B dimer from the NCP, the spectral signal must monitor only one transition (not a mixture of dimer and tetramer dissociation reactions) and the reaction must be reversible. The salt-induced unfolding of the NCP, monitored by both donor and acceptor FL, was sigmoidal, indicating a single, cooperative dissociation of the H2A-H2B dimers. The FL of the Trp donors on H3 and H4 were insensitive to all NCP structural transitions as was the FL of the H2A-108Cys-AEDANS.

Therefore, the changes in FL at 350 nm (and 490 nm for H2A-acceptor NCPs) can be attributed to FRET differences arising solely from dimer dissociation. These data can then be used to determine relative populations of NCP and free dimer plus tetramer-DNA complex. To accurately determine relative populations, well-defined pre- and post-transition baselines are required to determine the signals expected for fully folded and unfolded species. The greater stability of NCPs reconstituted with the 601 positioning sequence and the insensitivity of FRET to multiple transitions made this possible, in contrast to previous FRET systems (*13;20*). The criteria of reversibility was demonstrated by the >90% recovery of the expected FRET signals in samples refolded by dilution from salt concentrations at which the majority of the H2A-H2B dimers have dissociated.

Comparison to previously reported NCP FRET studies

Several studies have been published previously monitoring the dynamics of the nucleosome core particle with FRET systems (*13;20-22*). These systems have been used to monitor DNA-DNA, protein-DNA and protein-protein interactions with the majority of their focus on DNA dynamics. The system described in this report utilizes a much stronger NCP positioning element than previous studies and was designed to specifically monitor the changes in distance between the H2A-H2B dimer and the $(H3\text{-}H4)_2$ tetramer. The D-A pairs of previous FRET systems are somewhat larger in size, but more importantly are more hydrophobic than the Trp/Cys-AEDANS pair in this report. Large hydrophobic dyes have a tendency to sequester themselves from solvent (*65*), which may perturb macromolecular structures. This is suggested by the larger anisotropies of ~0.19 for labeled histones (compared to 0.05 for the free fluorophore) reported in previous NCP FRET studies (*13*). Additionally, the Förster distances of

these pairs are 50 to 60 Å, compared to 20 Å for the Trp/Cys-AEDANS pair. The molecular dimensions of the NCP are approximately 100 Å in diameter by 50 Å in width. It has been shown that under physiological conditions, H2A-H2B dimers are able to undergo rapid exchange between partially assembled NCPs and form unstable DNA-H2A-H2B complexes (*66*). It is possible that above 0.5 M NaCl, H2A-H2B dimers weakly associate with DNA in the DNA-tetramer complexes in a non-NCP conformation that would be detectable by FRET probes with R_0 values of ~50Å. With a shorter R_0, detection of non-specific interactions should be minimized.

An important difference between the current and previous FRET systems (*13;22*) is the use of the 601 artificial positioning sequence (*53*). This sequence yields a more stable and homogeneously positioned nucleosome (*18*) than the historically used 5S sequence and is becoming a standard positioning sequence in chromatin studies (for example, (*20;21;67;68*)). There are reports on the dynamics of DNA breathing in the context of the 601 sequence (*20*), but no biophysical studies describing its effects on histone dissociation. This is the first report characterizing the salt-induced dissociation of the dimer from 601-NCPs. Interestingly, it is the dimer that is stabilized by the 601 sequence, with little to no effect on the dissociation of the tetramer, as shown by Tyr FL, despite the central importance of the tetramer in NCP positioning (*69*).

Free energy of H2A-H2B dissociation from the NCP

A unique aspect of this FRET system is the ability to calculate a ΔG for dimer dissociation as discussed above. Thermodynamic studies have been limited to determinations of the relative binding free energy, $\Delta\Delta G$, between various DNA sequences, using competitive gel

shift assays (for review, (*55*)). Dilution-driven dissociation experiments to determine an absolute equilibrium constant, K_{EQ}, for complete nucleosome disassembly have been unsuccessful (*70;71*). The dissociation reactions were not reproducible unless destabilizing additives were used to prevent nonspecific loss of dissociated histones; even with additives, dilution-driven dissociation was not fully reversible, preventing the determination of histone-DNA binding free energies. The dilution driven experiments also demonstrated that a simple dissociation model (NCP ⇆ free DNA + histone octamer) is inappropriate, and that multiple non-nucleosomal histone-DNA species must be considered (*71*). Therefore, a first step toward a thermodynamic description of NCP assembly should begin with analysis of the ΔG for dimer dissociation, for which the current system is well-suited because of its reversibility and sensitivity to a single transition. Salt-induced dissociation, shown to be reversible, was employed instead of dilution-driven dissociation. However, analyses of perturbation experiments do require extrapolation to physiological ionic strengths.

Interestingly, two transitions were evident in the salt-dependence of the ΔG values for all FRET NCP combinations (Figure 3.11). The biphasic ΔG response is interpreted as dissociation of the first and then second H2A-H2B dimer, which have different affinities for the NCP. The population of hexameric nucleosome intermediates have been reported in the presence of histone chaperones (*72*) and depletion of chicken erythrocyte core particles under certain stringent conditions such as 3 or 4 M urea and 250 mM sodium chloride (*73*). The slope and ΔG values of the second transition are less than for the first, indicating positive cooperativity—the second dimer dissociates more easily than the first. The close proximity to each other of the H2A L1 loops has been proposed to play a role in the cooperative binding of the H2A/H2B dimers (*6;7*).

Free energy and cooperativity of H2A.Z-H2B dimer dissociation from the NCP

The structural and functional effects on the NCP upon incorporation of the variant H2A.Z has been the subject of controversy in the literature (for review, (*74*)). Based on comparison of H2A and H2A.Z NCP crystal structures, it was hypothesized that some interaction differences may stabilize H2A.Z NCPs, while others may be destabilizing (*39*). An initial report by Ausio and colleagues described a significant destabilization based on analytical ultracentrifugation studies of the ionic strength dependence of H2A and H2A.Z NCPs. A mixture of chicken erythrocytes histones and recombinant human H2A.1 or H2A.Z were reconstituted to form mononucleosome with random 146 bp DNAs and oligonucleosome arrays with the tandem 12mer repeat of the 208 bp sea urchin 5S sequence. Subsequently, a FRET based approach using recombinant *Xenopus laevis* histones and recombinant mouse H2A.Z was interpreted as a slight increase in stability for the H2A.Z NCPs reconstituted with the sea urchin 5S sequence (*13*). Three FRET pairs monitoring distances between the DNA ends, DNA-H2A and H2A-H4 showed an apparent shift of the H2A.Z dissociation to higher NaCl concentrations. However, the ΔC_M values of 36 to 66 mM are close to the experimental uncertainty of the experiments and may reflect effects of baseline choices in normalization. The slopes of the H2A and H2A.Z NCP transitions were very similar or slightly smaller for the DNA-DNA and DNA-H2A.Z FRET pairs, respectively. The latter change was attributed to either a decrease in cooperativity or a difference in the salt sensitivities of the H2A and H2A.Z molecular interactions with the NCP. A very recent study by Ausio and colleagues (*43*) contradicts their previous findings, attributing the differences to the histone source and purification method. The later study used histones isolated from chicken erythrocytes or bacterially expressed recombinant H2A.Z. Sedimentation analytical ultracentrifugation and sucrose gradient separation of nucleosomal components were

used to analyze the salt-dependence of NCP dissociation, and showed that H2A.Z NCPs exhibited a subtle increase in stabilty. It was also found that this stability difference was abolished by histone acetylation.

The results presented here confirm that the effects of H2A.Z on NCP stability are small and subtle. The transitions have similar midpoints (Figure 3.12), although the slopes are slightly steeper for H2A.Z, suggesting an enhanced cooperativity in dissociation relative the H2A. The salt dependence of the ΔG values is also steeper for H2A.Z, particularly for the second dimer dissociation—consistent with greater cooperativity than H2A. Accordingly, the ΔG values for the first dissociation are similar for H2A-H2B and H2A.Z-H2B dimers, while ΔG values for the second dissociation are lower for H2A.Z.

The effects observed for incorporation of H2A.Z highlight the importance of changes in the interactions of the L1 loops in H2A and H2A'. The H2A.Z NCP structure shows more hydrogen bonds and contacts between the L1-L1' loops than observed in the H2A-containing NCP structure (*39*). These differences were predicted to yield more cooperativity in the dissociation of the two dimers, which is the major effect observed in the current study. In contrast, the fewer hydrogen bonds in the H2A.Z-H3 interface were predicted to be destabilizing, but appear not to be a major determinant of stability for H2A and H2A.Z containing NCPs in the context of this study. It is apparent from the conflicting reports in the literature that the context of the NCPs must be considered as to the source of the histones, their modification state and the DNA sequence used. The current and previous FRET studies (*13;20*) used recombinant histones and homogeneous DNA fragments, although different from each other, while the Ausio studies used heterogeneous DNA isolated during preparation of histones from native chromatin. The extent to which the 601 DNA sequence mitigates the stabilization provided by the incorporation

of H2A.Z in other studies is unknown, but can be easily tested with this FRET system. Although positioning effects result primarily for tetramer-DNA interactions (*69*), the current results demonstrate that DNA sequence can alter the affinity of the H2A-H2B dimer for the NCP.

Implications and future applications of the FRET system

The nucleosome core particle is critical in the compaction of chromatin and in the regulation of DNA-templated chemistries. Significant chromatin remodeling must take place in order for the various replication, transcription, and repair machinery to access DNA bound by the nucleosomes. In addition to the inherent mobility of nucleosome core particles, various mechanisms such as ATP-dependent chromatin remodeling complexes, histone variants, and post-translation modifications are employed by the cell in order to provide access to DNA for these processes (*75*). The protein-protein FRET system described in this body of work will be used to dissect the kinetic properties of the NCP, in particular the rates of dimer association and dissociation. Additional work is also being done to further explore the effects of incorporating histone variants on nucleosome dynamics. Overall, FRET is a very powerful and versatile tool that will help provide a better understanding of how histone sequence and stability impacts nucleosome dynamics.

Acknowledgements:

The 601 nucleosome positioning sequence was kindly provided by the lab of Dr. Jonathan Widom (Northwestern University). The original expression plasmids for the major *Xenopus laevis* histones and the H2A.Z variant from *Mus muculus* were provided by Dr. Karolyn Luger (Colorado State University).

Table 3.1. Distances between the Cα atoms of the residues mutated for incorporation of the FRET donors and acceptors.[a]

FRET Pair	A/B to C; E/F to G[b]	A/B to G; E/F to C[b]	Global Ave, % FRET[c]
H3-78W to H2A-108Cys-AEDANS	32.9 33.0	35.5 35.5	**34.2,** **3.8%**
H4-49W to H2A-108Cys-AEDANS	30.2 30.5	28.4 28.4	**29.4,** **8.8%**
H4-60W to H2A-108Cys-AEDANS	23.8 24.7	26.6 26.9	**25.5,** **18.9%**
FRET Pair	**A/B to D; E/F to H**	**A/B to D; E/F to H**	**% FRET Efficiency[c]**
H3-78W to H2B-109Cys-AEDANS	19.6 20.0	48.9 49.6	**51.9 &** **0.5%**
H4-60W to H2B-109Cys-AEDANS	26.9 26.8	45.4 45.7	**14.2 &** **0.7%**

[a]Distances were estimated from the coordinates of the NCP X-ray crystal structure, 1AOI.pdb (*6*).

[b]The column headings refer to the pdb designations for the histone monomers in 1AOI.pdb (H3 as A and E; H4 as B and F; H2A as C and G; H2B as D and H).

[c]The Förster distance of the Trp:Cys-AEDANS FRET pair is ~20 Å (*76*). Expected FRET efficiency was calculated using Equation 3.

Table 3.2. Fitted parameters describing the equilibrium stability of the histones modified with FRET donors and acceptors.[a]

Histone Oligomer	$\Delta G^o(H_2O)$ (kcal mol^{-1})	*m* value (kcal mol^{-1}M^{-1})	Cm	Anisotropy (r)[b]		
				F	U	NCP
H2A-H2B dimer[c]						
WT-H2A•WT-H2B	11.8	2.9	1.7	NA	NA	NA
H2A-108Cys-AEDANS•WT-H2B	12.1 (0.3)	3.1 (0.1)	1.7	0.06	0.02	0.04
WT-H2A•H2B-109Cys-AEDANS	11.9 (0.3)	2.9 (0.2)	1.7	0.06	0.02	0.05
H3-H4 dimer[d]						
WT-H3•WT-H4	11.8	2.6	1.7	NA	NA	NA
H3-78W•WT-H4	11.7 (0.5)	2.9 (0.3)	1.5	0.10	0.06	0.11
WT-H3•H4-49W	12.8 (0.5)	3.1 (0.2)	1.7	0.08	0.04	0.07
WT-H3•H4-60W	11.7 (0.2)	2.5 (0.1)	1.8	0.08	0.04	0.14

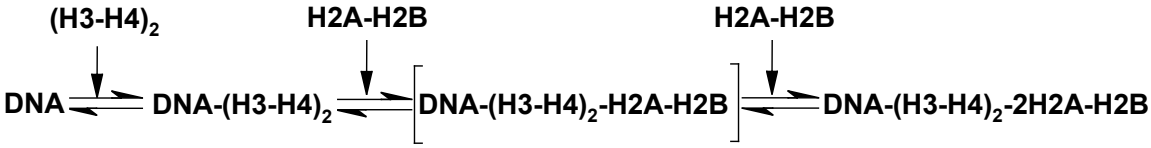
[a]The $\Delta G°(H_2O)$ values are the free energy of unfolding in the absence of denaturant at a standard state of 1 M oligomer; the *m* values reflect the slope of the unfolding transition (Eq. 7). C_M is the midpoint of the transition, the denaturant concentration at which the concentration of unfolded monomers equals $0.5•[Monomer]_{Total}$. The values in parentheses are the error at one standard deviation determined from rigorous error analysis of the global fits. Buffer conditions: 20 mM KPi pH7.2, 0.2 M KCl, 0.1 mM EDTA at 25ºC.

[b]The anisotropy r values are given for the folded histone oligomers, F; the unfolded monomers, U; and the NCP at 0 M NaCl. NA = not applicable as the relevant fluorophores are not present.

[c]The urea-induced unfolding equilibrium parameters for the WT H2A-H2B dimer are from (*48*). The C_m values are for transitions at 10 μM monomer.

[d]The values represent the stability of the H3-H4 dimer to GdmCl denaturation in the presence of 1 M TMAO. The C_m values are for transitions at 4 μM monomer. The equilibrium parameters for WT H3-H4 dimers are from (*46*).

A.



B.

Figure 3.1. **A.** Scheme of the nucleosome unfolding/folding pathway. Transient formation of hexameric intermediate, in brackets, is not directly detected by FRET methods. **B.** NCP structure with the mutated sites (Table 3.1) highlighted in various colors. Crystal structure (1kx5.pdb) was rendered in Pymol (*77*). Histones are colored: H2A, red; H2B, orange; H3, cyan; and H4, blue. Mutations: H2A Leu-108 (purple); H2B Ser-109 (green); H3 Phe-78 (pink); H4 Leu-49 (orange), and H4 Val-60 (magenta).

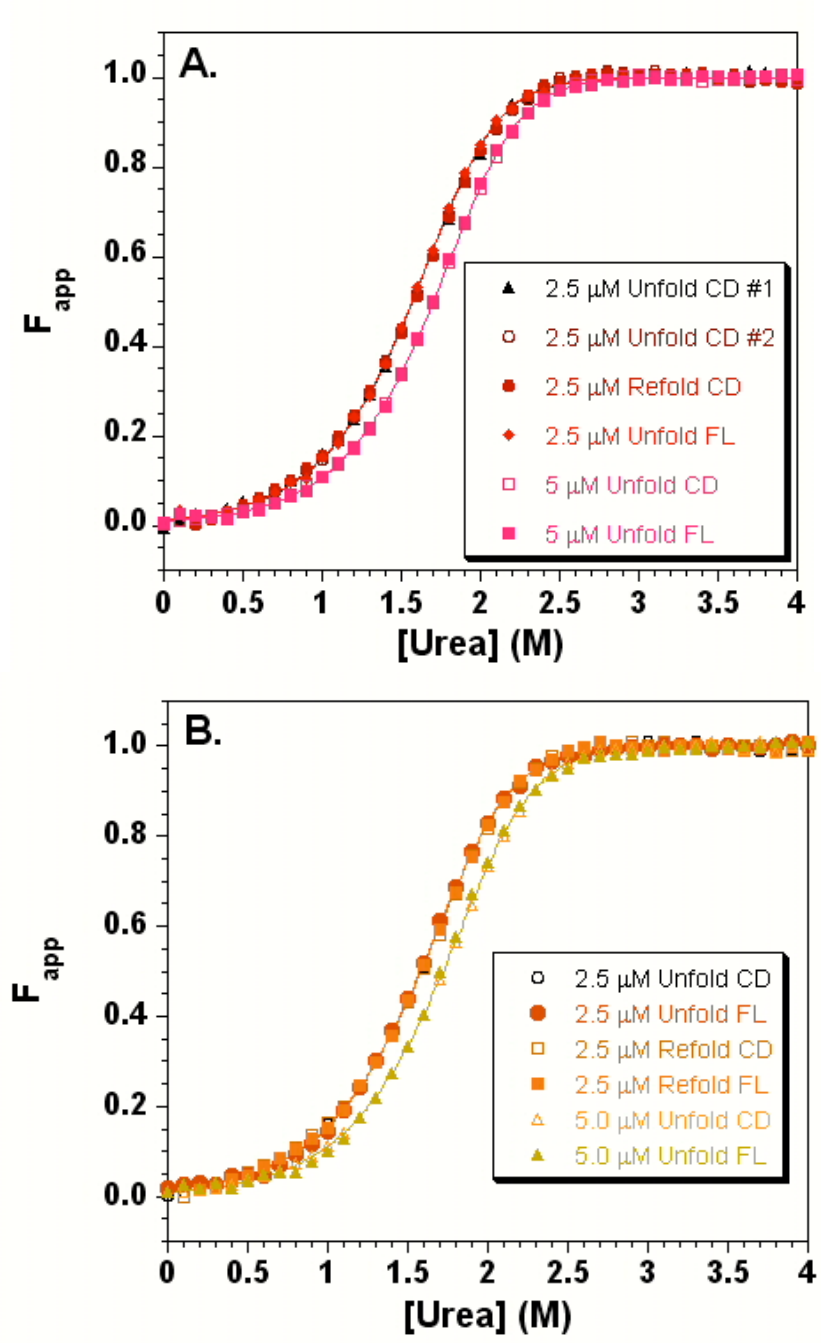


Figure 3.2. F_{app} plots for the urea equilibrium denaturation transitions of the Cys-AEDANS modified H2A-H2B dimers. **A.** H2A-108Cys-AEDANS•WT H2B **B.** WT H2A•H2B-109Cys-AEDANS 20 mM KPi pH 7.2, 200 mM KCl, 0.1 mM EDTA, 25°C.

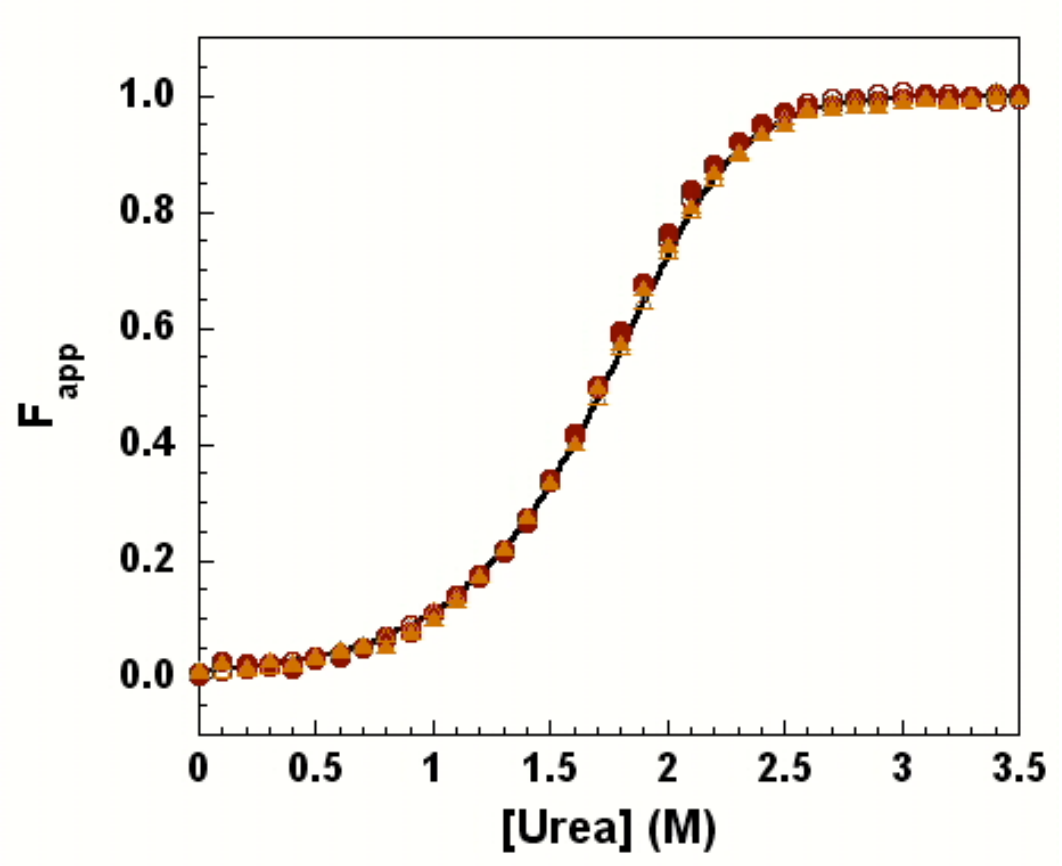


Figure 3.3. Apparent fraction of unfolded monomers (F_{app}) as a function of [Urea] for the unfolding transitions of WT and modified H2A•H2B dimers at 10 □M monomer. The solid black line represents the WT H2A-H2B equilibrium data. ○; H2A-108Cys-AEDANS•H2B CD, ●; H2A-108Cys-AEDANS•H2B FL, △; H2A•H2B-109Cys-AEDANS CD, ▲; H2A•H2B-109Cys-AEDANS FL. Buffer conditions are described in the legend of Figure 3.2.

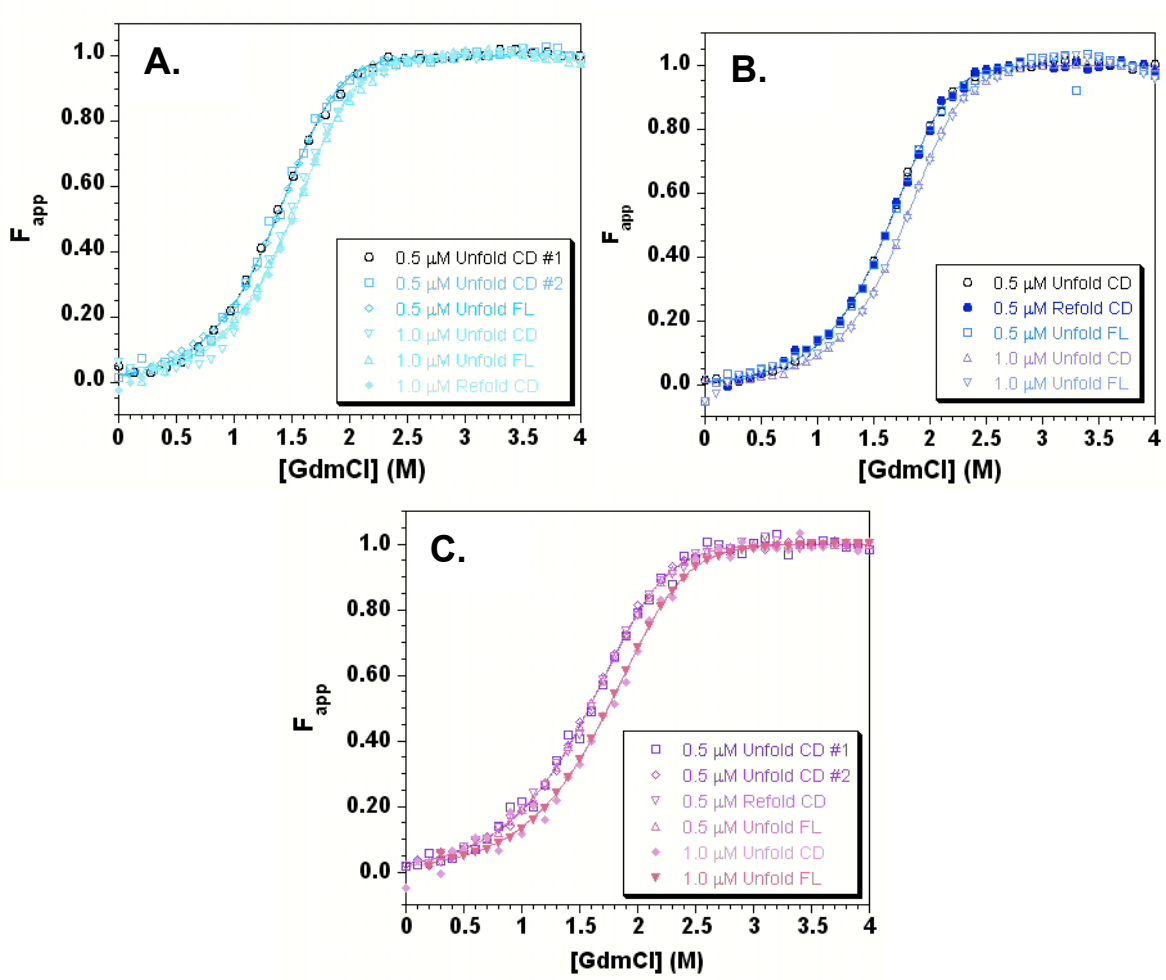


Figure 3.4. F_{app} plots for the GdmCl equilibrium denaturation transitions of the Trp engineered H3-H4 tetramers. **A.** H3-78W•H4 F_{apps} in 1M TMAO buffer as a function of GdmCl. **B.** H3•H4-49W F_{apps} in 1M TMAO buffer as a function of GdmCl. **C.** H3•H4-60W F_{apps} in 1M TMAO buffer as a function of GdmCl. Buffer Conditions: 20 mM KPi pH 7.2, 200 mM KCl, 0.1 mM EDTA, 25°C.

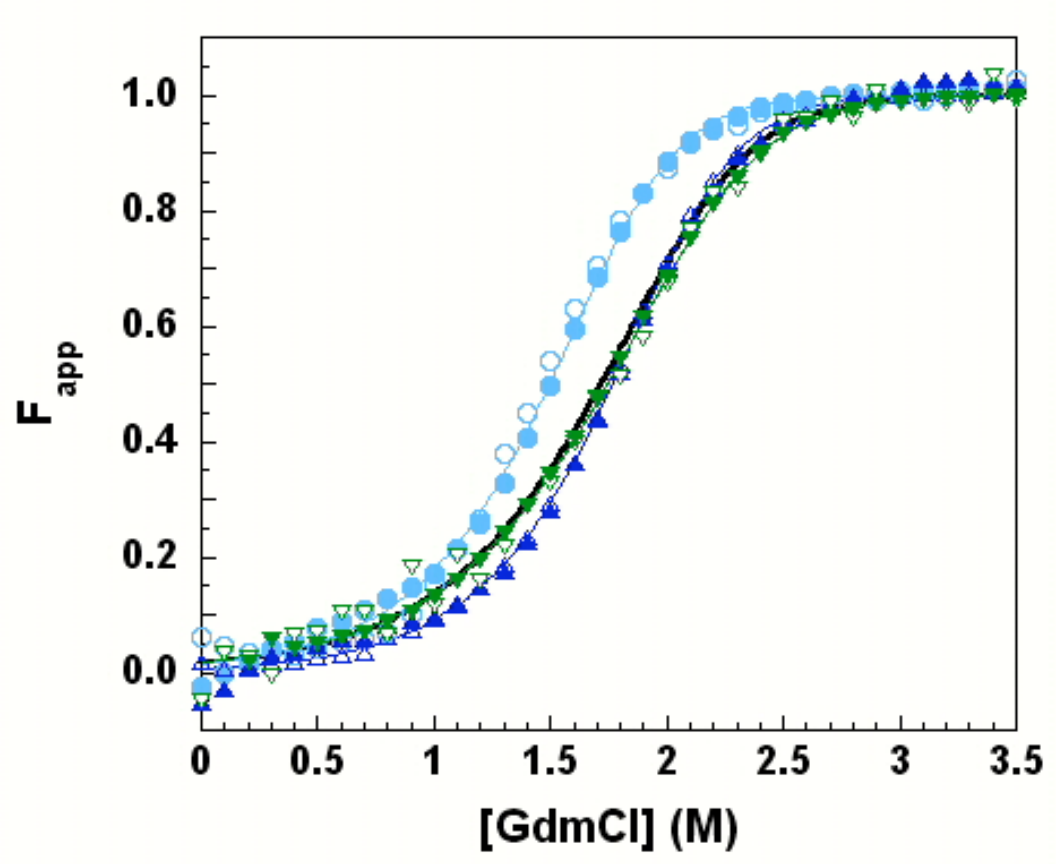


Figure 3.5. Apparent fraction of unfolded monomers (F_{app}) as a function of [GdmCl] for the unfolding transitions of the WT and Trp-containing H3-H4 oligomers at 4 □M monomer. The solid black line represents the two-state equilibrium unfolding of the WT H3•H4 dimer. ○ H3-78W•H4 CD; ● H3-78W•H4 FL; △ H3•H4-49W CD; ▲ H3•H4-49W FL; ▽ H3•H4-60W CD; ▼ H3•H4-60W FL. Conditions: 1 M TMAO, 20 mM KPi pH 7.2, 200 mM KCl, 0.1 mM EDTA, 25°C.

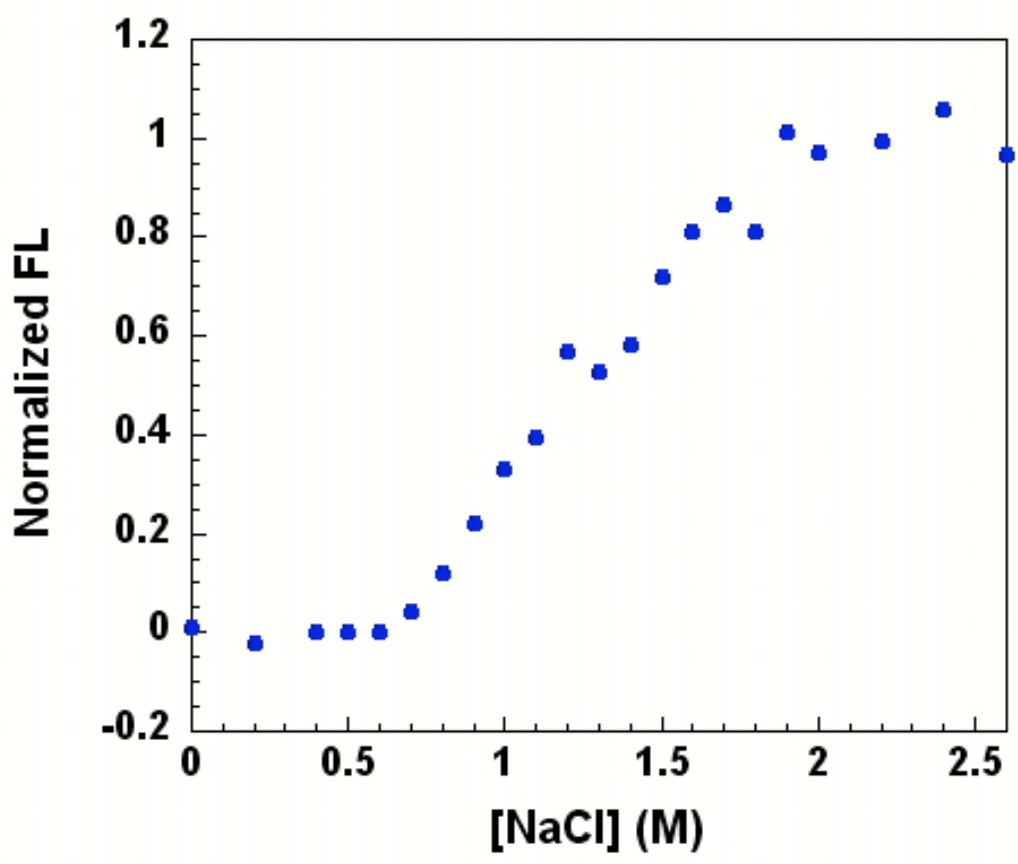


Figure 3.6. Normalized Tyr FL monitoring the salt-induced dissociation of NCPs reconstituted with the 601 DNA sequence and recombinant WT histones. WT 601-NCP as monitored by Tyr fluorescence. Conditions: 250 nM NCP, 20 mM Tris-Cl pH 7.6, 0.1 mM EDTA, 0.1 mM □-ME, 25°C.

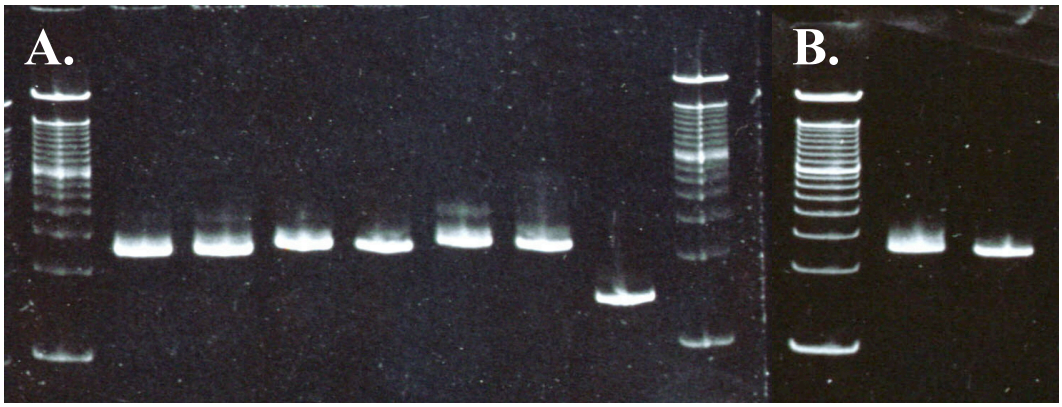


Figure 3.7. 5% acrylamide native gels stained with EtBr of NCPs reconstituted with the 149 bp 601 DNA fragment. **A.** Major H2A NCPs. From left to right: 100 bp ladder; H4-60W/H2A-108Cys-AEDANS NCP; H4-60W/H2B-109Cys-AEDANS NCP; H4-49W/H2A-108Cys-AEDANS NCP; H4-49W/H2B-109Cys-AEDANS NCP; H3-78W/H2A-108Cys-AEDANS NCP; H3-78W/H2B-109-CysAEDANS NCP; 149 bp 601 DNA. **B.** H2A.Z NCPs. From left to right: 100 bp ladder; H3-78W/H2B-109Cys-AEDANS NCP; H4-60W/H2B-109Cys-AEDANS NCP

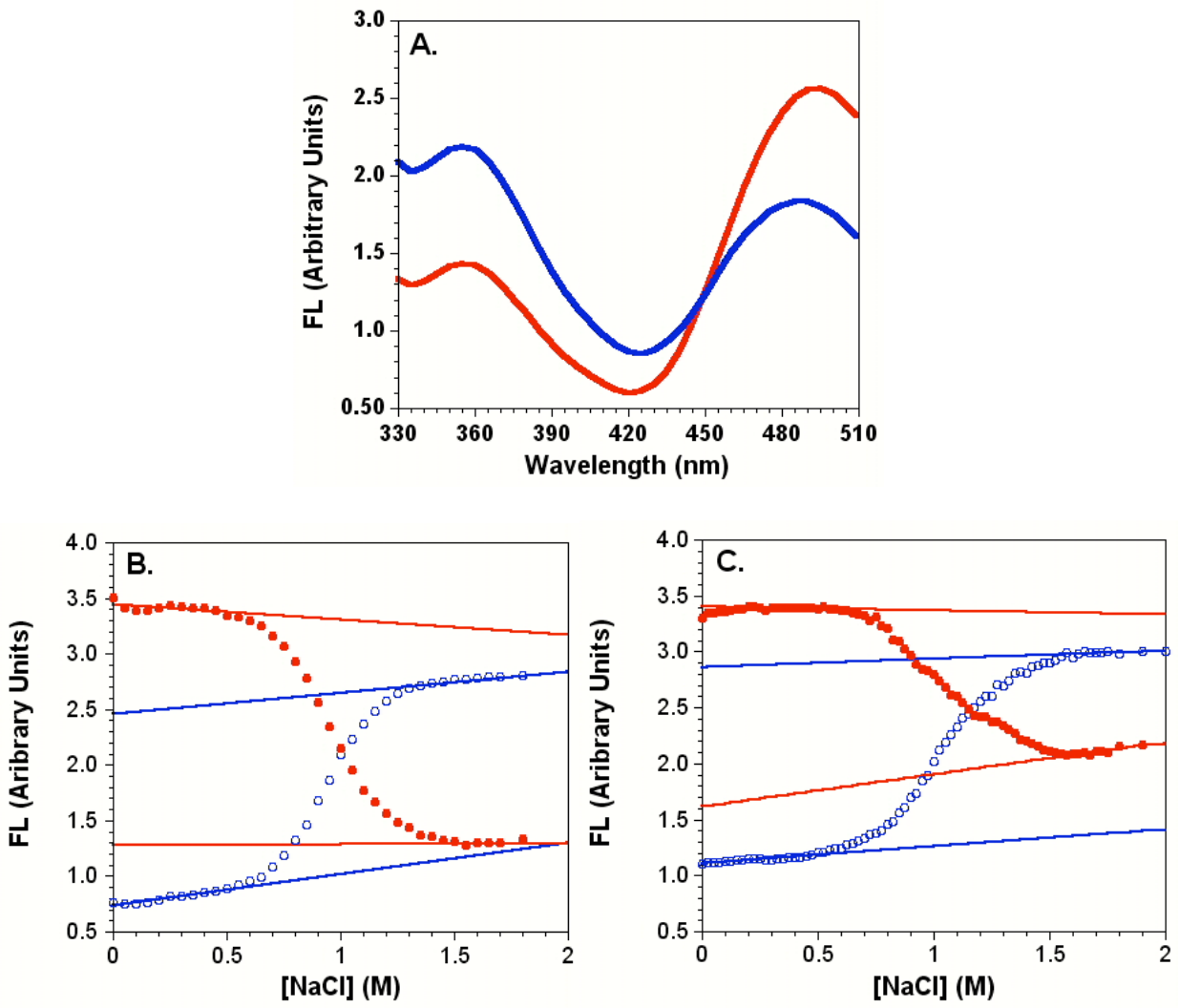


Figure 3.8. Representative FL data for the salt-induced dissociation of the FRET NCPs. **A.** FL emission spectra of the H4-49W/H2A-108Cys-AEDANS FRET NCP at 0 M (—) and 1.5 M NaCl concentrations (—). **B and C.** Salt dependence of the FL intensities for the H4-60W/H2A-108Cys-AEDANS and the H3-78W/H2B-109Cys-AEDANS NCPs, respectively. ○, Trp donor FL at 350 nm; ●, Cys-AEDANS FL at 490 nm. Conditions: 250 nM NCP; 20 mM Tris-Cl pH 7.6, 0.1 mM EDTA, 0.1 mM □-ME, 25°C.

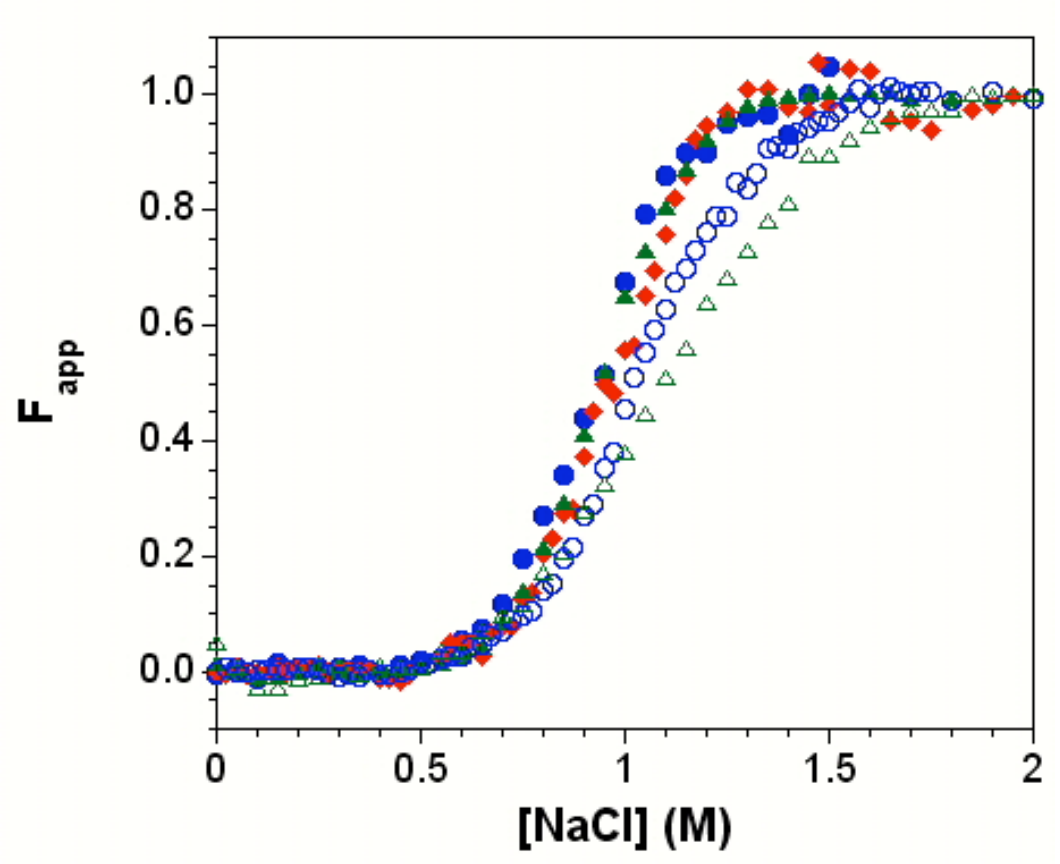


Figure 3.9. Normalized donor FL (F_{app}) data for the salt-induced dissociation of multiple FRET NCPs. Solid symbols represent FRET data for NCPs with the H2A-108Cys-AEDANS acceptor and the H3-78W (●), H4-49W (◆), and H4-60W (▲) donors. Open symbols represent FRET data for NCPs with the H2B-109Cys-AEDANS acceptor and the H3-W78 (○) and H4-W60 (△) donors. Conditions are given in the legend of Figure 3.8.

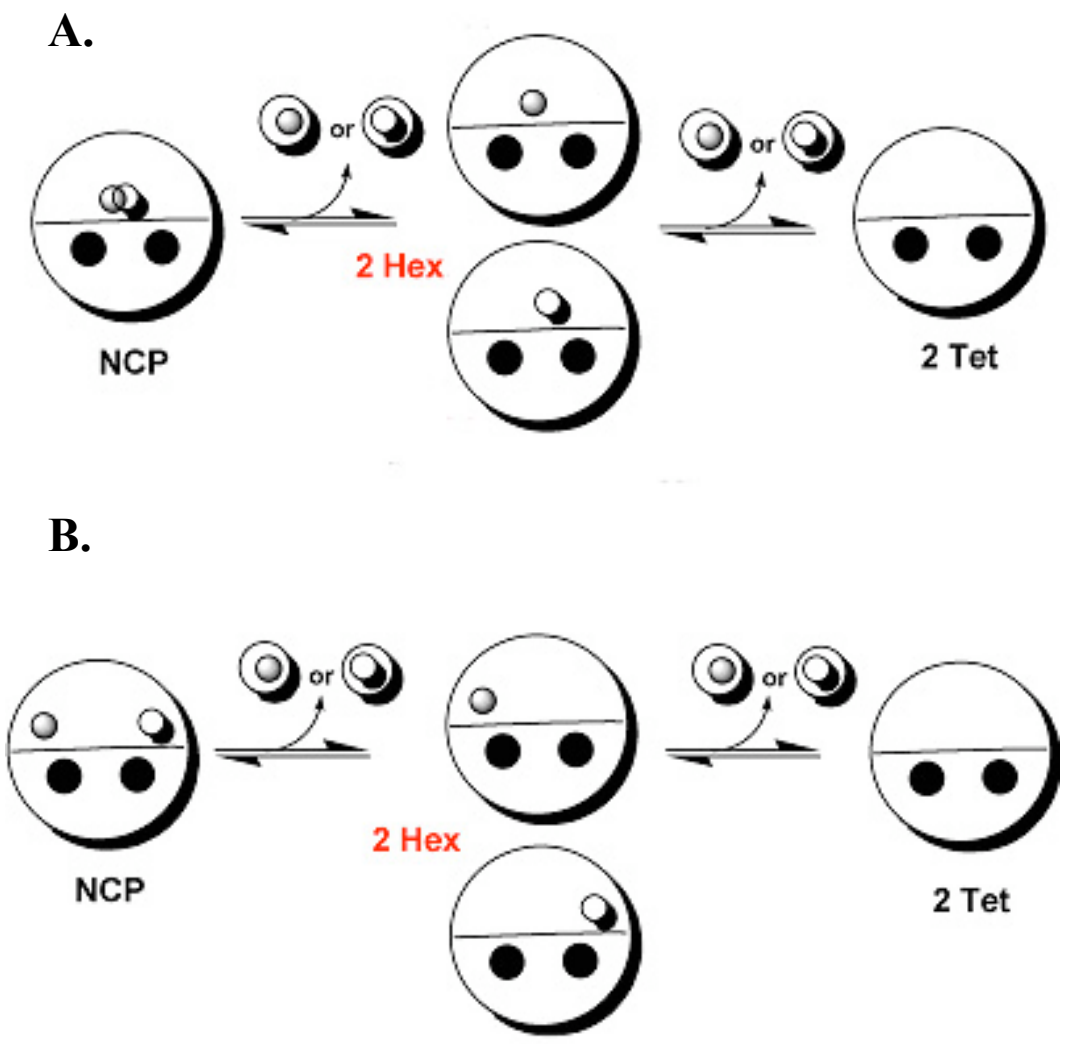


Figure 3.10. Schematic representation of NCP dissociation considering relative D-A and D-A' distances. Cys-AEDANS acceptors on the H2A-H2B dimers are represented by open circles and Trp donors on the $(H3\text{-}H4)_2$ tetramer, by closed circles. **A.** H2A-108Cys-AEDANS proximity to H3 and H4 Trp donors. In the hexameric NCP intermediates, the two donors FL are nearly identical as both remain within quenching distance after dissociation of the first dimer. **B.** H2B-109Cys-AEDANS proximity to H3 and H4 Trp donors. The two donors have different FL as one remains within quenching distance of the acceptor, while the other is not longer near an H2A-H2B dimer.

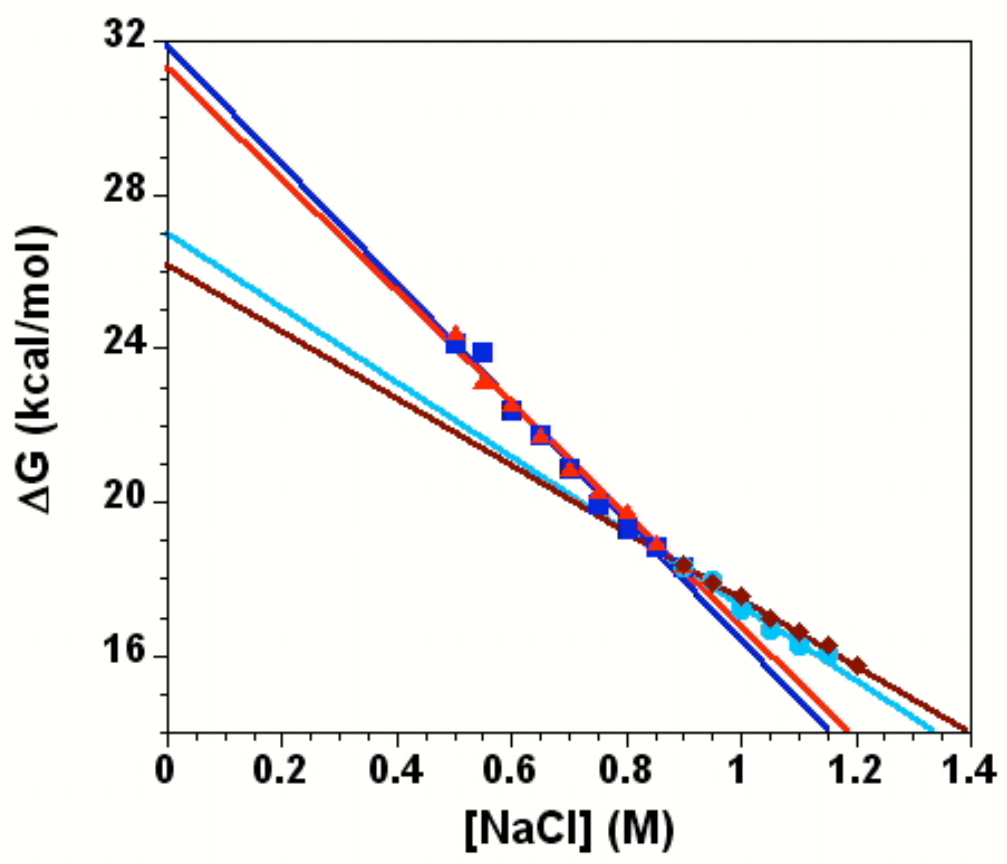


Figure 3.11. The salt dependence of the free energy of dimer dissociation from the H3-78W/H2A-108Cys-AEDANS 601-NCP. Values were calculated from FL intensities at 350 nm (■, ●) and 490 nm (▲, ◆) and were fit to two lines which intersect near the transition midpoint (Figure 3.9). Conditions are given in legend of Figure 3.8.

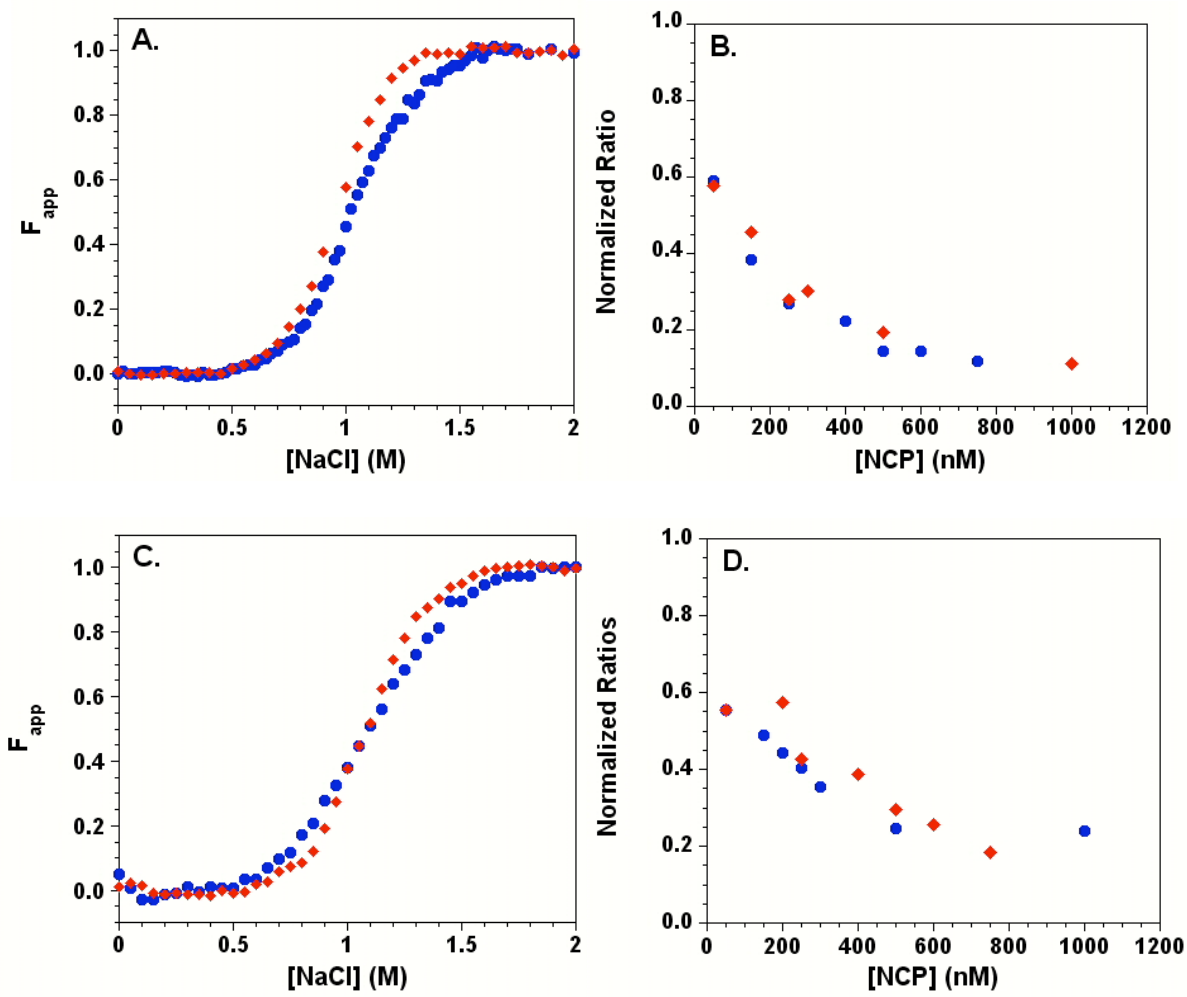


Figure 3.12. Equilibrium dissociation transitions for H2B-109Cys-AEDANS FRET NCPs with H2A (blue circles) and H2A.Z (red diamonds). Panels A and C are F_{app} curves describing the salt-dependence of the donor FL of H3-78W and H4-60W, respectively. Panels B and C are the normalized ratios of the FL at 350 and 490 nm as a function of NCP concentration at 1 M NaCl. Buffer conditions are given in the legend of Figure 3.8.

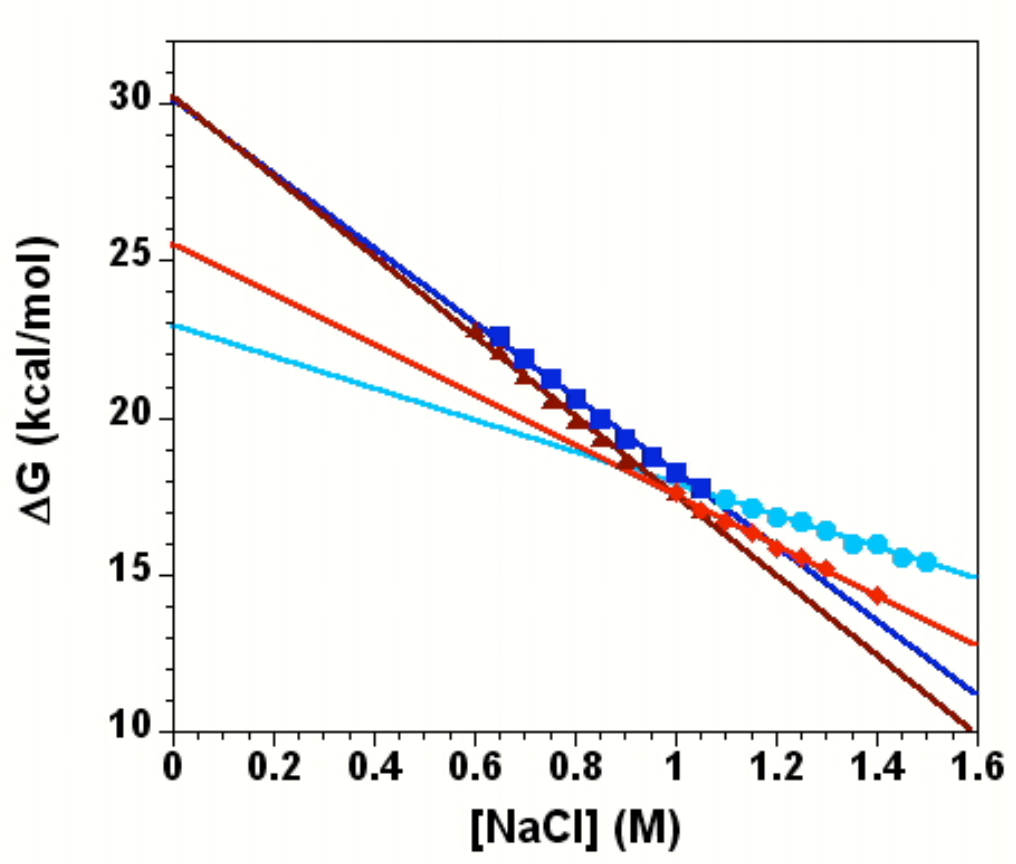


Figure 3.13. The salt dependence of the free energy of dimer dissociation for H3-W78/H2B-109Cys-AEDANS NCPs with H2A (■, ●) and H2A.Z (▲, ◆). Values were calculated from FL intensities at 350 nm and fit to two lines which intersect near the transition midpoint (Figure 3.12). Conditions are given in legend of Figure 3.8.

References

1. Kornberg, R. D. and Lorch, Y. (1991) Irresistible force meets immovable object: transcription and the nucleosome, *Cell 67*, 833-836.

2. Khorasanizadeh, S. (2004) The nucleosome: from genomic organization to genomic regulation, *Cell 116*, 259-272.

3. Luger, K. (2006) Dynamic nucleosomes, *Chromosome. Res. 14*, 5-16.

4. Workman, J. L. and Kingston, R. E. (1998) Alteration of nucleosome structure as a mechanism of transcriptional regulation, *Annu. Rev. Biochem. 67*, 545-579.

5. Arents, G. and Moudrianakis, E. N. (1995) The histone fold: A ubiquitous architectural motif utilized in DNA compaction and protein dimerization, *Proc. Natl. Acad. Sci. U.S.A. 92*, 11170-11174.

6. Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F., and Richmond, T. J. (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution, *Nature 389*, 251-260.

7. Arents, G., Burlingame, R. W., Wang, B. C., Love, W. E., and Moudrianakis, E. N. (1991) The nucleosomal core histone octamer at 3.1 Å resolution: a tripartite protein assembly and a left-handed superhelix., *Proc. Natl. Acad. Sci. U. S. A. 88*, 10148-52.

8. Eickbush, T. H. and Moudrianakis, E. N. (1978) The histone core complex: An octamer assembled by two sets of protein-protein interactions, *Biochemistry 17*, 4955-64.

9. Oohara, I. and Wada, A. (1987) Spectroscopic studies on histone-DNA interactions. II. Three transitions in nucleosomes resolved by salt-titration, *J. Mol. Biol. 196*, 399-411.

10. Smith, S. and Stillman, B. (1991) Stepwise assembly of chromatin during DNA replication in vitro, *EMBO J. 10*, 971-980.

11. Widom, J. (1998) Structure, dynamics, and function of chromatin in vitro, *Annu Rev Biophys Biomol Struct 27*, 285-327.

12. Akey, C. W. and Luger, K. (2003) Histone chaperones and nucleosome assembly, *Curr. Opin. Struct. Biol. 13*, 6-14.

13. Park, Y. J., Dyer, P. N., Tremethick, D. J., and Luger, K. (2004) A new fluorescence resonance energy transfer approach demonstrates that the histone variant H2AZ stabilizes the histone octamer within the nucleosome, *J. Biol. Chem. 279*, 24274-24282.

14. Polach, K. J. and Widom, J. (1995) Mechanism of protein access to specific DNA sequences in chromatin: a dynamic equilibrium model for gene regulation, *J. Mol. Biol. 254*, 130-149.

15. Polach, K. J. and Widom, J. (1996) A model for the cooperative binding of eukaryotic regulatory proteins to nucleosomal target sites, *J. Mol. Biol. 258*, 800-812.

16. Meersseman, G., Pennings, S., and Bradbury, E. M. (1992) Mobile nucleosomes--a general behavior, *EMBO J. 11*, 2951-2959.

17. Pennings, S., Meersseman, G., and Bradbury, E. M. (1991) Mobility of positioned nucleosomes on 5 S rDNA, *J. Mol. Biol. 220*, 101-110.

18. Thastrom, A., Lowary, P. T., Widlund, H. R., Cao, H., Kubista, M., and Widom, J. (1999) Sequence motifs and free energies of selected natural and non-natural nucleosome positioning DNA sequences, *J. Mol. Biol. 288*, 213-229.

19. Anderson, J. D., Thastrom, A., and Widom, J. (2002) Spontaneous access of proteins to buried nucleosomal DNA target sites occurs via a mechanism that is distinct from nucleosome translocation, *Mol. Cell Biol. 22*, 7147-7157.

20. Li, G. and Widom, J. (2004) Nucleosomes facilitate their own invasion, *Nat. Struct. Mol. Biol. 11*, 763-769.

21. Li, G., Levitus, M., Bustamante, C., and Widom, J. (2005) Rapid spontaneous accessibility of nucleosomal DNA, *Nat. Struct. Mol. Biol. 12*, 46-53.

22. Tomschik, M., Zheng, H., van, H. K., Zlatanova, J., and Leuba, S. H. (2005) Fast, long-range, reversible conformational fluctuations in nucleosomes revealed by single-pair fluorescence resonance energy transfer, *Proc. Natl. Acad. Sci. U.S.A. 102*, 3278-3283.

23. Ahmad, K. and Henikoff, S. (2002) Histone H3 variants specify modes of chromatin assembly, *Proc. Nat. Acad. Sci. U.S.A. 99 Suppl 4*, 16477-16484.

24. Ausio, J., Abbott, D. W., Wang, X., and Moore, S. C. (2001) Histone variants and histone modifications: a structural perspective, *Biochem Cell Biol. 79*, 693-708.

25. Ausio, J. and Abbott, D. W. (2002) The many tales of a tail: carboxyl-terminal tail heterogeneity specializes histone H2A variants for defined chromatin function, *Biochemistry 41*, 5945-5949.

26. Henikoff, S., Furuyama, T., and Ahmad, K. (2004) Histone variants, nucleosome assembly and epigenetic inheritance, *Trends Genet. 20*, 320-326.

27. Kamakaka, R. T. and Biggins, S. (2005) Histone variants: deviants?, 19 ed..

28. Redon, C., Pilch, D., Rogakou, E., Sedelnikova, O., Newrock, K., and Bonner, W. (2002) Histone H2A variants H2AX and H2AZ, 12 ed..

29. Sullivan, S. A. and Landsman, D. (2003) Characterization of sequence variability in nucleosome core histone folds, *Proteins 52*, 454-465.

30. Dhillon, N. and Kamakaka, R. T. (2000) A histone variant, Htz1p, and a Sir1p-like protein, Esc2p, mediate silencing at HMR, *Mol. Cell 6*, 769-780.

31. Fan, J. Y., Rangasamy, D., Luger, K., and Tremethick, D. J. (2004) H2A.Z alters the nucleosome surface to promote HP1alpha-mediated chromatin fiber folding, *Mol. Cell 16*, 655-661.

32. Guillemette, B., Bataille, A. R., Gevry, N., Adam, M., Blanchette, M., Robert, F., and Gaudreau, L. (2005) Variant histone H2A.Z is globally localized to the promoters of inactive yeast genes and regulates nucleosome positioning, *PLoS. Biol. 3*, e384.

33. Leach, T. J., Mazzeo, M., Chotkowski, H. L., Madigan, J. P., Wotring, M. G., and Glaser, R. L. (2000) Histone H2A.Z is widely but nonrandomly distributed in chromosomes of Drosophila melanogaster, *J. Biol. Chem. 275*, 23267-23272.

34. Meneghini, M. D., Wu, M., and Madhani, H. D. (2003) Conserved histone variant H2A.Z protects euchromatin from the ectopic spread of silent heterochromatin, *Cell 112*, 725-736.

35. Raisner, R. M., Hartley, P. D., Meneghini, M. D., Bao, M. Z., Liu, C. L., Schreiber, S. L., Rando, O. J., and Madhani, H. D. (2005) Histone variant H2A.Z marks the 5' ends of both active and inactive genes in euchromatin, *Cell 123*, 233-248.

36. Rangasamy, D., Berven, L., Ridgway, P., and Tremethick, D. J. (2003) Pericentric heterochromatin becomes enriched with H2A.Z during early mammalian development, *EMBO J. 22*, 1599-1607.

37. Stargell, L. A., Bowen, J., Dadd, C. A., Dedon, P. C., Davis, M., Cook, R. G., Allis, C. D., and Gorovsky, M. A. (1993) Temporal and spatial association of histone H2A variant hv1 with transcriptionally competent chromatin during nuclear development in Tetrahymena thermophila, *Genes Dev. 7*, 2641-2651.

38. Zhang, H., Roberts, D. N., and Cairns, B. R. (2005) Genome-wide dynamics of Htz1, a histone H2A variant that poises repressed/basal promoters for activation through histone loss, *Cell 123*, 219-231.

39. Suto, R. K., Clarkson, M. J., Tremethick, D. J., and Luger, K. (2000) Crystal structure of a nucleosome core particle containing the variant histone H2A.Z, *Nat. Struct. Biol. 7*, 1121-1124.

40. Davey, C. A., Sargent, D. F., Luger, K., Maeder, A. W., and Richmond, T. J. (2002) Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution, *J. Mol. Biol. 319*, 1097-1113.

41. Flaus, A., Rencurel, C., Ferreira, H., Wiechens, N., and Owen-Hughes, T. (2004) Sin mutations alter inherent nucleosome mobility, *EMBO J. 23*, 343-353.

42. Abbott, D. W., Ivanova, V. S., Wang, X., Bonner, W. M., and Ausio, J. (2001) Characterization of the stability and folding of H2A.Z chromatin particles: implications for transcriptional activation, *J. Biol. Chem. 276*, 41945-41949.

43. Thambirajah, A. A., Dryhurst, D., Ishibashi, T., Li, A., Maffey, A. H., and Ausio, J. (2006) H2A.Z stabilizes chromatin in a way that is dependent on core histone acetylation, *J. Biol. Chem. 281*, 20036-20044.

44. Lakowicz, J. R. (1999) Principles of Fluorescence Spectroscopy, 2nd ed., pp 367-394, Kluwer Academic/Plenum Publishers, New York.

45. Wu, P. and Brand, L. (1994) Resonance energy transfer: methods and applications, *Anal. Biochem. 218*, 1-13.

46. Banks, D. D. and Gloss, L. M. (2003) Equilibrium folding of the core histones: the H3-H4 tetramer is less stable than the H2A-H2B dimer, *Biochemistry 218*, 1-13.

47. Luger, K., Rechsteiner, T. J., Flaus, A. J., Waye, M. M., and Richmond, T. J. (1997) Characterization of nucleosome core particles containing histone proteins made in bacteria, *J. Mol. Biol. 272*, 301-311.

48. Gloss, L. M. and Placek, B. J. (2002) The effect of salts on the stability of the H2A-H2B histone dimer, *Biochemistry 41*, 14960-14968.

49. Placek, B. J., Harrison, L. N., Villers, B. M., and Gloss, L. M. (2005) The H2A.Z/H2B dimer is unstable compared to the dimer containing the major H2A isoform, *Protein Science 14*, 514-522.

50. Nelson, D. A. and Alonso, W. R. (1983) Extraction of histones H2A, H3 and H4 from yeast nuclei. Measurement of the extent of yeast histone acetylation following one-dimensional gel electrophoresis, *Biochim. Biophys. Acta 741*, 269-271.

51. Gill, S. C. and von Hippel, P. H. (1989) Calculation of protein extinction coefficients from amino acid sequence data, *Anal. Biochem. 182*, 319-326.

52. Keller, R. P. and Neville, M. C. (1986) Determination of total protein in human milk: comparison of methods, *Clin. Chem. 32*, 120-123.

53. Lowary, P. T. and Widom, J. (1998) New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning, *J. Mol. Biol. 276*, 19-42.

54. Dyer, P. N., Edayathumangalam, R. S., White, C. L., Bao, Y., Chakravarthy, S., Muthurajan, U. M., and Luger, K. (2004) Reconstitution of nucleosome core particles from recombinant histones and DNA, *Methods Enzymol. 375*, 23-44.

55. Thastrom, A., Lowary, P. T., and Widom, J. (2004) Measurement of histone-DNA interaction free energy in nucleosomes, *Methods 33*, 33-44.

56. Myers, J. K., Pace, C. N., and Scholtz, J. M. (1995) Denaturant *m* values and heat capacity changes: Relation to changes in accessble surface areas of protein folding., *Protein Science 4*, 2138-2148.

57. Bilsel, O., Zitzewitz, J. A., Bowers, K. E., and Matthews, C. R. (1999) Folding mechanism of the alpha-subunit of tryptophan synthase, an alpha/beta barrel protein: global analysis highlights the interconversion of multiple native, intermediate, and unfolded forms through parallel channels, *Biochemistry 38*, 1018-29.

58. Zitzewitz, J. A., Bilsel, O., Luo, J., Jones, B. E., and Matthews, C. R. (1995) Probing the folding mechanism of a leucine zipper peptide by stopped-flow circular dichroism spectroscopy, *Biochemistry 34,* 12812-12819.

59. Beechem, J. M. (1992) Global analysis of biochemical and biophysical data., *Methods in Enzymology 210*, 37-54.

60. Bolen, D. W. (2001) Protein stabilization by naturally occurring osmolytes, *Methods Mol Biol 168*, 17-36.

61. Bolen, D. W. and Baskakov, I. V. (2001) The osmophobic effect: Natural selection of a thermodynamic force in protein folding., *J. Mol. Biol. 310*, 955-963.

62. Wang, A. and Bolen, D. W. (1997) A naturally occurring protective system in urea-rich cells: mechanism of osmolyte protection of proteins against urea denaturation, *Biochemistry 36*, 9101-8.

63. Henkels, C. H., Kurz, J. C., Fierke, C. A., and Oas, T. G. (2001) Linked folding and anion binding of the Bacillus subtilis ribonuclease P protein, *Biochemistry 40*, 2777-89.

64. Record, M. T., Jr., Zhang, W., and Anderson, C. F. (1998) Analysis of effects of salts and uncharged solutes on protein and nucleic acid equilibria and processes: a practical guide to recognizing and interpreting polyelectrolyte effects, Hofmeister effects, and osmotic effects of salts, *Adv. Protein Chem. 51*, 281-353.

65. Randolph, J. B. and Waggoner, A. S. (1997) Stability, specificity and fluorescence brightness of multiply-labeled fluorescent DNA probes, *Nucleic Acids Res. 25*, 2923-2929.

66. Aragay, A. M., Diaz, P., and Daban, J.-R. (1988) Association of nucleosome core particle DNA with different histone oligomers: Transfer of histones between DNA-(H2A,H2B) and DNA-(H3,H4) complexes., *J. Mol. Biol. 204*, 141-154.

67. Dorigo, B., Schalch, T., Bystricky, K., and Richmond, T. J. (2003) Chromatin fiber folding: requirement for the histone H4 N-terminal tail, *J. Mol. Biol. 327*, 85-96.

68. Schalch, T., Duda, S., Sargent, D. F., and Richmond, T. J. (2005) X-ray structure of a tetranucleosome and its implications for the chromatin fibre, *Nature 436*, 138-141.

69. Dong, F. and van Holde, K. E. (1991) Nucleosome positioning is determined by the (H3-H4)$_2$ tetramer., *Proc. Natl. Acad. Sci. U. S. A. 88*, 10596-600.

70. Gottesfeld, J. M. and Luger, K. (2001) Energetics and affinity of the histone octamer for defined DNA sequences, *Biochemistry 40*, 10927-10933.

71. Thastrom, A., Gottesfeld, J. M., Luger, K., and Widom, J. (2004) Histone-DNA binding free energy cannot be measured in dilution-driven dissociation experiments, *Biochemistry 43*, 736-741.

72. Mazurkiewicz, J., Kepert, J. F., and Rippe, K. (2006) On the mechanism of nucleosome assembly by histone chaperone NAP1, *J. Biol. Chem. 281*, 16462-16472.

73. Read, C. M., Baldwin, J. P., and Crane-Robinson, C. (1985) Structure of subnucleosomal particles. Tetrameric (H3/H4)2 146 base pair DNA and hexameric (H3/H4)2(H2A/H2B)1 146 base pair DNA complexes, *Biochemistry 24*, 4435-4450.

74. Raisner, R. M. and Madhani, H. D. (2006) Patterning chromatin: form and function for H2A.Z variant nucleosomes, *Curr. Opin. Genet. Dev. 16*, 119-124.

75. Luger, K. (2006) Dynamic nucleosomes, *Chromosome. Res. 14*, 5-16.

76. Dalbey, R. E., Weiel, J., and Yount, R. G. (1983) Forster energy transfer measurements of thiol 1 to thiol 2 distances in myosin subfragment 1, *Biochemistry 22*, 4696-4706.

77. Delano, W. L. (2002) The PYMOL Molecular Graphics System, Delano Scientific, San Carlos, CA USA. http://www.pymol.org

CHAPTER IV

CONCLUSIONS AND FUTURE DIRECTIONS

The role of kinetic intermediates in protein folding

How a protein's native structure is dictated by the sequence of one or more polypeptide chains is referred to as the protein folding code. In addition to the native structure, the stability and kinetic pathway are encoded in the primary amino acid sequence. The focus of protein folding studies is to decipher this code with the ultimate goal of being able to predict the structure of proteins from genomic data alone and to design better proteins and enzymes for biotechnological and medical applications.

A focus of our lab is to elucidate the folding mechanisms of a group of protein oligomers that contain segment swapped structures. The group that we are studying is collectively called the intertwined, segment swapped alpha-helical DNA-binding dimers, or ISSADDs. These proteins include: 1) FIS, the bacterial Factor for Inversion Stimulation (*1*); 2) the prokaryotic Zn-repressor protein, SmtB; (Topping & Gloss, manuscript in preparation) 3) the archael histones, hPyA1 from a *Pyrococcus* strain, hFoB from *Methanobacterium formicicum*, and hMfB from *Methanothermus fervidus*; (*2*) and 4) the eukaryotic histone heterodimers, H2A-H2B and H3-H4 (*3;4*). The goal of studying these proteins is to: 1) understand the coordination of monomer 2° & 3° folding with development of 4° structure, as well as the relationships between folding landscapes and helical topology; and 2) gain insight into the role of kinetic intermediates in protein folding. Currently there is a debate in the field of protein folding on whether intermediates are productive species that contribute to efficient folding or kinetic traps that can lead to misfolding and aggregation. This argument was sparked by the observation that many small, monomeric, single-domain proteins are able to fold efficiently in a simple two-state manner with no detectable intermediates (*5*).

Chapter Two describes the folding mechanism of the ISSADD protein, FIS, a small homodimeric DNA binding protein found in enteric bacteria including *Escherichia coli*. Its structure is classified as a pseudo-domain swapped structure because it does not populate any monomeric equilibrium intermediates. FIS was determined to fold via an obligatory on-pathway dimeric intermediate that is formed in the dead time of the stopped-flow instrument, and therefore its rate of association is approaching the diffusion limit of 10^9 $M^{-1}s^{-1}$ (*6*). The dimeric intermediate then undergoes further folding through a first order reaction to the native dimer. This body of work can now be compared to the folding mechanisms of other members of the ISSADD proteins mentioned above. Interestingly, in this system there is a positive correlation between the rate of folding and the presence of intermediates. Therefore, ISSADDs may utilize intermediates as stepping-stones to help coordinate the secondary and tertiary structural elements of two polypeptide chains during dimerization.

Characterization and understanding the role of early folding intermediates is an important aspect of protein folding studies (*7*). The stopped-flow methods employed to study the folding mechanism of FIS were limited in sensitivity because of the lack of sensitive fluorescence probes and can only measure relaxation times on the timescale of milliseconds. There are four tyrosines per monomer (Y39, Y69, Y51, Y95) located throughout the three dimensional structure (*8*) and make different contributions to the total FL, with Y51 exhibiting the highest FL (*9*). However, in our studies, the unfolding transition measured by intrinsic tyrosine FL was not discernible from the folded and unfolded baselines. Therefore, only far-UV circular dichroism was used to monitor the folding reaction, thus limiting our understanding of the kinetic mechanism to the

coordination of secondary structure. One common strategy to improve FL sensitivity is to introduce an unique tryptophan into a protein of interest through site-directed mutagenesis. An advantage of tryptophan is that it has a typical quantum yield ~5-fold greater than tyrosine. However, care must be taken in choosing the mutation sites to avoid perturbations of protein structure and stability arising from the size and hydrophobicity of Trp.

Using site-directed mutagenesis, three unique tryptophans were introduced into FIS (Figure 4.1) at sites that would be sensitive to dimerization and folding. The locations of mutations were Ser30, Leu50 and Met65. The criteria for choosing these sites were: 1) located in regions partially excluded from solvent but not completely buried where there is little room to accommodate a larger residue; and 2) in or near the dimer interface. Since FIS is a naturally occurring *E. coli* DNA binding protein, present at high levels during certain growth conditions, a FIS deletion *E. coli* strain was obtained from Dr. Reid Johnson (Department of Biological Chemistry, UCLA) (*10*). The mutants were successfully overexpressed and purified using the methods described in Chapter Two. FL emission scans of all FIS mutants exhibited a significant change in the emission maximum upon unfolding in urea. Figure 4.2 shows the FL emission scans of FIS-S30W as an example.

The stability to urea-induced denaturation of the M65W Trp mutant was determined from titrations monitoring both far-UV CD and Trp FL. Unlike WT Tyr FL, a clear unfolding transition between relatively shallow unfolded and unfolded baselines was observed. Like WT FIS, the mutant folded via a reversible two-state mechanism with no evidence of hysteresis (Figure 4.3). CD and FL titrations were coincident and

protein concentration dependent as expected for a dimerization reaction with no detectable equilibrium intermediates. The $\Delta G°(H_2O)$ and *m* values were 14.7 ± 0.2 kcal mol^{-1} and 2.5 ± 0.1 kcal mol^{-1} M^{-1}, respectively. These values were similar to the WT values of 15.2 ± 0.8 kcal mol^{-1} and 2.9 ± 0.3 kcal mol^{-1} M^{-1}, respectively (Chapter 2; Table 2.1).

Kinetic refolding experiments were complicated by an unforeseen technical difficulty. Unlike WT FIS, the stocks of the Trp mutants couldn't be stored for more than a week. Aggregation and visible precipitation occurred upon storage at -20°C in the presence of 50% glycerol, conditions employed for the WT protein. When WT FIS was overexpressed in the FIS deletion strain, the purified protein was also aggregation prone.

The effects of the tendency to aggregate became evident in SF-FL refolding kinetics. The urea dependence of the refolding relaxation times were significantly different than those observed for WT (*1*). The kinetic traces were best fit by a single first order exponential (Figure 4.4A). The observed relaxation times as a function of denaturant are shown in Figure 4.4B, a negative $m^{\ddagger}$ value for each phase (Chapter 2, Equation 3). The slope is indicative of an unfolding reaction with exposure of solvent accessible surface area. Similar results were observed for SF-FL refolding reactions of the S30W and L50W mutants. The plausible explanation is that the refolding is dominated by the unfolding of a misfolded or aggregated off-pathway species (Chapter 2, Figure 2.6.B). By unknown mechanisms, this aggregation tendency appears to have been induced by the choice of *E. coli* expression strain. Given this inability to produce well-behaved FIS mutants, research focus necessarily shifted away from the folding mechanism of FIS to the design of a Forster energy transfer system to monitor the

dissociation of nucleosome core particles. This work was described in Chapter 3 and will be discussed further in the next section.

Future thermodynamic and kinetic NCP studies

The nucleosome core particle is an extremely dynamic complex that is critical in controlling the access of cellular machinery to DNA in processes such as replication, transcription and repair. ATP-dependent chromatin remodeling factors, chaperones, post-translational modifications and incorporation of histone variants can alter the structure of the nucleosome and higher order chromatin. Even in isolation, the nucleosome is inheritably mobile and able to "facilitate its own invasion" (*11*). While the structure and dynamics of chromatin structure have been studied intensely for more than three decades (*12;13*), much is still not understood, in particular the thermodynamics of the incorporation of histone variants and the kinetics of nucleosome assembly and disassembly.

The assembly/disassembly pathway of the NCP is a complex reaction that includes several competing structures and processes that make it difficult to determine an absolute free energy of dissociation (*14*). Various sub-nucleosomal structures such as hexameric intermediates (H2A-H2B/$(H3\text{-}H4)_2$-DNA) (*15*), and non-nucleosomal species (*16*) may exist under certain conditions. This is further complicated by the fact that H2A-H2B dimers can exchange between nucleosomes as well as weakly associate with free DNA (*17*). The mutiple protein-protein FRET system described in Chapter Three is the first to directly calculate a free energy of dissociation, ΔG, for H2A-H2B dimer, as well monitor the cooperative nature of dimer dissociation.

The incorporation of histone variants into the nucleosome core particle is one way chromatin structure can be regulated (for review, (*18*)). A very powerful and important use of FRET NCP systems is their ability to monitor the effect of variants on nucleosome disassembly (*19;20*). To date, the biophysical properties of H2A.Z NCPs are controversial (described in Chapter Three) and need to be further resolved. The FRET system employed in this study shows that the change in stability upon incorporation of H2A.Z is minimal with the largest difference occurring in the cooperativity of the dissociation of the dimers. In comparing H2A.Z and H2A NCPs, the results show that there is little difference in the dissociation of the first H2A- or H2A.Z-H2B dimer but the second H2A.Z-H2B dimer comes off the nucleosome more easily then the second H2A-H2B dimer. The cooperativity change observed may be explained by the differences in the L1-L1' loop interaction between the two H2A.Z dimers (*20*). Further evidence may lie in fact that the other H2A.1 variants, such as H2A.Bbd and macroH2A, also differ in this region (*18*). In future experiments we would like to further explore this possibly by swapping the L1 loop region on H2A.Z with the original sequence in WT. Additionally, we would like to further explore the effect variants have on nucleosome assembly, in particular, H2A.X.

H2A.X variants are found throughout most eukaryotic species, from animals to protists and share high sequence homology with H2A.1 (*21*). The majority of the sequence differences are found in the C-terminal domain. Its rather long C-terminal tail has been implicated in DNA double-stranded break repair. The extended C-terminal region contains an SQ phosphorylation target motif and is rapidly phosphorylated in response to DNA double-stranded breaks (for review, (*21*)). Interestingly, there is a

substitution of a His into the L1 region, similar to the other variants, although no other amino acids are changed. Using our FRET system, it would be interesting to examine whether any changes occur in nucleosome dynamics when H2A.X is incorporated (both unmodified and phosphorylated) or if it is used largely as a signaling marker.

Only equilibrium effects of histone variant incorporation on nucleosome assembly/disassembly have been studied thus far, with little work characterizing changes in NPC dynamics. The only report to date has studied the movement of the DNA ends away from the nucleosome core particle under physiological conditions (*11;22*). Stopped-flow FRET analysis of DNA breathing in nucleosomes determined that this process can occur on time scales of tens to hundreds of milliseconds (*22*). Preliminary kinetic experiments with this protein-protein FRET system demonstrated that the majority of the interactions in dimer dissociation occur faster than the timescale of manual mixing methods. Stopped-FL experiments will be employed to look at dimer association and dissociation rates of WT and variant NCPs. Another method to look at the dynamics of WT and variant NCPs is to monitor the physical exchange of dimers using our FRET system. NCPs containing Trp on either H3 or H4, and assembled with WT or variant dimers without Cys-AEDANS acceptors could be titrated with increasing amounts of their FRET counterpart (either the variant or WT FRET dimer) and the gain in FRET FL could be monitored over time. The opposite setup could also be performed, with the loss of FRET being observed over time. The experiments would be able to provide information on the k_{ON} and k_{OFF} rates and would correlate to the dimer equilibrium constant, K_{EQ}. This system is a very powerful and versatile tool that will

help provide a better understanding of how histone sequence and stability impacts nucleosome function in the regulation of the cell's DNA-templated chemistries.

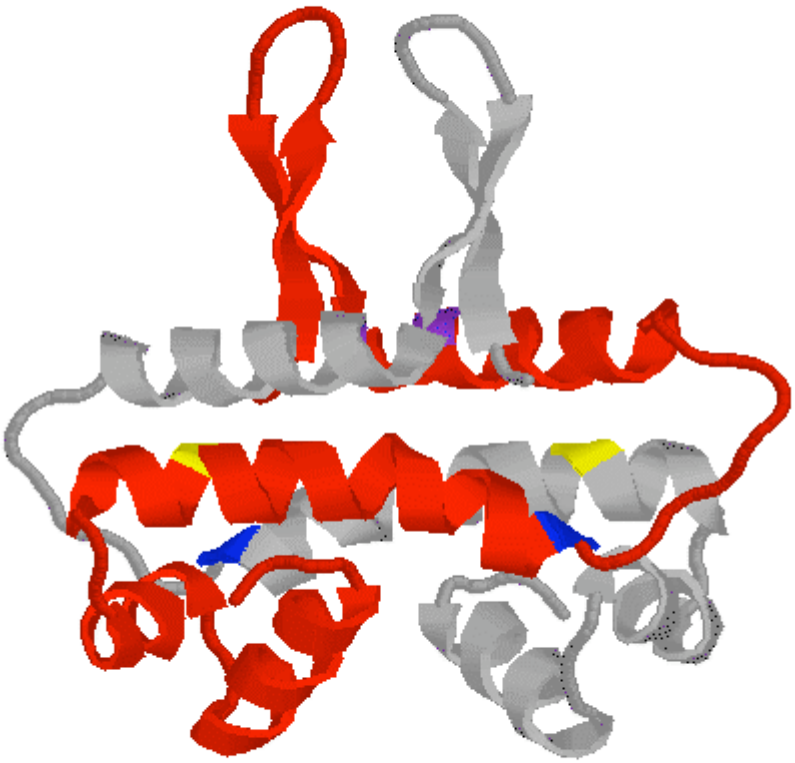


Figure 4.1. Ribbon model of a mutant FIS (K36E) in which the position of the N-terminal □-hairpins are resolved (pdb 1F36.pdb, (*23*)), depicting the locations of the Trp mutations. The two chains shown in red and gray and the locations of the Trp mutations are indicated: Ser-30, purple; Leu-50, blue; Met-65, yellow.

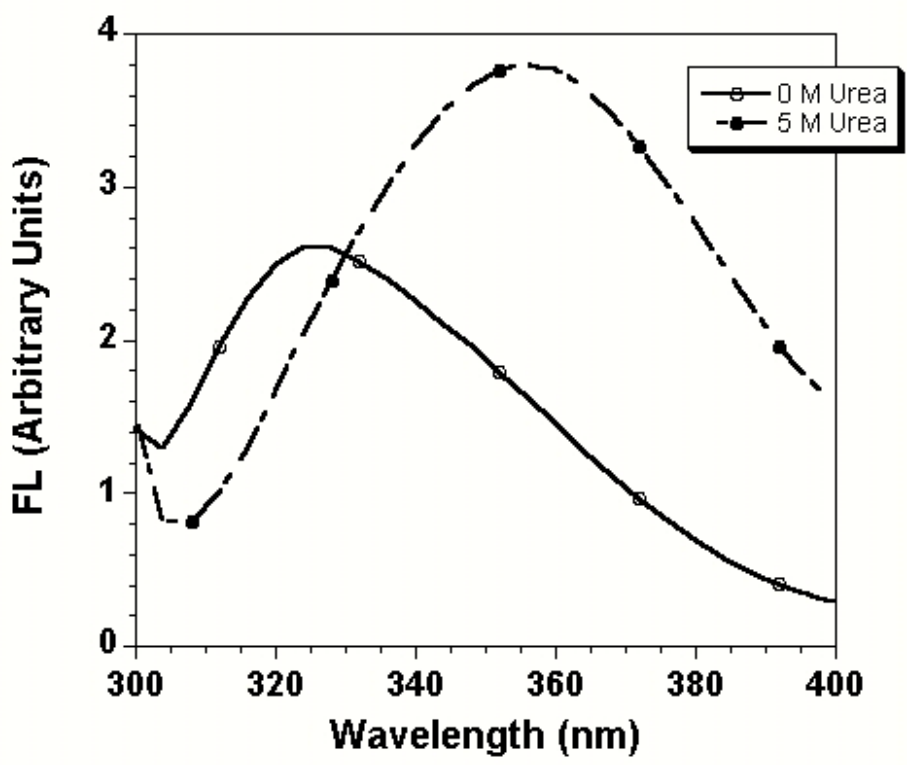


Figure 4.2. FL emission scans of folded and unfolded FIS S30W. Conditions: 200 mM KCl, 20 mM KPi, pH 7.2, 1 mM EDTA, 25 °C.

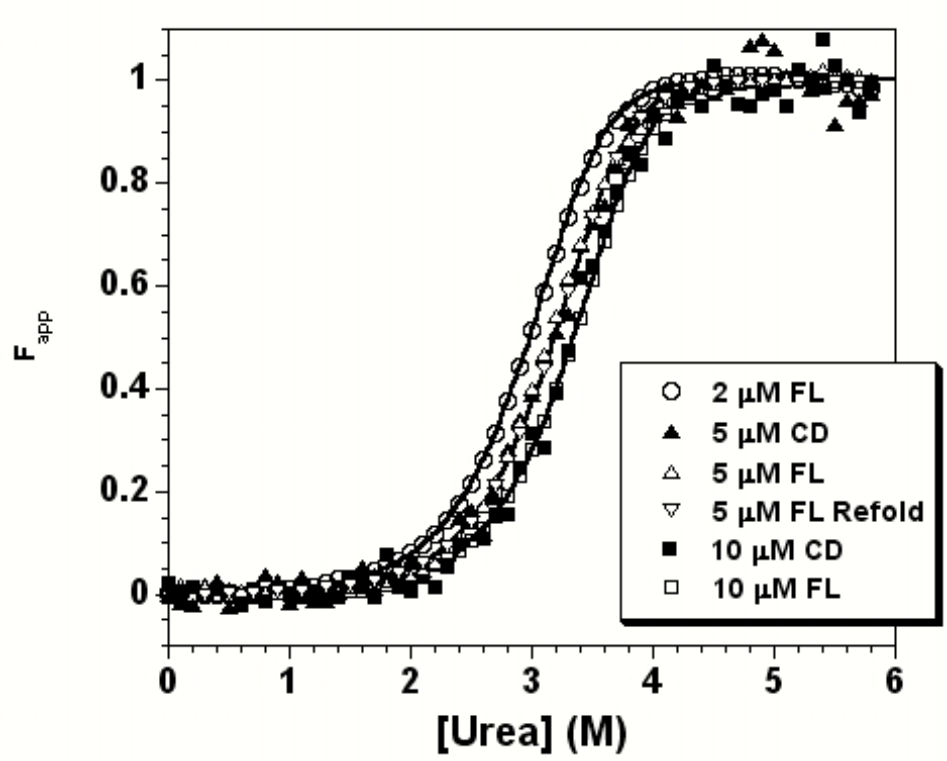


Figure 4.3. Equilibrium urea-induced unfolding transitions for FIS M65W monitored by far-UV CD and Trp FL. The solid lines represent global fits of the data to a two-state model for the equilibrium between native dimer and two unfolded monomers. Conditions: 200 mM KCl, 20 mM KPi, pH 7.2, 1 mM EDTA, 10 °C.

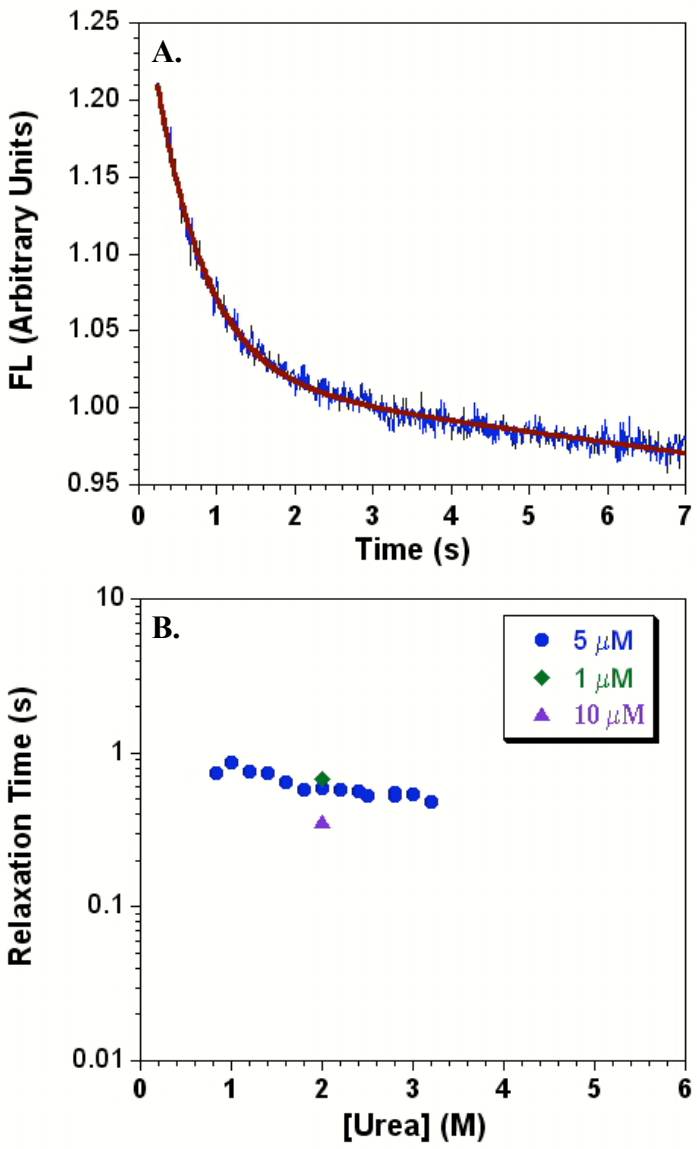


Figure 4.4. Urea dependence of SF-FL refolding kinetics for FIS M65W. **A.** Refolding kinetic trace to 1.6 M Urea. Solid line represents the local fit of the kinetic data to one first order exponential (Chapter 2, Equation 3) with an extra linear term to account for photobleaching. **B.** Relaxation times from local fits of the kinetic data. Conditions: 200 mM KCl, 20 mM KPi, pH 7.2, 1 mM EDTA, 10 °C.

Literature Cited

1. Topping, T. B., Hoch, D. A., and Gloss, L. M. (2004) Folding mechanism of FIS, the intertwined, dimeric factor for inversion stimulation, *J Mol Biol 335*, 1065-1081.

2. Topping, T. B. and Gloss, L. M. (2004) Stability and folding mechanism of mesophilic, thermophilic and hyperthermophilic archael histones: the importance of folding intermediates, *J Mol Biol 342*, 247-260.

3. Banks, D. D. and Gloss, L. M. (2004) Folding mechanism of the (H3-H4)2 histone tetramer of the core nucleosome, *Protein Sci 13*, 1304-1316.

4. Placek, B. J. and Gloss, L. M. (2005) Three-state kinetic folding mechanism of the H2A/H2B histone heterodimer: the N-terminal tails affect the transition state between a dimeric intermediate and the native dimer, *J Mol Biol 345*, 827-836.

5. Schindler, T., Herrler, M., Marahiel, M. A., and Schmid, F. X. (1995) Extremely rapid protein folding in the absence of intermediates, *Nature Struct. Biol. 2*, 663-673.

6. McCammon, J. A. (1996) A speed limit for protein folding, *Proc. Natl. Acad. Sci. U. S. A 93*, 11426-11427.

7. Roder, H. and Colon, W. (1997) Kinetic role of early intermediates in protein folding, *Curr. Opin. Struct. Biol. 7*, 15-28.

8. Kostrewa, D., Granzin, J., Stock, D., Choe, H. W., Labahn, J., and Saenger, W. (1992) Crystal structure of the factor for inversion stimulation FIS at 2.0 Å resolution, *J Mol Biol 226*, 209-26.

9. Boswell, S., Mathew, J., Beach, M., Osuna, R., and Colon, W. (2004) Variable contributions of tyrosine residues to the structural and spectroscopic properties of the factor for inversion stimulation, *Biochemistry 43*, 2964-2977.

10. Pan, C. Q., Finkel, S. E., Cramton, S. E., Feng, J. A., Sigman, D. S., and Johnson, R. C. (1996) Variable structures of Fis-DNA complexes determined by flanking DNA-protein contacts, *J Mol Biol 264*, 675-95.

11. Li, G. and Widom, J. (2004) Nucleosomes facilitate their own invasion, *Nat Struct. Mol. Biol. 11*, 763-769.

12. van Holde, K. (1989) Chromatin, (Rich, A., Ed.) Springer-Verlag, New York.

13. Luger, K. and Hansen, J. C. (2005) Nucleosome and chromatin fiber dynamics, *Curr. Opin. Struct. Biol. 15*, 188-196.

14. Thastrom, A., Gottesfeld, J. M., Luger, K., and Widom, J. (2004) Histone-DNA binding free energy cannot be measured in dilution-driven dissociation experiments, *Biochemistry 43*, 736-741.

15. Read, C. M., Baldwin, J. P., and Crane-Robinson, C. (1985) Structure of subnucleosomal particles. Tetrameric (H3/H4)2 146 base pair DNA and hexameric (H3/H4)2(H2A/H2B)1 146 base pair DNA complexes, *Biochemistry 24*, 4435-4450.

16. Royer, C. A., Ropp, T., and Scarlata, S. F. (1992) Solution studies of the interactions between the histone core proteins and DNA using fluorescence spectroscopy, *Biophys. Chem. 43*, 197-211.

17. Aragay, A. M., Diaz, P., and Daban, J.-R. (1988) Association of nucleosome core particle DNA with different histone oligomers: Transfer of histones between DNA-(H2A,H2B) and DNA-(H3,H4) complexes., *J. Mol. Biol. 204*, 141-154.

18. Kamakaka, R. T. and Biggins, S. (2005) Histone variants: deviants?, *Genes Dev 19*, 295-310.

19. Bao, Y., Konesky, K., Park, Y. J., Rosu, S., Dyer, P. N., Rangasamy, D., Tremethick, D. J., Laybourn, P. J., and Luger, K. (2004) Nucleosomes containing the histone variant H2A.Bbd organize only 118 base pairs of DNA, *EMBO J. 23*, 3314-3324.

20. Park, Y. J., Dyer, P. N., Tremethick, D. J., and Luger, K. (2004) A new fluorescence resonance energy transfer approach demonstrates that the histone variant H2AZ stabilizes the histone octamer within the nucleosome, *J Biol Chem 279*, 24274-24282.

21. Redon, C., Pilch, D., Rogakou, E., Sedelnikova, O., Newrock, K., and Bonner, W. (2002) Histone H2A variants H2AX and H2AZ, *Curr Opin Genet Dev 12*, 162-169.

22. Li, G., Levitus, M., Bustamante, C., and Widom, J. (2005) Rapid spontaneous accessibility of nucleosomal DNA, *Nat Struct. Mol. Biol. 12*, 46-53.

23. Safo, M. K., Yang, W. Z., Corselli, L., Cramton, S. E., Yuan, H. S., and Johnson, R. C. (1997) The transactivation region of the fis protein that controls site-specific DNA inversion contains extended mobile beta-hairpin arms, *EMBO J 16*, 6860-6873.