Kinetics of Coupled Binding and Conformational Change in Proteins and RNA

by

Kyle Gabriel Daniels

Department of Biochemistry
Duke University

Date:_______________________

Approved:

___________________________
Terrence G. Oas, Supervisor

___________________________
Michael C. Fitzgerald

___________________________
Kenichi Yokoyama

___________________________
Michael D. Been

___________________________
Scott C. Schmidler

Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biochemistry in the Graduate School of Duke University

2015
ABSTRACT

Kinetics of Coupled Binding and Conformational Change in Proteins and RNA

by

Kyle Gabriel Daniels

Department of Biochemistry
Duke University

Date:_______________________

Approved:

___________________________
Terrence G. Oas, Supervisor

___________________________
Michael C. Fitzgerald

___________________________
Kenichi Yokoyama

___________________________
Michael D. Been

___________________________
Scott C. Schmidler

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biochemistry in the Graduate School of Duke University

2015
Abstract

Ligand binding can modulate function of proteins and nucleic acids by changing both the populations of functionally distinct conformational states and the timescales on which they interconvert. For this reason, both thermodynamic and kinetic details of coupling can be important to proper function. How tightly does ligand bind to the different conformational states? What effect does ligand binding have on the conformational equilibrium and conformational kinetics? On what timescales and in what order do binding and conformational change occur? Using a combination of stopped-flow kinetics, isothermal titration calorimetry, and x-ray crystallography, we determine the mechanisms of coupled binding and conformational change in protein (Bacillus subtilis RNase P protein) and RNA (DP17 biosensor) systems.

The results demonstrate that rigorous kinetic analysis can be used to estimate the equilibrium and rate constants for conformational changes, as well as the affinities of ligands for different conformational states. A single ligand can bind to different conformational states of the same protein or nucleic acid with affinities that differ by orders of magnitude. This binding shifts the conformational equilibrium towards the higher affinity state through a combination of increasing rate constants for the forward conformational change and decreasing rate constants for the reverse conformational change. Using a flux-based analysis of the mechanisms we show that molecular recognition is kinetically partitioned between a number of pathways that differ by the order in which binding and conformational change occur. The absolute and relative flux through these pathways varies with ligand concentration, the affinities of the ligand for the various conformational states, and the ability of ligand to accelerate the conformational change. Together, the results give insights into how biological function depends on the kinetic and thermodynamic details of coupled binding and conformational change.
Dedication

For my mother, Anita, and my father, John, who have always loved and supported me.
# Contents

Abstract ........................................................................................................................................ iv

List of Tables .................................................................................................................................. x

List of Figures ............................................................................................................................... xi

List of Schemes .............................................................................................................................. xiii

List of Commonly Used Abbreviations .......................................................................................... xiv

Acknowledgements ........................................................................................................................ xv

1. Introduction ................................................................................................................................ 1

1.1 Coupled Binding and Conformational Change ................................................................. 1

1.2 Biological Importance of Coupled Binding and Conformational Change ...................... 2

1.2.1 State Machines & Allostery .............................................................................................. 2

1.2.2 Implications for Drug Design ............................................................................................. 5

1.3 Proteins and Nucleic Acids Are Regulated by Coupled Binding and Conformational Change .................................................................................................................. 6

1.3.1 Proteins ................................................................................................................................. 6

1.3.2 RNA ..................................................................................................................................... 8

1.3.3 DNA ................................................................................................................................... 10

1.4 Thermodynamic Description of Coupled Binding and Conformational Change .......... 10

1.4.1 Thermodynamic Description: Two-State One-Site Model ........................................... 10

1.4.2 Allostery and Cooperativity ................................................................................................. 13

1.5 Kinetic Description of Coupled Binding and Conformational Change .......................... 15

1.5.1 Kinetic Description: Two-State One-Site Model ............................................................ 15

1.5.2 The MWC and KNF Models ............................................................................................... 19

1.5.3 Molecular Recognition: Conformational Selection and Induced Fit ........................... 20
1.5.4 A Flux Description of Conformational Selection and Induced Fit .......................... 21
1.5.5 Kinetic Versus Thermodynamic Control .......................................................... 23
1.5.6 Kinetic Cooperativity ......................................................................................... 27
1.6 Structural, Thermodynamic, and Kinetic Studies of Coupled Binding and Conformational Change ................................................................. 30
1.6.1 Structural Studies ............................................................................................... 30
1.6.2 Thermodynamic Studies .................................................................................... 31
1.6.3 Kinetic Studies .................................................................................................... 32
1.6.4 Data Analysis ..................................................................................................... 33
2. Ligand Concentration Regulates the Pathways of Coupled Protein Folding and Binding ...... 36
2.1 Introduction ............................................................................................................. 36
2.1.1 Challenging the False Dichotomy of Conformational Selection versus Induced Fit ... 36
2.1.2 Bacillus subtilis RNase P Protein ........................................................................ 37
2.2 Results and Discussion ........................................................................................ 39
2.2.1 Structural and Functional Characterization of P-Pro ......................................... 39
2.2.2 U, I, and F Bind PPi with Distinct Affinities ......................................................... 43
2.2.3 PPi binding significantly alters folding and unfolding rate constants. ................. 51
2.2.4 Flux is Kinetically Partitioned Between the Pathways of Coupled Folding and Binding ................................................................................................. 52
2.2.5 Determinants of the Kinetic Partitioning of Flux ................................................ 55
2.3 Conclusion ............................................................................................................. 55
2.4 Materials and Methods ........................................................................................ 56
2.4.1 Expression and Purification of P-Pro ................................................................. 56
2.4.2 Crystallography ................................................................................................. 57
2.4.3 Enzymology ....................................................................................................... 57
2.4.4 Stopped-Flow Fluorescence ................................................................. 58
2.4.5 Analysis of Kinetic Data ........................................................................ 58
2.4.6 Flux Calculation ..................................................................................... 59
2.4.7 Isothermal Titration Calorimetry ........................................................... 60

3. Conformational Kinetics Reveal Affinities of Protein Conformational States ......................................................................................... 61
3.1 Introduction ................................................................................................ 61
3.2 Results and Discussion ............................................................................. 65
  3.2.1 Ligand-bound U, I, and F are populated in kinetic experiments, but not equilibrium experiments. .......................................................... 65
  3.2.2 Ligand binding affinity and specificity increase as P protein folds. .......... 75
  3.2.3 The magnitude of change in the folding and unfolding rate constants depends on κ. . 82
  3.2.4 The predominant P⁻Pro⁻ folding mechanism shifts from indirect to direct upon ligand binding ............................................................... 85
  3.2.5 The mechanism of molecular recognition is strongly dependent on κ and δ. .... 86
3.3 Conclusion ................................................................................................... 88
  3.3.1 Towards More Quantitative Descriptions of Molecular Recognition ....... 88
  3.3.2 Using Conformational Kinetics to Study Coupled Binding and Conformational Change ................................................................. 89
3.4 Materials and Methods ............................................................................. 90
  3.4.1 Stopped-Flow Kinetics ......................................................................... 90
  3.4.2 Analysis of Kinetic Data ....................................................................... 91
  3.4.3 Isothermal Titration Calorimetry .......................................................... 93
  3.4.4 Generation of Flux Contour Plots for Figure 3-9. ................................... 93
4. Kinetic Mechanism of cyclic di-GMP Sensing by DP17 RNA Biosensor ................................................................................................. 94
  4.1 Introduction .............................................................................................. 94
4.1.1 Biosensors ................................................................. 94
4.1.2 RNA Biosensors ........................................................... 94
4.1.3 Cyclic dinucleotide sensing in cells with VC2 and DP17 .......... 98
4.2 Results and Discussion .................................................. 99
4.2.1 Stopped-Flow Fluorescence ........................................... 99
4.2.2 Stopped-Flow Fluorescence – Reduced Photobleaching .......... 106
4.2.3 Allosteric Communication Between the c-di-G and DFBHI Binding Sites .... 109
4.2.4 DP17 Activation Mechanism .......................................... 115
4.2.5 Speeding Up Biosensor Activation ................................... 120
4.2.6 Use of a Covalently Attached Fluorophore ....................... 121
4.3 Conclusion .................................................................... 122
4.4 Materials and Methods .................................................. 122
4.4.1 Materials Preparation .................................................. 122
4.4.2 Stopped-Flow Fluorescence ........................................... 123
4.4.3 Data Analysis ............................................................. 124
Appendix A .................................................................. 125
Appendix B .................................................................. 131
References .................................................................. 135
Biography .................................................................. 143
List of Tables

Table 1-1: Interpreting δ value: effect of ligand on conformational rate constants .................. 18
Table 2-1: Crystallography statistics for P-Pro in complex with pyrophosphate .................... 41
Table 2-2: Dissociation constants for P-Pro and pyrophosphate ........................................ 46
Table 2-3: Conformational parameter value estimates from fit of stopped-flow data ............... 46
Table 2-4: Rate constant parameter value estimates from fit of stopped-flow data ................. 47
Table 2-5: Signal parameter value estimates from fit of stopped-flow data .......................... 47
Table 2-6: Dissociation constants for the P-Pro unfolded state and pyrophosphate ............... 48
Table 3-1: Parameter values from global fit of PPi and TMAO stopped-flow data .................. 68
Table 3-2: Parameter values from fit of P, stopped-flow data ............................................ 68
Table 3-3: ADP parameter values from global fit of NDP stopped-flow data ....................... 69
Table 3-4: GDP parameter values from global fit of NDP stopped-flow data ....................... 70
Table 3-5: UDP parameter values from global fit of NDP stopped-flow data ....................... 71
Table 3-6: CDP parameter values from global fit of NDP stopped-flow data ....................... 72
Table 3-7: Comparison of unliganded folding parameter values from global fits of PPi & TMAO data and NDP data .......................................................... 73
Table 4-1: Populations of C1 and C2 conformational states .............................................. 112
Table 4-2: Rate constants for C1 to C2 conformational changes ........................................ 112
Table 4-3: Association constants for C1 and C2 binding to c-di-G and DFHBI ...................... 112
Table 4-4: Fluorescence intensity of DFHBI ......................................................................... 113
Table 4-5: Association rate constants for C1 and C2 binding to c-di-G and DFHBI ............... 115
List of Figures

Figure 1-1: Proteins and nucleic acids are state machines .......................................................... 3
Figure 1-2: Coupled binding and conformational change ............................................................... 4
Figure 1-3: Kinetic control .............................................................................................................. 26
Figure 1-4: Kinetic cooperativity ................................................................................................. 29
Figure 2-1: P Pro coupled folding and binding scheme ................................................................. 38
Figure 2-2: Crystal structure of P Pro ............................................................................................ 40
Figure 2-3: Single and multiple turnover cleavage of fluorescent pre-tRNA by RNase P .......... 42
Figure 2-4: Kinetics of TMAO-induced and pyrophosphate-induced folding of P Pro ............... 44
Figure 2-5: TMAO- and PPi-induced folding of P Pro ................................................................. 45
Figure 2-6: Binding of pyrophosphate to P Pro and mimics of the P Pro N-terminus .................. 50
Figure 2-7: Total equilibrium flux for the U ⇌ FL2 reaction vs. ligand concentration .......... 53
Figure 2-8: Concentration-dependent mechanism of coupled folding and binding ............... 54
Figure 3-1: P Pro coupled folding and binding scheme ................................................................. 64
Figure 3-2: Kinetics of ligand-induced folding of P Pro ............................................................. 66
Figure 3-3: Kinetics of ligand-induced folding of P Pro ............................................................. 66
Figure 3-4: Populations of ligand-bound U, I, and F in kinetic and equilibrium experiments ..... 74
Figure 3-5: Affinities for U, I, and F for six ligands ..................................................................... 76
Figure 3-6: General scheme for coupled binding and conformational change ....................... 77
Figure 3-7: Isothermal titration calorimetry of P Pro with NDPs ............................................. 80
Figure 3-8: Ligand-induced perturbations of conformational kinetics ........................................ 83
Figure 3-9: Partitioning of flux through conformational selection and induced-fit pathways ..... 87
Figure 4-1: RNA Biosensors: Riboswitch aptamer and Spinach aptamer fusions .................... 95
Figure 4-2: Crystal structure of the Ve2 riboswitch aptamer .................................................. 96
Figure 4-3: Crystal structure of the Spinach aptamer bound to DFHBI ................................. 97
Figure 4-4: c-di-G concentration dependence of DP17 activation ......................................... 101
Figure 4-5: DFHBI concentration dependence of DP17 activation ....................................... 102
Figure 4-6: DP17 activation after pre-incubation with DFHBI ............................................. 104
Figure 4-7: Photobleaching of free DFHBI ........................................................................ 104
Figure 4-8: Photobleaching of free DFHBI and DP17 ternary complex ............................... 105
Figure 4-9: c-di-G concentration dependence of DP17 activation without photobleaching .... 107
Figure 4-10: DFHBI concentration dependence of DP17 activation without photobleaching .... 108
Figure 4-11: Fractional flux through DP17 activation pathways .......................................... 117
Figure 4-12: Fractional flux through DP17 activation pathways .......................................... 118
Figure 4-13: Time-dependent populations of species in DP17 activation ............................ 119
List of Schemes

Scheme 1-1: Thermodynamic Model of Two-State One-Site Coupled Conformational Change and Binding ................................................................. 12

Scheme 1-2: Kinetic Model of Two-State One-Site Coupled Conformational Change and Binding ........................................................................... 16

Scheme 1-3: Conformational Selection and Induced Fit ............................................... 22

Scheme 1-4: Coupled Binding and Conformational Change Under Kinetic Control ........ 25

Scheme 1-5: Coupled Binding and Conformational Change in an Enzyme .................... 28

Scheme 4-1: DP17 Activation Scenario I ........................................................................ 111

Scheme 4-2: DP17 Activation Scenario II ........................................................................ 111
List of Commonly Used Abbreviations

ADP – adenosine diphosphate

c-di-G – cyclic di-guanosine monophosphate

CDP – cytidine diphosphate

DFHBI – 3,5-difluoro-4-hydroxybenzylidene imidazolinone

DNA – deoxyribonucleic acid

GDP – guanosine diphosphate

NDP – nucleotide diphosphate

P₁ – inorganic phosphate

PP₁ – inorganic pyrophosphate


RNA – ribonucleic acid

UDP – uridine diphosphate
Acknowledgements

I would like to thank the many people who have provided me with intellectual, experimental, emotional, and financial support during my time at Duke University. My utmost thanks goes to my advisor Terry, who has given me both the guidance and space to grow as a scientist. I would also like to thank the Oas lab members who I have had the pleasure of working with over the past years. I want to thank all of Duke Biochemistry, including the administrators, students, postdocs, and faculty. Thank you to Richard Brennan for use of the ITC and for career guidance. A special thank you to my thesis committee: Michael Fitzgerald, Kenichi Yokoyama, Mike Been, and Scott Schmidler. I would also like to thank Hashim Al-Hashimi and Homme Hellinga for insightful discussion about my research.

I want to acknowledge my collaborators in the groups of Maria Schumacher, Carol Fierke, and Scott Schmidler. Thank you, Dorothy Beckett, for continued support and guidance throughout my time at Duke.

Finally, I’d like to thank my family and close friends for all they have done to help me become the scientist and person that I am today.
1. Introduction

1.1 Coupled Binding and Conformational Change

Most biological reactions rely on interplay between binding and changes in both macromolecular structure and dynamics. The two underlying phenomena—binding and conformational change—are themselves central to biology. Binding is among the most ubiquitous and important processes seen in biological systems. Binding of proteins, metals, water, small molecules, peptides, proteins and nucleic acids constitute most biological interactions. The biological functions of proteins and nucleic acids are intimately tied to their conformational heterogeneity. Many proteins and nucleic acids can occupy multiple discrete conformational states that differ by structural differences as small as residue rotamer changes and as large as global folding/unfolding events. Coupling of binding and conformational change is a powerful regulatory strategy exploited by many if not all biological systems. In this introduction, we present coupled binding and conformational change in simple thermodynamic and kinetic frameworks. I explain using the thermodynamic framework how coupled binding and conformational change can lead to thermodynamic cooperativity. I use the kinetic framework to explore mechanisms of molecular recognition and demonstrate how coupled binding and conformational change can generate kinetic cooperativity. Finally, we review important considerations for experiments and analyses used to study coupled binding and conformational change.
1.2 Biological Importance of Coupled Binding and Conformational Change

1.2.1 State Machines & Allostery

Since both conformational change and binding are critical to the functions of biological macromolecules, coupling of the two provides an effective means of functional regulation. Additionally, coupling of binding and conformational change plays an important role in driving catalysis. Both the functional regulation (usually through allostery) and catalysis associated with coupled binding and conformational change have similar molecular origins (1). To understand the utility of coupling binding and conformational change, it is useful to view proteins and nucleic acids as state machines.

Both proteins and nucleic acids are conformationally heterogeneous macromolecules whose functions are intimately tied to their conformations. Biological macromolecules have evolved conformational ensembles that comprise discrete states (2). These discrete states are collections of numerous conformations (3), which are often functionally distinct, making the discrete conformational states functionally distinct (Figure 1-1). The discrete states are separated by kinetic barriers (Figure 1-1) whose origins lie in the cooperativity of the conformational transitions between the states (3-5). As a consequence of the barriers, conformations in distinct states exchange with one another much more slowly than conformations within the same state. In other words, each distinct conformational state is a collection of rapidly interconverting structures that interconvert slowly with structures in other conformational states.

For example, folding of small proteins can often be described by a two-state folding model (5). The discrete unfolded and folded states are functionally distinct collections of many structures. The extremely large ensemble of unfolded structures rapidly interconvert with one another. Likewise, the structures in the folded ensemble interconvert rapidly with one another.
Figure 1-1: Proteins and nucleic acids are state machines

Proteins and nucleic acids have conformational ensembles that can be divided into states such as C1 and C2. Each state is a collection of structures separated by an energy barrier. Often states such as C1 and C2 are functionally distinct.
Figure 1-2: Coupled binding and conformational change

The functions of proteins and nucleic acids can be modulated by ligand concentration. Ligand binding can shift the conformational equilibrium between functionally distinct states, altering the overall level of function.
Interconversion between structures in the two states, folding and unfolding, occur on a much slower timescale than interconversion between structures within either of the two states.

This view of states can be used to explain proteins and nucleic acids as state machines. Proteins and nucleic acids are molecular machines that can transition between functionally distinct states in a switch-like manner. Shifting the populations of the different states modulates the functions of these molecular machines (2) (Figure 1-2). Populations of the different states, and therefore function, can be modulated through a number of different mechanisms. Covalent modifications such as phosphorylation/dephosphorylation are often used as regulatory mechanisms. Similarly, making or breaking covalent bonds in the macromolecule backbone (e.g., cleavage in a zymogen to form an active enzyme) or sidechain (e.g., disulfide bonds) are also exploited (6). Non-covalent interactions associated with binding are particularly common and important to the work described here (1, 2, 6). Practically speaking, binding can be treated as a temporary modification that drives conformational change. As we show later, the ability of ligand to shift the conformational equilibrium from one state to another state is important for generating thermodynamic and kinetic cooperativity.

**1.2.2 Implications for Drug Design**

In coupled binding and conformational change, there are two general/limiting mechanisms by which the ligand and macromolecule can recognize one another. In the induced fit mechanism, binding of the ligand precedes conformational change. In the conformational selection mechanism, the macromolecule undergoes a conformational change before binding ligand. The mechanism of molecular recognition—induced fit or conformational selection—is of particular interest in designing drugs that bind selectively to conformational states of proteins and RNAs (1, 7-10). One central question is: do drugs bind to conformations that already exist in their absence or do they cause their targets to adopt new conformations? If the conformations to which
drugs bind exist in the absence of the drugs, can they be structurally characterized? Presumably, conformations of interest could be characterized to allow for intelligent design of drugs. Practically speaking, this may be different because conformations of interest may be poorly populated and therefore difficult to study. Another important question is: on what timescale does the change to drug-bound conformation take place? The timescale of drug activity can be limited by the time required for a conformational change to a high-affinity conformational state favored by ligand binding. Our ability to circumvent the timescale constraints could depend on whether recognition occurs through induced fit or conformational selection. Assuming that binding is near-diffusion limited, the timescale for conformational selection is limited by the conformational change intrinsic to the macromolecule target. When binding to both conformational states is diffusion limited and the induced fit mechanism predominates, this conformational change is accelerated by initial drug binding, which in turn reduces the time required for drug action.

1.3 Proteins and Nucleic Acids Are Regulated by Coupled Binding and Conformational Change

The goal of this work is to study the mechanistic details of binding-coupled conformational change in an effort to better understand how evolved systems function and how synthetic systems can be improved. Below are examples that illustrate the importance of binding in regulation of biological functions of protein, DNA, and RNA. These examples are not intended to be exhaustive, but rather demonstrative of the range of interactions in which binding and conformational change are coupled.

1.3.1 Proteins

Binding-coupled conformational change plays a number of important roles in proteins. One of the most studied roles is allosteric control of proteins. Allostery, often referred to as “action at a distance”, occurs when a covalent or non-covalent modification at one site on a
macromolecule affects the properties of another site on the macromolecule. Allostery has been well studied in the oxygen-binding protein hemoglobin ((11) for review). Hemoglobin is a tetramer whose allosteric regulation manifests as cooperative binding for several ligands. Oxygen binds preferentially to hemoglobin’s relaxed state, driving conformational change in hemoglobin subunits from the tense state to the relaxed state. Communication between the four subunits in the tetramer leads to cooperative binding of oxygen within the tetramer: binding of oxygen to one subunit promotes binding of oxygen to the other subunits. Appropriate unloading of oxygen in the necessary tissues is facilitated by negative cooperativity between binding of oxygen and two other hemoglobin ligands: 2,3-bisphosphoglycerate (2,3-BPG) and protons (H⁺). Both 2,3-BPG and H⁺ bind preferentially to the tense state, which binds weakly to oxygen. Binding of 2,3-BPG and H⁺ therefore promote release of oxygen from hemoglobin.

Coupled binding and conformational change are also important in catalysis in a number of enzymes. DNA polymerases rely on ligand-induced conformational changes to discriminate between the correct and incorrect base to incorporate into growing DNA polymers(12-14). DNA-bound DNA polymerase in the open conformational state binds to incoming free nucleotide triphosphate. Binding of an incoming base complimentary to the current base promotes conversion of the DNA polymerase from the inactive open state to the active closed state, enhancing the fidelity of DNA polymerase action. Conformational changes coupled to binding of K⁺ have been implicated in the selectivity and high conduction speeds of the KcsA channel(15). Similarly, subtle conformational changes associated with a conformational selection mechanism are required for enzyme substrate recognition and catalysis in trypsin-like proteases such as thrombin and prothrombin-2(16).

Numerous proteins and protein segments are unfolded in the absence of the proper binding partner(17). Folding of these proteins and protein segments are driven by binding of
either surrogate or conjugate binding partners(18). The protein subunits of the *E. coli* and *B. subtilis* RNase P holoenzymes are unfolded in the absence of anionic ligands(19, 20). Folding can be induced by either the cognate RNA or practically any other anion at sufficient concentrations. The actin-sequestering protein thymosin B4 is unstructured when free in solution and folds upon binding to actin(21). DNA-binding basic regions of basic leucine zipper (bZIP) proteins and basic helix-loop-helix (bHLH) proteins are unstructured in the absence of DNA(22). These regions fold into α-helices that recognize specific half-sites in DNA major grooves. Researchers have postulated several functional advantages to being unstructured in the absence of a binding partner. These include fast turnover, relatively low binding affinity, increased association rates, and the ability to bind multiple partners(23). It is also possible that the unfolded nature of many intrinsically unfolded proteins is simply a non-deleterious and irrelevant consequence of the evolved sequence.

### 1.3.2 RNA

RNA molecules play important roles in biology, including catalysis, cell signaling, and transcriptional regulation. Not surprisingly, RNA molecules display a range of functional uses for coupled binding and conformational change. Ribozymes, riboswitches, and ribosomal RNAs have all been shown to undergo conformational changes linked to binding.

Riboswitches are metabolite-binding segments of mRNA that are involved in regulation of transcription and translation(24, 25). They function as typical state machines with at least two functionally distinct conformational states that drive a simple decision-making process. Riboswitches have two domains: an aptamer domain that binds ligand and an expression platform. In each riboswitch, there is an element called the switching sequence that can be incorporated in either the aptamer domain or the expression platform, depending on whether or not ligand is bound. In transcriptional regulation, ligand binding is coupled to an arrangement of
the switching sequence that either favors formation of a transcriptional terminator sequence to repress transcription or an anti-terminator sequence to promote transcription. In the regulation of translation, ligand binding is coupled to a conformational change between states in which the Shine-Dalgarno sequence is either sequestered or exposed for ribosomal binding and subsequent translation. Riboswitches have been shown to bind ligands with a wide range of structural characteristics. These ligands include ions such as fluoride (Fl) and magnesium (Mg$^{2+}$), amino acids, nucleotides, thiamine pyrophosphate, S-adenosyl-methionine (SAM), glucosamine-6-phosphate (GlcN6P), flavin mononucleotide (FMN), tetracycline, and cyclic dinucleotides. In addition to regulating numerous metabolic pathways in bacteria and higher organisms, riboswitch aptamers are quickly becoming a powerful platform for designing useful biosensors.

Ribozymes are RNA molecules that perform single or multiple turnover chemistry, either in *trans* or in *cis*. Naturally occurring ribozymes such as the hammerhead ribozyme, hairpin ribozyme, and tetrahymena group I ribozyme have been shown to undergo conformational changes coupled to binding of Mg$^{2+}$(27-29). The VS ribozyme was shown to be pH-dependent switch that results from a conformational change coupled to proton binding(30). Synthetic ribozymes created by fusing aptamers to hammerhead ribozymes display coupled binding and conformational change similar to riboswitches(31-35). Since the cleavage chemistry of the hammerhead ribozyme is dependent on its conformational state, the cleavage is coupled to ligand binding. Researchers have created fusions whose functions can be modulated by FMN, theophyllin, cGMP, cAMP, and ATP. Depending on the design of the construct, binding may either increase or decrease the observed rates of cleavage.

Ribosomal RNAs also undergo conformational changes coupled to binding of proteins and Mg$^{2+}$. The 16 S rRNA of *Bacillus stearothermophilus* binds both Mg$^{2+}$ and the ribosomal protein S15. Both Mg$^{2+}$ and S15 induce similar conformational changes in the three-helix junction
of the 16 S rRNA(36). This conformational change is believed to be necessary for assembly of the 30 S ribosomal subunit.

### 1.3.3 DNA

DNA bending, necessary for regulating cellular processes and packing DNA, is facilitated by binding of DNA bending proteins. These proteins may distort the DNA by as little as 10 degrees and as much as 720 degrees. In doing so, DNA bending proteins can condense DNA into inactive structures or bring distant sequence elements together for biologically important interactions. One protein that serves both purposes is the *E. coli* host integration factor (IHF)(37). Another protein HMG1, is a chromosomal protein that enhances affinity of other proteins for DNA by bending the DNA(38).

### 1.4 Thermodynamic Description of Coupled Binding and Conformational Change

#### 1.4.1 Thermodynamic Description: Two-State One-Site Model

The two-state one-site system depicted in Scheme 1-1 is a useful starting point for dissecting the principles of coupled binding and conformational change. As depicted in this scheme, a macromolecule M consists of two conformational states (C1 and C2) in equilibrium with one another. The macromolecule has one binding site such that each state could potentially bind to one molecule of ligand. C1 consists of C1 and C1L and C2 consists of C2 and C2L.

The relative populations of the unliganded C1 and C2 are described by the conformational equilibrium constant $K_{12}$. Likewise, the relative populations of the liganded C1L and C2L are described by the conformational equilibrium constant $K_{12L}$. It is worth noting that even though this scheme includes ligand binding, the conformational equilibria—between C1 and C2 and C1L and C2L—are each unimolecular. As mentioned before, ligand binding can be thought of as a temporary mutation; the liganded M behaves as a variant of the unliganded M.
The population ratios $C_1:C_2$ and $C_{1L}:C_{2L}$ depend only on the $K_{12}$ and $K_{12L}$, respectively. The absolute populations of $C_1$, $C_2$, $C_{1L}$, and $C_{2L}$ depend on the concentrations of the M and L, as well as the conformational equilibrium constants and binding constants. The affinities of $C_1$ and $C_2$ for L are described by the association constants $K_{A1}$ and $K_{A2}$. These association constants determine the ratios of $C_1:C_{1L}$ and $C_2:C_{2L}$ at any given ligand concentration. In Chapter 3 we define a useful quantity $k$, which is the ratio of $K_{A2} / K_{A1}$. $K_{A1}$ and $K_{A2}$ are the intrinsic affinities of $C_1$ and $C_2$ and should not be confused with the apparent affinity $K_{A,app}$ which describes the overall apparent affinity of M for L. The relationship between $K_{A,app}$, the conformational equilibrium constant, and the two intrinsic affinities is given by Equation 1-1. Inspection of Equation 1-1 reveals that the effect of $K_{12}$ depends on $k$. If $k = 1$, $K_{12}$ does not affect the $K_{A,app}$.

For $k > 1$, $K_{A,app}$ increases with $K_{12}$ and for $k < 1$, $K_{A,app}$ decreases as $K_{12}$ increases. All of these results are intuitive; M’s overall affinity for L is greatest when the conformational equilibrium favors the higher affinity conformational state.

$$K_{A,app} = (K_{A1} + K_{12} K_{A2})/(1 + K_{12})$$

(Equation 1-1)

$$K_{A1} \times K_{12L} = K_{12} \times K_{A2}$$

(Equation 1-2)

$$K_{12L} = K_{12} \times K_{A2} / K_{A1} = K_{12} \times k$$

(Equation 1-3)

Because Scheme 1-1 is a thermodynamic cycle, the conformational equilibrium constants and the binding constants are related by Equation 1-2. This can be rearranged to obtain Equation 1-3. Equation 1-3 tells us that the minimum requirement for coupling binding to conformational change is that $C_1$ and $C_2$ have different affinities for L. If $C_1$ and $C_2$ have the same affinity for ligand, $k = 1$ and $K_{12L} = K_{12}$. In such a scenario the relative populations of $C_1:C_2$ and $C_{1L}:C_{2L}$ are the same and ligand binding does not drive a conformational change. If $k > 1$, $K_{12L} > K_{12}$ and ligand binding drives the conformational equilibrium to $C_2$ by preferentially populating $C_{2L}$ over $C_{1L}$. Similarly, if $k < 1$, $K_{12L} < K_{12}$ and ligand drives the conformational equilibrium towards $C_1$.
Scheme 1-1: Thermodynamic Model of Two-State One-Site Coupled Conformational Change and Binding
by preferentially populating C1L over C2L. The larger the disparity in $K_{A1}$ and $K_{A2}$, the larger the effect binding has on shifting the populations of C1 and C2. For this reason, $\kappa$ can be viewed as a measure of how strongly ligand binding is coupled to conformational change.

When C1 represents the predominant conformation in the absence of ligand and binding is coupled to conformational change, $K_{12}$ is much less than 1 and $\kappa$ is much greater than 1. As a consequence, at any ligand concentration C1 is much more populated than the C2 and C2L is much more populated than C1L. The most populated species over the range of ligand concentrations that shift the conformational state from C1 to C2 are C1 and C2L. The intermediates C2 and C1L are generally poorly populated regardless of ligand concentration. This adds to the difficulty of studying these particular species.

1.4.2 Allostery and Cooperativity

In allostery, binding or mutation at one site on a molecule has an effect (such as altered binding affinity) at a distant site. The allosteric effect can be explained by the ability of the mutation or ligand to shift the conformational ensemble from one state to another state(2, 3).

Let us return to the example in Scheme 1-1 and assume that $K_{12} << 1$ and that $K_{A2} >> K_{A1}$. In this scenario, C1 is the dominant conformational state in the absence of ligand and C2 is the dominant state in the presence of ligand. Now consider a mutation to M at a site distant from the binding site that increases $K_{12}$ and shifts the conformational equilibrium towards C2 in the absence of ligand. The increase in $K_{12}$ increases the overall affinity ($K_{A,app}$) of M for L by shifting the conformational equilibrium towards the state with the higher affinity for L. If a mutation decreased $K_{12}$, it would decrease $K_{A,app}$ by shifting the equilibrium towards the state with the lower affinity for L. This change in affinity can be calculated using the equation for $K_{A,app}$, in which $K_{12}$ appears in both the numerator and denominator. Its effect on the denominator is much larger because it is multiplied by $K_{A2}$, which is typically a large number. The effect of such a mutation
comes from a change in the equilibrium constant, which redistributes the conformational ensemble. Note that there is no need to invoke any pathway or network of atoms through which the mutated site and the binding site communicate. The two sites communicate through a change in conformational state. Mutations such as the one discussed offer the advantage of not altering the binding site and have been exploited in nature and by researchers to modulate affinity and are an example of allostery due to sequence changes rather than the binding of a second ligand.(39).

A second binding site can also give rise to cooperative binding, and can be thought of as an intermolecular “mutation.” Let us assume again that $K_{12} \ll 1$ and that $K_{A2} \gg K_{A1}$ for ligand 1 (L1). If ligand 2 (L2) binds with higher affinity to C2 than to C1, then binding of L1 enhances the overall affinity of M for L2, because L1 shifts the conformational equilibrium towards the state that binds most tightly to L2. This leads to positive cooperativity. If L2 binds with higher affinity to C1 than to C2, then L1 decreases the overall affinity of M for L2 by shifting the conformational equilibrium towards the state that binds most weakly to L2. This leads to negative cooperativity. If L2 binds with the same affinities to C1 and C2, then binding of L1 does nothing to alter the affinity of M for L2 and there is no cooperativity. This framework of understanding also captures the reciprocity of cooperativity. If binding of L1 enhances the affinity of L2, then binding of L2 also enhances the affinity of L1 because L1 and L2 bind preferentially to the same state. Likewise, if binding of L1 decreases affinity for L2, then binding of L2 must also decrease the affinity for L1 because L1 and L2 bind preferentially to different states. Again, note that the sites communicate not through a pathway or network of atoms, but through a shift in conformational state.
1.5 Kinetic Description of Coupled Binding and Conformational Change

1.5.1 Kinetic Description: Two-State One-Site Model

Scheme 1-2 depicts the same two-state one-site model as Scheme 1-1, but with the kinetic details of the four elementary steps. While Scheme 1-1 is described by only 4 equilibrium constants, Scheme 1-2 is described 8 rate constants. The consequence of this is that there are an infinite number of kinetic scenarios that can give rise to the same thermodynamic scenario. Lengths of the half-headed arrows in Scheme 1-2 are meant to give a qualitative comparison of the associated rate constants.
Scheme 1-2: Kinetic Model of Two-State One-Site Coupled Conformational Change and Binding
The conformational rate constants give rise to the conformational equilibrium constants. For typical coupled binding and conformational change scenarios $k_{12} < k_{21}$ and $k_{12L} > k_{21L}$, so that $C_1$ and $C_2L$ are more populated than $C_2$ and $C_1L$. Many combinations of on and off rates can give rise to different affinities in $C_1$ and $C_2$. When binding to both $C_1$ and $C_2$ is diffusion limited, $k_{off,2} < k_{off,1}$ and dissociation of ligand from $CL_2$ is slower than dissociation of ligand from $CL_1$. The rate constants depicted in Scheme 1-2 are not concentration-dependent, but rates of the bimolecular association reactions are concentration-dependent. As a consequence, the overall rate of ligand-coupled conformational change is concentration-dependent.

From Scheme 1-2, it is apparent that for ligand binding to be coupled to conformational change, ligand must alter the conformational kinetics. A simple and worthwhile question to ask is what effect does ligand have on the conformational rate constants? The ligand-induced change in conformational kinetics can be quantified with the parameter $\delta$, which is calculated using Equation 1-4. The effect of binding on the conformational equilibrium constant depends on $\kappa$, and the change in rate constants must give rise to this change in equilibrium constant. As such, it makes sense to scale the change in rate constants to $\kappa$. A $\delta$ of 1 indicates that ligand increased $k_{12}$ but did not change $k_{21}$. A $\delta$ of 0 indicates that ligand decreased $k_{21}$, but did not change $k_{12}$. And a $\delta$ of 0.5 indicates that ligand increased $k_{12}$ and decreased $k_{21}$ by the same factor. For an exhaustive list of how to interpret $\delta$, see Table 1-1. The higher $\delta$, the better the ligand is at accelerating the conformational change from $C_1$ to $C_2$. As we will discuss later, this is an important determinant of the mechanism of molecular recognition and has implications for design of drugs and protein or RNA-based biosensors.

$$\delta = \log(k_{12L}/k_{12}) / \log(\kappa)$$  \hspace{1cm} (Equation 1-4)

Calculation of $\delta$ is similar to the calculation of $\phi(40)$, which quantifies the effect of amino acid substitutions on protein folding kinetics. $\phi$ analysis is used to probe transition states in
Table 1-1: Interpreting $\delta$ value: effect of ligand on conformational rate constants

<table>
<thead>
<tr>
<th>$\delta$ value</th>
<th>Effect of Ligand on $k_{12}$</th>
<th>Effect of Ligand on $k_{21}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta &lt; 0$</td>
<td>-</td>
<td>-----</td>
</tr>
<tr>
<td>$\delta = 0$</td>
<td>No Change</td>
<td>----</td>
</tr>
<tr>
<td>$0 &lt; \delta &lt; 0.5$</td>
<td>+</td>
<td>---</td>
</tr>
<tr>
<td>$\delta = 0.5$</td>
<td>++</td>
<td>--</td>
</tr>
<tr>
<td>$0.5 &lt; \delta &lt; 1$</td>
<td>+++</td>
<td>-</td>
</tr>
<tr>
<td>$\delta = 1$</td>
<td>+++</td>
<td>No Change</td>
</tr>
<tr>
<td>$\delta &gt; 1$</td>
<td>+++</td>
<td>+</td>
</tr>
</tbody>
</table>

- Indicates ligand binding decrease the rate constant.
+ Indicates ligand binding increases the rate constant.
The number of + or – indicates the relative magnitude of the change.
protein folding. The interpretations of $\delta$ given so far do not require that the conformational change from C1 to C2 use the same transition state with and without ligand. If the transition state is the same with and without ligand bound, more conclusions can be drawn from $\delta$. A $\delta$ of 0 indicates that the transition state has binding affinity equal to that of C1 and that the structure in the vicinity of the binding site is the same as that of C1, while a $\delta$ of 1 indicates that the transition state has a binding affinity and binding site structure the same as C2.

1.5.2 The MWC and KNF Models

For more than 50 years, discussion of the molecular recognition in coupled binding and conformational change has been dominated by two competing models. The Koshland-Nemethy-Filmer (KNF)(41) and Monod-Wyman-Changeux (MWC)(42) models were first proposed in the 1960s to describe the linkage between binding and conformational change. In the KNF model, a macromolecule exists in a particular conformational state in the absence of ligand. This state binds ligand, which then induces a conformational change in the macromolecule. The MWC model says that a macromolecule exists in a number of discrete conformational states, even in the absence of ligand. Any of these conformational states can bind ligand, which eventually shifts the population, selecting the conformational state that favors ligand binding.

There are two important differences between the KNF and MWC models. Perhaps the most important difference is whether or not the macromolecule of interest exists in one or more conformational states in the absence of ligand. The answer to this question has important implications for the functions of biological macromolecules and the strategies used in drug design, both of which will be discussed later. Another important difference between the KNF and MWC models is the order in which binding and conformational change occur. The “induced fit” proposed in the KNF model is inherently a kinetic description of the process, since binding occurs first to a single conformational state and then ligand induces a conformational change in the
macromolecule. The “conformational selection” proposed by the MWC model is, however, not inherently kinetic. At least two conformational states exist in the absence of ligand, and each can potentially bind ligand.

1.5.3 Molecular Recognition: Conformational Selection and Induced Fit

Induced fit and conformational selection as described by the KNF and MWC models are the predecessors of their modern kinetic namesakes. The mostly widely used modern definitions of induced fit and conformational selection refer specifically to the order of events in the coupled binding and conformational change process. According to the modern definition of induced fit, binding occurs first and is followed by a conformational change in the macromolecule. In modern conformational selection the macromolecule first undergoes a conformational change and then binds ligand. Throughout the remainder of this work we use the definitions of induced fit and conformational selection presented here. There are several important differences between these modern definitions and two models that preceded them. Both induced fit and conformational selection are kinetic in nature, describing an order of events. While the KNF model is kinetic in nature, the MWC model is not. Additionally, both induced fit and conformational selection can exist in the framework of the MWC model. The pre-existing equilibrium between conformational states in the MWC model is a prerequisite for modern conformational selection. This pre-existing equilibrium does not, however, preclude the possibility that coupled binding and conformational change occurs through induced fit. Similarly, a pre-existing equilibrium between conformational states does not prove that it occurs through conformational selection.

As stated earlier, the modern definitions of induced fit and conformational selection are kinetic, and identifying either as the mechanism of molecular recognition in coupled binding and conformational change requires appropriate kinetic evidence.
In some studies, the pre-existence of two conformational states in the absence of ligand has been used to infer that molecular recognition occurs through conformational selection. Typically, conformational selection and induced fit are distinguished from one another by performing kinetic experiments and obtaining the time constants ($k_{\text{obs}}$) for binding or conformational change. These $k_{\text{obs}}$ can then be analyzed to extract parameters for the pure conformational selection or induced fit schemes (43). Alternatively, the mechanism can be assigned based on qualitative assessment of the dependence of $k_{\text{obs}}$ on ligand concentration. Assignment of the mechanism in this manner is usually limited because there are many assumptions or limitations on the relative magnitudes of binding and conformational rate constants. Additionally, most simple analyses assume that experiments are performed with ligand in great excess of the macromolecule that undergoes conformational change.

1.5.4 A Flux Description of Conformational Selection and Induced Fit

Historically, studies of molecular recognition have been designed to determine whether it occurs through conformational selection or induced fit. This view of molecular recognition presents a false dichotomy, since conformational selection and induced fit are not mutually exclusive. Conformational selection and induced fit pathways (Scheme 1-3) have the same reactants and product, and can be thought of as two pathways competing in a kinetic race. The outcome of the kinetic race—the fraction of recognition that occurs through each pathway—depends on the rates of reaction through each pathway.

The rates of reactions through multistep pathways are readily calculated using the flux approach (44). Serial flux through successive steps in each pathway is calculated and compared. The flux approach is applied to the molecular recognition problem by calculating the flux through each pathway as a fraction of the total flux. The rates of the binding reactions in the conformational selection and induced fit pathways depend on ligand and macromolecule
Scheme 1-3: Conformational Selection and Induced Fit
concentrations. As a result, the flux through the two pathways is concentration dependent (45, 46). Increased ligand concentration relative to macromolecule concentration leads to increased flux through the induced fit pathway. Increased macromolecule concentration relative to ligand concentration leads to increased flux through the conformational selection pathway. As we show in Chapter 3, the fractional flux through each pathway also depends strongly on $\kappa$ and $\delta$. High $\kappa$ and low $\delta$ favor flux through conformational selection. Moderate to high $\kappa$ and high $\delta$ favor flux through induced fit.

Conformational selection and induced fit adequately describe only the simplest system: two conformational states and one binding site. Many macromolecules sample more than two conformational states and bind two or more ligands. The flux description extends well to more complex systems and adequately describes the mechanism of molecular recognition through competing pathways.

**1.5.5 Kinetic Versus Thermodynamic Control**

Considerable effort has been made to understand and study the mechanisms of molecular recognition, but why does it matter? In order for the mechanism of molecular recognition to be biologically important, there must be some difference in the biological utilities of conformational selection and induced fit. One consequential difference between conformational selection and induced fit lies in the timescales on which they occur. Because conformational change is often rate limiting in molecular recognition and ligand binding in induced fit can accelerate the conformational change, induced fit is typically faster than conformational selection. These timescales are of little consequence in biological systems under thermodynamic control, but are central to the function in systems under kinetic control.

For a biological system under thermodynamic control, the equilibrium populations of the different species determine the functional state of the system. Only the thermodynamic
characteristics of the system—affinities and conformational equilibrium constants—and concentrations of ligands and macromolecules matter. Take for example a two-state system in which one state is inactive and one state is active. The populations of these states—and therefore the overall function of the system—can be controlled by the concentration of an effector ligand whose binding is coupled to conformational change.

In a system under kinetic control, the timescale of two competing reactions is the deciding factor in the function of the system. Consider a system in which a macromolecule $M$ can bind two different molecules that determine the fate of $M$ (Scheme 1-4). Binding of $M$ to $A$ commits $M$ to a fate such as degradation, and binding of $M$ to $B$ commits $M$ to a fate such as transcription induction of a gene product (Scheme 1-4A). Assume that $C_1$ binds tightly to $A$ and $C_2$ binds tightly to $B$. The fate of $M$ is tied to its conformational state. An effector molecule $L$ can shift the conformational state to $C_2$ by populating $C_2L$ (Scheme 1-4B). Populating $C_2L$ through conformational selection is too slow to kinetically compete with binding of $C_1$ to $A$. With increasing concentrations of $L$, $C_2L$ is populated through the faster induced fit pathway and can bind $B$ on a timescale that can kinetically compete with binding of $A$. The system shown in Scheme 1-4 is under kinetic control and the functional state of the system depends on the concentration of the effector ligand $L$. Figure 1-3 shows the fate of $M$ as a function of the effector molecule concentration. For a $\delta$ of 0.68 (Figure 1-3 blue curve), ligand accelerates the conformational change from $C_1$ to $C_2$ and increasing concentrations of $L$ cause $M$’s fate to switch from binding to $A$ to binding to $B$. If $\delta$ is 0 (Figure 1-3 red curve), ligand binding does not accelerate the conformational change from $C_1$ to $C_2$ and increasing concentrations of $L$ do little to alter the fate of $M$. Although the thermodynamic parameters that describe the conformational changes and binding reactions in Figure 1-3 are identical, the outcomes are very different because the kinetic parameters differ and the system is under kinetic control.
Scheme 1-4: Coupled Binding and Conformational Change Under Kinetic Control
Figure 1-3: Kinetic control

The fate of M (Schemes 4A and 4B) depends on the concentration of ligand and the ability of ligand to accelerate the conformational change. The fate of M is different if ligand accelerates conformational change (blue curve) or does not accelerate conformational change (red curve).
1.5.6 Kinetic Cooperativity

The ability of ligand to catalyze the conformational change from the low-affinity to high-affinity state can lead to kinetic cooperativity in binding and catalysis. As with the cooperativity discussed earlier, this kinetic cooperativity is rooted in the ligand-linked redistribution of the ensemble. The kinetic cooperativity comes from the fact there are at least two pathways for conformational change: one slow and one fast. At lower ligand concentrations, the slower conformational change occurs through conformational selection before catalysis or binding of another ligand occurs. At higher ligand concentrations, the conformational change happens through the faster induced fit pathway and is followed by catalysis or binding of another ligand.

For kinetic cooperativity in binding, consider the following scenario: two ligands (L1 and L2) bind with low affinity to C1 and high affinity to state C2. The conformational change from state C1 to state C2 is slow in the absence of ligand. Conformational change driven by L1 has a high $\delta$, and is much faster than the unliganded conformational change. Conformational change driven by L2 has a low $\delta$, and is slow like the unliganded conformational change. In the absence of L1, saturation of the L2 binding site is limited by the slow conformational change, but in the presence of L1, the conformational change is fast and saturation of the L2 binding site is accelerated. This kinetic cooperativity in binding may play an important role in assembly of macromolecular complexes.

A monomeric enzyme in which the ligand that drives conformational change is also the substrate (Scheme 1-5) can display kinetic cooperativity when the ligand-induced conformational change is significantly faster than the unliganded conformational change. Notice that there is only one binding site and that the enzyme is monomeric. A system of this simplicity can produce the sigmoidal kinetics (Figure 1-4) characteristic of kinetic cooperativity. Such kinetic cooperativity
for a monomeric enzyme has been observed in human glucokinase, and which uses the sigmoidal response to glucose concentration to maintain proper glucose homeostasis (47).

Scheme 1-5: Coupled Binding and Conformational Change in an Enzyme
The rate of reaction for non-cooperative (blue) enzymes has a hyperbolic dependence on ligand concentration, while the rate of reaction for cooperative (red) enzymes has a sigmoidal dependence.
1.6 Structural, Thermodynamic, and Kinetic Studies of Coupled Binding and Conformational Change

The coupling of binding and conformational change can be viewed from structural, thermodynamic, and kinetic perspectives. Each perspective complements the others and combining all three perspectives gives the most complete understanding of the coupling and the role it may play in regulation. Although this work is most concerned with the kinetic and thermodynamic aspects of coupled binding and conformational change, it is important to understand some of the insights granted by structural studies.

1.6.1 Structural Studies

Structural biology tools such as X-ray crystallography, nuclear magnetic resonance, and SAXS provide rich information about the molecular details of binding-coupled conformational change. The information provided by crystallographic and NMR structures is often atomic level resolution. Because structural techniques give almost direct observations of conformational changes, they often provide the first observations of conformational change in a molecule. Alternatively, many structures confirm conformational changes suggested on the basis of other techniques.

A powerful advantage of structural techniques is that they allow observation of conformational changes at many distance and time scales and provide some clues as to how interaction with ligand promotes these conformational changes. The near-atomic resolution of some structural techniques allows researchers to determine the scale of conformational change. Common changes observed in structures are changes in sidechain rotamers, interhelical angles, interdomain angles, and local or global folding.

Structures can also provide useful information when designing thermodynamic and kinetic models for binding-coupled conformational change. In some cases, multiple
conformations of a molecule may be visualized. Important features such as oligomeric state and ligand binding stoichiometry may be revealed by structures. In some cases, φ or δ analysis can be used to infer the structures of transition state conformations. This information may be of use in initial design of models or in interpreting results of thermodynamic and kinetic experiments. Most importantly, structures can guide researchers in making useful perturbations such as residue substitutions, site-labels, and other chemical modifications.

1.6.2 Thermodynamic Studies

Equilibrium studies of coupled binding and conformational are typically carried out by varying ligand or co-solvent concentration or by varying temperature. These experiments can be used to obtain estimates of binding affinities or conformational equilibrium constants. For example, in the case of protein folding/unfolding the conformational equilibrium constants can often be determined by titrations with co-solvents such as urea, guanidium, or trimethylamine-N-oxide. Estimates of conformational equilibrium constants for conformational changes other than protein folding are difficult to obtain using equilibrium experiments. Most equilibrium experiments are unable to fully disentangle binding and conformational change, and are therefore unable to the intrinsic binding affinities of conformational states. One exception is when a conformational state can be isolated or trapped in the absence of ligand and then titrated with ligand. As we demonstrate in Chapter 3, conformational equilibrium constants and intrinsic affinities of conformational states can often be estimated from kinetic experiments. Often, equilibrium experiments can be used to support results from kinetic experiments or can be globally analyzed with kinetic data.
1.6.3 Kinetic Studies

Kinetic studies of coupled binding and conformational change typically take advantage of relaxation kinetics, in which the system is subjected to a change in conditions and monitored as it relaxes to equilibrium. Two types of relaxation kinetic experiments are generally used: transient kinetics and dynamic NMR. In transient kinetics the chemical environment of the system is perturbed by changing a variable, such as temperature or ligand concentration. The system is then monitored as it returns to equilibrium. During transient kinetic experiments, the populations of various chemical species change with time. In dynamic NMR experiments, the system is perturbed so that there is some excess excited state net magnetization. The decay of the magnetization is monitored over time as the system relaxes to equilibrium. Unlike in transient kinetic experiments, the populations of chemical species do not change in a dynamic NMR experiment.

Although there are many variables that can be manipulated in a relaxation kinetics experiment, the most important variable for the study of coupled binding and conformational change is the concentration of the ligand, followed closely by the concentration of the macromolecule to which it binds. Varying the ligand from subsaturating to saturating concentrations is important because it can potentially shift the mechanism of molecular recognition from conformational selection-like pathways to induced-fit like pathways. This provides kinetic information about elementary steps in the various pathways. Ideally, the ligand concentration can be increased until it nears the intrinsic $K_D$ of the lowest-affinity state, thereby increasing flux through the induced-fit like pathways. If experimental conditions—primarily concentrations and timescale—are properly chosen and the data is properly analyzed, kinetic and thermodynamic parameters can be estimated for all reactions through which there is significant flux.
1.6.4 Data Analysis

Kinetic data are typically analyzed in one of two ways. The data are either fit to a sum of exponentials to obtain time constants and amplitudes that are subsequently fit to a model, or the data are fit directly to a model. The first method is often simpler and faster than the second, but has significant drawbacks. When phases have similar time constants it can be difficult or impossible to resolve them. Additionally, amplitudes from fits to sums of exponentials are often not included in the analysis, despite the fact that they contain significant information. When time constants are properly fit but amplitudes cannot be fit to the model, it may mean that the model being used is incorrect. Another drawback is that the models used to fit time constants often have numerous assumptions that constrain the experimental design. These often include the assumptions that ligand is in excess of the macromolecule of interest and binding is much faster than conformational change. When phases are well-resolved and assumptions are satisfied, analysis of the time constants with or without amplitudes may be used to determine the rate constants for steps in coupled binding and conformational change. The time constants may also be analyzed to determine the mechanism of molecular recognition, provided the recognition occurs exclusively through either induced fit or conformational selection. This approach has been used to assign the mechanism of molecular recognition for many systems.

The alternative method, which we use in Chapters 2-4, is to fit the data directly to the mechanistic model. To achieve this, a model is first written using a system of differential equations that describes the kinetics of the experimental system. The model and data are bridged by an equation that relates the kinetics to the experimental observable. Data are fit to the model to estimate kinetic and thermodynamic parameters, as well as parameters related to the experimental observable. This analysis is readily done with software such as Mathematica or MATLAB. There are several advantages to using this approach. Second order kinetics are easily accounted for, so
ligand concentration does not need to be in excess. Time constants and amplitudes that cannot be resolved when fitting data to sums of exponentials are accounted for in the global fit of data to a model. The drawback of this approach is that the equation that relates the model to the experimental observable may introduce a large number of parameters that may leave the model underdetermined.

Data fitting is a critical part of interpreting thermodynamic and kinetic data in studies of coupled binding and conformational change. The goal of fitting is two-fold: To find parameter values that give the best agreement between the model and data, and to estimate uncertainties in those parameter values. Without uncertainties, the parameter estimates are ultimately meaningless because there is no way to tell how reliable the estimates are and to predict the magnitude of the effects of perturbations such as residue substitutions.

Determining the best parameter values requires a model that is sufficiently complex to explain the data, but not over-parameterized. Additionally, the data collected must have sufficient information content to produce a physically reasonable set of parameter values from the fitting process. Global fitting of the data can help ensure that the model being used is appropriate and also maximize the information content of the data. Failure to fit multiple kinetic data sets globally to one model or failure to fit kinetic and thermodynamic data globally to one model may indicate that the model is incorrect.

Fitting of the kinetic and thermodynamic data for coupled binding and conformational change often requires determining many parameter values. Several things should be considered to improve fitting and interpretation of fitting results. If the model is too complex for the given data, the system may be underdetermined; there may be infinitely many solutions. A trivial problem that can lead to this is fitting for too many parameters. One should ensure that as few parameters as possible are being fit for by fixing any parameters that are well-known and by expressing
parameters in terms of other parameters whenever possible. Normalizing data sets or making sure observable parameters (eg. spectroscopic signals) are consistent from experiment-to-experiment can also reduce the number of parameters that need to be fitted. Finally, fitting algorithms may become trapped in local minima and give answers that are not optimal. Care should be taken to ensure that the solutions obtained from fitting are not local minima. Using different starting parameter values for the fit is one way to test for local minima. If no global minimum can be reached, local minima should be compared to ensure that they are not significantly different from one another.
2. Ligand Concentration Regulates the Pathways of Coupled Protein Folding and Binding

2.1 Introduction

2.1.1 Challenging the False Dichotomy of Conformational Selection versus Induced Fit

Many biological systems couple protein conformational change to ligand binding. This coupling underlies all allosteric regulation. A mechanistic description, including the order in which binding and conformational change occur, is required for understanding the molecular basis of biological regulation and for rational drug design. Discussion of conformational coupling mechanism has focused on “conformational selection” and “induced fit” mechanisms, which represent limiting extremes of the order of events(18, 48, 49). In conformational selection, ligand binds directly to the poorly populated high-affinity conformational ensemble. In induced fit, ligand binds to the highly populated low-affinity ensemble followed by a transition to the high-affinity ensemble. These ensembles are distinguishable by the highly cooperative kinetic barriers that separate them. Numerous experimental studies have sought to determine which of these mechanisms predominates in various systems. Previous experimental studies have highlighted the necessity of performing kinetic experiments to distinguish between the mechanisms(50, 51). One such study treated conformational selection and induced fit mechanisms as mutually exclusive for a given protein-ligand pair and did not consider the possibility that a change in ligand concentration might cause a change in mechanism(50). Despite observation of combined conformational selection and induced fit mechanisms in silico(52-54), kinetic experiments have failed to yield mechanisms that describe the partitioning between induced fit and conformational selection. Here we use thermodynamic, structural, and kinetic data to obtain a detailed kinetic description of coupled folding and binding in Bacillus subtilis RNase P protein using the ligand
I perform a flux-based analysis\(^\text{(44)}\) of kinetic data obtained over a wide range of ligand concentrations and show that flux is kinetically partitioned between alternate reaction pathways described by both conformational selection and induced fit. The partitioning depends on ligand concentration. I also report binding affinities to the low- and high-affinity protein conformations and show that ligand binding redistributes the conformational ensembles by increasing folding rates and decreasing unfolding rates.

### 2.1.2 Bacillus subtilis RNase P Protein

Bacterial RNase P is a ribonucleoprotein complex that cleaves the 5’-leader sequence from pre-tRNA\(^\text{(55-57)}\). The P RNA subunit is catalytically active in vitro\(^\text{(58)}\), while the protein subunit enhances substrate specificity\(^\text{(59, 60)}\). The \textit{B. subtilis} protein subunit has three conformational sub-ensembles\(^\text{(61)}\) and two high affinity ligand binding sites\(^\text{(19)}\). The unfolded state predominates when the protein is not bound to anions. Folding of RNase P protein in the absence of ligand can be induced with osmolytes such as trimethylamine N-oxide (TMAO). Binding of small anions shifts the conformational equilibrium to populate the folded state. This ligand-induced conformational change, depicted in Figure 2-1, serves as a good model for studying how binding is coupled to conformational change in proteins.
Figure 2-1: \( \text{P}^{\text{Pro}} \) coupled folding and binding scheme

\( \text{P}^{\text{Pro}} \) exhibits three-state folding with unfolded (U), partially folded intermediate (I), and folded (F) states. Folding is strongly coupled to the binding of two pyrophosphate ligands (L). Pyrophosphate may be bound at the \( \alpha \)-site (L\( \alpha \)), \( \beta \)-site (L\( \beta \)), or both sites (L\( \alpha \beta \)) in any state. Unoccupied binding sites in U and I are reflected as undetectably weak affinities in the fits of the data. Species that remain unpopulated in 0 – 1 mM PP\( \text{i} \) — as revealed by fits of the data — and their associated transitions are indicated in gray. Conformational selection pathways are shown in blue, induced fit pathways in red, and mixed pathways in green.
2.2 Results and Discussion

2.2.1 Structural and Functional Characterization of P-Pro

In the present work, we have substituted two prolines with alanines to simplify the mechanism and have shown that the substitutions do not alter P protein structure or function. The F107W/P39A/P90A variant (referred to hereafter as P-Pro) was structurally characterized by X-ray crystallography. The structure of P-Pro is nearly superimposable with the structure of wildtype P protein (RMSD = 0.53 Å) (Figure 2-2A). PPi molecules occupy each of the two previously identified binding sites (Figure 2-2B). Electron density for a third PPi was observed proximal to Arg60. This binding site, referred to as the γ site, binds PPi too weakly to contribute significantly to the coupled folding and binding mechanisms at the concentrations used in our stopped-flow experiments. In agreement with the binding data, electron density maps of the three PPi sites shows clearly defined density for PPi of sites α and β while the electron density for the γ site only shows clear density for the bound phosphate with weak density surrounding the second unbound phosphate (Figure 2-2D-F). Furthermore, superposition of the P-Pro structure with the protein subunit of the T. maritima RNase P holoenzyme structure indicates that PPi binds to P-Pro at sites that bind P RNA or the 5’ leader of precursor tRNA (Figure 2-2C). Integrity of the P-Pro and the previously studied F107W variants was also assessed by their ability to form active RNase P holoenzyme. RNase P containing either variant was able to cleave fluorescently labeled pre-tRNA^AMP (Figure 2-3). Based on the minimal structural perturbations and retention of activity, we deemed the P-Pro reasonable to use for determining the mechanism of coupled binding and folding.
Figure 2-2: Crystal structure of P-Pro

(A) Alignment of crystal structures of wildtype (green, PDB 1A6F) and P-Pro (blue) B. subtilis RNase P protein subunit. Structures are aligned on the Cα backbone atoms of residues 2-114. (B) Crystal structure of P-Pro bound to pyrophosphate (red). The binding sites are designated α, β, and γ and have been assigned to specific species in the binding/folding mechanism as described in the text. (C) The structure of P-Pro (blue) in complex with pyrophosphate (red) was inserted into the crystal structure of Thermatoga maritima RNase P (PDB 3OKB) (64) in place of the T. maritima protein subunit. P-Pro was oriented by alignment with the Cα backbone of the T. maritima protein. The T. maritima RNase P RNA is shown in magenta and RNA 5’ leader sequence is shown in cyan. (D, E, F) Averaged kick omit maps of the three pyrophosphate binding sites. The 2Fobs – Fcalc map is shown in blue and is contoured at 1σ. The Fobs – Fcalc map is shown in green and is contoured at 3σ.
Table 2-1: Crystallography statistics for P-Pro in complex with pyrophosphate

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>P 6₁</td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>83.14, 83.14, 32.24</td>
</tr>
<tr>
<td>α, β, γ (°)</td>
<td>90, 90, 120</td>
</tr>
<tr>
<td>Rₚₒᵣₚ (%)ᵃ</td>
<td>4.8 (12.6)ᵃ</td>
</tr>
<tr>
<td>&lt;θ&gt;/&lt;θ₀&gt;</td>
<td>44.3 (11.9)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>94.4 (78.6)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>8.9 (7.6)</td>
</tr>
<tr>
<td>Total reflections (#)</td>
<td>49240</td>
</tr>
<tr>
<td>Unique reflections (#)</td>
<td>5545</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>24.02 (2.30)</td>
</tr>
<tr>
<td>Rₕₒᵣₑₑ/ₐₛₑₑ (%)ᵇ</td>
<td>21.18/23.66</td>
</tr>
</tbody>
</table>

Ramachandran statistics

<table>
<thead>
<tr>
<th>Category</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Favored</td>
<td>116 (98.2%)</td>
</tr>
<tr>
<td>Allowed</td>
<td>2 (1.8%)</td>
</tr>
<tr>
<td>Outlier</td>
<td>0 (0%)</td>
</tr>
</tbody>
</table>

RMSD

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bond lengths (Å)</td>
<td>0.010</td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>1.545</td>
</tr>
</tbody>
</table>

ᵃValues in parentheses are for highest-resolution shell.
ᵇ$R_{\text{work}} = \sum \left| \frac{F_{\text{obs}} - F_{\text{calc}}}{\sum F_{\text{obs}}} \right|$ and $R_{\text{free}} = \sum \left| \frac{F_{\text{obs}} - F_{\text{calc}}}{\sum F_{\text{obs}}} \right|$, where all reflections belong to a test set of 10% data randomly selected in PHENIX.
Figure 2-3: Single and multiple turnover cleavage of fluorescent pre-tRNA by RNase P

Structural integrity of P protein variants was assessed by their ability to form enzymatically active RNase P holoenzyme as described in Materials and Methods. (A) Under single-turnover conditions, mutations in P protein (F107W or P-Pro) do not affect the pH-dependent rate for RNase P holoenzyme-catalyzed cleavage of fluorescently labeled pre-tRNA<sup>Asp</sup>. (B) RNase P holoenzyme formed by P RNA and either wildtype (WT), F107W, or P-Pro proteins cleave fluorescently labeled pre-tRNA<sup>Asp</sup> under multiple turnover conditions. Data were fit to the Michaelis-Menten equation to calculate the following steady-state kinetic parameters: wildtype RNase P holoenzyme, $k_{cat} = 0.19 \pm 0.01$ s<sup>-1</sup>, $K_M = 110 \pm 16$ nM, and $k_{cat}/K_M = 1700 \pm 200$ mM<sup>-1</sup>s<sup>-1</sup>; F107W P protein, $k_{cat} = 0.103 \pm 0.003$ s<sup>-1</sup>, $K_M = 41 \pm 5$ nM, and $k_{cat}/K_M = 2500 \pm 300$ mM<sup>-1</sup>s<sup>-1</sup>; and P-Pro P protein, $k_{cat} = 0.122 \pm 0.003$ s<sup>-1</sup>, $K_M = 47 \pm 5$ nM, and $k_{cat}/K_M = 2600 \pm 300$ mM<sup>-1</sup>s<sup>-1</sup>.
2.2.2 U, I, and F Bind PPi with Distinct Affinities

I used stopped-flow fluorescence to monitor folding of P-Pro upon mixing with TMAO or PPi (Figures 2-4 and 2-5). Rather than fit to a series of exponentials, we used a Bayesian estimation method (Markov Chain Monte Carlo followed by sequential Monte Carlo) to globally fit all kinetic transients for both TMAO- and PPi-induced folding to a model containing thermodynamic, kinetic, and spectroscopic parameters (Tables 2-2 through 2-5) for the kinetic scheme depicted in Figure 2-1. To implement the model, we used a series of ordinary differential equations (see Appendix A) to describe the time-dependent free ligand concentration and populations of twelve microscopic protein species depicted in Figure 2-1. The time-dependent populations were then used to express the total fluorescence signal as a sum of population-weighted signals. This approach allowed us to analyze data collected under non-pseudo-first order conditions, in which total ligand concentrations ranged from 0.16 to 66 times the protein concentration and spanned the apparent K_D.

The α site in the folded state, implicated in the NMR PRE experiments and our stopped-flow experiments as being the highest affinity site, contains residues identified as the unfolded state α site. In the NMR PRE and stopped-flow experiments, the next highest affinity site is the β site. The γ site has the lowest affinity (confirmed by ITC) of the sites observed by NMR PRE experiments. Surprisingly, all three conformational ensembles – unfolded (U), partially folded (I), and folded (F) – have detectable affinity for PPi, based on the fits of stopped-flow data (Table 2-5). The measured affinities of U and I for PPi are nearly impossible to obtain from equilibrium experiments because the equilibrium populations of PPi-bound U and I do not exceed 1%. However, in our kinetic experiments we detected affinity (2.6 mM) in U for a single PPi at a site mechanistically linked to the highest affinity site in F. A peptide mimic of the fifteen N-terminal residues that we believe contain the U binding site of P-Pro binds PPi with a similar affinity as U
Figure 2-4: Kinetics of TMAO-induced and pyrophosphate-induced folding of P-Pro

Folding was monitored by stopped-flow fluorescence. Data points are the average of data points from three traces. The blue surfaces are the global best-fit of the data to the coupled folding and binding model.
Figure 2-5: TMAO- and PPi-induced folding of P-Pro

Folding was monitored by stopped-flow fluorescence. Two-dimensional projections of kinetic traces for TMAO-induced (A) and pyrophosphate-induced (B) folding of P-Pro as monitored by stopped-flow fluorescence. Data points are the average of data points from three traces. The red lines are the global best-fit of the data to the coupled folding and binding model. Data points for the equilibrium plots for TMAO-induced (C) and pyrophosphate-induced (D) folding are the final time points (30 seconds) of each trace in the stopped-flow experiments. The red lines are the ideal equilibrium plots generated using the parameter estimates obtained from fitting the stopped-flow data.
Table 2-2: Dissociation constants for P-Pro and pyrophosphate

<table>
<thead>
<tr>
<th>Conformational State</th>
<th>$\alpha$ site (M)</th>
<th>$\beta$ site (M)</th>
<th>$\gamma$ site (M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>U</td>
<td>$2.6 \times 10^{-3}$</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>I</td>
<td>$1.2 \times 10^{-4}$</td>
<td>$4.6 \times 10^{-4}$</td>
<td>ND</td>
</tr>
<tr>
<td>F</td>
<td>$7.6 \times 10^{-7}$</td>
<td>$2.3 \times 10^{-6}$</td>
<td>$4.0 \times 10^{-5}$</td>
</tr>
</tbody>
</table>

*Estimate (bold) and the 95% confidence interval (parentheses) were obtained from the fit of stopped-flow data as described in Materials and Methods.

Table 2-3: Conformational parameter value estimates from fit of stopped-flow data

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate (M)</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta G_{\text{UI}}$</td>
<td>$859$ (850, 864)</td>
<td>(cal mol$^{-1}$)</td>
</tr>
<tr>
<td>$\Delta G_{\text{IF}}$</td>
<td>$372$ (366, 382)</td>
<td>(cal mol$^{-1}$)</td>
</tr>
<tr>
<td>$m_{\text{UI}}$</td>
<td>$-984$ (-960, -965)</td>
<td>(cal mol$^{-1}$ M$^{-1}$)</td>
</tr>
<tr>
<td>$m_{\text{IF}}$</td>
<td>$-6960$ (-6950, -6950)</td>
<td>(cal mol$^{-1}$ M$^{-1}$)</td>
</tr>
<tr>
<td>$\beta_{\text{UI}}$</td>
<td>$0.41$ (0.39, 0.42)</td>
<td>(unitless)</td>
</tr>
<tr>
<td>$\beta_{\text{IF}}$</td>
<td>$0.566$ (0.563, 0.570)</td>
<td>(unitless)</td>
</tr>
<tr>
<td>$\beta_{\text{UF}}$</td>
<td>$0.88$ (0.86, 0.93)</td>
<td>(unitless)</td>
</tr>
</tbody>
</table>

*Estimate (bold) and the 95% confidence interval (parentheses) were obtained from the fit of stopped-flow data as described in Materials and Methods.
Table 2-4: Rate constant parameter value estimates from fit of stopped-flow data

<table>
<thead>
<tr>
<th>Parameter</th>
<th>((s^{-1})^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(k_{UI})</td>
<td>1.99 (1.97, 2.00)</td>
</tr>
<tr>
<td>(k_{IU}^b)</td>
<td>8.5 (8.4, 8.6)</td>
</tr>
<tr>
<td>(k_{U}^\alpha\alpha)</td>
<td>2.4 (2.0, 3.3)</td>
</tr>
<tr>
<td>(k_{I}^\alpha\alpha)</td>
<td>0.5 (0.4, 0.7)</td>
</tr>
<tr>
<td>(k_{IF})</td>
<td>1.06 (1.03, 1.09)</td>
</tr>
<tr>
<td>(k_{FI}^b)</td>
<td>2.00 (1.95, 2.05)</td>
</tr>
<tr>
<td>(k_{IF}^\alpha\alpha)</td>
<td>7 (6, 8)</td>
</tr>
<tr>
<td>(k_{FI}^\alpha\beta)</td>
<td>0.09 (0.06, 0.1)</td>
</tr>
<tr>
<td>(k_{IF}^\beta\beta)</td>
<td>22 (19, 25)</td>
</tr>
<tr>
<td>(k_{FI}^\beta\alpha\beta)</td>
<td>0.21 (0.18, 0.28)</td>
</tr>
<tr>
<td>(k_{IF}^\beta\alpha\beta)</td>
<td>120 (100, 140)</td>
</tr>
<tr>
<td>(k_{FI}^\alpha\beta\beta)</td>
<td>0.0072 (0.0068, 0.0078)</td>
</tr>
<tr>
<td>(k_{IF}^\alpha\beta)</td>
<td>0.036 (0.031, 0.043)</td>
</tr>
<tr>
<td>(k_{IF}^\beta\alpha)</td>
<td>0.29 (0.25, 0.35)</td>
</tr>
<tr>
<td>(k_{IF}^\beta\alpha)</td>
<td>9 (8, 10)</td>
</tr>
<tr>
<td>(k_{IF}^\alpha\beta)</td>
<td>0.021 (0.019, 0.024)</td>
</tr>
</tbody>
</table>

^aEstimate (bold) and the 95% confidence interval (parentheses) were obtained from the fit of stopped-flow data as described in Materials and Methods.

^bEstimate (bold) and the 95% confidence interval (parentheses) were derived using estimates and confidence intervals of fitted parameters.

Table 2-5: Signal parameter value estimates from fit of stopped-flow data

<table>
<thead>
<tr>
<th>Parameter</th>
<th>((AFU)^a)</th>
<th>Parameter</th>
<th>((AFU , M^{-1}b)^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Signal_{U})</td>
<td>8.758 (8.756, 8.761)</td>
<td>(Slope_{U})</td>
<td>0.255 (0.252, 0.259)</td>
</tr>
<tr>
<td>(Signal_{I})</td>
<td>7.60 (7.59, 7.62)</td>
<td>(Slope_{I})</td>
<td>-0.498 (-0.499, -0.495)</td>
</tr>
<tr>
<td>(Signal_{F})</td>
<td>7.555 (7.553, 7.556)</td>
<td>(Slope_{F})</td>
<td>-0.397 (-0.399, -0.394)</td>
</tr>
<tr>
<td>(Signal_{UBound})</td>
<td>9.5 (9.3, 9.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Signal_{IBound})</td>
<td>8.44 (8.40, 8.48)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Signal_{FBound})</td>
<td>7.474 (7.473, 7.475)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

^aEstimate (bold) and the 95% confidence interval (parentheses) were obtained from the fit of stopped-flow data as described in Materials and Methods.

^bAFU is arbitrary fluorescence units.
Table 2-6: Dissociation constants for the P$_{\text{Pro}}$ unfolded state and pyrophosphate

<table>
<thead>
<tr>
<th>Protein Form</th>
<th>$K_0$ (mM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unfolded Protein$^a$</td>
<td>2.6 (2.3, 2.7)</td>
</tr>
<tr>
<td>N-terminal Synthetic Peptide$^b$</td>
<td>1.0 (0.8, 1.2)</td>
</tr>
<tr>
<td>Peptide Mixture$^b$</td>
<td>1.0 (0.8, 1.2)</td>
</tr>
</tbody>
</table>

$^a$Estimate (bold) and the 95% confidence interval (parentheses) were obtained from the fit of stopped-flow data as described in Materials and Methods.
(Figure 2-6A-C and Table 2-6). I also detected affinity for the α and β binding sites in both I (120 μM and 460 μM) and F (0.76 μM and 2.3 μM). Affinity for the γ binding site (40 μM) was observed by ITC titration of 200 μM P^Pro with PP_i (Figure 2-6D). However, the affinity is too low for the site to contribute significantly to the coupled folding and binding mechanism at the concentrations used in our stopped-flow experiments. The affinities of U and I are particularly insightful for understanding how PP_i binding redistributes the ensemble towards F. From a kinetic standpoint, redistribution of the ensemble may result from increased folding rate constants, decreased unfolding rate constants, or both. Binding to U and I are prerequisites for increased folding rate constants. By comparing the conformational rate constants for free and PP_i-bound P^Pro, we show that PP_i redistributes the P^Pro ensemble by changing both folding and unfolding rate constants.
Figure 2-6: Binding of pyrophosphate to P^{Pro} and mimics of the P^{Pro} N-terminus

(A) Binding of pyrophosphate to the N-terminal peptide mimic AHLKKRNRLKKNEDW was measured by ITC. (B) Binding of pyrophosphate to a mixture of P^{Pro} fragments produced by formic acid cleavage (65) of P^{Pro} was measured by ITC. Red lines are the best-fit of the data to a single site binding isotherm. (C) Formic acid cleaves to the C-terminal side of aspartate residues. Cleavage sites are indicated by HCOOH markers. Formic acid cleavage was performed as described in materials and methods. (D) Binding of pyrophosphate to P^{Pro} was measured by ITC. The data were fit to a three-state, three-site coupled folding and binding model using parameters obtained from the fit of the kinetic data.
2.2.3 PP$_i$ binding significantly alters folding and unfolding rate constants.

The kinetic parameters of the best-fit model reveal a variety of effects of PP$_i$ binding on the kinetics of P$^\text{Pro}$ conformational changes. Estimates of rate constants for the conformational transitions shown in gray in Figure 2-1 were not obtained because U lacks a detectable β site, such that these transitions do not contribute to the coupled folding and binding mechanism. I analyzed rate constants for the remaining transitions shown in black (Table 2-4), and concluded that the folding and unfolding rate constants for P$^\text{Pro}$ depend on the number of PP$_i$ bound to P$^\text{Pro}$ (zero, one, or two) and the site at which the PP$_i$ is bound (α, or β, or both). In general, binding of PP$_i$ to P$^\text{Pro}$ increases folding rate constants and decreases unfolding rate constants, shifting the conformational equilibrium toward the more folded states. The relative contribution of folding and unfolding rate constants to the shift in equilibrium varies between the different conformational transitions. With PP$_i$ bound at the α site, the equilibrium constant for the U to I transition (K$_{UI}$) increases 21-fold but only 7% of the increase is due to a larger folding rate constant. In contrast, with PP$_i$ bound at the α site the equilibrium constant for the U to F transition (K$_{UF}$) increases 3400-fold and approximately 95% of the increase is due to a larger folding rate constant. The degree to which PP$_i$ binding changes the folding and unfolding rate constants can be interpreted in terms of the affinities of the transition states relative to U, I, and F. The U to I transition state has an α site that is similar to that of U, while the U to F transition state’s α site is similar to that of F. These transition states have no detectable β site. Binding of PP$_i$ perturbs the energies of the ground states and transition states of the conformational reactions, thereby altering the dynamics of P$^\text{Pro}$. The affinities of the transition states are different for each conformational reaction in the P$^\text{Pro}$ mechanism, such that binding of PP$_i$ redistributes the ensemble by increasing the folding rate constants and decreasing the unfolding rate constants differently for each conformational reaction.
2.2.4 **Flux is Kinetically Partitioned Between the Pathways of Coupled Folding and Binding**

The mechanism of coupled folding and binding in P-pro shows that both the population distribution of members within the ensemble and the preferred pathways by which these ensemble members interconvert depend on PPi concentration. To describe the interconversion of P-pro ensemble members we calculated the equilibrium flux through the 18 reaction pathways by which U and FL2 exchange. Four of the pathways follow the conformational selection mechanism and four follow the induced fit mechanism. The remaining ten pathways follow mixed mechanisms in which folding and binding alternate.

The total flux through all pathways between U and FL2 reaches a maximum of \(~30\%\) s\(^{-1}\) near the apparent Kd for PPi (\(~2\) μM) when half of the protein molecules are folded (Figure 2-7), then decreases substantially as the concentration of ligand increases or decreases from the apparent Kd. The flux is kinetically partitioned between the 18 pathways. Each pathway’s fractional contribution to the total flux depends on PPi concentration (Figure 2-8). At PPi concentrations at or below the apparent Kd of 2 μM, approximately 90% of the total flux is through the four conformational selection pathways. As PPi concentration increases, the contribution of the conformational selection pathways to the total flux decreases, and the contribution of the mixed mechanism pathways increases. In the presence of 20 μM PPi, 10-fold above the apparent Kd, 70% of the total flux is through seven of the mixed mechanism pathways. Approximately 12% of the total flux is through pathways in which U protein binds ligand before folding, and 58% of the flux is through pathways in which I binds ligand before folding. Most of the remaining 30% of the total flux is through the conformational selection pathways that dominate at lower ligand concentrations. As ligand concentration increases, the fractional flux through pathways in which U binds ligand before folding continues to increase. Since U does not
have a β site, the four fully induced-fit pathways in which U binds two molecules of PPi, before folding do not contribute significantly to the coupled folding and binding mechanism.

![Graph showing total equilibrium flux for the U ⇄ FL₂ reaction vs. ligand concentration](image)

**Figure 2-7: Total equilibrium flux for the U ⇄ FL₂ reaction vs. ligand concentration**

Flux was calculated from best-fit parameter values of the model describe in the text. Peak flux occurs at the apparent midpoint of the binding curve, where the populations of U and FL₂ are the same and falls off substantially at lower and higher PPi concentrations. Flux is defined as the percent of the P^{Pro} interconverting between U and FL₂ per second.
Figure 2-8: Concentration-dependent mechanism of coupled folding and binding

The mechanism of interconversion between U and FL₂ was assessed by calculating the fractional flux through each of 18 pathways. Species populations (spheres) and fractional fluxes (path lines) were calculated for 2.5 μM protein and the indicated pyrophosphate concentrations using parameter values derived from the global best-fit of TMAO- and PPI-induced folding stopped-flow data.
2.2.5 Determinants of the Kinetic Partitioning of Flux

The kinetic partitioning observed in P-Pro coupled folding and binding is not adequately described by the “conformational selection versus induced fit” dichotomy often used to describe coupled conformational change and binding reactions. In this false dichotomy the overall reaction is classified into only one of the two limiting mechanisms. An important feature of the P-Pro coupled folding and binding mechanism, and we would suggest most others, is that flux through the multiple distinct pathways is kinetically partitioned. The kinetic partitioning is strongly dependent on the rate constants for conformational change, the ligand binding affinities of the conformational states, and the free ligand concentration. Based on our assessment of the P-Pro mechanism and flux calculations we can draw general conclusions about dependencies of the partitioning. Fast folding kinetics in the unliganded, low affinity sub-ensemble, and low ligand concentration favor partitioning into conformational selection pathways. Slow folding kinetics in the unliganded protein and high affinity in the unfolded state favor partitioning into induced fit pathways. When the unfolded state is capable of binding ligand at a particular site, increasing ligand concentrations favors partitioning through pathways that begin with ligand bound to U at that site.

2.3 Conclusion

For many proteins coupled binding and conformational change may be kinetically partitioned between competing pathways, however, previous methodology and interpretation limited to the false dichotomy has precluded such observations. While the precise biological utility of a given coupled binding and conformational change mechanism is currently unknown, it is certain that biology and mechanism can only be correlated if the mechanism is properly determined. Kinetic experiments should be done over a wide range of ligand concentrations and
used to determine rate constants for elementary steps. This analysis is applicable to simple systems but also more complex systems, as long the data captures information about each of the elementary steps. Our study of \( P^{\text{Pro}} \) suggests two possible biological implications of a concentration-dependent mechanism. First, concentration-dependent partitioning gives rise to a potentially regulatory kinetic effect. The folding rate constants increase when ligand is bound, making induced fit faster than conformational selection. At low ligand concentrations where conformational selection dominates, folding is a slow step. At higher concentrations where induced fit dominates, proteins bypass this slow step. Second, the flux is kinetically partitioned between pathways with conformationally distinct intermediates. In some cases, these intermediates might participate in different biological processes (eg., binding different partners), and varying ligand concentration can be used to favor a particular pathway and a process associated with the intermediate in that pathway.

2.4 Materials and Methods

2.4.1 Expression and Purification of \( P^{\text{Pro}} \)

The gene for P protein variant P39A/P90A/F107W (\( P^{\text{Pro}} \)) used in this study was constructed from the previously described gene, F107W variant, by the QuickChange site-directed mutagenesis protocol (Stratagene). Both variants were overexpressed in \textit{Escherichia coli} BL21 DE3 pLysS and purified as previously described (66) with the following modification. The pooled fractions of P protein that eluted from the second CM-Sepharose column were dialyzed into water and sodium triphosphate (Sigma-Aldrich) was added to a final concentration of 10 mM to precipitate P protein. P protein was pelleted by centrifugation at 10,000g for 30 minutes. Pellets were resuspended in 6M Guanidine-HCl and exchanged into Guanidine-HCl several times using Pierce Protein Concentrators with 9K MWCO. Protein was stored in 6M Guanidine-HCl at -80°C until being extensively dialyzed against 20 mM sodium cacodylate pH 7.0 for use. In all
experiments, the protein concentration was determined using the Edelhoch(67) method using an extinction coefficient of 11460 M⁻¹ cm⁻¹ at 280 nm.

2.4.2 Crystallography

A 1 mL aliquot of RNase P protein (P<sub>Pro</sub>) in 6M Guanidine-HCl storage solution was dialyzed against water using a Pierce 3.5K MWCO Slide-Z-Lyzer dialysis cassette and concentrated using a 3K MWCO Amicon Ultra 0.5 mL centrifugal concentrator to 3.5 mg/mL in 10 mM Tris-HCl (pH 8.0) and mixed in a 1:1 ratio with crystallization solution (15% PEG 550 MME, 100 mM MES pH 6.5, 27.5 mM ZnSO₄ and 5 mM sodium pyrophosphate). Diffracting crystals were formed by hanging-drop vapor diffusion within three days. Cryo-preservation of the crystal was achieved with the addition of 18% PEG 550 MME to the crystallization condition. The crystals formed in the P6<sub>4</sub> space group with cell dimensions: a=b= 83.14 Å and c=32.24 Å. All data were collected at The Advanced Photon Source (APS) at Argonne National Laboratory, beamline 22-BM (SER-CAT). Data were processed and scaled with HKL3000(68). The structure was solved by molecular replacement using the <i>B. subtilis</i> RNase P protein structure as a starting model (62), and refined using Phenix(69). Table 2-1 contains statistics for the structure. PDB ID: 4JG4.

2.4.3 Enzymology

Wildtype <i>B. subtilis</i> P RNA and a fluorescently labeled <i>B. subtilis</i> pre-tRNA<sup>Asp</sup> (Fl-pre-tRNA) were prepared as described in (Rueda, Hsieh et al. 2005)(70). For single-turnover cleavage, 25 nM Fl-pre-tRNA and 500 nM RNase P holoenzyme (1:1 P protein to P RNA) were folded into 50 mM Tris/MES (pH varies), ~200 mM KCl, 10 mM CaCl₂ and 20 mM DTT. Multiple-turnover reactions contains 1 nM P RNA and 4 nM P protein in 50 mM Tris-HCl (pH 8), 100 mM KCl, 10 mM MgCl₂, 20 mM DTT and 0.01% NP-40. Cleavage reactions were
monitored by fluorescence polarization signal change using TECAN Infinite F500 plate reader (71).

2.4.4 Stopped-Flow Fluorescence

Stopped-flow fluorescence kinetic experiments were performed on an Applied Photophysics SX20 instrument at 25°C. Samples were mixed in an observation cell with a 2 mm pathlength and excited at 285 nm with a slit width of 1 mm. Emitted light was detected through a 320 nm high-pass filter. Experiments were performed by mixing protein with PP or TMAO solutions at a ratio of 1:5. The initial protein concentration was 15 μM and the final protein concentration was 2.5 μM. Final PP concentrations ranged from 0 to 166 μM and final TMAO concentrations ranged from 0 M to 0.93 M. Kinetic traces were collected over 30 seconds with logarithmic time sampling. At least three traces were collected and averaged for each concentration of TMAO or PP.

2.4.5 Analysis of Kinetic Data

The model used to analyze stopped-flow data assumes that each of the three conformational states of P-Pro contains up to two high affinity binding sites (Figure 2-4 for scheme) and that any conformational sub-ensemble can interconvert with any other sub-ensemble that has the same pattern of occupied ligand binding sites. Rate constants for folding and unfolding are assumed to be dependent on the microscopic liganding state of the molecule and the concentration of TMAO. Following standard practice, we assume that the logarithms of the microscopic rate constants are linearly dependent on co-solute (TMAO) concentration (72). The kinetics of ligand association and dissociation are too fast to be observed by stopped flow. In the analysis presented here, we have assumed that the second order association rate constant $k_{on}$ is diffusion limited and set it to $10^8$ M$^{-1}$s$^{-1}$. This assumption is consistent with our observation of a
burst phase increase in protein fluorescence upon ligand addition, which requires that $k_{on}$ be at least $10^7 \text{ M}^{-1}\text{s}^{-1}$. The model is insensitive to changes in $k_{on}$ between $10^7$ and $10^{10} \text{ M}^{-1}\text{s}^{-1}$. The dissociation rate constant $k_{off}$ for any individual binding reaction changes with $k_{on}$ because it is the product of the fixed $k_{on}$ and the fitted parameter for the association constant $K_A$.

Drift in fluorescence signal between collection of the TMAO-induced and PP$_7$-induced folding data was accounted for by addition of an offset parameter to the model. Fitted parameters were obtained by Bayesian estimation using uniform prior distributions defined on plausible parameter ranges. Posterior means and 95% credible intervals were computed by a sequential Monte Carlo technique for static data(73). 1000 particles were initialized using standard Markov chain Monte Carlo methods to draw from an initial posterior based on a subset of the data consisting of every ligand and osmolyte concentration at the first, middle, and last time points. Standard sequential Monte Carlo methods were used to propagate particles through a sequence of posterior distributions from this initial distribution to the full data posterior distribution. A small number of nearby maxima were identified, whose differences were statistically but not practically significant; we report only one.

### 2.4.6 Flux Calculation

Equilibrium flux for each reaction in the reaction scheme was calculated by multiplying the concentrations of products by the forward rate constants. Equilibrium flux for the 18 pathways of interconversion between U and FL$_2$ was calculated using the equation:

$$F_{\text{path}} = (\sum 1/F_i)^{-1}$$

where $F_{\text{path}}$ is the flux of the pathway and $F_i$ is the flux of the $i$th reaction in the pathway.
2.4.7 Isothermal Titration Calorimetry

Binding of PP$_i$ to the N-terminal peptide mimic AHLKRNRLKKNEDW was measured by ITC. The synthetic peptide AHLKRNRLKKNEDW was obtained from GenScript and purified by cation exchange chromatography and desalted by size exclusion chromatography. Crude peptide was resuspended in 100 mM NaCl, 5 mM EDTA, 50 mM sodium acetate pH 6.5 and loaded onto a CM Sepharose column. Peptide was eluted with a salt gradient of 100-800 mM NaCl, 5 mM EDTA, 50 mM sodium acetate pH 6.5. Fractions containing the peptide were pooled and the volume was reduced to 0.5 mL by SpeedVac. The 0.5 mL was loaded onto a Sephadex G-50 column and eluted with water. Eluted peptide was added to an equal volume of 40 mM sodium cacodylate pH 7.0 and frozen at -20°C until use. 170 μM peptide was titrated with PP$_i$ to a final PP$_i$ concentration of 3.5 mM. Binding of PP$_i$ to a mixture of P$^{Pro}$ fragments produced by formic acid cleavage of P$^{Pro}$ was measured by ITC. Formic acid cleavage was performed by incubating 25 μM P$^{Pro}$ in 2% formic acid for 2 hours in a boiling water bath. The presence of the predicted fragments (AHLKKRNRLNNED, FQKVFKHGTSVANRQFVLYTLD, QAEND, ELRVGSLSVSKKIGNAVMRNRIKRLIRQAFLEEKERLKEKD, and YIIIARKAASQLTYETKKSLQHLWRKSSLYKSSSK) was confirmed by MALDI-ToF mass spectrometry. P$^{Pro}$ fragments were recovered by lyophilizing the reaction mixture and resuspending in 20 mM sodium cacodylate pH 7.0. 20 μM peptide mixture was titrated with PP$_i$ to a final PP$_i$ concentration of 4.3 mM. Titrations were performed in 20 mM sodium cacodylate pH 7.0 at 25°C. Red lines in the Figure 2-7A and 2-7B are the best-fit of the data to a single site binding isotherm. 200 μM P$^{Pro}$ in 20 mM sodium cacodylate pH 7.0 at 25°C was titrated with PP$_i$ to a final PP$_i$ concentration of 740 μM. The data were fit to a three-state, three-site coupled folding and binding model using parameters obtained from the fit of the kinetic data.
3. Conformational Kinetics Reveal Affinities of Protein Conformational States

3.1 Introduction

The ensemble nature of proteins and nucleic acids, and their abilities to bind other molecules, are critical to their function(2). Often, the conformational ensembles of these macromolecules can be subdivided into kinetically distinguishable conformational subensembles or states in the thermodynamic sense. Thanks in part to the highly cooperative nature of biopolymer conformational changes(74-76), these conformational states are distinguishable by the kinetic barriers that separate them; conformers in separate states interconvert much more slowly than conformers within the same state. Thus, the reactions of such a system can be described as the interconversion of a small number of conformational states without ignoring the ensemble nature of each state(4, 5). To determine a mechanism that involves these conformational states, one must determine the timescale of their interconversion and their affinities for ligand. The practical definition of a conformational state is one whose delimiting kinetic barriers match the timescale of the experiment being used to measure its kinetics. In the case of the stopped-flow method used in this study, that timescale is 1 – 2 milliseconds or slower. Numerous studies have focused on biophysical characterization of conformational transitions coupled to binding(7, 18). None of these studies have quantified the affinities of multiple exchanging states for native ligands, despite the fact that affinities must be key determinants of the mechanism by which binding is coupled to conformational change.

A wide range of proteins and nucleic acids undergo conformational changes coupled to the binding of other biological macromolecules and small molecules. This coupled conformational change and binding is key to many biological processes including signaling(77), catalysis(78, 79), and assembly of macromolecular complexes(80). Additionally, the design of
small molecule allosteric effectors that selectively bind, and thereby populate, a desired conformational state of a protein or nucleic acid represents an attractive therapeutic strategy(7, 49, 81, 82) and potentially powerful research tool. A robust understanding of coupling on a level that would allow us to both understand and exploit existing biological systems and design new allosteric effectors or molecular switches requires experimental determination of two types of quantifiable parameters: the binding affinities of each conformational state and the forward and reverse rate constants for interconversion between those states.

Numerous methods exist to determine the affinities of macromolecules for their binding partners. Most of these methods rely on equilibrium measurements and are unable to determine the intrinsic ligand binding affinity of individual conformational states when more than one state is present. Standard equilibrium binding measurements are unable to deconvolute the affinities of the distinct states and the free energy of the conformational changes. To determine the affinity of an individual conformational state using equilibrium techniques the state must somehow be trapped or isolated. For many macromolecules, trapping of conformational states is prohibitively difficult or impossible. Lower-affinity states are often rapidly depopulated in the presence of ligand, which makes measuring their affinities by equilibrium methods nearly impossible. Measuring the affinities of individual states requires overcoming these difficulties.

Folding transitions are an important subclass of protein conformational changes and are often coupled to ligand binding (18). Recent folding studies suggest the utility of studying the kinetics of conformational changes that are coupled to binding as a way to determine the binding constants for the conformational states of proteins (46, 83). This approach contrasts with previous studies that used individually trapped states to estimate binding affinities of specific conformational states by equilibrium binding experiments (84-86). Unfortunately, neither of the recent kinetic studies explained why the affinities could be determined from kinetic experiments.
but not from equilibrium experiments. The authors also did not establish a conclusive link between the binding affinities of the conformational states and the effects of binding on conformational kinetics.

The protein subunit of bacterial RNase P has three conformational states: an unfolded state U, partially folded state I, and folded state F (61, 63). As shown in Figure 3-1, folding of the protein is coupled to binding of anionic ligands at two high-affinity sites denoted α and β(19). Estimates of binding affinities for U and I cannot be obtained from equilibrium binding experiments because U and I rapidly fold to F upon binding ligand. Folding of P protein upon addition of sufficient ligand concentrations occurs within 1-2 seconds. Because ligand-bound U and I are not significantly populated at equilibrium, fits of equilibrium binding data (isothermal titration calorimetry, intrinsic tryptophan fluorescence, circular dichroism) give no insight into the affinities of U and I. Useful binding information could be obtained if U, I, and F could be isolated under conditions compatible with binding experiments. U can be isolated by addition of molar concentrations of urea or guanidinium chloride. However, use of these denaturants is not sufficient to study binding to U because even high concentrations of urea cannot prevent folding of P protein by tight-binding anions, and guanidinium chloride electrostatically interferes with binding. As of yet, no method of isolating the partially folded intermediate I has been discovered. The folded state, F, can be isolated in the absence of ligand using the osmolyte trimethylamine N-oxide (TMAO), but use of TMAO with polyanionic ligands such as pyrophosphate induces aggregation of P protein. Because we cannot readily isolate U, I, or F for binding studies or obtain sufficient information from equilibrium binding experiments, we turned to kinetics to determine the binding affinities and conformational equilibria between U, I, and F.

In this study we aimed to better understand the ligand-binding affinity and specificity in U, I, and F and the effects of binding on conformational kinetics in RNase P protein. To do so,
Figure 3-1: $P^{\text{Pro}}$ coupled folding and binding scheme

$P^{\text{Pro}}$ exhibits three-state folding with unfolded (U), partially folded intermediate (I), and folded (F) states. Folding is strongly coupled to the binding of two pyrophosphate ligands (L). Pyrophosphate may be bound at the $\alpha$-site ($L_\alpha$), $\beta$-site ($L_\beta$), or both sites ($L_2$) in any state. Unoccupied binding sites in U and I are reflected as undetectably weak affinities in the fits of the data. Species that remain unpopulated in stopped-flow kinetic experiments as revealed by fits of the data and their associated transitions are indicated in gray.
we determined the kinetics of conformational change coupled to binding using six different phosphate derivatives. These were phosphate, pyrophosphate, and four nucleotide diphosphates: ADP, GDP, CDP and UDP. Global analysis of the nucleotide-diphosphate kinetic data allowed simultaneous determination of the ligand-binding affinities and rate constants of conformational change. U, I, and F all have measurable affinities for the studied ligands. The results indicate that ligand-bound forms of U and I are undetectable at equilibrium. In striking contrast, the ligand-bound forms of U, I, and F are all populated and detectable over the time course of kinetic experiments. The ratio of affinities of two conformational states in exchange — which we refer to as $\kappa$ — can be empirically related to the effect of ligand on conformational rate constants. The degree to which the change in dynamics is partitioned between changes in folding and unfolding rate constants—quantified in the parameter $\delta$—depends on the conformational state and ligand type. As a consequence of the $\kappa$ and $\delta$ values intrinsic to P protein, the dominant folding mechanism shifts from the indirect pathway (U—I—F) to the direct folding pathway (U—F) as ligand concentration rises. I also show that both $\kappa$ and $\delta$ are strong determinants of the mechanism of molecular recognition—conformational selection versus induced fit—in coupled conformational change and binding.

3.2 Results and Discussion

3.2.1 Ligand-bound U, I, and F are populated in kinetic experiments, but not equilibrium experiments.

The kinetic experiments described here were conducted with the P39A/P90A/F107W (PPro) variant of Bacillus subtilis ribonuclease P, a model system developed in previous studies (46, 61). To estimate the intrinsic affinities of U, I, and F, we used stopped-flow fluorescence to monitor folding of the protein upon mixing with varying concentrations of either P, ADP, GDP, UDP, or CDP. Stopped-flow kinetic data (Figures 3-2 & 3-3) were fit to the coupled folding and
Figure 3-2: Kinetics of ligand-induced folding of $P^{Pro}$

Folding was monitored by stopped-flow fluorescence. $P^{Pro}$ was mixed with either PP$_i$ (A), P$_i$ (B), ADP (C), GDP (D), UDP (E), or CDP (F). Data points are the average of data points from three traces. The blue surfaces are the best-fit of data to the coupled folding and binding models.

Figure 3-3: Kinetics of ligand-induced folding of $P^{Pro}$

Folding was monitored by stopped-flow fluorescence. $P^{Pro}$ was mixed with either PP$_i$ (A), P$_i$ (B), ADP (C), GDP (D), UDP (E), or CDP (F). Data points are the average of data points from three traces. The red lines are the best-fit of data to the coupled folding and binding models.
binding model described previously (Figure 3-1) to obtain estimates of the affinities, folding rate constants, and spectroscopic parameters (Tables 3-1 through 3-7). The time dependence of ligand-induced folding is determined by the populations of the 12 conformational and liganding states depicted in Figure 3-1 and their individual folding/unfolding kinetics. It follows that the rate of disappearance of U and I, and the formation of F upon the addition of ligand provides information about their conformational kinetics and affinities. The time-dependent populations of the 12 states in the coupled-folding and binding scheme are readily expressed as a system of differential equations that allow calculation of the observed fluorescence signal as a function of ligand concentration, protein concentration, and time (see Appendix A).

Because of the complexity of P^Pro (three conformational states and two binding sites) ligand-induced folding data for any single ligand is insufficient to determine a coupled binding and folding mechanism. In a previous study we addressed this problem by using the osmolyte trimethylamine N-oxide (TMAO) to induce folding in the absence of ligand and globally fit that data along with data obtained at various ligand (PPi) concentrations. In the current work, the same problem was overcome by globally fitting all four of the NDP-induced folding datasets. The global fit of the TMAO and PPi data and the global fit of the four NDP datasets yielded similar estimates of the conformational rate constants and equilibrium constants for unliganded P^Pro (Table 3-7). This result suggests that, for other complex systems whose conformational changes cannot be easily induced with osmolyte or denaturant, global analysis of kinetic data for ligands of varying affinities would include kinetic and equilibrium information about the unliganded states. Notably, each NDP ligand binds P^Pro with unique affinity and uniquely alters the conformational kinetics. Ligands with identical affinities or that induce identical changes in conformational kinetics would not allow determination of the coupled binding and conformational change mechanisms. In the case of macromolecules with only two conformational

67
Table 3-1: Parameter values from global fit of PP\textsubscript{i} and TMAO stopped-flow data

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>Standard Error</th>
<th>68% Confidence Interval</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal\textsubscript{U}</td>
<td>8.759</td>
<td>0.001</td>
<td>(8.757, 8.760)</td>
<td>AFU</td>
</tr>
<tr>
<td>Signal\textsubscript{I}</td>
<td>7.602</td>
<td>0.006</td>
<td>(7.596, 7.609)</td>
<td>AFU</td>
</tr>
<tr>
<td>Signal\textsubscript{F}</td>
<td>7.5549</td>
<td>0.0008</td>
<td>(7.5542, 7.5558)</td>
<td>AFU</td>
</tr>
<tr>
<td>Signal\textsubscript{l,Bound}</td>
<td>9.46</td>
<td>0.06</td>
<td>(9.40, 9.52)</td>
<td>AFU</td>
</tr>
<tr>
<td>Signal\textsubscript{r,Bound}</td>
<td>8.45</td>
<td>0.02</td>
<td>(8.43, 8.47)</td>
<td>AFU</td>
</tr>
<tr>
<td>K\alpha\textsubscript{U}</td>
<td>3.90</td>
<td>20</td>
<td>(370, 410)</td>
<td>M\textsuperscript{-1}</td>
</tr>
<tr>
<td>K\alpha\textsubscript{F}</td>
<td>1.32×10\textsuperscript{6}</td>
<td>200000</td>
<td>(1.29×10\textsuperscript{6}, 1.34×10\textsuperscript{6})</td>
<td>M\textsuperscript{-1}</td>
</tr>
<tr>
<td>K\alpha\textsubscript{I}</td>
<td>8300</td>
<td>400</td>
<td>(7900, 8600)</td>
<td>M\textsuperscript{-1}</td>
</tr>
<tr>
<td>K\beta\textsubscript{U}</td>
<td>3</td>
<td>3</td>
<td>(0.6)</td>
<td>M\textsuperscript{-1}</td>
</tr>
<tr>
<td>K\beta\textsubscript{F}</td>
<td>430000</td>
<td>10000</td>
<td>(420000, 440000)</td>
<td>M\textsuperscript{-1}</td>
</tr>
<tr>
<td>K\beta\textsubscript{I}</td>
<td>2200</td>
<td>200</td>
<td>(2000, 2400)</td>
<td>M\textsuperscript{-1}</td>
</tr>
<tr>
<td>k\textsubscript{uf}\alpha</td>
<td>2.4</td>
<td>0.3</td>
<td>(2.1, 2.8)</td>
<td>sec\textsuperscript{-1}</td>
</tr>
<tr>
<td>k\textsubscript{uf}\beta</td>
<td>7.3</td>
<td>0.8</td>
<td>(6.5, 8.1)</td>
<td>sec\textsuperscript{-1}</td>
</tr>
<tr>
<td>k\textsubscript{uf}\beta</td>
<td>13</td>
<td>7</td>
<td>(6, 20)</td>
<td>sec\textsuperscript{-1}</td>
</tr>
<tr>
<td>k\textsubscript{uf}\alpha</td>
<td>120</td>
<td>10</td>
<td>(110, 130)</td>
<td>sec\textsuperscript{-1}</td>
</tr>
<tr>
<td>k\textsubscript{uf}\alpha</td>
<td>9.1</td>
<td>0.5</td>
<td>(8.6, 9.7)</td>
<td>sec\textsuperscript{-1}</td>
</tr>
</tbody>
</table>

Table 3-2: Parameter values from fit of P\textsubscript{i} stopped-flow data

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>Standard Error</th>
<th>68% Confidence Interval</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal\textsubscript{U}</td>
<td>10.26</td>
<td>0.02</td>
<td>(10.24, 10.28)</td>
<td>AFU</td>
</tr>
<tr>
<td>Signal\textsubscript{I}</td>
<td>7.97</td>
<td>0.09</td>
<td>(7.89, 8.07)</td>
<td>AFU</td>
</tr>
<tr>
<td>Signal\textsubscript{F}</td>
<td>7.8</td>
<td>0.1</td>
<td>(7.7, 7.9)</td>
<td>AFU</td>
</tr>
<tr>
<td>Signal\textsubscript{l,Bound}</td>
<td>10.51</td>
<td>0.06</td>
<td>(10.46, 10.58)</td>
<td>AFU</td>
</tr>
<tr>
<td>Signal\textsubscript{r,Bound}</td>
<td>10.3</td>
<td>0.1</td>
<td>(10.2, 10.5)</td>
<td>AFU</td>
</tr>
<tr>
<td>K\alpha\textsubscript{U}</td>
<td>7.279</td>
<td>0.006</td>
<td>(7.273, 7.286)</td>
<td>AFU</td>
</tr>
<tr>
<td>K\alpha\textsubscript{F}</td>
<td>19</td>
<td>3</td>
<td>(17, 22)</td>
<td>sec\textsuperscript{-1}</td>
</tr>
<tr>
<td>K\alpha\textsubscript{I}</td>
<td>2060</td>
<td>90</td>
<td>(1970, 2160)</td>
<td>sec\textsuperscript{-1}</td>
</tr>
<tr>
<td>K\alpha\textsubscript{I}</td>
<td>130</td>
<td>20</td>
<td>(110, 150)</td>
<td>sec\textsuperscript{-1}</td>
</tr>
<tr>
<td>K\beta\textsubscript{U}</td>
<td>0.8</td>
<td>0.2</td>
<td>(0.5, 1.0)</td>
<td>sec\textsuperscript{-1}</td>
</tr>
<tr>
<td>K\beta\textsubscript{F}</td>
<td>2200</td>
<td>100</td>
<td>(2100, 2300)</td>
<td>sec\textsuperscript{-1}</td>
</tr>
<tr>
<td>K\beta\textsubscript{I}</td>
<td>120</td>
<td>20</td>
<td>(100, 120)</td>
<td>sec\textsuperscript{-1}</td>
</tr>
<tr>
<td>k\textsubscript{uf}\alpha</td>
<td>4.1</td>
<td>0.8</td>
<td>(3.3, 4.9)</td>
<td>sec\textsuperscript{-1}</td>
</tr>
<tr>
<td>k\textsubscript{uf}\beta</td>
<td>2.5</td>
<td>0.9</td>
<td>(1.6, 3.5)</td>
<td>sec\textsuperscript{-1}</td>
</tr>
<tr>
<td>k\textsubscript{uf}\beta</td>
<td>2.3</td>
<td>0.6</td>
<td>(1.6, 2.9)</td>
<td>sec\textsuperscript{-1}</td>
</tr>
<tr>
<td>k\textsubscript{uf}\alpha</td>
<td>45</td>
<td>4</td>
<td>(42, 49)</td>
<td>sec\textsuperscript{-1}</td>
</tr>
<tr>
<td>k\textsubscript{uf}\alpha</td>
<td>4.7</td>
<td>0.6</td>
<td>(4.1, 5.2)</td>
<td>sec\textsuperscript{-1}</td>
</tr>
</tbody>
</table>
Table 3-3: ADP parameter values from global fit of NDP stopped-flow data

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>Standard Error</th>
<th>68% Confidence Interval</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal_u</td>
<td>9.40</td>
<td>0.01</td>
<td>(9.39, 9.41)</td>
<td>AFU</td>
</tr>
<tr>
<td>Signal_i</td>
<td>4.66</td>
<td>0.05</td>
<td>(4.61, 4.71)</td>
<td>AFU</td>
</tr>
<tr>
<td>Signal_e</td>
<td>4.97</td>
<td>0.04</td>
<td>(4.93, 5.01)</td>
<td>AFU</td>
</tr>
<tr>
<td>Signal_u,Bound</td>
<td>6.4</td>
<td>0.6</td>
<td>(5.8, 6.9)</td>
<td>AFU</td>
</tr>
<tr>
<td>Signal_i,Bound</td>
<td>4.8</td>
<td>0.2</td>
<td>(4.5, 5.0)</td>
<td>AFU</td>
</tr>
<tr>
<td>Signal_e,Bound</td>
<td>4.23</td>
<td>0.01</td>
<td>(4.21, 4.24)</td>
<td>AFU</td>
</tr>
<tr>
<td>KoU</td>
<td>430</td>
<td>40</td>
<td>(390, 470)</td>
<td>M(^{-1})</td>
</tr>
<tr>
<td>KoF</td>
<td>960000</td>
<td>30000</td>
<td>(920000, 990000)</td>
<td>M(^{-1})</td>
</tr>
<tr>
<td>KoI</td>
<td>5900</td>
<td>800</td>
<td>(5100, 6700)</td>
<td>M(^{-1})</td>
</tr>
<tr>
<td>KpU</td>
<td>7</td>
<td>4</td>
<td>(4, 11)</td>
<td>M(^{-1})</td>
</tr>
<tr>
<td>KpF</td>
<td>540000</td>
<td>30000</td>
<td>(520000, 570000)</td>
<td>M(^{-1})</td>
</tr>
<tr>
<td>KpI</td>
<td>3000</td>
<td>400</td>
<td>(2600, 3400)</td>
<td>M(^{-1})</td>
</tr>
<tr>
<td>k_{0\alpha}</td>
<td>18</td>
<td>4</td>
<td>(14, 22)</td>
<td>sec(^{-1})</td>
</tr>
<tr>
<td>k_{1\alpha}</td>
<td>12</td>
<td>2</td>
<td>(9, 14)</td>
<td>sec(^{-1})</td>
</tr>
<tr>
<td>k_{F\beta}</td>
<td>14</td>
<td>5</td>
<td>(9, 19)</td>
<td>sec(^{-1})</td>
</tr>
<tr>
<td>k_{F1\alpha1}</td>
<td>130</td>
<td>30</td>
<td>(100, 160)</td>
<td>sec(^{-1})</td>
</tr>
<tr>
<td>k_{U\alpha}</td>
<td>16</td>
<td>3</td>
<td>(13, 19)</td>
<td>sec(^{-1})</td>
</tr>
<tr>
<td>Ifc</td>
<td>0.0003</td>
<td>0.0001</td>
<td>(0.0002, 0.0004)</td>
<td>AFU (\mu)M(^{-1})</td>
</tr>
</tbody>
</table>
Table 3-4: GDP parameter values from global fit of NDP stopped-flow data

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>Standard Error</th>
<th>68% Confidence Interval</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal_u</td>
<td>9.40</td>
<td>0.01</td>
<td>(9.39, 9.41)</td>
<td>AFU</td>
</tr>
<tr>
<td>Signal_i</td>
<td>4.66</td>
<td>0.05</td>
<td>(4.61, 4.71)</td>
<td>AFU</td>
</tr>
<tr>
<td>Signal_p</td>
<td>4.97</td>
<td>0.04</td>
<td>(4.93, 5.01)</td>
<td>AFU</td>
</tr>
<tr>
<td>Signal_{i,\text{Bound}}</td>
<td>4.9</td>
<td>0.4</td>
<td>(4.5, 5.2)</td>
<td>AFU</td>
</tr>
<tr>
<td>Signal_{p,\text{Bound}}</td>
<td>4.3</td>
<td>0.2</td>
<td>(4.1, 4.5)</td>
<td>AFU</td>
</tr>
<tr>
<td>Signal_{i,\text{F,Bound}}</td>
<td>4.01</td>
<td>0.01</td>
<td>(4.00, 4.03)</td>
<td>AFU</td>
</tr>
<tr>
<td>K_{\alpha}U</td>
<td>560</td>
<td>40</td>
<td>(510, 600)</td>
<td>M^{-1}</td>
</tr>
<tr>
<td>K_{\alpha}F</td>
<td>1.06 \times 10^6</td>
<td>50000</td>
<td>(1.01 \times 10^6, 1.11 \times 10^6)</td>
<td>M^{-1}</td>
</tr>
<tr>
<td>K_{\alpha}I</td>
<td>8200</td>
<td>800</td>
<td>(7400, 9000)</td>
<td>M^{-1}</td>
</tr>
<tr>
<td>K_{\beta}U</td>
<td>7</td>
<td>4</td>
<td>(3, 11)</td>
<td>M^{-1}</td>
</tr>
<tr>
<td>K_{\beta}F</td>
<td>570000</td>
<td>40000</td>
<td>(530000, 610000)</td>
<td>M^{-1}</td>
</tr>
<tr>
<td>K_{\beta}I</td>
<td>3400</td>
<td>500</td>
<td>(2900, 3900)</td>
<td>M^{-1}</td>
</tr>
<tr>
<td>k_{i,\alpha}</td>
<td>13</td>
<td>3</td>
<td>(10, 16)</td>
<td>sec^{-1}</td>
</tr>
<tr>
<td>k_{p,\alpha}</td>
<td>9</td>
<td>2</td>
<td>(7, 11)</td>
<td>sec^{-1}</td>
</tr>
<tr>
<td>k_{i,\beta}</td>
<td>9</td>
<td>5</td>
<td>(4, 14)</td>
<td>sec^{-1}</td>
</tr>
<tr>
<td>k_{p,\alpha\beta}</td>
<td>140</td>
<td>20</td>
<td>(120, 160)</td>
<td>sec^{-1}</td>
</tr>
<tr>
<td>k_{i,\beta\alpha}</td>
<td>14</td>
<td>2</td>
<td>(11, 16)</td>
<td>sec^{-1}</td>
</tr>
<tr>
<td>I_{fe}</td>
<td>0.0016</td>
<td>0.0001</td>
<td>(0.0015, 0.0017)</td>
<td>AFU \mu M^{-1}</td>
</tr>
</tbody>
</table>
Table 3-5: UDP parameter values from global fit of NDP stopped-flow data

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>Standard Error</th>
<th>68% Confidence Interval</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal\textsubscript{u}</td>
<td>9.40</td>
<td>0.01</td>
<td>(9.39, 9.41)</td>
<td>AFU</td>
</tr>
<tr>
<td>Signal\textsubscript{l}</td>
<td>4.66</td>
<td>0.05</td>
<td>(4.61, 4.71)</td>
<td>AFU</td>
</tr>
<tr>
<td>Signal\textsubscript{p}</td>
<td>4.97</td>
<td>0.04</td>
<td>(4.93, 5.01)</td>
<td>AFU</td>
</tr>
<tr>
<td>Signal\textsubscript{u,Bound}</td>
<td>6.1</td>
<td>0.7</td>
<td>(5.4, 6.8)</td>
<td>AFU</td>
</tr>
<tr>
<td>Signal\textsubscript{l,Bound}</td>
<td>5.5</td>
<td>0.3</td>
<td>(5.1, 5.8)</td>
<td>AFU</td>
</tr>
<tr>
<td>Signal\textsubscript{p,Bound}</td>
<td>4.14</td>
<td>0.02</td>
<td>(4.12, 4.16)</td>
<td>AFU</td>
</tr>
<tr>
<td>K\textsubscript{0}U</td>
<td>360</td>
<td>40</td>
<td>(320, 390)</td>
<td>M\textsuperscript{-1}</td>
</tr>
<tr>
<td>K\textsubscript{0}F</td>
<td>320000</td>
<td>20000</td>
<td>(300000, 340000)</td>
<td>M\textsuperscript{-1}</td>
</tr>
<tr>
<td>K\textsubscript{0}I</td>
<td>4900</td>
<td>600</td>
<td>(4300, 5500)</td>
<td>M\textsuperscript{-1}</td>
</tr>
<tr>
<td>K\textsubscript{p}U</td>
<td>7</td>
<td>3</td>
<td>(4, 10)</td>
<td>M\textsuperscript{-1}</td>
</tr>
<tr>
<td>K\textsubscript{p}F</td>
<td>240000</td>
<td>13000</td>
<td>(234000, 260000)</td>
<td>M\textsuperscript{-1}</td>
</tr>
<tr>
<td>K\textsubscript{p}I</td>
<td>1600</td>
<td>300</td>
<td>(1300, 1900)</td>
<td>M\textsuperscript{-1}</td>
</tr>
<tr>
<td>k\textsubscript{u}\alpha</td>
<td>18</td>
<td>5</td>
<td>(13, 23)</td>
<td>sec\textsuperscript{-1}</td>
</tr>
<tr>
<td>k\textsubscript{p}\alpha</td>
<td>9</td>
<td>2</td>
<td>(7, 11)</td>
<td>sec\textsuperscript{-1}</td>
</tr>
<tr>
<td>k\textsubscript{p}\beta</td>
<td>9</td>
<td>5</td>
<td>(5, 14)</td>
<td>sec\textsuperscript{-1}</td>
</tr>
<tr>
<td>k\textsubscript{p}\alpha\beta</td>
<td>140</td>
<td>30</td>
<td>(110, 170)</td>
<td>sec\textsuperscript{-1}</td>
</tr>
<tr>
<td>k\textsubscript{u}\alpha</td>
<td>16</td>
<td>2</td>
<td>(14, 18)</td>
<td>sec\textsuperscript{-1}</td>
</tr>
<tr>
<td>ife</td>
<td>0.0002</td>
<td>0.0001</td>
<td>(0.00016, 0.0004)</td>
<td>AFU µM\textsuperscript{-1}</td>
</tr>
</tbody>
</table>
Table 3-6: CDP parameter values from global fit of NDP stopped-flow data

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>Standard Error</th>
<th>68% Confidence Interval</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal$_u$</td>
<td>9.40</td>
<td>0.01</td>
<td>(9.39, 9.41)</td>
<td>AFU</td>
</tr>
<tr>
<td>Signal$_i$</td>
<td>4.66</td>
<td>0.05</td>
<td>(4.61, 4.71)</td>
<td>AFU</td>
</tr>
<tr>
<td>Signal$_p$</td>
<td>4.97</td>
<td>0.04</td>
<td>(4.93, 5.01)</td>
<td>AFU</td>
</tr>
<tr>
<td>Signal$_{i,\text{Bound}}$</td>
<td>5.4</td>
<td>0.6</td>
<td>(4.8, 5.9)</td>
<td>AFU</td>
</tr>
<tr>
<td>Signal$_{p,\text{Bound}}$</td>
<td>5.0</td>
<td>0.4</td>
<td>(4.6, 5.4)</td>
<td>AFU</td>
</tr>
<tr>
<td>Signal$_{f,\text{Bound}}$</td>
<td>4.07</td>
<td>0.02</td>
<td>(4.05, 4.09)</td>
<td>AFU</td>
</tr>
<tr>
<td>K$_{\alpha}U$</td>
<td>430</td>
<td>40</td>
<td>(390, 470)</td>
<td>M$^{-1}$</td>
</tr>
<tr>
<td>K$_{\alpha}F$</td>
<td>227000</td>
<td>9000</td>
<td>(218000, 236000)</td>
<td>M$^{-1}$</td>
</tr>
<tr>
<td>K$_{\alpha}I$</td>
<td>2800</td>
<td>400</td>
<td>(2400, 3200)</td>
<td>M$^{-1}$</td>
</tr>
<tr>
<td>K$_{\beta}U$</td>
<td>8</td>
<td>4</td>
<td>(4, 12)</td>
<td>M$^{-1}$</td>
</tr>
<tr>
<td>K$_{\beta}F$</td>
<td>179000</td>
<td>8000</td>
<td>(171000, 187000)</td>
<td>M$^{-1}$</td>
</tr>
<tr>
<td>K$_{\beta}I$</td>
<td>1100</td>
<td>200</td>
<td>(900, 1300)</td>
<td>M$^{-1}$</td>
</tr>
<tr>
<td>$k_{ii}\alpha$</td>
<td>27</td>
<td>5</td>
<td>(22, 32)</td>
<td>sec$^{-1}$</td>
</tr>
<tr>
<td>$k_{ii}\beta$</td>
<td>10</td>
<td>2</td>
<td>(8, 12)</td>
<td>sec$^{-1}$</td>
</tr>
<tr>
<td>$k_{ii}\alpha\beta$</td>
<td>12</td>
<td>5</td>
<td>(7, 17)</td>
<td>sec$^{-1}$</td>
</tr>
<tr>
<td>$k_{ii}\alpha\beta$</td>
<td>110</td>
<td>20</td>
<td>(90, 130)</td>
<td>sec$^{-1}$</td>
</tr>
<tr>
<td>$k_{ii}\alpha\beta$</td>
<td>14</td>
<td>2</td>
<td>(12, 16)</td>
<td>sec$^{-1}$</td>
</tr>
<tr>
<td>$k_{fe}$</td>
<td>0.0009</td>
<td>0.0001</td>
<td>(0.0008, 0.0010)</td>
<td>AFU $\mu$M$^{-1}$</td>
</tr>
</tbody>
</table>
Table 3-7: Comparison of unliganded folding parameter values from global fits of PPi & TMAO data and NDP data

<table>
<thead>
<tr>
<th>Parameter</th>
<th>PPi, &amp; TMAO</th>
<th>NDP</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta G_{U1}$ (cal mol$^{-1}$)</td>
<td>859 ± 4</td>
<td>884 ± 5</td>
</tr>
<tr>
<td>$\Delta G_{I1}$ (cal mol$^{-1}$)</td>
<td>372 ± 4</td>
<td>440 ± 10</td>
</tr>
<tr>
<td>$k_{U}$ (sec$^{-1}$)</td>
<td>1.994 ± 0.006</td>
<td>2.5 ± 0.2</td>
</tr>
<tr>
<td>$k_{IP}$ (sec$^{-1}$)</td>
<td>1.06 ± 0.01</td>
<td>0.84 ± 0.06</td>
</tr>
<tr>
<td>$k_{IP}$ (sec$^{-1}$)</td>
<td>0.036 ± 0.003</td>
<td>0.06 ± 0.01</td>
</tr>
</tbody>
</table>

*Parameter value estimates and standard errors were obtained from fits of stopped-flow data.*

states and one ligand binding site, kinetic experiments with one ligand is usually sufficient to determine the full coupled binding and conformational change mechanism. These experiments must have appropriate concentration and time ranges.

Kinetic and thermodynamic parameters from the fits of the stopped-flow data were used to calculate the populations of ligand-bound U, I, and F as a percentage of the total protein (Figure 3-4). Populations at equilibrium are shown as a function of ligand concentration. The ligand concentration spans from well below the apparent $K_D$ to well-above, and therefore ranges from non-saturating to saturating. Populations during a kinetic experiment (with ligand at 20 times the apparent $K_D$) are shown as a function of time. The population of liganded F in both equilibrium and kinetic experiments is substantial and well within the detection limit of intrinsic tryptophan fluorescence. The equilibrium populations of ligand-bound U and I are not significant at any concentration of ligand, including subsaturating ligand concentrations. Detection of ligand-bound U and I at equilibrium would require methods much more sensitive than standard optical or calorimetric binding assays. Ligand-bound U and I are only significantly populated during kinetic experiments. The populations of bound U and I in kinetic experiments are orders of magnitude greater than in equilibrium experiments, in this case are well within the detection limit of tryptophan fluorescence experiments.
Figure 3-4: Populations of ligand-bound U, I, and F in kinetic and equilibrium experiments

The populations of ligand bound P^{Pro} states as a percentage of the total protein are shown as a function of time for stopped-flow experiments and as a function of ligand concentration for equilibrium experiments. Populations were calculated using parameters from the fits of the data using a protein concentration of 5 μM. Time-dependent populations were calculated using a ligand concentration of 20 × K_{D,app}.
The information in Figure 3-4 suggests two reasons why kinetic experiments provide insights into affinities that equilibrium experiments do not. First, ligand-bound U and I are not significantly populated at equilibrium; they are populated in the mixture of protein and ligand before equilibrium is reached. Second, the change in fluorescence as a function of time reflects the transient populations of different species and their corresponding conformational rate constants. These populations are themselves dependent on the affinity of the corresponding state. Liganded species with unique conformational kinetics can only form if there is sufficient affinity, which is the property that makes them detectable in fits of the kinetic data.

3.2.2 Ligand binding affinity and specificity increase as P protein folds.

All three conformational states of P protein displayed significant affinity for P$_i$, PP$_i$, ADP, GDP, UDP, and CDP (Figure 3-5A). The protein’s affinities for all six ligands increase monotonically as the protein folds; the protein does not experience any decreases in affinity along the folding trajectory. This monotonic increase in affinity as the protein becomes more folded enhances the rate of ligand-induced folding. If the binding affinity of I were lower than that of U, the ligand binding would not favor formation of I relative to U, and folding via the intermediate would be slowed. It is unknown whether this pattern holds for other proteins or whether low affinity intermediates, with locally or globally disordered binding sites and thus lower affinity, may exist.

The relative affinities of two interconverting conformational states can be quantified by the ratio of their association constants, $\kappa$. $\kappa$ is equal to $K_{A2} / K_{A1}$, where $K_{A2}$ and $K_{A1}$ are the ligand association constants of two conformational states in exchange. Because the binding and conformational change reactions are coupled (see Figure 3-6), $\kappa$ is also the factor by which the conformational equilibrium constant changes as a result of ligand binding. More intuitively, $\kappa$ is a measure of how selective a ligand is for one of two conformational states. The ability of a ligand
Figure 3-5: Affinities for U, I, and F for six ligands

(A) Binding affinities for the \( \alpha \)-site of U, I, and F and \( \beta \)-site of I and F. (B) The affinity ratios (\( \kappa \)) for the binding sites U, I, and F. Error bars represent the 95% confidence interval.
Figure 3-6: General scheme for coupled binding and conformational change.

Two conformations $C_1$ and $C_2$ that bind a ligand $L$. $\kappa$ and $\delta$ are related to the equilibrium and rate constants as depicted.
to populate a macromolecule’s particular conformational state depends not only on the overall affinity of the ligand for the macromolecule, but also on $\kappa$. For two conformational states with a $\kappa$ of 1, ligand binding would not be coupled to conformational change and would not shift the conformational equilibrium. Throughout this manuscript, we designate the affinity of the more folded state $K_{A2}$ and the affinity of the less folded state $K_{A1}$. Figure 3-5B shows the $\kappa$ values for the six ligands used to study the $P^{Pro}$ folding mechanism. For all six ligands, the $U-F$ reaction has the largest $\kappa$ values.

As shown in Figure 3-5, $P_i$ binds to each state with the lowest affinities and $\kappa$ values of the six ligands. Consequentially, binding of $P_i$ is not as strongly coupled to folding as binding of $PP_i$ or the NDPs. Because of the lower $\kappa$ values associated with $P_i$ binding, the equilibrium and kinetic populations of ligand-bound $U$ and $I$ are higher for $P_i$ than for any of the other 5 ligands (Figure 3-4). Formation of these ligand-bound low-affinity states (such as $UL_{\alpha}$) is a prerequisite for the induced-fit mechanism of coupled binding and conformational change. As discussed later, $\kappa$ is an important determinant of how molecular recognition is partitioned between induced fit and conformational selection.

The extent of structure in a particular conformational state is the primary determinant of its binding specificity. For example, $U$ contains only one moderate-affinity binding site—$\alpha$—for the six ligands studied. This site can distinguish between $P_i$ and the other ligands, but cannot distinguish between $PP_i$ and the NDPs. In $U$ the $\alpha$ site is composed of highly basic, probably dynamic N-terminal residues(46) whose affinities for ligand appear to be determined primarily by the ligand’s charge. This lack of steric determinants due to poorly defined protein structure is presumably the origin of the low specificity in ligand recognition. Thus, the unfolded state $\alpha$ site of $P^{Pro}$ lacks structure necessary to distinguish the uncharged moieties of ADP, GDP, CDP, and UDP.
In contrast, the I state exhibits enhanced specificity commensurate with its accretion of folded structure. The affinities of the α and β sites in I indicate that it can partially recognize the uncharged moieties in the NDPs. The α site affinity of the NDPs are slightly decreased relative to PP_i, indicating that the additional structure at this site in I confers specificity that is lacking in U.

As depicted in Figure 3-5, F has higher ligand binding specificity than U or I. The F α site has similar affinities for PP_i, ADP, and GDP and approximately 3-5 fold lower affinities for UDP and CDP. The 3-5 fold reduction in affinity for UDP and CDP is modest compared to the ~1000-fold lower affinity for P_i. The increase in specificity as P protein folds strongly suggests structural elements that impart specificity are unformed in U, partially formed in I, and fully formed in F.

Isothermal titration calorimetry was used to further investigate differences in the F state’s affinity for purines and pyrimidines. P protein was titrated with either ADP, GDP, UDP, or CDP (Figure 3-7). The titrations confirmed that P protein binds ADP and GDP more tightly than UDP and CDP. Isotherms were fit to a coupled binding and folding model using the conformational equilibrium constants and binding affinities from the fit of stopped-flow data. The ITC isotherms
Figure 3-7: Isothermal titration calorimetry of P-Pre with NDPs

Binding isotherms from ITC titrations with ADP (red), GDP (orange), UDP (green), or CDP (magenta) were fit to a three-state coupled binding and folding model to obtain estimates of ΔH of binding. Affinities and conformational equilibrium constants obtained from the corresponding stopped-flow dataset were fixed. Solid lines represent the best fit.
provide no information about the affinities U or I because the liganded U and I are not significantly populated at equilibrium. The isotherms are consistent with the affinities obtained from fits of the stopped-flow data. ITC isotherms for ADP and GDP cannot be fit with the affinities for UDP and CDP. Likewise, isotherms for UDP and CDP cannot be fit using affinities for ADP and GDP. Both the kinetic and ITC data indicate that purines bind with higher affinity than pyrimidines to the folded state. The higher affinity of purines relative to pyrimidines is reflected in the observation that the binding isotherms for ADP and GDP saturate before the binding isotherms of UDP and CDP.

Either the ribose, base, or both have unfavorable steric interactions with the protein that decrease the affinity of UDP and CDP relative to PP$_i$. It is possible that ADP and GDP have fewer of these unfavorable interactions or more favorable interactions to compensate. The reduced affinity for pyrimidines, rather than an increased affinity for purines, may help P protein identify its specific binding site on P RNA. This site-dependent specificity may be an important determinant of holoenzyme assembly.

Based on a comparison with PP$_i$, the uncharged moieties of the NDPs decrease the affinity to binding sites in both I and F. This observation demonstrates that it is possible to parse the contributions of different ligand moieties to the affinity in the various conformational states. Although the selectivity for particular ligands here is modest (3-5 fold differences in affinities), other ligands and proteins could demonstrate much more selective binding. The kinetic strategy used here to measure affinities of conformational states may be of particular use in designing drugs that selectively bind and populate a particular state or pathway.
3.2.3 The magnitude of change in the folding and unfolding rate constants depends on $\kappa$.

Thus far, our focus has been on thermodynamic characteristics of ligand recognition. However, a kinetic approach also allows interpretation of the rate constants for conformational changes. Ligand binding shifts the conformational equilibrium toward the more folded state by increasing folding rate constants and decreasing unfolding rate constants. In $P^{\text{pro}}$, the magnitude of the change in the folding and unfolding rate constants depends on which conformational reaction, ligand, and binding site are involved. The multiple pathways depicted in Figure 3-1 allowed us to determine the rate constants for 5 unique ligand-bound folding reactions for each ligand, which amounts to 30 independent observations relating conformational kinetics to affinity ratios. Figure 3-8 shows the ligand-induced increase (relative to the unbound reactions) in folding (forward) rate constants (A) and decrease in unfolding (reverse) rate constants (B) plotted versus $\kappa$. Although there is no single monotonic relationship between these parameters, the trend suggests that the greater the ratio of affinities ($\kappa$) of the two states, the more ligand binding increases the folding rate constant and decreases the unfolding rate constant.

The positive correlation between the magnitude of the change in the rate constants and the magnitude of $\kappa$ underscores the fundamental link between $\kappa$ and the binding-induced change in conformational kinetics. It can be shown that $\kappa$ is equal to $k_{12}^L / k_{12} \times k_{21}^L / k_{21}$, where $k_{12}$ and $k_{21}$ are the forward reverse rate constants for conformational change without ligand bound and $k_{12}^L$ and $k_{21}^L$ are the rate constants with ligand bound (Figure 3-6). Conformational changes with larger $\kappa$ values have larger changes in the conformational equilibrium constants upon ligand binding. Accordingly, the largest changes in rate constants upon ligand binding are seen in the reactions with the largest $\kappa$ values. These constitute $U-F$ reactions with ligand bound at the $\alpha$
Figure 3-8: Ligand-induced perturbations of conformational kinetics.

Perturbations are enhanced in reactions with larger $\kappa$ values. The increase in folding rate constant (A) and decrease in unfolding rate constant (B) for the conformational reactions of $P^{Pro}$ depicted in Fig. 1 upon binding to six ligands are plotted versus the affinity ratio $\kappa$ of each reaction. Error bars represent the 95% confidence intervals in both parameters. The parameter $\delta$ (see Figure 5) quantifies the partitioning of the effect of ligand on conformational kinetics. When $\delta = 1$, ligand only accelerates folding; when $\delta = 0$, ligand only decelerates unfolding. Intermediate $\delta$ values correspond to perturbations of both rate constants. (C) $\delta$ values for folding reactions when bound to six ligands. Horizontal lines indicate average $\delta$s. (D) Fractional flux through the direct $U$ to $F$ folding pathway and the indirect $U$ to $I$ to $F$ pathway as a function of the concentration of six ligands normalized to their apparent $K_{D,app}$. 

83
site and I—F reactions with ligand bound at both the α and β sites. The effects of ligand on conformational rate constants demonstrate that binding can alter the timescales for conformational changes by orders of magnitude. Because $\kappa$ is equal to $k_{12L}/k_{12} \times k_{21L}/k_{21}$, the magnitude of both $k_{12L}/k_{12}$ and $k_{21L}/k_{21}$ are strongly correlated with $\kappa$. However, the relative effects of ligand on folding and unfolding rate constants will depend on the interaction of ligand with the transition state and this property can be very sensitive to conformation and ligand structure. A parameter that describes the partitioning of the ligand binding effect between forward and reverse conformational changes would provide useful mechanistic insights.

I investigated the degree to which changes in the conformational equilibrium constant upon binding were partitioned between changes in forward (folding) and reverse (unfolding) rate constants. This degree of partitioning, $\delta$, was calculated as shown in Figure 3-6. The calculation of $\delta$ is analogous to the calculation of $\phi$(40), which quantifies the extent of native structure in the transition state at the site of an amino acid substitution. Unlike studies of $\phi$ values, we do not presume that $\delta$ indicates the extent of folding in the transition state because ligand binding may significantly perturb the reaction coordinate of the transition state. Rather, $\delta$ provides a measure of how ligand binding changes conformational equilibrium. A $\delta$ of 0 indicates that ligand decreases the unfolding (reverse) rate constant has no effect on the folding (forward) rate constant. Conversely, a $\delta$ of 1 corresponds to a reaction in which the forward rate constant increases and the reverse rate constant is unchanged. Intermediate $\delta$ values describe reactions in which forward rate constants increase and reverse rate constants decrease. Calculated $\delta$ values for the coupled folding and binding reactions in P^Pro are shown in Figure 3-8C.
3.2.4 **The predominant P-Pro folding mechanism shifts from indirect to direct upon ligand binding.**

The P-Pro folding reaction can follow one of two pathways; either through a partially folded intermediate (I) or directly from U to F. As a consequence of the κ and δ values inherent to this system, ligand binding shifts the dominant pathway of conformational change from the U—I—F pathway to the U—F pathway (Figure 3-8D). At low ligand concentration the flux of the U—I—F pathway is 6-fold higher than that of the U to F pathway. In this model system, these two conformational pathways represent the various routes by which macromolecular conformational changes can occur. The presence of bound ligand can have a profound effect on the route of conformational change. Although the range of δ values for a given conformational transition is quite large, the average δ values for the U—F transition are higher than for the U—I and I—F transitions. In general, ligand binding shifts the increases the folding rate constants of the U—F equilibrium more than it decreases unfolding rate constants. In contrast, ligand binding shifts the equilibrium for the U—I and I—F reactions by increasing folding rate constants and decreasing unfolding rate constants about equally. For the ligands studied, both κ and δ for the U—F transition are large relative to those of the U—I and I—F transitions. As a consequence, ligand binding shifts the dominant pathway of conformational change from the U—I—F pathway to the U—F pathway (Figure 3-8D). As discussed before, a larger affinity ratio—κ—for two states corresponds to a larger increase in the conformational equilibrium constant upon ligand binding. The increase in the conformational equilibrium constant upon ligand binding is concomitant with a change in the conformational rate constants. Therefore, a larger κ naturally corresponds to larger changes in conformational kinetics. Of the three folding transitions, the U—F equilibrium constant is the most increased by ligand binding and, because of the high U—F δ value, this change mostly manifests as an increase in the folding rate constant. This effect greatly
enhances the relative flux through the U—F pathways. In other words, ligand binding lowers the activation energy of the U—F transition much more than that of the U—I and I—F transitions, leading to an increased flux through the U—F pathway.

In general, the ability of ligand binding to alter the pathway of conformational change in this way can allow biological systems to favor or disfavor sampling particular conformational states. Since conformational states can be functionally distinct, this can lead to an overall change in function. The strong link between κ and the conformational rate constants suggests that biological systems could achieve biologically necessary timescales and pathways for conformational changes by evolutionarily tuning the relative affinities of conformational states, thereby tuning κ.

### 3.2.5 The mechanism of molecular recognition is strongly dependent on κ and δ.

The parameters κ and δ both describe changes in conformational kinetics upon ligand binding. As such, we hypothesized that together they might be determinants of molecular recognition mechanisms. The extreme limits of these mechanisms are: 1) binding followed by conformational change (induced fit) and 2) conformational change followed by binding (conformational selection). Figure 3-9 shows the κ- and δ-dependence of fractional flux of the P

\[ P^{\text{Pro}} U+L \rightarrow I\cdot L, I+L \rightarrow F\cdot L \text{ and } U+L \rightarrow F\cdot L \] coupled transitions through the conformational selection and induced-fit mechanisms. Areas that are red favor conformational selection, while areas that are blue favor induced-fit. Conformational selection is favored for high κ and low δ. Induced-fit is favored at intermediate to high κ and high δ. There is significant flux through both mechanisms for low to intermediate κ and low to intermediate δ. Both κ and δ are important determinants of the mechanism of molecular recognition in coupled binding and conformational change. The exact influence of κ and δ on flux through conformational selection and induced-fit mechanisms
Figure 3-9: Partitioning of flux through conformational selection and induced-fit pathways

Flux through conformational selection (red) and induced-fit (blue) pathways as a function of $\delta$ and $\kappa$ for PP$_i$ binding to P$_{Pro}$. Conformational equilibrium constants and folding rate constants for the unliganded protein, as well as affinities for the low-affinity state were taken from fits of PP$_i$-induced folding monitored by stopped-flow fluorescence. Folding rate constants and affinities for the high-affinity state were calculated from $\delta$, $\kappa$ and the unliganded parameters. The fractional flux through conformational selection ($F_{cs}$) was calculated at the ligand concentration equal to $K_{D,app}$ for the given parameters.
depends on the intrinsic dynamics of the unbound conformational states and will therefore be
unique to each conformational transition.

For systems under kinetic control, a macromolecule has a limited time to undergo a
conformational change before being committed to a particular biological fate (eg. aggregation,
degradation, binding) determined by its conformation. The potentially different timescales of
conformational selection and induced-fit mechanisms could lead to different biological outcomes.
The timescale of conformational selection mechanism is limited by the rate of unliganded
conformational change. Molecular recognition that proceeds primarily through induced-fit is
inherently more amenable to acceleration because the conformational change in the ligand-bound
macromolecule can be orders of magnitude faster than in the unbound macromolecule. For this
reason, a ligand that drives conformational change primarily through an induced-fit pathway is
better suited to kinetically compete with biological processes that commit the macromolecule to a
particular biological fate. κ and δ—and therefore the mechanism and timescale of molecular
recognition—of a given macromolecule can be altered by changing the ligand or changing the
binding site through mutation. Optimization of κ is sufficient for populating a desired
conformational state for systems under thermodynamic control, however, optimization of both κ
and δ are necessary for populating a desired conformation for systems under kinetic control,
where the sequence of elementary steps determines the outcome of the overall reaction.

3.3 Conclusion

3.3.1 Towards More Quantitative Descriptions of Molecular Recognition

Literature discussion of conformational selection and induced-fit mechanisms in
biomolecular binding has often overlooked the critical roles that conformational state affinities
play in determining the mechanisms of coupled binding and conformational change(7, 18). I and
others have previously shown that such mechanisms depend on ligand concentration(44-46, 51). I
have shown here that the molecular origins of this ligand concentration dependence are in the intrinsic affinities of the conformational states and in the effects that ligand has on conformational kinetics. The importance of affinity in molecular recognition is demonstrated by the fact that relative flux between conformational selection and induced-fit mechanisms depends strongly on κ and δ. These two parameters quantify obvious features of molecular recognition: the propensity of each conformational state to bind ligand and the ability of the ligand to accelerate the conformational change to the high-affinity state. In general, induced fit is favored when ligand binds fairly weakly (intermediate κ) to the low-affinity state and greatly accelerates the conformational change (high δ). Conformational selection is favored when ligand binds extremely weakly (high κ) to the low-affinity state and does not accelerate conformational change (low δ). Conformational states of the same molecule can bind to a single ligand with affinities that vary by orders of magnitude. Interestingly, the kinetic studies with P^{pro} demonstrate that even relatively weak binding interactions in low-affinity states can be extremely important. These weak interactions are even more important when binding greatly accelerates conformational change and when ligand concentration approaches (<10%) the dissociation constant of the low-affinity state.

3.3.2 Using Conformational Kinetics to Study Coupled Binding and Conformational Change

Kinetics of conformational changes coupled to binding can be used to determine affinities of conformational states. In fields such as mechanistic enzymology and protein folding, there is a rich history of using kinetic experiments to study poorly populated intermediates and determine thermodynamic parameters that cannot be determined from equilibrium experiments. I have demonstrated that such an approach can also be applied to gain a mechanistic understanding of coupled binding and conformational change. The findings of this work can be generalized to any
conformational change in a protein or nucleic acid as long as the conformational states are separated by barriers and are detectable/distinguishable in a kinetic experiment. By estimating affinities of multiple ligands, one can estimate how much each chemical substituent of the ligand contributes to the affinity for the various conformational states. This information can be used to optimize the state/conformational selectivity of drugs that work by inducing conformational changes in the target protein, or to optimize synthetic molecular sensors that utilize coupled binding and conformational change. Recent advances in dynamic nuclear magnetic resonance techniques(87) that use chemical exchange to monitor conformational kinetics may provide the best estimates of affinities for conformational states.

3.4 Materials and Methods

3.4.1 Stopped-Flow Kinetics

Stopped-flow fluorescence kinetic experiments were performed on an Applied Photophysics SX20 instrument at 25°C. Samples were mixed in an observation cell with a 2 mm pathlength. Samples were excited at 285 nm for P$_i$ and 295 nm for ADP, GDP, CDP, and UDP. Slit width was 1mm. Emitted light was detected through a 320 nm high-pass filter. Experiments were performed by mixing protein with P$_i$, ADP, GDP, CDP, and UDP solutions at a volume ratio of 1:1. The initial protein concentration was 10 μM and the final protein concentration was 5 μM. Final P$_i$ concentrations ranged from 0.05 to 100 mM. Final ADP, GDP, CDP, and UDP concentrations ranged from 0 to 20 μM. Kinetic traces were collected over 30 seconds with logarithmic time sampling. Three traces were collected and averaged for each concentration of P$_i$, ADP, GDP, CDP, and UDP.
3.4.2 Analysis of Kinetic Data

The model used to analyze stopped-flow data assumes that each of the three conformational states of $P^{Pro}$ contains up to two high affinity binding sites (Figure 3-1 for scheme) and that any conformational state can interconvert with any other state that has the same pattern of occupied ligand binding sites. The time-dependent concentrations of the twelve species in the scheme in Figure 3-1 are represented as a system of differential equations. The fluorescence of the sample is expressed as the population-weighted fluorescence signals of free and bound $U$, $I$, and $F$. Rate constants for folding and unfolding are assumed to be dependent on the microscopic liganding state of the protein. Using mass balance constraints, we expressed many of the rate and equilibrium constants in terms of other rate and equilibrium constants in order to fit for as few parameters as possible and satisfy the principle of detailed balance. PP$_i$ data were globally fit with TMAO-induced folding data as previously described. P$_i$ data were fit using fixed values for unliganded conformational rate constants (obtained from the PP$_i$ and TMAO fit) in the unliganded protein. The nucleotide diphosphate-induced folding data were fit globally with no fixed parameters.

The kinetics of ligand association and dissociation are too fast to be observed by stopped flow. The tryptophan fluorescence increased within the mixing dead time (1.2 ms) in a ligand concentration-dependent manner when $P^{Pro}$ was mixed with either P$_i$ or PP$_i$, and a similar magnitude decrease was observed when $P^{Pro}$ was mixed with ADP, GDP, CDP, or UDP. The burst phase amplitude increased with ligand concentration and provided a lower bound ($10^7$ M$^{-1}$ s$^{-1}$) for the association rate constant $k_{on}$. I have assumed that the second order association rate constant $k_{on}$ is diffusion limited and set it to $10^8$ M$^{-1}$s$^{-1}$. This assumption is consistent with our observation of a burst phase change in protein fluorescence upon ligand addition, which requires $k_{on}$ to be at least $10^6$ M$^{-1}$s$^{-1}$. The model is insensitive to changes in $k_{on}$ between $10^7$ and $10^{10}$ M$^{-1}$s$^{-1}$. 

91
The dissociation rate constant $k_{\text{off}}$ for any individual binding reaction changes with $k_{\text{on}}$ because it is the product of the fixed $k_{\text{on}}$ and the fitted parameter for the association constant $K_A$.

Spectroscopic parameters from the fit of the stopped-flow data indicated that the burst phase is primarily due to a spectroscopic change that occurs when ligand binds to the $\alpha$ site of $U$.

Parameter estimates were obtained by fitting the stopped-flow data to the model described using the following approach. Fits of stopped-flow data were initialized by specifying initial guesses for the values and uncertainties of each parameter. 1000 parameter sets were generated by using the initial guesses for the parameter values as means and the initial guess for the uncertainties as standard deviations to sample from a normal distribution. At each iteration, a parameter set $(x)$ was chosen and a new was proposed using one of two methods: 1) the differences between two more parameter sets $(y$ and $z$) were multiplied by a scaling factor $a$ and added to parameter set $x$ or 2) the parameter values in parameter set $x$ were used as means to sample from normal distributions with standard deviations equal to the standard deviation of the parameters in all 1000 parameter sets multiplied by a scaling factor $b$. The proposed parameter set was either accepted as a replacement for parameter set $x$ or rejected based on the Metropolis criterion. Scaling factors $a$ an $b$ were adjusted to achieve proposal acceptance rates of approximately 20%. Fits were run for 1,000,000 iterations. Trajectories were examined to ensure that means and standard deviations for parameters had each converged. The final 1000 parameter sets constituted the posterior. 95% confidence intervals for each parameter were from the posterior distributions; After sorting, the 26th and 975th parameter value for each parameter were used as the upper and lower bounds of the 95% confidence interval. A Mathematica notebook implementing this code for analysis of the data is available upon request.
3.4.3 Isothermal Titration Calorimetry.

5 μM $P^\text{Pr}o$ was mixed with ADP, GDP, CDP, or UDP in a MicroCal VP-ITC instrument at 25°C with a stir speed of 307 rpm. Binding isotherms were fit in Mathematica to a three-state coupled binding and folding mechanism to obtain estimates of the enthalpies of binding for the α and β sites of the folded state. Affinities and conformational equilibrium constants of the unliganded protein were fixed at the mean estimates obtained from the global fit of ADP, GDP, CDP, and UDP stopped-flow kinetic experiments.

3.4.4 Generation of Flux Contour Plots for Figure 3-9.

Fluxes were calculated using the binding affinity of PP$i$ to the α site of the low-affinity state and the unliganded equilibrium constant and folding rate constant. The ligand-bound folding rate constant and the binding affinity for the higher-affinity state were calculated given $\kappa$ and $\delta$. Fluxes were calculated for the ligand concentration at which 50% of protein would be bound for the given thermodynamic parameters.
4. Kinetic Mechanism of cyclic di-GMP Sensing by DP17 RNA Biosensor

4.1 Introduction

4.1.1 Biosensors

Biosensors are proteins (88-91) and nucleic acids (26, 92, 93) used to detect and quantify small molecules. Some biosensors detect their small molecule ligands through a combination of binding and conformational change that produces a change in an observable signal such as fluorescence or color. Other biosensors work by performing enzymatic reactions on their ligands to make products that are detectable. In some cases biosensors are coupled to an electronic device to monitor and record detection. Proteins and nucleic acids are particularly useful as biosensors because they can be purified and used in vitro or produced in cells and use for in vivo detection or imaging of small molecules (94-97). Because proteins and nucleic acids have evolved to bind diverse small molecules, the potential range of targets detectable by biosensors is similarly diverse. Mutation to bases or amino acid residues can alter ligand specificity and affinity in biosensors (98). Rational design or high-throughput screening may be used to find biosensors with the desired specificity and affinity.

4.1.2 RNA Biosensors

Recently, RNA biosensors (Figure 4-1) have been produced by fusion of two RNA aptamers (99): a riboswitch aptamer (Figure 4-2) and the designed Spinach (100) (or Spinach2 (101)) aptamer (Figure 4-3). The riboswitch aptamer domain is capable of binding to the small molecule metabolite that it binds in the context of the riboswitch from which it is taken. The Spinach domain, created using SELEX (102), binds DFHBI, a small molecule mimic of the chromophore found in green fluorescent protein. DFHBI has relatively low fluorescence quantum
Figure 4-1: RNA Biosensors: Riboswitch aptamer and Spinach aptamer fusions

(A) RNA biosensors were designed by fusing a cyclic dinucleotide-binding riboswitch aptamer and the Spinach aptamer. (B) Putative scheme for RNA biosensor activity. Ligand binding enables binding of DFHBI.
Figure 4-2: Crystal structure of the Vc2 riboswitch aptamer

Crystal structure (PDB: 4YB1(98)) of the Vc2 riboswitch aptamer bound to cyclic-di-GMP. Vc2 (blue) binds to cyclid-di-GMP (red). Riboswitch aptamers such as Vc2 were used in the construction of RNA bionsensors.
Figure 4-3: Crystal structure of the Spinach aptamer bound to DFHBI

Crystal structure (PDB: 4TS2(103)) of the Spinach aptamer bound to DFHBI. Spinach (Red) binds to DFHBI (blue). The Spinach (or Spinach2) aptamer was used in the construction of RNA bionsensors.
yield when free in solution and becomes much more fluorescent upon binding to Spinach. The second generation of these RNA biosensors have Spinach RNA aptamer replaced with the Spinach2 RNA aptamer that has optimized folding properties to improve its function. A transducer stem connects the riboswitch aptamer to the Spinach2 aptamer. The purpose of the transducer stem is to allow structural communication between the metabolite-binding site in the riboswitch aptamer and the DFHBI binding site in the Spinach2 aptamer. In practice, a solution containing DFHBI and the RNA biosensor has low fluorescence in the absence of the metabolite and high fluorescence when metabolite is present. The fluorescence of the biosensor system shows a metabolite concentration dependence that is characteristic of simple binding.

The RNA biosensor system is a potentially powerful research tool for detecting and quantifying small molecules. DFHBI bound to the RNA biosensor is resistant to photobleaching(100), making it useful for long experiments. A suite of fluorophores similar to DFHBI have unique excitation and emission spectra that could be used for a variety of applications(104). The suite has excitation maxima ranging from 350-500 nm and emission maxima ranging from 450-600 nm. Because the DFHBI can cross cell membranes and RNA can be genetically encoded, the RNA biosensor system is useful for in-cell imaging to monitor concentration and location of small molecules(96, 99).

4.1.3 Cyclic dinucleotide sensing in cells with VC2 and DP17

So far, the RNA biosensors have been constructed using the aptamer domains of riboswitches that bind to cyclic dinucleotides. Cyclic dinucleotides are second messenger molecules involved in cell-signaling processes(105). Biosensors are useful for measuring the in-cell concentrations and spatial distribution of the cyclic-dinucleotides. Two cyclic dinucleotide-binding aptamers have been used to make the biosensors Vc2-Spinach and DP17-Spinach2. Vc2-Spinach and and DP17-Spinach2 contain cyclic di-GMP-binding aptamers from Vibrio cholerae.
and *Deinococcus proteolyticus*. The specificity of both riboswitches can be altered by base substitutions. DP17-Spinach2 is currently being used in studies of cyclic di-GMP (c-di-G) signaling in bacteria.

The current generation of RNA biosensors takes approximately 5-10 minutes to reach maximum fluorescence in the presence of the metabolite. This limits the use of the biosensors because cyclic dinucleotide signaling is both spacially and temporally regulated. The timescale on which cyclic dinucleotide levels change may sometimes be shorter than the 5-10 minutes required for detection by the RNA biosensors. A shorter activation time would improve the utility of the RNA biosensors. Currently, the mechanism of activation—the series of binding and conformational change reactions and the timescales on which they occur—is unknown. This makes assigning the rate-limiting step(s) impossible and makes speeding up the rate-limiting step(s) more difficult.

To improve the activation time of the biosensors, we need to understand the mechanism of activation. How does the binding of c-di-G lead to the formation of a fluorescent RNA-DFHBI complex? How do the c-di-G and DFHBI binding sites communicate with one another in the context of the RNA? On what timescales do binding and conformational change reactions occur? In what order do c-di-G and DFHBI bind to RNA? The goal of this work is to answer these questions and determine the mechanism of activation for DP17-Spinach2.

### 4.2 Results and Discussion

#### 4.2.1 Stopped-Flow Fluorescence

To determine the mechanism of activation for DP17-Spinach2 (DP17) we performed kinetic experiments using stopped-flow fluorescence. In the first set of experiments (Figure 4-4), DP17 was mixed with solutions containing a constant concentration of DFHBI and varying concentrations of c-di-G. In the second set of experiments (Figure 4-5), DP17 was mixed with
solutions containing varying concentrations of DFHBI and a constant concentration of c-di-G. Upon mixing the RNA with DFHBI and c-di-G, the fluorescence of the sample increases. The initial fluorescence signals observed in Figure 4-4A are variable. This variability has no dependence on the concentration of c-di-G, and the concentration of DFHBI is supposed to the same for each kinetic transient shown in 4.1. The initial fluorescence signals observed in Figure 4-5A vary linearly with DFHBI concentration and are in agreement with the fluorescence signals observed for DFHBI alone (Figure 4-5B). The linear dependence of the initial fluorescence in 4-5 on DFHBI concentration suggests that the variation in initial fluorescence in Figure 4-4A is not the result of a fluorescence change that occurs in the dead-time of the stopped-flow. Instead, it is the result of variations in DFHBI concentration. To improve fitting, concentration correction coefficients for c-di-G and DFHBI concentrations were floated during fitting.

Full fluorescence activation of the RNA biosensor takes approximately 10 minutes. There is an initial increase in fluorescence that is complete in approximately 1 second. This increase in fluorescence shows no dependence on the c-di-G concentration. In experiments with low c-di-GMP concentrations, there is a decrease in the fluorescence after the initial increase. For experiments with higher c-di-G concentrations, the remainder of the increase in fluorescence occurs on a much longer timescale. There are two simple ways that the markedly different timescales for the increases in fluorescence could occur. The simplest explanation is that DFHBI saturates its RNA binding site in approximately 1 second and becomes slightly more fluorescent. Full fluorescence activation resulting from the slow binding of c-di-G occurs on a longer timescale. The second explanation is that the RNA exists in two conformational states, which we will refer to as C1 and C2. DFHBI quickly saturates the RNA that pre-exists in the excited state and enhances the DFHBI fluorescence. Full fluorescence activation resulting from conversion of the remainder of the RNA from C1 to C2 occurs on a longer timescale. The conversion of the
**Figure 4-4: c-di-G concentration dependence of DP17 activation**

(A) Stopped-flow fluorescence traces were collected at c-di-G concentrations ranging from 0 to 20 μM. (B) The final time-point of each stopped-flow fluorescence transient is shown as a function of c-di-G concentration. Red lines represent the best global fit of the data shown in Figures 4-4 through 4-8.
Figure 4-5: DFHBI concentration dependence of DP17 activation

(A) Stopped-flow fluorescence traces were collected at DFHBI concentrations ranging from 0 to 20 μM. (B) The final time-point of each stopped-flow fluorescence transient (black points) and the initial fluorescence of free DFHBI (blue points) are shown as a function of DFHBI concentration. (C) The final time-point of each stopped-flow fluorescence transient is shown a function of DFHBI concentration. Red lines represent the best global fit of the data shown in Figures 4-4 through 4-8.
RNA from C1 to C2 is coupled to binding of DFHBI and/or c-di-G. Rigorous analysis of the data indicates that the second scenario—not the first—is true.

In another set of experiments (Figure 4-6), the RNA was incubated with DFHBI (>1 hour) before being mixed with varying concentrations of c-di-G. No initial increase in fluorescence was observed upon addition of c-di-G. At low c-di-G concentrations the fluorescence decreased with time. At higher c-di-G concentrations the fluorescence showed a slight decrease followed by a large increase with time.

Previous experiments showed that DFHBI bound to Spinach is resistant to photobleaching. The stopped-flow experiments conducted at low c-di-G concentrations show decreases in fluorescence consistent with photobleaching. This decrease in fluorescence is not evident in the experiments with higher c-di-G concentrations. Photobleaching of the DFHBI in the context of the RNA biosensor system was investigated with two experiments. In the first experiment (Figure 4-7), fluorescence of a solution containing only DFHBI was monitored under the same conditions as the previous stopped-flow experiments. The fluorescence of the solution decreased with approximately single exponential decay that resulted in a loss of about 20% of total fluorescence. In the second experiment (Figure 4-8), RNA was incubated with DFHBI and c-di-G (>1 hour) and the fluorescence was monitored under the same conditions as the previous stopped-flow experiments. The fluorescence decreased similarly to the fluorescence of the solution containing only DFHBI. Notably, the initial fluorescence of this solution was higher than the fluorescence final fluorescence expected based on the stopped-flow experiments shown in 4-4 and 4-5. Both experiments indicate that photobleaching significantly decreases the fluorescence of DFHBI during over the course of the stopped-flow experiments.
Figure 4-6: DP17 activation after pre-incubation with DFHBI

c-di-G concentration dependence of DP17 activation after pre-incubation with DFHBI. DFHBI and DP17 were pre-incubated for >1 hour before being mixed with c-di-G. Stopped-flow fluorescence traces were collected at c-di-G concentrations ranging from 0.1 to 20 μM. Red lines represent the best global fit of the data shown in Figures 4-4 through 4-8.

Figure 4-7: Photobleaching of free DFHBI

Stopped-flow fluorescence traces were collected with 0 μM (lower trace) and 10 μM (upper trace) DFHBI. Red lines represent the best global fit of the data shown in Figures 4-4 through 4-8.
DFHBI, c-di-G and DP17 were pre-incubated for >1 hour before being divided into two solutions and mixed 1:1 by stopped-flow. Red lines represent the best global fit of the data shown in Figures 4-4 through 4-8.
4.2.2 Stopped-Flow Fluorescence – Reduced Photobleaching

Accounting for the photobleaching in the stopped-flow experiments significantly complicates the data analysis. The photobleached DFHBI molecules have unique fluorescence when free and when bound to the RNA. Additionally, assumptions must be made about the affinity of the photobleached DFHBI for RNA and the effect the photobleached DFHBI has on the RNA conformational kinetics. Nonetheless, these photobleached DFHBI molecules exist in the experiments that utilize the RNA biosensors, so it is useful to understand the fluorescence properties of the photobleached DFHBI and how it interacts with the RNA. The effect of photobleaching can be better understood by doing experiments in which there is essentially no photobleaching. For that reason, another set of stopped-flow experiments similar to those in Figures 4-4 and 4-5 were performed. In these experiments (Figures 4-9 and 4-10), the shutter that allows excitation light into the sample cell was closed for the majority of the experiment. The shutter was opened for short periods (~0.1 seconds) to record fluorescence at approximately 20 time-points between 0 and 500 seconds. The transients from these experiments displayed both the fast small amplitude increase in fluorescence and slow large amplitude increase in fluorescence scene in the previous experiments. There was, however, no decrease in fluorescence in experiments with low c-di-G concentrations. The final fluorescence signals were also higher than those seen in Figures 4-4 and 4-5. Final fluorescence of this set of experiments is similar to the initial fluorescence seen in photobleaching of the ternary complex (Figure 4-8). Due to the small sample volumes used in these experiments and the difficulty of removing residual solutions from the stopped-flow instrument, the concentrations of c-di-G and DFHBI in these experiments deviate noticeably from the intended concentrations. The noise in the data is evident in Figures 4-9C and 4-10B. Concentration correction coefficients for c-di-G and DFHBI were floated during fitting of the data.
Figure 4-9: c-di-G concentration dependence of DP17 activation without photobleaching

Photobleaching was reduced by opening the excitation shutter for short periods (<0.1 sec) to collect fluorescence time-points. (A) Stopped-flow fluorescence traces were collected at c-di-G concentrations ranging from 0 to 20 μM. (B) Stopped-flow data in (A) is shown on linear timescale. (C) The final time-point of each stopped-flow fluorescence transient is shown as a function of c-di-G concentration. Red lines represent the best global fit of the data shown in Figures 4-9 and 4-10.
Figure 4-10: DFHBI concentration dependence of DP17 activation without photobleaching

(A) Stopped-flow fluorescence traces were collected at DFHBI concentrations ranging from 0 to 20 μM. (B) The final time-point of each stopped-flow fluorescence transient is shown a function of DFHBI concentration. Red lines represent the best global fit of the data shown in Figures 4-9 and 4-10.
4.2.3 Allosteric Communication Between the c-di-G and DFBHI Binding Sites

Two simple models (Schemes 4-1 and 4-2) were used as starting points to analyze the stopped-flow data. The stopped-flow data with or without photobleaching fit poorly to the model based on Scheme 4-1. Interestingly, many of the features observed in the transients of Figures 4-4 and 4-5 can be reproduced using the model based on Scheme 4-1. Many of the individual transients can be readily fit with this simple model. Attempts to globally fit the transients fail because the model cannot recapitulate the c-di-G and DFBH1 concentration dependence of the observed fluorescence.

In contrast, the stopped-flow data collected with minimal photobleaching is well-described by a model based on Scheme 4-2, in which the DFBH1 fluorescence depends on whether it is free, bound to RNA in state C1, or bound to RNA in state C2. The stopped-flow data that displays photobleaching fit well to a similar model that accounts treats photobleached DFBH1 as an additional ligand. It is not entirely surprising that the data fit well to models in which RNA conformational change is coupled to ligand binding. Riboswitches, such as the one from which the c-di-G binding aptamer were taken, function through metabolite-induced conformational change(106). Additionally, DP17 was constructed so that the binding of c-di-G would lead to a conformational change that promotes formation of structural features necessary for DFBH1 binding(99, 103). Equations 4-1 and 4-2 relate the concentrations of DFBH1-containing species to the observed fluorescence. SigD, SigD_C1, and SigD_C2 are the fluorescence signals of DFBH1 free in solution, bound to C1, and bound to C2. D(t), D_C1(t), D_C2(t) and are time-dependent concentrations (in µM) of free DFBH1, DFBH1 bound to C1, and DFBH1 bound to C2. Additional signal parameters and time-dependent concentrations for photobleached DFBH1 species are denoted in Equation 2 with a subscript ‘p.’ F0 is the background signal observed from scattering. Similar physical parameters for RNA conformational change, c-di-G binding, DFBH1
binding, and DFHBI fluorescence were obtained from the fits of the stopped-flow data sets with and without photobleaching. The results are summarized below.

\[ F(t) = \text{Sig}_{D_1} \times D_{1}(t) + \text{Sig}_{D_c} \times D_{c1}(t) + \text{Sig}_{D_c} \times D_{c2}(t) + F_0 \]  
(Equation 4-1)

\[ F(t) = \text{Sig}_{D_1} \times D_{1}(t) + \text{Sig}_{D_c} \times D_{c1}(t) + \text{Sig}_{D_c} \times D_{c2}(t) + \text{Sig}_{D_p} \times D_{p}(t) + \text{Sig}_{D_{c1p}} \times D_{c1p}(t) + \text{Sig}_{D_{c2p}} \times D_{c2p}(t) + F_0 \]  
(Equation 4-2)

DP17 exists in two conformational states (Table 4-1). In the absence of ligand, 94% of the RNA exists as C1 and 6% exists as C2. There is a highly cooperative transition between the two states, which are separated by a large kinetic barrier. Conformational change between the two states is extremely slow. Estimates for rate constants of conformational change are shown in Table 4-2. The most significant conformational reaction for which rate constants were determined is the C1—C2 reaction with c-di-G bound to the RNA. This reaction is the slowest step in the activation process through which there is significant flux. The data—for which c-di-G and DFHBI concentrations were widely varied—fit well to the two-state model described. The model does not account for any form of allostery besides that which is derived from conformational change coupled to binding. It is likely that within C1 and C2, the c-di-G and DFHBI binding sites do not communicate. Allosteric communication between the two binding sites is achieved through RNA conformational change.

Affinities for c-di-G and DFHBI binding to C1 and C2 are shown in Table 4-3. The conformational change from C1 to C2 is driven primarily by c-di-G binding. \( \kappa \) is the ratio of affinities of C2 and C1 (\( K_{A2}/K_{A1} \)). Comparing \( \kappa \) values is a simple way to compare the ability of two ligands to drive conformational change. c-di-G binds much more tightly to C2 than to C1 (\( \kappa \sim 20-25 \)). The high \( \kappa \) of c-di-G binding causes c-di-G to significantly shift the conformational equilibrium towards C2. With c-di-G bound, the RNA conformational equilibrium favors C2. DFHBI binds more similarly to C1 and C2 (\( \kappa \sim 4-5 \)). Even when saturated with DFHBI, the RNA
Scheme 4-1: DP17 Activation Scenario I

Symbols: N – c-di-G, D – DFHBI

Scheme 4-2: DP17 Activation Scenario II

Symbols: C1 – DP17 conformational state 1, C2 – DP17 conformational state 2, N – c-di-G, D – DFHBI.
Table 4-1: Populations of C1 and C2 conformational states

<table>
<thead>
<tr>
<th>Conformational State</th>
<th>Population (%)$^a$</th>
<th>Population (%)$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>94.4 ± 0.2</td>
<td>92.9 ± 0.3</td>
</tr>
<tr>
<td>C2</td>
<td>5.6 ± 0.2</td>
<td>7.1 ± 0.3</td>
</tr>
</tbody>
</table>

$^a$Parameter values and standard errors were estimated from fits of stopped-flow data in Figures 4-1 through 4-5.

$^b$Parameter values and standard errors were estimated from fits of stopped-flow data in Figures 4-6 and 4-7.

Table 4-2: Rate constants for C1 to C2 conformational changes

<table>
<thead>
<tr>
<th></th>
<th>$k_{12}$ (sec$^{-1}$)$^a$</th>
<th>$k_{12}$ (sec$^{-1}$)$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>0.0031 ± 0.0003</td>
<td>0.0015 ± 0.0001</td>
</tr>
<tr>
<td>C1 + c-di-G</td>
<td>0.0181 ± 0.0007</td>
<td>0.0174 ± 0.003</td>
</tr>
<tr>
<td>C1 + DFHBI</td>
<td>0.0047 ± 0.0007</td>
<td>0.0039 ± 0.0004</td>
</tr>
<tr>
<td>C1 + c-di-G + DFHBI</td>
<td>0.018 ± 0.002</td>
<td>0.045 ± 0.004</td>
</tr>
</tbody>
</table>

$^a$Parameter values and standard errors were estimated from fits of stopped-flow data in Figures 4-1 through 4-5.

$^b$Parameter values and standard errors were estimated from fits of stopped-flow data in Figures 4-6 and 4-7.

Table 4-3: Association constants for C1 and C2 binding to c-di-G and DFHBI

<table>
<thead>
<tr>
<th></th>
<th>$K_A$ (M$^{-1}$)$^a$</th>
<th>$K_A$ (M$^{-1}$)$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1 + c-di-G</td>
<td>3.4 (± 0.2) × 10$^9$</td>
<td>2.94 (± 0.07) × 10$^9$</td>
</tr>
<tr>
<td>C1 + DFHBI</td>
<td>8.6 (± 0.4) × 10$^4$</td>
<td>6.1 (± 0.2) × 10$^4$</td>
</tr>
<tr>
<td>C2 + c-di-G</td>
<td>8.0 (± 0.3) × 10$^6$</td>
<td>5.8 (± 0.3) × 10$^6$</td>
</tr>
<tr>
<td>C2 + DFHBI</td>
<td>3.7 (± 0.1) × 10$^3$</td>
<td>3.1 (± 0.1) × 10$^3$</td>
</tr>
</tbody>
</table>

$^a$Parameter values and standard errors were estimated from fits of stopped-flow data in Figures 4-1 through 4-5.

$^b$Parameter values and standard errors were estimated from fits of stopped-flow data in Figures 4-6 and 4-7.
Table 4-4: Fluorescence intensity of DFHBI

<table>
<thead>
<tr>
<th></th>
<th>Fluorescence (µM⁻¹)ᵃ</th>
<th>Fluorescence (µM⁻¹)ᵇ</th>
</tr>
</thead>
<tbody>
<tr>
<td>DFHBI</td>
<td>0.215 ± 0.004ᵈ</td>
<td>0.244 ± 0.001</td>
</tr>
<tr>
<td>DFHBI (C1)</td>
<td>0.4 ± 0.4</td>
<td>0.3 ± 0.2</td>
</tr>
<tr>
<td>DFHBI (C2)</td>
<td>56 ± 1</td>
<td>56.1 ± 0.1</td>
</tr>
<tr>
<td>pDFHBI</td>
<td>0.164 ± 0.001ₑ</td>
<td>−ᵣ</td>
</tr>
<tr>
<td>pDFHBI (C1)</td>
<td>1.1 ± 0.6</td>
<td>−ᵣ</td>
</tr>
<tr>
<td>pDFHBI (C2)</td>
<td>50 ± 1</td>
<td>−ᵣ</td>
</tr>
</tbody>
</table>

ᵃParameter values and standard errors were estimated from fits of stopped-flow data in Figures 4-1 through 4-5.
ᵇParameter values and standard errors were estimated from fits of stopped-flow data in Figures 4-6 and 4-7.
ᶜFluorescence parameters for photobleached DFHBI.
ᵈParameter value and standard error were estimated from fit of stopped-flow data in Figure 4-2B.
ᵉParameter value and standard error were estimated from fit of stopped-flow data in Figure 4-2B.
ᵣNo parameters for photobleached DFHBI were determined.
conformational equilibrium favors C1. With both DFHBI and c-di-G bound, the RNA
conformational equilibrium significantly favors C2. The low $\kappa$ for DFHBI binding prevents the
RNA from acting exclusively as a DFHBI sensor. When both ligands are bound, the combined $\kappa$
(the product of the two $\kappa$s) of nearly 100 counteracts the intrinsic C1—C2 equilibrium constant
($\sim0.05$) so that DP17 with both ligands bound favors C2 over C1 by a factor of 20. As a result,
approximately 95% of DP17 with both ligands bound exists as C2.

DFHBI fluorescence is the observable indicator for c-di-G binding. Estimates of the
fluorescence signals of DFHBI are shown in Table 4-4. As mentioned, the c-di-G communicates
with the DFHBI by driving RNA conformational change from C1 to C2. In C2, DFHBI
fluorescence is enhanced more than 100-fold over DFHBI bound to C1 or free in solution. In
other words, the c-di-G binding drives the conformational change that slightly increases DFHBI
affinity and greatly increases DFHBI fluorescence.

Photobleached DFHBI has approximately 80% of the fluorescence of DFHBI when free
and in solution. DFHBI seems only to be subject to photobleaching while free in solution.
Photobleaching of the RNA-bound DFHBI is negligibly slow. Photobleached DFHBI is
incorporated into the RNA as DFHBI associates and dissociates over the course of the
experiments. The model used to analyze the stopped-flow data for DP17 activation assumes that
DFHBI and photobleached DFHBI bind to RNA with the same affinities and have the same effect
on the DP17 conformational kinetics. This model describes the data well and the parameters
agree with those obtained from fitting the data collected with minimal photobleaching. Together,
these observations suggest that the assumptions about photobleached DFHBI are reasonable.
Photobleaching decreases the fluorescence of DFHBI, but does not significantly alter its
interactions with DP17.
The association kinetics (Table 4-5) in the DP17 biosensor are far below diffusion limited. Both c-di-G and DFHBI bind to DP17 with low association rate constants. The association rate constants are low for binding to both C1 and C2 forms of DP17. Interestingly, the association rate constants suggest that C1 associates slowly with DFHBI and C2 associates slowly with c-di-G. Future structural studies of DP17 may give some indication as to the reason for these slow association kinetics. Despite the slow association kinetics, the conformational change from C1 to C2 is still slow enough to be rate-limiting.

Table 4-5: Association rate constants for C1 and C2 binding to c-di-G and DFHBI

<table>
<thead>
<tr>
<th></th>
<th>$k_{on}$ (M$^{-1}$ sec$^{-1}$)$^b$</th>
<th>$k_{off}$ (M$^{-1}$ sec$^{-1}$)$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1 + c-di-G</td>
<td>$2.1 \pm 0.2 \times 10^4$</td>
<td>$1.25 \pm 0.05 \times 10^4$</td>
</tr>
<tr>
<td>C2 + DFHBI</td>
<td>$3 \pm 7$</td>
<td>$1 \pm 1$</td>
</tr>
<tr>
<td>C1 + c-di-G</td>
<td>$200 \pm 800$</td>
<td>$13 \pm 94$</td>
</tr>
<tr>
<td>C2 + DFHBI</td>
<td>$2.2 \pm 0.1 \times 10^5$</td>
<td>$2.1 \pm 0.1 \times 10^5$</td>
</tr>
</tbody>
</table>

$^a$ Parameter values and standard errors were estimated from fits of stopped-flow data in Figures 4-1 through 4-5.

$^b$ Parameter values and standard errors were estimated from fits of stopped-flow data in Figures 4-6 and 4-7.

4.2.4 DP17 Activation Mechanism

One goal of this work was to determine the order in which binding and conformational change events happen in DP17 activation. The answer is not necessarily simple, since there is more than one possible order of events. Activation may occur through multiple pathways, each with a unique order of events. Having determined the kinetic parameters for steps in the DP17 activation for which there is significant flux, it is possible to determine the flux through the various pathways by which DP17 can activate(44). There are six possible pathways through which C1 can form the C1•DFHBI•c-di-G complex. The flux through each of the six pathways can be calculated using the equation for serial flux. Most of the flux is through the pathway in which C1 binds c-di-G, undergoes the conformational change to C2, and then binds DFHBI.
Figure 4-11 shows the fraction of the total flux through each of the six pathways at 0.1, 1, and 10 µM c-di-G concentrations. At each of the three c-di-G concentrations, the dominant pathway is the same. Figure 4-12 demonstrates that this pathway dominates over a wide range of c-di-G. At low c-di-G concentrations, there is significant flux through the pathway in which DFHBI and c-di-G bind to C1 before it undergoes a conformational change to C2.

The activation mechanism of DP17 can also be understood by looking at the time-dependent populations of species formed during activation. Figure 4-13 shows the populations of the species formed over 1000 seconds when 10 µM c-di-G and 10 µM DFHBI are mixed with 100 nM DP17 RNA. The first event in activation is a decrease in the pre-existing C2 population as DFHBI binds to C2 to form C2•DFHBI. The C2•DFHBI population goes through a maximum at about 3 seconds. c-di-G binds to C1 to form C1•c-di-G, which reaches a maximum population at 10-20 seconds. At its maximum population, C1•c-di-G accounts for approximately 60% of the RNA. The high population of C1•c-di-G and its slow disappearance are the result of the rate-limiting conformational change from C1•c-di-G to C2•c-di-G. The population of C1•c-di-G decreases as it is converted into C2•c-di-G and subsequently C2•DFHBI•c-di-G. Notably, the population of C2•c-di-G is high even as the system comes to equilibrium. C2•c-di-G remains populated because DFHBI is not at saturating concentrations. Increasing the DFHBI concentration or increasing both the C1 and C2 affinity for DFHBI would decrease the population of C2•c-di-G by forming more C2•DFHBI•c-di-G. Such an improvement would increase the amplitude of fluorescence associated with c-di-G sensing.
Figure 4-11: Fractional flux through DP17 activation pathways

The fractional flux through six pathways of DP17 activation are shown. The dominant pathway at 0.1 μM (A), 1 μM (B), and 10 μM (C) c-di-G is the C1—C1•N—C2•N—C2•N•D. Symbols: C1 – DP17 conformational state 1, C2 – DP17 conformational state 2, N – c-di-G, D – DFHBI.
Figure 4-12: Fractional flux through DP17 activation pathways

The fractional flux through six pathways of DP17 activation are shown as a function of c-di-G concentration. The dominant pathway over the range of c-di-G concentrations shown is the C1—C1•N—C2•N—C2•N•D. Symbols: C1 – DP17 conformational state 1, C2 – DP17 conformational state 2, N – c-di-G, D – DFHBI.

<table>
<thead>
<tr>
<th>Pathway</th>
<th>Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_1 \rightleftharpoons C_2 \rightleftharpoons C_2\cdot N \rightleftharpoons C_2\cdot N\cdot D$</td>
<td>![Red]</td>
</tr>
<tr>
<td>$C_1 \rightleftharpoons C_2 \rightleftharpoons C_2\cdot D \rightleftharpoons C_2\cdot N\cdot D$</td>
<td>![Orange]</td>
</tr>
<tr>
<td>$C_1 \rightleftharpoons C_1\cdot N \rightleftharpoons C_2\cdot N \rightleftharpoons C_2\cdot N\cdot D$</td>
<td>![Purple]</td>
</tr>
<tr>
<td>$C_1 \rightleftharpoons C_1\cdot D \rightleftharpoons C_2\cdot D \rightleftharpoons C_2\cdot N\cdot D$</td>
<td>![Green]</td>
</tr>
<tr>
<td>$C_1 \rightleftharpoons C_1\cdot N \rightleftharpoons C_1\cdot N\cdot D \rightleftharpoons C_2\cdot N\cdot D$</td>
<td>![Blue]</td>
</tr>
<tr>
<td>$C_1 \rightleftharpoons C_1\cdot D \rightleftharpoons C_1\cdot N\cdot D \rightleftharpoons C_2\cdot N\cdot D$</td>
<td>![Black]</td>
</tr>
</tbody>
</table>
The populations of 8 species containing DP17 are shown as a function of time. Simulation was performed using parameters obtained from fits of data in Figures 4-4 through 4-8. Simulation conditions: DP17 RNA (100 nM) is mixed with a solution containing c-di-G (10 μM) and DFHBI (10 μM). For simplicity, the simulation assumes no photobleaching occurs.
4.2.5 Speeding Up Biosensor Activation

RNA biosensors are useful tools for studying cyclic dinucleotide mediated signaling in cells. So far, Vc2 and DP17 have been used to study the concentration and localization of c-di-G. Signaling by cyclic dinucleotides is known to be both spatially and temporally regulated. The slow activation time of DP17 in response to c-di-G limits its usefulness in studying c-di-G signaling. An important part of improving the RNA biosensors is reducing the activation time.

DP17 activates primarily through one pathway: C1 binds c-di-G, undergoes a conformational change to C2, and then binds DFHBI. The rate-limiting step in this activation pathway is the conformational change from C1 to C2 when c-di-G is bound. Experiments with RNase P protein showed that ligand can potentially accelerate conformational change by orders of magnitude. Conformational rate constants for the C1—C2 transition (Table 4-2) were estimated from analysis of the kinetic data. c-di-G has a slightly higher δ (0.6-0.8) than DFHBI (0.3-0.6), indicating c-di-G shifts the equilibrium primarily by accelerating the forward C1—C2 conformational change, while DFHBI does not. The significantly higher κ (20-25 for c-di-G versus 4-5 for DFHBI) and higher δ make c-di-G much better than DFHBI at accelerating the C1—C2 conformational change. Even so, c-di-G does not accelerate the conformational kinetics enough to make DP17 useful for sensing c-di-G in realtime.

The most reasonable way to reduce the activation time of DP17 would be through mutations. At least two reasonable strategies could be used to search for variants that respond more quickly to c-di-G: kinetic SELEX(102) and φ analysis(107). Kinetic SELEX would require generating a library of RNAs similar to DP17 and performing SELEX using immobilized c-di-G. The duration of the binding step would be shortened and the concentration/density of c-di-G reduced so that only RNAs that bind with high affinity to c-di-G in a short period of time would be enriched. This process could select for RNAs that quickly transition from C1 to C2.
Unfortunately, it could also select for RNAs that bind tightly to c-di-G while in C1. After several rounds of kinetic SELEX, enriched RNAs would be tested for sensing function including sensitivity to c-di-G and activation time.

ϕ analysis could be used to move towards intelligently designing RNA biosensors that activate in a shorter time. To perform ϕ analysis, mutations would be made and their effects on conformational kinetics would be determined from kinetic experiments. By making many mutations to probe the effects of adding or removing interactions throughout the RNA, the interactions formed in the transition state could be mapped. Increasing the number of transition state interactions that are already formed in C1 should increase the rate constant for the C1—C2 conformational change and reduce the activation time.

4.2.6 Use of a Covalently Attached Fluorophore

Current use of DP17 and similar biosensors requires the addition of DFHBI. DFHBI is only weakly coupled to the C1—C2 conformational change and the conformational change is necessary for the increase in DFHBI fluorescence. In principle, the current DFHBI could be replaced with a covalently attached fluorophore. A covalently attached fluorophore would likely not be genetically encoded, but rather added to the RNA using a chemistry that works in a cellular environment. Once an RNA molecule has been chemically modified with fluorophore, its activation time would be slightly reduced since association of DFHBI is complete. The primary advantage would be that the covalently attached fluorophore should be more resistant to photobleaching, since most of photobleaching in the present RNA biosensor system occurs when DFHBI is not bound to the RNA.
4.3 Conclusion

RNA biosensors are an attractive tool for *in vitro* and *in vivo* detection and quantification of small molecules. The c-di-G-sensing RNA DP17 was studied using stopped-flow kinetics to determine its mechanism of activation and discover thermodynamic and kinetic details of its function. c-di-G binding is strongly coupled to the C2—C2 conformational change in DP17. This conformational change slightly enhances the affinity for DFHBI, but greatly increases the fluorescence of DFHBI. The weak coupling of DFHBI binding and the C1—C2 conformational change prevents DP17 from simply acting as a DFHBI sensor. Sensing of c-di-G occurs primarily through binding of c-di-G to C1, followed by the C1—C2 conformational change, and binding of DFHBI. The rate-limiting step in activation is the C1—C2 conformational change. Increasing the rate constant for the C1—C2 conformational change would increase the utility of DP17 by making it useful in temporal studies of c-di-G signaling. The analysis used here could be useful for probing the functions of other protein and RNA biosensors.

4.4 Materials and Methods

4.4.1 Materials Preparation

DFHBI was purchased from Lucerna (New York, NY) and was prepared as a 31.6 mM stock solution in DMSO. Cyclic di-GMP was purchased from Axxora, LLC (Farmingdale, NY). DP17 RNA was transcribed from template DNA using T7 RNA polymerase in 40 mM Tris-HCl, pH 8.0, 6 mM MgCl₂, 2 mM spermidine, and 10 mM DTT. DP17 RNA was purified in a denaturing (7.5 M urea) 6% polyacrylamide gel and was extracted from gel pieces using Crush Soak buffer (10 mM Tris-HCl, pH 7.5, 200 mM NaCl and 1 mM EDTA, pH 8.0). RNAs were precipitated with ethanol, resuspended in TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA).
4.4.2 Stopped-Flow Fluorescence

Stopped-flow fluorescence experiments were performed on an Applied Photophysics SX-20 instrument at 37°C. Samples were excited at 448 nm and emission was collected using a 495 nm high-pass filter. All RNA, c-di-GMP, and DFHBI solutions were prepared in 125 mM KCl, 3 mM MgCl$_2$, 40mM HEPES pH 7.5. RNA was annealed by incubating at 70°C for 5 minutes before cooling at room temperature for 30 minutes and placed at 37°C. The final concentration of RNA in all experiments was 100 nM. DFHBI solutions were prepared fresh from the 31.6 mM stock each day of experiments.

For activation experiments, 200 nM RNA was mixed 1:1 with a solution containing both c-di-G and DFHBI. In experiments with varied c-di-G, DFHBI concentration was kept constant at 10 µM as c-di-G was varied from 0 to 20 µM. In experiments with varied DFHBI, c-di-G concentration was kept constant at 10 µM as DFHBI concentration was varied from 0 to 20 µM. To study pre-incubation with DFHBI, 200 nM RNA and 20 µM DFHBI were pre-incubated at 37°C for at least 1 hour before being mixed 1:1 with c-di-G to reach final c-di-G concentrations of 0 to 10 µM.

Photobleaching of the C2•DFHBI•c-di-G complex was tested by pre-incubating 15 µM c-di-G, 10 µM DFHBI, and 100 nM RNA for 1 hour. The solution was divided in half and mixed 1:1 so that no change in solution composition would occur. Photobleaching of DFHBI was carried out by mixing 20 µM DFHBI with buffer (125 mM KCl, 3 mM MgCl$_2$, 40 mM HEPES pH 7.5) in a 1:1 ratio.

Reduced photobleaching experiments were carried out by manually opening the excitation shutter for <0.1 seconds at a time at specific timepoints.
4.4.3 Data Analysis

Models to describe DP17 activation were written and implemented in Mathematica 9. Ordinary differential equations were used to calculate the time-dependent change in concentrations of RNA, ligands (DFHBI and c-di-G), and RNA-ligand complexes. Concentrations of DFHBI-containing species were related to observed fluorescence using the equations 4-1 and 4-2. All data were analyzed in Mathematica 9 using a parallel tempering MCMC algorithm. Data in figures 4-4 through 4-8 were globally fit. Data in figures 4-9 and 4-10 were globally fit. To improve fitting, concentration correction coefficients for c-di-G and DFHBI concentrations were floated during fitting.
Appendix A

Mathematical Description of P Protein Coupled Binding and Folding Model

Lt → total ligand concentration
Pt → total protein concentration
t → time
Lf → free ligand concentration

Kui → equilibrium constant for the U to I reaction
Kif → equilibrium constant for the I to F reaction
Kuf → equilibrium constant for the U to F reaction

kui → rate constant for the U to I transition
kif → rate constant for the I to F transition
kuf → rate constant for the U to F transition
kiu → rate constant for the I to U transition
kfi → rate constant for the F to I transition
kfu → rate constant for the F to U transition

kuiα → rate constant for the U to I transition with ligand bound at the α site
kifα → rate constant for the I to F transition with ligand bound at the α site
kufα → rate constant for the U to F transition with ligand bound at the α site
kiuα → rate constant for the I to U transition with ligand bound at the α site
kfiα → rate constant for the F to I transition with ligand bound at the α site
kfuα → rate constant for the F to U transition with ligand bound at the α site

kuiβ → rate constant for the U to I transition with ligand bound at the β site
kifβ → rate constant for the I to F transition with ligand bound at the β site
kufβ → rate constant for the U to F transition with ligand bound at the β site
kiuβ → rate constant for the I to U transition with ligand bound at the β site
kfiβ → rate constant for the F to I transition with ligand bound at the β site
kfuβ → rate constant for the F to U transition with ligand bound at the β site

kuiαβ → rate constant for the U to I transition with ligand bound at both the α and β sites
kifαβ → rate constant for the I to F transition with ligand bound at both the α and β sites
kufαβ → rate constant for the U to F transition with ligand bound at both the α and β sites
kiuαβ → rate constant for the I to U transition with ligand bound at both the α and β sites
kfiαβ → rate constant for the F to I transition with ligand bound at both the α and β sites
kfuαβ → rate constant for the F to U transition with ligand bound at both the α and β sites
\[ K_{\alpha u} \rightarrow \text{Binding constant for the } \alpha \text{ site in state } U \]
\[ K_{\alpha i} \rightarrow \text{Binding constant for the } \alpha \text{ site in state } I \]
\[ K_{\alpha f} \rightarrow \text{Binding constant for the } \alpha \text{ site in state } F \]
\[ K_{\beta u} \rightarrow \text{Binding constant for the } \beta \text{ site in state } U \]
\[ K_{\beta i} \rightarrow \text{Binding constant for the } \beta \text{ site in state } I \]
\[ K_{\beta f} \rightarrow \text{Binding constant for the } \beta \text{ site in state } F \]

\[ k_{\alpha u} \rightarrow \text{Association rate constant for ligand binding to the } \alpha \text{ site in state } U \]
\[ k_{\alpha i} \rightarrow \text{Association rate constant for ligand binding to the } \alpha \text{ site in state } I \]
\[ k_{\alpha f} \rightarrow \text{Association rate constant for ligand binding to the } \alpha \text{ site in state } F \]
\[ k_{\beta u} \rightarrow \text{Association rate constant for ligand binding to the } \beta \text{ site in state } U \]
\[ k_{\beta i} \rightarrow \text{Association rate constant for ligand binding to the } \beta \text{ site in state } I \]
\[ k_{\beta f} \rightarrow \text{Association rate constant for ligand binding to the } \beta \text{ site in state } F \]

\[ k_{d\alpha u} \rightarrow \text{Dissociation rate constant for the } \alpha \text{ site in state } U \]
\[ k_{d\alpha i} \rightarrow \text{Dissociation rate constant for the } \alpha \text{ site in state } I \]
\[ k_{d\alpha f} \rightarrow \text{Dissociation rate constant for the } \alpha \text{ site in state } F \]
\[ k_{d\beta u} \rightarrow \text{Dissociation rate constant for the } \beta \text{ site in state } U \]
\[ k_{d\beta i} \rightarrow \text{Dissociation rate constant for the } \beta \text{ site in state } I \]
\[ k_{d\beta f} \rightarrow \text{Dissociation rate constant for the } \beta \text{ site in state } F \]

\[ U(t) \rightarrow \text{Concentration of } U \text{ as a function of time} \]
\[ U_{\alpha}(t) \rightarrow \text{Concentration of } U \text{ with ligand bound at the } \alpha \text{ site as a function of time} \]
\[ U_{\beta}(t) \rightarrow \text{Concentration of } U \text{ with ligand bound at the } \beta \text{ site as a function of time} \]
\[ U_{\alpha\beta}(t) \rightarrow \text{Concentration of } U \text{ with ligand bound at the } \alpha \text{ and } \beta \text{ sites as a function of time} \]
\[ I(t) \rightarrow \text{Concentration of } I \text{ as a function of time} \]
\[ I_{\alpha}(t) \rightarrow \text{Concentration of } I \text{ with ligand bound at the } \alpha \text{ site as a function of time} \]
\[ I_{\beta}(t) \rightarrow \text{Concentration of } I \text{ with ligand bound at the } \beta \text{ site as a function of time} \]
\[ I_{\alpha\beta}(t) \rightarrow \text{Concentration of } I \text{ with ligand bound at the } \alpha \text{ and } \beta \text{ sites as a function of time} \]
\[ F(t) \rightarrow \text{Concentration of } F \text{ as a function of time} \]
\[ F_{\alpha}(t) \rightarrow \text{Concentration of } F \text{ with ligand bound at the } \alpha \text{ site as a function of time} \]
\[ F_{\beta}(t) \rightarrow \text{Concentration of } F \text{ with ligand bound at the } \beta \text{ site as a function of time} \]
\[ F_{\alpha\beta}(t) \rightarrow \text{Concentration of } F \text{ with ligand bound at the } \alpha \text{ and } \beta \text{ sites as a function of time} \]
The following parameters represent the first derivatives of the concentrations with respect to time:

- \( U'(t) \)
- \( U_{\alpha}'(t) \)
- \( U_{\beta}'(t) \)
- \( U_{\alpha\beta}'(t) \)
- \( I'(t) \)
- \( I_{\alpha}'(t) \)
- \( I_{\beta}'(t) \)
- \( I_{\alpha\beta}'(t) \)
- \( F'(t) \)
- \( F_{\alpha}'(t) \)
- \( F_{\beta}'(t) \)
- \( F_{\alpha\beta}'(t) \)

Signals and fractions of \( U, I \) & \( F \):

- \( \text{SigU} \rightarrow \text{Signal of } U \)
- \( \text{SigI} \rightarrow \text{Signal of } I \)
- \( \text{SigF} \rightarrow \text{Signal of } F \)
- \( \text{SigUBound} \rightarrow \text{Signal of } U \text{ with ligand bound} \)
- \( \text{SigIBound} \rightarrow \text{Signal of } I \text{ with ligand bound} \)
- \( \text{SigFBound} \rightarrow \text{Signal of } F \text{ with ligand bound} \)
- \( \text{FuFree} \rightarrow \text{Fraction of protein as } U \)
- \( \text{FuBound} \rightarrow \text{Fraction of protein as } U \text{ with ligand bound} \)
- \( \text{FiFree} \rightarrow \text{Fraction of protein as } I \)
- \( \text{FiBound} \rightarrow \text{Fraction of protein as } I \text{ with ligand bound} \)
- \( \text{FfFree} \rightarrow \text{Fraction of protein as } F \)
- \( \text{FfBound} \rightarrow \text{Fraction of protein as } F \text{ with ligand bound} \)

Kinetic and Thermodynamic Relationships:

\[ K_{uf} = K_{ui} K_{if} \]

On rates:

\[ k_{af \alpha} = k_{af \beta} = k_{ai \alpha} = k_{ai \beta} = k_{au \alpha} = k_{au \beta} = 1 \times 10^8 \text{ M}^{-1} \text{ s}^{-1} \]

List of off rate constants \( k_{dxyz} \) where \( x \) is the conformation and \( z \) is the site \( \alpha \) or \( \beta \):

- \( k_{d\alpha \alpha} = k_{au \alpha} / K_{u\alpha} \)
- \( k_{d\alpha \beta} = k_{au \beta} / K_{u\beta} \)
- \( k_{d\beta \alpha} = k_{ai \alpha} / K_{a\alpha} \)
- \( k_{d\beta \beta} = k_{ai \beta} / K_{a\beta} \)
- \( k_{df \alpha} = k_{af \alpha} / K_{a\alpha} \)
- \( k_{df \beta} = k_{af \beta} / K_{a\beta} \)
List of unfolding rate constants in all different liganding states:

\( \text{k}_{fi} = \frac{k_{if}}{K_{if}} \)
\( \text{k}_{fi}^\alpha = \frac{(k_{if}^\alpha K_{i\alpha})}{(K_{if} K_{af})} \)
\( \text{k}_{fi}^\beta = \frac{(k_{if}^\beta K_{i\beta})}{(K_{if} K_{bf})} \)
\( \text{k}_{fi}^{\alpha\beta} = \frac{(k_{if}^{\alpha\beta} K_{i\alpha} K_{\alpha i})}{(K_{if} K_{af} K_{bf})} \)
\( \text{k}_{iu} = \frac{k_{ui}}{K_{ui}} \)
\( \text{k}_{iu}^\alpha = \frac{(k_{ui}^\alpha K_{i\alpha})}{(K_{ui} K_{ai})} \)
\( \text{k}_{iu}^\beta = \frac{(k_{ui}^\beta K_{i\beta})}{(K_{ui} K_{bi})} \)
\( \text{k}_{iu}^{\alpha\beta} = \frac{(k_{ui}^{\alpha\beta} K_{i\alpha} K_{\alpha i})}{(K_{ui} K_{ai} K_{bi})} \)
\( \text{k}_{fu} = \frac{k_{uf}}{K_{uf}} \)
\( \text{k}_{fu}^\alpha = \frac{(k_{uf}^\alpha K_{u\alpha})}{(K_{uf} K_{af})} \)
\( \text{k}_{fu}^\beta = \frac{(k_{uf}^\beta K_{u\beta})}{(K_{uf} K_{bf})} \)
\( \text{k}_{fu}^{\alpha\beta} = \frac{(k_{uf}^{\alpha\beta} K_{u\alpha} K_{\alpha u})}{(K_{uf} K_{af} K_{bf})} \)
Free ligand concentration in terms of concentrations of total ligand and bound protein:

\[ L_f = (L_t - (U\alpha (t)+U\beta (t)+2\ U\alpha\beta (t)+I\alpha (t)+I\beta (t)+2\ I\alpha\beta (t)+F\alpha (t)+F\beta (t)+2\ F\alpha\beta (T))) \]

Differential Equations for Rate Laws Describing P Protein Coupled Binding and Folding:

\[ U' (t)= -kui U (t)+ kiu I (t)- kau\alpha U (t) Lf - kau\beta U (t) Lf + kdu\alpha U\alpha (t)+ kdu\beta U\beta (t)- kuf U (t)+ kfu F (T) \]

\[ U\alpha' (t)= -kdu\alpha U\alpha (t)+ kau\alpha U (t) Lf - kui\alpha U\alpha (t)+ kiu\alpha I\alpha (t)- kau\beta U\alpha (t) Lf + kdu\beta U\alpha\beta (t)- kuf\alpha U\alpha (t)+ kfu\alpha F\alpha (T) \]

\[ U\beta' (t)= -kdu\beta U\beta (t)+ kau\beta U (t) Lf - kui\beta U\beta (t)+ kiu\beta I\beta (t)- kau\alpha U\beta (t) Lf + kdu\alpha U\beta\beta (t)- kuf\beta U\beta (t)+ kfu\beta F\beta (T) \]

\[ U\alpha\beta' (t)= -kdu\alpha U\alpha\beta (t)- kdu\beta U\alpha\beta (t)+ kau\alpha U\beta (t) Lf + kau\beta U\alpha (t) Lf - kui\alpha\beta U\alpha\beta (t)+ kiu\alpha\beta I\alpha\beta (t)- kuf\alpha\beta U\alpha\beta (t)+ kfu\alpha\beta F\alpha\beta (T) \]

\[ I' (t)= -kii I (t)- kif I (t)+ kfi F (t)- kai\alpha I (t) Lf - kai\beta I (t) Lf + kdi\alpha I\alpha (t)+ kdi\beta I\beta (t) \]

\[ I\alpha' (t)= -kiu\alpha I\alpha (t)- kif\alpha I\alpha (t)+ kui\alpha U\alpha (t)+ kfi\alpha F\alpha (t)- kdi\alpha I\alpha (t)+ kai\alpha I (t) Lf - kai\beta I\alpha (t) Lf + kdi\beta I\alpha\beta (t) \]

\[ I\beta' (t)= -kiu\beta I\beta (t)- kif\beta I\beta (t)+ kui\beta U\beta (t)+ kfi\beta F\beta (t)- kdi\beta I\beta (t)+ kai\beta I\beta (t) Lf - kai\alpha I\beta (t) Lf + kdi\alpha I\alpha\beta (t) \]

\[ I\alpha\beta' (t)= -kiu\alpha\beta I\alpha\beta (t)- kif\alpha\beta I\alpha\beta (t)+ kui\alpha\beta U\alpha\beta (t)+ kfi\alpha\beta F\alpha\beta (t)- kdi\alpha I\alpha\beta (t)- kdi\beta I\alpha\beta (t)+ kai\alpha I\beta (t)+ kai\beta I\alpha (t) \]

\[ F' (t)= -kfi F (t)+ kif F (t)- kaf\alpha F (t) Lf - kaf\beta F (t) Lf + kdf\alpha F\alpha (t)+ kdf\beta F\beta (t)+ kfu F (T) \]

\[ F\alpha' (t)= -kdf\alpha F\alpha (t)+ kaf\alpha F (t) Lf - kfi\alpha F\alpha (t)+ kif\alpha I\alpha (t)- kaf\beta F\alpha (t) Lf + kdf\beta F\beta (t)+ kfu\alpha F\alpha (t) \]

\[ F\beta' (t)= -kdf\beta F\beta (t)+ kaf\beta F (t) Lf - kfi\beta F\beta (t)+ kif\beta I\beta (t)- kaf\alpha F\beta (t) Lf + kdf\alpha F\alpha (t)+ kfu\beta F\beta (T) \]

\[ F\alpha\beta' (t)= -kdf\alpha F\alpha\beta (t)- kdf\beta F\alpha\beta (t)+ kaf\alpha F\beta (t) Lf + kaf\beta F\alpha (t) Lf - kfi\alpha\beta F\alpha\beta (t)+ kfi\alpha\beta I\alpha\beta (t)+ kuf\alpha\beta U\alpha\beta (t)- kfu\alpha\beta F\alpha\beta (T) \]
Initial Conditions:
\[ U(0) = \frac{Pt}{(1 + KuI + Kif)} \]
\[ U(0) + U\alpha(0) + U\beta(0) + U\alpha\beta(0) + I(0) + I\alpha(0) + I\beta(0) + I\alpha\beta(0) + F(0) + F\alpha(0) + F\beta(0) + F\alpha\beta(0) = Pt \]
\[ U\beta(0) = 0 \]
\[ U\alpha\beta(0) = 0 \]
\[ I(0) = \frac{(Kui Pt)}{(1 + KuI + Kif)} \]
\[ I\alpha(0) = 0 \]
\[ I\beta(0) = 0 \]
\[ I\alpha\beta(0) = 0 \]
\[ F(0) = \frac{(Kui Kif Pt)}{(1 + KuI + Kif)} \]
\[ F\alpha(0) = 0 \]
\[ F\beta(0) = 0 \]
\[ F\alpha\beta(0) = 0 \]

For PPi and Pi:
\[ \text{Signal Observed} = (\text{SigU FuFree} + \text{SigUBound FuBound} + \text{SigI FiFree} + \text{SigI FiBound} + \text{SigF FfFree} + \text{SigF FfBound}) \]

For ADP, GDP, UDP, and CDP:
\[ \text{Signal Observed} = (\text{SigU FuFree} + \text{SigUBound FuBound} + \text{SigI FiFree} + \text{SigI FiBound} + \text{SigF FfFree} + \text{SigF FfBound}) - (ifeLt) \]

The ife parameter accounts for the inner filter effect caused by the nucleotide absorbance of excitation light.
Appendix B

Mathematical Description of RNA Biosensor Activation

\( N_t \rightarrow \) total cyclic dinucleotide concentration
\( D_t \rightarrow \) total DFHBI concentration
\( R_t \rightarrow \) total RNA concentration
\( N_f \rightarrow \) free cyclic dinucleotide concentration
\( D_f \rightarrow \) free DFHBI concentration
\( t \rightarrow \) time

\( K_{12} \rightarrow \) equilibrium constant for the C1 to C2 reaction

\( k_{12} \rightarrow \) rate constant for the C1 to C2 transition
\( k_{21} \rightarrow \) rate constant for the C2 to C1 transition

\( k_{12N} \rightarrow \) rate constant for the C1 to C2 transition with cyclic dinucleotide bound
\( k_{12D} \rightarrow \) rate constant for the C1 to C2 transition with DFHBI bound
\( k_{12ND} \rightarrow \) rate constant for the C1 to C2 transition with cyclic dinucleotide and DFHBI bound
\( k_{21N} \rightarrow \) rate constant for the C2 to C1 transition with cyclic dinucleotide bound
\( k_{21D} \rightarrow \) rate constant for the C2 to C1 transition with DFHBI bound
\( k_{21ND} \rightarrow \) rate constant for the C2 to C1 transition with cyclic dinucleotide and DFHBI bound
\( K_{A1N} \rightarrow \) Binding constant for cyclic dinucleotide binding to C1
\( K_{A1D} \rightarrow \) Binding constant for DFHBI binding to C1
\( K_{A2N} \rightarrow \) Binding constant for cyclic dinucleotide binding to C2
\( K_{A2D} \rightarrow \) Binding constant for DFHBI binding to C2

\( k_{a1N} \rightarrow \) Association rate constant for cyclic dinucleotide binding to C1
\( k_{a1D} \rightarrow \) Association rate constant for DFHBI binding to C1
\( k_{a2N} \rightarrow \) Association rate constant for cyclic dinucleotide binding to C2
\( k_{a2D} \rightarrow \) Association rate constant for DFHBI binding to C2

\( k_{d1N} \rightarrow \) Dissociation rate constant for cyclic dinucleotide dissociating from C1
\( k_{d1D} \rightarrow \) Dissociation rate constant for DFHBI dissociating from C1
\( k_{d2N} \rightarrow \) Dissociation rate constant for cyclic dinucleotide dissociating from C2
\( k_{d2D} \rightarrow \) Dissociation rate constant for DFHBI dissociating from C2

\( C1(t) \rightarrow \) Concentration of C1 as a function of time
\( C1N(t) \rightarrow \) Concentration of C1 with cyclic dinucleotide bound as a function of time
\( C1D(t) \rightarrow \) Concentration of C1 with DFHBI bound as a function of time
\( C1ND(t) \rightarrow \) Concentration of C1 with cyclic dinucleotide and DFHBI bound as a function of time
\( C2(t) \rightarrow \) Concentration of C2 as a function of time
\( C2N(t) \rightarrow \) Concentration of C2 with cyclic dinucleotide bound as a function of time
\( C2D(t) \rightarrow \) Concentration of C2 with DFHBI bound as a function of time
\( C2ND(t) \rightarrow \) Concentration of C2 with cyclic dinucleotide and DFHBI bound as a function of time
The following parameters represent the first derivatives of the concentrations with respect to time:
C1' (t)
C1N' (t)
C1D' (t)
C1ND' (t)
C2' (t)
C2N' (t)
C2D' (t)
C2ND' (t)

Kinetic and Thermodynamic Relationships:

List of dissociation rate constants written in terms of association rate constants and binding constants:

\[ k_{d1N} = \frac{k_{a1N}}{K_{A1N}} \]
\[ k_{d1D} = \frac{k_{a1D}}{K_{A1D}} \]
\[ k_{d2N} = \frac{k_{a2N}}{K_{A2N}} \]
\[ k_{d2D} = \frac{k_{a2D}}{K_{A2D}} \]

List of C2 to C1 rate constants written in terms for C1 to C2 rate constants, C1 to C2 equilibrium constants, and binding constants:

\[ k_{21} = \frac{k_{12ND}}{K_{12}} \]
\[ k_{21N} = \frac{(k_{12ND} K_{A1N})}{(K_{12} K_{A2N})} \]
\[ k_{21D} = \frac{(k_{12ND} K_{A1D})}{(K_{12} K_{A2D})} \]
\[ k_{21ND} = \frac{(k_{12ND} K_{A1N} K_{A1D})}{(K_{12} K_{A2N} K_{A2D})} \]
Free ligand concentration in terms of concentrations of total ligand and bound RNA:

\[ \text{Nf} = (\text{Nt} - (C_{1N}(t) + C_{1ND}(t) + C_{2N}(t) + C_{2ND}(t))) \]

\[ \text{Df} = (\text{Dt} - (C_{1D}(t) + C_{1ND}(t) + C_{2D}(t) + C_{2ND}(t))) \]

Differential Equations for Rate Laws Describing RNA Coupled Binding and Conformational Change:

\[
\begin{align*}
\text{C}_{1}'(t) &= -k_{12} \text{C}_1(t) + k_{21} \text{C}_2(t) - k_{a1N} \text{C}_1(t) \text{Nf}(t) - k_{d1N} \text{C}_1(t) \text{Nf}(t) + k_{d1N} \text{C}_1(t) + k_{d1D} \text{C}_1(t) \\
\text{C}_{1N}'(t) &= -k_{12N} \text{C}_1(t) + k_{21N} \text{C}_2(t) + k_{a1N} \text{C}_1(t) \text{Nf}(t) - k_{d1N} \text{C}_1(t) + k_{d1D} \text{C}_1(t) - k_{a1N} \text{C}_1(t) \text{Df}(t) + k_{d1N} \text{C}_1(t) + k_{d1D} \text{C}_1(t) \\
\text{C}_{1D}'(t) &= -k_{12D} \text{C}_1(t) + k_{21D} \text{C}_2(t) + k_{a1D} \text{C}_1(t) \text{Df}(t) - k_{d1D} \text{C}_1(t) + k_{d1D} \text{C}_1(t) - k_{a1D} \text{C}_1(t) + k_{d1D} \text{C}_1(t) \text{Nf}(t) + k_{d1D} \text{C}_1(t) + k_{d1D} \text{C}_1(t) \\
\text{C}_{1ND}'(t) &= -k_{12ND} \text{C}_1(t) + k_{21ND} \text{C}_2(t) - k_{d1D} \text{C}_1(t) \text{Df}(t) - k_{d1D} \text{C}_1(t) + k_{d1D} \text{C}_1(t) - k_{a1N} \text{C}_1(t) \text{Nf}(t) + k_{a1D} \text{C}_1(t) \text{Df}(t) + k_{a1D} \text{C}_1(t) + k_{a1D} \text{C}_1(t) \\
\text{C}_{2}'(t) &= k_{12} \text{C}_2(t) - k_{21} \text{C}_1(t) + k_{21N} \text{C}_2(t) + k_{21D} \text{C}_2(t) - k_{a2N} \text{C}_2(t) \text{Nf}(t) - k_{a2D} \text{C}_2(t) \text{Df}(t) + k_{d2N} \text{C}_2(t) + k_{d2D} \text{C}_2(t) \\
\text{C}_{2N}'(t) &= k_{12N} \text{C}_1(t) - k_{21N} \text{C}_2(t) + k_{a2N} \text{C}_2(t) \text{Nf}(t) - k_{d2N} \text{C}_2(t) + k_{d2D} \text{C}_2(t) - k_{a2D} \text{C}_2(t) \text{Df}(t) + k_{d2D} \text{C}_2(t) + k_{d2D} \text{C}_2(t) \\
\text{C}_{2D}'(t) &= k_{12D} \text{C}_1(t) - k_{21D} \text{C}_1(t) + k_{a2D} \text{C}_2(t) \text{Df}(t) + k_{a2D} \text{C}_2(t) \text{Df}(t) + k_{a2D} \text{C}_2(t) \text{Df}(t) - k_{d2D} \text{C}_2(t) + k_{d2D} \text{C}_2(t) - k_{a2D} \text{C}_2(t) \text{Df}(t) + k_{d2D} \text{C}_2(t) + k_{d2D} \text{C}_2(t) \\
\text{C}_{2ND}'(t) &= k_{12ND} \text{C}_1(t) - k_{21ND} \text{C}_2(t) - k_{d2D} \text{C}_2(t) - k_{d2D} \text{C}_2(t) + k_{a2N} \text{C}_2(t) \text{Nf}(t) + k_{a2D} \text{C}_2(t) \text{Df}(t) + k_{a2D} \text{C}_2(t) \text{Df}(t) + k_{a2D} \text{C}_2(t) \text{Df}(t)
\end{align*}
\]

Initial Conditions:

\[
\begin{align*}
\text{C}_1(0) &= \frac{\text{R}t}{(1 + K_{12})} \\
\text{C}_1N(0) &= 0 \\
\text{C}_1D(0) &= 0 \\
\text{C}_1ND(0) &= 0 \\
\text{C}_2(0) &= \frac{(K_{12} \text{R}t)}{(1 + K_{12})} \\
\text{C}_2N(0) &= 0 \\
\text{C}_2D(0) &= 0 \\
\text{C}_2ND(0) &= 0
\end{align*}
\]

Signal Observed = \((\text{SigDf} \text{Df}(t) + \text{SigD1} (\text{C1D}(t)+\text{C1ND}(t)) + \text{SigD2} (\text{C2D}(t)+\text{C2ND}(t)))\) where \(\text{SigDf}, \text{SigD1},\) and \(\text{SigD2}\) are fluorescence parameters.
References


Liu X (2013) Molecular Recognition of Inhibitors, Metal Ions and Substrates by Ribonuclease P. Ph.D. Dissertation (University of Michigan, Ann Arbor).


Biography
Kyle Gabriel Daniels was born on March 18, 1988 in Washington, D.C. and raised in nearby Prince George’s County Maryland. He attended the University of Maryland College Park where he worked as a Howard Hughes Medical Institute Undergraduate Research Fellow in the lab of Dr. Dorothy Beckett and graduated with a Bachelor of Science in Biochemistry in May of 2010. He began his doctoral work under the guidance of Terrence Oas at Duke University in August of 2010. While at Duke, Kyle received the National Science Foundation Graduate Research Fellowship and the Duke Dean’s Fellowship.

Publications:


Mosley PA, Daniels KG, Oas TG. (submitted to Biochemistry) “Electrostatic Energetics of Bacillus subtilis Ribonuclease P Protein by NMR-based Histidine pKa Measurements”.

