8-2016

Investigation of an energetic coupling between ligand binding and protein folding

Nathan W. Gardner
Purdue University

Graduate School Form
30 Updated 12/26/2015

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Nathan Gardner

Entitled

INVESTIGATION OF AN ENERGETIC COUPLING BETWEEN LIGAND BINDING AND PROTEIN FOLDING

For the degree of Doctor of Philosophy

Is approved by the final examining committee:

Chiwook Park
Chair

Daisuke Kihara

Carol B. Post

Elizabeth M. Topp

To the best of my knowledge and as understood by the student in the Thesis/Dissertation Agreement, Publication Delay, and Certification Disclaimer (Graduate School Form 32), this thesis/dissertation adheres to the provisions of Purdue University's "Policy of Integrity in Research" and the use of copyright material.

Approved by Major Professor(s): Chiwook Park

Approved by: Christine A. Hrycyna 7/6/2016

Head of the Departmental Graduate Program Date

INVESTIGATION OF AN ENERGETIC COUPLING BETWEEN LIGAND BINDING AND PROTEIN FOLDING

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Nathan W. Gardner

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

August 2016

Purdue University

West Lafayette, Indiana

To my supportive parents, who have enabled me to travel this far, and to my joyful son, who believes that I can conquer anything.

ACKNOWLEDGEMENTS

I would like to begin by thanking my advisor, Dr. Chiwook Park, whose guidance over the years has taught me how think about science and enabled me to transition from the field of physics to protein biochemistry. I especially want to thank Dr. Park as a mentor for allowing me to speak openly with him and for taking an interest in my well-being. I also want to thank my collaborators Dr. Daisuke Kihara, Lyman Monroe, Dr. Elizabeth Topp, and Dr. Jainik Panchal, without whom I could not have investigated my thesis work as deeply. A special thanks is due to Sarah McGuiness, an undergraduate researcher whose attention to detail and hard work were invaluable to me in running experiments for Chapters 3 and 4. Thank you also to Dr. Nathalie Declerck (Centre de Biochimie Structural, Montpellier, France), whose insightful suggestions upon reviewing our manuscript greatly steered my interpretation of our data. I also thank want to thank Chen Chen for her friendship and her useful comments and help in editing over the years. Thank you to John Schlebach, Joseph Kasper, and Mark Hinzman for their friendship and for inspiring me to pursue protein science. Lastly, I want to thank my parents, Scott and Cheryl Gardner for their unyielding support through the thick and thin of my academic endeavors.

TABLE OF CONTENTS

Page

Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

LIST OF ABBREVIATIONS

dPGM	*E. coli* cofactor-dependent phosphoglycerate mutase
AMP,	Adenosine monophosphate
ADP	Adenosine diphosphate
ATP	Adenosine triphosphate
ATPγS	Adenosine 5'-[γ-thio]triphosphate
CD	Circular dichroism
C_m	Midpoint of the unfolding transition
GTP	Guanosine triphosphate
HDX-MS	Hydrogen/deuterium exchange coupled with mass spectroscopy
IWAW	Intensity-weighted average wavelength
ITC	Isothermal titration calorimetry
$\Delta G°$	Gibbs free energy
$\Delta\Delta G°$	Change in the Gibbs free energy
SDS	Sodium dodecyl sulfate
PAGE	Polyacrylamide gel electrophoresis
$BS(PEG)_5$	Bis(succinimidyl) penta(ethylene glycol) crosslinker
EDTA	Ethylenediaminetetraacetic acid

ABSTRACT

Gardner, Nathan W. Ph.D., Purdue University, August 2016. Investigation of an Energetic Coupling Between Ligand Binding and Protein Folding. Major Professor: Chiwook Park.

The cellular environment presents a protein with many small molecules with which it may interact. Many novel interactions between proteins and non-substrate metabolites are being uncovered through proteome-wide screens. The homodimeric *Escherichia coli* cofactor-dependant phosphoglycerate mutase (dPGM) was identified as an ATP binding protein in a proteome-wide screen, but dPGM does not use ATP for catalysis. This dissertation elucidates the effect of ATP and other non-substrate metabolites on dPGM. Initial investigations revealed a partially unfolded, monomeric intermediate of dPGM that forms during equilibrium unfolding. ATP binding was found to occur at the active site of dPGM and to be energetically coupled with dimerization; ligand binding events reduce the population of intermediate. An investigation into the structure of the dPGM intermediate revealed a cooperative folding unit that couples the active site and dimer interface of dPGM. By coupling the two binding sites, the cooperative unit is responsible for conveying the allosteric effect observed between dimerization and ligand binding. We found that physiological salts reduce but do not prevent non-substrate metabolite binding at physiological concentrations. Further, anions bind specifically to dPGM and chloride

was found to bind to both of the energetically coupled sites on dPGM, the active site and dimer interface. Our findings illustrate how a cooperative link between a ligand binding site and oligomer interface can promote higher order oligomers and reduce intermediate populations. The physiological effect of the cooperative link and ligand binding to dPGM is a large enhancement in the stability of the dimer over the monomer intermediate and, possibly, competitive inhibition.

CHAPTER 1. INTRODUCTION

1.1 The thermodynamics of protein folding

Proteins are essential biopolymers that carryout chemical reactions, conduct chemical signaling, and provide structural support to the cell. Proteins are dynamic molecules and their function typically requires folding to the proper three dimensional conformations. In the early 1960s, pioneering work from by Christian B. Anfinsen and colleagues investigated the reversible folding of ribonuclease and poly-DL-alanyl trypsin and formed the basis of the thermodynamic hypothesis (Anfinsen, Haber, Sela, & White, 1961; Epstein & Anfinsen, 1962). The thermodynamic hypothesis is the idea that protein folding is governed by thermodynamics, and that proteins fold to the lowest energy conformation (Anfinsen, 1973). Levinthal showed that proteins must fold through certain routes due to kinetic barriers because a random search for the native state is not feasible given vast number of possible conformations (Levinthal, 1968).

Currently, protein folding is viewed as movement of the protein along the surface of a three-dimensional, funnel-shaped energy landscape (Dill & Chan, 1997). Every point along the funneled energy landscape represents a protein conformation. The vertical axis (z-axis) of the funnel corresponds to the internal energy of a conformation; the bottom of the funnel typically represents the low-energy, native state and the top of the funnel

represents the high-energy, globally unfolded form. The lateral axes, in the X−Y plane, define different conformational coordinates of the protein and reflect the numerous degrees of freedom of the protein chain (Dill & Chan, 1997). Protein folding is similar to a ball rolling down the funneled energy landscape; the protein will move to lower energy conformations and travel via the lowest energy routes. Protein are jostled about the energy landscape by Brownian motion so once the protein reaches the bottom of the funnel, the native state, the protein will continue to undergo some conformational fluctuation. The energy landscape can be smooth, which indicates a protein will fold cooperatively from the unfolded to native forms with no detectable intermediate; an energy landscape can also be rough and contain small wells, local minima, that represent low-energy intermediate conformations (Dill & Chan, 1997). Even under native conditions a protein exists as a conformational ensemble with the population of each conformation determined by the energy landscape (Dill & Chan, 1997; Frauenfelder, Sligar, & Wolynes, 1991; Hilser, Garcia-Moreno, Oas, Kapp, & Whitten, 2006). Conformational fluctuations manifest themselves as dynamic motions on the picosecond to millisecond timescales and are often necessary for the biochemical function of a protein (Hammes, 2002; Henzler-Wildman & Kern, 2007). The distribution of conformations in the ensemble of a protein gives rise to the macroscopic properties observed by traditional biophysical probes such as spectroscopy or calorimetry.

1.2 Cooperativity and allostery in protein

Although the energy landscape of a protein encodes a vast number of conformations, in practice, only a few conformations are adopted during folding (Englander, 2000; Hu et al.,

2013). Residues throughout a protein may be energetically coupled so that a perturbation to one, affects the other (Hilser, Dowdy, Oas, & Freire, 1998; T. Liu, Whitten, & Hilser, 2007). This coupling between residues can lead to them to fold and unfold in concert (positive cooperativity), or one residue may unfold when the other folds (negative cooperativity). The cooperative folding of proteins is an emergent property that arises from the energetic coupling between residues (Ernesto. Freire, Murphy, Sanchez-Ruiz, Galisteo, & Privalov, 1992; Gunasekaran, Ma, & Nussinov, 2004; T. Liu et al., 2007). Coupling between residues shapes the energy landscape and causes many proteins to fold as a hierarchical assembly of individual cooperative units (Bai, Sosnick, Mayne, & Englander, 1995; Chamberlain, Handel, & Marqusee, 1996; Hu et al., 2013). Folding cooperativity is important for reducing non-functional intermediate conformations that may misfold (Dobson, 2003) or aggregate (Horwich, 2002). Mutations in a protein can reshape the energy landscape and make intermediate forms more favorable or alter the folding pathway, ultimately promoting disease. The ΔF08 mutation in the cystic fibrosis receptor (CFTR) is a seminal example of how a single mutation can severely impair folding; with the ΔF08 mutation, CFTR does not fold to the native form and is degraded in the cytoplasm, giving rise to cystic fibrosis (Lukacs & Verkman, 2012).

Whether proteins truly fold through multiple pathways, as implied by the funnel-model (Dill & Chan, 1997; Dill, Ozkan, Shell, & Weikl, 2008) is debatable. Experiments that tend to support multiple folding routes (Mello & Barrick, 2004; Sridevi, Lakshmikanth, Krishnamoorthy, & Udgaonkar, 2004; Zhong, Rousseau, & Yeh, 2004) may be explained by a portion of the protein population misfolding before folding along a discrete pathway

(Englander, Mayne, & Krishna, 2007; Hu, Kan, Mayne, & Englander, 2016). The energy landscape of the Notch ankyrin repeat domain was experimentally mapped using a series of truncated mutants and appeared to show multiple low-energy folding routes (Mello & Barrick, 2004). When folding of the full length Notch ankyrin repeat domain was investigated, the protein was found to fold via one discrete pathway (Bradley & Barrick, 2006). Ribonuclease H (Hu et al., 2013) and Cytochrome c (Hu et al., 2016) also fold through one discrete pathway in which the cooperative subunits of the protein fold one at a time, and in the same order. The sequential folding of the subunits that make up Ribonuclease H and Cytochrome c was observed at near amino acid resolution by hydrogen/deuterium exchange coupled with mass spectrometry. From the proteins that were studied at an adequate level of detail thus far, protein appear to have evolved so the folding of the initial cooperative unit stabilizes the next cooperative unit and so on, which naturally gives rise to a linear folding pathway. However, these proteins are also small and are likely to fold well in comparison to much of the proteome. In fact, model proteins used for folding studies are highly biased towards well-folding, monomeric protein compared to the distribution of protein in the *E. coli* proteome (Braselmann, Chaney, & Clark, 2013).

The energetic coupling between residues in a protein is also responsible for allosteric regulation of a protein (Lockless & Ranganathan, 1999). Allostery is an fundamental feature of protein in which two binding sites, distal to one another, are energetically coupled so that binding at one site influences binding at the other (Gunasekaran et al., 2004; Hilser, Wrabl, & Motlagh, 2012). An allosteric response occurs when ligand

binding remodels a protein's energy landscape, causing a redistribution of the conformational ensemble that affects binding at a second site (Smock & Gierasch, 2009; Weber, 1972). In this way, binding of one ligand can promote conformations with greater or lesser affinity for the second ligand. Here, the ligand may also be another protein rather than a small molecule. For example, the dopamine and somatostatin receptors have been shown to form a heterodimer that exhibits greatly enhanced activity (Rocheville et al., 2000). The allosteric behavior of the dopamine and somatostatin receptors is driven by the protein−protein interaction. Allosteric coupling is crucial for control of molecular processes such as protein signaling (Chennubhotla, Yang, & Bahar, 2008; Ernesto Freire, 1999; Harford & Weitzman, 1975; Hynes, 2002; Yu et al., 2010) and autoinhibition (Yu et al., 2010). Non-allosteric protein have been found to have hidden allosteric sites which suggests that all dynamic protein are allosteric (Gunasekaran et al., 2004). Then, understanding allosteric effects and cooperativity within protein will facilitate an understanding of typical protein behavior.

The use of allosteric effects is a promising avenue for drug development (Berger et al., 1999; Kalid et al., 2010; Kenakin, 2002). Synthetic antibodies have been evolved through phage display to trap protein in specific conformations and could be used to promote active or inactive conformations of a protein without binding to the active site (Paduch et al., 2013). Pharmacological chaperones are small molecule drugs that either facilitate the proper folding of a protein or stabilize the protein in its folded form (Ringe & Petsko, 2009). Gaucher disease is a lysosomal storage disorder that is triggered by mutations in acid β-glucosidase and often by the mutation N370S. Acid β-glucosidase with the N370S

mutation has impaired folding and trafficking. The small molecule isofagomine is able to bind in the active site of N370S acid β-glucosidase and shift the protein ensemble to the active conformation, which restores proper cellular trafficking (Lieberman et al., 2007). Efforts are underway to develop a pharmacological chaperone for ΔF508 CFTR (Kalid et al., 2010; Robert et al., 2010; Van Goor et al., 2006); however, the chaperones are not yet effective at restoring CFTR activity (Lukacs & Verkman, 2012). Deciphering the basis of allostery and cooperativity within proteins is a key challenge for the understanding and prediction of protein behavior and the treatment of disease.

1.3 Energetics based proteomics

Ligand binding shifts a protein ensemble to binding competent forms. Because ligand binding enhances the conformational stability of the binding competent forms, the energetics of ligand binding can be determined from changes in protein stability (Sanchez-Ruiz, 2007; Schellman, 1976; Waldron & Murphy, 2003). Often the stabilizing effect of ligand binding protects protein against proteolytic degradation (Pace & Barrett, 1984; Sieber, Pluckthun, & Schmid, 1998; Walter et al., 1999). Either partial or global unfolding of a protein is required for a protease to cleave the polypeptide chain (Park & Marqusee, 2004b). By utilizing the change in the proteolytic susceptibility of a protein upon unfolding, Park and Marqusee developed the technique, pulse proteolysis to determine the global stability of proteins (Park & Marqusee, 2005). For pulse proteolysis, a high concentration of a non-specific protease is used to selectively digest unfolded protein in one short pulse of digestion. A short pulse length is essential to the technique because it keeps the protein from re-equilibrating during the course of the proteolysis

reaction. In this way, the unfolded proteins at a set point in time are selectively digested and the folded protein may be visualized (Park & Marqusee, 2005). Because proteins do not re-equilibrate during the proteolytic digestion, pulse proteolysis is analogous to taking a photograph of the folding protein population. When pulse proteolysis is conducted with a series of samples that span a range of urea concentrations, the amount of folded protein at each urea concentration is determined. From the dependence of the folded protein population on urea, the global stability can be determined (Greene & Pace, 1974; Pace, 1986). Through determination of the global stability of a protein in the absence and presence of a ligand, pulse proteolysis can report on changes in global stability incurred by ligand binding (Park & Marqusee, 2005). Pulse proteolysis is also extremely useful for monitoring protein stability in complex samples. When using pulse proteolysis, proteins can be separated after the reaction by techniques such as SDS-PAGE, eliminating the need for protein purification in order to quantify global stability. Through the use of western blotting, the global stability of human H-ras was compared and found to be the same in cell lysate from *Escherichia coli* and Jurkat T cells by pulse proteolysis (Kim, Song, & Park, 2009).

Unlike traditional methods based on spectroscopy or calorimetry, pulse proteolysis permits the study of protein unfolding for many proteins simultaneously on a proteomic scale (Adhikari & Fitzgerald, 2014; Chang, Schlebach, VerHeul, & Park, 2012; P.-F. Liu, Kihara, & Park, 2011). The ability to add protease to any sample, followed by gel-based visualization makes pulse proteolysis a simple tool to implement on a proteomic scale. Pulse proteolysis was first employed in a proteome-wide screen to screen for ATP

binding protein (P.-F. Liu et al., 2011). To identify ATP binding protein, samples of *E. coli* lysate were prepared in 3M urea in the presence and absence of 1.0 mM ATPγS. The concentration of free ATP in *E. coli* is thought be about 1 mM (Schneider & Gourse, 2004). Pulse proteolysis was carried out on each sample and the intact protein was separated and visualized by 2D-gel electrophoresis. Here, determination of global stability is not required to identify proteins that interact with ATP. A large difference in band intensity for the same protein in the absence and presence of ATP signifies that the protein interacts with ATP (Chang et al., 2012; P.-F. Liu et al., 2011). Ten ATP-binding proteins, including four novel ATP-binding proteins were identified (P.-F. Liu et al., 2011). To facilitate better visualization of protein and expand target-identification capabilities, the *E. coli* proteome was again probed for ATP binding but after the lysate was fractionated by ion exchange chromatography so proteins could be visualized by 1D SDS-PAGE (Chang et al., 2012). In this screen, the proteome was assayed by pulse proteolysis at multiple urea concentrations in the absence and presence of 1.0 mM ATPγS, which allowed the detection of additional ATP-binding proteins; thirty ATP-binding proteins, including nine novel ATP-binding proteins, were identified (Chang et al., 2012). One of the novel ATP-binding proteins was cofactor-dependent phosphoglycerate mutase (dPGM). At 2.0 M urea, dPGM was fully disgusted in the absence of ATP, but a significant population of folded protein remained when 1.0 mM ATPγS was present (Chang et al., 2012).

The denaturant dependence of methionine oxidation in the presence of hydrogen peroxide can also report on changes in protein stability (Strickland et al., 2013). Methionine can

only be oxidized when it is unfolded or solvent exposed so oxidation events should indicate unfolding. Changes in the oxidation rate in the presence of a ligand can be used to determine protein−ligand interactions (West et al., 2010). By combining stable isotope labeling in culture and mass spectrometry, changes in the methionine oxidation of a protein in the presence of a ligand can be determined in complex samples (Tran, Adhikari, & Fitzgerald, 2014). Methionine oxidation of the proteome of *Saccharomyces cerevisiae* was monitored in the presence and absence ATPγS to identify ATP binding protein (Tran et al., 2014). From the proteomic-screen, ATP was found to also stabilize *S. cerevisiae* dPGM. The dPGM from *E. coli* and *S. cerevisiae* both appear to be a novel ATP binding protein, as ATP has no known role in their catalytic activities. Then, ATP could have an allosteric effect upon the dPGMs and a novel binding site.

1.4 Cofactor-dependant phosphoglycerate mutases

Phosphoglycerate mutase proteins are found in two unrelated classes: cofactor-independent and cofactor-dependant. Most vertebrate, yeast, and bacteria possess a cofactor-dependant phosphoglycerate mutase (dPGM) (Fraser, Kvaratskhelia, & White, 1999; Grisolia & Joyce, 1959). The cofactor-dependency of phosphoglycerate mutase refers to the fact that mutase activity of dPGM requires phosphorylation of a conserved histidine (His10, *E. coli* numbering) within the active site by the cofactor 2,3-bisphosphoglycerate (2,3-BPG) (Fraser et al., 1999; Grisolia & Cleland, 1968). This histidine and a conserved catalytic core make dPGM a member of the histidine phosphatase superfamily, which includes enzymes such as fructose-2,6-bisphosphatase, outer membrane protein PhoE, and SixA (Rigden, 2008). In the activating step, dPGM

fulfils its role as a histidine phosphatase (EC 3.1.3.13) and catalyzes the transfer of a phosphoryl group from the cofactor, 2,3-BPG, to the active-site histidine, which yields a molecule of 2-phosphoglycerate or 3-phosphoglycerate. The phosphorylated dPGM is primed for its mutase activity (E.C. 5.4.2.11) and can catalyze the essential isomerization reaction between 2- and 3-phosphoglycerate in the glycolysis and gluconeogenesis pathways. Lastly, dPGM has been reported to function as a synthase (E.C. 5.4.2.4) and form 2,3-BPG in the presence of 1,3-bisphosphoglycerate and 3-phosphoglycerate (Laforet, Butterfield, & Alpers, 1974).

After *E. coli* dPGM was found to bind ATP in the proteomic screen, the stabilization of dPGM by ATP was confirmed by performing pulse proteolysis with purified dPGM (Chang et al., 2012). The *E. coli* dPGM is a homodimer consisting of 249-amino-acid monomers with a molecular weight of 28.4 kDa (Fraser et al., 1999). The crystallographic structures of both active and inactive *E. coli* dPGM have been solved (Bond, White, & Hunter, 2001, 2002). When dPGM is activated, the C-terminal tail folds over the active site entrance and forms interactions with residues along the active-site rim (Bond et al., 2001). Aside from the folding of the C-terminal tail, there are few structural differences between the structures of active and inactive dPGM. The active form of dPGM shows two sulfate ions bound within the active site (Bond et al., 2001), and the inactive form bound tetravanadate in the active site (Bond et al., 2002). The active site of dPGM appears rather accommodating, which raises the question of whether ATP can bind in the active site and whether such binding would be productive. The activity of phosphoglycerate mutases has been extensively studied over five decades (Clarke, Birch,

& Britton, 1974; Fraser et al., 1999; Grisolia & Cleland, 1968; Laforet et al., 1974; Rose & Dube, 1978; Rose & Kaklij, 1984; Vander Heiden et al., 2010; Winn, Watson, Harkins, & Fothergill, 1981). The phosphatase and mutase activities of dPGMs are sensitive to the presence of anions and especially divalent anions, which function as competitive inhibitors (Grisolia & Cleland, 1968; Kowalski, Nocon, Gamian, Kołodziej, & Rakus, 2012; Rose & Dube, 1978; Rose & Kaklij, 1984). ATP, however, has not been observed to play a role in the activity of a dPGM.

1.5 Investigating the energetics and of non-substrate metabolite binding and the folding of *E. coli* phosphoglycerate mutase

ATP was observed to enhance the stability of *E. coli* dPGM at a physiological concentration, which suggests that the interaction may occur in the cytosol. The stabilizing effect of ATP was determined by pulse proteolysis, which cannot discriminate between cleavable intermediate forms and the globally unfolded protein. Also, a protein could undergo conformational change without becoming susceptible to protease, which is invisible to pulse proteolysis. In this dissertation, the results show that proteolysis of dPGM occurred through a partially unfolded intermediate. We observe that ATP binding can induce a large shift in the conformational ensemble of dPGM and promote dimerization. The ligand binding and dimerization events are allosteric. Our results reveal how an allosteric effect can be intrinsic to the folding of an oligomeric protein and demonstrate the important role that the physiological milieu can have in reducing protein intermediates. We have also extended the use of equilibrium φ-analysis (Campos, Bueno, Lopez-Llano, Jiménez, & Sancho, 2004) to determine the structure of a partially unfolded

intermediate of an oligomer using only chemical denaturation. Determination of the structure of an intermediate by equilibrium φ-analysis can be useful to identify energetically coupled sites in a protein, which may assist the development of an allosteric drug.

In chapter 2, we investigate the effect of ligand binding on the equilibrium unfolding of dPGM to understand how ATP affects the stability of dPGM. A partially unfolded, monomeric intermediate of dPGM was found to be populated during equilibrium unfolding. Nucleotides, including ATP, and phosphates significantly enhance the stability of dimeric dPGM over the intermediate. Equilibrium unfolding of dPGM in the presence of 1.0 mM ATP appears cooperative. Through ligand docking and mutagenesis, ATP was determined to bind the active site of dPGM. Our results indicated that ligand binding at the active site is energetically coupled with dimerization such that they display positive allostery.

Chapter 3 reveals the structure of the dPGM intermediate to discern the cooperative network of secondary structural units that couple ligand binding at the active site and dimerization. Through extensive mutagenesis and equilibrium unfolding, equilibrium φ-values were determined for residues throughout the dPGM structure. The equilibrium φ-values report on whether side-chain contacts are maintained or lost in the intermediate state (Campos et al., 2004). Hydrogren/deuterium exchange (HDX) revealed differences in solvent exposure of the peptide backbone between native dPGM and the intermediate. Together, equilibrium φ-analysis and HDX uncover a cooperative folding unit in dPGM

that is unfolded in the intermediate state, and underlies the allostery observed between ligand binding at the active site and dimerization.

In chapter 4, we investigate the effect of non-substrate metabolite binding to dPGM in the physiological milieu to understand whether it has functional consequences. A range of metabolites are found to stabilize dPGM but the presence of salt, especially Mg^{2+}, reduces metabolite binding. The effect of salt itself on dPGM is stabilizing and promotes the dimer over the monomeric intermediate. The effect of salt on several dPGM variants reveals that chloride binds at either end of the cooperative network in dPGM: the active site and the dimer interface. For many metabolites and salts, the effects we observed occur at physiologically relevant concentrations. Our results suggest that the chemical environment of the *E. coli* cytosol would be highly stabilizing for dPGM and metabolites and anions may cause competitive inhibition.

CHAPTER 2. ENERGETIC COUPLING BETWEEN LIGAND BINDING AND DIMERIZATION IN *ESCHERICHIA COLI* PHOSPHOGLYCERATE MUTASE[1]

2.1 Summary

Energetic coupling of two molecular events in a protein molecule is ubiquitous in biochemical reactions mediated by proteins, such as catalysis and signal transduction. Here, we investigate energetic coupling between ligand binding and folding of a dimer using a model system that shows three-state equilibrium unfolding of an exceptional quality. The homodimeric Escherichia coli cofactor-dependent phosphoglycerate mutase (dPGM) was found to be stabilized by ATP in a proteome-wide screen, although dPGM does not require or utilize ATP for enzymatic function. We investigated the effect of ATP on the thermodynamic stability of dPGM using equilibrium unfolding. We found that, in the absence of ATP, dPGM populates a partially unfolded, monomeric intermediate during equilibrium unfolding. However, addition of 1.0 mM ATP drastically reduces the population of the intermediate by selectively stabilizing the native dimer. Using a computational ligand docking method, we predicted ATP binds to the active site of the enzyme using the triphosphate group. By performing equilibrium unfolding and isothermal titration calorimetry with active-site variants of dPGM, we confirmed that active-site residues are involved in ATP binding. Our findings show that ATP promotes dimerization of the protein by binding to the active site, which is distal from the dimer interface. This cooperativity suggests an energetic coupling between the active site and

[1] Adapted with permission from Gardner, N. W., Monroe, L. K., Kihara, D., & Park, C. (2016). Energetic Coupling between Ligand Binding and Dimerization in Escherichia coli Phosphoglycerate Mutase. *Biochemistry, 55*, 1711-1723. Copyright (2016) American Chemical Society.

the dimer interface. We also propose a structural link to explain how ligand binding to the active site is energetically coupled with dimerization.

2.2 Introduction

Cooperativity between two events occurring at two distant sites is a salient feature of protein molecules. Protein folding is a cooperative process where a group of structural units in a protein fold or unfold simultaneously (Ernesto Freire & Murphy, 1991). Due to the cooperative nature, conformational energy landscapes of proteins are composed of discrete partially unfolded forms rather than a continuum of possible conformations (Bai et al., 1995; Chamberlain et al., 1996; Hu et al., 2013). Ligand binding events at two different binding sites in a protein are frequently cooperative; binding of one ligand positively or negatively influences binding of the other ligand. The cooperative binding of two ligands is the basis of allostery (Hilser & Thompson, 2007; Motlagh, Wrabl, Li, & Hilser, 2014), which is commonly observed in many aspects of cellular processes, such as catalysis and signal transduction. The fundamental principle of cooperativity is the same whether observed in ligand binding or in protein folding (Luque, Leavitt, & Freire, 2002). Ligand binding as well as folding/unfolding of a region modulates the conformational energy landscape of a protein. The change in the energy and the population of various conformations results in the changes in macroscopic structural properties in proteins. The structural changes can be as subtle as changes in side-chain dynamics or as drastic as folding of unstructured domains (Motlagh et al., 2014). Understanding the energetic basis of the cooperativity in protein structure is one of the most important challenges in biochemistry.

From our previous proteomics study, we identified a protein that shows a remarkable energetic coupling between ligand binding and folding. Previously, we conducted an energetics-based proteome screen to identify ATP-binding proteins in the *E. coli* proteome (Chang et al., 2012; P.-F. Liu et al., 2011). This proteome screen exploits the fact that ligand binding stabilizes proteins (Sanchez-Ruiz, 2007; Schellman, 1976; Waldron & Murphy, 2003). We subjected proteins in an *E. coli* lysate to urea unfolding in the presence and absence of 1.0 mM ATPγS and compared the amount of proteins that remain intact after pulse proteolysis (Chang et al., 2012; P.-F. Liu et al., 2011). Pulse proteolysis is a way to monitor protein unfolding by selectively digesting the unfolded proteins through a brief incubation with a non-specific protease (Park & Marqusee, 2005). Different from traditional methods based on spectroscopy or calorimetry, pulse proteolysis allows us to study unfolding of many proteins simultaneously on a proteomic scale (Adhikari & Fitzgerald, 2014; Chang et al., 2012; P.-F. Liu et al., 2011). At 2.0 M urea, cofactor-dependent phosphoglycerate mutase (dPGM) was fully digested without ATP, but a significant population of the protein remained intact in the presence of 1.0 mM ATPγS (Chang et al., 2012). *E. coli* dPGM is a homodimer consisting of 249-amino-acid subunits with the molecular weight of 28.4 kDa. Interestingly, dPGM does not require ATP for its enzyme activity. However, we confirmed the stabilization of dPGM in the presence of 1.0 mM ATP by performing pulse proteolysis with purified dPGM (Chang et al., 2012). Later, the Fitzgerald group reported that ATP also stabilizes *S. cerevisiae* dPGM in an energetics-based proteome screen similar to ours (Tran et al., 2014).

Phosphoglycerate mutases (E.C. 5.4.2.11) catalyze the essential isomerization reaction between 3-phosphoglycerate and 2-phosphoglycerate in the glycolysis and gluconeogenesis pathways. Most vertebrate, yeast, and bacteria possess a cofactor-dependant phosphoglycerate mutase (dPGM) that must be activated by the cofactor 2,3-bisphosphoglycerate (2,3-BPG) (Bond et al., 2001). In the activating reaction, dPGM serves as a phosphatase (EC 3.1.3.13) for 2,3-BPG and converts 2,3-BPG into 2-phosphoglycerate or 3-phosphoglycerate while retaining a phosphate group on an active-site histidine residue. In the isomerization reaction, the phosphate group on the phosphohistidine of the activated enzyme is transferred to 2-phosphoglycerate or 3-phosphoglycerate, and the resulting 2,3-BPG is used to phosphorylate the histidine residue. Through this process, dPGM shuttles a phosphate group between 3-phosphoglycerate and 2-phosphoglycerate.

To quantitatively assess the effect of ATP on the stability of this enzyme, we monitored urea-induced equilibrium unfolding of *E. coli* dPGM using fluorescence and circular dichroism. In the absence of ATP, equilibrium unfolding of dPGM reveals two well-separated transitions with an equilibrium intermediate, indicating that dPGM unfolds in a three-state manner. Interestingly, we found that the equilibrium intermediate disappears in the presence of ATP; dPGM unfolds in a two-state manner without any observable equilibrium intermediate upon binding to ATP. Realizing that this protein is an excellent model system to understand the molecular mechanism for the energetic coupling between ligand binding and folding, we decided to investigate how ATP binding promotes dimerization of dPGM. We employed both experimental and computational approaches

to dissect the molecular details of the interaction between dPGM and ATP and the energetic coupling between ATP binding and dPGM folding. Based on our experimental data, we propose a structural model for the cooperativity between folding of the active site and the dimer interface.

2.3 Results

2.3.1 Equilibrium unfolding of dPGM in the presence and absence of ATP

We monitored equilibrium unfolding of dPGM in the absence and presence of 1.0 mM ATP by intrinsic tryptophan fluorescence (Figure 2.1A). From the fluorescence emission spectrum at each urea concentration, we calculate one intensity-weighted average wavelength (IWAW). When dPGM unfolds and its tryptophans become more solvent-exposed, IWAW shifts to longer wavelengths. Immediately apparent in Figure 2.1A is that equilibrium unfolding of dPGM occurs in a three-state manner in the absence of ATP. A shoulder occurs between 2 M and 3 M urea with the IWAW value of about 355 nm, signifying that an intermediate state with properties distinct from the native and unfolded states is populated. The first unfolding transition has the transition midpoint (C_m) of ~1.4 M urea and occurs with a signal change that is about 70 % of the total change in fluorescence. The second unfolding transition has a C_m value of ~3.5 M urea. We also monitored the equilibrium unfolding of dPGM by circular dichroism (Figure 2.2). The two transitions observed by CD occur with the same C_m values as those observed by fluorescence, suggesting that we are observing the same and only equilibrium intermediate. The first transition occurs with 35% of the total change in CD signal upon global unfolding. Unfolding to the intermediate state appears to involve some loss of

secondary structure as well as significant changes in the environment around tryptophan residues.

In the presence of 1.0 mM ATP, equilibrium unfolding of dPGM occurs in only one apparent transition with a C_m of about 3.6 M urea without any indication of an equilibrium intermediate (Figure 2.1A). The addition of 1.0 mM ATP not only stabilizes dPGM dramatically but also suppresses accumulation of the equilibrium intermediate. Due to strong UV absorption by ATP, equilibrium unfolding of dPGM in the presence of 1.0 mM ATP could not be monitored by CD.

2.3.2 Oligomeric state of the equilibrium intermediate

Developing a model to analyze the equilibrium unfolding of a dimeric protein with an equilibrium intermediate requires knowledge of whether the intermediate is a monomer or a dimer. First, we investigated the oligomeric state of the intermediate of dPGM through the dependence of equilibrium unfolding on the protein concentration. As the protein concentration is increased, dimeric states should be stabilized due to Le Chatelier's principle. Therefore, a transition that shows a dependence on the protein concentration involves a change in oligomeric state from dimer to monomer. When we monitored equilibrium unfolding of dPGM with three different concentrations of dPGM (10, 30 and 100 μg/mL), we clearly observed that, as the protein concentration increases, the apparent C_m value of the first transition gradually increases as well (Figure 2.3A). The second transition does not exhibit any observable dependency on the protein concentration. We also monitored equilibrium unfolding of 20, 30, and 100 μg/mL

dPGM by circular dichroism (CD), which confirmed the concentration dependence of the first transition (Figure 2.4). The concentration dependence of equilibrium unfolding best supports a monomer intermediate.

We further examined the oligomeric state of the intermediate by chemical cross-linking with $BS(PEG)_5$, which reacts selectively with amine groups. After dPGM was incubated with $BS(PEG)_5$ in varying concentrations of urea (0 – 5 M), the cross-linked samples were visualized by SDS-PAGE (Figure 2.3B). Without urea, the dPGM dimer and monomer bands are both visible and migrate with apparent molecular weights of ~60 kDa and ~30 kDa, respectfully. The dimer band intensity decreases as urea is increased and disappears at 2.5 M urea. For facile detection, we used in this experiment 0.20 mg/mL dPGM, which is 2-fold greater than the highest concentration we used in the equilibrium unfolding study. Still, it is obvious that the disappearance of the dimer band is coincidental with the first transition in the equilibrium unfolding of dPGM. This result confirms that the equilibrium intermediate is monomer.

We also monitored the change in the hydrodynamic radius of 30 μg/mL dPGM by size-exclusion chromatography under conditions favoring the native state (0 M urea) and the intermediate state (2.5 M urea) (Figure 2.3C). Without urea, one sharp peak is observed at an elution volume of 8.7 mL. In the presence of 2.5 M urea, one peak is observed at an elution volume of 9.5 mL. The change in the elution volume indicates that dPGM has a smaller radius at 2.5 M urea than at 0 M urea, suggesting that the intermediate is a monomer. The peak at 2.5 M urea is broader and, therefore, has a smaller peak height

than the peak at 0 M urea. Exchange between multiple conformations may cause the peak broadening observed at 2.5 M urea. The results of the concentration dependence of equilibrium unfolding, chemical cross-linking, and size exclusion experiments all support a monomer intermediate. With the oligomeric state determined, we use a three-state model with a monomer intermediate (Scheme 1) to analyze the equilibrium unfolding of dPGM.

2.3.3 Effect of ATP on the thermodynamic stability of dPGM

Using the model shown in Scheme 1, we quantified the effect of ATP on the equilibrium unfolding of dPGM (Figure 2.1 and Table 2.1). We fit globally the fluorescence monitored equilibrium unfolding data in the presence and absence of 1.0 mM ATP with the three-state model, despite the apparent two-state transition in equilibrium unfolding of dPGM in the presence of ATP. In the global fit, we reduced fitting parameters by sharing intermediate state and unfolded state signals, *m*-values, and the total protein concentration. In the absence of ATP, the free energy associated with the first transition ($N_2 \rightarrow 2I$; $\Delta G_{I\text{-}N}°$) and the second transition ($I \rightarrow U$; $\Delta G_{U\text{-}I}°$) were determined to be 11.5 ± 0.1 kcal/mol and 6.6 ± 0.3 kcal/mol, respectively (Table 2.1). Independent curve fitting of the equilibrium unfolding of dPGM monitored by CD yielded similar values (Table 2.1 and Figure 2.2). Note that the first transition is the unfolding of the dimer and that $\Delta G_{I\text{-}N}°$ is the free energy change at the standard condition where both $[N_2]$ and [I] are 1 M. To get a better sense of the stability of the protein under our experimental condition (1.0 μM dPGM), we determined f_N, f_I, and f_U using the equilibrium parameters from the curve-fitting (Figure 2.1B). According to the calculation, the effective stability (ΔG_{eff}) (Park &

Marqusee, 2004a), which is defined as $-\mathrm{RTln}\frac{1-f_N}{f_N}$ at a given protein concentration, is only 1.9 kcal/mol under native conditions; 1 in 25 monomeric units adopt non-native conformations (Figure 2.1B). 1.0 mM ATP increases $\Delta G_{I\text{-}N}°$ by 7.0 ± 0.2 kcal/mol, which corresponds to an increase in stability of 3.5 kcal/mol per monomeric unit. The effect of 1.0 mM ATP on $\Delta G_{U\text{-}I}°$ was insignificant (0.6 kcal/mol). In the presence of 1.0 mM ATP, the effective stability of 1.0 μM dPGM is increased to 5.4 kcal/mol. The resulting effect of ATP is that only 1 in 8,600 monomeric units adopt non-native conformations under native conditions (Figure 2.1B). Further, in the absence of ATP, the intermediate reaches a maximum fraction of just over 0.94 around 2.5 M urea. In the presence of 1.0 mM ATP, the intermediate only reaches a maximum fraction of 0.12 around 3.7 M urea. ATP dramatically suppresses accumulation of the equilibrium intermediate. Therefore, our analysis strongly suggests that ATP selectively binds to the native dimer and not to the monomeric intermediate.

The dependence of $\Delta G_{I\text{-}N}°$ and $\Delta G_{U\text{-}I}°$ on the urea concentration, $m_{I\text{-}N}$ and $m_{U\text{-}I}$, respectfully, signifies that both transitions involve substantial unfolding. The *m*-value is proportional to the change in a protein's solvent-accessible surface area during unfolding (Myers, Pace, & Scholtz, 1995). From the global fitting, we find that $m_{I\text{-}N}$ is 2.40 kcal $mol^{-1}M^{-1}$ and $m_{U\text{-}I}$ is 1.86 kcal $mol^{-1}M^{-1}$. The total *m*-value for global unfolding, $m_{U\text{-}N}$, is 3.06 kcal $mol^{-1}M^{-1}$ per monomeric unit ($m_{U\text{-}N} = m_{I\text{-}N}/2 + m_{U\text{-}I}$), which is in agreement with the *m*-value predicted for a dPGM monomer based on the number of residues (3.2 kcal $mol^{-1}M^{-1}$) (Myers et al., 1995). Curve-fitting of the equilibrium unfolding of dPGM

monitored by CD produced similar m-values: $m_{I\text{-}N}$ of 2.3 kcal $mol^{-1}M^{-1}$, $m_{U\text{-}I}$ of 1.80 kcal $mol^{-1}M^{-1}$, and $m_{U\text{-}N}$ of 2.95 kcal $mol^{-1}M^{-1}$ (Table 2.1). Using PISA(Krissinel & Henrick, 2007), we determined the area of the dimer interface in the crystallographic structure of dPGM (715 $Å^2$ per monomer), and we predicted from the area that the *m*-value for dimer dissociation would be only 0.1 kcal $mol^{-1}M^{-1}$ per monomer. The fact that $m_{I\text{-}N}/2$ (1.20 kcal $mol^{-1}M^{-1}$) is much greater than the value predicted for dimer dissociation suggests that the first transition involves not only dissociation of the dimer but also partial unfolding of the monomer. Further, $m_{U\text{-}I}$ (1.86 kcal $mol^{-1}M^{-1}$) is much less than the predicted *m*-value for unfolding of the monomer (3.2 kcal $mol^{-1}M^{-1}$), which clearly shows that the second transition cannot account for unfolding of an intact monomer. The spectroscopic signal of the intermediate state is distinct from the native state by both fluorescence and CD which also supports a partially unfolded intermediate (Figures 2.1 and 2.2).

2.3.4 Structural determinant of ATP binding to dPGM

To elucidate the structural determinant of ATP for the interaction with dPGM, we monitored equilibrium unfolding of dPGM in the presence of 1.0 mM adenosine, AMP, ADP, or GTP (Figure 2.5A). While the effect of adenosine is negligible, AMP, ADP, and GTP significantly shift the first unfolding transition to higher urea concentrations. We fit globally the equilibrium unfolding data in Figure 2.5A, including equilibrium unfolding of dPGM in the absence and presence of 1.0 mM ATP, using the three-state model (Table 2.2). The data sets in the global fitting share the same *m*-values and spectroscopic signals for native, intermediate, and unfolded states. AMP, ADP, and GTP all stabilize the native form exclusively; while $\Delta G_{I\text{-}N}°$ increases dramatically, $\Delta G_{U\text{-}I}°$ is overall insensitive to the

presence of the nucleotides. The change in $\Delta G_{I\text{-}N}°$ ($\Delta\Delta G_{I\text{-}N}°$) is clearly dependent on the number of phosphate groups (Table 2.2); the more phosphate groups a nucleotide contains, the larger $\Delta G_{I\text{-}N}°$ becomes. ATP and GTP have similar effects on dPGM, suggesting that the type of the nucleobase is not important in the interaction. The stabilizing effects of this series of nucleotides show that the phosphate groups of nucleotides are the primary structural determinant of the interaction with dPGM.

The critical role of the phosphate groups in the interactions between nucleotides and dPGM led us to examine whether the same effect could be achieved by phosphate alone. We carried out equilibrium unfolding of dPGM in the presence of 50 mM phosphate or 0.10 mM pyrophosphate (Figure 2.5B). Both salts dramatically stabilize dPGM, and the equilibrium unfolding appears two-state. We fit globally the equilibrium unfolding data in the presence of either phosphate or pyrophosphate in pair with the equilibrium unfolding data of dPGM without any ligand (Table 2.2). The stabilizing effects of both ions are similar in spite of the 500-fold difference in the concentration; 50 mM phosphate increased $\Delta G_{I\text{-}N}°$ by 5.9 kcal/mol, and 0.1 mM pyrophosphate increased $\Delta G_{I\text{-}N}°$ by 5.3 kcal/mol. Again, increasing the number of phosphates per molecule drastically increases the stability conveyed to dPGM. Interestingly, 0.1 mM pyrophosphate is as effective as 1.0 mM ADP in stabilizing dPGM, although pyrophosphate is present at a 10-fold lower concentration. The lack of the bulky nucleoside seems to enhance the binding affinity of pyrophosphate, which corroborates that the nucleoside group (nucleobase and ribose) does not have any significant role in the interaction. Rather, ATP and other nucleotides bind to dPGM as a mono-, di-, or triphosphate, probably through Coulombic interactions

with an extensive array of cationic residues of the protein. The potential binding site with this property is the active site of dPGM. As the crystal structure of dPGM reveals, the active site contains many basic residues in close proximity (Bond et al., 2001). Moreover, the crystallographic structure of a complex between inactive dPGM and tetravanadate has shown that tetravanadate binds to the active site of the enzyme (Bond et al., 2002). Therefore, binding of the triphosphate groups of ATP to the active site may reduce unfavorable Coulombic interactions in the active site and stabilize the native form of dPGM.

2.3.5 Prediction of the ATP-binding site by docking

To predict the ATP-binding site in dPGM, we docked ATP to dPGM using Maestro for GLIDE docking (Friesner et al., 2004). The docking grid was larger than one dPGM monomer to refrain from biasing the predicted binding site. As illustrated in Figure 2.6, ATP was predicted to bind to dPGM within the active site. Bound ATP is oriented with the adenosine moiety partially exposed to solvent and docked in the active site opening while the phosphate groups are sheltered within the active site. We identified potential hydrogen bond partners as donor-accepter pairs separated by d and with acceptable geometry, where 2.5 Å $\leq d \leq$ 3.3 Å. The adenine of ATP can only bond with Asn16 and the ribose forms one bond with Lys238. The phosphate groups of ATP are predicted to form ion-pair hydrogen bonds with multiple residues. The α phosphate group forms hydrogen bonds with Arg9, Lys99, Arg115, and Arg116. The β phosphate group forms hydrogen bonds with Arg89, Tyr91, Lys99, Arg115, Arg116 and Asn185. Located the

furthest into the active site pocket, the γ phosphate group can form hydrogen bonds with Gly23 and Lys99.

The predicted binding mode of ATP is analogous to the binding mode of the tetravanadate observed in the crystallographic structure of dPGM (Bond et al., 2002). Tetravanadate bound in the active site and the locations of the four vanadate groups were denoted as V1, V2, V3, and V4 (Bond et al., 2002). Remarkably, the α, β, and γ phosphates of ATP are predicted to be located approximately at V4, V3, and V2, respectively (Figure 2.6). Most of the predicted bonds between the phosphate groups of ATP and dPGM are also formed between tetravanadate and dPGM; only the α phosphate group is predicted to form bonds that are not observed with tetravanadate. This difference likely arises because three oxygen atoms of the vanadate moiety were modeled as zero occupancy due to poor electron density (Bond et al., 2002). The top four scoring poses from GLIDE were analyzed to test whether ATP is predicted to have multiple binding modes that are close in energy. Notably, the nucleoside moiety is predicted to adopt various conformations while the top two scoring poses show complete overlap in the positions of the phosphate groups (Figure 2.7). In the two less favorable poses, the α, β, and γ phosphates are located at V3, V2, and V1, respectively. The agreement between the predicted binding mode of ATP and that of tetravanadate in the crystallographic structure of *E. coli* dPGM supports that the predicted binding mode of ATP is viable.

2.3.6 Contribution of the active-site residues to ATP binding and stability

To test the involvement of active site residues in binding with ATP, we generated a series of dPGM variants with point mutation to alanine at Arg9, Lys99, Arg115, and Arg116 (Figure 2.6). If a residue interacts with ATP through its side chain, an alanine mutation will abolish the interaction and result in ATP conferring less stability to that variant than to wild-type dPGM. As a control, we also mutated Lys222, which is distal to the active site (Figure 2.6), to alanine. CD spectra show that all the variants retain the secondary structure of wild-type dPGM (Figure 2.8). We examined the equilibrium unfolding of each variant in the presence and absence of 1.0 mM ATP (Figure 2.9). To quantify the changes in the stability conveyed to dPGM by ATP ($\Delta G_{ATP}°$), we fit globally each variant in the presence and absence of 1.0 mM ATP using the three-state model represented in Scheme 1 (Table 2.3). Because ATP stabilizes only native dPGM, we define $\Delta G_{ATP}°$ as $\Delta G_{I\text{-}N}(ATP)° - \Delta G_{I\text{-}N}°$.

As with wild-type dPGM, the equilibrium unfolding of each variant appears three-state. Effects of R9A and K222A mutations on equilibrium unfolding were insignificant (Figures 2.9A and 2.9E). However, K99A, R115A, and R116A mutations all shift the first transition to higher urea concentrations than observed with wild-type dPGM. This indicates that K99A, R115A, and R116A mutations stabilize the native state alone. K99A (Figure 2.9B), R115A (Figure 2.9C), and R116A (Figure 2.9D) mutations increase $\Delta G_{I\text{-}N}°$ by 2.3± 0.2 kcal/mol, 3.1 ± 0.3 kcal/mol, and 1.2 ± 0.1 kcal/mol, respectively. No mutations alter the stability of **I** significantly.

In the presence of 1.0 mM ATP, the stability of all variants is dramatically increased (Figure 2.9). For each active site variant, $\Delta G_{ATP}°$ is less than that of wild type (7.0 kcal/mol). The R116A variant gains the most stability from ATP with a $\Delta G_{ATP}°$ of 5.4 ± 0.2 kcal/mol. R9A and K99A are the least stabilized by 1.0 mM ATP with a $\Delta G_{ATP}°$ of 4.2 ± 0.2 kcal/mol and 4.0 ± 0.4 kcal/mol, respectively. With R115A, $\Delta G_{ATP}°$ is 4.8 ± 0.4 kcal/mol. Notably, the intermediate is still apparent for R9A in the presence of 1.0 mM ATP due to the rightward shift of its second unfolding transition. The control mutation, K222A does not appear to significantly influence the stabilizing effect of ATP (Figure 2.9E). Because the active-site mutations reduce $\Delta G_{ATP}°$ while the control does not, ATP clearly interacts with the active-site residues.

The effect of the active site mutations on native dPGM ($\mathbf{N_2}$) and the intermediate form (**I**) betray conformational differences between the two forms. If the environment of the mutated active site resides was similar in both $\mathbf{N_2}$ and **I**, then removal of the side chain would stabilize **I** similar to $\mathbf{N_2}$. The K99A, R115A, and R116A mutations stabilize $\mathbf{N_2}$, as evident by the increase in $\Delta G_{I\text{-}N}°$, but had insignificant effects on the second transition (Table 2.3). The selective enhancement of $\Delta G_{I\text{-}N}°$ suggests that Lys99, Arg115, and Arg116 are in an unfavorable environment for $\mathbf{N_2}$ and that removal of the positive charges from the active site stabilizes the protein similar to phosphate binding. Because removal of Lys99, Arg115, or Arg116 does not alter the stability of **I**, the environment the residues experience in **I** is quite distinct from that in $\mathbf{N_2}$. Collectively, the effects of point mutations on the equilibrium unfolding of dPGM support a model where the active site is disrupted in the monomeric intermediate state.

2.3.7 Validation of the ATP-interacting residues by isothermal titration calorimetry

ATP binding by each dPGM variant was monitored using isothermal titration calorimetry (ITC) to obtain the dissociation equilibrium constant, K_d (Table 2.4). Wild-type dPGM was found to bind ATP with K_d of 0.32 ± 0.07 μM. R116A retained the greatest affinity for ATP out of the active site mutants with K_d of 1.5 ± 0.3 μM. The variants with the most jeopardized ATP binding are K99A and R115A, which have K_d of 7 ± 2 μM. For R9A, K_d is 5.5 ± 0.05 μM. As anticipated, K222A has a negligible impact on ATP binding with K_d of 0.4 ± 0.1 μM. The K_d value determined for each active-site mutant is increased relative to wild-type dPGM, which signifies the role of the residues in ATP binding. The effects of mutation on K_d confirm that each active-site residue chosen for mutation is involved in ATP binding while the control is not.

From the K_d values, we calculated the free energy change for binding, $\Delta G_{bind}°$, at 1.0 mM ATP (Table 2.4). The effect of a mutation on $\Delta G_{bind}°$ ($\Delta\Delta G_{bind}°$) should be the same as the effect of the mutation on $\Delta G_{ATP}°$ ($\Delta\Delta G_{ATP}°$) determined from equilibrium unfolding if our model (Scheme 1) is correct. As shown in Table 2.4, the $\Delta\Delta G_{bind}°$ and $\Delta\Delta G_{ATP}°$ values are similar for the active-site variants although $\Delta\Delta G_{bind}°$ values are consistently greater in magnitude than the $\Delta\Delta G_{ATP}°$ values. This difference may arise from the effect of urea on ATP binding because we calculated $\Delta\Delta G_{ATP}°$ values from equilibrium unfolding in urea. Overall, the effect of the mutations as determined by ITC corroborates what we determined from the equilibrium unfolding study. The agreement between the two techniques also supports that the three-state model applied to equilibrium unfolding is appropriate in the presence and absence of 1.0 mM ATP.

2.4 Discussion

We demonstrated that equilibrium unfolding of dPGM occurs in a three-state manner with a monomeric intermediate (Figure 2.1A). The *m*-values and the spectroscopic signals indicate that the monomeric intermediate is partially unfolded. In the presence of ATP, the monomeric intermediate is barely populated throughout equilibrium unfolding (Figure 2.1B). By increasing $\Delta G_{I\text{-}N}°$, 1.0 mM ATP nearly shifts the equilibrium unfolding of dPGM into a two-state folding system where the entire protein is either folded or unfolded. Using a computational modeling and site-directed mutagenesis, we also clearly demonstrated that ATP binds to the active site of the enzyme through the interactions between cationic residues in the active site and the triphosphate group of ATP. During catalysis, the active site of dPGM accommodates the substrates (2-phosphoglycerate and 3-phosphoglycerate) or the cofactor (2,3-bisphosphoglycerate). Apparently, the triphosphate group of ATP mimics the multianionic structures of the substrates and the cofactor. Though we identified dPGM as an ATP-binding protein in our previous proteomics screen, the interaction is not specific to ATP. Any ligand with repeating phosphate groups such as ADP, GTP, or pyrophosphate can bind to the active site with significant affinity. In this study, we investigated the binding of ATP to dPGM under a low salt condition to reliably assess the contribution of ligand binding to the folding of the protein. Because the interaction between ligands with phosphoryl groups and dPGM is mostly Coulombic, the affinity of the ligands to the protein is reduced under physiological salt conditions (Chapter 4). Still, the total concentration of the metabolites with diphosphate or triphosphate groups can be tens of millimolar in cellular environment, these functionally unrelated metabolites may significantly influence the stability of

dPGM. Further studies would be necessary to elucidate the role of the interaction between the metabolites and dPGM under physiological conditions

Ligand binding to dPGM selectively stabilizes the native dimer over a partially unfolded monomeric intermediate. Our analysis of equilibrium unfolding of dPGM in the presence of various ligands indicates that the ligands do not stabilize the monomeric intermediate; the effects of ligands on $\Delta G_{U\text{-}I}°$ are all negligible (Tables 2.1 and 2.2). The lack of stabilization by the ligands suggests that the active site is disrupted in the monomeric intermediate. The effect of active-site mutations on stability also suggests that the active site is unfolded in the intermediate. Therefore, dissociation of the dimer into the monomeric intermediate also induces unfolding of the active site. The active site of dPGM is distal from the dimer interface (Figure 2.6); yet, proper folding of the active site and the dimer interface appears to be cooperative.

A thermodynamic cycle clearly illustrates the cooperativity between ligand binding and dimerization in dPGM (Figure 2.10). In 2.5 M urea, where the cooperativity is most evident, dPGM is predominantly partially unfolded monomers (**I**) in the absence of ATP but ATP-bound dimers ($\mathbf{N_2{\bullet}L_2}$) in the presence of 1.0 mM ATP (Figure 2.1). The reaction from **I** to $\mathbf{N_2{\bullet}L_2}$ is thermodynamically favorable in the presence of 1.0 mM ATP. The free energy change associated with the reaction from **I** to $\mathbf{N_2{\bullet}L_2}$ in 2.5 M urea can be calculated using the thermodynamic parameters of equilibrium unfolding in Table 2.1 and Eq. 2 ($\Delta G_{N\text{-}I}°$ = –12.5 kcal/mol). Under our experimental condition (1 μM dPGM), the $\Delta G_{N\text{-}I}°$ value corresponds to the effective free energy change(Park & Marqusee, 2004a)

of –2.4 kcal/mol per monomer, which is consistent with the reaction from **I** to **N_2•L_2** being thermodynamically favorable. The predominant presence of **I** in the absence of ATP indicates that dimerization of the partially unfolded monomer (**I** to **N_2**) is thermodynamically unfavorable at 2.5 M urea. The free energy change associated with the reaction from **I** to **N_2** is again calculated using the thermodynamic parameters of equilibrium unfolding in Table 2.1 and Eq. 2 ($\Delta G_{N\text{-}I}° = –5.5$ kcal/mol). Under our experimental condition (1 μM dPGM), the effective free energy change(Park & Marqusee, 2004a) of the reaction is 1.1 kcal/mol per monomer, which again confirms that folding of **I** to **N_2** is unfavorable under this condition. Then, the thermodynamic relationship shows that binding of ATP to the native dimer (**N_2**) is associated with $\Delta G°$ of –3.5 kcal/mol in the presence of 1.0 mM ATP. The favorable binding energy by ATP (–3.5 kcal/mol per monomer) overcomes the unfavorable free energy associated with the folding and dimerization (1.1 kcal/mol per monomer). The affinity to ATP is enhanced significantly because the formation of **N_2** already pays the penalty of forming the binding-competent active site. This phenomenon is, therefore, an example of preorganization of the binding site.

The energetics of the same reaction can be analyzed through the pathway with ligand-bound folded monomer (**M•L**). As we cannot observe **M•L** under any circumstance, we do not have thermodynamic data for this pathway. However, the cooperative nature between ligand binding and dimerization can be explained with this pathway also. The dimerization of **M•L** to **N_2•L_2** must be thermodynamically favorable, because the free energy change associated with this reaction is the inverse of the energetic penalty (Δg_{int})

of exposing the dimer interface to the solvent. Therefore, ligand binding drives the folding of the monomer (**I** to **M•L**), which is followed by a favorable dimerization reaction. This scenario can be seen as a cooperative binding between two ligands. Binding of one ligand (L) promotes the folding of the other binding site (the dimer interface), which dramatically increases the affinity for the other ligand (another monomeric unit).

The ligand-induced dimerization of dPGM in 2.5 M urea is reminiscent of allostery observed in proteins with an intrinsically unstructured region. Maximal energetic coupling between two domains or regions is achieved when a segment of the protein is folded only in the presence of ligand and when abolishing the interaction between the domains causes a large energetic penalty (Hilser & Thompson, 2007; Wrabl et al., 2011). dPGM in 2.5 M urea satisfies these conditions. The active site and the dimer interface fold only when ATP or other stabilizing ligands are present. Also, exposing the dimer interface is significantly unfavorable. Also of note is that the energetic coupling in dPGM should enable ligand binding to one subunit to bolster the stability of the other subunit through the dimer interface, which will lead to favorable binding of ligand to the other subunit. Though the cooperativity is most dramatic in 2.5 M urea, the energetic coupling between the active site and the dimer interface exists under native conditions also. f_I is significantly decreased in the presence of 1.0 mM ATP (Figure 2.1B). We expect that phosphorylation of His10 and/or substrate binding in one subunit will alleviate columbic strain in the active site and stabilize the dimeric form, which will lead to the stabilization of the other subunit as well.

The crystallographic structure of dPGM offers clues as to how the active site may be structurally coupled to the dimer interface. A loop comprising residues 118-151 leads from the active site of dPGM to the dimer interface with portions of the loop making intersubunit contacts (Figure 2.11). The loop connects helix α5 (residues 108-117), which forms part of the active site rim and contains Arg115 and Arg116, to helix α6 (residues 152-167) (Figure 2.11). If the active site structure were disrupted, unfolding of helix α5 may increase the dynamics of the adjacent loop; likewise, ligand binding to residues Arg115 and Arg116 could anchor helix α5 and reduce the dynamics of the adjacent loop and, consequently, stabilize the dimer interface. Interestingly, Arg115 interacts with the C-terminal tail of dPGM, which folds over the active site, when the enzyme is phosphorylated (Bond et al., 2001). Arg115 and Arg116 have been postulated to play a main role in switching between inactive and active conformations due to substrate binding (Bond et al., 2001). These two substrate-binding residues appear to be a hotspot for energetically linking the active site to structural change throughout the protein.

The sequence, secondary structure, and active site structure of dPGM is well conserved across *E. coli, S. cerevisiae, S. pombe,* and human brain dPGM, with the key differences being localized to subunit interfaces (Bond et al., 2001). Despite the high sequence similarity between homologous dPGMs, *E. coli* and mammalian dPGMs are dimeric, *S. cerevisiae* dPGM is tetrameric, and *S. pombe* dPGM is monomeric (Bond et al., 2001). Like *E. coli* dPGM, oligomeric human and *S. cerevisiae* dPGMs have a “split” helix (α5/α6) with a long loop spanning from helix α5 to helix α6 that takes part in intersubunit

contacts (Rigden, Walter, Phillips, & Fothergill-Gilmore, 1999b; Yanli Wang et al., 2005). However, the monomeric *S. pombe* dPGM contains a drastically truncated loop, 25 residues shorter than the loop in *E. coli* dPGM, which would not be able to take part in the dimer interface (Figure 2.11) (J. Nairn et al., 1994; Uhrínová et al., 2001a). The absence of this long loop in the monomeric *S. pombe* dPGM suggests to us that the loop indeed provides a key structural link between the active site and subunit interface that drives the energetic coupling we observed. We are currently determining the structure of the monomeric intermediate of *E. coli* dPGM, which may elucidate the structural link between the active site and dimer interface.

The energetic coupling between ligand binding and the dimerization in dPGM is consistent with the findings from computational studies that active-site residues tend to have greater energetic coupling with other regions of a protein (T. Liu et al., 2007; Luque & Freire, 2000). A survey of sixteen proteins revealed that active sites tend to contain low-stability regions which rearrange in the presence of ligand and underlie high affinity binding (Luque & Freire, 2000). When the active sites unfold, conformational changes tend to propagate throughout the protein, not simply in the vicinity of the active site (Luque & Freire, 2000). These low-stability active-site regions are then energetically linked with a significant portion of the protein structure (Hilser et al., 1998). A rigorous analysis of the energetic coupling between regions in the structures of five monomeric proteins revealed that active sites contain the residues with the greatest amount of energetic coupling to other regions of the protein; active sites exert the greatest control over the conformational ensemble (T. Liu et al., 2007).

The effect of ATP and other stabilizing ligands on dPGM shows that cooperativity between ligand binding and dimerization may suppress dissociation of a dimer in the presence of ligand. This principle can be extended to multimeric proteins as well. In a multimeric protein, oligomerization can be energetically coupled with the folding of a distal active site and also ligand binding at the active site. The active site and the dimer interface of dPGM show positive cooperativity so that the two different regions fold and unfold in concert. The resulting effect of the positive cooperativity is that ligand binding stabilizes only the native multimeric forms and redistributes the ensemble populations away from partially unfolded monomeric forms. Increasing the energy gap between the native protein and partially unfolded conformations is paramount to avoiding risky partial unfolding, which can lead to proteolytic digestion (Hubbard, 1998), misfolding (Dobson, 2003), aggregation (Horwich, 2002), and may cause disease states (Thomas, Qu, & Pedersen, 1995). In principle, negative cooperativity is also possible in multimeric proteins. In this case, ligand binding is coupled with unfolding of the interface, which promotes the dissociation of multimeric proteins. One seminal example of the negative cooperativity is the release of the two active monomers of protein kinase A from the heterotetrameric complex with two regulatory subunits upon binding of cAMP to the regulatory subunits (Taylor, Buechler, & Yonemoto, 1990) Exploiting and imitating energetic coupling between ligand binding and multimerization would be valuable in designing and engineering novel proteins not only with desired structures but also with useful regulatory mechanisms.

2.5 Experimental Procedures

2.5.1 Preparation of purified dPGM

We overexpressed dPGM in *E. coli* BL21 (DE3)pLysS cells under the control of the T7 promoter as described previously (Chang et al., 2012). After resuspending the harvested cells in 20 mM TrisHCl-NaOH buffer (pH 8.0) containing 10 mM EDTA, we lysed the cells by sonication. We collected the supernatant after centrifugation and sterilized the lysate via filtration with a 0.45 μm filter. We loaded the lysate into a DEAE-sepharose FF column (GE Healthcare Life Sciences; Piscataway, NJ) pre-equilibrated with 20 mM TrisHCl-NaOH buffer (pH 8.0). Bound proteins were eluted with a linear NaCl gradient from 0.10 M to 0.35 M. We then pooled the fractions containing dPGM and dialyzed overnight against 20 mM TrisHCl-NaOH buffer (pH 8.0). We further purified the dialyzed solution using a Source 15Q column on an ÄKTA FPLC system (GE Healthcare Life Sciences; Pittsburgh, PA). Bound protein was eluted with a linear NaCl gradient from 0 M to 0.25 M. We checked the purity of dPGM by SDS-PAGE. We concentrated the purified protein with a Centriprep 10K centrifugal filter (EMD Millipore; Billerica, MA) and determined the concentration of dPGM by absorbance at 280 nm using an extinction coefficient of 56,380 M^{-1} cm^{-1}, which is estimated from the amino acid composition (Pace, Vajdos, Fee, Grimsley, & Gray, 1995).

QuikChange site-directed mutagenesis kits (Agilent Technologies, Santa Clara, CA) were used to produce all dPGM variants. The sequence of each variant was verified at the Purdue Genomics Core Facility (West Lafayette, IN). All dPGM variants were purified through the two rounds of anion exchange chromatography as described above.

2.5.2 Equilibrium unfolding and analysis

We prepared 30 μg/mL dPGM (~ 1.0 μM) in 20 mM TrisHCl-NaOH buffer (pH 8.0) containing varying concentrations of urea (0 – 5 M). We investigated the equilibrium unfolding dPGM under a low ionic strength condition where the application of three-state model was most reliable and the effect of ATP was maximal (Data not shown). The absolute urea concentrations were determined from the refractive index (Pace, 1986). Additional ligands varied on an experimental basis. The samples were equilibrated overnight prior to spectroscopy. The unfolding of dPGM was monitored by intrinsic tryptophan fluorescence or circular dichroism (CD). Fluorescence emission spectra were collected from 320 nm to 400 nm with a FlouroMax-3 fluorimeter (Horiba Jobin Yvon; Edison, NJ) using an excitation wavelength of 280 nm. Changes in molar ellipticity were monitored at 222 nm with a JASCO J-815 CD Spectrometer (JASCO; Easton, MD).

To analyze equilibrium unfolding by fluorescence, we first calculated the intensity-weighted average wavelength (IWAW):

$$\mathrm{IWAW} = \frac{\sum \lambda \cdot I}{\sum I} \tag{2.1}$$

where λ is the emission wavelength and I is the fluorescence intensity. Unlike the raw fluorescence intensity data, IWAW is not sensitive to variations in protein concentration and provids a reliable measure of protein unfolding. Equilibrium unfolding data were fit

to a three-state model with a monomer intermediate:(Gloss & Matthews, 1997; Mallam & Jackson, 2005)

$$N_2 \underset{}{\overset{K_1}{\rightleftarrows}} 2I \underset{}{\overset{K_2}{\rightleftarrows}} 2U$$

Scheme 2.1

The equilibrium constants are defined as: $K_1 = [I]^2/[N_2]$ and $K_2 = [U]/[I]$. By the linear extrapolation method (Greene & Pace, 1974), the dependence of the equilibrium constants on urea concentration are expressed as:

$$K_1 = \exp\left(\frac{m_{I\text{-}N} D - \Delta G^{\circ}_{I\text{-}N}}{RT}\right) \tag{2.2}$$

$$K_2 = \exp\left(\frac{m_{U\text{-}I} D - \Delta G^{\circ}_{U\text{-}I}}{RT}\right) \tag{2.3}$$

where D is the urea concentration, $\Delta G_{I\text{-}N}°$ is the free energy difference between N_2 and 2I at 0 M urea, and $\Delta G_{U\text{-}I}°$ is the free energy difference between U and I at 0 M urea. The $m_{I\text{-}N}$ and $m_{U\text{-}I}$ values are the dependence of $\Delta G_{I\text{-}N}°$ and $\Delta G_{U\text{-}I}°$, respectively, on the urea concentration. The spectroscopic signal, Y_o, is expressed as a summation of the contribution of each state:

$$Y_0 = Y_N f_N + Y_I f_I + Y_U f_U \tag{2.4}$$

Y_N, Y_I, and Y_U are the spectroscopic signals of N_2, I, and U, respectfully. f_N, f_I, and f_U are the fractions of the monomeric unit in N_2, I, and U, respectfully ($f_N + f_I + f_U = 1$). We expressed f_N, f_I, and f_U as functions of K_1, K_2, and P_t, the total protein concentration in the monomeric unit ($P_t = 2[N_2] + [I] + [U]$), as described elsewhere (Mallam & Jackson, 2005). We determined $m_{I\text{-}N}$, $m_{U\text{-}I}$, $\Delta G_{I\text{-}N}°$, and $\Delta G_{U\text{-}I}°$ by fitting equilibrium unfolding data with eq 2.4 using OriginPro 8.5.1 (OriginLab; Northampton, MA). For global fitting of multiple equilibrium unfolding data sets, we reduced parameters by sharing the spectroscopic signals for the intermediate state and unfolded state, both *m*-values, and P_t amongst all data sets. For global fittings in Figure 2.5, the spectroscopic signals were normalized to obtain the relative signal, Y_R, which is defined as:

$$Y_R = \frac{Y_0 - Y_N}{Y_U - Y_N} \tag{2.5}$$

where Y_0 is the spectroscopic signal at a given urea concentration.

When the stability of dPGM was determined in the presence of a ligand, we calculated the dissociation equilibrium constant K_d from the increase in the stability ($\Delta\Delta G_{unf}°$) using the following relationship (Schellman, 1975):

$$\Delta\Delta G_{unf}^{\circ} = RT \ln\left(1 + \frac{[L]}{K_d}\right) \tag{2.6}$$

where K_d is the dissociation equilibrium constant, and [L] is the ligand concentration.

2.5.3 Chemical cross-linking

To monitor dissociation of the dPGM dimer during equilibrium unfolding, we chemically cross-linked the dimer using bis(succinimidyl) penta(ethylene glycol) ($BS(PEG)_5$) (Thermo Fisher Scientific; Waltham, MA) at varying concentrations of urea. We dialyzed a concentrated dPGM solution against 20 mM HEPES-HCl buffer (pH 8.0). Using the dialyzed dPGM solution, we prepared 0.20 mg/mL dPGM in 20 mM HEPES-HCl buffer (pH 8.0) with varying concentrations of urea (0 – 5 M). We initiated cross-linking by adding $BS(PEG)_5$ to a final concentration of 2.0 mM. After incubating the reaction at 25°C for one hour, we mixed 80 µL of each sample with 20 µL of 5-fold concentrated SDS-PAGE sample buffer. We analyzed the SDS-PAGE samples with a 15% polyacrylamide gel. After staining the SDS-PAGE gel with Sypro Red Protein Gel Stain (Life Technologies; Grand Island, NY), we imaged the gel with a Typhoon scanner (GE Healthcare Life Sciences; Pittsburgh, PA) and quantified the band intensities using ImageJ (http://imagej.nih.gov/ij).

2.5.4 Size exclusion chromatography

We conducted size exclusion chromatography using a TSKgel G3000SWxL column (TOSOH Bioscience LLC; King of Prussia, PA) with a column volume of about 14.3 mL. The mobile phase was 20 mM TrisHCl-NaOH buffer (pH 8.0) containing 50 mM NaCl and 0 M or 2.5 M urea. We prepared 30 µg/mL dPGM in a respective mobile phase. For

each run, we injected 50 μL of dPGM solution into the column and ran at 0.5 mL/min while the absorbance at 280 nm was recorded.

2.5.5 ATP Docking using GLIDE

Docking of ATP to dPGM was carried out with the Maestro software suite (Schrödinger; New York, NY). The ATP ligand was prepared for docking using the LigPrep plugin included in Maestro. ATP was docked to inactive dPGM (PDB: 1E59) using GLIDE (grid-based ligand docking with energetics) docking (Friesner et al., 2004). For docking, the OPLS_2005 force field was used and ionization states for ATP were determined under pH 8.0 +/- 2.0 using Epik. A cubic docking grid was generated to fit the entire dPGM monomer so that the binding location was unbiased. Docking was performed using Glide's XP (extra precision) setting with the ligand treated flexibly. From the docking run the top 4 scoring ATP poses were retained.

2.5.6 Circular dichroism

CD spectra spanning 190 to 250 nm were collected on a JASCO J-815 CD spectrometer (JASCO, Easton, MD). dPGM samples were prepared from 10-fold dilutions of concentrated protein stock. The concentrations of dPGM were determined by absorbance at 280 nm. Samples were equilibrated at room temperature for at least two hours prior to the measurement.

2.5.7 Isothermal titration calorimetry

All dPGM variants were prepared at 42 μM (1.2 mg/mL) and dialyzed against 20 mM TrisHCl-NaOH (pH 8.0) for ITC. After dialysis, the dPGM concentration was determined by absorbance at 280 nm. ATP (Sigma-Aldrich; St. Louis, MO) was prepared at a concentration of 0.40 mM in the dialysis buffer. To monitor ATP binding to dPGM, the heat evolution was monitored with A MicroCal iTC200 calorimeter (GE Healthcare Life Sciences; Pittsburgh, PA) upon ATP titration. ATP was titrated through 21 injections of 0.40 mM ATP with 3-minute intervals. The resulting thermograms were automatically integrated by NITPIC (Keller et al., 2012) and fit with a model of 1:1 binding using SEDPHAT (https://sedfitsedphat.nibib.nih.gov) (Zhao, Piszczek, & Schuck, 2015). The fitted curves were plotted using the software Gussi (http://biophysics.swmed.edu/MBR/software.html).

Table 2.1 Thermodynamic parameters of dPGM in the presence and absence of ATP**.**

Probe	ATP (mM)	$\Delta G_{I-N}°$ (kcal/mol)	m_{I-N} (kcal $mol^{-1}M^{-1}$)	$\Delta G_{U-I}°$ (kcal/mol)	m_{U-I} (kcal $mol^{-1}M^{-1}$)	$\Delta G_{U-N}°$ (kcal/mol)
Fluorescence	0	11.5 ± 0.1	2.40 ± 0.06	6.6 ± 0.3	1.86 ± 0.09	24.7 ± 0.6
	1.0	18.5 ± 0.2	–	6.0 ± 0.4	–	30.5 ± 0.8
CD	0	11.9 ± 0.2	2.3 ± 0.1	6.2 ± 0.3	1.80 ± 0.07	24.3 ± 0.6

Table 2.2 Thermodynamic parameters of dPGM unfolding in the presence of nucleotides and phosphates.

Ligand	$\Delta G_{I-N}°$ (kcal/mol)	m_{I-N} (kcal mol^{-1}M^{-1})	$\Delta G_{U-I}°$ (kcal/mol)	m_{U-I} (kcal mol^{-1}M^{-1})	$\Delta\Delta G_{I-N}°$[d] (kcal/mol)	K_d[e] (μM)
–[a]	11.6 ± 0.1	2.56 ± 0.03	5.6 ± 0.2	1.62 ± 0.05	–	–
Adenosine[a]	11.7 ± 0.1	–	5.6 ± 0.2	–	0.09 ± 0.07	–
AMP[a]	13.6 ± 0.1	–	5.5 ± 0.2	–	1.96 ± 0.09	240 ± 40
ADP[a]	16.4 ± 0.1	–	5.2 ± 0.2	–	4.8 ± 0.1	18 ± 3
ATP[a]	19.3 ± 0.2	–	5.0 ± 0.3	–	7.7 ± 0.2	1.5 ± 0.5
GTP[a]	20.9 ± 0.8	–	4.3 ± 0.6	–	9.3 ± 0.8	0.4 ± 0.5
Phosphate[b]	17.3 ± 0.1	2.36 ± 0.03	5.6 ± 0.3	1.82 ± 0.08	5.9 ± 0.1	350 ± 60
Pyrophosphate[c]	16.8 ± 0.1	2.47 ± 0.04	5.2 ± 0.3	1.62 ± 0.09	5.3 ± 0.1	1.2 ± 0.2

[a]All data sets were fit with one global fit; all ligand concentrations were 1.0 mM.
[b]Equilibrium unfolding data with 50 mM Phosphate was fit globally along with equilibrium unfolding data without a ligand.
[c]Equilibrium unfolding data with 0.10 mM Pyrophosphate was fit globally along with equilibrium unfolding data without a ligand.
[d]$\Delta\Delta G_{I-N}° = \Delta G_{I-N}°$(ligand) – $\Delta G_{I-N}°$(no ligand).
[e]K_d values were calculated with eq 2.6.

Table 2.3 Thermodynamic parameters of dPGM variants in the presence and absence of ATP.

Variant	[ATP] (mM)	$\Delta G_{I-N}°$ (kcal/mol)	m_{I-N} (kcal mol^{-1}M^{-1})	$\Delta G_{U-I}°$ (kcal/mol)	m_{U-I} (kcal mol^{-1}M^{-1})	$\Delta G_{ATP}°^{a}$ (kcal/mol)
R9A	0	12.4 ± 0.1	2.98 ± 0.07	6.4 ± 0.5	1.6 ± 0.1	
	1.0	16.6 ± 0.2	–	6.3 ± 0.5	–	4.2 ± 0.2
K99A	0	13.8 ± 0.2	2.70 ± 0.08	7 ± 1	1.9 ± 0.3	
	1.0	18.5 ± 0.2	–	6 ± 1	–	4.0 ± 0.4
R115A	0	14.6 ± 0.3	2.7 ± 0.1	5 ± 1	1.5 ± 0.3	
	1.0	19.4 ± 0.3	–	5 ± 1	–	4.8 ± 0.4
R116A	0	12.71 ± 0.09	2.75 ± 0.05	6.0 ± 0.4	1.7 ± 0.1	
	1.0	18.1 ± 0.2	–	5.8 ± 0.4	–	5.4 ± 0.2
K222A	0	11.6 ± 0.1	2.48 ± 0.09	6.1 ± 0.4	1.9 ± 0.1	
	1.0	21 ± 2	–	4.4 ± 0.9	–	9 ± 2

[a] $\Delta G_{ATP}° = \Delta G_{I-N}(\text{ATP})° - \Delta G_{I-N}(\text{No ATP})°$

Table 2.4 ATP binding monitored by ITC.

Variant	K_d (x10^{-6} M)	$\Delta G_{bind}°$[a] (kcal/mol)	$\Delta\Delta G_{bind}°$[b] (kcal/mol)	$\Delta\Delta G_{ATP}°$[c] (kcal/mol)
WT	0.32 ± 0.07	4.7 ± 0.1		
R9A	5.5 ± 0.5	3.1 ± 0.05	–1.7 ± 0.1	–1.4 ± 0.1
K99A	7 ± 2	2.9 ± 0.2	–1.8 ± 0.2	–1.5 ± 0.2
R115A	7 ± 2	2.9 ± 0.2	–1.8 ± 0.2	–1.1 ± 0.2
R116A	1.5 ± 0.3	3.8 ± 0.1	–0.9 ± 0.2	–0.8 ± 0.1
K222A	0.4 ± 0.1	4.6 ± 0.1	–0.1 ± 0.2	1.2 ± 1

[a] $\Delta G_{bind}°$ was calculated with Eq. 6 using [ATP] = 1.0 mM.
[b] $\Delta\Delta G_{bind}° = \Delta G_{bind}°$(varient) – $\Delta G_{bind}°$(WT)
[c] $\Delta\Delta G_{ATP}° = \Delta G_{ATP}°$(variant) – $\Delta G_{ATP}°$(WT)

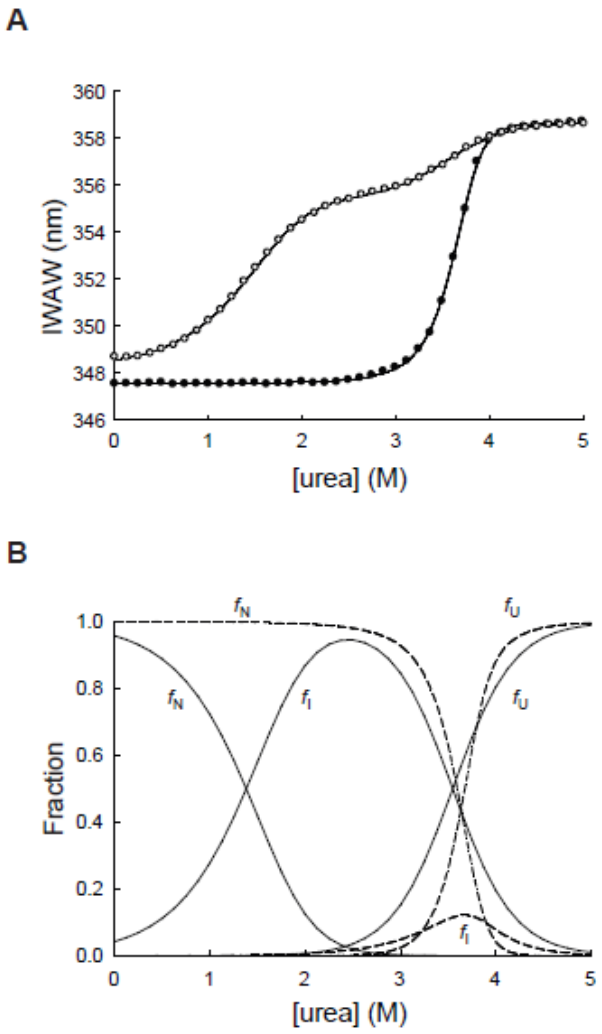


Figure 2.1 Effect of ATP on equilibrium unfolding of dPGM

(A) Equilibrium unfolding of 0.030 mg/mL dPGM was monitored by tryptophan fluorescence in the absence (○) and presence of 1.0 mM ATP (●). (B) Populations of native, intermediate, and unfolded dPGM in equilibrium unfolding. The fraction of native (f_N), intermediate (f_I) and unfolded (f_U) dPGM for 0.030 mg/mL dPGM in the presence (dashed line) and absence (solid line) of 1.0 mM ATP was calculated with the fitting parameters in Table 1.

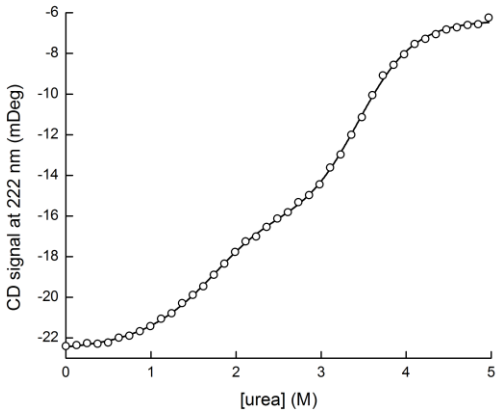
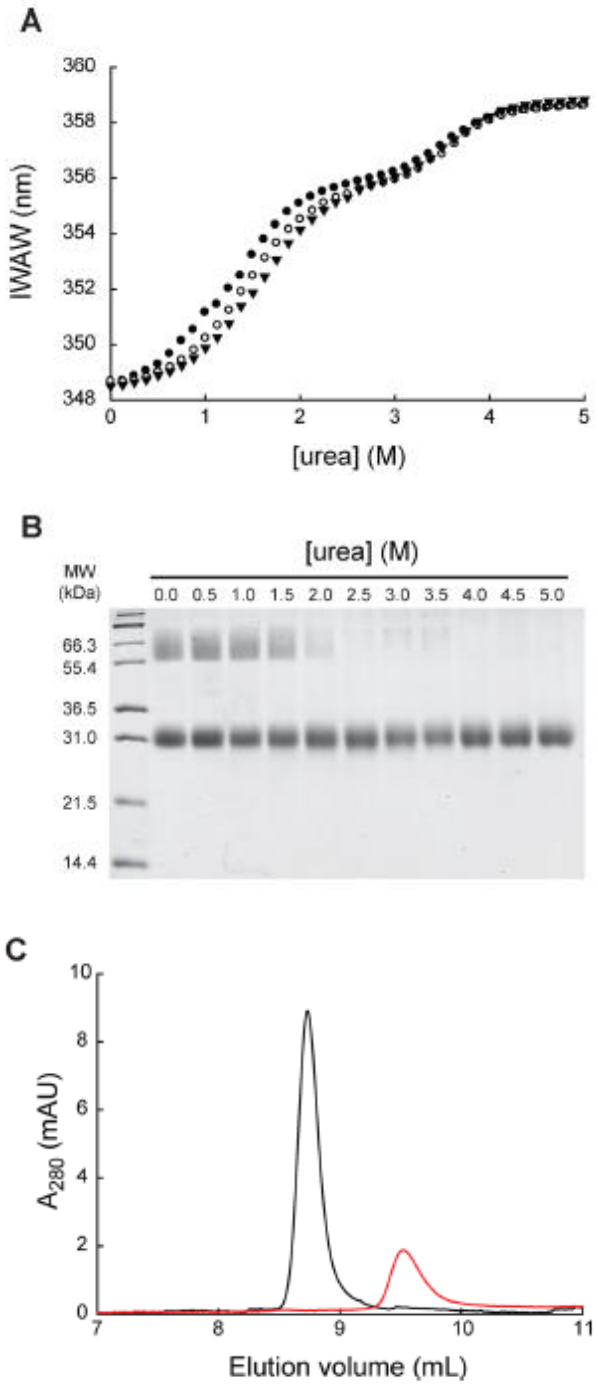


Figure 2.2 Equilibrium unfolding of dPGM monitored by circular dichroism.

Equilibrium unfolding of 0.030 mg/mL dPGM in varying concentration of urea was monitored by circular dichroism at 222 nm. Data were fit with Eq. 4 (solid line) and the resulting parameters are presented in Table 1.

Figure 2.3 Oligomeric state of the equilibrium intermediate.

(A) Equilibrium unfolding of dPGM at 0.01 mg/mL (●), 0.03 mg/mL (○) and 0.10 mg/mL (▼) monitored by tryptophan fluorescence. Intensity-weighted average wavelengths (IWAW) were calculated from fluorescence emission spectra. (B) Chemical cross-linking of 0.20 mg/mL dPGM in the presence of 0 to 5 M urea. MW indicates the molecular weights of standard proteins in the protein ladder. (C) Elution profiles of 0.03 mg/mL dPGM from size exclusion chromatography at 0 M urea (black) and 2.5 M urea (red). Elution was monitored by absorbance at 280 nm (A_{280}).


A
IWAW (nm)
360
358
356
354
352
350
348
0
1
2
3
4
5
[urea] (M)
B
[urea] (M)
MW
(kDa)
0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0
66.3
55.4
36.5
31.0
21.5
14.4
C
A_{280} (mAU)
10
8
6
4
2
0
7
8
9
10
11
Elution volume (mL)

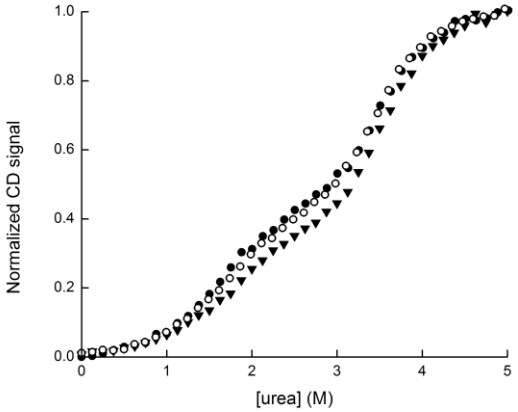


Figure 2.4 The concentration dependence of equilibrium unfolding of dPGM monitored by circular dichroism.

Equilibrium unfolding of 0.020 mg/mL (●), 0.030 mg/mL (○), and 0.10 mg/mL (▼) dPGM was monitored by circular dichroism at 222 nm. The CD signal was normalized for proper comparison.

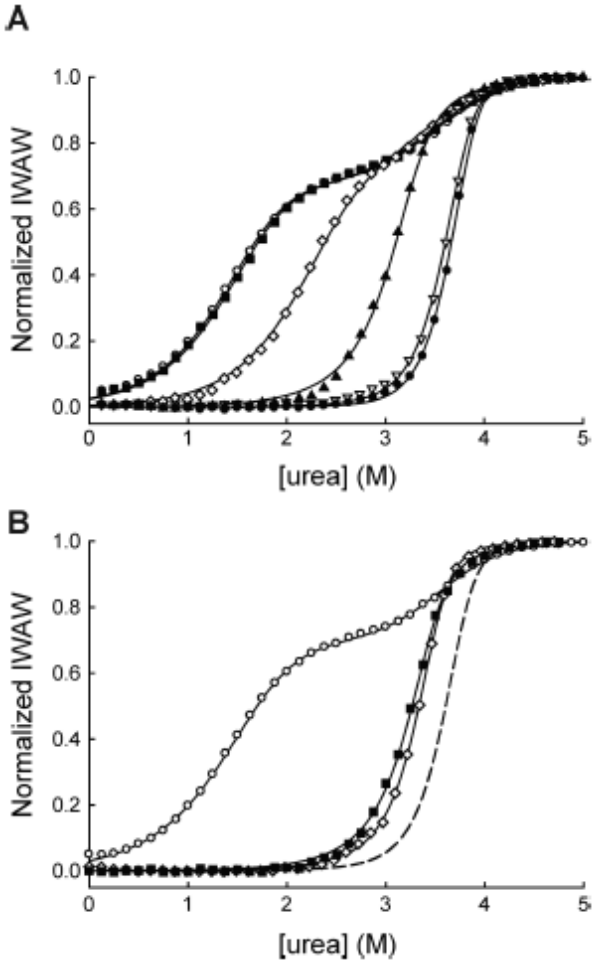


Figure 2.5 Effect of nucleotides on equilibrium unfolding of dPGM.

(A) Equilibrium unfolding of 0.03 mg/mL dPGM was examined in buffer only (○) and in the presence of adenosine (■), AMP (◇), ADP (▲), ATP (▽), and GTP (●) each at 1.0 mM. The data were globally fit to Eq. 4 (solid lines), and the fitting results are shown in Table 2. (B) Equilibrium unfolding of dPGM was also monitored in the presence of 0.050 M phosphate (◇) and 0.1 mM pyrophosphate (■). Both data sets and equilibrium unfolding data of dPGM without any ligand (○) were globally fit to Eq. 4 (solid lines), and the fitting results are shown in Table 2**.** The fitting results for dPGM unfolding in the presence of ATP is shown as a dashed line for reference. For proper comparison, intensity-weighted average wavelengths (IWAW) were normalized with Eq. 5.

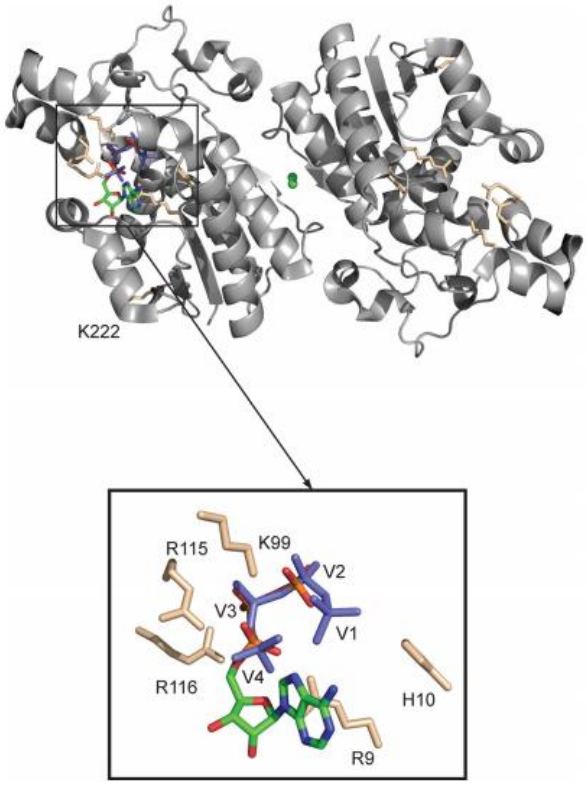


Figure 2.6 Predicted binding mode for ATP.

ATP was docked to dPGM (PDB entry 1E59) using GLIDE. ATP is shown docked in one active site in context of the full dPGM dimer. ATP is colored by element as follows: carbon (green), nitrogen (blue), phosphate (orange), oxygen (red). In the dimer structure, a chloride ion (green) is bound within the dimer interface. A magnified view of the active site reveals that the phosphates of ATP occupy the same positions as vanadate groups of the tetravanadate (light blue) that is bound in the crystallographic structure. The side chains are displayed for residues Arg9, Lys99, Arg115, and Arg116, which are predicted to hydrogen bond with ATP. The active-site histidine (His10) and Lys222 are also shown for reference.

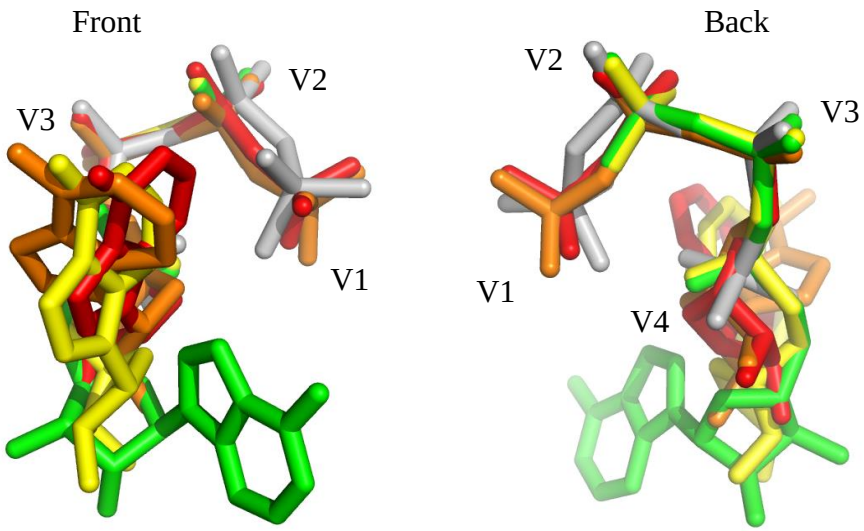


Figure 2.7 Comparison of the four most favorable poses of ATP docked to dPGM.

The top scoring poses for GLIDE docking of ATP to dPGM (PDB: 1E59) are displayed from the most favorable to least favorable as green, yellow, orange, and red, respectively. The tetravanadate found in the dPGM crystal structure is shown as grey, and vanadate positions, V1 through V4, are indicated for reference.

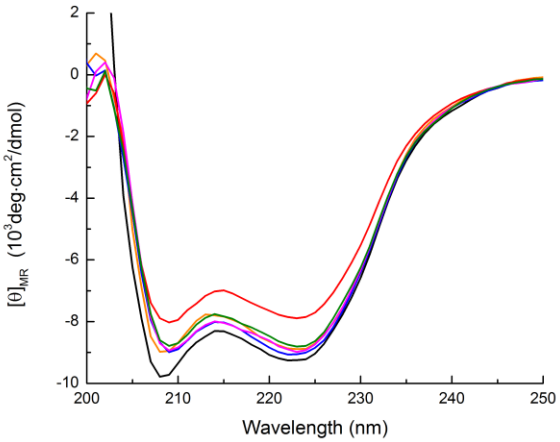


Figure 2.8 CD spectra of the dPGM variants.

The CD spectra of WT dPGM (black), R9A (orange), K99A (red), R115A (blue), R116A (magenta), K222A (green) dPGM at 25°C.

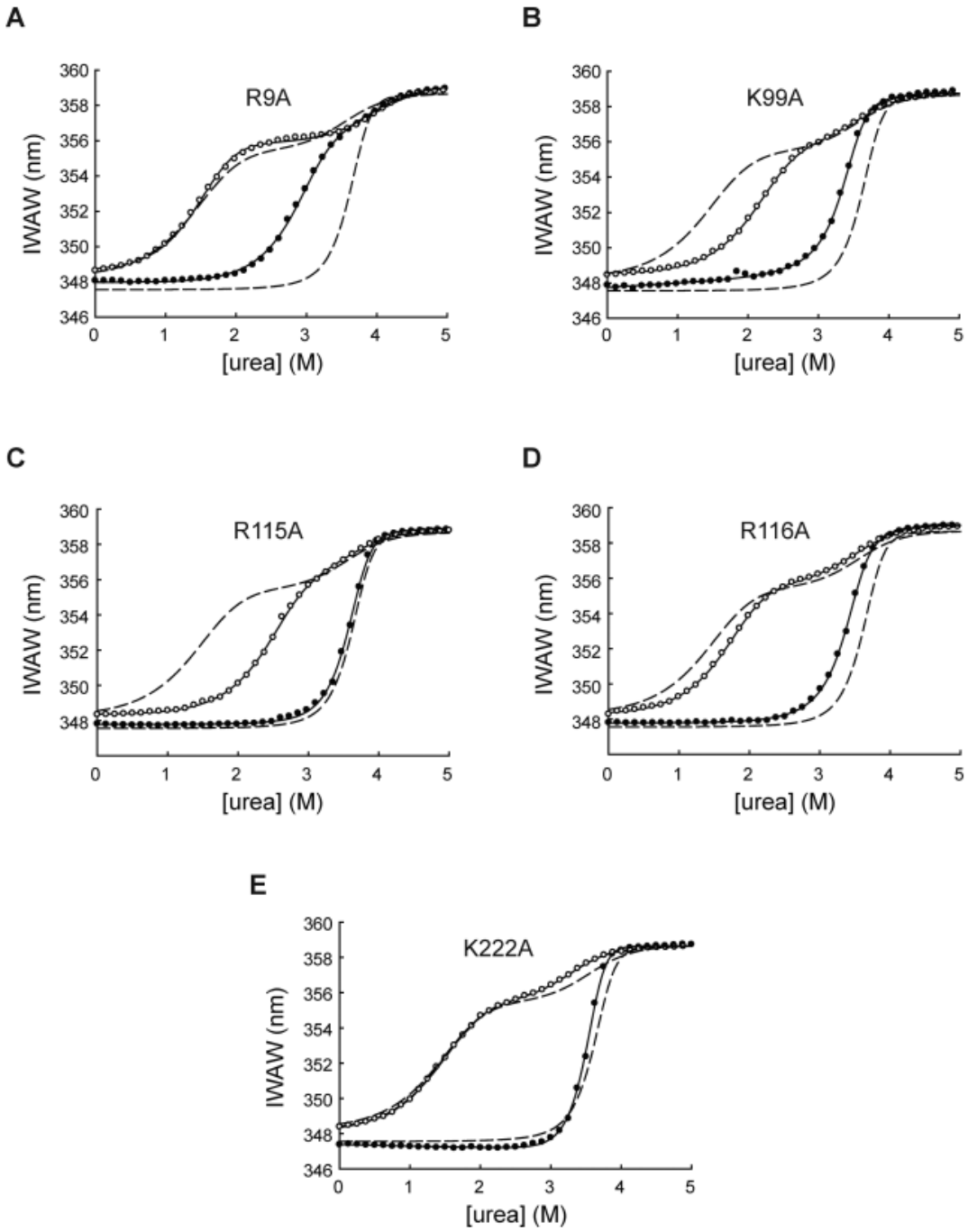


Figure 2.9 Effect of dPGM point mutations on the stabilizing effect of ATP.

Equilibrium unfolding of 0.03 mg/mL (A) R9A, (B) K99A, (C) R115A, (D) R116A and (E) K222A in the presence (●) and absence (○) of 1.0 mM ATP. The fit lines of wild type dPGM in the presence and absence of ATP are shown (dashed) for reference.

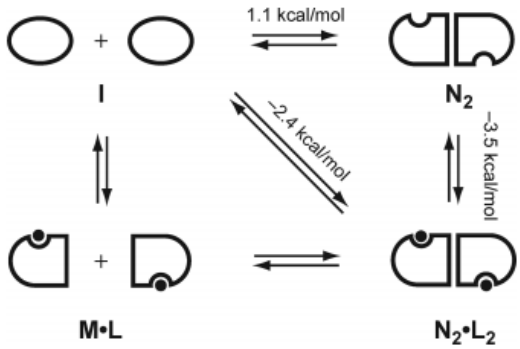


Figure 2.10 Mechanistic model for the energetic coupling of the dPGM active site and dimer interface

A simple thermodynamic cycle (A) illustrates the energetic coupling in dPGM. Apo dPGM exchanges between the intermediate state (I) and the native state (N_2). ATP binds to native dPGM (N_2•L_2) and reduces its free energy. If the dPGM active site was not energetically coupled to the interface, dPGM might form an ATP bound, folded monomer (M•L). The energetic penalty incurred for forming M•L ensures that dimerization and active site formation will occur together.

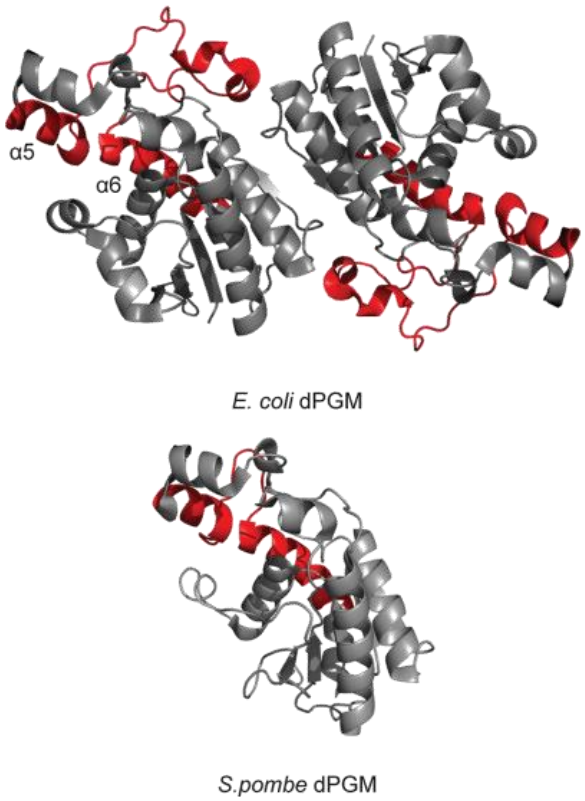


Figure 2.11 Putative structural link between the active site and the dimer interface.

The cartoon diagram of the crystallographic structure of *E. coli* dPGM (PDB entry 1E59) is shown with a plausible structural link (red) between the active site and the dimer interface. A long loop leads from active site helix α5 to the dimer interface and then to helix α6. For comparison, the equivalent region was shown in the cartoon diagram of the solution structure of *S. pombe* dPGM (PDB entry 1FZT), which is a monomeric protein.

CHAPTER 3. THE STRUCTURE OF A PARTIALLY UNFOLDED INTERMEDIATE OF *E. COLI* PHOSPHOGLYCERATE MUTASE

3.1 Summary

Previously, we determined that ligand binding to *Escherichia coli* cofactor-dependant phosphoglycerate mutase (dPGM), a homodimeric protein, is energetically coupled with dimerization. The equilibrium unfolding of dPGM revealed a stable, monomeric intermediate. Binding to several non-substrate metabolites stabilizes the dimeric native state over the monomeric intermediate, reducing the population of the intermediate. Both the active site and the dimer interface appear to be unfolded in the intermediate. We hypothesized that a loop containing residues 118−152 was responsible for the energetic coupling between the dimer interface and the distal active site and would be unfolded in the intermediate state. Here, we investigated the structure of the dPGM intermediate state by probing side-chain interactions and the peptide backbone. Equilibrium unfolding was used to assess the effect of truncating a side-chain on the stability of dPGM. By comparing the effect of a mutation on the global stability and the stability of the intermediate state, we determine an equilibrium φ-value to assess whether side-chain interactions are maintained or lost. Hydrogen/deuterium exchange coupled with mass spectrometry (HDX-MS) was used to investigate differences in the solvent exposure of the peptide backbone in the intermediate state and native state of dPGM. The results of

equilibrium φ-value analysis and HDX-MS reveal the least stable folding unit of dPGM that is unfolded in the intermediate state and links the active site to the dimer interface. The structure of the intermediate reveals how the cooperative network of residues in dPGM gives rise to the observed allostery between dimerization and ligand binding.

3.2 Introduction

Proteins fold as a group of interacting cooperative subunits to form conformational ensembles. The conformations that these ensembles contain range between fully folded and globally unfolded, with populations that are determined by the free energy of each conformation (Dill & Chan, 1997; Frauenfelder et al., 1991; Hilser et al., 2006). Allostery is an integral feature of protein where two binding sites, distant to one another, are energetically coupled so that binding at one site affects binding at the other (Gunasekaran et al., 2004; Hilser et al., 2012). It has been argued that all dynamic protein are allosteric or contain allosteric sites (Gunasekaran et al., 2004). Allosteric coupling is crucial for many molecular processes such as protein signaling (Chennubhotla et al., 2008; Ernesto Freire, 1999; Harford & Weitzman, 1975; Hynes, 2002; Yu et al., 2010) and has been exploited for drug development (Berger et al., 1999; Carlson & McCammon, 2000; Kenakin, 2002). A wide range of allosteric responses have been successfully modeled by considering the effects of ligand binding on the protein ensemble (Hilser & Thompson, 2007; Motlagh et al., 2014; Weber, 1972); binding at one site redistributes the conformational ensemble of the protein and shifts the population of conformations with high affinity for the second ligand, causing a change in the observed binding affinity for

the second ligand (Gunasekaran et al., 2004). The partitioning of protein into interacting, cooperative subunits is responsible for the allosteric effects on the conformational ensemble (Hilser & Thompson, 2007; Motlagh et al., 2014). Investigating the molecular basis of cooperativity and allostery within protein will better enable us to predict the effects of perturbations at a particular site in a protein on the conformational ensemble to accurately model, drug, and engineer proteins.

The cooperative nature of protein folding arises from the energetic coupling between residues (Ernesto. Freire et al., 1992; Gunasekaran et al., 2004; T. Liu et al., 2007). The coupling shapes a protein's energy landscape and causes the protein to fold as an assembly of individual cooperative units (Bai et al., 1995; Chamberlain et al., 1996; Hu et al., 2013). The energetic coupling between residues also enables allostery between distant sites. We previously found that the homodimeric protein *Escherichia coli* cofactor-dependant phosphoglycerate mutase (dPGM) shows an allosteric response between dimerization and ligand binding at a distal active site (Chapter 2). We found that dPGM populated a partially unfolded, monomeric intermediate at low urea concentrations, but in the presence of several non-substrate metabolites, the protein populations shifted to the native dimer (Chapter 2). Folding of the active site is cooperative with the dimerization of dPGM. The cooperative network of residues that mediates signal transmission between the active site of dPGM and the interface appears to be responsible for both the folding cooperativity and allostery between the two sites.

Previously, dPGM was identified from a proteome-wide screen for ATP binding protein in the *E. coli* proteome (Chang et al., 2012). This result was intriguing because dPGMs do not use ATP for their catalysis. Phosphoglycerate mutase proteins primarily function as a mutase (EC 5.4.2.11) and interconvert 3-phosphoglycerate and 2-phosphoglycerate in a reaction that is driven by equilibrium concentrations of the substrate. dPGMs take part in glycolysis and gluconeogenesis and the isomerization reaction proceeds in opposite directions in each pathway. The cofactor dependency of dPGM arises from the need to phosphorylate an active site histidine in order to function as a mutase. In its inactive form, dPGM is a phosphatase (EC 3.1.3.13) and reacts with the cofactor 2,3-bisphosphogylcerate to transfer a phosphate group to the active site histidine. An investigation into the effect of ATP on the energetics of dPGM revealed that some non-substrate metabolites (like ATP) can drastically stabilize the native protein over a partially unfolded intermediate (Chapter 2). By selectively stabilizing native dPGM, metabolite binding reduces the populations of intermediate and increases the conformational cooperativity of dPGM.

We hypothesized that a loop formed by residues 118-152, loop 118−152, that physically links the dPGM active site and dimer interface is the primary route of coupling between the two sites (Figure 2.11). Loop 118−152 is one of the least conserved regions amongst dPGM homologs that vary in quaternary structure between tetramer, dimer, and monomer (Bond et al., 2001). In the monomeric dPGM of *Schizosaccharomyces pombe*, the loop is truncated by 25 residues (J. Nairn et al., 1994; Uhrínová et al., 2001b). This suggests that

the loop plays a critical role in oligomerization. If loop 118−152 is energetically coupled with the active site and dimer interface, then it should be unfolded in the intermediate state. Allosteric effects are maximized when a region of the protein is unfolded or extremely dynamic in the absence of a bound ligand (Goodey & Benkovic, 2008; Hilser & Thompson, 2007; Motlagh, Li, Thompson, & Hilser, 2012). Then, to understand the cooperative and allosteric network in dPGM we could investigate the structure of the intermediate state to determine the unfolded regions. Several factors often complicate the study of partially unfolded protein intermediates: transient formation, low populations at equilibrium, indistinguishable spectroscopic signals, and the presence of highly dynamic or unstructured regions (Campos et al., 2004). The dPGM intermediate is relatively stable and readily apparent by tryptophan fluorescence which makes dPGM an interesting system to investigate the structural features that underlie allostery and its role in oligomerization.

Here, we test our hypothesis by elucidating structural differences between native dPGM and the partially unfolded, monomeric intermediate. The structural differences should reveal which regions of dPGM unfold and which remain folded in the intermediate state. By establishing which regions of dPGM fold upon dimerization we can assess how the structure of dPGM relates to the allostery between ligand binding and dimerization. To investigate the structure of the equilibrium intermediate of dPGM, we employ equilibrium φ-value (φ_{eq}) analysis (Campos et al., 2004; Sancho et al., 2002) and hydrogen/deuterium exchange coupled with mass spectrometry (HDX-MS). φ_{eq}-analysis

is similar to the traditional φ-value analysis, which was developed to probe the structure of transition states and kinetic intermediates (Fersht, Matouschek, & Serrano, 1992; Matouschek, Kellis, Serrano, Bycroft, & Fersht, 1990; Matouschek, Kellis, Serrano, & Fersht, 1989). Unlike traditional φ-value analysis, φ_{eq}-analysis uses only equilibrium information to quantify the effects of a mutation on the native state and intermediate state (Campos et al., 2004). By φ_{eq}-analysis we probe the side-chain contacts in the intermediate state and by HDX-MS we probe the solvent exposure of the peptide backbone. Our data reveal a cooperative folding unit that differentiates the native and intermediates states of dPGM and provide a framework to understand the observed allosteric effects.

3.3 Results

3.3.1 Probing the structure the of the dPGM intermediate through point mutations

To test our hypothesis that unfolding in the intermediate state was localized to loop 118-152, we probed the side-chain contacts of native dPGM and the intermediate state using point mutations (Table 3.1, Figure 3.1). Most mutations were designed to remove only a portion of the side chain and should result in the loss or weakening of several interactions. From the effect of a mutation on the stability of the native state, $\Delta\Delta G_{U-N}°$, and the intermediate state, $\Delta\Delta G_{U-I}°$, we should be able to determine whether the mutated residue maintains its native side-chain contacts in the intermediate state (folded) or losses them (unfolded) (Campos et al., 2004; Fersht et al., 1992). To quantify the effect of the

mutation we will calculate the ratio between $\Delta\Delta G_{U-I}°$ and $\Delta\Delta G_{U-N}°$, which is the equilibrium φ-value (Campos et al., 2004; Sancho et al., 2002):

$$\varphi_{eq} = \frac{\Delta\Delta G_{U-I}°}{\Delta\Delta G_{U-N}°} \tag{3.1}$$

where $\Delta\Delta G_{U-N}°$ is the change in free energy between the native state and unfolded state per monomer due to a mutation. The φ_{eq}-analysis is analogous to traditional φ-value analysis, which reports on the side-chain contacts in transition states (Fersht et al., 1992; Matouschek et al., 1990; Matouschek et al., 1989). The interpretation of φ_{eq} and φ-values is alike; a value of 1 means the native side-chain contacts are maintained in the intermediate and a value of 0 means the side-chain contacts are lost in the intermediate (Campos et al., 2004; Matouschek et al., 1989).

To obtain φ_{eq}-values, we perform equilibrium unfolding of the dPGM variants in order to accurately measure the stability of the native state and intermediate state. In 20 mM buffer only, dPGM populates the intermediate state at low urea concentrations. The mutations we made are designed to be destabilizing, which means the intermediate will be populated at lower urea concentrations that for wild-type dPGM and the native state may not be clearly discernible. To obtain the most robust thermodynamic parameters, the native state, intermediate state, and unfolded state should be well separated. Overlapping signals and the lack of a clear native state signal will result in erroneous values. We have

characterized the stabilizing effect NaCl on wild-type dPGM at concentrations of 0 M to 0.5 M (Figure 4.5), and found that Cl^- stabilizes the native state over the intermediate state. NaCl can be used to better differentiate the two states by increasing the free energy difference between them. We did not observe indications of structural change in wild-type dPGM induced by NaCl (Chapter 4), so we will assume that NaCl will not alter the conformation of the dPGM variants. When the native state signal cannot be well resolved in the equilibrium unfolding of a dPGM variant, we will utilize NaCl to stabilize the native state and separate the native state and intermediate states signals. The dPGM variants will be compared with wild-type under the same salt condition.

We generated nineteen dPGM variants with point mutations distributed throughout the protein (V8A, T22S, V27A, V34A, L43A, T56S, I63V, L66V, W67A, V80A, L86V, L94V, V111A, L125A, L147A, V160A, L169A, V191A, I235A) to probe most of the secondary structural units of dPGM (Table 3.1, Figure 3.1). The residue numbering and secondary structure assignment were determined by Bond, White, and Hunter (Bond et al., 2001). All assigned secondary structural units, ten α-helices, four 3_{10}-helices, and six strands of the core β-sheet, are involved in interactions with the residues that we chose for mutagenesis. Residues were determined to interact through Van der Waals interactions due to their proximity to one another (< 4 Å). Three mutations (I63V, W67A, and V80A) truncate side chains that interact with the opposing dPGM monomer. Trp67 can also form a hydrogen bond with a Cl^- buried within the dimer interface. Together, the mutations should provide a global survey of the structural elements of dPGM.

3.3.2 Characterization of the fold of the dPGM variants

All of the dPGM variants are able to form the secondary structure of the native state as determined by circular dichroism (CD) (Figure 3.2). Several dPGM variants were analyzed without added NaCl (V8A, T56S, L66V, and V80A dPGM), which reflects the variants under our least stabilizing condition (Figure 3.2A). Without salt, the CD spectra of the dPGM variants do not show any significant variance from wild-type dPGM. The remaining variants were analyzed under a more stabilizing condition with the presence of 0.10 M NaCl (Figure 3.2B); as expected the dPGM variants have spectra similar to wild-type dPGM in the presence of salt. Therefore, under the conditions we will investigate, the dPGM variants adopt the native secondary structure.

Size exclusion chromatography with the dPGM variants revealed that the mutations did not significantly affect the hydrodynamic radius or oligomeric state of dPGM (Figure 3.3). L86V and V111A dPGM were monitored without added salt, and they eluted at the same volume as wild-type dPGM (Figure 3.3A). The mobile phase contained 0.15 M NaCl for the other dPGM variants and the dPGM variants elute with similar volumes to wild-type. The average elution volume for all dPGM variants and wild-type dPGM in the presence of 0.15 M NaCl is 60.8 ± 0.9 mL. The small standard deviation of the average elution volume reflects the narrow distribution of elution volumes for the dPGM variants. No mutations appear to have disrupted dimer formation or substantially affected the hydrodynamic radius of dPGM; the dPGM variants all appear to fold to the native state.

3.3.3 Equilibrium unfolding of the dPGM variants

For equilibrium unfolding analysis the dPGM variants are divided into two groups; the variants in Group 1 have a $\Delta\Delta G_{I-N}^{\circ} \leq$ -0.8 kcal/mol in the presence of 0.10 M NaCl and the variants in Group 2 have a $\Delta\Delta G_{I-N}^{\circ} >$ -0.8 in the presence 0.10 M NaCl. The equilibrium unfolding of dPGM variants in Group 1 (V8A, V34A, L43A, T56S, I63V, L66V, V80A, L86V, V111A, V160A, V191A, and I235A dPGM) is analyzed in the presence of 0.10 M NaCl for φ_{eq}-analysis (Table 3.2, Figure 3.4). Due to the destabilizing nature of the mutations, addition of NaCl was necessary to resolve the native state (data not shown). The variants in Group 2 (T22S, V27A, W67A, L94V, L125A, and L147A dPGM) have large errors associated with their φ_{eq}-values due to uncertainty in $\Delta\Delta G_{U-I}^{\circ}$ (and ΔG_{U-I}°) and the small magnitude of $\Delta\Delta G_{U-I}^{\circ}$; changes in ΔG_{U-I}° within the fitting error result in huge changes to the φ_{eq}-value for dPGM variants in Group 2. Also, the W67A, L125A, and L147A mutations reduced the stability of the native state so that the intermediate state is already partially populated at 0 M urea in the presence of 0.10 M NaCl and the spectroscopic signal of the native state is not clear (Figure 3.5). To obtain more robust φ_{eq}-values we needed to clearly resolve the signal of the native state and the intermediate state and obtain accurate thermodynamic parameters; obtaining more accurate thermodynamic parameters could be achieved by repeating equilibrium unfolding in triplicate at an ideal NaCl concentration, but the concentration is likely to vary for each dPGM variant. The concentration of NaCl also affects the resolution of the two equilibrium unfolding transitions of wild-type dPGM (Figure 4.5), our reference state for φ_{eq}-value calculation. As [NaCl] is increased, the two equilibrium unfolding

transitions converge and the intermediate state becomes harder to discern. Because wild-type dPGM is more stable than the dPGM variants, the optimal NaCl concentration to resolve the native state and the intermediate state for a given variant is likely to differ from wild-type dPGM. For the dPGM variants in Group 2 and L169A dPGM of Group 1, we performed equilibrium unfolding under three difference NaCl concentrations and fit the data globally (Table 3.3, Figure 3.5). By varying the NaCl concentration we are able to incorporate conditions that resolve the native state and intermediate state well for both, the dPGM variant and wild-type dPGM.

The equilibrium unfolding of the dPGM variants in Group 1 in the presence of 0.10 M NaCl proceeds via two unfolding transitions like wild-type dPGM (Figure 3.4, 3.5G). All the mutations except L169A and I235A shift the apparent midpoints of the two unfolding transitions, C_{m1} and C_{m2}, to lower urea concentrations. L169A and I235A dPGM only show a reduced $C_{m2.}$ To obtain thermodynamic parameters, we fit the equilibrium unfolding of each dPGM variant with eq 2.4 (Table 3.2, 3.3), which is derived from the three-state model shown scheme 1 in the experimental procedures. When fitting was performed, we assumed that the *m*-value for the global unfolding of dPGM is not affected by mutation, but is always equal to the value of wild-type dPGM (6.5 kcal mol^{-1} M^{-1}). The *m*-value is the slope of the urea dependence of $\Delta G°$ and is proportional to the change in solvent exposed surface area upon folding or unfolding (Myers et al., 1995). Then, we make the assumption that the mutations did not significantly alter the amount of surface exposed upon global unfolding when compared to wild-type dPGM. The fitting shows

that most of the mutations in Group 1, V34A, L43A, T56S, I63V, L66V, V80A, V160A, V191A, I235A, have minimal effect on the free energy difference between native dPGM and the intermediate state, $\Delta G_{I-N}°$. The primary effect of the mutations is a reduction in free energy difference between the intermediate state and unfolded state, $\Delta G_{U-I}°$. Per monomer, the global stability ($\Delta G_{U-N}°$) of dPGM is $(\Delta G_{I-N}°)/2 + \Delta G_{U-I}°$. When $\Delta G_{I-N}°$ is minimally affected by mutation, a reduction in $\Delta G_{U-I}°$ affects $\Delta G_{U-N}°$ almost equally. All the dPGM variants in Group 1 have a lower global stability than wild-type dPGM. The mutations in Group 1 destabilized both the native state and the intermediate state with respect to the globally unfolded state of dPGM.

Several dPGM variants in Group 1 are outliers as they show a significant $\Delta\Delta G_{I-N}°$: V8A, L86V, V111A, and L169A dPGM. The effect of the L86V mutation was unique; L86V dPGM significantly reduced $\Delta G_{I-N}°$ ($\Delta\Delta G_{I-N}°$ = -1.43 kcal/mol). The intermediate state and native state are closest in energy for L86V dPGM, and L86V dPGM shows the lowest C_{m1} value (Figure 3.4) out of the variants in Group 1. The L86V mutation reduces $\Delta G_{U-N}°$ more than $\Delta G_{I-N}°$, which suggests that Val86 makes stronger interactions in the native state than the intermediate state. The V8A, V111A, and L169A mutations significantly increase $\Delta G_{I-N}°$, but they greatly reduce $\Delta G_{U-I}°$ so that $\Delta G_{U-N}°$ is still lower than for wild-type dPGM. Equilibrium unfolding of L169A dPGM (Figure 3.5G) appears to take place as one cooperative transition in the presence of 0.10 M NaCl. To separate the fluorescence signals of the intermediate state and the native state we investigated the equilibrium unfolding of L169A dPGM in the presence of 0.020 and 0.050 M NaCl also.

As the salt concentration is reduced, the two unfolding transitions for L169A dPGM become apparent (Figure 3.5G). Global fitting of the L169A dPGM resulted in thermodynamic parameters with low uncertainties while a single fit for equilibrium unfolding of L169A dPGM in the presence of 0.10 M NaCl did not (data not shown). The case of L169A dPGM highlights the use of a range of NaCl concentrations.

For each dPGM variant in Group 1, m_{I-N} is greater than the value for wild-type (2.61 kcal mol^{-1} M^{-1}) and m_{U-I} is less than the value of wild-type (1.94 kcal mol^{-1} M^{-1}). Unfolding of the native state appears to expose more surface area for the dPGM variants than for wild-type dPGM. The mutations may have caused further unfolding in the intermediate state. The large m_{I-N} for V8A, L86V, V111A, and V160A dPGM (m_{I-N} > 3.6 kcal mol^{-1} M^{-1}) strongly supports that these variants have a less folded intermediate state than wild-type dPGM. The smallest changes in *m*-values were observed with I235A dPGM, the variant with the smallest $\Delta\Delta G_{U-I}°$ out of the Group 1 dPGM variants. If removing a side chain increases the amount of unfolding in the intermediate state, then the wild-type side chain must form important interactions and remain folded in the intermediate state. The *m*-values of the dPGM variants in Group 1 suggest that the mutated residues remain folded in the intermediate state.

The equilibrium unfolding of the dPGM variants in Group 2 (T22S, V27A, W67A, L94V, L125A, and L147A dPGM) was monitored at three different NaCl concentrations (Figure 3.5). All the dPGM variants clearly show two equilibrium unfolding transitions under all

salt conditions. Notably, removal of a tryptophan from the W67A mutation shifted the fluorescence signal of native state from about 348 nm to 346 nm. For W67A, L125A, and L147A dPGM, the apparent C_{m1} in the presence of 0.10 M NaCl is less than 1.0 M urea. The three variants partially populate the intermediate state at 0 M urea, which obscures the fluorescence signal of the native state. Higher NaCl concentrations were used to increase $\Delta G_{I-N}°$ and resolve the native state signal of W67A, L125A, and L147A dPGM. Equilibrium unfolding of W67A, L125A, and L147A dPGM was conducted in the presence of 0.10 M, 0.20 M, and 0.50 M NaCl (Figure 3.5C, E, F). In the presence of 0.10 M NaCl, T22S, V27A, and L94V dPGM have an apparent C_{m1} of 1.5 M urea or greater. There was no need to better resolve the native state for these dPGM variants and so lower NaCl concentrations were used to better separate the unfolding transitions. The equilibrium unfolding of T22S, V27A, and L94V dPGM was monitored in the presence of 0.020 M, 0.050 M, and 0.10 M NaCl (Figure 3.5A, B, D). In contrast to Group 1, the dPGM variants in Group 2 tend to show no change in C_{m2}, only a reduced C_{m1}.

Global fitting was performed on each variant to obtain the thermodynamic parameters (Table 3.3). L125A dPGM was consistently fit with a much lower m_{I-N} than wild-type and when the global *m*-value is fixed at 6.5 kcal mol^{-1} M^{-1}, m_{U-I} and $\Delta G_{U-I}°$ appeared to be overestimated. Equilibrium unfolding of L125A dPGM was fit globally without fixing any parameters (Table. 3.3, Figure 3.5E). Global fitting revealed that the dPGM variants in Group 2 all reduce $\Delta G_{I-N}°$ by at least 1 kcal/mol and have small or insignificant effects on $\Delta G_{U-I}°$ (Table 3.3). The W67A mutation reduced $\Delta G_{U-I}°$ by about 0.8 kcal/mol, the

greatest effect of the mutations in Group 2. The average $\Delta\Delta G_{I-N}°$ for W67A, L125A, and L147A dPGM are -4.1 ± 0.2, -4.9 ± 0.1, and -3.6 ± 0.2 kcal/mol, respectfully; these mutations show a large destabilizing effect on the native state. The L125A mutation reduced $\Delta G_{I-N}°$ more than any other mutation that we tested. For T22S, V27A, and L94V dPGM the average $\Delta\Delta G_{I-N}°$ is -1.26 ± 0.06, -1.1 ± 0.2, and -2.1 ± 0.2 kcal/mol, respectfully. The T22S, V27A, and L94V mutations clearly destabilize the native state, but to a lesser degree than those probed by the W67A, L125A, and L147A mutations.

The dPGM variants in Group 2 were fit with *m*-values similar to wild-type, except for the most destabilized dPGM variant: L125A dPGM. For L125A dPGM, m_{I-N} is 1.97 ± 0.4 kcal mol^{-1} $M^{-1}$and m_{U-I} is 2.05 ± 0.8 kcal mol^{-1} M^{-1}. While m_{U-I} is not significantly affected by the mutation, the m_{I-N} for L125A dPGM is about 25% lower than that of wild-type, suggesting that less surface is buried in the native conformation of L125A dPGM. The interactions made by Leu125 appear critical for the native form of dPGM but of little consequence to the intermediate. For T22S and L94V dPGM the *m*-values are the same as wild-type dPGM, within error. Collectively, the *m*-values support that the dPGM variants in Group 2 undergo a similar change in surface area upon unfolding as wild-type dPGM. Unlike the mutations in Group 1, the T22S, V27A, W67A, L94V, L125A, and L147A mutations did not appear to promote further unfolding of the dPGM intermediate.

3.3.4 φ_{eq}-value analysis

With thermodynamic parameters determined for the dPGM variants, φ_{eq}-analysis was used to assess whether side-chain contacts are lost or maintained in the intermediate state. The φ_{eq}-value as defined in eq 3.1 is cast in terms of $\Delta\Delta G_{U-N}°$. Curve fitting returns values for $\Delta G_{I-N}°$ and $\Delta G_{U-I}°$ and in practice, φ_{eq}-values were calculated with $\Delta\Delta G_{I-N}°$ and $\Delta\Delta G_{U-I}°$ following equation 3.2 (Table 3.2, 3.3):

$$\varphi_{eq} = \frac{\Delta\Delta G_{U-I}°}{\frac{\Delta\Delta G_{I-N}°}{2} + \Delta\Delta G_{U-I}°} \tag{3.2}$$

The φ_{eq}-values for the dPGM variants in Group 1 fall near 1.0 (Table 3.3, 3.4); φ_{eq}-values for the dPGM variants in Group 2 fall near 0 (Table 3.4). A φ_{eq}-value of 0 or 1 indicates the loss or preservation of side-chain contacts in the intermediate state, respectfully.

Of the dPGM variants in Group 1, V8A, V34A, L43A, T56S, I63V, L66V, V80A, V160A, V191A, and I235A have φ_{eq}-values between 0.8 and 1.2. L86V dPGM is unique with a fractional φ_{eq} of 0.73 ± 0.09. V111A and L169A dPGM have φ_{eq}-values of 1.3 ± 0.2 and 1.43 ± 0.02, respectfully. The low uncertainty for the φ_{eq}-value of L169A dPGM is achieved by averaging values from the global curve fitting (Table 3.3). The uncertainty in the φ_{eq}-value of I235A dPGM was the largest (± 0.5) and is a result of the mutation having a small effect on the stability of dPGM. In general, the uncertainty of the φ_{eq}-values decreases as the absolute value of $\Delta\Delta G_{U-I}°$ increases. The high φ_{eq}-values of

V111A and L169A dPGM signify that the mutated residues convey more stability to the intermediate state than the native state. Val111 and Leu169 could form additional side-chain contacts in the intermediate state which would explain the high φ_{eq}-values. The fractional φ_{eq}-value of L86V dPGM suggests either that a fraction of the native side-chain contacts are lost or that the residue experiences conformational heterogeneity in the intermediate state (Fersht et al., 1992). Leu86 makes van der Waals interactions with the peptide backbone of Tyr55 and Thr56 and interacts with Tyr163. T56S dPGM has a φ_{eq}-value of 1.1 ± 0.2. Given the φ_{eq}-value, Thr56 appears to remain folded in the intermediate state and likely, Tyr55 as well. V160A dPGM makes van der Waals interactions with Lys85 and its φ_{eq}-value (0.9 ± 0.1) reveals that Val160 should be folded in intermediate state. Tyr163, which is located in the same α-helix as Val160, is also expected to be folded in the intermediate state. The secondary structural elements that Leu86 interacts with appear to remain folded in the intermediate state, but the 3_{10}-helix 1, containing Leu86, does lose contacts with the bound dPGM monomer in the intermediate state. The 3_{10}-helix may be more dynamic in the intermediate state of L86V dPGM than wild-type dPGM as a result. We think increased conformational heterogeneity accounts for the small deviation in the φ_{eq}-value of L86V dPGM from a value of 1. The φ_{eq}-values of all the dPGM variants in Group are near 1 and suggest that the mutated residues maintain side-chain contacts and remain folded in the intermediate state.

The six dPGM variants in Group 2, T22S, V27A, W67A, L94V, L125A, and L147A dPGM all have φ_{eq}-values close to zero ($\varphi_{eq} \leq 0.25$) as shown in Table 3.4. The φ_{eq}-value

of L125A dPGM is the most negative at -0.20 ± 0.05 and is less than 0 because the L125A mutation slightly enhanced $\Delta G_{U-I}°$ while reducing global stability. However, the $\Delta\Delta G_{U-I}°$ for L125A dPGM under different NaCl concentrations are not significantly different from 0. Despite the very large $\Delta\Delta G_{I-N}°$ for L125A dPGM, the φ_{eq}-value is still sensitive to small changes in $\Delta G_{U-I}°$ that fall within the uncertainty in $\Delta G_{U-I}°$. For proteins with intermediate states that are less discernible than the intermediate of dPGM, determining robust, low φ_{eq}-values by urea-induced equilibrium unfolding would be extremely challenging. The low φ_{eq}-values of the dPGM variants in Group 2 signify that side-chain contacts are lost in the intermediate state and suggest that the mutated residues are unfolded in the intermediate state. However, there are indications that Trp67 is not unfolded in the intermediate state despite its low φ_{eq}-value. W67A dPGM has the largest φ_{eq}-value (0.25 ± 0.02) of the dPGM variants in Group 2, indicative of the fact that the W67A mutation did reduce the stability of the intermediate state slightly. Analysis of Trp67 in the crystal structure of dPGM reveals that the side chain does make van der Waals interactions with Ile63 and His64 within the same α-helix (Table 3.1). I63V dPGM has a φ_{eq}-value of 0.9 ± 0.2 and L66V dPGM had a φ_{eq}-value of 1.0 ± 0.1; the high φ_{eq}-values of these two dPGM variants supports that α-helix 3 is folded in the intermediate state and, consequently, Trp67 should maintain a portion of its side-chain contacts in the intermediate state. Still, Trp67 loses contacts that are formed between dPGM monomers and may lose the ability to form hydrogen bond with Cl^- in the intermediate state which gives rise to the low φ_{eq}-value. Other than Trp67, the dPGM variants in Group 2 do appear to probe residues that are unfolded in the intermediate state of dPGM.

Mapping the φ_{eq}-values of all the dPGM variants to the crystal structure of dPGM (Figure 3.7A, B) clearly reveals a cluster of low φ_{eq}-values around loop 118−152 and a cluster of high φ_{eq}-values in the monomer core and along the dimer interface. In addition to loop 118−152, loop 18−30 and loop 89−98 contain residues with φ_{eq}-values near 0. The loss of side-chain contacts indicated by the φ_{eq}-values suggests that these loops are unfolded in the intermediate state. Interestingly, W67A is the only mutation near the core of the dimer interface that resulted in a low φ_{eq}-value. Both I63V and V80A mutations resulted in the loss of interactions between dPGM monomers (Table 3.1) but still yield φ_{eq}-values near 1. Apparently helices 3 and strand C, which form the majority of the dimer interface, remain folded upon monomer dissociation. The β-sheet running through the dPGM core and α-helices 2, 4, 5, 6, 7, 8, and 10 all contain high φ_{eq}-value mutations and appear to be folded in the intermediate state. No mutations were made in Strands E and F, but the L43A, V191A, and I235A mutations all probed side chains that interact with the Strands (Table 3.1) and result in φ_{eq}-values near 1. The structure of dPGM reveals that 3_{10}-helix 1, in which Leu86 resides, is on the border between the loop region with low φ_{eq}-values and the core with high φ_{eq}-values (Figure 3.7A, B). The positioning of 3_{10}-helix 1 reveals that a significant portion of the interactions made by the 3_{10}-helix are formed with the loops that appear to be unstructured in intermediate state. The structural analysis supports the idea that conformational heterogeneity in the intermediate state gives rise to the fractional φ_{eq}-value of L86V dPGM. The view that arises from φ_{eq}-analysis is that unfolding in the intermediate state of dPGM is localized to one region that is formed by several interacting loops linking α-helices 4 and 5 of the active site to the dimer interface.

3.3.5 HDX-MS

In addition to probing side-chain interactions, we utilized hydrogen/deuterium exchange coupled with mass spectrometry (HDX-MS) to monitor changes in the peptide back bone of dPGM in the intermediate state at the peptide level. Peptides that report little exchange are protected from solvent and should be folded while a large amount of exchange indicates unfolding. Samples of wild-type dPGM were prepared under conditions to maximize the population of the native state, intermediate state, or the unfolded state. With native dPGM, HDX-MS was conducted in the presence and absence (data not shown) of 1.0 mM ATP; we detected little difference in deuterium uptake between the two conditions and used dPGM in the presence of ATP as the condition to maximize the population of the native state. Each state of the protein was placed in a deuterated buffer to allow deuterium exchange for two minutes. Samples were quenched by addition of quench buffer and flash freezing them with liquid nitrogen. Quenched samples were thawed and immediately analyzed using LC/MS with online pepsin digestion to obtain high-resolution information at the peptide level. The deuterium incorporation was quantified from differences in peptide mass relative to an undeuterated control (Figure 3.6). The peptides we analyzed were found in all samples and cover 93% of the protein sequence. HDX-MS was repeated in triplicate for dPGM in each state and the resulting % deuterium uptake of each peptide was averaged to obtain the mean % deuterium uptake (Figure 3.6A).

The mean % deuterium uptake of dPGM in the native, intermediate, and unfolded state is much lower for the native state in the majority of the peptides (Figure 3.6A). The unfolded state samples show deuterium incorporation that is greater than or equal to the native state and intermediate state samples for all peptides. Overall, the mean % deuterium uptake is greatest near the C-terminus for dPGM in each state. Peptide 114−125, a portion of loop 118−152, and peptides 199−203 and 235−249 show high deuterium uptake in all three states of dPGM. This suggests that the residues covered by these peptides are located in dynamic regions in all states of dPGM. To better analyze the differences between the native and intermediate states we calculated the relative deuterium uptake of each peptide (Figure 3.6B) with respect to the peptide in the unfolded state:

$$f_{\%D} = \frac{\%D_x}{\%D_U} \tag{3.3}$$

where $\%D_x$ is the mean % deuterium uptake of a peptide from state x and $\%D_U$ is the mean % deuterium uptake of the peptide from the unfolded state. An $f_{\%D}$ of 1 signifies deuterium uptake equal to the unfolded state and a value of 0 would signify no relative uptake. Several peptides from the intermediate state of dPGM show protection from deuterium exchange with a relative uptake of less than 0.6: peptides 1−7, 169−186, 188−197, and 203−216. Peptide 216−233 also shows some protection with an $f_{\%D}$ of 0.66. The mean % deuterium uptake of the five peptides is at least 20% lower than in the

unfolded state. The corresponding secondary structural units that show some protection from HDX in the intermediate are α-helices 1 and 8 and strands A, D, E, and F. The protected residues are predominantly in the dPGM core. The remaining peptides from the intermediate state show a high $f_{\%D}$. In the native state, several peptides are highly protected from HDX and have an $f_{\%D}$ of less than 0.25: peptides 30−52, 53−67, 70−77, 71−94, 147−163, 169−186, and 188−197. Another group of peptides from the native state has an $f_{\%D}$ between 0.25 and 0.50 (peptides 1−7, 91−113, 126−142, and 203−216). The peptides from the native state showing a high deuterium uptake with an $f_{\%D}$ greater than 0.5 are peptides 6−24, 114−125, 216−233, 199−203, and 235−249.

Several peptides show a large change in the relative deuterium uptake ($\Delta f_{\%D}$) between the native state and the intermediate state, $\Delta f_{\%D} = f_{\%D}(\text{native}) - f_{\%D}(\text{intermediate})$. A large $\Delta f_{\%D}$ suggests that the residues covered by the peptide experience much different solvent exposure in the intermediate state than in the native state. Residues that are unfolded in the intermediate state are expected to fall within a peptide showing a large $\Delta f_{\%D}$. The largest $\Delta f_{\%D}$ observed is -0.81 for peptide 53−67. For the six peptides spanning residues 30 to 113 and peptides 126−142 and147−163 the value of $\Delta f_{\%D}$ is greater in magnitude than 0.55; all of the peptides showing a large difference in $\Delta f_{\%D}$ also show a dramatically lower mean % deuterium uptake for the native state than for the intermediate and unfolded states. With the exception of peptide 114−125, which shows high deuterium uptake in every state of dPGM, the peptides showing a $\Delta f_{\%D} > 0.55$ span a continuous sequence.

To understand the structural differences between the native state and intermediate state of dPGM that are captured by HDX-MS, the $\Delta f_{\%D}$ values are mapped to the dPGM crystal structure as shown in Figure 3.7C and D. The peptides with a large $\Delta f_{\%D}$ are localized to the dimer interface (α-helices 2 and 3, strand C), the active site (α-helices 4 and 5), loops 89−98 and 118−152, α-helix 6. A portion of loop 118−152 covered by peptide 114−125 has a low $\Delta f_{\%D}$ due to the fact that deuterium uptake by peptide 114−125 was high for all states of dPGM. The secondary structures covered by peptides with a large $\Delta f_{\%D}$ and peptide 114−125 form a large, continuous region of dPGM that is either unfolded in the intermediate state or experiences a large change in solvent exposure upon dissociation. Our φ_{eq}-analysis corroborates that loops 89−98 and 118−152 are unfolded in the intermediate state. Most of the region covered by peptides with a large $\Delta f_{\%D}$ was found to retain side-chain contacts in the intermediate state (α-helices 2, 3, 4, 5, and 6, loop 83−88, strand C), suggesting that the region is not unfolded. The secondary structural units involved most likely undergo a large change in solvent exposure or dynamics upon transitioning the intermediate state. The core of the dPGM monomer containing the majority of the β-sheet and α-helices 1, 8, 9, and 10 shows low $\Delta f_{\%D}$-values, which supports the result from φ_{eq}-analysis that the core of dPGM is folded in the intermediate state.

3.4 Discussion

Through use of φ_{eq}-analysis and HDX-MS we describe a low resolution structure of the partially unfolded monomeric intermediate of dPGM (Figure 3.8). The intermediate state

structure can be grouped into three discrete regions. One region contains the secondary structural features that appear unfolded in the intermediate state and is the least stable folding unit of dPGM (red in Figure 3.8). The other two regions both appear folded in the intermediate state, but one shows a large $\Delta f_{\%D}$ (orange in Figure 3.8) due to changes in solvent exposure while the other shows a low $\Delta f_{\%D}$ (gray in Figure 3.8). We have previously found that the active site of dPGM is energetically coupled with the dimer interface and both appear to be disrupted in the intermediate state (Chapter 2). We hypothesized that loop 118−152 was responsible for linking the active site and dimer interface and would unfold in the transition from the native state to the intermediate state. Here, we showed that loop 118−152 in addition to two other intervening loops link the active site and dimer interface through a network of tertiary interactions. Three loops (loop 19−29, 88−98, and 118−152) and three α-helices (α-helix 1, 4, and 5) form the least stable, cooperative folding unit of dPGM. Interactions between these secondary structural features physically and energetically link the active site (α-helices 4 and 5) and dimer interface such that a portion of each will fold and unfold in unison. The energetic linkages between the residues involved in the least stable folding unit underscore the allostery observed between ligand binding and dimerization of dPGM (Chapter 2).

In contrast to the dPGM core, the least stable cooperative folding unit is primarily composed of loops. Our data suggest that three α-helices (helices 1, 4, and 5) are involved in the folding unit but are not fully unfolded in the intermediate state; the helices appear to lose their native interactions with other structural features while

retaining helical structure. Many factors suggest α-helix 4 and 5 lose interactions with other secondary structural units in the intermediate state. The two helices form many interactions with the three loops that unfold in the intermediate state and only helix 5 forms a few interactions with helices 8 and 9 of the dPGM core. HDX results showed that α-helices 4 and 5 are only protected from deuterium exchange in the native state. Residue Lys99 of α-helix 4 and residues Arg115 and Arg116 of α-helix 5 were previously investigated by mutagenesis (Chapter 2); the three residues are involved in ligand binding and appear to be unfolded in the intermediate state (Chapter 2). However, α-helices 4 and 5 interact with one another significantly to form the kinked active site rim (Figure 3.8). Val111 seems to form important contacts between α-helices 4 and 5 in the intermediate state as removal of the side chain appears to promote further unfolding in the intermediate state (Table 3.1, 3.2). Together, the results suggest that α-helices 4 and 5 no longer form part of the active site rim and may be dynamic in the intermediate state, but the two helices remain in contact with one another. The residues in α-helix 1 interact with one another and with loops 19−29, and 89−98, which are both unfolded in the intermediate state. Both Thr22 and Val27 interact with α-helix 1 (residues 13-19) and both residues lose their native contacts in the intermediate state. Peptide 6−24 includes α-helix 1 and it shows low deuterium incorporation even in the unfolded state (Figure 3.6A). Similar to α-helices 4 and 5, α-helix 1 likely becomes dynamic due to the unfolding of adjacent loops but the helix appears to remain structured due to internal interactions. Residual helical structure in unfolded protein is not uncommon (Ahmad et al., 2015;

Leyrat et al., 2011; Yi Wang & Shortle, 1997). Full unfolding appears restricted to the network of loops in the least stable folding unit, which we term the *loop region*.

Our results suggest that the folded portion of the dPGM intermediate is comprised of two regions with distinct properties (Figure 3.7). The first folded region has a small $\Delta f_{\%D}$ and forms the core of the dPGM monomer and we refer to it as the *stable core.* The stable core contains strands A, D, E, and F as well as α-helices 8, 9, and 10. His10, the site of phosphorylation by the cofactor of dPGM, is also included in the stable core. Previously, we characterized R9A and K222A dPGM, both of which contain a point mutation in the stable core (Chapter 2). The equilibrium unfolding of both variants occurs with the transition from the native state to the intermediate state coincident with wild-type. While other mutations that remove a positive charge from the active site significantly shift the first unfolding transition to higher urea concentrations, R9A does not (Chapter 2). Investigation of R9A and K222A dPGM support that the core and a portion of the active site remain folded in the intermediate state. The second folded region, which shows a large $\Delta f_{\%D}$, we denote as the *interface region*, because most of the secondary structural units in this region form the dimer interface (α-helix 3, strand C, 3_{10}-helix 1) or make extensive interactions with the structural units that form the interface (α-helix 2 and 7, strand B). The last secondary structural unit included in the interface region is α-helix 6, which interacts with loop 118−152. The large $f_{\%D}$ of the interface region in the intermediate state can be explained by the increased solvent exposure upon dimer dissociation and altered dynamics, such increased local fluctuation, resulting from the

many lost interactions. So long as native side-chain contacts are maintained in the majority of the intermediate state conformational ensemble, our φ_{eq}-analysis will not report on the region's altered dynamics. The stable core of the dPGM intermediate and the exposed interface region may be discrete folding units as colored in Figure 3.7, though we did not directly observe separate folding of the two.

The structure of the dPGM intermediate and partitioning of dPGM into interacting folding units (Figure 3.9) clarifies the energetic coupling and allostery between ligand binding and dimerization of dPGM. Recall that allostery occurs when a protein contains two distal binding sites and binding of ligand at one site influences the binding of the second ligand at the distal binding site (Gunasekaran et al., 2004; Hilser et al., 2012). In the case of dPGM, one binding site is the dimer interface formed by the loop region and the interface region, and the ligand is another dPGM monomer; the second binding site is the dPGM active site formed by the loop region and the stable core. Many anions appear able to bind at the active site (Chapter 2, Chapter 4). Protein active sites tend to be split into regions of high and low structural stability with the catalytic machinery located in the high stability region and most of the substrate binding residues located in the low stability region (Ernesto Freire, 1999; Luque & Freire, 2000). The presence of a dynamic, low stability region is thought to enable allostery and transmit information between the active site and other regions of the protein via conformational change (Ernesto Freire, 1999; T. Liu et al., 2007). Both the dimer interface of dPGM and the active site clearly show such a duel character of high and low stability regions. Presumably, the stable

dimer interface of dPGM is responsible for recognition and initial binding between the two dPGM monomers because most of the binding site is preformed. Preformed structure facilitates the binding of intrinsically disordered protein before folding to the final bound form occurs (Fuxreiter, Simon, Friedrich, & Tompa, 2004; Knott & Best, 2012); preformed secondary structural units interact in a high-energy transition state and drive folding to the low-energy bound form (Daggett & Fersht, 2003; Fuxreiter et al., 2004; Zitzewitz, Ibarra-Molero, Fishel, Terry, & Robert Matthews, 2000). Interaction between the interface regions of two dPGM monomers likely occurs in a transition state followed by folding of the loop regions. Folding of the loop region completes the dimer interface and propagates information about the dimerization event throughout the loop region and active site. The dPGM active site contains most of the catalytic machinery (Bond et al., 2002), residues Arg9, His10, Arg61, and His183, in the stable core while two residues involved in catalysis, Asn16 and Glu88, as well as substrate binding residues, Thr22, Gly23, Arg89, Tyr91, Lys99, Arg115, and Arg116, are located in the dynamic loop region. The intermediate state of dPGM has a partially formed active site but most substrate binding residues are only found to be folded in the native dimer. Dimerization of dPGM will radically enhance ligand binding affinities due to the folding of the loop region. Conversely, if the preformed active site of the intermediate state can interact with a ligand through a high-energy transition state, it will facilitate folding of the remaining portion of the active site and the loop region. Then, encountering a ligand in the intermediate state will result in enhanced dimerization. The dynamic loop region of dPGM is critical for the allosteric effects observed. The allostery observed with dPGM,

in which one ligand is the protein itself, is an interesting case in which the function appears to be an enhancement of conformational stability rather than relaying a signal. However, the energetic coupling between the active site and dimer interface should energetically couple the active sites of both bound dPGM monomers. Binding events or phosphorylation at one active site may have an effect on binding and phosphorylation at the second active site.

The partitioning of dPGM into rather independent folding units enables the protein to have a stable core while also maintaining a large relevant stability, ($\Delta G_{I-N}°$). The relevant stability of a protein is the energy gap between the native protein and the lowest energy non-functional intermediate (Sancho et al., 2002). This energy gap is the most relevant to retain native, active protein because the intermediate is the most probable non-native conformation to be sampled. The core of dPGM is stable enough to form an intermediate and any stabilization of the core should stabilize the intermediate state and native state. Although the loop region is critical for ligand binding and dimerization of dPGM, it does not appear to be significantly coupled with the stable core. The tendency of the φ_{eq}-values for dPGM to fall near 0 or 1 supports the idea that the dPGM core is not significantly coupled with the loop region. The structure of an equilibrium intermediate of *Anabaena* apoflavodoxin was also determined through extensive φ_{eq}-analysis. Of our twenty dPGM variants, the φ_{eq}-values for seventeen variants fall within 0.3 of 0 or 1; for apoflavodoxin, only four out of eighteen of the determined φ_{eq}-values fall within 0.3 of 0 or 1 (Campos et al., 2004). The fractional φ_{eq}-values throughout the folded core of the apoflavodoxin

intermediate appeared to arise from the weakening of side-chain interactions throughout the core relative to the native state (Campos et al., 2004). The weakened bonds throughout the core of the apoflavodoxin intermediate reveal that the unfolded region of the intermediate is coupled with the folded region and folding of both is required for full strength interactions. The tendency of high φ_{eq}-values for dPGM to fall near 1 reveals that the side-chain contacts in the folded region of the dPGM intermediate are as favorable as in the native state; the side-chain interactions are not significantly affected by unfolding of the loop region. In contrast to apoflavodoxin, the folded core of the dPGM intermediate appears to function as an independent folding unit. Apoflavodoxin was rationally engineered to exhibit cooperative, two-state unfolding by enhancing the stability of the least stable folding unit (Lamazares, Clemente, Bueno, Velázquez-Campoy, & Sancho, 2015). The independence of the dPGM core allows ligand binding to only enhance the relevant stability ($\Delta G_{I-N}°$) and increase the folding cooperativity of dPGM analogous to the engineering efforts with Apoflavodoxin.

The cooperative network and allostery observed with dPGM enables the protein to greatly increase the stability of the native form over a partially unfolded intermediate by binding metabolites (Chapter 2) or salts (Chapter 4) that are prevalent in the cellular environment. The low stability of the loop region and high deuterium uptake imply that the region is dynamic. The cooperative network and allosteric response are also likely to be intimately coupled with the dynamics of the loop region (Chennubhotla et al., 2008; Wrabl et al., 2011). Homologous dPGMs possess varying quaternary structures (tetramer,

dimer, monomer) and show the greatest sequence divergence in the loop region (Bond et al., 2001). The pivotal role of the loop region on the energetics and allosteric behavior of *E. coli* dPGM suggest that this divergent region is fine tuned for the proper behavior of each dPGM in its physiological milieu. The dPGM of *Saccharomyces cerevisiae* is a tetramer and may access an active dimeric and even monomeric intermediate (Hermann, Jaenicke, & Price, 1985; White, Fothergill-Gilmore, Kelly, & Price, 1993a). An active site mutation, H181A, causes of *S. cerevisiae* to dissociate into dimers upon cofactor binding, showing negative cooperativity (White, Fothergill-Gilmore, Kelly, & Price, 1993b). The link between ligand binding and oligomerization appears much different for *S. cerevisiae* dPGM than for *E. coli* dPGM. Allosteric behavior and cooperative networks have been found to be conserved amongst homologous proteins in some cases (Law, Fuentes, & Lee, 2009; Suel, Lockless, Wall, & Ranganathan, 2003; Wrabl et al., 2011) and to be divergent in other cases (Coyle, Flores, & Lim, 2013; Kuriyan & Eisenberg, 2007; Morra, Genoni, & Colombo, 2014). Colocalization of proteins may greatly facilitate the evolution of allosteric behavior and lead to differences in behavior for homologous protein with essentially the same structure (Kuriyan & Eisenberg, 2007). An investigation into the cooperative network and allosteric behavior of dPGM homologs may reveal how divergence in the loop region regulates the cooperative network and quaternary structure of the dPGM proteins.

3.5 Experimental Procedures

3.5.1 Protein expression and purification

E. coli BL21 (DE3) PLysS cells were used to over-express dPGM as previously described (Chapter 2). All dPGM variants were constructed with QuikChange site-directed mutagenesis kits (Agilent Technologies, Santa Clara, CA) and plasmids were sequenced at the Purdue Genomics Core Facility (West Lafayette, IN) to confirm that the mutations were made. BL21 cells were lysed by sonication and all dPGM variants were purified through two rounds of anion exchange chromatography (Chapter 2). Purified dPGM was concentrated with the Centriprep 10K filter (EMD Millipore, Billerica, MA) and stored at 4°C or -80°C for prolonged storage. The purity of each dPGM variant was determined to be greater than 95% by analysis of SDS-PAGE gels with Coomassie blue staining. The final concentration of each dPGM variant was determined by absorbance at 280 nm using an extinction coefficient estimated from the amino acid composition (Pace et al., 1995).

3.5.2 Circular dichroism

CD spectra spanning 200 to 250 nm were collected on a JASCO J-815 CD spectrometer (JASCO, Easton, MD) for all dPGM variants. 0.30 mg/mL dPGM samples were prepared from 10-fold dilutions of the protein stock into 20 mM TrisHCl−NaOH (pH 8.0) with or without 0.10 M NaCl as shown in Figure 3.2. The concentrations of dPGM in the 10-fold dilutions were determined by absorbance at 280 nm and were used to calculate the dPGM concentration in each CD sample. Samples were equilibrated at room temperature for at

least two hours prior to spectroscopy. The CD signal of the buffer was subtracted before the raw CD signal of the dPGM variants was converted into mean-residue molar ellipticity.

3.5.3 Size exclusion chromatography

A HiLoad 16/600 Superdex 75 pg column (GE Healthcare Life Sciences; Piscataway, NJ) was used to perform size exclusion chromatography with each dPGM variant. The mobile phase was 20 mM TrisHCl−NaOH buffer (pH 8.0) alone (Figure 3.3A) or buffer with 0.15 M NaCl (Figure 3.3B,C,D). We prepared 1.0 mg/mL dPGM in the respective mobile phase and injected 1.0 mL of the sample for each run. Size exclusion was conducted with a flow rate 1.6 mL/min while the absorbance at 280 nm was recorded.

3.5.4 Equilibrium unfolding

The urea-induced equilibrium unfolding of dPGM was monitored by intrinsic tryptophan fluorescence. For all experiments, 0.030 mg/mL dPGM was prepared in 20 mM TrisHCl−NaOH buffer (pH 8.0) containing varying concentrations of urea (0 – 5 M) and NaCl. The urea concentrations were determined from the refractive index of a stock solution (Pace, 1986). The equilibrium unfolding samples were equilibrated at room temperature overnight prior to spectroscopy. Fluorescence emission spectra spanning 320 to 400 nm were collected with a FlouroMax-3 fluorimeter (Horiba Jobin Yvon; Edison New Jersey, NJ) using an excitation wavelength of 280 nm. During data collection, the samples were maintained at 25°C.

The equilibrium unfolding data is analyzed as intensity-weighted average wavelength (IWAW) as previously described (Chapter 2). The data was fit with a three-state model (eq 2.4) detailed in scheme 1:

$$N_2 \rightleftarrows 2I \rightleftarrows 2U$$

Scheme 1

Curve fitting was carried out as previously described using OriginPro 8.5.1 (OriginLab; Northampton, MA) with eq 2.4 to determine the *m*-values and the free energy differences between states, $\Delta G_{I-N}°$and $\Delta G_{U-I}°$ (Chapter 2). However, for all dPGM variants except L125A dPGM, the global *m*-value was fixed at the value of wild-type dPGM (6.5 kcal mol^{-1} M^{-1}) determined from a global fitting of wild-type dPGM in the presence of 0 to 0.50 M NaCl (Chapter 4). The global *m*-value was fixed by defining m_{U-I} as $(6.5 - m_{I-N})/2$ in eq 2.3. For global fits, the fluorescence signal of the native state and intermediate state, the dPGM concentration, and m_{I-N} are shared amongst all data sets. The φ_{eq}-values, as defined by eq 3.1 (Campos et al., 2004; Sancho et al., 2002), were calculated from the effect of a mutation on the stability of dPGM per monomer following eq 3.2 as described in the Results. Note that $\Delta\Delta G_{U-N}°$ in eq 3.1 is equivalent to $\Delta\Delta G_{I-N}°/2 + \Delta\Delta G_{U-I}°$ in eq 3.2. When global fitting was performed, an average φ_{eq}-value was calculated from the individual φ_{eq}-values for equilibrium unfolding at each NaCl concentration included in the fitting. The φ_{eq}-values were mapped to the crystallographic

structure (PDB code: 1E59) of dPGM manually with the software PyMol (Schrödinger LLC, New York, NY).

3.5.5 Hydrogen/deuterium exchange coupled with mass spectrometry

For HDX-MS, dPGM samples were prepared at a concentration of 100 μg/mL in 20 mM TrisHCl−NaOH (pH 8.0) at three urea concentrations: 0 M, 2.5 M, and 5.0 M. One dPGM sample was also prepared containing 1.0 mM ATP to stabilize the native state. The dPGM samples were equilibrated overnight at room temperature. Again, the urea concentrations were determined from the refractive index of a stock solution (Pace, 1986). Deuterated buffers were prepared by lyophilizing TrisHCl−NaOH (pH 8.0) at three urea concentrations (0 M, 2.5 M, and 5.0 M) and reconstituting it with D_2O. Four rounds of lyophilization were carried out, each time reconstituting the buffer with D_2O. Solutions for the control were prepared by reconstituting buffer containing no urea using H_2O.

To initiate deuterium exchange, 4.0 μL of a dPGM sample was mixed with 36.0 μL of the respective deuterated buffer or the control buffer. HDX pulse labeling was carried out in triplicate for two minutes. The reaction was then quenched using 80.0 μL quench buffer (0.2% FA + 5% methanol, 4.5 M urea at pH 2.5) followed by flash freezing with liquid nitrogen. Samples were analyzed by LC/MS (Agilent 1200 LC, 6520 ESI-QTOF, Agilent Technologies, Santa Clara, CA). Each sample was eluted with a mobile phase of 0.1 % formic acid in LC/MS-grade water for isocratic elution; gradient elution was achieved using 0.1 % formic acid in LC/MS-grade water (solvent A) and 0.1 % formic acid in

LC/MS-grade acetonitrile (solvent B). An isocratic flow was maintained for 1.5 minutes followed by a gradient from 25% solvent B to 70% solvent B over 6 minutes. dPGM was digested online by a column containing immobilized pepsin beads (Thermo Fisher Scientific, Waltham, MA). The Zorbax C18 column used for peptide level analysis was procured from Agilent Technologies.

A theoretical mass list using MassHunter (Agilent Technologies) was prepared with the undeuterated control sample and was used for further analysis. Deuterium uptake for all peptides was calculated by HDExaminer software (Sierra Analytics, Modesto, CA). Briefly, the change in centroid of the peptide mass envelope relative to the control, as calculated by the software, was monitored for all samples. The % deuterium uptake was calculated from the ratio of the change in centroid mass to the number of exchangeable amides per peptide. Only the peptides that appeared in all the samples with good intensity were selected for further analysis. The % deuterium uptake for the triplicate samples of dPGM under each condition were averaged to obtain the mean % deuterium uptake.

We defined the relative deuterium uptake as $f_{\%D}$ in eq 3.2 to assess differences between the native state and the intermediate state of dPGM. The difference in $f_{\%D}$ between native dPGM and the intermediate state is defined as $\Delta f_{\%D} = f_{\%D}(\text{native}) - f_{\%D}(\text{intermediate})$ and was determined to visualize differences in HDX between the native form of dPGM and

intermediate relative to the globally unfolded form. $\Delta f_{\%D}$ was mapped to the crystallographic structure (PDB code: 1E59) of dPGM manually using PyMol (Schrödinger LLC).

Table 3.1 Structural description the mutations implemented in *E. coli* dPGM

Variant	Position	Van der Waals contacts[a]	Structural elements involved
V8A	Strand A	L7, R9, I180, A181	Strand A, strand D
T22S	loop	Y25, V27, R61	loop, helix 3
V27A	loop	S13, Q14, W15, T22, D28, R61	Helix 1, loop, helix 3
V34A	Helix 2	L29, S30, K38, N68	Helix 2, helix 3
L43A	Helix 2	L44, E47, L229, G230	Helix 2, strand F
T56S[b]	Strand B	I63, L66, K82	helix 3, strand C
I63V	Helix 3	V58, A75[d], W76[d]	loop, helix 3[d]
L66V	Helix 3	A54, L70, V80	Strand B, helix 3, strand C
W67A[c]	Helix 3	I63, H64, D70[d], A76[d]	Helix 3, helix 3[d]
V80A	Strand C	L66, L78, K82, K82[d]	Helix 3, strand C, strand C[d]
L86V	3_{10} helix 1	Y55, T56, Y163	Strand B, helix 6
L94V	3_{10} helix 2	Y91, W114	3_{10} helix 2, helix 5
V111A	Helix 5	T102, A103, R115	Helix 4, helix 5
L125A	loop	H90, G92, A93, D129, R131, Y132, L148	3_{10} helix 2, loop
L147A	loop	Y132, P133, G134, Y139, L142, S143	3_{10} helix 3, 3_{10} helix 4, loop
V160A	Helix 6	K85, T156, R159, S186	3_{10} helix 1, helix 6, helix 8
L169A	Helix 7	W164, P170, K173	Helix 6, helix 7
V191A	Helix 8	L187, K192, D195, L205, I207, Y215	Helix 8, helix 9, strand E
V214A	Strand D	L5, Y49, Y227, L229	Strand A, strand E, strand F
I235A	Helix 10	P208, V211, Y228, L229, G230, N231	loop, strand F, helix 10

[a] Determined to be residues within 4 Å of only the atoms that are removed by mutation.
[b] Forms a hydrogen bond with A181.
[c] Forms a hydrogen bond with a Cl^- ion in the dimer interface.
[d] Contact located in the opposing monomeric unit of the dPGM dimer.

Table 3.2 Thermodynamic parameters and φ-values of dPGM variants in the presence of 0.10 M NaCl

Variant	m_{I-N} (kcal mol^{-1}M^{-1})	$\Delta G_{I-N}°$[b] (kcal/mol)	$\Delta\Delta G_{I-N}°$ (kcal/mol)	m_{U-I}[c] (kcal mol^{-1}M^{-1})	$\Delta G_{U-I}°$ (kcal/mol)	$\Delta\Delta G_{U-I}°$[d] (kcal/mol)	φ_{eq}[e]
V8A[a]	4.32 ± 0.06	15.3 ± 0.1	1.3 ± 0.1	1.09 ± 0.03	2.62 ± 0.09	-3.9 ± 0.2	1.2 ± 0.1
V34A	3.18 ± 0.08	14.0 ± 0.1	0.0 ± 0.1	1.66 ± 0.04	5.4 ± 0.2	-1.1 ± 0.3	1.0 ± 0.4
L43A	3.04 ± 0.07	14.2 ± 0.1	0.2 ± 0.1	1.73 ± 0.04	5.5 ± 0.1	-1.0 ± 0.2	1.1 ± 0.4
T56S	3.40 ± 0.06	14.4 ± 0.1	0.4 ± 0.1	1.55 ± 0.03	4.9 ± 0.1	-1.6 ± 0.2	1.1 ± 0.2
I63V	3.36 ± 0.02	13.79 ± 0.05	-0.23 ± 0.06	1.57 ± 0.01	4.80 ± 0.05	-1.7 ± 0.2	0.9 ± 0.2
L66V	3.45 ± 0.04	14.01 ± 0.08	-0.01 ± 0.09	1.53 ± 0.02	4.43 ± 0.05	-2.1 ± 0.2	1.0 ± 0.1
V80A	3.45 ± 0.01	13.60 ± 0.02	-0.42 ± 0.04	1.53 ± 0.01	4.78 ± 0.04	-1.7 ± 0.2	0.9 ± 0.1
L86V	3.64 ± 0.01	12.59 ± 0.06	-1.43 ± 0.07	1.43 ± 0.01	4.58 ± 0.01	-1.9 ± 0.2	0.73 ± 0.09
V111A[a]	3.8 ± 0.1	15.3 ± 0.3	1.3 ± 0.3	1.36 ± 0.06	3.6 ± 0.1	-2.9 ± 0.2	1.3 ± 0.2
V160A	3.7 ± 0.1	13.7 ± 0.2	-0.3 ± 0.2	1.41 ± 0.05	3.9 ± 0.2	-2.6 ± 0.3	0.9 ± 0.1
V191A	3.07 ± 0.03	14.4 ± 0.8	0.4 ± 0.8	1.72 ± 0.02	4.69 ± 0.09	-1.8 ± 0.2	1.1 ± 0.3
I235A	2.80 ± 0.06	14.15 ± 0.09	0.13 ± 0.1	1.85 ± 0.03	5.7 ± 0.1	-0.8 ± 0.2	1.1 ± 0.5

[a] Fitting performed with the signal of the intermediate, Y_I, fixed at 355 nm.
[b] $\Delta\Delta G_{I-N}° = \Delta G_{I-N}°(\text{variant}) - \Delta G_{I-N}°(\text{wild-type})$
[c] $m_{U-I} = (6.5 - m_{I-N})/2$
[d] $\Delta\Delta G_{U-I}° = \Delta G_{U-I}°(\text{variant}) - \Delta G_{U-I}°(\text{wild-type})$
[e] φ_{eq}-values are calculated following eq 3.2.

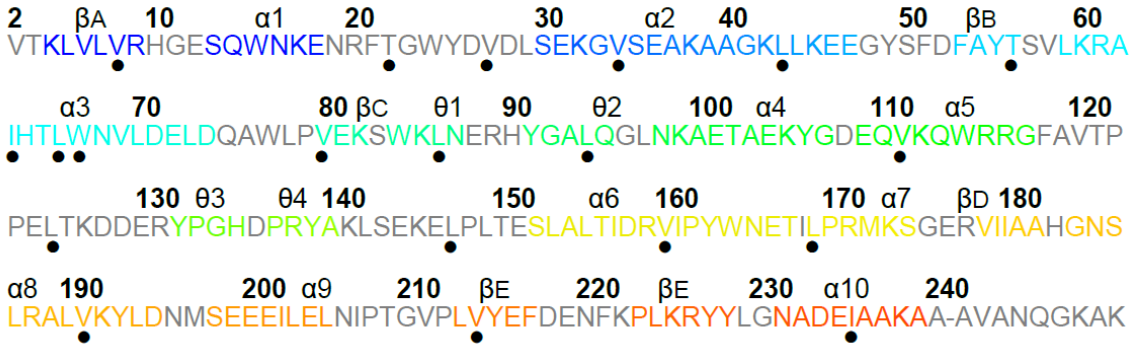
Table 3.3 Thermodynamic parameters and φ-values of dPGM variants at three salt conditions.

Variant	[NaCl] (M)	m_{I-N} (kcal mol^{-1}M^{-1})	$\Delta G_{I-N}°$[a] (kcal/mol)	$\Delta\Delta G_{I-N}°$ (kcal/mol)	m_{U-I}[b] (kcal mol^{-1}M^{-1})	$\Delta G_{U-I}°$ (kcal/mol)	$\Delta\Delta G_{U-I}°$[c] (kcal/mol)	φ_{eq}[d]	Average φ_{eq}[e]
	0.020	2.60±0.04	11.30±0.05	-1.19±0.06	1.95±0.02	6.72±0.08	0.1±0.2	-0.3±0.5	
T22S	0.050	–	12.05±0.06	-1.30±0.07	–	6.62±0.08	0.1±0.2	-0.2±0.4	-0.19±0.08
	0.10	–	12.71±0.07	-1.3±0.1	–	6.56±0.08	0.1±0.2	-0.1±0.4	
	0.020	2.74±0.04	11.55±0.05	-0.9±0.1	1.88±0.02	6.66±0.08	0.1±0.2	-0.1±0.5	
V27A	0.050	–	12.15±0.06	-1.2±0.1	–	6.46±0.08	0.0±0.2	0.1±0.3	0.0±0.1
	0.10	–	12.91±0.07	-1.1±0.1	–	6.45±0.07	0.0±0.2	0.1±0.4	
	0.10	2.89±0.02	10.12±0.02	-3.90±0.04	1.81±0.01	5.93±0.04	-0.6±0.2	0.23±0.08	
W67A	0.20	–	10.50±0.02	-4.13±0.05	–	5.85±0.02	-0.8±0.2	0.27±0.07	0.25±0.02
	0.50	–	11.01±0.02	-4.29±0.05	–	6.23±0.04	-0.8±0.2	0.26±0.07	
	0.020	2.56±0.04	10.59±0.05	-1.90±0.06	1.97±0.02	6.81±0.07	0.2±0.2	-0.3±0.3	
L94V	0.050	–	11.22±0.05	-2.13±0.06	–	6.60±0.07	0.1±0.2	-0.1±0.2	-0.1±0.1
	0.10	–	11.71±0.06	-2.31±0.06	–	6.50±0.07	0.0±0.2	0.0±0.2	
	0.10	1.97±0.04	9.25±0.04	-4.77±0.06	2.05±0.08	6.9±0.3	0.4±0.4	-0.2±0.2	
L125A[f]	0.20	–	9.70±0.05	-4.93±0.07	–	6.9±0.3	0.3±0.4	-0.1±0.2	-0.20±0.05
	0.50	–	10.25±0.07	-5.05±0.09	–	7.5±0.3	0.5±0.4	-0.2±0.2	

	0.10	2.77±0.01	10.10±0.01	-3.92±0.04	1.87±0.01	5.86±0.01	-0.6±0.2	0.20±0.08	
L147A	0.20	–	11.10±0.02	-3.53±0.05	–	6.19±0.03	-0.4±0.2	0.20±0.09	0.21±0.03
	0.50	–	11.83±0.02	-3.47±0.05	–	6.56±0.03	-0.4±0.2	0.20±0.09	
	0.020	2.91±0.04	13.22±0.07	0.73±0.08	1.80±0.02	5.37±0.09	-1.2±0.2	1.4±0.4	
L169A	0.050	–	14.10±0.08	0.75±0.09	–	5.24±0.08	-1.3±0.2	1.4±0.4	1.43±0.02
	0.10	–	14.89±0.09	0.9±0.1	–	5.10±0.07	-1.4±0.2	1.5±0.4	

[a] $\Delta\Delta G_{I-N}°$= $\Delta G_{I-N}°$(variant) - $\Delta G_{I-N}°$(wild-type)
[b] $m_{U-I} = (6.5 - m_{I-N})/2$
[c] $\Delta\Delta G_{U-I}°$= $\Delta G_{U-I}°$(variant) – $\Delta G_{U-I}°$(wild-type)
[d] φ_{eq}-values are calculated following eq 3.2.
[e] Average φ_{eq} is calculed from the three φ_{eq}-values for each dPGM variant.
[f] Fitting was performed without constraining m_{U-N}.

A



B

C

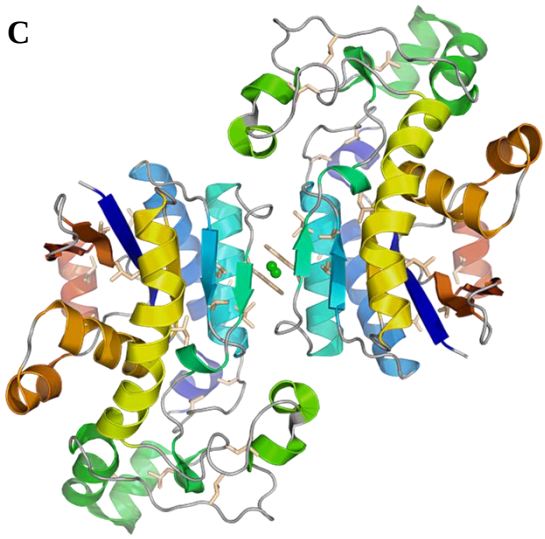


Figure 3.1 Structure of dPGM showing the location of point mutations

The sequence of dPGM (A) is displayed with the location of point mutation denoted by a dot beneath the corresponding residue. A dash follows the last residue resolved in the crystal structure of dPGM (Ala240). The secondary structural units are labeled above the corresponding residues and residues are colored to match the secondary structural features as displayed in the crystal structure of the dPGM dimer (B,C) (PDB entry 1E59). The secondary structural features of dPGM are rainbow colored (blue to red) from the N- to C-terminus and loops are colored gray. The orientation of dPGM in (C) is rotated 180° relative to (B). The side chains of the residues chosen for mutation are shown. A Cl^- ion bound in the dimer interface of dPGM is shown as a green sphere.

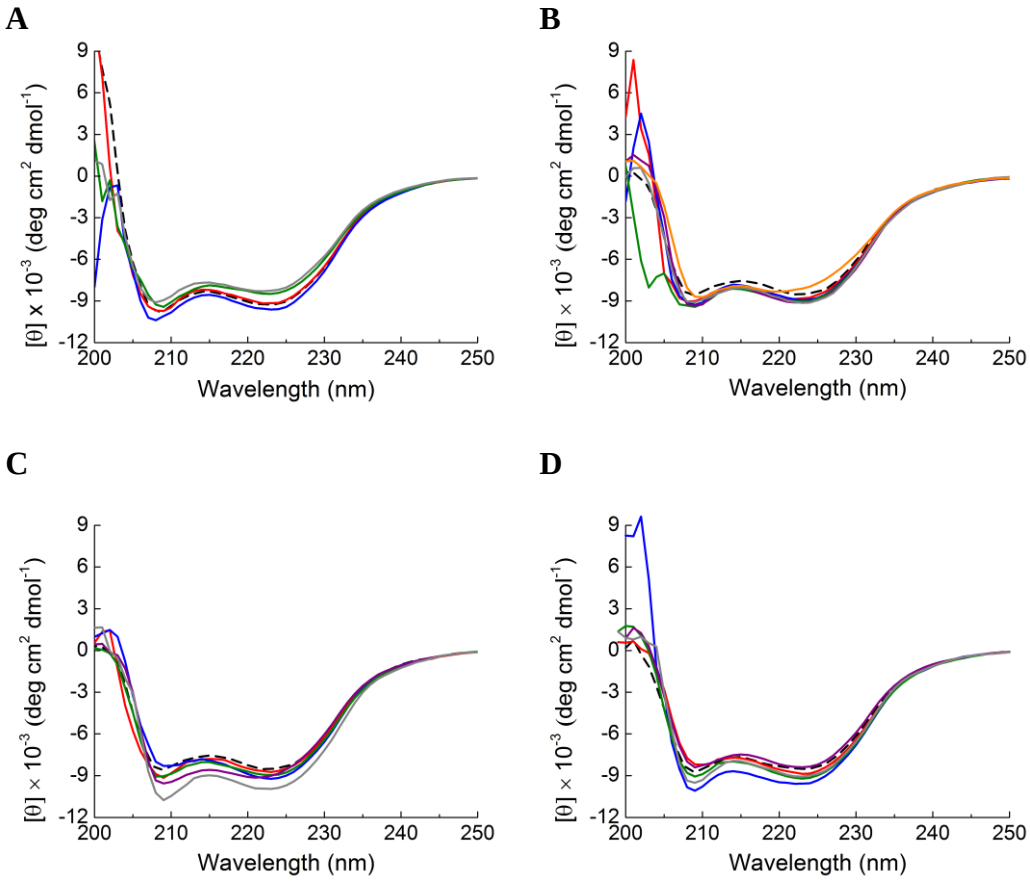


Figure 3.2 Effect of mutation on the CD spectra dPGM

The CD spectra of (A) V8A (—), T56S (—), L66V (—), and V80A (—) dPGM were examined in 20 mM TrisHCl−NaOH (pH 8.0). The spectra of wild-type dPGM (— —) is shown for reference. The CD spectra of the other variants were examined in the presence of 0.10 M NaCl as follows: (B) T22S (—), V27A (—), V34A (—), L43A (—), I63V (—), and W67A (—) dPGM; (C) L86V (—), L94V (—), V111A (—), L125A (—), and L147A (—) dPGM; (D) V160A (—), L169A (—), V191A (—), V214A (—), and I235A (—) dPGM. The spectrum of wild-type dPGM is shown (— —) for reference.

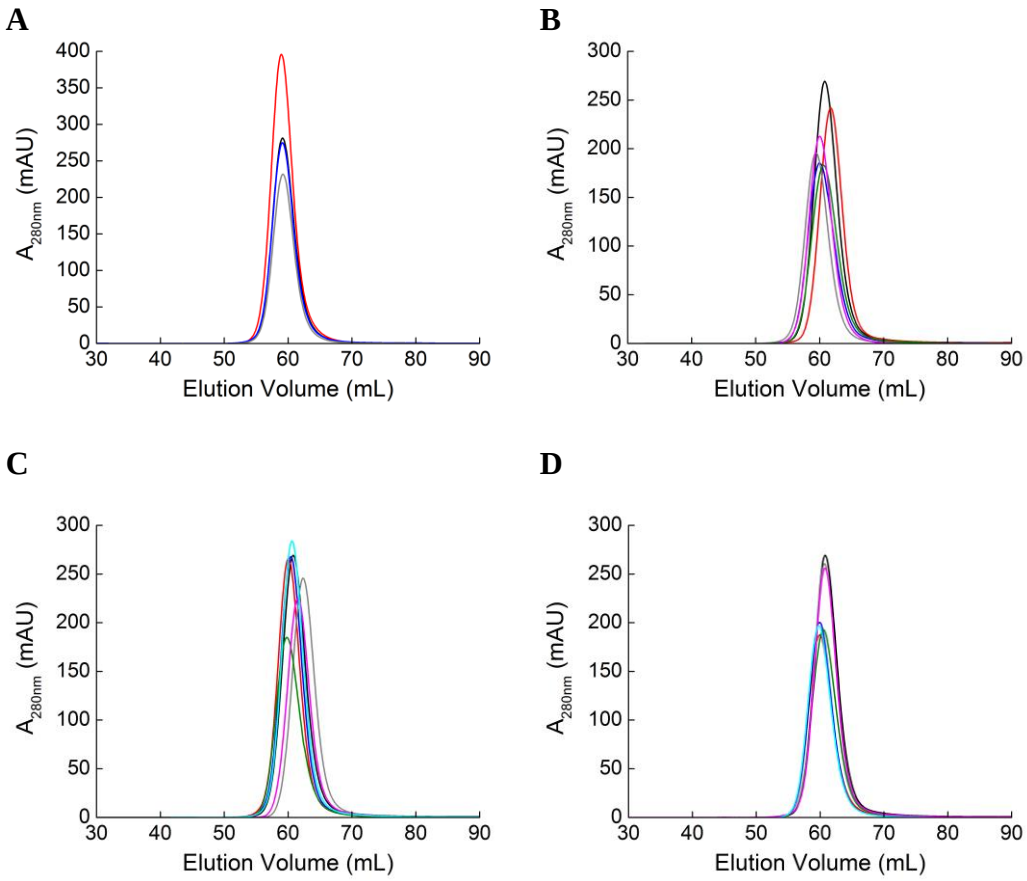


Figure 3.3 Effect of mutation on the elution of dPGM from size exclusion chromatography

The absorbance at 280 nm was monitored during size exclusion chromatography of (A) wild-type dPGM (—), L86V (—), V111A (—), and V214A (—) dPGM without added salt. (B) Size exclusion was also monitored in the presence of 0.15 M NaCl for wild-type dPGM (—) and the following variants: (B) V8A (—), T22S (—), V27A, (—) V34A (—), and L43A (—) dPGM; (C) T56S (—), I63V (—), L66V (—), W67A (—), V80A (—), and L94V (—) dPGM; (D) L125A (—), L147A (—), V160A, (—) L169A (—), V191A (—), and I235A (—) dPGM.

Figure 3.4 Effect of point mutations on the equilibrium unfolding of dPGM

The equilibrium unfolding of the following dPGM variants was monitored by fluorescence in the presence of 0.10 M NaCl: (A) V8A (●), V34A (■), L43A (□), and T56S (○) dPGM; (B) I63V (●), L66V (□), V80A (■), and L86V (○) dPGM; (C) V111A (○), V160A (●), V191A (■), and I235A (□) dPGM. The data were fit with eq 2.4 (solid line), and the fitting results are listed in Table 3.2. The fitting result for wild-type dPGM unfolding in the presence of 0.10 M NaCl is shown (dashed line) for reference.

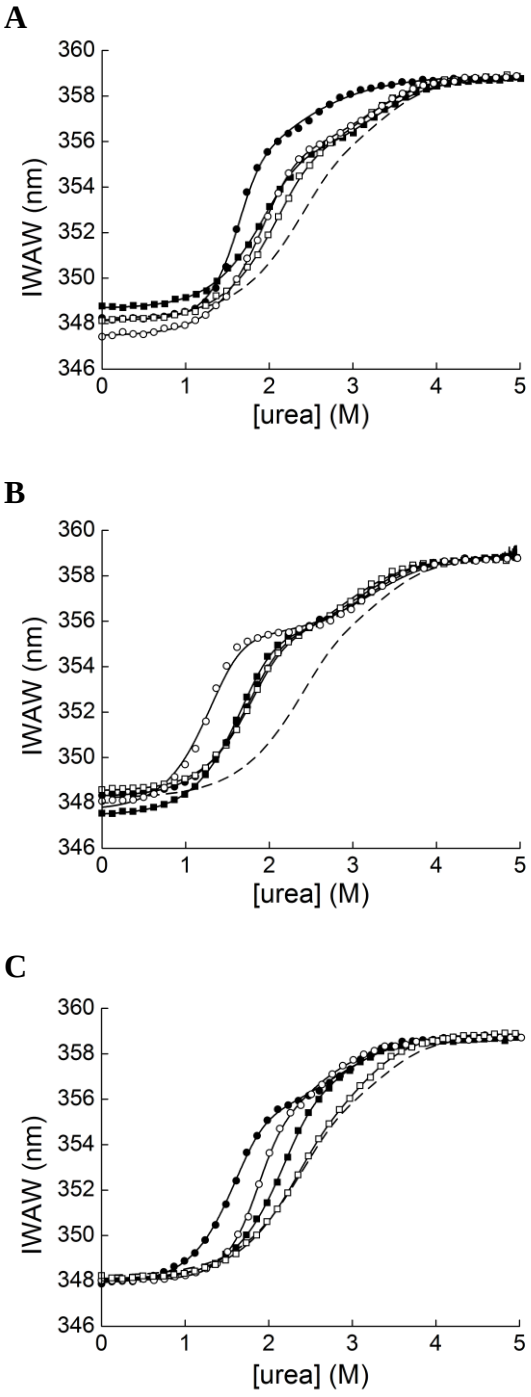
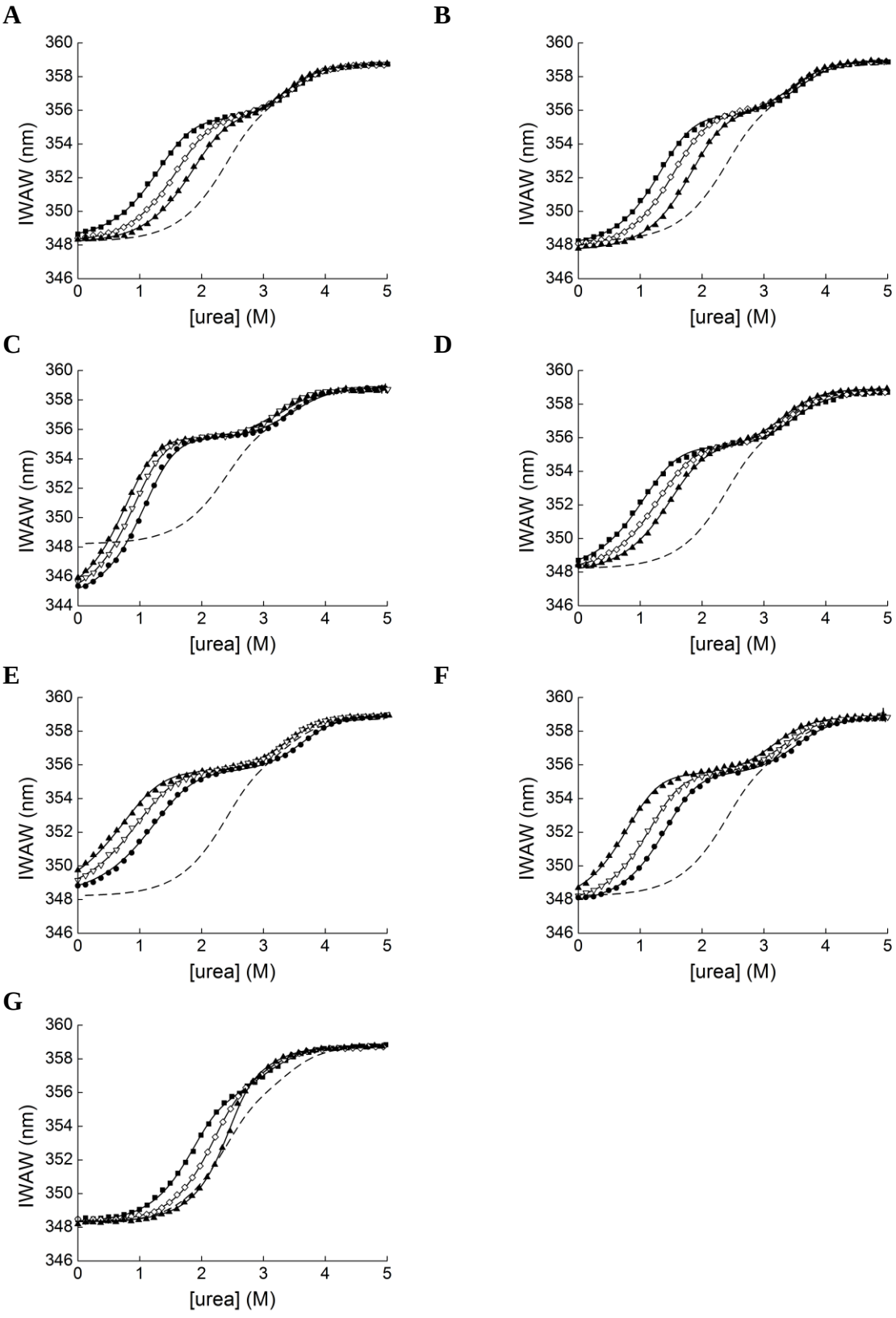

A
B
C
IWAW (nm)
[urea] (M)
360
358
356
354
352
350
348
346
0
1
2
3
4
5

Figure 3.5 Effects mutation upon dPGM in the presence of NaCl

The equilibrium unfolding of (A) T22S, (B) V27A, (C) W67A, (D) L94V, (E) L125A, (F) L147A, and (G) L169A dPGM was monitored by tryptophan fluorescence under three different NaCl concentrations. The concentration of NaCl is denoted as follows: 0.02 M (■), 0.05 M (◇), 0.10 M (▲), 0.20 M (▽) and 0.50 M (●). The equilibrium unfolding of each variant was fit globally by eq 2.4 (solid lines), and the resulting thermodynamic parameters are listed in Table 3.3. The fitting result for wild-type dPGM unfolding in the presence of 0.10 M NaCl is shown (dashed line) for comparison.


A
B
C
D
E
F
G
IWAW (nm)
[urea] (M)

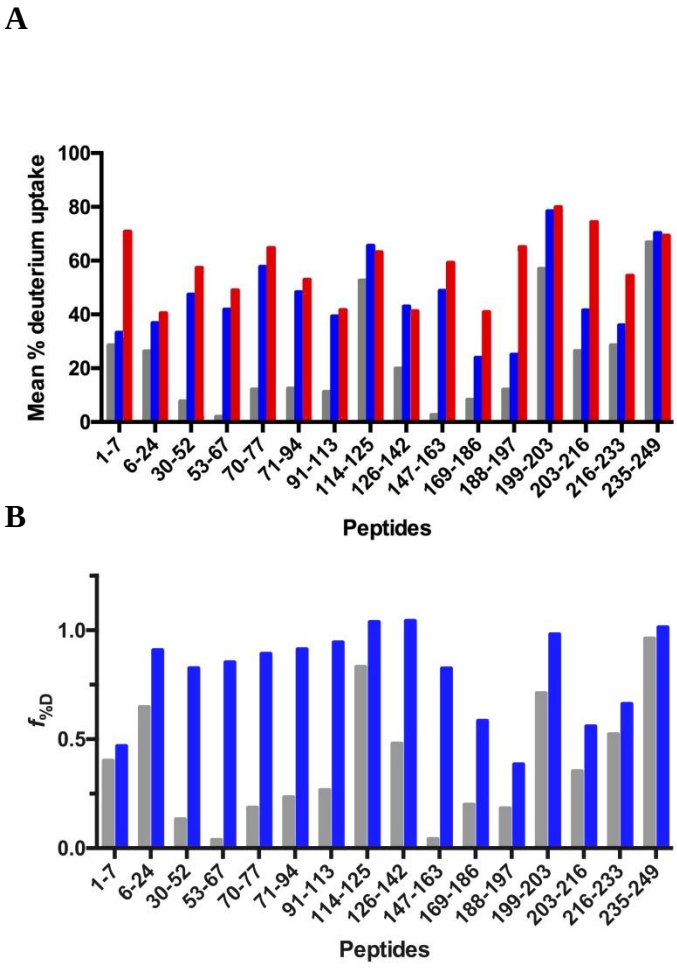
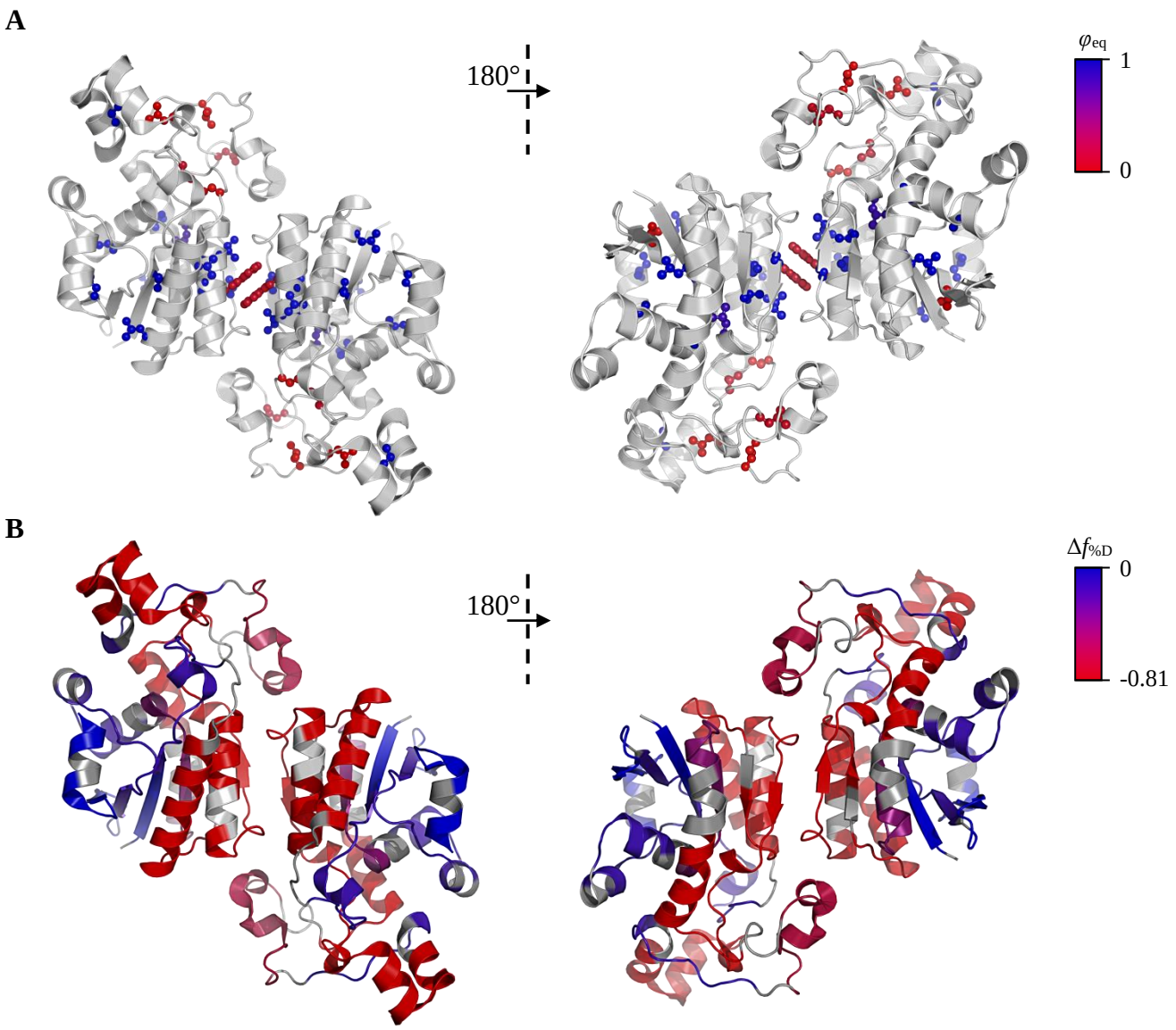


Figure 3.6 Deuterium uptake of the dPGM dimer, intermediate, and unfolded form

(A) The mean % deuterium uptake for peptides from dPGM in the native state (gray), intermediate state (blue), and unfolded state (red). (B) The fraction of relative deuterium uptake, $f_{\%D,}$ for each peptide from the native state (gray) and the intermediate state (blue), where 1.0 signifies the same amount of deuterium uptake as the unfolded state. The $f_{\%D}$ was determined following eq 3.3.

Figure 3.7 A structural view of the differences between native dPGM and the intermediate state.

(A, B) The φ_{eq}-values (Tables 3.2, 3.3) are mapped to the crystal structure of dPGM (PDB entry 1E59) with the mutated side chains displayed and colored blue to red for φ_{eq}-values of 1 to 0, respectfully. (C, D) The difference in relative deuterium uptake between the native state and the intermediates state, $\Delta f_{\%D}$, is mapped to the crystal structure of dPGM with the values of $\Delta f_{\%D}$ ranging from 0 to -0.81 colored as blue to red, respectfully. The orientation of dPGM in (B) and (D) is rotated 180° about the vertical axis relative to (A) and (C) as denoted.


A
180°
φ_{eq}
1
0
B
180°
$\Delta f_{\%D}$
0
-0.81

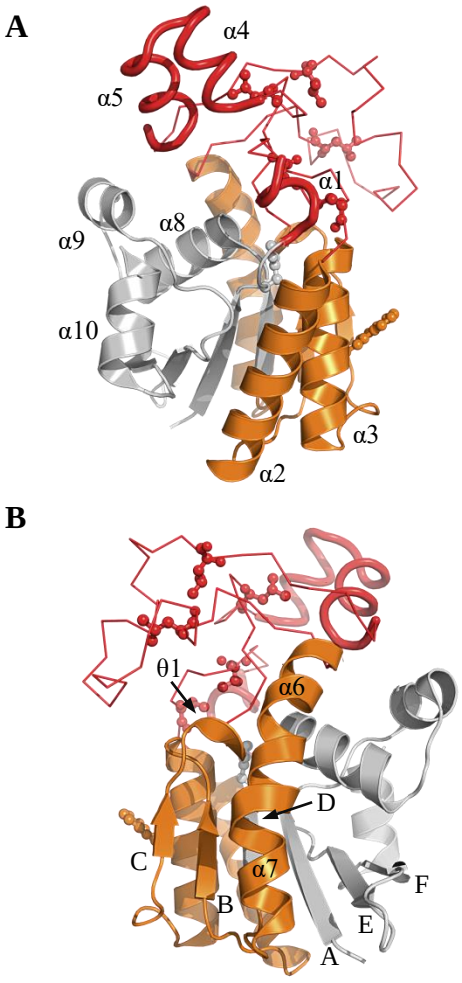


Figure 3.8 The structure of the dPGM intermediate.

The structure of the dPGM intermediate is mapped to the crystal structure of the native monomer (PDB entry 1E59) with the unfolded loop region denoted as red and its backbone depicted as a ribbon. α-helices 1, 4, and 5 (red) are depicted as tubes and appear to lose contact with other structural units in the intermediate state but retain some local structure. The orange interface region and gray stable core regions, displayed as a cartoon, both remain folded in the intermediate. The interface region loses many contacts from the unfolding of the loop region and dissociation of the dimer. The structure in (B) is rotated 180° about the vertical axis relative to (A). The α-helices (α), 3_{10}-helicies (θ), and strands (letters) of the β-sheet are labeled.

CHAPTER 4. EFFECT OF A PHYSIOLOGICAL MILEAU ON THE ENERGETICS OF METABOLITE BINDING BY *ESCHERICHIA COLI* PHOSPHOGLYCERATE MUTASE

4.1 Summary

The chemical environment of the cell offers protein a multitude of potential binding partners. *Escherichia coli* cofactor-dependant phosphoglycerate mutase (dPGM) has been demonstrated to bind non-substrate metabolites such as ATP and other nucleotides *in vitro*. dPGM is a homodimeric protein and was found to populate a partially unfolded, monomeric intermediate during equilibrium unfolding. Partially unfolded protein are often more accessible than the globally unfolded form and can misfold, aggregate, or trigger disease states. Binding to ligand selectively promotes the dPGM dimer and reduces monomer populations. Here, we seek to understand the physiological relevance of the interactions between dPGM and the chemical environment. We investigated how physiological salt conditions impact the stabilizing effect of metabolites and ATP binding. Salt reduces the interactions between dPGM and metabolites, but also selectively enhances the stability of the dimer. Through mutagenesis we identify three residues that appear to bind chloride: Trp67, Lys99, and R115A. Many anions and anionic metabolites interact with dPGM at physiologically relevant concentrations. We find that the presence of 0.10 M NaCl, 5.0 mM $MgCl_2$, and 1.0 mM ATP more than doubles the effective free energy of dPGM compared to the protein in buffer alone. The cooperative linkage

between ligand binding and dimerization in dPGM allows the enzyme to utilize a multitude of abundant anions to drastically increase the energy gap between a partially unfolded intermediate and the native state in a physiological milieu.

4.1 Introduction

Often protein folding is regarded as a cooperative, two-state transition and is investigated using small, monomeric protein. However, the majority of the proteome is comprised of oligomers and in *Escherichia coli* dimeric proteins are, by far, the most populous form (Braselmann et al., 2013). Often, dimeric proteins are found to populate partially unfolded intermediates *in vitro* (Rumfeldt, Galvagnion, Vassall, & Meiering, 2008). In the cellular environment, significant partial unfolding can lead to proteolytic digestion (Hubbard, 1998), misfolding (Dobson, 2003), and aggregation (Horwich, 2002). Both misfolding and aggregation may causes disease states (Thomas et al., 1995). With partial unfolding being so risky, many dimeric proteins likely avoid partial unfolding in their physiological milieu. Substrate binding offers a facile way to promote binding-competent protein conformations, but substrate may not always be maintained at saturating levels and not all protein have a substrate. A theoretical analysis of the evolutionary effects of random mutation on a protein in the presence of various small molecules has predicted that proteins may frequently evolve nonfunctional binding sites that aid in stability (Manhart & Morozov, 2015). By stabilizing the native protein, nonfunctional interactions actually bolster enzyme activity and can, therefore, be indirectly selected for by natural selection (Manhart & Morozov, 2015). Large-scale screens for weak protein-metabolite interactions are now becoming feasible and have revealed numerous novel interactions

with no known function (Li, Gianoulis, Yip, Gerstein, & Snyder, 2010; Orsak et al., 2011). These screens suggests that nonfunctional binding interactions are common and may play an integral role in stabilizing folded protein over partially unfolded intermediates.

The cofactor-dependent phosphoglycerate mutase (dPGM) enzymes take part in glycolysis and gluconeogenesis. The dPGM enzymes are dependent on the cofactor 2,3-bisphosphoglycerate for activation. As a phosphatase (EC 3.1.3.13), dPGM converts 2,3-bisphosphoglycerate into 2-phosphoglycerate or 3-phosphoglycerate while retaining a phosphate group on a conserved, active-site histidine. Once phosphorylated, dPGMs catalyze the reversible isomerization reaction between 2-phosphoglycerate and 3-phosphoglycerate, fulfilling their primary role as a mutase (EC 5.4.2.11). In proteome wide screens, both *E. coli* dPGM and *Saccharomyces cerevisiae* dPGM have been identified as ATP-binding protein (Chang et al., 2012; Tran et al., 2014). However, the enzymatic activity of dPGM does not require ATP, and assays that monitor the mutase activity have been observed in the presence and absence of ATP with no notice of an effect from the presence of ATP (Fraser et al., 1999).

Previously, we determined that ATP can bind in the active site of the homodimeric *E. coli* dPGM; the active site of dPGM is energetically coupled with the dimer interface such that ligand binding and dimerization promote one another (Chapter 2). dPGM was found to readily populate a partially unfolded, monomeric intermediate but ligand binding selectively increases the stability of the native dimer over the intermediate and

reduces the population of the intermediate. The stabilizing effect of ATP on dPGM appears to be shared by other nucleotides and phosphates (Chapter 2). Further, the crystal structures of active and inactive *E. coli* dPGM have been solved with the active site occupied by sulfate and tetravanadate, respectfully (Bond et al., 2001, 2002). We utilized ATP to elucidate the energetic coupling in dPGM under low ionic strengths (Chapter 2, Chapter 3). However, the activity of dPGM has been observed to be highly sensitive to salts and ionic strength with its maximal activity occurring below 0.1 M ionic strength (Grisolia & Cleland, 1968; Rose & Dube, 1978; Rose & Kaklij, 1984). Because metabolites appear to interact with dPGM through Coulombic interactions in the active site, ligand binding may also be highly dependent upon salt concentrations. *E. coli* dPGM presents a valuable three-state system to assess changes in protein energetics because it can highly populate a stable intermediate with a distinctive spectroscopic signal and the two unfolding transitions can be well separated. This allows for accurate determination of global stability and the energy gap between the native state and partially unfolded intermediate, which bears more physiological relevance (Sancho et al., 2002).

Here, we seek to understand the physiological relevance of interactions between ligand and dPGM and the relevance of the energetic coupling between ligand binding and dimerization. We utilize equilibrium unfolding and isothermal titration calorimetry to assess the influence of metabolites, salts, and ionic strength on the stability of dPGM. In general, salt appears to reduce interactions between metabolites and dPGM while still acting to reduce the accessibility of the intermediate. Through site-directed mutagenesis we investigate the location and mechanism of interactions between chloride and dPGM.

Our results highlight the sensitivity of dPGM to its environment and its ability to harness the chemical environment to dramatic increase stability of the native dimer over the intermediate. The energetic coupling between ligand binding and dimerization in dPGM appears pivotal for governing the enzyme's stability under a physiological milieu.

4.1 Results

4.1.1 Effect of PEP and citrate on the thermodynamic stability of dPGM

Previously, dPGM was shown to be stabilized by nucleotides through interactions with the phosphate groups (Chapter 2). The cellular environment contains many more anionic metabolites and here we examined the equilibrium unfolding of dPGM in the presence of 1.0 mM phosphoenolpyruvate (PEP) and 1.0 mM citrate (Figure 4.1) to understand whether the stabilizing effect observed with the nucleotides is shared by other anionic metabolites. The equilibrium unfolding of dPGM was monitored by intrinsic tryptophan fluorescence. As shown in Figure 4.1, equilibrium unfolding of dPGM appears cooperative with a transition midpoint, C_m, of approximately 3.3 M urea in the presence of both citrate and PEP. Global fitting of the equilibrium unfolding data in the absence and presence of citrate and PEP allowed us to apply the three-state model to all data sets to obtain thermodynamic parameters (Table 4.1). 1.0 mM Citrate increases $\Delta G_{I-N}°$ by 6.1 ± 0.3 kcal/mol and 1.0 mM PEP increases $\Delta G_{I-N}°$ by 6.7 ± 0.3 kcal/mol. PEP confers greater stability to dPGM than the more negatively charged citrate or ADP, but PEP is less stabilizing than ATP (Table 2.2). PEP is close in structure to the dPGM substrate 2-phosphoglycerate, which apparently causes PEP to display a larger stabilizing effect than more negatively charge ADP and citrate. Unlike the other metabolites we have examined,

citrate does not contain phosphate, yet it still dramatically stabilizes dPGM just like the phosphorylated metabolites.

4.1.2 Effect of salt on the stabilizing effect of anionic metabolites

In the cellular environment Mg^{2+} chelation and ionic strength likely affect the Coulombic interactions between anionic metabolites and dPGM. To examine these two factors, we performed equilibrium unfolding of dPGM in the presence of 1.0 mM ATP, and in the presence of 1.0 mM citrate under the following salt conditions: 5.0 mM $MgCl_2$ (Figure 4.2A), 0.10 M NaCl (Figure 4.2B), 5.0 mM $MgCl_2$ and 0.10 M NaCl (Figure 4.2C). Under each salt condition, the midpoint of the first unfolding transition, C_{m1}, is shifted to higher urea concentrations while the midpoint of the second unfolding transition is slightly reduced. In the presence of 0.10 M NaCl, C_{m1} is increased by about 1.0 M urea. In the presence of 5.0 mM $MgCl_2$, C_{m1} is increased by about 0.2 M urea. Under each salt condition, ATP and citrate increase C_{m1} less than observed without salt. With both salts present, ATP and citrate only increase C_{m1} by about 0.3 M urea. Notably, 5.0 mM $MgCl_2$ reduces the interaction between dPGM and the metabolites more than 0.10 M NaCl and effectively enough for equilibrium unfolding to remain clearly three-state in the presence of either ATP or citrate.

To obtain thermodynamic parameters, the equilibrium unfolding of dPGM without a metabolite, with 1.0 mM ATP, and with 1.0 mM citrate were fit globally to the three-state model using eq. 2.4 under each salt condition (Table 4.2). Under all three salt conditions, $\Delta G_{I-N}°$ is increased relative to dPGM without added salt. The smallest increase in $\Delta G_{I-N}°$

was observed in the presence of 5.0 mM $MgCl_2$ ($\Delta\Delta G_{I-N}° = 1.5$ kcal/mol) and the largest increase in $\Delta G_{I-N}°$ was found in the presence of both salts ($\Delta\Delta G_{I-N}° = 3.0$ kcal/mol). With no added salt, 1.0 mM ATP increases $\Delta G_{I-N}°$ by 7.0 kcal/mol (Chapter 2) and 1.0 mM citrate increases $\Delta G_{I-N}°$ by 6.7 kcal/mol. With 5.0 mM $MgCl_2$ present, the stabilizing effect of both metabolites is sharply reduced with 1.0 mM ATP and 1.0 mM citrate raising $\Delta G_{I-N}°$ by only 2.6 ± 0.1 and 2.7 ± 0.1 kcal/mol, respectively. The presence of 0.10 M NaCl has a less dramatic effect on the stabilizing effect of ATP than on that of citrate; the stabilizing effect of citrate is halved while the effect of ATP is hardly attenuated. In the presence of both salts, 1.0 mM ATP and 1.0 mM citrate minimally enhance $\Delta G_{I-N}°$; under this condition, the two salts stabilize dPGM three-fold more than the metabolites. The stability of the intermediate state ($\Delta G_{U-I}°$) does not show a dependence on the presence of salt or metabolite.

The *m*-value is proportional to the change in solvent accessible surface area upon unfolding or the surface area buried upon folding (Myers et al., 1995). For dPGM without added salt, m_{I-N} is 2.40 kcal·mol^{-1}M^{-1}, m_{U-I} is 1.86 kcal·mol^{-1}M^{-1}, and m_{U-N} is 3.06 kcal·mol^{-1}M^{-1} per monomer (Chapter 2). The presence of 0.1 M NaCl shows no effect on m_{I-N}, but when $MgCl_2$ is present, m_{I-N} is greater than without salt. In the presence of 5.0 mM $MgCl_2$, both unfolding transitions are clearly visible (Figure 4.2A) in all data sets, and this results in the most robust fit and *m*-value determination; here, m_{I-N} is 3.09 ± 0.05 kcal·mol^{-1}M^{-1} (1.55 kcal·mol^{-1}M^{-1} per monomer) and m_{U-I} is 1.65 ± 0.03 kcal·mol^{-1}M^{-1}, which results in a m_{U-N} of 3.2 ± 0.04 kcal·mol^{-1}M^{-1} per monomer. In the presence of 5.0 $MgCl_2$, the *m*-value for global unfolding is the same as the predicted *m*-value, which is

based off the number of amino acids in the protein (Myers et al., 1995). The value of m_{U-I} decreases when only $MgCl_2$ is present and increases when NaCl is also present (Table 4.2). The differences in m_{U-I} most likely result from the poor resolution of the intermediate state signal and the second unfolding transition in the equilibrium unfolding of dPGM under both conditions including NaCl. Our results suggest the presence of Mg^{2+} allows the native state of dPGM to adopt a slightly more compact form than without salt and the presence of NaCl does not affect the solvent exposed surface area of dPGM.

4.1.3 The effect of salt on ATP binding

To directly assess the effect of $MgCl_2$ and NaCl on metabolite binding, we determined the dissociation constant (K_d) for ATP binding by dPGM under all three salt conditions using isothermal titration calorimetry (ITC) (Figure 4.3). We previously determined the K_d for dPGM-ATP binding without added salt (0.32 μM). When 5.0 mM $MgCl_2$ was included in the assay, K_d was increased to 37 ± 8 μM. When 0.10 M NaCl was included, K_d was determined to be 2.9 ± 0.2 μM. In the presence of both 5.0 mM $MgCl_2$ and 0.10 M NaCl, no binding was observed. During the ITC experiments 0.40 mM ATP was injected into the sample and only a K_d at least seven-fold lower (~60 μM) should be detected by ITC (Duff, Grubbs, & Howell, 2011). From the stabilizing effect of ATP in the presence of $MgCl_2$ and NaCl ($\Delta\Delta G_{I-N}° = 0.9$ kcal/mol), determined from equilibrium unfolding, the K_d for ATP binding in the presence of both salts was estimated to be 0.76 mM following eq. 2.6. The K_d for ATP binding in the presence of 5.0 mM $MgCl_2$ or 0.10 M NaCl calculated from the results of equilibrium unfolding are 0.13 mM and 0.010 mM, respectfully. The K_d determined from equilibrium unfolding results are about 3.5-fold

greater than those determined by ITC, which could be due to the high urea concentrations used to unfold dPGM. A 3.5-fold reduction in the K_d for ATP binding in the presence of both salts, as determined by equilibrium unfolding, is 0.22 mM. Overall, the ITC results corroborate the effects of $MgCl_2$ and NaCl observed in the equilibrium unfolding of dPGM.

Naturally occurring ATP usually exists as in a metal-chelated form; Mg^{2+} binds to ATP with a K_d of about 0.10 mM, and Na^+ binds to ATP with a K_d of about 77 mM (Wilson & Chin, 1991). The strong suppression of ATP binding by 5.0 mM $MgCl_2$ suggests that Mg^{2+} and dPGM compete for free ATP. If dPGM and Mg^{2+} compete for ATP, the apparent K_d should follow the competitive binding relationship shown in eq 4.1 and derived in the Experimental Procedures:

$$K_d^{app} = K_d\left(\frac{[B]}{K_{dB} - K_d}\right) \tag{4.1}$$

where K_d^{app} is the apparent dissociation constant for dPGM·ATP in the presence of Mg^{2+}, K_d is the dissociation constant for dPGM·ATP in without added salt, [B] is the Mg^{2+} concentration, and K_{dB} is the dissociation constant for Mg·ATP. Under our condition, 5.0 mM $MgCl_2$, the expected K_d^{app} is 16 μM, which is in close agreement with the value determined by ITC (37 μM). Competition between dPGM and Na^+ for ATP is only expected to give a K_d^{app} of 0.42 μM. Rather than compete with dPGM for ATP binding, NaCl most likely reduces the affinity of dPGM for ATP through charge screening or due

to competitive binding between chloride and ATP at the dPGM active site. Citrate is also likely to bind to magnesium as the K_d for $MgCit^-$ is 0.28 mM at pH 7.5 (Walser, 1961). Following eq 2.6, the K_d for dPGM·Citrate without added salt and in the presence of 5.0 mM $MgCl_2$ is 5.8 μM and 0.11 mM, respectfully. If Mg^{2+} and dPGM compete for citrate binding, then following eq 4.1 with a K_d for dPGM·Citrate of 5.8 μM, the K_d^{app} for dPGM·Citrate in the presence of 5.0 mM $MgCl_2$ is expected to be 0.11 mM, in agreement with K_d determined from equilibrium unfolding in the presence of 5.0 mM $MgCl_2$. Under our experimental conditions, Mg^{2+} appears to compete with dPGM for ATP and citrate binding.

4.1.4 Ion-specific effects on the thermodynamic stability of dPGM

We have observed that a range on anionic molecules selectively stabilize native dPGM with varying strengths. To clarify ion-specific effects, we monitored the equilibrium unfolding of dPGM with the following physiological salts at an ionic strength of 0.10 M: KCl, $MgCl_2$, NaBr, NaCl, NaF, Na_2SO_4 (Figure 4.4). All of the salts caused an increase in C_{m1} with $MgCl_2$ having the least effect and Na_2SO_4 causing the most dramatic increase. In contrast to the other salts, the presence of NaF caused a clear increase in C_{m2}. F^- appears to form stabilizing interactions with the intermediate state. The monovalent salts all shift the apparent C_{m1} to about 2.0 – 2.3 M urea. $MgCl_2$ and Na_2SO_4 are clear outliers in their effect on the first unfolding transition while the monovalent salts exhibit similar behavior.

Curve fitting with the three-state model (eq 2.4) reveals that all the salts increased $\Delta G_{I-N}°$, but the magnitude of the stability increase is determined by the anion (Table 4.3). The increase in stability of native dPGM in the presence of $MgCl_2$ ($\Delta\Delta G_{I-N}°$ = 1.97 kcal/mol) is actually similar to the stability of dPGM in the presence of NaCl ($\Delta\Delta G_{I-N}°$ = 1.79 kcal/mol, Table 4.2) and KCl ($\Delta\Delta G_{I-N}°$ = 2.19 kcal/mol) despite a significant difference in the apparent C_{m1}. With chloride as the anion, the choice of cation does not appear to influence the stabilizing effect of the salt on dPGM. However, the presence of 0.10 M NaBr, NaCl, or NaF results in significantly different values of $\Delta\Delta G_{I-N}$ (Table 4.3). The most stabilizing salt, Na_2SO_4, bears the most charged anion and increased $\Delta G_{I-N}°$ by 4.5 ± 0.1 kcal/mol. The stabilizing effects of the salts suggest that the cation does not influence dPGM and the observed stabilizing effects of salts are due to interactions with the anions. Also, the stabilizing effects of anions on $\Delta G_{I-N}°$ appear to be highly specific to the ion type and, therefore, not due to ionic strength effects.

We again find that $MgCl_2$ significantly increases m_{I-N} and KCl, NaBr, and NaF also increase m_{I-N} compared to dPGM without added salt. The high *m*-value observed in the presence of $MgCl_2$ causes dPGM to unfold at lower urea concentrations that anticipated given the stabilizing effect of $MgCl_2$. The presence of NaBr or NaF raised m_{I-N} by more than 0.35 kcal mol^{-1} M^{-1} compared to dPGM without added salt, which is a 10% increase. The effect of the salts on m_{I-N} supports the idea that native dPGM is more expanded in the low ionic strength buffer without added salt. In the presence of NaF, both m_{U-I} and $\Delta G_{U-I}°$ are much greater than in buffer alone; this actually causes the global stability of dPGM to be greater in the presence of NaF than in the presence of Na_2SO_4, although the

intermediate is populated at lower urea concentrations in the presence of NaF. The primary effect of salt on the *m*-values of dPGM appears to be a small increase in $m_{I–N}$ which could signify that that native dimer is more compact in the presence of added salt than in buffer alone.

4.1.5 Effect of NaCl concentration on the thermodynamic stability of dPGM

In the cellular context, dPGM is likely to experience a range of salt concentrations. To clarify the dependence of dPGM stability on salt concentration, we examined the equilibrium unfolding of dPGM in presence of NaCl ranging from 0.020 M to 0.50 M NaCl (Figure 4.5). As the concentration of NaCl is increased, the first unfolding transition of dPGM shifts to higher urea concentrations, while the second transition remains stable. At 0.50 M NaCl, the midpoints of both transitions are close in value, and equilibrium unfolding of dPGM appears cooperative. The equilibrium unfolding of dPGM at each NaCl concentration, including in buffer alone, was fit globally with Eq. 2.4 to obtain thermodynamic parameters (Table 4.4). As anticipated, $\Delta G_{I–N}°$ shows a direct dependence upon the NaCl concentration while $\Delta G_{U–I}°$ does not. The effect of the highest NaCl concentrations, 0.50 M NaCl, is a $\Delta\Delta G_{I–N}°$ of 3.59 kcal/mol, which is still about 1 kcal/mol less than the effect of 0.33 mM Na_2SO_4. The effect of 0.10 M NaCl on $\Delta G_{I–N}°$ as reported by the global fit ($\Delta\Delta G_{I–N}°$ = 2.31 kcal/mol) is close in value to the effect of 0.10 KCl. The *m*-values for the NaCl series are $m_{I–N}$ = 2.61 ± 0.02 kcal·mol^{-1}M^{-1}, $m_{U–I}$ = 1.94 ± 0.04 kcal·mol^{-1}M^{-1}, and $m_{U–N}$ = 3.25 kcal·mol^{-1}M^{-1} per monomer. The *m*-values determined by the global fitting are more robust and are slightly greater in value than those determined in the presence of 0.10 M NaCl (Table 4.1). The global *m*-value in

the presence of NaCl is in agreement with value determined in the presence of 5.0 mM $MgCl_2$ and the theoretical value based on the dPGM sequence. The primary effect increasing NaCl concentrations is an enhancement of $\Delta G_{I-N}°$.

4.1.6 The effect of point mutations on the chloride-induced stabilization of dPGM.

Two regions of dPGM are prime candidates for interactions between dPGM and chloride ions: the active site and the dimer interface. The crystallographic structure of dPGM reveals an extremely basic active site and a chloride ion bound in the dimer interface between Trp67 and Lys82 from each monomer (Figure 4.6) (Bond et al., 2001, 2002). We have found that ATP binds in the dPGM active site (Chapter 2) and here, we have shown that salts can stabilize dPGM similar to ATP. To probe the involvement of the dimer interface and the active site of dPGM in binding to chloride, we analyzed the effect of NaCl on three dPGM Variants: W67A, K99A, and R115A dPGM (Figure 4.7A, B, C). W67A dPGM was previously characterized in the presence of 0.10 M to 0.50 M NaCl (Chapter 3) and found to be greatly destabilized, but still able to fold to the native state. Lys99 and Arg115 are located in the active site and were previously shown to be involved in ATP binding (Chapter 2). The effect of NaCl on the stability of the three dPGM variants is compared with wild-type and K222A dPGM (Figure 4.7D). The K222A mutation will test the effect of removing a charged side chain distal to the active site and dimer interface (Figure 4.6). If a side chain is involved in forming stabilizing interactions with chloride, then an alanine mutation should result in a less favorable interaction between chloride and dPGM. A decrease in the favorability of the interaction will be apparent if the $\Delta\Delta G_{I-N}°$ due to NaCl is less for a dPGM variant than for wild-type

dPGM. To compare the stabilizing effect of chloride upon each variant, we conducted equilibrium unfolding at several NaCl concentrations for each variant (Figure 4.7).

For all dPGM variants, equilibrium unfolding at each NaCl concentration and without added salt were fit globally using eq 2.4 (Table 4.5). We find that the W67A mutation reduced $\Delta G_{I-N}°$ by 3.22 kcal/mol relative to wild-type dPGM without added salt. Using the thermodynamic parameters, the fractional population of the W67A dPGM intermediate state was determined to be 0.47 at 0 M urea without added salt. In the presence of 0.5 M NaCl, the population of the W67A dPGM intermediate state is reduced to 0.07 with the rest being native dPGM. For each dPGM variant, the $\Delta G_{I-N}°$ increases alongside the NaCl concentration. At each concentration of NaCl, the $\Delta\Delta G_{I-N}°$ for W67A dPGM is significantly reduced compared to wild-type dPGM, where $\Delta\Delta G_{I-N}° = \Delta G_{I-N}°(\text{NaCl}) - \Delta G_{I-N}°(\text{no salt})$. In the presence of 0.20 M NaCl, the $\Delta\Delta G_{I-N}°$ for wild-type is dPGM is 0.8 kcal/mol more than the $\Delta\Delta G_{I-N}°$ for W67A dPGM. Clearly some stabilizing interactions with chloride are lost from the W67A mutation. Both K99A and R115A dPGM are more stable than wild-type at each NaCl concentration. For K99A dPGM, $\Delta\Delta G_{I-N}°$is significantly less than that of wild-type dPGM at NaCl concentrations of 0.050 mM or greater. R115A dPGM shows a $\Delta\Delta G_{I-N}°$ that is significantly less than wild-type dPGM at NaCl concentrations of 0.10 M or greater (Table 4.5). Both active-site mutations reduce $\Delta\Delta G_{I-N}°$ in the presence of NaCl, but much less than the W67A mutation. For K222A dPGM, the effect of NaCl appears nearly identical to the effect on wild-type dPGM (Table 4.5). W67A, K99A, and R115A dPGM are less stabilized by NaCl than WT and K222A dPGM. The reduced effect of NaCl on W67A, K99A, and

R115A dPGM supports the role of the dimer interface and active site in binding chloride. The W67A mutation was the most efficacious mutation in reducing $\Delta\Delta G_{I-N}°$ which suggests that chloride binding at the dimer interface is highly dependent upon Trp67. The dPGM active site contains many positively charged residues and the effect of NaCl on K99A and R115A dPGM reveals that the involvement of any one residue is likely not critical.

4.2 Discussion

We have demonstrated that the native state of dPGM is selectively stabilized over a partially unfolded, monomeric intermediate by many abundant, anionic metabolites and salts that dPGM will encounter in the cellular environment. For example, the intracellular citrate and PEP concentrations in glucose-fed *E. coli* are about 2.0 mM and 0.18 mM, respectfully (Bennett et al., 2009). The concentration of citrate and PEP that we used (1.0 mM) appears relevant to the physiological milieu. Although several anionic metabolites show a tight binding to dPGM in buffer, the effect of physiological salts, especially Mg^{2+}, will greatly reduce metabolite binding (Figure 4.2). However, we observed that citrate and ATP still bind to dPGM in the presence of excess $MgCl_2$ and NaCl which suggests that they may bind dPGM in the *E. coli* cytosol. While salt reduces the ability of metabolites to stabilize dPGM, small anions like chloride compensate for the reduction and selectively enhance the stability of native dPGM. Previously, we described an energetic linkage between the active site of dPGM and the dimer interface (Chapter 2). Investigation into the structure of the equilibrium intermediate of dPGM revealed the cooperative unit that links the active site and dimer interface (Chapter 3). The

cooperativity between the two sites gives rise to allostery between dimerization and ligand binding. Interestingly, the effects of NaCl on W67A, K99A, and R115A dPGM suggest that chloride binds to both of the allosteric sites.

For protein with an equilibrium intermediate, the energy gap between the native protein and the first inactive intermediate, $\Delta G_{I-N}°$, is the most relevant to conformational stability as the intermediate is the most likely non-native confirmation to be populated (Sancho et al., 2002). To put the stabilizing effect of salts and metabolites on dPGM in perspective, we determined the effective stability, $\Delta G_{eff} = -RT \ln[(1 - f_N)/f_N]$, (Park & Marqusee, 2004a) of dPGM at 0 M urea from the ratio of nonnative to native protein, $(1 - f_N)/f_N$. The effective stability is comparable amongst protein of varying oligomeric state, and it takes into account the relevant stability. Without added salt, the effective stability of 1 μM dPGM is 1.9 kcal/mol; in the presence of 5.0 mM MgCl2, 0.10 M NaCl, and 1.0 M ATP, the effective stability of 1 μM dPGM is 3.9 kcal/mol. The presence of salt and a metabolite doubled the effective stability of dPGM. Without urea and in only 20 mM buffer, 1 in 25 dPGM monomers should adopt the intermediate conformation. With a ΔG_{eff} of 3.9 kcal/mol, about 1 in every 750 dPGM monomers should adopt the intermediate conformation. *E. coli* dPGM appears to have evolved to interact with abundant anions and metabolites to achieve a proper stability in its physiological milieu.

The interactions between protein and salt can involve any combination of preferential solvation, Hofmeister effects, electrostatic interactions arising from the ionic strength as described by the Debye–Hückel approximation, and specific binding (de los Rios &

Plaxco, 2005; Record Jr, Zhang, & Anderson, 1998). With dPGM we observed specificity for the anion type and that ionic strength did not explain the effects of salt on $\Delta\Delta G_{I-N}°$ (Table 4.3). Therefore, we expected that chloride forms specific binding interactions with dPGM. The effect of NaCl on the three dPGM variants, W67A, K99A, and R115A dPGM confirmed that chloride forms specific interactions with dPGM to stabilize the native state over the intermediate. Because the crystallographic structure of dPGM shows that Trp67 forms interactions with a buried chloride ion (Bond et al., 2001, 2002), removal of the Trp67 side chain should reduce the stabilizing effect of NaCl. Compared to wild-type dPGM the $\Delta\Delta G_{I-N}°$ due to NaCl for W67A dPGM is, on average, 26% lower. W67A dPGM is still greatly stabilized in the presence of NaCl, which suggests that the majority of the stabilizing effect of NaCl is due to chloride binding within the active site. The presence of chloride is more stabilizing for K99A and R115A dPGM than for W67A dPGM. The large active site of dPGM contains many positively charged side chains (Bond et al., 2001), and we expect that multiple chloride ions can bind, especially considering that the active site accommodates substrate bearing multiple phosphate groups. The lesser reduction in the $\Delta\Delta G_{I-N}°$ of K99A and R115A dPGM compared to wild-type dPGM supports the idea that multiple chloride ions bind within the active site. The investigators who solved the structure of the homologous human dPGM postulated that low concentrations of chloride primarily stabilize the quaternary structure of dPGM while higher concentrations primarily stabilize the active site (Yanli Wang et al., 2005). Our results suggest that chloride binds with greater affinity at the dimer interface but the positive cooperativity between the active site and dimer interface

of dPGM causes binding at either site to stabilize the quaternary structure and the active site (Chapter 3).

Although we demonstrated the stabilizing effects of metabolites and salts on dPGM *in vitro*, the observed effects seem to be relevant to the stability of dPGM in the *E. coli* cytosol. First, about 1.4×10^4 copies of dPGM exists per *E. coli* cell (Ishihama et al., 2008), which amounts to about 14 μM after accounting for cell volume (Kubitschek & Friske, 1986). This concentration is 14-fold greater than that in most of our experiments and will stabilize the native dimer relative to our conditions. During our investigation, we kept the concentrations of metabolites within physiologically relevant ranges. Here and in our energetics-based proteomics screen where dPGM was identified (Chang et al., 2012), we used 1.0 mM ATP to approximate the concentration of free ATP in the *E. coli* cytosol (Schneider & Gourse, 2004). We observed that GTP as well as ATP stabilizes dPGM extensively (Chapter 2), which suggests that any nucleotides in cells can contribute to the stability of the protein. The total intracellular pool of triphosphate nucleotides (NTPs) in glucose-fed *E. coli* is about 27 mM (Bennett et al., 2009). Though a large fraction of the NTPs are bound to proteins and the concentration of free NTPs would be significantly lower than 27 mM, the intracellular pool of NTPs, plus the diphosphate and monophosphate nucleotides, should convey more stability to the protein than we observed with 1.0 mM ATP. As mentioned, the intracellular citrate and PEP concentrations are also relevant. We have shown that just 0.1 mM pyrophosphate increases dPGM stability such that equilibrium unfolding appears cooperative (Chapter 2), and pyrophosphate is maintained at a concentration of about 0.3 mM throughout the

E. coli growth cycle (Jukka & Harold, 1988). Therefore, several of the metabolites we examined were present at physiologically relevant concentrations.

Anions compete with the metabolites to stabilize dPGM primarily through interactions in the active site. We observed that both ATP and citrate binding to dPGM was drastically attenuated by 0.1 M NaCl and 5.0 mM $MgCl_2$ (Table 4.2). The intracellular chloride concentration of *E. coli* has a linear dependence on the extracellular concentration and does not appear to be tightly regulated (Schultz, Wilson, & Epstein, 1962). Many strains of *E. coli* can grow with up to 0.5 M extracellular NaCl (Kunin, Hua, Van Arsdale White, & Villarejo, 1992); the entire range of NaCl concentrations used in this study could be experienced by *E. coli*. Free Mg^{2+} concentrations are thought to be low, on the order of 1 to 2 mM (Alatossava, Jütte, Kuhn, & Kellenberger, 1985), which is less than the concentration in our experiments. Although pin pointing how much each metabolite or anion contributes to the stability of dPGM in *E. coli* cells is not feasible, our data strongly suggest that salts will greatly reduce, but not eliminate, interactions between dPGM and metabolites. Based on the effect of ATP and Citrate on the stability of dPGM in presence of salt (Table 4.2), salts likely confer more stability to dPGM than metabolites under a physiological milieu. Still, our results suggest that dPGM is able to bind the metabolites in the cytosol. Collectively, the metabolites and anions in the cellular environment should stabilize native dPGM over the intermediate significantly more than the conditions we examined here.

Because most interactions between metabolites or salts and dPGM appear to occur in the active site of dPGM, metabolites and salts may inhibit catalysis by dPGM. However, as long as the inhibition is competitive, a sufficient amount of the substrate would overcome the inhibition. Activity studies on chicken breast dPGM have demonstrated that several anions competitively inhibit substrate binding to dPGM (Rose & Dube, 1978; Rose & Kaklij, 1984). The total 3-phosphoglycerate (3-PG) concentration in glucose-fed *E. coli* has been found to be 1.5 mM (Bennett et al., 2009). The reported K_m values for 3-PG and 2-PG are 0.20 ± 0.03 mM and 0.19 ± 0.04 mM, respectfully (Fraser et al., 1999). The main catalytic activity of dPGM is driven by the equilibrium constant between 3-PG and 2-PG and must function in opposite directions for glycolysis and gluconeogenesis. A low K_m value would suggest that product is released rapidly, but a consequence would be rapid release of substrate in the reverse reaction (Bennett et al., 2009). Therefore, reversible reactions tend to have large K_m values so substrate binding is productive and to maintain sensitivity to substrate flux (Bennett et al., 2009). The K_m value for 3-PG determined with chicken breast dPGM was found to increase with the square of the KCl concentration (Rose & Dube, 1978). K_m values determined for many enzymes are below the concentrations of their substrates in *E. coli*, which should cause enzyme saturation and insensitivity to metabolic flux; however, competitive inhibition from other metabolites or salts may bring up the K_m value for the substrate to an adequate level for enzymes to be sensitized to the substrate concentration (Bennett et al., 2009). An ordinary-differential-equation-based modeling of the central nitrogen metabolism in *E. coli* predicts that competition for saturated enzyme active sites maintains sensitivity to metabolic flux (Yuan et al., 2009); competitive inhibition of glutamine synthetase was

even predicted to occur from non-functional or non-substrate metabolites (Yuan et al., 2009). The non-substrate metabolites we examined may competitively inhibit dPGM in the *E. coli* cytosol. If high ATP concentrations inhibit the mutase activity of dPGM, ATP binding could serve as a negative feedback loop to slow glycolysis. The effects of anions on dPGM, that we observed, suggest that physiological anions and metabolites will greatly increase K_m values for 3-PG and 2-PG compared to those determined for *E. coli* dPGM *in vitro*.

Under the right conditions, could dPGM have activity for several of the metabolites (ATP, citrate, and pyrophosphate) we have tested? The metabolites appear to bind at relevant concentrations. The crystallographic structure of the homologous human dPGM-B was solved with citrate bound in the active site (Yanli Wang et al., 2005). The binding mode of citrate is similar to the tetravanadate found in *E. coli* dPGM and was predicted to be similar to the binding mode of 3-phosphoglycerate (Yanli Wang et al., 2005). Further, human dPGM-B was also found to convert PEP to pyruvate and retain the phosphoryl group on His11 (human dPGM numbering) in rapidly proliferating cells (Vander Heiden et al., 2010). This suggests that a change in pH, metabolite concentration, or post translational modification could alter the activity of dPGM. We tested whether dPGM was able to hydrolyze ATP by monitoring ATP in the presence of dPGM by ^{32}P NMR. The three peaks of ATP shift in the presence of dPGM relative to ATP alone, which indicates binding, but the peaks did not change over the period of 24 hours (data not shown); we did not detect any hydrolysis of ATP to ADP. We observed that dPGM appears to compete with Mg^{2+} for ATP binding, which is clearly distinct from ATP

binding observed for kinases,(Johnson, Lowe, Noble, & Owen, 1998; Karthikeyan et al., 2003) ATP-binding cassettes,(Davidson, Dassa, Orelle, & Chen, 2008) and P-type ATPases, (Banci et al., 2010; Bublitz, Morth, & Nissen, 2011) which bind Mg^{2+}-chelated ATP. Therefore, ATP binding to dPGM is not for the common uses of ATP in biochemical reactions such as hydrolysis or phosphotransfer. We only observed stabilizing effects of the metabolites on *E. coli* dPGM and the most probable physiological effect of the interactions appears to be competitive inhibition.

E. coli dPGM has evolved to bind many abundant anions and enhance its conformational cooperativity by promoting the dimeric form over a partially unfolded, monomeric intermediate. A theoretical model of the evolutionary effects of random mutation predicts that proteins may frequently evolve nonfunctional binding sites that aid in stability (Manhart & Morozov, 2015). Even when fitness is determined solely by biochemical activities of a protein, more stable proteins are selected through evolution because promoting the native state promotes activity. Marginally stable proteins are predicted to evolve more interactions with abundant solutes to increase the stability of the protein in its physiological milieu and ultimately, promote the protein's function (Manhart & Morozov, 2015). Mutations that favor intrinsic stability tend to result in a loss of function (Beadle & Shoichet, 2002; Meiering, Serrano, & Fersht, 1992; Shoichet, Baase, Kuroki, & Matthews, 1995). Non-substrate ligand binding offers an alternative way to gain conformational stability without suppressing enzyme function other than an increase in K_m when binding is competitive with substrate. As discussed above, the increase in K_m may also be beneficial. Extrapolating to the cellular environment is not clear; protein can

be substantially destabilized in the cell environment in a sequence-dependent manor (Danielsson et al., 2015; Miklos, Sarkar, Wang, & Pielak, 2011). Destabilization appears to occur because many proteins can bind portions of the unfolded sequence and promote the unfolded state (Danielsson et al., 2015). Simulations are beginning to take on the task of answering how *in vitro* results translate to *in vivo* conditions with a focus on the effect of crowding (Frembgen-Kesner & Elcock, 2013; McGuffee & Elcock, 2010). In addition to crowding and interacting with other proteins, our results and a number of screens for protein-metabolite interactions (Li et al., 2010; Link, Kochanowski, & Sauer, 2013; Orsak et al., 2011; Roelofs, Wang, Sintim, & Lee, 2011), highlight the impact non-substrate metabolite and ion binding can have in governing protein stability in the cellular environment.

4.3 Experimental Procedures

4.3.1 Preparation of purified dPGM

dPGM was over-expressed in *E. coli* BL21 (DE3) pLysS as previously described (Chapter 2). All dPGM variants were purified through two rounds of anion exchange chromatography with dialysis before the second round, as previously detailed (Chapter 2) (Chapter 2). The purity of dPGM was determined by SDS-PAGE. The final concentration of dPGM was determined by absorbance at 280 nm using an extinction coefficient estimated from the amino acid composition (Pace et al., 1995). Mutagenesis to produce the dPGM variants was carried out previously (Chapter 2, Chapter 3). Protein stocks were stored at 4°C or at -80°C for extended storage.

4.3.2 Equilibrium unfolding

Urea-induced equilibrium unfolding of dPGM was monitored by intrinsic tryptophan fluorescence. For all experiments 0.030 mg/mL dPGM was prepared in 20 mM TrisHC−NaOH buffer (pH 8.0) containing varying concentrations of urea (0 – 5 M). The urea concentrations were determined from the refractive index (Pace, 1986). Samples were filtered through a 0.2 µm filter and equilibrated overnight at 25°C prior to spectroscopy. Fluorescence emission spectra spanning 320 to 400 nm were obtained with a FlouroMax-3 fluorimeter (Horiba Jobin Yvon; Edison New Jersey, NJ) using an excitation wavelength of 280 nm. During data collection the samples were maintained at 25°C.

The equilibrium unfolding data is analyzed as intensity-weighted average wavelength (IWAW) as previously described (Chapter 2). The data was fit with a three-state model detailed in scheme 1:

$$N_2 \rightleftarrows 2I \rightleftarrows 2U$$

Scheme 1

Fitting was carried out as previously described using OriginPro 8.5.1 (OriginLab; Northampton, MA) to determine *m*-values and the free energy difference between states, $\Delta G_{I-N}°$ and $\Delta G_{U-I}°$ (Chapter 2). For global fitting, the fluorescence signal of the

intermediate state, the dPGM concentration, and both *m*-values are shared amongst all data sets.

4.3.3 Isothermal titration calorimetry

For ITC, the dPGM samples were prepared at 42 μM (1.2 mg/mL) and dialyzed against 20 mM TrisHCl−NaOH buffer (pH 8.0) under several salt conditions: 5.0 mM $MgCl_2$, 0.10 M NaCl, 5.0 $MgCl_2$ and 0.10 M NaCl. After dialysis, the dPGM concentration was determined by absorbance at 280 nm. ATP (Sigma-Aldrich; St. Louis, MO) was prepared at a concentration of 2.0 mM in the dialysis buffer. The heat evolution was monitored with A MicroCal iTC200 calorimeter (GE Healthcare Life Sciences; Piscataway, NJ) upon ATP titration. ATP was titrated through 21 injections of 0.40 mM ATP with 3-minute intervals. The resulting thermograms were automatically integrated by NITPIC (Keller et al., 2012) and fit using SEDPHAT (https://sedfitsedphat.nibib.nih.gov).

4.3.4 Estimating the apparent K_d for dPGM−metabolite binding in the presence of a competitive cation

We derived a simple model to estimate the apparent K_d (K_d^{app}) that would be observed if dPGM and Mg^{2+} compete with one another to bind ATP (eq 4.1). When dPGM, ATP, and Mg^{2+} are present, the equilibrium is defined by constants K_{d1} and K_{d2}:

$$K_{d1} = \frac{[P][ATP]}{[P \cdot ATP]} \tag{4.2}$$

$$K_{d2} = \frac{[Mg^{2+}][ATP]}{[Mg^{2+} \cdot ATP]} \tag{4.3}$$

where [P], [ATP], and [Mg^{2+}] are the concentrations of free dPGM, ATP, and Mg^{2+}, respectfully. [P·ATP] and [Mg^{2+}·ATP] are the concentrations of ATP-bound dPGM and Mg^{2+}, respectfully. Mass conservation dictates that the total concentrations of P, ATP, and Mg^{2+} (P_t, $[ATP]_0$; and $[Mg^{2+}]_0$, respectfully) are defined as follows:

$$P_t = [P] + [P \cdot ATP] \tag{4.4}$$

$$[ATP]_0 = [ATP] + [Mg^{2+} \cdot ATP] + [P \cdot ATP] \tag{4.5}$$

$$[Mg^{2+}]_0 = [Mg^{2+}] + [Mg^{2+} \cdot ATP] \tag{4.6}$$

Because the concentration of dPGM is low (42 μM during ITC), even if all dPGM are saturated by ATP, the concentration of free ATP is 10-fold larger and is not greatly reduced by binding to dPGM. Therefore, we assume that the term [P·ATP] is negligible in eq 4.5. Then, [Mg^{2+}·ATP] = $[ATP]_0$ − [ATP]. Because Mg^{2+} is in excess compared to ATP and present at a concentration 50-fold greater than the K_d for Mg^{2+}·ATP, we also make the assumption that all ATP is bound so that [Mg^{2+}·ATP] ≈ $[ATP]_0$. Given this assumption, eq 6 can be solved for [Mg^{2+}]:

$$[Mg^{2+}] = [Mg^{2+}]_0 - [ATP]_0 \tag{4.7}$$

By solving eq 4.4 for [P] and substitution into eq 4.2, [ATP] can be defined by

$$[ATP] = \frac{[P \cdot ATP]K_{d1}}{P_t - [P \cdot ATP]} \tag{4.8}$$

Substitution of eq 4.7, 4.8, and the approximation for [Mg^{2+}·ATP] into eq 4.3 yields the definition

$$K_{d2} = \frac{\left([Mg^{2+}]_0 - [ATP]_0\right)\left(\frac{[P \cdot ATP]K_{d1}}{P_t - [P \cdot ATP]}\right)}{[ATP]_0} \tag{4.9}$$

Simplifying eq 4.9 and solving it for [P·ATP] gives the expression

$$[P \cdot ATP] = \frac{2P_t K_{d2}[ATP]_0}{K_{d1}[Mg^{2+}]_0 + [ATP]_0(K_{d2} - K_{d1})} \tag{4.10}$$

K_d^{app} is defined as the concentration of ATP where half of the dPGM monomers appear to be bound. When $[ATP]_0 = K_d^{app}$, the fraction of bound dPGM (f_B) is defined by

$$f_B = \frac{[P \cdot ATP]}{P_t} = \frac{1}{2} \tag{4.11}$$

Substitution of eq 4.11 into eq 4.10, with $[ATP]_0 = K_d^{app}$, results in the competitive binding relationship given in eq 4.1.

Table 4.1 The effect of PEP and Citrate on dPGM stability.

Metabolite[a]	m_{N-I} (kcal mol^{-1}M^{-1})	$\Delta G_{I-N}°$ (kcal/mol)	$\Delta\Delta G_{I-N}°$[b] (kcal/mol)	m_{I-U} (kcal mol^{-1}M^{-1})	$\Delta G_{U-I}°$ (kcal/mol)
−	2.49 ± 0.08	11.6 ± 0.1	−	1.9 ± 0.1	6.6 ± 0.5
Citrate	−	17.7 ± 0.3	6.1 ± 0.3	−	5.9 ± 0.5
PEP	−	18.3 ± 0.3	6.7 ± 0.3	−	5.7 ± 0.5

[a]All metabolite concentrations are 1.0 mM.
[b]$\Delta\Delta G_{I-N}° = \Delta G_{I-N}(\text{metabolite})° - \Delta G_{I-N}(\text{no metabolite})°$

Table 4.2 The effect of a metabolite on the stability of dPGM in salt.

Metabolite[a]	[NaCl] (M)	[$MgCl_2$] (mM)	$m_{I–N}$ (kcal $mol^{-1}M^{-1}$)	$\Delta G_{I–N}°$ (kcal/mol)	$\Delta\Delta G_{I–N}°$[b] (kcal/mol)	$m_{U–I}$ (kcal $mol^{-1}M^{-1}$)	$\Delta G_{U–I}°$ (kcal/mol)
–	0	5.0	3.09 ± 0.05	13.0 ± 0.1	–	1.65 ± 0.03	5.5 ± 0.1
ATP	0	5.0	–	15.6 ± 0.1	2.6 ± 0.1	–	5.5 ± 0.1
Citrate	0	5.0	–	15.7 ± 0.1	2.7 ± 0.1	–	5.5 ± 0.1
–	0.10	0	2.5 ± 0.1	13.5 ± 0.3	–	1.5 ± 0.1	4.6 ± 0.5
ATP	0.10	0	–	19.0 ± 0.4	5.5 ± 0.5	–	3.8 ± 0.7
Citrate	0.10	0	–	16.9 ± 0.3	3.4 ± 0.4	–	4.2 ± 0.5
–	0.10	5.0	2.88 ± 0.04	14.5 ± 0.1	–	2.1 ± 0.1	6.7 ± 0.4
ATP	0.10	5.0	–	15.5 ± 0.1	1.0 ± 0.1	–	6.8 ± 0.4
Citrate	0.10	5.0	–	15.3 ± 0.1	0.8 ± 0.1	–	6.7 ± 0.4

[a] All metabolite concentrations are 1.0 mM.
[b] $\Delta\Delta G_{I–N}° = \Delta G_{I–N}(\text{metabolite})° - \Delta G_{I–N}(\text{no metabolite})°$

Table 4.3 The effect of salts on the stability of dPGM at an ionic strength of 0.10 M.

Salt	m_{I-N} (kcal mol^{-1}M^{-1})	$\Delta G_{I-N}°$ (kcal/mol)	$\Delta\Delta G_{I-N}°$[a] (kcal/mol)	m_{U-I} (kcal mol^{-1}M^{-1})	$\Delta G_{U-I}°$ (kcal/mol)	$\Delta G_{U-N}°$[b] (kcal/mol)
$MgCl_2$[c]	2.99 ± 0.02	13.68 ± 0.4	1.97 ± 0.05	1.95 ± 0.06	6.1 ± 0.2	25.9 ± 0.4
NaCl[d]	2.61 ± 0.02	14.02 ± 0.04	2.31 ± 0.05	1.94 ± 0.04	6.5 ± 0.2	27.0 ± 0.4
KCl[d]	2.62 ± 0.03	13.90 ± 0.06	2.19 ± 0.07	2.2 ± 0.2	7.4 ± 0.8	29 ± 2
NaBr[d]	2.82 ± 0.06	14.6 ± 0.2	2.9 ± 0.2	2.1 ± 0.1	6.4 ± 0.6	27 ± 1
NaF[d]	2.77 ± 0.05	14.18 ± 0.09	2.47 ± 0.09	2.3 ± 0.1	8.0 ± 0.5	30 ± 1
Na_2SO_4[c,e]	2.48 ± 0.04	16.2 ± 0.1	4.5 ± 0.1	1.59 ± 0.07	5.0 ± 0.3	26.2 ± 0.6

[a] $\Delta\Delta G_{I-N}° = \Delta G_{I-N}°(\text{Salt}) - \Delta G_{I-N}°(\text{no salt})$

[b] $\Delta G_{U-N}° = \Delta G_{I-N}° + 2(\Delta G_{U-I}°)$

[c] Equilibrium unfolding was performed with 33.3 mM salt.

[d] Equilibrium unfolding was performed with 0.10 M salt.

[e] Equilibrium unfolding was fit globally along with dPGM in buffer alone.

Table 4.4 Thermodynamic parameters of dPGM in the presence of 0 – 0.5 M NaCl.

[NaCl] (mM)	m_{I-N} (kcal mol^{-1}M^{-1})	$\Delta G_{I-N}°$[a] (kcal/mol)	$\Delta\Delta G_{I-N}°$ (kcal/mol)	m_{U-I} (kcal mol^{-1}M^{-1})	$\Delta G_{U-I}°$ (kcal/mol)	$\Delta G_{U-N}°$[b] (kcal/mol)
0	2.61 ± 0.02	11.71 ± 0.03	–	1.94 ± 0.04	6.8 ± 0.2	25.3 ± 0.4
20	–	12.49 ± 0.03	0.78 ± 0.04	–	6.6 ± 0.2	25.7 ± 0.4
50	–	13.35 ± 0.04	1.64 ± 0.05	–	6.5 ± 0.2	26.4 ± 0.4
100	–	14.02 ± 0.04	2.31 ± 0.05	–	6.5 ± 0.2	27.0 ± 0.4
200	–	14.63 ± 0.05	2.92 ± 0.06	–	6.6 ± 0.2	27.8 ± 0.4
500	–	15.30 ± 0.05	3.59 ± 0.06	–	7.0 ± 0.2	29.3 ± 0.4

[a] $\Delta\Delta G_{I-N}° = \Delta G_{I-N}°(\text{NaCl}) - \Delta G_{I-N}°(\text{no salt})$

[b] $\Delta G_{U-N}° = \Delta G_{I-N}° + 2(\Delta G_{U-I}°)$

Table 4.5 The effect of NaCl on the stability of dPGM variants.

Variant	[NaCl] (M)	m_{I-N} (kcal mol^{-1}M^{-1})	$\Delta G_{I-N}°$ (kcal/mol)	$\Delta\Delta G_{I-N}°$[a] (kcal/mol)	m_{U-I} (kcal mol^{-1}M^{-1})	$\Delta G_{U-I}°$ (kcal/mol)
W67A	0	2.83 ± 0.02	8.28 ± 0.02	–	1.66 ± 0.01	5.65 ± 0.01
	0.10	–	10.03 ± 0.03	1.75 ± 0.04	–	5.39 ± 0.03
	0.20	–	10.40 ± 0.03	2.12 ± 0.04	–	5.40 ± 0.04
	0.50	–	10.90 ± 0.03	2.63 ± 0.04	–	5.76 ± 0.04
K99A	0	2.79 ± 0.03	14.06 ± 0.06	–	2.09 ± 0.08	7.5 ± 0.3
	0.020	–	14.82 ± 0.07	0.76 ± 0.09	–	7.2 ± 0.3
	0.050	–	15.43 ± 0.07	1.37 ± 0.09	–	7.1 ± 0.3
	0.10	–	15.81 ± 0.07	1.75 ± 0.09	–	7.1 ± 0.3
	0.20	–	16.38 ± 0.08	2.3 ± 0.1	–	7.3 ± 0.3
R115A	0	2.89 ± 0.04	15.2 ± 0.1	–	2.2 ± 0.1	8.0 ± 0.5
	0.020	–	16.1 ± 0.1	0.9 ± 0.1	–	7.7 ± 0.5
	0.050	–	16.7 ± 0.1	1.5 ± 0.1	–	7.5 ± 0.4
	0.10	–	17.2 ± 0.1	2.0 ± 0.1	–	7.6 ± 0.4
	0.20	–	17.6 ± 0.1	2.4 ± 0.1	–	7.6 ± 0.4
K222A	0	2.61 ± 0.02	11.79 ± 0.03	–	1.87 ± 0.02	6.01 ± 0.08
	0.020	–	12.75 ± 0.04	0.96 ± 0.05	–	5.81 ± 0.07
	0.050	–	13.40 ± 0.04	1.61 ± 0.05	–	5.76 ± 0.07
	0.10	–	14.21 ± 0.05	2.42 ± 0.06	–	5.74 ± 0.07
	0.20	–	14.84 ± 0.05	3.05 ± 0.06	–	5.80 ± 0.07

[a] $\Delta\Delta G_{I-N}° = \Delta G_{I-N}°(\text{NaCl}) - \Delta G_{I-N}°(\text{no salt})$

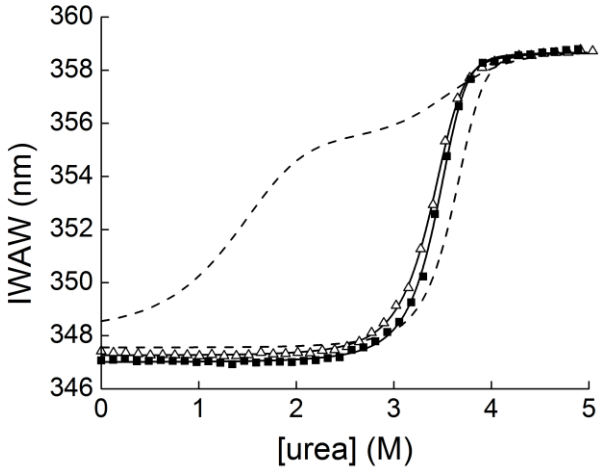


Figure 4.1 Effect of citrate and PEP on the equilibrium unfolding of dPGM

Equilibrium unfolding of dPGM was performed in the presence of 1.0 mM Citrate (△) and 1.0 mM PEP (■) and the data was fit globally along with equilibrium unfolding of dPGM without ligand using eq 2.4 (solid lines) (Table 4.1). Fitting results for the equilibrium unfolding of dPGM without ligand and in the presence of 1.0 mM ATP are shown for reference (dashed lines).

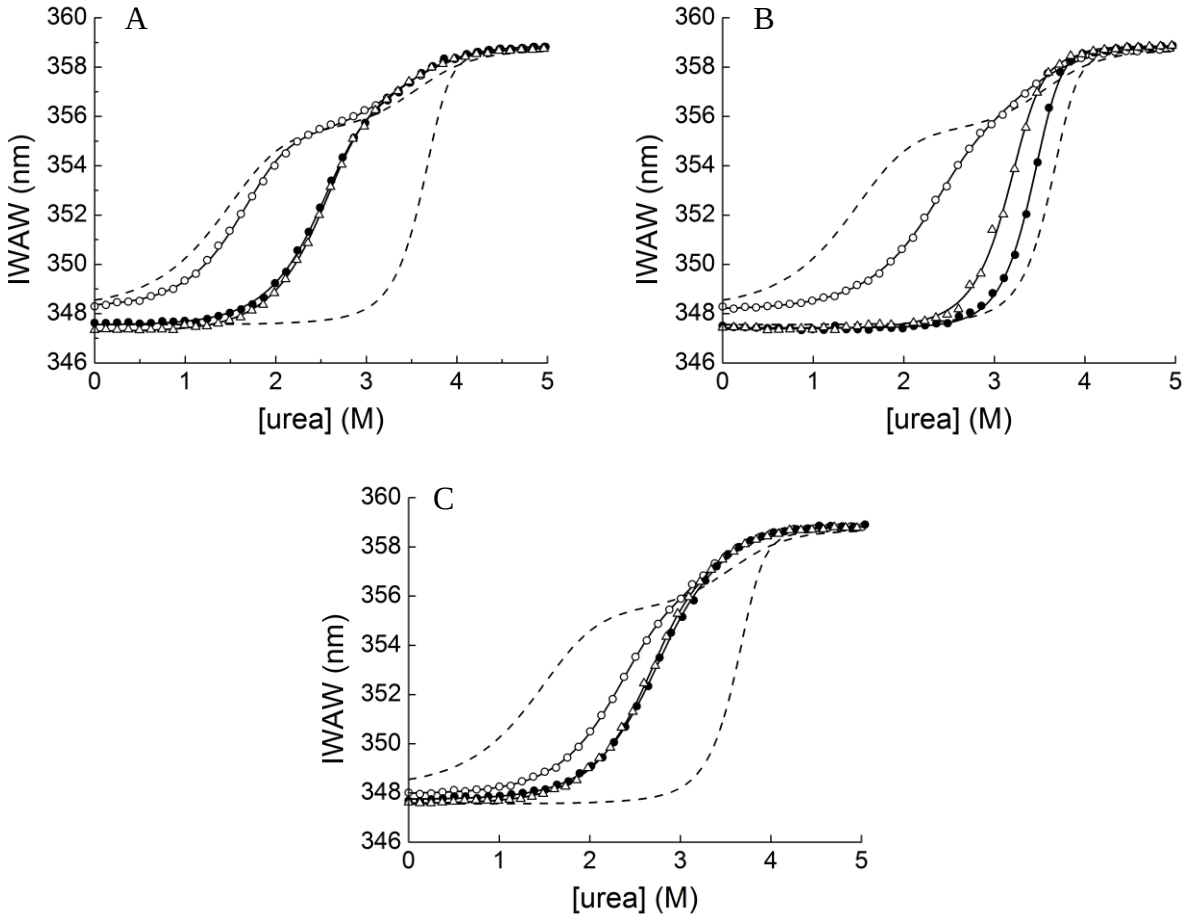


Figure 4.2 Effect of a metabolite on the equilibrium unfolding of dPGM in the presence of salt.

Equilibrium unfolding of dPGM was performed in the presence and absence of 1.0 mM citrate (△) or 1.0 mM ATP (●) under three salt conditions: (A) 5.0 mM $MgCl_2$, (B) 0.10 M NaCl, (C) 5.0 mM $MgCl_2$ and 0.10 M NaCl. Under each salt condition data were fit globally along with the equilibrium unfolding of dPGM without ligand using eq 2.4 (solid lines), and the results are listed in Table 4.2. Fitting results for the equilibrium unfolding of dPGM without added salt, in the presence and absence of 1.0 mM ATP, are shown (dashed lines) for reference.

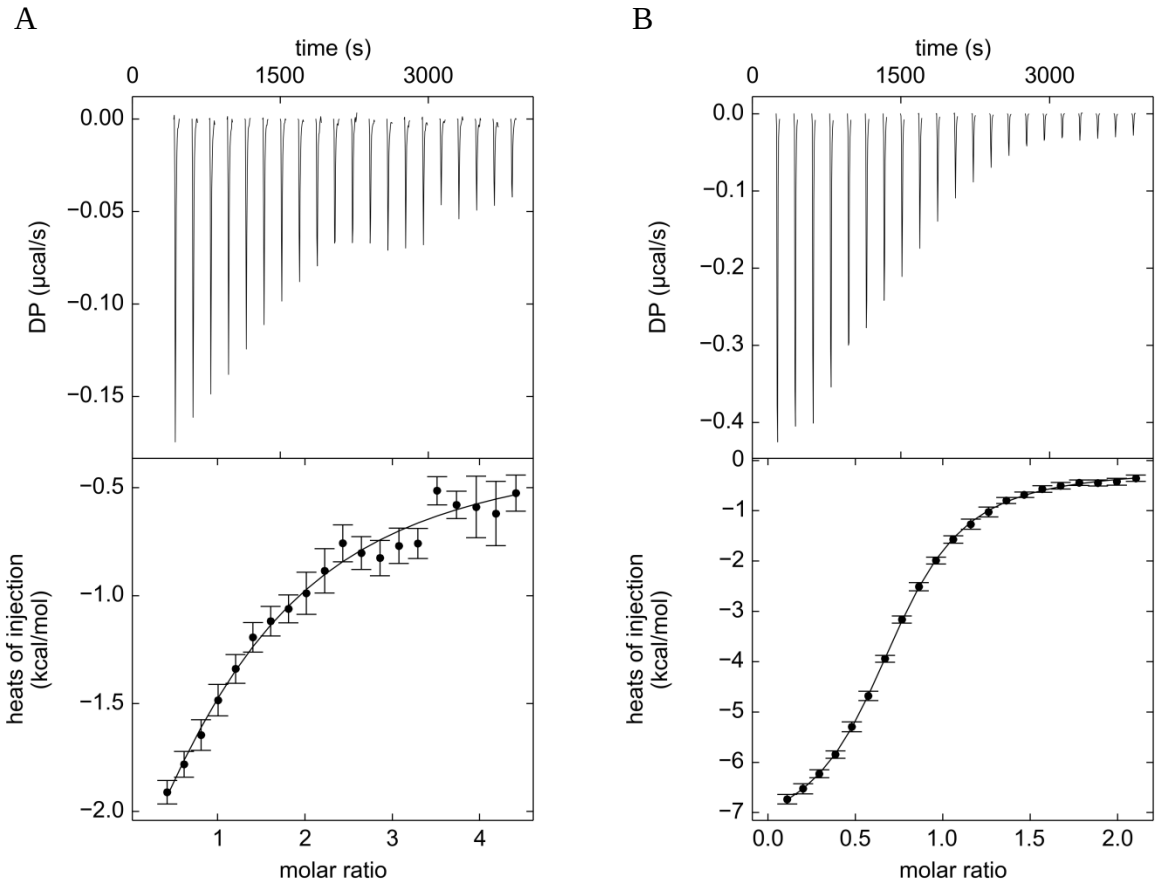


Figure 4.3 The effect of salt on ATP binding to dPGM

The heat evolution upon ATP binding to dPGM was monitored by isothermal titration calorimetry in the presence of 5.0 mM $MgCl_2$ (A) and 0.10 M NaCl (B). ATP was titrated in to the ITC cell containing dPGM in 21 injections and the reconstructed thermograms and isotherms shown. The isotherms were fit with a 1:1 binding model (─).

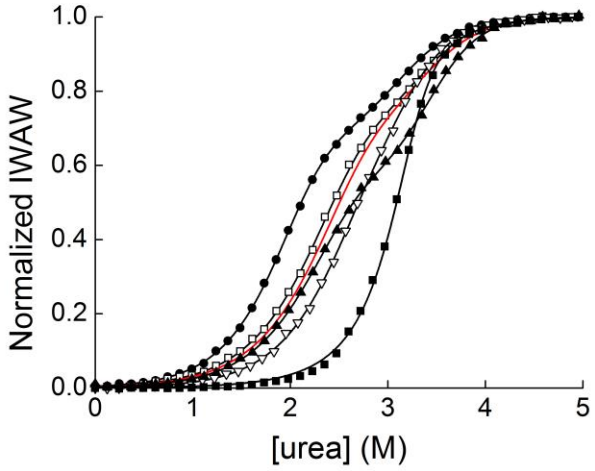


Figure 4.4 Effect of anions on the equilibrium unfolding of dPGM at an ionic strength of 0.10 M.

Equilibrium unfolding of dPGM was monitored in the presence of 33.3 mM $MgCl_2$ (●), 0.10 M KCl (□), 0.10 M NaF (▲), 0.10 M NaBr (▽), and 33.3 mM Na_2SO_4 (■). Data were fit with eq 2.4 (─) and normalized using eq 2.5 for proper comparison. For a three-state fitting of the equilibrium unfolding of dPGM in the presence of Na_2SO_4, the data were fit globally along with equilibrium unfolding of dPGM without added salt (Chapter 2). The fitting results for the equilibrium unfolding of dPGM in the presence of 0.10 M NaCl (Figure 4.2B) is shown for reference (─). Fitting results are listed in Table 4.3.

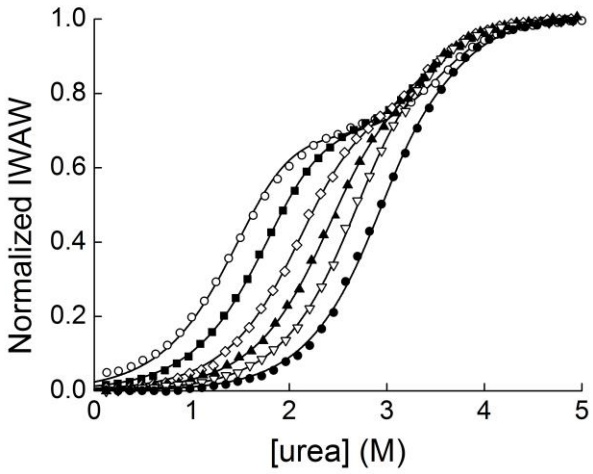


Figure 4.5 Effect of NaCl on the equilibrium unfolding of dPGM.

Equilibrium unfolding of dPGM was monitored with NaCl at 0 M (○), 0.02 M (■), 0.05 M (◇), 0.10 M (▲), 0.20 M (▽) and 0.50 M (●). Data were fit globally with eq 2.4 (─), and the fitting results are listed in Table 4.4. For proper comparison, IWAWs were normalized with eq 2.5.

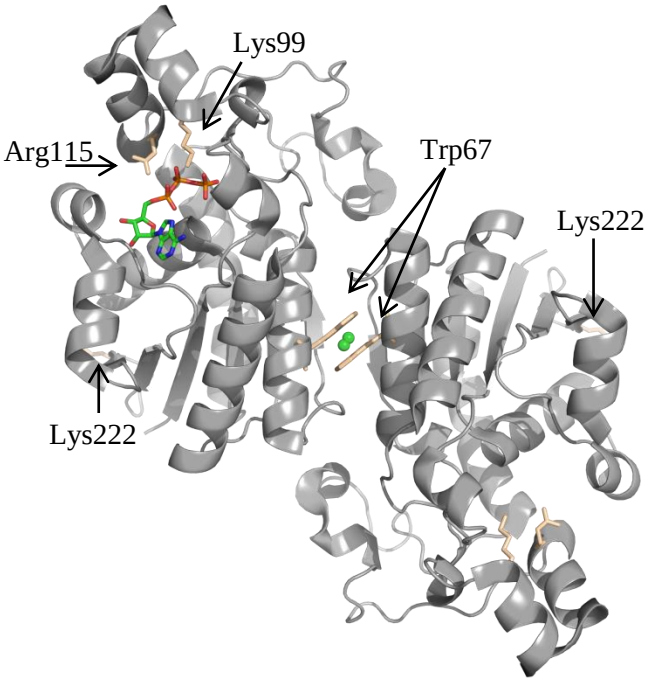


Figure 4.6 Structure of dPGM showing chloride binding and sites of mutation

The crystallographic structure of dPGM (PDB entry 1E59) is shown with the side chains that were probed by for mutation displayed (tan) and labeled. A chloride ion (green) is buried in the dimer interface and hydrogen bonds with Trp67. Lys99 and Arg115 form interactions with ATP, which was docked in the dPGM active site (Chapter 2).

.

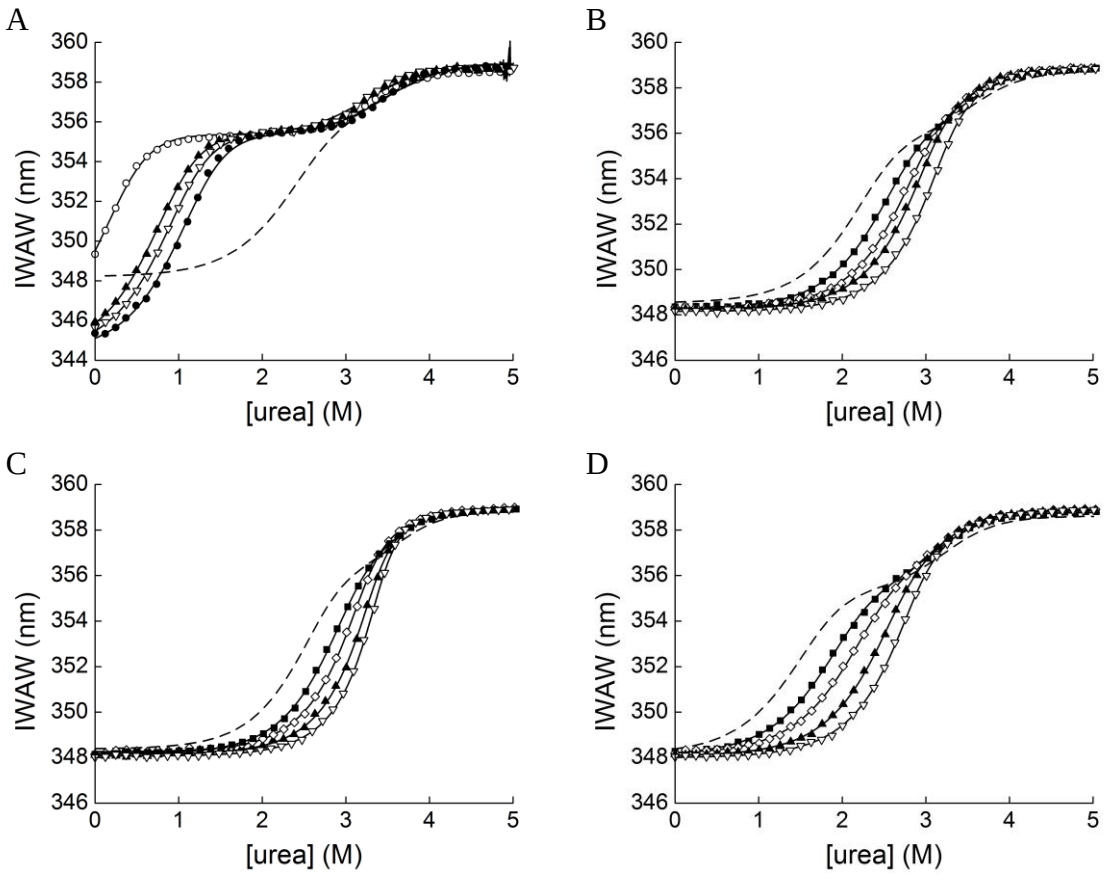


Figure 4.7 Effect of NaCl of the equilibrium unfolding of dPGM variants.

(A) Equilibrium unfolding of W67A dPGM was monitored without added salt (○) and with 0.10 M NaCl (▲), 0.20 M NaCl (▽), and 0.50 M NaCl (●). The equilibrium unfolding of K99A dPGM (B), R115A dPGM (C), and K222A dPGM (D) was monitored in presence of NaCl at 0.020 M (■), 0.050 M (◇), 0.10 M (▲), and 0.20 M (▽). The data for each dPGM variant were fit globally (solid lines) with eq 2.4 and the results are listed in Table 4.5. The fitting results for the equilibrium unfolding of wild-type dPGM in the presence of 0.10 M NaCl are shown (dashed line) for comparison.

CHAPTER 5. IMPLICATIONS OF THE INVESTIGATION

5.1 Cofactor-dependant phosphoglycerate mutases

Our investigation with *E. coli* dPGM has revealed that the protein can populate a partially unfolded intermediate, but many metabolites and anions found within the cytosol suppress partial unfolding. Because the intermediate is less accessible in a physiological milieu, dPGM seems to have evolved to avoid populating the intermediate. If the intermediate is undesirable, than it is intriguing that dPGM evolved with such a stable core. The stable core of dPGM, which remains folded in the intermediate state, may provide an advantage during protein folding. We observed that a majority of the dimer interface is formed in the intermediate state. If dimerization and folding of the loop region is rate-limiting, the stability of the dPGM core may prevent unfolding from occurring before dimerization. Prior to dimerization, dPGM is readily digested by protease, as the monomeric intermediate was digested when the equilibrium unfolding of dPGM was monitored by pulse proteolysis (Chang et al., 2012). Clearly, the fully folded is required to ensure the structural integrity of dPGM. Future investigations on the effect of ligand binding and mutation on the kinetics of dPGM would be essential to understand the rate limiting step in folding and how the process of folding is influenced by the cooperativity and allostery observed with dPGM.

The secondary structure and active site of dPGM is well conserved across *E. coli*, *S. cerevisiae*, *S. pombe,* and human brain dPGM, with the key differences being localized to subunit interfaces (Bond et al., 2001). The residues we chose for mutation are strictly conserved amongst the four species. Active site residues Lys99, Arg115, and Arg116 (Jacqueline Nairn et al., 2000; Rigden et al., 1999b) are involved in substrate binding and Arg9 is implicated in catalysis (Bond et al., 2002; Jacqueline Nairn et al., 2000). The native function of these residues is to interact with phosphate and carboxylate groups of 2,3-BPG, 2-PG, and 3-PG. The conservation between the active sites of all the homologs suggests that all four dPGMs could bind ATP, and other anionic metabolites, through the same residues. In addition to *E. coli* dPGM, ATP has been shown to bind to *S. cerevisiae* dPGM in cell lysate (Tran et al., 2014). Several polyanionic inhibitors were demonstrated to compete with 2,3-BPG binding to both *S. pombe* and *S. cerevisiae* dPGM and were shown to occupy the active site cleft of *S. cerevisiae* dPGM (Rigden, Walter, Phillips, & Fothergill-Gilmore, 1999a); the inhibitors interacted with residues Arg115 and Arg116 (*E. coli* numbering) of *S. cerevisiae*. The crystallographic structure of human dPGM has been solved with citrate occupying the active site (Yanli Wang et al., 2005). The binding mode of citrate is similar to that of the tetravanadate and ATP (Figure 2.6) with citrate also forming hydrogen bonds with Lys99 and Arg115 (*E. coli* numbering) (Yanli Wang et al., 2005). Therefore, ATP, other anionic metabolites, and anions, can likely bind within the conserved active site of homologous dPGMs as we observed with *E. coli* dPGM.

Homologous dPGMs may show activity for some of the anionic metabolites under the right conditions. For example, human dPGM has been shown to convert PEP to pyruvate in rapidly proliferating cells (Vander Heiden et al., 2010). Post-translational modifications are also used to regulate dPGM activity and may alter interactions with metabolites. Phosphorylation of Tyr26 (Tyr25 by *E. coli* numbering) in the loop region of human brain dPGM increases the stability and activity of dPGM. Our investigation shows that the phosphorylation event is likely to only form new interactions in the loop region and should increase $\Delta G_{I-N}°$. If a dPGM does display low activity for a metabolite, a post-translational modification could stabilize the conformations with greatest activity for the metabolite. This event would increase the apparent activity of dPGM for the metabolite. A simple function of some of the metabolites could be negative feedback by competitive inhibition to halt glycolysis. Both ATP and PEP can bind in the dPGM active site and are produced during glycolysis, downstream of dPGM. If the metabolite concentrations are high, additional glycolysis is not an immediate concern for the cell, and the metabolites may out-compete 3-phosphoghlycerate binding and inactivate dPGM, slowing glycolysis. Additional studies with homologous dPGMs could shed light on this topic as well as illuminating whether the energetic linkage between ligand binding and oligomerization is conserved by homologous dPGMs.

5.2 dPGM as a model system for folding studies

E. coli dPGM presents a wonderful model system to investigate the folding of oligomeric protein. Current model proteins are highly biased in favor of small monomeric proteins, which fold well and tend to show no folding intermediates (Braselmann et al., 2013).

Dimeric proteins are the most abundant form in the E coli proteome (Braselmann et al., 2013) and many dimeric proteins have been found to populate intermediates (Rumfeldt et al., 2008). The model proteins do not reflect the abundance of proteins in the proteome, many of which do not appear to fold reversibly. Our investigation on dPGM shows that non-substrate metabolite binding can be useful to resist unfolding or partial unfolding. However, dPGM has even more to offer as a model system for investigating the energetics and kinetics of protein folding. Because the intermediate state of dPGM is relatively stable and clearly distinguishable by optical spectroscopy, its properties can be studied more readily than most intermediates. dPGM is also fairly easy to purify and the protein folds well, for we never detected signs of aggregation. The sensitivity of dPGM to its chemical environment allows the experimentalist to modify $\Delta G_{I-N}°$ with ease. These factors make dPGM a great model system to probe the folding of oligomeric proteins. Using dPGM, for example, the folding pathway of an oligomeric protein could be examined to help clarify the debate over whether folding proceeds via multiple routes or a single pathway. Further, dPGM would allow the investigators to analyze the dependency of the folding pathway of the chemical environment or on the magnitude of $\Delta G_{I-N}°$.

5.3 Implications for oligomeric protein design

The partitioning of the cooperative subunits of dPGM and the energetic linkage between ligand binding and dimerization can be used as guidelines for oligomeric protein design. Two concepts seem important to the energetics of dPGM: partitioning between the stable and unstable cooperative subunits and partitioning of the active site and oligomer

interface between stable and unstable cooperative subunits. The least stable subunit of dPGM (the loop region) folds independently from the stable subunits and mutations in the least stable subunit, do not alter the stability of the stable subunits. Therefore, there is little interaction between the least stable subunit and the stable subunits. This partitioning allows the least stable subunit to change conformations and to function as a switch, while the stable subunits ensure that the protein remains relatively stable. Partitioning the active site and oligomeric interface between stable and unstable subunits ensures that a portion of the binding site will be pre-formed or stable, while another portion is dynamic or unfolded. The pre-formed, stable portion of the binding site promotes initial interactions with the ligand or binding partner. After association, a change in the dynamics or fold of the unstable portion of the binding site yields high affinity binding and transmits information about the binding event to the residues that are coupled with the unstable region. For dPGM, both the active site and dimer interface are partitioned between the stable core and the unstable loop region.

To mimic the energetics of dPGM, a protein could be designed with a stable core and an unstable binding region. Novel activities and active sites have been designed in a protein; thirty two *de novo* Retro-adolases were designed with varying protein folds (Jiang et al., 2008). That is, different structures gave rise to the same activity. Then, simple proteins could be designed with an active site and an oligomer interface with no regard for partitioning. The topology of the designed proteins could be analyzed and then the proteins can be rethreaded to connect the secondary structural units with a novel topology (Agah, Poulos, Yu, Kucharska, & Faham, 2016). The new connections should partition

the proteins into two portions with minimal interactions between each. One region, the least stable, will include most of the substrate binding residues and the least stable portion of the oligomer interface. Then, folding of the stable subunit will prime the proteins for oligomerization and ligand binding and either event should promote the other. Such a design feat is currently incredibly complicated. However, computational tools are rapidly developing to design well folded *de novo* protein (Mitra, Shultis, & Zhang, 2013). By analyzing the energetics of naturally occurring proteins, such as dPGM, we will hopefully be able to engineer *de novo* oligomeric protein with functions that are sensitive to signaling events.

LIST OF REFERENCES

LIST OF REFERENCES

Adhikari, J., & Fitzgerald, M. C. (2014). SILAC-pulse proteolysis: A mass spectrometry-based method for discovery and cross-validation in proteome-wide studies of ligand binding. *J Am Soc Mass Spec, 25*, 2073-2083.

Agah, S., Poulos, S., Yu, A., Kucharska, I., & Faham, S. (2016). Protein rethreading: A novel approach to protein design. *Sci Rep, 6*, 26847.

Ahmad, B., Muteeb, G., Alam, P., Varshney, A., Zaidi, N., Ishtikhar, M., Badr, G., Mahmoud, M. H., & Khan, R. H. (2015). Thermal induced unfolding of human serum albumin isomers: Assigning residual α helices to domain II. *Int J Biol Macromol, 75*, 447-452.

Alatossava, T., Jütte, H., Kuhn, A., & Kellenberger, E. (1985). Manipulation of intracellular magnesium content in polymyxin B nonapeptide-sensitized Escherichia coli by ionophore A23187. *J Bacteriol, 162*, 413-419.

Anfinsen, C. B. (1973). Principles that govern the folding of protein chains. *Science, 191*, 223-230.

Anfinsen, C. B., Haber, E., Sela, M., & White, F. (1961). The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. *Proc Natl Acad Sci, 47*, 1309-1314.

Bai, Y., Sosnick, T. R., Mayne, L., & Englander, S. W. (1995). Protein folding intermediates: Native-state hydrogen exchange. *Science, 269*, 192-197.

Banci, L., Bertini, I., Cantini, F., Inagaki, S., Migliardi, M., & Rosato, A. (2010). The binding mode of ATP revealed by the solution structure of the N-domain of human ATP7A. *J Biol Chem, 285*, 2537-2544.

Beadle, B. M., & Shoichet, B. K. (2002). Structural Bases of Stability–function Tradeoffs in Enzymes. *J Mol Biol, 321*, 285-296.

Bennett, B., Kimball, E., Gao, M., Osterhout, R., Van Dien, S., & Rabinowitz, J. (2009). Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nat Chem Biol, 5*, 593-599.

Berger, C., Weber-Bornhauser, S., Eggenberger, J., Hanes, J., Plückthun, A., & Bosshard, H. R. (1999). Antigen recognition by conformational selection. *FEBS Lett, 450*, 149-153.

Bond, C., White, M., & Hunter, W. (2001). High resolution structure of the phosphohistidine-activated form of *Escherichia coli* cofactor-dependent phosphoglycerate mutase. *J Biol Chem, 276*, 3247-3253.

---. (2002). Mechanistic implications for *Escherichia coli* cofactor-dependent phosphoglycerate mutase based on the high-resolution crystal structure of a vanadate complex. *J Mol Biol, 316*, 1071-1081.

Bradley, C. M., & Barrick, D. (2006). The notch ankyrin domain folds via a discrete, centralized pathway. *Structure, 14*, 1303-1312.

Braselmann, E., Chaney, J., & Clark, P. (2013). Folding the proteome. *Trends Biochem Sci, 38*, 337-344.

Bublitz, M., Morth, J. P., & Nissen, P. (2011). P-type ATPases at a glance. *J Cell Sci, 124*, 2515-2519.

Campos, L. A., Bueno, M., Lopez-Llano, J., Jiménez, M. Á., & Sancho, J. (2004). Structure of Stable Protein Folding Intermediates by Equilibrium ϕ-Analysis: The Apoflavodoxin Thermal Intermediate. *J Mol Biol, 344*, 239-255.

Carlson, H. A., & McCammon, J. A. (2000). Accommodating Protein Flexibility in Computational Drug Design. *Mol Pharmacol, 57*, 213-218.

Chamberlain, A. K., Handel, T. M., & Marqusee, S. (1996). Detection of rare partially folded molecules in equilibrium with the native conformation of RNase H. *Nat Struct Biol, 3*, 782-787.

Chang, Y., Schlebach, J., VerHeul, R., & Park, C. (2012). Simplified proteomics approach to discover protein-ligand interactions. *Protein Sci, 21*, 1280-1287.

Chennubhotla, C., Yang, Z., & Bahar, I. (2008). Coupling between global dynamics and signal transduction pathways: a mechanism of allostery for chaperonin GroEL. *Mol Biosyst, 4*, 287-292.

Clarke, J. B., Birch, M., & Britton, H. G. (1974). The equilibrium constant of the phosphoglyceromutase reaction. *Biochem J, 139*, 491-497.

Coyle, Scott M., Flores, J., & Lim, Wendell A. (2013). Exploitation of Latent Allostery Enables the Evolution of New Modes of MAP Kinase Regulation. *Cell, 154*, 875-887.

Daggett, V., & Fersht, A. R. (2003). Is there a unifying mechanism for protein folding? *Trends Biochem Sci, 28*, 18-25.

Danielsson, J., Mu, X., Lang, L., Wang, H., Binolfi, A., Theillet, F.-X., Bekei, B., Logan, D. T., Selenko, P., & Wennerström, H. (2015). Thermodynamics of protein destabilization in live cells. *Proc Natl Acad Sci, 112*, 12402-12407.

Davidson, A. L., Dassa, E., Orelle, C., & Chen, J. (2008). Structure, function, and evolution of bacterial ATP-binding cassette systems. *Microbiol Mol Biol Rev 72*, 317.

de los Rios, M. A., & Plaxco, K. W. (2005). Apparent Debye−Huckel Electrostatic Effects in the Folding of a Simple, Single Domain Protein. *Biochemistry, 44*, 1243-1250.

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

Dill, K., Ozkan, S., Shell, M., & Weikl, T. (2008). The protein folding problem. *Annu Rev Biophys, 37*, 289-316.

Dobson, C. M. (2003). Protein folding and misfolding. *Nature, 426*, 884-890.

Duff, M. R., Grubbs, J., & Howell, E. E. (2011). Isothermal Titration Calorimetry for Measuring Macromolecule-Ligand Affinity. *J Vis Exp*, 2796.

Englander, S. W. (2000). Protein folding intermediates and pathways studied by hydrogen exchange. *Annu Rev Biophys Biomol Struct, 29*, 213-238.

Englander, S. W., Mayne, L., & Krishna, M. (2007). Protein folding and misfolding: mechanism and principles. *Q Rev Biophys, 40*, 287-326.

Epstein, C. J., & Anfinsen, C. B. (1962). Reversible reduction and reoxidation of poly-DL-alanyl trypsin. *J Biol Chem, 237*, 3464-3467.

Fersht, A. R., Matouschek, A., & Serrano, L. (1992). The folding of an enzyme. *J Mol Biol, 224*, 771-782.

Fraser, H. I., Kvaratskhelia, M., & White, M. F. (1999). The two analogous phosphoglycerate mutases of *Escherichia coli*. *FEBS Lett, 455*, 344-348.

Frauenfelder, H., Sligar, S. G., & Wolynes, P. G. (1991). The energy landscapes and motions of proteins. *Science, 254*, 1598-1603.

Freire, E. (1999). The propagation of binding interactions to remote sites in proteins: Analysis of the binding of the monoclonal antibody D1.3 to lysozyme. *Proc Natl Acad Sci, 96*, 10118-10122.

Freire, E., Murphy, K., Sanchez-Ruiz, J., Galisteo, M., & Privalov, P. (1992). The molecular basis of cooperativity in protein folding. Thermodynamic dissection of interdomain interactions in phosphoglycerate kinase. *Biochemistry, 31*, 250-256.

Freire, E., & Murphy, K. P. (1991). Molecular basis of co-operativity in protein folding. *J Mol Biol, 222*, 687-698.

Frembgen-Kesner, T., & Elcock, A. H. (2013). Computer Simulations of the Bacterial Cytoplasm. *Biophys Rev, 5*, 109-119.

Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., Mainz, D. T., Repasky, M. P., Knoll, E. H., Shelley, M., Perry, J. K., Shaw, D. E., Francis, P., & Shenkin, P. S. (2004). Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J Med Chem, 47*, 1739-1749.

Fuxreiter, M., Simon, I., Friedrich, P., & Tompa, P. (2004). Preformed Structural Elements Feature in Partner Recognition by Intrinsically Unstructured Proteins. *J Mol Biol, 338*, 1015-1026.

Gloss, L. M., & Matthews, C. R. (1997). Urea and thermal equilibrium denaturation studies on the dimerization domain of *Escherichia coli* Trp repressor. *Biochemistry, 36*, 5612-5623.

Goodey, N. M., & Benkovic, S. J. (2008). Allosteric regulation and catalysis emerge via a common route. *Nat Chem Biol, 4*, 474-482.

Greene, R. F., & Pace, C. N. (1974). Urea and guanidine-hydrochloride denaturation of ribonuclease, lysozyme, α-chymotrypsin, and β-lactoglobulin. *J Biol Chem, 249*, 5388-5393.

Grisolia, S., & Cleland, W. W. (1968). Influence of salt, substrate, and cofactor concentrations on the kinetic and mechanistic behavior of phosphoglycerate mutase. *Biochemistry, 7*, 1115-1121.

Grisolia, S., & Joyce, B. K. (1959). Distribution of Two Types of Phosphoglyceric Acid Mutase, Diphosphoglycerate Mutase, and d-2,3-Diphosphoglyceric Acid. *J Biol Chem, 234*, 1335-1337.

Gunasekaran, K., Ma, B., & Nussinov, R. (2004). Is allostery an intrinsic property of all dynamic proteins? *Proteins: Struct, Funct, Bioinf, 57*, 433-443.

Hammes, G. G. (2002). Multiple Conformational Changes in Enzyme Catalysis. *Biochemistry, 41*, 8221-8228.

Harford, S., & Weitzman, P. (1975). Evidence of isosteric and allosteric nucleotide inhibition of citrate synthease from multiple-inhibition studies. *Biochem J, 151*, 455-458.

Henzler-Wildman, K., & Kern, D. (2007). Dynamic personalities of proteins. *Nature, 450*, 964-972.

Hermann, R., Jaenicke, R., & Price, N. C. (1985). Evidence for active intermediates during the reconstitution of yeast phosphoglycerate mutase. *Biochemistry, 24*, 1817-1821.

Hilser, V. J., Dowdy, D., Oas, T. G., & Freire, E. (1998). The structural distribution of cooperative interactions in proteins: Analysis of the native state ensemble. *Proc Natl Acad Sci, 95*, 9903-9908.

Hilser, V. J., Garcia-Moreno, E. B., Oas, T. G., Kapp, G., & Whitten, S. T. (2006). A statistical thermodynamic model of the protein ensemble. *Chem Rev, 106*, 1545-1558.

Hilser, V. J., & Thompson, E. B. (2007). Intrinsic disorder as a mechanism to optimize allosteric coupling in proteins. *Proc Natl Acad Sci, 104*, 8311-8315.

Hilser, V. J., Wrabl, J. O., & Motlagh, H. N. (2012). Structural and Energetic Basis of Allostery. *Annu Rev Biophys, 41*, 585-609.

Horwich, A. (2002). Protein aggregation in disease: a role for folding intermediates forming specific multimeric interactions. *J Clin Investig, 110*, 1221-1232.

Hu, W., Kan, Z.-Y., Mayne, L., & Englander, S. W. (2016). Cytochrome c folds through foldon-dependent native-like intermediates in an ordered pathway. *Proc Natl Acad Sci, 113*, 3809-3814.

Hu, W., Walters, B., Kan, Z.-Y., Mayne, L., Rosen, L., Marqusee, S., & Englander, S. W. (2013). Stepwise protein folding at near amino acid resolution by hydrogen exchange and mass spectrometry. *Proc Natl Acad Sci U S A, 110*, 7684-7689.

Hubbard, S. J. (1998). The structural aspects of limited proteolysis of native proteins. *Biochim Biophys Acta, 1382*, 191-206.

Hynes, R. O. (2002). Integrins: Bidirectional, Allosteric Signaling Machines. *Cell, 110*, 673-687.

Ishihama, Y., Schmidt, T., Rappsilber, J., Mann, M., Hartl, F. U., Kerner, M. J., & Frishman, D. (2008). Protein abundance profiling of the *Escherichia coli* cytosol. *BMC Genomics, 9*, 102.

Jiang, L., Althoff, E. A., Clemente, F. R., Doyle, L., Röthlisberger, D., Zanghellini, A., Gallaher, J. L., Betker, J. L., Tanaka, F., Barbas, C. F., Hilvert, D., Houk, K. N., Stoddard, B. L., & Baker, D. (2008). De Novo Computational Design of Retro-Aldol Enzymes. *Science, 319*, 1387-1391.

Johnson, L. N., Lowe, E. D., Noble, M. E., & Owen, D. J. (1998). The structural basis for substrate recognition and control by protein kinases. *FEBS Lett, 430*, 1-11.

Jukka, K. H., & Harold, L. D. (1988). Comparative assessment of inorganic pyrophosphate and pyrophosphatase levels of Escherichia coli, Clostridium pasteurianum , and Clostridium thermoaceticum. *FEMS Microbiol Lett, 52*.

Kalid, O., Mense, M., Fischman, S., Shitrit, A., Bihler, H., Ben-Zeev, E., Schutz, N., Pedemonte, N., Thomas, P. J., Bridges, R. J., Wetmore, D. R., Marantz, Y., & Senderowitz, H. (2010). Small molecule correctors of F508del-CFTR discovered by structure-based virtual screening. *J Comput Aided Mol Des, 24*, 971-991.

Karthikeyan, S., Zhou, Q., Mseeh, F., Grishin, N. V., Osterman, A. L., & Zhang, H. (2003). Crystal structure of human riboflavin rinase Reveals a β barrel fold and a novel active site arch. *Structure, 11*, 265-273.

Keller, S., Vargas, C., Zhao, H., Piszczek, G., Brautigam, C. A., & Schuck, P. (2012). High-precision isothermal titration calorimetry with automated peak-shape analysis. *Anal Chem, 84*, 5066-5073.

Kenakin, T. (2002). Efficacy at g-protein-coupled receptors. *Nat Rev Drug Discov, 1*, 103-110.

Kim, M. S., Song, J., & Park, C. (2009). Determining protein stability in cell lysates by pulse proteolysis and Western blotting. *Protein Sci, 18*, 1051-1059.

Knott, M., & Best, R. B. (2012). A Preformed Binding Interface in the Unbound Ensemble of an Intrinsically Disordered Protein: Evidence from Molecular Simulations. *PLoS Comput Biol, 8*, e1002605.

Kowalski, W., Nocon, D., Gamian, A., Kołodziej, J., & Rakus, D. (2012). Association of C-terminal region of phosphoglycerate mutase with glycolytic complex regulates energy production in cancer cells. *J Cell Physiol, 227*, 2613-2621.

Krissinel, E., & Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *J Mol Biol, 372*, 774-797.

Kubitschek, H. E., & Friske, J. A. (1986). Determination of bacterial cell volume with the Coulter Counter. *J Bacteriol, 168*, 1466-1467.

Kunin, C. M., Hua, T. H., Van Arsdale White, L., & Villarejo, M. (1992). Growth of Escherichia coli in human urine: role of salt tolerance and accumulation of glycine betaine. *J Infect Dis, 166*, 1311-1315.

Kuriyan, J., & Eisenberg, D. (2007). The origin of protein interactions and allostery in colocalization. *Nature, 450*, 983-990.

Laforet, M. T., Butterfield, J. B., & Alpers, J. B. (1974). The biosynthesis of 2,3-diphosphoglycerate by monophosphoglycerate mutase from muscle and erythrocytes. *Arch Biochem Biophys, 165*, 179-187.

Lamazares, E., Clemente, I., Bueno, M., Velázquez-Campoy, A., & Sancho, J. (2015). Rational stabilization of complex proteins: a divide and combine approach. *Sci Rep, 5*, 9129.

Law, A. B., Fuentes, E. J., & Lee, A. L. (2009). Conservation of Side-Chain Dynamics Within a Protein Family. *J Am Chem Soc, 131*, 6322-6323.

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

Leyrat, C., Jensen, M. R., Ribeiro, E. A., Gérard, F. C. A., Ruigrok, R. W. H., Blackledge, M., & Jamin, M. (2011). The N(0)-binding region of the vesicular stomatitis virus phosphoprotein is globally disordered but contains transient α-helices. *Protein Sci, 20*, 542-556.

Li, X., Gianoulis, T. A., Yip, K. Y., Gerstein, M., & Snyder, M. (2010). Extensive *in vivo* metabolite-protein interactions revealed by large-scale systematic analyses. *Cell, 143*, 639-650.

Lieberman, R. L., Wustman, B. A., Huertas, P., Powe, A. C., Pine, C. W., Khanna, R., Schlossmacher, M. G., Ringe, D., & Petsko, G. A. (2007). Structure of acid [beta]-glucosidase with pharmacological chaperone provides insight into Gaucher disease. *Nat Chem Biol, 3*, 101-107.

Link, H., Kochanowski, K., & Sauer, U. (2013). Systematic identification of allosteric protein-metabolite interactions that control enzyme activity in vivo. *Nat Biotech, 31*, 357-361.

Liu, P.-F., Kihara, D., & Park, C. (2011). Energetics-based discovery of protein-ligand interactions on a proteomic scale. *J Mol Biol, 408*, 147-162.

Liu, T., Whitten, S. T., & Hilser, V. J. (2007). Functional residues serve a dominant role in mediating the cooperativity of the protein ensemble. *Proc Natl Acad Sci, 104*, 4347-4352.

Lockless, S. W., & Ranganathan, R. (1999). Evolutionarily Conserved Pathways of Energetic Connectivity in Protein Families. *Science, 286*, 295-299.

Lukacs, G. L., & Verkman, A. S. (2012). CFTR: folding, misfolding and correcting the ΔF508 conformational defect. *Trends Mol Med, 18*, 81-91.

Luque, I., & Freire, E. (2000). Structural stability of binding sites: Consequences for binding affinity and allosteric effects. *Proteins: Struct, Funct, Bioinf, 41*, 63-71.

Luque, I., Leavitt, S. A., & Freire, E. (2002). The linkage between protein folding and functional cooperativity: Two sides of the same coin? *Annu Rev Biophys Biomol Struct, 31*, 235-256.

Mallam, A., & Jackson, S. (2005). Folding studies on a knotted protein. *J Mol Biol, 346*, 1409-1421.

Manhart, M., & Morozov, A. V. (2015). Protein folding and binding can emerge as evolutionary spandrels through structural coupling. *Proc Natl Acad Sci U S A, 112*, 1797-1802.

Matouschek, A., Kellis, J. T., Serrano, L., Bycroft, M., & Fersht, A. R. (1990). Transient folding intermediates characterized by protein engineering. *Nature, 346*, 440-445.

Matouschek, A., Kellis, J. T., Serrano, L., & Fersht, A. R. (1989). Mapping the transition state and pathway of protein folding by protein engineering. *Nature, 340*, 122-126.

McGuffee, S. R., & Elcock, A. H. (2010). Diffusion, crowding & protein stability in a dynamic molecular model of the bacterial cytoplasm. *PLoS Comput Biol, 6*, 1000694.

Meiering, E. M., Serrano, L., & Fersht, A. R. (1992). Effect of active site residues in barnase on activity and stability. *J Mol Biol, 225*, 585-589.

Mello, C. C., & Barrick, D. (2004). An experimentally determined protein folding energy landscape. *Proc Natl Acad Sci U S A, 101*, 14102-14107.

Miklos, A. C., Sarkar, M., Wang, Y., & Pielak, G. J. (2011). Protein crowding tunes protein stability. *J Am Chem Soc, 133*, 7116-7120.

Mitra, P., Shultis, D., & Zhang, Y. (2013). EvoDesign: de novo protein design based on structural and evolutionary profiles. *Nucleic Acids Res, 41*, W273-W280.

Morra, G., Genoni, A., & Colombo, G. (2014). Mechanisms of Differential Allosteric Modulation in Homologous Proteins: Insights from the Analysis of Internal Dynamics and Energetics of PDZ Domains. *J Chem Theory Comput, 10*, 5677-5689.

Motlagh, H. N., Li, J., Thompson, E. B., & Hilser, Vincent J. (2012). Interplay between allostery and intrinsic disorder in an ensemble. *Biochem Soc Trans, 40*, 975-980.

Motlagh, H. N., Wrabl, J. O., Li, J., & Hilser, V. J. (2014). The ensemble nature of allostery. *Nature, 508*, 331-339.

Myers, J. K., Pace, C. N., & Scholtz, J. M. (1995). Denaturant *m* values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci, 4*, 2138-2148.

Nairn, J., Duncan, D., Price, N. E., Kelly, S. M., Fothergill-Gilmore, L. A., Uhrinova, S., Barlow, P. N., Rigden, D. J., & Price, N. C. (2000). Characterization of active-site mutants of Schizosaccharomyces pombe phosphoglycerate mutase. *Eur J Biochem, 267*, 7065-7074.

Nairn, J., Price, N. C., Fothergill-Gilmore, L. A., Walker, G. E., Fothergill, J. E., & Dunbar, B. (1994). The amino acid sequence of the small monomeric phosphoglycerate mutase from the fission yeast Schizosaccharomyces pombe. *Biochem J, 297*, 603-608.

Orsak, T., Smith, T. L., Eckert, D., Lindsley, J. E., Borges, C. R., & Rutter, J. (2011). Revealing the allosterome: Systematic identification of metabolite–protein interactions. *Biochemistry, 51*, 225-232.

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

Pace, C. N., & Barrett, A. J. (1984). Kinetics of tryptic hydrolysis of the arginine-valine bond in folded and unfolded ribonuclease T1. *Biochem J, 219*, 411-417.

Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., & Gray, T. (1995). How to measure and predict the molar absorption coefficient of a protein. *Protein Sci, 4*, 2411-2423.

Paduch, M., Koide, A., Uysal, S., Rizk, S. S., Koide, S., & Kossiakoff, A. A. (2013). Generating conformation-specific synthetic antibodies to trap proteins in selected functional states. *Methods, 60*, 3-14.

Park, C., & Marqusee, S. (2004a). Analysis of the stability of multimeric proteins by effective ΔG and effective *m*-values. *Protein Sci, 13*, 2553-2558.

---. (2004b). Probing the High Energy States in Proteins by Proteolysis. *J Mol Biol, 343*, 1467-1476.

---. (2005). Pulse proteolysis: a simple method for quantitative determination of protein stability and ligand binding. *Nat Meth, 2*, 207-212.

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

Rigden, D. J. (2008). The histidine phosphatase superfamily: structure and function. *Biochem J, 409*, 333-348.

Rigden, D. J., Walter, R. A., Phillips, S. E. V., & Fothergill-Gilmore, L. A. (1999a). Polyanionic Inhibitors of Phosphoglycerate Mutase: Combined Structural and Biochemical Analysis. *J Mol Biol, 289*, 691-699.

---. (1999b). Sulphate ions observed in the 2.12 Å structure of a new crystal form of S. cerevisiae phosphoglycerate mutase provide insights into understanding the catalytic mechanism1. *J Mol Biol, 286*, 1507-1517.

Ringe, D., & Petsko, G. A. (2009). Q&A: What are pharmacological chaperones and why are they interesting? *J Biol, 8*, 1-4.

Robert, R., Carlile, G. W., Liao, J., Balghi, H., Lesimple, P., Liu, N., Kus, B., Rotin, D., Wilke, M., de Jonge, H. R., Scholte, B. J., Thomas, D. Y., & Hanrahan, J. W. (2010). Correction of the Delta phe508 cystic fibrosis transmembrane conductance regulator trafficking defect by the bioavailable compound glafenine. *Mol Pharmacol, 77*, 922-930.

Rocheville, M., Lange, D. C., Kumar, U., Patel, S. C., Patel, R. C., & Patel, Y. C. (2000). Receptors for Dopamine and Somatostatin: Formation of Hetero-Oligomers with Enhanced Functional Activity. *Science, 288*, 154-157.

Roelofs, K. G., Wang, J., Sintim, H. O., & Lee, V. T. (2011). Differential radial capillary action of ligand assay for high-throughput detection of protein-metabolite interactions. *Proc Natl Acad Sci, 108*, 15528-15533.

Rose, Z. B., & Dube, S. (1978). Phosphoglycerate mutase. Kinetics and effects of salts on the mutase and bisphosphoglycerate phosphatase activities of the enzyme from chicken breast muscle. *J Biol Chem, 253*, 8583-8592.

Rose, Z. B., & Kaklij, G. S. (1984). The effects of anions on phosphoglycerate mutase. *Biochem Biophys Res Commun, 121*, 834-841.

Rumfeldt, J. A., Galvagnion, C., Vassall, K. A., & Meiering, E. M. (2008). Conformational stability and folding mechanisms of dimeric proteins. *Prog Biophys Mol Biol, 98*, 61-84.

Sanchez-Ruiz, J. M. (2007). Ligand effects on protein thermodynamic stability. *Biophys Chem, 126*, 43-49.

Sancho, J., Bueno, M., Campos, L. A., Fernandez-Recio, J., Iran, M. P., Lopez, J., Machicado, C., Pedroso, I., & Toja, M. (2002). The Relevant Stability of Proteins with Equilibrium Intermediates. *Scientific World J, 2*.

Schellman, J. A. (1975). Macromolecular binding. *Biopolymers, 14*, 999-1018.

---. (1976). The effect of binding on the melting temperature of biopolymers. *Biopolymers, 15*, 999-1000.

Schneider, D., & Gourse, R. (2004). Relationship between growth rate and ATP concentration in *Escherichia coli*: a bioassay for available cellular ATP. *J Biol Chem, 279*, 8262-8268.

Schultz, S. G., Wilson, N. L., & Epstein, W. (1962). Cation transport in Escherichia coli. II. Intracellular chloride concentration. *J Gen Physiol, 46*, 159-166.

Shoichet, B. K., Baase, W. A., Kuroki, R., & Matthews, B. W. (1995). A relationship between protein stability and protein function. *Proc Natl Acad Sci, 92*.

Sieber, V., Pluckthun, A., & Schmid, F. X. (1998). Selecting proteins with improved stability by a phage-based method. *Nat Biotech, 16*, 955-960.

Smock, R. G., & Gierasch, L. M. (2009). Sending Signals Dynamically. *Science, 324*, 198-203.

Sridevi, K., Lakshmikanth, G. S., Krishnamoorthy, G., & Udgaonkar, J. B. (2004). Increasing stability reduces conformational heterogeneity in a protein folding intermediate ensemble. *J Mol Biol, 337*, 699-711.

Strickland, E. C., Geer, M. A., Tran, D. T., Adhikari, J., West, G. M., DeArmond, P. D., Xu, Y., & Fitzgerald, M. C. (2013). Thermodynamic analysis of protein-ligand binding interactions in complex biological mixtures using the stability of proteins from rates of oxidation. *Nat Protocols, 8*, 148-161.

Suel, G. M., Lockless, S. W., Wall, M. A., & Ranganathan, R. (2003). Evolutionarily conserved networks of residues mediate allosteric communication in proteins. *Nat Struct Mol Biol, 10*, 59-69.

Taylor, S. S., Buechler, J. A., & Yonemoto, W. (1990). cAMP-dependent protein kinase: framework for a diverse family of regulatory enzymes. *Annu Rev Biochem, 59*, 971-1005.

Thomas, P. J., Qu, B.-H., & Pedersen, P. L. (1995). Defective protein folding as a basis of human disease. *Trends Biochem Sci, 20*, 456-459.

Tran, D. T., Adhikari, J., & Fitzgerald, M. C. (2014). Stable Isotope Labeling with Amino Acids in Cell Culture (SILAC)-Based Strategy for Proteome-Wide Thermodynamic Analysis of Protein-Ligand Binding Interactions. *Molecular & Cellular Proteomics : MCP, 13*, 1800-1813.

Uhrínová, S., Uhrín, D., Nairn, J., Price, N. C., Fothergill-Gilmore, L. A., & Barlow, P. N. (2001a). Solution structure and dynamics of an open β-sheet, glycolytic enzyme, monomeric 23.7 kDa phosphoglycerate mutase from *Schizosaccharomyces pombe*. *J Mol Biol, 306*, 275-290.

---. (2001b). Solution structure and dynamics of an open β-sheet, glycolytic enzyme, monomeric 23.7 kDa phosphoglycerate mutase from Schizosaccharomyces pombe1. *J Mol Biol, 306*, 275-290.

Van Goor, F., Straley, K. S., Cao, D., Gonzalez, J., Hadida, S., Hazlewood, A., Joubran, J., Knapp, T., Makings, L. R., Miller, M., Neuberger, T., Olson, E., Panchenko, V., Rader, J., Singh, A., Stack, J. H., Tung, R., Grootenhuis, P. D., & Negulescu, P. (2006). Rescue of DeltaF508-CFTR trafficking and gating in human cystic fibrosis airway primary cultures by small molecules. *Am J Physiol Lung Cell Mol Physiol, 290*, 27.

Vander Heiden, M. G., Locasale, J. W., Swanson, K. D., Sharfi, H., Heffron, G. J., Amador-Noguez, D., Christofk, H. R., Wagner, G., Rabinowitz, J. D., Asara, J. M., & Cantley, L. C. (2010). Evidence for an Alternative Glycolytic Pathway in Rapidly Proliferating Cells. *Science, 329*, 1492-1499.

Waldron, T. T., & Murphy, K. P. (2003). Stabilization of proteins by ligand binding: Application to drug screening and determination of unfolding energetics. *Biochemistry, 42*, 5058-5064.

Walser, M. (1961). Ion association v. dissociation constants for complexes of vitrate with sodium, potassium, calcium, and magnesium ions. *J Phys Chem A, 65*, 159-161.

Walter, R. A., Nairn, J., Duncan, D., Price, N. C., Kelly, S. M., Rigden, D. J., & Fothergill-Gilmore, L. A. (1999). The role of the C-terminal region in phosphoglycerate mutase. *Biochem J, 337*, 89-95.

Wang, Y., & Shortle, D. (1997). Residual helical and turn structure in the denatured state of staphylococcal nuclease: analysis of peptide fragments. *Fold Des, 2*, 93-100.

Wang, Y., Wei, Z., Liu, L., Cheng, Z., Lin, Y., Ji, F., & Gong, W. (2005). Crystal structure of human B-type phosphoglycerate mutase bound with citrate. *Biochem Biophys Res Commun, 331*, 1207-1215.

Weber, G. (1972). Ligand binding and internal equilibiums in proteins. *Biochemistry, 11*, 864-878.

West, G. M., Tucker, C. L., Xu, T., Park, S. K., Han, X., Yates, J. R., & Fitzgerald, M. C. (2010). Quantitative proteomics approach for identifying protein–drug interactions in complex mixtures using protein stability measurements. *Proc Natl Acad Sci U S A, 107*, 9078-9082.

White, M. F., Fothergill-Gilmore, L. A., Kelly, S., & Price, N. C. (1993a). Dissociation of the tetrameric phosphoglycerate mutase from yeast by a mutation in the subunit contact region. *Biochem J, 295*, 743-748.

---. (1993b). Substitution of His-181 by alanine in yeast phosphoglycerate mutase leads to cofactor-induced dissociation of the tetrameric structure. *Biochem J, 291*, 479-483.

Wilson, J. E., & Chin, A. (1991). Chelation of divalent cations by ATP, studied by titration calorimetry. *Anal Biochem, 193*, 16-19.

Winn, S. I., Watson, H. C., Harkins, R. N., & Fothergill, L. A. (1981). Structure and Activity of Phosphoglycerate Mutase. *Philos Trans R Soc Lond B Biol Sci, 293*, 121-130.

Wrabl, J. O., Gu, J., Liu, T., Schrank, T. P., Whitten, S. T., & Hilser, V. J. (2011). The role of protein conformational fluctuations in allostery, function, and evolution. *Biophys Chem, 159*, 129-141.

Yu, B., Martins, I., Li, P., Amarasinghe, G., Umetani, J., Fernandez-Zapico, M., Billadeau, D., Machius, M., Tomchick, D., & Rosen, M. (2010). Structural and energetic mechanisms of cooperative autoinhibition and activation of Vav1. *Cell, 140*, 246-256.

Yuan, J., Doucette, C. D., Fowler, W. U., Feng, X.-J., Piazza, M., Rabitz, H. A., Wingreen, N. S., & Rabinowitz, J. D. (2009). Metabolomics-driven quantitative analysis of ammonia assimilation in E. coli. *Molecular Systems Biology, 5*, 302-302.

Zhao, H., Piszczek, G., & Schuck, P. (2015). SEDPHAT – A platform for global ITC analysis and global multi-method analysis of molecular interactions. *Methods, 76*, 137-148.

Zhong, S., Rousseau, D. L., & Yeh, S. R. (2004). Modulation of the folding energy landscape of cytochrome C with salt. *J Am Chem Soc, 126*, 13934-13935.

Zitzewitz, J. A., Ibarra-Molero, B., Fishel, D. R., Terry, K. L., & Robert Matthews, C. (2000). Preformed secondary structure drives the association reaction of GCN4-p1, a model coiled-coil system1. *J Mol Biol, 296*, 1105-1116.

VITA

VITA

Nathan William Gardner
Purdue University

Education

2011–2015 Purdue University West Lafayette, IN
PhD completed August 2016
Purdue University Interdisciplinary Life Science Program (PULSe)
Research advisor: Dr. Chiwook Park
Thesis title: Investigation of an energetic coupling between ligand binding and protein folding

2007–2011 Lawrence Technological University (LTU) Southfield, MI
B.S. in Physics
Thesis title: Employing a novel infrared camera to monitor heat flux in the NSTX diverter
Graduated as a member of the Honor Society and the Lambda Iota Tau honor society

Research Experience

2011–2016 Graduate Research Assistant Purdue University, West Lafayette, IN
Advisor: Dr. Chiwook Park, Medicinal Chemistry and Molecular Pharmacology
Energetic coupling in E. coli phosphoglycerate mutase
Rapid K_d determination of protein–drug interactions in cell lysate by Pulse and Western

Summer 2010 Research Intern Princeton Plasma Physics Laboratory, Princeton, NJ
Advisors: Dr. Adam McLean and Dr. Rajesh Maingi
Tracking heat flux from plasma to the diverter with a novel, dual-band, fast IR camera

Summer 2009 Research Intern Argonne National Laboratory, Lemont, IL
Advisor: Dr. Rik Yoshida
Assessing the performance of ATLAS to track and measure cosmic muons

Publications

Energetic coupling between ligand binding and dimerization in *E. coli* phosphoglycerate mutase
Nathan W. Gardner, Lyman, K. Monroe, Daisuke Kihara and Chiwook Park. *Biochemistry* 2016, 55, 1711-1723.