RESIDUE LEVEL QUANTIFICATION OF PROTEIN STABILITY IN LIVING CELLS

William B. Monteith

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Chemistry.

Chapel Hill
2014

Approved by:

Gary J. Pielak

Nancy L. Thompson

Sharon L. Campbell

Linda L. Spremulli

Max L. Berkowitz

ABSTRACT

William B. Monteith: Residue Level Quantification of Protein Stability in Living Cells
(Under the direction of Gary J. Pielak)

The intracellular milieu differs from the dilute conditions in which most biophysical and biochemical studies are performed. This difference has led both experimentalists and theoreticians to tackle the challenging task of understanding how the intracellular environment affects the properties of biopolymers. Despite a growing number of in-cell studies, there is a lack of quantitative, residue-level information about equilibrium thermodynamic protein stability under non-perturbing conditions. My dissertation addresses this void by describing the use of NMR-detected hydrogen-deuterium exchange of quenched cell lysates to measure individual opening free energies of the 56-amino acid B1 domain of protein G (GB1) in living *Escherichia coli* cells without adding destabilizing co-solutes or heat. Comparisons to dilute solution data (pH 7.6 and 37 °C) show that opening free energies increase by as much as 1.14 ±0.05 kcal/mol in cells. Importantly, I also show that homogeneous protein crowders destabilize GB1, highlighting the challenge of recreating the cellular interior. These findings are discussed in terms of hard-core excluded volume effects, charge-charge GB1-crowder interactions and other factors. The quenched lysate method is applied further in mutational studies of GB1 to make the first quantification of non-specific protein-protein interactions in cells. I show that a surface mutation in GB1 is 10-times more destabilizing in *E. coli*

than in buffer. The results indicate that quinary interactions between surface exposed residues and cytoplasmic proteins can play a key role in determining the native stability of a protein, whereas such a role is absent in buffer alone. The methods developed and applied throughout this work should prove useful for extension to other globular proteins in efforts to gain a more complete understanding of the effects of the intracellular environment on protein chemistry.

To the two most patient and loving people I will ever know,

Sharon Ann Gerwig and Gary Vincent Monteith

ACKNOWLEDGEMENTS

Within one week of rotating in the laboratory of my mentor, Prof. Gary Pielak, I knew I had found my home for the foreseeable future. It wasn't because of the science (I thought all *E. coli* was dangerous to be around prior to starting my PhD), but because I had found a highly functional and intelligent human being with some of the most bizarre neuroticisms I had ever encountered. I was intrigued. Over the years, he never ceased to amaze and amuse with his one-liners, nuttiness and general lack of tact (most of which I have kept an electronic record of and will pass down to current Pielakers when I leave the lab). I am lucky to have the distinct pleasure of sharing an apartment for three months with Gary and his wife, Elizabeth, in Galway Ireland working with collaborators at NUI-Galway (many thanks to Elizabeth for being such a wonderful cook and sustaining me not only during our time in Ireland, but also the many times at the Pielak Lab gatherings at 3 Martha Lane).

I have the utmost gratitude and respect for Gary. He motivated me out of numerous local minima throughout the years and instilled in me the perseverance to keep pushing in the laboratory. Gary has always given me all the credit for developing the 'quenched-lysate' method for measuring in-cell protein stability, but the truth is that I wanted to walk away from that project many times in the early years, and I could not have turned it into a success without his encouragement,

wisdom, support and espresso. His mental database for scholarly literature is freakish and unparalleled – ask him a question in protein chemistry and he returns the seminal paper, including author, journal and year almost immediately. I guess that's par for the course for a guy who translates Morse code at 5 am each day. Thanks for the memories and the education, GJP.

I would like to acknowledge the Department of Chemistry at East Carolina University for an outstanding education and excellent research opportunities prior to beginning my doctoral work. I am indebted to Prof. Allison Danell for teaching me to think like an analytical chemist, and I thank Profs. Rickey Hicks and Colin Burns for sparking my fascination with nuclear magnetic resonance. A big thank you to the department for inviting Gary and me to give a seminar in March 2014 at ECU. It was truly an honor and the most fun I have ever had presenting my work.

Many thanks to Prof. Peter Crowley and his laboratory at NUI-Galway for hosting me during the Walton fellowship. Special thanks to Ciara Kyne for aiding in backbone amide chemical shift assignments at low pH, and generally for showing me such great craic in beautiful Ireland.

I owe a great deal of appreciation to Dr. Marc ter Horst in the NMR facility for giving the Pielak group first priority on the Varian 600 MHz instrument and for rapidly addressing spectrometer maintenance issues as they arose. I also thank Dr. Ashutosh Tripathy in the Mac-In-Fac for teaching me how to use the microcalorimeter and for assistance with DSC experiments.

I feel fortunate for being involved in such a wonderful and fun lab family, from previous members to current Pielakers. I am particularly grateful for the initial

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND SYMBOLS

$\Delta G_{op}^{\circ'}$ standard, modified Gibbs free energy of opening

$\Delta\Delta G_{op}^{\circ'}$ change in standard, modified Gibbs free energy of opening

$\delta\Delta\Delta G_{op,int}^{\circ'}$ standard, modified interaction Gibbs free energy of opening

$\Delta H_{vH}^{\circ'}$ standard, modified van't Hoff enthalpy of denaturation

$\Delta H_{cal}^{\circ'}$ standard, modified calorimetric enthalpy of denaturation

$\Delta S^{\circ'}$ standard, modified entropy of denaturation

°C   degree Celcius

μ micro- (10^{-6})

x g centrifugal force, times gravity

BSA   bovine serum albumin

CI2   chymotrypsin inhibitor 2

DSC   differential scanning calorimetry

EDTA   ethylene diamine tetraacetic acid

FPLC   fast protein liquid chromatography

g   gram

GB1   B1 domain of streptococcal protein G

h   hour

H/D   hydrogen/deuterium

HSQC   heteronuclear single quantum coherence

IPTG   isopropyl β-D-1-thiogalactopyranoside

k_{cl} closing rate constant

k_{int}	intrinsic rate constant of hydrogen-deuterium exchange
k_{obs}	observed rate constant of hydrogen-deuterium exchange
k_{op}	opening rate constant
K_{eq}	equilibrium constant of denaturation
K_{op}	equilibrium constant of backbone amide opening
kcal/mol	kilocalories per mole
kDa	kilodalton
L	liter
LB	Luria Bertani medium
M	molar
m	milli- (10^{-3})
min	minute
NMR	nuclear magnetic resonance
OD_{600}	optical density at 600 nanometers
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
pI	isoelectric point
s	second
SDS	sodium dodecyl sulfate
T_m	melting temperature
wt	wildtype
Z_{net}	net charge

CHAPTER 1: PROTEIN STABILITY – HISTORICAL PERSPECTIVES AND A LOOK AHEAD

1.1 Introduction

The structure-function paradigm states that the amino acid sequence of a globular protein encodes its singular, minimum-energy and biologically active conformation. It is clear from Levinthal's paradox[1] and the short timescale of protein folding that biology does not rely on a random search through conformational space for a protein to find its functional form, but rather that the protein folding 'problem' (*i.e.*, the inability to predict the three dimensional structure of a native globular protein from its primary structure alone) should be solvable by invoking the principles of physical chemistry. Knowledge of the specific energetics of folding – namely the driving and opposing forces that result in net *stabilization* – from primary structure will open the door to innumerable opportunities in enzyme design, biotechnology and the engineering of protein-based therapeutics.

Understanding how proteins fold[2] traces back nearly a century when Wu proposed that protein denaturation was, in fact, due to unfolding of the amino acid chain held together through noncovalent linkages.[3] Importantly, the unfolding process was shown to be thermodynamically reversible,[4-6] giving rise to the pivotal idea that the equilibrium between folded and unfolded states could be monitored biophysically.

What forces are responsible for controlling this equilibrium? Early theories[7,8] suggested a primary role for electrostatic interactions between charged side chains, but this idea was disproven based on inconsistencies in observed 'electrostricted' volumes of model compounds.[9] Then, following the discovery of the α-helix and β sheets by Mirsky and Pauling,[10] it was widely held that peptide hydrogen bonds dominated. Kauzmann, however, argued that because H-bonds could form in both the folded and unfolded states, these interactions, albeit stabilizing, could not *drive* folding.[11] Instead, he shifted the focus outside of the protein: water. Hydrophobic species were known to have low affinity in aqueous media, because water molecules form complex, entropically disfavored[12] structures that encapsulate the hydrophobe. Kauzmann's idea of 'hydrophobic bonding'[11] led to a new view, where collapse of hydrophobic side chains into a compact, globular structure minimized the ordering of water molecules. As more definitive evidence against hydrogen bonding surfaced,[13,14] evidence for the hydrophobic effect mounted,[15-17] solidifying it as the principle mechanism responsible for folding.

Stabilization of the native state occurs by overcoming, primarily, the loss of conformational entropy associated with the unfolded state *via* hydrophobic collapse, van der Waals forces, coulombic attractions, H-bonding, and disulfide linkages. This remarkable balancing of free energy cannot be understated: the total free energy associated with the folded and unfolded states of small proteins (not including covalent bonds) is hundreds of kcal/mol, whereas most globular proteins are stabilized by only 5 – 15 kcal/mol,[18] energetically equivalent to the free energy of one or two typical hydrogen bonds in water. Nevertheless, it is precisely this small

margin of stabilization that defines the functional role of a protein. The following

sections focus on some of the the experimental methods used to measure such

subtle differences in free energy.[1]

1.2 Conformational protein stability

Anfinsen's pioneering work with ribonuclease A, which resulted in his

'thermodynamic hypothesis' (and a Nobel Prize), states that the global minimum in

free energy is attained in the folded, native state.[5] Although it must be appreciated

that exceptions have been identified, for instance when folding is kinetically

controlled,[20] my dissertation focuses on the wealth of proteins that undergo

reversible, two-state folding (*i.e.*, no intermediates are significantly populated at

equilibrium) under equilibrium thermodynamic control.

More specifically, the emphasis is on the conformational stability of globular

proteins, defined as the difference in standard state Gibbs free energy under

physiological conditions between the native (N), structurally-defined and biologically

active folded state and the thermodynamic ensemble of denatured (D), unfolded

states (Figure 1.1).

$$N \rightleftharpoons D \qquad \qquad \textbf{[1]}$$

$$\Delta G^{\circ'} = G_D^{\circ'} - G_N^{\circ'} \qquad \qquad \textbf{[2]}$$

---

[1]Equilibrium measurements will be emphasized throughout this work; kinetic measurements of equilibrium stability obtained *via* chevron analysis[19] are not discussed.

1.3 Measurement of K_{eq}

Protein stability can be measured by determining the relative populations of N and D at equilibrium with a variety of spectroscopic tools.[21]

$$\Delta G^{o'} = -RT\ln\frac{[D]}{[N]} = -RT\ln K_{eq} \qquad [3]$$

In terms of population distributions, stabilization by ~7 kcal/mol corresponds to 10^5 molecules in the native state for every one of the unfolded species at 37 °C. Hence, making stability measurements *in vitro* under native conditions can be impractical because calculations of K_{eq} are only accurate in the range of 0.1 – 10 using most techniques. The lone exception, amide exchange, is discussed below. Instead, the equilibrium shown in Eq. **1** is typically perturbed by adding heat or a denaturing co-solute (such as urea or guanidinium chloride) to populate the unfolded state and the response is measured by various spectroscopic methods (*e.g.*, electronic absorption, circular dichroism, and fluorescence spectroscopies). Tracking the denaturation results in a progress curve with a native and a denatured baseline (left panel, Figure 1.2). Assigning these baselines allows the fractions of the denatured state and native state to be determined as a function of denaturant concentration or T in the steep transition region. In co-solute unfolding experiments, K_{eq} is calculated using Eq. **3** and the resulting plot of $\Delta G^{o'}$ as a function of [denaturant] is then extrapolated to zero denaturant concentration, $\Delta G^{o'}(H_2O)$.[22] For thermal denaturation, K_{eq} is calculated and a van't Hoff plot of $\ln(K_{eq})$ versus T^{-1} yields $\Delta H_{vH}^{o'}$ values at temperatures around the T_m of the protein.[6] If the change in heat capacity (ΔC_p) upon unfolding is known, a stability curve[23] can be extrapolated to yield $\Delta G^{o'}$ at

temperatures beyond the experimentally determined range using the Gibbs-Helmholtz equation (see below).

Stability measurements can also be made under non-perturbing conditions by pairing hydrogen-deuterium (H/D) exchange[24,25] with NMR spectroscopy[26] or mass spectrometry.[27] NMR-detected H/D exchange is the primary method for measuring stability in the Pielak lab and in this dissertation. Since the advent of two-dimensional methods in the 1980s,[28] stability information can be resolved to the level of individual backbone amide protons. The trade off is the requirement for more protein sample with NMR experiments. In amide H/D exchange, exposure of backbone N-H groups to solvent through local fluctuations of the native state and higher-energy global unfolding events results in irreversible H/D exchange when the protein is resuspended in D_2O. With the validity of several assumptions (see Chapter 3), the ratio of the observed rate of N-H peak decay (k_{obs}) to the intrinsic rate of exchange (k_{int}) obtained from data on unstructured peptides[29] defines the equilibrium constant for the opening (*i.e.*, unfolding) reaction of each N-H group (Figure 1.3).

$$\Delta G_{op}^{\circ\prime} = -RT \ln \frac{k_{obs}}{k_{int}} = -RT \ln K_{op} \qquad \textbf{[4]}$$

In this way, an opening free energy ($\Delta G_{op}^{\circ\prime}$) is obtained for each (non-proline) backbone amide. The largest values of $\Delta G_{op}^{\circ\prime}$ approximate $\Delta G^{\circ\prime}$ obtained from global denaturation experiments.[30] The mechanism of H/D exchange and applications are described in further detail in Chapters 3 and 4.

1.4 Calorimetric measurement of $\Delta G^{\circ\prime}$

Differential scanning calorimetry (DSC) directly measures the enthalpy of denaturation and, therefore, can be used with the Gibbs-Helmholtz equation to obtain $\Delta G^{\circ\prime}$.

$$\Delta G^{\circ\prime} = \Delta H^{\circ\prime} - T\Delta S^{\circ\prime} \tag{5}$$

The development of DSC by Privalov[16] allowed determination of the excess heat capacity of the protein as a function of temperature, resulting in two key advances. First, the discovery that ΔC_p on denaturation is nonzero and (nearly) constant[16] reaffirmed the role of the hydrophobic effect[11] in protein stability. Second, as outlined below, DSC provides the best evidence for two-state behavior.

Knowing that ΔC_p for denaturation is positive and assuming the reaction involves only two thermodynamic states leads to the following form of Eq. **5**.[23,31]

$$\Delta G^{\circ\prime}(T) = \Delta H^{\circ\prime}_{cal}\left(1 - \frac{T}{T_m}\right) + \Delta C_p\left[(T - T_m) + T\ln\frac{T}{T_m}\right] \tag{6}$$

$\Delta H^{\circ\prime}_{cal}$ is the enthalpy change at the midpoint of the melting transition (T_m), where $\Delta G^{\circ\prime} = 0$ and $\Delta S^{\circ\prime}_m = \frac{\Delta H^{\circ\prime}_{cal}}{T_m}$. Precise measurement of ΔC_p, T_m and $\Delta H^{\circ\prime}_{cal}$ (Figure 1.4) allow protein stability curves[23] of $\Delta G^{\circ\prime}$ as a function of T to be calculated.

Denaturation can also be followed in a model-dependent manner by integrating the heat capacity as a function of temperature, giving the progress curve in Figure 1.2. This treatment of calorimetric data allows the van't Hoff enthalpy to be

obtained as described above for spectroscopic measurements. The major

advantage of calorimetric determinations of stability is that the measurement of $\Delta H_{cal}^{o'}$

does not assume a two-state process, which is implicit in the other methods.

Therefore, the best evidence for an equilibrium two-state folding process is when

$\Delta H_{cal}^{o'}$ measured calorimetrically matches the model-dependent van't Hoff $\Delta H_{vH}^{o'}$.

1.5 Macromolecular crowding and protein stability

Traditional biochemistry and biophysics examine the properties of proteins,

including stability, in simple buffered solutions far removed from the complex,

crowded native environment within cells (Figure 1.5).[32] Initial theories[33-35] concerning

the effects of macromolecular crowding on protein stability treated the crowd as inert

spheres, giving rise to repulsive interactions between the test protein and myriad

macromolecules occupying the cell. Experiments followed that employed synthetic

polymers such as dextran, Ficoll, and polyethylene glycols as crowders to mimic the

cellular interior.[36-39] The results were mostly consistent; crowding stabilized proteins

compared to buffer alone. Furthermore, the mechanism of stabilization seemed

clear: hard-core excluded volume effects shift the equilibrium toward the compact,

native structure by destabilizing the denatured state ensemble, a purely entropic

effect.

Although insightful, natural progression in the field required moving away from

studies using synthetic polymers and the treatment of crowding molecules as inert,

hard spheres. After all, cells are full of proteins and nucleic acids, not synthetic

polymers.[32,40] The importance of transient, non-specific protein-protein interactions[41]

to the equilibrium partitioning of molecules in the native and denatured states was realized by studies with more physiologically relevant crowders, such as proteins and reconstituted cytosol.[42-45] These 'soft interactions'[46] have been found to modulate the stability in a context dependent fashion, where attractive chemical interactions tend to destabilize test proteins by lowering the free energy of the denatured state ensemble, similar to how urea destabilizes proteins.[47] On the other hand, repulsive interactions are expected to favor the native state and enhance hard-core stabilizing effects.[48] Moreover, Minton[49] and Zhou[50] have independently suggested temperature as an important factor in modulating hard and soft effects. In summary, the net effect of crowding on protein stability will be a combination of steric repulsions and non-specific chemical interactions that depend on the chemical nature of the test protein and crowding species.

1.6 Protein stability in cells

To gain the most physiologically relevant and comprehensive understanding of how macromolecular crowding affects protein stability, quantitative information gleaned from studies performed inside living cells needs to be combined with in-cell molecular dynamics simulations.[51] The seminal experiments of Ghaemmaghami and Oas[52] on λ repressor showed that its stability was unchanged in cells relative to buffer alone using H/D exchange and mass spectrometry. Later, Gierasch and workers used a fluorescent tag in cellular retinoic acid-binding protein I to report destabilization in cells.[53,54] These experimental results were corroborated *in silico* only when both hard and soft interactions were taken into consideration.[51] More recently, the Gruebele group has added significantly to the breadth of knowledge

about in-cell folding and stability with fluorescence imaging of phosphoglycerate kinase and the cell surface protein VIsE.[55-58] Their results demonstrate that protein stability can be modulated *via* cellular compartmentalization, tying together hypotheses about the strength and type of soft interactions that a protein experiences in a given environment.

However, the in-cell studies above have relied on urea denaturation[52-54] or thermal modification[55-58] to quantify global stability. Such perturbations are known to trigger stress responses inside cells.[59,60] Chapter 3 describes the first residue-level measurement of protein stability under non-perturbing conditions using native state H/D exchange and NMR spectroscopy.[48] I found that the B1 domain of protein G (GB1) is stabilized in cells, which is attributed to repulsive soft interactions. I augment these results (Chapter 4) by reporting the first quantification of soft interactions between the cytoplasm of *E. coli* and a variant of GB1.

1.7 Summary

The more details we amass from experiments performed in cells, the closer we approach a fundamental knowledge of protein chemistry in the cellular environment. Because the functional role of proteins is defined by the equilibrium between the structural native state and the denatured state ensemble, this endeavor starts with understanding the forces stabilizing proteins in cells. In addition to furthering our fundamental knowledge of biochemistry, advancements in the field of macromolecular crowding will drive efforts to stabilize protein-based therapeutics. Nevertheless, it must be appreciated how the complex and heterogeneous cellular

environment hinders in-cell biophysical measurements. For example, calorimetry will never be able to distinguish the enthalpy of unfolding of a test protein in a sea of macromolecules and is therefore limited to purified protein samples in buffer or in simple, unphysiological, crowded conditions. In this vein, the purpose of Chapter 2 is to illustrate the problems associated with making quantitative protein stability measurements using traditional H/D experiments and in-cell NMR.[61] The remainder of my dissertation (Chapters 3 and 4) describes applications of the method I developed to circumvent these limitations.

1.8 Figures
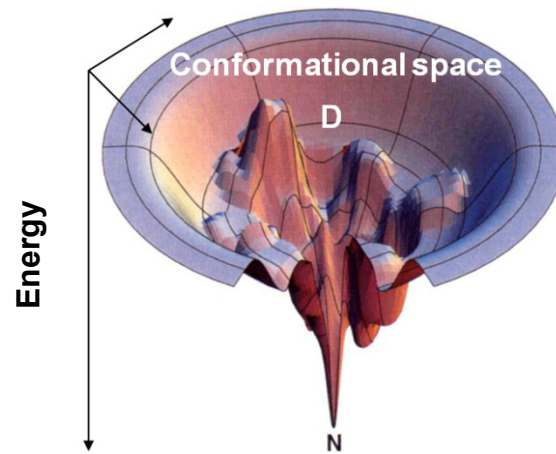
Figure 1.1 Folding landscape for a globular protein.[62] The presence of intermediates and molten-globule states can make the landscape rugged, as depicted. Species in these kinetic traps are not significantly populated for proteins adhering to two-state equilibrium unfolding.
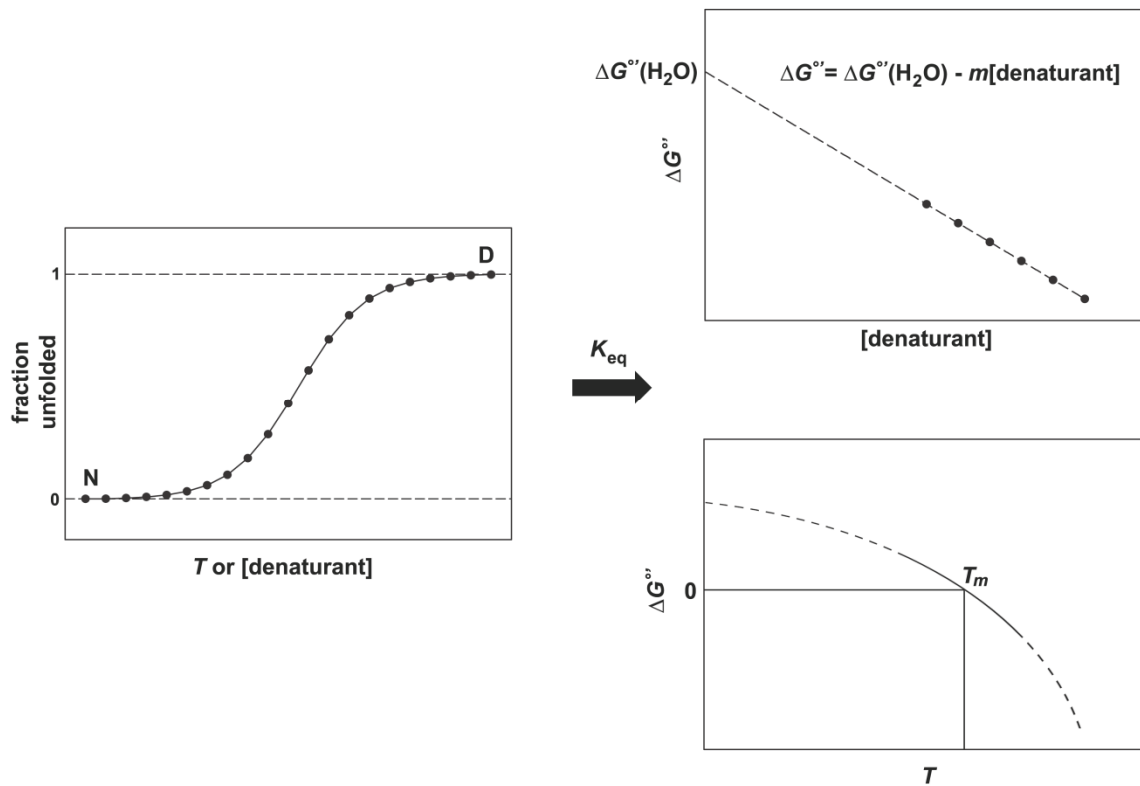
Figure 1.2 Determination of K_{eq} and $\Delta G^{\circ\prime}$ from co-solute (upper right panel) or

thermal (lower right panel) denaturation experiments. The progress curve in the left

panel relies on the assumption of a two-state folding model to track K_{eq}. The slope of

the line in the upper right panel is the *m*-value, a measure of the dependence of

$\Delta G^{\circ\prime}$ on denaturant concentration for a given protein.
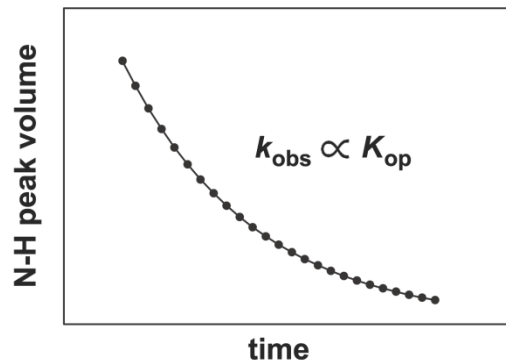
Figure 1.3 Backbone amide H/D exchange rates can provide residue-level

information on protein stability without the use of heat or co-solute perturbation.
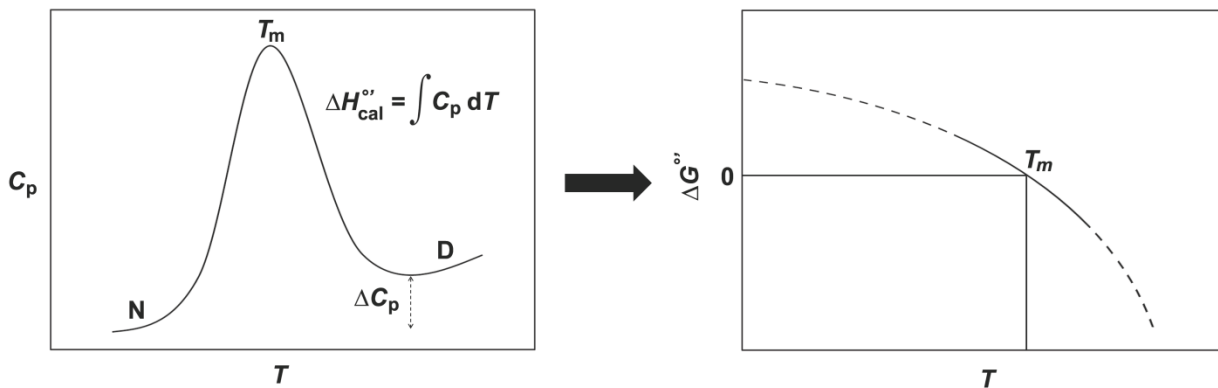
Figure 1.4 Differential scanning calorimetry (DSC) directly measures

thermodynamic parameters of the protein unfolding transition.
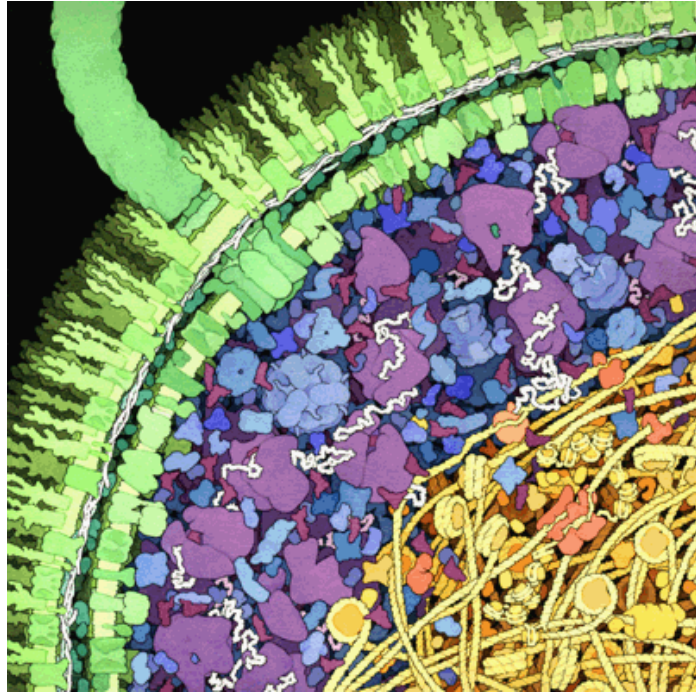
Figure 1.5 Cross-sectional illustration of the *Escherichia coli* interior. The cell wall is shown in green. The cytoplasmic area is colored blue and purple. The large purple molecules are ribosomes and the small, maroon molecules and white strands are tRNA and mRNA, respectively. Enzymes are blue. The nucleoid region containing DNA and DNA polymerases is in yellow and orange. Illustration used with permission from David S. Goodsell at Scripps Research Institute.

CHAPTER 2: DIFFERENTIAL ROTATIONAL DYNAMICS OF GLOBULAR PROTEINS IN CELLS[1]

2.1 Introduction

In-cell nuclear magnetic resonance spectroscopy[63] utilizing the $^1H - {}^{15}N$ heteronuclear single quantum coherence (HSQC) experiment yields high quality spectra when applied to intrinsically disordered proteins in *Escherichia coli*. For globular proteins, however, protein signals from inside the cell are not observed.[64,65] Here, we show in a simple and direct way that protein dynamics determines the quality of the in-cell $^1H - {}^{15}N$ HSQC spectrum by fusing the globular protein, ubiquitin, to the disordered protein, α-synuclein.

Most knowledge about protein structure and dynamics is gleaned from dilute solution experiments. The native intracellular environment of proteins, where macromolecular concentrations may exceed 300 g/L,[32] presents a different set of conditions. Furthermore, it is known that crowding can impact protein stability and dynamics.[66,67]

The $^1H - {}^{15}N$ HSQC experiment[68] is commonly used to characterize proteins in dilute solution. However, the ability to obtain high resolution spectra from overexpressed ^{15}N-enriched proteins in cells yields mixed results. Disordered

---

[1] The material in this chapter was published in ChemBioChem. The original citation is as follows: Barnes CO, Monteith WB, Pielak GJ (2011) Internal and global protein motion assessed with a fusion construct and in-cell NMR spectroscopy. ChemBioChem 12:390-391. WB Monteith and GJ Pielak wrote the paper, and permission was received from CO Barnes to include the work in this dissertation.

proteins, such as α-synuclein[64] and FlgM,[69] exhibit high quality spectra inside cells. Most globular proteins, on the other hand, fail to produce useful in-cell spectra.[64,65] Here, we present in a single experiment compelling evidence that this difference in detectability is caused by their different rotational dynamics.

2.2 Results

2.2.1 Production of the ubiquitin-α-synuclein fusion construct

We produced a histidine-tagged fusion protein from the globular protein, ubiquitin, and the disordered protein, α-synuclein (Figure 2.1). Ligation independent cloning[70] resulted in a structural gene comprising a N-terminal six-histidine (H_6) segment and a short disordered linker followed by the UBQ-αSN sequence. The composition of the intra protein linker was chosen to ensure flexibility between the constituent proteins.

The fusion protein was visualized by using SDS-PAGE after 1.5 h of expression in *E. coli* (Figure 2.4). Its apparent molecular weight (29 kDa) is the sum of the apparent molecular weights of purified H_6-ubiquitin (9.5 kDa) and wildtype α-synuclein (19 kDa), proving that the fusion protein is intact in cells [α-synuclein (calculated MW, 14 kDa) migrates abnormally slowly].[71]

2.2.2 In-cell NMR

The in-cell $^1H - {}^{15}N$ HSQC spectrum of UBQ-αSN (Figure 2.2A) matches that of α-synuclein in cells.[72] The narrow ^1H chemical shift range observed is due to the partially ordered structure adopted by α-synuclein. Crosspeaks associated with

ubiquitin,[65] which span a large chemical shift range owing to a well defined structure, are absent. No protein spectrum is observed in the supernatant of the cell sample (Figure 2.5), indicating that the fusion protein does not leak from cells during the experiment. This finding is consistent with work showing that leakage does not occur for either protein.[65]

Upon lysing the cells, we observe not only α-synuclein crosspeaks, but also crosspeaks corresponding to the ubiquitin portion of the fusion protein (Figure 2.2B). To confirm the ubiquitin signals, ^{1}H – ^{15}N HSQC spectra of cell lysates from H$_6$-ubiquitin and α-synuclein expressing cells were overlaid with the UBQ-αSN cell lysate spectrum (Figure 2.3). The contour levels in Figure 2.3 were chosen for clarity. Detailed analysis confirms that each H$_6$-ubiquitin and α-synuclein crosspeak overlaps with a crosspeak from the fusion protein. This overlap, together with the results shown in Figure 2.2, provide unequivocal evidence that the in-cell spectrum is from the C-terminal portion of the fusion protein, which corresponds to α-synuclein, and that the N-terminal ubiquitin is only observed in the cell lysate.

2.3 Discussion

The ability to obtain high resolution NMR spectra depends on the rotational diffusion of the target protein. As rotational correlation times increase, peak broadening prevents detection. Inside cells, however, the effects of crowding (excluded volume, increased viscosity, and binding to other species) impede global rotation, making most globular proteins invisible in ^{1}H – ^{15}N HSQC spectra.[64,65,73]

Our data on the fusion protein provides a simple and direct example of the impact of differential rotational dynamics on the ability to detect proteins via in-cell ^{1}H – ^{15}N-based NMR experiments. The intrinsically disordered portion from α-synuclein, with its ensemble of interconverting dihedral angle combinations, possesses sufficient internal motion even under the crowded conditions in the cell. Put another way, the transverse relaxation is slow enough to yield a high resolution ^{1}H – ^{15}N HSQC spectrum. Conversely, the rotational motion of the N-terminal ubiquitin portion is dominated by slower, global tumbling that is made even slower in cells. We have confirmed this result *in vitro* with proteins as crowding agents.[73] The crowding in these environments causes the nuclei to relax quickly, broadening the crosspeaks beyond detection. Only when the cells are lysed and the lysate diluted are the effects mitigated, providing the globular portion with enough rotational freedom for its spectrum to be observed.

Localization studies indicate that the concentration of the fusion protein is nearly equal in the cytoplasm and periplasm (data not shown). Thus any differences between the crowding conditions of the cytoplasm and periplasm are not large enough to dramatically impact the rotational dynamics in either location.

2.4 Summary

We have shown that the intrinsically disordered component of UBQ-αSN exhibits a high resolution spectrum, whereas the globular portion is only observed when the cells are lysed and diluted. These findings demonstrate succinctly the importance of a protein's rotational diffusion properties on the feasibility of obtaining a high quality ^{1}H – ^{15}N HSQC spectrum in cells. Although spectra from globular

proteins in cells can be obtained by using ^{19}F NMR,[65] acquiring useful ^{15}N spectra remains a significant challenge.

2.5 Materials and methods

2.5.1 Production of the fusion construct

The genes encoding human ubiquitin and human α-synuclein were amplified by using the polymerase chain reaction. The ~230 bp fragment for ubiquitin was amplified with a mixture of the following primers: Forward 1 – 5' GAC GAC GAC AAG ATG GCA ATC TTC GTC AAG ACG 3', Forward 2 – 5' GAC GAC GAC AAG ATG GCA ATC TTC GTC AAG ACG TTA ACC GG 3', Reverse – 5' CAT CTT GTC GTC GTC GCA ACC ACC TCT TAG TCT TAA GAC 3', with the reverse primer encoding a mutation (underlined) that changes ubiquitin's stop codon to a cysteine. The ~430 bp α-synuclein gene was amplified with a mixture of the following primers: Forward – 5' GAC GAC GAC AAG ATG GAT GTA TTC ATG AAA GGA 3', Reverse 1 – 5' TGA GGA GAA GCC CGG TTA CGC CTC AGG TTC GTA GTC 3', Reverse 2 – 5' TGA GGA GAA GCC CGG TTA GGC CTC AGG TTC GTA GTC 3'. PCR products were treated with 1 μL of 20 U/μL DPN I (Biolabs) for 1 h at 37 °C. The PCR product was purified on a spin column (SpinPrep^{TM} PCR Clean-Up Kit, Qiagen).

Compatible overhangs were created by treating the purified PCR product with T4 DNA Polymerase as described in the protocol for the Ek/LIC Cloning Kit (Novagen). The treated inserts were annealed by mixing 5 ng of ubiquitin insert with 10 ng of α-synuclein insert at 22 °C for 10 min. After 10 min, EDTA was added to a

final concentration of 7 mM, and the solution incubated at 22 °C for an additional 5 min. The pET-46 Ek/LIC vector was annealed to the product from the previous step as described in the protocol for the Ek/LIC Cloning Kit (Novagen). The annealing reaction was transformed into NovaBlue GigaSingles™ Competent Cells (Novagen) and XL21 Supercompetent Cells (Stratagene). The samples were plated against 0.1 mg/mL ampicillin before incubating overnight at 37 °C. The sequences of the purified DNA from multiple colonies were determined at the UNC-CH Genome Analysis Facility. DNA encoding the fusion protein was subjected to site-directed mutagenesis to change the cysteine residue at the end of ubiquitin to a glycine (underlined) with the following primers: 5' AGA CTA AGA GGT GGT GGC GAC GAC GAC AAG 3', along with its reverse complement: 5' CTT GTC GTC GTC GCC ACC ACC TCT TAG TCT 3'. Products were treated with 1 μL of 20 U/μL DPN I for 1 h at 37 °C before being transformed into competent DH5α cells (Invitrogen). The sequence of the desired construct, called UBQ-αSN, was confirmed.

2.5.2 Protein expression and in-cell NMR

Plasmids containing the UBQ-αSN gene were transformed into BL21 (DE3) Gold cells (Stratagene) and selected with ampicillin (0.1 mg/mL). Cells were grown and subjected to NMR experiments as described by Li *et al.*,[65] after 1.5 h of expression. Data were acquired for 4 h at 37 °C, although a complete spectrum could be observed in 20 min. Protein detection was carried out by using SDS-PAGE (18% Tris-HCl Criterion™ gels, Bio-Rad) with Coomassie blue staining. Protein location was determined by osmotic shock as described by Slade *et al.*[74] NMR data

were acquired at the UNC Biomolecular NMR facility on a Varian Inova 600 MHz

NMR spectrometer. Data were processed and visualized with NMRPipe[75] and

NMRViewJ,[76] respectively.

2.6 Figures

$$H_6VDDDDK-\boxed{UBQ}-DDDK-\boxed{\alpha SN}$$

Figure 2.1 Cartoon representation of the UBQ-αSN fusion protein.
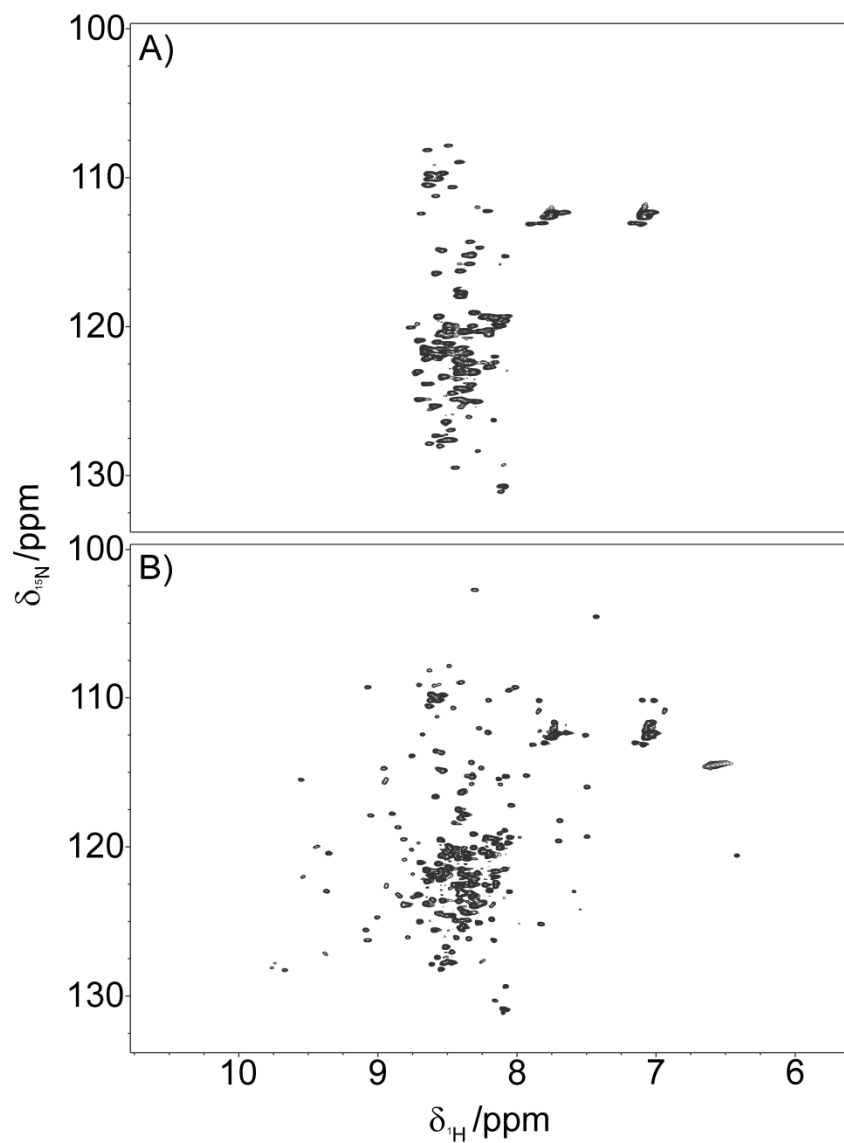
Figure 2.2 A) In-cell, ^1H – ^{15}N HSQC spectrum of *E. coli* expressing UBQ-αSN after 1.5 h. B) Spectrum of the UBQ-αSN cell lysate.
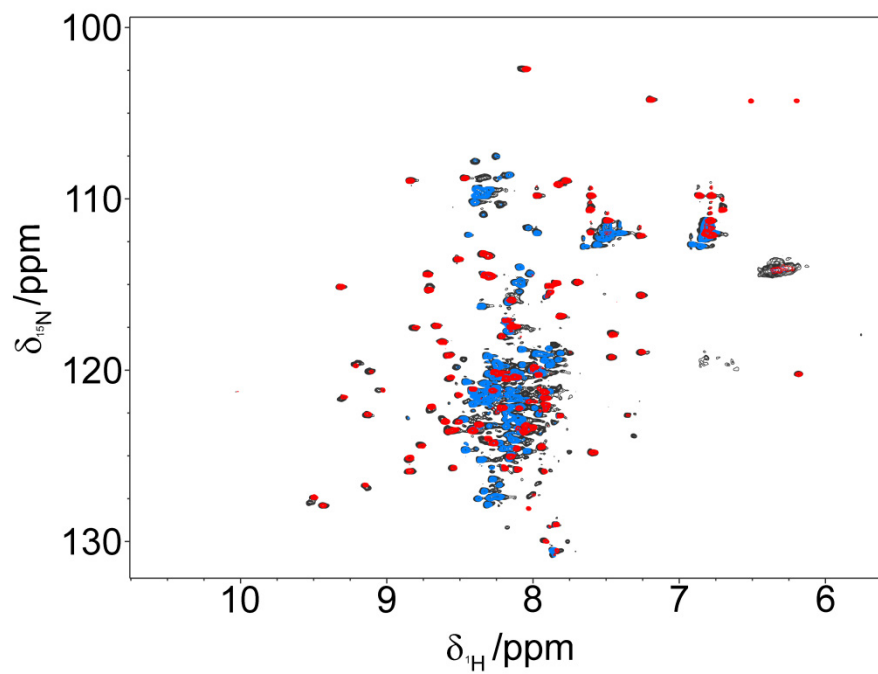
Figure 2.3 Overlaid spectra of UBQ-αSN, α-synuclein, and His$_6$-ubiquitin cell lysates after 1.5 h expression in *E. coli*.
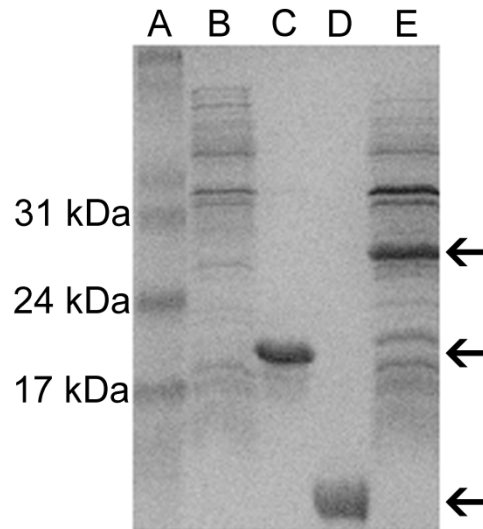
Figure 2.4 SDS-PAGE of (A) molecular weight markers, (B) *E. coli* lysate from cells expressing the UBQ-αSN fusion protein prior to induction, (C) 12 μL of 1 μg/mL purified wildtype α-synuclein, (D) 12 μL of 1 μg/mL purified H$_6$-ubiquitin, and (E) *E. coli* culture from (B) 1.5 h after induction with 1 mM IPTG. Arrows indicate H$_6$-ubiquitin, wildtype α-synuclein, and UBQ-αSN.
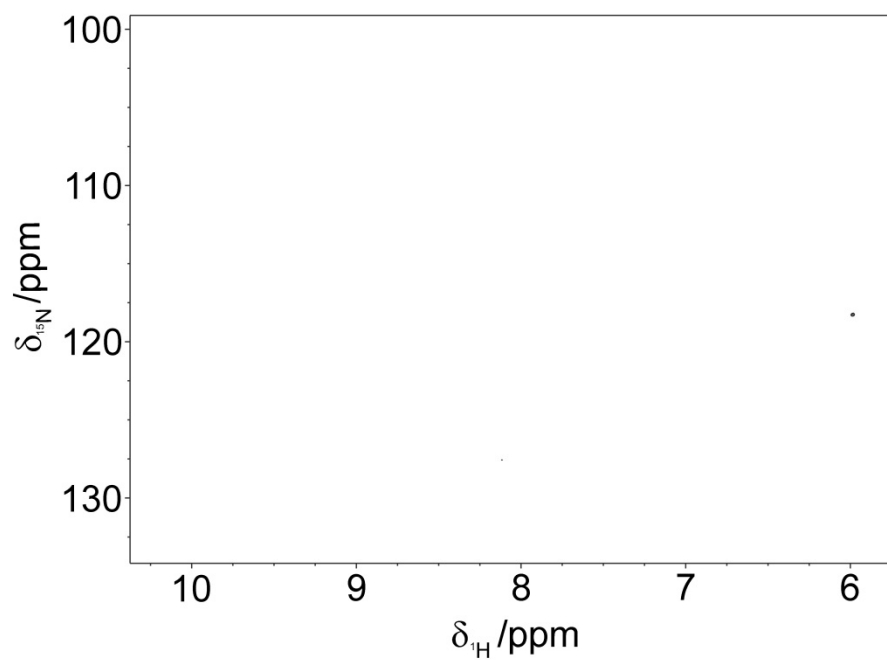
Figure 2.5 ^{1}H – ^{15}N HSQC spectrum of supernatant acquired immediately after acquiring the in-cell UBQ-αSN spectrum. Data was acquired using the same parameters as those used to acquire the data in Figure 2.2.

CHAPTER 3: RESIDUE LEVEL QUANTIFICATION OF PROTEIN STABILITY IN LIVING CELLS[1]

3.1 Introduction

Proteins function in a heterogeneous and crowded intracellular environment. Macromolecules comprise 20 – 30% of the volume of an *Escherichia coli* cell and reach concentrations of 300 – 400 g/L.[32,77] Theory predicts that the properties of proteins and nucleic acids can be significantly altered in cells compared to buffer alone.[78,79] Nevertheless, most biochemical and biophysical studies are conducted under dilute (<10 g/L macromolecules) conditions. Here, we augment the small but growing list of reports probing the equilibrium thermodynamic stability of proteins in living cells,[52,54-56,58] and provide the first measurement of residue-level stability under non-perturbing conditions.

Until recently, the effects of macromolecular crowding on protein stability were thought to be caused solely by hard-core, steric repulsions arising from the impenetrability of matter.[33,36,79] The expectation was that crowding enhances stability by favoring the compact native state over the ensemble of denatured states. Increased attention to transient, non-specific protein-protein interactions[41,80-82] has

---

[1] The material in this chapter has been accepted for publication in PNAS. The original citation is as follows: Monteith WB, Pielak GJ (2014) Residue level quantification of protein stability in living cells. Proc Natl Acad Sci U S A. WB Monteith and GJ Pielak wrote the paper.

led both experimentalists[46,49,83,84] and theoreticians[43,51,85] to recognize the effects of chemical interactions between crowder and test protein when assessing the net effect of macromolecular crowding. These weak, non-specific interactions can reinforce or oppose the effect of hard-core repulsions, resulting in increased or decreased stability depending on the chemical nature of the test protein and crowder.[42,44,45,86]

We chose the B1 domain of streptococcal protein G[87] as our test protein because its structure, stability and folding kinetics have been extensively studied in dilute solution.[88-98] Its small size (56 amino acids; 6.2 kDa) and high thermal stability make GB1 well suited for studies by NMR spectroscopy.

3.1.1 NMR-detected backbone amide H/D exchange

Quantifying the equilibrium thermodynamic stability of proteins relies on determining the relative populations of native and denatured states. Because the denatured state ensemble of a stable protein is sparsely populated under native conditions, stability is usually probed by adding heat or a co-solute to promote unfolding so that the concentration ratio of the two states can be determined.[99] However, stability can be measured without these perturbations by exploiting the phenomenon of backbone amide H/D exchange[24] detected by NMR spectroscopy.[100] The observed rate of amide proton (N-H) exchange, k_{obs}, is related to equilibrium stability by considering a protein in which each N-H exists in an open (exposed, exchange-competent) state, or a closed (protected, exchange-incompetent) state.[24,25]

$$\text{closed (N-H)} \underset{k_{cl}}{\overset{k_{op}}{\rightleftharpoons}} \text{open (N-H)} \xrightarrow{k_{int}} \text{open (N-D)} \underset{k_{op}}{\overset{k_{cl}}{\rightleftharpoons}} \text{closed (N-D)} \qquad [1]$$
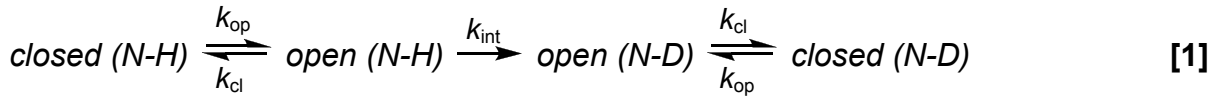
Each position opens and closes with rate constants, k_{op} and k_{cl} (where $K_{op} = k_{op}/k_{cl}$), and exchange from the open state occurs with intrinsic rate constant, k_{int}. Values for k_{int} are based on exchange data from unstructured peptides.[29,101] If the test protein is stable (*i.e.*, $k_{cl} \gg k_{op}$), the observed rate becomes:

$$k_{obs} = \frac{k_{op}k_{int}}{k_{cl} + k_{int}} \qquad [2]$$

Exchange occurs within two limits.[25] At the EX1 limit, closing is rate determining, and $k_{obs} = k_{op}$. This limit is usually observed for less stable proteins and at basic pH.[102] Most globular proteins undergo EX2 kinetics, where exchange from the open state is rate limiting (*i.e.*, $k_{cl} \gg k_{int}$), and k_{obs} values can be converted to equilibrium opening free energies, $\Delta G_{op}^{\circ'}$.[26]

$$k_{obs} = \frac{k_{op}}{k_{cl}}k_{int} = K_{op}k_{int} \qquad\qquad \Delta G_{op}^{\circ'} = -RT \ln \frac{k_{obs}}{k_{int}} \qquad [3,4]$$

The backbone amides most strongly involved in H-bonded regions of secondary structure exchange only from the fully unfolded state, yielding a maximum value of $\Delta G_{op}^{\circ'}$.[30,103,104] For these residues $\Delta G_{op}^{\circ'}$ approximates the free energy of denaturation, $\Delta G_{den}^{\circ'}$, providing information on global stability. Lower amplitude fluctuations of the native state can give rise to partially unfolded forms,[105] resulting in residues with $\Delta G_{op}^{\circ'}$ values less than those of the global unfolders.

In summary, NMR-detected H/D exchange can measure equilibrium thermodynamic stability of a protein at the level of individual amino acid residues under non-perturbing conditions. Inomata *et al.*[106] employed this technique to measure k_{obs} values in human cells for four residues in ubiquitin, but experiments confirming the exchange mechanism were not reported and opening free energies were not quantified. Our results fill this void and provide quantitative residue-level protein stability measurements in living cells under non-perturbing conditions.

3.2 Results

3.2.1 Exchange in cells

We attempted to measure GB1 stability directly in cells by pairing H/D exchange with in-cell NMR[107] and conventional serial ^{15}N-^{1}H HSQC analysis,[108] because GB1 is one of the few proteins that gives reasonable spectra in *E. coli*.[80,81,109,110] The signal-to-noise ratios of the spectra, however, were insufficient for quantification. We overcame this problem by modifying the approach of Ghaemmaghami and Oas[52] for measuring stability in 'discrete', quenched cell lysates (Figure 3.1 and Section 3.4.3). Briefly, the cells are transferred and washed into D_2O, where they remain viable for the course of the experiment (Figure 3.11). An aliquot of cell slurry is removed at defined times, the cells are lysed, exchange is quenched and the lysate is analyzed by NMR. The dead time is ~1 h. Representative HSQC spectra of the initial and final lysates of wt GB1 illustrate the decrease in N-H crosspeak volume for 17 backbone amides due to exchange

(Figure 3.2). Profiles for representative residues are shown in Figure 3.3. Values of k_{obs} are tabulated in Table 3.1.

The decay of the T18 cross peak illustrates the upper limit for measuring exchange. Quantification of k_{obs} for T18 required a lower contour level and fitting to fewer times than the 17 more slowly exchanging residues. At contour levels lower than those shown in Figure 3.2, resonances from six additional residues (K10, A20, A24, T25, Q32, N35) are detectable in the spectrum from the initial time point, but decay is too rapid to obtain k_{obs}. Rates for these residues are listed as $>k_{obs,T18}$.

Crosspeaks from 24 backbone amides do not appear to exchange. We conclude that these residues are 'quench-labeled'. That is, they are least protected from solvent and, therefore, are labeled with protons immediately prior to quenching, when the proton concentration increases 10^4-fold. This conclusion arises from two considerations. First, the side chain amides, which are solvent exposed, behave similarly. Second, as discussed below, quench-labeling is not observed in the serially-acquired (*i.e.,* no quench step) dilute solution data. Exchange rates for quench-labeled residues are also listed as $>k_{obs,T18}$, even though their rates are probably even larger than those for the six residues described above. We attempted to assign the remaining residues, but the lysate was not stable enough for acquisition of three-dimensional NMR data.

In summary, 48 of the 56 residues provide information on exchange. For the 17 slowly exchanging residues, k_{obs} values were converted to free energies of opening, $\Delta G_{op}^{o'}$, by using Eq. **4** (Figure 3.9 and Table 3.1). The value for T18 was not

included because its rate in cells was obtained from limited data. Elevated rates of intrinsic exchange ($10 - 100$ s^{-1}) under our conditions (pH$_{corr}$ 7.6, 37 °C) prevented quantification of 30 residues. Based on our results for T18, we conclude that the k_{obs} values for these 30 residues are $>7 \times 10^{-4}$ s^{-1} in cells and $>3 \times 10^{-4}$ s^{-1} in buffer (see below).

3.2.2 Exchange in dilute solution

To compare the in-cell rates to those acquired in dilute solution (pH$_{corr}$ 7.6, 37 °C), we mimicked the discrete sampling method (Figure 3.1) using purified GB1 instead of the GB1-containing lysate (Section 3.4.5). The exchange behavior is similar under both conditions; *i.e.*, the same residues exchange slowly, the same residues exchange too rapidly to quantify, and the same quench-labeling is observed. However, quantifying the exchange of T18 was possible because the dead time is <5 min, compared to ~1 h for the in-cell studies (Figure 3.3). Similar to the in-cell data, T18 is the fastest exchanging quantifiable residue. The k_{obs} values are tabulated in Table 3.1, and the concomitant $\Delta G^{o\prime}_{op,buff}$ values were used to calculate the changes in stability (Figures 3.4 and 3.5, Table 3.1) caused by crowding in cells ($\Delta\Delta G^{o\prime}_{op,cell} = \Delta G^{o\prime}_{op,cell} - \Delta G^{o\prime}_{op,buff}$).

To validate the discrete dilute solution protocol, conventional H/D exchange experiments[108] involving serially acquired spectra of a single lyophilized sample were performed in buffer (Section 3.4.6). The $\Delta G^{o\prime}_{op,buff}$ values from discrete and serial acquisitions are the same within the uncertainty of the measurement (Figure 3.9 and Table 3.2). Thus, the serial method was employed for subsequent *in vitro*

studies. These data also show that lyophilization of GB1 does not affect our results. In addition, the 24 residues that do not appear to exchange in the discrete, quenched measurements are completely exchanged by the initial time point in the serial measurements, consistent with our conclusion about quench labeling.

3.2.3 *In vitro* crowding by BSA and lysozyme

To assess the effect of individual protein crowders[42,111] on GB1 stability, we acquired exchange data in 100 g/L solutions of either bovine serum albumin (BSA) or lysozyme (Section 3.4.7). Of the 17 common GB1 residues quantified in cells and in buffer, 13 yielded measurable rates in BSA; the others exchanged too quickly (Table 3.3). The corresponding opening free energies, $\Delta G^{o\prime}_{op,BSA}$, were compared to those obtained in buffer and in cells (Figure 3.6). BSA destabilizes GB1 compared to dilute solution, whereas the protein is stabilized in cells. In lysozyme, exchange rates are so large that the backbone N-H signals have completely decayed by the first acquisition (~20 min). We conclude that lysozyme destabilizes GB1 by >1 kcal/mol compared to buffer alone.

3.2.4 Exchange with the I6L variant

Recently, we used a thermodynamic cycle comprising k_{obs} measurements of residues in chymotrypsin inhibitor 2 and a variant, in buffer and in reconstituted cytosol, to confirm the EX2 mechanism.[45] We repeated this strategy with the destabilized GB1 variant, I6L.[112] Comparisons of $\Delta G^{o\prime}_{op,buff,I6L}$ values were possible for 12 residues (Tables 3.4, 3.5). We made three comparisons. First, we calculated the change in residue-level stability of the variant in cells compared to buffer

49

($\Delta\Delta G^{\circ\prime}_{op,cell,I6L} = \Delta G^{\circ\prime}_{op,cell,I6L} - \Delta G^{\circ\prime}_{op,buff,I6L}$). Second, we calculated the change caused

by the mutation ($\Delta\Delta G^{\circ\prime}_{op,mut} = \Delta G^{\circ\prime}_{op,I6L} - \Delta G^{\circ\prime}_{op,wt}$) in cells. Third, we calculated the

effect of the mutation in dilute solution (Figures 3.7 and 3.10). We use these data to

assess the thermodynamic cycle in the Discussion.

3.2.5 Calorimetry

We used differential scanning calorimetry to quantify the free energy of

denaturation, $\Delta G^{\circ\prime}_{den}$ (Section 3.4.8). Due to the high thermal stability of the wild-type

protein (T_m = 79.0 °C at pH_{corr} 7.6), it is difficult to obtain adequate post-transitional

baselines for robust fitting while maintaining reversibility.[113] To solve this problem,

we used the calorimetric enthalpy, $\Delta H^{\circ\prime}_{cal}$, of the destabilized I6L variant and the T_m

values of the two proteins to calculate $\Delta\Delta G^{\circ\prime}_{den,mut}$ with the equation, $\Delta\Delta G^{\circ\prime}_{den,mut} =$

$\Delta H^{\circ\prime}_{cal}(T_{m,I6L} - T_{m,wt})/T_{m,I6L}$.[23] The value of $\Delta\Delta G^{\circ\prime}_{den,mut}$ (-0.68 ±0.06 kcal/mol) is

consistent with the average $\Delta\Delta G^{\circ\prime}_{op,mut}$ (-0.8 ±0.1 kcal/mol) from dilute solution H/D

exchange experiments.

3.3 Discussion

We obtained, in triplicate, backbone amide exchange rates in cells and buffer

for 18 residues distributed throughout GB1: **Y3, K4, L5, I6** and L7 in β1; **T18** in β2;

A26, E27, K28, V29, **K31,** Y33 and **A34** in the α-helix; **T44** and **D46** in β3; and **T51,**

F52 and **T53** in β4. The set includes 14 (in bold) of the 16 residues suggested to

exchange *via* global unfolding (*i.e.*, $\Delta G^{\circ\prime}_{op} \approx \Delta G^{\circ\prime}_{den}$) in dilute solution[93] (the other two,

F30 and V54, are unassigned). We obtained k_{obs} values under all four conditions (wt

protein and I6L variant in buffer and in cells) for the 12 underlined residues.

To interpret the effect of the intracellular environment, we must first determine

the meaning of the opening free energies. If they arise from globally exchanging

residues, then we expect constant values of $\Delta G_{op}^{o\prime}$ and $\Delta\Delta G_{op,cell}^{o\prime}$ across the primary

structure (Figures 3.4 and 3.9). Although there is deviation, the range is <1 kcal/mol.

For this reason we remain cautious about over interpreting these deviations because

data from 20 proteins[114] show that global unfolders yield $\Delta G_{op}^{o\prime}$ values within 1

kcal/mol of $\Delta G_{den}^{o\prime}$ obtained from thermal or co-solute denaturation. In addition, if

some residues exchange by local unfolding and others by global unfolding we might

expect a correlation between $\Delta G_{op}^{o\prime}$ and $\Delta\Delta G_{op,cell}^{o\prime}$, but this is not the case. In

summary, we believe these residues exchange through global unfolding or high

energy fluctuations that are energetically indistinguishable from global unfolding

under physiological conditions. With this caveat in mind, we suggest two additional

sources of deviation. First, although most evidence points to equilibrium two-state

folding of GB1,[96-98] there is evidence of complex kinetic pathways, so we cannot rule

out the possibility that intermediates may be populated at equilibrium in cells or in

buffer. Second, despite our knowledge that k_{int} values do not change under crowded

conditions,[115] deviations could arise because intrinsic rates are derived from model

peptides, not the specific primary structure of GB1.

3.3.1 Exchange mechanism in buffer

As stated in the Introduction, to convert k_{obs} values to $\Delta G_{op}^{o\prime}$ values, the test

protein must be stable (*i.e.,* $k_{cl} >> k_{op}$) and k_{int} must be rate determining (the EX2

limit). GB1 is highly stable in dilute solution,[88,89] and intrinsic exchange rates in

buffer are known.[29,101] Proof that intrinsic exchange is rate determining for GB1 in

buffer comes from two sources. First, stopped-flow measurements provide a lower

limit of $\sim 10^3$ s^{-1} for k_{cl},[90] whereas k_{int} values are $< 10^2$ s^{-1}. In addition, H/D

measurements at two pH values can be used to assess the exchange mechanism,

because intrinsic exchange is base catalyzed above pH 4.[116] Specifically, if k_{int} is

rate determining, changing the pH by one unit should change k_{obs} by a factor of 10.

Consistent with this idea, a plot of log k_{obs} *versus* log k_{obs} for GB1 residues in buffer

at pH 7.6 and 6.7 (Figure 3.12) has a slope of 0.9 ±0.1 and an intercept (-1.1 ±0.3)

equal to the difference in pH. Lastly, because our exchange experiments yield data

for residues involved in global unfolding, the $\Delta G_{op}^{o\prime}$ values should approximate $\Delta G_{den}^{o\prime}$

from calorimetry. This approximation holds for both wt GB1 and the I6L variant

(Section 3.4.8).

3.3.2 Exchange mechanism in cells

The fact that the ^{15}N-^{1}H HSQC spectrum of GB1 can be overlaid with the

spectrum in buffer[80] indicates the protein is stable in cells. Further, k_{int} values do not

change significantly in reconstituted cytosol compared to buffer.[115] Unfortunately, we

cannot test the requirement that k_{int} is rate determining by changing the pH because

we cannot accurately manipulate the intracellular pH. For this reason we turned to

the thermodynamic cycle,[45] mentioned in Results (Figure 3.10). Briefly, if k_{int} is rate determining in cells, the change in $\Delta G_{op}^{o'}$ caused by a mutation ($\Delta\Delta G_{op,mut}^{o'}$) should be the same in buffer and cells (Figure 3.7). Of the 12 comparable residues, three (K4, T51, T53) of the $\Delta\Delta G_{op,mut}^{o'}$ values agree within one standard deviation of the mean, and another three (Y3, A26, A34) within two standard deviations. The remaining six residues (K28, V29, K31, T44, D46, F52) differ by more than two standard deviations. We were puzzled that not all the residues satisfied the condition $\Delta\Delta G_{op,mut,buff}^{o'} \approx \Delta\Delta G_{op,mut,cell}^{o'}$, because, as discussed above, all 12 residues are exposed only on global unfolding and possess similar protection factors. We hypothesize that the exceptions arise because the thermodynamic cycle neglects the possibility that mutations introduce new interactions (with respect to wild-type) between GB1 and the cytoplasm that are absent in buffer. We provide additional support for this hypothesis in Chapter 4.

To provide further, albeit indirect, evidence that we are measuring free energies, we estimated the effect the intracellular environment would need to impose on GB1 to move exchange to the EX1 limit ($k_{obs} \approx k_{op}$). Using a k_{cl} of 10^3 s^{-1} from Alexander et $al.$[90] and the average value of $-RT \ln\left(\frac{k_{obs}}{k_{int}}\right)$ for the 17 residues quantified in cells, the cytoplasm would have to decrease k_{op} $10^2 - 10^3$-fold and decrease k_{cl} by an order of magnitude compared to dilute solution to force exchange into the regime where k_{cl} is rate determining. Such drastic effects are unlikely and have never been observed in cells.[54-57] In summary, the data are consistent with the assumption that we are measuring free energies of opening in cells.

3.3.3 GB1 structure in cells

Although the folding kinetics[54-57,106] and equilibrium thermodynamic stability[52,54-56,58] of globular proteins can be influenced by crowding, their tertiary structures should remain unchanged[106,109,117] because the packing densities of globular proteins approximate those for ideal packing of hard spheres.[17] As discussed above, the ability to overlay the in-cell spectrum with that from dilute solution is consistent with this expectation.

Furthermore, the exchange data show similar patterns along the primary structure in both cells and buffer, supporting the conclusion that the tertiary structure is unchanged. More specifically, the pattern of solvent accessible surface area (SASA) along the sequence has an approximate inverse relationship with $\Delta\Delta G_{op}^{o\prime}$ (Figure 3.4). The average SASA for the 17 residues with quantifiable exchange rates (excluding T18), the seven residues that decay too quickly, and the 24 quench-labeled residues are 0.6 ±0.3 $Å^2$, 1.3 ±0.9 $Å^2$, and 2 ±1 $Å^2$, respectively, where the uncertainties are the sample standard deviations. Hence, the quench-labeled and rapidly exchanging residues are more likely to have larger solvent exposure. These exposed backbone amide nitrogen atoms are found in loops, the outer strands of the four-stranded sheet (β2 and β3), and the ends of the helix. The observations about SASA are consistent with dilute solution studies of GB1 structure,[88,92] dynamics[91] and H/D exchange.[93] We conclude that the cellular interior does not change the structure of GB1 compared to buffer.

3.3.4 Cellular environment and GB1 stability

The cytoplasm of *E. coli* stabilizes GB1 residues by 0.43 ±0.06 to 1.14 ±0.05 kcal/mol compared to buffer at the same pH and temperature (Figures 3.4 and 3.5). Recent advances in both the experimental and theoretical aspects of macromolecular crowding allow this stabilization to be rationalized in terms of the properties of GB1 and the *E. coli* cytoplasm.

The net effect of macromolecular crowding arises from the relative effects of hard-core repulsions, which are always stabilizing, and chemical interactions, which may be stabilizing or destabilizing.[46,86] GB1 has a pI of 4.8, similar to that of the majority of *E. coli* proteins,[118] and a net charge of -4 at pH 7.6. These properties are expected to result in a large number of charge-charge repulsions in cells. Indeed, it has been suggested that these repulsions are what allow GB1 to tumble freely in the cell and yield high quality in-cell ^{15}N-^{1}H HSQC spectra.[80,81,109,110,119,120] These soft repulsive interactions enhance the volume excluded by hard interactions in cells, thus favoring the compact native state and resulting in the observed stabilization. Our results, together with those from others[52,53,55,56,58] show that protein stability in cells can be increased, decreased, or unaffected compared to buffer alone, demonstrating that physiologically-relevant crowding effects are context dependent, with the type and strength of quinary interactions[41] playing a key role.

3.3.5 Effect of protein crowders *in vitro*

Contrary to the stabilization of GB1 in cells, individual protein crowders destabilize the protein compared to buffer alone (Figure 3.6). GB1 is destabilized to

such an extent in 100 g/L lysozyme that quantification was not possible. This destabilization can be understood by the prevalence of weak, attractive interactions between positively charged lysozyme (pI = 11.3) and anionic GB1. The attractive interactions are destabilizing because the unfolded state possesses more reactive surface than the folded state, lowering the free energy of the denatured state ensemble relative to the native state.

A similar explanation for destabilization by BSA is less straightforward. Based on our rationale for in-cell stabilization, we expected stabilization of GB1 in 100 g/L BSA (pI = 4.7) compared to buffer alone because both GB1 and BSA have anionic surfaces. However, this destabilization is in agreement with our observations[42,45,111,121] for chymotrypsin inhibitor 2 (CI2, pI = 6.0), supporting the hypothesis that non-specific, attractive backbone interactions can overcome charge-charge effects and hard-core repulsions. Moreover, the fact that cells are not crowded with only one protein complicates such simplistic comparisons. Nevertheless, the effect of BSA and other protein crowders can be rationalized *via* Zhou's realization that despite the presence of stabilizing, repulsive soft interactions between a test protein and a crowder, there exists a temperature above which crowding will be stabilizing.[50] Given our data, we expect a crossover above 37 °C for the BSA-GB1 pair, which is reasonable because the crossover for CI2-BSA is 37 °C.[50]

3.3.6 Contributions to protein stability in cells

Recent work has shown that the effects of macromolecular crowding on globular protein stability depend on the nature of the crowder.[44,45,51,121] Synthetic polymers tend to act as inert spheres and are stabilizing. Physiologically relevant crowders (*e.g.* proteins, cytoplasm) modulate the hard-core effect through soft chemical interactions: attractive forces between the crowder and test protein favor destabilization, and repulsive interactions enhance stability. However, this idea may be too simple, as indicated by our observation on the effect of BSA and the temperature dependence of crowding discussed by Zhou.[50] Another complication is the role the cell has in modulating stability *via* compartmentalization, as highlighted by the Gruebele group.[56,58] While the present study enriches our knowledge of the forces stabilizing proteins under native conditions, more studies are necessary to bring a comprehensive understanding of the effects of cellular crowding on protein stability.

3.4 Materials and methods

Unless otherwise stated, pH readings are uncorrected for the deuterium isotope effect.[122] Intrinsic rate constants from the online program SPHERE[101] were calculated for exchange at 37 °C and pH 7.2. Experiments were performed in triplicate. Uncertainties are the standard deviation of the mean.

3.4.1 Vector

The pET11a plasmid containing the gene encoding T2Q GB1 was provided by the Spicer lab at Duke University. The T2Q mutation prevents N-terminal

deamidation,[94] and we refer to this form as wild-type (wt). The I6L variant was produced by site-directed mutagenesis (QuikChange; Agilent) with the following primers: forward 5' CAT ATG CAG TAC AAG CTT <u>CTG</u> CTG AAC GGT AAA ACC C 3', reverse 5' G GGT TTT ACC GTT CAG <u>CAG</u> AAG CTT GTA CTG CAT ATG 3', where the isoleucine to leucine mutation is underlined. The DNA sequence of coding region was confirmed by chain termination sequencing[123] at Eton Bioscience, Inc.

3.4.2 In-cell H/D exchange: growth, protein expression and initiation

The plasmid encoding GB1 was transformed into competent BL21 (DE3) *Escherichia coli* cells and spread onto Luria Broth agar plates containing 100 μg/mL ampicillin, followed by incubation overnight at 37 °C. A single colony was used to inoculate a 50 mL overnight culture in M9 minimal medium (50 mM Na_2HPO_4, 20 mM KH_2PO_4, 9 mM NaCl, 5 g/L $^{15}NH_4Cl$ as the sole nitrogen source and supplemented with 4 g/L glucose, 2 mM $MgSO_4$, 10 mg/L thiamine HCl, 10 mg/L biotin, 100 μM $CaCl_2$, and 100 μg/mL ampicillin). This culture was incubated overnight with shaking (New Brunswick Scientific) at 37 °C and 225 rpm. The next morning, the 50 mL culture was poured into 950 mL of supplemented M9. This 1 L culture was grown with shaking at 37 °C until the optical density at 600 nm reached 0.6. Protein expression was then induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG). After 2 h, expression was halted by adding chloramphenicol to a final concentration of 50 μg/mL. The cells were harvested for 30 min at 4,000 x *g*.

To ensure exchange into >99% D_2O, the medium was carefully decanted and the cell pellet was resuspended in 200 mL of sodium phosphate buffered saline (PBS; 8.5 mM Na_2HPO_4, 1.8 mM NaH_2PO_4, 150 mM NaCl) prepared from 99.9% D_2O and containing 50 μg/mL chloramphenicol. A timer was initiated to mark the start of exchange. The resuspended cells were gently pelleted (~2,500 x g, 10 min), the supernatant removed and the pellet carefully resuspended with a minimal volume (2 – 3 mL) of PBS in D_2O containing 50 μg/mL chloramphenicol. The cell slurry in D_2O was returned to the 37 °C shaker and equilibrated for 10 min prior to preparing the first lysate sample.

3.4.3 In-cell H/D exchange: quenched cell lysates and NMR spectroscopy

At specified times after initiation, a 500 μL aliquot was removed from the cell slurry and transferred to a 2-mL tube containing 335 mg of 0.1 mm silica beads (Lysing Matrix B; MP Biomedicals). The aliquot was pelleted *via* centrifugation at 4,000 x g at 4 °C and the supernatant carefully removed. Quench buffer (350 μL 100 mM sodium citrate, pH 3.0, 4 °C) was added and the cells lysed *via* bead vortexing[124] on a Maxi Mix II vortexer (Thermo Scientific) for 5 min at 4 °C. The lysed cells were centrifuged at 17,000 x g for 5 min at 4 °C, and the cleared, quenched lysate was transferred to an NMR tube and stored on ice prior to NMR analysis.

For wt GB1, aliquots were removed at approximately 1, 2, 3, 5, 8, 13, and 22 h after initiating exchange. For the destabilized I6L variant, four samples were prepared between 45 min and 2 h. Additional time points were not acquired for two reasons. First, the variant's stability is lower than that of wt, and thus exchange was

complete in a shorter time. Second, the expression level of the variant is lower than that of wt GB1, leading to a lower concentration of I6L GB1 in the NMR samples.

For each quenched lysate, a ^{15}N-^{1}H HSQC spectrum was acquired at 5 °C on a Varian Inova 600 MHz spectrometer equipped with a triple resonance HCN probe using sweep widths of 2500 Hz in the ^{15}N dimension and 12001 Hz in the ^{1}H dimension. Each spectrum was acquired in ~20 min and comprised 64 increments in the ^{15}N dimension with eight scans per increment.

Processing was performed with nmrPipe.[75] The N-H crosspeak volumes were obtained from NMRViewJ.[76] Backbone amide assignments are based on published work[88,125] and personal communication with the Crowley lab at NUI-Galway. Crosspeak volumes were plotted against time and fit to an exponential decay function (SigmaPlot) to obtain k_{obs} values for the assigned residues that exhibit appropriate decay profiles.

The pH of each quenched lysate was measured after spectral acquisition and ranged between pH 3 and 4 for all experiments. Typical deviations were ≤0.1 pH unit between lysates from the same experiment. The concentration of wt GB1 in quenched lysates was 1 – 2 mM based on comparing peak volumes from a lysate of a non-exchanged sample (prepared as above, but without exchange into D_2O; Figure 3.8) to volumes from a purified sample of wt GB1.

3.4.4 Protein purification for dilute solution and *in vitro* crowding studies

Isolation and purification of ^{15}N-enriched wt GB1 was based on the protocol of Lindman *et al.*[113] Cells from a 1 L growth in ^{15}N M9 minimal medium and 2 – 3 h induction were harvested and poured into 30 mL of lysis buffer (10 mM Tris-HCl, 1 mM EDTA, pH 7.5) preheated to ~85 °C, whereupon the temperature dropped to 70 – 75 °C. The sample was stirred and heated until the temperature reached 80 °C. The lysed cells were cooled on ice for 10 min and then centrifuged at 28,000 x *g* for 30 min.

The supernatant was filtered through a 0.45 μm syringe-driven unit (Millex) and purified *via* anion exchange chromatography at 4 °C on an AKTA FPLC (GE Healthcare) by using diethylaminoethyl (DEAE) cellulose resin. Buffer A (20 mM Tris-HCl, pH 7.5) was used to load the crude lysate onto the column and elute impurities. Buffer B (20 mM Tris-HCl, 1 M NaCl, pH 7.5) was used to produce a linear gradient of 0 – 400 mM NaCl. Fractions were assessed by using sodium dodecyl sulfate polyacrylamide gel electrophoresis (4 – 20% Criterion TGX gels; Biorad) with Coomassie Brilliant Blue R-250 staining. Fractions containing GB1 were pooled and concentrated for further purification by size exclusion chromatography (Superdex 75) at 4 °C with a running buffer of 20 mM potassium phosphate, 50 mM NaCl, pH 6.0. The pure fractions were pooled, dialyzed against water, frozen and lyophilized (Labconco).

Temperature induced lysis, although appropriate for wt GB1 because of its high melting temperature,[88,89] leads to poor yields of the destabilized I6L variant. For

I6L, we employed sonic dismembration. Cells from a 1 L growth in ^{15}N M9 minimal medium and 2 – 3 h induction were harvested and resuspended in 10 mL of lysis buffer. The cells were lysed by sonication [500 W dismembrator equipped with 1/8 in. tip, 15% amplitude (Fisher Scientific)] for 10 min using a 2 s on, 2 s off pulse program. The lysed cells were centrifuged for 30 min at 28,000 x g. Streptomycin sulfate (10 mg/mL) was added to the supernatant, followed by stirring at 4 °C for 30 min. The mixture was centrifuged at 28,000 x g for 30 min and the final supernatant filtered through a 0.45 μm membrane. The resulting lysate was purified as described above.

3.4.5 Mock in-cell H/D exchange

To mimic the quenched lysate protocol used for the in-cell experiments, studies on purified wt GB1 in buffer alone were performed with samples prepared with a quench step at discrete times. For these experiments, ~25 mg of pure ^{15}N-enriched GB1 was resuspended to a final pH of 7.2 in 2.0 mL of PBS prepared in 95% D_2O. The D_2O content mimics the situation in cells exchanged into 99% D_2O, where a basal proton pool exists because of rapidly exchanging sites in proteins and other cellular components. The pH corresponds to a ^{2}H-corrected pH of 7.6, the pH in *E. coli* cells.[126,127] A timer was initiated at resuspension to mark the start of exchange, and the sample was shaken at 37 °C. At approximately 5, 15, 30, 60, 90, 180, and 360 min after initiation, a 250 μL aliquot was removed, quenched with 350 μL quench buffer and transferred to an NMR tube on ice prior to NMR analysis. Data collection and processing were performed as described above for the quenched

lysate samples. The pH of each quenched dilute solution sample was measured after NMR analysis and found to be 3.7.

3.4.6   Conventional *in vitro* H/D exchange in buffer

To validate the protocol for measuring dilute solution exchange rates at discrete times, the traditional method[108] involving serial HSQC acquisitions on a single exchange sample without a quench step was performed. A 1 – 2 mM sample of wt GB1 was prepared by resuspending purified GB1 in PBS prepared in 95% D_2O, pH 7.2 and immediately transferred to an NMR tube and placed into the spectrometer at 37 °C. After equilibration and shim adjustment, serial ^{15}N-^1H HSQC spectra were collected for ~6 h (~20 min per spectrum), at which time the exchange reaction was complete. The sample was removed from the magnet, inspected to ensure no precipitate had formed, and the pH confirmed to be 7.2. Data processing was performed with nmrPipe.[75] The Rate Analysis tool within NMRViewJ[76] was used to plot crosspeak volumes against time and fit the exponential decays to obtain k_{obs} values. For the destabilized I6L variant in buffer, data acquisition was meaningful only for ~2 h after initiation of exchange. To collect a sufficient number of data points for curve fitting, the number of scans per increment was reduced to four, resulting in an acquisition time of ~10 min per spectrum.

3.4.7   Conventional *in vitro* H/D exchange under crowded conditions

To study the effect of protein crowders on wt GB1 stability, bovine serum albumin (BSA; Sigma Aldrich) and lysozyme (Sigma Aldrich), were exchanged into D_2O. Briefly, 2 – 3 g of the crowder were resuspended in 50 mL of D_2O adjusted to

pH 10. The solution was heated for >5 h. The sample was then frozen and

lyophilized, and the process repeated once. To initiate the H/D exchange

experiment, a 100 g/L protein crowder sample prepared in PBS (95% D_2O) was

used to resuspend purified ^{15}N-enriched wt GB1 to a final concentration of ~1 mM

and final pH of 7.2. The sample was quickly mixed, applied to a 0.45 μm filter, and

inserted into the spectrometer at 37 °C. After thermal equilibration and shim

adjustment, serial spectra were acquired as described above. Data processing and

analysis were performed as described for serial, dilute solution H/D exchange. In

solutions of both BSA and lysozyme, GB1 exchange was complete within 1 – 2 h.

3.4.8 Differential scanning calorimetry (DSC)

DSC experiments were performed on a MicroCal VP-DSC calorimeter and the

data analyzed in Origin 7. Protein samples (0.5 – 1 mM) were prepared in the same

buffer as dilute solution exchange experiments. An upward scan from 20 °C to 95 °C

at a scan rate of 60 °C/h was performed, followed by cooling to 20 °C. A second

upward scan to 95 °C was performed to assess reversibility. Baselines were

corrected by subtracting the buffer/buffer scan from the protein/buffer scan.

Experiments were performed in triplicate with the standard deviation of the mean

indicated (except for measurements of $\Delta\Delta G^{o'}_{den,mut}$, where the uncertainty is from

propagation of error analysis).

T_m values were measured at pH_{corr} 7.6 for wt (79.0 ±0.2 °C) and the I6L

variant (75.6 ±0.2 °C). The free energy of denaturation for I6L ($\Delta G^{o'}_{den,I6L}$ = 6.30 ±0.01

kcal/mol) was extrapolated to 37 °C using Equation 4 from the work of Becktel and

Schellman[23] and the value of ΔC_p from the work of Alexander, *et al.*[89] The ratio of the

van't Hoff enthalpy to the calorimetric enthalpy was unity for the unfolding transition

of the variant, consistent with a two-state folding mechanism.[16] An accurate stability

curve could not be constructed for wt GB1 due to inadequate post-transitional

baselines. However, the global change in free energy of denaturation caused by the

I6L mutation ($\Delta\Delta G_{den,mut}^{o'}$ = -0.68 ±0.06 kcal/mol) was calculated based on Equation

20 from the work of Becktel and Schellman.[23] This value is consistent with the

change in stability measured *via* NMR ($\Delta\Delta G_{den,mut}^{o'}$ = -0.8 ±0.1 kcal/mol) using the

average $\Delta G_{op}^{o'}$ values of the measurable, slowly exchanging residues in dilute

solution for wt (17 residues; $\Delta G_{op,avg}^{o'}$ = 7.01 ±0.11 kcal/mol) and I6L (12 residues;

$\Delta G_{op,avg}^{o'}$ = 6.25 ±0.09 kcal/mol).

3.4.9 *E. coli* viability

At the start and after 24 h of H/D exchange, 10^7-fold dilutions of the cell slurry

were made into Luria Bertani (LB) medium. Twenty μL were spread on LB-agar

plates containing 100 μg/mL ampicillin and incubated overnight at 37 °C (Figure

3.11). The number of colonies formed on each plate was counted and indicates

minimal loss of viability over the duration of the exchange experiments, consistent

with other studies.[128,129]
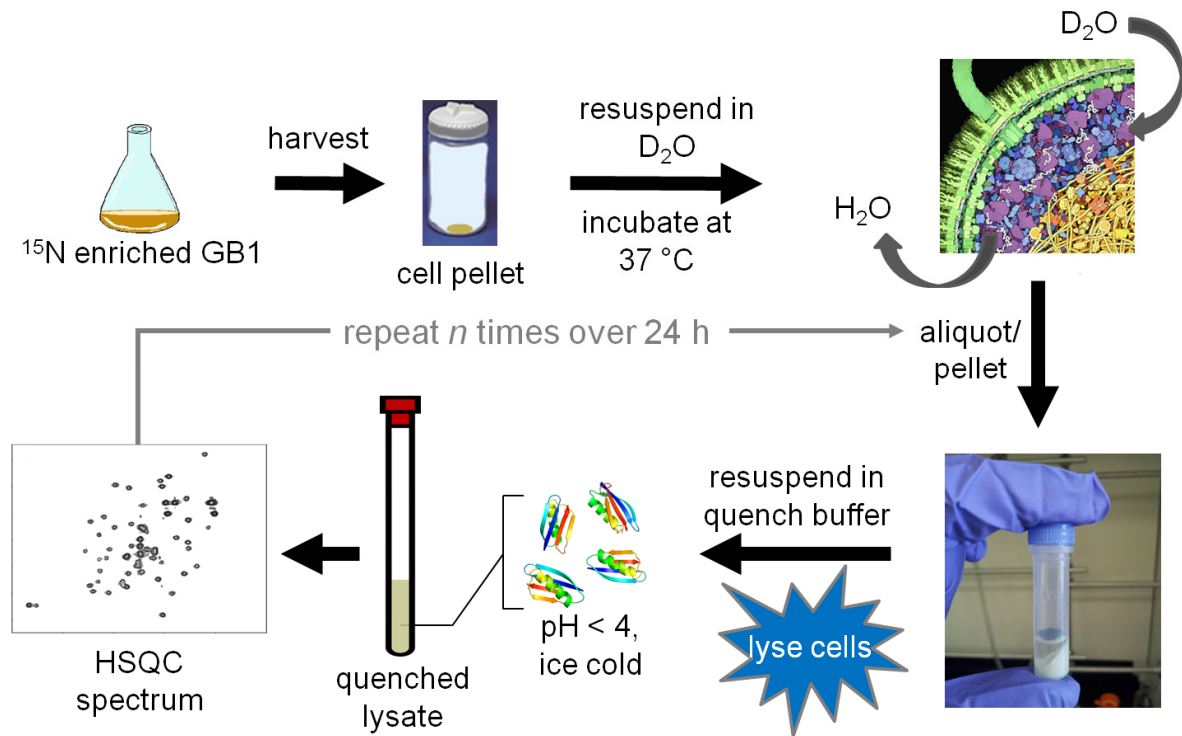
Figure 3.1 In-cell H/D exchange protocol. The cross sectional illustration of an *E. coli* cell is used with permission of David S. Goodsell (Scripps Research Institute).
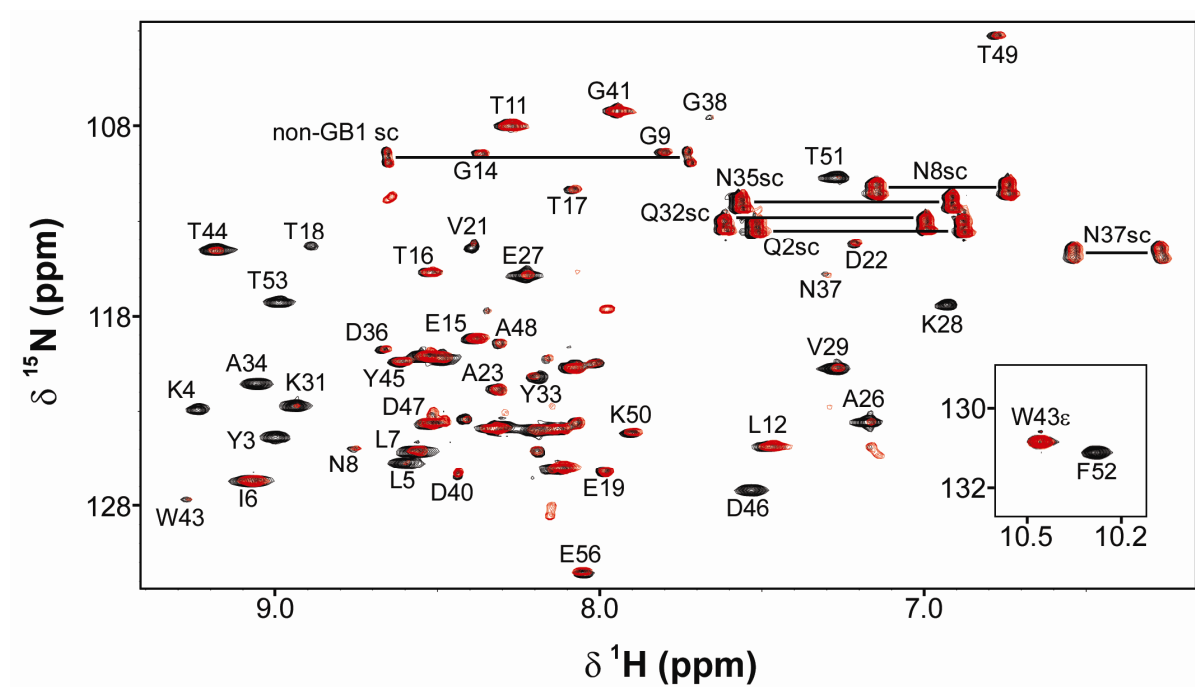
Figure 3.2 Overlaid ^{15}N-^{1}H HSQC spectra with assignments (sc; side chain) of the initial (black; 1 h exchange) and final (red; 22 h exchange) quenched lysates of an in-cell H/D exchange experiment on GB1. Assignments are based on published work[88,125] and personal communication with the Crowley lab at NUI-Galway.
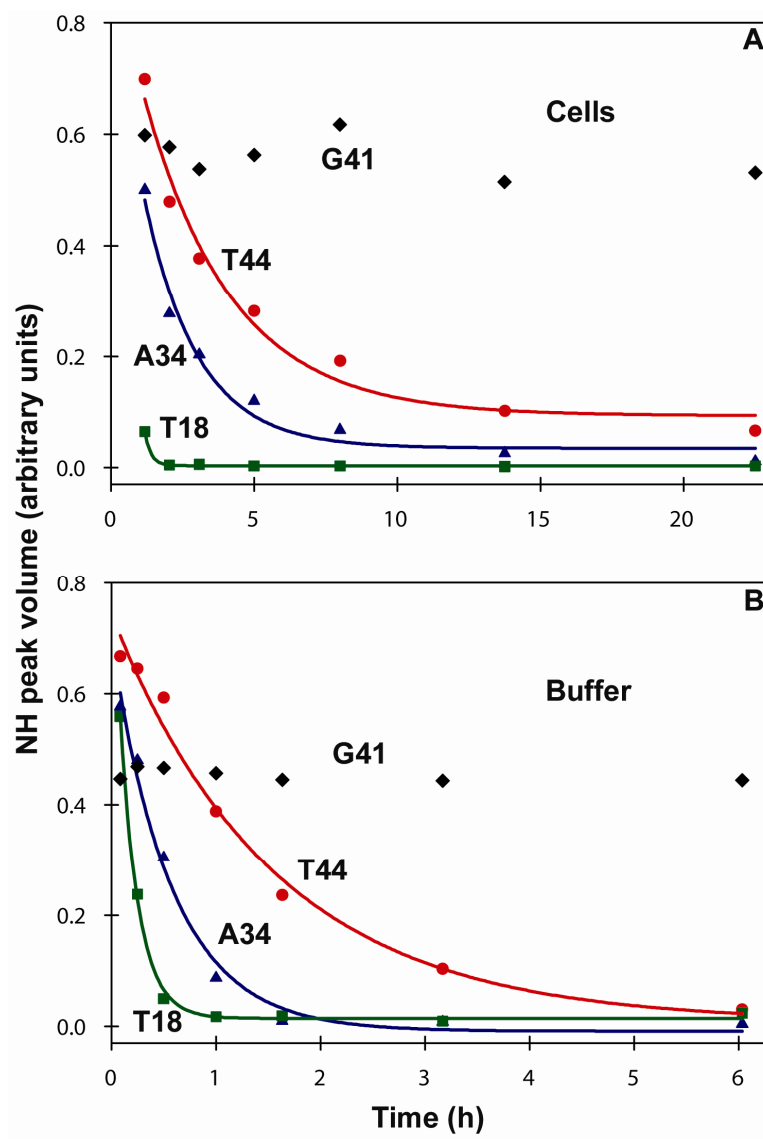
Figure 3.3 Backbone amide H/D decay profiles with corresponding best fits for T44, A34, and T18 in A) cells and in B) buffer (PBS, pH 7.6, 37 °C). Data for G41 are included to illustrate quench-labeling.
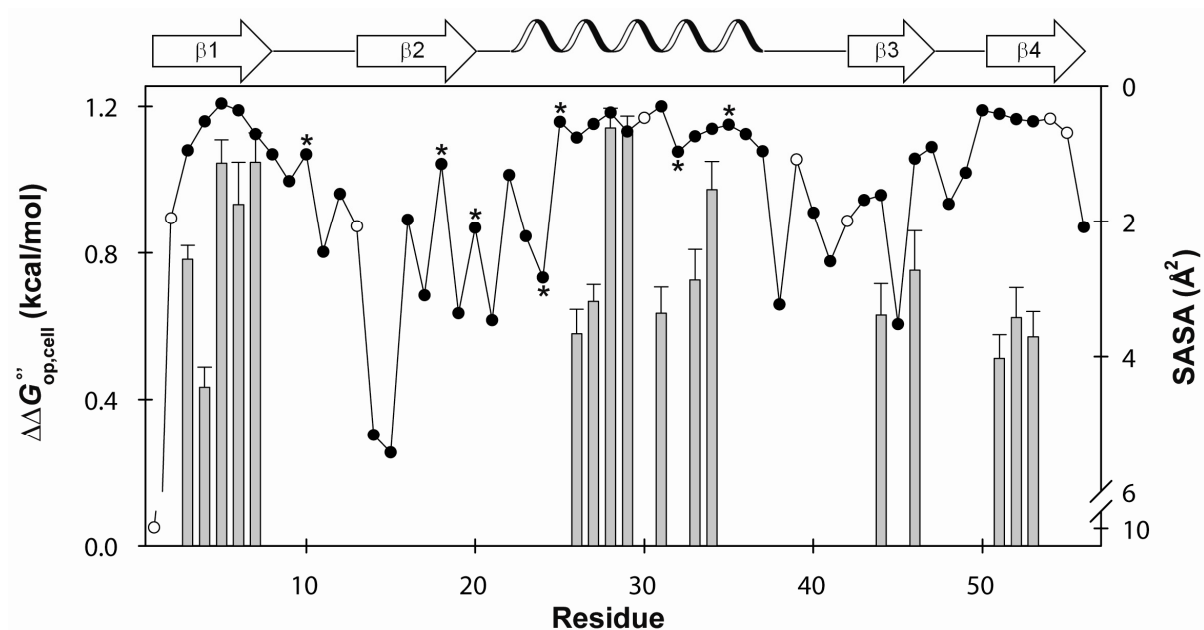
Figure 3.4 $\Delta\Delta G^{o'}_{op,cell}$ ($\Delta G^{o'}_{op,cell} - \Delta G^{o'}_{op,buff}$; left axis, gray bars) values for wt GB1

residues that give quantifiable decay rates in cells and in buffer (pH 7.6, 37 °C) and

(right axis) the solvent accessible surface area (SASA) for each backbone amide

(scatter plot). Error bars represent the standard deviation of the mean. Quench-

labeled residues are indicated by filled circles without $\Delta\Delta G^{o'}$ values. Residues that

decay too rapidly for accurate measurement are labeled with an asterisk.

Unassigned residues are shown as open circles. The SASA for each backbone

nitrogen atom was computed using the POPS* server[130] and PDB file 1PGB.
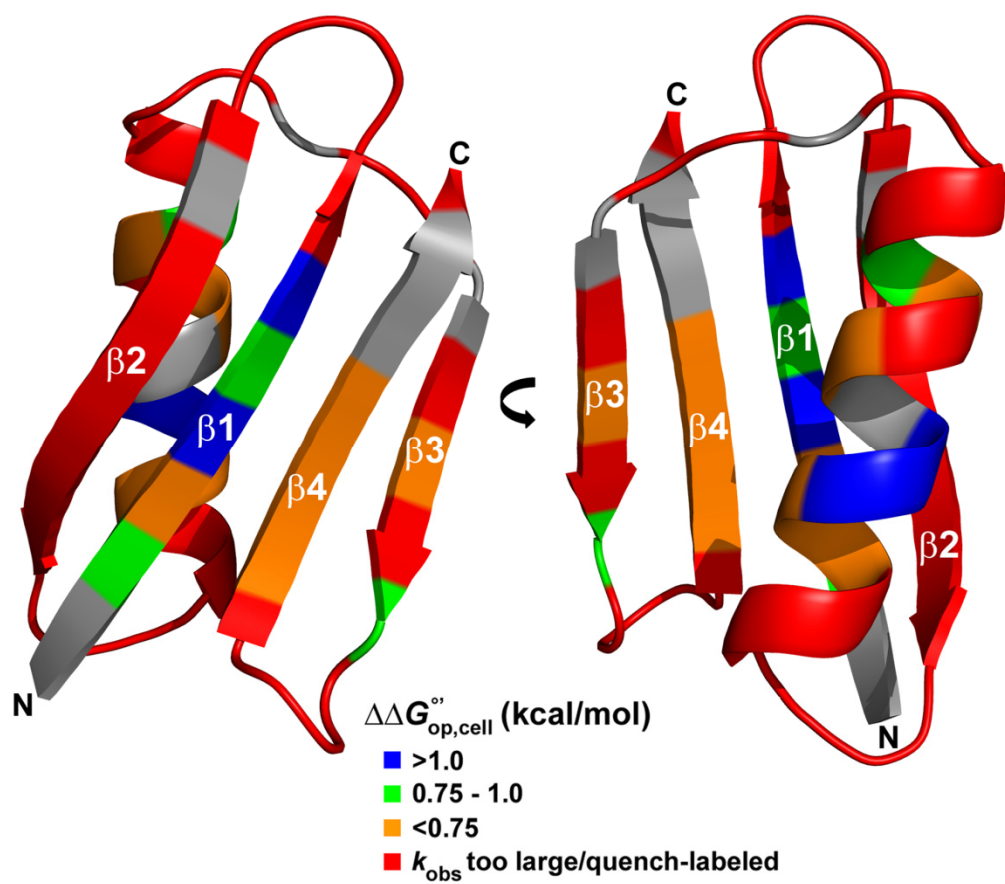
Figure 3.5 GB1 (1PGB) is stabilized in cells. Residues are colored by the magnitude of $\Delta\Delta G^{\circ'}_{\text{op,cell}}$. Gray residues are unassigned.
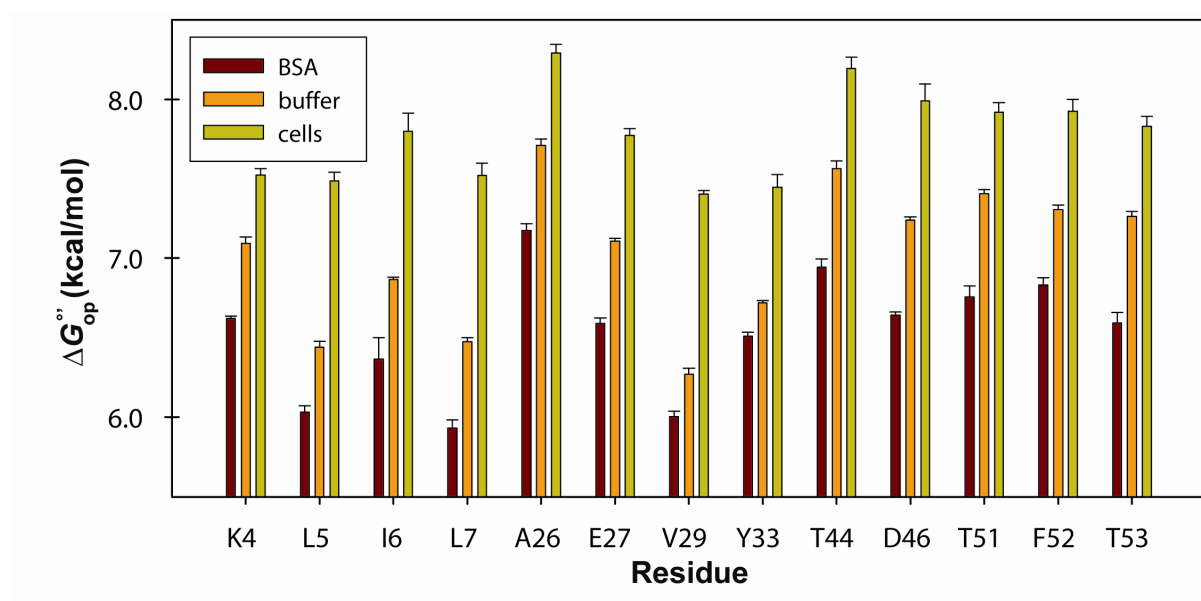
Figure 3.6 $\Delta G_{op}^{\circ'}$ values for wt GB1 residues in 100 g/L BSA, buffer and cells. Error bars represent the standard deviation of the mean from three trials.
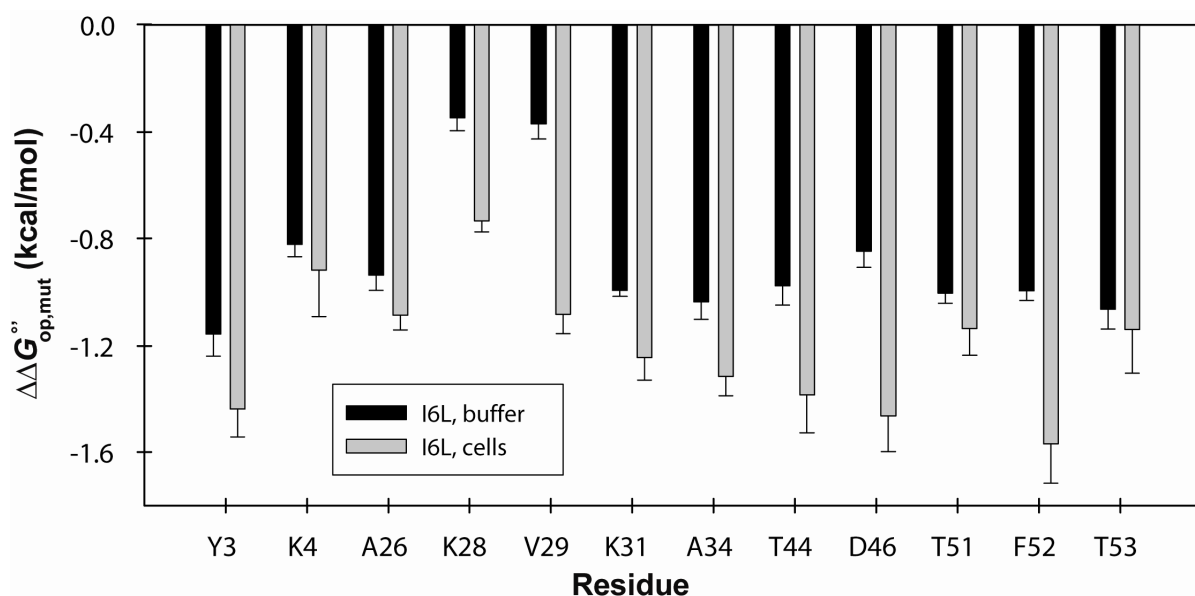
Figure 3.7 $\Delta\Delta G^{\circ'}_{op,mut}$ ($\Delta G^{\circ'}_{op,I6L}$ − $\Delta G^{\circ'}_{op,wt}$) caused by the I6L mutation in cells and in

buffer. Error bars represent the standard deviation of the mean from three trials.
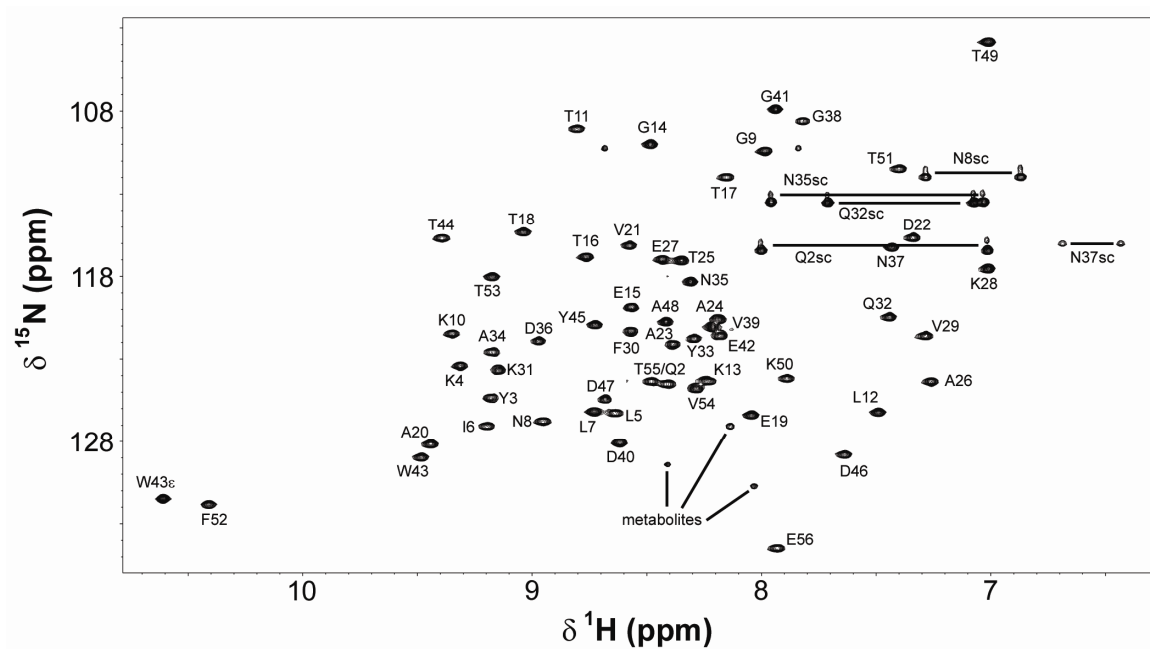
Figure 3.8 ^{15}N-^{1}H HSQC spectrum with assignments (sc; side chain) for GB1 in a non-exchange sample (90/10 v/v H$_2$O/D$_2$O) under quenched lysate conditions (100 mM sodium acetate, pH 4.7, 5 °C).
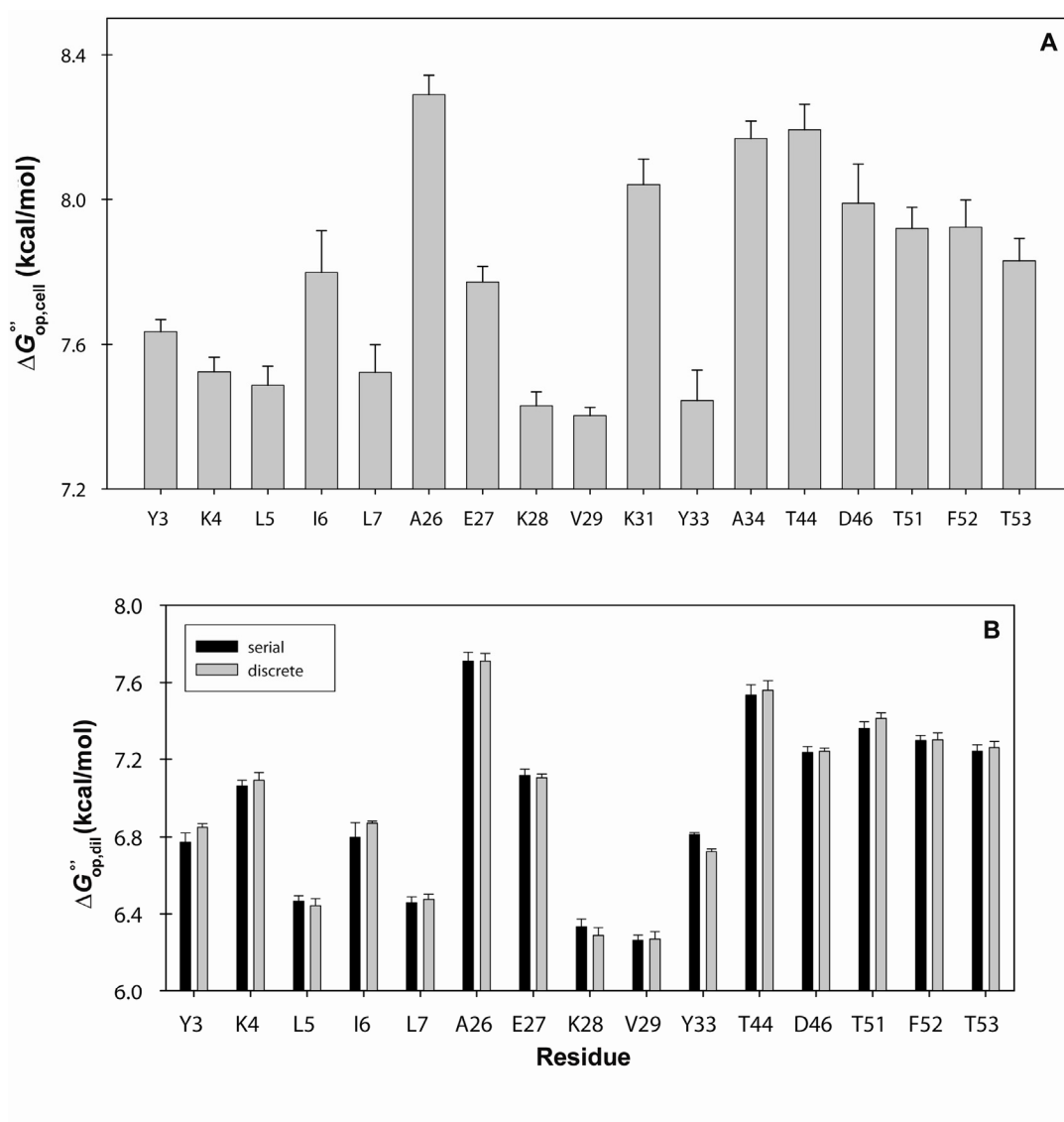
Figure 3.9 $\Delta G_{op}^{\circ'}$ values for GB1 residues that give measurable decay rates at pH

7.6, 37 °C *(A)* in cells and *(B)* in buffer from serial HSQC measurements (black bars)

and discrete *(i.e., quenched)* measurements (gray bars). Resonances from K31 and

A34 overlap in serial measurements and could not be resolved to obtain k_{obs} values.

Error bars represent the standard deviation of the mean from three trials.
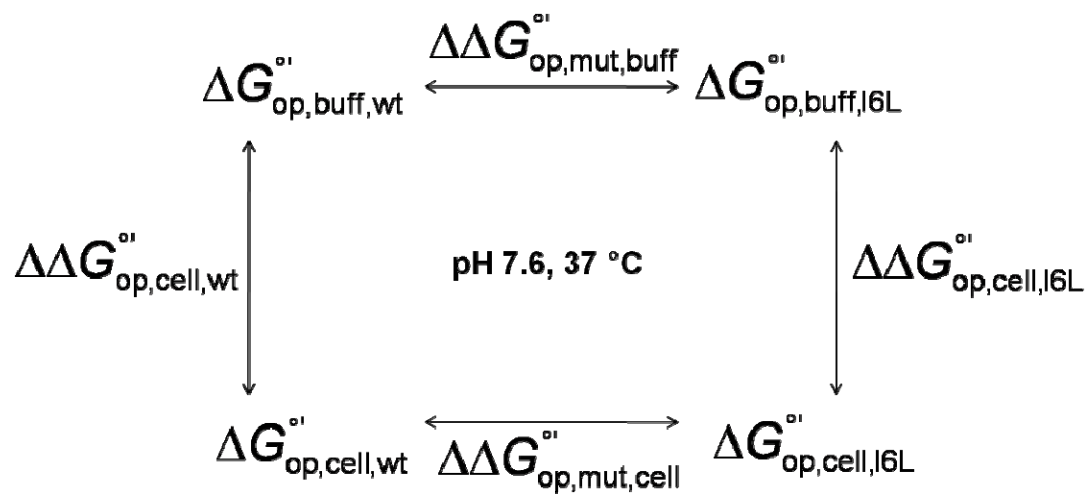
Figure 3.10 Thermodynamic cycle proposed to confirm EX2 mechanism of exchange in cells for GB1.
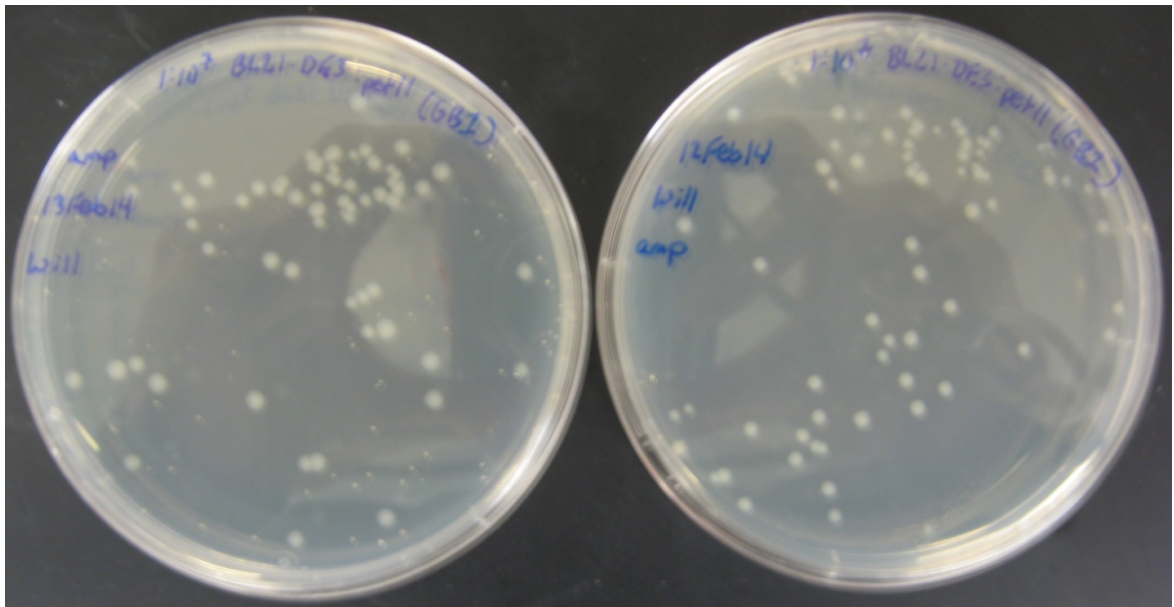
Figure 3.11 Twenty-four hour viability of *E. coli*. Dilutions of 1:10^7 were made of the cells from an in-cell exchange experiment of GB1 and 20 μL were plated at the beginning of (right plate) and 24 h after (left plate) the experiment. Eighty colony forming units (CFU) were counted on the initial plating and 67 CFU were counted on the 24-h plating.
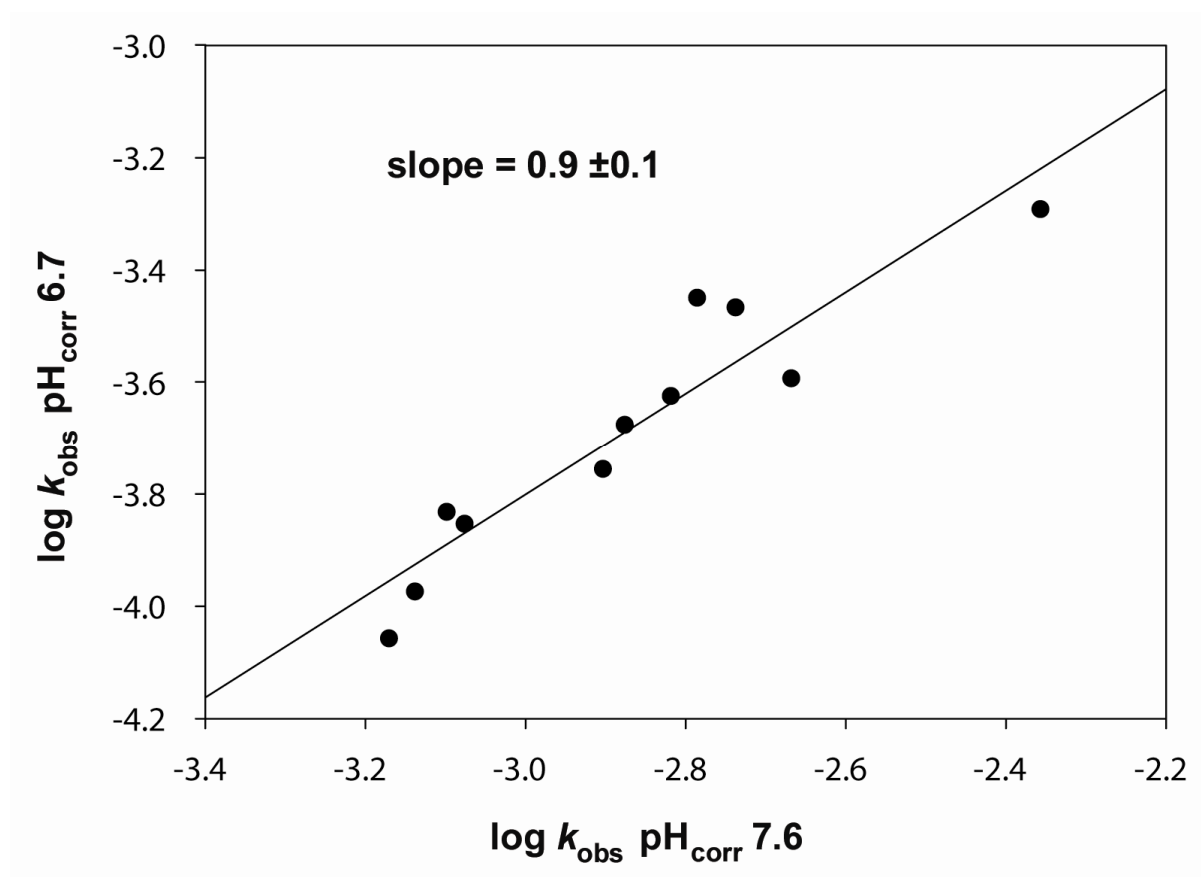
Figure 3.12 The pH dependence of H/D exchange in buffer for I6L GB1 residues. The log k_{obs} values for 10 residues were comparable and plotted against each other at pH_{corr} 7.6 and pH_{corr} 6.7. Exchange is dependent on pH in the EX2 limit, but independent of pH in the EX1 limit. A best fit line with unitary slope indicates exchange occurs in the EX2 limit, whereas a slope of zero indicates EX1 exchange. Units of k_{obs} are s^{-1}.

3.6 Tables

Table 3.1 Backbone amide proton exchange rates (k_{obs}, s^{-1}) and corresponding $\Delta G_{op}^{o'}$ (kcal/mol) values for wt GB1 in cells and in buffer (PBS, pH$_{corr}$ 7.6, 37 °C). [a]Green indicates residues for which exchange is slow enough to quantify, yellow indicates residues that decay too rapidly to quantify, blue indicates 'quench-labeled' residues (see text), red indicates ambiguous assignments in lysate. [b]SDM: standard deviation of the mean from three trials. [c]Exchange rate in cells for T18 was obtained using limited data and represents an upper limit (see text). [d]$\Delta\Delta G_{op,cell}^{o'} = \Delta G_{op,cell}^{o'} - \Delta G_{op,buff}^{o'}$. [e]From propagation of error analysis of $\Delta\Delta G_{op,cell}^{o'}$.

residue	$k_{obs,cells}$[a]	SDM$_{cells}$[b]	$k_{obs,buff}$[a]	SDM$_{buff}$[b]
M1	-		-	
Q2	-		-	
Y3	2.3E-04	±0.1E-04	8.2E-04	±0.2E-04
K4	3.3E-04	±0.2E-04	6.7E-04	±0.4E-04
L5	1.2E-04	±0.1E-04	6.5E-04	±0.4E-04
I6	2.5E-05	±0.5E-05	1.08E-04	±0.02E-04
L7	5.1E-05	±0.6E-05	2.8E-04	±0.1E-04
N8	>6.7E-04		>2.8E-03	
G9	>6.7E-04		>2.8E-03	
K10	>6.7E-04		>2.8E-03	
T11	>6.7E-04		>2.8E-03	
L12	>6.7E-04		>2.8E-03	
K13	-		-	
G14	>6.7E-04		>2.8E-03	
E15	>6.7E-04		>2.8E-03	
T16	>6.7E-04		>2.8E-03	
T17	>6.7E-04		>2.8E-03	
T18[c]	6.7E-04	±0.6E-04	2.8E-03	±0.4E-04
E19	>6.7E-04		>2.8E-03	
A20	>6.7E-04		>2.8E-03	
V21	>6.7E-04		>2.8E-03	
D22	>6.7E-04		>2.8E-03	
A23	>6.7E-04		>2.8E-03	

residue	$k_{obs,cells}$[a]	SDM_{cells}[b]	$k_{obs,buff}$[a]	SDM_{buff}[b]
A24	>6.7E-04		>2.8E-03	
T25	>6.7E-04		>2.8E-03	
A26	1.5E-04	±0.1E-04	3.8E-04	±0.2E-04
E27	6.7E-05	±0.5E-05	1.98E-04	±0.06E-04
K28	2.4E-04	±0.1E-04	1.6E-03	±0.1E-03
V29	1.04E-04	±0.04E-04	6.5E-04	±0.4E-04
F30	-		-	
K31	1.5E-04	±0.2E-04	4.1E-04	±0.1E-04
Q32	>6.7E-04		>2.8E-03	
Y33	3.2E-04	±0.4E-04	1.02E-03	±0.02E-03
A34	1.3E-04	±0.1E-04	6.3E-04	±0.6E-04
N35	>6.7E-04		>2.8E-03	
D36	>6.7E-04		>2.8E-03	
N37	>6.7E-04		>2.8E-03	
G38	>6.7E-04		>2.8E-03	
V39	-		-	
D40	>6.7E-04		>2.8E-03	
G41	>6.7E-04		>2.8E-03	
E42	-		-	
W43	>6.7E-04		>2.8E-03	
T44	7.3E-05	±0.9E-05	2.0E-04	±0.2E-04
Y45	>6.7E-04		>2.8E-03	
D46	9E-05	±2E-05	2.91E-04	±0.08E-04
D47	>6.7E-04		>2.8E-03	
A48	>6.7E-04		>2.8E-03	
T49	>6.7E-04		>2.8E-03	
K50	>6.7E-04		>2.8E-03	
T51	1.9E-04	±0.2E-04	4.4E-04	±0.2E-04
F52	1.6E-04	±0.2E-04	4.3E-04	±0.2E-04
T53	1.9E-04	±0.2E-04	4.9E-04	±0.2E-04
V54	-		-	
T55	-		-	
E56	>6.7E-04		>2.8E-03	

residue	$\Delta G^{o'}_{op,cell}$	SDM_{cells}[b]	$\Delta G^{o'}_{op,buff}$	SDM_{buff}[b]	$\Delta\Delta G^{o'}_{op,cell}$[d]	uncertainty[e]
Y3	7.63	±0.03	6.85	±0.02	0.78	±0.04
K4	7.52	±0.04	7.09	±0.04	0.43	±0.06
L5	7.49	±0.05	6.44	±0.04	1.05	±0.06
I6	7.8	±0.1	6.87	±0.01	0.9	±0.1
L7	7.52	±0.08	6.47	±0.03	1.05	±0.08

residue	$\Delta G^{o'}_{op.cell}$	SDM_{cells}[b]	$\Delta G^{o'}_{op.buff}$	SDM_{buff}[b]	$\Delta\Delta G^{o'}_{op.cell}$[d]	uncertainty[e]
A26	8.29	±0.05	7.71	±0.04	0.58	±0.07
E27	7.77	±0.04	7.11	±0.02	0.67	±0.05
K28	7.43	±0.04	6.29	±0.04	1.14	±0.05
V29	7.40	±0.02	6.27	±0.04	1.13	±0.04
K31	8.04	±0.07	7.41	±0.02	0.64	±0.07
Y33	7.45	±0.08	6.72	±0.01	0.73	±0.08
A34	8.17	±0.05	7.20	±0.06	0.97	±0.08
T44	8.19	±0.07	7.56	±0.05	0.63	±0.09
D46	8.0	±0.1	7.24	±0.02	0.8	±0.1
T51	7.92	±0.06	7.41	±0.03	0.51	±0.06
F52	7.92	±0.07	7.30	±0.04	0.62	±0.08
T53	7.83	±0.06	7.26	±0.03	0.57	±0.07

Table 3.2 $\Delta G_{op}^{o'}$ (kcal/mol) values for wt GB1 in buffer (PBS, pH_{corr} 7.6, 37 °C), obtained from serial and discrete (*i.e.*, quenched) measurements. [a]SDM: standard deviation of the mean from three trials.

residue	$\Delta G_{op.serial}^{o'}$	SDM_{serial}[a]	$\Delta G_{op.discrete}^{o'}$	$SDM_{discrete}$[a]
Y3	6.77	±0.05	6.85	±0.02
K4	7.06	±0.03	7.09	±0.04
L5	6.46	±0.03	6.44	±0.04
I6	6.80	±0.08	6.87	±0.01
L7	6.46	±0.03	6.47	±0.03
A26	7.71	±0.05	7.71	±0.04
E27	7.12	±0.03	7.11	±0.02
K28	6.33	±0.04	6.29	±0.04
V29	6.26	±0.03	6.27	±0.04
Y33	6.81	±0.01	6.72	±0.01
T44	7.54	±0.05	7.56	±0.05
D46	7.23	±0.03	7.24	±0.02
T51	7.36	±0.03	7.41	±0.03
F52	7.29	±0.03	7.30	±0.04
T53	7.24	±0.03	7.26	±0.03

Table 3.3 Backbone amide proton exchange rates (k_{obs}, s^{-1}) and corresponding $\Delta G_{op}^{o'}$ (kcal/mol) values for wt GB1 in 100 g/L BSA (PBS, pH$_{corr}$ 7.6, 37 °C). [a]SDM: standard deviation of the mean from three trials.

residue	$k_{obs,BSA}$	SDM$_{kobs}$[a]	$\Delta G_{op,BSA}^{o'}$	SDM$_{\Delta G}$[a]
K4	1.80E-03	±0.08E-03	6.62	±0.02
L5	1.58E-03	±0.02E-03	6.03	±0.04
I6	3.1E-04	±0.6E-04	6.4	±0.1
L7	8.3E-04	±0.2E-04	5.93	±0.05
A26	1.14E-03	±0.06E-03	7.17	±0.04
E27	5.7E-04	±0.4E-04	6.59	±0.03
V29	126E-05	±1E-05	6.00	±0.04
Y33	1.8E-03	±0.1E-03	6.51	±0.03
T44	6.9E-04	±0.5E-04	6.94	±0.05
D46	9.6E-04	±0.9E-04	6.64	±0.02
T51	1.59E-03	±0.08E-03	6.75	±0.07
F52	1.2E-03	±0.2E-03	6.83	±0.05
T53	1.79E-03	±0.08E-03	6.59	±0.07

Table 3.4 Backbone amide proton exchange rates (k_{obs}, s^{-1}) and corresponding $\Delta G_{op}^{o'}$ (kcal/mol) values for I6L GB1 in cells and in buffer (PBS, pH$_{corr}$ 7.6, 37 °C). [a]SDM: standard deviation of the mean from three trials. [b]$\Delta\Delta G_{op,cell,I6L}^{o'} = \Delta G_{op,cell,I6L}^{o'} - \Delta G_{op,buff,I6L}^{o'}$. [c]From propagation of error analysis of $\Delta\Delta G_{op,cell,I6L}^{o'}$.

residue	$k_{obs,cells}$	SDM$_{cells}$[a]	$k_{obs,buff}$	SDM$_{buff}$[a]
Y3	2.4E-03	±0.4E-03	3.4E-03	±0.5E-03
K4	1.6E-03	±0.4E-03	1.61E-03	±0.05E-03
L6	6E-04	±1E-04	5.8E-04	±0.5E-04
A26	8.7E-04	±0.1E-04	1.10E-03	±0.08E-03
K28	8.0E-04	±0.2E-04	1.73E-03	±0.07E-03
V29	6.1E-04	±0.7E-04	7.5E-04	±0.5E-04
K31	1.11E-03	±0.08E-03	1.30E-03	±0.03E-03
A34	1.22E-03	±0.08E-03	2.38E-03	±0.05E-03
T44	7E-04	±1E-04	6.2E-04	±0.6E-04
D46	9E-04	±1E-04	7.4E-04	±0.7E-04
T51	1.2E-03	±0.2E-03	1.42E-03	±0.06E-03
F52	2.0E-03	±0.4E-03	1.35E-03	±0.02E-03
T53	1.3E-03	±0.3E-03	1.7E-03	±0.2E-03

residue	$\Delta G_{op,cell,I6L}^{o'}$	SDM$_{cells}$[a]	$\Delta G_{op,buff,I6L}^{o'}$	SDM$_{buff}$[a]	$\Delta\Delta G_{op,cell,I6L}^{o'}$[b]	uncertainty[c]
Y3	6.2	±0.1	5.70	±0.08	0.5	±0.1
K4	6.6	±0.2	6.27	±0.02	0.3	±0.2
A26	7.20	±0.01	6.77	±0.04	0.43	±0.04
K28	6.70	±0.02	5.94	±0.03	0.75	±0.03
V29	6.32	±0.07	5.90	±0.04	0.42	±0.08
K31	6.80	±0.05	6.41	±0.02	0.38	±0.05
A34	6.86	±0.05	6.16	±0.03	0.70	±0.06
T44	6.8	±0.1	6.59	±0.05	0.2	±0.1
D46	6.52	±0.08	6.39	±0.06	0.1	±0.1
T51	6.78	±0.08	6.40	±0.03	0.38	±0.08
F52	6.4	±0.1	6.31	±0.01	0.0	±0.1
T53	6.7	±0.2	6.20	±0.07	0.5	±0.2

Table 3.5 $\Delta\Delta G^{o'}_{op,mut}$ ($\Delta G^{o'}_{op,I6L} - \Delta G^{o'}_{op,wt}$, kcal/mol) caused by the I6L mutation in cells and in buffer (PBS, pH$_{corr}$ 7.6, 37 °C). [a]From propagation of error analysis.

residue	$\Delta\Delta G^{o'}_{op,mut,cell}$	uncertainty$_{cell}$[a]	$\Delta\Delta G^{o'}_{op,mut,buff}$	uncertainty$_{buff}$[a]
Y3	-1.4	±0.1	-1.16	±0.08
K4	-0.9	±0.2	-0.82	±0.05
A26	-1.09	±0.06	-0.94	±0.06
K28	-0.73	±0.04	-0.35	±0.05
V29	-1.08	±0.07	-0.37	±0.06
K31	-1.24	±0.08	-0.99	±0.02
A34	-1.31	±0.07	-1.04	±0.07
T44	-1.4	±0.1	-0.98	±0.07
D46	-1.5	±0.1	-0.85	±0.06
T51	-1.1	±0.1	-1.00	±0.04
F52	-1.6	±0.1	-0.99	±0.04
T53	-1.1	±0.2	-1.06	±0.08

CHAPTER 4: QUINARY STRUCTURE MODULATES PROTEIN STABILITY IN CELLS[1]

4.1 Significance

The interactions stabilizing the secondary, tertiary and quaternary structure of globular proteins are well defined,[2] but little is known about quinary structure, which organizes the proteome and is key to metabolism and signal transduction.[41,131] This gap in knowledge arises because proteins function in a crowded cellular environment quite different from the dilute conditions where they are traditionally studied. Macromolecular crowding can significantly alter the biophysical landscape of proteins, including their equilibrium thermodynamic stability.[48,54-56,58] Experimental and computational efforts establish that crowding effects arise from a combination of short range (steric) repulsions and longer range interactions (referred to as 'soft' interactions in Chapter 3) between the test protein and surrounding macromolecules.[44-46,51,78] Despite a growing number of in-cell studies, details about the energetics of soft, quinary interactions remain elusive. Here we show that a surface mutation in the B1 domain of protein G (GB1) is 10-fold more destabilizing in *Escherichia coli* than in buffer, a surprising result that establishes the crucial importance of intermolecular electrostatic interactions in cells. Using a double mutant cycle,[132] we quantify quinary interactions for the first time. Remarkably, the

---

[1] The material in this chapter is being submitted for publication. If accepted, the original citation will be as follows: Monteith WB, Guzman Cisneros E, Pielak GJ (2014) Quinary structure modulates protein stability in cells. WB Monteith and GJ Pielak wrote the paper.

energetics of intermolecular interactions between the cytoplasm and protein surface are as large as those of specific protein-protein interactions.[133] Our results provide quantitative evidence that there remains much to be learned about the effects of surface mutations and post-translational modifications in natural proteins under physiologically relevant conditions. This realization will drive the challenging and critical task of implementing quinary interactions into models for understanding and manipulating the biological role of proteins.[131]

4.2 Introduction

It has been known for more than 50 years that amide proton exchange experiments can measure the free energy required to expose individual backbone amide protons to solvent.[24] For GB1 at a physiological pH of 7.6, these opening free energies, $\Delta G_{op}^{o'}$, equal $-RT\ln\left(\frac{k_{obs}}{k_{int}}\right)$, where R is the gas constant, T is the absolute temperature, k_{obs} is the observed rate of exchange, and k_{int} is the rate in an unstructured peptide.[48] Importantly, we know that the cytoplasm does not affect k_{int}. We also know that the experimentally accessible $\Delta G_{op}^{o'}$ values for GB1 at pH 7.6 approximate those required to denature the protein.[48] Therefore, such data provide a thermodynamically rigorous measure of equilibrium global protein stability.

In Chapter 3, we quantified the stability of GB1 at the residue level in living *E. coli* using NMR-detected backbone amide hydrogen/deuterium exchange experiments[30] in quenched cell lysates.[48] GB1 is stabilized in cells relative to buffer. We attributed the stabilization to repulsive electrostatic interactions between anionic GB1 (Z_{net} = -4) and the bulk anionic composition of the *E. coli* proteome.[118]

Our implementation of a thermodynamic cycle (Figure 4.1) in that study suggested the presence of intermolecular interactions between the cytosol and the I6L variant that were absent in buffer. Such thermodynamic cycles[132] have played an important role in protein chemistry since they were developed to assess the free energy of specific protein-protein interactions[132,133] and interactions between side chains in globular protein stability.[134,135] Briefly, the individual energetic effects of two single-site amino acid changes are compared to the combined effect of both mutations. If the sites interact, the sum of the contributions from the single-site changes will not equal the contribution from the double mutant. The disparity between these two values measures the strength of the interaction.

During the course of our in-cell studies, we realized that transferring a single mutant (denoted by the subscript 'var') from buffer ('buff') to cells ('cell') is analogous to making a second mutation (Figure 4.1). Discrepancies in the horizontal and vertical sides of Figure 4.1 define the coupling, or interaction, free energy ($\delta\Delta\Delta G_{\mathrm{int}}^{\circ\prime}$) associated with the combined perturbations:

$$\delta\Delta\Delta G_{\mathrm{int}}^{\circ\prime} = (\Delta G_{\mathrm{cell,var}}^{\circ\prime} - \Delta G_{\mathrm{cell,wt}}^{\circ\prime}) - (\Delta G_{\mathrm{buff,var}}^{\circ\prime} - \Delta G_{\mathrm{buff,wt}}^{\circ\prime})$$

$$= \Delta\Delta G_{\mathrm{mut,cell}}^{\circ\prime} - \Delta\Delta G_{\mathrm{mut,buff}}^{\circ\prime}$$

$$= \Delta\Delta G_{\mathrm{cell,var}}^{\circ\prime} - \Delta\Delta G_{\mathrm{cell,wt}}^{\circ\prime} \qquad [1]$$

where a negative value of $\delta\Delta\Delta G_{\mathrm{int}}^{\circ\prime}$ indicates attractive interactions introduced (relative to wt) by transferring the mutant from buffer to cells.

4.3 Results

Differences in the residue-level stability change caused by the I6L mutation ($\Delta\Delta G^{\circ\prime}_{op,mut}$) in cells and in buffer (Figure 4.2a) were used to calculate $\delta\Delta\Delta G^{\circ\prime}_{op,int}$ (Figure 4.3). The variation in $\Delta\Delta G^{\circ\prime}_{op,mut}$ and $\delta\Delta\Delta G^{\circ\prime}_{op,int}$ for a given protein across the primary structure probably arises from a convolution of two effects. First, there is an inherent uncertainty in the exact values of k_{int}, because these values come from model peptides, not GB1. Second, the variation could reflect subtle differences in the free energy required to expose the proton. At present, these factors cannot be separated.

The interaction free energies for I6L are small or negligible (Figures 4.3 and 4.4). This observation is reasonable because the substitution only shifts the position of a surface exposed, non-polar (methyl) group,[94] such that interactions with the cytoplasm are similar to those experienced in wt GB1. This result led us to test the hypothesis with a variant predicted to produce a larger effect.

As stated above, the increased stability of the wt protein in cells compared to buffer is attributed to repulsive interactions between GB1 and the cytoplasm. We reckoned that an appropriate charge reversal might reduce the repulsions, *decreasing* the stability in cells, but maintain wt stability in buffer. We chose D40 for two reasons. First, it is the only acidic residue in a surface exposed loop. Second, it lacks intraprotein side chain hydrogen bonds.[88] As predicted, the D40K mutation, which decreases the net negative charge of the protein by two units, has a nearly negligible effect in buffer (Figure 4.2b, average $\Delta\Delta G^{\circ\prime}_{op,mut,buff}$ of 0.11 ±0.08 kcal/mol)

but a 10-fold larger destabilizing effect in cells (average $\Delta\Delta G^{o'}_{op,mut,cell}$ of -1.1 ±0.2

kcal/mol). Thus, a mutation with an innocuous effect in buffer experiences new and

significant interactions in cells. These quinary interactions in cells stabilize both the

native state and the denatured state ensemble of D40K compared to those of wt

(Figure 4.5). However, the denatured state is more stabilized because the compact

native state provides less available surface for charge-charge interaction, resulting in

net destabilization of D40K in cells (Figures 4.6 and 4.7).

Most importantly, the interaction energies of the D40K and the I6L mutants

are strikingly different (Figures 4.3 and 4.4). The average interaction free energy

caused by the I6L change is -0.3 ±0.2 kcal/mol, whereas the average value for D40K

is -1.2 ±0.1 kcal/mol (Figure 4.3). To gain further support for our ideas, we studied

the D40N variant, which changes the charge by one unit. As expected, its interaction

free energies are between those of I6L and D40K (data not shown).

We attempted to measure interaction free energies with two other charge

reversals of surface exposed side chains, E19K and E56K, but their stabilities in

cells and buffer were too low for quantification. The most likely reason for their

instability is that these side chains are known to form intramolecular hydrogen

bonds.[92] This observation highlights the necessity for using an 'uninvolved' (at least

in buffer) residue, such as D40, to quantify interaction free energies.

The most important assumption in our analysis is that the structures of the

native state and denatured state ensemble of the two proteins are unchanged in

buffer and in cells. This assumption appears to be valid because patterns of H/D

exchange data in buffer and in cells are similar (*i.e.*, protection factors are consistent) and the HSQC spectra in quenched cell lysates and buffer can be overlaid with the GB1 in-cell spectra.[48] In addition, the presence of molecular chaperones has been suggested to confound equilibrium measurements of protein stability. Given the availability of chaperones in cells, however, this argument appears to be valid for only a small portion (≤5%) of the proteome.[136] Small, rapid folders such as GB1 most likely reach the native state without assistance in *E. coli*.

4.4 Discussion

Despite several initial observations,[7,8] electrostatic interactions were generally dismissed as important players in protein stabilty.[9] This situation eventually led to the recognition of the hydrophobic effect[11] as the dominant stabilizing factor.[2] Since Richards' demonstration that the core of a native globular protein is as well packed as crystals of small organic compounds,[17] the focus has been put squarely on intramolecular interactions in the core. These ideas were synthesized by Lattman and Rose[137] who concluded that changes in the core alter the stability of a protein, but not its fold.

In the present work, our ability to assess protein stability in cells leads us to reassess the role of electrostatic interactions in modulating globular protein stability. Both large and small stability changes have been observed for *intramolecular* ion-pair interactions in buffer,[138,139] but our findings demonstrate that such conclusions are not necessarily applicable inside cells, where *intermolecular* interactions between macromolecules abound. Furthermore, the effect can be large; the D40K

90

interaction free energy (~1 kcal/mol) accounts for nearly one-fifth of the total stability (~7 kcal/mol), and is in the range of interaction free energies observed in *specific* protein-protein complexes.[133] The implications for larger proteins with multiple sites akin to D40 in GB1, and for the effects of charge altering post-translational modifications required for signal transduction (*e.g.*, phosphorylation, acetylation, myristoylization, sulfation), cannot be ignored. We posit that these surface residues are as important as core residues to folding and stability, but this role can only be recognized by studies in native cellular environments. We are not, however, questioning the key role of the hydrophobic effect and core packing because, as stated above, the structure of all the proteins studied appears to be the same in buffer and in cells.

4.5 Conclusions

It is useful to cast our discussion in terms of Anfinsen's thermodynamic hypothesis, which states that "the native conformation is determined by the totality of interatomic interactions…in a given environment. [T]his idea emphasize[s] the fact that a protein molecule only makes stable, structural sense when it exists under conditions similar to those for which it was selected – the so-called physiological state".[5] The emphasis on the 'physiological state' has, until recently, been ignored; nearly all studies have been conducted with purified protein in simple buffered solutions. We, and others, are beginning to fill this void by showing the differences between folding in buffer and folding in cells.[48,54-56,58] Here, we have taken the next step by demonstrating that non-specific, quinary interactions with the cytoplasmic milieu can be modulated in a way that alters protein stability, whereas the stability is

essentially unchanged in buffer. Hence, intermolecular interactions can significantly impact folding and stability in cells, revealing a new role for surface residues that will aid in understanding native protein function. Although such interactions are precisely stated in Anfinsen's hypothesis, they have been hitherto unrecognized because we have been looking in the wrong place: buffer, instead of cells.

4.6 Materials and methods

Unless otherwise indicated, pH readings are uncorrected for the isotope effect.[122] Intrinsic rates were calculated at 37 °C, pH 7.2 using SPHERE.[101] Observed exchange rates were converted to equilibrium unfolding free energies as described in Chapter 3.[48] Tables of rates and free energies are found in section 4.8 and Chapter 3.

4.6.1 Vector

The pET11a plasmid containing the T2Q GB1 (wt) gene and its I6L variant were described in section 3.4.1. The D40K variant was produced by site-directed mutagenesis (QuikChange; Agilent) with the following primers: forward 5' C GAC AAC GGT GTT <u>AAA</u> GGT GAA TGG ACC 3', reverse 5' GGT CCA TTC ACC <u>TTT</u> AAC ACC GTT GTC G 3' (mutation underlined).

4.6.2 Protein purification

Isolation and purification of ^{15}N-enriched wt GB1 and its variants was described in section 3.4.4. However, the D40K variant does not bind to the ion exchange column owing to a less negative charge (Z_{net} = -2 compared to Z_{net} = -4 for

wt and I6L GB1). Fractions containing D40K GB1 were collected in the wash step of

anion exchange chromatography and further purified *via* size exclusion

chromatography as detailed in section 3.4.4.

4.6.3 In-cell H/D exchange

The protocol for measuring H/D exchange in *E. coli* with quenched cell

lysates was reported in sections 3.4.2 and 3.4.3. For wt GB1, aliquots were removed

approximately 1, 2, 3, 5, 8, 13, and 22 h after initiating exchange. Four samples

prepared between 45 min and 3 h were sufficient to capture decay profiles of the

destabilized I6L and D40K variants.

4.6.4 *In vitro* H/D exchange

Wildtype GB1 exchange rates measured using a quench step with discrete

samples (to mimic the in-cell protocol) yield the same values as those measured

with the traditional method[108] involving serial HSQC acquisitions on a single

exchange sample. Consequently, the serial method was employed here.

Experimental details and data analysis are found in section 3.4.6.

4.6.5 DSC

T_m values were measured at pH_{corr} 7.6 for wt (79.0 ±0.2 °C), the I6L variant

(75.6 ±0.2 °C) and the D40K variant (80.2 ±0.1 °C) as described in section 3.4.8.

Errors are the standard deviation of the mean from three trials.
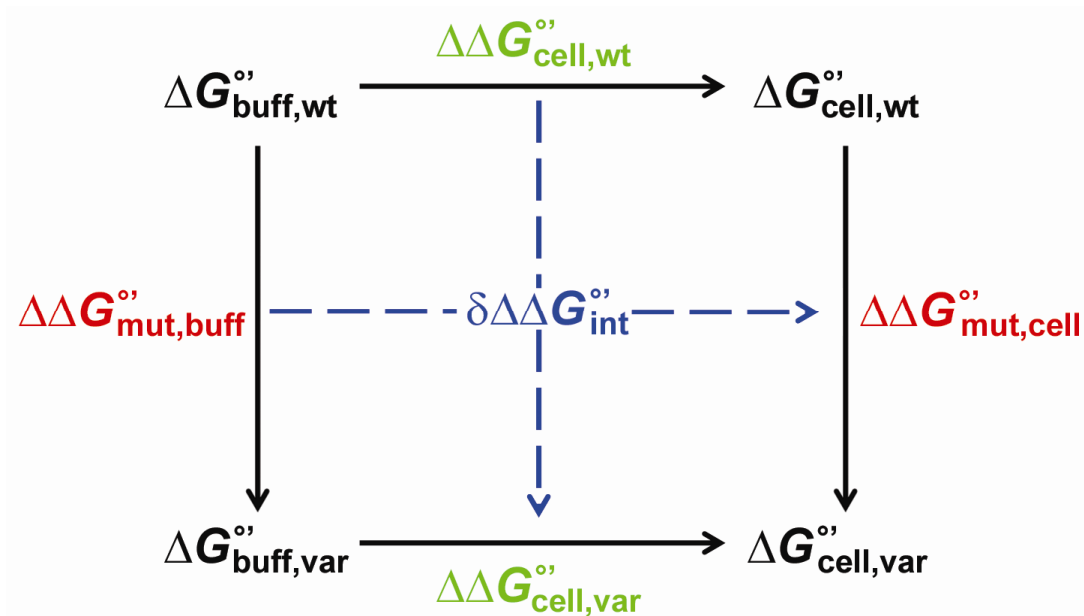
Figure 4.1 Thermodynamic cycle quantifying intracellular interactions introduced by mutating GB1.
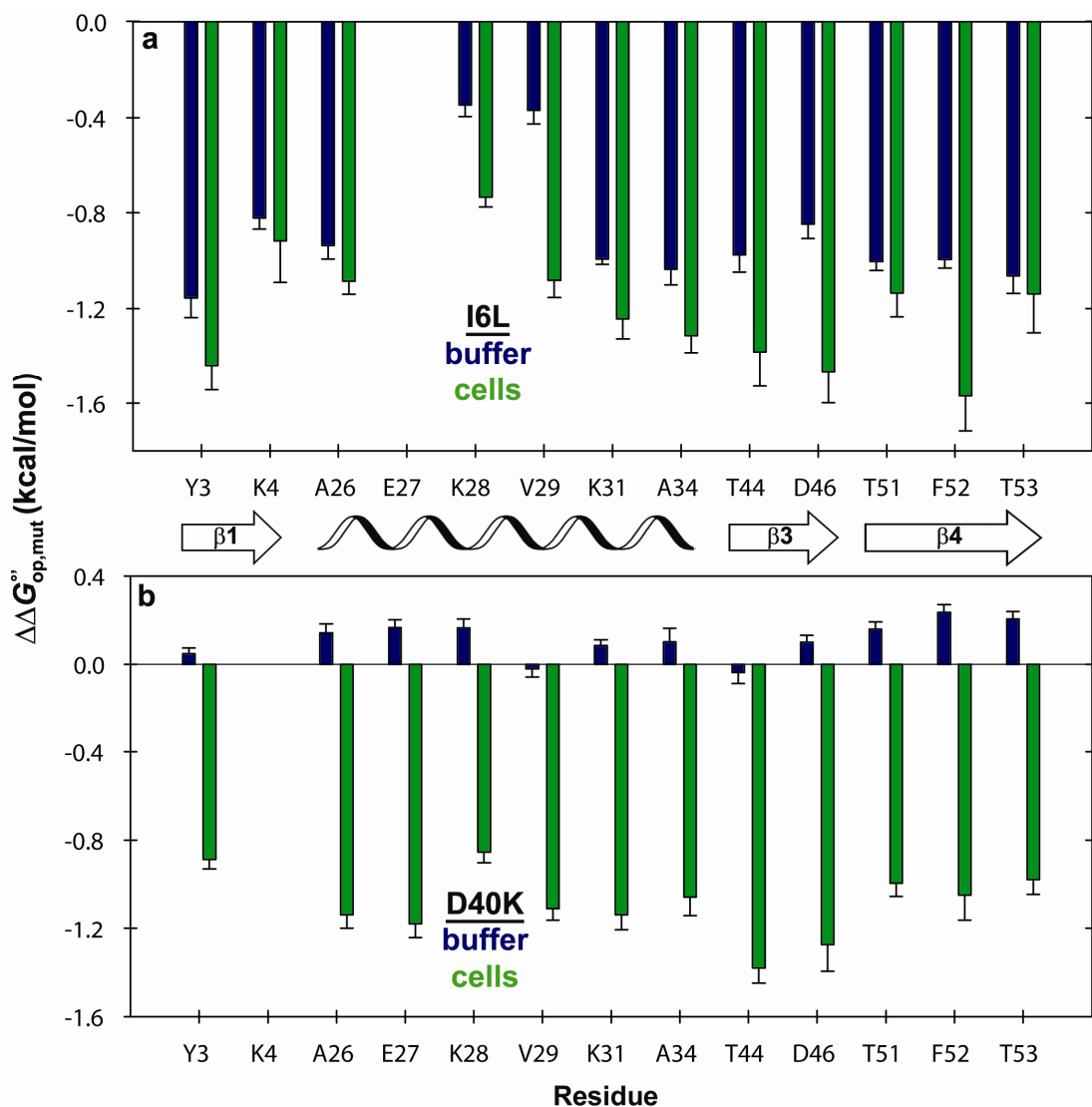
Figure 4.2 Stability changes ($\Delta\Delta G^{o\prime}_{op,mut} = \Delta G^{o\prime}_{op,var} - \Delta G^{o\prime}_{op,wt}$) caused by the (a) I6L and (b) D40K mutations in cells (green) and in buffer (blue). Error bars represent the uncertainty propagated from triplicate measurements. GB1 secondary structure elements associated with the residues are indicated between plots (a) and (b).
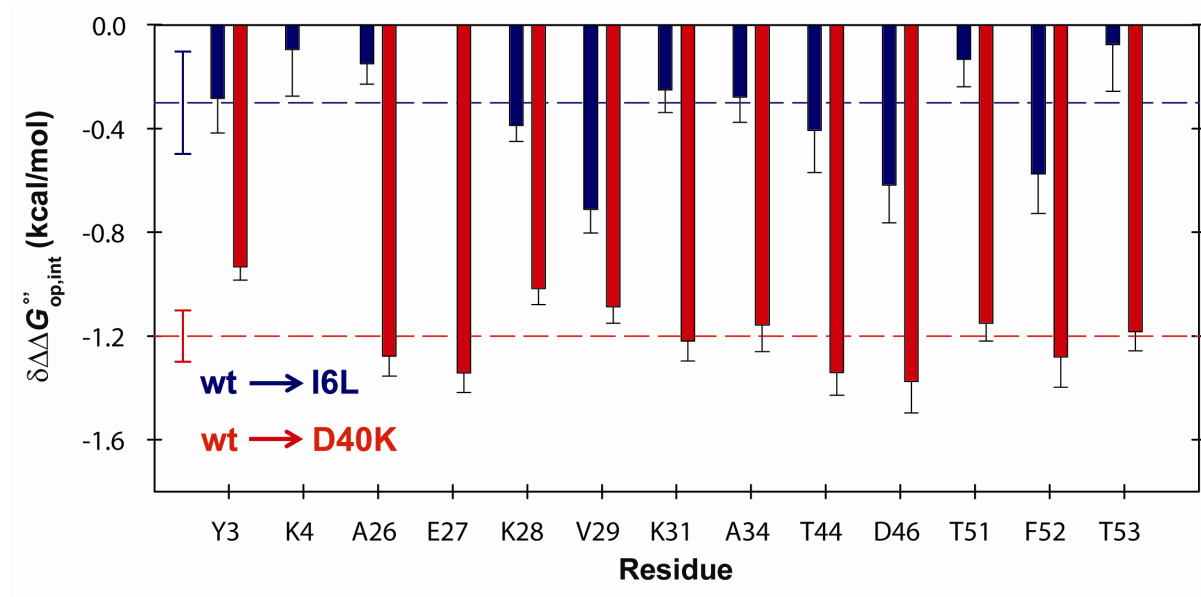
Figure 4.3 Interaction free energies ($\delta\Delta\Delta G^{\circ\prime}_{op,int} = \Delta\Delta G^{\circ\prime}_{op,mut,cell} - \Delta\Delta G^{\circ\prime}_{op,mut,buff}$) of the I6L (blue) and D40K (red) variants with the cytosol. Error bars represent the uncertainty propagated from triplicate measurements. Dashed lines and associated error bars are the average value and standard deviation for each mutation. K4 and E27 crosspeak volumes were insufficient for quantification in the D40K and I6L variants, respectively.
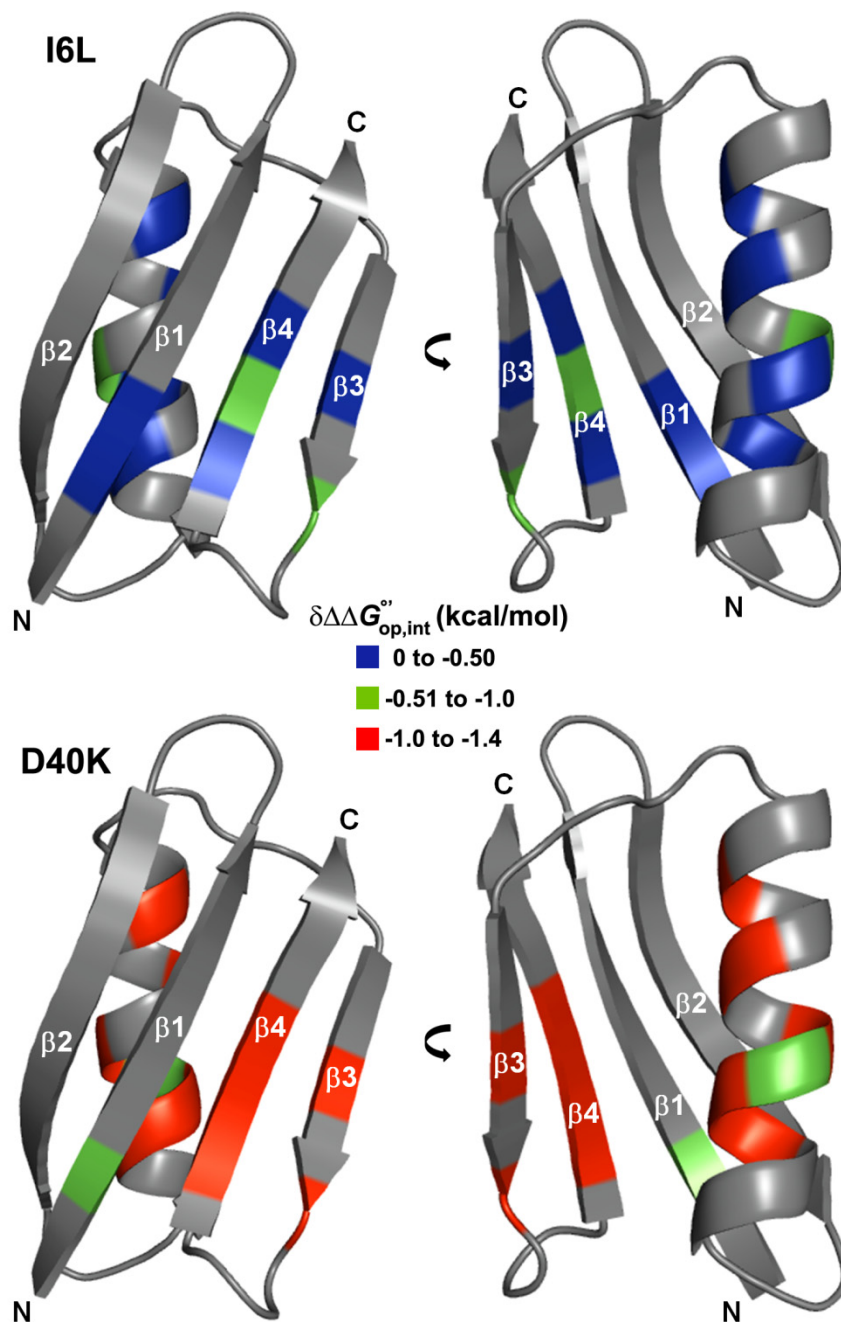
Figure 4.4 GB1 (1pgb) variants colored by interaction free energies ($\delta\Delta\Delta G^{o\prime}_{op,int}$). The coupled effect of mutating GB1 in cells is significantly more destabilizing to D40K, where the mutation involves a charged surface residue. Gray residues yield no data.
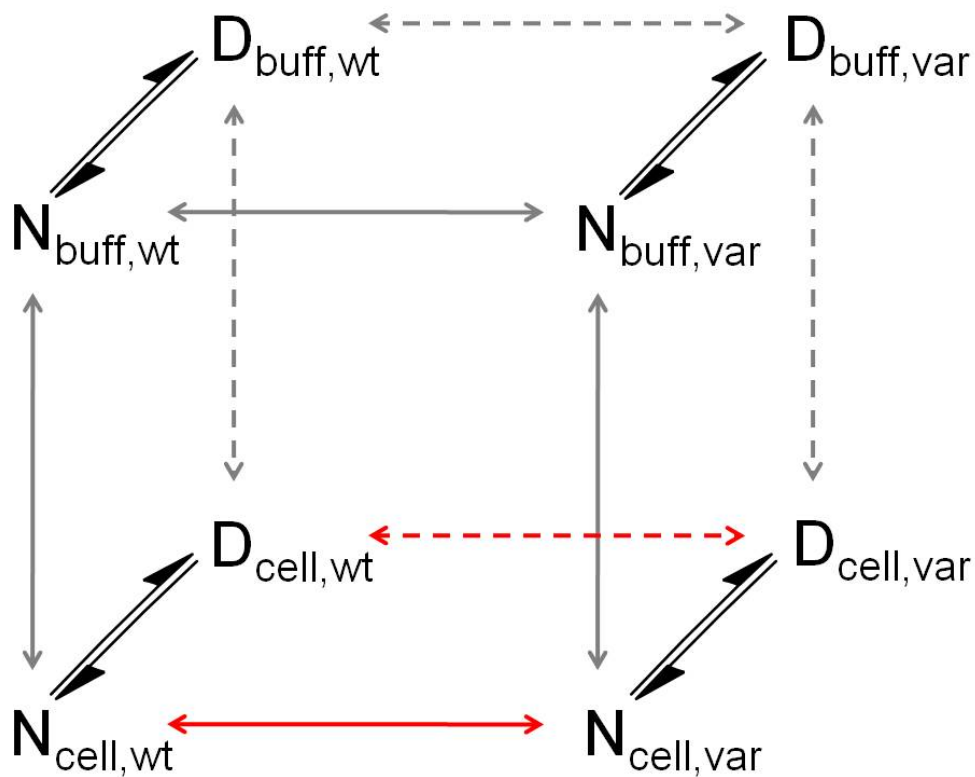
Figure 4.5 Expanded thermodynamic cycle from Figure 4.1, depicting equilibria between native (N) and denatured (D) states of wt GB1 and its variants. Intracellular interactions cause deviations in $\Delta\Delta G^{°'}_{op,mut,cell}$ and $\Delta\Delta G^{°'}_{op,mut,buff}$ (red).
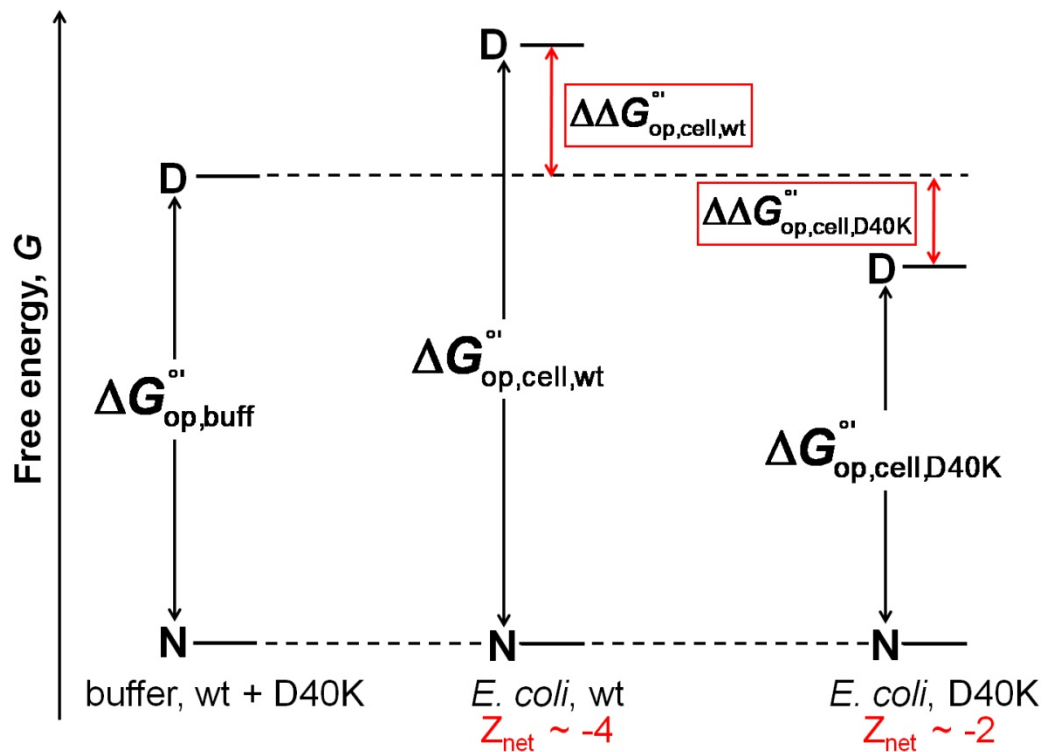
Figure 4.6 Free energy diagrams for wt and D40K GB1 in buffer and cells. In-cell

stabilization of wt GB1 is due to charge-charge repulsions with the cytoplasm

(middle diagram). Destabilization of D40K in cells is due to weakened repulsions

with the cytoplasm owing to a less negative charge with respect to wt (right-most

diagram). For clarity, the native states have been arbitrarily set to the same absolute
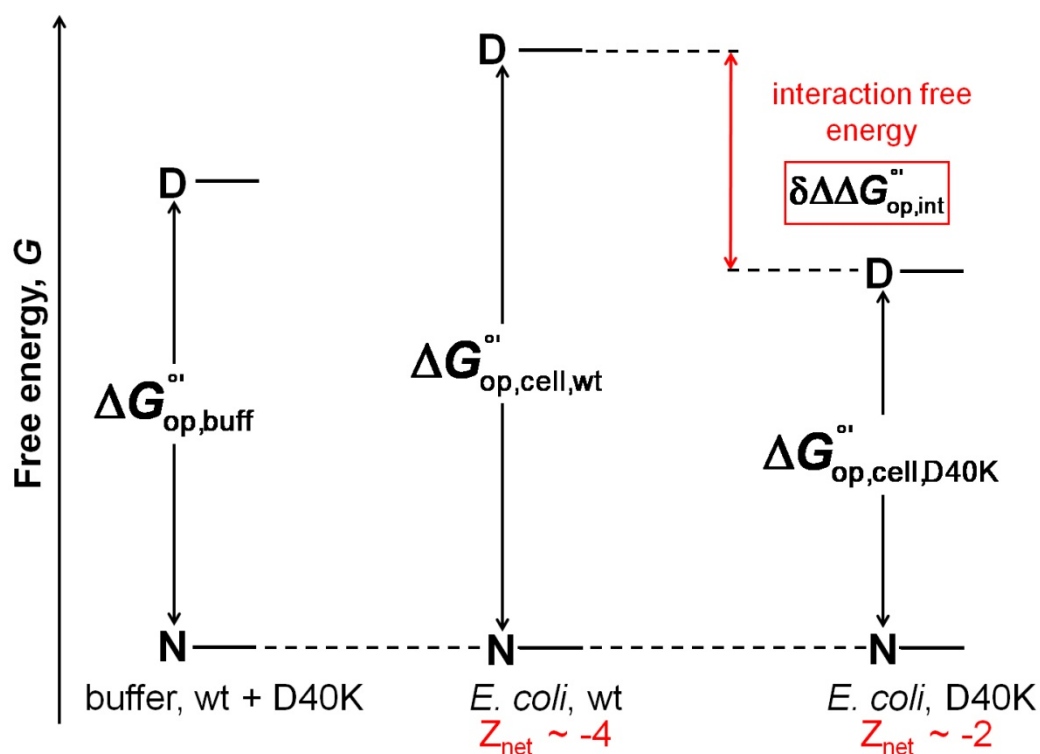
free energy.

Figure 4.7 Visualization of the interaction free energy ($\delta\Delta\Delta G^{\circ'}_{op,int}$) for D40K GB1 as the difference $\Delta\Delta G^{\circ'}_{op,cell,D40K} - \Delta\Delta G^{\circ'}_{op,cell,wt}$ from Figure 4.6. For clarity, the native states have been arbitrarily set to the same absolute free energy.

4.8 Tables

Table 4.1 Backbone amide proton exchange rates (k_{obs}, s^{-1}) and corresponding $\Delta G^{o'}_{op}$ (kcal/mol) values for D40K GB1 in cells and in buffer (PBS, pH$_{corr}$ 7.6, 37 °C). [a]SDM: standard deviation of the mean from three trials. [b]$\Delta\Delta G^{o'}_{op,cell,D40K} = \Delta G^{o'}_{op,cell,D40K} - \Delta G^{o'}_{op,buff,D40K}$. [c]From propagation of error, $\sigma = \sqrt{(\delta_{\Delta G^{o'}_{op,cell}})^2 + (\delta_{\Delta G^{o'}_{op,buff}})^2}$.

residue	$k_{obs,cells}$	SDM$_{cells}$[a]	$k_{obs,buff}$	SDM$_{buff}$[a]
Y3	9.7E-04	±0.4E-04	7.6E-04	±0.2E-04
A26	9.4E-04	±0.5E-04	3.03E-04	±0.07E-04
E27	4.5E-04	±0.4E-04	1.51E-04	±0.08E-04
K28	9.7E-04	±0.5E-04	1.19E-03	±0.02E-03
V29	6.3E-04	±0.5E-04	6.71E-04	±0.06E-04
K31	9.3E-04	±0.2E-04	3.6E-04	±0.1E-04
A34	7.2E-04	±0.8E-04	5.25E-04	±0.05E-04
T44	6.77E-04	±0.08E-04	2.13E-04	±0.03E-04
D46	6.8E-04	±0.6E-04	2.5E-04	±0.1E-04
T51	9.6E-04	±0.2E-04	3.4E-04	±0.1E-04
F52	9E-04	±1E-04	2.91E-04	±0.02E-04
T53	9.4E-04	±0.3E-04	3.49E-04	±0.07E-04

residue	$\Delta G^{o'}_{op,cell,D40K}$	SDM$_{cells}$[a]	$\Delta G^{o'}_{op,buff,D40K}$	SDM$_{buff}$[a]	$\Delta\Delta G^{o'}_{op,cell,D40K}$[b]	uncertainty[c]
Y3	6.75	±0.03	6.90	±0.02	-0.15	±0.03
A26	7.15	±0.04	7.85	±0.01	-0.70	±0.04
E27	6.60	±0.05	7.27	±0.03	-0.67	±0.06
K28	6.58	±0.03	6.45	±0.01	0.12	±0.03
V29	6.30	±0.05	6.25	±0.01	0.04	±0.05
K31	6.90	±0.01	7.49	±0.02	-0.58	±0.03
A34	7.11	±0.07	7.30	±0.01	-0.18	±0.07
T44	6.82	±0.01	7.53	±0.01	-0.71	±0.01
D46	6.71	±0.05	7.34	±0.03	-0.62	±0.06
T51	6.93	±0.01	7.57	±0.02	-0.64	±0.02
F52	6.88	±0.08	7.53	±0.00	-0.66	±0.08
T53	6.85	±0.02	7.47	±0.01	-0.61	±0.02

Table 4.2 $\Delta\Delta G^{o\prime}_{op,mut}$ ($\Delta G^{o\prime}_{op,D40K} - \Delta G^{o\prime}_{op,wt}$, kcal/mol) caused by the D40K mutation in cells and in buffer (PBS, pH$_{corr}$ 7.6, 37 °C). [a]From propagation of error. $\Delta G^{o\prime}_{op}$ values for wt GB1 are found in Chapter 3, Table 3.1.

residue	$\Delta\Delta G^{o\prime}_{op,mut,cell}$	uncertainty$_{cell}$[a]	$\Delta\Delta G^{o\prime}_{op,mut,buff}$	uncertainty$_{buff}$[a]
Y3	-0.89	±0.04	0.05	±0.02
A26	-1.14	±0.06	0.14	±0.04
E27	-1.18	±0.07	0.17	±0.04
K28	-0.85	±0.05	0.16	±0.04
V29	-1.11	±0.05	-0.02	±0.04
K31	-1.14	±0.07	0.08	±0.03
A34	-1.06	±0.08	0.10	±0.06
T44	-1.38	±0.07	-0.04	±0.05
D46	-1.28	±0.12	0.10	±0.03
T51	-0.99	±0.06	0.16	±0.03
F52	-1.05	±0.11	0.23	±0.04
T53	-0.98	±0.07	0.21	±0.03

Table 4.3 $\delta\Delta\Delta G^{o'}_{op,int}$ ($\Delta\Delta G^{o'}_{op,mut,cell} - \Delta\Delta G^{o'}_{op,mut,buff}$, kcal/mol) values for I6L and D40K

GB1. [a]From propagation of error. $\Delta\Delta G^{o'}_{op,mut}$ values for I6L GB1 are found in Chapter

3, Table 3.5.

residue	$\delta\Delta\Delta G^{o'}_{op.int.I6L}$	uncertainty$_{I6L}$[a]	$\delta\Delta\Delta G^{o'}_{op.int.D40K}$	uncertainty$_{D40K}$[a]
Y3	-0.3	±0.1	-0.93	±0.05
K4	-0.1	±0.2	-	-
A26	-0.15	±0.08	-1.28	±0.08
E27	-	-	-1.34	±0.08
K28	-0.39	±0.06	-1.02	±0.06
V29	-0.71	±0.09	-1.09	±0.06
K31	-0.25	±0.09	-1.22	±0.08
A34	-0.3	±0.1	-1.2	±0.1
T44	-0.4	±0.2	-1.34	±0.09
D46	-0.6	±0.2	-1.4	±0.1
T51	-0.1	±0.1	-1.15	±0.07
F52	-0.6	±0.2	-1.3	±0.1
T53	-0.1	±0.2	-1.18	±0.07

REFERENCES

1. Levinthal C. How to fold graciously. In: Debrunner P, Tsibris JCM, Münck E, Eds. (1969) Mossbauer Spectroscopy in Biological Systems: Proceedings of a meeting held at Allerton House. University of Illinois Press, Urbana, Monticello, IL, pp. 22-24.

2. Dill K (1990) Dominant forces in protein folding. Biochemistry 29:7133-7155.

3. Wu H (1929) A theory of denaturation and coagulation of proteins. Am J Physiol 90:562-563.

4. Harrington WF, Schellman JA (1956) Evidence for the instability of hydrogen-bonded peptide structures in water, based on studies of ribonuclease and oxidized ribonuclease. C R Trav Lab Carlsberg, Ser Chim 30:21-43.

5. Anfinsen CB (1973) Principles that govern the folding of protein chains. Science 181:223-230.

6. Hermans J, Scheraga HA (1961) Structural studies of ribonuclease. V. Reversible change of configuration. J Am Chem Soc 83:3283-3292.

7. Mirsky A, Pauling L (1936) On the structure of native, denatured, and coagulated proteins. Proc Natl Acad Sci U S A 22:439-447.

8. Eyring H, Stearn A (1939) The application of the theory of absolute reaction rates to proteins. Chem Rev 24:253-270.

9. Jacobsen C, Linderstrøm-Lang KU (1949) Salt linkages in proteins. Nature 164:411-412.

10. Pauling L, Corey RB, Branson HR (1951) The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain. Proc Natl Acad Sci U S A 37:205-211.

11. Kauzmann W (1959) Some factors in the interpretation of protein denaturation. Adv Protein Chem 14:1-63.

12. Frank HS, Evans MW (1945) Free volume and entropy in condensed systems III. Entropy in binary liquid mixtures; partial molal entropy in dilute solutions; structure and thermodynamics in aqueous electrolytes J Chem Phys 13:507-532.

13. Tanford C, Hauenstein JD, Rands DG (1955) Phenolic hydroxyl ionization in proteins. II. Ribonuclease. J Am Chem Soc 77:6409-6413.

14. Singer SJ (1963) The properties of proteins in nonaqueous solvents. Adv Protein Chem 17:1-68.

15. Perutz MF (1965) Structure and function of haemoglobin I. A tentative atomic model of horse oxyhaemoglobin. J Mol Biol 13:646-668.

16. Privalov P, Khechinashvili N (1974) A thermodynamic approach to the problem of stabilization of globular protein structure: A calorimetric study. J Mol Biol 86:665-684.

17. Richards FM (1977) Areas, volumes, packing, and protein structure. Annu Rev Biophys Bioeng 6:151-176.

18. Pace CN (1990) Conformational stability of globular proteins. Trends Biochem Sci 15:14-17.

19. Chan HS, Dill KA (1998) Protein folding in the landscape perspective: chevron plots and non-Arrhenius kinetics. Proteins: Struct, Funct, Genet 30:2-33.

20. Baker D, Agard DA (1994) Kinetics versus thermodynamics in protein folding. Biochemistry 33:7505-7509.

21. Pace CN, Shirley BA, Thomson JA. Measuring the conformational stability of a protein. In: Creighton TE, Ed. (1989) Protein Structure: A Practical Approach. IRL Press, Oxford, pp. 311-329.

22. Pace CN, Tanford C (1968) Thermodynamics of the unfolding of β-lactoglobulin A in aqueous urea solutions between 5 and 55°. Biochemistry 7:198-208.

23. Becktel WJ, Schellman JA (1987) Protein stability curves. Biopolymers 26:1859-1877.

24. Linderstrøm-Lang KU. Deuterium exchange and protein structure. In: Neuberger A, Ed. (1958) Symposium on Protein Structure. Methuen, London, pp. 23-34.

25. Hvidt AA, Nielsen SO (1966) Hydrogen exchange in proteins. Adv Prot Chem 21:287-386.

26. Englander SW, Mayne L (1992) Protein folding studied using hydrogen-exchange labeling and two-dimensional NMR. Annu Rev Biophys Biomol Struct 21:243-265.

27. Ghaemmaghami S, Fitzgerald MC, Oas TG (2000) A quantitative, high-throughput screen for protein stability. Proc Natl Acad Sci U S A 97:8296-8301.

28. Wagner G, Wüthrich K (1982) Amide proton exchange and surface conformation of the basic pancreatic trypsin inhibitor in solution: Studies with two-dimensional nuclear magnetic resonance. J Mol Biol 160:343-361.

29. Bai Y, Milne JS, Mayne L, Englander SW (1993) Primary structure effects on peptide group hydrogen exchange. Proteins: Struct, Funct, Genet 17:75-86.

30. Englander SW, Kallenbach NR (1983) Hydrogen exchange and structural dynamics of proteins and nucleic acids. Q Rev Biophys 16:521-655.

31. Brandts JF (1964) The thermodynamics of protein denaturation. II. A model of reversible denaturation and interpretations regarding the stability of chymotrypsinogen. J Am Chem Soc 86:4302-4314.

32. Zimmerman SB, Trach SO (1991) Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of *Escherichia coli*. J Mol Biol 222:599-620.

33. Minton AP (1983) The effect of volume occupancy upon the thermodynamic activity of proteins: some biochemical consequences. Mol Cell Biochem 55:119-140.

34. Zhou H-X (2004) Loops, linkages, rings, catenanes, cages, and crowders: entropy-based strategies for stabilizing proteins. Acc Chem Res 37:123-130.

35. Minton AP (2005) Models for excluded volume interaction between an unfolded protein and rigid macromolecular cosolutes: Macromolecular crowding and protein stability revisted. Biophys J 88:971-985.

36. Laurent TC (1963) The interaction between polysaccharides and other macromolecules. 5. The solubility of proteins in the presence of dextran. Biochem J 89:253-257.

37. Sasahara K, McPhie P, Minton AP (2003) Effect of dextran on protein stability and conformation attributed to macromolecular crowding. J Mol Biol 326:1227-1237.

38. Tokuriki N, Kinjo M, Negi S, Hoshino M, Goto Y, Urabe I, Yomo T (2004) Protein folding by the effects of macromolecular crowding. Protein Sci 13:125-133.

39. Charlton LM, Barnes CO, Li C, Orans J, Young GB, Pielak GJ (2008) Residue-level interrogation of macromolecular crowding effects on protein stability. J Am Chem Soc 130:6826-6830.

40. Laurent TC (1995) An early look at macromolecular crowding. Biophys Chem 57:7-14.

41. McConkey E (1982) Molecular evolution, intracellular organization, and the quinary structure of proteins. Proc Natl Acad Sci U S A 79:3236-3240.

42. Miklos AC, Sarkar M, Wang Y, Pielak GJ (2011) Protein crowding tunes protein stability. J Am Chem Soc 133:7116-7120.

43. Feig M, Sugita Y (2012) Variable interactions between protein crowders and biomolecular solutes are important in understanding cellular crowding. J Phys Chem B 116:599-605.

44. Harada R, Tochio N, Kigawa T, Sugita Y, Feig M (2013) Reduced native state stability in crowded cellular environment due to protein-protein interactions. J Am Chem Soc 135:3696-3701.

45. Sarkar M, Smith AE, Pielak GJ (2013) Impact of reconstituted cytosol on protein stability. Proc Natl Acad Sci U S A 110:19342-19347.

46. Sarkar M, Li C, Pielak GJ (2013) Soft interactions and crowding. Biophysical Reviews 5:187-194.

47. Makhatadze GI, Privalov PL (1992) Protein interactions with urea and guanidinium chloride: A calorimetric study. J Mol Biol 226:491-505.

48. Monteith WB, Pielak GJ (2014) Residue level quantification of protein stability in living cells. Proc Natl Acad Sci U S A In press.

49. Minton AP (2013) Quantitative assessment of the relative contributions of steric repulsion and chemical interactions to macromolecular crowding. Biopolymers 99:239-244.

50. Zhou HX (2013) Polymer crowders and protein crowders act similarly on protein folding stability. FEBS Lett 587:394-397.

51. McGuffee SR, Elcock AH (2010) Diffusion, crowding, & protein stability in a dynamic molecular model of the bacterial cytoplasm. PLoS Comput Biol 6:e1000694.

52. Ghaemmaghami S, Oas TG (2001) Quantitative protein stability measurement *in vivo*. Nat Struct Biol 8:879-882.

53. Ignatova Z, Gierasch LM (2004) Monitoring protein stability and aggregation *in vivo* by real-time fluorescent labeling. Proc Natl Acad Sci U S A 101:523-528.

54. Ignatova Z, Krishnan B, Bombardier JP, Marcelino AMC, Hong J, Gierasch LM (2007) From the test tube to the cell: exploring the folding and aggregation of a β-clam protein. Biopolymers 88:157-163.

55. Ebbinghaus S, Dhar A, McDonald JD, Gruebele M (2010) Protein folding stability and dynamics imaged in a living cell. Nat Methods 7:319-323.

56. Dhar A, Girdhar K, Singh D, Gelman H, Ebbinghaus S, Gruebele M (2011) Protein stability and folding kinetics in the nucleus and endoplasmic reticulum of eucaryotic cells. Biophys J 101:421-430.

57. Guo M, Xu Y, Gruebele M (2012) Temperature dependence of protein folding kinetics in living cells. Proc Natl Acad Sci U S A 109:17863-17867.

58. Guzman I, Gelman H, Tai J, Gruebele M (2014) The extracellular protein VlsE is destabilized inside cells. J Mol Biol 426:11-20.

59. Lindquist S (1986) The heat-shock response. Annu Rev Biochem 55:1151-1191.

60. Withman B, Gunasekera T, Beesetty P, Agans R, Paliy O (2013) Transcriptional responses of uropathogenic *Escherichia coli* to increased environmental osmolality caused by salt or urea. Infect Immun 81:80-89.

61. Barnes CO, Monteith WB, Pielak GJ (2011) Internal and global protein motion assessed with a fusion construct and in-cell NMR. ChemBioChem 12:390-391.

62. Dill KA, Chan HS (1997) From Levinthal to pathways to funnels. Nat Struct Mol Biol 4:10-19.

63. Serber Z, Ledwidge R, Miller S, Dötsch V (2001) Evaluation of parameters critical to observing proteins inside living *Escherichia coli* by in-cell NMR spectroscopy. J Am Chem Soc 123:8895-8901.

64. Li C, Charlton LM, Lakkavaram A, Seagle C, Wang G, Young GB, Macdonald JM, Pielak GJ (2008) Differential dynamical effects of macromolecular crowding on an intrinsically disordered protein and a globular protein: implications for in-cell NMR spectroscopy. J Am Chem Soc 130:6310-6311.

65. Li C, Wang Y, Creager-Allen R, Lutz EA, Scronce H, Slade KM, Ruf RAS, Mehl RA, Pielak GJ (2010) Protein ^{19}F NMR in *Escherichia coli*. J Am Chem Soc 132:321-327.

66. Elcock AH (2010) Models of macromolecular crowding effects and the need for quantitative comparisons with experiment. Curr Opin Struct Biol 20:196-206.

67. Pielak GJ, Li C, Miklos AC, Schlesinger AP, Slade KM, Wang G, Zigoneanu IG (2009) Protein nuclear magnetic resonance under physiological conditions. Biochemistry 28:226-234.

68. Bodenhausen G, Ruben DJ (1980) Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. Chem Phys Lett 69:185-189.

69. Dedmon MM, Patel CN, Young GB, Pielak GJ (2002) FlgM gains structure in living cells. Proc Natl Acad Sci U S A 99:12681-12684.

70. Aslanidis C, de Jong PJ (1990) Ligation-independent cloning of PCR products (LIC-PCR). Nucleic Acids Res 18:6069-6074.

71. Moussa C, Wersinger C, Rusnak M, Tomita Y, Sidhu A (2004) Abnormal migration of human wild-type α-synuclein upon gel electrophoresis. Neurosci Lett 371:239-243.

72. McNulty BC, Young GB, Pielak GJ (2006) Macromolecular crowding in the *Escherichia coli*. J Mol Biol 355:893-897.

73. Wang Y, Li C, Pielak GJ (2010) Effects of proteins on protein diffusion. J Am Chem Soc 132:9392-9397.

74. Slade KM, Baker R, Chua M, Thompson NL, Pielak GJ (2009) Effects of recombinant protein expression on green fluorescent protein diffusion in *Escherichia coli*. Biochemistry 48:5083-5089.

75. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6:277-293.

76. Johnson BA, Blevins RA (1994) NMR View: A computer program for the visualization and analysis of NMR data. J Biomol NMR 4:603-614.

77. Ellis RJ (2001) Macromolecular crowding: an important but neglected aspect of the intracellular environment. Curr Opin Struct Biol 11:114-119.

78. Zhou HX, Rivas GN, Minton AP (2008) Macromolecular crowding and confinement: Biochemical, biophysical, and potential physiological consequences. Annu Rev Biophys 37:375-397.

79. Ellis RJ (2001) Macromolecular crowding: obvious but underappreciated. Trends Biochem Sci 26:597-604.

80. Crowley PB, Chow E, Papkovskaia T (2011) Protein interactions in the *Escherichia coli* cytosol: an impediment to in-cell NMR spectroscopy. ChemBioChem 12:1043-1048.

81. Wang Q, Zhuravleva A, Gierasch LM (2011) Exploring weak, transient protein-protein interactions in crowded *in vivo* environments by in-cell nuclear magnetic resonance spectroscopy. Biochemistry 50:9225-9236.

82. Srere PA (2000) Macromolecular interactions: tracing the roots. Trends Biochem Sci 25:150-153.

83. Knowles DB, LaCroix AS, Deines NF, Shkel I, Record MT (2011) Separation of preferential interaction and excluded volume effects on DNA duplex and hairpin stability. Proc Natl Acad Sci U S A 108:12699-12704.

84. Latham MP, Kay L (2013) Probing non-specific interactions of Ca^{2+}-calmodulin in *E. coli* lysate. J Biomol NMR 55:239-247.

85. Zhou HX (2013) Influence of crowded cellular environments on protein folding, binding, and oligomerization: Biological consequences and potentials of atomistic modeling. FEBS Lett 587:1053-1061.

86. Gershenson A (2014) Deciphering protein stability in cells. J Mol Biol 426:4-6.

87. Fahnestock S, Alexander P, Nagle J, Filpula DR (1986) Gene for an immunoglobulin-binding protein from a group G streptococcus. J Bacteriol 167:870-880.

88. Gronenborn AM, Filpula DR, Essig NZ, Achari A, Whitlow M, Wingfield PT, Clore GM (1991) A novel, highly stable fold of the immunoglobulin binding domain of streptococcal protein G. Science 253:657-661.

89. Alexander P, Fahnestock S, Lee T, Orban J, Bryan P (1992) Thermodynamic analysis of the folding of the streptococcal protein G IgG-binding domains B1 and B2: why small proteins tend to have high denaturation temperatures. Biochemistry 31:3597-3603.

90. Alexander P, Orban J, Bryan P (1992) Kinetic analysis of folding and unfolding the 56 amino acid IgG-binding domain of streptococcal protein G. Biochemistry 31:7243-7248.

91. Barchi Jr. JJ, Grasberger B, Gronenborn AM, Clore GM (1994) Investigation of the backbone dynamics of the IgG-binding domain of streptococcal protein G by heteronuclear two-dimensional ^{1}H-^{15}N nuclear magnetic resonance spectroscopy. Protein Sci 3:15-21.

92. Gallagher T, Alexander P, Bryan P, Gilliland G (1994) Two crystal structures of the B1 immunoglobulin-binding domain of streptococcal protein G and comparison with NMR. Biochemistry 33:4721-4729.

93. Orban J, Alexander P, Bryan P, Khare D (1995) Assessment of stability differences in the protein G B1 and B2 domains from hydrogen-deuterium exchange: comparison with calorimetric data. Biochemistry 34:15291-15300.

94. Smith CK, Withka JM, Regan L (1994) A thermodynamic scale for the β-sheet forming tendencies of the amino acids. Biochemistry 33:5510-5517.

95. Kuszewski J, Clore GM, Gronenborn AM (1994) Fast folding of a prototypic polypeptide: The immunoglobulin binding domain of streptococcal protein G. Protein Sci 3:1945-1952.

96. Park S-H, Shastry MCR, Roder H (1999) Folding dynamics of the B1 domain of protein G explored by ultrarapid mixing. Nat Struct Mol Biol 6:943-947.

97. Krantz BA, Mayne L, Rumbley J, Englander SW, Sosnick TR (2002) Fast and slow intermediate accumulation and the initial barrier mechanism in protein folding. J Mol Biol 324:359-371.

98. Morrone A, Giri R, Toofanny Rudesh D, Travaglini-Allocatelli C, Brunori M, Daggett V, Gianni S (2011) GB1 is not a two-state folder: identification and characterization of an on-pathway intermediate. Biophys J 101:2053-2060.

99. Pace CN (1990) Measuring and increasing protein stability. Trends Biotechnol 8:93-98.

100. Dempsey CE (2001) Hydrogen exchange in peptides and proteins using NMR spectroscopy. Prog Nucl Magn Reson Spectrosc 39:135-170.

101. SPHERE:http://www.fccc.edu/research/labs/roder/sphere/sphere.html.

102. Ferraro DM, Lazo ND, Robertson AD (2004) EX1 hydrogen exchange and protein folding. Biochemistry 43:587-594.

103. Raschke T, Marqusee S (1998) Hydrogen exchange studies of protein structure. Curr Opin Biotechnol 9:80-86.

104. Englander SW, Sosnick TR, Englander JJ, Mayne L (1996) Mechanisms and uses of hydrogen exchange. Curr Opin Struct Biol 6:18-23.

105. Bai Y, Sosnick TR, Mayne L, Englander SW (1995) Protein folding intermediates: native-state hydrogen exchange. Science 269:192-197.

106. Inomata K, Ohno A, Tochio H, Isogai S, Tenno T, Nakase I, Takeuchi T, Futaki S, Ito Y, Hiroaki H, Shirakawa M (2009) High-resolution multi-dimensional NMR spectroscopy of proteins in human cells. Nature 458:106-109.

107. Serber Z, Keatinge-Clay AT, Ledwidge R, Kelly AE, Miller SM, Dötsch V (2001) High-resolution macromolecular NMR spectroscopy inside living cells. J Am Chem Soc 123:2446-2447.

108. Miklos AC, Li C, Pielak GJ (2009) Using NMR-detected backbone amide ^{1}H exchange to assess macromolecular crowding effects on globular-protein stability. Methods Enzymol 466:1-18.

109. Reardon PN, Spicer LD (2005) Multidimensional NMR spectroscopy for protein characterization and assignment inside cells. J Am Chem Soc 127:10848-10849.

110. Xu G, Ye Y, Liu X, Cao S, Wu Q, Cheng K, Liu M, Pielak GJ, Li C (2014) Strategies for protein NMR in *Escherichia coli*. Biochemistry 53:1971-1981.

111. Wang Y, Sarkar M, Smith AE, Krois A, Pielak GJ (2012) Macromolecular crowding and protein stability. J Am Chem Soc 134:16614-16618.

112. Gronenborn AM, Frank MK, Clore GM (1996) Core mutants of the immunoglobulin binding domain of streptococcal protein G: stability and structural integrity. FEBS Lett 398:312-316.

113. Lindman S, Xue W-F, Szczepankiewicz O, Bauer MC, Nilsson H, Linse S (2006) Salting the charged surface: pH and salt dependence of protein G B1 stability. Biophys J 90:2911-2921.

114. Huyghues-Despointes BMP, Scholtz JM, Pace CN (1999) Protein conformational stabilities can be determined from hydrogen exchange rates. Nat Struct Mol Biol 6:910-912.

115. Smith AE, Sarkar M, Young GB, Pielak GJ (2013) Amide proton exchange of a dynamic loop in cell extracts. Protein Sci 22:1313-1319.

116. Molday RS, Englander SW, Kallen RG (1972) Primary structure effects on peptide group hydrogen exchange. Biochemistry 11:150-158.

117. Sakakibara D, Sasaki A, Ikeya T, Hamatsu J, Hanashima T, Mishima M, Yoshimasu M, Hayashi N, Mikawa T, Walchli M, Smith BO, Shirakawa M, Guntert P, Ito Y (2009) Protein structure determination in living cells by in-cell NMR spectroscopy. Nature 458:102-105.

118. Spitzer J, Poolman B (2009) The role of biomacromolecular crowding, ionic strength, and physicochemical gradients in the complexities of life's emergence. Microbiol Mol Biol Rev 73:371-388.

119. Selenko P, Serber Z, Gadea B, Ruderman J, Wagner G (2006) Quantitative NMR analysis of the protein G B1 domain in *Xenopus laevis* egg extracts and intact oocytes. Proc Natl Acad Sci U S A 103:11904-11909.

120. Li C, Liu M (2013) Protein dynamics in living cells studied by in-cell NMR spectroscopy. FEBS Lett 587:1008-1011.

121. Sarkar M, Lu J, Pielak GJ (2014) Protein crowder charge and protein stability. Biochemistry 53:1601-1606.

122. Glasoe PK, Long FA (1960) Use of glass electrodes to measure acidities in deuterium oxide. J Phys Chem B 64:188-190.

123. Sanger F, Nicklen S, Coulson A (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A 74:5463-5467.

124. Benov L, Al-Ibraheem J (2002) Disrupting *Escherichia coli*: a comparison of methods. J Biochem Mol Biol 35:428-431.

125. Gronenborn AM, Clore GM (1996) Rapid screening for structural integrity of expressed proteins by heteronuclear NMR spectroscopy. Protein Sci 5:174-177.

126. Booth IR (1985) Regulation of cytoplasmic pH in bacteria. Microbiol Rev 49:359-378.

127. Navon G, Ogawa S, Shulman RG, Yamane T (1977) High-resolution ^{31}P nuclear magnetic resonance studies of metabolism in aerobic *Escherichia coli* cells. Proc Natl Acad Sci U S A 74:888-891.

128. Bryant JE, Lecomte JT, Lee AL, Young GB, Pielak GJ (2005) Protein dynamics in living cells. Biochemistry 44:9275-9279.

129. Hicks WM, Kotlajich MV, Visick JE (2005) Recovery from long-term stationary phase and stress survival in *Escherichia coli* require the L-isoaspartyl protein carboxyl methyltransferase at alkaline pH. Microbiology 151:2151-2158.

130. Cavallo L, Kleinjung J, Fraternali F (2003) POPS: a fast algorithm for solvent accessible surface areas at atomic and residue level. Nucleic Acids Res 31:3364-3366.

131. Wirth AJ, Gruebele M (2013) Quinary protein structure and the consequences of crowding in living cells: leaving the test-tube behind. BioEssays 35:984-993.

132. Carter PJ, Winter G, Wilkinson AJ, Fersht AR (1984) The use of double mutants to detect structural changes in the active site of the tyrosyl-tRNA synthetase (Bacillus stearothermophilus). Cell 38:835-840.

133. Schreiber G, Fersht AR (1995) Energetics of protein-protein interactions: analysis of the barnase-barstar interface by single mutations and double mutant cycles. J Mol Biol 248:478-486.

134. Horovitz A, Fersht AR (1992) Co-operative interactions during protein folding. J Mol Biol 224:733-740.

135. Green SM, Shortle D (1993) Patterns of nonadditivity between pairs of stability mutations in staphylococcal nuclease. Biochemistry 32:10131-10139.

136. Lorimer GH (1996) A quantitative assessment of the role of the chaperonin proteins in protein folding *in vivo*. FASEB J 10:5-9.

137. Lattman E, Rose G (1993) Protein folding - what's the question? Proc Natl Acad Sci U S A 90:439-441.

138. Pace CN, Alston RW, Shaw KL (2000) Charge-charge interactions influence the denatured state ensemble and contribute to protein stability. Protein Sci 9:1395-1398.

139. Spector S, Wang M, Carp SA, Robblee J, Hendsch ZS, Fairman R, Tidor B, Raleigh DP (2000) Rational modification of protein stability by the mutation of charged surface residues. Biochemistry 39:872-879.