Macromolecular Crowding and Protein Chemistry: Views from Inside and Outside Cells

Yaqiang Wang

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
2012

Approved by:

Nancy Allbritton, Ph.D.

Christopher Fecko, Ph.D.

Gary Pielak, Ph.D.

Linda Spremulli, Ph.D.

Nancy Thompson, Ph.D.

Abstract

YAQIANG WANG: Macromolecular Crowding and Protein Chemistry: Views from Inside and Outside Cells
(Under the direction of Professor Gary J. Pielak, Ph.D.)

The cytoplasm is crowded, and the concentration of macromolecules can reach ~ 300 g/L, an environment vastly different from the dilute, idealized conditions usually used in biophysical studies. Macromolecular crowding arise from two phenomena, excluded volume and nonspecific chemical interactions, until recently, only excluded volume effect has been considered. Theory predicts that this macromolecular crowding can have large effects. Most proteins, however, are studied outside cells in dilute solution with macromolecule concentrations of 10 g/L or less. In-cell NMR provides a means to assess protein biophysics at atomic resolution in living cells, but it remains in its infancy, and several potential challenges need to be addressed. One challenge is the inability to observe ^{15}N-^{1}H NMR spectra from many small globular proteins.

^{19}F NMR was used to expand the application of in-cell NMR. This work suggests that high viscosity and weak interactions in the cytoplasm can make routine ^{15}N enrichment a poor choice for in-cell NMR studies of globular proteins in *Escherichia coli*. To gain insight into this problem, I turned to *in vitro* experiments where conditions can be controlled with precision. Using both synthetic polymers and globular proteins, I studied the effects of crowding on the

diffusion of the test protein, chymotrypsin inhibitor 2. The results not only pinpoint the source of the problem – nonspecific chemical interactions – but also suggest that proteins are more suitable mimics of the intracellular environment.

I also measured the stability of ubiquitin in solutions crowded with synthetic polymers or globular proteins to further elucidate the role of nonspecific chemical interactions under crowded conditions. The increased stability observed in synthetic crowders was consistent with a dominant entropic role for excluded volume, but the effect of protein crowders depended on charge. Protein-induced crowding increased stability when the sign of the net charge of the crowder was the same as that of ubiquitin, but decreased stability when the proteins were oppositely charged. The results indicate that synthetic polymers do not provide physiologically relevant insights and that the overall effect of macromolecular crowding depends on the winner of the near stalemate between excluded volume and nonspecific interactions.

To my beautiful wife, Xiaohui Fang, for all your love, patience and advice over the years.

Acknowledgements

The road of science is full of challenges. To pursue my dream, I was confused, sad, as well as excited and cheerful. Writing this dissertation is something like drafting a short memoir reviewing this tough and valuable journey. In the past five years, I have certainly been enriched by my graduate study and research experience. Here, near the end of this journey, I would like to express my sincere acknowledgement to my professors, colleagues, friends and family.

First of all, I gratefully and sincerely thank my advisor, Professor Gary Pielak. His solid knowledge and persistent attitude toward scientific research, as well as his unwavering support, has made him an excellent advisor. His office was always open whenever I had questions. I do not know how many times I sat in his office with frustration and walked out with insightful explanations and efficient solutions. I always enjoyed those funny stories he told me. I also greatly appreciate his strong support of my career development and generous help in my personal life. I am fortunate to be his student.

I own many thanks to Dr. Conggang Li for his help on my research. We worked closely on several projects. I learned from him a great deal of not only fundamental and practical NMR knowledge, but also how to approach and solve scientific problems. I also appreciate Dr. Guifang Wang's help with my experiments. I earned a lot of practical biochemistry knowledge from her.

I feel so grateful and lucky to have such helpful colleagues who provided a constant source of knowledge and inspiration. Drs. Lisa Charlton, Rebecca Sutton and Kristin Slade taught me how to express and purify proteins. Dr. Andrew Miklos, Ms. Mohona Sarkar, Mr. Austin Smith and I worked on the protein stability project. We had a good time and discussed a lot about crowding and protein stability. Drs. Alexander Schlesinger, Xavier Tadeo and I collaborated on the Protein L project and enjoyed working together. I thank Dr. Imola Zigoneanu and Ms. Laura Benton for their help on the disordered protein diffusion project. I thank Mssrs. Alexander Krois and Vishavpreet Singh for their help on sample preparation. Also, I want to remember those folks in the rest of the Pielak tribe: Ms. Jillian Tyrrell, Mr. Will Monteith, Mr. Joe Lu, Mr. Christopher Barnes, Ms. Naima Sharaf, Ms. Yuri Yang, Ms. Kristen Black, Ms. Heidi Scronce, Mr. Evan Lutz, and Mr. Emmanuel Chan.

I also thank my committee, Professors Nancy Allbritton, Christopher Fecko, Linda Spremulli, and Nancy Thompson, for directing my Ph.D. study. I thank Drs. Marc ter Horst and Gregory Young for spectrometer maintenance. I also want to express my thanks to Ms. Elizabeth Pielak for taking time to read and comment on my manuscripts.

Finally and most importantly, I would like to thank my family, especially my wife, Xiaohui Fang, who has contributed all her time to take care of the family and support my study. I thank my parents, Qingchun Wang and Yufeng Zhang, my sisters Yaling Wang and Yahong Wang, and my brother Yawei Wang for their unconditional love and support.

Table of Contents

List of Tables

List of Figures

List of Abbreviations and Symbols

3FY	3-fluorotyrosine
A	angstrom
A.A.	amino acid
AcGFP	Aequorea coerulescens green fluorescent protein
AMP	ampicillin
avg.	average
BSA	bovine serum albumin
^{o}C	degree Celsius
cm	centimeter
CAM	calmodulin
CI2	chymotrypsin inhibitor 2
D_2O	deuterium oxide
DEER	double electron–electron resonance
DNA	deoxyribonucleic acid
EDTA	ethylenediaminetetraacetic acid
EPR	electron paramagnetic resonance
FPLC	fast protein liquid chromatography
g	gram
x *g*	times gravity
GB1	protein G B1 domain
GFP	green fluorescent protein
h	hour
HDH	histidinol dehydrogensase

His	histidine
HMQC	heteronuclear multiple-quantum coherence
HSQC	heteronuclear single-quantum coherence
Hz	hertz
IPTG	isopropyl-β-D-1-thiogalactopyranoside
kcal	kilocalorie
k_{cl}	rate of closing
kDa	kilodalton
kHz	kilohertz
k_{int}	intrinsic rate of exchange
k_{obs}	overall rate of exchange
$K^{o'}_{op'}$	equilibrium constant for opening
k_{op}	rate of opening
L	liter
LB	Luria broth
LIC	ligation-independent cloning
M	Molar
m	meter
mA	milliampere
MAS	magic angle spinning
mg	milligram
MHz	megahertz
min	minute
mL	milliliter
mm	millimeter

mM	millimolar
mol	mole
mOsm	osmolality (mmol/kg)
MWCO	molecular weight cut off
ng	nanogram
Ni-NTA	nickel-nitrilotriacetic acid
nL	nanoliter
nm	nanometer
NmerA	N-terminal metal-binding domain of mercuric ion reductase
NMR	nuclear magnetic resonance
OD_{600}	optical density at 600 nm
Osm	osmolality (mol/kg)
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDZ3	the third PDZ domain
PEG	polyethylene glycol
PH	polyhedron
p*I*	Isoelectric point
pL	picoliter
PMSF	phenylmethanesulphonylfluoride
ppm	chemical shift in parts per million
ProtL	protein L
PVP	polyvinylpyrrolidone
RMSD	root mean square deviation
RNase	A ribonuclease A

R_1	longitudinal relaxation rate
R_2	transverse relaxation rate
R_H	hydrodynamic radius
rpm	revolutions per minute
s	second
S^2	order parameter
SE	Stokes-Einstein
SED	Stokes-Einstein-Debye
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SPHERE	server program for hydrogen exchange rate estimation
T_1	longitudinal relaxation time
T_2	transverse relaxation time
TE	tris ethylenediaminetetraacetic acid
tfmF	trifluoromethyl-L-phenylalanine
TROSY	transverse relaxation optimized spectroscopy
UBQ	ubiquitin
V	volt
v/v	volume/volume
w/v	weight/volume

Greek-based Symbols

αSYN	α-synuclein
$\Delta G^{o'}$	standard unfolding free energy
$\Delta H^{o'}$	standard enthalpy change

$\Delta S^{o'}$	standard entropy change
μL	microgram
μL	microliter
μm	micrometer

Chapter 1 - Introduction

The material in this chapter is a review paper from:

Wang Y, Li C, and Pielak GJ, In-cell magnetic resonance spectroscopy, *Chinese Journal of Magnetic Resonance*, in press

1.1 Introduction

The environment inside cells is exceptionally complex and contains macromolecules at concentrations exceeding 300 g/L and volume occupancies of 30%,[1] vastly different from the dilute, idealized conditions usually used in biophysical studies. Until recently, however, most protein studies were still performed in vitro and in dilute solutions, conditions that can provide beautiful, but perhaps physiologically irrelevant, data. New techniques for examining protein biophysics under physiological conditions are needed to help us fully understand protein function. In-cell NMR provides a means to investigate proteins in their native environment.[2]

Protein resonances in living cells were first observed four decades ago,[3] but in-cell NMR has enjoyed wide spread attention only in the last ten years.[4, 5] The main reason for the revived interest is the adoption of isotopic enrichment and labeling techniques that allow the protein of interest – the test protein – to be easily distinguished from other intracellular components. The most popular methods for accumulating these proteins in prokaryotic cells such as *Escherichia* *coli* is overexpression.[6-21] In contrast, however, expression levels in eukaryotic cells are too low to be useful for NMR. Thanks to translocation[22] and microinjection,[23-31] isotopically labeled protein can be accumulated in these cells, although microinjection is limited to relatively large cells such as the oocytes (~ 1 mm diameter[32]).

In-cell NMR is the only technique that provides atomic-level information about protein biophysics in living cells. It has been used for structure determination,[21] protein folding studies,[14, 33] drug screening,[34] and for assessing macromolecular interactions[6, 11, 35] and post-translational modifications,[36-38] and even for examining nucleic acids.[15, 39] These applications have been recently reviewed.[40-43] Here, we focus on the limitations and potential pitfalls of in-cell NMR, and then discuss recent advances.

1.2 Limitations

Surprisingly, few globular proteins have yielded decent in-cell NMR spectra.[12, 42, 44, 45] The quality of high resolution NMR spectra depends on the line width of the resonances, because narrow resonances are easy to detect. The degree of line broadening depends on both the homogeneity of the sample and the global tumbling of the test protein.

1.2.1 Homogeneity

Line broadening arises from the inhomogeneous nature of the in-cell samples. First, the nonuniform cell distributions in these densely packed samples degrade the homogeneity of the magnetic field, resulting in shim/pulse imperfections.[23, 46] Second, the cytoplasm is a highly anisotropic and organized environment.[47, 48] This “cellular anisotropy” also leads to inhomogeneous broadening that can vary from resonance to resonance in the test protein.[7]

It is difficult, however, to separate the contribution of heterogeneity to those from chemical exchange and viscosity. Fortunately, the difference in resonance widths between ^{15}N transverse relaxation optimized spectra (TROSY) and anti-TROSY spectra, $\Delta\Delta\upsilon^{TAT}$, is independent of chemical exchange and sample inhomogeneity.[49] The Gierasch group performed glycerol titrations on the purified protein G B1 domain (GB1, 6 kDa) to examine the viscosity dependence of $\Delta\Delta\upsilon^{TAT}$ and found a linear relationship.[44] The apparent intracellular viscosity in the *E. coli* cell can be estimated from this linear relationship. Viscosity also can be estimated from $^{1}H^{N}$ linewidths.[50] Their data indicated that the viscosity estimated from $\Delta\Delta\upsilon^{TAT}$ is 30% lower than the viscosity estimated from the $^{1}H^{N}$ linewidths. This result implies that heterogeneity arising from the intracellular environment accounts for <30% of the resonance broadening. Therefore, tumbling-related effects appear to dominate line broadening.

1.2.2 Global Tumbling

Factors that affect the rate of global tumbling include molecular weight, viscosity, and nonspecific interactions.[44, 51-54]

1.2.2.1 Molecular Weight

One potential obstacle to solution NMR is molecular weight. The smaller the protein, the easier it is to detect by NMR because smaller proteins tumble more quickly than larger proteins. The globular protein tumbling rate is also known as global correlation time, τ_c. When expressed in ns, τ_c in dilute solution is approximately 0.6 times of the molecular weight (kDa) of the protein.[50] Proteins less than 30 kDa can provide quite decent ^{1}H-^{15}N heteronuclear single quantum coherence (HSQC) spectra in dilute solution, but not in cells. On the other hand, disordered proteins, such as α-synuclein (14 kDa) and FlgM (10 kDa), give high-quality spectra inside cells.[14, 55] Barnes *et al.* recently produced a histidine-tagged fusion of the globular human ubiquitin and the disordered human α-synuclein in *E. coli*. Although the apparent molecular weight of the fusion is 29 kDa, the authors obtained high-quality in-cell spectra, but only from the α-synuclein portion of the fusion.[8]

Globular proteins, however, tell a different story. The N-terminal metal-binding domain of mercuric ion reductase (NmerA, 7 kDa), the third PDZ domain (PDZ3, 7 kDa), barley chymotrypsin inhibitor 2 (CI2, 8 kDa), ubiquitin (9 kDa), and cytochrome *c* (12 kDa) are invisible in cells.[8, 12, 16, 44] In contrast, the structure of TTHA1718 (7 kDa) within *E. coli* was determined by in-cell NMR.[21, 56] Additionally, HSQC spectra of GB1 (6 kDa), and even the GB1-GB1 fusion protein have been observed inside cells.[12, 44, 57]

Taken together, these results suggest that molecular weight is not the reason why most globular proteins are undetectable in cells. As discussed in the next two sections, the intrinsic properties of intracellular environment explain the undetectability of most globular proteins.

1.2.2.2 Viscosity

The key difference between intracellular and dilute solution conditions is the high concentration of macromolecules, including proteins, nucleic acids and ribosomes in cells.[1] The intracellular viscosity that arises from these high concentrations has been reported to be between 2 and 10 times that of water.[44, 58, 59]

To examine the effect of viscosity on protein resonances, glycerol, synthetic polymers and proteins were used to mimic the viscous cellular conditions.[53, 54] A typical high-quality CI2 spectrum was obtained in solution containing 350 g/L glycerol. The result was not surprising because the viscosity of the glycerol solution is only 2.9 cP. High-quality CI2 spectra, however, were also observed in 300 g/L solutions of the synthetic polymers polyvinylpyrrolidone 40 (PVP) and Ficoll 70 (Ficoll), whose viscosities are 54 cP and 24 cP, respectively. Even though the viscosities of synthetic polymer solutions are 10 times larger than those of glycerol at similar g/L-concentrations, the line widths of CI2 resonances in synthetic polymers do not dramatically increase.[54]

Using proteins as crowding agents led to remarkably different results. The viscosities of 300 g/L solutions of lysozyme, ovalbumin, and bovine serum albumin (BSA) are all less than 5 cP. The spectra, however, were severely degraded in protein solutions, although a few backbone resonances were detected in 300 g/L BSA. A similar result was observed in cell lysates containing 200 g of protein per liter. Synthetic polymers are “inert” macromolecules, while proteins have charge on their surface. The data suggest not only that viscosity alone cannot explain the undetectability of most globular proteins, but also that the weak, nonspecific chemical interactions between CI2 and protein crowding agents fundamentally affect the protein resonances.

1.2.2.3 Nonspecific Interactions

To test the idea that weak, nonspecific interactions cause globular test proteins to be undetectable, we examined backbone ^{15}N relaxation of CI2 in the presence of different crowding agents. The longitudinal relaxation time, T_1, and the transverse relaxation time, T_2, are affected by both viscosity, τ_c, and temperature, but the product of $1/T_1$ and $1/T_2$ (R_1R_2) is constant at a given temperature and magnetic field strength when the product of the Larmor frequency and the τ_c is much greater than unity.[60] Thus, the R_1R_2 is a useful tool for assessing weak intermolecular interactions.[52]

For unbound CI2, R_1R_2 should equal 19.6 s^{-2} at 600 MHz.[52] Larger values indicate its involvement in larger assemblies. The average values of R_1R_2 for data acquired in glycerol and synthetic polymers implied that the intermolecular interactions are weak. In contrast, values from experiments with protein crowders and the cell lysate indicated that these crowders interact more strongly with CI2. These results pointed to nonspecific chemical interactions as the source of the difference between CI2 in solutions crowded by synthetic polymers and in solutions crowded by proteins. Importantly, these results also strongly suggest that nonspecific interactions limit the detectability in in-cell NMR.

Several independent observations from computational and experimental studies support this idea. Feig and Sugita examined the crowding effect by using molecular dynamics simulations of CI2 in the presence of BSA and lysozyme.[61]

Their data confirmed the experimental observation[54] that CI2 interacts with BSA and lysozyme. Another computational simulation of protein diffusion in the *E. coli* cytoplasm also highlighted the importance of protein-protein interactions.[62]

Crowley *et al.* showed that cytochrome *c* (13 kDa, *p*I 10) is undetectable by in-cell NMR.[12] Size-exclusion chromatography results indicated that the apparent molecular weight of the protein in lysates is >150 kDa, a value much too large for conventional NMR. The data suggest that cytochrome *c* interacts with *E. coli* cytolic proteins. These nonspecific interactions can be eliminated by elevated concentrations of NaCl. In addition, inverting the surface charge on cytochrome *c* allowed observation of ^{1}H-^{15}N HSQC spectrum in cells.

In summary, it appears that nonspecific interactions between test proteins and intracellular proteins are the dominant factors leading to the undetectability of most globular proteins in cells. The average isoelectric point of proteins in *E. coli* is around 6,[47] which mean that most proteins are polyanions at physiological pH (~7.6)[63]. For GB1 (*p*I 4.5), which is a negatively charged at intracellular pH, the resulting repulsive interactions enable GB1 to behaves like a monomer, and so can be observed by in-cell NMR.[12, 44, 57] For positively charged protein, such as cytochrome *c* (*p*I 10), an attractive interaction is expected. It is therefore not surprising that cytochrome *c* is undetectable because its "stickiness" increases its apparent size, increasing τ_c.[12] The introduction of charge-inversion mutations converted the attractive interactions to repulsive interactions, thereby making cytochrome *c* spectra visible in the *E. coli* cytoplasm.[12]

1.3 Potential Pitfalls

Overexpression is the most widely used approach for accumulating the millimolar concentrations of the test proteins that are required to obtain interpretable in-cell NMR spectra. As discussed above, the crowded cellular environment, however, may cause such severe resonance broadening that the test protein is undetectable despite its high concentration. Under these conditions, even small amounts of have leaked protein will cause artifacts.[45, 64, 65] In another words, the observed "in-cell resonances" might come from test protein that leaked into the cell culture media. Therefore, it is always necessary after acquiring an in-cell NMR spectrum to separate carefully the cells and the cell media by centrifugation and examine the media for the presence of the test protein.

To investigate the connection between protein leakage and in-cell NMR, Barnes and Pielak studied four proteins, human α-synuclein, *E. coli* HdeA, CI2, and human ubiquitin using *E. coli* strain BL-21(DE3).[66] The cell slurry supernatants were examined after 1.5 h and 3.0 h of induction by using the ^{1}H-^{15}N band-Selective Optimized Flip-Angle Short-Transient heteronuclear multiple quantum coherence (SOFAST-HMQC)[67] experiment. The results showed that α-synuclein and ubiquitin do not leak. HdeA and CI2 spectra, however, were visible 3.0 h postinduction, suggesting they leaked. The intracellular concentration data showed that the expression levels of HdeA and CI2 are significant higher than these of α-synuclein and ubiquitin. The results indicated that leakage becomes a problem when the test protein is expressed at concentration exceeding 50 fg/cell, which correspond to ~20% of total intracellular protein.[68]

Leakage can be avoided by using alternative expression systems and *E. coli* strains. Take CI2 for example. We did not observe CI2 resonances in cells when the protein is under the control of the araBAD promoter rather than T7 promoter.[16] Expressing CI2 using the less efficient trifluoromethyl-L-phenylalanine expression system and expressing CI2 in *E. coli* strain DH10B also prevented leakage.[16]

Leakage can also be a problem upon storage of the cells. Freezing cells prior to in-cell NMR studies has been recommended.[10] We prepared a ^{15}N-enriched ubiquitin sample for in-cell experiments then stored the sample at -20 ^{o}C overnight. The cells were thawed and used to collect an "in-cell" spectrum. The results showed that the HSQC spectrum of ubiquitin, which is undetectable in a fresh sample, is visible after storage.[16] Storing the sample at -80 ^{o}C gave the same result. Adding 10% (v/v) glycerol decreased, but did not always prevent, leakage. Cells should not be stored prior to in-cell experiments.

1.4 Future Directions

As discussed above, the complex and crowded intracellular environment slows the tumbling of test proteins, broadening their resonances. Nevertheless, the utility of in-cell NMR continues to expand thanks to new isotope enrichment and labeling techniques and magic angle spinning. Additionally, electron paramagnetic resonance (EPR) with its high sensitivity and low background has opened a new door to understanding protein biophysics in cells.

1.4.1 Specific Labeling and Enrichment

To overcome the interference of strong background signals and the line broadening caused by slow tumbling of test protein inside cells, selective labeling and enrichment strategies as well as new pulse sequences to enhance the size limitation for solution NMR are especially needed to develop for in-cell NMR. In principal, existing labeling and enrichment techniques and pulse sequences designed for studying large macromolecules can be used for in-cell NMR. ^{13}C-^{1}H HSQC spectra of [methyl-^{13}C] methionine enriched proteins have been successfully used to study calmodulin and FKBP in the bacterium cytoplasm.[19] [Methyl-^{13}C] enrichment of methyl containing amino acids and deuterium enrichment with TROSY based pulse sequences used to investigate the structure and dynamics of large protein complexes and molecular machinery should be applicable for in-cell NMR. In addition to [methyl-^{13}C] enrichment, ^{13}C-detection NMR experiments of ^{13}C, ^{15}N enriched protein also provide valuable complementary information about protein structure and dynamics in cells.[9]

^{19}F is another attractive nucleus for in-cell NMR because of its high natural abundance, high detectability, and the fact that fluorine is not found in native proteins. One dimensional ^{19}F spectra of various proteins have been obtained in bacteria[16, 69] and yeast,[70-72] even when they cannot be observed by using the conventional ^{15}N-^{1}H HSQC experiment. Given the sensitivity of its chemical shift and relaxation to conformational and dynamical changes, ^{19}F NMR will be a sensitive probe for monitoring protein-protein interactions and chemical reactions in cells.[15, 16, 20, 33]

1.4.2 Magic Angle Spinning

When target proteins are water insoluble (*e.g.*, membrane proteins and protein fibrils) or their tumbling rate is too slow for solution NMR study, solid state magic angle spinning (MAS) NMR may be a good choice for obtaining high resolution spectra in complex environments. As discussed above, ~30% of line broadening for in-cell GB1 spectra can be attributed to contribution from chemical exchange and inhomogeneity.[44] For inhomogeneous samples, line broadening caused by magnetic susceptibility and residual anisotropic interactions can be removed by MAS. To date, MAS NMR has been used to observe test protein resonances in inclusion bodies and in native cell membranes without purification.[73, 74] The feasibility of MAS NMR for *in situ* detection of the human LR11 transmembrane domain in native *E. coli* membranes were demonstrated by using ^{13}C-^{13}C homonuclear correlation experiments.[73] There was little interference from lipids and other *E. coli* membrane proteins, and approximately 50% of the resonances from transmembrane residues could be assigned. Fgp41 was also studied by MAS NMR in lyophilized whole cells by amino acid type ^{13}CO and ^{15}N enrichment of recombinant protein in inclusion bodies.[74] In both cases, advanced labeling and enrichment strategies and multi-dimensional heteronuclear MAS NMR were required to further characterize the structural and dynamical properties of these proteins in whole cells.[73, 74]

1.4.3 In-cell EPR

Electron paramagnetic resonance (EPR), also called electron spin resonance, was first observed by Zavoisky, who used it to examine transition metal complexes.[75] Since that time, especially in conjunction with spin labeling strategies, EPR has become a powerful tool for studying the structure and dynamics of nucleic acids,[76] peptides,[77] proteins[78] and viruses[79] in solution.

The non-invasive nature of EPR spectroscopy makes it an ideal approach for investigating living systems. In-cell EPR has two main advantages over in-cell NMR. First, lower concentrations can be used because EPR is much more sensitive per spin than NMR. Often μM or lower concentrations are useful, compared to mM concentrations for NMR. Second, since EPR only detects unpaired electron spins, there is no background from diamagnetic molecules.[80-83] The latter is particular important because, as mentioned above, in cell experiments are often hampered by the presence of many different cellular components.

Double electron–electron resonance (DEER) is a pulsed, two-frequency EPR technique for the determination of relatively long distances (from 1.5 to 8.0 nm) between electron spin centers, and has been used to study proteins.[84, 85] Igarashi *et al.* demonstrated the feasibility of DEER experiments for distance measurements of site-directed spin labeled proteins in *Xenopus laevis* oocytes.[82] Recently, in-cell EPR has been expanded to the study of nucleic acids, both RNA and DNA, in *Xenopus laevis* oocytes.[81, 83]

The applicability of in-cell EPR, however, is limited to relatively large cells such as the oocytes because microinjection is difficult for smaller cells. One way around this problem would be to use alternatives methods, such as cell penetrating peptides,[22] for transporting the labeled protein into cells, although the short half-lives of spin labels in the cellular environment may be limiting. Fortunately, Azarkh *et al.* evaluated the reduction kinetics of two structurally different spin labels, the five membered heterocyclic ring nitroxide PCA (3-carboxy-2,2,5,5-tetramethylpyrrolidinyl-1-oxy) and its six membered ring analog TOAC (2,2,6,6-tetramethylpiperidine-N-oxyl-4-carboxilic acid) in oocyte cell extracts.[80] The results indicated that PCA is more stable than TOAC and that the latter is a suitable spin label for in-cell EPR.

1.5 Conclusions

We are in the post-reductionist era of biochemistry,[86] where the ultimate goal of biologists is to study biomolecules in their natural cellular environment. In-cell NMR can help us reach this goal. Although there are concerns and limitations, the recent developments described here will expand the application of in-cell NMR.

Chapter 2 – In-cell ^{19}F NMR

The material in this chapter is from:

Li C, Wang GF, Wang Y, Creager-Allen R, Lutz EA, Scronce H, Slade KM, Ruf RA, Mehl RA, Pielak GJ, Protein ^{19}F NMR in Escherichia coli, *Journal of the American Chemical Society*, 2010, 132(1):321-327

(CL, GFW, YW and GJP designed research; CL, GFW, YW, RCA, EAL, HS, KMS and RAR performed research; CL, GFW, YW and GJP analyzed data; RAM provided reagents; CL, GFW, YW, and GJP wrote the paper.)

2.1 Introduction

Most proteins function inside cells under crowded and complex conditions, where the concentration of macromolecules can reach ~400 g/L.[1, 66, 85] Studying proteins in their cellular context, although difficult, is important for understanding how environments affect functions. In-cell NMR provides a means to assess protein structure, function, and interactions with other proteins, DNA, and small molecules at atomic resolution in living cells.[2, 6, 10, 14, 19, 22, 23, 31, 34, 42, 86-91] Recently, the high-resolution NMR structure of a small, 66-residues protein in the cytosol of *Escherichia coli* has been reported.[21]

The success of in-cell NMR experiments depends on overcoming several obstacles. As currently practiced, in-cell NMR in *E. coli* requires protein over expression, which may lessen its biological significance. Current practice also requires growth on nutrients enriched in NMR-active nuclei, usually ^{15}N or ^{13}C.

Normal metabolism of these nutrients causes a background spectrum that obscures signals from the protein being studied. Proteins that leak from the cell also cause artifacts.[45] Furthermore, the crowded intracellular environment broadens resonances from globular proteins, lowering the sensitivity of NMR experiments. For instance, specific methyl labeling, a technique typically used only for large proteins, was required to obtain sufficient long-range structure restraints for the small protein cited above. Augustus *et al.* showed that the repressor protein MetJ is completely undetectable in *E. coli* because of weak, non-specific DNA binding,[6] and indetectability has been reported for other globular proteins.[42]

Because natural proteins contain no fluorine, this 100% abundant spin-½ nucleus with its high sensitivity (83% of ^{1}H), spectral simplicity, and large chemical shift range is attractive for protein NMR *in vitro* and in cells.[92-95] ^{19}F in-cell NMR was first applied to detect protein mobility in the yeast *Saccharomyces cerevisiae*,[68-70] and a preliminary study has been reported in *E. coli*.[67] Here we describe detailed studies using this bacterium.

We examine one disordered and five globular proteins containing both ^{15}N and/or ^{19}F. More specifically, we incorporate the fluorinated amino acid analogue, 3-fluorotyrosine (3FY), and trifluoromethyl-L-phenylalanine (tfmF) into proteins ranging in size from 7 to 100 kDa.

2.2 Materials and Methods

2.2.1 Expression Systems

The ubiquitin (UBQ),[10] calmodulin (CAM),[96] and α-synuclein (αSYN) expression systems were gifts from Alexander Shekhtman (State University of New York at Albany), Anthony Persechini (University of Missouri, Kansas City), and Peter Lansbury (Harvard), respectively. pET28a plasmids (Novagen) containing the gene for the truncated chymotrypsin inhibitor 2 (CI2)[97] or the PDZ3[98] domain were a gift from Andrew Lee (University of North Carolina at Chapel Hill). The GFP and histidinol dehydrogensase (HDH) expression system has been described.[58, 67] For ^{15}N enrichment, the plasmids were transformed into BL-21(DE3-Gold) competent cells. The CAM, CI2, and PDZ3 transformants were spread onto Luria broth agar plates containing 60 μg/mL kanamycin, and the others were spread onto plates containing 60 μg/mL ampicillin.

2.2.2 ^{15}N Enrichment and 3FY Labeling

The procedure was similar to that described by Khan *et al.*[99] and Li *et al.*[94] Ten mL of Luria-Bertani (LB) media (10 g Bacto-Tryptone, 5 g Bacto-yeast extract and 10 g NaCl in 1 L of H_2O) containing the appropriate antibiotic were inoculated with a single colony and incubated overnight at 37 ^oC with shaking at 250 rpm. These overnight cultures were added to 100 mL of Tryptone-Yeast media (16 g/L Bacto-tryptone, 10 g/L Bacto-yeast extract, 5 g/L NaCl, 1 mM NaOH) containing antibiotic. These pre-cultures were grown with shaking at 37 ^oC until the absorbance at 600 nm (A_{600}) reached between 0.8 and 1.0. The pre-cultures were pelleted at 25 ^oC for 10 min at 1,600*g*. One L of ^{15}N-enriched M9 media[100] plus 1 mL of 1 mg/L thiamine HCl was used to resuspend the cell pellet. This culture was grown with shaking at 37 ^oC to an A_{600} of 0.4. Seventy mg of 3-

fluoro-*D*, *L*-tyrosine (96%, Lancaster), 60 mg of *L*-phenylalanine (Sigma), 60 mg of *L*-tryptophan (Sigma) and 0.5 g of N-(phosphonomethyl)glycine (96%, Sigma) were dissolved in 1 L of media. This mixture was added 30 min before induction. The induced culture was grown overnight with shaking at 37 ^{o}C.

2.2.3 tfmF Labeling

Amber stop codons (TAG) were were incorporated at the sites for the tfmF labeling by using site-directed mutagenesis (QuickChange, Stratagene) of the target genes, which are present in the arabinose-inducible expression vector, pBAD. The labeling procedure was similar to that described by Hammill *et al.*[101] A single DH10B colony containing both the appropriate pBAD and pDule-tfm-Phe vectors was picked from an ampicillin/tetracycline plate and used to inoculate 50 mL of LB media containing 100 mg/L ampicillin and 25 mg/L tetracycline. The culture was grown overnight at 37 °C with shaking at 250 rpm. A 2.5 mL sample of the saturated overnight cultures was added to 500 mL of warm arabinose autoinduction medium[101] and the culture was shaken (250 rpm) at 37 °C for 1 h. tfmF was added to a final concentration of 1 mM after 30 min from a 100 mM stock solution prepared by dissolution in 20 mM NaOH. The cultures were shaken at 37°C for additional 40 h (A_{600} ~5). Cells were harvested by centrifuging at 1,200*g* for 20 min. For in-cell NMR studies, 100 mL cultures were centrifuged at 1,200*g* for 20 min, washed twice with 100mL of LB and resuspended in 1 mL of LB.

2.2.4 Purification

α-Synuclein. The protein was purified as described,[102] except that the freeze-thaw step was eliminated. Purity was assessed by using SDS-PAGE and its expected molecular weight was confirmed with mass spectrometry (NanoESI-MS).

PDZ3. The ^{15}N-enrichment procedure was similar to that described by Serber *et al.*[103] Luria Bertani (100 mL) media containing 60 μg/mL kanamycin was inoculated with a single colony and incubated overnight at 37 ^{o}C with shaking at 250 rpm. The overnight culture was pelleted for 10 min at 1600*g* (Sorvall RC-5B, GSA rotor). One L of ^{15}N-enriched Spectra 9 media (Cambridge Isotope Laboratories, Inc.) containing the antibiotic was used to resuspend the cell pellets. This culture was grown with shaking at 37 ^{o}C to an A_{600} nm of 0.8. Expression was induced with isopropyl-β-D-thiogalactoside at a final concentration of 1 mM, and allowed to proceed for four h.

2.2.5 Preparing for in-cell NMR

Cultures (usually ~100 mL) were centrifuged at 1,200*g* for 30 min at room temperature. The cell pellets were resuspended in 2 mL of LB media. The samples, comprising 90:10 mixtures of cell slurry:D_2O, were placed in 5 mm NMR tubes for data acquisition. Supernatants were collected by centrifugation (Eppendorf, model 5418, 2,000*g* for 10 min) after the experiments to assess leakage.[45] The pellets were resuspended in buffer (50 mM Tris, pH 8.0) to a final volume of 1 mL. Lysates were made from the resuspended pellets by sonication (Fisher Scientific, Sonic Dismembrator Model 500) on ice for 10 min with a duty cycle of 2 s on, 5 s off. The lysate was collected after centrifugation at 16,000*g* for 10 min. Viscosities were measured with a Viscolite 700 viscometer (Hydramotion Ltd., England).

2.2.6 NMR

^{15}N-^{1}H-HSQC spectra were acquired on a cold-probe equipped Varian Inova 500 MHz spectrometer at 25 ^{o}C. The ^{1}H dimension had a sweep width of 8 401 Hz and comprised 1024 complex points. The ^{15}N dimension has a sweep width of 2200 Hz and comprised 64 complex points. The data were processed with NMRPipe[104] and NMRDraw.[105] ^{19}F spectra were acquired at 37 ^{o}C on a Varian Inova 600 MHz spectrometer equipped with a 5 mm $^{19}F(H)$ z-gradient probe. The spectra comprised 128 to 2048 transients, a 30 kHz sweep width, and a 2 s delay before acquisition. ^{19}F chemical shift are referenced to trifluoroethanol at 0 ppm.

2.2.7 Protein Concentration

Purified proteins were used as standards. The concentration of each pure protein was measured spectrophotometrically [ubiquitin, ε_{280nm} = 1280 $cm^{-1}M^{-1}$;[106] PDZ3, $\varepsilon_{280\ nm}$ = 2560 $cm^{-1}M^{-1}$ [107] calmodulin, ε_{276nm} = 3300 $cm^{-1}M^{-1}$;[108] GFP, $\varepsilon_{475\ nm}$= 32500 $cm^{-1}M^{-1}$ (as reported by the manufacturer)].

For each culture, 1 mL aliquots were centrifuged at 16,000*g* for 10 min after induction. The pellets were resuspended in 20 mM potassium phosphate buffer (pH 7.5). The proteins in lysates and standards were resolved by electrophoresis on 10-20% gradient SDS polyacrylamide gels (Criterion, Bio-Rad) for 65 min at 200 V. Gels were analyzed by Coomassie staining with a VersaDoc

MP imager (Bio-Rad). Quantity-One software (Bio-Rad) was used to quantify the band intensities.

The concentration of the protein under study in the NMR tube, C_{tube}, was determined from the SDS PAGE experiment described above. Cell densities in the NMR tube, C, were determined by serial dilution and plating. The protein concentration in cells, C_{cell}, was calculated from the equation:

$$C_{cell} = \frac{C_{tube}}{C \ast V_{cell}}$$

V_{cell} is the volume of an *E. coli* cell [1×10^{-15} L[109]]. Measurements were performed in triplicate.

2.2.8 Protein Localization

Two methods, osmotic shock[110, 111] and osmotic shock plus lysozyme,[112] were used to determine the location of expressed protein.

2.3 Results

2.3.1 3FY-Labeled, ^{15}N-Enriched αSYN

There are four tyrosines in this 140 residue, intrinsically disordered protein, one at position 39 and three near the C-terminus, at positions 125,133 and 136. We labeled all these residues with 3FY. As shown in Figure 2.1A, the ^{19}F spectrum of the cell slurry shows a broad protein resonance at ~-60 ppm and a sharp resonance from free 3FY at -59.6 ppm. The assignment of the αSYN resonance was confirmed by comparison to the spectrum of the purified protein. The assignment of the free 3FY resonance was confirmed by comparison to the supernatant spectrum. The ^{15}N-^{1}H HSQC spectrum of the cell slurry (Figure

2.1B) shows numerous αSYN crosspeaks, consistent with previous work.[55, 88] To check for leakage, we subjected the cell slurry to centrifugation and examined the supernatant. The presence of only the free 3FY resonance in the ^{19}F spectrum (Figure 2.1C) and the near absence of crosspeaks in the HSQC spectrum (Figure 2.1D) indicate that little or no αSYN had leaked. The cells were then lysed by sonication, the cellular debris removed by centrifugation, and the clear supernatant examined by NMR. The ^{19}F resonances sharpened (Figure 2.1E), revealing three protein peaks that shifted upfield by ~0.1 ppm. The ^{19}F spectrum has been assigned.[94] The middle peak comprises 3FY resonances from residues 39 and 125. The crosspeaks in the HSQC spectrum of the lysate are sharper than those from the cell slurry but the spectrum is essentially unchanged. The limited chemical shift dispersion of the ^{19}F and ^{1}H resonances show that αSYN is disordered in cells, consistent with other work.[55, 88]

2.3.2 tfmF-Labeled αSYN

To overcome the incomplete resolution of the four ^{19}F resonances from 3FY labeled tyrosines in cells (Figure 2.2A), we labeled the protein with tfmF at three of the four tyrosines by using an orthogonal aminoacyl synthase system.[67] Before performing in-cell NMR experiments, we assessed the system by purifying the labeled protein (Figure 2.2B) and using mass spectrometry to confirm the expected 52 Da increase in mass, from 14461 to 14513 Da. The peak at 14555 is labeled and acetylated protein.

The in-cell ^{19}F spectra for proteins labeled at positions 39, 125 and 133 are shown as green traces in Figure 2.2D. The tfmF 39 resonance is broader than the tfmF 125 and 133 resonances in the cell slurry. The resonances from the lysates and from the purified proteins are narrower and shift upfield by ~0.1 compared to those from the cell slurry. Only a free tmfF resonance is observed in the supernatants, showing that the protein does not leak.

2.3.3 ^{19}F-Labeled, ^{15}N-Enriched UBQ

This 8 kDa globular protein has one tyrosine. Figure 2.3A shows the ^{19}F spectrum of the cell slurry. The spectrum contains a sharp free 3FY resonance and a broad protein resonance. The identity of these resonances was confirmed by comparisons to spectra of the purified protein and supernatant. Figure 2.3B shows the HSQC spectrum from the slurry. Only metabolite signals[63] are observed. Figure 2.3C and 3D show the ^{19}F and HSQC spectra from the supernatant collected immediately after the in-cell NMR experiment. Only the free 3FY resonance is present in the ^{19}F spectrum, and the HSQC spectrum is nearly blank. These observations show that UBQ has not leaked from the cells. The cells were then lysed. The ^{19}F lysate spectrum (Figure 2.3E) shows a single sharp protein resonance and the HSQC spectrum (Figure 2.3F) closely resembles that of pure UBQ.[113]

We also collected HSQC data on the globular, 11 kDa PDZ3 domain of PSD95[98, 114] in cells. Like UBQ, the PDZ3 domain is expressed at mM levels (Figure 2.4 and Table 2.1) but its HSQC spectrum cannot be obtained from the cell slurry. The protein signals, however, appear upon lysis (Figure 2.5).

We do not understand our inability to reproduce the published results on UBQ, which has been reported to yield high resolution spectra in *E. coli*.[10, 38] Our use of a different growth medium is not the reason because we obtain similar results to those shown in Figure 2.3 when we use the media described in the publications. We also tried expressing the protein at different temperatures without success. Our studies were conducted on a cold-probe equipped 500 MHz instrument. Lack of sensitivity does not explain our inability to detect UBQ in cells because we obtain the same results with a cold-probe equipped 700 MHz spectrometer.

Freezing cells prior to in-cell NMR studies has been suggested.[10] We prepared another ^{15}N-enriched UBQ sample for in-cell experiments but stored the sample at -20 ^{o}C overnight. The sample was thawed and used to collect an in-cell spectrum. The spectrum of native UBQ,[113] which is not observed in a fresh sample (Figure 2.3B), is visible in the previously frozen sample (Figure 2.6). Storing the sample at -80 ^{o}C gives the same result. Adding 10% (v/v) glycerol decreases, but does not always prevent, leakage. We conclude that cells should not be frozen if they are to be studied by using in-cell NMR.

2.3.4 ^{19}F-Labeled, ^{15}N-Enriched CI2

This 7 kDa globular protein has one tyrosine. Figure 2.7A shows the ^{19}F spectrum of 3FY-labeled cell slurry. Three resonances are observed. The sharpest resonance is from free 3FY. The other two resonances are from CI2. Both have a chemical shift of -59.2 ppm. One protein resonance is broad, with a width at half height of ~1.5 ppm. The other resonance is sharper and superimposed on the broad resonance. The HSQC spectrum of the cell slurry (Figure 2.7B) shows a spectrum almost identical to that of purified CI2.[97] The spectrum from the supernatant collected immediately after the in-cell experiment contains a resonance from both free 3FY and 3FY-labeled CI2 (Figure 2.7C). The HSQC spectrum of the supernatant (Figure 2.7D) is almost identical to the spectrum from the cell slurry (Figure 2.7A). These data show that CI2 has leaked from the cells, consistent with previous work.[55] After lysis (Figure 2.7E), only free 3FY and a single sharp resonance from the labeled protein is observed. The HSQC spectrum of the lysate is identical to the HSQC spectrum from the cell slurry. Comparing the three ^{19}F spectra suggests that the broad resonance at -59.2 ppm in the cell slurry is intracellular CI2 and the superimposed sharper resonance is from CI2 that has leaked from the cells.

2.3.5 tfmF-Labeled CI2

Figure 2.8 shows the ^{19}F spectra of CI2 labeled at positions 18 and 42 in cells and lysates. The protein resonances have a width at half height of ~0.20 ppm in cells. They shift upfield by 0.15-0.20 ppm and narrow to ~0.03 ppm upon lysis. There are no protein signals from the supernatants collected after the NMR experiments, indicating that tfmF-labeled CI2 does not leak. This result is surprising considering the results obtained from the 3FY-labeled protein (Figure 2.7). As discussed below, a lower expression level may explain the absence of leakage. The small signals near the free tfmF may be a degradation product of labeled CI2 or a tfmF metabolite.

2.3.6 ^{19}F-Labeled, ^{15}N-Enriched CAM

This 16 kDa two-lobed globular protein has two tyrosine residues. Figure 2.9A shows the ^{19}F spectrum of the 3FY-labeled cell slurry. Three resonances are evident. The sharp resonance is from free 3FY. The other two, one on either side of the 3FY resonance, are from the protein. The HSQC spectrum from the slurry (Figure 2.9B) shows only metabolite signals.[55, 63] Figure 2.9C and D show the ^{19}F spectrum and the HSQC spectrum from the supernatant collected immediately after the in-cell NMR experiment. Only free 3FY is observed in the ^{19}F spectrum, and the HSQC spectrum is nearly devoid of crosspeaks, indicating that CAM does not leak from the cells. Figure 2.9E shows the ^{19}F spectrum of the clear lysate. The broad protein resonances observed in the cell slurry narrow on lysis but the width at half height for the broadest resonance is still >0.5 ppm. The observation of CAM crosspeaks[115] in the HSQC spectrum of the lysate (Figure 2.9F) proves that detectable amounts of the protein are present. The HSQC spectrum of the lysate also show that CAM is not fully Ca^{2+} loaded.[116]

2.3.7 3FY-Labeled, ^{15}N-Enriched GFP

This 27 kDa globular protein contains 12 tyrosines. The resonances in the ^{19}F and ^{15}N-^{1}H HSQC spectra are too broad to detect in cells and lysates (Figure 2.10) but SDS-PAGE analysis and the fluorescence of the samples show that the protein is overexpressed (Figure 2.4 and Table 2.1).

2.3.8 tfmF-Labeled GFP

tfmF might be a better label for larger proteins because the trifluromethyl group adds rotational motion that is independent of molecular tumbling. The green traces in Figure 2.11 show the ^{19}F spectra of GFP labeled at position 39 and 221 in cells. The ^{19}F resonances from the two proteins in cells are broad, with widths at half height of ~0.4 ppm, but observable. The corresponding resonances from the purified protein are narrower, with widths of <0.1 ppm. The only resonance in the supernatant from the cell slurry is from free tfmF, which shows that labeled GFP does not leak. Lysis caused an upfield shift of 0.10 - 0.15 ppm.

2.3.9 tfmF-Labeled HDH

We applied the tfmF labeling method to this 98 kDa homodimer. The ^{19}F spectra are presented in Figure 2.12. As shown by comparisons to spectra for the purified protein and the supernatant, the sharp resonance in the cell slurry spectrum is from free tfmF and the broad resonance (width at half height of ~ 1.0 ppm) is from HDH. The only resonance in the supernatant is from free tfmF, showing that labeled HDH does not leak.

2.4 Discussion

We used NMR to study six proteins enriched in ^{15}N and/or labeled with ^{19}F in *E. coli* cells. The proteins are present in the cytoplasm [although some αSYN[58] and CI2 (Figure 2.13) is periplasmic]. Two ^{19}F labeling strategies were used. One strategy, incorporating 3FY in place of tyrosine, was accomplished by expressing the protein in ^{15}N-enriched minimal media containing 3FY, phenylalanine, tryptophan, and N-(phosphomonoethyl) glycine.[99] The other strategy involved an orthogonal tRNA synthase system[67] to replace residues with tfmF.

A 1D ^{19}F spectrum can be acquired in minutes (compared to an hour for ^{15}N-^{1}H HSQC spectra shown here), which allows the study of proteins near their physiological concentrations. We estimate an intracellular concentration of the tfmF-labeled proteins of 50 to 100 μM from the areas of the free tfmF and the protein resonances, the tfmF concentration in the media, and the fact that the cells occupy half the slurry volume of NMR samples. This concentration equals that of the most abundant soluble *E. coli* proteins.[117, 118] Furthermore, these ^{19}F experiments can be performed as a function of time to obtain data on signal transduction and metabolism.

^{19}F labeling is well suited to assess leaking. Controls must be performed to ensure the protein of interest is inside the cells during the NMR experiment.[45] For CI2 (Figure 2.7), we see a sharp 3FY resonance from leaked protein and a broad resonance from intracellular protein. By comparing the signal intensity in the supernatant (leaked CI2) to that in the lysate (total CI2) we estimate that 5 - 10% of the protein leaks from the cells. Importantly, this small fraction of leaked CI2 accounts for 100% of the CI2 signal in the HSQC spectrum of the cell slurry.

To assess the effect of the expression system on leakage, we repeated the experiments in the same *E. coli* strain [BL21(DE3)] with CI2 under control of the araBAD promoter[58] rather than the T7 promoter. We did not observe CI2 resonances in the HSQC spectrum from the cell slurry, but we did observe CI2 resonances in the lysate (Figure 2.14). These experiments confirm the indetectability of CI2 spectra in cells and suggest that the expression system affects protein leakage. For our pBAD experiment, however, leaked protein might not be detected because CI2 expression was also lower [~0.2 mM compared to 1 mM in BL21(DE3)]. We repeated the experiment using strain

DH10B, but expression was so low that CI2 crosspeaks were not observable even in the lysate. Additionally, in opposition to what has been recommended,[10] storing cells in the freezer should be avoided. As we have shown for UBQ (Figure 2.6), the freeze-thaw cycle disrupts a fraction of the cells, spilling the enriched protein into the surrounding dilute solution. We have shown elsewhere that encapsulating the cells controls leaking.[55] In summary, although more experiments are required to deconvolute the effects of the expression system, expression level, and strain, ^{19}F provides a straightforward assay for leakage.

^{19}F labeling extends the utility of in-cell NMR for studying intrinsically disordered proteins. For αSYN, we observe both backbone ^{15}N and side chain^{19}F signals in cells, although the resonances from the cells are broader than those from lysates and dilute samples (Figure 2.1). Because of the limited chemical shift dispersion of disordered proteins, tfmF labeling (Figure 2.2) is preferred over 3FY labeling (Figure 2.1) because any natural, ribosomally encoded, amino acid can be replaced with tfmF. Moreover, tfmF labeling provides dynamic information. For αSYN in cells, the tfmF 39 resonance is broader than the C-terminal tfmF resonances. This observation indicates constrained motion at position 39, consistent with reports that position 39 has residual structure while the C-terminal region is completely disordered.[94, 119, 120] By the same reasoning, the increased width of the tfmF 39 resonance in GPF compared to the tfmF 221 resonance (Figure 2.11) indicates that the side chain at position 221 is more mobile than the side chain at position 39. Such dynamic information is masked in dilute solution studies of purified proteins because the difference in the intrinsic line width is small in dilute solution.

Counter to the utility of ^{15}N enrichment for in-cell studies of disordered proteins, we do not observe ^{15}N signals from globular proteins in cells. Similar problems have been observed elsewhere,[6, 42] and Sakakibara *et al.* report the instance of a 7 kDa globular protein that is amenable to in-cell NMR in one *E. coli* strain, but not another.[21] We can exclude several causes for our failure to detect the HSQC spectra of the globular proteins studied here. It is not insufficient expression. The data in Table 2.1 show that the ^{15}N enriched proteins are expressed at mM levels, which should allow detection. We can also rule out overexpression because the low concentration of tfmF labeled proteins in cells (50-100 μM) relative to ^{15}N-enriched proteins (mM) still leads to broad ^{19}F resonances. Augustus *et al.*[6] showed that DNA binding explains the absence of an in-cell HSQC spectrum from the MetJ protein, but this is not a reasonable explanation for our results because the proteins are not DNA binders and all have *p*I values of 6.5 or less. We can rule out insolubility because the proteins are found in the supernatant of the lysates, not in the pellets. Strong membrane binding is also excluded because membranes are found in the pellets. The proteins also appear to be mostly, and perhaps completely, in their native states because HSQC spectra from lysates of cells expressing UBQ, CI2, CAM, apocytochrome b_5[45] and PDZ3 (Figure 2.5) are like those of the native proteins in dilute solution. Even though the HSQC spectrum of GFP is not observed in cell slurries or lysates, we know that GFP is in its native state because the samples fluoresce.

Increased viscosity in cells is one reason for our inability to observe HSQC spectra of globular proteins in *E. coli*. High viscosity slows molecular tumbling, increasing the breadth of crosspeaks, which decreases their detectability.[55] The measured viscosity of the clear supernatants from the lysates is 2 - 4 fold times that of water. Since we did not add liquid during lysis, we can use the protein concentration in the cells and the lysate (Table 2.1) to estimate that the cytoplasm is diluted 1.5 - 3.0 fold in the lysates. Combining these ranges, and assuming a direct relationship between viscosity and concentration, gives a crude estimated intracellular viscosity of 3 - 12 times that of water.

The Stokes-Einstein-Debye equation[121] predicts a direct linear correlation between viscosity and the apparent molecular size of globular proteins. If we assume this equation is valid inside cells, the apparent molecular weight of UBQ, the smallest protein studied here, would be 24 - 96 kDa. The lower value is compatible with the detection by the NMR methods we used, but the upper value is too large to yield an HSQC spectrum of the protein. This increase in apparent molecular weight will be even greater for the larger proteins. In summary, if our assumptions are valid, the increased viscosity in cells can explain our inability to detect globular proteins. At least one assumption, however, is suspect.

The Stokes-Einstein-Debye relationship breaks down in cells and in lysates because the definition of the viscosity assumes that the species increasing the viscosity (the viscogen) is infinitely smaller than the test molecule.

This definition is valid in systems comprising small viscogens (like glycerol) and a globular test protein, but it is invalid in cells where the viscogens and the test protein are approximately the same size. Such macromolecular crowding can cause negative deviation from the Stokes-Einstein-Debye law, at least when synthetic polymers are used as crowding agents.[53] That is, increases in viscosity will decrease the tumbling rate by a smaller amount than is predicted by the equation. Given this negative deviation, the apparent molecular sizes will be less than the estimates given above, providing confidence that if viscosity were the only factor, we should have observed at least the smallest protein, UBQ, in cells.

We suggest that nonspecific interactions also contribute to our inability to detect globular protein HSQC spectra in *E. coli* cells. There is precedence for the idea that weak, nonspecific interactions are a feature of the cellular interior. It was suggested in the 1930s that the cells might be highly organized, and in the 1940s the complete enzymic repertoire of the Krebs cycle was isolated as a whole.[122] Recent NMR studies indicate that 50% of the proteins in bacterial cells are completely immobile[123] and that nonspecific interactions occur *in vitro* when proteins are used as crowding agents.[52] In summary, we suggest that a combination of increased viscosity and nonspecific protein interactions explains our inability to obtain high-quality solution-state NMR spectra from globular proteins in *E. coli*.

^{19}F labeling not only facilitates leakage detection, but also overcomes the problem of detecting small globular proteins in cells (e.g., Figure 2.3, Figure 2.8). Even though the 3FY resonances in UBQ, CI2 and CAM are broad, they are detectable in cells. We attribute this detectability to the low background and high sensitivity afforded by ^{19}F and the limited number of labels in the proteins. The width of 3FY resonances in UBQ, CI2, and CAM were used to estimate rotational correlation times of ~40 -100 ns for the rotational correlation times.[99] These values represent a 10-fold increase in correlation time compared to dilute solution. Such an increase is also consistent with our inability to detect crosspeaks from ^{15}N-enriched globular proteins in cell slurries.

For larger globular proteins like GFP, even the ^{19}F resonances from 3FY are too broad to observe (Figure 2.10). A clue to overcoming this problem came from our work on disordered proteins. NMR spectra of disordered proteins are observable in cells because the disorder facilitates internal protein motions.[55] These motions are almost completely damped by the inherent order of globular proteins, such that global protein motion of globular proteins determines the line width of their resonances,[124] and hence their detectability.

We reasoned that the independent internal motion of the trifluoromethyl group of tfmF would sharpen the ^{19}F resonances, thereby facilitating the detection of larger globular proteins. This prediction is borne out. Resonances from tfmF-labeled GFP and HDH are observed in cells (Figure 2.11 and Figure 2.12). Furthermore, as expected, the ^{19}F peak width of tfmF is molecular weight dependent, increasing from 0.2 ppm for CI2 to 0.5 ppm for GFP, to ~0.6 ppm for the 48 kDa homodimer, nitroreductase,[67, 101] and ~1 ppm for the 98 kDa homodimer, HDH[67, 101] (Figure 2.12). This idea of increased sensitivity via increased internal motion is used in the labeling of methyl groups with ^{13}C,[19] but tfmF offers the advantage that the labeled compound is less susceptible to metabolic scrambling.

The chemical shift of ^{19}F is sensitive to its environment. This sensitivity is readily seen upon cell lysis. Lysis causes a ~0.1 ppm upfield shift of resonances from the intracellular protein for every protein investigated. An upfield shift was also observed for small ^{19}F containing molecules by Xu *et al.*[125] By comparing the ^{31}P to ^{19}F shifts for compounds containing both nuclei, these authors showed that the difference between extracellular and intracellular compounds arises because of differences in protein hydration inside and outside of cells. These differences in hydration may help explain the changes in protein stability in cells compared to the dilute solution.[126, 127]

2.5 Conclusion

The high viscosity and weak interactions in the cytoplasm can make routine ^{15}N enrichment a poor choice for in-cell NMR studies of globular proteins in *E. coli*. We demonstrated that ^{19}F labeling is a suitable labeling method for studying not only globular proteins but also disordered proteins in cells with NMR. The ^{19}F chemical shift and line width provides site-specific structural and dynamics information in cells. In addition, we have shown that the increased motion of tfmF expands the application of in-cell NMR to larger globular proteins. Finally, the decreased rotational motion of globular proteins suggests that high resolution magic angle spinning[128] might be well suited for in-cell NMR.

2.6 Tables

Table 2.1 Cell and protein concentration

Protein	MW, kDa	Cell concentration, cells/mL x 10^{-11}	Protein concentration, mM NMR tube	cells
ubiquitin	8.5	6.3	1.8	2.9
PDZ3	10.8	3.1	1.3	4.2
calmodulin	16.8	4.9	2.3	4.7
GFP	26.9	5.9	1.2	2.0

2.7 Figures

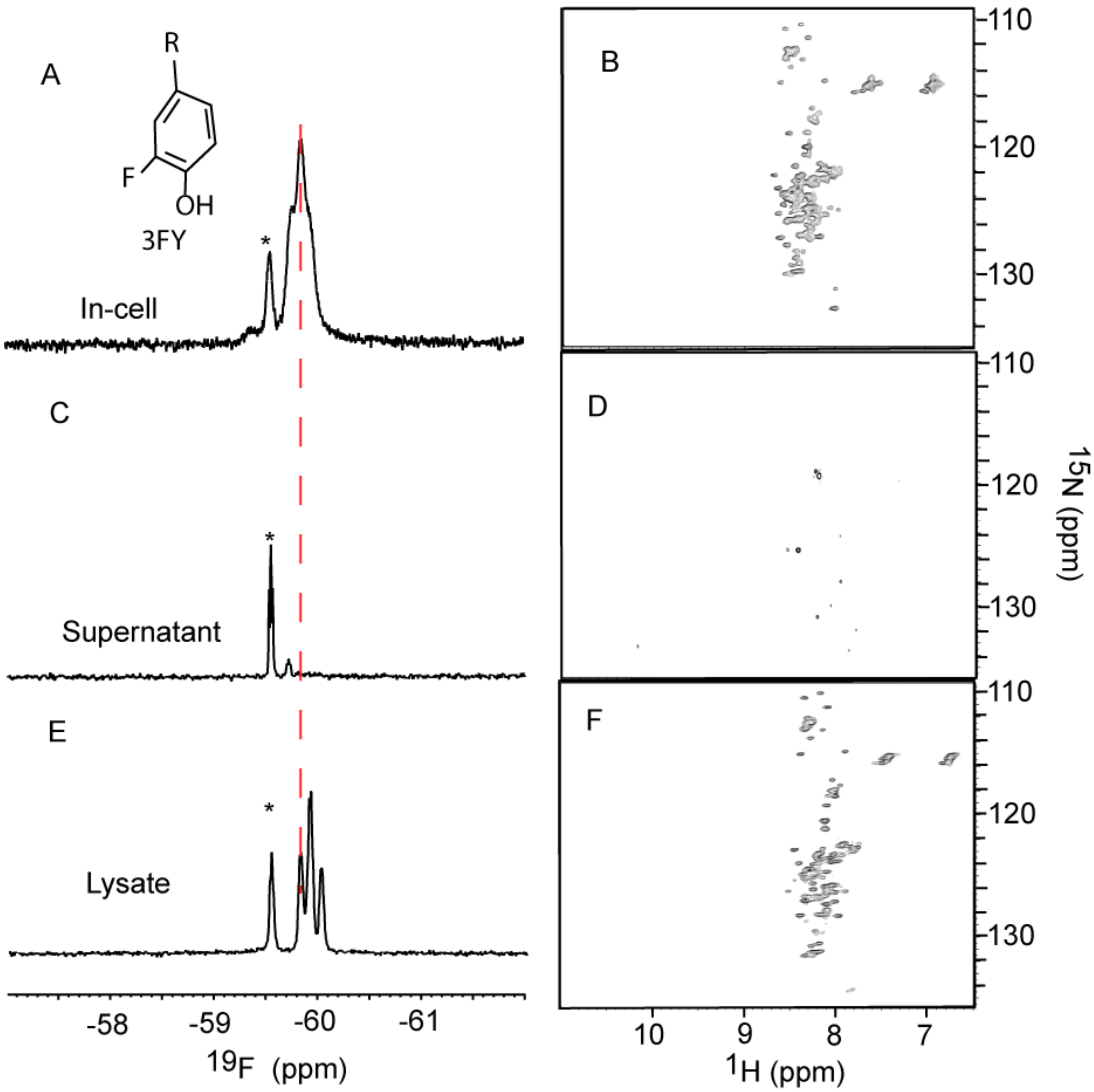

Figure 2.1 ^{19}F- and ^{1}H-^{15}N HSQC- spectra of ^{15}N-enriched, 3FY labeled α-synuclein.

Panels A and B show in-cell spectra. The inset in panel A shows the structure of 3FY. Panels C and D show spectra of supernatants collected immediately after completing the in-cell spectra. Panels E and F show spectra of supernatants from the clear lysates. The asterisks indicate the free 3FY resonances. The dashed vertical line shows the upfield shift on cell lysis.

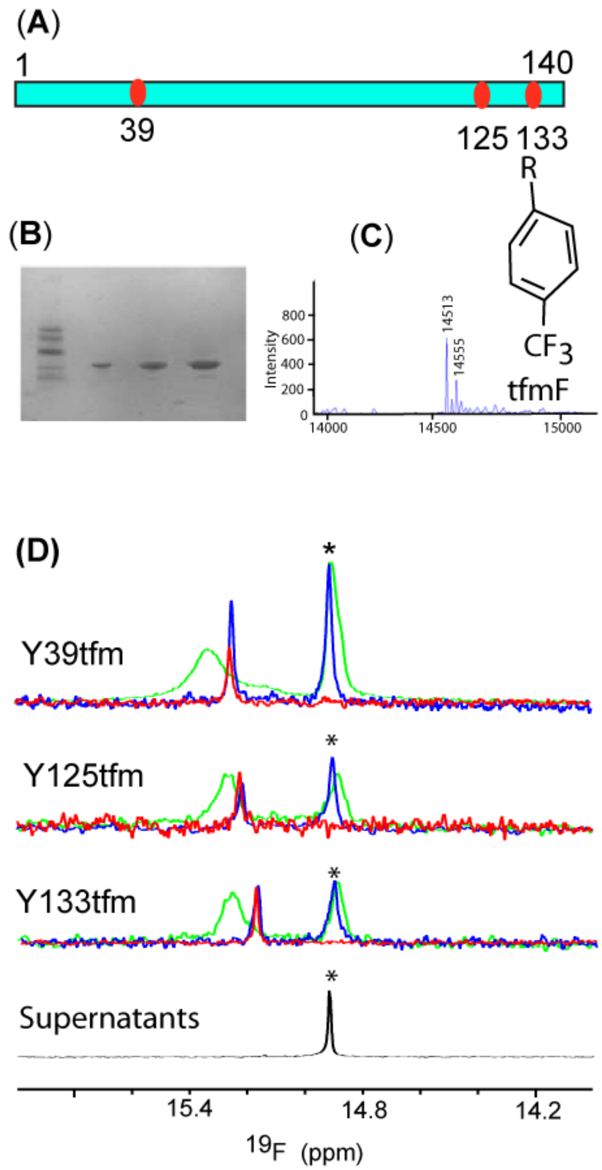

Figure 2.2 tfmF-labeled α-synuclein.

Sites of tfmF incorporation in a-synuclein (A). SDS-PAGE of the three purified tfmF labeled α-synucleins (B). ESI-mass spectrum of tfmF39 labeled α-synuclein (C). The inset shows the structure of tfmF. ^{19}F spectra of labeled synuclein (D). Spectra from cell slurries are shown in green. Spectra from clear lysates are shown in blue. Spectra from purified tfmF-labeled proteins are shown in red. Spectra from supernatants collected immediately after the in-cell NMR experiments are shown in black. The asterisks indicate the free tfmF resonances.

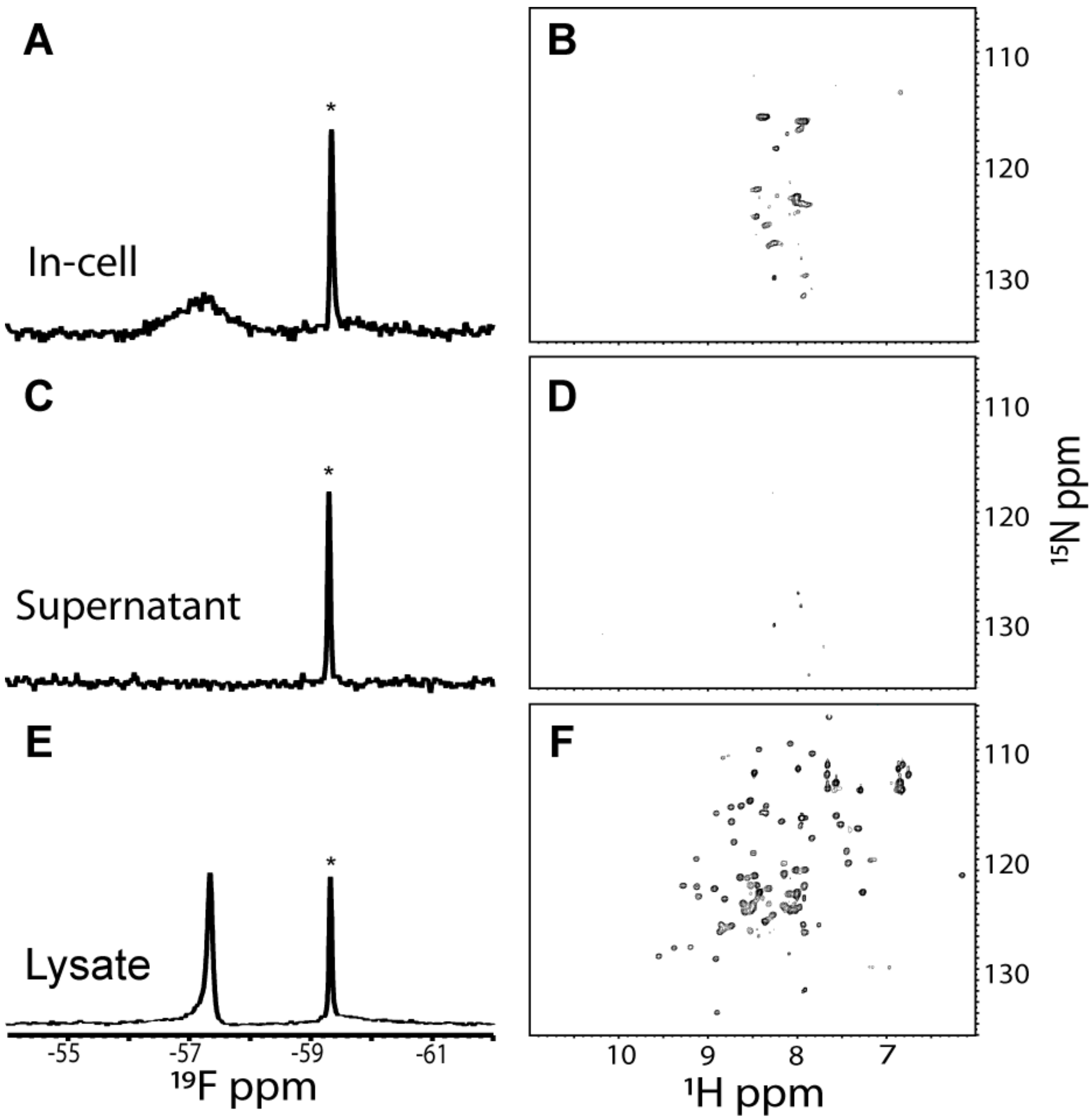

Figure 2.3 ^{19}F- and ^{1}H-^{15}N HSQC- spectra of ^{15}N-enriched, 3FY-labeled ubiquitin.

The panels are labeled as described in the caption of Figure 2.1.

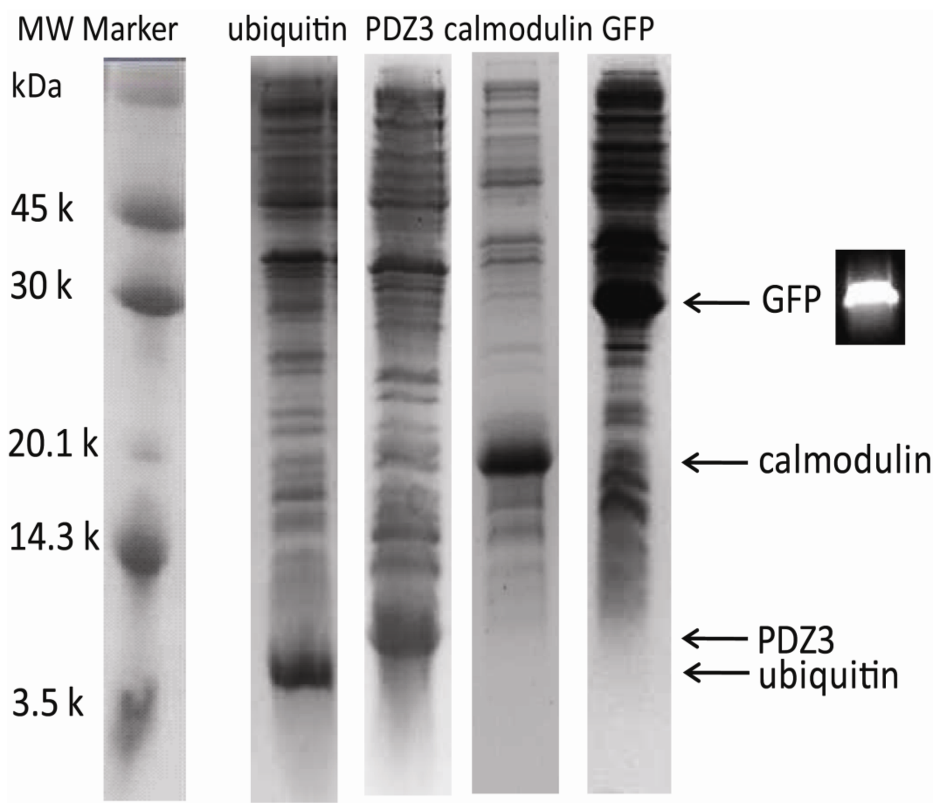

Figure 2.4 SDS-PAGE of protein expression level in cells.

Cell lysates were separated on an 18% gel and visualized with Coomassie staining. GFP was also visualized by using fluorescence.

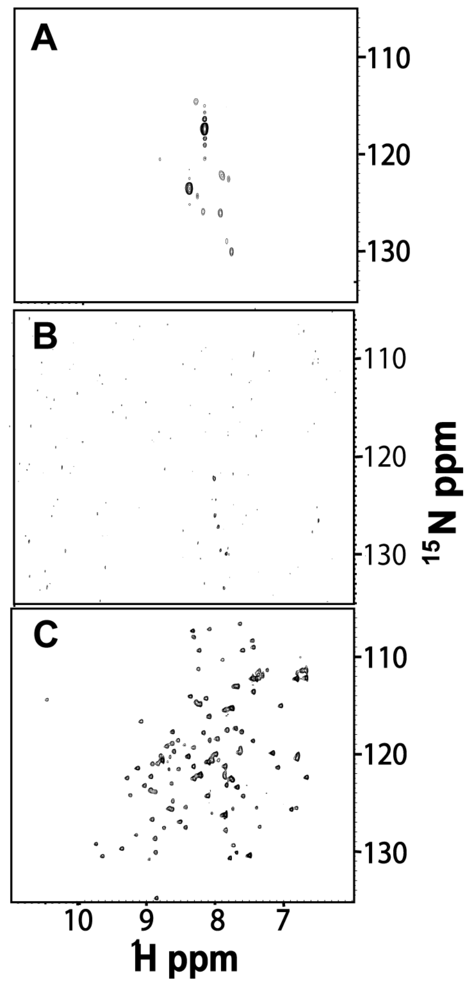

Figure 2.5 ^{1}H-^{15}N HSQC- spectra of ^{15}N-enriched PDZ3

Cell slurry (A). Supernatant collected immediately after completing the in-cell spectrum (B). Supernatant from the cell lysate (C).

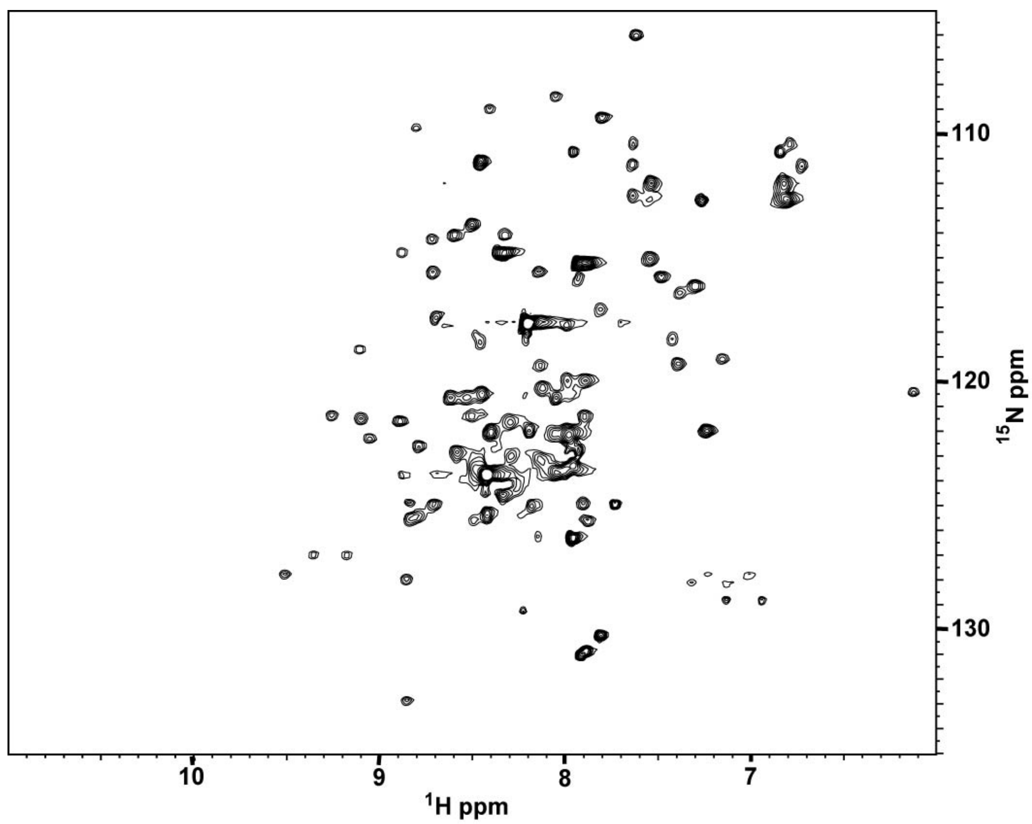

Figure 2.6 ^{1}H-^{15}N HSQC- spectra of an in-cell ubiquitin sample after storage at -20 ^{o}C overnight.

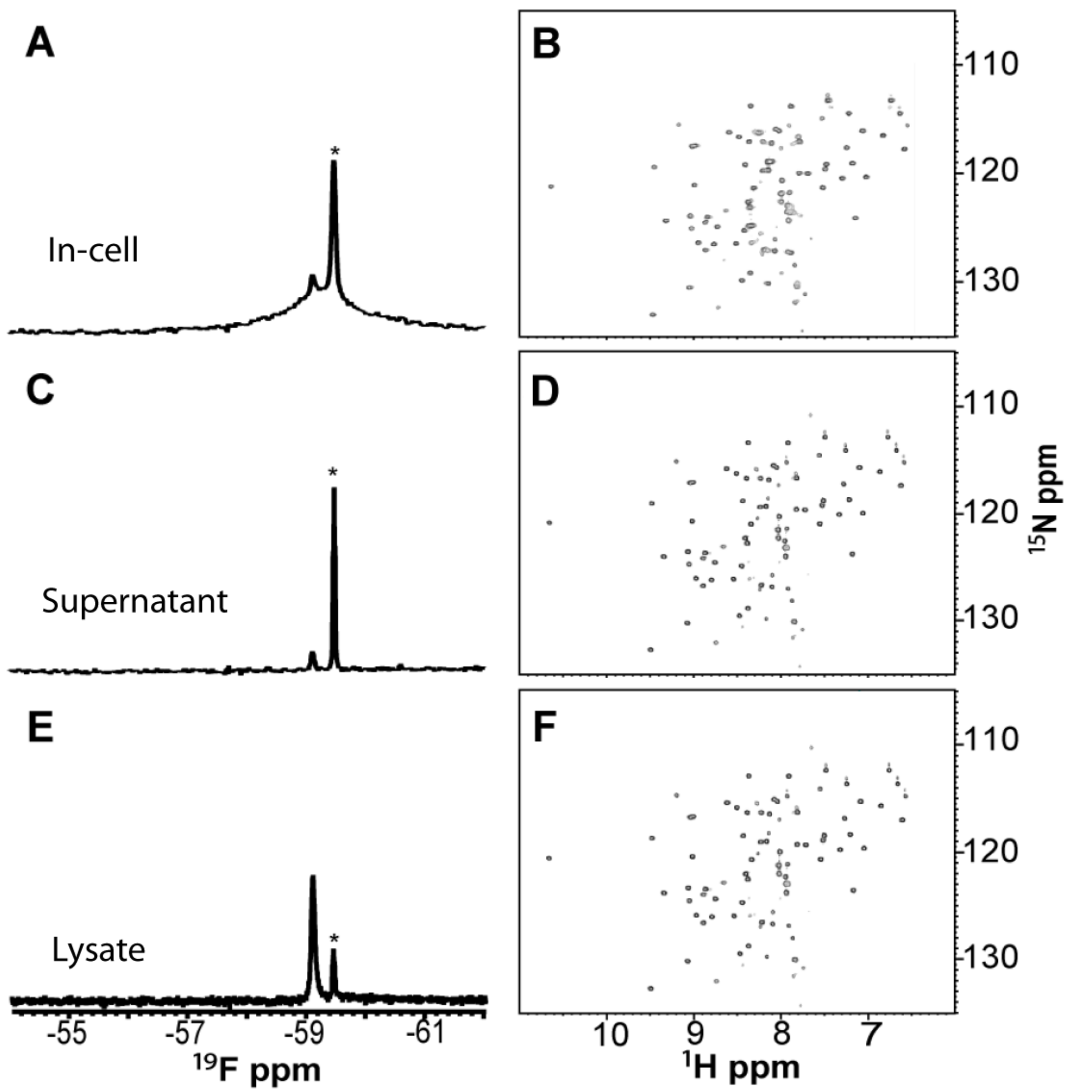

Figure 2.7 ^{19}F- and ^{1}H-^{15}N HSQC- spectra of ^{15}N-enriched, 3FY-labeled CI2. The panels are labeled as described in the caption of Figure 2.1.

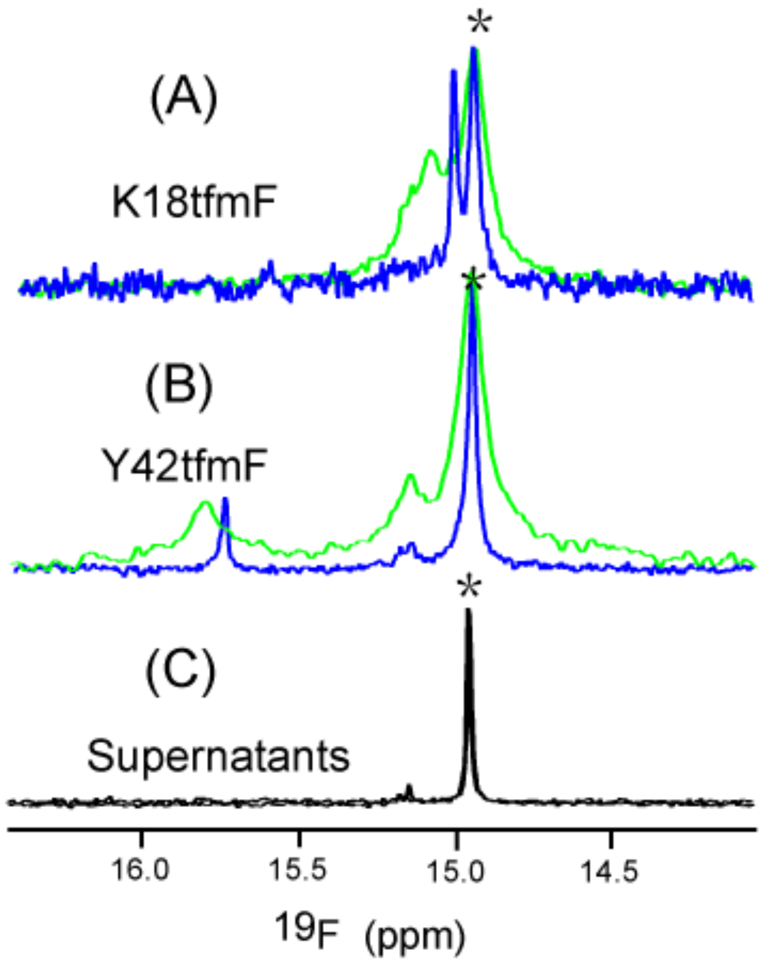

Figure 2.8 ^{19}F spectra of tfmF labeled CI2.

^{19}F spectra of K18tfmF CI2 in cells (green) and lysates (blue)(A), K42tfmF CI2 in cells and in lysates (B), and supernatant collected after the in-cell NMR experiments (C). The asterisks indicate the free tfmF resonances.

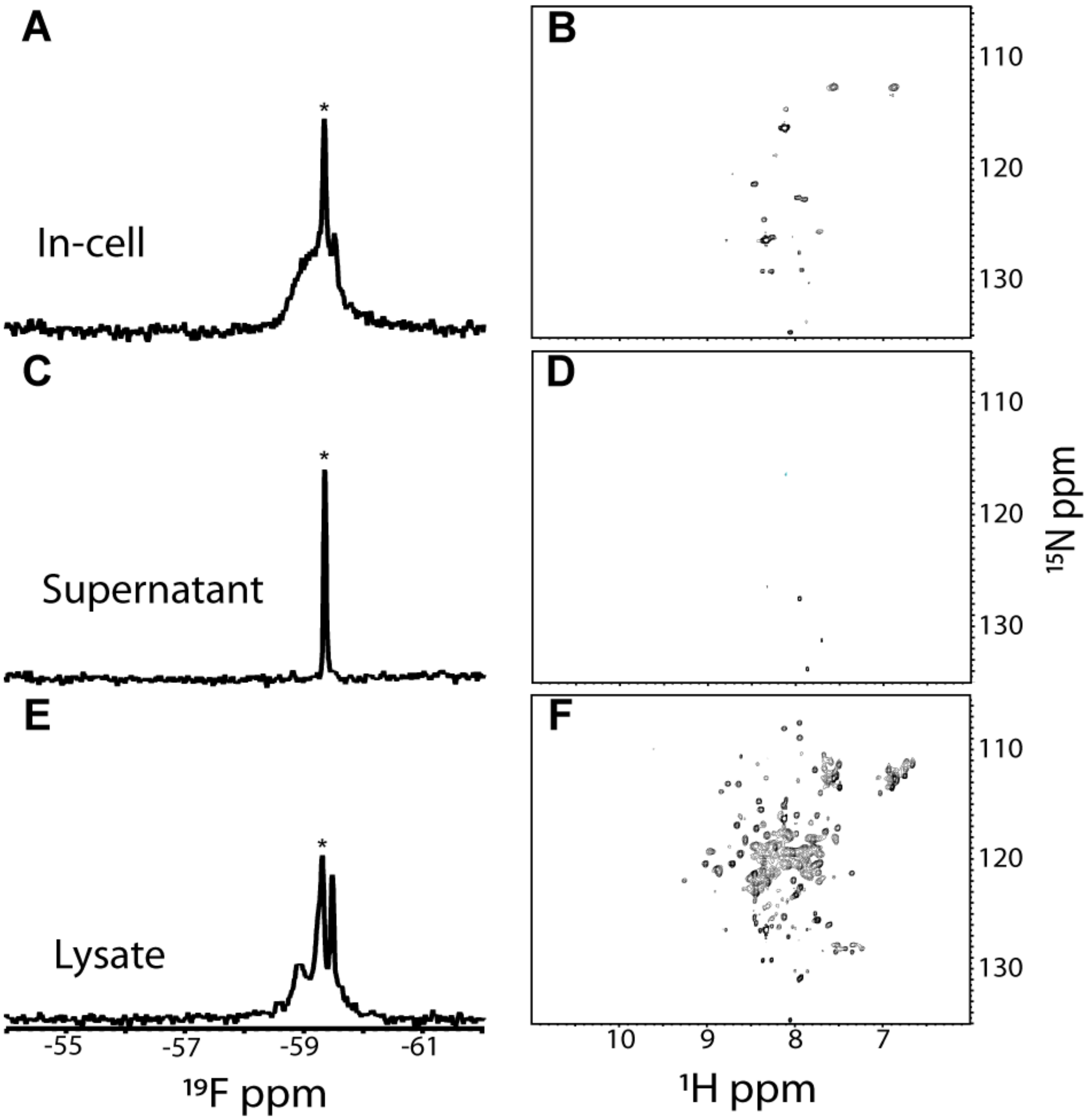

Figure 2.9 ^{19}F- and ^{1}H-^{15}N HSQC- spectra of ^{15}N-enriched, 3FY-labeled CAM. The panels are labeled as described in the caption of Figure 2.1.

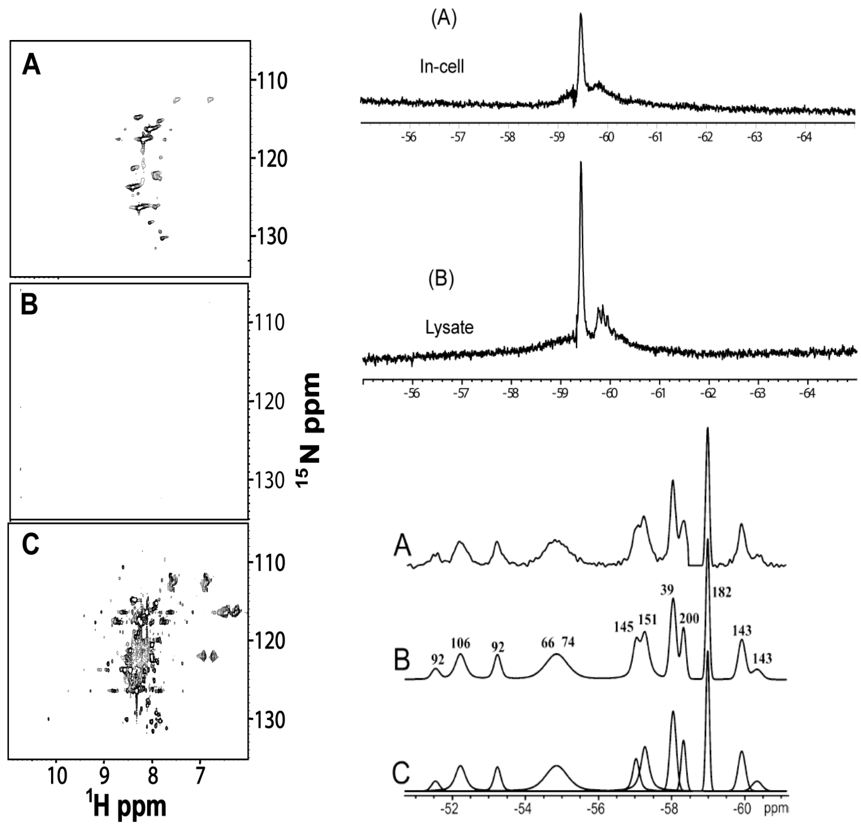

Figure 5. (A) ^{19}F spectrum (564 MHz) of GFP with Lorentzian-to-Gaussian resolution enhancement. (B) Line-fitting. (C) Individual lines in the fitting. The narrow resonance of free 3-fluorotyrosine at −58.7 ppm has been zeroed to facilitate least squares line-fitting. The assignments of Y39 and Y200 are tentative, as described in the text.

Figure 2.10 GFP data.

Left column: ^{1}H-^{15}N HSQC- spectra of ^{15}N-enriched GFP. Cell slurry (A). Supernatant collected immediately after completing the in-cell spectrum (B). Supernatant from the cell lysate (C).
Right column: ^{19}F spectra of 3FY-labeled GFP in cells (A) and the cell lysate (B). The in vitro spectrum of 3FY GFP [from[99]] is shown below panel B.

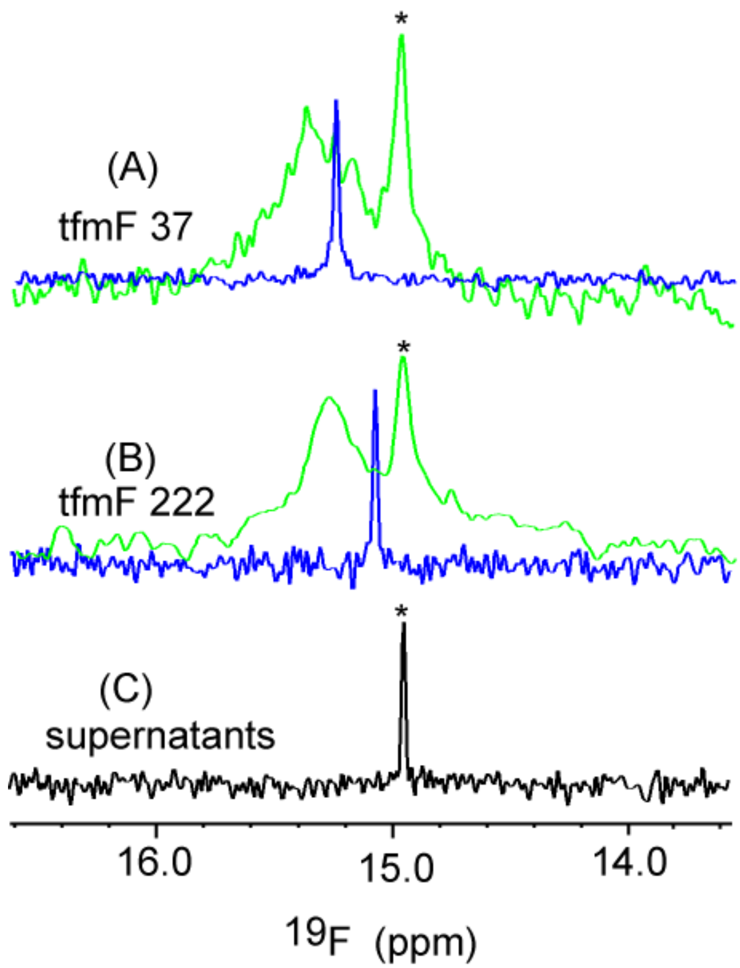

Figure 2.11 ^{19}F spectra of tfmF-labeled GFP.

^{19}F spectra of tfmF 39 labeled GFP in cells (green) and in purified protein in solution (blue) (A), tfmF 221 labeled GFP in cells and purified protein in solution (B), and in the supernatants collected after the in-cell NMR experiments (C). The asterisks indicate the free tfmF resonances.

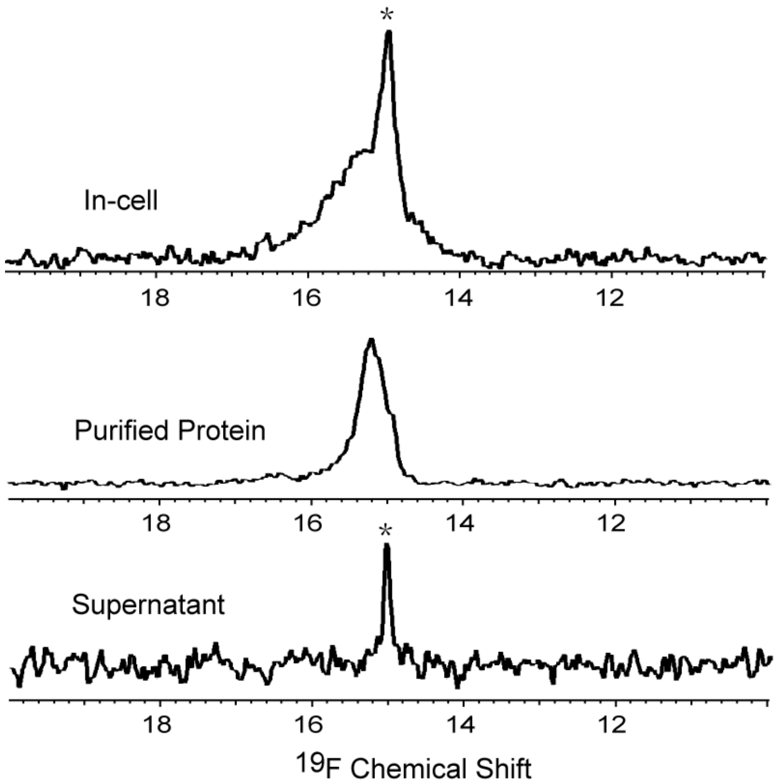

Figure 2.12 ^{19}F spectra of tfmF-labeled histidinol dehydrogenase.

^{19}F spectra of L225tfmF histidinol dehydrogenase. In-cell sample (A), purified protein (B), supernatant collected after the in-cell NMR experiments (C). The asterisks indicate the resonance from free tfmF.

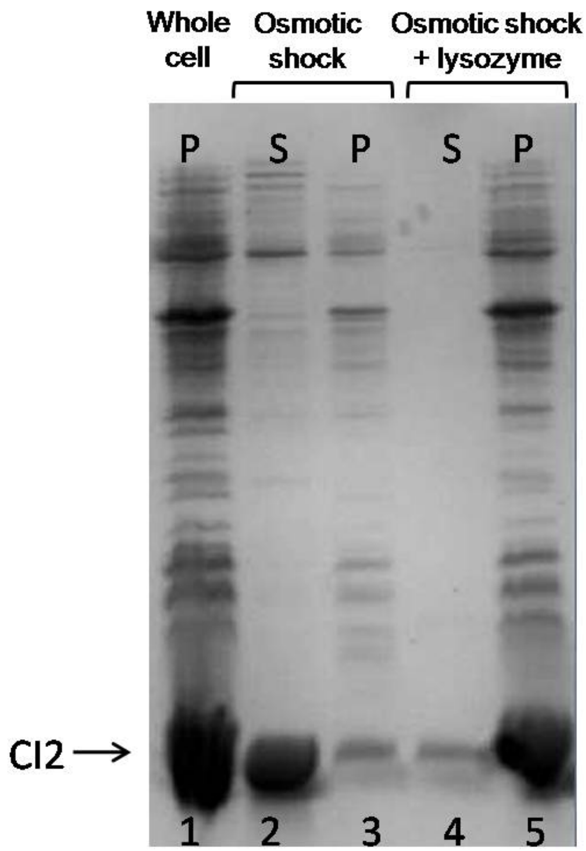

Figure 2.13 Protein location of CI2.

Aliquots of *E. coli* [BL-21 (DE3)] expressing CI2 were centrifuged and the pellets exposed to osmotic shock (lanes 2, 3) and osmotic shock plus lysozyme (lanes 4, 5). The pellets (P) and supernatants (S) were resolved by SDS-PAGE (18% gel) with Coomassie staining. Lane 1 is the untreated cell lysate. Proteins in the supernatants are periplasmic. Proteins in the pellets are cytoplasmic.

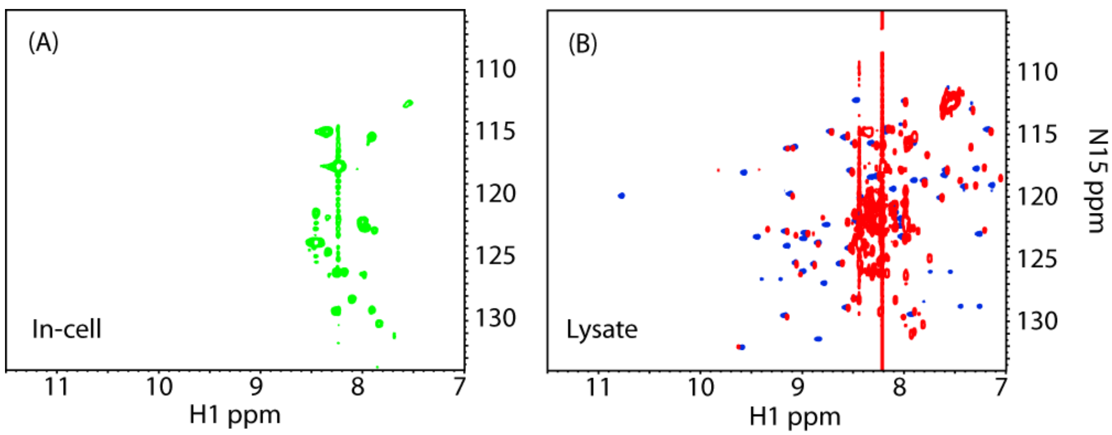

Figure 2.14 ^{1}H-^{15}N HSQC- spectra of ^{15}N-enriched CI2 expressed from the pBAD promoter in BL21(DE3) cells.

Cell slurry (A). Cell lysate (Red) and Purified CI2 (blue) (B).

Chapter 3 – Effects of Proteins on protein diffusion

The material in this chapter is from:

Wang Y, Li C, Pielak GJ, Effects of Proteins on protein diffusion, Journal of the American Chemical Society, 2010, 132(27):9392-9397

(YW and GJP designed research; YW performed research; YW, CL and GJP analyzed data; YW and GJP wrote the paper.)

3.1 Introduction

Protein diffusion affects many aspects of cell biology, from metabolism to signal transduction. The intracellular environment, however, is complex and difficult to study directly. Most work is performed in solutions where the total protein concentration is less than 10 g/L. These dilute solutions give optimal signals, but may lack biological relevance. Macromolecules occupy up to 30% of a cell's volume and reach concentrations of 100 to 400 g/L.[66] Such large volume occupancies affect protein stability,[97] folding,[129, 130] and aggregation,[131] but only recently has attention been directed to the effects of macromolecular crowding on protein diffusion.[132, 133] Furthermore, many studies of macromolecular crowding use synthetic polymers rather than natural proteins.

Diffusion is described by the Stokes-Einstein Law, $D_t = \kappa T/6\pi\eta r$, and the Stokes-Einstein-Debye law, $D_r = \kappa T/8\pi\eta r^3$, where D_t is the translational diffusion coefficient, D_r is the rotational diffusion coefficient, η is the solution viscosity, k is the Boltzmann constant, and r is the radius of protein being studied. These relationships are based on the assumption that the protein is much larger than the molecule used to increase the viscosity.[121, 134, 135] High concentrations of macromolecules are expected to cause deviations from the Stokes Laws as the macromolecules approach the size of the test protein. Deviations come in two forms. Negative deviation means that increased viscosity decreases diffusion less than predicted, and positive deviation means that increased viscosity decreases diffusion more than predicted. Studies of protein diffusion with synthetic polymers as crowding agents show negative deviation for both translational and rotational diffusion.[136-141] For protein diffusion in protein solutions, most efforts have focused on translation, where both positive deviation[142, 143] and negative deviation[136] have been observed.

The ability to detect a protein by using NMR spectroscopy depends on its rotational dynamics, which are reflected in the protein's rotational correlation time (τ_c). Increasing the viscosity or the protein size increases the τ_c, resulting in a longer longitudinal relaxation time, T_1, and a shorter transverse relaxation time, T_2. Long T_1 values decrease the sensitivity of experiments and short T_2 values broaden the resonances.[16, 42, 55] Since D_r is proportional to $1/\tau_c$, rotational motion is reflected in the width of its resonances.

Here, we use NMR spectroscopy to quantify both the rotational and translational diffusion of a 7.4 kDa ^{15}N-enriched globular protein, chymotrypsin inhibitor 2 (CI2), as a function of crowder concentration. These crowders include the glycerol, synthetic polymers, globular proteins, and *Escherichia coli* cell lysates. We find that proteins and synthetic polymers have dramatically different effects on CI2 diffusion. The difference is caused by weak interactions between the proteins that dramatically decrease the rotational motion of CI2. The results not only provide new information about protein diffusion under physiologically relevant conditions but also explain the difficulty in obtaining in-cell NMR spectra of globular proteins[16, 42] and suggest that synthetic polymers are not suitable systems for assessing the biological effects of crowding.

3.2 Materials and Methods

^{15}N-enriched CI2 was expressed and purified as described.[55, 97] Chicken lysozyme, chicken ovalbumin, bovine serum albumin (BSA), Ficoll 70 (Ficoll) and polyvinylpyrrolidone 40 (PVP) were purchased from Sigma-Aldrich and used without further purification. Viscosities were measured with a Viscolite 700 viscometer (Hydramotion Ltd., England). Glycerol, PVP and Ficoll were dissolved in 50 mM sodium acetate (pH 5.4). A more concentrated buffer was required for proteins crowders. Lysozyme, ovalbumin and BSA were dissolved in 200 mM sodium acetate (pH 5.4).

3.2.1 *E. coli* Lysates

Cultures of strain BL21 (DE3) Gold (Stratagene) containing an empty pET28a plasmid (Novagen) were grown at 37 ^{o}C with shaking in a New Brunswick Scientific I26 incubator at 250 rpm in 12, 250-mL Erlenmeyer flasks, each containing 100 mL of Luria-Bertani (LB) media (10 g Bacto-Tryptone, 5 g Bacto-yeast extract, and 10 g NaCl in 1 L of H_2O) and 50 μg/mL kanamycin.

Each overnight culture was diluted into 1 L of LB media containing 50 μg/mL kanamycin. After 12 h at 37 °C with shaking at 250 rpm, the cultures were harvested by centrifugation at 1200*g* (Sorvall RC-3B, H6000A) for 30 min at 4 °C. The pellets were stored at -20 °C overnight. Pellets were resuspended in 10 mL of distilled and deionized water. The suspensions were sonicated (Fisher Scientific, Sonic Dismembrator Model 500) on ice for 10 min with a duty cycle of 2 s on, 2 s off. The lysate was collected after centrifugation at 15000*g* (Sorvall RC-5B, SS-34) for 30 min and lyophilized (Labconco, 7740020). The protein concentration in the re-dissolved lysates (pH 7.4) was determined with a modified Lowry assay (Thermo Scientific).

3.2.2 Relaxation and Diffusion

The experiments were performed on a 600 MHz Varian Inova spectrometer equipped with a standard triple resonance HCN probe with three axis gradients at 25 °C. The relaxation and diffusion experiments were performed as described.[52, 53] Briefly, translational diffusion was measured by using a heteronuclear stimulated echo sequence.[144] Gradient strengths ranged from 1.2 G/cm to 58.0 G/cm. Rotational diffusion was assessed from the ^{15}N T_1/T_2 ratio acquired with pulse sequences from the Biopack software supplied with the instrument.[145] The ^{1}H dimension was acquired with a sweep width of 12000 Hz and comprised 1024 complex points. The ^{15}N dimension was acquired with a sweep width of 2500 Hz and comprised 64 complex increments. For T_1 measurements in solutions of 50 and 100 g/L crowders, the relaxation delays were 0.01, 0.4, 0.6, 0.7, 0.9, and 1.2 s. Delays of 0.01, 0.3, 0.4, 0.6, 0.9, 1.2, and 1.5 s were chosen for 200 g/L, and delays of 0.01, 0.4, 0.6, 0.9, 1.3, and 1.8 s were used for 300 g/L. For T_2 measurements in solution of 50 and 100 g/L, the delays were 0.01, 0.03, 0.07, 0.09, 0.15, and 0.21 s. Delays of 0.01, 0.03, 0.07, 0.09, 0.11, and 0.19 s were used for the 200 g/L. Delays of 0.01, 0.04, 0.05, 0.07, 0.09, and 0.11 s were used for 300 g/L. Eight transients were acquired per spectrum. The data were processed with NMRPipe[104] and NMRView.[105]

3.3 Results

3.3.1 Crowders

The properties of CI2 and the crowders are given in Table 3.1. The synthetic polymers comprise PVP and Ficoll. PVP is a random coil polymer.[146] Its backbone structure is shown in Figure 3.1. Ficoll, a cross-linked and branched derivative of sucrose, is more globular.[147] The proteins include BSA, ovalbumin, and lysozyme.

3.3.2 Spectra

^{15}N-^{1}H heteronuclear single quantum correlation (HSQC) spectra of CI2 were acquired in aqueous solutions containing 350 g/L glycerol and 300 g/L synthetic polymers, proteins, and in rehydrated *E. coli* lysate. Different crowders have different effects on the spectra. A typical high quality spectrum[52] was obtained in glycerol (Figure 3.1A). High-quality spectra were also observed in 300 g/L solutions of the synthetic polymers PVP and Ficoll (Figure 3.1B and C). The effect of protein crowders of increasing size (Table 3.1) is shown in Figure 3.1D-F. Low quality spectra were obtained in 300 g/L BSA, and only side-chain resonances from mobile asparagines and glutamines were observed in lysozyme, ovalbumin and cell lysate (Figure 3.1G).

3.3.3 Diffusion Data

The pulsed-field gradient experiment used to quantify D_t[144] makes no assumption about CI2 size. The method to assess rotational diffusion [i.e., T_1/T_2[145]] relies on the assumptions that CI2 is rigid and can be treated as a sphere. The first assumption is known to be valid.[148] Inspection of the structure shows that CI2 has the shape of a typical globular protein,[149] and, as discussed below, NMR data indicate it can be treated as a sphere.

Figure 3.2 shows the ratios of the diffusion coefficient in buffer (D_b) to that under crowded conditions (D_c) as a function of the relative viscosity for various crowders. In these plots, large y-values reflect a large impediment to diffusion. As expected, translational diffusion and rotational diffusion of CI2 decrease with increasing viscosity. The behavior in terms of the Stokes Laws, however, depends on the crowder. As observed previously,[53] both rotational and translational diffusion follow the Stokes Laws in glycerol (Figure 3.2A). Dividing the Stokes-Einstein-Debye equation by the Stokes-Einstein equation yields $D_r/D_t = 3/4r^{-2}$, where r is the apparent CI2 radius. Consistent with the Stokes Laws, the radius from the glycerol data, 1.7 nm, is independent of glycerol concentration and compares favorably with the 1.4 nm estimated from the molecular weight and partial specific volume of CI2. This similarity provides confidence that CI2 can be treated as a sphere. Next, we examine the effects of macromolecular crowders where diffusion can deviate from the Stokes Laws.

The synthetic polymers generate negative deviation for both translational and rotational diffusion (Figure 3.2B and C). That is, diffusion is affected less than predicted by the Stokes Laws. Furthermore, the polymers impede CI2's translational motion more than its rotational motion. Proteins have the opposite effect (Figure 3.2D-F). They cause positive deviation for rotational diffusion and either positive or no deviation for translation. Also in opposition to observations on synthetic polymers, rotational diffusion is impeded more than translational diffusion. Consistent with our conclusion that protein crowders severely impede rotation, we are unable to acquire rotational diffusion data in 300 g/L solutions of lysozyme, ovalbumin and lysates because the resonances broaden beyond detection.

To our knowledge, there is only one report on the rotational diffusion of a protein in solutions crowded with proteins.[143] In that report, the test protein apomyoglobin shows negative deviation, which is opposite to what we observe. If negative deviation were general, we would expect to observe high-quality HSQC spectra in solutions crowded with globular proteins and in cells. This expectation, however, is not fulfilled; solutions crowded by globular proteins yield poor-quality or no spectra (Figure 3.1), and none of the five globular proteins we have studied by in-cell NMR yield useful spectra.[16] Others report findings similar to ours.[6, 84] Perhaps apomyoglobin is not a good model protein because it is not completely globular.[150]

Figure 3.2G shows that diffusion in cell lysates is similar to diffusion in solutions crowded by proteins. This similarity suggests that concentrated proteins solutions are physiologically relevant models.

3.3.4 Relaxation Data

The average ^{15}N line width [$1/(\pi T_2)$] of backbone CI2 resonances in different crowders was assessed from relaxation data (Figure 3.3). The average width increases with glycerol concentration. The resonances broaden in PVP and Ficoll. The widths are larger in solutions crowded by proteins, and similar to the widths obtained in cell lysates. Linewidth, however, is affected by both viscosity and binding. The product of longitudinal relaxation rate R_1 ($1/T_1$) and transverse relaxation rate R_2 ($1/T_2$) can be made independent of viscosity (see Discussion) and is hence a good method for assessing weak binding.[52] A histogram of the average R_1R_2 values for various crowders is shown in Figure 3.4. Smaller average values are observed for glycerol and synthetic polymers than for protein crowders and the cell lysates.

3.4 Discussion

3.4.1 CI2 is Invisible in HSQC Spectra in Cells, at High Protein Concentrations, and in Cell Lysates

Even in a 350 g/L glycerol (93 Da) solution, which has a relative macroscopic viscosity of 2.9, the CI2 spectrum looks like it does in dilute solution (Figure 3.1A). The viscosities of synthetic polymer solutions are 10 times larger than those of glycerol at similar g/L-concentrations (Table 3.2), yet we still obtain typical CI2 spectra (Figure 3.1B and C).

Using proteins as crowding agents leads to dramatically different results. The spectral quality is extremely low in concentrated protein solutions (Figure 3.1D-F), despite the fact that these solutions have viscosities similar to those of the glycerol samples, and 10-fold lower than those of the synthetic polymers. The spectra are so severely degraded in BSA that only CI2 glutamine and asparagine side-chain resonances and a few backbone resonances are detected. Backbone resonances are completely absent in spectra acquired with lysozyme and ovalbumin. The side-chain resonances are observed because they have internal motion that is independent of overall rotational motion.[55] Importantly, we observe the same effect with cell lysates (Figure 3.1G), suggesting that our results are biologically relevant. Our results are also consistent with those from in-cell NMR experiments, where resonances become too broad to give useful HSQC spectra.[6, 42, 52, 53]

We cannot blame bulk viscosity for the poor quality of the spectra in protein solutions because the viscosities are far lower than those of the synthetic polymers. We also can rule out inhomogeneity as a factor because the solutions are homogenous. To understand the difference between the effects of synthetic polymers and proteins, we used NMR to quantify CI2 diffusion.

3.4.2 Synthetic Polymers and Proteins Have Opposite Effects

The synthetic polymers PVP and Ficoll are much larger than CI2 (Table 3.1). At the concentrations used here (≥100 g/L), molecules of these polymers overlap to form a mesh.[151] If the chemical interactions between the polymers and CI2 are extremely weak, we expect CI2 to experience less than the macroscopic viscosity. This expectation is borne out (Figure 3.1B and C). We also note that PVP and Ficoll slow CI2's rotational diffusion less than its translational diffusion. This result is expected because rotation in the mesh should be easier than translation through the mesh. It is interesting to compare the PVP results to the Ficoll results. The stronger deviation observed in Ficoll is expected because its molecule weight is larger (Table 3.1). It is also of interest to estimate the apparent size of CI2 from D_r/D_t as described above for glycerol solution. In 200 g/L solution of synthetic polymers, the apparent radius is 1.1 nm in PVP and 1.0 nm in Ficoll, which, assuming a partial specific volume of 0.73 mL/g, corresponds to apparent molecular weights of 4.7 and 3.6 kDa, respectively. Thus, CI2 acts like a smaller protein in solutions of synthetic polymers.

Assuming that nonspecific, noncovalent chemical interactions between the proteins and CI2 are extremely weak, the concentrated solutions of globular proteins should act like a collection of spheres. Negative deviation is also expected for these systems as long as the protein remains mobile. Inert spheres should remain mobile up to near the close-packing limit, which for practical purposes occurs at a volume occupancy of ~64%.[152] The volume occupancy here is only ~21% at the highest concentrations (300 g/L). Nevertheless, we observe not the expected negative deviation but positive deviation for rotational diffusion and positive or negligible deviation for translational diffusion for proteins solutions (Figure 3.2D-F) and in cell lysates (Figure 3.2G). This strong attenuation of rotational diffusion does not depend on the size or charge of the protein (Table 3.1), suggesting the generality of our results. We suggest that the dramatically different effects of synthetic polymers and proteins arise because of nonspecific, noncovalent chemical interactions between the proteins and CI2. We also estimated the effective size of CI2 under these conditions from D_r/D_t. In 200 g/L protein solutions, the apparent radius of CI2 is 2.5 nm in BSA and 2.4 nm in lysozyme, corresponding to apparent molecular weights of 56.7 kDa and 45.4 kDa, respectively. These apparent molecular weights are more than 7 times those calculated from CI2's amino acid sequence. The increase in size suggests that CI2 interacts with other proteins in solutions. Put another way, even weak favorable interactions between CI2 and the protein crowders should lead to the observed larger effects on rotation compared to translation because rotational diffusion depends on volume, r^3, while translational diffusion depends only on size, r.

3.4.3 Relaxation Data Indicate Nonspecific, Noncovalent Chemical Interactions Involving Proteins

NMR is useful for investigating weak protein interactions in dilute solution[153] and under crowded conditions.[52] The simplest quantitative experiment is to examine the average resonance widths under different conditions. We used T_2 data to assess line widths [$1/(\pi T_2)$]. Favorable interactions between CI2 and the crowders will broaden resonances by impeding rotation. The data in Figure 3.3 show not only that widths increase with crowder concentration, but also that protein crowders have the most dramatic effect. Lysozyme, ovalbumin, and lysates have such a strong effect that we can only estimate the widths at the highest concentration. The data are consistent with the presence of favorable CI2-crowder interactions, especially between the protein crowders and CI2. Unfortunately, width also increases with viscosity, so this method alone cannot provide definitive information on CI2-crowder interactions.

T_1 and T_2 are affected by viscosity, global correlation time, and temperature, but Kneller *et al.*[60] showed that the product of $1/T_1$ and $1/T_2$ (R_1R_2) is constant when the product of the Larmor frequency and the global correlation time is much greater than unity at a given temperature and magnetic field. In addition, the protein must lack extensive millisecond internal motion, which is known to be true for CI2.[148] This viscosity independence makes R_1R_2 a useful tool for assessing intermolecular interactions.[52]

The R_1R_2 data are shown in Figure 3.4. Provided CI2 has a rotational correlation time >7 ns (assured by the viscosity of all our samples), R_1R_2 should equal 19.6 s^{-2} at 600 MHz for unbound CI2.[52] As we have shown, R_1R_2 values from 19.6 s^{-2} to 24.0 s^{-2} are consistent with CI2 dimerization.[52] Larger values indicate involvement in larger assemblies, most likely with the crowding molecules.[52]

The average value of R_1R_2 data for glycerol and the synthetic polymers (Figure 3.4) are consistent with extremely weak interactions with CI2. Nevertheless this sensitive method indicates that interactions in PVP are stronger than interactions in Ficoll. We cannot state with certainty that these are exclusively CI2-PVP interactions, but NMR pulsed-field gradient experiments indicate that CI2 can be no more than a dimer in solutions containing 300 g/L of 40 kDa PVP.[97]

Protein crowders give different results and show that they interact more strongly with CI2. The R_1R_2 values in concentrated protein solutions and in lysates exceed those for monomeric or dimeric CI2 and depend strongly on crowder protein concentration. In summary, the data point to a nonspecific affinity of proteins for one another as the source of the difference between the diffusion of CI2 in solutions crowded with synthetic polymers and proteins. The chemical origin of these noncovalent interactions may reside in the local distribution of complementary CI2-protein charges and in the repeating nature of polypeptide amide nitrogen H-bond donors and carbonyl oxygen acceptors.

Although proteins interact more strongly with CI2 than do synthetic polymers, results of previous work show that the dissociation constant for CI2-protein complexes is large, 10 mM or greater.[52] Another indication that these are weak interactions is that the value of R_1R_2 does not depend in a predictable way on the charge of the crowding protein (Table 3.1). Most importantly, our data show that even weak protein-protein interactions severely impede rotation.

3.5 Conclusions

The intracellular environment is crowded and inhomogeneous, and weak interactions are a special and critical feature of living cells.[122] For instance, weak interactions are thought to organize metabolic paths and protein-protein interaction networks.[154, 155] The importance of weak protein-protein interactions under crowded conditions has also been highlighted in a recent computational study and a recent review of the crowding literature.[62, 156] Our study provides quantitative data supporting these hypotheses and methods for assessing weak but physiologically important interactions.

From a practical point of view, the results explain why ^{15}N-^{1}H HSQC spectra of globular proteins are difficult to detect in cells.[42, 46] Although we focused on a single protein, our difficulty in observing in-cell HSQC spectra of five globular proteins suggests that weak interactions are universal.[16] Augustus *et al.* also suggest that weak interactions between proteins and DNA result in disappearance of the MetJ spectrum in ^{15}N HSQC experiments.[6] The fact that synthetic polymer crowders and globular proteins have such different effects on diffusion suggests that synthetic polymers may not be the best choice for modeling the effects of the intracellular environment on protein diffusion.

3.6 Tables

Table 3.1 Properties of CI2 and crowders

Molecule	Molecular Weight, kDa	pI	Charge at pH 5.4
CI2	7	6.5	Cation
Glycerol	0.09	NA*	Neutral
PVP	40	NA	Neutral
Ficoll	70	NA	Neutral
Lysozyme	15	11.0	Cation
Ovalbumin	45	4.6	Anion
BSA	66	4.7	Anion

*NA: Not Applicable

Table 3.2 Translational and rotational diffusion coefficients for CI2

Crowder Crowder	g/L	Viscosity cP	Translation 10^{-7} cm^2/sec	Rotation 10^{7} rad^2/sec	R_1R_2 s^{-2}
Buffer	-	1.0	15.51	4.07	13.8
Glycerol	350	2.9	5.17	1.36	16.2
Glycerol	420	3.8	4.13	1.07	17.1
PVP	100	7.6	4.24	2.03	18.4
PVP	200	21.5	1.87	1.14	22.1
PVP	300	53.8	0.95	0.58	29.5
Ficoll	100	2.5	5.33	2.49	15.6
Ficoll	200	9.8	2.59	1.89	18.0
Ficoll	300	24.3	1.32	1.22	20.7
Lysozyme	100	1.3	5.12	0.65	33.3
Lysozyme	200	1.6	3.86	0.52	40.3
Lysozyme	300	3.9	1.39	-	-
Ovalbumin	100	1.4	9.02	1.06	23.1
Ovalbumin	200	2.0	6.76	0.82	28.8
Ovalbumin	300	4.5	3.14	-	-
BSA	100	1.5	8.18	1.05	22.3
BSA	200	2.5	5.51	0.64	36.1
BSA	300	4.8	2.66	0.36	-
Lysate	100	2.5	7.07	1.17	24.3
Lysate	200	3.6	3.97	0.58	38.7

Conditions for glycerol, PVP and Ficoll: 50 mM acetate buffer, pH 5.4, 25 ^{o}C.
Conditions for BSA, ovalbumin, lysozyme: 200 mM acetate buffer, pH 5.4, 25 ^{o}C.
The PVP data have been published.[53]

3.7 Figures

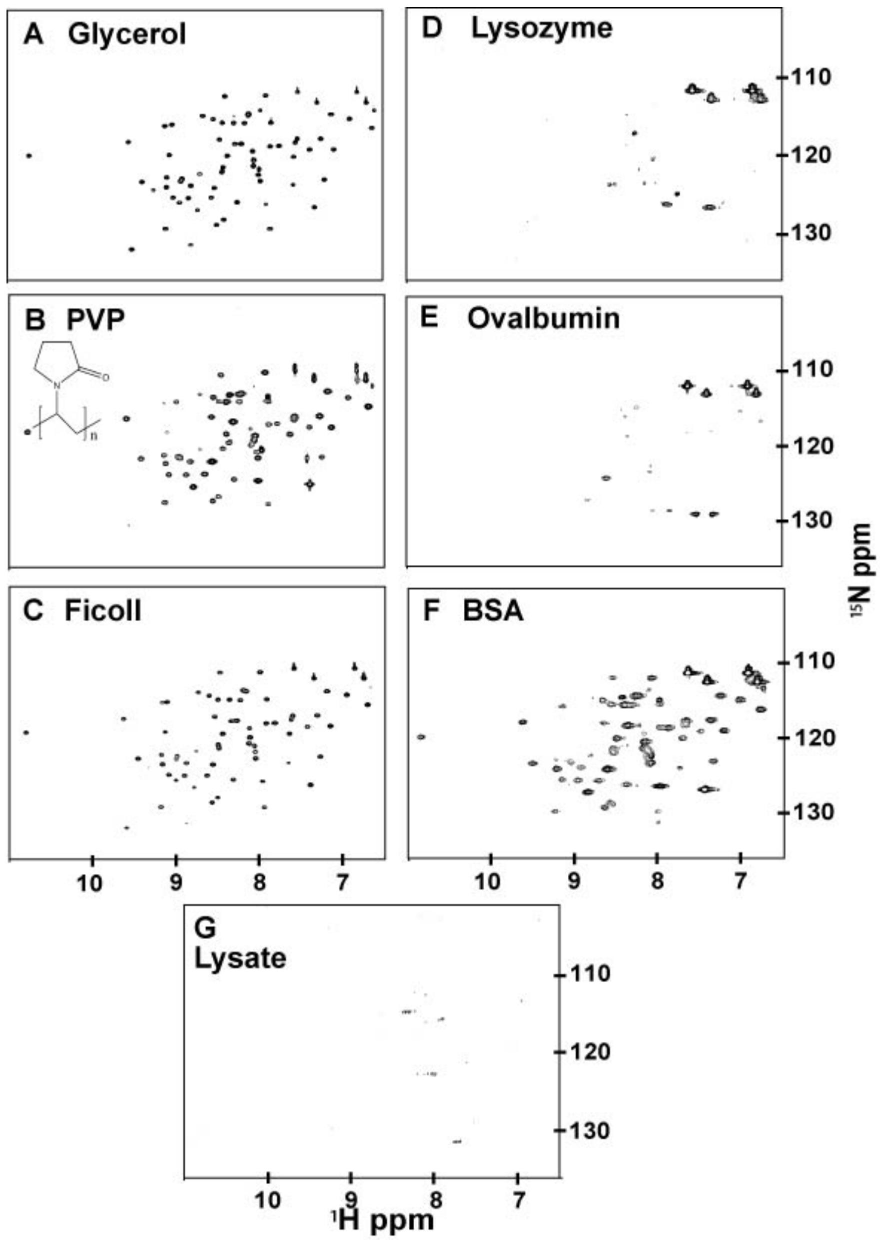

Figure 3.1 ^{15}N-^{1}H HSQC- spectra of CI2 solutions under crowded conditions.

^{15}N-^{1}H HSQC- spectra of CI2 solutions (1 mM, 25 ^{o}C, pH 5.4) containing 350 g/L glycerol (A) and 300 g/L PVP (B), Ficoll (C), lysozyme (D), ovalbumin (E), BSA (F), and *E. coli* lysate (G). The backbone structure of PVP is shown in panel B.

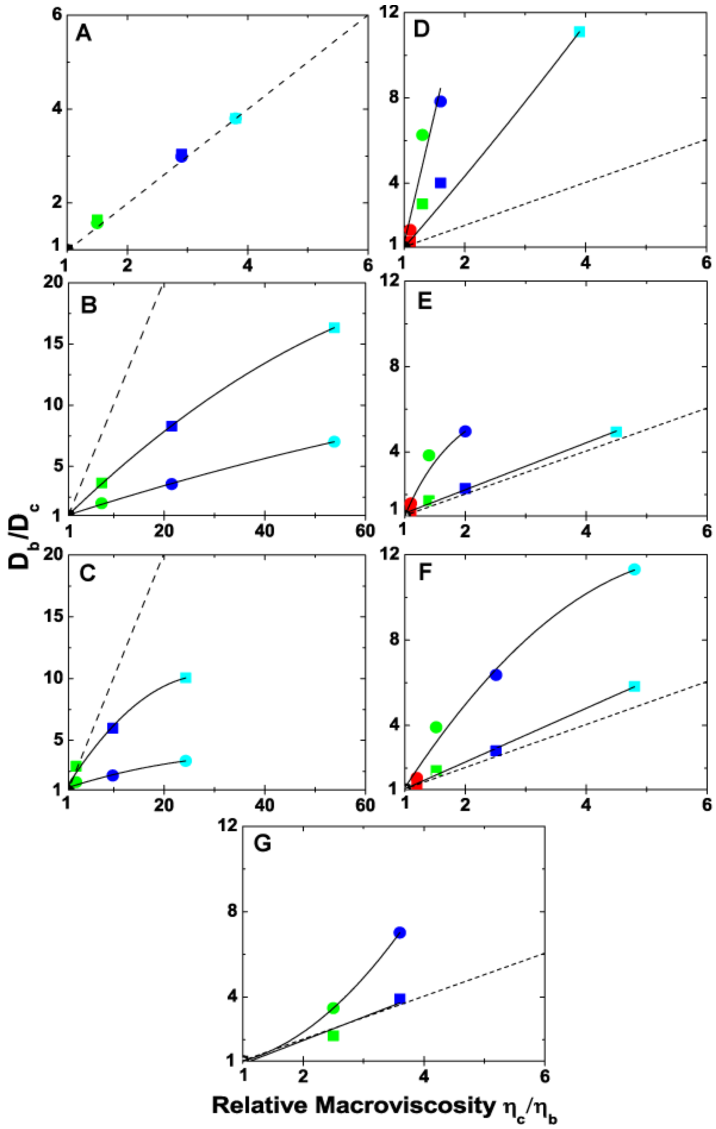

Figure 3.2 Ratio of translational and rotational diffusion coefficients of CI2 under crowded conditions.

Ratio of translational (squares) and rotational (circles) diffusion coefficients of CI2 in dilute buffer (D_b) to its diffusion coefficients in crowded solutions (D_c) (25 ^oC, pH 5.4) containing glycerol (A), PVP (B), Ficoll (C), lysozyme (D), ovalbumin (E), BSA (F), and *E. coli* lysate (G) as a function of relative viscosity (Glycerol: green, 200 g/L; blue, 350 g/L, cyan, 420 g/L. Other crowders: red, 50 g/L; green, 100 g/L; blue, 200 g/L; cyan, 300 g/L.). The smooth curves are polynominal fits of no theoretical significance. The dashed lines illustrate the unitary slope and origin-intercept expected for Stokes Laws. Points below and above dashed line indicate negative deviations and positive deviations, respectively. The uncertainties are smaller than the symbols. The PVP data have been published.[42, 53]

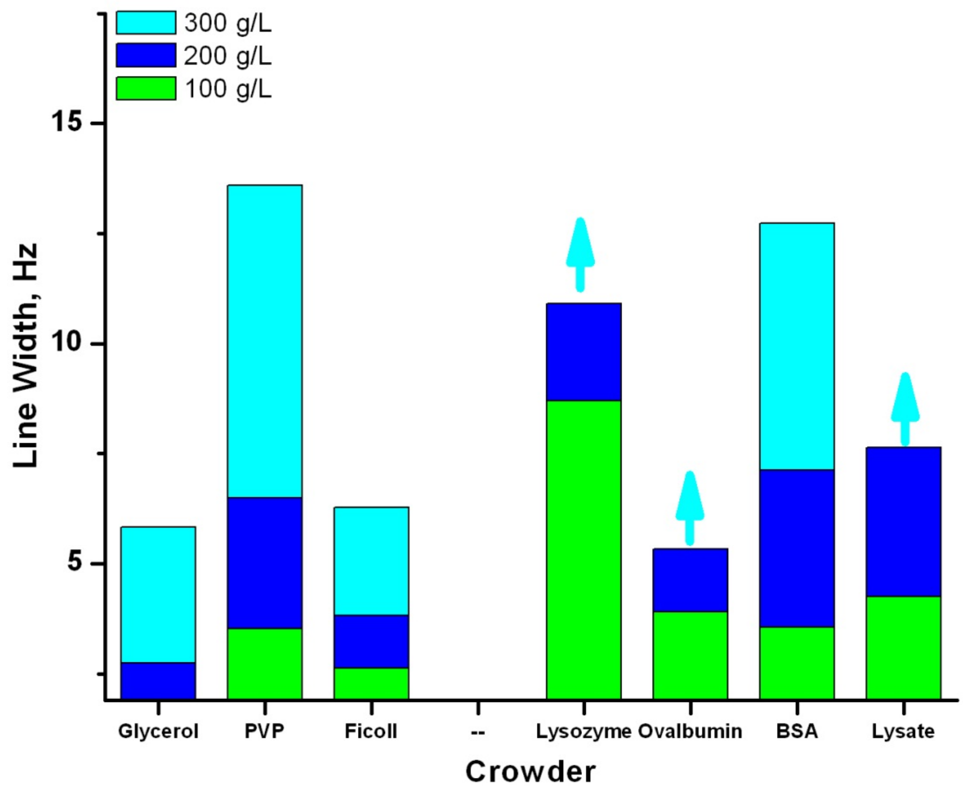

Figure 3.3 Average widths of CI2 backbone amide ^{15}N resonances under crowded conditions.

Average widths of CI2 backbone amide ^{15}N resonances (25 ^{o}C, pH 5.4) derived from T_2 measurements [line width = $1/(\pi T_2)$]. The starting point of Y-axis represents the average line width in dilute solution. The arrows indicate that the widths in 300 g/L are too broad to observe. Glycerol concentrations are given in the legend to Figure 3.2.

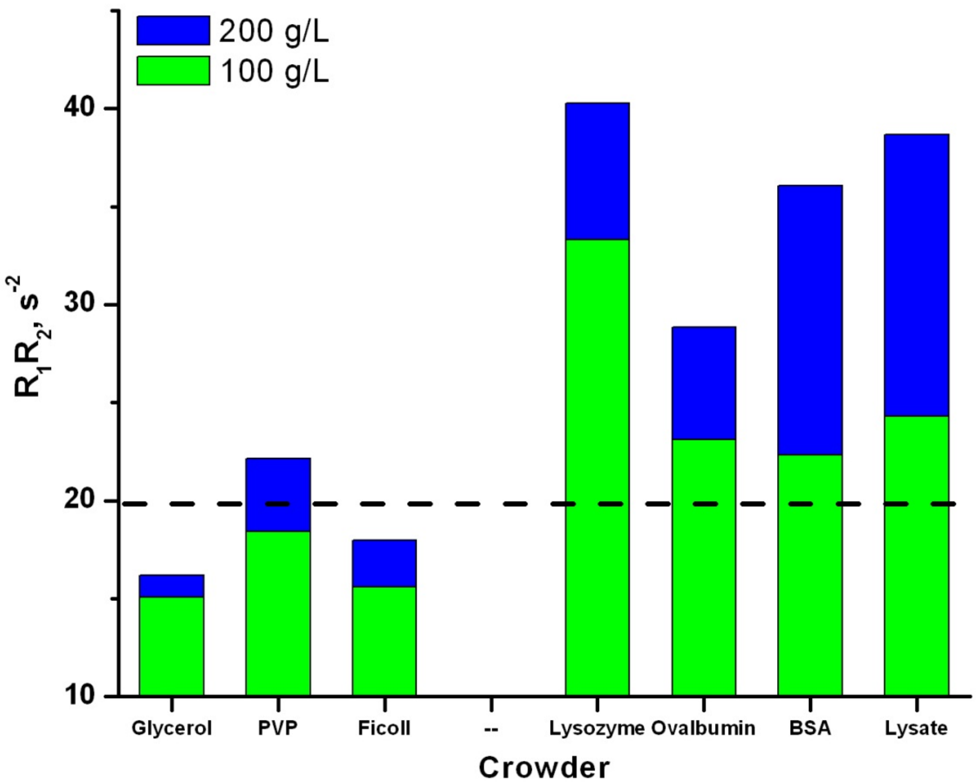

Figure 3.4 Histograms of average R_1R_2 values for CI2 under crowded conditions.

Histograms of average R_1R_2 values for CI2 in solutions of glycerol, synthetic polymers, globular proteins, and *E. coli* lysates (25 °C, pH 5.4). The dashed line is the theoretical maximum value for monomeric CI2 in the absence of conformational exchange.[52] Glycerol concentrations are given in the legend to Figure 3.2.

Chapter 4 – Macromolecular Crowding and Protein Stability

The material in this chapter is from:

Wang Y, Sarkar M, Smith AE, Krois AS, and Pielak GJ, Macromolecular crowding and protein stability, Submitted

(YW and GJP designed research; YW and SAE performed research; YW, MS, and GJP analyzed data; ASK helped prepare the samples; YW and GJP wrote the paper.)

4.1 Introduction

The cellular interior is exceptionally complex and contains macromolecules at concentrations exceeding 300 g/L and volume occupancies of 30%.[1] This crowded environment is vastly different from the dilute, idealized conditions usually used in biophysical studies. The consequences of macromolecular crowding[157] on globular protein stability arise from two phenomena:[158] the excluded volume effect and nonspecific chemical interactions. First, we describe the equilibrium thermodynamics of stability and then discuss the parameters in terms of crowding.

The stability of globular proteins can be defined as the standard-state free energy change, $\Delta G^{o'}{}_{D}$, of the reaction[159]

$$N \rightleftarrows D \tag{1}$$

where N is the biologically active native state and D is the denatured state. $\Delta G^{o'}_{D}$ can be dissected into its enthalpic, $\Delta H^{o'}_{D}$, and entropic, $\Delta S^{o'}_{D}$, components

$$\Delta G^{o'}_{D,T} = \Delta H^{o'}_{D,T} - T\Delta S^{o'}_{D,T} \qquad (2)$$

where T represents the absolute temperature. For globular proteins, $\Delta H^{o'}_{D,T}$ and $\Delta S^{o'}_{D,T}$ are temperature dependent[160] such that

$$\Delta H^{o'}_{D,T} = \Delta H^{o'}_{D,T_{ref}} + \Delta C_p\left(T - T_{ref}\right) \qquad (3)$$

$$\Delta S^{o'}_{D,T} = \Delta S^{o'}_{D,T_{ref}} + \Delta C_p \ln\left(\frac{T}{T_{ref}}\right) \qquad (4)$$

where T_{ref} is the reference temperature and ΔC_p is the heat capacity change upon denaturation. Substituting equations (3) and (4) into equation (2) gives

$$\Delta G^{o'}_{D,T} = \Delta H^{o'}_{D,T_{ref}} - T\Delta S^{o'}_{D,T_{ref}} - \Delta C_p\left[\left(T_{ref} - T\right) + T\ln\left(\frac{T}{T_{ref}}\right)\right] \qquad (5)$$

where $\Delta G^{o'}_{D}$ is zero at the temperature T_m where the concentrations of N and D are equal. Inspection of equation (2) shows that at T_m, $\Delta S^{o'}_{D,Tm}$ equals $\Delta H^{o'}_{D,Tm}/T_m$, such that equation (5) can be converted to

$$\Delta G^{o'}_{D,T} = \Delta H^{o'}_{D,T_m}\left(1 - \frac{T}{T_m}\right) - \Delta C_p\left[\left(T_m - T\right) + T\ln\left(\frac{T}{T_m}\right)\right] \qquad (6)$$

Figure 4.1 shows a plot of $\Delta G^{o'}_{D}$ versus T for the small globular protein. The curvature arises because ΔC_p is non-zero (increasing ΔC_p narrows the curve), leading to a temperature of maximum stability, T_{max}, and two values for T_m. The higher T_m is the more pertinent one because the lower T_m is usually below the freezing point of the solution. Changing the parameters affects curve's shape and location. Increasing or decreasing $\Delta S^{o'}{}_{D}$ translates the curve down or up, respectively, whereas the opposite is true for $\Delta H^{o'}{}_{D}$. Of course, these translations also affect T_{m}.

Macromolecular crowding may affect $\Delta H^{o'}{}_{D}$, $\Delta S^{o'}{}_{D}$ or both. The excluded volume component of macromolecular crowding arises because the crowding molecules decrease the space available to the protein being studied. Application of Le Chatelier's principle shows that volume exclusion favors N because this form occupies less space than D. In its purest form, volume exclusion is entirely entropic because it involves only the arrangement of molecules, not their interaction. A crowder that acts solely by excluded volume decreases $\Delta S^{o'}{}_{D}$, which translates the curve of the $\Delta G^{o'}{}_{D}$ versus T plot up, increasing T_{m}. The original formulation of macromolecular crowding theory[158] and, until recently, most work has stressed only excluded volume.[8, 12, 44, 54, 62, 161, 162] The other key phenomenon, nonspecific chemical interactions, can be attractive or repulsive. Repulsive interactions will be stabilizing because repulsion increases the apparent excluded volume. Nonspecific attractions, because they involve not only the formation of non-covalent bonds but also the associated changes in solvation, could either increase or decrease $\Delta H^{o'}{}_{D}$ and $\Delta S^{o'}{}_{D}$.[163]

Little is known about how crowding actually affects $\Delta H^{o'}{}_{D}$ and $\Delta S^{o'}{}_{D}$. To fill this gap, we used NMR-detected amide ^{1}H exchange experiments[164] to obtain these parameters under crowded conditions. We chose ubiquitin (pI 6.4) as the test protein because it folds in a two-state manner,[106] and its unusually high T_{max}

allowed us to estimate ΔC_p. For synthetic crowders, we chose the uncharged polymers polyvinylpyrrolidone (PVP) and Ficoll because their effects on proteins are known to arise from their macromolecular natures.[165, 166] For more biologically relevant crowders, we chose two globular proteins, bovine serum albumin (BSA) and lysozyme. PVP, Ficoll, and the two proteins were used at concentrations of 100 g/L, which, although lower than the macromolecular solute concentration in cells, is the highest concentration that allows acquisition of high-quality data.

4.2 Methods

4.2.1 Protein Expression and Purification

The pET-46 plasmid (Novagen) containing the gene for histidine-tagged ubiquitin[8] was transformed into BL-21 (DE3-Gold) competent *Escherichia coli* cells (Stratagene). The transformants were spread onto Luria Broth agar plates containing 0.1 g/L ampicillin. Liquid Luria-Bertani (LB) media (100 mL containing 1 g Bacto-Tryptone, 0.5 g Bacto-yeast extract, and 1 g NaCl in H_2O) containing 0.1 g/L ampicillin was inoculated with a single colony of ubiquitin-expressing *E. coli* cells and incubated overnight at 310 K with shaking at 250 rpm. The next morning, this pre-culture was pelleted (Sorvall RC-3B, H6000A rotor,1600 *g*). One L of ^{15}N enriched M9 media (6 g Na_2HPO_4, 2 g glucose, 3 g KH_2PO_4, 0.5 g NaCl, 1 g $^{15}NH_4Cl$, 2 mM $MgSO_4$,10 μM $CaCl_2$) containing thiamine HCl (1 mg/L) and ampicillin (0.1 g/L) was used to resuspend the cell pellet. This culture was incubated at 310 K with shaking until its optical density at 600 nm reached 0.8.

Induction was then initiated by adding isopropyl-β-D-1-thiogalactopyranoside to a final concentration of 1 μM. Induction was allowed to proceed for 4 h, whereupon the culture was centrifuged at 1600*g* and the pellet frozen.

The pellet was resuspended in 20 mL of buffer (50 mM Na_2HPO_4, 500 mM NaCl, 30 mM imidazole, pH 7.6). Cells were lysed by sonic dismembration for 10 min (Fisher Scientific, Sonic Dismembrator Model 500, 14% amplitude, 2 s pulse, 3 s rest). The lysate was centrifuged at 14000*g* for 30 min, and the supernatant retained. Streptomycin sulfate (0.2 g) was added with stirring on ice for 30 min, followed by centrifugation at 14000*g* for 30 min. The supernatant was forced through a sterilized 0.22-μm filter. The ubiquitin was purified by Ni^{2+}-affinity chromatography on an AKTA FPLC (GE Healthcare). The column was washed with 60 mL of low imidazole buffer (50 mM Na_2HPO_4, 500 mM NaCl, 30 mM imidazole, pH 7.6), and then eluted with 80 mL of high imidazole buffer (50 mM Na_2HPO_4, 500 mM NaCl, 500 mM imidazole, pH 7.6). The pure fractions (as assessed by SDS-PAGE) were pooled, dialyzed against H_2O and subjected to size exclusion chromatography (Superdex 200 10/300) using water as eluent. The protein was then lyophilized.

4.2.2 NMR

Amide proton exchange experiments were performed as described by Miklos *et al.*[164] on a 600 MHz Varian Inova spectrometer equipped with a standard triple-resonance HCN probe and three axis gradients. The ^{1}H dimension was acquired with a sweep width of 12000 Hz and comprised 1024 complex points. The ^{15}N dimension was acquired with a sweep width of 2500 Hz and comprised 64 complex increments. Each experiment required two samples, an optimization sample and an exchange sample. Optimization samples of 1 mM ubiquitin in 50 mM sodium phosphate, pH 5.4, with 15% D_2O were used for shim adjustment and pulse width calibration. pH values were obtained from direct meter readings, uncorrected for the isotope effect.[167] Exchange samples contained 1 mM ubiquitin and 50 mM sodium acetate buffer, pH 5.4, and were made with 99.9% D_2O.

Ficoll, BSA, and lysozyme were exchanged in D_2O prior to use (PVP has no exchangeable protons). One gram of each was suspended in 10 mL of D_2O. Exchange for 36 h at 310 K was followed by lyophilization overnight. The dried samples were again suspended in 10 mL of D_2O and the process repeated.

Twenty to twenty-four consecutive HSQC spectra[168, 169] were acquired per exchange sample. Processing was performed with NMRPipe.[104] Assignments have been described.[113] Crosspeak volumes were plotted against time and fit to exponential decays by using NMRViewJ[105] to yield values of k_{obs}, the rate of exchange for the a particular residue.

4.2.3 Amide ^{1}H Exchange and Protein Stability

Exchange occurs via the scheme shown in equation (7):[170, 171]

$$cl - {}^1H \underset{k_{cl}}{\overset{k_{op}}{\rightleftarrows}} op - {}^1H \xrightarrow{k_{int}} op - {}^2H \underset{k_{op}}{\overset{k_{cl}}{\rightleftarrows}} cl - {}^2H \quad (7)$$

where cl - ^{1}H is the amide proton in N, which opens and closes with rate constants k_{op} and k_{cl}. op - ^{1}H is the open, exchange competent state. For ubiquitin at pH values < 8.5, k_{cl} is much larger than k_{int},[172, 173] such that the free energy required to expose an amide proton ($\Delta G^{o'}_{op}$) can be determined using the equation

$$\Delta G^{o'}_{op} = -RT\ln K_{op} = -RT\ln \frac{k_{op}}{k_{cl}} = -RT\ln \frac{k_{obs}}{k_{int}} \quad (8)$$

where R is the gas constant, k_{obs} is the observed rate constant of exchange, and k_{int} is the rate constant for an amide proton in a peptide lacking stable structure[174]. Values of k_{int} were calculated using the program SPHERE[174]. The crowders do not change k_{int}.[164, 166]

The standard deviation of $\Delta G^{o'}_{op}$, ~2.0 kcal/mol, was not from random error; the reproducibility, determined by repeating one condition (298 K) three times, was within 0.1 kcal/mol (Table 4.9), consistent with our previous efforts.[164, 166] Instead, the uncertainty arose from systematic error. Specifically, the values of kint, which are determined using peptides,[174, 175] are probably not exactly correct for the denatured state of any particular protein. Support for this idea comes from the observation that the standard deviation dropped 10-fold for DD$G^{o'}_{D}$ between solution in the absence of crowding agent and in 100 g/L PVP at 323 K, which does not involve the use of k_{int}. Fits to equation (6) were performed with Origin (OriginLab, Northampton, MA, USA).

4.3 Results

Amide ^{1}H exchange data can be analyzed to yield the free energy required to expose an amide ^{1}H to exchange with solvent, $\Delta G^{o'}_{op}$.[171] We focus on the global protein stability, $\Delta G^{o'}_{D}$, which we define as the average of $\Delta G^{o'}_{op}$

values from 10 residues that are exposed only on global unfolding. These residues were identified previously by combining data from stopped flow and NMR experiments.[172, 173, 176] All experiments were performed at pH 5.4 in 50 mM sodium acetate buffer. The datasets are given in the Supporting Information.

The $\Delta G^{o'}_{D}$ versus temperature data for ubiquitin are plotted in Figure 4.1. Table 4.1 shows the parameters calculated from the least-squares fit of these data to equation (6). The value of ΔC_p, 1.4 kcal/mol·K, exactly matches the value obtained previously from calorimetry.[106, 177] The T_m in dilute solution, 381 K, is close to that obtained from calorimetry (373 K) at neutral pH.[178, 179] The slight deviation is due to the difference in conditions – D_2O increases T_m[180] –, our use of a his-tagged protein, and the extrapolation from the temperatures at which the data were acquired.

We first examined the stability of ubiquitin in 100 g/L solutions of 40 kDa PVP. The $\Delta G^{o'}_{D}$ versus temperature curve is shown in Figure 4.2, and the fitted parameters are given in Table 4.1. The main effect of PVP was to increase T_m. Increasing the PVP concentration shifted the curve even further to the right (Figure 4.3). The data from analysis of ubiquitin stability in 70 kDa Ficoll (Figure 4.2) support the idea that the increased T_m is a general result for synthetic polymers. The broader curve in Ficoll compared to that in the absence of crowding arises from the diminution of ΔC_p (Table 4.1).

To study the effect of more biologically relevant crowders, we examined the stability of ubiquitin in 100 g/L solutions of lysozyme (15 kDa, pI 11.0) and BSA (67 kDa, pI 4.7). To avoid denaturing BSA ($T_m \approx 338$ K)[181] experiments were performed only up to 323 K. As shown in Figure 4.4, crowding by lysozyme increased T_m, $\Delta S^{o'}_{D,Tm}$, $\Delta H^{o'}_{D,Tm}$, and ΔC_p compared to the parameters in buffer alone. BSA also increased $\Delta S^{o'}_{D,Tm}$, $\Delta H^{o'}_{D,Tm}$ and ΔC_p, but decreased T_m.

4.4 Discussion

The simplest interpretation of theory predicts that macromolecular crowding will always stabilize proteins because crowders enhance the representation of compact forms of the protein in the denatured state ensemble. This enhancement decreases $\Delta S^{o'}_D$ and increases both $\Delta G^{o'}_D$ and T_m. Inspection of Figure 4.2 and Figure 4.4 show that the real situation is much more complex. Crowding can be stabilizing or destabilizing depending on both the nature of the crowding agent and the temperature.

4.4.1 Analysis at a Common Temperature

Uncovering the origin of these effects requires analysis of the temperature dependence of protein stability in terms of well established equilibrium thermodynamic principles.[182] Until now, studies have focused on only the high temperature portion of the melting curve, well above T_{max}. This narrow focus obviated assessment of ΔC_p, and, hence, $\Delta S^{o'}_D$ and $\Delta H^{o'}_D$ could not be compared at a common temperature. Ubiquitin's high T_{max} offered the opportunity to observe the curvature in stability-versus-temperature plots and, hence, the estimation of ΔC_p. Using equation (6) we then calculated $\Delta G^{o'}_D$, $\Delta H^{o'}_D$, and $T\Delta S^{o'}_D$ values under crowded conditions at a common temperature, T_m of ubiquitin in the absence of crowders. The values are shown in Table 4.2.

4.4.2 Excluded Volume Appears to Dominate for Uncharged Synthetic Polymers

Synthetic polymers are the most widely used crowding agents. Our analysis of the PVP and Ficoll data is consistent with the prediction that crowding decreases $\Delta S^{o'}_{D}$, yet not all of the decrease is reflected in the increased $\Delta G^{o'}_{D}$. The fact that NMR relaxation data indicate that PVP and Ficoll are relatively inert toward protein[54] lead us to suggest that the compensatory changes in $\Delta H^{o'}_{D}$ arise from intramolecular native state interactions that persist in the crowder-compacted denatured state ensemble; a conclusion consistent with others work.[183]

4.4.3 Effects of Protein Crowders Depend on Charge

Proteins are more biologically relevant than synthetic polymers. To test the effect of a positively charged protein on ubiquitin stability, we used lysozyme as a crowder. Both lysozyme (pI 11.0) and ubiquitin (pI 6.4) are polycations under our conditions, which means the molecules repel each other. As shown in Table 4.2, lysozyme's effect on the stability arose from the same combination of changes in $\Delta H^{o'}_{\mathrm{D}}$ and $\Delta S^{o'}_{\mathrm{D}}$ as observed for PVP and Ficoll. We conclude that nonspecific repulsive chemical interactions support the excluded volume effect, which in turn favors a more compact ensemble of denatured states.

Results obtained in the presence of negatively charged BSA contradict the idea that macromolecular crowding always increases stability (Table 4.2). The decrease in ubiquitin stability we observed in the presence of BSA is consistent with our study of chymotrypsin inhibitor 2[161] and studies of protein stability in cells.[126, 127] Our ability to study the temperature dependence of $\Delta G^{o'}{}_D$ allowed us to identify the source of this destabilization. As shown in Table 4.2, the result was not subtle; changes in $\Delta H^{o'}{}_D$ and $\Delta S^{o'}{}_D$ in BSA were of the opposite sign compared to those for PVP, Ficoll, and lysozyme. The BSA-induced destabilization cannot arise from a lack of an excluded volume effect; all the crowders used here occupy 7-8% of the solution volume.[166, 184, 185] The decrease in stability due to the presence of BSA, therefore, must result from cancellation of the excluded volume effect.

Under the conditions used here, ubiquitin is a polycation, BSA (pI 4.7) is a polyanion, and therefore the two proteins should attract each other. We suggest that the inherent nonspecific attractive interactions cancel the effect of volume exclusion. Three observations support this idea. NMR relaxation studies[52, 54] and molecular dynamics simulations[61] studies show that BSA interacts with chymotrypsin inhibitor 2. Third, the destabilization of chymotrypsin inhibitor 2 in BSA can be alleviated by increasing the concentration of NaCl.[161] In summary, nonspecific attractive chemical interactions can overcome the stabilization induced by excluded volume. These results show that the charge of the macromolecular crowding agent plays a key role in determining the effect of the crowder on protein stability.

4.5 Summary and Biological Implications.

Both excluded volume and nonspecific chemical interactions must be considered when predicting the stability of a protein under crowded conditions.

The overall effect depends on the winner of the nearly evenly matched battle between excluded volume and nonspecific chemical interactions. Our results indicate that synthetic polymers, which are widely used to mimic the crowded cellular environment, are unsuitable for providing insight into the biological effects of crowding. The results also have important implications for understanding cellular chemistry. The stability of a protein inside cells can be tuned by the charge and size of surrounding proteins. Quoting Spitzer and Poolman,[48] “the cytoplasm is a highly anisotropic and structured environment, in which many proteins carry out their functions as multimeric complexes at specific subcellular locations.” Given the tight competition between excluded volume and nonspecific chemical interactions, altering the intracellular environment at certain “addresses” could be used to regulate key protein functions such as transcription, translation, replication, and segregation.[186-189]

4.6 Tables

Table 4.1 Thermodynamic parameters and T_m values

Co-solute[a]	T_m K	$\Delta H^{o'}_{D,Tm}$ kcal/mol	$\Delta S^{o'}_{D,Tm}$ kcal/mol·K	ΔC_p kcal/mol·K
None	381 ± 17	95 ± 13	0.25 ± 0.05	1.4 ± 0.5
PVP	389 ± 27	95 ± 16	0.25 ± 0.06	1.4 ± 0.6
Ficoll	435 ± 52	89 ± 12	0.21 ± 0.05	0.9 ± 0.5
Lysozyme	395 ± 16	113 ± 9	0.29 ± 0.03	1.6 ± 0.4
BSA	374 ± 2	155 ± 2	0.42 ± 0.01	3.3 ± 0.2

[a] Co-solute concentrations were 100 g/L.

Table 4.2 Change in thermodynamic parameters at the T_m in the absence of crowder

Co-solute[a]	$\Delta\Delta G^{o'}_{D,c\text{-}d}$ **kcal/mol**	$\Delta\Delta H^{o'}_{D,c\text{-}d}$ **kcal/mol**	$T\Delta\Delta S^{o'}_{D,c\text{-}d}$ **kcal/mol**
PVP	2	-11	-13
Ficoll	8	-53	-61
Lysozyme	4	-5	-9
BSA	-3	85	88

[a]Co-solute concentrations were 100 g/L.
Change in thermodynamic parameters: crowded minus no crowding agent.
T_m in the absence of crowder is 381 K.

Table 4.3 $\Delta G^{o'}_{op}$ (kcal/mol) values for globally exchanging ubiquitin residues in **dilute solution**

Residue	288 K	298 K	308 K	318 K	323 K
V5	4.78	6.09	7.19	7.91	7.47
L15	6.24	6.99	7.78	6.91	7.75
V17	5.29	6.86	6.83	7.54	7.20
D21	6.49	7.97	8.66	9.18	8.15
V26	4.49	3.46	7.19	8.07	7.75
K27	7.19	8.42	7.88	8.61	8.14
A28	6.07	7.23	8.34	9.29	-
K29	7.60	8.00	8.21	9.18	8.50
I30	3.59	5.66	7.95	7.55	-
I44	-	5.65	6.23	6.97	6.84
$\boldsymbol{\Delta G^{o'}_{D}}$	5.8 ± 1.3	6.8 ± 1.5	7.6 ± 0.8	8.1 ± 0.9	7.7 ± 0.5

$\Delta G^{o'}_{D}$ is the average $\Delta G^{o'}_{op}$ values, and its uncertainty is the standard deviation of the mean.
Condition: 50 mM sodium acetate, pH 5.4.

Table 4.4 $\Delta G^{o'}_{op}$ (kcal/mol) values for globally exchanging ubiquitin residues in **100 g/L PVP**

Residue	288 K	298 K	308 K	313 K	323 K
V5	4.28	6.42	6.49	7.53	7.68
L15	6.24	6.64	6.65	7.27	8.02
V17	-	6.13	6.12	7.49	7.76
D21	5.99	7.63	8.39	9.16	8.48
V26	5.12	-	6.86	7.80	8.25
K27	-	6.88	7.45	8.66	8.14
A28	-	7.26	8.49	9.27	-
K29	4.70	7.16	7.74	8.72	8.47
I30	4.56	5.37	6.05	7.39	8.07
I44	4.32	5.60	-	7.04	7.35
$\boldsymbol{\Delta G^{o'}_{D}}$	5.0 ± 0.8	6.6 ± 0.8	7.1 ± 0.9	8.0 ± 0.8	8.0 ± 0.9

The $\Delta G^{o'}_{D}$ is the average $\Delta G^{o'}_{op}$ values, and its uncertainty is the standard deviation of the mean.
Condition: 50 mM sodium acetate, pH 5.4.

Table 4.5 $\Delta G^{o'}_{op}$ (kcal/mol) values for globally exchanging ubiquitin residues in **200 g/L PVP**

Residue	298 K	308 K	313 K	318 K	323 K	328 K
V5	-	5.96	6.75	8.82	7.98	7.82
L15	5.69	6.50	6.87	7.67	8.44	7.95
V17	5.30	6.33	6.65	7.12	7.79	7.73
D21	6.98	8.28	8.08	8.87	8.76	7.51
V26	9.91	6.50	7.62	8.76	8.39	7.81
K27	6.83	7.28	7.64	8.17	8.79	8.41
A28	7.26	8.39	8.19	8.97	8.86	7.62
K29	-	7.39	8.00	8.61	8.63	7.57
I30	2.95	4.50	6.61	8.87	8.00	7.27
I44	4.61	5.23	6.00	8.39	7.41	7.49
$\boldsymbol{\Delta G^{o'}_{D}}$	6.2 ± 2.1	6.6 ± 1.2	7.2 ± 0.8	8.4 ± 0.6	8.3 ± 0.5	7.7 ± 0.3

The $\Delta G^{o'}_{D}$ is the average $\Delta G^{o'}_{op}$ values, and its uncertainty is the standard deviation of the mean.
Condition: 50 mM sodium acetate, pH 5.4.

Table 4.6 $\Delta G^{o'}_{op}$ (kcal/mol) values for globally exchanging ubiquitin residues in **100 g/L Ficoll**

Residue	288 K	298 K	308 K	318 K	323 K	328 K
V5	4.64	5.85	7.37	8.47	9.06	9.22
L15	6.24	7.18	7.17	6.35	6.91	6.71
V17	1.91	7.16	6.93	8.30	8.68	9.34
D21	6.40	7.76	8.58	8.61	9.63	9.57
V26	5.12	4.62	9.25	8.45	9.44	9.71
K27	-	3.29	8.59	8.85	9.62	9.68
A28	5.70	7.45	8.43	9.35	9.74	9.67
K29	9.61	8.88	8.77	9.56	9.50	9.67
I30	4.31	4.31	8.26	8.43	8.99	9.03
I44	2.57	5.62	6.55	8.04	8.32	8.61
$\boldsymbol{\Delta G^{o'}_{D}}$	5.2 ± 2.3	6.2 ± 1.8	8.0 ± 0.9	8.4 ± 0.9	9.0 ± 0.9	9.1 ± 0.9

The $\Delta G^{o'}_{D}$ is the average $\Delta G^{o'}_{op}$ values, and its uncertainty is the standard deviation of the mean.
Condition: 50 mM sodium acetate, pH 5.4.

Table 4.7 $\Delta G^{o'}_{op}$ (kcal/mol) values for globally exchanging ubiquitin residues in **100 g/L lysozyme**

Residue	288 K	298 K	308 K	318 K	323 K	328 K
V5	-	6.70	6.78	8.03	9.06	8.54
L15	-	6.51	6.65	8.21	9.14	8.99
V17	5.66	4.84	6.44	7.98	8.58	8.59
D21	3.51	7.32	8.28	9.70	9.96	9.30
V26	4.47	5.08	8.83	8.84	9.33	8.93
K27	2.67	5.96	8.07	9.43	9.88	9.41
A28	6.89	6.78	8.34	9.81	10.07	9.41
K29	2.58	10.90	9.06	9.44	9.65	9.07
I30	3.93	4.57	-	8.12	8.89	8.92
I44	3.58	3.78	5.49	8.04	8.69	8.61
$\boldsymbol{\Delta G^{o'}_{D}}$	4.2 ± 1.5	6.2 ± 2.0	7.6 ± 1.2	8.8 ± 0.8	9.3 ± 0.5	9.0 ± 0.3

The $\Delta G^{o'}_{D}$ is the average $\Delta G^{o'}_{op}$ values, and its uncertainty is the standard deviation of the mean.
Condition: 50 mM sodium acetate, pH 5.4.

Table 4.8 $\Delta G^{o'}_{op}$ (kcal/mol) values for globally exchanging ubiquitin residues in **100 g/L BSA**

Residue	298 K	308 K	318 K	323 K
V5	-	6.29	8.13	8.95
L15	5.90	6.23	7.85	9.00
V17	0.99	6.23	7.81	8.42
D21	7.09	8.28	9.68	9.15
V26	5.25	6.43	8.45	9.23
K27	2.41	7.18	8.91	9.62
A28	3.64	7.78	9.78	9.26
K29	2.06	7.57	9.11	9.44
I30	5.10	-	7.91	8.60
I44	2.14	4.70	7.23	8.62
$\boldsymbol{\Delta G^{o'}_{D}}$	3.8 ± 2.1	6.7 ± 1.1	8.5 ± 0.9	9.0 ± 0.4

The $\Delta G^{o'}_{D}$ is the average $\Delta G^{o'}_{op}$ values, and its uncertainty is the standard deviation of the mean.
Condition: 50 mM sodium acetate, pH 5.4.

Table 4.9 $\Delta G^{o'}_{op}$ (kcal/mol) for three trials for globally exchanging ubiquitin residues in **dilute solution** at 298 K

Residue	Trial 1	Trial 2	Trial 3	Average
V5	6.26	6.38	5.64	6.09 ±
L15	6.99	7.26	6.70	6.99 ±
V17	6.01	7.40	7.16	6.86 ±
D21	7.84	8.14	7.93	7.97 ±
V26	-	2.85	4.08	3.46 ±
K27	7.17	8.69	9.40	8.42 ±
A28	7.38	7.42	6.88	7.23 ±
K29	7.23	8.53	8.23	8.00 ±
I30	7.45	4.65	4.88	5.66 ±
I44	5.70	5.59	5.66	5.65 ±
$\boldsymbol{\Delta G^{o'}_{D}}$	6.9 ± 0.7	6.7 ± 1.9	6.7 ± 1.6	6.8 ± 0.1

$\Delta G^{o'}_{D}$ is the average $\Delta G^{o'}_{op}$ values, and its uncertainty is the standard deviation of the mean.
Condition: 50 mM sodium acetate, pH 5.4.

4.7 Figures

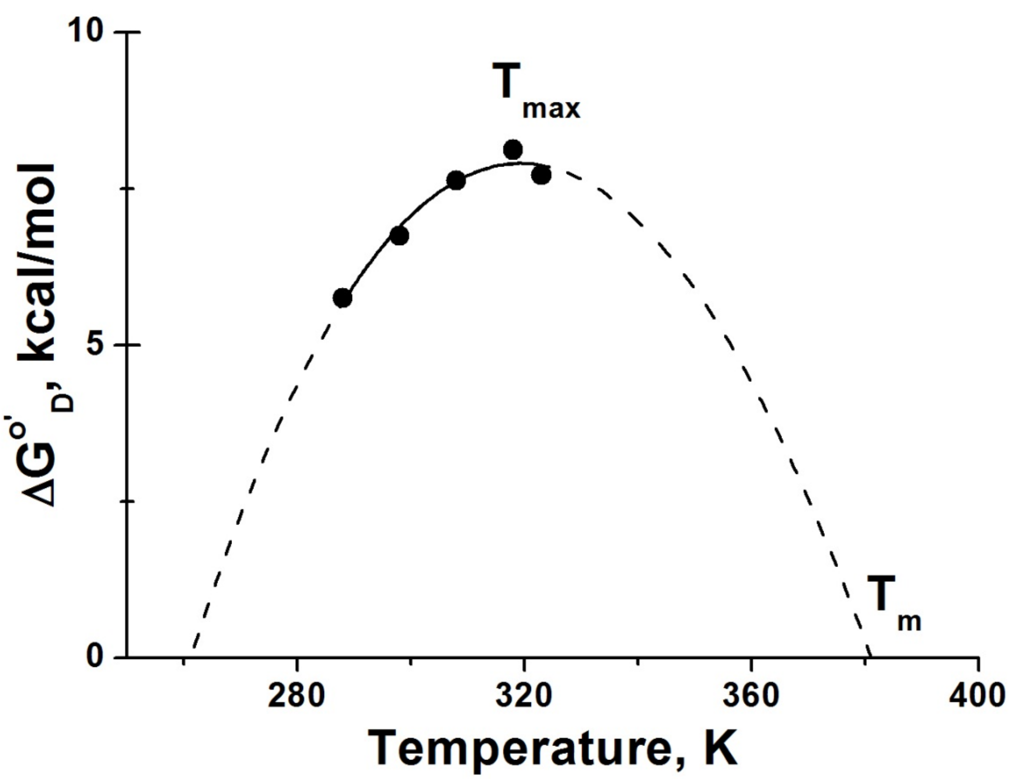

Figure 4.1 Stability of ubiquitin as a function of temperature.

The curve is a fit of the data to equation (6). The solid curve indicates the range measurable by NMR-detected amide proton exchange. Experiments were performed in 50 mM sodium acetate, pH 5.4.

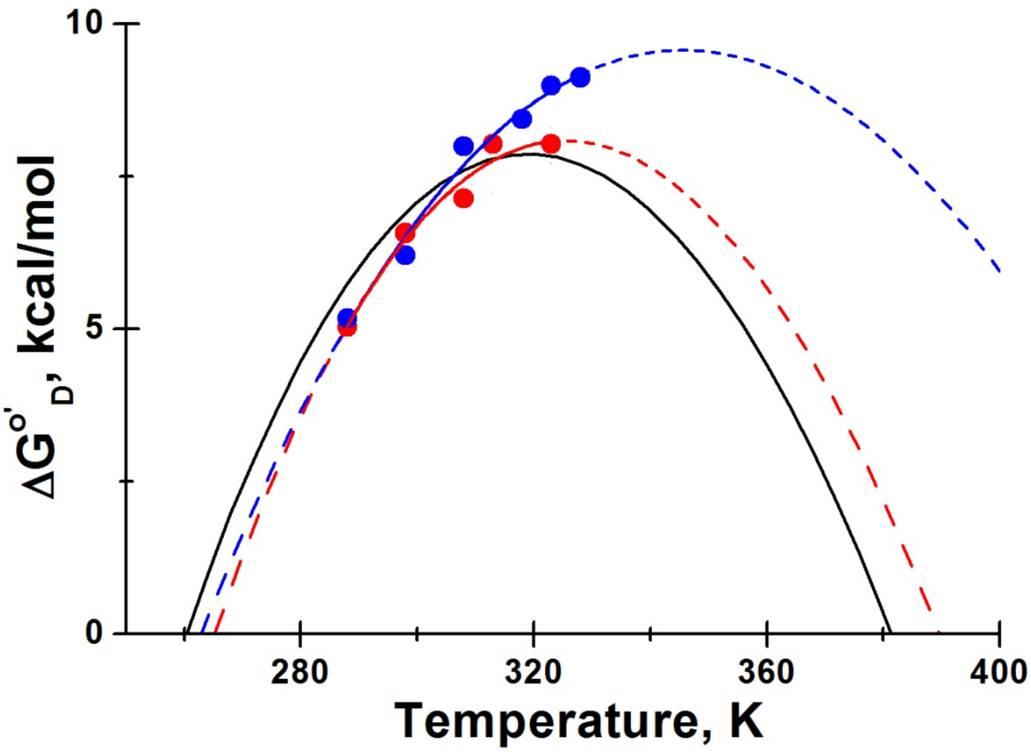

Figure 4.2 Thermal stability curves in solutions containing 100 g/L PVP and 100 g/L Ficoll.

The black curve shows the stability in buffer without crowding agent (from Figure 4.1). Red, 100 g/L PVP; blue, 100 g/L Ficoll. Experiments were performed in 50 mM sodium acetate, pH 5.4.

Figure parameters

Co-solute	T_m K	$\Delta H^{o'}_{D, Tm}$ kcal/mol	$\Delta S^{o'}_{D, Tm}$ kcal/mol·K	ΔC_p kcal/mol·K
200 g/L PVP	394 ± 63	96 ± 42	0.24 ± 0.15	1.4 ± 1.6

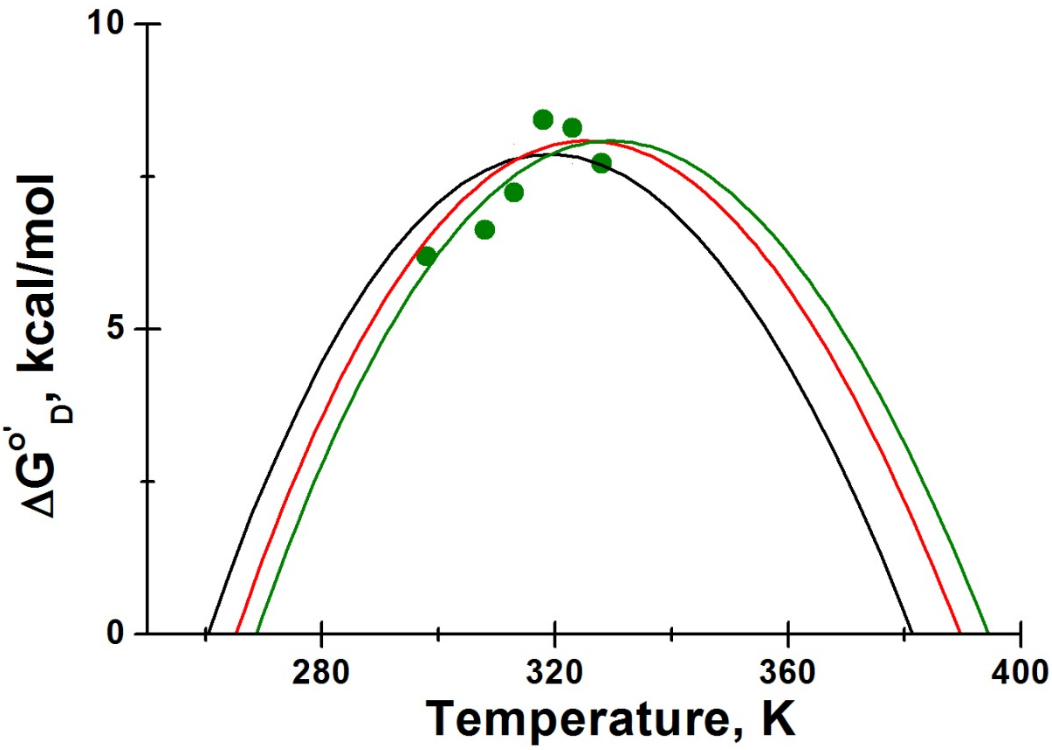

Figure 4.3 Thermal stability of ubiquitin in 200 g/L PVP.

The black and red curves show the stabilities in dilute solution and in 100 g/L PVP, respectively (from Figure 4.1 and Figure 4.2). Olive, 200 g/L PVP.

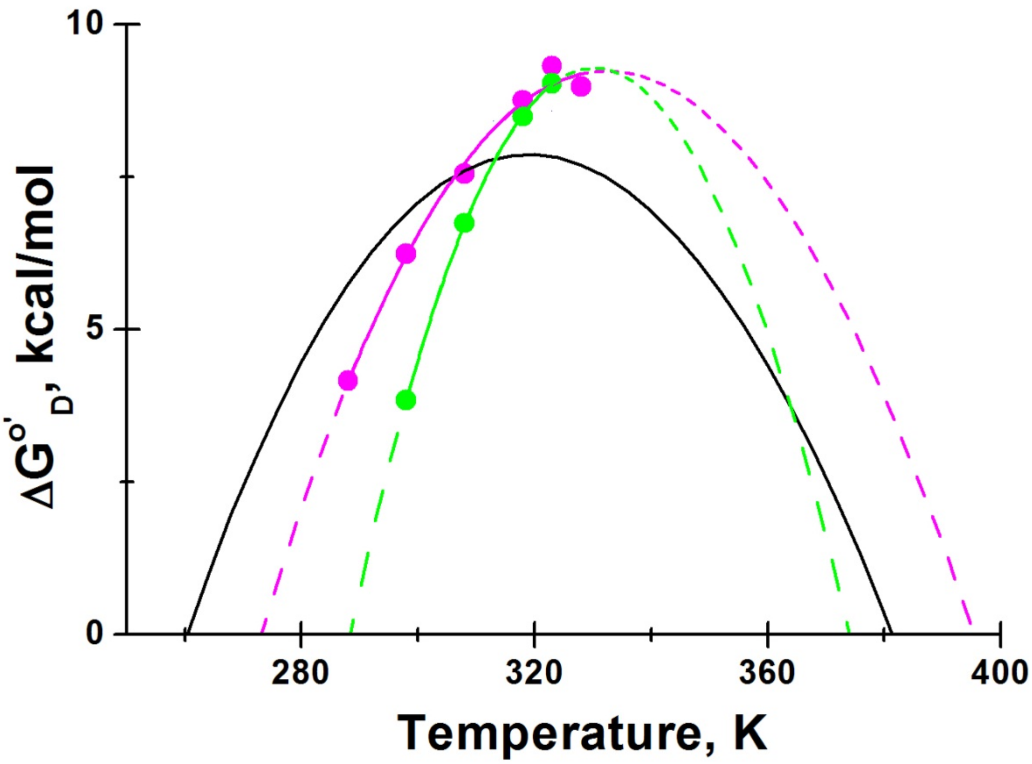

Figure 4.4 Thermal stability curves in solutions containing 100 g/L bovine serum albumin and 100 g/L lysozyme.

The black curve shows the stability in buffer without crowding agent (from Figure 4.1). Green, 100 g/L bovine serum albumin; magenta, 100 g/L lysozyme. Experiments were performed in 50 mM sodium acetate, pH 5.4.

References

1. Zimmerman, S.B., and Trach, S.O. (1991) Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of *Escherichia coli*, *J. Mol. Biol. 222*, 599-620.

2. Serber, Z., and Dötsch, V. (2001) In-cell NMR spectroscopy, *Biochemistry 40*, 14317-14323.

3. Llinas, M., Wüthrich, K., Schwotzer, W., and Von Philipsborn, W. (1975) ^{15}N nuclear magnetic resonance of living cells, *Nature 257*, 817-818.

4. Reckel, S., Hansel, R., Lohr, F., and Dotsch, V. (2007) In-cell NMR spectroscopy, *Prog. Nucl. Magn. Reson. Spectrosc. 51*, 91-101.

5. Sharaf, N.G., Barnes, C.O., Charlton, L.M., Young, G.B., and Pielak, G.J. (2010) A bioreactor for in-cell protein NMR, *J. Magn. Reson. 202*, 140-146.

6. Augustus, A.M., Reardon, P.N., and Spicer, L.D. (2009) MetJ repressor interactions with DNA probed by in-cell NMR, *Proc. Natl. Acad. Sci. U. S. A. 106*, 5065-5069.

7. Banci, L., Barbieri, L., Bertini, I., Cantini, F., and Luchinat, E. (2011) In-cell NMR in *E. coli* to monitor maturation steps of hSOD1, *PLoS ONE 6*, e23561.

8. Barnes, C.O., Monteith, W.B., and Pielak, G.J. (2011) Internal and global protein motion assessed with a fusion construct and in-cell NMR spectroscopy, *ChemBioChem 12*, 390-391.

9. Bertini, I., Felli, I.C., Gonnelli, L., Kumar, M.V.V., and Pierattelli, R. (2011) ^{13}C direct-detection biomolecular NMR spectroscopy in living cells, *Angew. Chem. Int. Ed. 50*, 2339-2341.

10. Burz, D.S., Dutta, K., Cowburn, D., and Shekhtman, A. (2006) In-cell NMR for protein-protein interactions (STINT-NMR), *Nat. Protoc. 1*, 146-152.

11. Burz, D.S., Dutta, K., Cowburn, D., and Shekhtman, A. (2006) Mapping structural interactions using in-cell NMR spectroscopy (STINT-NMR), *Nat. Meth. 3*, 91-93.

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

13. Cruzeiro-Silva, C., Albernaz, F.P., Valente, A.P., and Almeida, F.C. (2006) In-cell NMR spectroscopy: Inhibition of autologous protein expression reduces *Escherichia coli* lysis, *Cell Biochem. Biophys. 44*, 497-502.

14. Dedmon, M.M., Patel, C.N., Young, G.B., and Pielak, G.J. (2002) FlgM gains structure in living cells, *Proc. Natl. Acad. Sci. U. S. A. 99*, 12681-12684.

15. Gochin, M., James, T.L., and Shafer, R.H. (1984) In vivo ^{19}F-NMR of 5-fluorouracil incorporation into RNA and metabolites in *Escherichia coli* cells, *Biochim. Biophys. Acta 804*, 118-124.

16. Li, C., Wang, G.F., Wang, Y., Creager-Allen, R., Lutz, E.A., Scronce, H., Slade, K.M., Ruf, R.A., Mehl, R.A., and Pielak, G.J. (2010) Protein ^{19}F NMR in *Escherichia coli*, *J. Am. Chem. Soc. 132*, 321-327.

17. Serber, Z., Lai, H.C., Yang, A., Ou, H.D., Sigal, M.S., Kelly, A.E., Darimont, B.D., Duijf, P.H., Van Bokhoven, H., McKeon, F., and Dotsch, V. (2002) A C-terminal inhibitory domain controls the activity of p63 by an intramolecular mechanism, *Mol. Cell. Biol. 22*, 8601-8611.

18. Serber, Z., Richter, C., Moskau, D., Bohlen, J.M., Gerfin, T., Marek, D., Haberli, M., Baselgia, L., Laukien, F., Stern, A.S., Hoch, J.C., and Dotsch, V. (2000) New carbon-detected protein NMR experiments using CryoProbes, *J. Am. Chem. Soc. 122*, 3554-3555.

19. Serber, Z., Straub, W., Corsini, L., Nomura, A.M., Shimba, N., Craik, C.S., Ortiz de Montellano, P., and Dötsch, V. (2004) Methyl groups as probes for proteins and complexes in in-cell NMR experiments, *J. Am. Chem. Soc. 126*, 7119-7125.

20. Wang, G.F., Li, C., and Pielak, G.J. (2010) ^{19}F NMR studies of alpha-synuclein-membrane interactions, *Protein Sci. 19*, 1686-1691.

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

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

23. Selenko, P., Serber, Z., Gadea, B., Ruderman, J., and 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.

24. Bodart, J.F., Wieruszeski, J.M., Amniai, L., Leroy, A., Landrieu, I., Rousseau-Lescuyer, A., Vilain, J.P., and Lippens, G. (2008) NMR observation of Tau in *Xenopus* oocytes, *J. Magn. Reson. 192*, 252-257.

25. Schnizler, K., Kuster, M., Methfessel, C., and Fejtl, M. (2003) The roboocyte: Automated cDNA/mRNA injection and subsequent TEVC recording on *Xenopus oocytes* in 96-well microtiter plates, *Receptors & channels 9*, 41-48.

26. Gard, D.L., and Kirschner, M.W. (1987) Microtubule assembly in cytoplasmic extracts of *Xenopus* oocytes and eggs, *J. Cell Biol. 105*, 2191-2201.

27. Smet, C., Leroy, A., Sillen, A., Wieruszeski, J.M., Landrieu, I., and Lippens, G. (2004) Accepting its random coil nature allows a partial NMR assignment of the neuronal Tau protein, *Chembiochem 5*, 1639-1646.

28. Lippens, G., Wieruszeski, J.M., Leroy, A., Smet, C., Sillen, A., Buee, L., and Landrieu, I. (2004) Proline-directed random-coil chemical shift values as a tool for the NMR assignment of the tau phosphorylation sites, *ChemBioChem 5*, 73-78.

29. Sillen, A., Barbier, P., Landrieu, I., Lefebvre, S., Wieruszeski, J.M., Leroy, A., Peyrot, V., and Lippens, G. (2007) NMR investigation of the interaction between the neuronal protein tau and the microtubules, *Biochemistry 46*, 3055-3064.

30. Sakai, T., Tochio, H., Inomata, K., Sasaki, Y., Tenno, T., Tanaka, T., Kokubo, T., Hiroaki, H., and Shirakawa, M. (2007) Fluoroscopic assessment of protein leakage during *Xenopus oocytes* in-cell NMR experiment by co-injected EGFP, *Anal. Biochem. 371*, 247-249.

31. Sakai, T., Tochio, H., Tenno, T., Ito, Y., Kokubo, T., Hiroaki, H., and Shirakawa, M. (2006) In-cell NMR spectroscopy of proteins inside *Xenopus laevis* oocytes, *J. Biomol. NMR 36*, 179-188.

32. Charlton, L.M., and Pielak, G.J. (2006) Peeking into living eukaryotic cells with high-resolution NMR, *Proc. Natl. Acad. Sci. U. S. A. 103*, 11817-11818.

33. Schlesinger, A.P., Wang, Y., Tadeo, X., Millet, O., and Pielak, G.J. (2011) Macromolecular crowding fails to fold a globular protein in cells, *J. Am. Chem. Soc. 133*, 8082-8085.

34. Xie, J., Thapa, R., Reverdatto, S., Burz, D.S., and Shekhtman, A. (2009) Screening of small molecule interactor library by using in-cell NMR spectroscopy (SMILI-NMR), *J. Med. Chem. 52*, 3516-3522.

35. Arnesano, F., Banci, L., Bertini, I., Felli, I.C., Losacco, M., and Natile, G. (2011) Probing the interaction of Cisplatin with the human copper chaperone atox1 by solution and in-cell NMR spectroscopy, *J. Am. Chem. Soc. 133*, 18361-18369.

36. Vogt, P.K. (2001) PI 3-kinase, mTOR, protein synthesis and cancer, *Trends Mol. Med. 7*, 482-484.

37. Selenko, P., Frueh, D.P., Elsaesser, S.J., Haas, W., Gygi, S.P., and Wagner, G. (2008) In situ observation of protein phosphorylation by high-resolution NMR spectroscopy, *Nat. Struct. Mol. Biol. 15*, 321-329.

38. Burz, D.S., and Shekhtman, A. (2008) In-cell biochemistry using NMR spectroscopy, *PLoS ONE 3*, e2571.

39. Hänsel, R., Foldynová-Trantírková, S., Löhr, F., Buck, J., Bongartz, E., Bamberg, E., Schwalbe, H., Dötsch, V., and Trantírek, L. (2009) Evaluation of parameters critical for observing nucleic acids inside living *Xenopus laevis* oocytes by in-cell NMR spectroscopy, *J. Am. Chem. Soc. 131*, 15761-15768.

40. Burz, D.S., and Shekhtman, A. (2010) The STINT-NMR method for studying in-cell protein-protein interactions, *Curr. Protoc. Protein Sci. Chapter 17*, Unit 17.11.

41. Maldonado, A.Y., Burz, D.S., and Shekhtman, A. (2011) In-cell NMR spectroscopy, *Prog. Nucl. Magn. Reson. Spectrosc. 59*, 197-212.

42. Pielak, G.J., Li, C., Miklos, A.C., Schlesinger, A.P., Slade, K.M., Wang, G.F., and Zigoneanu, I.G. (2009) Protein nuclear magnetic resonance under physiological conditions, *Biochemistry 48*, 226-234.

43. Robinson, K.E., Reardon, P.N., and Spicer, L.D. (2012) In-cell NMR spectroscopy in *Escherichia coli*, *Methods Mol. Biol. 831*, 261-277.

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

45. Pielak, G.J. (2007) Retraction, *Biochemistry 46*, 8206.

46. Serber, Z., Corsini, L., Durst, F., and Dotsch, V. (2005) In-cell NMR spectroscopy, *Methods Enzymol. 394*, 17-41.

47. Spitzer, J., and 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.

48. Spitzer, J.J., and Poolman, B. (2005) Electrochemical structure of the crowded cytoplasm, *Trends Biochem. Sci. 30*, 536-541.

49. Palmer, A.G., 3rd, Kroenke, C.D., and Loria, J.P. (2001) Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules, *Methods Enzymol. 339*, 204-238.

50. Cavanagh, J., Fairbrother, W.J., Palmer, A.G., Skelton, N.J., and Rance, M. (2006) *Protein NMR spectroscopy: Principles and practice*, 2nd ed., Academic Press, London.

51. Crowley, P.B., Brett, K., and Muldoon, J. (2008) NMR spectroscopy reveals cytochrome c-poly(ethylene glycol) interactions, *ChemBioChem 9*, 685-688.

52. Li, C., and Pielak, G.J. (2009) Using NMR to distinguish viscosity effects from nonspecific protein binding under crowded conditions, *J. Am. Chem. Soc. 131*, 1368-1369.

53. Li, C., Wang, Y., and Pielak, G.J. (2009) Translational and rotational diffusion of a small globular protein under crowded conditions, *J. Phys. Chem. B 113*, 13390-13392.

54. Wang, Y., Li, C., and Pielak, G.J. (2010) Effects of proteins on protein diffusion, *J. Am. Chem. Soc. 132*, 9392-9397.

55. Li, C., Charlton, L.M., Lakkavaram, A., Seagle, C., Wang, G., Young, G.B., Macdonald, J.M., and Pielak, G.J. (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.

56. Ikeya, T., Sasaki, A., Sakakibara, D., Shigemitsu, Y., Hamatsu, J., Hanashima, T., Mishima, M., Yoshimasu, M., Hayashi, N., Mikawa, T., Nietlispach, D., Walchli, M., Smith, B.O., Shirakawa, M., Guntert, P., and Ito, Y. (2010) NMR protein structure determination in living *E. coli* cells using nonlinear sampling, *Nat. Protoc. 5*, 1051-1060.

57. Reardon, P.N., and Spicer, L.D. (2005) Multidimensional NMR spectroscopy for protein characterization and assignment inside cells, *J. Am. Chem. Soc. 127*, 10848-10849.

58. Slade, K.M., Baker, R., Chua, M., Thompson, N.L., and Pielak, G.J. (2009) Effects of recombinant protein expression on green fluorescent protein diffusion in *Escherichia coli*, *Biochemistry 48*, 5083-5089.

59. Hedrick, W.R., Mathew, A., and Zimbrick, J.D. (1979) Intracellular viscosity of lymphocytes determined by a N-15 spin label probe, *J. Magn. Reson. 36*, 207-214.

60. Kneller, J.M., Lu, M., and Bracken, C. (2002) An effective method for the discrimination of motional anisotropy and chemical exchange, *J. Am. Chem. Soc. 124*, 1852-1853.

61. Feig, M., and Sugita, Y. (2012) Variable Interactions between Protein Crowders and Biomolecular Solutes Are Important in Understanding Cellular Crowding, *J. Phys. Chem. B 116*, 599-605.

62. McGuffee, S.R., and Elcock, A.H. (2010) Diffusion, crowding & protein stability in a dynamic molecular model of the bacterial cytoplasm, *PLoS Comput. Biol. 6*, e1000694.

63. Bryant, J.E., Lecomte, J.T., Lee, A.L., Young, G.B., and Pielak, G.J. (2005) Protein dynamics in living cells, *Biochemistry 44*, 9275-9279.

64. Bryant, J.E., Lecomte, J.T., Lee, A.L., Young, G.B., and Pielak, G.J. (2006) Cytosol has a small effect on protein backbone dynamics, *Biochemistry 45*, 10085-10091.

65. Barnes, C.O., and Pielak, G.J. (2011) In-cell protein NMR and protein leakage, *Proteins 79*, 347-351.

66. Luby-Phelps, K. (2000) Cytoarchitecture and physical properties of cytoplasm: volume, viscosity, diffusion, intracellular surface area, *Int. Rev. Cytol. 192*, 189-221.

67. Jackson, J.C., Hammill, J.T., and Mehl, R.A. (2007) Site-specific incorporation of a (19)F-amino acid into proteins as an NMR probe for characterizing protein structure and reactivity, *J. Am. Chem. Soc. 129*, 1160-1166.

68. Haggie, P.M., and Brindle, K.M. (1999) Mitochondrial citrate synthase is immobilized in vivo, *J. Biol. Chem. 274*, 3941-3945.

69. Williams, S.P., Fulton, A.M., and Brindle, K.M. (1993) Estimation of the intracellular free ADP concentration by ^{19}F NMR studies of fluorine-labeled yeast phosphoglycerate kinase in vivo, *Biochemistry 32*, 4895-4902.

70. Williams, S.P., Haggie, P.M., and Brindle, K.M. (1997) ^{19}F NMR measurements of the rotational mobility of proteins in vivo, *Biophys. J. 72*, 490-498.

71. Fu, R., Wang, X., Li, C., Santiago-Miranda, A.N., Pielak, G.J., and Tian, F. (2011) In situ structural characterization of a recombinant protein in native *Escherichia coli* membranes with solid-state magic-angle-spinning NMR, *J. Am. Chem. Soc. 133*, 12370-12373.

72. Vogel, E.P., Curtis-Fisk, J., Young, K.M., and Weliky, D.P. (2011) Solid-state nuclear magnetic resonance (NMR) spectroscopy of human immunodeficiency virus gp41 protein that includes the fusion peptide: NMR detection of recombinant fgp41 in inclusion bodies in whole bacterial cells and structural characterization of purified and membrane-associated fgp41, *Biochemistry 50*, 10013-10026.

73. Zavoisky, Y. (1945) Spin-magnetic resonance in paramagnetics, *J. Phys. USSR 9*, 245-249.

74. Schiemann, O., Piton, N., Plackmeyer, J., Bode, B.E., Prisner, T.F., and Engels, J.W. (2007) Spin labeling of oligonucleotides with the nitroxide TPA and use of PELDOR, a pulse EPR method, to measure intramolecular distances, *Nat. Protoc. 2*, 904-923.

75. Kohler, S.D., Weber, A., Howard, S.P., Welte, W., and Drescher, M. (2010) The proline-rich domain of TonB possesses an extended polyproline II-like conformation of sufficient length to span the periplasm of Gram-negative bacteria, *Protein Sci. 19*, 625-630.

76. Robotta, M., Braun, P., van Rooijen, B., Subramaniam, V., Huber, M., and Drescher, M. (2011) Direct evidence of coexisting horseshoe and extended helix conformations of membrane-bound alpha-synuclein, *Chemphyschem : a European journal of chemical physics and physical chemistry 12*, 267-269.

77. Usselman, R.J., Walter, E.D., Willits, D., Douglas, T., Young, M., and Singel, D.J. (2011) Monitoring structural transitions in icosahedral virus protein cages by site-directed spin labeling, *J. Am. Chem. Soc. 133*, 4156-4159.

78. Azarkh, M., Okle, O., Eyring, P., Dietrich, D.R., and Drescher, M. (2011) Evaluation of spin labels for in-cell EPR by analysis of nitroxide reduction in cell extract of *Xenopus laevis* oocytes, *J. Magn. Reson. 212*, 450-454.

79. Azarkh, M., Okle, O., Singh, V., Seemann, I.T., Hartig, J.S., Dietrich, D.R., and Drescher, M. (2011) Long-range distance determination in a DNA model system inside *Xenopus laevis* oocytes by in-cell spin-label EPR, *ChemBioChem 12*, 1992-1995.

80. Igarashi, R., Sakai, T., Hara, H., Tenno, T., Tanaka, T., Tochio, H., and Shirakawa, M. (2010) Distance determination in proteins inside *Xenopus laevis* oocytes by double electron-electron resonance experiments, *J. Am. Chem. Soc. 132*, 8228-8829.

81. Krstic, I., Hansel, R., Romainczyk, O., Engels, J.W., Dotsch, V., and Prisner, T.F. (2011) Long-range distance measurements on nucleic acids in cells by pulsed EPR spectroscopy, *Angew. Chem. Int. Ed. 50*, 5070-5074.

82. Borbat, P.P., McHaourab, H.S., and Freed, J.H. (2002) Protein structure determination using long-distance constraints from double-quantum coherence ESR: study of T4 lysozyme, *J. Am. Chem. Soc. 124*, 5304-5314.

83. Altenbach, C., Kusnetzow, A.K., Ernst, O.P., Hofmann, K.P., and Hubbell, W.L. (2008) High-resolution distance mapping in rhodopsin reveals the pattern of helix movement due to activation, *Proc. Natl. Acad. Sci. U. S. A. 105*, 7439-7444.

84. Gierasch, L.M., and Gershenson, A. (2009) Post-reductionist protein science, or putting Humpty Dumpty back together again, *Nat. Chem. Biol. 5*, 774-777.

85. Malmstrom, J., Beck, M., Schmidt, A., Lange, V., Deutsch, E.W., and Aebersold, R. (2009) Proteome-wide cellular protein concentrations of the human pathogen *Leptospira interrogans*, *Nature 460*, 762-765.

86. Brindle, K.M., Williams, S.-P., and Boulton, M. (1989) ^{19}F NMR detection of a fluorine-labelled enzyme in vivo, *FEBS Lett. 255*, 121-124.

87. Hubbard, J.A., MacLachlan, L.K., King, G.W., Jones, J.J., and Fosberry, A.P. (2003) Nuclear magnetic resonance spectroscopy reveals the functional state of the signalling protein CheY in vivo in Escherichia coli, *Mol. Microbiol. 49*, 1191-1200.

88. McNulty, B.C., Young, G.B., and Pielak, G.J. (2006) Macromolecular crowding in the *Escherichia coli* periplasm maintains alpha-synuclein disorder, *J. Mol. Biol. 355*, 893-897.

89. Selenko, P., and Wagner, G. (2006) NMR mapping of protein interactions in living cells, *Nat. Meth. 3*, 80-81.

90. Serber, Z., Keatinge-Clay, A.T., Ledwidge, R., Kelly, A.E., Miller, S.M., and Dötsch, V. (2001) High-resolution macromolecular NMR spectroscopy inside living cells, *J. Am. Chem. Soc. 123*, 2446-2447.

91. Serber, Z., Selenko, P., Hansel, R., Reckel, S., Lohr, F., Ferrell, J.E., Jr., Wagner, G., and Dötsch, V. (2006) Investigating macromolecules inside cultured and injected cells by in-cell NMR spectroscopy, *Nat. Protoc. 1*, 2701-2709.

92. Cellitti, S.E., Jones, D.H., Lagpacan, L., Hao, X., Zhang, Q., Hu, H., Brittain, S.M., Brinker, A., Caldwell, J., Bursulaya, B., Spraggon, G., Brock, A., Ryu, Y., Uno, T., Schultz, P.G., and Geierstanger, B.H. (2008) *In vivo* incorporation of unnatural amino acids to probe structure, dynamics, and ligand binding in a large protein by nuclear magnetic resonance spectroscopy, *J. Am. Chem. Soc. 130*, 9268-9281.

93. Danielson, M.A., and Falke, J.J. (1996) Use of ^{19}F NMR to probe protein structure and conformational changes, *Annu. Rev. Biophys. Biomol. Struct. 25*, 163-195.

94. Li, C., Lutz, E.A., Slade, K.M., Ruf, R.A., Wang, G., and Pielak, G.J. (2009) ^{19}F-NMR studies of α-synuclein conformation and fibrillation, *Biochemistry 48*, 8578-8584.

95. Li, H., and Frieden, C. (2007) Observation of sequential steps in the folding of intestinal fatty acid binding protein using a slow folding mutant and ^{19}F NMR, *Proc. Natl. Acad. Sci. U. S. A. 104*, 11993-11998.

96. Pervushin, K.V., Wider, G., Riek, R., and Wüthrich, K. (1999) The 3D NOESY-[(1)H,(15)N,(1)H]-ZQ-TROSY NMR experiment with diagonal peak suppression, *Proc. Natl. Acad. Sci. U. S. A. 96*, 9607-9612.

97. Charlton, L.M., Barnes, C.O., Li, C., Orans, J., Young, G.B., and Pielak, G.J. (2008) Residue-level interrogation of macromolecular crowding effects on protein stability, *J. Am. Chem. Soc. 130*, 6826-6830.

98. Petit, C.M., Zhang, J., Sapienza, P.J., Fuentes, E.J., and Lee, A.L. (2009) Hidden dynamic allostery in a PDZ domain, *Proc. Natl. Acad. Sci. U. S. A. 106*, 18249-18254.

99. Khan, F., Kuprov, I., Craggs, T.D., Hore, P.J., and Jackson, S.E. (2006) ^{19}F NMR studies of the native and denatured states of green fluorescent protein, *J. Am. Chem. Soc. 128*, 10729-10737.

100. Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor University Press, New York.

101. Hammill, J.T., Miyake-Stoner, S., Hazen, J.L., Jackson, J.C., and Mehl, R.A. (2007) Preparation of site-specifically labeled fluorinated proteins for ^{19}F-NMR structural characterization, *Nat. Protoc. 2*, 2601-2607.

102. Conway, K.A., Lee, S.J., Rochet, J.C., Ding, T.T., Williamson, R.E., and Lansbury, P.T., Jr. (2000) Acceleration of oligomerization, not fibrillization, is a shared property of both α-synuclein mutations linked to early-onset Parkinson's disease: implications for pathogenesis and therapy, *Proc. Natl. Acad. Sci. U. S. A. 97*, 571-576.

103. Serber, Z., Ledwidge, R., Miller, S.M., and 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.

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

105. Johnson, B.A., and Blevins, R.A. (1994) NMR View: A computer program for the visualization and analysis of NMR data, *J. Biomol. NMR 4*, 603-614.

106. Wintrode, P.L., Makhatadze, G.I., and Privalov, P.L. (1994) Thermodynamics of ubiquitin unfolding, *Proteins 18*, 246-253.

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

108. McCabe, T.J., Fulton, D., Roman, L.J., and Sessa, W.C. (2000) Enhanced electron flux and reduced calmodulin dissociation may explain "calcium-independent" eNOS activation by phosphorylation, *J. Biol. Chem. 275*, 6123-6128.

109. Koch, A.L. (1987) *The Variability and Individuality of the Bacterium*, American Society for Microbiology, Washington DC.

110. Huang, C., Ren, G., Zhou, H., and Wang, C.C. (2005) A new method for purification of recombinant human alpha-synuclein in Escherichia coli, *Protein Expr. Purif. 42*, 173-177.

111. Shevchik, V.E., Condemine, G., and Robert-Baudouy, J. (1994) Characterization of DsbC, a periplasmic protein of Erwinia chrysanthemi and Escherichia coli with disulfide isomerase activity, *EMBO J. 13*, 2007-2012.

112. Birdsell, D.C., and Cota-Robles, E.H. (1967) Production and ultrastructure of lysozyme and ethylenediaminetetraacetate-lysozyme spheroplasts of Escherichia coli, *J. Bacteriol. 93*, 427-437.

113. Schneider, D.M., Dellwo, M.J., and Wand, A.J. (1992) Fast internal main-chain dynamics of human ubiquitin, *Biochemistry 31*, 3645-3652.

114. Law, A.B., Fuentes, E.J., and Lee, A.L. (2009) Conservation of side-chain dynamics within a protein family, *J. Am. Chem. Soc. 131*, 6322-6323.

115. Jaren, O.R., Kranz, J.K., Sorensen, B.R., Wand, A.J., and Shea, M.A. (2002) Calcium-induced conformational switching of Paramecium calmodulin provides evidence for domain coupling, *Biochemistry 41*, 14158-14166.

116. Kitevski-LeBlanc, J.L., Evanics, F., and Prosser, R.S. (2009) Approaches for the measurement of solvent exposure in proteins by ^{19}F NMR, *J. Biomol. NMR 45*, 255-264.

117. Pedersen, S., Bloch, P.L., and Neidhardt, F.C. (1978) Patterns of protein expression in E. coli: a catalog of the amount of 140 individual proteins at different growth rates, *Cell 14*, 179-190.

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

119. Eliezer, D., Kutluay, E., Bussell, R., Jr., and Browne, G. (2001) Conformational properties of α-synuclein in its free and lipid-associated states, *J. Mol. Biol. 307*, 1061-1073.

120. Wu, K.P., Kim, S., Fela, D.A., and Baum, J. (2008) Characterization of conformational and dynamic properties of natively unfolded human and mouse α-synuclein ensembles by NMR: implication for aggregation, *J. Mol. Biol. 378*, 1104-1115.

121. Debye, P.J. (1929) *Polar Molecules*, Chemical Catalog Company, New York.

122. Srere, P.A. (2000) Macromolecular interactions: tracing the roots, *Trends Biochem. Sci. 25*, 150-153.

123. Persson, E., and Halle, B. (2008) Cell water dynamics on multiple time scales, *Proc. Natl. Acad. Sci. U. S. A. 105*, 6266-6271.

124. Jarymowycz, V.A., and Stone, M.J. (2006) Fast time scale dynamics of protein backbones: NMR relaxation methods, applications, and functional consequences, *Chem. Rev. 106*, 1624-1671.

125. Xu, A.S.L., Waldeck, R., and Kuchel, P.W. (1993) Transmembrane ^{19}F NMR chemical shift difference of fluorinated solutes in liposomes, erythrocytes and erythrocyte ghosts, *NMR Biomed. 6*, 136-143.

126. Ghaemmaghami, S., and Oas, T.G. (2001) Quantitative protein stability measurements *in vivo*, *Nat. Struct. Mol. Biol. 8*, 879-882.

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

128. Lindon, J.C., Beckonert, O.P., Holmes, E., and Nicholson, J.K. (2009) High-resolution magic angle spinning NMR spectroscopy: Application to biomedical studies, *Prog. Nucl. Magn. Reson. Spectrosc. 55*, 79-100.

129. Ai, X., Zhou, Z., Bai, Y., and Choy, W.Y. (2006) ^{15}N NMR spin relaxation dispersion study of the molecular crowding effects on protein folding under native conditions, *J. Am. Chem. Soc. 128*, 3916-3917.

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

131. Munishkina, L.A., Ahmad, A., Fink, A.L., and Uversky, V.N. (2008) Guiding protein aggregation with macromolecular crowding, *Biochemistry 47*, 8993-9006.

132. Elowitz, M.B., Surette, M.G., Wolf, P.E., Stock, J.B., and Leibler, S. (1999) Protein mobility in the cytoplasm of Escherichia coli, *J. Bacteriol. 181*, 197-203.

133. Konopka, M.C., Shkel, I.A., Cayley, S., Record, M.T., and Weisshaar, J.C. (2006) Crowding and confinement effects on protein diffusion in vivo, *J. Bacteriol. 188*, 6115-6123.

134. Einstein, A. (1906) On the theory of the Brownian movement, *Annalen der Physik 324*, 371-381.

135. Einstein, A. (1956) *Investigations on the theory of the Brownian movement*, Dover Publications, New York.

136. Banks, D.S., and Fradin, C. (2005) Anomalous diffusion of proteins due to molecular crowding, *Biophys. J. 89*, 2960-2971.

137. Dauty, E., and Verkman, A.S. (2004) Molecular crowding reduces to a similar extent the diffusion of small solutes and macromolecules: measurement by fluorescence correlation spectroscopy, *J. Mol. Recognit. 17*, 441-447.

138. Goins, A.B., Sanabria, H., and Waxham, M.N. (2008) Macromolecular crowding and size effects on probe microviscosity, *Biophys. J. 95*, 5362-5373.

139. Kozer, N., Kuttner, Y.Y., Haran, G., and Schreiber, G. (2007) Protein-protein association in polymer solutions: from dilute to semidilute to concentrated, *Biophys. J. 92*, 2139-2149.

140. Kuttner, Y.Y., Kozer, N., Segal, E., Schreiber, G., and Haran, G. (2005) Separating the contribution of translational and rotational diffusion to protein association, *J. Am. Chem. Soc. 127*, 15138-15144.

141. Lavalette, D., Hink, M.A., Tourbez, M., Tetreau, C., and Visser, A.J. (2006) Proteins as micro viscosimeters: Brownian motion revisited, *Eur. Biophys. J. 35*, 517-522.

142. Muramatsu, N., and Minton, A.P. (1988) Tracer diffusion of globular proteins in concentrated protein solutions, *Proc. Natl. Acad. Sci. U. S. A. 85*, 2984-2988.

143. Zorrilla, S., Hink, M.A., Visser, A.J., and Lillo, M.P. (2007) Translational and rotational motions of proteins in a protein crowded environment, *Biophys. Chem. 125*, 298-305.

144. Ferrage, F., Zoonens, M., Warschawski, D.E., Popot, J.L., and Bodenhausen, G. (2003) Slow diffusion of macromolecular assemblies by a new pulsed field gradient NMR method, *J. Am. Chem. Soc. 125*, 2541-2545.

145. Kay, L.E., Torchia, D.A., and Bax, A. (1989) Backbone dynamics of proteins as studied by ^{15}N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease, *Biochemistry 28*, 8972-8979.

146. Molyneux, P. (1983) *Water-soluble synthetic polymers : properties and behavior*, CRC Press, Boca Raton.

147. Fissell, W.H., Manley, S., Dubnisheva, A., Glass, J., Magistrelli, J., Eldridge, A.N., Fleischman, A.J., Zydney, A.L., and Roy, S. (2007) Ficoll is not a rigid sphere, *Am. J. Physiol. Renal Physiol. 293*, F1209-1213.

148. Whitley, M.J., Zhang, J., and Lee, A.L. (2008) Hydrophobic core mutations in CI2 globally perturb fast side-chain dynamics similarly without regard to position, *Biochemistry 47*, 8566-8576.

149. McPhalen, C.A., Svendsen, I., Jonassen, I., and James, M.N. (1985) Crystal and molecular structure of chymotrypsin inhibitor 2 from barley seeds in complex with subtilisin Novo, *Proc. Natl. Acad. Sci. U. S. A. 82*, 7242-7246.

150. Hughson, F.M., Barrick, D., and Baldwin, R.L. (1991) Probing the stability of a partly folded apomyoglobin intermediate by site-directed mutagenesis, *Biochemistry 30*, 4113-4118.

151. Rubinstein, M., and Colby, R.H. (2003) *Polymer Physics*, Oxford University Press, USA.

152. Aste, T., and Weaire, D. (2008) *The pursuit of perfect packing*, 2nd ed., Taylor & Francis, New York.

153. Vaynberg, J., and Qin, J. (2006) Weak protein-protein interactions as probed by NMR spectroscopy, *Trends Biotechnol. 24*, 22-27.

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

155. Durek, P., and Walther, D. (2008) The integrated analysis of metabolic and protein interaction networks reveals novel molecular organizing principles, *BMC Syst. Biol. 2*, 100.

156. Elcock, A.H. (2010) Models of macromolecular crowding effects and the need for quantitative comparisons with experiment, *Curr. Opin. Struct. Biol. 20*, 196-206.

157. Minton, A.P., and Wilf, J. (1981) Effect of macromolecular crowding upon the structure and function of an enzyme: glyceraldehyde-3-phosphate dehydrogenase, *Biochemistry 20*, 4821-4826.

158. Minton, A.P. (1981) Excluded volume as a determinant of macromolecular structure and reactivity, *Biopolymers 20*, 2093-2120.

159. Hermans, J., and Scheraga, H.A. (1961) Structural studies of ribonuclease. V. reversible change of configuration, *J. Am. Chem. Soc. 83*, 3283-3292.

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

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

162. Zhou, H.-X., Rivas, G., and Minton, A.P. (2008) Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences, *Annu. Rev. Biophys. 37*, 353-373.

163. Lumry, R., and Rajender, S. (1970) Enthalpy-entropy compensation and phenomena in water solutions of proteins and small molecules: A ubiquitous property of water, *Biopolymers 9*, 1125-1127.

164. Miklos, A.C., Li, C., and Pielak, G.J. (2009) Using NMR-detected backbone amide ^{1}H exchange to assess macromolecular crowding effects on globular-protein stability, *Methods Enzymol. 466*, 1-18.

165. Dhar, A., Samiotakis, A., Ebbinghaus, S., Nienhaus, L., Homouz, D., Gruebele, M., and Cheung, M.S. (2010) Structure, function, and folding of phosphoglycerate kinase are strongly perturbed by macromolecular crowding, *Proc. Natl. Acad. Sci. U. S. A. 107*, 17586-17591.

166. Miklos, A.C., Li, C., Sharaf, N.G., and Pielak, G.J. (2010) Volume exclusion and soft interaction effects on protein stability under crowded conditions, *Biochemistry 49*, 6984-6991.

167. Schowen, K.B., and Schowen, R.L. (1982) Solvent isotope effects of enzyme systems, *Methods Enzymol. 87*, 551-606.

168. Kay, L.E., Keifer, P., and Saarinen, T. (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity, *J. Am. Chem. Soc. 114*, 10663-10665.

169. Bodenhausen, G., and Ruben, D.J. (1980) Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy, *Chem. Phys. Lett. 69*, 185-189.

170. Berger, A., and Linderstrom-Lang, K. (1957) Deuterium exchange of poly-DL-alanine in aqueous solution, *Arch. Biochem. Biophys. 69*, 106-118.

171. Englander, S.W., and Kallenbach, N.R. (1983) Hydrogen exchange and structural dynamics of proteins and nucleic acids., *Q. Rev. Biophys. 16*, 521-655.

172. Benitez-Cardoza, C.G., Stott, K., Hirshberg, M., Went, H.M., Woolfson, D.N., and Jackson, S.E. (2004) Exploring sequence/folding space: folding studies on multiple hydrophobic core mutants of ubiquitin, *Biochemistry 43*, 5195-5203.

173. Sivaraman, T., Arrington, C.B., and Robertson, A.D. (2001) Kinetics of unfolding and folding from amide hydrogen exchange in native ubiquitin, *Nat. Struct. Biol. 8*, 331-333.

174. Zhang, Y.-Z. (1995) Protein and peptide structure and interactions studied by hydrogen exchange and NMR, In *Structural Biology and Molecular Biophysics*, University of Pennsylvania, Philadelphia.

175. Bai, Y., Milne, J.S., Mayne, L., and Englander, S.W. (1993) Primary structure effects on peptide group hydrogen exchange, *Proteins 17*, 75-86.

176. Gladwin, S.T., and Evans, P.A. (1996) Structure of very early protein folding intermediates: New insights through a variant of hydrogen exchange labelling, *Fold Des 1*, 407-417.

177. Ibarra-Molero, B., Loladze, V.V., Makhatadze, G.I., and Sanchez-Ruiz, J.M. (1999) Thermal versus guanidine-induced unfolding of ubiquitin. An analysis in terms of the contributions from charge-charge interactions to protein stability, *Biochemistry 38*, 8138-8149.

178. Makhatadze, G.I., Lopez, M.M., Richardson, J.M., 3rd, and Thomas, S.T. (1998) Anion binding to the ubiquitin molecule, *Protein Sci. 7*, 689-697.

179. Grasso, D.M., Sciacca, M.F.M., Milardi, D., Pappalardo, M., and La Rosa, C. (2006) Role of electrostatics in the thermal stability of ubiquitin - a combined DSC and MM study, *J. Therm. Anal. Calorim. 86*, 311-314.

180. Makhatadze, G.I., Clore, G.M., and Gronenborn, A.M. (1995) Solvent isotope effect and protein stability, *Nat. Struct. Biol. 2*, 852-855.

181. Michnik, A., Michalik, K., and Drzazga, Z. (2005) Stability of bovine serum albumin at different pH, *J. Therm. Anal. Calorim. 80*, 399-406.

182. Becktel, W.J., and Schellman, J.A. (1987) Protein stability curves, *Biopolymers 26*, 1859-1877.

183. Hong, J., and Gierasch, L.M. (2010) Macromolecular crowding remodels the energy landscape of a protein by favoring a more compact unfolded state, *J. Am. Chem. Soc. 132*, 10445-10452.

184. Lavrenko, P.N., Mikriukova, O.I., and Okatova, O.V. (1987) On the separation ability of various Ficoll gradient solutions in zonal centrifugation, *Anal. Biochem. 166*, 287-297.

185. Squire, P.G., and Himmel, M.E. (1979) Hydrodynamics and protein hydration, *Arch. Biochem. Biophys. 196*, 165-177.

186. Gralla, J.D. (2005) Escherichia coli ribosomal RNA transcription: regulatory roles for ppGpp, NTPs, architectural proteins and a polymerase-binding protein, *Mol. Microbiol. 55*, 973-977.

187. Toro, E., and Shapiro, L. (2010) Bacterial chromosome organization and segregation, *Cold Spring Harbor perspectives in biology 2*, a000349.

188. Spitzer, J. (2011) From water and ions to crowded biomacromolecules: in vivo structuring of a prokaryotic cell, *Microbiol. Mol. Biol. Rev. 75*, 491-506.

189. Shapiro, L., and Losick, R. (1997) Protein localization and cell fate in bacteria, *Science 276*, 712-728.