Protein NMR in Insect and Bacterial Cells

Alexander P. Schlesinger

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 Biochemistry and Biophysics.

Chapel Hill
2011

Approved by:

Gary Pielak, Ph.D.

Andrew Lee, Ph.D.

Michael Jarstfer, Ph.D.

Matthew Redinbo, Ph.D.

Ashutosh Tripathy, Ph.D.

Abstract

Alexander P. Schlesinger: Protein NMR in Insect and Bacterial Cells
(Under the direction of Gary J. Pielak, Ph.D.)

Proteins perform their functions in cells where macromolecular solutes can reach concentrations of >300 g/L and occupy >30% of the volume. The volume excluded by these macromolecules will stabilize globular proteins because the native state of a globular protein occupies less space than the denatured state. The idea that volume exclusion should fold globular proteins is tested in *Escherichia coli* by using protein L Kx7E, a 7-kDa unstable globular protein variant of wild-type protein L. This variant requires high concentrations of monovalent salts to fold. The standard unfolding free energy of Kx7E in the absence of NaCl is -1.0 kcal/mol as determined by using *in vitro* ^{1}H-^{15}N nuclear magnetic resonance spectroscopy. In-cell nuclear magnetic resonance spectroscopy with ^{19}F-labeled Kx7E is used to show that the crowded cytoplasm of *E. coli* cannot overcome even this modest free energy deficit and fold the variant. The data are consistent with the idea that non-specific interactions between cytoplasmic components can overcome the excluded volume effect. Evidence for these interactions is provided by the observation that adding simple salts folds Kx7E in dilute solution, but increasing the salt concentration inside *E. coli* does not fold the variant.

An attempt was made to acquire nuclear magnetic resonance spectra of the intrinsically disordered protein α-synuclein and the green fluorescent protein inside insect cells. Infection of *Spodoptera frugiperda* insect cells with recombinant baculoviruses containing a late basic protein promoter produces sufficient recombinant protein while maintaining cell viability for spectral acquisition. However, incorporation of the nuclear magnetic resonance active isotopes ^{15}N or ^{19}F into the recombinant proteins through the addition of $^{15}NH_4Cl$ or 5-fluorotryptophan to the cell culture media proved unsuccessful. The similarity of the spectra for α-synuclein, green fluorescent protein, and a fusion protein suggests isotope incorporation into intracellular metabolites. Thus, protein nuclear magnetic resonance spectra cannot be acquired in insect cells by using these experimental conditions.

Acknowledgements

There are many people who deserve credit for helping me throughout my graduate career. I would not have been able to finish without their guidance. I would like to thank my fellow graduate and undergraduate students in the Pielak lab. I appreciate that this intelligent collection of individuals is always willing to help each other. In particular, I want to thank Yaqiang Wang for helping me numerous times with NMR data acquisition and figure preparation. You always helped me when I needed it, and this dissertation could not have been completed without you.

I want to thank Xavier Tadeo and Oscar Millet for providing the basis for much of my dissertation. Thank you Michael Chua for providing laser scanning confocal microscopy training. Thank you Marc ter Horst and Gregory Young for maintaining the NMR spectrometers. I also want to express my thanks to Elizabeth Pielak for taking the time to read and comment on my manuscripts.

I want to thank Matthew Redinbo and Andrew Lee for taking the time to be on my graduate committee and for providing helpful guidance. I also want to thank Michael Jarstfer and Ashutosh Tripathy for agreeing to be a part of my committee. Thank you Mike for giving me the opportunity to rotate in your lab. Thank you Ash for teaching me how to use the instruments at the Macromolecular Interactions Facility. I want to thank Gary Pielak for giving me

the opportunity to join his lab. You provided me with the advice and guidance I needed to succeed. Thank you for letting me attempt several interesting and challenging projects. Even though I made many mistakes, your understanding and enthusiasm motivated me to do my best.

Finally, I want to thank my parents Jo-Anne and Siegfried Schlesinger for providing an environment conducive to learning and for always listening to my day-to-day graduate school experiences, good and bad.

Table of Contents

List of Tables

List of Figures

List of Abbreviations and Symbols

Å	angstrom
A.A.	amino acid
AcGFP	*Aequorea coerulescens* green fluorescent protein
AcMNPV	*Autographa californica* multiple nuclear polyhedrosis virus
AMP	ampicillin
avg.	average
BLAST	basic local alignment search tool
BSA	bovine serum albumin
°C	degree Celsius
cm	centimeter
D_2O	deuterium oxide
DNA	deoxyribonucleic acid
EDTA	ethylenediaminetetraacetic acid
EGFP	enhanced green fluorescent protein
Ek	enterokinase
FBS	fetal bovine serum
3-FY	3-fluorotyrosine
5-FW	5-fluorotryptophan
g	gram
x g	times gravity
GEM	gentamicin
GFP	green fluorescent protein

h	hour
His	histidine
hr5	homologous region 5
HRP	horse radish peroxidase
HSP27	human heat shock protein 1
hsp70	heat shock protein 70
HSQC	heteronuclear single-quantum coherence
Hz	hertz
IE1	immediate early 1
IgG	immunoglobulin G
IPTG	isopropyl-β-D-1-thiogalactopyranoside
kcal	kilocalorie
kDa	kilodalton
kHz	kilohertz
L	liter
LB	Luria broth
LIC	ligation-independent cloning
M	Molar
m	meter
mA	milliampere
MCS	multiple cloning site
mg	milligram
MHz	megahertz

min	minute
mL	milliliter
mm	millimeter
mM	millimolar
MOI	multiplicity of infection
mol	mole
mOsm	osmolality (mmol/kg)
MWCO	molecular weight cutoff
ng	nanogram
Ni-NTA	nickel-nitrilotriacetic acid
nL	nanoliter
nm	nanometer
NMR	nuclear magnetic resonance
OD_{600}	optical density at 600 nm
Osm	osmolality (mol/kg)
p.i.	post infection
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEG	polyethylene glycol
pfu	plaque forming unit
PH	polyhedrin
pL	picoliter
PMSF	phenylmethanesulphonylfluoride

ppm	chemical shift in parts per million
ProtL	protein L
PVDF	polyvinylidene difluoride
PVP	polyvinylpyrrolidone
RMSD	root mean square deviation
RNase A	ribonuclease A
rpm	revolutions per minute
s	second
S2	Schneider 2
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
Sf9	*Spodoptera frugiperda* 9
SFM	serum-free medium
$TCID_{50}$	50% tissue culture infectious dose method
TE	tris ethylenediaminetetraacetic acid
V	volt
v/v	volume/volume
w/v	weight/volume
α-syn	α-synuclein
$\Delta G^{\circ\prime}$	standard unfolding free energy
$\Delta G_0{}^{\circ\prime}$	standard unfolding free energy in the absence of NaCl
µL	microliter
µm	micrometer

µM	micromolar
π	pi
%	percent

Chapter 1

Macromolecular Crowding

1.1 Macromolecular Crowding Changes Protein Properties

Almost all protein studies have been conducted in dilute solutions where the total concentration of macromolecules is less than 10 g/L. Proteins, however perform their biological functions in cells where macromolecules are present at hundreds of g/L and occupy about 30% of the available volume (1-4).

The effect of high concentrations of macromolecules on equilibrium properties arises from two phenomena. The first, excluded volume, is the result of the impenetrable nature of atoms. The volume excluded by the crowding molecules is unavailable to the test protein. The native, folded state of a globular protein takes up less space than the denatured state. Application of Le Chatelier's principle leads to the conclusion that volume exclusion always favors the native state because it occupies less space (5).

The other phenomenon involves chemical interactions, which can be specific or non-specific. If the crowding molecule interacts with only the native state, the effect is stabilizing. If the crowder has an affinity for protein surface in general, however, the effect is destabilizing. The opposing chemical effects are reminiscent of ligand binding and urea denaturation. Ligand binding pulls the equilibrium between the native and denatured state toward the native state

because the crowder binds this state. Urea pulls the equilibrium toward the denatured state because the denatured state exposes more surface area. Although high concentrations of urea also introduce an excluded volume effect, this contribution is smaller than that from chemical interactions with the denatured state (6). Thus, one cannot know, a priori, how macromolecular crowding will affect globular protein stability.

Predictions from theory suggest this excluded volume will cause large changes in protein properties, including structure, folding rate, equilibrium constants of association, and reaction equilibria (1), compared to dilute solutions. Macromolecular crowding is also expected to affect biological function (7, 8). For example, crowding-induced reduction of protein diffusion is expected to reduce the rate of diffusion-limited interactions, thus potentially affecting metabolic pathways (1). Additionally, crowding is predicted to enhance the collapse of polypeptide chains into functional proteins and assist in the assembly of oligomeric structures (1).

Experiments support these predictions both *in vitro* (8-14) and in cells (2, 15-19). DNA polymerase is functional under an extended range of inhibitory conditions in the presence of *in vitro* crowders due to increased binding of the polymerase to DNA (20, 21). Studies in cells show that the disordered part of FlgM gains structure in *Escherichia coli* (17), and these structural changes suggest that the protein has a low inherent stability which may affect its transport into the extracellular medium (22).

1.2 Macromolecular Crowding Should be Studied in Cells

With the exception of a few in-cell studies (17-19, 23-28), most of the information about macromolecular crowding effects on proteins comes from *in vitro* crowding studies which are often conducted by using a single synthetic crowder with limited polydispersity (29). Cells, however, contain many different proteins, nucleic acids, and other macromolecules of different sizes, interactions, and localizations within the cytoplasm (30, 31).

The complexity in cells contributes to the macromolecular crowding effect and can alter protein properties compared to *in vitro* crowded solutions. *In vitro* experiments attempting to probe the complexity of macromolecular crowding in cells, including differences in macromolecule sizes and localizations, show that these factors affect protein properties. For example, a model cytoplasm separated into two protein phases was demonstrated to differentially distribute newly introduced proteins and nucleic acids (31). A simulation comprising 50 of the most abundant macromolecules in *E. coli* at physiological concentrations, and representing 85% of the characterized cytoplasmic contents by weight, demonstrated changes in protein stability compared to values calculated by volume exclusion alone (32). These results suggest that effects in addition to volume exclusion must be considered when studying proteins *in vivo*. In another study, denatured RNase A refolded into a native-like structure in the presence of high concentrations of 20 kDa polyethylene glycol (PEG) but did not refold with 0.2 kDa PEG (14), suggesting that size plays a role in macromolecular crowding.

Despite the useful information from *in vitro* experiments, some protein properties cannot be explored *in vitro*, arguing for protein study in cells. Elowitz et al. showed that the diffusion coefficient of the green fluorescent protein (GFP) in *E. coli* is 11 times less than that in water (2), but only 3 times less using *in vitro* protein crowder concentrations comparable to the total cellular protein concentration (200 g/L) (33). In another experiment, the reduced form of lysozyme was unable to refold properly in the presence of *in vitro* crowded solutions (12). However, reduced lysozyme refolds efficiently in cells, indicating that additional factors contribute to macromolecular crowding in cells such as chaperones which assist in protein folding (1).

Furthermore, cells are not just crowded bags of molecules, but are highly organized, which aids in cellular functions (34). For example, the five enzymes involved in converting D-mandelate to benzoate in *Pseudomonas putida* were shown to form a complex in cells (35). Glycosomal enzyme complexes were purified from the protist *Trypanosoma brucei* by cross-linking with reagents that span just 11-12 Å, suggesting that the enzymes are in close proximity and possibly explaining the high rate of glycolysis in this organism (36). Additionally, fluorescence microscopy experiments show colocalization of enzymes involved in purine synthesis in HeLa cells (37).

It is important to study proteins under physiologically relevant conditions to gain physiologically relevant data. Because of the difficulty mimicking the crowded cellular environment *in vitro*, studying proteins directly in cells or in lysates provides the best way to obtain physiologically relevant protein

information. These studies highlight differences in protein properties in cells versus *in vitro* crowded conditions and argue for the study of proteins in cells.

1.3 NMR is Ideally Suited to Study Macromolecular Crowding in Cells

Some techniques exist for studying protein properties both *in vitro* and *in vivo*, but most of them are unsuitable for direct study of proteins in live cells. A common method is to attach a reporter molecule, such as GFP, to the protein of interest. The reporter molecule provides a convenient means of detection, but potentially alters properties such as stability and diffusion. Other techniques like analytical ultracentrifugation provide information about proteins in solution but cannot be used to study these properties in live cells.

Nuclear magnetic resonance (NMR) spectroscopy is a noninvasive technique that provides a wide range of atomic resolution information about macromolecules in aqueous solutions or under physiologically relevant, in-cell or *in vitro*, conditions (23). "In-cell NMR spectroscopy" (23, 38) is the observation of macromolecules directly in living cells. In-cell NMR has been used to determine protein structure (17, 39, 40), post-translational modification (27, 41), signal transduction (42), and binding (43, 44). Requirements for acquisition of high resolution protein in-cell spectra include high protein concentration (39, 45), the absence of test protein leakage from the cells (46), cell viability over the course of the experiment, and isotopic enrichment of the test protein (39, 47).

Common in-cell NMR experiments in *E. coli* include the ^{1}H-^{15}N heteronuclear single-quantum coherence (HSQC) experiment, as implemented by the Dötch group (23). In cells, this experiment is limited to disordered proteins.

Globular proteins are not observable, probably due to restricted rotational motion in *E. coli* (48, 49). ^{19}F NMR, a more sensitive technique (50, 51), is also used for protein in-cell NMR experiments. Brindle et al. demonstrated incorporation of ^{19}F containing amino acids into 3-phosphoglycerate kinase in yeast cells (24). ^{19}F NMR has also been used with success in *E. coli*, and its greater sensitivity allows detection of globular proteins in these cells (52).

In this dissertation, in-cell NMR is utilized to study proteins inside of live cells. In Chapter 2, the fold of an unstable globular protein is assessed in *E. coli* cells. The standard unfolding free energy of this unstable globular protein is determined in dilute solution. These two pieces of information are used to answer the question whether the cytoplasm can provide the free energy required to fold the protein. Chapter 3 describes the development of an insect expression system optimized for in-cell NMR. Quantification of recombinant protein yields, optimization of cell viability, formulation of insect media for isotopic labeling of recombinant proteins, and NMR results are presented.

Chapter 2

Macromolecular Crowding Fails to Fold a Globular Protein in Bacteria

2.1 Introduction

Until recently, the stabilization afforded by volume exclusion was thought to dominate both *in vitro* and in cells, although there were hints of compensation (15, 16, 32). Wang et al. have shown that non-specific, non-covalent interactions between crowder proteins and a test protein can compete with excluded volume effects to affect diffusion (48). Here, this idea is tested in terms of protein stability in crowded *in vitro* solutions and in *Escherichia coli*.

The Immunoglobulin G binding domain of protein L from *Streptococcus magnus* was chosen as the test protein. Ever since its introduction (53), Protein L (ProtL) has been used as a model for globular protein folding (54). This well studied 7-kDa protein exhibits reversible unfolding at 25 °C in dilute solution via a two-state reaction with a stability of 4.3 kcal/mol (55). Changing seven of its lysine residues to glutamic acids lowers the stability to -0.4 kcal/mol, such that the majority of the protein molecules are in the denatured state in dilute buffer. The destabilization arises from the variant's decreased hydrophobic surface area, not the increase in negative charge (55). This Kx7E variant, however, folds on adding sodium (55) or potassium (see 2.3.6). If excluded volume effects dominate the crowding effect, then the fraction of the variant molecules in the

native state should increase in cells relative to dilute solution. This idea was tested by using NMR.

2.2 Materials and Methods

2.2.1 Minimal Media Formulations

^{15}N-Enriched Minimal Media

One L of ^{15}N-enriched minimal medium is prepared by adding 100 mL of 10x salts (60 g/L Na_2HPO_4, 30 g/L KH_2PO_4, 5 g/L NaCl, and 7 g/L $^{15}NH_4Cl$ [Cambridge Isotope Laboratories, Inc.; Andover, MA], pH 7.0), 1 mL of 1 mg/mL thiamine HCl, 2 mL of 1 M $MgS0_4$, 10 mL of a 20% (v/v) glucose solution, and 1 mL of 100 mg/mL ampicillin to deionized water and diluting to 1 L. The ampicillin solution was sterile filtered with 0.22 µm polyvinylidene difluoride (PVDF) filters (EMD Millipore; Billerica, MA). All other solutions were sterilized by autoclaving for 20 min in a Sterilmatic Sterilizer (Market Forge Industries, Inc.; Everett, MA) set to the liquid cycle. An unenriched form of this minimal medium was prepared as described above but with the addition of NH_4Cl instead of $^{15}NH_4Cl$. Unless otherwise noted, all chemicals are from Thermo Fisher Scientific, Inc. (Fair Lawn, NJ).

^{19}F-Labeled Minimal Media

One L of ^{19}F-labeled minimal media is prepared as described for ^{15}N-enriched minimal media with the addition of 0.5 g *N*-(phosphonomethyl)-glycine, 60 mg L-phenylalanine, 60 mg L-tryptophan, 70 mg *m*-fluoro-D,L-tyrosine (3-FY), and 10 mg biotin (Thermo Fisher Scientific, Inc.; Fair Lawn, NJ). Unless otherwise noted, the above components are from Sigma-Aldrich Co. (St. Louis,

MO). Unlabeled tyrosine minimal media were prepared with an unenriched 10x salts solution and L-tyrosine instead of 3-FY. The osmolality is 0.16 Osm as determined with a Vapro 5520 Vapor Pressure Osmometer (Wescor, Inc.; Logan, UT). Hyperosmotic media (56) were formulated by adding NaCl to a final concentration of 0.5 M in ^{19}F-labeled minimal media. The pH value was readjusted to 7.0 with 8 M NaOH. The osmolality is 1.05 Osm as determined with the vapor pressure osmometer.

2.2.2 ProtL Production

pET-15b ProtL Vector Transformations

Four pET-15b vectors (EMD Biosciences, Inc.; San Diego, CA [Fig. 2.1]) containing the wild-type ProtL gene or the ProtL variants Kx6Q, Kx6E, or Kx7E were provided by Xavier Tadeo of the Millet lab in the Structural Biology Unit of the Centro de Investigación Cooperativa bioGUNE, 48160 Derio, Spain. The vectors were transformed into BL21(DE3) *Escherichia coli* cells (Agilent Technologies, Inc.; Cedar Creek, TX), streaked onto Luria broth (LB; 1% [w/v] Bacto Tryptone [BD; Franklin Lakes, NJ], 0.5% [w/v] yeast extract [Thermo Fisher Scientific, Inc.; Fair Lawn, NJ], and 1% [w/v] NaCl) plates containing 100 µg/mL ampicillin (LB_{AMP}), and incubated overnight at 37 °C.

^{15}N Enrichment and ^{19}F Labeling of ProtL

Single colonies were used to inoculate 100-mL liquid cultures of LB_{AMP}. The cultures were incubated for 12-16 h at 37 °C with 225-rotations per minute (rpm) shaking. The cultures were centrifuged in a Sorvall RC-5B centrifuge with a Sorvall GSA rotor (Thermo Fisher Scientific, Inc.; Waltham, MA) at 4,000 rpm for

30 min at 25 °C. The supernatants were decanted, and the pellets resuspended in 1 L of ^{15}N-enriched minimal media, 1 L of ^{19}F-labeled minimal media, or 1 L of hyperosmotic media. The cultures were incubated at 37 ° C with 225-rpm shaking. The inducer, isopropyl-β-D-1-thiogalactopyranoside (IPTG, 1 mM final concentration; Thermo Fisher Scientific, Inc.; Fair Lawn, NJ), was added when the cultures reached an optical density at 600 nm (OD_{600}) of 0.6 in ^{15}N-enriched minimal media or 0.8 in ^{19}F-labeled minimal media and hyperosmotic media. The induced cultures were incubated at 37 °C with 225-rpm shaking for 4 h in ^{15}N-enriched minimal media or 20-24 h in ^{19}F-labeled minimal media and hyperosmotic media. As a control, *E. coli* without an expression vector were grown to an OD_{600} of 0.6 in ^{15}N-enriched minimal media and then incubated for an additional 4 h.

Cultures were prepared for in-cell NMR experiments or lysed for protein purification. Wild-type ProtL cultures were induced for 1 h in ^{15}N-enriched media for in-cell NMR experiments.

2.2.3 Osmotic Shock

The location of wild-type ProtL, Kx6Q, Kx6E, and Kx7E in *E. coli* cells was determined by adapting an osmotic shock protocol (57). Cultures producing wild-type ProtL or Kx7E were induced in the unenriched form of the minimal media used for ^{15}N enrichment. Aliquots (1 mL) were collected 0-4 h post induction for cultures producing wild-type ProtL or Kx7E. Aliquots (1 mL) were collected 4 h or 7 h post induction for cultures producing Kx6Q or Kx6E, respectively. The cultures were centrifuged in a Centrifuge 5418 benchtop centrifuge (Eppendorf

North America; Hauppauge, NY) at 8,000 x g for 10 min. The pellets were resuspended in 100 µL of an ice cold 40% (v/v) sucrose, 30 mM Tris-HCl solution (pH 7.2) containing 2 mM ethylenediaminetetraacetic acid (EDTA). The resuspended pellets were incubated at room temperature for 10 min and then centrifuged in the benchtop centrifuge at 10,000 x g for 10-20 min. The pellets were resuspended in 90 µL of ice cold water and 10 µL of 20 mM $MgCl_2$. The resuspended pellets were incubated on ice for 3 min and then centrifuged in the benchtop centrifuge at 12,000 x g for 10-20 min. The supernatants and pellets contain periplasmic and cytoplasmic proteins, respectively.

Protein loading dye, 5x, (300 mM Tris-HCl, pH 6.8, 50% [v/v] glycerol, 10% [w/v] sodium dodecyl sulfate [SDS], 0.5% [w/v] bromophenol blue [Sigma-Aldrich Co.]) was added to the the supernatants and pellets (resuspended in water) to a final 1x concentration. The samples (20 µL each) were loaded onto 18% Tris-HCl SDS polyacrylamide gels (Bio-Rad Laboratories, Inc.; Hercules, CA) and electrophoresed at 200 V for 50-60 min. The gels were stained with Brilliant Blue R-250 (Thermo Fisher Scientific, Inc.; Pittsburgh, PA) and destained in aqueous solutions containing 40% (v/v) methanol and 10% (v/v) acetic acid followed by final destaining in a 10% methanol, 10% acetic acid solution. Gel images were acquired with a VersaDoc 4000 MP Imaging System (Bio-Rad Laboratories, Inc.).

2.2.4 ProtL Purification

Induced cultures producing wild-type ProtL, Kx6Q, Kx6E, or Kx7E were centrifuged in a Sorvall RC-3B centrifuge with a Sorvall H6000A rotor (Thermo

Fisher Scientific, Inc.; Waltham, MA) at 4,000 rpm for 30 min at 4 °C. The pellets were resuspended in 30 mL of 2 M NaCl, 20 mM phosphate buffer (20 mM sodium phosphate, pH 6.0) containing 1 mM phenylmethanesulphonylfluoride (PMSF). The cell slurries were sonicated on ice with a Sonic Dismembrator 500 (Thermo Fisher Scientific, Inc.; Pittsburgh, PA) at a 14% amplitude and a 2 s on, 2 s off duty cycle for a total of 10 min. The lysates were centrifuged in the Sorvall RC-5B centrifuge with a Sorvall SS-34 rotor (Thermo Fisher Scientific, Inc.; Waltham, MA) at 14,000 rpm for 30 min at 4 °C. Lysa tes containing Kx7E were heated at 100 °C for 10 min followed by centrifugat ion in the Sorvall RC-5B centrifuge with the Sorvall SS-34 rotor at 14,000 rpm for 30 min at 4 °C. The resultant supernatants were concentrated to 5 mL by using a 3,000 molecular weight cutoff (MWCO) centrifugal filter unit (EMD Millipore). The lysates were loaded onto a Superdex 75 size exclusion column (GE Healthcare; Piscataway, NJ), which was equilibrated in and eluted with 2 M NaCl, 20 mM phosphate buffer, pH 6.0. Fractions were loaded onto 18% Tris-HCl SDS polyacrylamide gels and electrophoresed at 200 V for 50-60 min. The gels were stained, destained, and imaged as described in 2.2.3.

Fractions containing ProtL were pooled and dialyzed against water overnight at 4 °C by using Pierce 3,500 MWCO SnakeSki n Pleated Dialysis Tubing (Thermo Fisher Scientific, Inc.; Rockford, IL). The water was exchanged twice, and the samples were dialyzed >3 h after each exchange. The proteins were lyophilized with a FreeZone 1 Liter Benchtop Freeze Dry System (Labconco Corp.; Kansas City, MO) and stored at -80 °C.

2.2.5 NMR

^{1}H -^{15}N heteronuclear single-quantum coherence (HSQC) spectra were acquired at 25 °C on a Varian Inova 600 MHz spectrome ter (Palo Alto, CA) equipped with a triple resonance HCN probe. The ^{1}H dimension sweep width was 8401 Hz and comprised 1024 complex points. The ^{15}N dimension sweep width was 2200 Hz and comprised 64 complex points. The data were processed with NMRPipe (58) and NMRView (59).

^{19}F spectra were acquired at 37 °C on the Varian Inova 600 MHz spectrometer equipped with a 5 mm ^{19}F $\{^{1}H\}$ z-gradient probe. The spectra comprised 128 to 2048 transients with a 30 kHz sweep width and a 2 s acquisition delay. Trifluoroethanol was used as a reference at 0 ppm.

2.2.6 Dilute Solution NMR NaCl Titrations

^{15}N-Enriched Kx7E NaCl Titration

Buffer solutions containing purified ^{15}N-enriched Kx7E in 20 mM phosphate buffer with 0 or 1 M NaCl (pH 6.0) were combined to achieve Kx7E solutions containing 0 to 1 M NaCl. The Kx7E concentration and pH values were the same in all samples. Samples (450 µL each) were pipetted into S500 NMR tubes (Norell, Inc.; Landisville, NJ), and HSQC (60) spectra were acquired as described in 2.2.5 after adding 10% (v/v) D_20.

A titration curve was produced (Fig. 2.2) by plotting the fraction of folded protein versus NaCl concentration (Table 2.1). The data were fit to a two-state model of the form $fraction\ folded = \frac{(Nx+n)+(Dx+d)e^{-m(x-C)/RT}}{1+e^{-m(x-C)/RT}}$ where x is the concentration of NaCl, RT is the gas constant times the absolute temperature,

$m = \frac{\partial \Delta G^{\circ\prime}}{\partial x}$ where $\Delta G^{\circ\prime}$ is the standard unfolding free energy, $C = -\frac{\Delta {G_0}^{\circ\prime}}{m}$ where $\Delta {G_0}^{\circ\prime}$ is $\Delta G^{\circ\prime}$ in the absence of NaCl, n and d are the linear folded and unfolded baselines defined by $n = n^{\circ} + Nx$ and $d = d^{\circ} + Dx$, respectively, where n° and d° are the intercepts, and N and D are the slopes of the folded region and the unfolded region, respectively (61).

^{19}F-Labeled Kx7E NaCl Titration

A 4 M NaCl, 20 mM phosphate buffer, pH 6.0, solution was added to purified ^{19}F-labeled Kx7E in 20 mM phosphate buffer, pH 6.0, in increasing amounts to achieve NaCl concentrations of 0.1, 0.2, 0.3, 0.5, and 1 M. ^{19}F spectra of ^{19}F-labeled Kx7E were acquired after adding 10% (v/v) D_20.

2.2.7 *In Vitro* NMR with Synthetic Polymers, Proteins, and Osmolytes

Purified ^{19}F-labeled Kx7E was dissolved in 20 mM phosphate buffer, pH 6.0 to a final concentration of 5 mM. Purified ^{19}F-labeled Kx7E was dissolved to a final concentration of 2 mM in 20 mM phosphate buffer, pH 6.0, solutions containing the following components at final concentrations of 300 g/L: 40-kDa polyvinylpyrrolidone (PVP; Sigma-Aldrich Co.), bovine serum albumin (BSA; Thermo Fisher Scientific, Inc.; Fair Lawn, NJ), 8-kDa polyethylene glycol (PEG; Sigma-Aldrich Co.), 70-kDa Ficoll (Sigma-Aldrich Co.), D-glucose, or dextran (Sigma-Aldrich Co.). Purified ^{19}F-labeled Kx7E was dissolved to a final concentration of 2 mM in 20 mM phosphate buffer, pH 6.0, solutions containing 585 g/L sucrose or 72% (v/v) glycerol. Purified ^{19}F-labeled Kx7E was dissolved to a final concentration of 3 mM in a 20 mM phosphate buffer, pH 6.0, solution containing 600 g/L betaine (Sigma-Aldrich Co.). A final concentration of 10% (v/v)

D_2O and 1 mM PMSF was added to each sample. The samples were pipetted into S500 NMR tubes, and ^{19}F spectra were acquired as described in 2.2.5. Glucose solutions at final concentrations of 370, 440, and 610 g/L were prepared by successively adding more glucose to the 300 g/L glucose sample after NMR acquisitions. Unless otherwise noted, all chemicals are from Thermo Fisher Scientific, Inc. (Fair Lawn, NJ).

2.2.8 In-Cell NMR

Inducing production in ^{15}N-enriched minimal media for 4 h results in leakage of wild-type ProtL as determined by detection of its spectrum in the cell supernatant. Reducing the time to 1 h prevented leakage, in agreement with previous results (46). Induction in ^{19}F-labeled minimal media does not result in leakage because production levels are depressed in this medium.

E. coli producing wild-type ProtL, Kx6Q, Kx6E, or Kx7E in ^{15}N-enriched minimal media, *E. coli* producing wild-type ProtL or Kx7E in ^{19}F-labeled minimal media, *E. coli* producing Kx7E in hyperosmotic media, and *E. coli* without an expression vector in ^{15}N-enriched minimal media were pelleted by centrifugation in the RC-3B centrifuge with the H6000A rotor at 4,000 rpm for 30 min at 25 °C. Pellets of *E. coli* producing Kx7E in hyperosmotic media were resuspended in 1 mL of hyperosmotic media per liter of cell culture. All other pellets were resuspended in 1 mL of LB media per liter of cell culture.

D_2O was added to the cell slurries to a final concentration of 10% (v/v), and the samples were pipetted into S500 NMR tubes. HSQC and ^{19}F spectra were acquired as described in 2.2.5. After acquisition, the slurries were removed

from the NMR tubes and centrifuged in the benchtop centrifuge at 8,000 x g for 10 min. Spectra of the supernatants were acquired. The pellets of *E. coli* induced in ^{15}N-enriched minimal media or ^{19}F-labeled media were washed twice with 20 mM phosphate buffer, pH 6.0, and each pellet was resuspended in 20 mM phosphate buffer, pH 6.0, containing PMSF (1 mM final concentration). The cell slurries were sonicated on ice with the Sonic Dismembrator 500 at a 14% amplitude and a 2 s on, 2 s off duty cycle for a total of 10 min. The lysates were centrifuged in the benchtop centrifuge at 17,000 x g for 10 min. A final concentration of 10% (v/v) D_2O was added to the cleared lysates for HSQC and ^{19}F NMR acquisition as described in 2.2.5. The ^{19}F spectrum of the Kx7E cell lysate was acquired by adding guanidine chloride (3 M final concentration) to the insoluble lysate fraction. NaCl or KCl (1 M final concentration) was added to the cleared lysates of cells induced in ^{15}N-enriched minimal media, and the HSQC spectra were acquired.

E. coli *Producing ^{19}F-Labeled Kx7E in Hyperosmotic Media*

E. coli cultures producing Kx7E were induced with IPTG for 26 h in hyperosmotic media (formulation described in 2.2.1). After induction, the cultures were prepared for in-cell ^{19}F NMR as described above. Only the in-cell spectrum was acquired.

2.2.9 3-FY ^{19}F ProtL Assignments

ProtL pET-15b Mutagenesis

The wild-type ProtL gene from the pET-15b vector was mutated by using a site-directed mutagenesis kit (Agilent Technologies, Inc.; Santa Clara, CA) with

primers designed to introduce random amino acids at tyrosine positions 34, 36, and 56. The primers are shown in Table 2.2. The mutagenesis reactions were transformed into XL-1 Blue *E. coli* cells (Agilent Technologies, Inc.; Santa Clara, CA). The cells were plated onto LB_{AMP} plates and incubated for >16 h at 37 °C. Colonies were picked from the plates, inoculated into 5-mL liquid cultures of LB_{AMP}, and grown for 12-16 h at 37 °C with 250-rpm shaking. ProtL pET-15b plasmid DNA was isolated by using a miniprep kit (Qiagen, Inc.; Valencia, CA) and sequenced (University of North Carolina at Chapel Hill [UNC-CH] Genome Analysis Facility) to identify constructs with codons encoding amino acids other than tyrosine at positions 34, 36, and 56.

ProtL Variant Production in LB Media

Positive clones were transformed into BL21(DE3) *E. coli* cells, streaked onto LB_{AMP} plates, and incubated overnight at 37 °C. Colonies were picked from the plates and used to inoculate 5 mL of liquid LB_{AMP}. The cultures were incubated at 37 °C with 250-rpm shaking. IPTG (1 mM final concentration) was added when the cultures reached an OD_{600} of 0.6. The induced cultures were incubated for 4 h at 37 °C with 250-rpm shaking. One-mL aliquots of the cultures were centrifuged in the benchtop centrifuge at 8,000 x g for 10 min. The pellets were resuspended in 80 µL of water and 20 µL of 5x protein loading dye and heated for 10 min at 100 °C. Cell lysates (15 µL each) were loaded onto 10-20% Tris-HCl SDS polyacrylamide gels (Bio-Rad Laboratories, Inc.) and electrophoresed at 200 V for 65 min. The gels were stained, destained, and imaged as described in 2.2.3.

ProtL Variant Production in Unlabeled Tyrosine Minimal Media

Colonies were picked from the same plates described above and inoculated into 5-mL of liquid LB_{AMP}. The cultures were incubated at 37 °C with 250-rpm shaking to an OD_{600} of 0.6. The cultures were centrifuged at full speed in an IEC HN-SII centrifuge (Thermo Fisher Scientific, Inc.; Pittsburgh, PA) for 10 min at room temperature. The supernatants were decanted and the cell pellets resuspended each in 5 mL of unlabeled tyrosine minimal media. The cultures were incubated at 37 °C with 250-rpm shaking. IPTG (1 mM final concentration) was added when the cultures reached an OD_{600} of 0.8. The induced cultures were incubated for 20-24 h at 37 °C with 250-rpm shaking. Cells were harvested, lysed, and visualized for protein production as described above.

ProtL Variant Production in ^{19}F-Labeled Minimal Media for ^{19}F NMR Assignments

Colonies of BL21(DE3) *E. coli* cells producing the ProtL Y34A, Y36H, or Y56N variants were picked from LB_{AMP} plates and used to inoculate 100 mL of liquid LB_{AMP}. The cultures were incubated overnight at 37 °C with 250-rpm shaking. Each culture (25 mL) was centrifuged at full speed in the IEC HN-SII centrifuge for 10 min at room temperature to pellet the cells. Each pellet was resuspended in 250 mL of ^{19}F-labeled minimal media. The cultures were incubated at 37 °C with 250-rpm shaking. IPTG (1 mM final concentration) was added when the cultures reached an OD_{600} of 0.8. The induced cultures were incubated for 24 h at 37 °C with 250-rpm shaking. The cells were pelleted by centrifugation in the RC-3B centrifuge with the H6000A rotor at 4,000 rpm for 10 min at 25 °C. The supernatants were decanted, and the pellets were washed

twice with LB media. Each pellet was resuspended in 250 mL of LB media containing PMSF (1 mM final concentration). The cell slurries were sonicated on ice with the Sonic Dismembrator 500 at a 14% amplitude and a 2 s on, 2 s off duty cycle for a total of 10 min. The lysates were centrifuged in the benchtop centrifuge at 17,000 x g for 10 min. A final concentration of 10% (v/v) D_2O was added to the cleared lysates. The cleared lysates were pipetted into S500 NMR tubes, and ^{19}F spectra were acquired as described in 2.2.5.

Assignment Rationale and Results

Unfortunately, there is no current way to relate the chemical shifts from fluorine-containing amino acids to the protein environment as acquired by X-ray crystallography or NMR spectroscopy. Therefore, assignments are achieved by using site directed mutagenesis (62). Wild-type ProtL contains tyrosines at positions 34, 36, and 56. Creating three variants of the protein, where the tyrosine residue at each position is substituted for a different amino acid, allows assignment of the ^{19}F resonances. Each variant gives rise to a ^{19}F spectrum with just two resonances from the protein.

When making mutations, common sense suggests that the amino acid size and net charge should be perturbed as little as possible so that the protein structure is not significantly altered. Thus, mutagenesis was performed on the pET-15b vector to mutate the codons encoding the three tyrosines in wild-type ProtL to encode phenylalanines, which are similar in size and have the same net neutral charge. BL21 *E. coli* cells producing these variants, however, were unable to grow in ^{19}F-labeled minimal media. Inexplicitly, the variant caused the

apparent cellular toxicity because BL21 *E. coli* cells producing wild-type ProtL have no difficulty growing in ^{19}F-labeled minimal media.

A second strategy was developed to perform the assignments where no bias in the amino acid substitution was introduced. Primers were designed (Table 2.2) with random nucleotide substitutions at the codons encoding the wild-type ProtL tyrosines at positions 34, 36, and 56. By using these random primers, multiple pET-15b mutants are generated. Transformation of these vectors into *E. coli* allows screening for the cultures that can grow in ^{19}F-labeled minimal media and produce the variants. Additionally, a basic local alignment search tool (BLAST) was used to find naturally occurring ProtL variants at postions 34, 36, or 56. The rationale is that naturally occurring ProtL variants are less likely to be toxic to *E. coli* than genetically modified variants utilizing random amino acid substitutions. The BLAST search yielded a ProtL variant from the bacterium *Finegoldia magna* where tyrosine 56 is changed to asparagine (63). Therefore, primers were also designed to mutate the codon encoding amino acid position 56 to asparagine (Table 2.2). DNA sequencing confirmed the generation of pET-15b mutants encoding the ProtL variants Y34A, Y34P, Y34L, Y36P, Y36H, Y36Q, Y56Q, Y56S, Y56V, and Y56N. Results from SDS-PAGE (Fig. 2.3 and Fig. 2.4) show overproduction of the ProtL variants, with the exception of the Y34P variant. The production levels of the variants are similar in LB media and unlabeled tyrosine minimal media (Fig. 2.3 and Fig. 2.4) and in ^{19}F-labeled minimal media (Fig. 2.3*b*). Importantly, the production levels of the variants in all tested media are comparable to the production level of wild-type ProtL in LB

media. The ^{19}F lysate spectrum is easily acquired for *E. coli* producing wild-type ProtL (Fig. 2.5*e*) Thus, any of these ProtL variants, excluding the Y34P variant, should provide suitable production levels in ^{19}F-labeled minimal media.

The ^{19}F spectra of the cell lysates producing the Y34A, Y36H, or Y56N variants in ^{19}F-labeled minimal media (Fig. 2.5*b-d*) each display two protein resonances and one resonance from unincorporated 3-FY. Taken together, the spectra allow assignment of the three tyrosines in wild-type ProtL (Fig. 2.5*a*). Interestingly, while the protein resonances from the Y56N variant are strong and comparable to those from the wild-type ProtL cell lysate spectrum (Fig. 2.5*e*), the protein resonances from the Y34A and Y36H variants are weaker (Fig. 2.5*b, c*). The production levels of these two variants in ^{19}F-labeled minimal media (Fig. 2.3*b*) are the same as the wild-type protein in LB media (Fig. 2.3*a*). The reduction in signal for the Y34A and Y36H variants might be caused by aggregation or by binding of the recombinant proteins to cellular components. This hypothesis, while not tested here, could be examined by adding denaturants such as guanidine chloride to the samples and checking for recovery of strong protein resonances in the ^{19}F spectra.

2.3 Results

2.3.1 Wild-type ProtL, Kx6Q, Kx6E, and Kx7E are Located in the Cytoplasm

Detection of wild-type ProtL and its variants in the cytoplasm or periplasm of *E. coli* was performed by osmotically shocking *E. coli* cells to separate the cytoplasm and periplasm (57). SDS polyacrylamide gel electrophoresis (SDS-PAGE) results (Fig. 2.6) show an increase in the production of wild-type ProtL

and the Kx7E variant over 4 h of induction in minimal media. Kx7E is located almost entirely in the cytoplasm through 4 h. The wild-type protein is primarily located in the cytoplasm from 0 to 1 h. Later, it is present in similar quantities in both compartments. At all time points the production level of the wild-type protein is greater than that of Kx7E. These observations are consistent with the idea that increasing production causes cytoplasmic proteins to migrate to the periplasm (46). For all studies of the wild-type protein, the production time was limited such that the protein is in the cytoplasm.

SDS-PAGE results (Fig. 2.7) show that Kx6Q and Kx6E are located primarily in the cytoplasm after 4 h and 7 h of induction in minimal media, respectively. Kx6E is also present in the periplasmic fraction, possibly because the culture was induced for 7 h rather than 4 h. The production level of these two variants appears higher than that of Kx7E, most likely due to the introduction of one fewer mutation.

2.3.2 Wild-Type ProtL is Folded in Dilute Solution

The fold was monitored by using the ^{1}H-^{15}N HSQC experiment. The spectrum shows that the protein is folded (Fig. 2.8*a*) as indicated by the large dispersion of ^{1}H chemical shifts. The fold was also assessed by using ^{19}F NMR. Wild-type ProtL contains three tyrosines, so the substitution of 3-FY for tyrosine (64) is expected to yield three resonances in the ^{19}F spectrum. This prediction is borne out (Fig. 2.9*a*). The resonances are well separated as expected for a folded protein.

2.3.3 Kx7E Requires Salt to Fold in Dilute Solution

Purified, ^{15}N-enriched Kx7E was dissolved in 20 mM phosphate buffer solutions, pH 6.0, containing 0 to 1 M NaCl. The limited ^{1}H chemical shift dispersion observed in the HSQC spectrum (Fig. 2.8*b*) indicates that the variant is unfolded in buffer alone (55, 65-67). Upon increasing the NaCl concentration to 0.3 M, the dispersion increases to that expected for a folded protein, although crosspeaks from the unfolded form are also visible (Fig. 2.8*c*). In 0.8 M NaCl, almost all of the variant is folded (Fig. 2.8*d*) and the crosspeak pattern is similar to that of the wild-type protein (Fig. 2.8*a*).

HSQC spectra were acquired from 0 to 1 M NaCl in 0.1 M intervals (Fig. 2.10), and the average intensities of the unfolded and folded Kx7E crosspeaks at each NaCl concentration (Table 2.1) were used to produce a titration curve (Fig. 2.2). The results show that the transition is 50% complete at 0.3 M NaCl and reaches a maximum at 0.8 M NaCl. Even at 0.8 M NaCl, a few crosspeaks from the unfolded protein remain (Fig. 2.10*i*). The data were fit to a two-state model that assumes linear native and denatured state baselines and a linear dependence of the standard unfolding free energy on salt concentration (61). The standard unfolding free energy at 0 M NaCl, $\Delta G_0{}^{\circ\prime}$, is -1.0 kcal/mol (RMSD of 2%). This value represents the free energy required to fold half the variant molecules in the absence of NaCl. The salt-induced folding is reversible (55).

The fold of the Kx7E variant was then monitored by using ^{19}F NMR. Purified ^{19}F-labeled Kx7E was dissolved in 20 mM phosphate buffer solutions, pH 6.0, containing NaCl at concentrations from 0 to 1 M. The spectrum acquired at

37 °C in the absence of added NaCl (Fig. 2.9 *b*) shows the three expected resonances in a narrow chemical shift range, indicative of an unfolded protein. As the NaCl concentration increases, the Y34 and Y36 resonances from the folded protein appear downfield, and the intensities of the resonances attributed to the unfolded protein decrease (Fig. 2.9*c*-*f*). At 1 M NaCl (Fig. 2.9*g*), most of the Kx7E variant is folded, although unfolded protein resonances remain at low intensities. The chemical shifts from the variant in 1 M NaCl are consistent with those of wild-type ProtL (Fig. 2.9*a*). The increasing fraction of folded protein with increasing NaCl concentration is consistent with the HSQC data (Fig. 2.10), although not enough data points were acquired for quantification.

2.3.4 Kx7E Does Not Fold Under *In Vitro* Crowded Conditions

The ability of synthetic polymers and proteins to fold the Kx7E variant was monitored by using ^{19}F NMR. Purified and lyophilized ^{19}F-labeled Kx7E was resuspended in 20 mM phosphate buffer, pH 6.0, solutions containing 300-g/L final concentrations of PEG, PVP, Ficoll, or BSA. The ^{19}F spectra in the synthetic polymers and BSA (Fig. 2.11*b*-*e*) remain unchanged with respect to Kx7E in buffer alone (Fig. 2.11*a*). The BSA sample displays some broadening (Fig. 2.11*e*), suggesting that the crowder chemically interacts with Kx7E (48). The results show that these macromolecular crowders are unable to fold Kx7E.

High concentrations of several sugars and osmolytes were also tested. The ^{19}F spectra of Kx7E in 72% (v/v) glycerol and 300 g/L dextran (Fig. 2.12*b, c*) are unchanged from that of the variant in buffer (Fig. 2.11*a*). The ^{19}F spectra of Kx7E in 585 g/L sucrose and 600 g/L betaine (Fig. 2.12*a*, *d*) show a single

downfield resonance that may be from the folded protein. The ^{19}F spectrum of Kx7E in 300 g/L glucose (Fig. 2.13), however, shows the downfield Y34 and Y36 resonances from the folded protein. The sample was removed from the NMR tube so that additional glucose could be added to acquire ^{19}F spectra at glucose concentrations of 370, 440, and 610 g/L (Fig. 2.13). Each successively higher glucose concentration causes an increase in the intensity of the Y34 and Y36 resonances from the folded protein and a corresponding decrease in the unfolded protein resonances. Glucose remains the only tested candidate, other than NaCl, to fold Kx7E. Next, in-cell NMR was used to determine whether the crowded conditions in the *E. coli* cytoplasm could fold the protein.

2.3.5 Wild-Type ProtL is Folded in *E. coli*

In-cell NMR was used to assess wild-type ProtL in *E. coli*. The HSQC spectrum of the cell slurry containing the ^{15}N-enriched wild-type protein is shown in Fig. 2.14*a*. The spectrum is the same as that of cells without an expression vector (Fig. 2.15); the spectrum of the protein is unobservable. The only crosspeaks in Fig. 2.14*a* are from ^{15}N-enriched metabolites (68). These observations are consistent with results showing that ^{1}H-^{15}N HSQC spectra from globular proteins are not detectable in *E. coli* (49, 52, 69, 70). After lysing the cells, the wild-type ProtL spectrum is detected (Fig. 2.14*c*) and is unchanged in the presence of 1 M NaCl (Fig. 2.14*d*). The spectrum of the cell supernatant shows no protein crosspeaks (Fig. 2.14*b*), confirming that the protein is in the cells. To detect wild-type ProtL in cells, ^{19}F NMR was utilized (70).

The in-cell experiments were repeated with ^{19}F-labeled wild-type ProtL. The high sensitivity of ^{19}F, the more favorable relaxation properties of ^{19}F in a mobile side chain, the low background of fluorine in biological molecules, and the limited number of labels per protein allow detection of globular proteins in *E. coli* (52). Three broad resonances from the protein corresponding to Y34, Y36, and Y56 (Fig. 2.5) were detected in the ^{19}F spectrum of the cell slurry (Fig. 2.16*a*). The fourth resonance, from unincorporated 3-FY, is sharper and overlaps the Y56 resonance. The protein resonances are well separated and similar to those from the *in vitro* spectrum of wild-type ProtL (Fig. 2.9*a*), indicating that the wild-type protein is folded in the cytoplasm of *E. coli*. The spectrum of the supernatant also shows a sharp resonance from unincorporated 3-FY (Fig. 2.16*b*). Three additional small resonances are detected. The positions of these resonances match those of wild-type ProtL in dilute solution (Fig. 2.9*a*), suggesting that a small amount of protein (less than 10%) leaks from the cells, but not enough to obscure the spectrum of the cell slurry. The cleared cell lysate (Fig. 2.16*c*) shows three sharp resonances corresponding to Y34, Y36, and Y56 from the protein, proving that the spectrum of the cell slurry (Fig. 2.16*a*) is from wild-type ProtL inside cells. These data demonstrate that wild-type ProtL remains folded in *E. coli* and can be detected by ^{19}F NMR.

2.3.6 Kx7E is Unfolded in *E. coli*

Disordered proteins are detectable in cells by HSQC NMR because of their internal motion (49, 52, 69). The HSQC spectrum of the Kx7E cell slurry exhibits a narrow chemical shift dispersion in the ^{1}H dimension (Fig. 2.14*e*),

characteristic of the unfolded protein (Fig. 2.8*b*). The ability to detect the HSQC spectrum of the variant in cells also shows that the protein is soluble in the *E. coli* cytoplasm. The spectrum of the cell supernatant shows no protein crosspeaks (Fig. 2.14*f*), indicating that the spectrum shown in Fig. 2.14*e* arises from the variant inside cells. The variant remains unfolded upon cell lysis (Fig. 2.14*g*) but folds upon increasing the NaCl (Fig. 2.14*h*) or KCl (Fig. 2.17*b*) concentration to 1 M. These results demonstrate that, although Kx7E is foldable, the *E. coli* cytoplasm is unable to fold the variant.

The HSQC results were confirmed by using ^{19}F NMR. The ^{19}F spectrum of the Kx7E cell slurry displays an envelope of broad resonances in a narrow chemical shift range and a sharp peak attributed to unincorporated 3-FY (Fig. 2.16*d*). The position of the broad envelope is inconsistent with the resonances from the folded wild-type protein (Fig. 2.9*a* and Fig. 2.16*a*) but consistent with the position of the resonances from the purified unfolded Kx7E variant (Fig. 2.9*b*). The absence of the downfield Y34 resonance from the folded variant (Fig. 2.9*g*) is especially evident. No resonances from the Kx7E variant are detected in the spectrum of the supernatant, indicating that the protein does not leak (Fig. 2.16*e*). Almost no protein signal is detected in the spectrum of the cleared lysate (Fig. 2.18*a*). To gain insight, the pellet retained from cell lysis was resuspended and used to acquire a ^{19}F spectrum. Broad resonances, most likely from Kx7E, were detected (Fig. 2.18*b*). Sharp resonances, however, were recovered after adding guanidine chloride (Fig. 2.16*f*). The spectrum of Kx7E cannot be observed in the cleared lysate because the protein binds the insoluble fraction.

These results corroborate the in-cell HSQC spectra, demonstrating that Kx7E remains unfolded in cells.

2.3.7 Kx6Q and Kx6E NMR Results in *E. coli*

In addition, the folds of the Kx6Q and Kx6E variants were assessed in *E. coli* by using HSQC NMR. The unfolding free energies for the Kx6Q and Kx6E variants are 4.00-4.15 and 0.50-0.60 kcal/mol, respectively, in dilute solution as determined by guanidine chloride and urea titration experiments with circular dichroism detection (55). Therefore, both variants are stable globular proteins. Both cell slurry spectra (Fig. 2.19*a*, *e*) resemble the spectrum of intracellular metabolites (Fig. 2.15). No crosspeaks from Kx6Q or Kx6E were detected in the spectra of the supernatants (Fig. 2.19*b*, *f*). Upon lysis, the Kx6Q spectrum is detected and is characteristic of a folded protein (Fig. 2.19*g*). No protein is detected in the Kx6E cell lysate spectrum (Fig. 2.19*c*), suggesting aggregation or lack of Kx6E production in *E. coli*. Adding 1 M NaCl has no effect on either spectrum (Fig. 2.19*d*, *h*), which is expected for the stable globular protein Kx6Q and in the absence of soluble Kx6E. To check the reproducibility of the Kx6E result, another in-cell NMR experiment was performed with *E. coli* producing Kx6E. The spectrum of the cell slurry resembles the spectrum of intracellular metabolites (Fig. 2.20*a*), but also contains broad crosspeaks that are widely dispersed along the ^{1}H dimension. Recovery of protein crosspeaks in the spectrum upon cell lysis confirms the presence of Kx6E (Fig. 2.20*b*), and the wide chemical shift dispersion along the ^{1}H dimension indicates that Kx6E is folded as expected for a stable globular protein.

2.3.8 Increasing the Salt Concentration in Cells Fails to Fold Kx7E

An attempt was made to fold the Kx7E variant in *E. coli* by increasing the intracellular concentration of K^+ by growth and induction of *E. coli* in hyperosmotic (1.05 Osm) media (56, 71, 72). The ^{19}F spectrum of the Kx7E cell slurry in hyperosmotic media (Fig. 2.16*g*), however, shows no resonances from the folded protein. The increase in linewidth and the upfield shift of the resonance envelope of the unfolded protein is probably caused by the intracellular increase in K^+ concentration.

2.4 Discussion

Many *in vitro* studies of macromolecular crowding focus on volume exclusion. In these experiments, the environment comprises only one or a few different types of crowders. Furthermore, it is usually assumed that the crowding molecules have little interaction with the test protein unless their van der Waals surfaces attempt to overlap, at which point the repulsive forces increase in a highly exponential manner. Recent work, however, demonstrates the importance of non-specific intermolecular interactions in both experiments and simulations (32, 48, 73, 74). Thus, one cannot ignore the effects of chemical interactions in studies of macromolecular crowding.

The unstable Kx7E variant of the globular protein L was examined. The variant folds in the presence of high concentrations of salt (55), prompting the question whether the excluded volume provided by the *E. coli* cytoplasm would have a similar effect.

In vitro HSQC spectra of the variant were acquired in increasing concentrations of NaCl. The spectra allowed construction of a folding titration curve. The standard unfolding free energy of Kx7E in the absence of NaCl, $\Delta G_0{}^{\circ\prime}$, is -1.0 kcal/mol. This NMR derived value compares favorably to those derived by Tadeo et al. (-0.4 to -0.5 kcal/mol) from guanidine chloride and urea titration experiments with circular dichroism detection (55). Therefore, the *E. coli* cytoplasm would have to deliver between 0.4 and 1.0 kcal/mol of stabilization to fold 50% of the protein molecules.

Attempts to fold Kx7E *in vitro* were made by addition of synthetic polymers, proteins, sugars, and osmolytes at high concentrations (≥300 g/L). None of the tested crowders, with the exception of glucose, were able to fold the variant. The volume exclusion provided by these crowders may not be able to fold Kx7E into the thermodynamically favored native globular state because the crowders bind weakly to Kx7E (48, 74, 75). The ability of the small molecule glucose to fold the variant points to an effect other than macromolecular crowding induced volume exclusion.

In-cell NMR data (Fig. 2.14 and Fig. 2.16) show that wild-type ProtL is folded in *E. coli* while the Kx7E variant remains unfolded. The *E. coli* cytoplasm does not provide sufficient free energy to fold Kx7E. More specifically, the volume exclusion provided by the highly crowded intracellular environment, which always favors the native state of globular proteins, is insufficient to overcome the unfavorable free energy of folding (-1.0 kcal/mol). Predictions from theory show that volume exclusion could provide 2.5 kcal/mol of stabilization for

a protein the size of ProtL (76). The fact that Kx7E is unfolded in the *E. coli* cytoplasm suggests the presence of a destabilizing effect.

Experiments were conducted to test whether increasing the K^+ concentration in cells could fold the Kx7E variant. Although NaCl was used in the *in vitro* experiments (Figs. 2.9-2.11), addition of 1 M KCl to *E. coli* lysates producing Kx7E demonstrates that KCl also folds the variant (Fig. 2.17*b*). The cytoplasmic K^+ concentration in *E. coli* is ~0.2 M (71, 72) for cells grown in minimal media (0.1 Osm). Nearly all of the cytoplasmic K^+ in *E. coli* grown in minimal media, however, is associated with polyanions such as DNA and RNA (56). Thus, not enough salt is available to fold the variant in *E. coli*.

The cytoplasmic concentration of K^+ was increased by increasing the osmolality of the growth medium from 0.16 Osm to 1.05 Osm with NaCl. The *E. coli* cell membrane is impermeable to NaCl, so the cells compensate for the increased external osmolality by increasing the amounts of K^+, glutamate$^-$, trehalose, proline, and glycine betaine in the cytoplasm (72). Under these conditions the cytoplasmic K^+ concentration reaches between 0.5 and 0.9 M (71, 72). Importantly, at least half of this upshift in K^+ is due to an increase in free cytoplasmic K^+ rather than K^+ that is locally accumulated with polyanions (71). Therefore, the concentration of free cytoplasmic K^+ under hyperosmotic conditions is at least 0.3 M, which leads to easily detectable levels of the folded protein *in vitro* (Fig. 2.2 and Fig. 2.9). Even under these conditions the folded form is not detected in cells (Fig. 2.16*g*). Furthermore, it is unlikely that the unfolded form is associated with chaperones because the observation of

crosspeaks in Fig. 2.14*e* is inconsistent with the molecular weight of a Kx7E-chaperone or proteasome complex (77).

Our results suggest that any stabilizing excluded volume effect, which must exist because of the high concentration of macromolecules in the cytoplasm, is more than offset by a destabilizing effect. The destabilization probably arises from non-specific protein-protein interactions. Wang et al. show that rotational diffusion of the globular protein chymotrypsin inhibitor 2 is slowed beyond that expected by viscosity alone in crowded protein solutions and *E. coli* lysates, and that the difference is due to non-specific interactions between the crowders and test protein (48). These non-specific intermolecular interactions in cells can lead to irreversible denaturation (78), explaining why the Kx7E variant is unfolded in cells. Such non-specific protein-protein interactions also explain why the variant does not fold completely *in vitro* even at high salt concentrations (Fig. 2.2, Fig. 2.8*d*, and Fig. 2.9*g*).

Although volume exclusion is expected to stabilize the native state of globular proteins, it must be borne in mind that the *E. coli* cytoplasm comprises biologically active molecules. These molecules chemically interact with the protein under study, and these interactions can be stabilizing or destabilizing. Therefore, it is not clear, a priori, how macromolecular crowding in cells, which includes both volume exclusion and chemical interactions, will affect protein stability (29). The few available in-cell studies of globular protein stability indicate no change or a slight destabilization in cells compared to dilute solution (15, 16). Although dilute solution data may be physiologically relevant for stable globular

proteins, the data presented here show that non-specific intermolecular interactions can overcome volume exclusion induced compaction, causing globular proteins to remain unfolded. The situation is different for intrinsically disordered proteins where the crowded environment in cells results in compaction (17). The dissimilar behavior is probably caused by the different amino acid compositions of these two protein classes (79).

Given that salt folds the variant *in vitro*, it is striking that increasing the intracellular salt concentration fails to have the same effect. This discrepancy indicates that non-specific intermolecular interactions may prevent refolding of globular proteins even when solution conditions favor it, leading to the suggestion that the rate of unfolding in cells deserves attention.

2.5 Conclusions

Wild-type ProtL and the Kx6Q, Kx6E, and Kx7E variants were purified from *E. coli* for use in *in vitro* and in-cell NMR experiments. Conditions were found such that all the proteins were located in the cytoplasm of *E. coli* after induction.

In vitro HSQC and ^{19}F NMR experiments show that Kx7E is unfolded in the absence of NaCl, 50% folded at 0.3 M NaCl, and reaches the maximum folded fraction at 0.8 M NaCl (90%). The standard unfolding free energy of Kx7E in the absence of NaCl, $\Delta G_0{}^{\circ\prime}$, is -1.0 kcal/mol. The cell must provide this amount of free energy to fold half the variant molecules. NMR spectra show that wild-type ProtL is folded in the absence of NaCl.

In-cell HSQC and ^{19}F NMR results are consistent with wild-type ProtL remaining folded in *E. coli*. In-cell HSQC and ^{19}F NMR results show that Kx7E is unfolded in *E. coli*. Therefore, the *E. coli* cytoplasm does not provide sufficient free energy (-1.0 kcal/mol) to fold Kx7E. The data support the conclusion that the volume exclusion provided by macromolecular crowding is unable to overcome the non-specific intermolecular interactions in *E. coli*, thus maintaining Kx7E in its unfolded form.

An attempt was made to fold Kx7E by increasing the intracellular concentration of K^+ in *E. coli*. *E. coli* producing Kx7E were grown in hyperosmotic media to increase the free cytoplasmic K^+ to at least 0.3 M, which in dilute solution leads to easily detectable levels of the folded protein. The spectrum of the cell slurry, however, shows a spectrum similar to that obtained in normal osmolality media, indicating that Kx7E remains unfolded. Thus, macromolecular crowding and increased K^+ concentrations are unable to overcome the non-specific intermolecular interactions present in the *E. coli* cytoplasm.

These results emphasize the importance of examining non-specific intermolecular interactions in cells. While the concentration of macromolecules in the *E. coli* cytoplasm can exceed 300 g/L, providing significant volume exclusion expected to favor the native state of globular proteins, these biological molecules are not inert. They include proteins, nucleic acids, osmolytes, and other molecules, and they chemically interact with the protein under study. These interactions can be stabilizing or destabilizing, making predictions of protein stability in cells difficult. The inability of the *E. coli* cytoplasm to fold Kx7E

suggests the presence of destabilizing non-specific intermolecular interactions that volume exclusion is unable to overcome. The fact that even increasing the free cytoplasmic K^+ to at least 0.3 M fails to fold Kx7E supports the idea that these non-specific intermolecular interactions prevent refolding even though the solution conditions favor it. Therefore, the study of protein unfolding rates in cells provides an interesting avenue for further investigation.

2.6 List of Tables

Table 2.1 Kx7E dilute solution NaCl titration

NaCl (M)	Number of Unfolded Crosspeaks	Number of Folded Crosspeaks	Ratio of Average Folded/Unfolded Volumes	Fraction Folded
0.00	67	0	0.00	0
0.10	55	53	0.12	0.10
0.15	54	63	0.20	0.19
0.20	57	68	0.41	0.33
0.25	57	69	0.53	0.39
0.30	58	66	0.88	0.50
0.40	56	67	1.36	0.62
0.50	39	105	0.98	0.73
0.60	19	107	0.99	0.85
0.70	22	106	1.26	0.86
0.80	7	97	0.65	0.90

Crosspeaks attributed to the folded and unfolded forms were counted in each HSQC spectrum. Backbone amide crosspeaks present in the 0 M NaCl spectrum were attributed to the unfolded protein. Crosspeaks that appeared upon adding NaCl were attributed to the folded protein. For each spectrum, the average volume of the crosspeaks attributed to the folded protein was divided by the average volume of the crosspeaks attributed to the unfolded protein. The fraction of the folded protein at each concentration was calculated by using the following equation:

$$\text{Fraction Folded} = \frac{(\text{Number}_{\text{folded}} * \text{Avg. Volume}_{\text{folded}})}{(\text{Number}_{\text{folded}} * \text{Avg. Volume}_{\text{folded}} + \text{Number}_{\text{unfolded}} * \text{Avg. Volume}_{\text{unfolded}})}$$

Table 2.2 ProtL pET-15b mixed primers used to introduce random amino acids at positions 34, 36, and 56 in wild-type ProtL

Name	Sequence
Forward Y34Any	5' GCAACTAGTGAAGCTNNNGCATATGCAGATACTTTGAAGAAAGAC 3'
Reverse Y34Any	5' GTCTTTCTTCAAAGTATCTGCATATGCNNNAGCTTCACTAGTTGC 3'
Forward Y36Any	5' GCAACTAGTGAAGCTTATGCANNNGCAGATACTTTGAAGAAAGAC 3'
Reverse Y36Any	5' GTCTTTCTTCAAAGTATCTGCNNNTGCATAAGCTTCACTAGTTGC 3'
Forward Y56Any	5' GTCGACGTTGCAGATAAAGGTNNNACTTTAAATATTAAATTTGCTGG 3'
Reverse Y56Any	5' CCAGCAAATTTAATATTTAAAGTNNNACCTTTATCTGCAACGTCGAC 3'
Forward Y56N	5' GTCGACGTTGCAGATAAAGGTAACACTTTAAATATTAAATTTGCTGG 3'
Reverse Y56N	5' CCAGCAAATTTAATATTTAAAGTGTTACCTTTATCTGCAACGTCGAC 3'

The altered codons are underlined

2.7 List of Figures

pET-15b sequence landmarks	
T7 promoter	463-479
T7 transcription start	452
His•Tag coding sequence	362-380
Multiple cloning sites	
(*Nde* I - *Bam*H I)	319-335
T7 terminator	213-259
lacI coding sequence	(866-1945)
pBR322 origin	3882
bla coding sequence	4643-5500

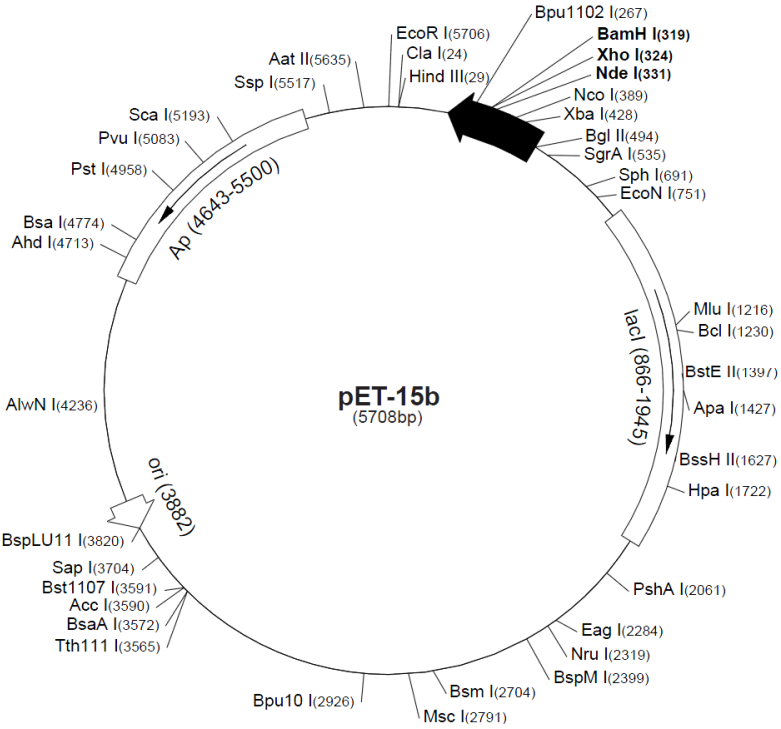


Figure 2.1 pET-15b vector map.

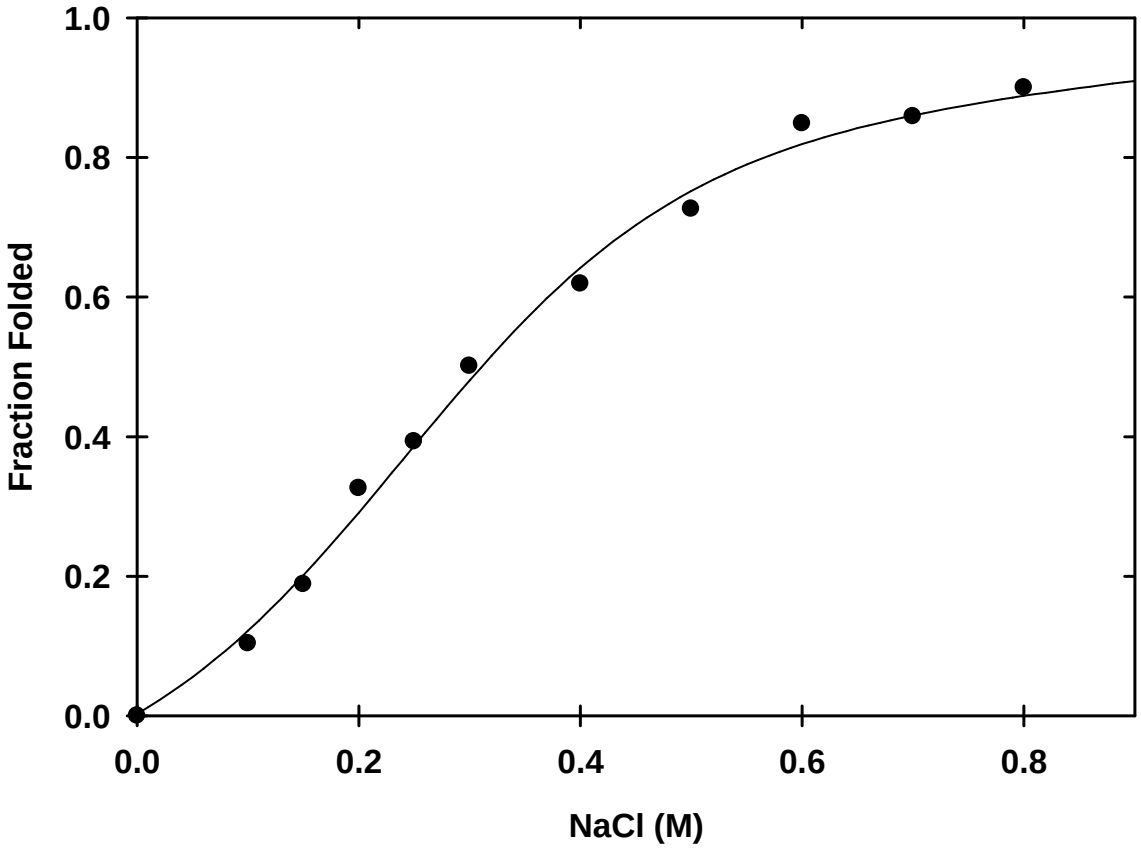


Figure 2.2 Titration curve showing the fraction of folded Kx7E (20 mM phosphate buffer, pH 6.0, 25 °C) as a function of NaCl concentr ation. The data were fit to a two-state model (61).

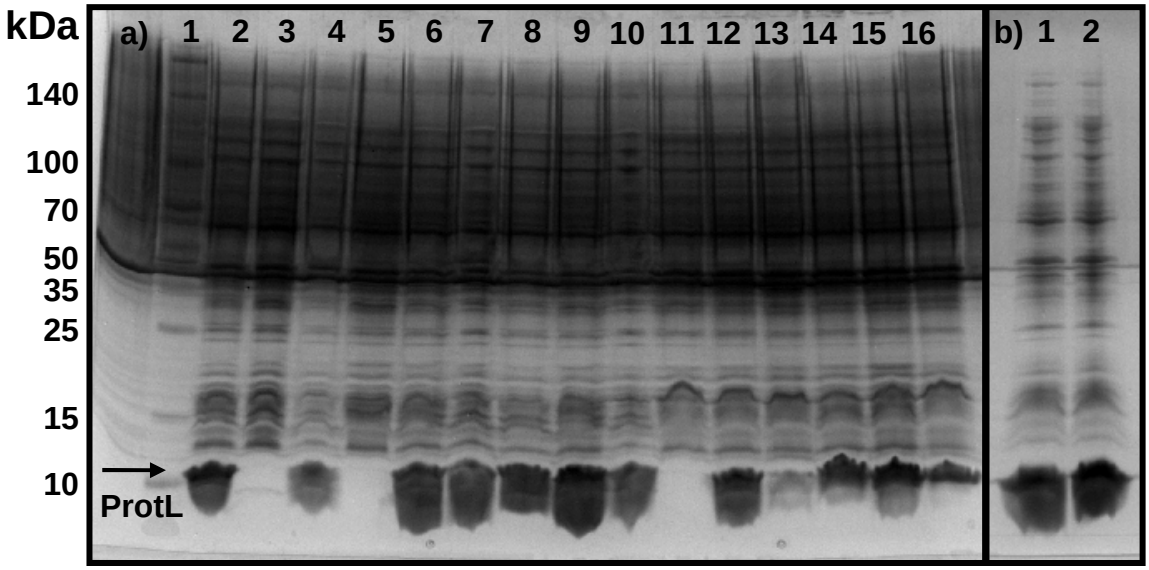


Figure 2.3 *E. coli* lysates producing ProtL Y34 and Y36 variants visualized by Brilliant Blue SDS-PAGE. a) Lysates of cells grown and induced in LB media and unlabeled tyrosine minimal media. Lanes 5-10 contain lysates of cells grown and induced in LB media to produce the variants Y34P (lane 5), Y34A (lane 6), Y34L (lane 7), Y36P (lane 8), Y36H (lane 9), and Y36Q (lane 10). Lanes 11-16 contain lysates of cells grown and induced in unlabeled tyrosine minimal media to produce the variants Y34P (lane 11), Y34A (lane 12), Y34L (lane 13), Y36P (lane 14), Y36H (lane 15), and Y36Q (lane 16). Lane 1 contains a protein molecular weight marker. Lanes 2 and 3 contain lysates of cells producing the wild-type ProtL and the Kx7E variant in LB media, respectively. Lane 4 contains the lysate of cells producing the Y56N variant in LB media. b) Lysates of cells grown and induced in ^{19}F-labeled minimal media to produce the variants Y34A (lane 1) and Y36H (lane 2).

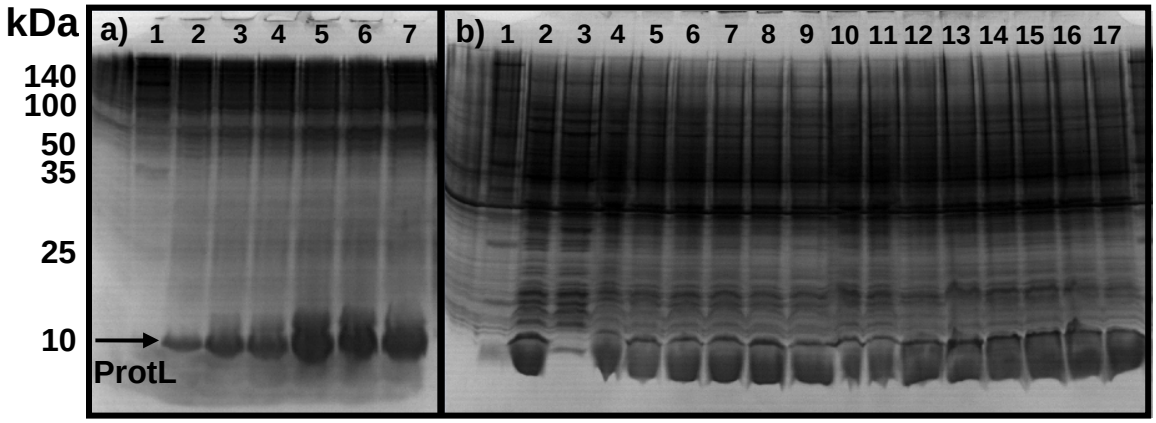


Figure 2.4 *E. coli* lysates producing ProtL Y56 variants visualized by Brilliant Blue SDS-PAGE. a) Lysates of cells grown and induced in LB media. Lane 1 contains a protein molecular weight marker followed by the variants Y56N (lanes 2-3), Y56Q (lane 4), Y56S (lane 5), and Y56V (lanes 6-7). b) Lysates of cells grown and induced in unlabeled tyrosine minimal media. Lane 1 contains a protein ladder followed by the variants Y56N (lanes 4-9), Y56Q (lanes 10-11, Y56S (lanes 12-13), and Y56V (lanes 14-17). Lanes 2 and 3 contain lysates of cells producing wild-type ProtL and the Kx7E variant in LB media, respectively.

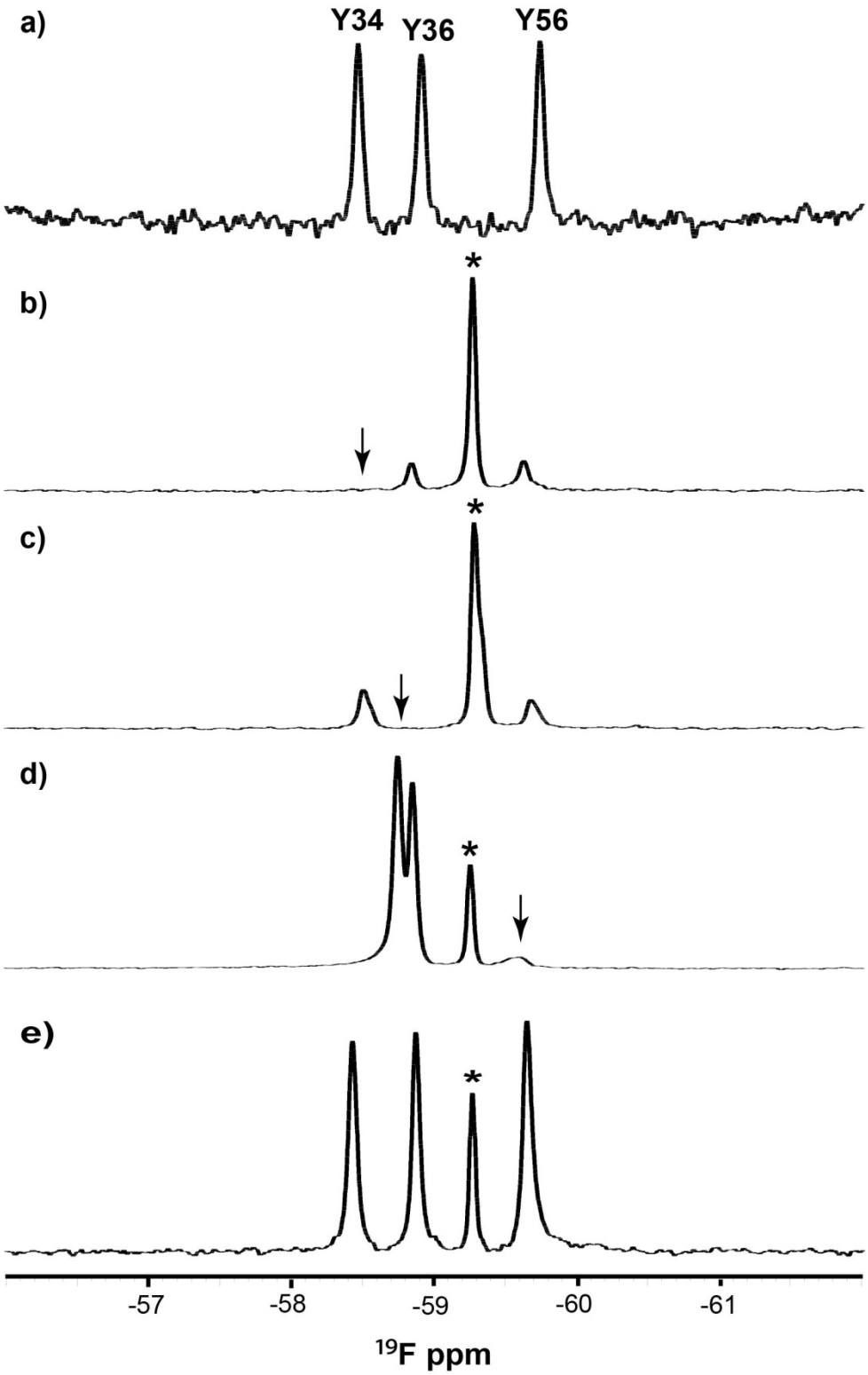


Figure 2.5 ^{19}F assignment spectra (37 °C) of ProtL. a) Purified wild- type ProtL (20 mM phosphate buffer, pH 6.0, 37 °C). Lysates of *E. coli* cells producing the variants b) Y34A, c) Y36H, and d) Y56N. e) Lysate of *E. coli* cells producing wild-type ProtL. Asterisks denote the resonance from free 3-FY.

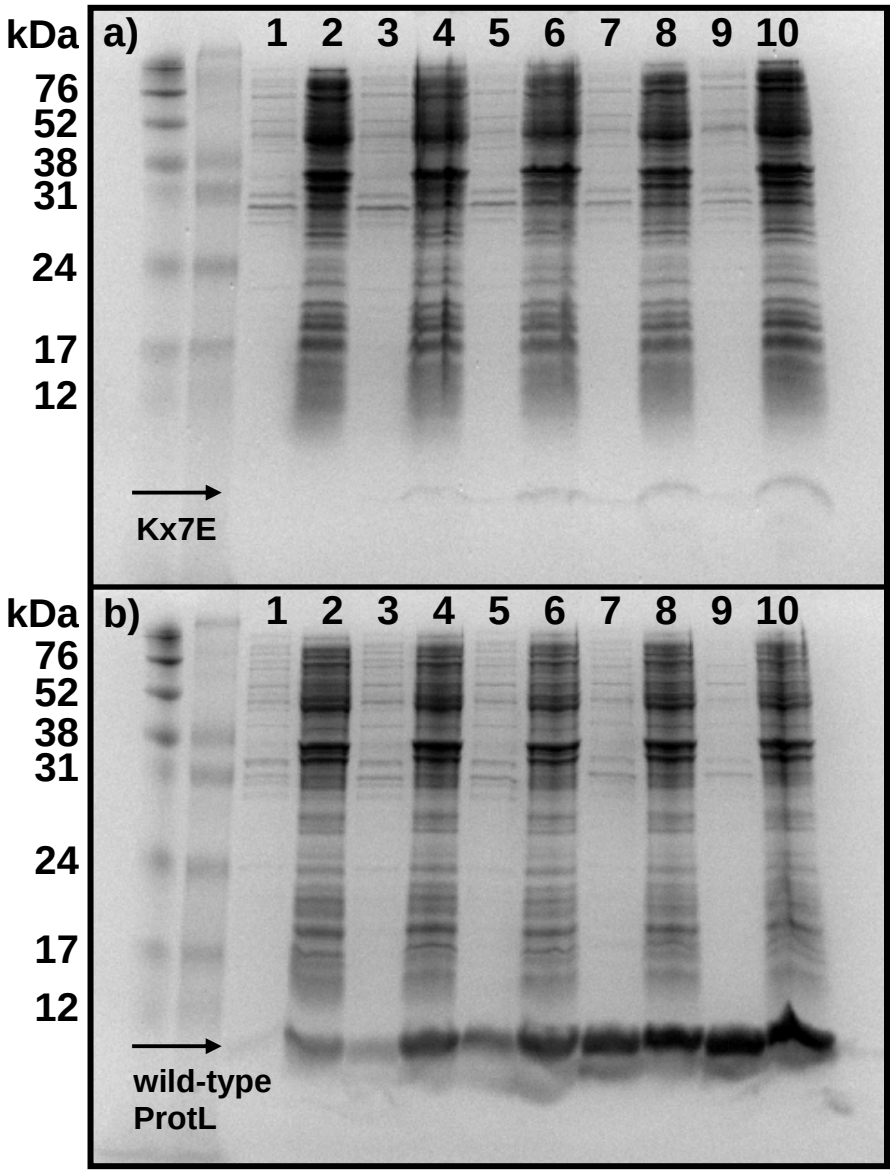


Figure 2.6 ProtL location determined by osmotic shock. Brilliant Blue SDS-PAGE results of periplasmic and cytoplasmic fractions of *E. coli* producing a) the Kx7E variant and b) the wild-type protein. Cells were harvested and lysed 0-4 h after induction in minimal media. Lanes 1, 3, 5, 7, and 9 correspond to the periplasm. Lanes 2, 4, 6, 8, and 10 correspond to the cytoplasm. Lanes 1-2, 3-4, 5-6, 7-8, and 9-10 correspond to induction times of 0 h, 0.5 h, 1 h, 2 h, and 4 h, respectively.

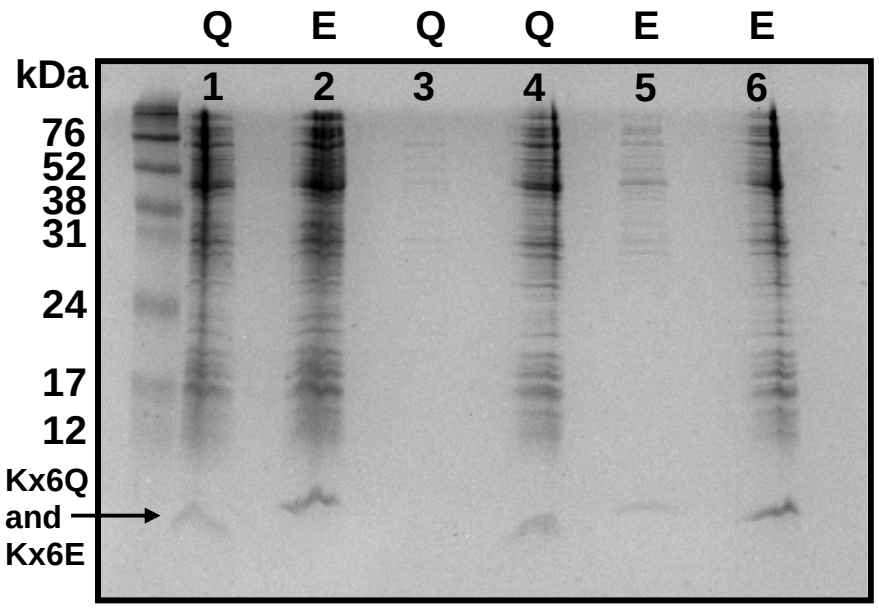


Figure 2.7 Brilliant Blue SDS-PAGE results of periplasmic and cytoplasmic fractions of osmotically shocked *E. coli* cells producing Kx6Q and Kx6E in minimal media. Lanes 1 and 2 contain Kx6Q and Kx6E whole cell lysates, respectively. Lanes 3 and 4 contain periplasmic and cytoplasmic fractions of Kx6Q, and lanes 5 and 6 contain periplasmic and cytoplasmic fractions of Kx6E, respectively. *E. coli* Kx6Q induction time was 4 h while *E. coli* Kx6E induction time was 7 h.
Q: ProtL Kx6Q
E: ProtL Kx6E

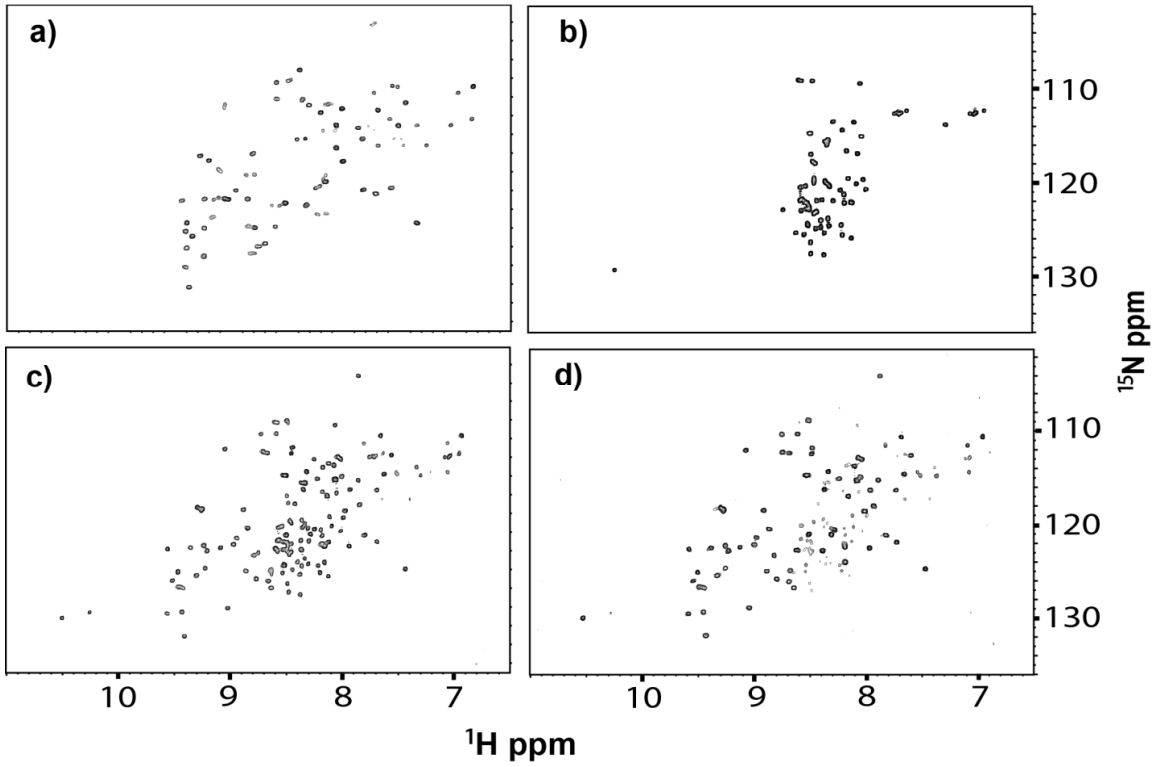


Figure 2.8 HSQC spectra (20 mM phosphate buffer, pH 6.0, 25 °C) of ^{15}N-enriched a) wild-type ProtL and ^{15}N-enriched Kx7E in b) 0 M, c) 0.3 M, and d) 0.8 M NaCl.

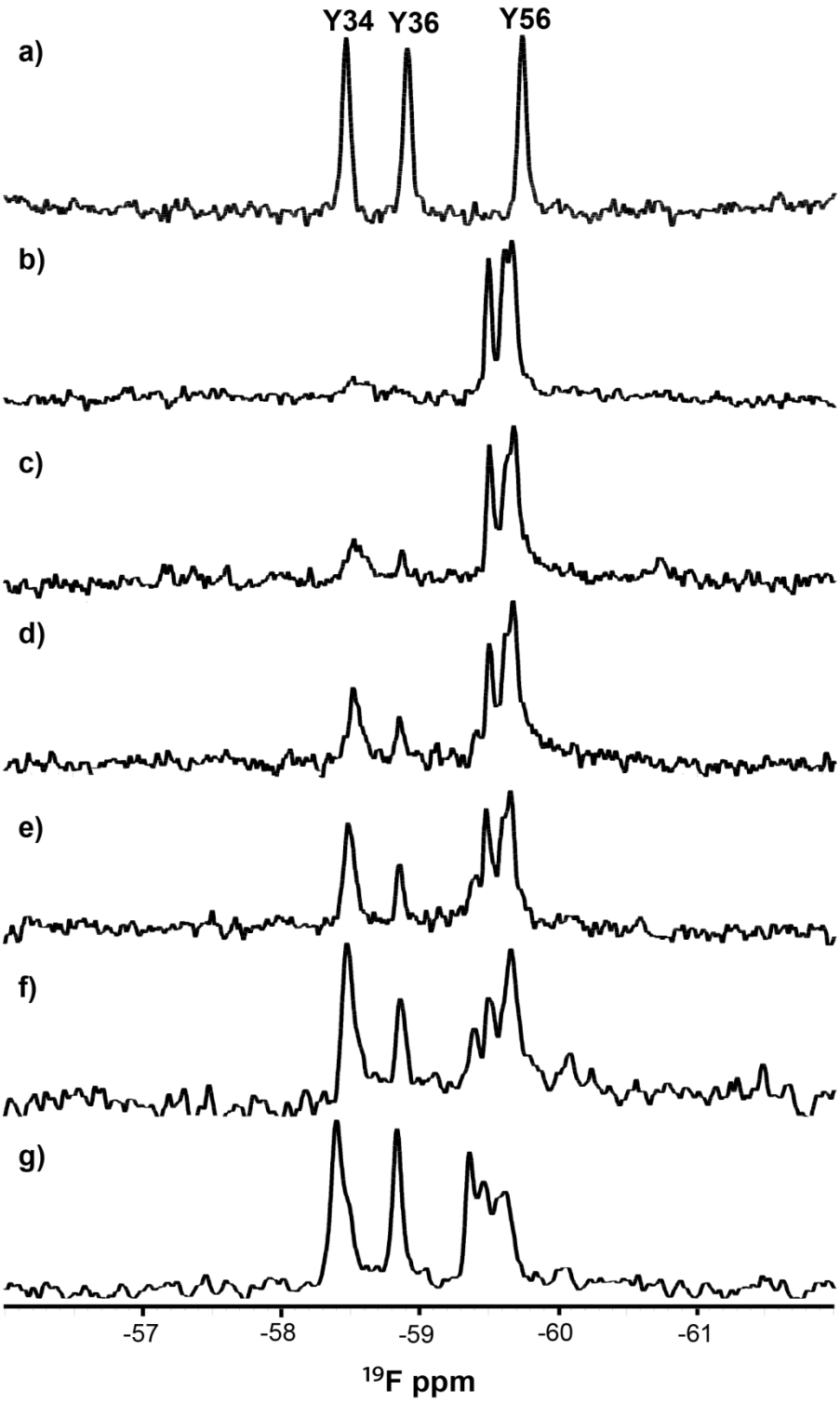


Figure 2.9 ^{19}F spectra (20 mM phosphate buffer, pH 6.0, 37 °C) of ^{19}F-labeled a) wild-type ProtL and ^{19}F-labeled Kx7E in b) 0 M, c) 0.1 M, d) 0.2 M, e) 0.3 M, f) 0.5 M, and g) 1 M NaCl. Assignments were made by using mutagenesis (Table 2.2 and Fig. 2.5).

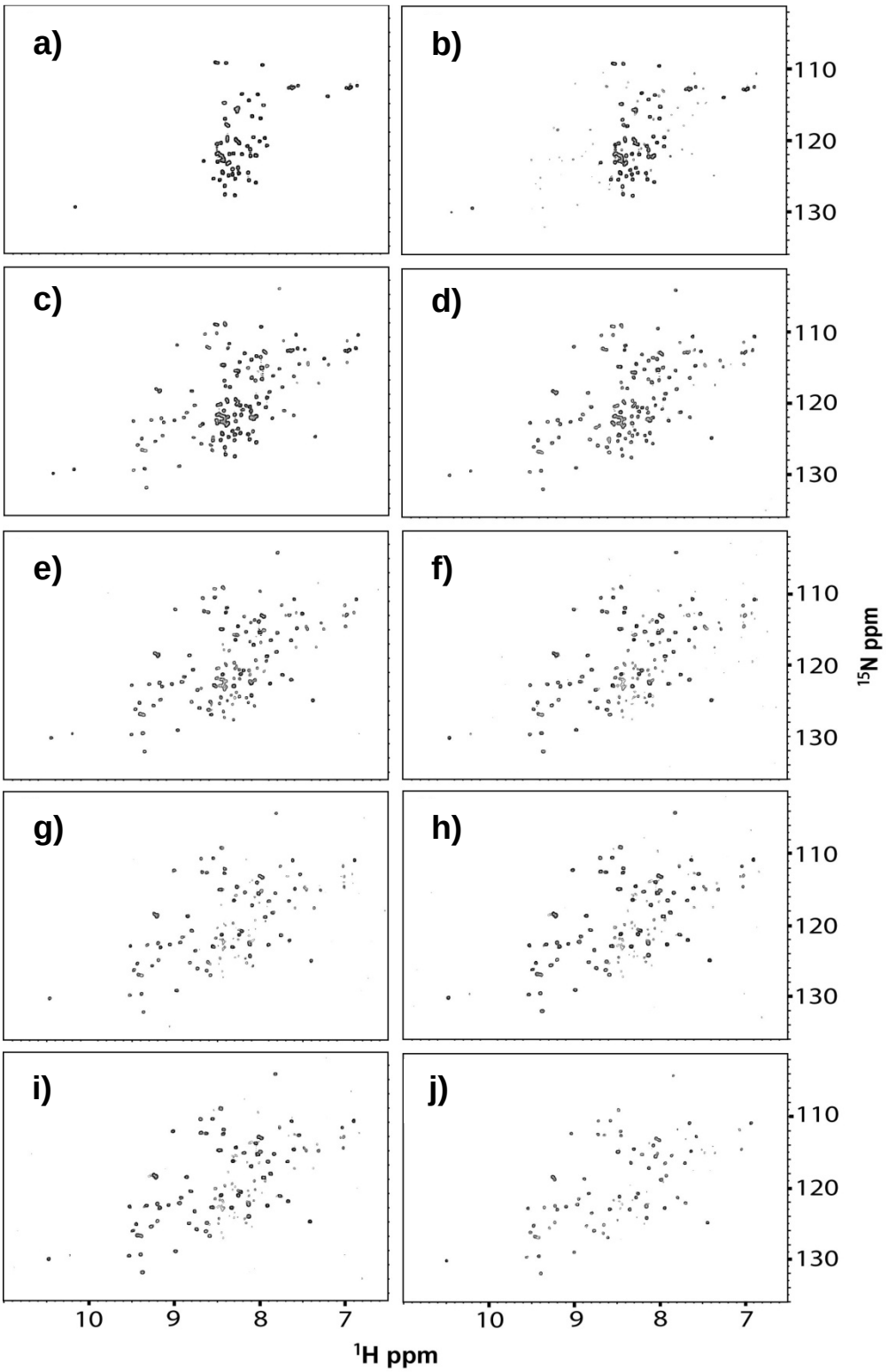


Figure 2.10 HSQC spectra (20 mM phosphate buffer, pH 6.0, 25 °C) of ^{15}N-enriched Kx7E in a) 0 M, b) 0.1 M, c) 0.2 M, d) 0.3 M, e) 0.4 M, f) 0.5 M, g) 0.6 M, h) 0.7 M, i) 0.8 M, and j) 1 M NaCl.

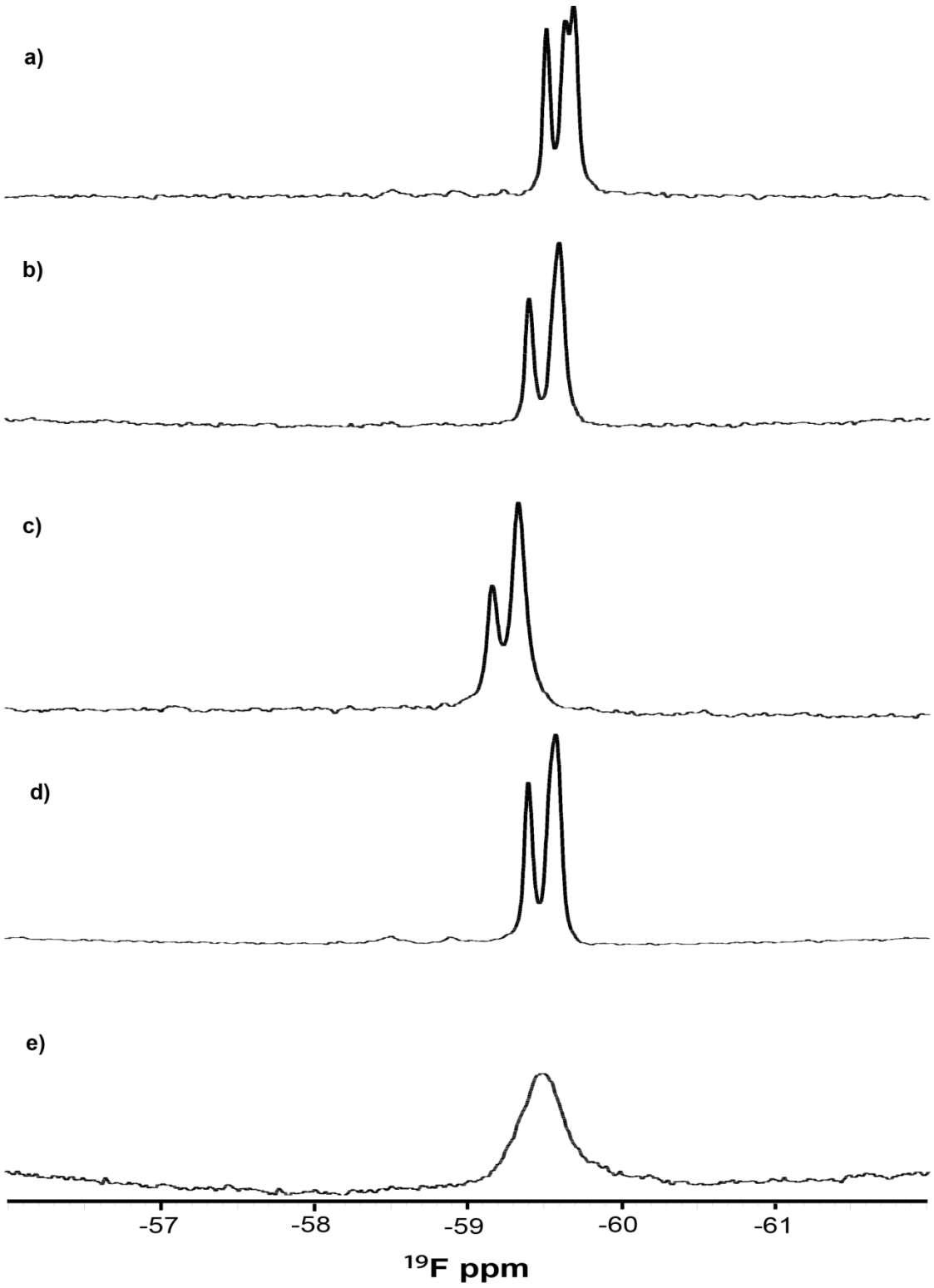


Figure 2.11 ^{19}F spectra (20 mM phosphate buffer, pH 6.0, 37 °C) of ^{19}F-labeled Kx7E in a) buffer and 300 g/L b) PEG, c) PVP, d) Ficoll, and e) BSA.

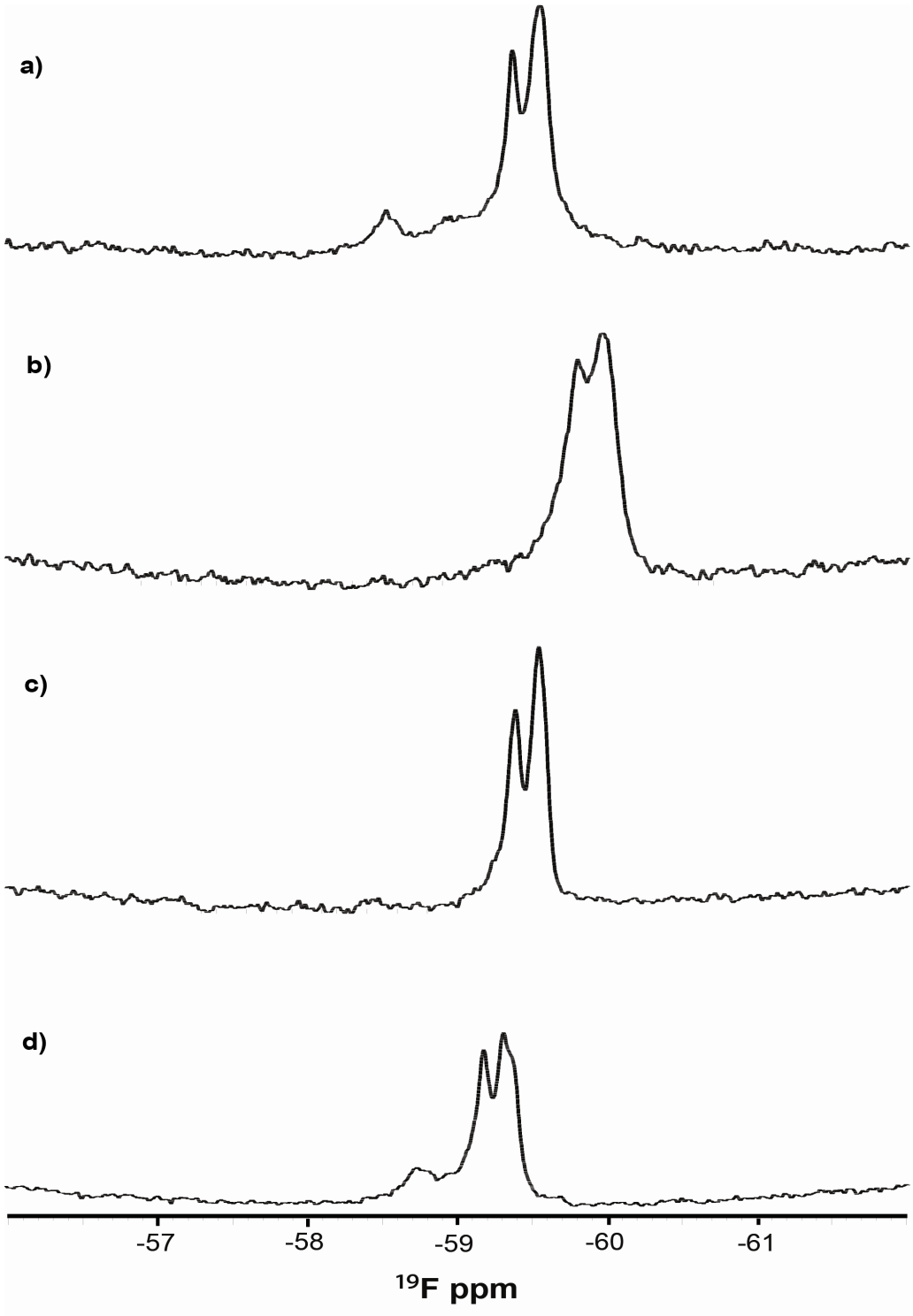


Figure 2.12 ^{19}F spectra (20 mM phosphate buffer, pH 6.0, 37 °C) of ^{19}F-labeled Kx7E in a) 585 g/L sucrose, b) 72% (v/v) glycerol, c) 300 g/L dextran, and d) 600 g/L betaine.

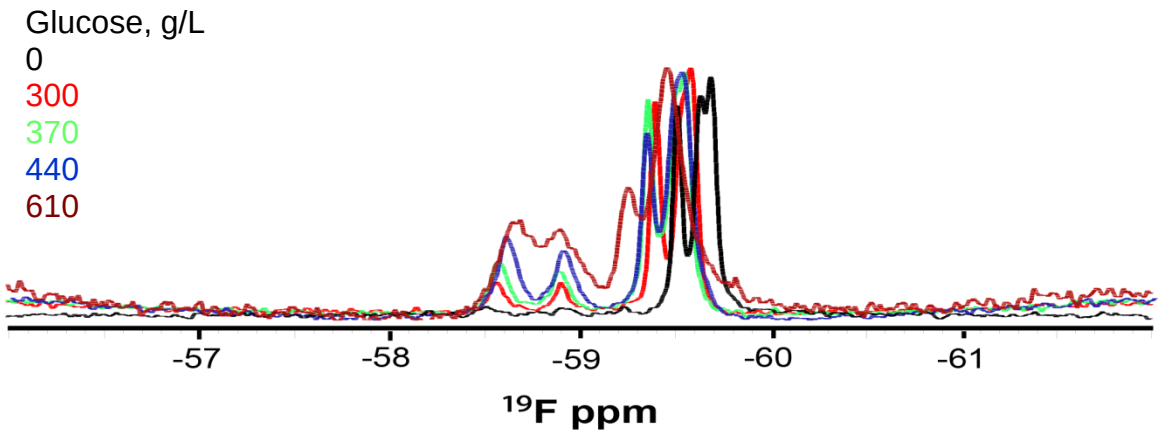


Figure 2.13 ^{19}F spectrum (20 mM phosphate buffer, pH 6.0, 37 °C) o f ^{19}F-labeled Kx7E in 300-610 g/L glucose.

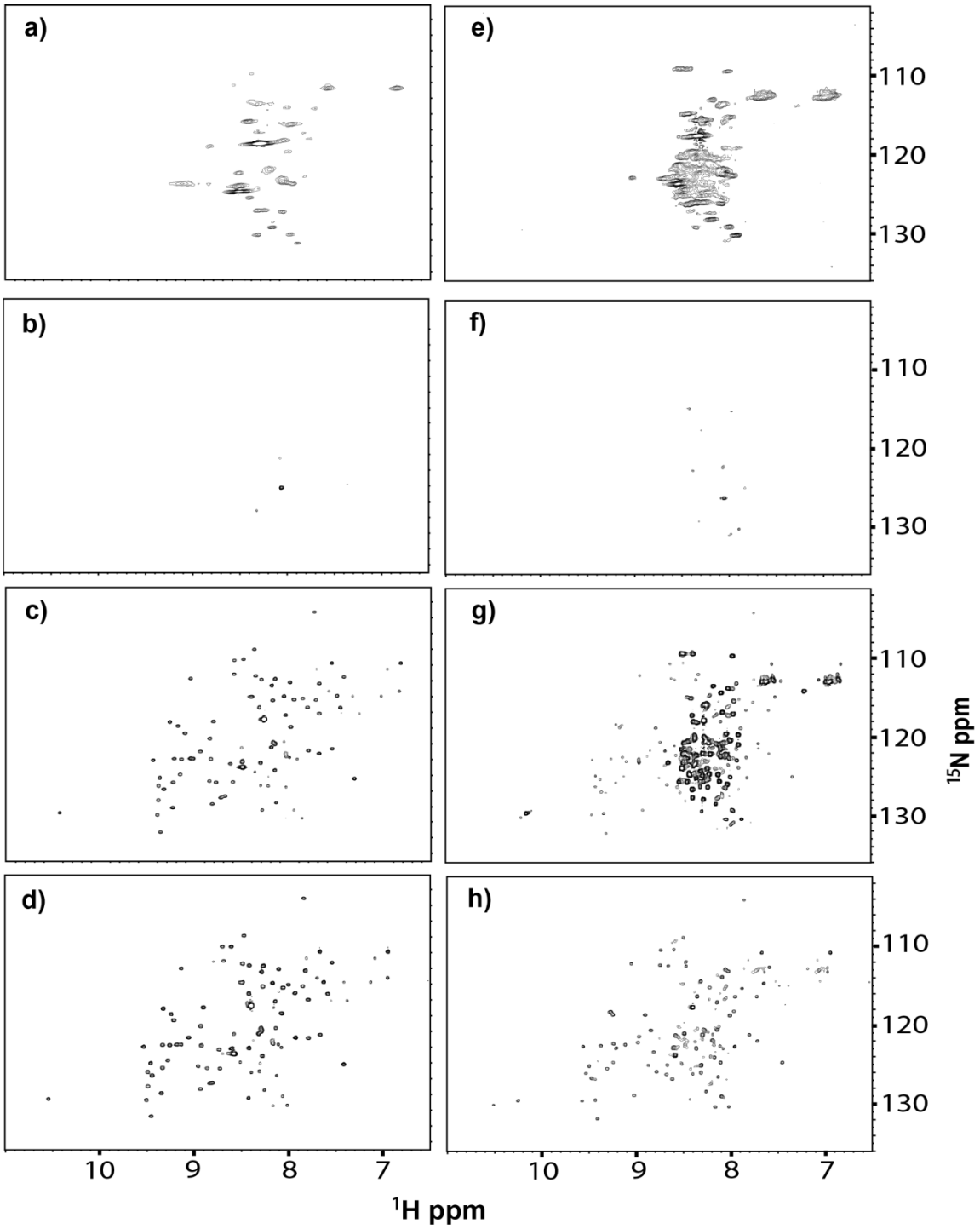


Figure 2.14 HSQC spectra (25 °C) of ^{15}N-enriched wild-type ProtL and Kx7E in *E. coli*. a) Cell slurry producing the wild-type protein. b) Wild-type ProtL cell supernatant. c) Lysate of cells from a). d) Lysate of cells shown in a) upon adding NaCl to a final concentration of 1 M. e) Cell slurry producing the Kx7E variant. f) Kx7E cell supernatant. g) Lysate of cells from e). h) Lysate of cells shown in e) upon adding NaCl to a final concentration of 1 M.

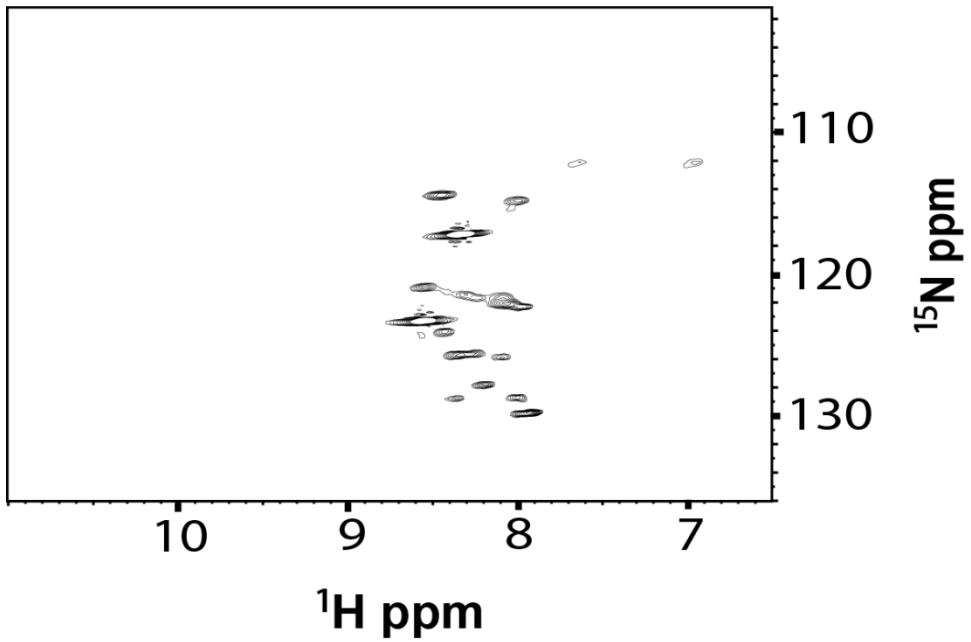


Figure 2.15 HSQC spectrum (25 °C) of *E. coli* without an expression vector.

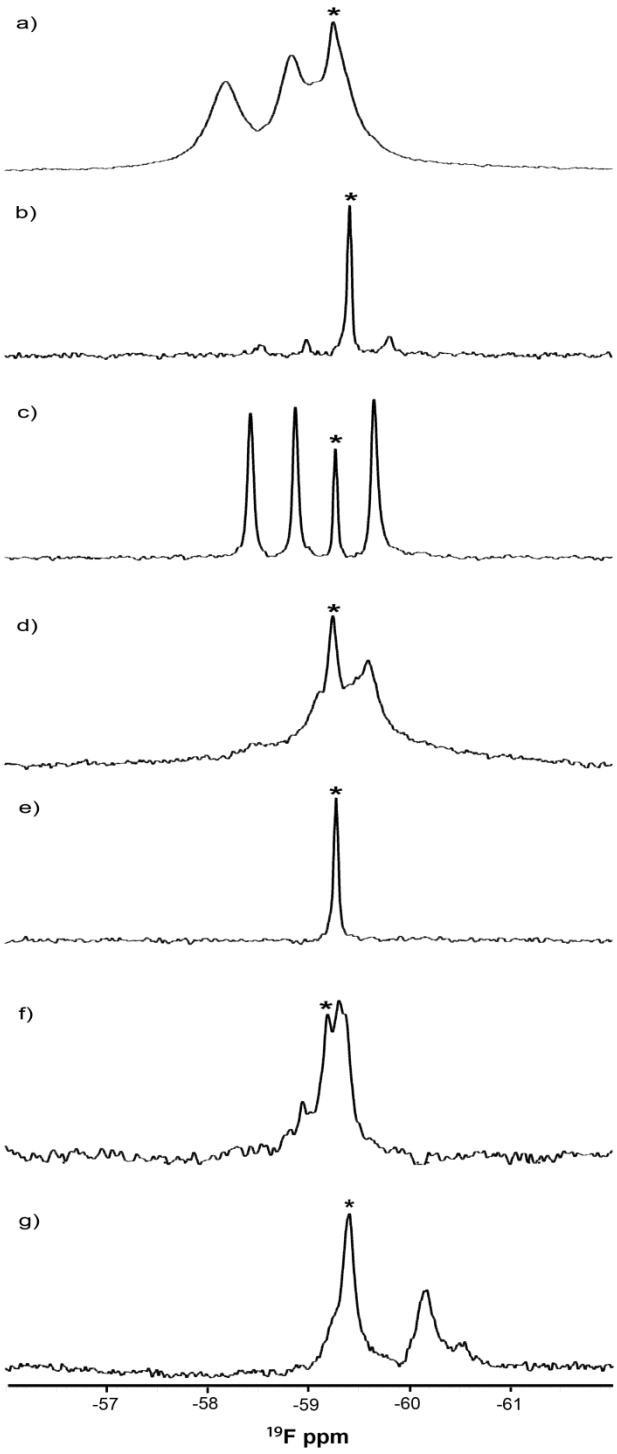


Figure 2.16 ^{19}F spectra (37 °C) of ^{19}F-labeled wild-type ProtL and Kx7E in *E. coli*. a) Cell slurry producing wild-type protein. b) Supernatant from the sample used in a). c) Lysate of cells from a). d) Cell slurry producing the Kx7E variant. e) Supernatant from the sample used in d). f) Insoluble lysate fraction of cells from d) with 3 M guanidine chloride. g) Cell slurry producing the Kx7E variant grown and induced in hyperosmotic media. Asterisks denote the free 3-FY resonance.

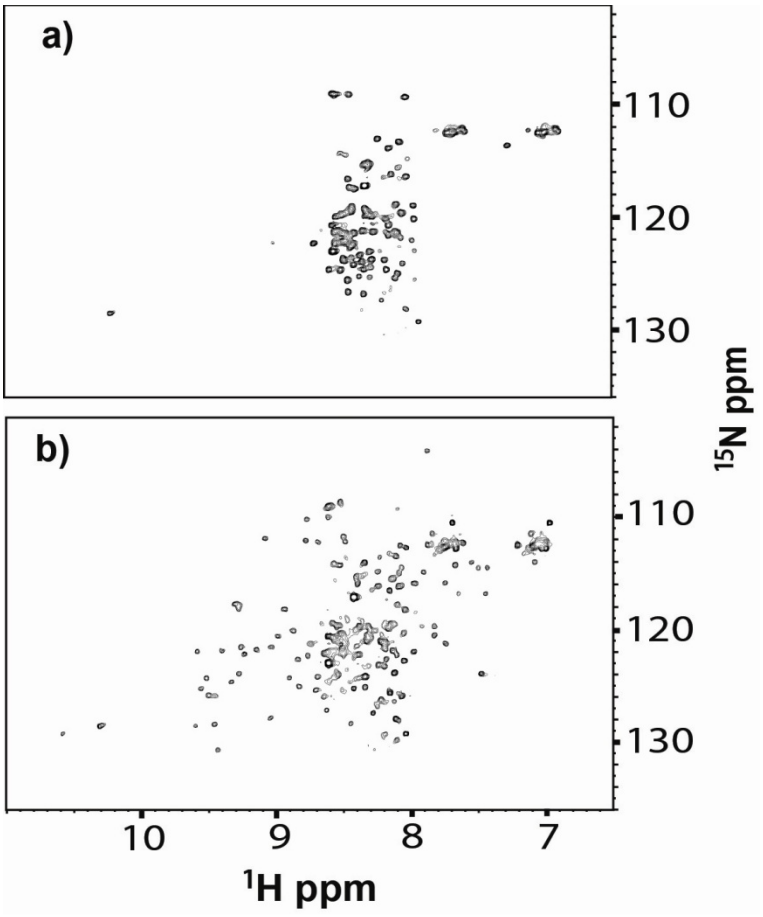


Figure 2.17 HSQC spectra (25 °C) of ^{15}N-enriched Kx7E from *E. coli*. a) Lysate of cells producing the Kx7E variant. b) Lysate of cells shown in a) upon adding KCl to a final concentration of 1 M.

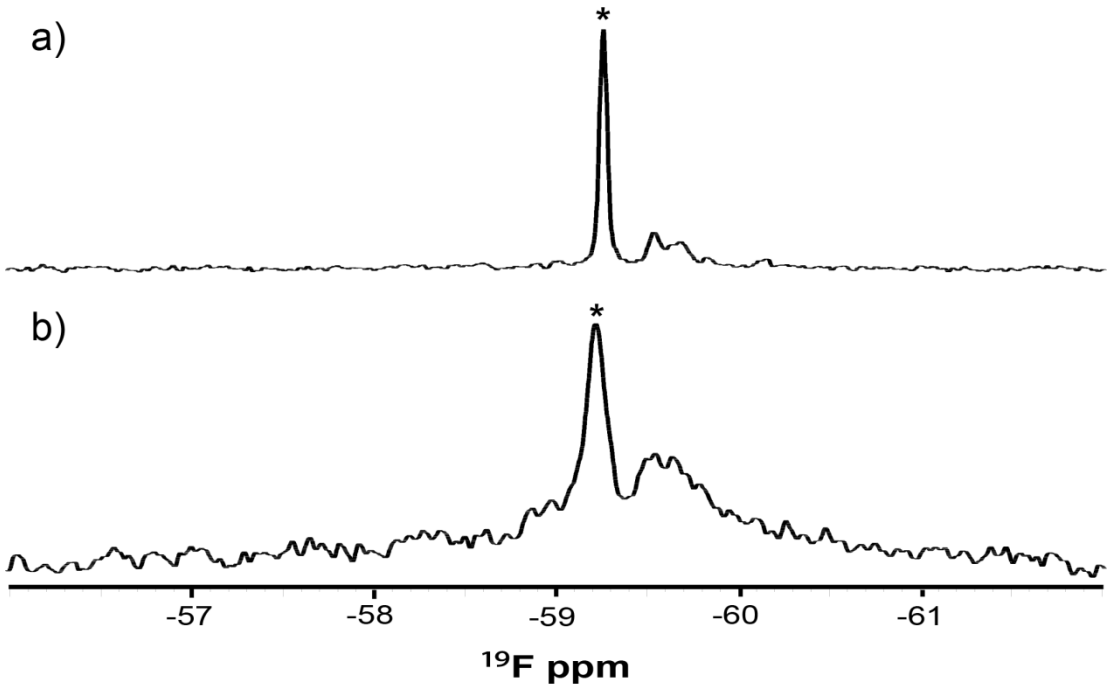


Figure 2.18 ^{19}F spectra (37 °C) of ^{19}F-labeled Kx7E from *E. coli* lysates. a) Cleared cell lysate. b) Insoluble lysate fraction resuspended in 20 mM phosphate buffer, pH 6.0. Asterisks denote the resonance from free 3-FY.

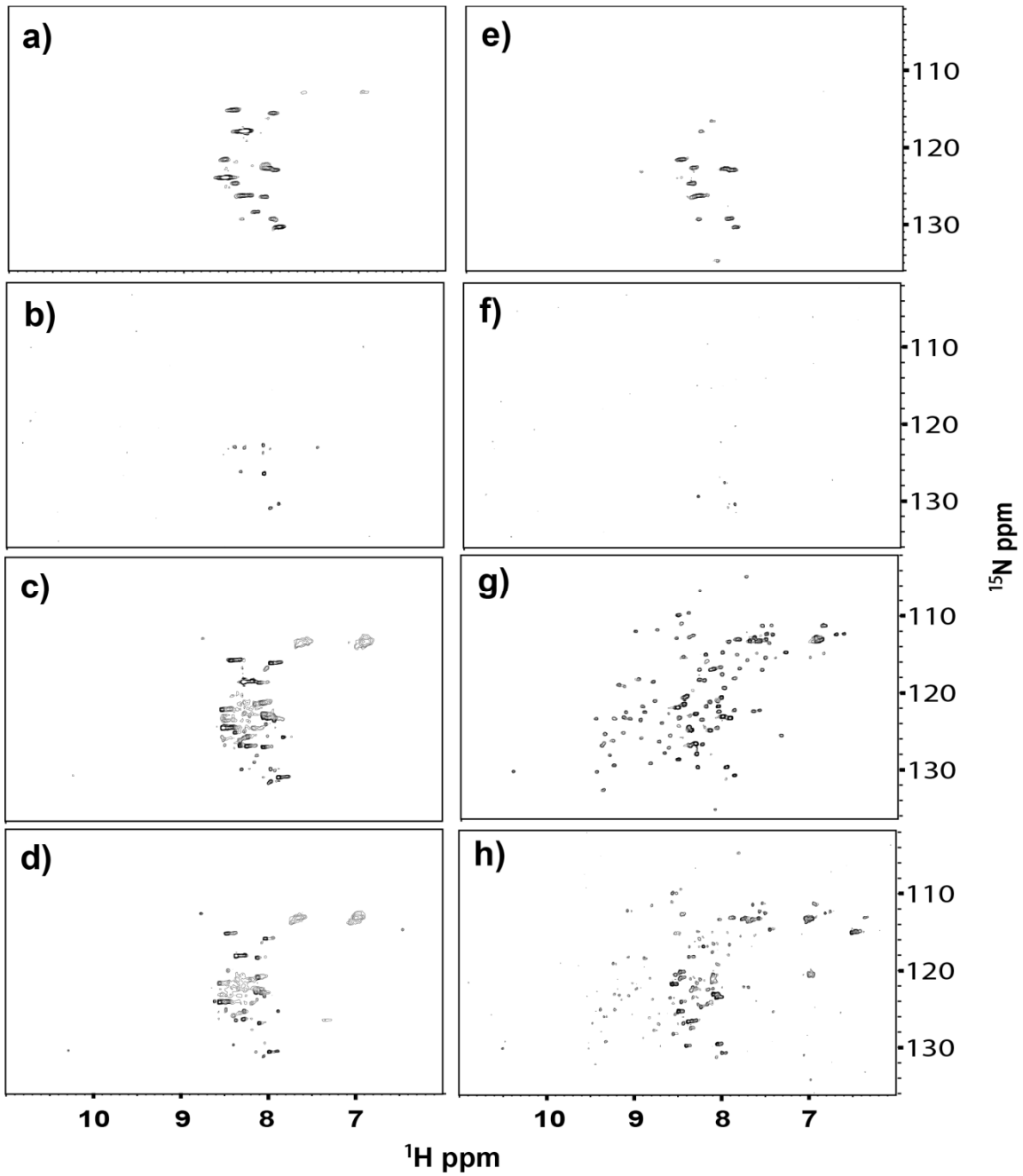


Figure 2.19 HSQC spectra (25 °C) of ^{15}N-enriched Kx6E and Kx6Q in *E. coli*. a) Cell slurry producing Kx6E. b) Kx6E cell supernatant. c) Lysate of cells from a). d) Lysate of cells shown in a) upon adding NaCl to a final concentration of 1 M. e) Cell slurry producing Kx6Q. f) Kx6Q cell supernatant. g) Lysate of cells from e). h) Lysate of cells shown in e) upon adding NaCl to a final concentration of 1 M.

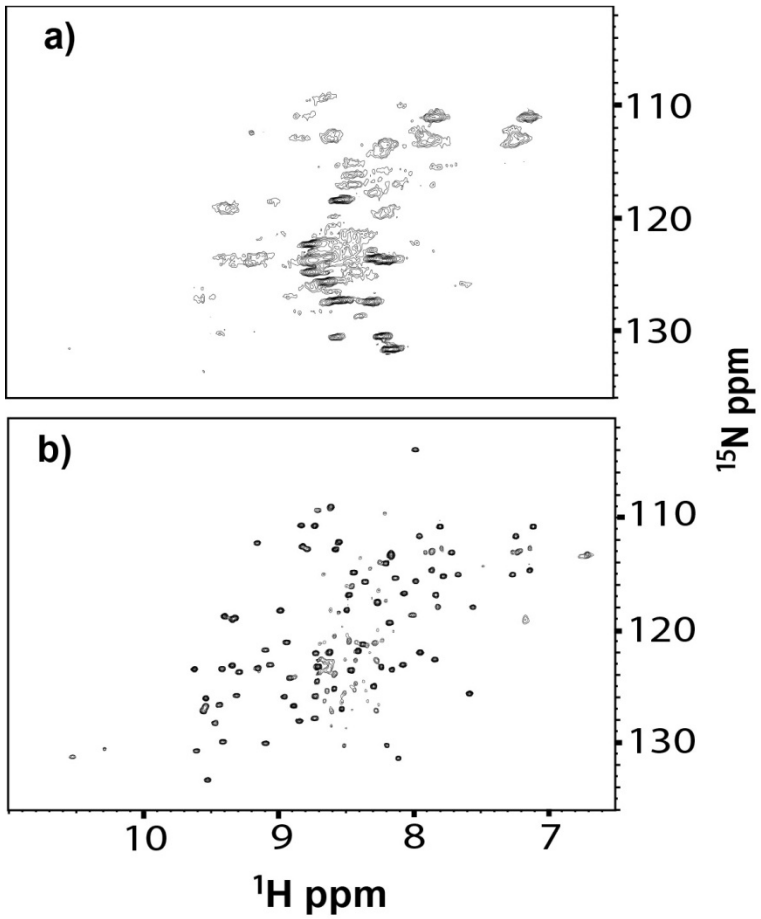


Figure 2.20 HSQC spectra (25 °C) of ^{15}N-enriched Kx6E in *E. coli*. a) Cell slurry producing Kx6E. b) Lysate of cells from a).

Chapter 3

Insect Cell Recombinant Protein Production for In-Cell NMR

3.1 Introduction

The goal of this project is to examine properties of α-synuclein (α-syn) in higher eukaryotic cells by using in-cell NMR. The Pielak group has already obtained information about α-syn in *Escherichia coli*, so the next step is to observe α-syn in higher eukaryotic cells. By using in-cell nuclear magnetic resonance (NMR) spectroscopy, properties such as structure, diffusion, and dynamics can be assessed in higher eukaryotic cells and compared to data obtained in *E. coli*. As α-syn is a human protein, the rationale behind moving its study into higher eukaryotic cells is to provide an environment that is as physiologically relevant as possible. In this way, the observed α-syn properties will be applicable to its behavior and function in humans, thus potentially elucidating its pathogenic role in Parkinson's disease.

NMR has been used to study proteins in *E. coli* and yeast and has provided valuable information about many recombinant proteins, but few in-cell NMR experiments have been performed in higher eukaryotic cells (70). Techniques for performing higher eukaryotic in-cell NMR are required to obtain physiologically relevant information about human proteins. Higher eukaryotic proteins often have additional requirements to function, such as posttranslational

modifications, which only higher eukaryotic cells can provide. The composition of higher eukaryotic cells is fundamentally different from that of prokaryotic cells, and thus human proteins must be studied in these higher order cells to obtain the most physiologically relevant data.

Higher eukaryotic cells vary in composition from prokaryotic cells. They contain organelles, carry out post-translational modifications, and are not as crowded as prokaryotic cells (3, 19). The degree of crowding is expected to affect protein properties such as structure, binding affinity, and mobility (80). For example, the diffusion coefficient for the green fluorescent protein (GFP) is significantly higher in the cytoplasm of eukaryotic cells than in the *E. coli* cytoplasm (27 $\mu m^2/s$ versus 7.7 $\mu m^2/s$, respectively) (2, 81). Additionally, the extent of crowding varies in some eukaryotic organelles, thus affecting protein properties. The rat liver mitochondrial matrix is more crowded than the cytosol (82). However, the diffusion coefficient of GFP in the mitochondrial matrix and cytoplasm are similar (2, 83), suggesting that some organization exists in the matrix to reduce crowding effects (83). These results demonstrate that significant differences exist between prokaryotic and higher eukaryotic cells and that these differences can change protein properties.

Observation of proteins by in-cell NMR has been performed in a few higher eukaryotic cells. The first experiments were performed in *Xenopus laevis* oocytes (19, 26). Routine in-cell NMR studies, however, are not practical for several reasons. The test protein must be injected into each oocyte and each experiment requires many oocytes (~200). Consistency and efficiency in the

injection process also requires specialized equipment. Moreover, their large size makes maintaining cellular structural integrity, and thus viability, difficult over the duration of NMR experiments, although embedding the cells in Ficoll solutions helps to maintain oocyte integrity and prevent protein leakage (28). Other expression systems include mammalian cell lines, but these are difficult to develop and may not produce high enough levels of protein for in-cell NMR experiments. Despite these constraints, isotopically enriched recombinant proteins have been inserted into human HeLa cells by using cell penetrating peptides, and two-dimensional NMR spectra were obtained (84).

Insect cell expression systems utilizing a recombinant baculovirus to produce the protein of interest provide a cheaper, safer alternative to many mammalian cells lines. Insect expression systems are relatively quick to set up and can produce high levels of recombinant protein (85-90). Additionally, insect cells provide a eukaryotic environment conducive to proper post-translational modifications (89). Therefore, insect cells were chosen as the higher eukaryotic cell type for production of human α-syn.

The intrinsically disordered human protein α-syn is used as a model because intrinsically disordered proteins are observable by in-cell NMR in prokaryotes (17, 18, 49), and their physical properties are affected by macromolecular crowding (17, 18, 49, 91). α-Syn is also a relevant protein because of its involvement in Parkinson’s disease (92). Parkinson's disease is a neurological disorder resulting in tremors, rigid movement, and slowing of voluntary body motions that afflicts approximately 1-2% of individuals over 65

years of age (93). α-Syn forms fibrillar aggregates in cytoplasmic inclusions known as Lewy bodies, which are the clinical hallmark of Parkinson's disease (92, 94). Additionally, the Pielak group has experience producing α-syn and performing *in vitro* NMR (95) and in-cell NMR in *E. coli* (18, 49, 52).

The Pielak group has studied the effects of macromolecular crowding on α-syn structure in dilute solution, under crowded conditions *in vitro*, and in *E. coli* (18, 91, 96). Crowding *in vitro* and in *E. coli* prevents a temperature induced conformational change observed in dilute solution (Fig. 3.1) (18). α-Syn is disordered and collapsed in *in vitro* crowded solutions and in *E. coli* at 35 °C but is extended and gains some structure in dilute solution at 35 °C.

The α-syn ^{1}H-^{15}N heteronuclear single-quantum coherence (HSQC) spectrum in dilute solution at 10 °C contains about 70 crosspeaks (Fig. 3.1*a*) and is representative of a disordered protein, while the spectrum at 35 °C contains about 30 crosspeaks (Fig. 3.1*b*). This difference has been attributed to an intermediate reversible conformational exchange that occurs on the chemical shift time scale where the first 100 amino acid residues of α-syn gain secondary structure (91). The remaining crosspeaks in dilute solution at 35 °C are from the C-terminal disordered region. The α-syn disorder observed in dilute solution at 10 °C is maintained in the *in vitro* crowded bovine serum albumin solution at 35 °C (Fig. 3.1*d*) and in *E. coli* at 35 °C (Fig. 3.1 *c*). Thus, these crowded environments prevent the temperature induced α-syn conformational change observed in dilute solution at 35 °C. Additionally, the nearly identical spectra in *E. coli* at 35 °C and

in vitro crowded solution at 35 °C suggest that macromolecular cr owding alone is sufficient to maintain α-syn disorder.

The time required for α-syn protofibril and mature fibril formation is shortened under *in vitro* crowded conditions (97, 98). These experiments suggest that changes in cell cytoplasmic concentration could promote formation of α-syn protofibrils and stabilize them (97). Some evidence shows that the total concentration of macromolecules in cells changes as individuals age. For example, cell volumes decrease in aged rats (99), and protein degradation is inhibited in brains of patients with Parkinson's disease (100). Both situations could lead to an effective increase in the concentration of α-syn, promoting α-syn fibril formation. Furthermore, the reduced cellular viscosity in higher eukaryotes compared to prokaryotes may accelerate disease causing α-syn aggregates (101) due to increased diffusion (98). Therefore, studying the effects of macromolecular crowding on α-syn properties in higher eukaryotic cells should yield important insights into the molecular basis of Parkinson's disease.

No NMR information about α-syn in higher eukaryotic cells is available. Other intrinsically disordered proteins, such as the C-terminal activation domain of c-Fos and the kinase-inhibition domain of $p27^{Kip1}$, have been shown not to undergo significant conformational changes in the presence of high concentrations of dextrans and Ficoll 70 as assessed by circular dichroism and fluorescence spectroscopy (102). These observations, coupled with data showing that *E. coli* maintains α-syn disorder, lead to the prediction that α-syn will also remain unfolded in insect cells. It is possible, however, that an increase

in α-syn diffusion and dynamics might be observed because higher eukaryotic cells are less crowded than *E. coli*.

To observe α-syn by in-cell NMR in higher eukaryotic cells, the baculovirus expression system is used to generate a recombinant baculovirus containing the human α-syn gene for production of high levels of isotopically labeled α-syn in *Spodoptera frugiperda* 9 (Sf9) insect cells. Many recombinant proteins have been isotopically labeled in Sf9 cells by using the baculovirus expression system. In one of the earliest examples, recombinant rhodopsin was produced and labeled in Sf9 cells with deuterated tyrosines and analysed by Fourier transform infrared spectroscopy (103). Rhodopsin was also the first ^{15}N-enriched eukaryotic membrane protein to be analysed by solid state NMR, and it was produced and enriched in Sf9 cells (104). Brüggert et al. demonstrated selective ^{15}N enrichment of purified glutathione-S-transferase protein produced in Sf9 insect cells (105).

Previous work shows that α-syn can be produced in Sf9 insect cells (106). Production of sufficient concentrations of isotopically labeled α-syn for insect in-cell NMR experiments, however, requires overcoming some technical hurdles. The first hurdle is tracking the progress of insect baculovirus infection so that α-syn production may be followed. To overcome this hurdle, GFP is produced in insect cells either in parallel or as a fusion protein with α-syn. GFP has been successfully produced in baculovirus infected Sf9 cells (107) and is a convenient reporter protein due to its fluorescence; baculovirus infection and protein production can be followed under a fluorescent microscope (108). Furthermore,

the solution NMR spectra of 3-fluorotyrosine (3-FY)-labeled GFP (64) and ^{15}N-enriched GFP (109) are available, potentially aiding in the assessment of GFP spectra observed in Sf9 cells.

The second hurdle is finding a culture media for the isotopic labeling of α-syn and GFP in Sf9 cells. Commercial insect media for the isotopic labeling of recombinant proteins are prohibitively expensive. For example, ^{15}N-enriched Bioexpress-2000 media from Cambridge Isotope Laboratories, Inc. is $7,500 per liter. Thus, custom media based on the formulation presented by Walton et al. (110) are prepared to isotopically enrich or label the recombinant proteins in Sf9 cells at a fraction of the cost. The methods used to produce isotopically enriched or labeled α-syn and GFP in Sf9 cells and findings from in-cell NMR experiments are presented here.

3.2 Materials and Methods

3.2.1 Insect Cell Culture Growth and Maintenance

Sf9 Cell Monolayer Culture

Spodoptera frugiperda 9 (Sf9) cells (Lineberger Comprehensive Cancer Center Tissue Culture Facility, University of North Carolina at Chapel Hill [UNC-CH]) were provided in 25-cm^2 cell culture flasks in Grace's Medium and 10% heat inactivated fetal bovine serum (FBS). Cells were grown and passaged by dilution in Sf-900 II SFM (serum-free medium; Invitrogen Corp.; Carlsbad, CA) without antibiotics in 25-cm^2 cell culture flasks (Corning, Inc. - Life Sciences; Wilkes-Barre, PA). Cells were grown at 27 °C under ambient air conditions in an Isotemp Undercounter Refrigerated Incubator (Thermo Fisher Scientific, Inc.;

Dubuque, IA). Cells were passaged by dilution 1:10 with fresh media once monolayer confluency reached 80%-100%. Cell density and viability were determined with a hemocytometer (Hausser Scientific Co.; Horsham, PA) and 0.4% trypan blue (Sigma-Aldrich Co.; St. Louis, MO) staining using an Axiovert 40 CFL microscope (Carl Zeiss, Inc.; Thornwood, NY). All Sf9 cell culture work was performed in a Purifier Class II Biosafety Cabinet (Labconco Corp.; Kansas City, MO).

Sf9 Cell Suspension Culture

Sf9 cells (Lineberger Comprehensive Cancer Center Tissue Culture Facility, UNC-CH) were grown and passaged in Sf-900 II SFM without antibiotics. Cells were grown at 27 °C under ambient air conditio ns in 125-mL and 250-mL spinner flasks (Corning, Inc. - Life Sciences) on a Thermolyne Type 45700 Cellgro stirrer (Thermo Fisher Scientific, Inc.; Waltham, MA) at 70 rotations per minute (rpm). Cells were seeded at 0.5 x 10^6 cells/mL and passaged by dilution with fresh media to 0.5 x 10^6 cells/mL when cell densities reached >2 x 10^6 cells/mL.

Sf9 cells (Expression Systems, LLC; Woodland, CA) were grown and passaged in ESF 921 insect cell medium (Expression Systems, LLC) without antibiotics. Cells were grown as 20-mL suspension cultures in 125-mL Erlenmeyer shake flasks (Corning, Inc. - Life Sciences) with the solid caps vented half-a-turn. Cells were grown at 27 °C under ambient air conditions on an Orbit 1900 Heavy Duty Shaker (Labnet International, Inc.; Edison, NJ) at 150 rpm. Cells were seeded at 0.5 x 10^6 cells/mL and passaged by dilution with fresh

media to 0.5 x 10^6 cells/mL when cell densities reached 6-8 x 10^6 cells/mL, generally after 5 days. Sf9 cells in suspension were passaged up to twelve times. Cell density and viability were determined with the hemocytometer and 0.4% trypan blue staining. All Sf9 cell culture work was performed in the Purifier Class II Biosafety Cabinet.

Drosophila *S2 Cell Monolayer Culture*

Drosophila Schneider 2 (S2) cells (Invitrogen Corp.) were grown in 25-cm^2 cell culture flasks in Schneider's medium (Invitrogen Corp.) supplemented with 10% heat inactivated FBS (Atlanta Biologicals, Inc.; Lawrenceville, GA). No antibiotics were added. S2 cells were grown at 27 °C u nder ambient air conditions and passaged by dilution 1:10 with fresh media once monolayer confluency reached 80%-100%. Cells were passaged up to twelve times. Cell density and viability were determined with the hemocytometer and 0.4% trypan blue staining. All S2 cell culture work was performed in the Purifier Class II Biosafety Cabinet.

3.2.2 Recombinant Baculovirus Production

EGFP pFastBac Dual Vector

The pFastBac Dual vector (Invitrogen Corp.) containing the enhanced green fluorescent protein (EGFP) gene under the p10 promoter was provided by Dr. Michael Miley of the Redinbo lab in the Department of Chemistry, UNC-CH.

α-Syn pFastBac 1 Vector

The α-syn gene from the PT7-7 vector (18) was mutated using a site-directed mutagenesis kit (Agilent Technologies, Inc.; Santa Clara, CA) to remove

the internal *EcoRI* site and amplified by using polymerase chain reaction (PCR) with primers 5' CTTTAAGAAGGAGAATTCATATGGATGTATTCATG 3' and 5' GAACATCTGTCAGCAGAATTCAAGAAACTGGGAGC 3' (111) and Vent DNA polymerase (New England Biolabs, Inc.; Ipswich, MA) in an Eppendorf Mastercycler (Eppendorf North America; Hauppauge, NY). Both primers contain the *EcoRI* restriction site. The PCR reaction products were separated on 1% (w/v) agarose (Sigma-Aldrich Co.) gels electrophoresed at 100 V for 30 min. The amplified α-syn gene was excised and purified by using a gel extraction kit (Qiagen, Inc.; Germantown, MD). The α-syn PCR product and the pFastBac 1 vector (Invitrogen Corp.) were digested with *EcoRI* (New England Biolabs, Inc.) and purified by agarose gel electrophoresis followed by gel extraction. The digested products were mixed and ligated with Quick T4 DNA ligase (New England Biolabs, Inc.) and transformed into DH5α *E. coli* cells (Invitrogen Corp.). Cells were streaked onto Luria broth (LB) plates containing 100 µg/mL ampicillin (LB_{AMP}; Thermo Fisher Scientific, Inc.; Fair Lawn, NJ) and incubated overnight at 37 °C. Colonies were picked from the plates and used to inoculate 5-mL liquid cultures of LB_{AMP} and grown for 12-16 h at 37 °C with 250-rpm shaking . α-Syn pFastBac 1 DNA was isolated by using a miniprep kit (Qiagen, Inc.; Valencia, CA) and sequenced (UNC-CH Genome Analysis Facility) to identify constructs with the α-syn gene inserted into the vector in the proper orientation.

α-Syn pUltraBac-1 and α-Syn-EGFP pUltraBac-1 Vectors

The pUltraBac-1 vector containing the EGFP gene under a late basic protein promoter (P_{Basic}) was provided by Felix Freuler (Novartis International AG;

Basel, Switzerland). To generate the α-syn pUltraBac-1 vector, the α-syn gene from the PT7-7 vector was amplified by using PCR with primers 5' GAAATAATTT TGTTTAACCTCGAGAAGGAGATATACATATGGATG 3' and 5' GAAACTGGGA GCAATGCATATTTCTTAGGCTTCAGGTTC 3' containing *XhoI* and *NsiI* restriction sites, respectively. To generate the α-syn-EGFP pUltraBac-1 vector, the α-syn gene from the PT7-7 vector was amplified by using PCR with primers 5' GAAATAATTTTGTTTAACCTCGAGAAGGAGATATACATATGGATG 3' and 5' GTAACACTCGAGAGGCTTCAGGTTCGTAGTCTTGATACC 3'. Both contain the *XhoI* restriction sites. The PCR reaction products were separated on 1% (w/v) agarose gels electrophoresed at 100 V for 55 min. The amplified α-syn genes were excised and purified by using the gel extraction kit. The pUltraBac-1 vector and the α-syn PCR product for α-syn pUltraBac-1 vector generation were digested with *XhoI* and *NsiI* (New England Biolabs, Inc.). The pUltraBac-1 vector and the α-syn PCR product for α-syn-EGFP pUltraBac-1 vector generation were digested with *XhoI*. The digested products were separated by agarose gel electrophoresis and purified by using the gel extraction kit. The digested products were mixed and ligated with T4 DNA ligase (New England Biolabs, Inc.) overnight at 16 °C and transformed into DH5 α *E. coli*. Cells were streaked onto LB_{AMP} plates and incubated overnight at 37 °C. Liquid cu ltures of LB_{AMP} (5 mL) were inoculated with colonies from the LB_{AMP} plates and grown for 12-16 h at 37 °C with 250-rpm shaking. The α-syn pUltraBac-1 and α-syn-EGFP pUltraBac-1 vector DNA was isolated by using the miniprep kit and checked by restriction

enzyme mapping and DNA sequence analysis to identify constructs containing the α-syn gene inserted into the vector in the proper orientation.

Recombinant Bacmids

Recombinant bacmids and baculoviruses were generated as described in the instructions for the Bac-to-Bac Baculovirus Expression System (Invitrogen Corp.). The EGFP pFastBac Dual, α-syn pFastBac 1, α-syn pUltraBac-1, or α-syn-EGFP pUltraBac-1 vector DNA were transformed into DH10Bac *E. coli* cells (Invitrogen Corp.) to allow transposition into the DH10Bac bacmid. Cells were streaked onto LB plates containing 50 µg/mL kanamycin (Thermo Fisher Scientific, Inc.; Fair Lawn, NJ), 7 µg/mL gentamicin (MP Biomedicals, LLC; Solon, OH), 10 µg/mL tetracycline (Thermo Fisher Scientific, Inc.; Fair Lawn, NJ), 100 µg/mL Bluo-gal (Invitrogen Corp.), and 40 µg/mL isopropyl-β-D-1-thiogalactopyranoside (IPTG; Thermo Fisher Scientific, Inc.; Fair Lawn, NJ) for blue/white colony selection. Positive colonies were selected and the recombinant bacmids purified as described by the manufacturer.

EGFP pFastBac Dual and α-Syn pFastBac 1 Recombinant Baculoviruses

Transfection mixtures comprising 1 µg of EGFP pFastBac Dual recombinant bacmid or 1 µg of α-syn pFastBac 1 recombinant bacmid and 6 µL of Cellfectin (Invitrogen Corp.) were added to each well of a 6-well plate (Corning, Inc. - Life Sciences) seeded with 1 x 10^6 Sf9 cells/well. After 5 h of incubation, the DNA:lipid complexes were removed. Two mL of Sf-900 II SFM were added to each well of the 6-well plates. The plates were incubated at 27 °C for 72 h. Cell culture media were collected and centrifuged in a Centrific Model

228 centrifuge with a 232 rotor (Thermo Fisher Scientific, Inc.; Pittsburgh, PA) at 1,380 x g for 5 min at room temperature to remove cell debris. The supernatants are the P1 baculovirus stocks. The P1 stocks were added to Sf9 cells in suspension (2 x 10^6 cells/mL) at a multiplicity of infection (MOI, number of infectious particles per cell) of 0.1, and the resulting cell cultures were incubated at 27 °C with 150-rpm shaking. The desired amount of virus (in mL) is calculated by multiplying the desired MOI by the total number of cells and then dividing by the titer of the baculovirus stock. Cell culture media were collected 48 h post infection (p.i.) to generate the P2 stocks. P3 stocks were generated from the P2 stocks in the same way. The P1, P2, and P3 stocks were not titered, but the concentrations were assumed to be 5 x 10^6 plaque forming units (pfu) per mL, 5 x 10^7 pfu/mL, and 1 x 10^8 pfu/mL, respectively, as suggested by the manufacturer. The baculovirus stocks were stored at 4 ° C.

α-Syn pUltraBac-1 and α-Syn-EGFP pUltraBac-1 Recombinant Baculoviruses

Transfection mixtures comprising 15 µg of α-syn pUltraBac-1 recombinant bacmid or 15 µg of α-syn-EGFP pUltraBac-1 recombinant bacmid and 70 µL of Cellfectin II (Invitrogen Corp.) were added to 75-cm^2 cell culture flasks (Corning, Inc. - Life Sciences) seeded with 1.3 x 10^7 Sf9 cells per culture flask. After 5 h, the DNA:lipid complexes were removed, and 15 mL of ESF 921 media were added to each flask. The flasks were incubated at 27 °C for 72 h. Cell culture media were collected and centrifuged in the Centrific Model 228 centrifuge at 1,380 x g for 3 min at room temperature to remove cell debris. The supernatants are the P1 baculovirus stocks. The P1 stocks were added to 25 mL of Sf9 cells in

suspension (2 x 10^6 cells/mL) at a MOI of 0.1, and the cell cultures were incubated at 27 °C with 150-rpm shaking. The P1 stock s were not titered but assumed to be 5 x 10^6 pfu/mL as suggested by the manufacturer. Cell culture media were collected 72 h p.i. when the majority (83%) of Sf9 cells infected with α-syn-EGFP pUltraBac-1 recombinant baculovirus fluoresced green. The α-syn-EGFP pUltraBac-1 P2 stock was titered in a 96-well plate (Corning, Inc. - Life Sciences) based on the number of green-fluorescing Sf9 cells by using the 50% tissue culture infectious dose method ($TCID_{50}$) (112). The titer of the α-syn pUltraBac-1 P2 stock was assumed to be the same. Each P2 stock was used to infect five 50-mL Sf9 suspension cultures (3 x 10^6 cells/mL) at a MOI of 0.1 to generate the P3 stocks. The cell suspensions were harvested 48 h p.i. and centrifuged in the Centrific Model 228 centrifuge at 1,380 x g for 5 min to pellet the cells. The media were sterile filtered with 0.22 µm polyvinylidene difluoride (PVDF) filters (EMD Millipore; Billerica, MA) into Nalgene Oak Ridge high speed centrifuge tubes (Thermo Fisher Scientific, Inc.; Rochester, NY) and centrifuged in a Sorvall RC-5B centrifuge with a Sorvall SS-34 rotor (Thermo Fisher Scientific, Inc.; Waltham, MA) at 19,000 rpm for 1 h at 4 °C. The supernatants were decanted and the baculovirus pellets resuspended in a total of 1.2 mL Dulbecco's Phosphate Buffered Saline (PBS), without addition of $CaCl_2$ and $MgCl_2$ (Sigma-Aldrich Co.). The titer of the α-syn-EGFP pUltraBac-1 P3 stock was determined to be 8.47 x 10^9 pfu/mL as described above. The titer of the α-syn pUltraBac-1 P3 stock was assumed to be the same. The baculovirus stocks were stored at 4 °C.

3.2.3 Insect Cell Recombinant Protein Production and Visualization

EGFP Production and Visualization

Sf9 suspension cells were grown to 2 x 10^6 cells/mL in Sf-900 II SFM using 250-mL spinner flasks and infected with P3 EGFP pFastBac Dual recombinant baculovirus stock (1 x 10^8 pfu/mL) at a MOI of 10. Fluorescent cells were visualized with a Zeiss 510 Meta Laser Scanning Confocal Microscope (Michael Hooker Microscopy Facility, UNC-CH). Cells were harvested 72 h p.i. by centrifugation in the Centrific Model 228 centrifuge at 1,380 x g for 2 min and resuspended in 2% sodium dodecyl sulfate (SDS) lysis buffer (2% [w/v] SDS, 62.5 mM Tris-HCl, pH 6.7, 15% [v/v] glycerol, 0.1% [w/v] bromophenol blue [Sigma-Aldrich Co.]). Lysates were loaded onto 18% Tris-HCl SDS polyacrylamide gels (Bio-Rad Laboratories, Inc.; Hercules, CA). Gels were electrophoresed at 200 V for 50 min. EGFP fluorescence was detected by using a VersaDoc 4000 MP Imaging System (Bio-Rad Laboratories, Inc.) and correlated to protein concentration by measuring band densities with the built-in software. After scanning, gels were stained with Brilliant Blue R-250 (Thermo Fisher Scientific, Inc.; Pittsburgh, PA). Gels were then destained in aqueous solutions containing 40% (v/v) methanol and 10% (v/v) acetic acid followed by final destaining in a 10% methanol, 10% acetic acid solution.

α-Syn Production and Visualization

Sf9 suspension cells were grown to 1.5 x 10^6 cells/mL in Sf-900 II SFM by using 125-mL spinner flasks and infected with P3 α-syn pFastBac 1 recombinant baculovirus stock (1 x 10^8 pfu/mL) at a MOI of 10. Cells were harvested every 24

h p.i. for 8 days, centrifuged in the Centrific Model 228 centrifuge at 1,380 x g for 5 min, and resuspended in 2% SDS lysis buffer. Lysates were loaded onto 18% Tris-HCl SDS polyacrylamide gels and electrophoresed at 200 V for 65 min. Western blotting was performed by transferring gel contents to PVDF membranes (Bio-Rad Laboratories, Inc.) and electrophoresing at 150 mA for 120 min. The membranes were blocked with 5% (w/v) non-fat dry milk (LabScientific, Inc.; Livingston, NJ) solutions. A 2000x dilution of α-syn sc-12767 mouse monoclonal antibody (Santa Cruz Biotechnology, Inc.; Santa Cruz, CA) was added to bind α-syn. Excess antibody was washed off, and a 4000x dilution of goat anti-mouse immunoglobulin G horse radish peroxidase (IgG-HRP sc-2005, Santa Cruz Biotechnology, Inc.) was added to the membranes. HRP chemiluminescence was generated by using a Pierce ECL Western Blotting Substrate kit (Thermo Fisher Scientific, Inc.; Rockford, IL) and detected with the VersaDoc 4000 MP Imaging System. α-Syn levels were quantified with the built-in software.

α-Syn and α-Syn-EGFP Production and Visualization

Sf9 cells in suspension were seeded in 125-mL spinner flasks to a final density of 1.9 x 10^6 cells/mL in ESF 921 media (9 mL total suspension volume) and infected with P2 α-syn pUltraBac-1 or P2 α-syn-EGFP pUltraBac-1 recombinant baculovirus stock (5.85 x 10^6 pfu/mL) at a MOI of 1. Fluorescent cells were visualized with the Axiovert 40 CFL microscope equipped with a HBO 50 ac mercury lamp (Carl Zeiss, Inc.). One-mL aliquots of cell suspensions were collected 48 h and 72 h p.i. and centrifuged in a Centrifuge 5418 benchtop

centrifuge (Eppendorf North America) at 1,400 x g for 2 min. Cell density and viability were determined with the hemocytometer and 0.4% trypan blue staining. The pellets were resuspended in Triton X-100 lysis buffer (25 mM sodium phosphate buffer, pH 7.0, 150 mM NaCl, 1% Triton X-100 [Sigma-Aldrich Co.]). Lysates were loaded onto 18% Tris-HCl SDS polyacrylamide gels and electrophoresed at 200 V for 65 min. EGFP fluorescence detection followed by western blotting for α-syn detection were performed as described above.

3.2.4 Sf9 Custom Media for α-Syn Production

Custom Insect Media Preparation

Algal amino acid mix and Cell Free amino acid mix were from Cambridge Isotope Laboratories, Inc. (Andover, MA). Dialyzed yeastolate and ESF NMR medium lacking amino acids, NaCl, and yeastolate were from Expression Systems, LLC. L-Glutamine, L-glutamate, L-cysteine, L-tryptophan, L-asparagine, NaCl, and NH_4Cl were from Thermo Fisher Scientific, Inc. (Fair Lawn, NJ). The ESF NMR medium has been discontinued by the manufacturer.

The custom insect medium comprised the following components at their final concentrations in ESF NMR medium: 1.5 g/L algal amino acid mix, 0.5 g/L L-glutamine, 2 g/L L-glutamate, 0.1 g/L L-cysteine, 0.1 g/L L-tryptophan, 4% (v/v) dialyzed yeastolate, and 4.3 g/L NaCl. Additional custom media were prepared as described above but with the addition of final concentrations of 0.1 g/L L-asparagine or 0.8 g/L NH_4Cl or both. Another custom medium comprised the following components at their final concentrations in ESF NMR medium: 1.5 g/L

Cell Free amino acid mix, 4% (v/v) dialyzed yeastolate, and 4.3 g/L NaCl. Custom media were prepared at 100-mL volumes.

The osmolalities of the custom media were measured with a Vapro 5520 Vapor Pressure Osmometer (Wescor, Inc.; Logan, UT) and adjusted to 350-400 mOsm with NaCl. The pH values were adjusted to 6.2-6.3 with 4 M NaOH. Using aseptic technique, the custom media were filtered through 0.22 µm PVDF filters in the Purifier Class II Biosafety Cabinet.

Sf9 Cell Growth and Viability in Custom Insect Media

Sf9 suspension cells maintained in ESF 921 media and 125-mL Erlenmeyer shake flasks were pelleted by centrifugation in the Centrific Model 228 centrifuge at 1,380 x g for 2 min and resuspended in 15 mL of custom media to a final cell density of 1 x 10^6 cells/mL. Cells were incubated at 27 °C with 100-rpm shaking in 125-mL Erlenmeyer shake flasks. Cell aliquots of 100 µL were removed about every 24 h for 6 days. Cell density and viability were determined with the hemocytometer and 0.4% trypan blue staining.

α-Syn Production in Custom Media Upon Recombinant Baculovirus Infection

Sf9 suspension cells maintained in ESF 921 media in 125-mL Erlenmeyer shake flasks were pelleted by centrifugation in the Centrific Model 228 centrifuge at 1,380 x g for 2 min and resuspended in 17 mL of each of the five custom media described above and ESF 921 media to a final cell density of 1 x 10^6 cells/mL. The suspension cultures were infected with P3 α-syn pFastBac 1 recombinant baculovirus stock (1 x 10^8 pfu/mL) at a MOI of 3 and incubated at 27 °C with 100-rpm shaking in 125-mL Erlenmeyer shake fl asks. Cell aliquots of

100 µL were removed about every 24 h p.i. for 6 days, and cell density and viability were determined with the hemocytometer and 0.4% trypan blue staining. Additional 1-mL aliquots for each cell suspension were removed every day, centrifuged in the Centrific Model 228 centrifuge at 1,380 x g for 3 min, and resuspended in 2% SDS lysis buffer. α-Syn levels from whole cell lysates were determined by western blotting as described in 3.2.3. Additional infections were performed at a MOI of 3 with P4 α-syn recombinant baculovirus at 1.8 and 3.6 x 10^8 pfu/mL. The titers were determined by using a FastPlax Titer Kit (EMD Biosciences, Inc.; San Diego, CA). The custom medium for these infections comprised the following components at their final concentrations in ESF NMR medium: 1 g/L algal amino acid mix, 0.5 g/L L-glutamine, 2 g/L L-glutamate, 0.1 g/L L-cysteine, 0.1 g/L L-tryptophan, 0.1 g/L L-asparagine, 0.8 g/L NH_4Cl, 20% (v/v) dialyzed yeastolate, and 4.5 g/L D-glucose. Fifty mL of the custom medium was prepared. Osmolality quantification, pH determination, and sterile filtering were performed as described above. Sf9 cells were resuspended in 15 mL of the custom medium to a final cell density of 2 x 10^6 cells/mL and transferred to 125-mL Erlenmeyer shake flasks. The suspension cultures were infected and incubated at 27 °C with 125-rpm shaking. Aliquots of 4 mL were removed on days 3 and 4 p.i. for detection of α-syn by western blotting as described above.

3.2.5 InsectDirect

pIEx-7 Vector Construction and Purification

Insect GeneJuice Transfection Reagent, InsectDirect pIEx-7 enterokinase ligation-independent cloning (Ek/LIC) Vector Kit (including NovaBlue GigaSingles

competent cells and T4 DNA polymerase LIC-qualified), and Insect Robopop Ni-NTA (nickel-nitrilotriacetic acid) His-Bind Purification Kit were from EMD Biosciences, Inc. The IE1 promoter and AS S-tag primers used for DNA sequence analysis are shown in Table 3.1 and were ordered from Sigma-Aldrich Co. The EndoFree Plasmid Maxi Kit was from Qiagen, Inc.

Primers for PCR amplification of human wild-type cytochrome *c*, human wild-type α-syn, *Aequorea coerulescens* GFP (AcGFP), ubiquitin, and human heat shock protein 1 (HSP27) genes were designed according to directions in the pIEx-7 Ek/LIC Vector Kit and ordered from Sigma-Aldrich Co. (Table 3.2). PCR reactions were performed with the primers and template vectors (Table 3.2) by using Vent DNA polymerase and the Eppendorf Mastercycler. PCR reaction products were separated on 1% (w/v) agarose gels electrophoresed at 100 V for 50 min. The amplified Ek/LIC gene inserts were excised and purified with the gel extraction kit. Ek/LIC PCR amplified inserts were treated with T4 DNA polymerase and annealed into the pIEx-7 vector according to the manufacturer's directions. The annealing reactions were transformed into NovaBlue GigaSingles competent cells, streaked onto LB_{AMP} plates, and incubated overnight at 37 °C. Colonies were picked from the plates and used to inoculate 5-mL liquid cultures of LB_{AMP} and grown for 12-16 h at 37 °C with shaking at 250 rpm. Plasmid DNA was isolated by using the miniprep kit. The DNA was checked by restriction enzyme mapping and DNA sequence analysis using pIEx-7 sequencing primers (Table 3.1) to verify insertion of the target genes into the pIEx-7 vector.

Positive colonies were used to inoculate 5-mL liquid cultures of LB_{AMP}. The cultures were grown for 8 h at 37 °C with shaking at 250 rpm and used (0.5 mL) to inoculate 100 mL of fresh LB_{AMP} in 500-mL baffled flasks. These cultures were grown overnight at 37 °C with shaking at 250 rpm, transferred to 250-mL centrifuge bottles, and centrifuged in the Sorvall RC-5B centrifuge with a Sorvall GSA rotor (Thermo Fisher Scientific, Inc.; Waltham, MA) at 6,000 rpm for 10 min at 4 °C. Plasmid DNA was purified from the cell pellets by using the EndoFree Plasmid Maxi Kit and resuspended in 0.5 mL of Tris ethylenediaminetetraacetic acid (EDTA) buffer (TE buffer). DNA yields were generally 1-2 µg/µL.

InsectDirect Sf9 Recombinant pIEx-7 Transfection and Protein Production

Ten-mL aliquots of Sf9 suspension cultures were transfected with pIEx-7 purified DNA containing the gene of interest by using the Insect GeneJuice Transfection Reagent according to the manufacturer's directions. Suspension cultures were incubated at 27 °C with shaking at 150 rpm for 48 h. Suspension cultures were lysed and the recombinant proteins purified with the Insect Robopop Ni-NTA His-Bind Purification Kit according to the manufacturer's directions. The purified recombinant proteins were loaded onto 18% Tris-HCl SDS polyacrylamide gels and electrophoresed at 200 V for 65 min. Protein yields were determined in triplicate by using a Pierce Modified Lowry Protein Assay Kit (Thermo Fisher Scientific, Inc.; Rockford, IL) with cytochrome *c* as a standard.

3.2.6 *Drosophila* S2 Cell Infection

pUAST-AcGFP and pUltraBac-1 Vector Construction

The AcGFP gene was PCR amplified (primers are presented in Table 3.3) by using the template pAcGFP1 vector (Clontech; Mountain View, CA), Vent DNA polymerase, and the Eppendorf Mastercycler. PCR reaction products were separated on 1% (w/v) agarose gels electrophoresed at 100 V for 55 min. The amplified AcGFP gene was excised and purified by using the gel extraction kit. The AcGFP PCR product and the pUAST vector containing the *Drosophila* heat shock protein 70 (hsp70) promoter (*Drosophila* Genomics Resource Center; Bloomington, IN) were digested with *KpnI* and *XbaI* (New England Biolabs, Inc.). The reaction products were separated by agarose gel electrophoresis and the relevant species isolated by gel extraction. The digested products were mixed and ligated with T4 DNA ligase overnight at 16 °C a nd transformed into DH5α *E. coli*. Cells were streaked onto LB_{AMP} plates and incubated overnight at 37 °C. Colonies were picked from the plates and used to inoculate 5-mL liquid cultures of LB_{AMP} and grown for 12-16 h at 37 °C with shaking at 250 rpm. pUAST-AcGFP DNA was isolated with the miniprep kit and checked by restriction enzyme mapping and DNA sequence analysis to identify constructs with the AcGFP gene inserted into the hsp70 promoter multiple cloning site (MCS) of the pUAST vector.

Mutagenesis reactions (primers are presented in Table 3.4) to introduce *HindIII* and *BamHI* restriction sites into the pUltraBac-1 vector were performed by using the site-directed mutagenesis kit. The mutagenesis reactions were

transformed into XL 10-Gold *E. coli* cells (Agilent Technologies, Inc.) and grown overnight in 5 mL of LB_{AMP} at 37 °C with 250-rpm shaking. Plasmid DNA was isolated with the miniprep kit. The identities of the mutants were confirmed by restriction enzyme mapping and DNA sequence analysis.

pUAST-AcGFP and pUltraBac-1 were digested with *HindIII* and *BamHI* (New England Biolabs, Inc.). The reaction products were separated by agarose gel electrophoresis and the relevant species isolated by gel extraction as described above. The hsp70-AcGFP fragment and the digested pUltraBac-1 vector were mixed and ligated with T4 DNA ligase overnight at 16 °C and transformed into DH5α *E. coli*. Cells were streaked onto LB plates containing 100 µg/mL ampicillin and 7 µg/mL gentamicin ($LB_{AMP/GEM}$) and incubated overnight at 37 °C. Positive colonies were picked and used to inocu late 5-mL liquid cultures of $LB_{AMP/GEM}$. The cultures were grown for 12-16 h at 37 °C with 250-rpm shaking. pUltraBac-1-hsp70-AcGFP DNA was isolated by using the miniprep kit. The identity of the constructs was confirmed by restriction enzyme mapping and DNA sequence analysis.

pUltraBac-1-hsp70-AcGFP Recombinant Baculovirus Generation

pUltraBac-1-hsp70-AcGFP DNA was transformed into DH10Bac *E. coli* cells to allow transposition into the DH10Bac bacmid. The cells were streaked onto LB plates containing 50 µg/mL kanamycin, 7 µg/mL gentamicin, 10 µg/mL tetracycline, 100 µg/mL Bluo-gal (Invitrogen Corp.), and 40 µg/mL IPTG for blue/white selection. Positive colonies were selected and the recombinant bacmid purified as described by the manufacturer.

A transfection mixture comprising 15 µg of pUltraBac-1-hsp70-AcGFP recombinant bacmid and 70 µL of Cellfectin II (Invitrogen Corp.) was added to a 75-cm^2 cell culture flask seeded with 5 x 10^6 Sf9 cells. After 5 h, the DNA:lipid complex was removed and 15 mL of ESF 921 medium was added. The flask was incubated at 27 °C for 72 h. The culture medium was collected and centrifuged in the Centrific Model 228 centrifuge at 1,380 x g for 2 min to remove cell debris. The supernatant is the P1 baculovirus stock. The P1 stock was titered in a 96-well plate based on the number of green-fluorescing Sf9 cells from the $TCID_{50}$. The P1 stock was added at a MOI of 0.1 to 10 mL of Sf9 suspension cells at 2 x 10^6 cells/mL in ESF 921 medium. The medium was collected 48 h p.i. to generate the P2 stock and was titered as described above. The P2 stock was used to infect five 50-mL Sf9 suspension cultures (2 x 10^6 cells/mL) at a MOI of 0.1 to generate the P3 stock. The cell suspensions were harvested 48 h p.i. and centrifuged in the benchtop centrifuge at 500 x g for 5 min to pellet the cells. The medium was sterile filtered with 0.22 µm PVDF filters into Nalgene Oak Ridge high speed centrifuge tubes and centrifuged in the Sorvall RC-5B centrifuge with the Sorvall SS-34 rotor at 19,000 rpm for 2 h at 4 °C. The medium was decanted and the baculovirus pellet resuspended in a total of 1.2 mL PBS. The titer of the P3 stock was determined to be 5.85 x 10^9 pfu/mL as described above. All baculovirus stocks were stored at 4 °C.

Drosophila *S2 Cell Infection*

S2 cells were seeded in a 24-well tissue culture plate (Corning, Inc. - Life Sciences) at 1 x 10^6 cells/well and infected with P3 pUltraBac-1-hsp70-AcGFP

recombinant baculovirus at MOIs of 1, 10, 50, 100, and 200. Cells were harvested 36 h p.i., and cell viability was checked with the hemocytometer and 0.4% trypan blue staining. Cells were centrifuged in the Centrific Model 228 centrifuge at 1,380 x g for 2 min and resuspended in 2% SDS lysis buffer. Lysates were loaded onto 18% Tris-HCl SDS polyacrylamide gels and electrophoresed at 200 V for 50 min. AcGFP fluorescence detection, gel staining, and gel destaining were performed as described in 3.2.3.

3.2.7 EGFP Production in Sf9 Cells with pUltraBac-1-hsp70-AcGFP

Custom Insect Media

The custom algal medium comprised the following components at their final concentrations in ESF NMR medium: 3 g/L algal amino acid mix, 2 g/L L-glutamine, 0.4 g/L L-glutamate, 0.1 g/L L-cysteine, 0.1 g/L L-tryptophan, 0.1 g/L L-asparagine, 7.3 g/L D-glucose, and 20% (v/v) dialyzed yeastolate. Forty mL of this custom medium was prepared. The twenty amino acid (20 A.A.) custom medium comprised the following components at their final concentrations in ESF NMR medium: 3 g/L Cell Free amino acid mix, 7.3 g/L D-glucose, and 20% (v/v) dialyzed yeastolate. Forty mL of this custom medium was prepared. Variable yeastolate custom algal media were prepared as described above for the custom algal medium but with the addition of final concentrations of 0%, 2%, or 10% (v/v) dialyzed yeastolate and 2.16 g/L mannose, 3.61 g/L glucose, 1.20 g/L galactose, 0.38 g/L fructose, and 1.75 g/L sucrose instead of 7.3 g/L glucose to more accurately mimic the sugar composition of commercial ESF 921 media (113). Ten mL of each of the three variable yeastolate media were prepared. “No

yeastolate, no tryptophan" custom algal medium was prepared as described above for the variable yeastolate custom algal media but without addition of dialyzed yeastolate and tryptophan, respectively. Ten mL of this custom medium was prepared. Amino acid custom media containing all 20 ribosomally encoded amino acids (20 individual A.A.), 19 individual A.A., or no A.A. comprised individual amino acids and sugars (final concentrations are given in Table 3.5) in ESF NMR medium. The amino acid concentrations for 20 and 19 individual A.A. media mimicked the composition of 3 g/L algal amino acid mix. Twenty mL of each medium was prepared. The individual amino acids are from Sigma-Aldrich Co. The sugars are from Thermo Fisher Scientific, Inc. (Fair Lawn, NJ). Materials and further information on preparation of custom media, including osmolality quantification and sterile filtering, are described in 3.2.4.

EGFP Production in Commercial and Custom Insect Media

Routinely maintained Sf9 suspension cultures in ESF 921 media were pelleted by centrifugation in the Centrific Model 228 centrifuge at 1,380 x g for 2 min. Cell pellets were resuspended to a final density of 2 x 10^6 cells/mL in ESF 921 medium, custom algal medium with or without 5% ESF 921 medium, 20 A.A. custom medium with or without 5% ESF 921 medium, variable yeastolate custom algal media, "no yeastolate, no tryptophan" custom algal medium, 20 individual A.A., 19 individual A.A., and no A.A. custom media. Ten mL of each Sf9 suspension culture were added to 125-mL Erlenmeyer shake flasks and infected with P3 pUltraBac-1-hsp70-AcGFP recombinant baculovirus at a MOI of 1 (MOI of 3 for 20 individual A.A., 19 individual A.A., and no A.A. custom media).

Baculovirus titers were determined from the $TCID_{50}$. Cell aliquots of 100 µL were removed through 72 h p.i. to determine cell density and viability with the hemocytometer and 0.4% trypan blue staining. Cells were centrifuged in the Centrific Model 228 centrifuge at 1,380 x g for 2 min and resuspended in 1% lysis buffer (50 mM Tris-HCl, pH 7.8, 150 mM NaCl, 1% Triton X-100). Lysates were loaded onto 18% Tris-HCl SDS polyacrylamide gels and electrophoresed at 200 V for 50 min. EGFP fluorescence detection was performed as described in 3.2.3.

Sf9 Cell Viability in the NMR Tube

Routinely maintained Sf9 suspension cultures in ESF 921 medium were centrifuged in the Centrific Model 228 centrifuge at 1,380 x g for 2 min. The medium was decanted and the cell pellet (200 µL) added to a S500 NMR tube (Norell, Inc.; Landisville, NJ) at room temperature. Cell viability was checked after 3 h, 6 h, 21 h, and 25 h by removing 100 µL aliquots of cells and counting on the hemocytometer using a Live/Dead Viability/Cytotoxicity kit (Invitrogen Corp.).

5-Fluorotryptophan Custom Algal Amino Acid Media Preparation

5-Fluorotryptophan (5-FW; Sigma-Aldrich Co.) was used to prepare ^{19}F-labeled custom algal medium as described in 3.2.7 except that 0.1 g/L tryptophan was replaced by 0.1 g/L 5-FW, and 7.3 g/L glucose was replaced by 2.16 g/L mannose, 3.61 g/L glucose, 1.20 g/L galactose, 0.38 g/L fructose, and 1.75 g/L sucrose to more accurately mimic the sugar composition of commercial ESF 921 media (113). Forty mL of this custom medium was prepared.

Sf9 EGFP Production in 5-FW Custom Algal Media

Routinely maintained Sf9 suspension cultures in ESF 921 media were pelleted by centrifugation in the Centrific Model 228 centrifuge at 1,380 x g for 2 min. Cell pellets were resuspended to a final density of 2 x 10^6 cells/mL in 5-FW custom algal media (9 mL) and infected with P3 pUltraBac-1-hsp70-AcGFP recombinant baculovirus at a MOI of 3. Baculovirus titers were determined from the $TCID_{50}$. Cell aliquots of 100 µL were removed prior to infection and 24 h, 48 h, and 72 h p.i. to determine the cell density and the viability with the hemocytometer and 0.4% trypan blue staining. To detect EGFP, cells were centrifuged in the Centrific Model 228 centrifuge at 1,380 x g for 2 min and resuspended in 1% lysis buffer. Lysates were loaded onto 18% Tris-HCl SDS polyacrylamide gels and electrophoresed at 200 V for 50 min. EGFP fluorescence detection was performed as described in 3.2.3.

Production of 5-FW-Labeled EGFP in Sf9 Cells for In-Cell NMR

A 20-mL Sf9 suspension culture at 3.8 x 10^6 cells/mL in ESF 921 medium was infected with P3 pUltraBac-1-hsp70-AcGFP recombinant baculovirus at a MOI of 1 and placed into a 125-mL Erlenmeyer shake flask. Baculovirus titers were determined from the $TCID_{50}$. Cells were incubated at 27 °C with 150-rpm shaking. After 14 h, the cells were pelleted by centrifugation in the Centrific Model 228 centrifuge at 1,380 x g for 2 min. Pellets were resuspended in 20 mL of 5-FW custom algal media. Cell density and viability were determined 48 h p.i. with the hemocytometer and 0.4% trypan blue staining. After 48 h p.i., cells were pelleted by centrifugation in the Centrific Model 228 centrifuge at 1,380 x g for 2

min. Cells were washed with ESF 921 media and centrifuged again at 1,380 x g for 2 min to repellet cells. A final concentration of 10% (v/v) D_2O (Cambridge Isotope Laboratories, Inc.) was added to the cell pellet. The cell slurry was pipetted into a Shigemi NMR tube (Shigemi, Inc.; Allison Park, PA). ^{19}F spectra were acquired as described in 3.2.10. Cells were lysed with 1% lysis buffer and centrifuged in the benchtop centrifuge at 17,000 x g for 30 min. ^{19}F spectra of the cleared lysate were acquired as described in 3.2.10.

3.2.8 Sf9 EGFP Production in ^{15}N-Enriched Custom Media

$^{15}NH_4Cl$ custom medium comprised components at the final concentrations shown in Table 3.5 under the heading ESF NMR medium. $^{15}NH_4Cl$ was from Cambridge Isotope Laboratories, Inc. All other medium components, osmolality quantification, and sterile filtering are described in 3.2.4.

In-Cell NMR

Routinely maintained Sf9 suspension cultures in ESF 921 media were centrifuged in the Sorvall RC-5B centrifuge with the SS-34 rotor at 1,500 rpm for 10 min at 25 °C to pellet the cells. The cells were r esuspended in the $^{15}NH_4Cl$ custom medium to a final density of 2.9 x 10^6 cells/mL in 125-mL Erlenmeyer shake flasks (50-mL culture volumes). Cultures were infected with P3 pUltraBac-1-hsp70-AcGFP recombinant baculovirus at a MOI of 1. Baculovirus titers were determined from the $TCID_{50}$. Cells were incubated at 27 °C with 150-rpm shaking. Cell density, viability, and fluorescence were determined under the microscope 48 h p.i. with the hemocytometer and 0.4% trypan blue staining. After 48 h p.i., the cells were pelleted by centrifugation in the Sorvall RC-5B centrifuge

with the SS-34 rotor at 1,500 rpm for 10 min at 25 °C. The cells were washed with PBS and pelleted again. D_2O was added to attain a final D_2O concentration of 10% (v/v), and the cell slurry was pipetted into the S500 NMR tube. HSQC (60) spectra were acquired as described in 3.2.10. Cells were lysed with 1% lysis buffer and centrifuged in the benchtop centrifuge at 17,000 x g for 30 min. HSQC spectra of the lysate were acquired as described in 3.2.10.

In Vitro *NMR*

Sf9 cell lysates were centrifuged in the Sorvall RC-5B centrifuge with the Sorvall SS-34 rotor at 16,000 rpm for 30 min at 15 °C to remove cell debris. Lysates were loaded onto a Superdex 75 size exclusion column (GE Healthcare; Piscataway, NJ), which was equilibrated in and eluted with 20 mM sodium phosphate buffer, pH 6.0. Fractions were loaded onto 18% Tris-HCl SDS polyacrylamide gels and electrophoresed at 200 V for 50 min. Fluorescent fractions were pooled and concentrated by using a 3,000 molecular weight cutoff (MWCO) centrifugal filter unit (EMD Millipore). The EGFP concentration was checked by using a Nanodrop ND-1000 Spectrophotometer (Thermo Fisher Scientific, Inc.; Wilmington, DE). D_2O was added to attain a final D_2O concentration of 10% (v/v) for samples with and without a final concentration of 200 mM NaCl. HSQC spectra were acquired as described in 3.2.10.

3.2.9 Sf9 α-Syn and α-Syn-EGFP Production in ^{15}N-Enriched Custom Media

The $^{15}NH_4Cl$ custom medium is described in 3.2.8. ^{15}N-enriched custom IPL-41 media were prepared according to the chemically defined IPL-41 formulation available from Invitrogen Corp., but without the addition of amino

acids (Table 3.6). Chemically Defined Lipid Concentrate, 100x (Invitrogen Corp.) was added to the IPL-41 media at a final concentration of 1% (v/v), and ^{15}N-enriched algal amino acid mix (Cambridge Isotope Laboratories, Inc.) was added to attain a final concentration of 1 g/L. The ^{15}N-enriched algal amino acid mix lacks cysteine, tryptophan, glutamine, and asparagine. These amino acids were added to the custom media in unenriched form (Table 3.6). Yeastolate Ultrafiltrate, 50x (Invitrogen Corp.) was added to half of the prepared media at a final concentration of 2% (v/v). The pH values and osmolalities were adjusted to 6.2 and 375 mOsm, respectively, and the media were sterile filtered as described in 3.2.4. The individual amino acids are from Sigma-Aldrich Co. The sugars are from Thermo Fisher Scientific, Inc. (Fair Lawn, NJ). Vitamin B12 is from MP Biomedicals, LLC. $FeSO_4 \cdot 7H_2O$ is from Allied Chemical Corp. (Morristown, NJ). Unless otherwise noted, the ^{15}N-enriched custom IPL-41 media components are from Thermo Fisher Scientific, Inc. (Fair Lawn, NJ).

ESF insect medium lacking L-phenylalanine, L-isoleucine, L-leucine, L-valine, L-methionine, L-tyrosine, and L-threonine (Expression Systems, LLC) was used to prepare a custom NMR medium containing ^{15}N-enriched phenylalanine. All seven amino acids absent from the ESF media were added at the concentrations present in commercial ESF 921 media (see Table 3.5 under $^{15}NH_4Cl$ Custom Media). L-Phenylalanine was added in ^{15}N-enriched form (Cambridge Isotope Laboratories, Inc.). The unenriched amino acids were from Thermo Fisher Scientific, Inc. (Fair Lawn, NJ). The pH of the medium was

adjusted to 6.2 followed by sterile filtering as described in 3.2.4. The osmolality of the medium (423 mOsm), determined as in 3.2.4, was not adjusted.

In-Cell NMR

Routinely maintained Sf9 suspension cultures in ESF 921 media were centrifuged in the Sorvall RC-5B centrifuge with the SS-34 rotor at 1,500 rpm for 10 min at 25 °C to pellet the cells. Cells were resuspe nded in 30 mL of $^{15}NH_4Cl$ custom medium to a final density of 5.3 x 10^6 cells/mL in 125-mL Erlenmeyer shake flasks. Cells were resuspended in 40 mL of ^{15}N-phenylalanine custom medium or ^{15}N-enriched custom IPL-41 media to final densities of 4.5 x 10^6 cells/mL in 125-mL Erlenmeyer shake flasks. Two cultures in $^{15}NH_4Cl$ custom media were infected with either P3 α-syn pUltraBac-1 or P3 α-syn-EGFP pUltraBac-1 recombinant baculovirus at a MOI of 1. Two cultures in ^{15}N-phenylalanine custom media were infected with either P3 α-syn pUltraBac-1 or P3 α-syn-EGFP pUltraBac-1 recombinant baculovirus at the same MOI. One culture in ^{15}N-enriched custom IPL-41 medium was infected with P3 α-syn-EGFP pUltraBac-1 recombinant baculovirus at that MOI. The cultures were incubated at 27 °C with 150-rpm shaking. Cell density, viability, and fluorescence were determined under the microscope 43-44 h p.i. with the hemocytometer and 0.4% trypan blue staining. After 43-44 h p.i., the cells were pelleted by centrifugation in the Sorvall RC-5B centrifuge with the SS-34 rotor at 1,500 rpm for 10 min at 25 °C. The supernatants were decanted, and D_2O was added to the pellets to attain a final D_2O concentration of 10% (v/v). The cell slurries were pipetted into S500 NMR tubes, and HSQC spectra were acquired as described in 3.2.10. After

acquisition, cells were lysed with Triton X-100 lysis buffer (25 mM sodium phosphate buffer, pH 7.0, 150 mM NaCl, 1% Triton X-100) and centrifuged in the benchtop centrifuge at 17,000 x g for 30 min. HSQC spectra of the lysates were acquired as described in 3.2.10.

3.2.10 NMR

^{19}F spectra were acquired at 37 °C on a Varian Inova 60 0 MHz spectrometer (Palo Alto, CA) equipped with a 5 mm ^{19}F {^{1}H} z-gradient probe. The spectra comprised 128 to 2048 transients with a 30 kHz sweep width and a 2 s acquisition delay. Trifluoroethanol at 0 ppm was used to reference the ^{19}F chemical shifts.

^{1}H -^{15}N HSQC spectra were acquired at 25 °C on the Varian I nova 600 MHz spectrometer equipped with a triple resonance HCN probe. The ^{1}H dimension sweep width was 8401 Hz and comprised 1024 complex points. The ^{15}N dimension sweep width was 2200 Hz and comprised 64 complex points. The data were processed with NMRPipe (58) and NMRView (59).

3.3 Results and Discussion

3.3.1 EGFP Detection

As proof of technique, recombinant baculovirus was generated by using the pFastBac Dual vector (Fig. 3.2) containing the EGFP gene. The production of EGFP, a fluorescent protein, provides a convenient means of detecting baculovirus infection (114). Infected Sf9 cells were visualized by laser scanning confocal microscopy (Fig. 3.3). Sf9 cells emitted green fluorescence, indicating production of EGFP in the cells (Fig. 3.3*a, b*). Although the fraction of cells

fluorescing green was not quantified, the majority of cells fluoresce green. The total cellular proteins were subjected to sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE; Fig. 3.4). No bands corresponding to EGFP were detected by Brilliant Blue staining (Fig. 3.4*a*). However, fluorescence imaging shows bands corresponding to EGFP in the whole cell lysates (Fig. 3.4*b*). Visual confirmation of EGFP production in infected Sf9 cells provided the technical expertise and confidence required to attempt α-syn production, a protein that cannot be followed by fluorescence.

3.3.2 α-Syn Detection

The human wild-type α-syn gene was inserted under the control of the *Autographa californica* multiple nuclear polyhedrosis virus (AcMNPV) polyhedrin (PH) promoter in the pFastBac 1 vector (Fig. 3.5). α-Syn pFastBac 1 recombinant baculovirus was generated and used to infect Sf9 cells in suspension culture. Western blot results (Fig. 3.6) show the presence of α-syn in the whole cell lysates.

The estimated concentration of α-syn in 1 L of cell suspension at 2 x 10^6 cells/mL 1-8 days p.i. (Table 3.7) was calculated as follows. Cell aliquots were pelleted and resuspended in lysis buffer to a final density of 1 x 10^7 cells/mL. Multiplication of the lysate cell density by the volume of lysate loaded onto SDS polyacrylamide gels yields the number of cells in each lane. The amount of α-syn present in the western blots was determined with the VersaDoc software by comparison of lysate band intensities to the intensities of known amounts of loaded α-syn control (1.9 µg per lane). Dividing the amount of α-syn present in

the lysate by the number of cells yields the amount of α-syn per cell. The amount of α-syn per cell was divided by the molecular weight of α-syn (14.4 kDa) to yield the number of moles of α-syn per cell. Multiplication of the number of moles by 2 x 10^9 cells (2 x 10^6 cells/mL x 1000 mL) provides an estimation of the α-syn concentration in a 1-L Sf9 cell suspension at 2 x 10^6 cells/mL.

The estimated intracellular concentration of α-syn was calculated as follows. The amount of α-syn in a single cell was divided by the molecular weight of α-syn to yield the number of moles of α-syn per cell. The average Sf9 cell diameter is 15 µm (89), so the volume of a single cell was calculated to be $4/3*\pi(7.5 \times 10^{-6}\ m)^3 = 1.8 \times 10^{-15}\ m^3$. Conversion from m^3 to L yields 1.8 pL for the volume of a single Sf9 cell. The number of moles of α-syn per cell was divided by the volume of a single Sf9 cell to yield the estimated intracellular concentration of α-syn in Molar.

The maximum α-syn protein concentration in the cell suspension at a cell density of 2 x 10^6 cells/mL is 3.4 ±0.2 µM, which is achieved on day 4 p.i. (Table 3.7). It must be noted, however, that cell viability on day 4 p.i. is 57 ±3%, too low for in-cell NMR. Previous work shows that insect cells may be grown in suspension culture at 3 x 10^7 cells/mL while maintaining >90% viability (115). This higher Sf9 cell density corresponds to maximum α-syn protein concentrations of ~50 µM in cell suspension. Protein concentrations in the 10^2 µM range, however, are required for the acquisition of high resolution NMR data (39, 45). Although the α-syn concentration is estimated to be too low for NMR at 3 x 10^7 cells/mL, Sf9 cell densities can be increased by pelleting the cells through

centrifugation. Pelleting Sf9 cells from suspension culture and counting by hemocytometer yields 2.6 x 10^8 cells/mL, the maximum achievable density of Sf9 cells. The cell pellet can be placed in the NMR tube to achieve the maximum concentration of α-syn for NMR experiments. Estimation of the concentration of α-syn in the NMR tube was determined by multiplying the estimated α-syn concentrations at 2 x 10^6 cells/mL by a factor of 130 (2.6 x 10^8 cells/mL / 2 x 10^6 cells/mL = 130). At 4 days p.i., the maximum α-syn concentration expected in the NMR tube is 447 ±31 μM, which is sufficient for in-cell NMR (Table 3.7). These results demonstrate that the α-syn protein concentration achieved in suspension culture has the potential to yield acceptable concentrations for protein in-cell NMR when cells are pelleted and placed in the NMR tube. The next step is to isotopically enrich the protein.

3.3.3 Custom Sf9 Insect Media

Commercially available media for isotopic enrichment of recombinant proteins in insect cells are prohibitively expensive (Bioexpress-2000 media ^{15}N 98% from Cambridge Isotope Laboratories, Inc. is $7,500 per liter). As a result, only a single study using this media to uniformly label a recombinant protein in Sf9 cells has been published (116). Therefore, attempts were made to produce custom media by using a basal medium and adding less expensive growth components. Walton et al. developed a custom medium for a fraction of the cost of commercially available media (110). The medium utilizes a relatively cheap amino acid mixture derived from algae to supply most of the ^{15}N-enriched amino acids. By infecting Sf9 cells in this custom medium with recombinant baculovirus

to produce the avian Thy-1 protein, the investigators observed 60% of the production level achieved by using commercial ^{15}N insect media. Furthermore, 80% of the anticipated resonances for Thy-1 were observed in ^{1}H-^{15}N HSQC spectra of the purified protein.

Custom insect media were prepared by following the published protocol of Walton et al. (110). Custom media were initially prepared in unenriched form to assess cell growth and viability. The algal amino acid mix lacks four amino acids: glutamine, asparagine, cysteine, and tryptophan. These amino acids were added to the custom algal media preparations as described in 3.2.4. Custom media prepared by Walton et al. did not call for the addition of asparagine. As a precaution in this study, asparagine was included in some preparations. ^{15}N-glutamine is one of the most expensive amino acids, and it has been shown that insect cells can produce glutamine from glutamate and NH_4Cl (117). Therefore, custom algal media were also prepared with the additions of glutamate, NH_4Cl, asparagine, cysteine, and tryptophan as described in 3.2.4. Later, addition of $^{15}NH_4Cl$ will allow isotopic enrichment of glutamine. Additionally, Cell Free amino acid mix, containing all 20 amino acids, was used to prepare custom media without the need to supplement with additional amino acids as described in 3.2.4.

Cells were grown in custom media, and aliquots (100 µL) were removed about every 24 h for 6 days to determine the cell density and viability. Figures 3.7 and 3.8 show Sf9 cell growth and viability data in custom media, respectively. Sf9 cells in custom media grow from an initial density of 1 x 10^6 cells/mL to 2-3 x 10^6 cells/mL after 6 days (142 h), indicating that cells undergo about 1 round of cell

division (Fig. 3.7). Sf9 cell viability remains greater than 80% through 5 days (117 h) of incubation in all custom media compared to 91% in commercial ESF 921 medium (Fig. 3.8). Although cell growth is slowed in custom media, cell viability remains acceptable for in-cell NMR.

Next, suspension cultures of Sf9 cells in custom media were infected with α-syn pFastBac 1 recombinant baculovirus. Aliquots (100 µL) were removed about every 24 h p.i. for 6 days to determine cell density and viability. Aliquots (1 mL) were removed about every 24 h p.i. for 6 days to determine α-syn production. Infected Sf9 cells in custom media grew from an initial density of 1 x 10^6 cells/mL to 1.5-3 x 10^6 cells/mL 6 days (140 h) p.i. (Fig. 3.9*a*). Cell viability remained ≥76% 3 days (68 h) p.i. in custom media (Fig. 3.9*b*).

Western blots of whole cell lysates detected virtually no α-syn (data not shown). α-Syn production may have been low due to the unknown α-syn baculovirus titer or a sugar or yeastolate limitation in the media. Custom medium was prepared with larger amounts of yeastolate and glucose. Sf9 cells were infected in this new custom medium (see 3.2.4) at a MOI of 3 with α-syn pFastBac 1 recombinant baculovirus of known titer. However, western blots of whole cell lysates from 72 h and 96 h p.i. show that almost no α-syn (Fig. 3.10) was produced. The estimated α-syn concentration is 0.1 µM 96 h p.i. at a cell density of 2 x 10^6 cells/mL. This level of α-syn production is insufficient for in-cell NMR even when pelleting the cells into the NMR tube (2.6 x 10^8 cells/mL, 13 µM α-syn). Furthermore, cell viability after just 72 h p.i. is 55%, unacceptable for in-

cell NMR. These results demonstrate that Sf9 production of α-syn by baculovirus infection is insufficient for in-cell NMR using the developed custom media.

3.3.4 InsectDirect

After the unsuccessful attempts to produce sufficient α-syn and maintain viability in infected Sf9 cells for in-cell NMR, another expression system was employed. The InsectDirect system (EMD Biosciences, Inc.; Madison, WI) produces high levels of recombinant protein without the need to create a recombinant baculovirus (118). Results from the company show that 8 mg of HSP27 was produced in a 100-mL Sf9 suspension culture (118). This production level corresponds to a HSP27 concentration of 30 µM, enough for in-cell NMR when pelleting the cells into the NMR tube (calculated as described in 3.3.2 using 27 kDa as the molecular weight of HSP27). The InsectDirect vectors employ a homologous region 5 (hr5) enhancer and immediate early 1 (IE1) promoter combination to recruit endogenous insect cell transcription machinery to directly produce the protein of interest (119). This shortcut avoids the cytotoxic effects of baculovirus infection. Additionally, the vectors facilitate ligation-independent cloning (LIC) to insert the desired genes (120), enabling quick screening of multiple target genes. The system is also fast; vector construction to protein production can be accomplished in one week.

The rapid generation of vectors and quick turnaround to protein production allowed the testing of several target genes. The cytochrome *c*, α-syn, AcGFP, ubiquitin, and HSP27 genes were cloned into InsectDirect pIEx-7 Ek/LIC vectors (Fig. 3.11) and purified to µg quantities. The recombinant pIEx-7 vectors were

transfected into Sf9 cells in suspension, and the cultures were harvested after 48 h for recombinant protein purification. The pIEx-7 Ek/LIC vector encodes a histidine tag (His-tag) at the N-terminus of the protein of interest, allowing easy purification using the Insect Robopop Ni-NTA His-Bind Purification Kit.

Protein yields were determined by using the Lowry assay (121). All recombinant proteins are detectable by this method, but the estimated yields are too low for in-cell NMR (Table 3.8). AcGFP and HSP27 are the two highest yielding proteins at 1.5 mg/L (0.06 µM) and 2.7 mg/L (0.10 µM) at 2 x 10^6 cells/mL, respectively, and are detectable by SDS-PAGE (Figures 3.12 and 3.13). However, α-syn, cytochrome *c*, and ubiquitin were not detected by SDS-PAGE with Brilliant Blue staining. Gel results show that the major HSP27 product is an aggregated form that runs around 60 kDa (Fig. 3.12), too large for in-cell NMR detection with the ^{1}H-^{15}N HSQC experiment (122, 123). AcGFP was detectable by fluorescence (Fig. 3.13*b*) and ran on the gel at 27 kDa, but a non-fluorescent higher molecular weight band was also present (Fig. 3.13*a*). Since recombinant protein concentrations in the 10^2 µM range are required for in-cell NMR, AcGFP and HSP27 yields are not sufficient for in-cell NMR even when Sf9 cells are maximally compacted into a NMR tube (8 µM for AcGFP and 13 µM for HSP27, calculated as described in 3.3.2 using 27 kDa as the molecular weight for both proteins). Furthermore, transfected Sf9 cell viabilities after 48 h were <50% (Table 3.8), unsuitable for in-cell NMR, compared to 81% for mock infected cells. These results show that the InsectDirect system neither produced sufficient

quantities of recombinant protein nor sustained cell viability for in-cell NMR experiments.

3.3.5 *Drosophila* S2 Cell Infection

Since attempts to produce sufficient levels of recombinant protein and maintain cell viability for in-cell NMR were unsuccessful, an alternative strategy was employed. Baculoviruses are capable of infecting not only lepidopteran cells but also nonnative host cells (124-130). To drive production of the protein of interest, the virus is modified to contain a promoter that is active in the nonnative host cell (131). Because the baculovirus is present in a nonnative host cell, the virus is incapable of replicating, and thus the cytotoxic effects and cell lysis associated with baculovirus infection are avoided.

Drosophila S2 cells were selected as the nonnative host for a modified recombinant baculovirus because these cells grow under conditions similar to those used for Sf9 cells (room temperature ambient air conditions), and these cells are relatively easy to culture (132). Lee et al. showed infection of *Drosophila* S2 cells with a recombinant baculovirus containing EGFP under the control of the *Drosophila* heat shock protein 70 (hsp70) promoter produced 3.2 mg/L of EGFP at 8 x 10^6 cells/mL and maintained 93% cell viability 72 h p.i. (130). S2 cells occupy about one-fifth the volume of Sf9 cells (130), so it is expected that the cells could be placed in a NMR tube at densities of ~10 x 10^8 S2 cells/mL (maximum Sf9 density is 2.6 x 10^8 cells/mL). These S2 cell densities yield an estimated EGFP concentration of 400 µM, which is suitable for in-cell NMR.

A strategy to produce AcGFP in S2 cells was developed by using the hsp70 promoter to drive production of AcGFP in recombinant baculovirus infected S2 cells (130, 133). The AcGFP gene from the pAcGFP1 vector was cloned into the MCS of the pUAST vector (Fig. 3.14) under the control of the hsp70 promoter. The hsp70-AcGFP cassette was excised by restriction enzyme digestion and cloned into the pUltraBac-1 vector (Fig. 3.15). The pUltraBac-1 vector is a modified version of pFastBac Dual where the p10 promoter has been replaced with the late basic protein promoter (P_{Basic}) and the EGFP gene, under the control of P_{Basic} (133). EGFP production occurs sooner under P_{Basic} than p10 control, allowing faster titer determination when amplifying the recombinant baculovirus in Sf9 insect cells. *HindIII* and *BamHI* restriction sites were added to the pUltraBac-1 vector upstream of the PH promoter and downstream of the PH promoter MCS, respectively. In the modified vector, restriction enzyme digestion removes the lepidopteran specific PH promoter and replaces it with the S2 host specific hsp70 promoter-AcGFP cassette.

pUltraBac-1-hsp70-AcGFP recombinant baculovirus generation was performed in a similar manner to that described in 3.2.2 except that the P2 stock was used to infect a total culture volume of 250 mL to generate large amounts of P3 recombinant baculovirus. Because baculoviruses are incapable of replicating in nonnative host cells, recombinant protein production is based on the number of viral particles in each cell upon initial infection. Thus, it is important to infect S2 cells at a high MOI (e.g. 100) to produce high yields of recombinant protein (130). The P3 baculovirus was concentrated from the culture media by centrifugation at

high speeds as has been previously reported (133). The P3 stock titer was 5.85 x 10^9 pfu/mL, in the range expected from this concentration method.

S2 cells were seeded in 24-well tissue culture plates and infected at varying MOIs up to 200 with P3 pUltraBac-1-hsp70-AcGFP recombinant baculovirus. Cell viability was confirmed to be 95% by trypan blue staining prior to infection. After 36 h, cells were inspected under the microscope. Cell viability appeared low due to the shriveled shape of most of the cells, and no cells fluoresced green. No AcGFP was detected in fluorescence scans and with Brilliant Blue staining of SDS gels containing S2 whole cell lysates. Lack of AcGFP production appears to be tied to the low viability of S2 cells upon infection with recombinant baculovirus at high MOIs. These results are not consistent with reports showing that baculovirus infection of S2 cells has little negative effect on S2 cell viability (130).

3.3.6 EGFP Production in Sf9 Cells with pUltraBac-1-hsp70-AcGFP

Production of the reporter EGFP protein under P_{Basic} promoter control appeared to be high during baculovirus amplification of pUltraBac-1-hsp70-AcGFP in Sf9 insect cells due to the large number of fluorescent cells observed under the microscope (data not shown). To quantify EGFP production levels, Sf9 cells were infected with P3 pUltraBac-1-hsp70-AcGFP baculovirus at a MOI of 1 in commercial ESF 921 media. Cells were harvested 48 h and 72 h p.i., lysed, and electrophoresed on a SDS gel to detect fluorescence (Fig. 3.16). Examination of cells 48 h and 72 h p.i. revealed that nearly all cells fluoresce green 48 h p.i. Initial cell viability was 96% prior to infection, 85% 48 h p.i., and

54% 72 h p.i. (Table 3.9). Quantification of EGFP fluorescence from Sf9 whole cell lysates 48 h and 72 h p.i. shows that the EGFP concentration is 2.1 µM and 2.2 µM at 2 x 10^6 cells/mL, respectively (Table 3.9). Thus, EGFP production reaches its maximum 48 h p.i. when cell viability is still acceptable for in-cell NMR. EGFP concentration 48 h p.i. is estimated to be 278 µM when cells are pelleted to 2.6 x 10^8 cells/mL in the NMR tube. This EGFP concentration is sufficient for in-cell NMR. Calculations were performed as described in 3.3.2 using 27 kDa as the molecular weight of EGFP.

The pUltraBac-1 vector contains the EGFP gene under the control of the lepidopteran active P_{Basic} promoter to allow faster titering of recombinant baculoviruses in Sf9 insect cells (133). The P_{Basic} promoter produces recombinant proteins at earlier times (12 h p.i.) than the frequently used PH promoter (18 h p.i. or later), and protein production levels are comparable between the promoters (134). Earlier production of recombinant protein by the P_{Basic} promoter is advantageous for in-cell NMR because maximum protein production takes place before the cytotoxic effects associated with baculovirus infection reduce cell viability.

EGFP production and Sf9 cell viability upon pUltraBac-1-hsp70-AcGFP recombinant baculovirus infection were assessed by using custom media to reduce the cost of isotopic enrichment for NMR as discussed in 3.3.3. Sf9 cells were infected at a MOI of 1 in commercial ESF 921 media, custom algal media, 20 A.A. media, and the two custom media supplemented with 5% ESF 921 media. The 5% ESF 921 media were added to stimulate cell growth. Samples

were collected to determine cell density and viability and lysed for fluorescence detection on SDS gels 48 h and 72 h p.i. (Fig. 3.17). Nearly all cells fluoresce green 48 h p.i. under the microscope, and cells appear healthy as noted by their round shape and lack of cell debris. Cell viabilities were ≥89% in custom media 48 h p.i. as assessed by trypan blue staining (Table 3.10). EGFP concentrations 48 h and 72 h p.i. were determined as described in 3.3.2 using the fluorescence results (Fig. 3.17) and 27 kDa as the molecular weight of EGFP. At 2.6 x 10^8 cells/mL, the estimated maximum Sf9 cell density in the NMR tube, the EGFP concentrations are ≥185 µM in all custom media 48 h p.i. (Table 3.10). These EGFP concentrations are sufficient for protein in-cell NMR. Furthermore, maximum EGFP production is achieved 48 h p.i. in custom media when cell viability remains ≥89% (Table 3.10). EGFP production in algal and 20 A.A. custom media is nearly as high as production in commercial ESF 921 media. It is evident that the early protein production conferred by the P_{Basic} promoter provides an advantage for in-cell NMR; waiting until 72 h p.i. before collecting cells yields cell viabilities ≤61% in custom media, unacceptable for in-cell NMR (Table 3.10). These results are also more favorable than those obtained with the α-syn pFastBac 1 recombinant baculovirus where the α-syn gene is under the PH promoter. In that system, maximum α-syn production was not observed until about day 4 p.i. when cell viability was 57% (Table 3.7), unacceptable for in-cell NMR.

Sf9 cell viability in the NMR tube was assessed to determine the feasibility of performing in-cell NMR experiments. Sf9 suspension cultures were centrifuged

to pellet the cells, and the cells were pipetted into a NMR tube. The NMR tube was left at room temperature, and cell aliquots (100 µL) were removed after 3 h, 6 h, 21 h, and 25 h to determine cell viability. Initial Sf9 cell viability was nearly 100% and was measured to be 93%, 90%, 84%, and 13% after 3 h, 6 h, 21 h, and 25 h, respectively. These results demonstrate that Sf9 cell viability inside the NMR tube remains acceptable for in-cell NMR experiments for at least 6 h.

5-Fluorotryptophan (5-FW) was selected as the means to label EGFP with a NMR active nucleus. 5-FW is relatively inexpensive, and ^{19}F labeling should provide better sensitivity than ^{15}N enrichment (70). Labeling sites, however, are limited to tryptophan residues only. EGFP contains one tryptophan at position 57, and *in vitro* EGFP ^{19}F NMR data from protein produced in *E. coli* are available (135). Additionally, studies show that 5-FW incorporation does not affect protein structure (136). Sf9 cell viability experiments were performed to assess whether 5-FW containing media had a deleterious effect on Sf9 cell growth and viability. Sf9 suspensions in commercial ESF 921 media were pelleted by centrifugation and resuspended in 5-FW custom algal media. Trypan blue staining showed Sf9 cell viability remained high (97%) after 48 h of incubation. Next, EGFP production was tested in Sf9 infected cells using 5-FW custom algal media. Cell viability and EGFP concentrations were assessed 24 h, 48 h, and 72 h p.i. (Table 3.11). EGFP concentrations were determined as described in 3.3.2 using 27 kDa as the molecular weight of EGFP. Results show 48 h p.i. an estimated EGFP concentration of 183 µM at 2.6 x 10^8 cells/mL, high enough for in-cell NMR. Additionally, cell viability is 87% 48 h p.i., acceptable for in-cell NMR. These

results are similar to those obtained from infected Sf9 cells 48 h p.i. in custom algal media containing tryptophan (228 µM and 89% viability, Table 3.10) indicating that 5-FW exerts no harmful effects on cell viability or protein production.

Sf9 EGFP production in 5-FW custom algal media and viability in the NMR tube is acceptable for in-cell NMR experiments. A Sf9 cell infection using P3 pUltraBac-1-hsp70-AcGFP recombinant baculovirus was performed in commercial ESF 921 media. After 14 h, the suspension culture was centrifuged, and the Sf9 cell pellet was resuspended in 5-FW custom algal media. Switching to labeling media was done 14 h p.i. when EGFP production by the P_{Basic} promoter is maximized, a strategy expected to reduce background labeling of other proteins and metabolites (137). Cells were pelleted by centrifugation 48 h p.i., when sufficient EGFP for in-cell NMR is expected, and placed in a NMR tube for ^{19}F NMR acquisition of the cell slurry. Cell viability, as assessed by trypan blue staining, was 89% 48 h p.i. and nearly all cells fluoresced green under the microscope, indicating production of EGFP. The ^{19}F spectrum of the cell slurry shows a single resonance (Fig. 3.18*a*). NMR acquisition of the cleared cell lysate shows the same resonance (Fig. 3.18*b*). Comparison to the ^{19}F spectrum of 5-FW alone (data not shown) reveals that the cell slurry and lysate resonances are from 5-FW.

As EGFP production appears to be sufficient (cells fluoresce green), absence of protein ^{19}F resonances may be due to the lack of 5-FW incorporation into EGFP by the Sf9 cells. The 5-FW custom algal medium was checked for

sources of natural tryptophan. Yeastolate, an undefined mixture of amino acids, peptides, sugars, vitamins, and minerals is often added to insect cell media to promote growth (138, 139). Yeastolate may be a source of tryptophan. Therefore, 5-FW custom algal media were prepared without yeastolate. The EGFP concentration is estimated to be 190 µM 48 h p.i. at 2.6 x 10^8 cells/mL for Sf9 cells infected in custom algal media containing 0% yeastolate, and the cells remain 87% viable (Table 3.12). Addition of yeastolate to the media did not increase the production of EGFP (Table 3.12). Thus, yeastolate is not required for cell growth and EGFP production in infected Sf9 cells.

A control experiment was performed in unlabeled custom algal media without dialyzed yeastolate and tryptophan to remove all sources of tryptophan. Sf9 cells were infected with pUltraBac-1-hsp70-AcGFP recombinant baculovirus in this custom medium lacking yeastolate and tryptophan. It was anticipated that cells would be unable to grow or produce proteins, including EGFP, because Sf9 cells cannot synthesize tryptophan (140). Results show that the EGFP concentration is 215 µM at 2.6 x 10^8 cells/mL and Sf9 cells are 96% viable 48 h p.i. in custom algal media containing no yeastolate and no tryptophan (Table 3.13). The production of EGFP is only slightly reduced from that of Sf9 infected cells grown in the presence of tryptophan (270 µM 48 h p.i. at 2.6 x 10^8 cells/mL). Additionally, nearly all Sf9 cells fluoresced green in custom media without tryptophan when viewed 48 h p.i. under the microscope. The results show that infected Sf9 cells produce EGFP in the absence of tryptophan.

A final attempt was made to control sources of tryptophan in the custom media. Rather than using the algal amino acid mix as the source of amino acids, custom media were prepared where each amino acid except tryptophan was added from separate sources (Table 3.5). There should be no source of tryptophan is this medium. Media were prepared with all 20 amino acids, 19 amino acids (no tryptophan), and without addition of any amino acids. Sf9 cells were infected with pUltraBac-1-hsp70-AcGFP recombinant baculovirus in these custom media. The data in Table 3.14 show that Sf9 cell viability remains high 48 h p.i. even in the absence of any amino acid source (≥82%). EGFP production was not quantified. However, the majority of cells grown in 20 individual A.A. and 19 individual A.A. media fluoresced green 72 h p.i. and about 50% of cells grown in media lacking amino acids fluoresced green 72 h p.i. as assessed under the microscope. These results demonstrate that infected Sf9 cells are able to maintain cell viability for at least 48 h and produce recombinant proteins in media lacking tryptophan or any amino acids. Since Sf9 cells are able to obtain tryptophan by some unknown means, it is unlikely that the cells would utilize the unnatural 5-FW provided in the custom media for EGFP production. Thus, protein in-cell NMR with 5-FW labeling is not possible in Sf9 insect cells under these conditions.

3.3.7 Sf9 EGFP Production in ^{15}N-Enriched Custom Media

It is possible that Sf9 cells did not utilize 5-FW because it is an unnatural amino acid. An alternative strategy for isotopic enrichment in Sf9 cells may be the use of ^{15}N-containing amino acids or other ^{15}N compounds because these

components differ from their natural counterparts by just one atomic mass unit. Use of $^{15}NH_4Cl$ as a source of ^{15}N amino acids for enriching overproduced proteins in Sf9 cells is attractive due to its low cost. It is known that Sf9 cells can adapt to and grow in media supplemented with NH_4Cl at concentrations up to about 15 mM (110, 141). Sf9 cells incorporate nitrogen from NH_4Cl into glutamic acid and alanine backbone amides and glutamine and asparagine side chain amides (110).

$^{15}NH_4Cl$ custom medium was prepared as shown in Table 3.5. $^{15}NH_4Cl$ was added to a final concentration of 15 mM. Individual amino acids were added to the medium to mimic the amino acid composition of commercial ESF 921 media (113) rather than the algal amino acid mix used previously. Glutamine was omitted and additional glutamate was added because Sf9 cells can synthesize glutamine from NH_4Cl and glutamate (117).

Sf9 cells were infected with pUltraBac-1-hsp70-AcGFP recombinant baculovirus in $^{15}NH_4Cl$ custom medium. Cell viability prior to infection was 97% and remained high 48 h p.i. (96%). Forty percent of Sf9 cells fluoresced green 48 h p.i. The lower fraction of green fluorescing cells 48 h p.i. compared to results in custom media without NH_4Cl may be due to a reduction of EGFP production caused by NH_4Cl inhibition of baculovirus infection (110). The HSQC spectrum of the cell slurry shows just three crosspeaks, two of which appear in the protein side chain amide region (Fig. 3.19*a*). These results show ^{15}N enrichment of one or more cellular proteins, but cannot be definitively attributed to EGFP. Sf9 lysates prepared 96 h p.i. show a similar crosspeak pattern (Fig. 3.19*b*). EGFP

was purified by size exclusion chromatography (Fig. 3.20) and concentrated for use in *in vitro* HSQC acquisition. The total protein concentration in the pooled fractions containing EGFP was 420 µM. This *in vitro* EGFP sample was used to acquire a HSQC spectrum (Fig. 3.21*a*). Only two crosspeaks associated with protein side chain amides were observed. These results show ^{15}N incorporation into one or more proteins. Loss of the single crosspeak from the spectrum of the cell slurry, associated potentially with a protein amide backbone nitrogen, suggests that the enriched protein in the *in vitro* sample had aggregated. The two crosspeaks associated with protein side chain amides might still be present in an aggregated form of the protein because the side chains have greater mobility (52). A review of the literature revealed that GFP forms dimers at concentrations ≥100 µM (142). These dimers are too large to detect by HSQC NMR. Adding NaCl, however, may break up the hydrophilic contacts stabilizing the dimer. A spectrum of the purified EGFP in 200 mM NaCl was acquired (Fig. 3.21*b*). The spectrum, however, still only contains the two amide side chain crosspeaks.

It cannot be definitively stated that EGFP was ^{15}N-enriched in Sf9 cells because *in vitro* spectra were inconclusive, perhaps due to EGFP dimerization or aggregation. However, Sf9 cells appear to have incorporated ^{15}N into proteins through $^{15}NH_4Cl$ in the custom media. Thus, the possibility to ^{15}N enrich other recombinant proteins exists. The production of an intrinsically disordered protein such as α-syn in Sf9 cells using $^{15}NH_4Cl$ custom media may allow the acquisition of in-cell HSQC spectra because intrinsically disordered proteins are easier to detect than globular proteins in cells (49).

3.3.8 Sf9 α-Syn and α-Syn-EGFP Production in ^{15}N-Enriched Custom Media

Detection of protein amide side chain crosspeaks in the cell slurry, lysate, and *in vitro* HSQC spectra of Sf9 cells infected with pUltraBac-1-hsp70-AcGFP recombinant baculovirus suggests that ^{15}N enrichment of EGFP may have occurred. However, the globular structure of EGFP prevents detection of the protein backbone crosspeaks by HSQC NMR (69). Furthermore, EGFP dimers (142) or higher order aggregates may have formed in the purified sample, preventing attempts to detect the protein backbone crosspeaks by *in vitro* HSQC NMR. It was hypothesized that ^{15}N enrichment of an intrinsically disordered recombinant protein in Sf9 cells would allow detection of the protein's backbone crosspeaks by in-cell HSQC NMR because disordered proteins are detectable in cells (49). The intrinsically disordered protein α-synuclein was selected due to prior experience successfully producing the recombinant protein in Sf9 cells.

The α-syn gene was cloned into the pUltraBac-1 vector under the control of the P_{Basic} promoter. Two constructs were created, one containing α-syn and another containing both the α-syn and EGFP genes to generate a fusion protein. The rationale in creating the fusion construct was to allow fluorescence visualization of Sf9 cell infection. To generate the α-syn pUltraBac-1 vector, the EGFP gene under P_{Basic} control was removed from the pUltraBac-1 vector by restriction enzyme digestion to allow insertion of the α-syn gene under P_{Basic}. To generate the α-syn-EGFP pUltraBac-1 vector, the α-syn gene was amplified with primers designed to remove the stop codon and maintain the reading frame for EGFP. The pUltraBac-1 vector was then digested with a single restriction

enzyme at the multiple cloning site downstream of the P_{Basic} promoter to allow insertion of the α-syn gene. The resulting construct under P_{Basic} control contained the α-syn gene followed by the EGFP gene. The two genes were separated by a nine nucleotide linker sequence (5’ TCTCGAGCC 3’) predicted to produce the peptide α-syn-Ser-Arg-Ala-EGFP. DNA sequence analysis of the completed vector revealed that the linker sequence is 5’ TCTCGAGGG 3’ resulting in a linker region comprising Ser-Arg-Gly. The discrepancy probably arises from inaccuracies in the pUltraBac-1 sequencing data.

Initial Sf9 infection experiments were performed with α-syn pUltraBac-1 and α-syn-EGFP pUltraBac-1 recombinant baculoviruses to quantify production levels of the recombinant proteins in commercial ESF 921 media. Inspection of cells 48 h and 72 h p.i. under the microscope showed evidence of infection and acceptable viability through 48 h p.i. (Table 3.15). Under the microscope, 90% of Sf9 cells infected with the α-syn-EGFP pUltraBac-1 recombinant baculovirus fluoresced green 48 h p.i., indicating successful infection. Sf9 cells infected with α-syn pUltraBac-1 recombinant baculovirus did not fluoresce green as expected, but it was assumed that the cells were infected because both infections were performed in parallel. Additionally, enlargement of most of the cells inoculated with the α-syn pUltraBac-1 recombinant baculovirus suggested infection had occurred. Importantly for in-cell NMR, Sf9 cell viability 48 h p.i. was ≥90% for Sf9 cells infected with α-syn pUltraBac-1 and α-syn-EGFP pUltraBac-1 recombinant baculoviruses (Table 3.15).

Cell aliquots were collected 48 h and 72 h p.i. and the cellular proteins separated by SDS-PAGE. EGFP was detected by fluorescence, and α-syn was detected by western blotting. Fluorescence was observed at the molecular weight corresponding to EGFP (27 kDa) and at a higher molecular weight in the lanes corresponding to Sf9 cells infected with α-syn-EGFP pUltraBac-1 recombinant baculovirus (Fig. 3.22*a*). The fluorescence of the higher molecular weight species is more intense than that of the lower molecular weight band. Thus, the higher molecular weight species is interpreted as the primary recombinant α-syn-EGFP fusion protein, and the lower molecular weight species appears to be a degradation product, most likely free EGFP.

Quantification of EGFP fluorescence 48 h and 72 h p.i. yields an estimated intracellular α-syn-EGFP concentration of 5000 µM and 5080 µM, respectively (Table 3.15). Thus, EGFP production reaches its maximum 48 h p.i. when cell viability is still acceptable for in-cell NMR (94%).The α-syn-EGFP concentration is estimated to be 2340 µM when pelleting the cells into the NMR tube 48 h p.i. This α-syn-EGFP concentration is sufficient for in-cell NMR.

Detection of α-syn by western blotting corroborates the fluorescence data. The lanes corresponding to cells infected with α-syn-EGFP pUltraBac-1 recombinant baculovirus display multiple bands running at higher molecular weights than the corresponding α-syn control (Fig. 3.22*b*). The highest running and most intense band is interpreted as the α-syn-EGFP fusion protein. The lower molecular weight bands are assumed to be degradation products containing the α-syn epitope of the antibody. The α-syn-EGFP intracellular and

2.6 x 10^8 cells/mL concentrations 48 h p.i. are estimated to be 3970 µM and 1860 µM, respectively (Table 3.15), sufficient for in-cell NMR and in close agreement with estimates from fluorescence detection.

Western blot results also show the presence of α-syn in the lysates of cells infected with α-syn pUltraBac-1 recombinant baculovirus (Fig. 3.22*b*). Lanes corresponding to cells infected with α-syn pUltraBac-1 recombinant baculovirus display bands running at the molecular weight corresponding to the α-syn control (~20 kDa). Quantification of α-syn 48 h and 72 h p.i. yields estimated intracellular concentrations of 1210 µM and 916 µM, respectively (Table 3.15). Cell viability remains 90% 48 h p.i. The estimated α-syn concentration at 2.6 x 10^8 cells/mL 48 h p.i. is 565 µM (Table 3.15), sufficient for in-cell NMR.

These results demonstrate that the α-syn pUltraBac-1 and α-syn-EGFP pUltraBac-1 recombinant baculoviruses were successfully generated for the production of α-syn and α-syn-EGFP in Sf9 cells. Quantification of band intensities yields estimated α-syn and α-syn-EGFP concentrations sufficient for in-cell NMR experiments. Furthermore, α-syn and α-syn-EGFP levels are similar 48 h and 72 h p.i. (Fig. 3.22), indicating that cells can be harvested 48 h p.i. for in-cell NMR experiments when cell viabilities are ≥90% (Table 3.15). The next step is to isotopically enrich the proteins with ^{15}N.

Sf9 cells were infected in $^{15}NH_4Cl$ custom medium with α-syn pUltraBac-1 or α-syn-EGFP pUltraBac-1 recombinant baculoviruses to isotopically enrich recombinant α-syn and α-syn-EGFP. Cell viability prior to infection was 99% and remained high at 96% upon harvesting the cells 43 h p.i. (Table 3.16). Only 24%

of cells infected with α-syn-EGFP pUltraBac-1 recombinant baculovirus fluoresced green 43 h p.i. as assessed under the microscope. The low percentage of green fluorescing cells, due to inhibition of baculovirus infection by NH_4Cl (110), is consistent with prior observations of infected Sf9 cells producing EGFP in the same media (see 3.3.7).

The HSQC spectra of the cell slurries producing α-syn or α-syn-EGFP show the same three-crosspeak pattern (Fig. 3.23*a*, *c*). Two of the crosspeaks correspond to the protein side chain amide region. The third crosspeak may be from the backbone of the recombinant proteins. This broad crosspeak becomes two sharper crosspeaks in the spectra of the cell lysates (Fig. 3.23*b*, *d*). However, α-syn contains six glutamines and three asparagines, so at least nine protein backbone crosspeaks are expected. Furthermore, the α-syn and α-syn-EGFP cell slurry spectra are nearly identical to the EGFP cell slurry spectrum (Fig. 3.19). The similarity of the cell slurry spectra for three different recombinant proteins suggests that the signals arise from the enrichment of intracellular metabolites. The inability of Sf9 cells to utilize $^{15}NH_4Cl$ for uniform enrichment of recombinant proteins and the inhibiting effects of NH_4Cl on baculovirus infection prompted the search for a new custom media formulated without NH_4Cl to uniformly enrich recombinant proteins.

Limitations in the availability of ESF NMR media without amino acids prompted the search for other insect media suitable for in-cell NMR. Absence of amino acids in the media was required so that they could be added into the media in ^{15}N-enriched form. The search for commercially available media was

halted when the formulation of a chemically defined Sf9 medium was discovered. The composition of this medium, known as IPL-41, has been available for several decades and supports normal growth and recombinant protein production in Sf9 cells (143, 144). The formulation comprises amino acids, vitamins, inorganic salts, sugars, and other organic compounds. A chemically defined mixture of lipids is added to provide additional components required for growth and protection from shear stress while in suspension culture. Finally, yeastolate is often added to provide additional growth factors. The chemically defined IPL-41 formulation recipe available from Invitrogen Corp. was used as the basis for the custom IPL-41 NMR media (Table 3.6). Amino acids were provided in ^{15}N-enriched form by the addition of ^{15}N-enriched algal amino acid mix. The algal amino acid mix lacks tryptophan, cysteine, asparagine, and glutamine, so these amino acids were added to the custom media in unenriched form. The lipid mixture was added to the media, and yeastolate was also added to half of the media. Yeastolate promotes Sf9 cell growth but contains amino acids, so it was not added to all of the custom IPL-41 NMR media in the event that its inclusion prevented ^{15}N enrichment of the recombinant proteins.

Sf9 suspension cultures were pelleted and resuspended in ^{15}N-enriched custom IPL-41 media with or without yeastolate. The cultures were infected with α-syn-EGFP pUltraBac-1 recombinant baculovirus and incubated for 44 h. Cell viabilities were ≥94% for both samples prior to infection and remained ≥90% 44 h p.i. (Table 3.16). The percentage of Sf9 cells fluorescing green was not quantified, but nearly all cells fluoresced green as observed under the

microscope. The near 100% fluorescence suggests that NH_4Cl inhibited baculovirus infection in the previous experiment. Production levels of α-syn-EGFP were not determined by SDS-PAGE but assumed to be sufficient for in-cell NMR experiments due to the near uniform fluorescence observed and previous infection experiments in custom media (Table 3.10). The HSQC spectrum of the Sf9 cell slurry infected with α-syn-EGFP pUltraBac-1 recombinant baculovirus in ^{15}N-enriched custom IPL-41 media without yeastolate displays a similar three-crosspeak pattern observed in all other ^{15}N enrichment experiments (Fig. 3.24*a*; compare to Fig. 3.19 and Fig. 3.23). A few additional crosspeaks are present in the spectrum of the cell slurry, but not as many as expected for near uniformly labeled recombinant α-syn-EGFP (140 A.A. - 5 Pro - 3 Asn - 6 Gln - 0 Cys - 0 Trp = 126 backbone crosspeaks expected for α-syn). The spectrum of the cell lysate displays a similar crosspeak pattern (Fig. 3.24*b*). The similarities of these spectra to previous ^{15}N enrichment experiments points again to ^{15}N enrichment of intracellular metabolites and not the recombinant proteins of interest.

Another ^{15}N enrichment experiment in custom ESF media was performed concurrently with the experiments in ^{15}N-enriched custom IPL-41 media. Although ESF NMR media without amino acids is no longer available from the manufacturer, an alternative medium is available. This ESF medium lacks phenylalanine, isoleucine, leucine, valine, methionine, tyrosine, and threonine. Using this ESF medium, ^{15}N-phenylalanine custom medium was prepared by adding ^{15}N-phenylalanine and the remaining six amino acids in unenriched form (see 3.2.9). Infection of Sf9 cells in this custom medium was rationalized as a

simple method to determine ^{15}N enrichment of recombinant proteins through ^{15}N-phenylalanine protein backbone detection by NMR.

Sf9 suspension cultures in ^{15}N-phenylalanine custom media were infected with α-syn pUltraBac-1 or α-syn-EGFP pUltraBac-1 recombinant baculoviruses and incubated for 44 h. Viabilities of cultures infected with either baculovirus were ≥98% 44 h p.i. (Table 3.16). Nearly all cells fluoresced green in the α-syn-EGFP pUltraBac-1 recombinant baculovirus infected culture as observed under the microscope. Production levels of α-syn and α-syn-EGFP were not determined by SDS-PAGE but assumed to be sufficient for in-cell NMR experiments based on previous results. The HSQC spectrum of the Sf9 cell slurry infected with α-syn pUltraBac-1 recombinant baculovirus and acquired 44 h p.i. displays two crosspeaks in the protein amide side chain region and several crosspeaks that could be from protein backbone amides (Fig. 3.25). However, α-syn contains just two phenylalanines, so only two crosspeaks are expected in the spectrum of the cell slurry. The spectrum of the Sf9 cell slurry infected with α-syn-EGFP pUltraBac-1 recombinant baculovirus and both lysate spectra were not acquired. The greater than expected number of crosspeaks and apparent enrichment of asparagine and glutamine side chain amides renders the ^{15}N-phenylalanine enrichment experiments difficult to interpret. It appears that Sf9 cells incorporate ^{15}N-phenylalanine and its breakdown products into other amino acids and metabolites. Identification of crosspeaks associated with α-syn becomes difficult with the high degree of background enrichment. Thus, under the current

experimental conditions, it is not possible to acquire Sf9 in-cell HSQC spectra of ^{15}N-enriched disordered recombinant proteins.

3.4 Conclusions

The baculovirus expression system was successfully utilized to produce sufficient quantities of recombinant proteins while maintaining cell viability for in-cell NMR experiments. Results show that Sf9 cell infection with recombinant baculovirus containing a P_{Basic} promoter produces sufficient EGFP for in-cell NMR. The P_{Basic} promoter produces high levels of EGFP at earlier times (prior to cell lysis) than the more commonly used PH promoter, and thus cell viability remains acceptable for in-cell NMR. However, in-cell ^{19}F NMR studies proved unsuccessful due to the lack of ^{19}F incorporation into EGFP in the form of 5-FW. Sf9 cells can metabolize NH_4Cl, and so addition of $^{15}NH_4Cl$ to Sf9 media provided a potential method to isotopically enrich EGFP. Results were inconclusive because EGFP formed dimers or aggregated, and either form is too large to detect in cells by using the HSQC experiment. However, ^{15}N enrichment was detected in the side chains of one or more cellular proteins, indicating that Sf9 cells use $^{15}NH_4Cl$ for protein production. These findings suggested that ^{15}N enrichment of a disordered protein in Sf9 cells by these methods might yield successful protein in-cell NMR results. The NMR results for infected Sf9 cells producing α-syn and α-syn-EGFP in ^{15}N-containing custom media were inconclusive. Similarity in the spectra of the Sf9 cell slurries producing α-syn, EGFP, and α-syn-EGFP in $^{15}NH_4Cl$ containing media and the absence of uniform enrichment of the proteins in ^{15}N algal amino acid containing media suggests the

enrichment of other intracellular components. Therefore, the acquisition of disordered recombinant proteins in Sf9 cells by HSQC NMR is not possible under the current experimental conditions. Commerically available ^{15}N-enriched Bioexpress-2000 media (Cambridge Isotope Laboratories, Inc.) was not utilized in these experiments. This media has been shown to uniformly enrich a recombinant protein (116) in insect cells. Thus, this media remains a potential, albeit expensive, candidate for future isotopic enrichment experiments.

3.5 List of Tables

Table 3.1 pIEx-7 sequencing primers for target gene insert

Name	Sequence
IE1 Promoter Primer	5’ TGGATATTGTTTCAGTTGCAAG 3’
AS S-Tag Primer	5’ GTCCGTGTGCTGGCGTTC 3’

Table 3.2 Primers for PCR amplification of target genes into the pIEx-7 Ek/LIC vector

Gene	Vector	Sequence (forward and reverse complement)
α-Syn	pT7-7[1]	5’ GACGACGACAAGATGGATGTATTCATGAAAGGAC 3’
		5’ GAGGAGAAGCCCGGTTAGGCTTCAGGTTCGTAGTC 3’
Cytochrome *c*	pBTR[2]	5’ GACGACGACAAGATGGGCGACGTGGAAAAAGGCA 3’
		5’ GAGGAGAAGCCCGGTTATCATTCGTTCGTCGCCTT 3’
AcGFP	pAcGFP1[3]	5’ GACGACGACAAGATGGTGAGCAAGGGCGCCGA 3’
		5’ GAGGAGAAGCCCGGTCACTTGTACAGCTCATCCAT 3’
Ubiquitin	pASK[4]	5’ GACGACGACAAGATGGCAATCTTCGTCAAGACGTT 3’
		5’ GAGGAGAAGCCCGGTCAACCACCTCTTAGTCTTAA 3’
HSP27	pCMV6-AC[5]	5’ GACGACGACAAGATGACCGAGCGCCGCGTCCCCTT 3’
		5’ GAGGAGAAGCCCGGTTACTTGGCGGCAGTCTCATC 3’

[1] pT7-7 vector with the α-syn gene is from (18)
[2] pBTR vector with the human cytochrome *c* gene is from (145)
[3] pAcGFP1 is from Clontech Laboratories, Inc. (Mountain View, CA)
[4] pASK-Ubq was provided by Alexander Shekhtman (146)
[5] pCMV6-AC vector with the human HSP27 gene is from Origene Technologies, Inc. (Rockville, MD)

Table 3.3 pAcGFP1 primers for PCR amplification of the AcGFP gene

Name	Sequence
pAcGFP1 forward	5’ CCGGCTCGTATGTTGTGTGGAATTGTGAGCGGAT 3’
pAcGFP1 reverse	5’ TGTCAGAGGTTTTCACCGTCATCACCGAAACGCG 3’

Table 3.4 pUltraBac-1 mutagenesis primers to introduce *BamHI* and *HindIII* sites

Name	Sequence
Forward *HindIII*	5’ CTCGTCTTGGTCGTTTGAAGCTTTGTTGCTGTGTTTCC 3’
Reverse *HindIII*	5’ GGAAACACAGCAACAAAGCTTCAAACGACCAAGACGAG 3’
Forward *BamHI*	5’ CACTGCTTGAGCCTAGGGGATCCGAACCAGATAAG 3’
Reverse *BamHI*	5’ CTTATCTGGTTCGGATCCCCTAGGCTCAAGCAGTG 3’

The *HindIII* and *BamHI* recognition sites are underlined

Table 3.5 Amounts of sugars, amino acids, and $^{15}NH_4Cl$ used to create custom media

20 Individual A.A. Custom Media[1]		$^{15}NH_4Cl$ Custom Media[2]
Ingredient	Concentration (g/L)	
Glucose	4.81	3.60
Galactose	1.20	1.20
Mannose	2.16	2.16
Fructose	0.38	0.38
Sucrose	1.75	1.75
	Concentration (mg/L)	
Aspartate	28.5	159
Glutamate	372	2446
Asparagine	201	359
Serine	11.1	185
Glutamine	999	0
Glycine	18.6	117
Histidine	1.8	178
Arginine	14.4	381
Threonine	14.7	204
Alanine	354	251
Proline	15	338
Tyrosine	6.6	303
Valine	22.5	607
Methionine	6.3	902
Cysteine	201	100
Isoleucine	18.9	762
Leucine	324	632
Phenylalanine	9.3	597
Lysine	21.3	193
Tryptophan	201[3, 4]	100
$^{15}NH_4Cl$	-	800

[1] Amino acid content mimics 3 g/L algal amino acid mix (113)
[2] Amino acid content mimics commercial ESF 921 media (113) except no glutamine was added, an additional 2 g/L glutamate was added, and cysteine and tryptophan were added at 100 mg/L
[3] Tryptophan was omitted in 19 individual A.A. custom media
[4] All amino acids were omitted in no A.A. custom media
Custom media components were added to ESF NMR media

Table 3.6 ^{15}N-enriched custom IPL-41 media

Ingredient	Concentration[1] (mg/L)
Biotin	0.16
Choline chloride	20
D-Calcium pantothenate	0.008
Folic Acid	0.08
Nicotinic acid	0.16
Para-Aminobenzoic acid	0.32
Pyridoxine hydrochloride	0.4
Riboflavin	0.08
Succinic acid	4.8
Thiamine hydrochloride	0.08
Vitamin B12	0.24
Isoinositol	0.4
$(NH_4)_6Mo_7O_{24}\cdot 4H_20$	0.04
$CaCl_2$	500
$CoCl_2\cdot 6H_2O$	0.05
$CuCl_2\cdot 2H_2O$	0.2
$FeSO_4\cdot 7H_2O$	0.55
$MgSO_4$	918
$MnCl_2\cdot 4H_2O$	0.02
KCl	1200
$NaHCO_3$	350
NaCl	2850
$NaH_2PO_4\cdot H_2O$	1160
$ZnCl_2$	0.04
α-Ketoglutaric acid	29.6
D-Glucose	2500
Fumaric acid	4.4
Malic acid	53.6
Maltose	1000
Sucrose	1650
Cysteine	120
Tryptophan	100
Glutamine	1000
Asparagine	1300
^{15}N-enriched algal amino acid mix	1000
Chemically Defined Lipid Concentrate, 100x	1% (v/v)
Yeastolate Ultrafiltrate, 50x	4% (v/v)[2]

[1] The media components are listed at their final concentrations in deionized water. Cysteine, tryptophan, glutamine, and asparagine were added at the concentrations present in the IPL-41 formulation recipe (Invitrogen Corp.)
[2] ^{15}N-enriched custom IPL-41 media were prepared with or without Yeastolate Ultrafiltrate, 50x

Table 3.7 Viabilities and α-syn production in Sf9 cells infected with α-syn pFastBac 1 recombinant baculovirus 1-8 days p.i.

Day p.i.	Viability (%)[1]	α-Syn Concentration (μM)[2]		
		2 x 10^6 cells/mL	2.6 x 10^8 cells/mL	Intracellular
1	92 ±5	0.28 ±0.02	36 ±2	79 ±5
2	87 ±6	0.29 ±0.05	37 ±7	81 ±14
3	63 ±6	2.3 ±1.0	293 ±130	637 ±283
4	57 ±3	3.4 ±0.2	447 ±31	973 ±68
5	39 ±6	3.4 ±0.5	440 ±64	958 ±139
6	31 ±12	1.9 ±0.1	250 ±17	545 ±37
7	26 ±16	0.45 ±0.04	58 ±5	126 ±11
8	15 ±14	0.40 ±0.03	52 ±4	113 ±9

[1] Sf9 viability prior to infection was 93 ±3%
[2] Estimated concentrations calculated as described in 3.3.2
The uncertainty is the standard error of three separate experiments

Table 3.8 Recombinant protein yields from Sf9 cells 48 h post transfection with pIEx-7 vectors

Protein	Number of Measurements	Yield[1] (mg/L)	Viability
Mock Infected	2	0	81 ±1
α-Syn	2	1.3 ±0.1	33 ±1
Cytochrome *c*	3	0.5 ±0.1	48 ±9
AcGFP	5	1.5 ±0.2	33 ±7
Ubiquitin	2	0.7 ±0.1	37 ±6
HSP27	4	2.7 ±0.4	44 ±4

[1] Protein yields normalized to 2 x 10^6 cells/mL and reported with standard errors

Table 3.9 Viabilities and EGFP concentrations in Sf9 cells infected with pUltraBac-1-hsp70-AcGFP recombinant baculovirus in commercial ESF 921 media

Hours p.i.	Viability (%)[1]	EGFP Concentration (μM)[2]		
		2 x 10^6 cells/mL	2.6 x 10^8 cells/mL	Intracellular
48	85	2.1	278	604
72	54	2.2	282	614

[1] Sf9 viability prior to infection was 96%
[2] Estimated concentrations calculated as described in 3.3.2

Table 3.10 Viabilities and EGFP concentrations in Sf9 cells infected with pUltraBac-1-hsp70-AcGFP recombinant baculovirus in custom media

Hours p.i.	ESF 921	Algal Media	Algal+5% ESF 921	20 A.A. Media	20 A.A.+5% ESF 921
			Viability (%)[1]		
48	95	89	92	96	93
72	86	61	56	55	61
			Concentration, 2 x 10^6 cells/mL (µM)[2]		
48	1.88	1.75	1.89	1.63	1.42
72	2.49	1.83	1.81	1.44	1.43
			Concentration, 2.6 x 10^8 cells/mL (µM)		
48	244	228	246	212	185
72	324	238	235	187	186
			Intracellular Concentration (µM)		
48	533	496	535	461	402
72	703	518	512	408	405

[1] Sf9 viability prior to infection was 98%
[2] Estimated concentrations calculated as described in 3.3.2

Table 3.11 Viabilities and EGFP concentrations in Sf9 cells infected with pUltraBac-1-hsp70-AcGFP recombinant baculovirus in 5-FW custom algal media

	ESF 921 Media		5-FW Custom Algal Media	
	Viability (%)	[EGFP] (µM)	Viability (%)	[EGFP] (µM)
Initial	91	n.d.	94	n.d
24 h p.i.	99	0.85[1]/111[2]/239[3]	97	0.69/90/195
48 h p.i.	88	1.94/252/549	87	1.41/183/398
72 h p.i.	72	2.04/265/578	59	1.18/153/333

n.d. = not determined
[1] EGFP concentration at 2 x 10^6 cells/mL
[2] EGFP concentration at 2.6 x 10^8 cells/mL
[3] EGFP intracellular concentration
Estimated concentrations calculated as described in 3.3.2

Table 3.12 Viabilities and EGFP concentrations in Sf9 cells infected with pUltraBac-1-hsp70-AcGFP recombinant baculovirus in variable yeastolate custom algal media

	0% Yeastolate	2% Yeastolate	10% Yeastolate
		Viability (%)	
Initial	99	99	99
24 h p.i.	96	95	97
48 h p.i.	87	84	93
72 h p.i.	54	60	74
		[EGFP] (µM)	
24 h p.i.	0.55[1]/72[2]/155[3]	0.47/61/134	0.29/38/81
48 h p.i.	1.46/190/414	1.44/187/408	0.94/122/267
72 h p.i.	1.31/170/372	1.41/183/398	1.13/147/318

[1] EGFP concentration at 2 x 10^6 cells/mL
[2] EGFP concentration at 2.6 x 10^8 cells/mL
[3] EGFP intracellular concentration
Estimated concentrations calculated as described in 3.3.2

Table 3.13 Viabilities and EGFP concentrations in Sf9 cells infected with pUltraBac-1-hsp70-AcGFP recombinant baculovirus in “no yeastolate” custom algal media

	“No Yeastolate, No Tryptophan”		“No Yeastolate” and 0.1 g/L Tryptophan	
	Viability (%)	[EGFP] (µM)	Viability (%)	[EGFP] (µM)
48 h p.i.	96	1.65[1]/215[2]/467[3]	94	2.08/270/589
72 h p.i.	84	1.77/230/502	79	2.59/337/734

[1] EGFP concentration at 2 x 10^6 cells/mL
[2] EGFP concentration at 2.6 x 10^8 cells/mL
[3] EGFP intracellular concentration
Estimated concentrations calculated as described in 3.3.2

Table 3.14 Viabilities of Sf9 cells infected with pUltraBac-1-hsp70-AcGFP recombinant baculovirus in 20 individual A.A., 19 individual A.A., and no A.A. custom media

	20 Individual A.A.	19 Individual A.A.	No A.A.
		Viability (%)	
24 h p.i.	91	89	83
48 h p.i.	86	85	82

Table 3.15 Viabilities and recombinant protein concentrations in Sf9 cells infected with α-syn pUltraBac-1 or α-syn-EGFP pUltraBac-1 recombinant baculovirus in commercial ESF 921 media

	α-Syn pUltraBac-1 Baculovirus		α-Syn-EGFP pUltraBac-1 Baculovirus			
	Viability (%)	Conc. (μM) (western blot)	Viability (%)	Conc. (μM) (fluorescence)	Conc. (μM) (western blot)	Conc. (μM) (average)
Initial	96	n.d.	96	n.d.	n.d.	n.d.
48 h p.i.	90	4.34[1]/565[2]/1210[3]	94	18.0/2340/5000	14.3/1860/3970	16.2/2100/4490
72 h p.i.	74	3.30/429/916	74	18.3/2380/5080	14.9/1940/4150	16.6/2160/4610

n.d. = not determined
[1] estimated protein concentration at 2 x 10^6 cells/mL
[2] estimated protein concentration at 2.6 x 10^8 cells/mL
[3] estimated intracellular protein concentration

Table 3.16 Viabilities and fluorescence of Sf9 cells infected with α-syn pUltraBac-1 and α-syn-EGFP pUltraBac-1 recombinant baculoviruses in ^{15}N-enriched custom media for in-cell NMR

	$^{15}NH_4Cl$ Media		^{15}N-Enriched IPL-41 Media		^{15}N-Phe Media	
	α-Syn	α-Syn-EGFP	α-Syn-EGFP - Yeastolate	α-Syn-EGFP + Yeastolate	α-Syn	α-Syn-EGFP
			Viability (%)			
0 h p.i.	99	99	100	94	99	100
43-44 h p.i.	96	96	98	90	98	98
Fluoroscence[1]	No	24%	Yes	Yes	No	Yes

[1] Percent fluorescing cells was determined only for Sf9 cells infected with α-syn-EGFP pUltraBac-1 recombinant baculovirus in $^{15}NH_4Cl$ custom media. Nearly all cells fluoresced green in the samples labeled “Yes” as observed under the microscope.

3.6 List of Figures

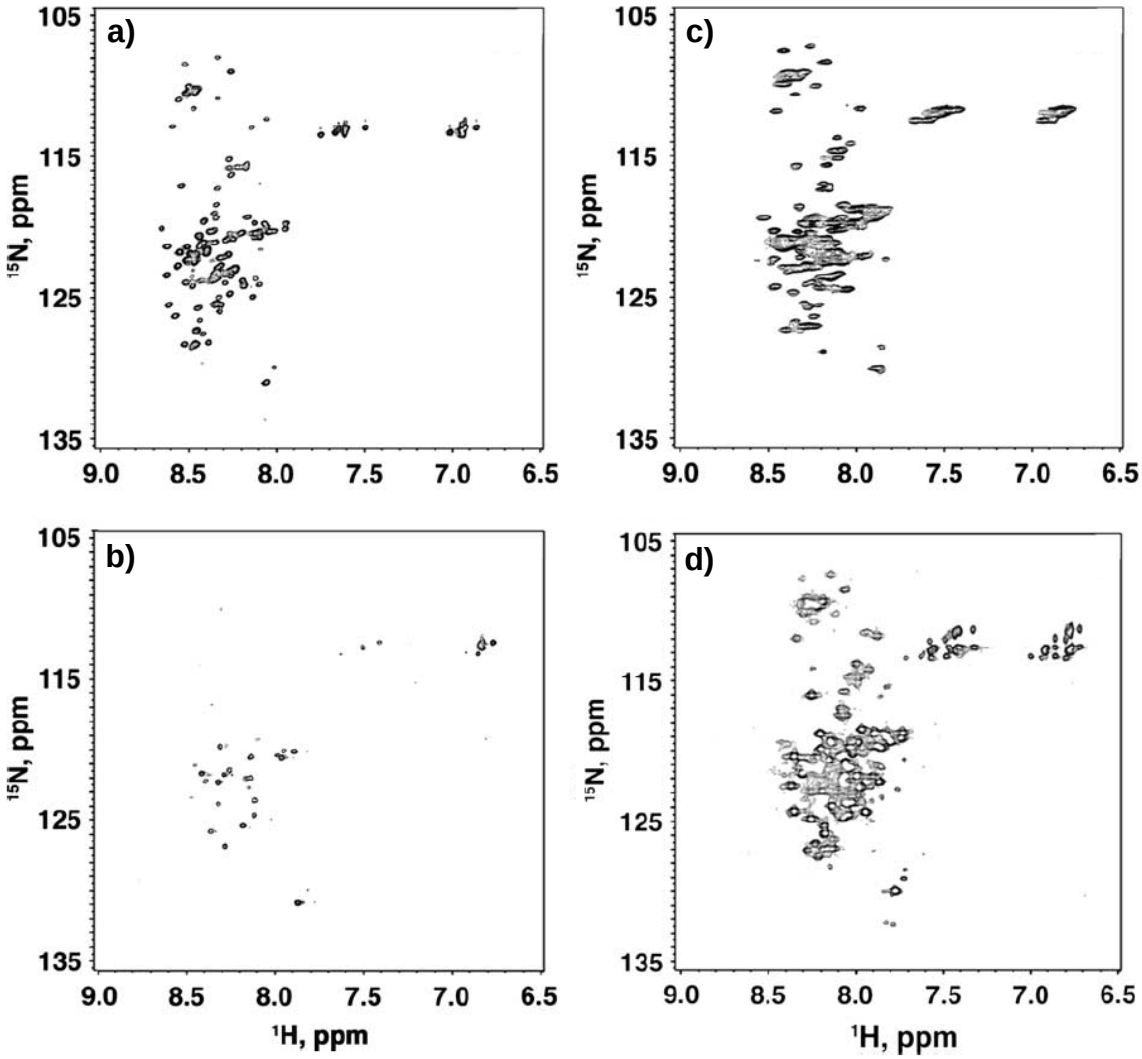
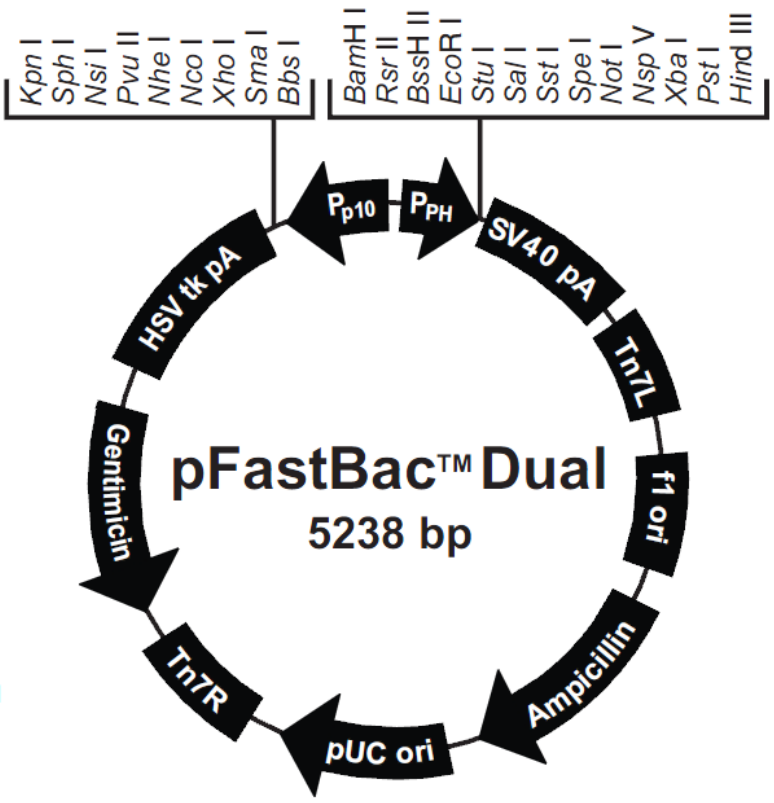


Figure 3.1 HSQC spectra of wild-type α-syn a) in dilute solution at 10 °C, b) dilute solution at 35 °C, c) in *E. coli* at 35 °C, and d) in a 300 g/L solution of bovine serum albumin at 35 °C. Images adapted from (18).



Comments for pFastBac™Dual
5238 nucleotides

f1 origin: bases 102-557
Ampicillin resistance gene: bases 689-1549
pUC origin: bases 1694-2367
Tn7R: bases 2611-2835
Gentamicin resistance gene: bases 2902-3435 (complementary strand)
HSV tk polyadenylation signal: bases 3992-4274 (complementary strand)
Multiple cloning site: bases 4274-4337 (complementary strand)
p10 promoter (P_{p10}): bases 4338-4459 (complementary strand)
Polyhedrin promoter (P_{PH}): bases 4478-4606
Multiple cloning site: bases 4606-4704
SV40 polyadenylation signal: bases 4722-4962
Tn7L: bases 4991-5156

Figure 3.2 pFastBac Dual vector map (Invitrogen Corp.). After transposition into a baculovirus bacmid and insect cell transfection, this vector allows coexpression of two genes under the control of the polyhedrin (PH) and p10 promoters.

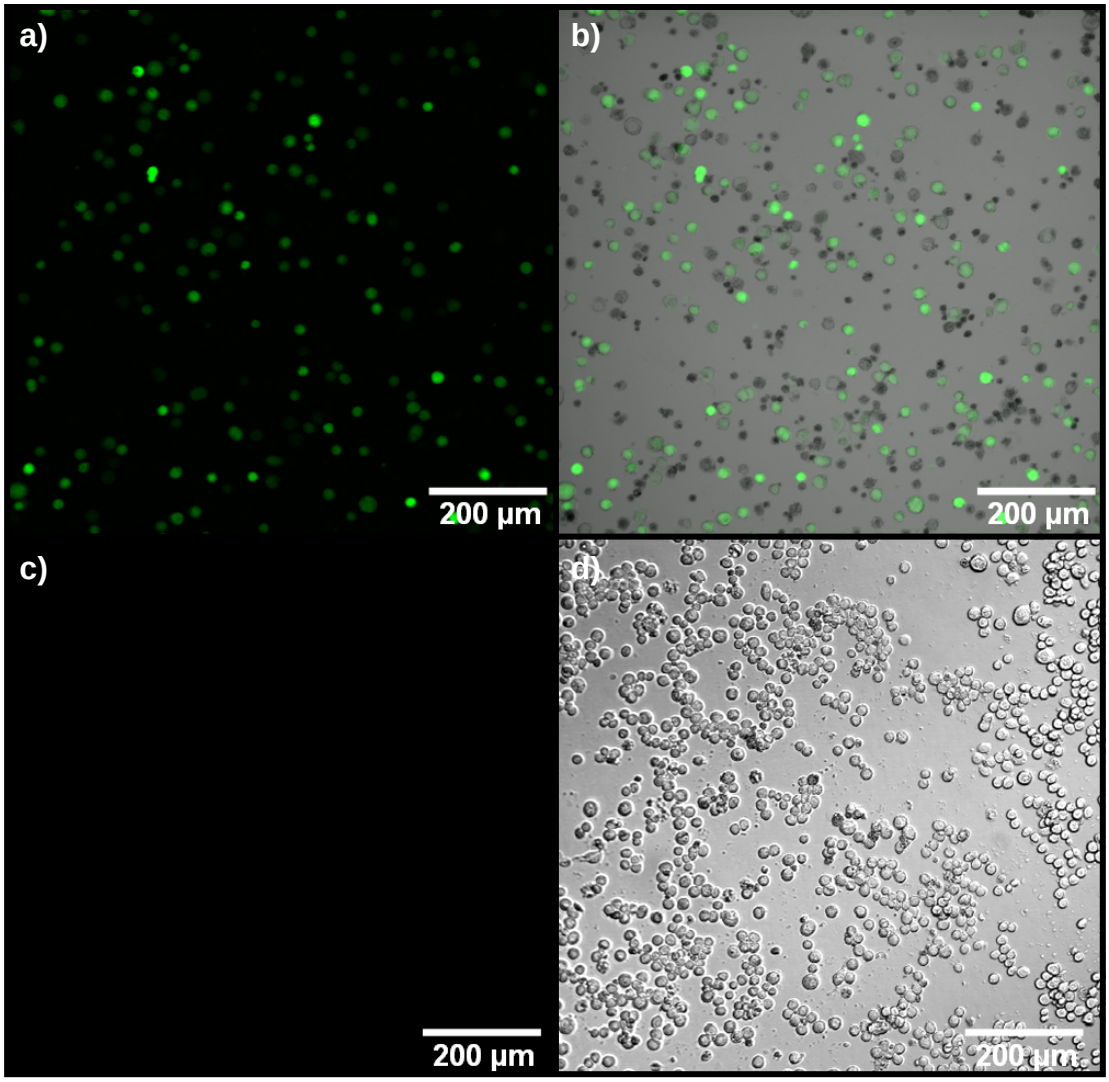


Figure 3.3 Fluorescence images of Sf9 cells infected with EGFP pFastBac Dual recombinant baculovirus. a) Fluorescence image of Sf9 cells 8 days p.i. b) Fluorescence and transmitted light overlay of Sf9 cells 8 days p.i. c) Fluorescence image of uninfected Sf9 cells. d) Fluorescence and transmitted light overlay of uninfected Sf9 cells. Images were acquired with the Zeiss 510 Meta Laser Scanning Confocal Microscope using a 488 nm excitation wavelength.

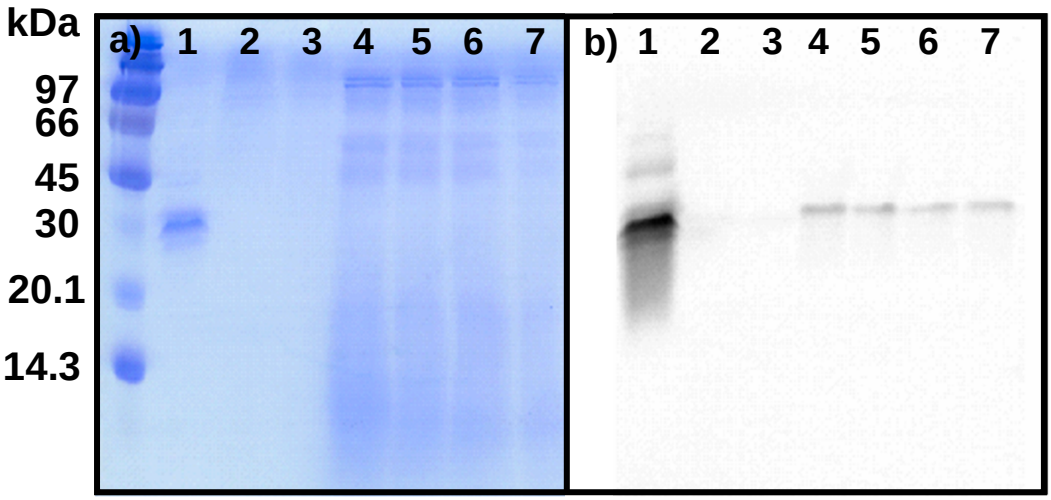
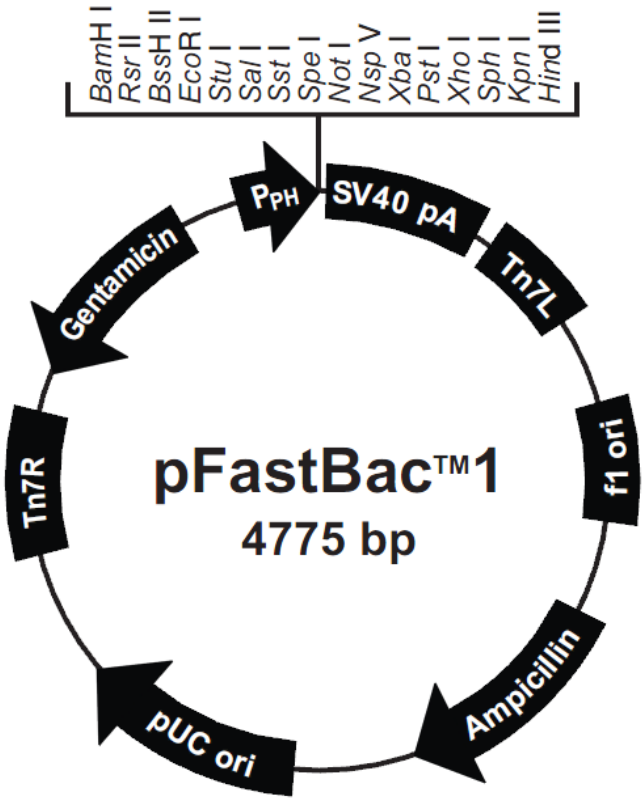


Figure 3.4 SDS-PAGE images of whole cell lysates from Sf9 cells infected with EGFP pFastBac Dual recombinant baculovirus visualized with a) Brilliant Blue staining and b) fluorescence. Lane 1 contains an EGFP positive control, lanes 2-3 contain spent media (to detect protein leakage), and lanes 4-7 contain identical whole cell lysates. Lane identifications are the same in a) and b).



Comments for pFastBac™1
4775 nucleotides

f1 origin: bases 2-457
Ampicillin resistance gene: bases 589-1449
pUC origin: bases 1594-2267
Tn7R: bases 2511-2735
Gentamicin resistance gene: bases 2802-3335 (complementary strand)
Polyhedrin promoter (P_{PH}): bases 3904-4032
Multiple cloning site: bases 4037-4142
SV40 polyadenylation signal: bases 4160-4400
Tn7L: bases 4429-4594

Figure 3.5 pFastBac 1 vector map (Invitrogen Corp.). After transposition into a baculovirus bacmid and insect cell transfection, this vector allows expression of the gene of interest under the control of the PH promoter.

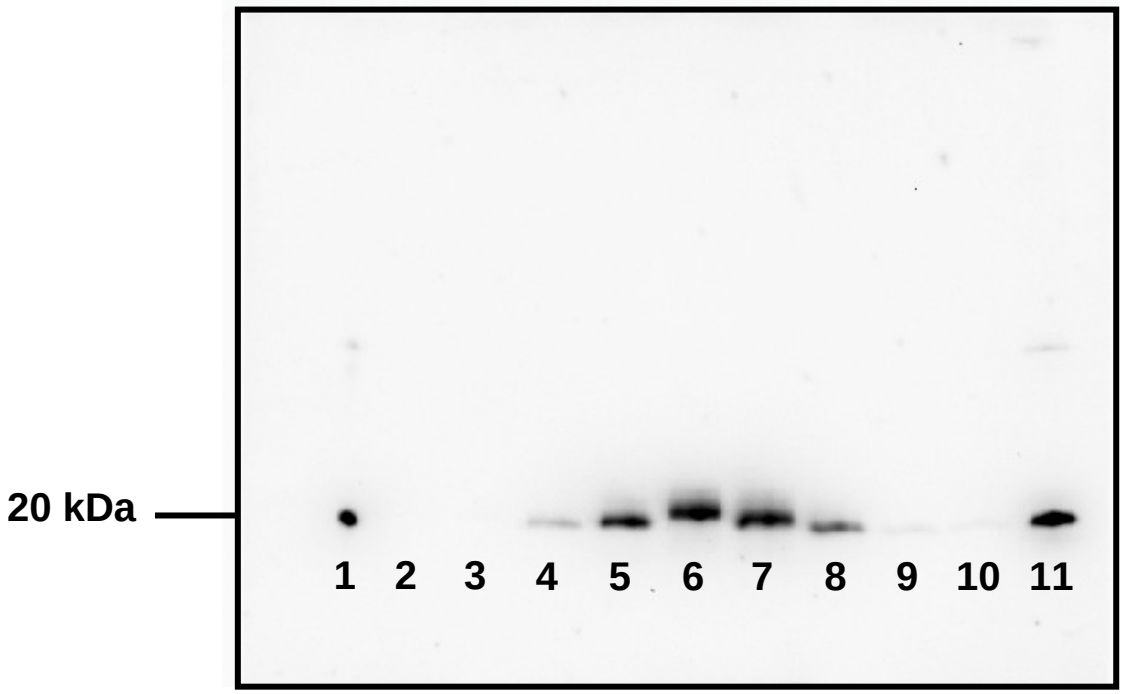


Figure 3.6 Western blot of whole cell lysates of Sf9 cells infected with α-syn pFastBac 1 recombinant baculovirus. Lanes 1 and 11 contain α-syn positive controls (each lane loaded with 1.9 µg). Lane 2 contains uninfected Sf9 whole cell lysate. Lanes 3-10 correspond to infected Sf9 suspension cells harvested 1-8 days p.i., respectively.

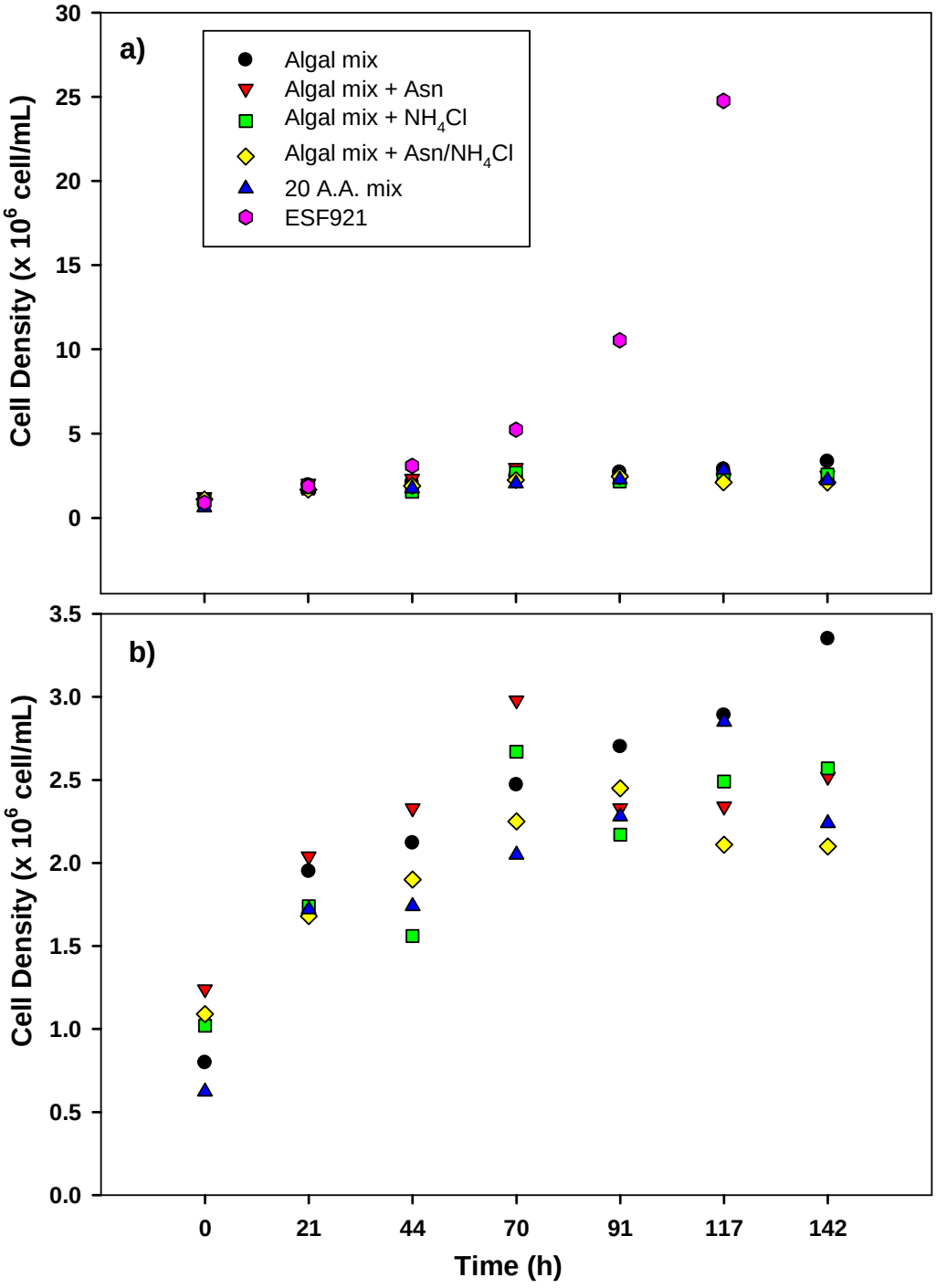


Figure 3.7 Sf9 cell growth in custom media. a) Cell growth in custom media and commercial ESF 921 medium. b) Expanded view of Sf9 cell growth in custom media only.

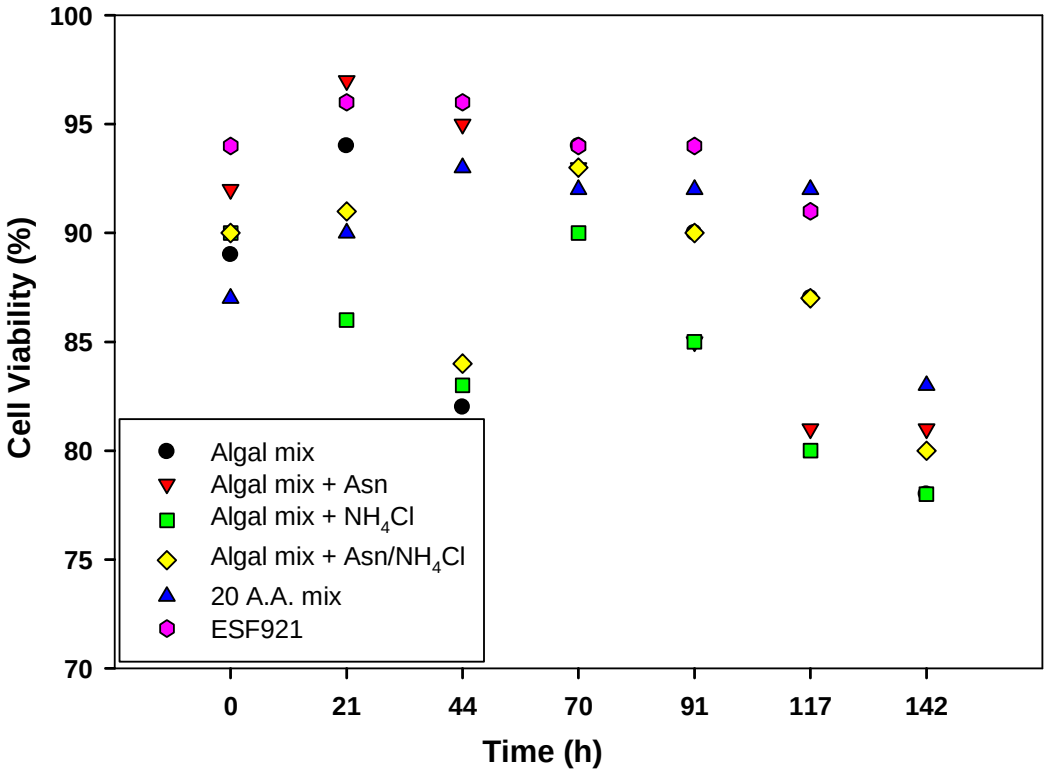


Figure 3.8 Sf9 cell viability in custom media.

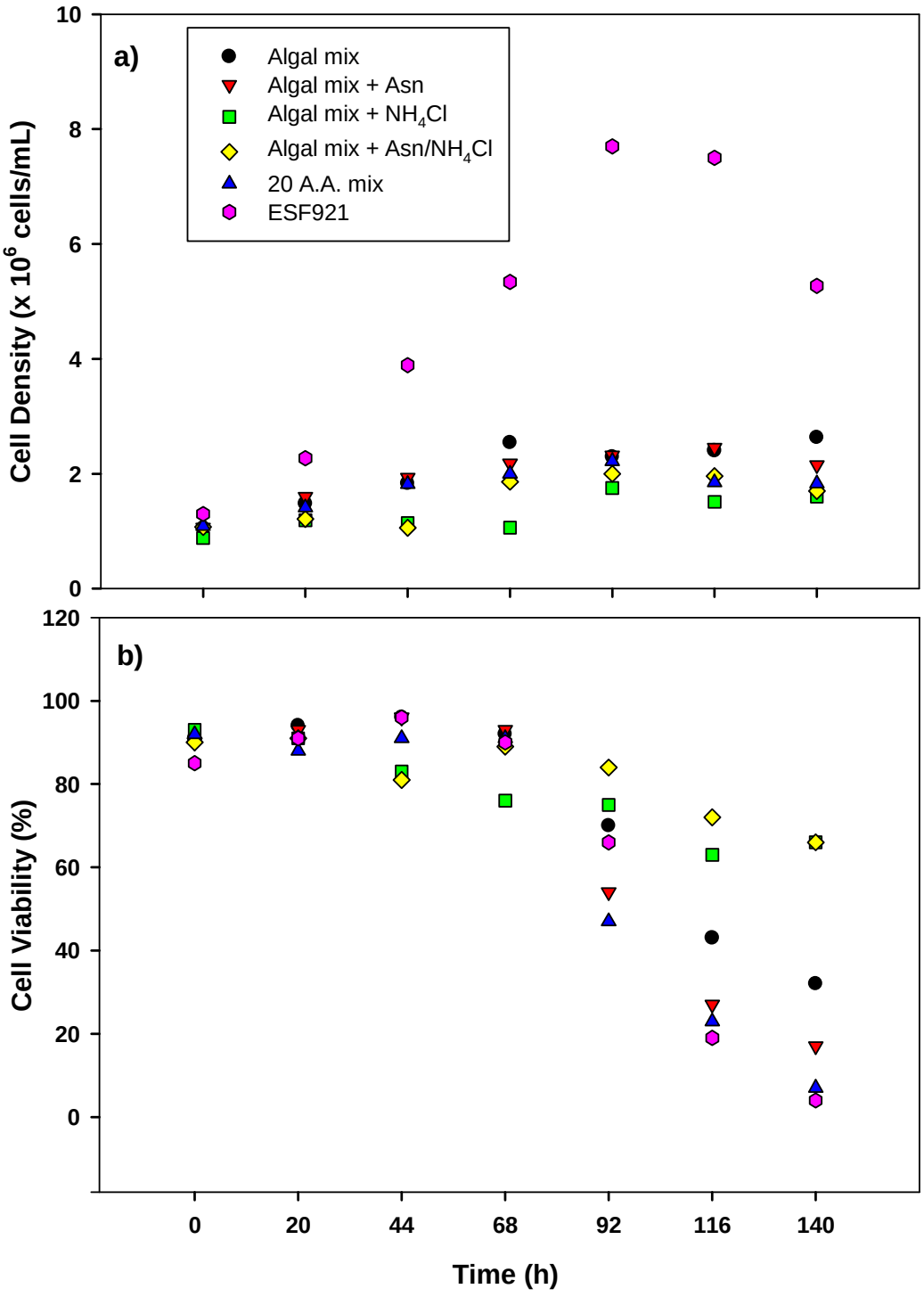


Figure 3.9 a) Growth and b) viability of α-syn pFastBac 1 recombinant baculovirus infected Sf9 cells in custom media.

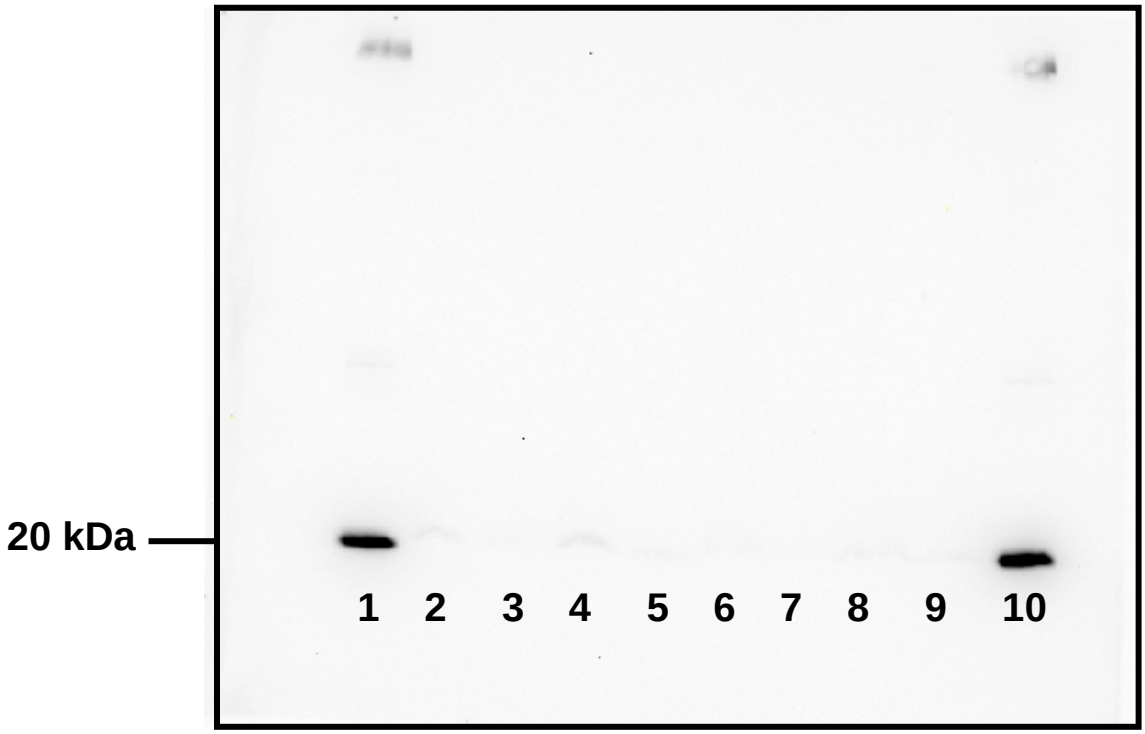


Figure 3.10 Western blot of whole cell lysates of Sf9 cells infected with α-syn pFastBac 1 recombinant baculovirus of known titer at a MOI of 3 in custom media containing larger amounts of yeastolate and glucose (see 3.2.4). Lanes 2, 4, 6, and 8 are loaded with increasing amounts of whole cell lysate 72 h p.i. Lanes 3, 5, 7, and 9 are loaded with increasing amounts of whole cell lysate 96 h p.i. Lanes 1 and 10 are α-syn controls (each lane loaded with 1.9 µg).

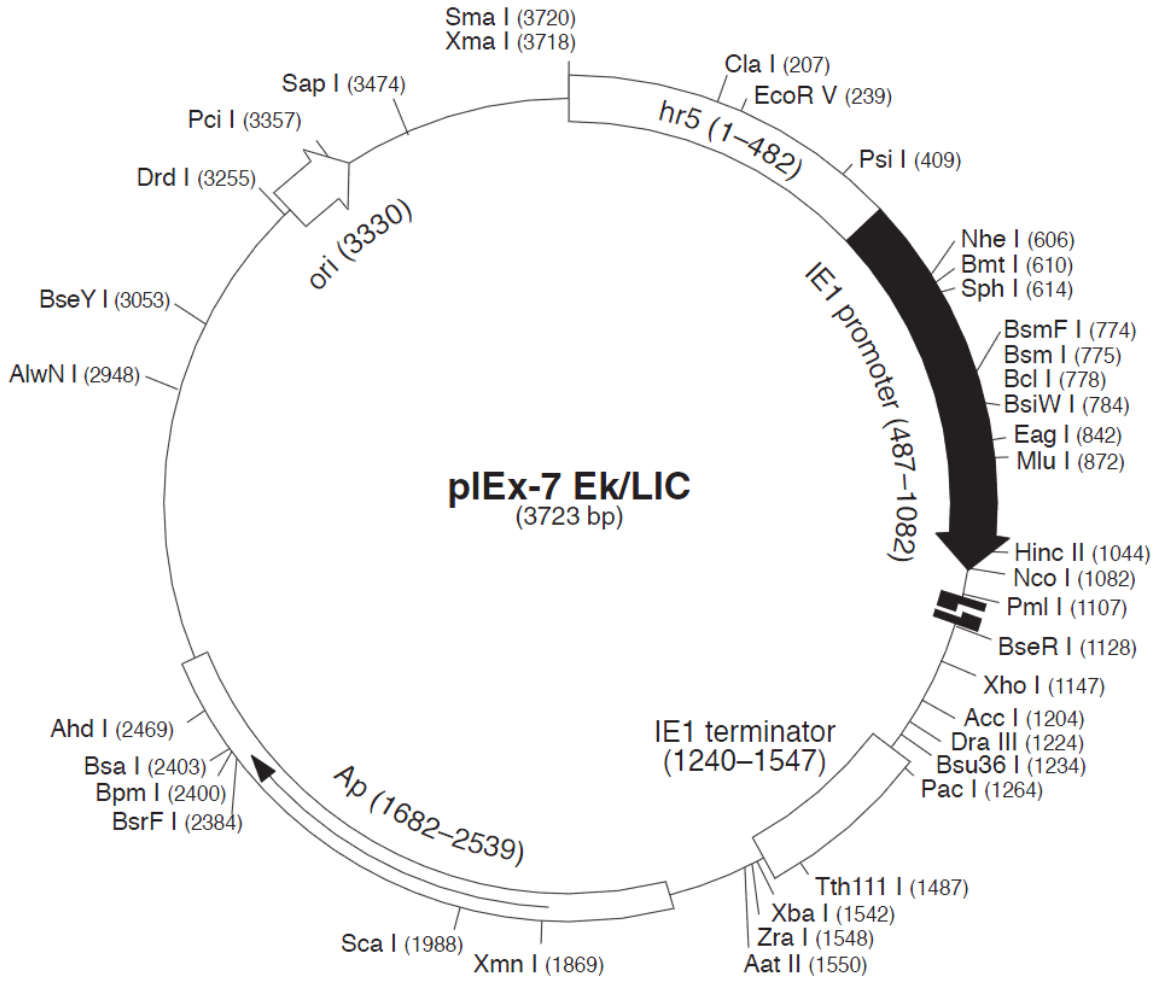


Figure 3.11 pIEx-7 Ek/LIC vector map (EMD Biosciences, Inc.). The vector employs an hr5 enhancer and IE1 promoter combination to recruit endogenous insect cell transcription machinery to directly produce the recombinant protein of interest. The vector employs ligation-independent cloning to insert genes, enabling quick screening of multiple targets.

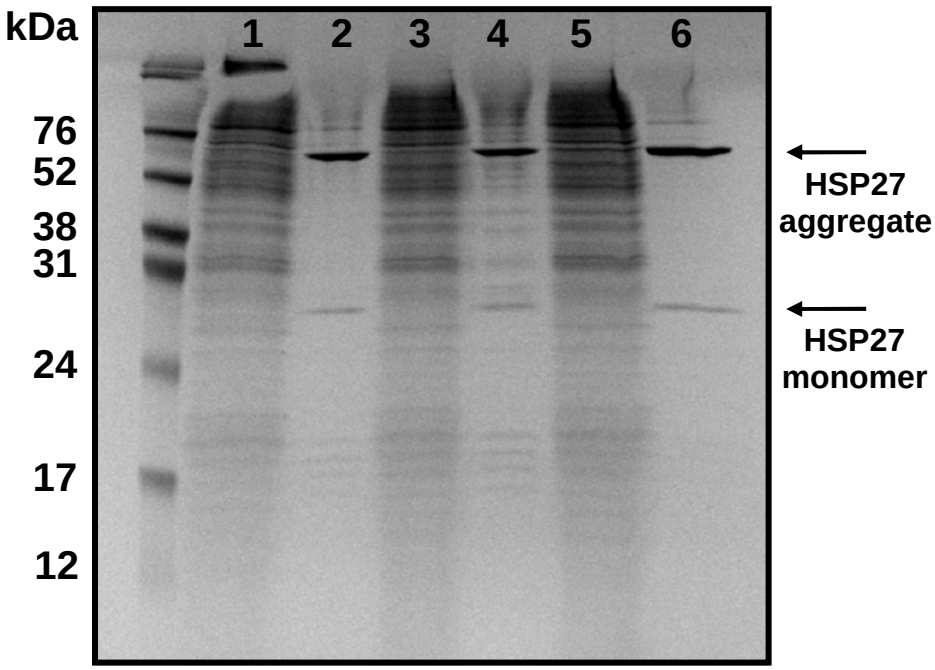


Figure 3.12 SDS-PAGE results for recombinant human HSP27 purified from pIEx-7 HSP27 transfected Sf9 cells using the InsectDirect transfection system. Lanes 1, 3, and 5 contain identical lysates of pIEx-7 HSP27 transfected Sf9 cells. Lanes 2, 4, and 6 contain identical samples of Ni-NTA purified HSP27. The monomer runs at about 27 kDa while a higher molecular weight product runs at about 60 kDa.

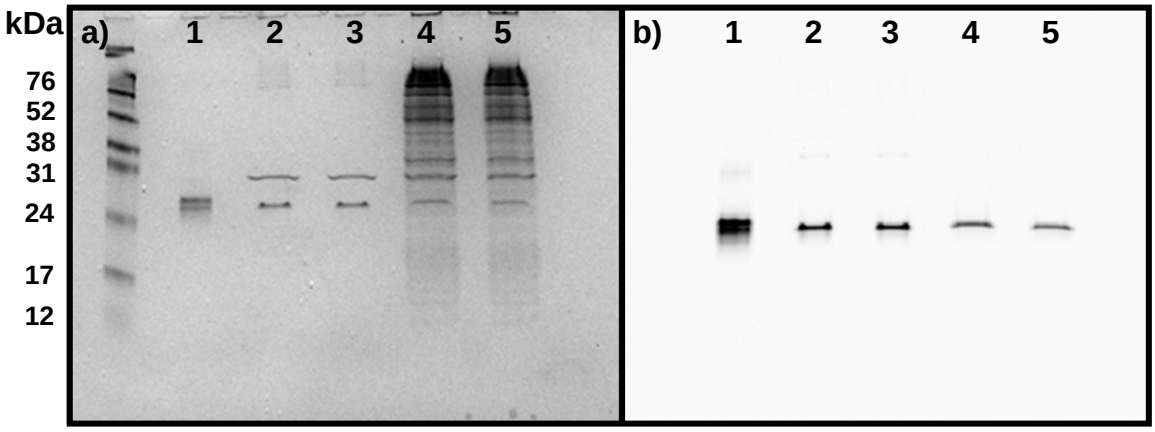


Figure 3.13 SDS-PAGE results for recombinant AcGFP purified from Sf9 cells transfected with pIEx-7 AcGFP. Lane 1 contains the AcGFP control, lanes 2-3 contain Ni-NTA purified AcGFP, and lanes 4-5 contain whole cell lysates. a) Brilliant Blue staining. b) Fluorescence detection prior to Brilliant Blue staining. Lane identifications are the same in a) and b).

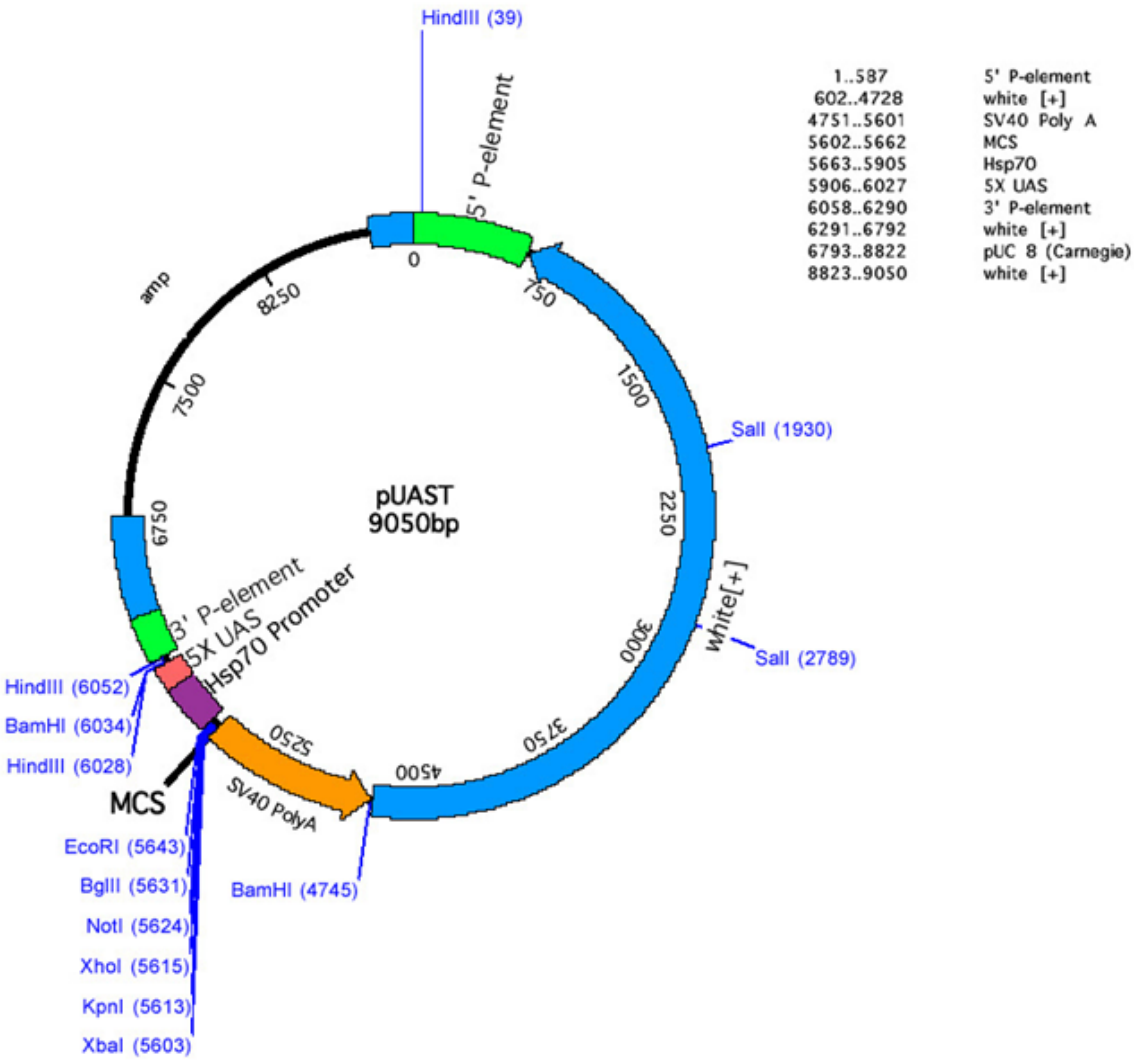
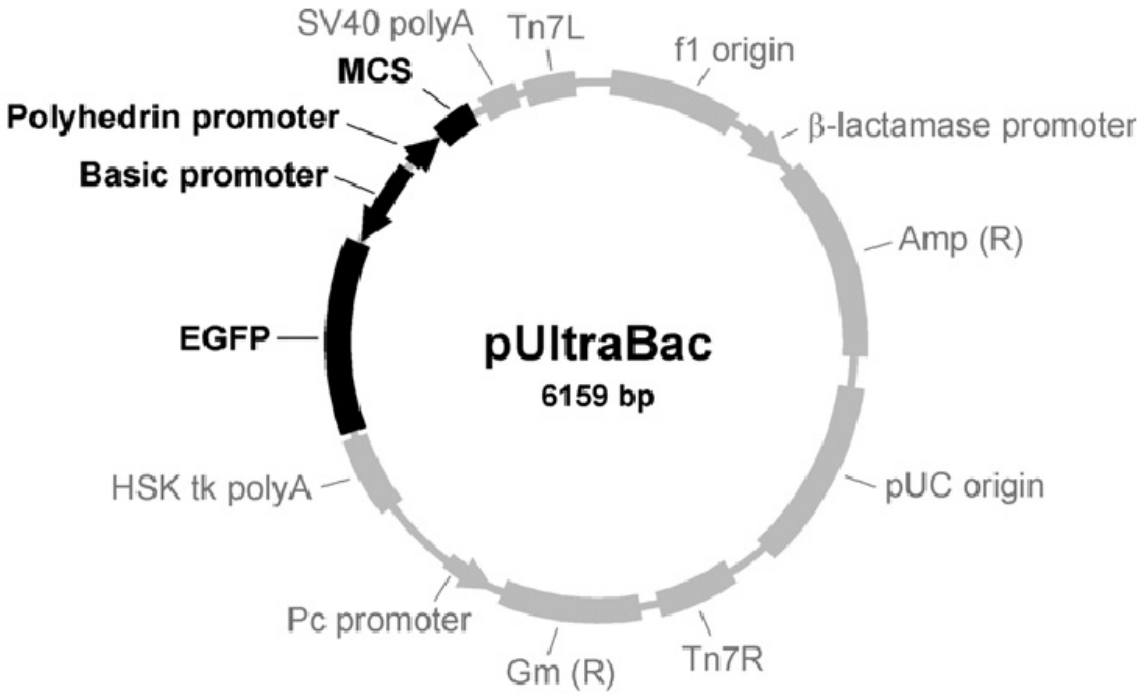


Figure 3.14 pUAST vector map. This vector contains a *Drosophila* hsp70 promoter. Image adapted from the *Drosophila* Genomics Resource Center (http://dgrc.cgb.indiana.edu).



Multiple cloning site (MCS)

```
                                BamHI
                    NcoI        RsrII       BssHII    EcoRI     StuI        SalI
CACCATCGGG CGCCATGGAT CCCGGTCCGA AGCGCGCGGA ATTCAAAGGC CTACGTCGAC
GTGGTAGCCC GCGGTACCTA GGGCCAGGCT TCGCGCGCCT TAAGTTTCCG GATGCAGCTG

          SpeI                              XbaI
SacI                NotI        BstBI               PstI          HindIII
GAGCTCACTA GTCGCGGCCG CTTTCGAATC TAGAGCCTGC AGTCTCGACA AGCTTGTCGA
CTCGAGTGAT CAGCGCCGGC GAAAGCTTAG ATCTCGGACG TCAGAGCTGT TCGAACAGCT
```

Figure 3.15 pUltraBac-1 vector map. This vector contains the EGFP gene under the control of the lepidopteran active late basic protein promoter (P_{Basic}) to allow fast titering of recombinant baculoviruses in Sf9 insect cells. Adapted from (133).

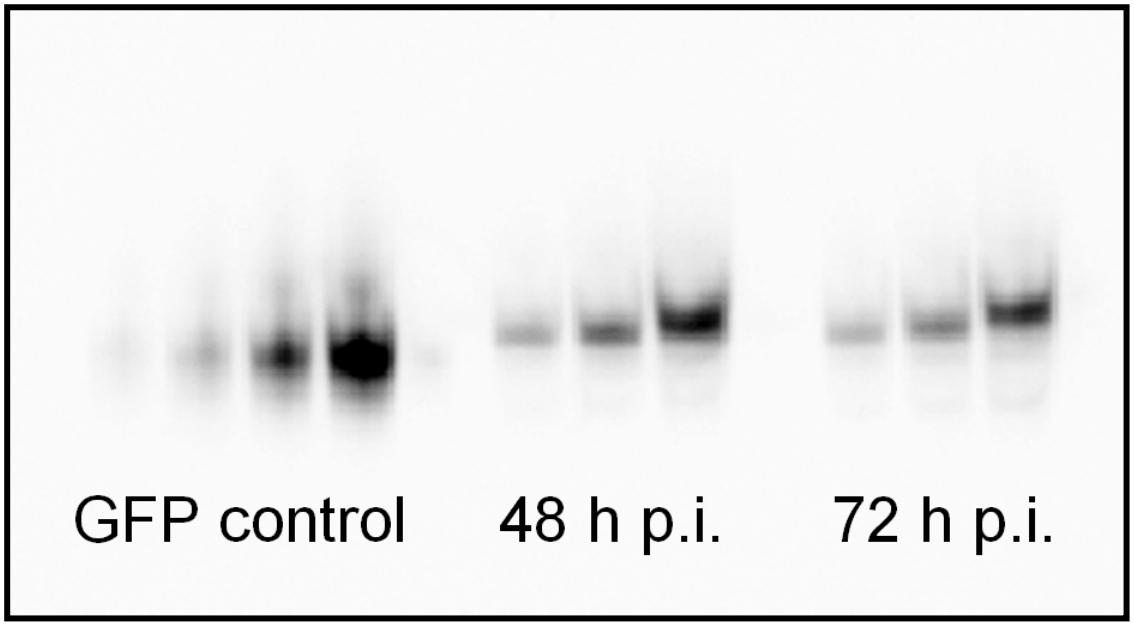


Figure 3.16 Fluorescence SDS-PAGE results showing EGFP production in Sf9 cells infected with pUltraBac-1-hsp70-AcGFP recombinant baculovirus in commercial ESF 921 media. The four GFP control lanes correspond to 1, 2, 4, and 8 µg of GFP. Forty-eight h p. i. and 72 h p.i. correspond to Sf9 whole cell lysates harvested 48 h and 72 h p.i. and loaded in triplicate in increasing amounts.

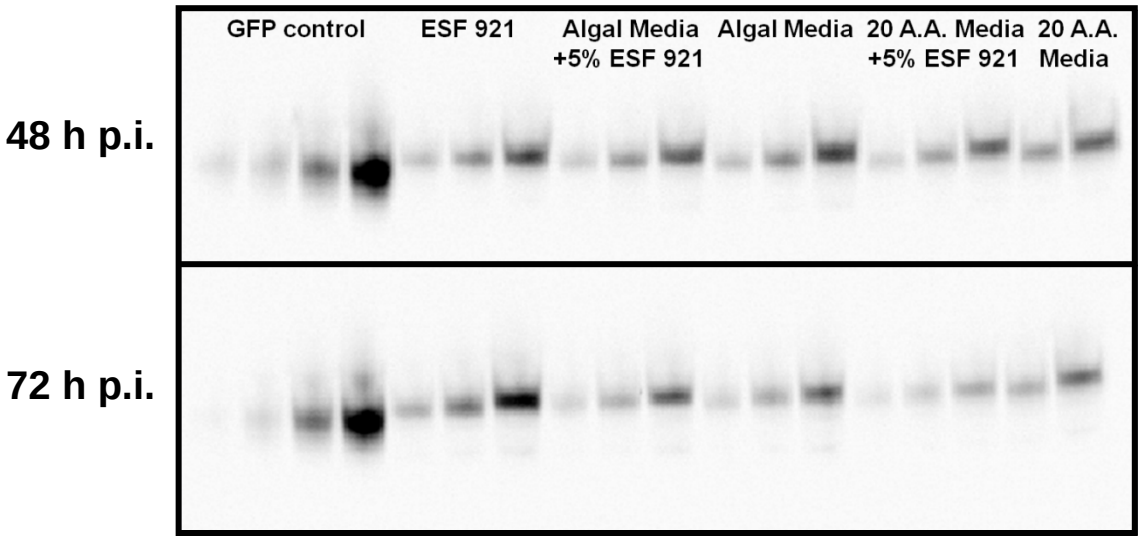


Figure 3.17 Fluorescence SDS-PAGE results showing EGFP production in Sf9 cells infected with pUltraBac-1-hsp70-AcGFP recombinant baculovirus in custom media. The four GFP control lanes correspond to 1, 2, 4, and 8 µg of GFP. Forty-eight h p.i. and 72 h p.i. correspond to Sf9 whole cell lysates harvested 48 h and 72 h p.i. and loaded in triplicate in increasing amounts.

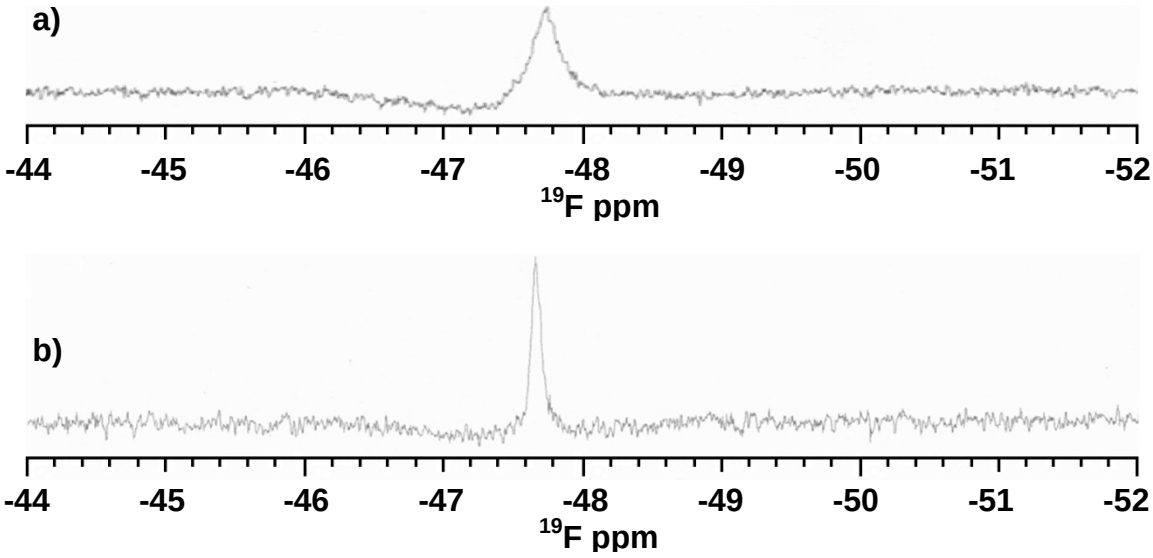


Figure 3.18 ^{19}F spectra of pUltraBac-1-hsp70-AcGFP recombinant baculovirus infected Sf9 cells producing EGFP in 5-FW custom algal medium 48 h p.i. a) Spectrum of the cell slurry. b) Spectrum of the Sf9 cleared lysate.

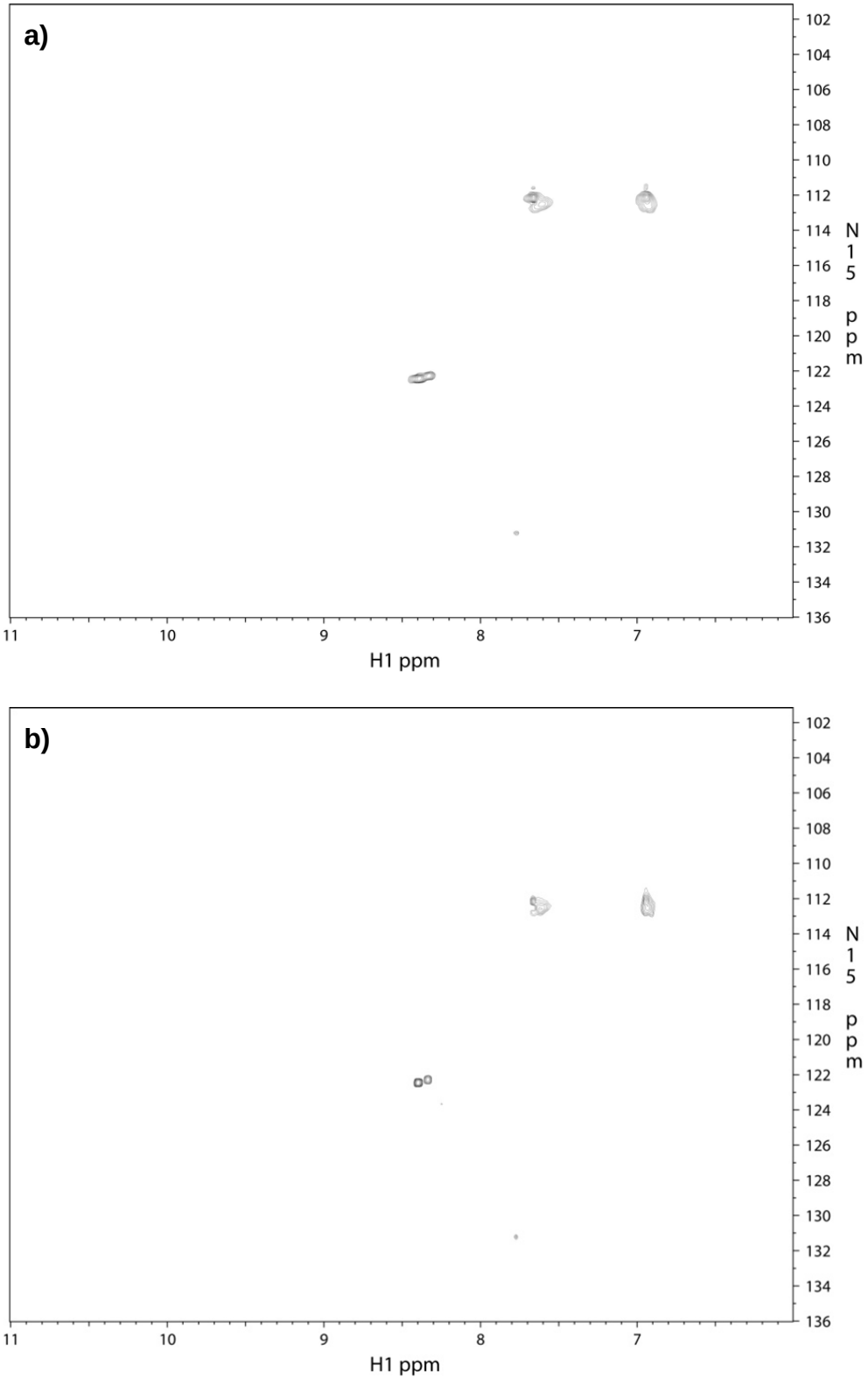


Figure 3.19 HSQC spectra of pUltraBac-1-hsp70-AcGFP recombinant baculovirus infected Sf9 cells producing EGFP in $^{15}NH_4Cl$ custom medium. a) Spectrum of the cell slurry 48 h p.i. b) Spectrum of the Sf9 cleared lysate 96 h p.i.

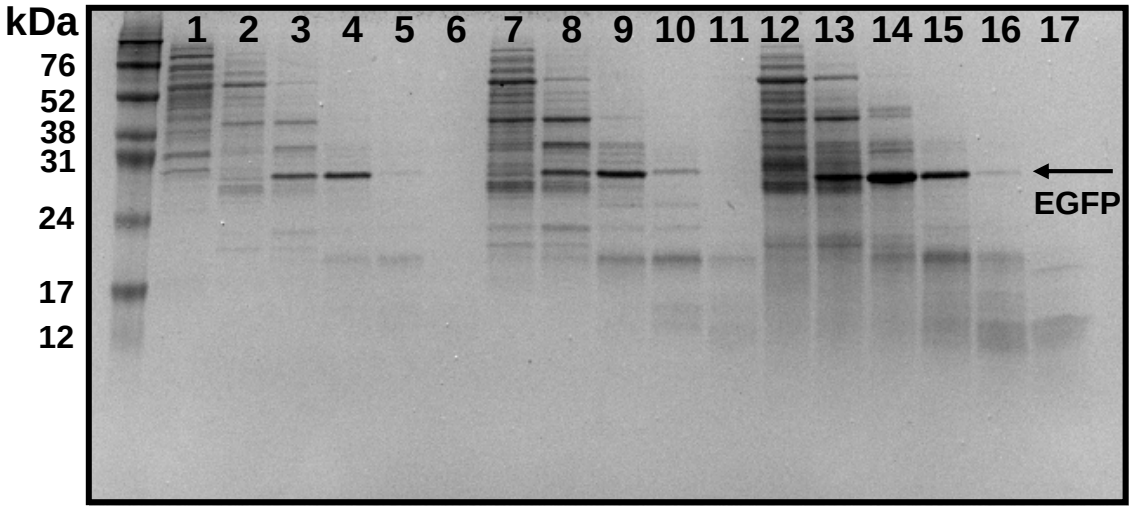


Figure 3.20 SDS-PAGE results of EGFP purified by size exclusion chromatography from lysates of Sf9 cells infected with pUltraBac-1-hsp70-AcGFP recombinant baculovirus and harvested 48 h p.i. Cells were grown in $^{15}NH_4Cl$ custom media. Lanes 1-6, 7-11, and 12-17 contain fractions from three separate purifications. Fractions containing EGFP were pooled.

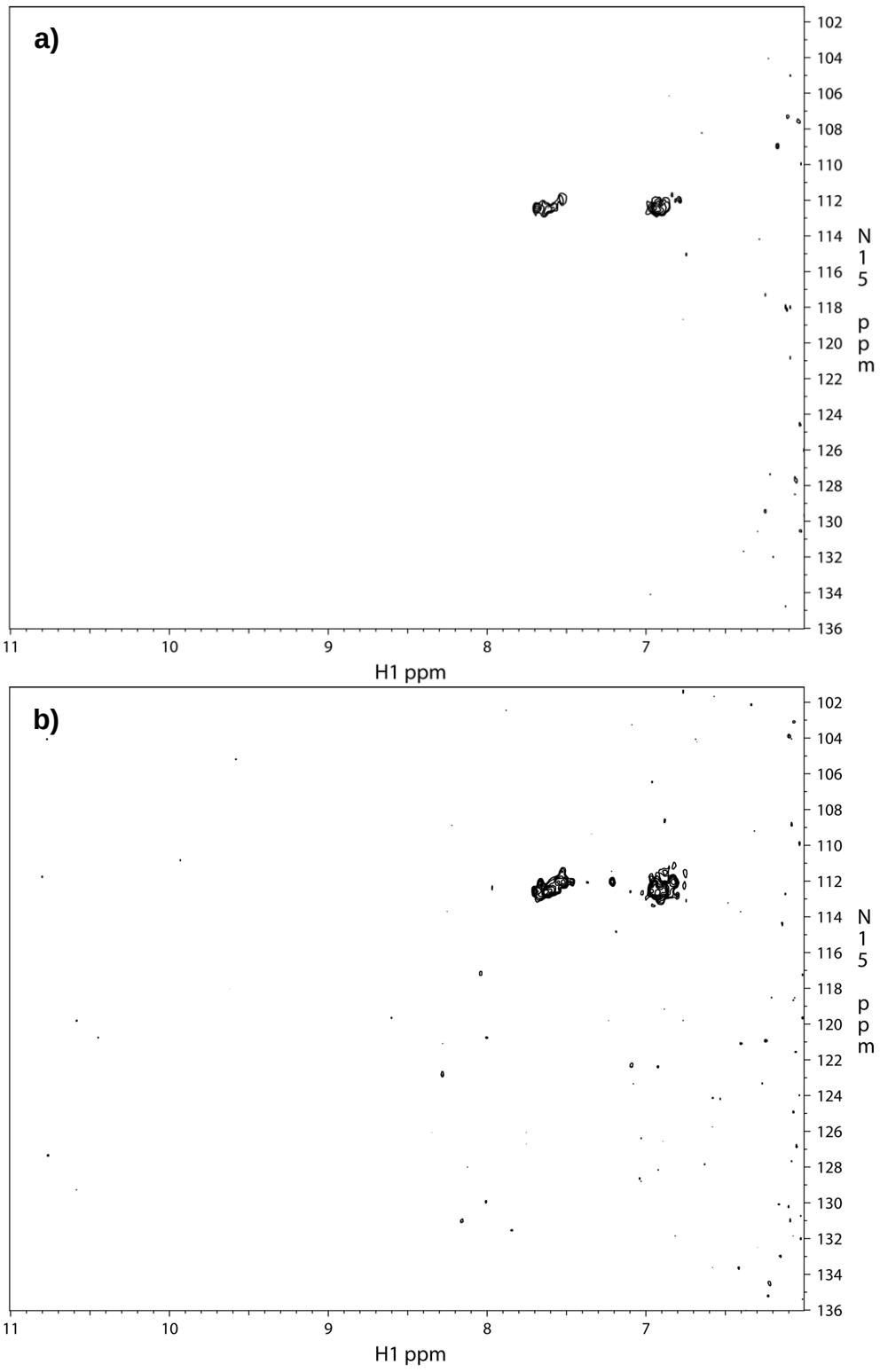


Figure 3.21 HSQC spectra of purified EGFP from Sf9 cells infected with pUltraBac-1-hsp70-AcGFP recombinant baculovirus in a) 20 mM sodium phosphate buffer, pH 6.0, and in b) 200 mM NaCl, 20 mM sodium phosphate buffer, pH 6.0.

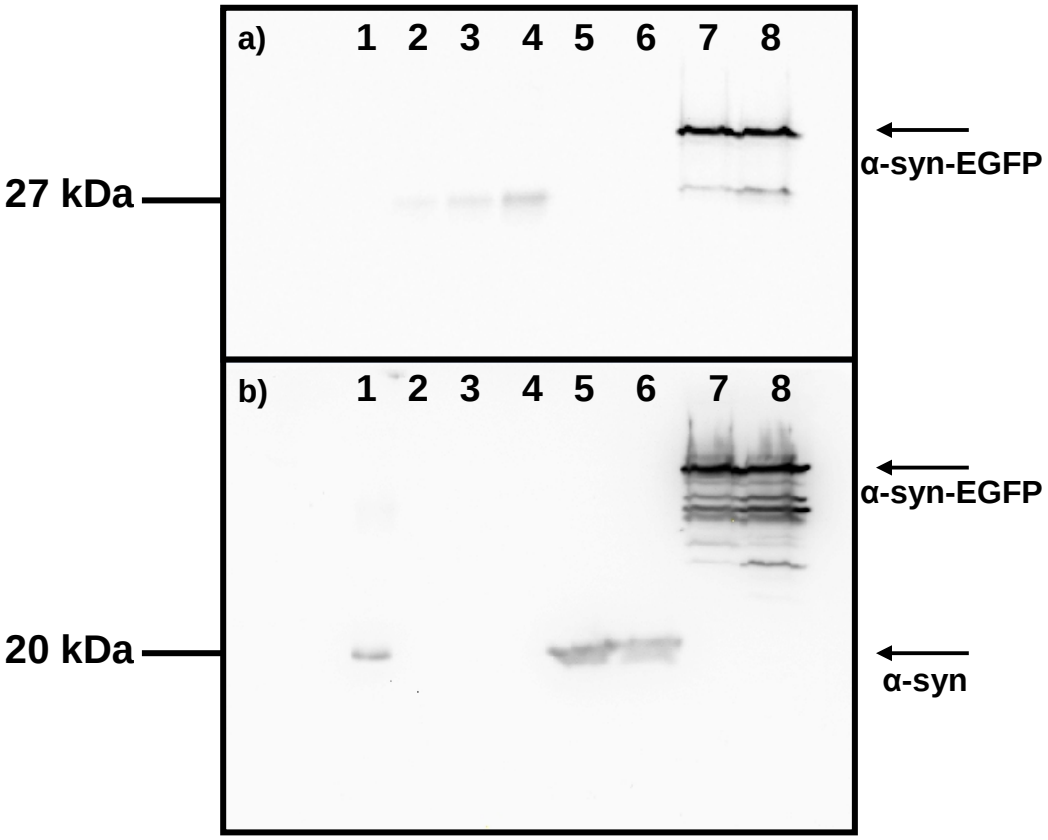


Figure 3.22 Detection of α-syn and α-syn-EGFP from whole cell lysates of Sf9 cells infected with α-syn pUltraBac-1 or α-syn-EGFP pUltraBac-1 recombinant baculoviruses in commercial ESF 921 media. a) Fluorescence image of SDS polyacrylamide gel for detection of EGFP. Lane 1 contains 1.9 µg α-syn positive control. Lanes 2-4 contain AcGFP positive controls loaded at 2, 4, and 8 µg. Lanes 5 and 6 contain lysates of Sf9 cells infected with α-syn pUltraBac-1 recombinant baculovirus and harvested 48 h and 72 h p.i., respectively. Lanes 7 and 8 contain lysates of Sf9 cells infected with α-syn-EGFP pUltraBac-1 recombinant baculovirus and harvested 48 h and 72 h p.i., respectively. b) Western blot of gel imaged in a) to detect the presence of α-syn. Lane identifications are the same as in a).

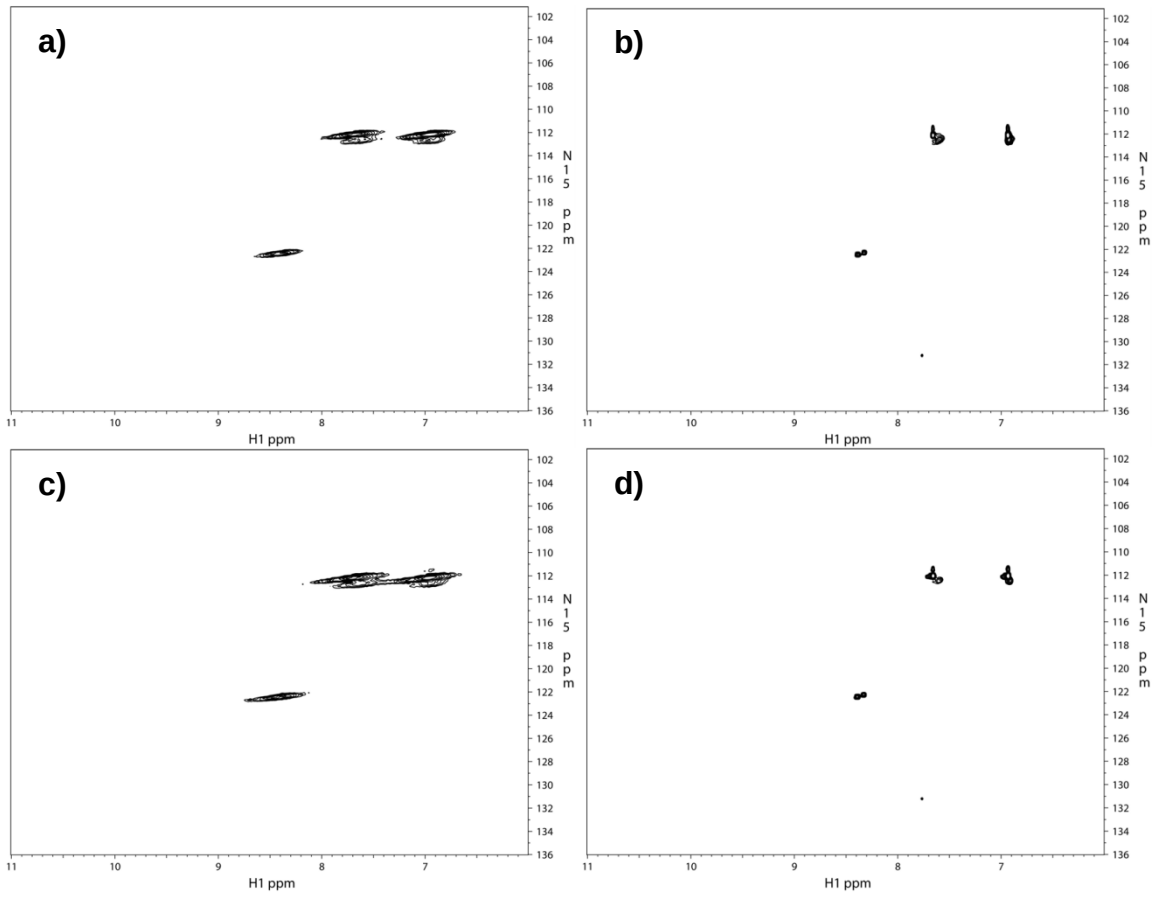


Figure 3.23 HSQC spectra of Sf9 cells infected with α-syn pUltraBac-1 or α-syn-EGFP pUltraBac-1 recombinant baculoviruses and grown in $^{15}NH_4Cl$ custom media. a) Spectrum of the α-syn pUltraBac-1 recombinant baculovirus infected cell slurry 43 h p.i. b) Lysate of the spectrum shown in a). c) Spectrum of the α-syn-EGFP pUltraBac-1 recombinant baculovirus infected cell slurry 43 h p.i. d) Lysate of the spectrum shown in c).

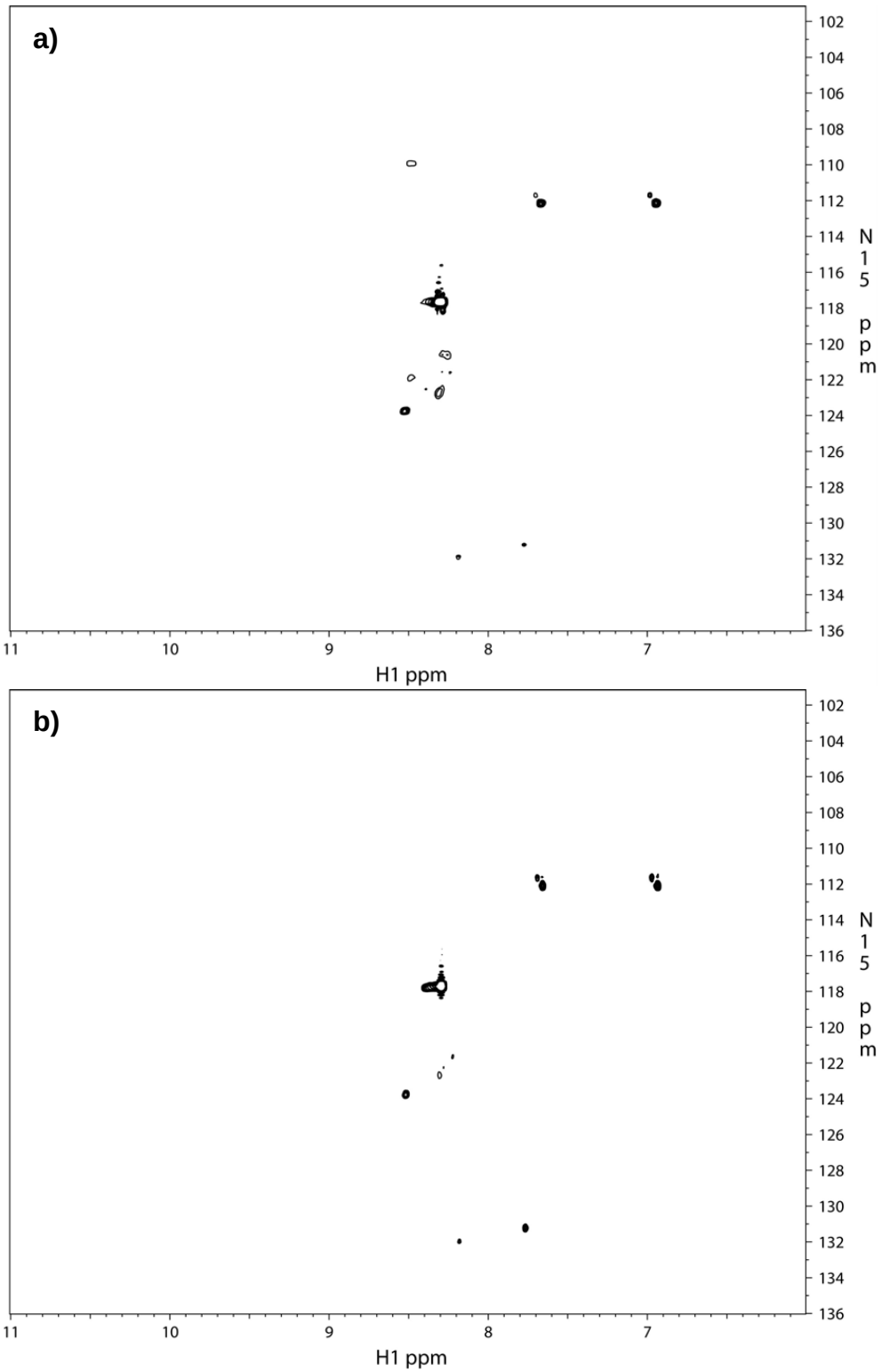


Figure 3.24 HSQC spectra of Sf9 cells infected with α-syn-EGFP pUltraBac-1 recombinant baculovirus and grown in ^{15}N-enriched custom IPL-41 media without yeastolate. a) Spectrum of the cell slurry 44 h p.i. b) Lysate of the spectrum shown in a).

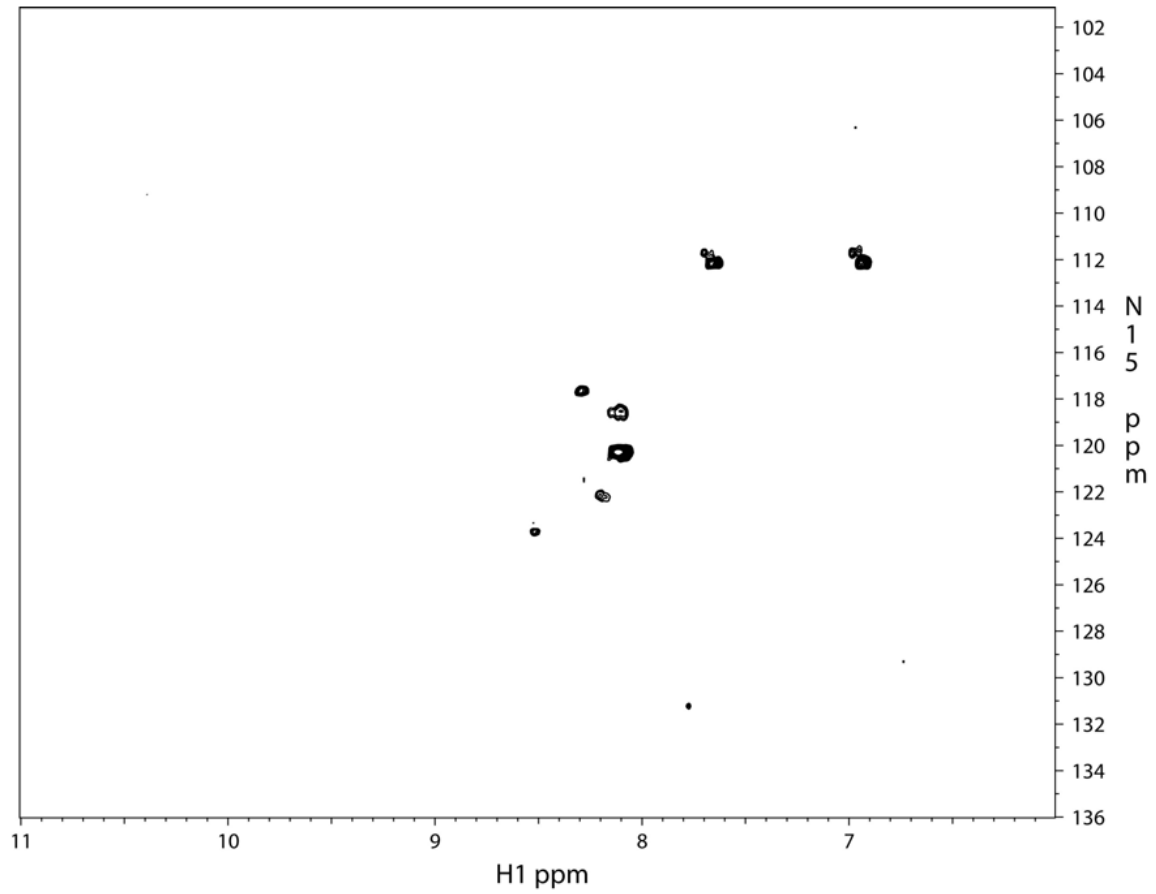


Figure 3.25 HSQC spectrum of Sf9 cells infected with α-syn pUltraBac-1 recombinant baculovirus and grown in ^{15}N-phenylalanine custom medium. The spectrum of the cell slurry was acquired 44 h p.i.

References

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

2. Elowitz MB, Surette MG, Wolf PE, Stock JB, Leibler S (1999) Protein mobility in the cytoplasm of *Escherichia coli*. *J Bacteriol* 181:197-203.

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

4. Fulton AB (1982) How crowded is the cytoplasm? *Cell* 30:345-347.

5. Cheung MS, Klimov D, Thirumalai D (2005) Molecular crowding enhances native state stability and refolding rates of globular proteins. *Proc Natl Acad Sci USA* 102:4753-4758.

6. Schellman JA (2003) Protein stability in mixed solvents: a balance of contact interaction and excluded volume. *Biophys J* 85:108-125.

7. Zimmerman SB, Minton AP (1993) Macromolecular crowding: biochemical, biophysical, and physiological consequences. *Annu Rev Biophys Biomol Struct* 22:27-65.

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

9. Shearwin KE, Winzor DJ (1990) Thermodynamic nonideality as a probe of reversible protein unfolding effected by variations in pH and temperature: studies of ribonuclease. *Arch Biochem Biophys* 282:297-301.

10. Baskakov I, Bolen DW (1998) Forcing thermodynamically unfolded proteins to fold. *J Biol Chem* 273:4831-4834.

11. Timasheff SN (1993) The control of protein stability and association by weak interactions with water: how do solvents affect these processes? *Annu Rev Biophys Biomol Struct* 22:67-97.

12. van den Berg B, Ellis RJ, Dobson CM (1999) Effects of macromolecular crowding on protein folding and aggregation. *EMBO J* 18:6927-6933.

13. Perham M, Stagg L, Wittung-Stafshede P (2007) Macromolecular crowding increases structural content of folded proteins. *FEBS Lett* 581:5065-5069.

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

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

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

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

18. McNulty BC, Young GB, Pielak GJ (2006) Macromolecular crowding in the *Escherichia coli* periplasm maintains α-synuclein disorder. *J Mol Biol* 355:893-897.

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

20. Zimmerman SB, Harrison B (1987) Macromolecular crowding increases binding of DNA polymerase to DNA: an adaptive effect. *Proc Natl Acad Sci USA* 84:1871-1875.

21. Zimmerman SB, Trach SO (1988) Macromolecular crowding extends the range of conditions under which DNA polymerase is functional. *Biochim Biophys Acta* 949:297-304.

22. Daughdrill GW, Chadsey MS, Karlinsey JE, Hughes KT, Dahlquist FW (1997) The C-terminal half of the anti-sigma factor, FlgM, becomes structured when bound to its target, σ^{28}. *Nat Struct Biol* 4:285-291.

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

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

25. Burz DS, Dutta K, Cowburn D, Shekhtman A (2006) Mapping structural interactions using in-cell NMR spectroscopy (STINT-NMR). *Nat Methods* 3:91-93.

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

27. Selenko P, Frueh DP, Elsaesser SJ, Haas W, Gygi SP, Wagner G (2008) *In situ* observation of protein phosphorylation by high-resolution NMR spectroscopy. *Nat Struct Mol Biol* 15:321-329.

28. Bodart JF, Wieruszeski JM, Amniai L, Leroy A, Landrieu I, Rousseau-Lescuyer A, Vilain JP, Lippens G (2008) NMR observation of tau in *Xenopus* oocytes. *J Magn Reson* 192:252-257.

29. Zhou HX, Rivas G, Minton AP (2008) Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences. *Annu Rev Biophys* 37:353-373.

30. Walter H, Brooks DE (1995) Phase separation in cytoplasm, due to macromolecular crowding, is the basis for microcompartmentation. *FEBS Lett* 361:135-139.

31. Long MS, Jones CD, Helfrich MR, Mangeney-Slavin LK, Keating CD (2005) Dynamic microcompartmentation in synthetic cells. *Proc Natl Acad Sci USA* 102:5920-5925.

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

33. Muramatsu N, Minton AP (1988) Tracer diffusion of globular proteins in concentrated protein solutions. *Proc Natl Acad Sci USA* 85:2984-2988.

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

35. Halpin RA, Hegeman GD, Kenyon GL (1981) Carbon-13 nuclear magnetic resonance studies of mandelate metabolism in whole bacterial cells and in isolated, *in vivo* cross-linked enzyme complexes. *Biochemistry* 20:1525-1533.

36. Aman RA, Kenyon GL, Wang CC (1985) Cross-linking of the enzymes in the glycosome of *Trypanosoma brucei*. *J Biol Chem* 260:6966-6973.

37. An S, Kumar R, Sheets ED, Benkovic SJ (2008) Reversible compartmentalization of *de novo* purine biosynthetic complexes in living cells. *Science* 320:103-106.

38. Selenko P, Wagner G (2007) Looking into live cells with in-cell NMR spectroscopy. *J Struct Biol* 158:244-253.

39. Serber Z, Ledwidge R, Miller SM, 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.

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

41. Landrieu I, Lacosse L, Leroy A, Wieruszeski JM, Trivelli X, Sillen A, Sibille N, Schwalbe H, Saxena K, Langer T, Lippens G (2006) NMR analysis of a tau phosphorylation pattern. *J Am Chem Soc* 128:3575-3583.

42. Hubbard JA, MacLachlan LK, King GW, Jones JJ, Fosberry AP (2003) Nuclear magnetic resonance spectroscopy reveals the functional state of the signalling protein CheY *in vivo* in *Escherichia coli*. *Mol Microbiol* 49:1191-1200.

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

44. Misra TK, Brown NL, Haberstroh L, Schmidt A, Goddette D, Silver S (1985) Mercuric reductase structural genes from plasmid R100 and transposon Tn*501*: functional domains of the enzyme. *Gene* 34:253-262.

45. Reckel S, Löhr F, Dötsch V (2005) In-cell NMR spectroscopy. *ChemBioChem* 6:1601-1606.

46. Barnes CO, Pielak GJ (2011) In-cell protein NMR and protein leakage. *Proteins: Struct, Funct, Bioinf* 79:347-351.

47. Samuel-Landtiser M, Zachariah C, Williams CR, Edison AS, Long JR (2007) Incorporation of isotopically enriched amino acids. *Curr Protoc Protein Sci* 26:3.1-3.49.

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

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

50. Gerig JT (1994) Fluorine NMR of proteins. *Prog Nucl Magn Reson Spectrosc* 26:293-370.

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

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

53. Wikström M, Drakenberg T, Forsén S, Sjöbring U, Björck L (1994) Three-dimensional solution structure of an immunoglobulin light chain-binding domain of protein L. Comparison with the IgG-binding domains of protein G. *Biochemistry* 33:14011-14017.

54. Yi Q, Baker D (1996) Direct evidence for a two-state protein unfolding transition from hydrogen-deuterium exchange, mass spectrometry, and NMR. *Protein Sci* 5:1060-1066.

55. Tadeo X, López-Méndez B, Trigueros T, Laín A, Castaño D, Millet O (2009) Structural basis for the aminoacid composition of proteins from halophilic *Archea*. *PLoS Biol* 7:e1000257.

56. Cayley S, Lewis BA, Guttman HJ, Record MT (1991) Characterization of the cytoplasm of *Escherichia coli* K-12 as a function of external osmolarity: implications for protein-DNA interactions *in vivo*. *J Mol Biol* 222:281-300.

57. Shevchik VE, Condemine G, 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.

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

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

60. Bodenhausen G, Ruben DJ (1980) Natural abundance ^{15}N NMR by enhanced heteronuclear spectroscopy. *Chem Phys Lett* 69:185-189.

61. Santoro MM, Bolen DW (1988) Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl *α*-chymotrypsin using different denaturants. *Biochemistry* 27:8063-8068.

62. Frieden C, Hoeltzli SD, Bann JG (2004) The preparation of ^{19}F-labeled proteins for NMR studies. *Methods Enzymol* 380:400-415.

63. Murphy JP, Trowern AR, Duggleby CJ (1994) Nucleotide sequence of the gene for peptostreptococcal protein L. *DNA Seq* 4:259-265.

64. Khan F, Kuprov I, Craggs TD, Hore PJ, Jackson SE (2006) ^{19}F NMR studies of the native and denatured states of green fluorescent protein. *J Am Chem Soc* 128:10729-10737.

65. Zhang O, Forman-Kay JD (1995) Structural characterization of folded and unfolded states of an SH3 domain in equilibrium in aqueous buffer. *Biochemistry* 34:6784-6794.

66. Dyson HJ, Wright PE (1996) Insights into protein folding from NMR. *Annu Rev Phys Chem* 47:369-395.

67. McDonald CC, Phillips WD (1969) Proton magnetic resonance spectra of proteins in random-coil configurations. *J Am Chem Soc* 91:1513-1521.

68. Sharaf NG, Barnes CO, Charlton LM, Young GB, Pielak GJ (2010) A bioreactor for in-cell protein NMR. *J Magn Reson* 202:140-146.

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

70. Pielak GJ, Li C, Miklos AC, Schlesinger AP, Slade KM, Wang GF, Zigoneanu IG (2009) Protein nuclear magnetic resonance under physiological conditions. *Biochemistry* 48:226-234.

71. Richey B, Cayley DS, Mossing MC, Kolka C, Anderson CF, Farrar TC, Record MT, Jr. (1987) Variability of the intracellular ionic environment of *Escherichia coli*: differences between *in vitro* and *in vivo* effects of ion concentrations on protein-DNA interactions and gene expression. *J Biol Chem* 262:7157-7164.

72. Konopka MC, Weisshaar JC, Record MT, Jr. (2007) Methods of changing biopolymer volume fraction and cytoplasmic solute concentrations for *in vivo* biophysical studies. *Methods Enzymol* 428:487-504.

73. Li C, Pielak GJ (2009) Using NMR to distinguish viscosity effects from non-specific protein binding under crowded conditions. *J Am Chem Soc* 131:1368-1369.

74. Miklos AC, Li C, Sharaf NG, Pielak GJ (2010) Volume exclusion and soft interaction effects on protein stability under crowded conditions. *Biochemistry* 49:6984-6991.

75. Crowley PB, Brett K, Muldoon J (2008) NMR spectroscopy reveals cytochrome *c*-poly(ethylene glycol) interactions. *ChemBioChem* 9:685-688.

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

77. Ruschak AM, Religa TL, Breuer S, Witt S, Kay LE (2010) The proteasome antechamber maintains substrates in an unfolded state. *Nature* 467:868-871.

78. Plaza del Pino IM, Ibarra-Molero B, Sanchez-Ruiz JM (2000) Lower kinetic limit to protein thermal stability: a proposal regarding protein stability *in vivo* and its relation with misfolding diseases. *Proteins* 40:58-70.

79. Dunker AK, Silman I, Uversky VN, Sussman JL (2008) Function and structure of inherently disordered proteins. *Curr Opin Struct Biol* 18:756-764.

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

81. Swaminathan R, Hoang CP, Verkman AS (1997) Photobleaching recovery and anisotropy decay of green fluorescent protein GFP-S65T in solution and cells: cytoplasmic viscosity probed by green fluorescent protein translational and rotational diffusion. *Biophys J* 72:1900-1907.

82. Srere PA (1980) The infrastructure of the mitochondrial matrix. *Trends Biochem Sci* 5:120-121.

83. Partikian A, Ölveczky B, Swaminathan R, Li Y, Verkman AS (1998) Rapid diffusion of green fluorescent protein in the mitochondrial matrix. *J Cell Biol* 140:821-829.

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

85. Smith GE, Fraser MJ, Summers MD (1983) Molecular engineering of the *Autographa californica* nuclear polyhedrosis virus genome: deletion mutations within the polyhedrin gene. *J Virol* 46:584-593.

86. Smith GE, Summers MD, Fraser MJ (1983) Production of human beta interferon in insect cells infected with a baculovirus expression vector. *Mol Cell Biol* 3:2156-2165.

87. Pennock GD, Shoemaker C, Miller LK (1984) Strong and regulated expression of *Escherichia coli* β-galactosidase in insect cells with a baculovirus vector. *Mol Cell Biol* 4:399-406.

88. Maeda S, Kawai T, Obinata M, Fujiwara H, Horiuchi T, Saeki Y, Sato Y, Furusawa M (1985) Production of human *α*-interferon in silkworm using a baculovirus vector. *Nature* 315:592-594.

89. O'Reilly DR, Miller LK, Luckow VA (1992) Baculovirus expression vectors: a laboratory manual. WH Freeman and Company (New York, NY).

90. Gierse JK, Luckow VA, Askonas LJ, Duffin KL, Aykent S, Bild GS, Rodi CP, Sullivan PM, Bourner MJ, Kimack NM, Krivi GG (1993) High-level expression and purification of human leukotriene A_4 hydrolase from insect cells infected with a baculovirus vector. *Protein Expression Purif* 4:358-366.

91. McNulty BC, Tripathy A, Young GB, Charlton LM, Orans J, Pielak GJ (2006) Temperature-induced reversible conformational change in the first 100 residues of α-synuclein. *Protein Sci* 15:602-608.

92. Spillantini MG, Schmidt ML, Lee VM, Trojanowski JQ, Jakes R, Goedert M (1997) α-Synuclein in Lewy bodies. *Nature* 388:839-840.

93. Lansbury PT, Jr., Brice A (2002) Genetics of Parkinson's disease and biochemical studies of implicated gene products. *Curr Opin Genet Dev* 12:299-306.

94. Spillantini MG, Crowther RA, Jakes R, Hasegawa M, Goedert M (1998) α-Synuclein in filamentous inclusions of Lewy bodies from Parkinson's disease and dementia with Lewy bodies. *Proc Natl Acad Sci USA* 95:6469-6473.

95. Li C, Lutz EA, Slade KM, Ruf RA, Wang GF, Pielak GJ (2009) ^{19}F NMR studies of α-synuclein conformation and fibrillation. *Biochemistry* 48:8578-8584.

96. Morar AS, Olteanu A, Young GB, Pielak GJ (2001) Solvent-induced collapse of α-synuclein and acid-denatured cytochrome *c*. *Protein Sci* 10:2195-2199.

97. Shtilerman MD, Ding TT, Lansbury PT, Jr. (2002) Molecular crowding accelerates fibrillization of α-synuclein: could an increase in the cytoplasmic protein concentration induce Parkinson's disease? *Biochemistry* 41:3855-3860.

98. Uversky VN, Cooper EM, Bower KS, Li J, Fink AL (2002) Accelerated α-synuclein fibrillation in crowded milieu. *FEBS Lett* 515:99-103.

99. Nagy IZ, Nagy K, Lustyik G (1982) Protein and water contents of aging brain. *Exp Brain Res* Suppl 5:118-122.

100. McNaught KS, Olanow CW, Halliwell B, Isacson O, Jenner P (2001) Failure of the ubiquitin-proteasome system in Parkinson's disease. *Nat Rev Neurosci* 2:589-594.

101. Volles MJ, Lansbury PT, Jr. (2003) Zeroing in on the pathogenic form of α-synuclein and its mechanism of neurotoxicity in Parkinson's disease. *Biochemistry* 42:7871-7878.

102. Flaugh SL, Lumb KJ (2001) Effects of macromolecular crowding on the intrinsically disordered proteins c-Fos and $p27^{Kip1}$. *Biomacromolecules* 2:538-540.

103. DeLange F, Klaassen CHW, Wallace-Williams SE, Bovee-Geurts PHM, Liu XM, DeGrip WJ, Rothschild KJ (1998) Tyrosine structural changes detected during the photoactivation of rhodopsin. *J Biol Chem* 273:23735-23739.

104. Creemers AFL, Klaassen CHW, Bovee-Geurts PHM, Kelle R, Kragl U, Raap J, de Grip WJ, Lugtenburg J, de Groot HJM (1999) Solid state ^{15}N NMR evidence for a complex Schiff base counterion in the visual G-protein-coupled receptor rhodopsin. *Biochemistry* 38:7195-7199.

105. Brüggert M, Rehm T, Shanker S, Georgescu J, Holak TA (2003) A novel medium for expression of proteins selectively labeled with ^{15}N-amino acids in *Spodoptera frugiperda* (Sf9) insect cells. *J Biomol NMR* 25:335-348.

106. Ellis CE, Schwartzberg PL, Grider TL, Fink DW, Nussbaum RL (2001) *α*-Synuclein is phosphorylated by members of the Src family of protein-tyrosine kinases. *J Biol Chem* 276:3879-3884.

107. Reilander H, Haase W, Maul G (1996) Functional expression of the *Aequorea victoria* green fluorescent protein in insect cells using the baculovirus expression system. *Biochem Biophys Res Commun* 219:14-20.

108. King LA, Thomas CJ, Wilkinson N, Possee RD (1999) Expression of green fluorescent protein using baculovirus vectors. *Methods Enzymol* 302:394-408.

109. Khan F, Stott K, Jackson S (2003) ^{1}H, ^{15}N and ^{13}C backbone assignment of the green fluorescent protein (GFP). *J Biomol NMR* 26:281-282.

110. Walton WJ, Kasprzak AJ, Hare JT, Logan TM (2006) An economic approach to isotopic enrichment of glycoproteins expressed from Sf9 insect cells. *J Biomol NMR* 36:225-233.

111. Slade KM (2009) Protein diffusion in *Escherichia coli*. *Ph.D. dissertation*: University of North Carolina at Chapel Hill, NC.

112. Murhammer DW, Ed. (2007) Baculovirus and insect cell expression protocols, 2nd ed. *Methods in molecular biology 388*: Humana Press (Totowa, NJ); Chapter 10.

113. Walton WJ (2006) Isotopic enrichment and receptor-binding analysis of insect and Lec1 cell expressed avian Thy-1. *Ph.D. dissertation*: Florida State University, FL.

114. Eriksson S, Raivio E, Kukkonen JP, Eriksson K, Lindqvist C (1996) Green fluorescent protein as a tool for screening recombinant baculoviruses. *J Virol Methods* 59:127-133.

115. Zhang J, Collins A, Chen M, Knyazev I, Gentz R (1998) High-density perfusion culture of insect cells with a biosep ultrasonic filter. *Biotechnol Bioeng* 59:351-359.

116. Strauss A, Bitsch F, Fendrich G, Graff P, Knecht R, Meyhack B, Jahnke W (2005) Efficient uniform isotope labeling of Abl kinase expressed in Baculovirus-infected insect cells. *J Biomol NMR* 31:343-349.

117. Drews M, Doverskog M, Öhman L, Chapman BE, Jacobsson U, Kuchel PW, Häggström L (2000) Pathways of glutamine metabolism in *Spodoptera frugiperda* (Sf9) insect cells: evidence for the presence of the nitrogen assimilation system, and a metabolic switch by $^{1}H/^{15}N$ NMR. *J Biotechnol* 78:23-37.

118. Loomis KH, Yaeger KW, Batenjany MM, Mehler MM, Grabski AC, Wong SC, Novy RE (2005) InsectDirect system: rapid, high-level protein expression and purification from insect cells. *J Struct Funct Genomics* 6:189-194.

119. Jarvis DL, Weinkauf C, Guarino LA (1996) Immediate-early baculovirus vectors for foreign gene expression in transformed or infected insect cells. *Protein Expr Purif* 8:191-203.

120. Li MZ, Elledge SJ (2007) Harnessing homologous recombination *in vitro* to generate recombinant DNA via SLIC. *Nat Methods* 4:251-256.

121. Lowry OH, Rosebrough NJ, Farr AL, Randall RJ (1951) Protein measurement with the Folin phenol reagent. *J Biol Chem* 193:265-275.

122. Li M (2010) Can we determine a protein structure quickly? *J Comput Sci Technol* 25:95-106.

123. Yu HT (1999) Extending the size limit of protein nuclear magnetic resonance. *Proc Natl Acad Sci USA* 96:332-334.

124. Boyce FM, Bucher NL (1996) Baculovirus-mediated gene transfer into mammalian cells. *Proc Natl Acad Sci USA* 93:2348-2352.

125. Hofmann C, Sandig V, Jennings G, Rudolph M, Schlag P, Strauss M (1995) Efficient gene transfer into human hepatocytes by baculovirus vectors. *Proc Natl Acad Sci USA* 92:10099-10103.

126. Condreay JP, Witherspoon SM, Clay WC, Kost TA (1999) Transient and stable gene expression in mammalian cells transduced with a recombinant baculovirus vector. *Proc Natl Acad Sci USA* 96:127-132.

127. Liang CY, Wang HZ, Li TX, Hu ZH, Chen XW (2004) High efficiency gene transfer into mammalian kidney cells using baculovirus vectors. *Arch Virol* 149:51-60.

128. Sandig V, Hofmann C, Steinert S, Jennings G, Schlag P, Strauss M (1996) Gene transfer into hepatocytes and human liver tissue by baculovirus vectors. *Hum Gene Ther* 7:1937-1945.

129. Sarkis C, Serguera C, Petres S, Buchet D, Ridet JL, Edelman L, Mallet J (2000) Efficient transduction of neural cells *in vitro* and *in vivo* by a baculovirus-derived vector. *Proc Natl Acad Sci USA* 97:14638-14643.

130. Lee DF, Chen CC, Hsu TA, Juang JL (2000) A baculovirus superinfection system: efficient vehicle for gene transfer into *Drosophila* S2 cells. *J Virol* 74:11873-11880.

131. Condreay JP, Kost TA (2007) Baculovirus expression vectors for insect and mammalian cells. *Curr Drug Targets* 8:1126-1131.

132. Baum B, Cherbas L (2008) *Drosophila* cell lines as model systems and as an experimental tool. *Methods Mol Biol* 420:391-424.

133. Philipps B, Rotmann D, Wicki M, Mayr LM, Forstner M (2005) Time reduction and process optimization of the baculovirus expression system for more efficient recombinant protein production in insect cells. *Protein Expr Purif* 42:211-218.

134. Hill-Perkins MS, Possee RD (1990) A baculovirus expression vector derived from the basic protein promoter of *Autographa californica* nuclear polyhedrosis virus. *J Gen Virol* 71:971-976.

135. Seifert MH, Ksiazek D, Azim MK, Smialowski P, Budisa N, Holak TA (2002) Slow exchange in the chromophore of a green fluorescent protein variant. *J Am Chem Soc* 124:7932-7942.

136. Campos-Olivas R, Aziz R, Helms GL, Evans JN, Gronenborn AM (2002) Placement of ^{19}F into the center of GB1: effects on structure and stability. *FEBS Lett* 517:55-60.

137. Serber Z, Corsini L, Durst F, Dötsch V (2005) In-cell NMR spectroscopy. *Methods Enzymol* 394:17-41.

138. Shen CF, Kiyota T, Jardin B, Konishi Y, Kamen A (2007) Characterization of yeastolate fractions that promote insect cell growth and recombinant protein production. *Cytotechnology* 54:25-34.

139. Ikonomou L, Bastin G, Schneider YJ, Agathos SN (2001) Design of an efficient medium for insect cell growth and recombinant protein production. *In Vitro Cell Dev Biol-Animal* 37:549-559.

140. Doverskog M, Ljunggren J, Öhman L, Häggström L (1997) Physiology of cultured animal cells. *J Biotechnol* 59:103-115.

141. Bedard C, Tom R, Kamen A (1993) Growth, nutrient consumption, and end-product accumulation in Sf9 and BTI-EAA insect cell cultures: insights into growth limitation and metabolism. *Biotechnol Prog* 9:615-624.

142. Phillips GN, Jr. (1997) Structure and dynamics of green fluorescent protein. *Curr Opin Struct Biol* 7:821-827.

143. Weiss SA, Smith GC, Kalter SS, Vaughn JL (1981) Improved method for the production of insect cell-cultures in large volume. *In Vitro* 17:495-502.

144. Maiorella B, Inlow D, Shauger A, Harano D (1988) Large-scale insect cell-culture for recombinant protein-production. *Nat Biotechnol* 6:1406-1410.

145. Olteanu A, Patel CN, Dedmon MM, Kennedy S, Linhoff MW, Minder CM, Potts PR, Deshmukh M, Pielak GJ (2003) Stability and apoptotic activity of recombinant human cytochrome *c*. *Biochem Biophys Res Commun* 312:733-740.

146. Burz DS, Shekhtman A (2008) In-cell biochemistry using NMR spectroscopy. *PLoS One* 3:e2571.