EFFECTS OF ANTIMICROBIAL PRESERVATIVES ON THE PARTIAL UNFOLDING AND AGGREGATION OF MODEL AND PHARMACEUTICAL PROTEINS

by

REGINA L. BIS

B.S., Colorado School of Mines, 2007

A thesis submitted to the
Faculty of the Graduate School of the
University of Colorado in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Pharmaceutical Sciences Program
2014
This thesis for the Doctor of Philosophy degree by

Regina L. Bis

has been approved for the

Pharmaceutical Sciences Program

by

David L. Bain, Chair
Krishna M.G. Mallela, Advisor
John F. Carpenter
Thomas J. Anchordoquy
Theodore W. Randolph

Date: 04/30/14
Bis, Regina L. (Ph.D., Pharmaceutical Sciences)

Effects of Antimicrobial Preservatives on the Partial Unfolding and Aggregation of Model and Pharmaceutical Proteins

Thesis directed by Assistant Professor Krishna M.G. Mallela

ABSTRACT

Multi-dose protein therapeutics compose approximately one-third of all protein-based drugs on the worldwide market. These formulations include at least one antimicrobial preservative (AP) to combat the growth of bacteria and other microbes during repeated contact between the solution and syringe needle. While multi-dose therapeutics are beneficial in terms of patient compliance and economics, the necessary inclusion of APs has been linked to protein aggregation in the liquid state. However, the mechanism of AP-induced protein aggregation remains unclear. In this work we examine the effects of commonly used APs on the unfolding and aggregation of a model protein cytochrome c (Cyt c) and a pharmaceutical protein interferon alpha-2a (IFNA2). We demonstrate that APs cause Cyt c unfolding and aggregation in the specific pattern m-cresol (CR) > phenol (PH) > benzyl alcohol (BA) > phenoxyethanol (PE) > chlorobutanol (CB). The aggregation reaction was preceded by the unfolding a local region of Cyt c that was shown to be an aggregation “hot-spot”. By stabilizing this region through specific modification, aggregation was inhibited. To translate these results to a pharmaceutically relevant system, we chose to study AP effects on IFNA2. We developed a novel bacterial expression and purification scheme to produce functional, correctly folded IFNA2 in-
house in order to investigate AP-induced aggregation. We discovered that APs lead to IFNA2 aggregation in the same pattern as shown with Cyt c, and the extent of aggregation was dependent upon the AP. Additionally, we observed that BA, the AP used in IFNA2 liquid formulation, perturbed a local region of IFNA2 that may lead to the formation of an aggregation-prone species. This “hot-spot” was identified as an area of high aggregation propensity. Through these studies, we show that the mechanism of AP-induced aggregation is independent of protein identity, and confirm that local or partial unfolding is the first step in the AP-induced aggregation reaction.

The form and content of this abstract are approved. I recommend its publication.

Approved: Krishna M.G. Mallela
ACKNOWLEDGMENTS

First, thank you to my advisor Krishna Mallela, for allowing me the opportunity to work, learn, and grow as a scientist. For their years of support, guidance, and conversation, I thank my committee members: Drs. Dave Bain, John Carpenter, Ted Randolph, LaToya Jones Braun, and Tom Anchordoquy. Thank you to my labmates for making the lab a place of family and enjoyment: Surinder, Swati, Jus, Dinen, Javier.

To Mr. Mike Thomas, for opening my eyes to the possibilities of science, and teaching me so many lessons that I have used both inside and out of the laboratory: God speaks physics, not law; and when in doubt, write a symbol.

To my friends who have known me from the beginning of this adventure and always believed that there would be an end, thank you, especially to BFAM, JB, MAD, my HA, Speck, Dr. Tron, HLP, NMP, BF³ and TMC. To the kids at College Inn, thank you for letting me write this thing at *my* table.

To my parents, words cannot express my gratitude for all you have done for me, for all of your encouraging words and smiles and love, and for giving me the chance to pursue crazy dreams.

Last, and in no way least, thank you to my husband Steve, for his unending love, confidence, and laughter. Without you, I would be lost.

This work was supported by the National Institutes of Health (Leadership training grant in Pharmaceutical Biotechnology, T32GM008732) and the Pharmaceutical Research and Manufacturers of America Foundation (Pre-doctoral fellowship in Pharmaceutical Sciences, AWD-120487).
# TABLE OF CONTENTS

## CHAPTER

### I. INTRODUCTION

Importance and Mechanisms of Biological and Biotechnological Aggregation .......................... 2

Aggregation in Natural Systems .............................................................................. 2

Aggregation in Biotechnology ............................................................................. 4

Preservative Effects on Proteins ........................................................................... 7

Cytochrome c as a Model Protein ........................................................................ 11

Interferon alpha-2a as a Pharmaceutical Protein .................................................. 13

### II. ANTIMICROBIAL PRESERVATIVES CAUSE PARTIAL PROTEIN UNFOLDING AND AGGREGATION OF CYTOCHROME C

Introduction ........................................................................................................ 17

Materials and Methods ....................................................................................... 20

Materials ............................................................................................................. 20

Preservative Efficacy Test .................................................................................... 21

Size Exclusion Chromatography ........................................................................ 21

Isothermal Incubation Experiments ................................................................... 22

Thermal Scanning Method ................................................................................... 23

Denaturant Melts ............................................................................................... 24

Nuclear Magnetic Resonance ............................................................................. 24

Hydrogen Exchange .......................................................................................... 25

Results ............................................................................................................... 25

Antimicrobial Activity of APs ............................................................................ 25
Design of the IFNA2 Expression Construct _______________________________ 53
Expression and Purification of Recombinant Human IFNA2 _________________ 53
Minor Loss Observed Following Freeze-Thaw of IFNA2 Aliquots ____________ 54
Confirmation of IFNA2 Identity Using Biophysical and Structural Methods____ 56
IFNA2 Retains Antiviral and Anti-proliferative Activity ____________________ 61
Discussion __________________________________________________________ 62

IV. ANTIMICROBIAL PRESERVATIVES INDUCE AGGREGATION OF INTERFERON ALPHA-2A: THE ORDER IN WHICH PRESERVATIVES INDUCE PROTEIN AGGREGATION IS INDEPENDENT OF THE NATURE OF THE PROTEIN _____________ 66

Introduction _________________________________________________________ 66
Materials and Methods ________________________________________________ 67
Materials __________________________________________________________ 67
Size Exclusion Chromatography _______________________________________ 67
Isothermal Incubation Experiments _____________________________________ 68
Thermal Scanning Method ____________________________________________ 69
Results _____________________________________________________________ 70
APs Cause IFNA2 Aggregation ________________________________________ 70
IFNA2 Aggregation Increases with AP Concentration ______________________ 71
Aggregation Temperature Decreases Linearly with AP Concentration_______ 73
Discussion __________________________________________________________ 74

V. THE ROLE OF BENZYL ALCOHOL IN THE UNFOLDING AND AGGREGATION OF INTERFERON ALPHA-2A ____________________________ 78

Introduction _________________________________________________________ 78
Materials and Methods ________________________________________________ 79
Materials __________________________________________________________ 79
APPENDIX

A. INVESTIGATION OF SOLUTION PARAMETERS ON CYTOCHROME C AGGREGATION IN THE PRESENCE OF ANTIMICROBIAL PRESERVATIVES 121

Simple Alcohols to Model Antimicrobial Preservatives 121

Using Fluorescent Dyes to Monitor Microenvironment Solution Parameters 125

Dialanine as a Model for the Peptide Bond 127

B. ISOTHERMAL CALORIMETRY SHOWS NO M-CRESOL BINDING SITES ON CYTOCHROME C 129

C. CHLOROBUTANOL EFFECTS ON INTERFERON ALPHA-2A 130

D. ACRYLAMIDE QUENCHING OF INTERFERON ALPHA-2A FLUORESCENCE IN THE PRESENCE OF BENZYL ALCOHOL 132
LIST OF TABLES

Table

2.1 Antimicrobial preservatives used in this study ________________________________ 20

2.2 Parameters obtained by globally fitting the 695nm melting curves shown in Figure 2.7, assuming the same m values for all the curves ______________________ 35

2.3 Comparison of the Tm^Agg of reduced and oxidized forms of Cyt c ___________ 41

4.1 Antimicrobial preservatives used in this study ________________________________ 69

5.1 Parameters obtained from fitting denaturant melt data in Figure 5.5 to a 2-state unfolding model ____________________________________________________ 89

C.1 Parameters obtained from fitting denaturant melt data to a 2-state unfolding model ___________________________________________________________ 130

D.1 Stern-Volmer constants calculated from the slopes in Figure D.2 _____________ 134
LIST OF FIGURES

Figure

1.1 Molecular structure of cytochrome c ......................................................... 12
1.2 Molecular structure of interferon alpha-2a ________________________________ 14
2.1 Molecular structure of cytochrome c .......................................................... 22
2.2 Preservative efficacy test of APs on DH5 alpha E. coli cells ___________________ 26
2.3 Effect of APs on Cyt c aggregation under isothermal conditions ..................... 28
2.4 Variation in the optical density at 695nm as a function of increasing temperature at different concentrations of phenoxyethanol ...................................................... 31
2.5 Decrease in the $T_m^{\text{Agg}}$ as a function of concentration of the five APs ........... 33
2.6 Correlation between the melting temperatures of the Met80 region, $T_m^{695}$, calculated from the initial decrease in the 695nm absorbance in thermal scanning and the aggregation temperatures, $T_m^{\text{Agg}}$ ........................................... 35
2.7 Destabilization of the local protein region around Met80 as monitored by denaturant melts at room temperature ............................................................... 36
2.8 Effect of APs on Cyt c solution structure and dynamics .................................... 38
2.9 Aggregation of reduced cytochrome c in comparison with that of the oxidized form .................................................................................................................. 41
3.1 Post-translational fusion complex and corresponding protein sequence .......... 53
3.2 Schematic of expression and purification scheme optimized to obtain soluble IFNA2 from E. coli ................................................................. 55
3.3 Purification of His6-SUMO-IFNA2 fusion complex ........................................... 56
3.4 Cleavage of His6-SUMO-IFNA2 and purification of IFNA2 ......................... 57
3.5 Mass spectrophotometry analysis of IFNA2 ................................................. 58
3.6 Biophysical characterization of IFNA2 ......................................................... 61
3.7 Comparison of IFNA2 2D-NMR structure with published assignments _________ 62

3.8 Antiviral and anti-proliferative activity of IFNA2 in comparison to known standards __________________________________________________________ 63

4.1 Molecular structure of interferon alpha-2a __________________________________________________________ 68

4.2 Effects of APs on IFNA2 monomer concentration under isothermal conditions _________________________________________________________ 71

4.3 Isothermal kinetics of IFNA2 aggregation in the presence of APs ____________ 72

4.4 Variation in the optical density at 450nm as a function of increasing temperature at different concentrations of phenol __________________________ 73

4.5 Dependence of $T_m^{\text{Agg}}$ on AP concentration __________________________________________________________ 75

5.1 Molecular structure of interferon alpha-2a __________________________________________________________ 82

5.2 Isothermal kinetics of IFNA2 aggregation in the presence of BA __________ 84

5.3 Variation in the optical density at 450nm as a function of increasing temperature and varying concentrations of benzyl alcohol ______________________ 85

5.4 Decrease in $T_m^{\text{Agg}}$ as a function of benzyl alcohol concentration __________ 86

5.5 Destabilization of IFNA2 as a function of BA concentration in the presence of the denaturant GdmHCl at room temperature ___________________________ 88

5.6 Intermediate detection as a function of BA concentration in the presence of the denaturant GdmHCl at room temperature _________________________________ 90

5.7 IFNA2 chemical shift changes as a function of BA concentration in the $^{15}$N dimension __________________________________________________________ 91

5.8 Hydrogen exchange rates in IFNA2 as measured by 2D-NMR plotted as a function of residue position ___________________________________________ 92

5.9 Amino acid contacts within IFNA2 three-dimensional structure _______________ 93

5.10 Aggregation propensity analyses by various software programs ______________ 96

A.1 Increase in Cyt c optical density at 695nm as a function of methanol concentration observed during thermal scanning experiments____________________ 122

A.2 Variation of $T_m^{\text{Agg}}$ as a function of methanol concentration _____________ 122
A.3 Correlation between Tm695, which shows the unfolding of the red loop, and Tm^Agg, the aggregation temperature 123
A.4 Change in thermal scanning trace between oxidized and reduced Cyt c measured at 695 nm 123
A.5 Calculated slopes of denaturation plotted as a function of known solution parameters 124
A.6 Basic structure of a molecular rotor 126
A.7 DCVJ fluorescence measured at the emission wavelength 493 nm as a function of CR concentration in the presence of 300μM Cyt c 126
A.8 Representative traces of FT-IR of dialanine in the absence and presence of 0.8% v/v CR 128
B.1 Isothermal calorimetry of Cyt c and m-Cresol binding 129
C.1 Isothermal incubation of 10μM IFNA2 at 50°C in the presence of 0.3% v/v AP 130
C.2 IFNA2 unfolding in the presence of 1% CB using the denaturant GdmHCl 131
D.1 Locations of tryptophan residues in the three-dimensional structure of IFNA2 132
D.2 Stern-Volmer plot of acrylamide quenching of IFNA2 in the presence of BA 133
CHAPTER I
INTRODUCTION

The use of protein-based pharmaceutics has become increasingly commonplace with the advent of recombinant DNA technology, first introduced as a successful method of obtaining pure protein in the early 1970s [1, 2]. From the initial production of recombinant insulin in 1983 [3], the market for protein drugs has expanded dramatically. In 2011, approximately 350 parenteral drugs existed on the worldwide market [4], with over half of these containing at least one protein as the active ingredient [5]. Protein-based pharmaceutics offer attractive treatment options for numerous diseases for which no therapies are available or small molecule drugs are ineffective. As such, the study of the stability of protein drugs is of utmost importance.

Major concerns in regards to protein stability consistently have been the unfolding of the active protein molecule, lowering its effectiveness, and assembly of protein molecules into higher order structures in solution. In its simplest form, the mechanism of protein unfolding is two-state, meaning that there are only two species extant in the unfolding reaction: the folded species and the unfolded species. Each of these states consists of an ensemble of states, which is accessible to the protein molecule in either the folded or unfolded state. A multitude of variations of this mechanism exist, such as the inclusion of one or more intermediate steps. Protein unfolding is a unimolecular reaction: a single protein molecule can transition from the folded to the unfolded state without the presence of another protein molecule or other external agents.
In contrast, protein aggregation is a multimolecular reaction. In order for a successful aggregation event to occur, a particular number of protein molecules must be present for initiation. If these molecules associate, they form a critical nucleus which can then seed the aggregation of more protein molecules. Consistently, protein aggregation depends on protein concentration. Protein aggregation has been characterized as the formation of soluble and insoluble agglomerates as a result of a particular stress, and these agglomerates may or may not have a functional component. This phenomenon can occur in one of two ways: the first, through oligomer formation, as is seen in many biological processes; the second, via aggregate formation, which has been shown to have various negative affects in the biotechnology field as well as triggers the onset of numerous diseases, such as Alzheimer’s disease and Parkinson’s disease. Understanding the underlying mechanisms of protein aggregation in general has far-reaching effects, not only for the pharmaceutical industry but also for the improved treatment of these diseases.

**Importance and Mechanisms of Biological and Biotechnological Protein Aggregation**

**Aggregation in Natural Systems.**

A number of critical biological processes are characterized by monomeric units assembling into oligomeric structures. Globular subunits of the multifunctional protein actin, G-actin, polymerize to construct F-actin, which is filamentous in nature. Without this ability to transition between two states, actin would be unable to engage in its
abundance of protein-protein interactions including essential functions such as motility and muscle contraction [6, 7]. Microtubules, the cytoskeletal polymer of tubulin, are a collection of agglomeration events, and are responsible for maintaining cell structure as well as assisting in intracellular transport. The two monomeric species of tubulin, α-tubulin and β-tubulin, dimerize to form protofilaments [8]. A number of these fibrous protofilaments combine to produce the cylindrical, hollow microtubule. Many other examples exist, including collagen, the most abundant protein found in humans [8, 9], and hemoglobin, a multi-functional protein known for its oxygen transport capability [10].

Despite these naturally occurring oligomerization events, a number of non-native aggregates have been associated with serious health complications and debilitating diseases. These conditions are commonly referred to as amyloidoses. For example, patients who suffer from renal failure often undergo hemodialysis treatment, which results in platelet and leukocyte microaggregate deposits on dialysis membranes [11]. A number of neurological diseases are characterized by non-native protein aggregation events that lead to decreases in neuronal function. Self-aggregates of amyloid-β, a cleavage product of amyloid precursor protein, has been linked to the onset and progression of Alzheimer’s disease [12]. The mutation and subsequent aggregation of the protein huntingtin has been implicated in the development of Huntington’s disease [13]. While therapies for amyloidoses are available, they only lessen specific symptoms and have varying efficacies from patient to patient. Currently, there are no cures for amyloidosis diseases.
It is interesting to note that amyloidogenic proteins have also been shown to have non-pathological and functional components in a number of biological systems. For example, the yeast prion Sup35p, a protein involved in mRNA transcription regulation, undergoes inactivation via polymerization, which results in a variety of phenotypes [14]. The fungal prion, HET-s, self-assembles to initiate a complex cell death mechanism [15]. In mammals, aggregates of the membrane protein Pmel17 accumulate within melanosomes and provide a surface on which melanin granules develop [16]. Pmel17 aggregates are characterized by their fibrous nature, which closely resembles the fibrils seen in amyloid deposits.

**Aggregation in Biotechnology.**

Protein-based formulations require a shelf-life stability of 18-24 months [17]. Even with proper storage and handling, protein degradation may occur through various pathways, such as chemical modifications of oxidation or deamidation, or through aggregation [18, 19]. Damage to the protein product can decrease the effective dose of delivered drug, raising economic and patient health concerns. Specifically, protein aggregates in therapeutics have been linked to an increase in toxic and immunogenic responses in patients, including the development of neutralizing antibodies that render the administered drug useless [20-26].

An important facet of drug stability is the conditions in which the product is formulated. Every aspect of the formulation process, from excipient addition to container choice, can affect the stability and aggregation of proteins in solution [27].
The solution pH plays an essential role in determining the colloidal and conformational stability of the drug product. If the solution pH is far away from the isoelectric point of the protein, colloidal stability is enhanced through increased charge-charge repulsion between protein molecules. However, the pH of the system affects protein conformational stability by changing the charge density on the surface of the molecule and the native electrostatic interactions, potentially hindering the protein from assuming its correct folded conformation. A formulation pH near the protein isoelectric point may favor aggregation reactions through the lack of electrostatic repulsion. A narrow pH range, therefore, is necessary to facilitate conformational as well as colloidal stability. This phenomenon has been demonstrated with a number of proteins. Granulocyte colony stimulating factor (GSCF) demonstrated significantly different aggregation propensities across a wide pH range following incubation even though it retains its native structure [28, 29]. In buffers with conditions ranging from GCSF’s pI (pH 6.1) to pH 7, with and without sucrose, GCSF monomer loss was observed over the course of 5 days of incubation at 37°C. At pH 3.5, aggregation was not seen except when sodium chloride was added to the solution [29]. In the case of the novel therapeutic fusion protein abatacept, rapid oligomer formation was observed at pH 6 but aggregation was significantly delayed at pH 7.5 following incubation at 40°C. Secondary structural characteristics were identical across this pH range, with slight perturbations of tertiary structure [30].

The use of additives and excipients in liquid formulation is critical in order to maintain product stability as well as to ensure maximum drug efficacy. These compounds range from salts added for ionic strength to preferentially excluded molecules like
sucrose for their protective effects. The choice and concentration of these compounds is critical. For example, salt ions preferentially interact with protein molecules in a concentration-dependent manner. Low salt concentrations can stabilize proteins through electrostatic charge-shielding [31], leading to changes in colloidal and conformational stability. High concentrations, however, can cause either stabilization or denaturation, depending on the protein and the salt [32]. The same salt can affect proteins in different ways as well. For example, the inclusion of NaCl caused an increase in GCSF aggregation rate at a pH where aggregation was not favored [29]. The opposite phenomenon was seen for factor VII SQ, which had maximum stability at high and low pH in the presence of 85mM NaCl as seen by differential scanning calorimetry [33]. UV-Visible spectrophotometric analysis of keratinocyte growth factor (KGF) stored at 45°C showed diminished aggregation with an increase in NaCl concentration [34]. In contrast, the presence of varying concentrations of ammonium sulfate increased the amount of KGF aggregates following storage. Due to its unique ability affect protein solubility, ammonium sulfate gradients have been used on protein-rich solutions to specifically precipitate a protein of interest. The critical salt concentration required depends upon the nature and concentration of the protein and various solution conditions such as pH and temperature [35]. Even the addition of sucrose at certain concentrations loses its protective nature. Sucrose imparts protection to protein molecules against aggregation via preferential exclusion, in which it is isolated from the protein surface and favors a compact, folded protein [36, 37]. In the case of the model system trypsinogen, sucrose-protein contacts were shown to improve stability in the solid state; however, during spray-drying in the presence of high sucrose concentration, sucrose was excluded from
the protein surface in such a way that inhibited sucrose-protein contacts [38]. Additionally, surfactants are often included in protein drug formulations in order to inhibit protein interactions at interfaces, such as the air-water interface or the surface of the storage container [39]. However, non-ionic surfactants such as Tween 20 have been shown to increase aggregation in acidic fibroblast growth factor (AFGF) at pH 7.4 in phosphate buffer [40]. Pegylated GCSF (PEG-GCSF) aggregation was mitigated in the presence of Tween 20 during agitation; but without agitation increasing concentrations of Tween 20 induced PEG-GSCF aggregation when stored at 29°C over the course of five months [41].

Protein-based therapeutics are available in a variety of dosage forms. A popular choice in terms of economics and patient compliance are multi-dose products. Economically, production of multi-dose formulations can decrease manufacturing, medical waste disposal, and storage costs. From a patient perspective, the ability to administer therapeutics at home rather than at a physician’s office is an attractive option. These formulations include at least one antimicrobial preservative (AP) to combat the growth of bacteria and other microbes during repeated contact between the drug solution and a syringe needle. However, these necessary compounds have been linked to the unfolding and aggregation of proteins in the liquid state.

Preservative Effects on Proteins

In one of the first reports, the presence of various phenolic compounds (many of which are commonly used preservatives) greatly affected the aggregation of recombinant
human growth hormone (rhGH). For example, upon the addition of pharmaceutically relevant concentrations of m-cresol (CR), soluble rhGH monomer was depleted within the first day of incubation at 50°C. Additionally, low temperature incubation experiments were conducted which demonstrated a causal link between phenolic compound presence and rhGH aggregation [42]. These results are especially interesting considering that protein-based drug systems are stored between 2-8°C in the liquid state and have lyophilized analogs.

Other recent reports further demonstrate that the presence of APs influence the stability and aggregation of various protein systems. A number of these studies focused on the influence of benzyl alcohol (BA) on protein systems, since BA is the most prevalently used preservative in protein therapeutics [43]. BA caused complete loss of tertiary structure in interferon-γ, with no perturbation observed in secondary structure [44]. Tertiary structure was lost more rapidly with an increase in protein concentration. The rate of tertiary structure loss was slowed in acetate buffer compared to succinate buffer in the presence of 0.9% BA, demonstrating the sensitivity of interferon-γ to the ionic strength of solution. Interleukin-1 receptor antagonist (IL-1RA) aggregation occurred during reconstitution with BA even in the presence of formulation stabilizers, such as sucrose [45]. Isothermal incubation of IL-1RA at room temperature in the presence of 0.9% BA did not result in an increase of aggregation compared to controls over eight days [45], whereas incubation at 37°C caused complete loss of soluble monomer within five days [46]. While BA alone did not perturb the secondary structure or cause aggregation of GCSF at low pH, the presence of BA induced anti-parallel β-sheet formation upon aggregation of GCSF at pH 7 [28]. BA perturbed the tertiary
structure of GCSF at acidic and neutral pH, but was seen more significantly at 37°C than room temperature. Additionally, hydrogen-deuterium exchange of GCSF increased in the presence of BA, indicating that BA induced a structural change in GCSF that affected the compact, protected nature of the native state.

Similarly, interleukin-1 receptor (IL-1R) has been shown to be sensitive to BA as well as CR and phenol (PH) through the loss of monomer content following seven and 60 days of storage at 37°C. IL-1R melting temperature (Tm) also decreased in the presence of all three preservatives. BA caused the greatest amount of monomer loss as well as decrease in Tm, followed by CR, then PH; however, the concentration of BA used was nine times greater than CR and nearly 14 times greater than PH, making direct comparisons difficult. The change in Tm with each AP was evaluated with respect to control values, which led to the determination of the normalized preservative force imparted by each AP on IL-1R. Compared in this way, CR influenced IL-1R the most, followed by PH and finally by BA [47].

A monoclonal antibody incubated in the presence of CR lost all soluble monomer over the course of two days at 50°C, where PH induced the formation of both soluble and insoluble aggregates. BA caused significant monomer loss at concentrations above 2%, but only 7% monomer loss was observed at 0.5% BA. Chlorobutanol (CB) was shown to affect antibody monomer content dramatically; at 0.5% CB, 97% of monomer was lost. The presence of BA and CB also reduced the Tm measured by differential scanning calorimetry to the same extent [48].

It has been hypothesized that APs, being hydrophobic in nature, impart their influence on proteins via hydrophobic interactions that mimic weak binding events. This
interpretation states that APs preferentially bind to hydrophobic patches on the protein surface [46], driving unfolding and populating aggregation-prone species. Recently, however, it was suggested that BA interacts with hydrophilic residues through a cation-π mechanism, as was seen with IL-1RA [49]. Additionally, it has been suggested that AP-induced aggregation results from the local unfolding of a small region of a protein rather than a global unfolding event [44, 46, 50].

Considering the available data regarding preservative effects on protein stability and aggregation, there are a few key points to note. First, APs decrease the melting temperature of proteins. Secondly, APs consistently cause soluble protein monomer loss in accelerated stability studies. Thirdly, BA affects protein structure on a tertiary, but not secondary, level. This phenomenon is in support of the hypothesis that APs cause partial unfolding prior to aggregation.

It remains unclear how exactly these preservatives mechanistically affect the stability and aggregation of a protein. Are there particular regions within a protein’s tertiary structure that determine the extent of AP-induced perturbation? If so, do these regions act as aggregation “hot-spots”; i.e., does the unfolding of a specific local region correlate to an aggregation event? Do all proteins contain “hot-spots” that allow for APs to cause partial unfolding and aggregation? And finally, do individual APs cause protein unfolding and aggregation in different ways? The few comparative studies that exist demonstrate that CR generally has the most deleterious effect on protein aggregation. How do the other commonly used APs compare to each other in terms of inducing aggregation?
In an attempt to elucidate answers to these questions, we have used a model protein, Cyt c, and a pharmaceutical protein, interferon α-2a, and we assessed the effect of APs on the stability and aggregation of these systems. The results of these investigations will aid in the understanding of AP-induced aggregation as well as in the development of strategies to prevent aggregation. These findings will not only impact the pharmaceutical industry, but may also have applications in mitigating aggregation seen in disease.

**Cytochrome c as a Model Protein.**

In general, pharmaceutical proteins are not well characterized in terms of their solution behavior or biophysics. Therefore, the use of a model protein to monitor AP effects is advantageous. Cyt c (Figure 1.1) offers a number of benefits as a model system. Cyt c contains three α-helices and three Ω-loops, and plays a critical role in the electron transport chain during cellular respiration [51] as well as apoptosis [52]. This protein has been studied extensively and much is understood in regards to its stability and dynamics. Amino acid-resolved folding and unfolding pathways have been determined using hydrogen exchange nuclear magnetic resonance (HX NMR); hence, the contributions of individual residues to the overall stability of the protein are known. The folding pathway of Cyt c consists of six localized regions, called foldons, which assemble cooperatively into the native state [53, 54]. This knowledge allows for the use of robust methods, such as NMR, to monitor small changes in dynamics that are influenced by external (pH, excipients, etc.) or internal (chemical modifications, mutations) factors. Previous experiments conducted by myself and others in our laboratory have indicated that CR and
BA cause the unfolding a specific region of model protein Cytochrome c (Cyt c), populating a partially unfolded species that is highly prone to aggregate [55, 56], which formed a foundation for this thesis work. This local region, identified as the 70s-80s loop, is known to be the last region of Cyt c to fold, as well as the first to unfold. A section of amino acids within the 70s-80s loop has also been identified to be aggregation-prone, due to six hydrophobic residues adjacent to one another in this stretch.

FIGURE 1.1 Molecular structure of cytochrome c (Cyt c; 1HRC.pdb).

Eight APs are commonly used in protein-based drugs, and five of these are included in liquid formulations. The effects of these five APs (CR, BA, PH, phenoxyethanol (PE) and chlorobutanol (CB)) on Cyt c were evaluated and compared in Chapter II. APs lead to Cyt c aggregation in a concentration-dependent manner, and the extent of these effects varied by the AP. CR caused Cyt c to aggregate the most, followed
by PH, BA, PE, and CB. Additionally, the same specific region of Cyt c, the 70s-80s loop, was shown to unfold rapidly in the presence of CR and PH as compared to control. BA and PE also influenced this loop but to a lesser extent than CR and PH. This phenomenon indicated that 1) the unfolding of the \( \Omega \)-loop initiates aggregation and 2) the order in which these APs induce the unfolding of this loop is CR > PH > BA > PE > CB. This pattern of AP-induced unfolding was consistent with the order in which APs caused Cyt c aggregation.

**Interferon \( \alpha \)-2a as a Pharmaceutical Protein.**

The interferon family of proteins belongs to an important aspect of the innate immune system responsible for antiviral and anti-proliferative functions during viral assault [57, 58]. Interferon transcription is initiated following the recognition of viral factors on cell surface receptors, and results in self-propagating protein production [59-61]. As such, the interferon family has been utilized to treat various viral disorders such as hepatitis C and hairy cell leukemia [62-64].

Interferon-\( \alpha \) has at least 13 subtypes found in humans. All interferon-\( \alpha \) subtypes contain 165-166 amino acids and predominantly \( \alpha \)-helical structure [65]. Two synthetic (interferon \( \alpha \)-2b, Intron®, Merck; interferon alfacon-1, Infergen®, Amgen) and two naturally occurring forms of interferon-\( \alpha \) (interferon \( \alpha \)-2a, Roferon®, Roche; and interferon \( \alpha \)-2c) have been developed for therapeutic use, with pegylated versions (e.g., PEG-Interferon \( \alpha \)-2a, Pegasys®, Genentech) also on the market. Interferon \( \alpha \)-2a (IFNA2) (Figure 1.2) is available as a multi-dose therapeutic and contains BA as a preservative.
FIGURE 1.2 Molecular structure of interferon α-2a (IFNA2; 1ITF.pdb)

Early experiments demonstrated that IFNA2 aggregates in the solution state, and these aggregates have been linked to an increase in neutralizing antibodies in clinical trials [66-68]. Characterization of the aggregates showed the presence of IFNA2-IFNA2 species. Human serum albumin (HSA) was included in the formulation as an inert stabilizer, but did not suppress IFNA2 aggregation; (HSA)-IFNA2 aggregates were observed as well. Following this study, HSA was removed in order to inhibit heterogeneous aggregation [69].

In order to examine the effects of APs on IFNA2 stability and aggregation, one must first obtain pure protein for experimentation. Historically, IFNA2 has been obtained directly from human leukocytes [70] or by using recombinant expression in bacterial
inclusion bodies [71]. In order to obtain active and pure protein using these methods, extensive purification [72, 73] and subsequent refolding was required [74]. **Chapter III** describes an in-house expression and purification scheme we developed for the high yield production of soluble IFNA2 in *Escherichia coli*. This method utilizes a fusion complex, SUMO, which increases the soluble expression of recombinant proteins in bacteria [75-77]. Two-dimensional NMR and circular dichroism verified protein identity as well as correct conformation [78]. Activity analysis using both antiviral and anti-proliferative assays demonstrated that our purified IFNA2 retained satisfactory functionality in comparison to known EC50 values.

The role of APs in IFNA2 aggregation was investigated in **Chapter IV**. APs caused the aggregation of IFNA2 through the loss of soluble monomer. The extent of aggregation was dependent on the concentration and type of the AP used. The relationship between AP concentration and aggregation followed the pattern CR > PH > BA > PE. These data matched exactly with previous results using Cyt c and rhGH. Additionally, these data validated the use of Cyt c as a model protein system to predict the effects of APs on pharmaceutically relevant proteins.

In **Chapter V**, the specific biophysical and structural interactions between BA and IFNA2 were examined. BA caused a concentration-dependent increase in IFNA2 aggregation. The free energy of unfolding, $\Delta G$, decreased in the presence of BA, and a decrease in unfolding $m$-value as a function of BA concentration indicated the possible presence of an intermediate. The $m$-value is a representation of the accessible surface area that is exposed as a protein moves from the folded to unfolded states, and a changing $m$-value suggests a deviation from two-state behavior. Upon further investigation, BA
appeared to preferentially populate a partially unfolded intermediate as observed using ANS fluorescence. Interestingly, the presence of BA influenced local regions in IFNA2. HX and 2D NMR experiments suggest that BA interrupts the interhelical contacts between residues located in Helix A and Helix C, populating an aggregation-prone species by either unfolding these helices or inducing domain swapping.
CHAPTER II

ANTIMICROBIAL PRESERVATIVES CAUSE PARTIAL PROTEIN UNFOLDING AND AGGREGATION OF CYTOCHROME C¹

Introduction

Protein-based pharmaceuticals comprise a significant portion of drug formulations. More than 350 parenteral formulations are available worldwide, with almost 150 protein-based pharmaceutical drugs commercially available in the United States [4, 19]. One-third of these parenteral products are multi-dose [5], which are advantageous in terms of patient compliance as well as economics. Protein formulations require a shelf-life stability of 18-24 months [17]. To maintain product viability, multi-dose formulations require the presence of antimicrobial preservatives (APs) [5, 43] to inhibit the growth of microbes and bacteria during repeated contact between the solution and a syringe needle [5]. APs are also found in topical ointments [79], and in multi-dose delivery systems such as mini-pumps for continuous infusion [80] and injection pens [81].

The use of APs in protein formulations is of recent concern because of the discovery that these small molecules cause protein aggregation. In one of the first reports, Maa et al. [42] demonstrated that the addition of aromatic and aliphatic alcohols results in the aggregation of human growth hormone. Similar results were found with APs, most of which are alcohols, and other proteins. For example, benzyl alcohol, the most widely

used AP, induces the aggregation of interferon-γ [82], interleukin-1 receptor antagonist [50], and human granulocyte colony stimulating factor [28]; m-cresol induces the aggregation of human growth hormone [42, 83] and interleukin-1 receptor [47]; and phenol induces the aggregation of an antibody [48]. Such protein aggregates in formulations can decrease the efficacy of the delivered drug as well as stimulate undesirable toxic and immunologic responses in patients [20, 21, 23-26, 84-89]. Therefore, an understanding of the molecular mechanisms underlying AP-induced protein aggregation and the development of strategies to minimize such aggregation are of paramount importance in developing stable multi-dose drugs.

Out of the eight commonly used APs [5, 43], five (listed in Table 2.1) are specifically used in liquid protein formulations. However, it has not been clear whether all five APs result in protein aggregation, and to what extent. Preliminary evidence from Maa et al. [42] indicates that various alcohols induce the aggregation of human growth hormone to different extents, implying that protein aggregation may depend on the nature of the alcohol. Such comparative information on APs is not available, which will be helpful in choosing the right AP when designing a stable therapeutic formulation so that the formation of protein aggregates and the resultant immunogenic and toxic effects can be minimized.

The ideal method of understanding aggregation mechanisms is to study how various APs induce the aggregation of a pharmaceutical protein of interest using biophysical methods. However, most pharmaceutical proteins are relatively less understood in terms of their biophysics and solution behavior. Therefore, we used a model protein, cytochrome c (Cyt c; Figure 2.1), which has been well-characterized in the
literature. This protein offers many spectroscopic probes due to its covalently linked heme chromophore, in addition to the traditional probes such as circular dichroism and fluorescence of aromatic residues used to probe protein structure. A particularly useful tool to probe the partial unfolding of Cyt c is an absorption band at 695 nm that reports on the unfolding of one of the least stable regions in the protein (Red-colored $\Omega$-loop in Fig. 1). In Cyt c, the methionine residue at position 80 (Met80) that is part of this Red loop is covalently linked to the ferric iron of the heme group. The 695 nm band originates due to charge transfer from the sidechain sulfur of Met80 to the ferric iron. When the local region around Met80 unfolds, this absorption band is absent because of the breakage of the Met80 to heme bond. The identification of the $\Omega$-loop containing Met80 as being a least stable region, or weakest link in the protein, came from earlier residue-resolved hydrogen exchange (HX) experiments measured using 2D-NMR [53, 90-92], protein unfolding monitored in response to various stresses [93-95], and from studies on ultrafast protein dynamics [96]. Our earlier computational analysis predicts that the six residues around Met80 acts as an aggregation ‘hot-spot’ whose unfolding may lead to Cyt c aggregation [56], and experimental results confirm this hypothesis [56]. Subsequent determination of the molecular structures of Cyt c oligomers using X-ray methods and their characterization using biophysical techniques indicate that the aggregates have the same native structure as that of monomeric Cyt c with no change in the secondary structure (identical circular dichroism spectrum), except that the C-terminal helices are domain-swapped between the monomers and the Met80 region is unfolded [97]. These earlier results on Cyt c which determined the nature of partial protein unfolding that leads to aggregation provides a unique opportunity to test the effect and relative strength of
different APs on protein unfolding and aggregation. We have also tested whether stabilizing the local protein region that acts as the aggregation ‘hot-spot’ will reduce the AP-induced protein aggregation.

**TABLE 2.1** Antimicrobial preservatives used in this study.

<table>
<thead>
<tr>
<th>AP</th>
<th>Molecular Structure</th>
<th>Molecular Weight (Da)</th>
<th>Source</th>
<th>Purity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benzyl alcohol (BA)</td>
<td><img src="image1" alt="structure" /></td>
<td>108.1</td>
<td>Merck</td>
<td>97</td>
</tr>
<tr>
<td>4-Chloro-1-butanol (CB)</td>
<td><img src="image2" alt="structure" /></td>
<td>108.6</td>
<td>Sigma</td>
<td>85</td>
</tr>
<tr>
<td>m-Cresol (CR)</td>
<td><img src="image3" alt="structure" /></td>
<td>108.1</td>
<td>Sigma</td>
<td>99</td>
</tr>
<tr>
<td>Phenol (PH)</td>
<td><img src="image4" alt="structure" /></td>
<td>94.1</td>
<td>Sigma</td>
<td>99.5</td>
</tr>
<tr>
<td>2-Phenoxyethanol (PE)</td>
<td><img src="image5" alt="structure" /></td>
<td>138.2</td>
<td>Fluka</td>
<td>99.5</td>
</tr>
</tbody>
</table>

**Materials and Methods**

**Materials.**

Equine Cyt c (Type VI) was obtained from the Sigma Chemical Company (St. Louis, MO, USA). Prior to experiments, the protein was oxidized using potassium ferricyanide (Fisher Chemicals, Fair Lawn, NJ, USA) to remove any traces of the reduced form, dialyzed extensively against the desired buffer, and filtered through a 0.22 μm
filter. The purity of the protein was determined from the ratio of the absorbance values at 409 nm (heme Soret band) and 280 nm (aromatic absorption band), which was 4.6 as expected for a pure oxidized protein [98]. In addition, the absence of the reduced form in the sample was confirmed from the absorbance values at 339, 526.5, 541.74, and 550 nm using a previously described method [98, 99]. For experiments with the reduced form, sodium dithionite (Sigma Chemical Company) was used to reduce Cyt c [100]. The five APs used in this study are listed in Table I.

**Preservative Efficacy Test.**

To confirm the antimicrobial activity of APs, a simplified preservative efficacy test was performed [101]. A primary culture of *Escherichia coli* DH5α cells was incubated overnight at 37°C in a shaker. Aliquots of 0.5 ml were transferred into six 50 ml culture flasks containing either no preservative (control), or one of the five APs. Cultures were incubated at 37°C with shaking for six hours, and optical density at 600 nm was used to measure the cell count.

**Size Exclusion Chromatography (SEC).**

To monitor the effect of benzyl alcohol (BA) on protein aggregation, Cyt c (2 mM in 0.1 M sodium phosphate, 0.15 M sodium chloride, pH 7) was incubated at 37°C on a rotator (Thermo Scientific Labquake Shaker Rotisserie) and samples were taken at desired intervals. Concentration of the monomer was estimated by injecting 5 μL onto a TSKgel 5 μm G3000SWxl column on an Agilent 1100 HPLC. The mobile phase used for this column was 0.1 M sodium phosphate, 0.1 M sodium sulfate, pH 6.7 at a flow rate of
0.7 mL min\(^{-1}\). Absorbance at 280 nm was used to estimate the protein concentration. The average of 0% v/v BA triplicates on day 0 was used to normalize the peak area of subsequent sample sets.

**FIGURE 2.1** Molecular structure of cytochrome c (Cyt c; IHRC.pdb). The protein consists of three \(\alpha\)-helices and three \(\Omega\)-loops. Individual cooperatively unfolding regions detected in Cyt c are colored in terms of their increasing stability: Infrared (Nested-yellow), Red, Yellow, Green, and Blue [53]. The unfolding of Cyt c proceeds either through the Red loop or through the bottom Infrared loop [53]. The figure also shows the covalently linked heme, and its two axial ligands His18 and Met80. The structure was generated using the MOLSCRIPT program.

**Isothermal Incubation Experiments.**

Cyt c solution was incubated at the desired temperature, and the changes in optical density at 695 nm and 800 nm were measured as a function of the incubation
time. For these experiments, the cuvette of buffer solution was initially equilibrated at the desired temperature, and Cyt c (300 μM final concentration) was added to the cuvette. The aggregation kinetics was monitored until the signal reached a plateau. At longer incubation times, the aggregates started to settle down to the bottom of the cuvette, resulting in decreased optical density. At that point, the experiment was stopped.

**Thermal Scanning Method.**

The aggregation temperature (Tm\textsuperscript{Agg}) of the protein was measured using a thermal scanning method on a Chirascan Plus spectrometer (Applied Photophysics, Surrey, UK). The temperature was increased at a rate of 1°C/step followed by 2 min equilibration, and changes in the optical density at 800 nm were recorded. Tm\textsuperscript{Agg} was determined using Global Analysis T-Ramp software (Pro-Data Global3 v1.1.0), provided by Applied Photophysics using a single transition and double baseline correction. For these experiments, 300 μM Cyt c in 0.1 mM sodium phosphate, 0.15 mM sodium chloride, pH 7 was used. The reduced Cyt c samples contained 10X sodium dithionite to maintain the protein in reduced form during thermal scanning.

Temperature unfolding of the Met80 region was performed by monitoring changes in the optical density at 695 nm using the above described thermal scanning method. The optical density initially decreased followed by an increase. From the initial decrease, the melting temperature (Tm\textsuperscript{695}) was determined as the temperature at which the absorbance is half of the difference between the absorbance values of the native and unfolded states. Absorbance of the unfolded state at high temperature was independently measured by denaturing 300 μM Cyt c with 6 M Urea.
Denaturant Melts.

Guanidinium Chloride (GdmCl) was used as the denaturant. Protein solutions at varying GdmCl concentrations were prepared and equilibrated overnight before measuring changes in optical signals as a function of the denaturant concentration. Concentration of the denaturant was determined using refractive index measurements [102]. For measuring changes in the 695 nm absorbance, 300 µM Cyt c in 0.1 M sodium phosphate, 0.15 M sodium chloride, pH 7 was used. The ΔG values were determined by fitting the changes in optical signals at different denaturant concentrations to a 2-state unfolding model [103, 104].

Nuclear Magnetic Resonance (NMR).

2D gradient COSY experiments were run to monitor changes in Cyt c amide crosspeaks as a function of the AP using a Varian Inova 600 MHz NMR instrument equipped with a cryoprobe. For these experiments, 3 mM Cyt c (0.1 M sodium phosphate, 0.15 M sodium chloride, pH 7) and deuterated APs (0.8% v/v) were used. Deuterated chlorobutanol was not available commercially, and hence the NMR experiments were performed with the other four deuterated APs (Isotec Inc). The NMR spectra were collected with 8000 Hz spectral width and 512 points in each direction. The spectra were processed in magnitude mode using the nmrPipe software (Delaglio) with zero filling to twice their real points, apodization with nonshifted sine multiplication, exponential broadening, and Gaussian transformation. The spectra were baseline corrected in both dimensions. Changes in the peak positions (chemical shifts) and peak volumes were calculated using the nmrDraw package. Cyt c residue assignments
available in the literature [105] were used for this purpose. 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) was used as a standard to reference the chemical shifts as well as to normalize the measured changes in crosspeak volumes with the addition of APs.

**Hydrogen Exchange (HX).**

Cyt c (3 mM) whose amides were protons was lyophilized in deionized water. It was dissolved in the appropriate buffer in D$_2$O just before running the HX experiment and filtered through 0.22 μm filter. A series of 2D gradient COSY was recorded back to back using a Varian Inova 600 MHz NMR instrument equipped with a cryoprobe to monitor the HX of various amide hydrogens. NMR spectra were processed using nmrPipe software and the changes in peak volumes due to exchange were measured using nmrDraw software. The changes in peak volumes due to exchange were normalized with the intensities of five non-exchangeable crosspeaks (C$_6$H-C$_7$H & C$_4$H-C$_5$H of Trp59, C$_2$H-C$_3$H of Tyr74, C$_7$H - C$_8$H$_3$ of Leu64, and C$_7$H - C$_8$H$_3$ of Leu98).

**Results**

**Antimicrobial Activity of APs.**

To demonstrate the antimicrobial efficacy of the five APs, we tested their effect on the growth of *Escherichia coli* bacteria [101]. For this purpose, we used DH5α cells in LB media and monitored the cell count by measuring changes in the optical density at 600 nm as a function of the growth time. Without APs, the growth curve showed an exponential increase (Figure 2.2). However, no increase in the optical density was
observed with the addition of any of the five APs, indicating that these molecules inhibited bacterial growth.

**FIGURE 2.2** Preservative efficacy test of APs on DH5α E. coli cells. The filled circles represent the data untreated with APs, whereas the other symbols represent the data treated with various APs: 1% (v/v) benzyl alcohol (BA), 0.3% (v/v) m-cresol (CR), 0.5% (v/v) phenol (PH), 0.5% (v/v) phenoxyethanol (PE), and 0.5% (v/v) chlorobutanol (CB). These concentrations correspond to those used in protein formulations. The cell count was monitored by measuring the relative optical density at 600 nm as a function of the growth time. Addition of APs inhibited bacterial growth.

**APs Induce Cyt c Aggregation.**

Earlier studies have indicated that APs cause aggregation of pharmaceutical proteins over a course of several days and months when stored close to room temperature [28, 42, 47, 48, 50, 82, 83]. Monitoring the effect of different APs on such aggregation and at different solution conditions is not feasible due to long incubation times. Therefore, to probe the aggregation mechanisms on a convenient laboratory scale, we
accelerated the aggregation kinetics by performing isothermal incubation studies at higher temperatures. We first showed that APs induce Cyt c aggregation at physiological temperatures by following the concentration of the Cyt c monomer in solution as a function of the incubation time, using benzyl alcohol (BA) as the AP (Figure 2.3A). Over a time course of four days, 3% v/v BA aggregated Cyt c by 50%, whereas no such aggregation was observed in the absence of AP. We further examined the effect of temperature on the kinetics of AP-induced protein aggregation (Figure 2.3B). With the increase in solution temperature from 65°C to 85°C, the rate of Cyt c aggregation induced by 0.8% v/v BA increased. Based on this data, we chose 75°C (Figure 2.3C) and 80°C (Figure 2.3D) so that the effect of APs on Cyt c aggregation can be conveniently monitored on the timescale of hours. The rate and extent of aggregation depends on the nature of AP. The five APs followed the order m-cresol (CR) > phenol (PH) > phenoxyethanol (PE) > BA > chlorobutanol (CB). Under identical conditions, no aggregation was observed for Cyt c in the absence of APs. For these comparison experiments, a concentration of 0.8% v/v was chosen for all the APs, because some of the APs, in particular CR, were not soluble at higher AP concentrations.

**APs Decrease the Temperature at which Cyt c Aggregates.**

An alternative method used in the literature to probe the effects of various solution conditions on protein aggregation is to measure their influence on the temperature at which the protein aggregates during thermal scanning [28, 46, 50, 55, 56]. We have previously shown that both methods yield similar information, i.e., the solution conditions that resulted in faster aggregation of Cyt c also exhibited a lower aggregation
FIGURE 2.3 Effect of APs on Cyt c aggregation under isothermal conditions.
FIGURE 2.3 Effect of APs on Cyt c aggregation under isothermal conditions. (a) Fraction of soluble monomer remaining in solution after incubation at 37°C with shaking, as determined by size-exclusion chromatography gel filtration. The black and grey bars represent the value without and with 3% (v/v) benzyl alcohol (BA), respectively. (b) Aggregation kinetics at 0.8% (v/v) BA as a function of the solution temperature. (c and d) Aggregation kinetics at 75°C and 80°C, respectively, at 0.8% (v/v) concentration of the five APs, m-cresol (CR), phenol (PH), phenoxyethanol (PE), benzyl alcohol (BA), and chlorobutanol (CB).
temperature [55, 56]. We used a temperature scanning method to monitor the effects of five APs on the aggregation temperature of Cyt c. For this purpose, we measured the optical density at 695 nm to simultaneously monitor the melting of the least stable structure in Cyt c (Red loop containing Met80 region in Figure 2.1) in addition to protein aggregation. In buffer (0% trace in Figure 2.4), the optical density initially decreased, indicating unfolding of the local region surrounding Met80. Subsequently, the signal increased until a plateau was reached, which might be because of protein aggregation. We confirmed this by measuring the optical density at 800 nm using the same thermal scanning experiment (Figure 2.4 Inset). At this wavelength, the protein and the buffer do not absorb, and hence the observed changes in the optical density can be attributed solely to protein aggregation. The increase in optical density at 800 nm as a function of solution temperature exactly matched the latter part of the 695 nm curve (Figure 2.4 Inset), indicating that the increase in 695 nm signal at higher temperatures is due to aggregation, and not due to the increased absorbance of the Met80 to heme charge transfer band. At much higher temperatures beyond the plateau region, optical density decreased as preformed protein aggregates start settling down to the bottom of the cuvette. We performed this thermal scanning experiment at varying concentrations of the AP PE (Figure 2.4) to measure its effect on the aggregation temperature. With the addition of PE, the midpoint temperature of aggregation ($T_{m^{Agg}}$) decreased, as determined from the 695 nm as well as 800 nm optical density curves. In the absence of PE, the $T_{m^{Agg}}$ was 83.6°C, whereas the addition of 2% PE decreased the $T_{m^{Agg}}$ to 69.5°C, indicating that the presence of AP enhanced Cyt c aggregation.
FIGURE 2.4 Variation in the optical density at 695 nm as a function of increasing temperature at different concentrations of phenoxyethanol (PE). Inset shows the optical density curve at 695 nm in comparison with that of 800 nm for 0% PE concentration, indicating that the increase in optical density at higher temperatures is due to protein aggregation.

Aggregation Temperature Decreases Linearly with Increasing AP Concentration.

Similar to PE, the other four APs decreased the $T_m^{\text{Agg}}$ of Cyt c (Figure 2.5). A linear correlation was observed between the concentration of each AP and the $T_m^{\text{Agg}}$. This correlation is significant in that it is a qualitative measure of how effective each preservative is in inducing the aggregation of Cyt c. By comparing the relative slopes, the five APs induced the aggregation in the order CR > PH > BA > PE > CB. The AP that is most effective in causing protein aggregation is CR. The corresponding slope indicates that for every percent concentration change of CR, the $T_m^{\text{Agg}}$ of Cyt c decreases by 17.5°C. In contrast, CB is the least aggregation-causing AP with a slope of 3.3°C/%v/v.
This trend is very similar to that observed from isothermal incubation studies (Figure 2.3), except the reversal of the effects of BA and PE.

**Unfolding of the Met80 Region Correlates with the Aggregation Temperature.**

To determine the unifying mechanism underlying AP-induced aggregation, in particular whether the same local protein region acts as an aggregation ‘hot-spot’ for all the APs, we examined the correlation between the unfolding of the local protein region around the residue Met80, which was predicted to be the aggregation ‘hot-spot’ [56], and Cyt c aggregation. As discussed above, the unfolding temperature of the Met80 region (T\textsubscript{m}\textsuperscript{695}) was determined from the initial decrease in the absorbance at 695 nm with increase in temperature as the value at which the absorbance is half of the absorbance values of the native and unfolded states (measured independently). A decrease in the absorbance at 695 nm signifies the disruption of the Met80 to heme bond, indicating destabilization of the local structure around Met80 (Figure 2.1). This decrease occurred prior to protein aggregation as can be observed by comparing the absorbance traces at 695 nm and 800 nm (Figure 2.4 Inset). With the increase in the concentration of the AP, the temperature at which the Met80 region melted decreased (Figure 2.4), indicating that the presence of AP destabilized the local structure around Met80. The decrease in T\textsubscript{m}\textsuperscript{695} is very similar to that of T\textsubscript{m}\textsuperscript{Agg}.

We determined the T\textsubscript{m}\textsuperscript{695} of all the 16 samples (control with no AP, and three concentrations each of the five APs) from the initial decrease in the optical density at 695 nm (for example, Figure 2.4 for PE) and plotted them against T\textsubscript{m}\textsuperscript{Agg} (Figure 2.6). The relationship between T\textsubscript{m}\textsuperscript{695} and T\textsubscript{m}\textsuperscript{Agg} was linear (r\textsuperscript{2} = 0.79), indicating that the destabiliz-
FIGURE 2.5 Decrease in the $T_m^{\text{Agg}}$ as a function of concentration of the five APs. Individual panels show the slopes of such variation.

Formation of the Met80 region by the presence of APs may be the critical event that triggers Cyt c aggregation. However, the $T_{m}^{\text{Agg}}$ is higher than $T_{m}^{695}$ by approximately 10-20°C, possibly because of the concentration dependence of protein aggregation. In general, protein unfolding is a unimolecular reaction whereas protein aggregation is a multimolecular reaction, and hence aggregation is expected to depend on the protein concentration. Consistently, we did not observe Cyt c aggregation when the aggregation
experiments were performed at 10 μM protein concentration [55]. With the increase in protein concentration, Cyt c starts aggregating, and its T_m^Agg shifts to low temperatures, confirming the strong dependence of Cyt c aggregation on its concentration. All the experiments presented in this paper were done at a protein concentration of 300 μM.

**APs Destabilize the Local Protein Region Around Met80.**

To further confirm that the presence of APs destabilized the local protein region around Met80, we measured the changes in its stability by monitoring the changes in the 695 nm absorbance with the addition of the denaturant GdmCl (Figure 2.7). Based on the data in Figure 2.5, we chose a concentration of 0.8% v/v APs for these studies which is common to all the APs. With the addition of any of the five APs, the decrease in 695 nm shifted to lower denaturant concentration when compared with that in the absence of AP (Figure 2.7). These denaturant melts qualitatively indicate that the APs which aggregate Cyt c most show a decreased C_m value and a shallower m-value, indicating a correlation between partial protein unfolding and aggregation. The decrease in m-values may indicate the presence of an intermediate, however, the presence of slopy baselines and the absence of a second, cooperative transition did not result in unique fits to a 3-state unfolding model. The sloped baselines in the denaturant melts also resulted in large errors in obtained ∆G values when the individual curves were fit to a 2-state model. Therefore, to analyze the effect of APs on protein stability, we assumed that the m-value of the partial protein unfolding measured by the decrease in 695 nm absorbance is the same for all the five APs, and all the six melts were globally fit to a 2-state model (Figure 2.7). The obtained fit parameters are listed in Table 2.2. The stability decreased in the order
CR > PH > BA > PE > CB, which matches the order of the effectiveness of APs in inducing Cyt c aggregation (Figure 2.5).

**FIGURE 2.6** Correlation between the melting temperatures of the Met80 region, $T_m^{695}$, calculated from the initial decrease in the 695 nm absorbance in thermal scanning (Figure 2.4) and the aggregation temperatures, $T_m^{Agg}$. The plot includes data from 16 sample sets (Cyt c in buffer with no AP, and with each of five APs at three concentrations each).

**TABLE 2.2** Parameters obtained by globally fitting the 695 nm melting curves shown in Figure 2.7, assuming the same $m$ value for all the curves.

<table>
<thead>
<tr>
<th>AP</th>
<th>$\Delta G$ (kcal/mol)</th>
<th>$m$ (kcal/mol/ M[GdmCl])</th>
</tr>
</thead>
<tbody>
<tr>
<td>No AP</td>
<td>7.13 ± 0.39</td>
<td>-2.67 ± 0.15</td>
</tr>
<tr>
<td>0.8% CB</td>
<td>6.88 ± 0.38</td>
<td>-2.67 ± 0.15</td>
</tr>
<tr>
<td>0.8% PE</td>
<td>6.86 ± 0.38</td>
<td>-2.67 ± 0.15</td>
</tr>
<tr>
<td>0.8% BA</td>
<td>6.64 ± 0.38</td>
<td>-2.67 ± 0.15</td>
</tr>
<tr>
<td>0.8% PH</td>
<td>6.46 ± 0.40</td>
<td>-2.67 ± 0.15</td>
</tr>
<tr>
<td>0.8% CR</td>
<td>5.95 ± 0.39</td>
<td>-2.67 ± 0.15</td>
</tr>
</tbody>
</table>
To further demonstrate that the APs enhance the unfolding of the Met80 region, we performed amide hydrogen exchange (HX) experiments using 2D NMR (Figure 2.8). We first monitored the changes in the protein structure using 2D NMR COSY spectra recorded in the absence and presence of the five APs (Figure 2.8A). No significant changes were observed in the positions of individual amide crosspeaks, indicating that the APs did not perturb the overall global protein structure. The chemical shifts of mainchain amide and Cα protons of all residues were within 0.1 ppm. We subsequently measured the stability of the Met80 region. For this purpose, we used HX in combination with 2D NMR, which measures the stability of structures around individual amino acids [106]. For an amide proton to be exchanged with the solvent, it needs to be exposed to
the solvent, and the exchange rate is proportional to the stability of the structure protecting the amide against exchange. Although individual amide protons exchange through three different types of unfolding mechanisms [106], earlier HX experiments on Cyt c indicated that the amides of the two residues Tyr74 and Ile75 exchange with the solvent only upon complete unfolding of the Met80 region [90]. In the absence of APs, these two residues exchange with the rate constants 0.57/hr and 0.27/hr, respectively. With the addition of APs, the exchange rate of these two residues increased significantly (Figure 2.8B). The amide protons exchanged at a much faster rate in the presence of PE (Tyr74: 1.73/hr and Ile75: 1.21/hr) and BA (Tyr74: 2.60/hr and Ile75: 1.19/hr), whereas they exchanged within the deadtime of the experiment (~10 min) in the presence of PH and CR. This order in which the APs enhance the amide HX of the Met80 region (Figure 2.8B) matches with that of the effectiveness of APs in inducing Cyt c aggregation (Figure 2.5).

**Stabilizing the Met80 Region Decreases Protein Aggregation.**

The above results indicate that the Met80 region may be the weakest link whose destabilization triggers Cyt c aggregation. To confirm this and to demonstrate a possible strategy of decreasing protein aggregation, we stabilized the weakest link by reducing the iron in the heme from Fe$^{+3}$ (ferric) to Fe$^{+2}$ (ferrous). Reduction of the heme increases the stability of the bond between Met80 and the heme by 3.2 kcal/mol with no change in the protein structure [107, 108]. This is very similar to introducing a site-specific stabilizing mutation in the Met80 region, without the hassle of cloning, expressing, and purifying the mutant protein. The aggregation temperature $T_{m}^{Agg}$ of Cyt c in its reduced (ferrous) state
FIGURE 2.8 Effect of APs on Cyt c solution structure and dynamics.
FIGURE 2.8 Effect of APs on Cyt c solution structure and dynamics. (a) Changes in the 2D NMR COSY fingerprint region with the addition of APs. Black, Red, Blue, Green, and Yellow colors represent the spectra with no AP, 0.8% (v/v) phenoxyethanol (PE), 0.8% (v/v) benzyl alcohol (v/v), 0.8% (v/v) phenol (PH), and 0.8% (v/v) m-cresol (CR), respectively. (b and c) Changes in the HX rates of residues Tyr74 and Ile75 with the addition of APs. The exchange of these two “marker” residues represent the partial unfolding of the Met80 region.
was determined from the change in optical density at 800 nm with increasing solution temperature. No 695 nm absorbance band is seen in the reduced form due to the absence of the charge transfer between Met80 and ferrous iron. Reduction of Cyt c increased the $T_m^{\text{Agg}}$ for all five APs and at all AP concentrations. For example, in the presence of 2% PE, oxidized Cyt c has a $T_m^{\text{Agg}}$ of 69.5°C, whereas reduced Cyt c has a $T_m^{\text{Agg}}$ above 90°C (Figure 2.9). In addition, the difference in $T_m^{\text{Agg}}$ between the oxidized and reduced forms was nearly identical for different APs (Table 2.3), indicating that the stabilization of the Met80 region may be the major factor responsible for reducing protein aggregation, consistent with the correlation seen in Figure 2.6. For most reduced samples, full aggregation curves could not be obtained due to the allowed maximum temperature values in the thermal scanning method, and hence exact $T_m^{\text{Agg}}$ values could not be determined.

**Discussion**

Multi-dose protein formulations require APs to prevent the accidental growth of microbes during repeated use [5, 43]. However, APs have been shown to cause protein aggregation [5, 28, 42, 43, 47, 48, 50, 82, 83], and the underlying physical mechanisms are poorly understood. Here, we used a model protein Cyt c to examine the effect of five APs commonly used in liquid protein formulations on protein aggregation (Figure 2.3 & 2.5), and determined the nature of partial protein unfolding that leads to protein aggregation (Figure 2.6). We also showed that stabilizing the weakest link, whose unfolding leads to protein aggregation, reduces aggregation (Figure 2.9). These results al-
FIGURE 2.9 Aggregation of reduced cytochrome c (triangles) in comparison with that of the oxidized form (circles). The AP concentrations used were 0.8% (v/v) m-cresol (CR), 1% (v/v) phenol (PH), 3% (v/v) benzyl alcohol (BA), 2% (v/v) phenoxyethanol (PE), and 1% (v/v) chlorobutanol (CB). For all APs, the reduced form aggregated at higher temperatures than the oxidized form.

TABLE 2.3 Comparison of the $T_m^{Agg}$ of Reduced and Oxidized Forms of Cyt c. There is an approximate constant difference between the $T_m^{Agg}$ values indicating that the increase in $T_m^{Agg}$ is predominantly because of the stabilization of the Met80 region due to heme reduction.

<table>
<thead>
<tr>
<th>AP</th>
<th>$T_m^{Agg}$ (Oxidized) (°C)</th>
<th>$T_m^{Agg}$ (Reduced) (°C)</th>
<th>$\Delta T_m^{Agg}$ (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3% (v/v) BA</td>
<td>57.3</td>
<td>79.9</td>
<td>22.6</td>
</tr>
<tr>
<td>2% (v/v) PE</td>
<td>69.5</td>
<td>91.1</td>
<td>21.6</td>
</tr>
<tr>
<td>3% (v/v) PE</td>
<td>61.4</td>
<td>84.4</td>
<td>23.0</td>
</tr>
</tbody>
</table>
so demonstrate that the same region acts as the aggregation ‘hot-spot’ for all APs, and therefore, any strategies developed to reduce protein aggregation induced by one AP works for all other APs.

The effect of APs on the extent of protein aggregation depends on the nature of the AP. For Cyt c, CR is the AP which causes the most aggregation, whereas CB is the AP which causes the least aggregation (Figure 2.5). Therefore, if Cyt c were a pharmaceutical protein, CB would be the right choice for the AP in its formulation rather than CR. Alternatively, lower concentrations of CR have to be used compared to that of CB, so that CR does not cause protein aggregation. It would be interesting to determine whether the APs follow the same order in terms of inducing aggregation of other proteins, in particular pharmaceutical proteins. In one case where the effect of phenolic compounds, which include three of the APs examined here, on the aggregation of recombinant human growth hormone was examined [42], the three APs followed a similar trend as that observed here for Cyt c (Figure 2.5): CR > PH > BA. In another study, CR and PH caused an antibody to precipitate in formulation, while CB showed little to no effect on protein stability [48]. The ability of APs to aggregate proteins may depend on the nature of their interactions with proteins, which needs to be determined. Initial studies in this direction indicated that APs do not have strong binding sites on proteins [50, 56]. Most of the physical properties which might control protein aggregation, such as hydrophobicity and dielectric constant, have not been characterized for the five APs examined here, and it is therefore not possible at this time to determine which specific physical parameter governs the aggregation reaction (see Appendix A).
Based on the results presented here, we propose screening of various APs or combinations of APs before selecting the right AP for a protein formulation. The ideal choice for an AP would be the one that causes the least protein aggregation yet possesses the desired antimicrobial effect. At the same time, identifying the weakest link in the protein and measuring the effect of APs on its unfolding might help in reducing protein aggregation, either by introducing site-specific mutations or by modulating solvent properties to stabilize the weakest link in a protein. In addition, one needs to ensure that the addition of AP and/or protein modifications do not compromise the function of the therapeutic protein of interest.

Stabilizing the weakest links in pharmaceutical proteins to reduce their aggregation and to improve their shelf-life has been gaining considerable attention in recent years. This is especially important in the case of multi-dose formulations that need to be stored over long periods of time and preferably at room temperatures in underdeveloped countries. For example, in therapeutic insulin, three amino acid residues that cause its aggregation have been identified, and mutating these residues decreased aggregation and increased the shelf-life of the formulation at room temperature [109]. In another study, aggregation-prone regions in therapeutic antibodies were identified and were stabilized to reduce aggregation [110]. Our results presented here show that a similar strategy of stabilizing the weakest links in a protein might also decrease its AP-induced aggregation.
CHAPTER III

HIGH YIELD SOLUBLE BACTERIAL EXPRESSION AND STREAMLINED PURIFICATION OF RECOMBINANT HUMAN INTERFERON ALPHA-2A

Introduction

The interferon family of cytokines was discovered in the late 1950s, when soluble antiviral factors conferred protection to membranes during live influenza virus challenge [111]. The interferon family is classified by its antiviral, antiproliferative, and immunoregulatory functions. These proteins are critical members of the innate immune response whose induction is initiated by the binding of various viral components to cell surface recognition receptors. Downstream transcription factors are activated via kinase pathways that lead to interferon production. Upon release from the cell, interferon binds to ubiquitously expressed cell surface receptors which trigger a signaling cascade involving activation of antiviral and antiproliferative factors as well as up-regulating their own expression. Thus, interferon induction leads to a positive feedback loop characterized by high antiviral activity [57, 58, 60, 61].

Human interferons comprise two categories: Type I and Type II. These classifications refer to the cellular receptor that is used for interferon function as well as amino acid sequence [65]. Type I interferons consist of α, β, and ω subtypes, and bind to the IFNα/β (IFNAR) receptor. Type I interferons are ubiquitously induced by viral

infection in many cell types. Type II interferon has a single subtype, \(\gamma\), which interacts with the IFN\(\gamma\) receptor. T cells and natural killer cells produce interferon \(\gamma\) following activation by early cytokines [57, 61].

Interferon \(\alpha\) consists of 13 functional subtypes of highly conserved sequence and structural homology (75-99%) found in humans [65]. Currently, two native and two PEGylated interferon \(\alpha\) products exist in the pharmaceutical market [112]. Interferon \(\alpha\)-2 products, such as interferon \(\alpha\)-2a (IFNA2) are used to treat multiple diseases such as hairy cell leukemia and hepatitis c [60, 62-64]. These drugs are available in a variety of dosage forms including multi-dose and single use syringes.

As the biopharmaceutics industry continues to grow, the study of protein-based pharmaceuticals and their formulation stability becomes increasingly important. It is often preferable to obtain proteins from an in-house expression and purification scheme due to cost or other factors. IFNA2 was initially derived directly from human leukocytes and purified for pharmaceutical use [70]. However, because IFNA2 is not glycosylated \textit{in vivo} [113], it is a perfect candidate for expression using bacteria. In most attempts of using \textit{Escherichia coli} (\textit{E. coli}), IFNA2 was regularly expressed in inclusion bodies. [71, 73, 74]. Our aim in this work was to develop and optimize an expression and purification scheme of IFNA2 in \textit{E. coli} that would not only result in high yield soluble protein, but also would require simplified purification steps in order to obtain highly pure protein.
Materials and Methods

Plasmid Construction and Screening.

E. coli codon-optimized cDNA for human interferon α-2a, corresponding to full-length human IFNA2 (Protein DataBank ID 1ITF), was synthesized by Operon (Huntsville, Alabama) and delivered in the TOPO vector. It was obtained as a lyophilized powder and reconstituted in 10 mM Tris, pH 7.0, to a concentration of 120 ng/ml. The cDNA was transformed into BL21(DE3) cells (1 µl cDNA to 1 ml competent cells), plated on kanamycin/ampicillin antibiotic agar plates, and incubated at 37°C overnight. Colonies were transferred to LB media and grown overnight at 37°C in the presence of kanamycin and ampicillin. DNA was extracted using a miniprep kit (QIAGEN) and quantified on a Nanodrop (Wilmington, Delaware). The TOPO vector contained recognition sequences for BamH1 (forward) and Xho1 (reverse) restriction sites. The IFNA2 DNA sequence was amplified with specific forward (CATATGGGATCCTGCGATTTACCGCAAACC) and reverse (GGTGCTCGAGTTACTCTTTCGAACGCAG) primers, followed by gel extraction and digestion with BamH1 and Xho1 restriction enzymes. The amplified DNA was purified using gel extraction and inserted into our vector of choice. We used the modified pET-28b vector, pET-SUMO (a generous gift from Christopher Lima (Sloan-Kettering Institute)). Ligation of the IFNA2 gene into pET-SUMO was achieved with identical forward and reverse primers listed above. Following ligation, pET-SUMO-IFNA2 was transformed into competent DH5α E. coli cells and plated on kanamycin antibiotic agar plates. Resultant colonies were screened for successful ligation using a miniprep kit and subsequent DNA sequencing (University of
Colorado Anschutz Medical Campus). A positive clone was transformed into competent BL21(DE3) *E. coli* cells for expression.

**Expression of IFNA2.**

In order to express IFNA2 in the soluble fraction of *E. coli*, the following scheme was used. A primary 5 ml culture of LB media was inoculated with 5 μL kanamycin and incubated overnight at 37°C with shaking. Primary culture was transferred to secondary (100 ml), inoculated with kanamycin and incubated at 37°C with shaking for 2.5 hours. The secondary culture was transferred to full scale (1L), treated with kanamycin, and incubated at 37°C with shaking for 3 hours. At OD\textsubscript{600}=1.0, temperature was reduced to 16°C and protein expression was induced with 0.2mM IPTG for 16 hours with shaking. Cells were harvested using centrifugation at 4200 g for 10 min. Cell pellets were stored at -80°C.

**Purification of Soluble SUMO-IFNA2.**

The cell pellet obtained in the previous step was thawed at room temperature and resuspended in 50 mM Tris, 100 mM NaCl, pH 7.5 with two dissolved protease inhibitor cocktail tablets (complete Mini EDTA-free, Roche). Cell lysis was achieved using probe sonication (16 cycles of one min on/off, half-power). Lysate material was removed via centrifugation at 17,000 g for 45 min at 4°C. Supernatant was loaded onto a 5 ml Nickel Sepharose 6 Fast Flow column (GE Healthcare Life Sciences, Pittsburgh, Pennsylvania) (equilibration buffer 50 mM Tris, 100 mM NaCl, 40 mM imidazole, pH 7.5) at 1 ml/min using a centrifugal pump. The column was washed with 50 mM Tris, 100 mM NaCl, 50
mM imidazole, pH 7.5. Protein was eluted using a linear gradient of imidazole (50-500 mM) in 50 mM Tris, 100 mM NaCl, pH 7.5. Fractions containing protein were identified via SDS-PAGE and dialyzed into 50 mM Tris, 200 mM NaCl, pH 8. Often, dilution of positive fractions was required to prevent the highly concentrated SUMO-IFNA2 from precipitating during dialysis. SUMO-IFNA2 (33.3 kDa) concentration was determined using the extinction coefficient calculated from ExPASy (www.expasy.org) of $\varepsilon_{280} = 19.54$ mM$^{-1}$.

**Removal of His$_6$-SUMO tag.**

To remove the His$_6$-SUMO tag, SUMO-IFNA2 was digested with yeast protease Ulp1 (purified in-house) (1µg Ulp1:0.25mg SUMO-IFNA2) and the cleavage reaction was incubated at room temperature for 16 hrs. Separation of SUMO tag and IFNA2 was achieved on a 5ml Nickel Sepharose 6 Fast Flow column using a step gradient of imidazole (70-500 mM) in 50 mM Tris pH 7.5, 100 mM NaCl. Pure IFNA2 fractions, collected in the flow-through and wash steps, were identified using SDS-PAGE and dialyzed into formulation buffer (10 mM ammonium acetate, 120 mM NaCl, pH 5.0). IFNA2 (19.2 kDa) concentration was determined using an extinction coefficient of $\varepsilon_{280} = 18.7$ mM$^{-1}$. Aliquots of 100 µl of 100µM IFNA2 were transferred to individual eppendorf tubes and flash frozen using liquid nitrogen and stored at -80°C.

**Deletion of Amino-Terminal Serine.**

The pET-SUMO plasmid leaves a serine residue at the N-terminus of IFNA2 following cleavage with Ulp1. This amino acid was included in order to enhance
cleavage of the protein of interest if the first N-terminal residue is a proline, lysine, valine or leucine [114]. The first N-terminal residue in IFNA2 is a cysteine; therefore the serine residue was unnecessary. In order to eliminate this amino acid, the entire pET-SUMO-IFNA2 plasmid was subjected to polymerase chain reaction using specific forward (CAG AGA ACA GAT TGG TGG ATG CGA TTT ACC GC) and reverse (GCG GTA AAT CGC ATC CAC CAA TCT GTT CTC TG) primers. These primers were designed to delete the TCC codon corresponding to serine. Following amplification, DNA was purified via gel extraction and transformed into BL21(DE3) cells. Colonies were transferred to LB media and grown overnight at 37°C in the presence of kanamycin. DNA was extracted from *E. coli* cultures and submitted for DNA sequencing. A successful sequencing result, in which the TCC codon corresponding to the excess serine residue was deleted, was transformed into BL21(DE3) cells, plated on a kanamycin antibiotic agar plate, and incubated overnight at 37°C. One resultant colony was grown in LB media overnight at 37°C and was used to create a glycerol stock. All mentions of IFNA2 in this report refer to the purified protein with a matching sequence to that of the pharmaceutical product (PDB ID 1ITF).

**Thawing Procedure.**

Prior to any experimentation, individual aliquots of IFNA2 were processed in the following manner. Flash frozen aliquots were thawed in room temperature water for a maximum of five minutes followed by centrifugation at 21,000 g for 10 min. The supernatant was transferred to a fresh eppendorf tube and the concentration was determined using an extinction coefficient of $\varepsilon_{280} = 18.54$ mM$^{-1}$. 
Analytical Ultracentrifugation.

In order to establish the assembly state of IFNA2 at relevant concentrations, sedimentation velocity experiments were conducted using a Beckman XL-A analytical centrifuge equipped with absorbance optics and an An-60 Ti rotor (Beckman Coulter, Brea, California). IFNA2 concentrations ranging from 10 µM to 50 µM in formulation buffer were tested. Samples were spun at 50,000 rpm at 4°C and absorbance was monitored at 280 nm. In each experiment, a two-channel Epon centerpiece was used. IFNA2 assembly, reflected by the sedimentation coefficient, was assessed using SedFit (www.analyticalultracentrifugation.com). The sedimentation coefficient distribution, c(s), was corrected to 20°C and water ($s_{20,w}$) using established methods [115]. Partial specific volume of IFNA2 and buffer viscosity and density were calculated using Sednterp (www.unh.edu).

Circular Dichroism (CD).

To determine IFNA2 secondary structure, circular dichroism was used. Far-UV spectra were obtained at room temperature on a Chirascan Plus spectrometer (Applied Photophysics, Surrey, United Kingdom). IFNA2 stock was diluted into water to a concentration of 10 µM and scanned between 190 nm and 250 nm.

Mass Spectrometry.

In order to confirm the molecular weight and sequence of purified IFNA2, intact protein monitoring as well as a tryptic digest in concert with mass spectrometry was done. The tryptic digest was carried out as previously described [116]. Samples were
submitted for mass spectrometry analysis (University of Colorado Denver Anschutz Medical Campus). Intact protein samples were injected on an Agilent liquid chromatography system (Santa Clara, California) in conjunction with a Waters Q-TOF2 mass spectrophotometer (Millford, Massachusetts). Tryptic digest samples were prepared using iodoacetamide and analyzed for amino-terminal residue identity using an Applied Biosystems MALDI (Carlsbad, California). Theoretical digest fragment molecular weights were obtained from Protein Prospector (www.ucsf.edu).

**Nuclear Magnetic Resonance (NMR).**

2D $^{15}$N-$^1$H heteronuclear single quantum coherence (HSQC) NMR experiments were run to confirm the structure of IFNA2. IFNA2 was singly labeled with $^{15}$N using M9 minimal media and expressed using cold induction for 16 hrs. For these experiments, 100 μM IFNA2 (50 mM acetic acid, 0.02% sodium azide, pH 3.5) was used. The NMR spectra were obtained using a Varian 900 MHz instrument equipped with a cryoprobe (Rocky Mountain NMR Facility). Data were processed using nmrPipe software [117]. IFNA2 residue assignments available in the literature [78] were used to confirm IFNA2 solution structure.

**Antiviral and Anti-Proliferative Activity Assays.**

For the antiviral assay, IFNA2 samples were tested in duplicate in a viral challenge using Encephalomyocarditis (EMC) virus on human A549 cells. Plates were stained with crystal violet, a visual cytopathic effect inhibition (CPE) was performed and the dye was then solubilized and absorbance at 570 nm was read. The data were analyzed
using a sigmoidal fit with a variable slope. PBL interferon alpha-2a (*E. coli* expressed) was run side-by-side for comparison. All data for titer and specific activity determination (U/ml and U/mg, respectively) were calculated using the 2nd WHO International Standard for interferon alpha-2a, NIBSC 95/650 (63,000 U/ml). The IFNA2 sample was also tested in triplicate in anti-proliferation assays using human ovarian cancer cells, Ovcar-3. Anti-proliferation was quantified using Promega Cell Titer 96 Aqueous Nonradioactive Cell Proliferation assay reagent (MTS, Cat# G-5430) (Madison, Wisconsin) and absorbance at 490 nm was read. The data were analyzed using a sigmoidal fit with a variable slope. The bottom portion of the anti-proliferation curve was constrained to zero (% maximum).

PBL interferon alpha-2a (*E. coli* expressed) and the 2nd WHO International Standard for interferon alpha-2a (NBSC 95/650) were run in parallel for comparison.

**Results**

Prior to the optimization of this protocol, we made numerous attempts to obtain soluble IFNA2 from *E. coli*. Soluble bacterial expression of IFNA2 was previously achieved using pET-28b and Codon-Plus *E. coli* cells (data not shown). This vector introduced a His<sub>6</sub> tag, allowing for purification on a nickel column. However, thrombin is used to remove the tag following purification, which leaves amino acids on the fusion terminus of the protein. In order to ensure that the purified protein exactly matches the pharmaceutical protein on the market, the pET-SUMO plasmid was pursued as an expression vector.
Design of the IFNA2 Expression Construct.

The plasmid for protein expression used in this study was a modified pET-28b vector, pET-SUMO. This vector contains a polyhistidine tag (His$_6$) followed by a SUMO (small ubiquitin related modifier) moiety (Figure 3.1A), a yeast-derived component whose tertiary structure is recognized specifically by the yeast cysteine protease Ulp1 [118]. Expressed IFNA2 was amino-terminally linked directly to the SUMO moiety, as shown in Figure 3.1A. The codon-optimized synthesized cDNA was designed to result in the expression of a protein with the sequence shown in Figure 3.1B. The expression system described would result in a protein consisting of His$_6$-SUMO-IFNA2.

![Diagram](Image)

**FIGURE 3.1** Post-translation fusion complex and corresponding protein sequence. A) Pictorial representation of the translated His$_6$-SUMO-IFNA2 fusion complex with the Ulp1 cleavage site indicated by an arrow. B) The primary protein sequence of IFNA2 following cleavage of the fusion tag from IFNA2. (PDB 1ITF)

Expression and Purification of Recombinant Human IFNA2.

The expression and purification scheme described in this study is shown as a flowchart in Figure 3.2. The inclusion of the His$_6$ tag allowed for simplified purification of the lysate supernatant using a nickel column. The His$_6$-SUMO-IFNA2 complex
resulted in a predominant species with a molecular weight of approximately 33.3 kDa (Eluate 3 lane in Figure 3.3).

Cleavage of the recombinant protein complex was achieved using the yeast protease Ulp1. A successful cleavage reaction resulted in two distinct species observed using SDS-PAGE (Post-treatment lane in Figure 3.4). The upper band, running at approximately 20 kDa, corresponded to the His<sub>6</sub>-SUMO tag, which is known to migrate slowly during electrophoresis resulting in a higher molecular weight estimate than its calculated value (~14 kDa) [75]. Based on the location of SDS-PAGE migration, it is likely that the impurities observed prior to cleavage were His<sub>6</sub>-SUMO-INFA2 expression fragments. The lower band seen post-cleavage corresponds to IFNA2 at a molecular weight of 19.2 kDa.

The purification protocol introduced here yielded 16 mg of purified IFNA2 per 1 liter of expression media. This yield was achieved reproducibly with the application of the steps shown in Figure 3.2.

**Minor Loss Observed Following Freeze-Thaw of IFNA2 Aliquots.**

Numerous protein systems have been shown to be sensitive to freeze-thaw damage via aggregation during freezing or thawing [119, 120]. Because our purified protein was flash frozen and subsequently stored in the frozen state, it was possible that damage could occur either during the freezing process or as a result of thawing. To test the sensitivity of our purified protein to freeze-thaw stress, individual aliquots were thawed in room temperature water and centrifuged in order to remove any potential agg-
FIGURE 3.2 Schematic of expression and purification scheme optimized to obtain soluble IFNA2 from *E. coli*.
FIGURE 3.3 Purification of His₆-SUMO-IFNA2 fusion complex. The lysate fraction was loaded onto nickel resin and the flow-through was collected. A low concentration imidazole wash allowed for further depletion of any bacterial contaminants. Eluates 2-5 show the fusion complex eluting from the nickel resin, with overexpression of the complex evident in Eluate 3 and Eluate 4 at ~33kDa. Approximate molecular weight is shown on the far left.

regates. No visible aggregates were observed. The concentration of the resulting supernatant consistently was within 10% of the approximate 100μM aliquot concentration.

Confirmation of IFNA2 Identity Using Biophysical and Structural Methods.

In order to confirm the identity of the protein obtained using our purification protocol, numerous biophysical and structural techniques were employed.
FIGURE 3.4 Cleavage of His$_6$-SUMO-IFNA2 and purification of IFNA2. Pre-treatment refers to elution fractions from Figure 3.3 prior to addition of Ulp1 protease. Following the cleavage reaction, the post-treatment lane shows two distinct bands representing the His$_6$-SUMO tag and untagged IFNA2. Pure IFNA2 is collected in the flow-through. A high concentration imidazole elution fraction shows the His$_6$-SUMO tag.

To determine the molecular weight of our purified protein with high precision, LC-MS was utilized. The main protein peak was observed at 19.2 kDa, consistent with the calculated molecular weight obtained from ExPASy (Figure 3.5A).

A benefit of using the pET-SUMO vector is the lack of additional amino acids on the protein of interest as a result of cleavage with Ulp1. To ensure that our purified protein retained no residual tag amino acids on its amino-terminus, tryptically digested gel fragments were subjected to mass spectrometry analysis. A theoretical tryptic digest was conducted using Protein Prospector, in which the protein primary sequence is used to calculate all possible digest fragments with corresponding molecular weights. Initial
FIGURE 3.5 Mass spectrophotometry analysis of IFNA2.
FIGURE 3.5 Mass spectrophotometry analysis of IFNA2. A) Intact full-length protein injection shows a predominant peak at 19237 Da. B) Tryptically digested IFNA2 initially showed a residual serine on the N-terminus of the protein, indicated by the 1457 m/z peak. The inset text shows the N-terminal fragment corresponding to this m/z value. The upper and lower panels show pre- and post-deletion of the N-terminal serine, respectively. C) Tryptically digested IFNA2 shows the presence of the 1513 m/z peak, corresponding to N-terminal cysteine. The inset text shows the N-terminal fragment at this m/z value. The upper and lower panels show pre- and post-deletion of the N-terminal serine, respectively.
analysis indicated that a residual serine preceded the expected amino-terminal fragment, at 1457 m/z (Figure 3.5B). Because the Ulp1 protease does not require the presence of serine for successful cleavage of IFNA2, the serine codon was removed using a reconstructed vector. Subsequent tryptic digest analysis demonstrated an amino-terminal digest fragment corresponding to the theoretical fragment without serine at 1526 m/z, confirming the complete removal of the His6-SUMO tag and no additional amino acids on the purified protein (Figure 3.5C). These m/z values reflect carbamidomethyl modifications on the cysteine residue, due to sample treatment with iodoacetamide [116], with an overall charge value of 1.

Proteins within the interferon family are characterized by predominantly α-helical structure [58]. Thus, the use of circular dichroism to determine secondary structural characteristics of our purified protein was employed. Stock protein was diluted in water to a concentration of 10 μM and the resultant far-UV spectrum is shown in Figure 3.6A. The CD spectrum showed peaks at 208 and 222 nm, consistent with well-folded α-helical structure. The MRE (mean residue ellipticity) obtained in this study agrees with previously documented CD spectra for IFNA2 [121], with minor deviations attributed to differences in buffer conditions and pH.

Additionally, sedimentation velocity experiments were used to estimate the protein molecular weight and to determine the assembly state of purified IFNA2. These results confirmed that the purified protein was a monomer in solution (Figure 3.6B). The corresponding molecular weight for this species was 19.5 kDa, which is in good agreement with theoretical calculations (19.2 kDa) and MS results (19.2 kDa).
The solution structure of IFNA2 has been solved using 2D-NMR techniques [78]. We performed 2D-NMR HSQC experiments at pH 3.5 to confirm the structure of our purified protein with no tag. This pH condition was chosen to reflect the solution conditions used to determine IFNA2 structure previously. The overlay of the spectrum with the published assigned crosspeaks is shown in Figure 3.7. The assigned peaks match exactly with the spectrum obtained at pH 3.5, confirming the identity and structure of our purified protein.

![Graphs](image)

**FIGURE 3.6** Biophysical characterization of IFNA2. A) Far-UV circular dichroism spectrum of IFNA2 shows that this protein is predominantly α-helical. B) Sedimentation velocity data collected using analytical ultracentrifugation demonstrates that IFNA2 is a monomer in solution at pharmaceutically relevant concentrations.

**IFNA2 Retains Antiviral and Anti-Proliferative Activity.**

To ensure that our purified protein maintained functional activity, antiviral and anti-proliferation assays were conducted (Figure 8). The cytopathic effect inhibition (CPE) assay, used to verify IFNA2 antiviral activity against EMC viral challenge on human A549 cells, showed good agreement between IFNA2 and known controls (Figure
8A, Figure 8C), with an IFNA2 activity of 3.6E8 U/mg compared to IFNA2 standard activity of 1.9E8 U/mg. The resultant half maximal effective concentration (EC50) of IFNA2 was 1.09E-13 M. IFNA2 anti-proliferative function was assessed using human ovarian cancer cells (Figure 8B, Figure 8C). In this case, comparable anti-proliferative activities were observed for IFNA2 and known standards, with an IFNA2 EC50 of 1.16E-10 M.

**FIGURE 3.7** Comparison of IFNA2 2D-NMR structure with published assignments. The blue peaks correspond to IFNA2 data collected at pH 3.5. Overlaid numerical values correspond to peak assignments as determined by Klaus et al [78].

**Discussion**

The use of protein based pharmaceuticals is becoming increasingly important in the treatment of numerous diseases for which effective small molecule drugs are not avai-
FIGURE 3.8 Antiviral and anti-proliferative activity of IFNA2 in comparison to known standards. (A) IFNA2 antiviral activity following encephalomyocarditis (EMC) viral challenge on human A549 cells. (B) IFNA2 anti-proliferative function against human ovarian cancer cells. (C) Calculated EC50 [M] values for the CPE assay in (A) and the anti-proliferative assay in (B). Interferon α-2a (IFNA2); *E. coli*-derived interferon α-2a standard (PBL IFNα2A); 2nd International WHO standard for interferon α-2a (NIBSC 95/650).

The interferon family, comprising a group of immunologic factors that are upregulated during viral or microbial onslaught, constitutes an important class of protein-based formulations. IFNA2 is one of 13 interferon-α subtypes found in humans, and is used pharmaceutically to treat various diseases including hepatitis C and leukemia [60, 62-64]. Other pharmaceutically available interferon drugs include interferon-α 2b and interferon-β.
Human IFNA2 is not glycosylated or otherwise modified post-translationally [113], indicating that the bacterial expression system of *E. coli* can be utilized. It has been demonstrated in the literature that IFNA2 can be obtained from *E. coli*. However, previously published expression and purification schemes require a) numerous purification steps in order to obtain pure, homogenous protein [72]; or b) purification from inclusion bodies and subsequent refolding [71, 73, 74]. A purification protocol necessitating multiple steps introduces a number of problems, such as enhancing protein stability concerns as well as simply increasing the time required for purification. Additionally, multiple purification steps reduce the final yield [72]. Finally, refolding following purification may result in soluble protein losses due to membrane interactions and other factors. To our knowledge, ours is the first published instance of obtaining soluble IFNA2 from *E. coli* while still obtaining an improved overall yield of protein.

The use of fusion partners to increase solubility has been demonstrated in the literature before [75-77], and interferon α-2c has been purified in the soluble fraction using maltose binding protein [122]. In particular, the SUMO fusion partner has been shown to dramatically increase the solubility of multiple unrelated protein systems, including green fluorescent protein and myostatin, in comparison to other traditional solubility tags [77]. With the addition of the SUMO tag, overexpression of IFNA2 in the soluble fraction was obtained. Another benefit of the SUMO tag lies in its structural recognition by the protease Ulp1 [118]. In many widely used fusion systems, such as His<sub>6</sub>, cleavage of the tag requires a specific amino acid sequence that is recognized by a protease. Successful cleavage removes the majority of the tag residues, but generally leaves non-native amino acids on the fusion terminus of the protein of interest [123]. In
the case of IFN2C, a non-native methionine residue was retained on the fusion terminus of the protein following cleavage of the tag [122]. However, because the SUMO moiety is structurally identified by the protease Ulp1, no residual amino acids are retained on the fusion terminus of the protein. Because of this, the protein obtained following cleavage is sequentially pure. The removal of the entire tag from purified IFNA2 was confirmed using mass spectrometry methods in conjunction with tryptic digestion.

We used cold induction to express SUMO-IFNA2 in the soluble fraction. Cold induction has been a method for encouraging the soluble expression of recombinant proteins for a number of years [124, 125]. In the case of SUMO-IFNA2, inducing at 37°C or 25°C resulted in inclusion body expression of IFNA2 and soluble expression of numerous fragmented species. By decreasing the temperature to 16°C, fragmentation was mitigated and SUMO-IFNA2 was brought into the soluble fraction. Additionally, a lower concentration of inducing agent, IPTG, was used. This allowed for slower protein expression, which promoted better folding of the protein [126].

To the best of our knowledge, this is the first instance that IFNA2 has been expressed in the soluble fraction of *E. coli* and purified in a single step to homogeneity and purity. Additionally, in comparison to other published methods using *E. coli* as an expression system, we show here a reproducible and simplified method to achieve significantly improved yields. It should be noted that this expression and purification scheme is best suited for *in vitro* applications such as formulation stability studies.
CHAPTER IV

ANTIMICROBIAL PRESERVATIVES INDUCE AGGREGATION OF
INTERFERON ALPHA-2A: THE ORDER IN WHICH PRESERVATIVES
INDUCE PROTEIN AGGREGATION IS INDEPENDENT OF THE NATURE OF
THE PROTEIN

Introduction

Multi-dose protein formulations comprise approximately one third of protein-based pharmaceuticals available on the global market [5]. These formulations are beneficial in terms of economics and patient compliance, and require the inclusion of at least one antimicrobial preservative (AP) in order to inhibit the growth of microbes and bacteria during administration [5, 43].

It has become increasingly important to study APs in protein pharmaceuticals because these necessary compounds have been linked to protein aggregation in the liquid state. One of the earliest reports demonstrated that the addition of various aromatic compounds induced the aggregation of recombinant human growth hormone [42]. Numerous studies have been published showing the ability of APs to cause destabilization and aggregation of many proteins [45, 46, 48, 82]. Aggregates in these formulations cause a decrease in the effective concentration of delivered drug as well as result in toxic immunogenic responses in patients [20-26]. In order to minimize AP-induced protein aggregation, an understanding of the interactions between APs and proteins is critical.
Previous work from our laboratory demonstrated a specific relationship between individual APs and the aggregation of a model protein Cytochrome c (Cyt c) [127]. The aim of this work is to study the effects of various APs on a pharmaceutically relevant protein and validate the results obtained using the model protein. Interferon α-2a (IFNA2) (Figure 4.1) has been shown to aggregate in its formulation state [66-69]. APs have been used in IFNA2 multi-dose liquid formulations. However, it is unknown whether other APs also cause IFNA2 aggregation. Here we demonstrate that IFNA2 is prone to aggregation in the presence of APs and that the extent and order of these effects match exactly to what we observed in the model protein studies [127].

**Materials and Methods**

**Materials.**

IFNA2 synthesized DNA was obtained from Operon (Huntsville, Alabama) and expressed using the pET-SUMO plasmid (a generous gift from Christopher Lima, Sloan-Kettering Institute). Soluble protein was purified on a Nickel Sepharose 6 Fast Flow column (GE Healthcare Life Sciences, Pittsburgh, Pennsylvania). Preservatives were obtained in their highest available purity and shown in Table 1. All experiments were performed in 10 mM ammonium acetate pH 5.0, 120 mM NaCl, unless otherwise noted.

**Size Exclusion Chromatography.**

To monitor the effects of various APs on protein aggregation, IFNA2 (10μM in 0.01M ammonium acetate, 0.12M sodium chloride, pH 5) was incubated at 50°C and
samples were taken at desired intervals. Concentration of monomer was estimated by injecting 70μL onto a TSKgel 5μM G3000SWxl column (Tosoh Bioscience LLC, San Francisco, California) on an Agilent 1100 HPLC (Santa Clara, California). The mobile phase used was 0.01M ammonium acetate, 0.12M sodium chloride, pH 5, at a flow rate of 1 mL min⁻¹. Absorbance at 280 nm was used to determine the monomer content.

![Molecular structure of interferon α-2a (IFNA2; 1ITF.pdb). The protein is α-helical in nature with helices organized according to color: helix A, residues 11-21 (orange); helix B, residues 52-68 (green); helix B’, residues 70-75 (yellow); helix C, residues 78-100 (purple); helix D, residues 110-132 (cyan); helix E, residues 137-157 (blue). Residues 22-51 comprise the AB-loop, with residues 40-43 usually found in a 3₁₀ helix (red).](image)

**FIGURE 4.1** Molecular structure of interferon α-2a (IFNA2; 1ITF.pdb). The protein is α-helical in nature with helices organized according to color: helix A, residues 11-21 (orange); helix B, residues 52-68 (green); helix B’, residues 70-75 (yellow); helix C, residues 78-100 (purple); helix D, residues 110-132 (cyan); helix E, residues 137-157 (blue). Residues 22-51 comprise the AB-loop, with residues 40-43 usually found in a 3₁₀ helix (red).

**Isothermal Incubation Experiments.**

IFNA2 (10μM) was incubated at the desired temperature with various APs, and the changes in optical density at 350 nm were measured as a function of the incubation time [128, 129]. Buffer and protein do not absorb at this wavelength. The aggregation
kinetics were monitored until the signal reached a plateau. At longer incubation times, the aggregates started to settle down to the bottom of the cuvette, resulting in decreased optical density. At that point, the experiment was stopped.

**TABLE 4.1** Antimicrobial preservatives used in this study.

<table>
<thead>
<tr>
<th>AP</th>
<th>Molecular Structure</th>
<th>Molecular Weight (Da)</th>
<th>Source</th>
<th>Purity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benzyl alcohol (BA)</td>
<td><img src="image" alt="Structure" /></td>
<td>108.1</td>
<td>Merck</td>
<td>97</td>
</tr>
<tr>
<td>m-Cresol (CR)</td>
<td><img src="image" alt="Structure" /></td>
<td>108.1</td>
<td>Sigma</td>
<td>99</td>
</tr>
<tr>
<td>Phenol (PH)</td>
<td><img src="image" alt="Structure" /></td>
<td>94.1</td>
<td>Sigma</td>
<td>99.5</td>
</tr>
<tr>
<td>2-Phenoxyethanol (PE)</td>
<td><img src="image" alt="Structure" /></td>
<td>138.2</td>
<td>Fluka</td>
<td>99.5</td>
</tr>
</tbody>
</table>

**Thermal Scanning Method.**

The aggregation temperature ($T_{m}^{Agg}$) of the protein was measured on an UV-Visible spectrophotometer (Agilent Technologies, Santa Clara, California). The temperature was increased at a rate of 1°C/step followed by 90 sec equilibration, and changes in the optical density at 450 nm were recorded [130, 131]. $T_{m}^{Agg}$ was determined as the temperature at half the maximum optical density. For these experiments, 10μm IFNA2 in formulation buffer was used with varying AP concentration.
Results

**APs Cause IFNA2 Aggregation.**

In previous studies, it has been shown that APs induce protein aggregation over a period of days or months [28, 42, 47, 48, 50, 82, 83]. In order to test the effects of different APs on protein aggregation on a convenient laboratory timescale, we accelerated aggregation kinetics by conducting isothermal incubation studies at elevated temperatures. Initially, we demonstrated that APs induce IFNA2 aggregation at 50°C by monitoring the concentration of IFNA2 monomer in solution as a function of time (Figure 4.2). After a period of one day in the absence of AP, approximately 80% IFNA2 monomer remained in solution. No monomer was detected after eight hours in the presence of 0.3% v/v m-cresol (CR) and after 12 hours in the presence of 0.3% v/v phenol (PH). Both BA and phenoxyethanol (PE) exerted a marginal effect on monomer loss, losing approximately 20% monomer content over the course of 24 hours. To differentiate the impact of BA and PE on IFNA2 monomer loss, an isothermal incubation study was performed using 0.9% v/v AP concentration (Figure 4.2 inset). In 0.9% BA, all IFNA2 monomer disappeared within eight hours at 50°C, while monomer was detected up to eight hours in PE samples. These results follow the order CR > PH > BA > PE in terms of effects on IFNA2 destabilization. We also performed isothermal incubation studies at 50°C while monitoring the change in optical density in order to correlate monomer loss with protein aggregation (Figure 4.3). In these cases, the same pattern of AP-induced destabilization was observed: CR > PH > BA > PE. No IFNA2 aggregation
FIGURE 4.2 Effects of APs on IFNA2 monomer concentration under isothermal conditions. Soluble monomer remaining in solution following 24 hours of incubation at 50°C in the presence and absence of 0.3% (v/v) AP as measured by size exclusion chromatography. The vertical bars represent each sample set: black – native (no AP); red – m-cresol (CR); green – phenol (PH); yellow – benzyl alcohol (BA); blue – phenoxyethanol (PE). Error bars indicate triplicate data. Individual time points are normalized with respect to No AP at Hour 0.

was observed in the absence of APs. Isothermal incubation experiments using Cyt c demonstrated an identical order of AP-induced aggregation: CR > PH > BA > PE [127].

IFNA2 Aggregation Increases with AP Concentration.

Another method used in the literature to monitor solution condition effects on protein stability and aggregation is thermal scanning. We utilized this technique to deter-
mine the effects of preservative concentration on IFNA2 aggregation, measured by changes in the optical density at 450 nm. This wavelength was selected because neither the protein nor any solution components absorb in this range, and the changes in optical density can be attributed to protein aggregation. In buffer alone, the optical density began at baseline until the signal increased as temperature was raised, and the signal reached a plateau. At higher temperatures beyond the plateau region, aggregated protein particles began to settle to the bottom of the cuvette, causing an observable decrease in the optical density. We performed these thermal scanning experiments in the presence of varying concentrations of each BA (Figure 4.4). With the inclusion of BA, the midpoint temperature of aggregation ($T_{m, Agg}$) decreased. In the absence of BA, the $T_{m, Agg}$ of IFNA2 was 63.9±0.9°C. The addition of 1.5% BA decreased the $T_{m, Agg}$ to 50.1±0.5°C, indicating
FIGURE 4.4 Variation in the optical density at 450 nm as a function of increasing temperature at different concentrations of phenol (PH).

that the presence of BA induced the aggregation of IFNA2. Similarly, Cyt c aggregation was observed to be dependent upon AP concentration [127].

Aggregation Temperature Decreases Linearly with AP Concentration.

As seen with BA, the other APs also caused a decrease in $T_{m}^{\text{Agg}}$ of IFNA2 (Figure 4.5). Interestingly, a linear correlation was observed between the $T_{m}^{\text{Agg}}$ and the concentration of AP. This relationship is significant in that it is a qualitative measurement of how efficiently each AP causes IFN aggregation, which is represented in their individual slope values. This correlation is similar to the relationship seen in $\Delta G$ and
denaturant concentration. Comparison of these slopes demonstrates an effective order of CR > PH > BA > PE, with CR being the most effective AP in causing protein aggregation. These experiments indicate that for every percent concentration change of CR, the $T_{m}^{agg}$ of IFNA2 decreases by 22°C. In contrast, PE has a slope of 8.2°C/%v/v, nearly three-fold less in its effectiveness to aggregation IFNA2 than CR. This trend is identical to that observed in the isothermal incubation studies (Figure 4.3).

Protein therapeutics are marketed in a variety of dosage forms, including multi-dose formulations, which require the inclusion of at least one antimicrobial preservative in order to combat the growth of microbes and bacteria during repeated contact between the solution and a syringe needle [5, 17, 43]. However, it has been shown that these preservatives may cause protein aggregation [28, 42, 47, 48, 50, 82, 83]. We have demonstrated previously that APs used in liquid protein formulations lead to protein destabilization and aggregation using the model protein Cyt c [55, 56, 127]. The extent of this effect was dependent upon the nature of the AP, and the pattern of destabilization observed was CR > PH > BA > PE.

**Discussion**

In this study, we tested the conclusions drawn from model studies on a pharmaceutically relevant protein interferon α-2a. IFNA2 belongs to a family of cytokines that play crucial roles in the innate immune response, and is one of numerous interferon-α subtypes found in humans. A number of interferon-α products exist in the pharmaceutical market and are used to treat various debilitating diseases including hairy
FIGURE 4.5 Dependence of $T_m^{\text{Agg}}$ on AP concentration. Individual panels demonstrate the slope of this variation.

IFNA2 multi-dose formulations have been shown to aggregate in the liquid state [66-69] and contain APs. In order to validate the conclusions drawn from our studies on a model protein Cyt c, we studied the effect of multiple APs on IFNA2 aggregation using various biophysical techniques. The influence of APs on IFNA2 stability and aggregation was assessed with numerous biophysical techniques. Isothermal incubation studies showed that APs cause the loss of IFNA2 monomer (Figure 2), and the increase in optical density due to protein aggregation (Figure 3). The order in which APs induce IFNA2 aggregation was $\text{CR} > \text{PH} > \text{BA} > \text{PE}$, which is the same order we observed earlier in the case of the model protein Cyt c [127]. Further, the addition of AP caused a decrease in $T_m^{\text{Agg}}$. The slopes indicate that CR was the most efficient at aggregating IFNA2, whereas
PE was the least. Again, this order exactly matched the pattern we observed in our earlier studies on model protein Cyt c [127]. Based on these results, PE appears to be the best preservative choice for pharmaceutical formulations in comparison with the commonly used BA.

Similar comparisons, although not as extensive as demonstrated in the case of IFNA2 or Cyt c, were made earlier on other proteins. Recombinant human growth hormone (rhGH) aggregation was monitored during freezing, high-temperature incubation, and agitation using changes in optical density and the percentage monomer loss by SEC. In this case, the three APs followed a similar order in causing rhGH aggregation: CR > PH > BA [42]. In the case of a monoclonal antibody (IgG) [48], aggregation was monitored during isothermal incubation using visual inspection of the samples, percent monomer loss using SEC, and light scattering. The three APs examined induced protein aggregation in the order: CR > PH > BA. Changes in interleukin-1 receptor (IL-1R) melting temperature, assessed by DSC, and monomer content, measured using SEC, in the presence of preservatives also showed a similar relationship [47].

Taken together, the data on these proteins (Cyt c, IFNA2, rhGH, IgG, IL-1R), suggest that the order in which APs induce protein aggregation is independent of the nature of protein. Cyt c, IFNA2, and rhGH are α-helical proteins, whereas the IgG is primarily β-sheet and IL-1R contains both α-helix and β-sheet components. The proteins also differ in their function: Cyt c plays a role in electron transport; IFNA2, IgG, and IL-1R have critical functions in the immune system; and rhGH is a multi-functional protein with emphasis on growth and regeneration.
The above observations indicate that the effect of APs on the extent of protein aggregation depends on the nature of the AP and how it interacts with proteins. It is apparent that APs induce conformational changes and partial protein unfolding, but the mechanism of this interaction remains unclear. Initial studies indicated that APs do not have strong binding sites on proteins [42, 50, 132]. However, hydrogen bonding between the peptide backbone and benzyl alcohol has been shown with polyproline by monitoring changes in amide I, amide II, and hydroxyl bands using infrared spectroscopy [133]. The potential for APs to hydrogen bond with the peptide group depends heavily on the structure of the AP as well as solution conditions. It was recently suggested that benzyl alcohol interacts with a hydrophilic region of interleukin-1 receptor [49], rather than through hydrophobic interactions as was previously speculated [46]. Again, this interaction can be influenced significantly by solution pH and other components. Perhaps the effects of APs on proteins are a combination of hydrogen bonding, hydrophobic contacts, and electrostatics, but further examination of the exact mechanism is necessary in order to develop stable and effective pharmaceuticals.
CHAPTER V
THE ROLE OF BENZYL ALCOHOL IN THE UNFOLDING AND
AGGREGATION OF INTERFERON ALPHA–2A

Introduction

Protein-based pharmaceuticals comprise nearly half of all parenteral products available worldwide [4, 19]. One third of these products are available as multi-dose formulations [5], which are advantageous for patient compliance as well as economics. In addition to a shelf-life stability requirement of 18-24 months [17], multi-dose formulations also necessitate the inclusion of antimicrobial preservatives (APs) [5, 43] to combat the growth of bacteria and other microbes during repeated contact between the solution and a syringe needle. APs are also found in topical ointments [79] and in multi-dose delivery systems [81, 134].

In recent years, the study of APs in protein pharmaceutics has become increasingly important due to the fact that these compounds have been shown to cause protein aggregation. One of the first reports demonstrated that the addition of various aromatic and aliphatic alcohols caused aggregation of human growth hormone [42]. In recent studies, APs have been shown to induce the aggregation of numerous protein systems, including interleukin-1 receptor antagonist [45, 46], interferon-γ [82] and an antibody [48]. Such protein aggregates in formulations can affect the amount and efficacy of the delivered drug as well as promote undesirable toxic and immunologic responses in patients [20-26, 84, 85, 88]. Therefore, an understanding of the molecular mechanisms
underlying AP-induced protein aggregation and the development of strategies to minimize such aggregation are critical in developing stable multi-dose drugs.

Previously published work from our laboratory demonstrated that local protein unfolding, rather than global unfolding, leads to aggregation. This local unfolding identified an aggregation “hot-spot” that could be modified to increase stability and decrease aggregation [55, 56, 127]. It is important to test the principles learned from model protein studies on a pharmaceutically relevant protein to determine if these behaviors are general to AP-induced aggregation. For this purpose, we chose to study interferon α-2a (IFNA2) (Figure 5.1). IFNA2 has been shown in the literature to aggregate under formulation conditions [66-69], which include the preservative benzyl alcohol (BA). Whether BA causes IFNA2 aggregation has not been explored. We show here that IFNA2 is susceptible to aggregation in the presence of multiple APs, and that the aggregation mechanism proceeds through a partially unfolded intermediate.

Materials and Methods

Materials.

IFNA2 synthesized DNA was obtained from Operon (Huntsville, Alabama) and expressed using the pET-SUMO plasmid (a generous gift from Christopher Lima, Sloan-Kettering Institute). Protein was expressed in Escherichia coli BL21(DE3) cells, and the soluble protein was purified using a Nickel Sepharose 6 Fast Flow column (GE Healthcare Life Sciences, Pittsburgh, Pennsylvania). The SUMO tag was cleaved using the Ulp1 protease, leaving no additional amino acids. The final protein sequence is
identical to that of the pharmaceutical protein, which was confirmed by mass spectrometry.

**Isothermal Incubation Experiments.**

IFNA2 (10 µM) in formulation buffer (0.01 M ammonium acetate, 0.12 M sodium chloride, pH 5) was incubated at the desired temperature, and the changes in optical density at 350 nm were measured as a function of the incubation time [128, 129]. Buffer and protein do not absorb at this wavelength. The aggregation kinetics were monitored until the signal reached a plateau. At longer incubation times, the aggregates started to settle down to the bottom of the cuvette, resulting in decreased optical density. At that point, the experiment was stopped.

**Thermal Scanning Method.**

The aggregation temperature (T$_{m}^{\text{Agg}}$) of the protein was measured on an UV-Visible spectrophotometer (Agilent Technologies, Santa Clara, California). The temperature was increased at a rate of 1°C/step followed by 90 second equilibration, and changes in the optical density at 450 nm were monitored [130, 131]. T$_{m}^{\text{Agg}}$ was determined as half the maximum optical density. For these experiments, 10 µm IFNA2 in formulation buffer buffer (0.01 M ammonium acetate, 0.12 M sodium chloride, pH 5) was used with varying BA concentration.
**Denaturant Melts.**

Guanidinium Chloride (GdmCl) was used as the denaturant. Protein solutions at varying GdmCl concentrations were prepared and equilibrated for one hour at room temperature before measuring changes in fluorescence signals as a function of the denaturant concentration. Concentration of the denaturant was determined using refractive index measurements [102]. For measuring changes in the 280 nm (aromatic) fluorescence, 3 µM IFNA2 in formulation buffer buffer (0.01 M ammonium acetate, 0.12 M sodium chloride, pH 5) was used with varying concentrations of BA. Fluorescence spectra were obtained using a Quantamaster Spectrofluorometer (Photon Technology International, Birmingham, New Jersey) with 4nm slits. The ΔG values were determined by fitting the changes in optical signals at different denaturant concentrations to a two-state unfolding model [103, 104].

**ANS Fluorescence.**

1-Anilinonaphthalene-8-Sulfonic Acid (ANS) (Sigma, St. Louis, MO) was used to monitor partial unfolding of IFNA2. The same experimental setup described above for denaturant melts was employed for ANS fluorescence experiments. Fluorescence intensity was measured using a Quantamaster Spectrofluorometer. Excitation and emission wavelengths used were 360 nm and 480 nm, respectively, and slit widths were 4 nm. All samples were normalized to the fluorescence standard rhodamine.
FIGURE 5.1 Molecular structure of interferon α-2a (IFNA2; 1ITF.pdb). The protein is α-helical in nature with helices organized according to color: helix A, residues 11-21 (orange); helix B, residues 52-68 (green); helix B’, residues 70-75 (yellow); helix C, residues 78-100 (purple); helix D, residues 110-132 (cyan); helix E, residues 137-157 (blue). Residues 22-51 comprise the AB-loop, with residues 40-43 usually found in a $3_{10}$ helix (red).

Nuclear Magnetic Resonance.

BA Titration.

2D $^{15}$N-$^1$H heteronuclear single quantum coherence (HSQC) experiments were run to monitor changes in IFNA2 amide crosspeaks as a function of BA concentration using a Varian Inova 600 MHz NMR instrument equipped with a cryoprobe (Palo Alto, California). IFNA2 was singly labeled with $^{15}$N using M9 minimal media and expressed using cold induction for 16 hours. For these experiments, 100 μM IFNA2 (0.01 M ammonium acetate, 0.12 M sodium chloride, pH 5) and a gradient of BA concentrations were used. Samples included 0%, 0.3%, 0.6% and 0.9% v/v BA. HSQC spectra were collected over eight hours (single scan). Changes in the peak positions (chemical shifts)
and peak volumes were calculated using the nmrDraw package (Frank Delaglio, National Institutes of Health). IFNA2 residue assignments available in the literature [78] were used for this purpose.

**Hydrogen Exchange.**

2D $^{15}$N-$^1$H heteronuclear single quantum coherence (HSQC) experiments were run to monitor changes in IFNA2 amide crosspeaks as a function of BA using a Varian Inova 900 MHz NMR instrument equipped with a cryoprobe (Palo Alto, California). For these experiments, 200 µM IFNA2 (0.05 M D-acetic acid, 0.02% sodium azide pH 3.5) was used. A time-zero HSQC of IFNA2 was recorded in aqueous buffer (0.05M acetic acid, 0.002% sodium azide, pH 3.5) prior to exchange. Exchange was carried out on a Q-Sepharose size-exclusion column (bed volume 4.5ml) for two minutes and HSQC spectra were recorded every hour. Changes in peak positions and peak volumes were calculated using the nmrDraw package.

**Results**

**BA Induces IFNA2 Aggregation.**

Previous work has demonstrated that BA cause aggregation of pharmaceutical proteins over a course of several days and months when stored close to room temperature [28, 42, 47, 48, 50, 82, 83]. Screening the effects of BA-induced aggregation is not feasible due to long incubation times. Therefore, to probe aggregation mechanisms on a convenient laboratory scale, we accelerated aggregation kinetics by performing
isothermal incubation studies at higher temperatures. We examined the effect of BA concentration on the kinetics of protein aggregation. With the increase of BA concentration at 42°C, the isothermal kinetics of IFNA2 aggregation increased (Figure 5.2).

![Graph](image)

**FIGURE 5.2** Isothermal kinetics of IFNA2 aggregation in the presence of BA. Aggregation kinetics of IFNA2 at 42°C with varying concentrations of BA: 2% v/v (yellow), 1.5% v/v BA (green), 0.9% v/v BA (red), and 0% v/v BA (black).

**BA Decreases the Temperature at Which IFNA2 Aggregates.**

An alternative method used in the literature to probe the effects of various solution conditions on protein aggregation is to monitor their influence on the temperature at which the protein aggregates during thermal scanning [28, 46, 50, 55, 56]. For this purpose, we measured the optical density at 450 nm [130, 131]. At this wavelength, the protein and the buffer do not absorb, and hence the observed changes in
the optical density can be attributed solely to protein aggregation. Initially, the signal increased with increasing temperature and a plateau was reached. At higher temperatures beyond the plateau, the optical density decreased as protein aggregates began to settle to the bottom of the cuvette. We performed this thermal scanning experiment in the presence of varying concentrations of BA to measure its effect on the midpoint temperature of aggregation, $T_{m}^{Agg}$, of IFNA2 (Figure 5.3). With the addition of BA, the $T_{m}^{Agg}$ decreased, as determined from the 450 nm optical density curves. In the absence of BA, the $T_{m}^{Agg}$ was 63.9°C ±0.9, whereas the addition of 2% BA decreased the $T_{m}^{Agg}$ to 42.1°C ±0.3, indicating that the presence of BA enhanced IFNA2 aggregation.

FIGURE 5.3 Variation in the optical density at 450 nm as a function of increasing temperature and varying concentrations of benzyl alcohol.
**Aggregation Temperature Decreases Linearly with BA Concentration.**

Following thermal scanning experiments, the aggregation temperature was plotted as a function of BA concentration. This is qualitatively similar to how ∆G or k_u varies linearly with the addition of denaturant. A linear correlation was observed between the concentration of BA and the $T_{m^{\text{Agg}}}$ of IFNA2. This correlation is significant in that it is a qualitative measure of how effective BA is in inducing aggregation of IFNA2, reflected in the value of the slope. The corresponding slope for BA indicates that for every percent change of BA, the $T_{m^{\text{Agg}}}$ of IFNA2 decreases by 10.5°C (Figure 5.4).

![Graph showing the decrease in $T_{m^{\text{Agg}}}$ as a function of benzyl alcohol concentration. The calculated slope is a qualitative measure of BA’s ability to aggregate IFNA2.](image)

**FIGURE 5.4** Decrease in $T_{m^{\text{Agg}}}$ as a function of benzyl alcohol concentration. The calculated slope is a qualitative measure of BA’s ability to aggregate IFNA2.

**BA Causes Destabilization of IFNA2.**

In order to monitor the effect of BA on IFNA2 stability, we followed changes in aromatic fluorescence as a function of BA and denaturant concentration (Figure 5.5A). With the addition of increasing concentrations of BA, the calculated ∆G decreased.
significantly (Figure 5.5B). These denaturant melts indicate that BA causes protein destabilization. Changes in the corresponding \( m \)-values were also observed (Table 5.1). The \( m \)-value is a representation of the accessible surface area that is exposed as a protein moves from the folded to the unfolded state. The decrease in \( m \)-value while maintaining the same midpoint (\( C_m \)) may indicate the presence of an intermediate [135].

**BA Preferentially Populates an Intermediate.**

In order to further test the possible formation of an intermediate in the presence of BA, we used the hydrophobic fluorescent dye ANS. ANS fluorescence has been shown to detect partially unfolded intermediate states in the presence of GdmCl [136]. The denaturant concentration at which maximum ANS fluorescence was observed was used for comparative purposes. With increasing concentrations of BA, the maximum ANS intensity shifted to lower denaturant concentrations, indicating an earlier onset of partial protein unfolding and intermediate formation (Figure 5.6). For example, in the presence of 0% BA, maximum ANS fluorescence was observed at 3.7M GdmCl. In contrast, 1.5% BA shifted the maximum ANS fluorescence to 3.0M GdmCl.

**BA Perturbs IFNA2 Locally.**

To better understand the specific effects of BA on IFNA2 structure, a BA titration was performed using 2D \(^1\text{H}-^{15}\text{N}\) HSQC NMR. With an increase in BA concentration, numerous \(^1\text{H}-^{15}\text{N}\) peaks were perturbed, which were monitored through chemical shift and volume changes. These changes were used to identify the local regions in IFNA2 that were most susceptible to influences by BA (Figures 5.7 and 5.8). All BA concentrations
FIGURE 5.5 Destabilization of IFNA2 as a function of BA concentration in the presence of the denaturant GdmHCl at room temperature. (A) IFNA2 denaturant melts in the presence of varying BA concentrations. (B) Change in the free energy of unfolding ($\Delta\Delta G$) as a function of BA relative to 0% v/v BA.
TABLE 5.1 Parameters obtained from fitting denaturant melt data in Figure 5.5 to a 2-state unfolding model. The decrease in $m$ value with no significant change in $C_m$ implies the presence of an intermediate.

<table>
<thead>
<tr>
<th>BA</th>
<th>$\Delta G$ (kcal/mol)</th>
<th>$m$ (kcal/mol/M [GdnHCl])</th>
<th>$C_m$ (M [GdnHCl])</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0%</td>
<td>10.2 ± 0.6</td>
<td>-2.6 ± 0.2</td>
<td>3.9 ± 0.02</td>
</tr>
<tr>
<td>0.9%</td>
<td>8.4 ± 0.4</td>
<td>-2.1 ± 0.1</td>
<td>4.0 ± 0.09</td>
</tr>
<tr>
<td>1.5%</td>
<td>5.5 ± 0.9</td>
<td>-1.4 ± 0.2</td>
<td>3.9 ± 0.02</td>
</tr>
<tr>
<td>2.0%</td>
<td>3.8 ± 0.3</td>
<td>-0.9 ± 0.1</td>
<td>4.2 ± 0.16</td>
</tr>
</tbody>
</table>

showed $^1$H chemical shift changes less than 0.1ppm. However, multiple residues demonstrated significant chemical shift alterations (i.e., greater than 0.1ppm) in the $^{15}$N dimension. The most substantial changes are observed in residues within the AB-loop and Helix A (Figures 5.1 and 5.7). For example, between 0% and 0.3% v/v BA, the $^{15}$N chemical shift of F27 increased by nearly 0.4 ppm. As BA concentration was increased, changes in chemical shifts were observed across the entire protein, indicating the propagation of unfolding to other regions of the protein.

Hydrogen exchange experiments in the presence and absence of BA were conducted in order to monitor the effects of BA on individual amino acid stability (Figure 5.8). Nearly all of the residues in the AB-loop exchanged within the deadtime of the experiment, including residue F27, demonstrating that the AB-loop is the most dynamic region of IFNA2. Residues L17, A19, M21 and L30 demonstrated significantly increased exchange rates in the presence of 0.9% v/v BA. These residues are located within the terminus of Helix A and the AB-loop. Additionally, W76, D82, and K83 in Helix C show
FIGURE 5.6 Intermediate detection as a function of BA concentration in the presence of the denaturant GdmHCl at room temperature. (A) Change in denaturant midpoint at maximum ANS intensity as a function of BA concentration. (B) Increase in fluorescence emission intensity of ANS as a function of BA concentration. Inset shows normalized ANS intensity curves as a function of BA and GdmHCl concentrations.
FIGURE 5.7 IFNA2 chemical shift changes as a function of BA concentration in the $^{15}$N dimension. The change in crosspeak represents the absolute change from 0% v/v BA. Top panel: 0.3% v/v BA; middle panel: 0.6% v/v BA; bottom panel: 0.9% v/v BA. The black rectangle shows the region with the highest chemical shift changes compared to 0% v/v BA.
FIGURE 5.8 Hydrogen exchange rates in IFNA2 as measured by 2D-NMR plotted as a function of residue position. The rate values represent the relative increase of individual amino acid exchange rate in the presence of 0.9% v/v BA as compared to 0% v/v BA. Colored bars indicate helices: orange (Helix A), red (3_10 Helix), green (B Helix), yellow (B’ Helix), purple (Helix C), light blue (Helix D), dark blue (Helix E). Black rectangles show maximum HX rate changes compared to 0% v/v BA.

increased exchange rates with the addition of BA. In the three-dimensional structure of IFNA2, L17 docks against K83 and has a contact area of 43.5 Å². Residue M21 and W76 have a similar relationship, with a contact area of 33.4 Å² (Figure 5.9).

In light of the HX results that show partial protein unfolding, we used computational programs available in the literature to identify regions within IFNA2 that may act as “hot-spots” for aggregation. Protein aggregation is a multi-molecular reaction...
FIGURE 5.9 Amino acid contacts within IFNA2 three-dimensional structure. (A) Critical residues in the AB-loop are not protected through intra-protein contacts and are highly labile in the solution state. (B) Residue L17 in Helix A shares a contact area of 43.5 Å² with K83 in Helix C. (C) Residue M21 in Helix A shares a contact area of 33.4 Å² with W76 in Helix C.

that is governed by hydrophobic interactions, secondary structural elements, electrostatics, and other physical properties [137-139]. Aggregation prediction programs, such as TANGO [138] and AGGRESCAN [140], take these parameters into account and determine specific regions that are aggregation-prone. The 3D analysis program ZipperDB [141] uses the available crystal structure of a protein to determine its propensity for assuming cross-β sheet structure, the most common structure found in protein aggregates. These programs were used in conjunction to identify three
aggregation hot-spots in IFNA2 (Figure 5.10). One of these predicted hot-spots is within the highly dynamic AB-loop, with another located within Helix A. BA appears to affect multiple areas of IFNA2 to varying degrees, but imparts significant influence on the AB-loop and within Helix A, evidenced by significant changes observed in chemical shifts and HX rates for many amino acids in this region (Figures 5.7). The third predicted hot-spot is within Helix D, which did not show significant changes in crosspeak position or volume in either the titration experiment or hydrogen exchange, demonstrating the importance of conducting experiments such as NMR rather than relying upon computational predictions.

### Discussion

Multi-dose formulations require the inclusion of APs to combat the growth of bacteria and other microbes during administration [5, 17, 43]. However, recent studies have suggested a link between APs and protein aggregation [5, 17, 28, 42, 43, 47, 48, 50, 82, 83]. The ideal method of understanding aggregation mechanisms is to study how various APs induce the aggregation of a pharmaceutical protein of interest using biophysical methods. In our previous work, we used a model protein, Cyt c, which has been well-characterized in the literature, to monitor preservative-induced protein aggregation in response to various stresses [55, 56, 127]. We demonstrated that BA, the most commonly used preservative in liquid formulation, caused protein aggregation, and the aggregation was preceded by a local unfolding event rather than the global unfolding of the entire protein. Additionally, we showed that stabilization of this region mitigated aggregation in the presence of BA. These results were significant in that they were the
first to identify the interaction of BA with proteins is mechanistically local in nature, rather than global. Using NMR and additional high resolution techniques provided a clear advantage to other AP-induced aggregation studies in the literature.

In this study, we tested the conclusions obtained from our model system on the pharmaceutically relevant protein interferon α-2a. IFNA2 is one of 13 interferon-α subtypes found in humans and belongs to the cytokine family of proteins, which plays critical roles in the innate immune response. Interferons are naturally produced as a result of viral or microbial infections and activate a non-specific, self-perpetuating response by binding to cell surface receptors to initiate a downstream cascade of anti-proliferative and anti-viral factors [57-61]. As a result, interferon-based pharmaceuticals are used to treat numerous debilitating diseases including various leukemias and hepatitis C [62-64].

IFN2A therapeutic formulations have been known to aggregate for nearly two decades [66-69]. Numerous studies have been performed in order to demonstrate the occurrence of IFNA2 aggregation in formulation. For example, in the late 1990s Hochuli tested numerous IFNA2 formulations for aggregates and determined that the inclusion of human serum albumin increased aggregation propensity of the entire system. As a result, HSA was removed from the formulation design and replaced with polysorbate 80 as a stabilizer [69]. However, IFNA2 formulations still contain a measurable population of aggregates and contain BA as a preservative. In this work, we have tested whether BA cause IFNA2 aggregation. The 2D-NMR solution structure for IFNA2 has been solved [78], which expedited NMR studies and allowed us to monitor changes on an amino acid level induced by BA.
FIGURE 5.10 Aggregation propensity analyses by various software programs. Highlighted sections (within black rectangles) indicate overlapping regions that were predicted to be aggregation hot spots by all three programs.
The influence of BA on IFNA2 stability and aggregation was monitored with numerous biophysical and structural techniques. It was shown that BA leads to IFNA2 aggregation in a concentration-dependent manner. IFNA2 $\Delta G$ decreased in the presence of increasing concentrations of BA, with a constant $C_m$ but a decreasing $m$-value. The $m$-value is related to the accessible surface area of the protein between the folded and unfolded states. Changes in this value suggest a deviation from two-state behavior, and may indicate the presence of an intermediate in the unfolding transition [135]. Using the fluorescent dye ANS, we demonstrated that IFNA2 contains a partially unfolded intermediate that is preferentially populated in the presence of BA.

Multiple aggregation propensity programs were utilized to determine potential aggregation “hot-spots” in IFNA2. Two of the predicted “hot-spots” were significantly perturbed in the presence of BA (Figures 5.7 and 5.8), and these regions show share surface contacts in the folded protein structure (Figure 5.9). It is possible that unfolding of one or both of these regions leads to the formation of a partially unfolded intermediate. BA might be influencing this region by simply disrupting the contacts between these critical residues and initiating a local unfolding event that populates an aggregation-prone species. These data indicate that the AB-loop and the terminus of Helix A may play major roles in initiating IFNA2 aggregation.

Another possible interpretation is the disruption of these contacts causes a domain swap, leading to new interhelical connections in a native-like configuration. Significant chemical shift changes were observed in Helix A and the AB-loop under exchanging conditions, indicating that the amino acids in this region move with the addition of BA. There are a number of other examples demonstrating the relationship between domain
swapping and the formation of protein aggregates. Domain swapping has been linked to oligomer formation and is an important step in the aggregation pathway of cystatin c [142] as well as RNase A [143] and T7 endonuclease I [144]. Additionally, cytochrome c (Cyt c) has been shown to assemble into higher order species through the successive domain swapping of helices from one monomer to another monomer [97].

Modifying the aggregation “hot-spots” of IFNA2 via specific mutations may confer stability against BA-induced aggregation. However, introducing mutations into a pharmaceutical protein must be done with caution. Previous work using alanine scanning demonstrated that a number of residues located in the AB-loop are critical for receptor binding, including F27, L30, and K31 [74, 145, 146]. On the other hand, IFNA2 has been pegylated in a formulation in order to increase drug lifetime by protecting the protein from proteolytic degradation, and has been shown to inhibit aggregation [147]. Pegylation occurs in six locations in IFNA2 including at K31 in the AB-loop. [148] This positional isomer has very low affinity to the receptor compared to unmodified IFNA2 (520 nM versus 5 nM, respectively) [149]. It is possible that mutating residues within the AB-loop may confer stability at the cost of potency, but further studies are necessary to determine if the potential loss of activity is comparable to other modified systems. Mutating a pharmaceutical protein to minimize aggregation must be done carefully to ensure that the modified drug product maintains efficacy and safety [150].

Understanding the specific effects of preservatives on proteins is critical to the development of safer and more potent protein pharmaceuticals. We have shown that the BA influences model proteins and pharmaceutical proteins in a similar manner. In-depth studies on other protein formulations in the presence of APs are required in order to
further test the principles we have demonstrated here and on Cyt c [127]. Probing the specific mechanism of interaction between APs and proteins will also be useful in the development of stable pharmaceuticals. Recently, it was suggested that benzyl alcohol interacts with a hydrophilic region of interleukin-1 receptor antagonist [49], rather than via hydrophobic interactions as was speculated previously [46]. At this time, it remains unclear if all APs interact with proteins in the same manner, and why APs influence protein stability and aggregation to different degrees.

**Conclusions**

BA, the most commonly used AP in liquid pharmaceutical formulations, caused the aggregation of the pharmaceutically relevant protein IFNA2. BA induced IFNA2 aggregation in a similar manner as was observed in previously reported model protein studies using Cyt c, namely by perturbing a local region of the protein that correlated with an increase in aggregation. Further studies are required in order to determine the specific mechanism of interaction between BA and proteins, whether it is due to direct binding, influence of solution parameters such as hydrophobicity, or other physical factors (see Appendix A).
CHAPTER VI
SUMMARY AND FUTURE DIRECTIONS

As protein- and peptide-based drugs continue to grow in the pharmaceutical market, an understanding of the mechanistic interactions between solution components and proteins becomes increasingly important. The results presented in this thesis demonstrate novel findings regarding the influence of APs on protein stability and aggregation. In Chapter II, APs were shown to cause local unfolding of the model protein Cyt c, and this unfolding event was an initiating factor in AP-induced aggregation. This chapter demonstrated for the first time the nature of the aggregation “hot-spot,” which was determined using structure-based studies in the presence of multiple APs. The APs induced Cyt c aggregation in the pattern CR > PH > BA > PE > CB. Stabilizing this “hot-spot” region conferred protection against aggregation. In order to validate these results on the model protein Cyt c, the effects of APs on the pharmaceutical protein IFNA2 were examined. An original IFNA2 expression and purification scheme was developed in-house using E. coli, which is discussed in Chapter III. This method improved soluble IFNA2 fraction expression using the SUMO fusion partner which was fully cleaved using the specific protease Ulp1, leaving no residual amino acids on the fusion terminus. An increase in the overall protein yield was observed in comparison to previous methods, which consistently resulted in bacterial expression in inclusion bodies and required multiple purification steps. Additionally, the pure protein was structurally identical to the pharmaceutically available drug and maintained antiviral and anti-proliferative activity compared to the World Health Organization international
standards. The influence of APs on IFNA2 was assessed and demonstrated to be identical to what was observed in the case of the model protein Cyt c, as was shown in Chapter IV to be CR > PH > BA > PE. Because BA is the AP of choice in IFNA2 liquid formulation, its specific effects on IFNA2 were explored in Chapter V. BA appears to preferentially populate an IFNA2 partially unfolded intermediate. Furthermore, BA significantly perturbs a specific local region in IFNA2 structure that may initiate aggregation. This is the first instance of using various biophysical and structural techniques to identify the aggregation “hot-spots” in IFNA2.

Based on these data, it is apparent that APs influence proteins in a similar fashion, as was seen with Cyt c and IFNA2. Few other studies have examined AP effects on protein stability and aggregation: rhGH, a monoclonal antibody, and IL-1R demonstrated the same pattern of AP-induced aggregation (CR > PH > BA) as has been shown here. For both Cyt c and IFNA2, it appears that APs influence protein stability and aggregation via disruption of local interactions that leads to the formation of an aggregation-prone species.

Despite these key findings, new questions have arisen. Namely, how do these conclusions relate to biological systems and the pharmaceutical industry at large? Also, based on these results, where do we go from here?

**Biological Relevance**

The mechanism of AP-induced aggregation of IFNA2 might be initiated by disruption of interhelical contacts between Helix A and Helix C, demonstrated by 2D
NMR and HX experiments (Chapter V). These residues (Leu17 and Lys83, Met21 and Trp76) may form critical contacts that are required to maintain the native, active conformation; however, the folding pathway of IFNA2 has not been elucidated. If the addition of AP interferes with stabilizing interhelical contacts, one of two things may occur. First, the helices may unfold, leading to a partially unfolded species that is aggregation-prone. Second, the helices may move within the three-dimensional structure of the protein in order to make new contacts. This movement is described as domain swapping, and has been observed before. Cyt c, for example, has been shown to polymerize, and the polymerization pathway involves a swapping of the C-terminal helix in one monomer with the corresponding region of another Cyt c molecule [97]. However, the C-terminal helix is not the same region which unfolds and leads to aggregation in the presence of APs, but it neighbors the 70s-80s loop and may be influenced by its unfolding.

This phenomenon is also seen in disease. Amyloidogenic diseases and disorders, such as Alzheimer’s disease and Parkinson’s disease, are characterized by the aggregation and deposition of proteins within cells. The current paradigm for the amyloidogenic pathway begins with a partially unfolded or misfolded protein with exposed hydrophobic residues acting as nuclei for aggregation. These small, native-like molecules can then assemble into filaments, which form fibrils and finally plaque depositions. One of the suggested modes of fibril formation is through domain swapping (see Chapter V Discussion and [151]).

It is possible that the conclusions regarding AP-induced aggregation can assist in better understanding of protein function as well as disease. In an extensive review, David
Eisenberg’s group summarized a number of proteins that have been documented to participate in domain swapping, and suggested this assembly mechanism could be favored by nature to establish stable oligomeric species [152]. These examples included proteins from bacteria [153], yeast [14], invertebrates [154], and mammals [16]. Additionally, domain swapping has been suggested as a mechanism by which amyloid aggregates are formed, which has been well documented [142-144, 155]. If APs influence disparate proteins in the same way, and AP-induced aggregation proceeds via domain swapping, then well-understood, well-behaved, and easily obtainable proteins (such as Cyt c, for example) could be extensively studied to identify exact aggregation mechanisms that may apply to multiple systems.

**Biotechnology Relevance**

Protein formulation is often approached from an experiential perspective; i.e., if it works in one system, it will be applied to other systems. This tactic applies for salt concentration, stabilizing excipients, and AP inclusion. However, as has been previously discussed (Chapters II, IV, and V), these formulation compounds may have significant deleterious effects on protein stability and aggregation, which has been often ignored. In these cases, empirical analysis is beneficial in order to determine the best formulation design for the protein in question.

The results presented here describe AP influences on a variety of proteins, and demonstrate that APs perturb proteins in the same way (i.e., CR > PH > BA > PE > CB). Less extensive studies on rhGH [42], IL-R1 [47], and monoclonal antibody [48]
aggregation in the presence of APs demonstrated the same pattern, CR > PH > BA. These data suggest that AP effects are universal, or that protein identity is irrelevant. If this interpretation is correct, empirical evidence must be collected to confirm the pattern of AP-induced aggregation, as has been demonstrated in this work, and then historical formulation design can again be pursued.

The results presented in this thesis suggest that screening of APs or a combination of APs would be useful in formulation design. Also, these data demonstrate that CB may be the best AP for therapeutic use, due to its minimal ability to induce aggregation for both Cyt c and IFNA2. CR and PH should only be used with caution, because even at pharmaceutically relevant concentrations, these preservatives induced protein aggregation the most severely.

**Future Directions**

Future work can include understanding the aggregation mechanism induced by APs as well as protection strategies against these aggregation events.

An aspect of formulation that is often overlooked in terms of preservative effects on proteins are innate solution characteristics such as hydrophobicity, viscosity, and dielectric constants of APs (see Appendix A). Each of these parameters can influence protein stability and aggregation. A hydrophobic solvent would diminish the necessity to bury hydrophobic residues within the interior of the protein and could favor unfolded protein species. Preservatives are alcoholic in nature, with distinctly hydrophobic regions. How do APs affect the microenvironment hydrophobicity surrounding a protein surface?
Specifically, does the hydrophobic nature of APs cause protein unfolding? Nonpolar solvents may also promote aggregation by populating unfolded, aggregation-prone species that will associate due to their exposed hydrophobic patches. On the other hand, organic solvents have been shown to increase the functionality of enzymes in comparison to aqueous environments. The dielectric constant is related to hydrophobicity, in that it refers to how well a particular medium stores charge. The more hydrophilic a solvent, the larger capacity it has for polar molecules, and its dielectric constant will be very high. Water, for example, has a dielectric constant of 80.4 at 20°C compared to 33.6 at 25°C for methanol [156], meaning methanol has less potential for hydrogen bonding and electrostatic stabilization than water. Conversely, a hydrophobic solvent will be less capable of storing charge, evidenced by lower dielectric constant values. Ethanol, which has a higher octanol-water partition coefficient than methanol [157], also has a lower dielectric constant of 24.3 at 25°C [156]. In terms of viscosity, the more viscous a solution, the lower the degree of freedom a molecule has. Viscosity could influence possible aggregation events by limiting the movement of individual molecules through the solvent, lowering the collision frequency. An understanding of how these basic solution parameters of APs contribute to protein unfolding and aggregation would be of enormous benefit.

Nearly all available literature demonstrating AP effects on proteins, including Cyt c and IFNA2, has been done on α-helical systems, with the only exceptions being a monoclonal antibody and IL-1R. It would be useful to examine the influence of APs on proteins with varying secondary and tertiary structural elements. For example, monitoring AP effects on a random coil or natively disordered protein could confirm or deny protein
identity as a critical factor in AP-induced protein aggregation. Natively disordered protein aggregation has been implicated in amyloidoses [152], so perhaps AP-induced aggregation would broaden mechanistic knowledge of this relationship.

The interhelical connections that are disrupted in IFNA2 with BA addition must be explored further. Are these residues critical in the folding or unfolding pathway of IFNA2? Stopped-flow NMR studies in the presence of a denaturant such as urea would clarify the roles of these contacts. Without an insight into the unfolding pathway of IFNA2, it is not possible to quantitatively determine the importance of these amino acid contacts in BA-induced unfolding and aggregation.

In an effort to inhibit aggregation induced by APs, mutational studies on the pharmaceutical protein IFNA2 could be conducted. With the model protein Cyt c, reduction of the heme was sufficient to increase the stability of the protein as a whole, and inhibited the formation of the aggregation-prone species. Most other protein systems do not have a convenient means by which to confer stability, and thus more research in this area is necessary. As was discussed in Chapter V, pegylation of IFNA2 resulted in increased drug half-life but reduced overall activity in comparison to unmodified protein. In the case of chymotrypsin, pegylation inhibited aggregation [147]; however, it is unknown whether pegylation would provide protection against AP-induced aggregation in all cases. Additionally, introducing a mutation into IFNA2 may grievously affect protein efficacy while inhibiting protein aggregation. As such, any mutational study must be undertaken with activity assays done in parallel.
REFERENCES


APPENDIX A

INVESTIGATION OF SOLUTION PARAMETERS ON CYTOCHROME C
AGGREGATION IN THE PRESENCE OF ANTIMICROBIAL
PRESERVATIVES

As discussed in Chapters I-V, antimicrobial preservatives induce protein instability and aggregation. However, the mechanism of interaction between APs and proteins has not been deduced. Additionally, many preservatives used in pharmaceutical formulations are alcoholic in nature and as such these compounds can be used to mimic AP effects. In order to probe the specific influence that APs impart on proteins, we used simple alcohols to monitor the effects of solution parameters on protein stability and aggregation. Solution characteristics such as viscosity and hydrophobicity are readily available for simple alcohols whereas they are limited for APs.

Simple Alcohols to Model Antimicrobial Preservatives

In order to monitor the effects of simple alcohols on protein stability and aggregation, three straight chain alcohols (methanol, ethanol, propanol) and one branched chain alcohol (t-butanol) were used. In all cases, an increase in alcohol concentration led to Cyt c aggregation and a corresponding decrease in $T_m^{Agg}$. Figure A.1 shows this phenomenon with increasing concentrations of methanol. At 0% v/v alcohol concentration, the $T_m^{Agg}$ of Cyt c was 83.3 °C. In the presence of 30% v/v Methanol, the $T_m^{Agg}$ decreased to 56.4 °C, indicating that methanol destabilized Cyt c and induced aggregation at a lower temperature compared to control. These results are representative of all four alcohols used. Additionally, the calculated slopes of denaturation for each of the alcohols used in this study demonstrated an increased ability to cause Cyt c aggregation as a function of alcohol concentration (Figure A.2) In the case of methanol, the slope was approximately 9°C/%/v/v. This value indicates that for every increase in volume percent of methanol, the $T_m^{Agg}$ of Cyt c will decrease by 9°C. $T_m^{Agg}$ variation in the presence of ethanol, propanol, and t-butanol resulted in slopes of 11.4°C/%v/v, 19.9°C/%v/v, and 13.6°C/%v/v, respectively. With the exception of t-butanol, which is a branched alcohol, the slope of denaturation obtained for the other simple alcohols increased with increasing chain length.

In our previous work with APs, we demonstrated that aggregation in Cyt c was initiated by the unfolding of the red loop. This loop has been shown to unfold first and to have the lowest stability in comparison to other regions of the protein. It also contains a critical methionine residue at position 80, which participates in a charge-transfer reaction with the iron heme. To test whether simple alcohols induced aggregation in a similar way as APs, the $T_m^{Agg}$ were compared to the $T_m^{695}$ (Figure A.3). The resultant relationship showed a linear correlation, indicating that the unfolding of the red loop may be the critical event in simple alcohol-induced aggregation. In order to inhibit the unfolding of
this loop and subsequent aggregation of Cyt c in the presence of APs, we reduced the iron heme. This approach successfully decreased aggregation of Cyt c induced by APs, quantitatively shown by an increase in $T_m^{Agg}$ compared to the oxidized form. In the case of simple alcohols, this relationship between oxidation state of the heme and decreased aggregation was also true (Figure A.4). For example, the reduction of Cyt c increased the $T_m^{Agg}$ from 62.5°C to 80.6°C in the presence of 20% v/v ethanol.

**FIGURE A.1** Increase in Cyt c optical density at 695 nm as a function of methanol concentration observed during thermal scanning experiments.

**FIGURE A.2** Variation of $T_m^{Agg}$ as a function of methanol concentration.
FIGURE A.3 Correlation between $T_m^{695}$, which shows the unfolding of the red loop, and $T_m^{Agg}$, the aggregation temperature.

FIGURE A.4 Change in thermal scanning trace between oxidized and reduced Cyt c measured at 695 nm. The blue trace represents oxidized Cyt c, and the red shows reduced Cyt c.

Various solution parameters for each alcohol were plotted against the individual calculated slopes of denaturation (Figure A.5a) (values obtained from [157]). The relationship between each alcohol’s ability to induce Cyt c aggregation and hydrophobicity demonstrated a positive correlation, which is intuitive. With an increase in hydrophobicity of the solvent, protein unfolding becomes more energetically favorable due to the fact that hydrophobic residues no longer need to be buried in the protein
interior. However, when the same slopes were plotted against viscosity, an interesting result was observed (Figure A.5b). As alcohol chain length increased, the resultant viscosity and slope of denaturation also increased. However, an increase in viscosity should decrease the propensity for aggregation in two ways. First, if aggregation is diffusion-limited, a more viscous solution will decrease the rate of diffusion, lowering the collision frequency. Second, the viscosity of solution potentially could impact the breathing motions of the protein at equilibrium, restricting the population of aggregation-prone species. t-Butanol did not follow the same trend with viscosity as the other simple alcohols, indicating that a branched chain alcohol may have other influencing factors that lead to aggregation.

Based on these data, a few conclusions can be drawn. Simple alcohols influence Cyt c in the same way as APs, in that there was a concentration-dependent decrease in $T_{m}^{\text{Agg}}$ and aggregation propensity is dependent upon oxidation state of the heme. $T_{m}^{\text{Agg}}$ and $T_{m}^{695}$ demonstrated a linear correlation in both cases, suggesting that the same partially unfolded species initiates aggregation induced by APs as well as simple alcohols. Additionally, simple alcohols demonstrated a positive correlation between inducing Cyt c aggregation and hydrophobicity. However, the relationship between viscosity and aggregation is the opposite of what was expected. This result indicates that solution parameters are multidimensional and difficult to tease apart, at least at the concentrations used in these experiments. Due to the high concentrations of alcohol that were utilized here, it is clear that simple alcohols are useful in monitoring bulk effects on protein stability and aggregation. These bulk influences may not accurately relate to microenvironment changes, which can be studied using extrinsic fluorescent dyes.

**FIGURE A.5** Calculated slopes of denaturation plotted as a function of known solution parameters. A) Slopes as a function of logP, a measure of hydrophobicity. B) Slopes as a function of viscosity.
Using Fluorescent Dyes to Monitor Microenvironment Solution Parameters

Numerous dyes have been detailed in the literature, falling into both covalent and non-covalent dyes. These dyes have been introduced to characterize solvent parameters [158, 159], measure plasma viscosity [160], monitor protein aggregation [161], and protein structural changes [159]. We chose to use non-covalent fluorescent dyes to monitor and characterize the influence of antimicrobial preservatives (APs) on the microenvironment of Cyt c. These dyes include 1-anilinonaphthalene-8-sulfonate (ANS), nile red (NR) and 9-(dicyanovinyl)-julolidine (DVCJ).

Based on previous work, we determined the order of AP influence on protein stability as a function of concentration and temperature: CR > PH > BA > PE > CB, where CR caused the highest decrease in protein stability while CB resulted in the lowest. However, in order to translate this information to formulation design, it is necessary to correlate our conclusions with specific solution parameters that can serve as predictive conditions for stability. Solution parameters such as hydrophobicity values, dielectric constants, and viscosity are not readily available for antimicrobial preservatives in general. Thus, we chose to monitor our model system using extrinsic dyes in order to define any possible correlations between our previous conclusions about APs and their solution parameters.

Each of these dyes has been well-documented in the literature. ANS is a naphthalene derivative that fluoresces in organic or apolar environments, but exhibits little to no fluorescence in aqueous solutions [162]. Additionally, ANS and its related derivatives are sensitive to viscosity and temperature. As the dielectric constant of the environment decreases, an increase in quantum yield is observed as well as a blue shift in the emission maximum [163].

However, it has been suggested that protein signal may interfere with ANS fluorescence due to its near-UV excitation [164]. Because of this, we chose to include NR as a secondary polarity-sensitive dye. NR fluorescence is induced in apolar environments and undergoes a blue shift as solvent polarity is decreased [165]. Neither ANS nor NR showed significant fluorescence changes in comparison to controls during titration experiments with CR or BA, most likely due to the very low concentrations of AP added (0.8% v/v CR, 3% v/v BA) (data not shown).

Molecular rotors have been shown in the literature to be accurate predictors of microenvironment solution characteristics such as viscosity. These fluorescent molecules exhibit FRET-like behavior, with a donor region, a linker, and an acceptor region of the molecule. In the excited state, these dyes undergo an intramolecular rotation that corresponds directly to the sensitivity of the dye toward its environment [166]. The rotor DCVJ demonstrates increased fluorescence emission intensity as solution viscosity is increased due to its high sensitivity to rigidity [167]. A slight blue shift can be observed in the presence of organic solvents. In the presence of high concentrations of benzyl alcohol (3% v/v), the fluorescence emission intensity of DCVJ was not different from control. However, at 0.8% v/v m-cresol, the emission intensity dramatically increased. As a result of these preliminary results, DCVJ fluorescence intensity was observed as a function of m-cresol concentration (Figure A.7). We observed a CR-concentration-dependent increase in DCVJ fluorescence emission intensity, indicating an increased
rigidity of the environment with increasing concentrations of CR. This implies that even at low % v/v concentrations, CR significantly influences the microenvironment viscosity surrounding Cyt c. However, the increase in fluorescence intensity demonstrates an increase in microenvironment viscosity, which should theoretically inhibit aggregation. In our previous work, we showed that CR severely affects Cyt c stability and aggregation in comparison to all other APs studied. Taken together, it is possible that CR influences Cyt c aggregation due to a manipulation of the microenvironment viscosity immediately surrounding the protein surface.

FIGURE A.6 Basic structure of a molecular rotor. “D” represents the donor region, “A” shows the acceptor region. These are linked by the hinge region, pictured here as a blue square [166].

FIGURE A.7 DCVJ fluorescence (5μM) measured at the emission wavelength 493nm as a function of m-cresol (CR) concentration in the presence of 300μM Cyt c.
Dialanine as a Model for the Peptide Bond

We have demonstrated in our lab that APs influence different protein systems in the same manner; i.e., CR causes the most severe destabilization and aggregation, whereas CB resulted in the least. In order to explain this phenomenon, we sought to identify the commonality between multiple protein systems. One universal characteristic between all protein systems is the peptide bond. Literature evidence suggests that the peptide bond may be perturbed in the presence of alcoholic compounds including benzyl alcohol [133]. Our aim was to determine if the peptide bond acted as a binding site for alcoholic preservatives in a simplified system of dialanine. Previously published data demonstrated that benzyl alcohol influenced the isomerization and stability of polyproline by forming hydrogen bonds between the hydroxyl group of the alcohol and the carboxyl group of the peptide, which was detected by FT-IR [133]. Following a modified protocol, we incubated dialanine in the presence and absence of benzyl alcohol and monitored changes in the Amide I and II regions (Figure A.8). We were unable to resolve the alcohol hydroxyl band. Despite numerous experimental set-ups, we were unsuccessful in detecting any reproducible and concrete interactions between the alcohol hydroxyl group and the peptide carboxyl group. It is highly likely that our experimental design may explain this. The previously published data operated in the solid state with samples being dried directly on a window of the detector while under vacuum. Our experiments were in the liquid state without vacuum. We also attempted to monitor any possible changes in dialanine cross-peak volume and position using 1D NMR. No changes were observed in our spectra (data not shown). Based on these results, we do not see significant binding of benzyl alcohol to a simple peptide like dialanine. This data is in agreement with previously published work using ITC ([46] and Appendix B); if there had been a specific binding site (i.e., peptide bond) for alcoholic compounds on a protein surface, there would have been a large change in signal as a function of alcohol concentration.

Formulation design is approached from a historical as well as empirical perspective. In this work, we attempted to discern the solution parameter(s) that is responsible for compromising protein stability and aggregation propensity in an effort to accelerate formulation design. Our results indicate that hydrophobicity plays a role in aggregation on a bulk level, but this result is to be expected. Microenvironment observations demonstrate that viscosity of excipients may be critical as well. From a mechanistic perspective, it is unlikely that APs (and other alcohols) preferentially bind to the peptide bond of a protein. What is clear is that solution characteristics are multidimensional and their influence on protein stability and aggregation is highly dynamic.
FIGURE A.8 Representative traces of FT-IR of dialanine in the absence (blue, purple) and presence (green, yellow) of 0.8% v/v CR.
FIGURE B.1 Isothermal calorimetry of Cyt c and m-cresol binding. Top panel: Isothermal calorimetry titration of m-Cresol (0.8%) into Cyt c (2mM). Bottom panel: Fit of data using a sequential binding site model demonstrates weak binding with multiple sites, as has been seen previously [46].
APPENDIX C

CHLOROBUTANOL EFFECTS ON INTERFERON ALPHA-2A

FIGURE C.1 Isothermal incubation of 10μM IFNA2 at 50°C in the presence of 0.3% v/v AP (modified from Figure 4.2). Note the minor change in optical density induced by CB.

TABLE C.1 Parameters obtained from fitting denaturant melt data to a 2-state unfolding model. The CB data represents a single experiment; the reported error is that of the fit.

<table>
<thead>
<tr>
<th></th>
<th>ΔG [kcal/mol]</th>
<th>m</th>
<th>C_m</th>
</tr>
</thead>
<tbody>
<tr>
<td>0% AP</td>
<td>10.2±0.6</td>
<td>-2.6±0.2</td>
<td>3.9</td>
</tr>
<tr>
<td>0.9% BA</td>
<td>8.4±0.4</td>
<td>-2.1±0.1</td>
<td>4.3</td>
</tr>
<tr>
<td>1% CB</td>
<td>10.8±2</td>
<td>-2.4±0.5</td>
<td>4.5</td>
</tr>
</tbody>
</table>

IFNA2 in the presence of 1% CB did not show changes compared to 0% CB in thermal scanning experiments. A denaturant melt in the presence of 1% CB returned 2-state unfolding fit parameters that are within error of 0% AP data (Table C.1), which suggests that CB does not affect IFNA2 in the same way as other APs. Due to the similarity in ΔG values between 0% AP and 1% CB, and the fact that thermal scanning experiments yielded no aggregation in the presence of CB, it is possible that while CB may not be stabilizing IFNA2, its particular effects are so minimal as to be negligible by the methods used here.
FIGURE C.2 IFNA2 unfolding in the presence of 1% CB using the denaturant GdmHCl as monitored using 280 nm excitation and 328 nm emission fluorescence. Data have been adjusted to show complete IFNA2 unfolding at 5M GdmHCl, based on 0% v/v denaturant melts.
APPENDIX D

ACRYLAMIDE QUENCHING OF INTERFERON ALPHA-2A FLUORESCENCE IN THE PRESENCE OF BENZYL ALCOHOL

IFNA2 contains two tryptophan residues located at positions 76 and 140. These Trp residues are located near each other in the three-dimensional structure but are within different environments (Figure C.1). W76 is between Helix B’ and Helix C on a small amino acid stretch that is unstructured, while W140 is within Helix E. Both of these residues have low solvent accessibility as determined by Accelrys by Discovery Studio.

FIGURE D.1 Locations of tryptophan residues in the three-dimensional structure of IFNA2. Purple highlights show W76 between Helix B’ and Helix C and W140 in Helix E.

In order to monitor benzyl alcohol’s effects on intrinsic tryptophan fluorescence, acrylamide quenching in the presence of varying BA concentrations was conducted. Excitation wavelength was set at 295 nm to monitor Trp fluorescence and an emission wavelength of 332 nm, to avoid any interference arising from BA. Using the fluorescent intensities in the presence and absence of quencher, Stern-Volmer constants can be calculated at each concentration of acrylamide and BA (Figure D.2, Table D.1) as shown in Equation D.1:

$$\frac{I_0}{I} = 1 + K_{SV} [Q]$$

Equation D.1
where \( I_0 \) is the fluorescence intensity in the absence of quencher, \( I \) is the fluorescence intensity in the presence of quencher, and \([Q]\) is the quencher concentration. Stern-Volmer constants (\(K_{SV}\)) can be obtained as the slope of the line \([Q] \) versus \(I_0/I\) [168]. The data presented here represents a single experimental set, using 3mm IFNA2 in 10 mM ammonium acetate, 120 mM sodium chloride, pH 5.

\(K_{SV}\) magnitudes give an indication of the Trp environment. Larger values suggest an easily accessible Trp residue, which would require a lower concentration of acrylamide to quench [168]. Taking 0% BA as a baseline, the addition of 0.3% BA increases the accessibility of one or both of the Trp residues. In 0.6% BA, the Stern-Volmer constant decreases, implying a burial of the Trp residues under these conditions. With 0.9% BA, the Stern-Volmer constant is much higher than either 0.3% or 0.6%, demonstrating a significant change in the protective environment of one or both of the Trp residues. Additionally, lower concentrations of acrylamide were needed to achieve observable quenching in the case of 0.9% BA; at 0.1M acrylamide, the measured \(I_0/I\) value was higher than in all other cases.

Based on these data, it is apparent that BA changes the accessibility of IFNA2 Trp residues to solvent, as indicated by an increase in \(K_{SV}\) with increasing BA concentration.

**FIGURE D.2** Stern-Volmer plot of acrylamide quenching of IFNA2 in the presence of BA.
TABLE D.1 Stern-Volmer constants calculated from the slopes in Figure D.2. Error represents error of the linear fit.

<table>
<thead>
<tr>
<th>[BA]</th>
<th>$K_{SV}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00%</td>
<td>0.5819±0.02</td>
</tr>
<tr>
<td>0.30%</td>
<td>0.7714±0.02</td>
</tr>
<tr>
<td>0.60%</td>
<td>0.6847±0.03</td>
</tr>
<tr>
<td>0.90%</td>
<td>0.8346±0.05</td>
</tr>
</tbody>
</table>