The Development of Novel Excipients for the Stabilization of Proteins Against Aggregation

by

Curtiss P. Schneider
Colorado School of Mines, 2004

Submitted to the Department of Chemical Engineering in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2011

Author .....
Department of Chemical Engineering
December 13, 2010

Certified by .....
Bernhardt L. Trout
Professor of Chemical Engineering
Thesis Supervisor

Accepted by .....
William M. Deen
Professor of Chemical Engineering
Chairman, Committee for Graduate Students

The Development of Novel Excipients for the Stabilization of Proteins Against Aggregation

by

Curtiss P. Schneider

Submitted to the Department of Chemical Engineering on December 13, 2010, in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

Abstract

Although protein based therapeutics is the fastest growing sector of the pharmaceutical industry, production costs remain incredibly high and rapid commercialization of new protein drug candidates are not being fully realized due to the presence of many barriers, namely the physical and chemical instabilities of proteins. Of these degradation pathways, protein aggregation is arguably the most common and troubling manifestation of protein instability, occurring in almost all phases of development. Protein aggregates are usually nonnative in structure, may exhibit reduced biological activity, and can remain soluble and/or precipitate from solution. In addition to reducing efficacy, if administered to a patient, aggregates can cause adverse reactions, such as immune response, sensitization, or even anaphylactic shock. Therefore, if even a small amount of aggregates form during formulation or storage, a product can be rendered unacceptable. Moreover, for the practical application of traditional and novel drug delivery techniques, protein based therapeutics must be formulated at relatively high concentrations and must remain stable for extended periods of time. The structural differences among various proteins are so significant, that the application of a universal stabilization strategy has not yet been successful, though the effects of common excipients are generally universal. The current approach toward stabilizing protein drugs against aggregation is by trial-and-error testing of different combinations of cosolutes (*e.g.* salts, sugars, surfactants, amino acids, *etc.*) using empirically derived heuristics. While ubiquitously used, this approach is inefficient and does not always enable the discovery of stable protein solution formulations. In response to this major problem, we have developed and tested a new class of excipients that has the potential for wide spread application as a universal stabilizer of protein therapeutics. When compared to other commonly used excipients, our novel excipients offer more than an order of magnitude improvement at suppressing the aggregation of a model protein. As a result, if used in formulations, the shelf life of a protein drug, at room or refrigerated temperatures, may be extended from a few weeks to several months or years. Furthermore, these excipients will likely be useful during production and purification for improving yield and lowering downstream purification costs.

Thesis Supervisor: Bernhardt L. Trout
Title: Professor of Chemical Engineering

Acknowledgments

The work I accomplished during the time I spent at MIT, not to mention the success that led me to this institution, would never have happened without the devoted efforts of so many important people in my life.

I must acknowledge my parents first because none of my achievements in life would be possible without them. They have provided me with an abundance of support and guidance throughout my life, in addition to teaching me the value of hard work and academic achievement. Both of my parents strongly encouraged me to follow my dreams and did everything they could to make sure that those dreams came true. My mother has devoted her entire life to doing whatever she can, sometimes at great lengths, for me and my siblings. My father has always put our family first, working tirelessly to provide for us what most parents could not. I am truly blessed to have parents who were determined to make sure I started out in life on the right foot.

Of course, I would not be at this point in my education and my experience at MIT without the supervision and support from my thesis advisor, Prof. Bernhardt Trout. I should acknowledge that in the beginning, I was somewhat of a risk considering that I was his first experimental student. But he gave me a chance and was quite patient and helpful as I set up the lab and slowly got started with a challenging project. I am thankful for his guidance throughout my time here and I am proud that we achieved the goals we set out to accomplish at the beginning of the project. I should note that reaching those goals was greatly aided by the guidance of my thesis committee. I am indebted to Prof. K. Dane Wittrup, Prof. Jonathan King, and Prof. Alexander Klibanov for their expert advice and motivation throughout the course of the project. I am further indebted to the John C. Haas Fellowship, the Chemical Engineering Department, the National Institutes of Health and the Singapore-MIT Alliance for financial support during my studies.

I would like to thank the Trout Group as a whole for making my time at MIT fun and adventurous. Specifically, I would like to thank Dr. Vladimir Voynov for being such a great and supportive officemate, Dr. Gregg Beckham for being such a good group leader, and Dr. Jie Chen for making group outings fun and interesting. Most importantly, I must acknowledge Diwakar Shukla, who is the computational counterpart to my experimental work. Without the molecular

perspective insight he was able to elucidate, the theoretical work he was able to accomplish, and the many hours of discussion and planning he devoted toward this project, we would never have been able to accomplish the main goals of the project.

I would like to acknowledge two professors from the Colorado School of Mines, my undergraduate institution, because without them, I would never have attended MIT. I would like to thank my undergraduate research advisor, Prof. David W. M. Marr, who taught me what it means to do academic research and Prof. E. Dendy Sloan, who not only taught me what it means to be a chemical engineer but was also the person who advised me to attend graduate school. I would also like to acknowledge all of the devoted teachers and professors from throughout my educational experience who helped to shape my young mind, introduced to me the enjoyment of learning, and prepared me for a successful career.

And last but not least, I would like to give out a very special acknowledgement to my wife, Rachel (we were recently wedded on October 10, 2010), who has been incredibly supportive and understanding as I worked to finish my Ph.D, not to mention making life worth living.

Contents

List of Figures

List of Tables

Chapter 1

1. Introduction

1.1. Background on Protein Aggregation

Protein based therapeutics (*e.g.* monoclonal antibodies) have become indispensable in the treatment of a wide variety of diseases and disorders. In comparison with small chemical drugs, protein pharmaceuticals have high specificity and activity at relatively low concentrations, in addition to eliciting minimal side effects [1, 2]. Moreover, recent advancements in recombinant DNA technology have allowed for the development of potent drug candidates and new high-throughput methodologies have allowed such products to be produced on a large scale [3]. As a result, protein drug development has become the fastest growing sector of the pharmaceutical industry, leading to over one hundred protein based drugs brought to market in the last decade alone, with hundreds more in some phase of development [4].

Regardless of the rapid growth and increasing demand, production costs continue to remain high and the commercialization of many protein drug candidates have not been fully realized due to the presence of many barriers, namely the physical and chemical instabilities of proteins. Of these degradation pathways, protein aggregation (the assembly of protein molecules into multi-meric states, (see Figure 1-1)) is arguably the most common and troubling manifestation of protein instability, occurring in almost all phases of development [5]. Protein aggregates are usually nonnative in structure, may exhibit reduced biological activity, and can remain soluble and/or precipitate from solution. In addition to reducing efficacy, if administered to a patient, aggregates can cause adverse reactions, such as an immune response, sensitization, or even anaphylactic shock [6]. Therefore, if even a small amount of aggregates form during formulation or storage, a product can be rendered unacceptable (which can easily derail the product development process for a biotech company). To make matters worse, for the practical application of traditional and novel drug delivery techniques, protein based therapeutics must be formulated at relatively high concentrations and must remain stable for extended periods of time [7].

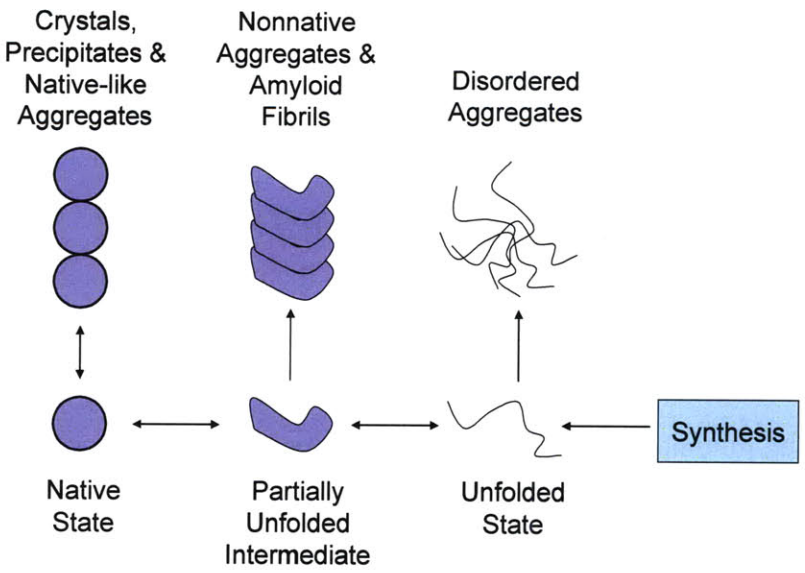

Figure 1-1: An overview of protein aggregation pathways. Proteins in the native state can form native like oligomers, fibrils, and crystals. Partially unfolded and completely unfolded proteins can form nonnative aggregates and amyloid fibrils. All forms of the associated state are undesirable and potentially harmful.

1.2. Stabilization Methods

As mentioned above, the administration of therapeutic proteins is restricted to injections, either intravenously or subcutaneously, because any protein taken in pill form would be digested before it could be absorbed into the blood stream (that is unless novel and revolutionizing delivery methods are introduced). For injections, both doctors and patients desire doses of therapeutic proteins to be dissolved in a minimal volume of aqueous solution because the intravenous injection of larger volumes requires a longer time in the clinic and large-volume subcutaneous injections often cause patient discomfort at the site of injection. To minimize the injection volume of a biopharmaceutical solution, the protein must be dissolved at a relatively high concentration in an aqueous media. However, as mentioned above, aqueous solutions are intrinsically unstable due to aggregation, especially at high concentrations.

One of the simplest methods for preventing or slowing aggregation (besides preparing dilute solutions) is to freeze the solution or store it at refrigerated conditions. However, this is not always effective and often, is not a viable option. Therefore, the most prevalent method for combating protein aggregation, in addition to tuning buffer conditions, is to add to the solution a low molecular weight additive (commonly referred to as a cosolute, cosolvent or excipient), such as salts, sugars, polyols, *etc.*, which have the propensity to deter aggregation [2, 8]. Unfortunately, due to the diversity in protein chemistry and structure, additives that work well

for a particular protein may not work universally. Industrial researchers typically select these excipients via a trial-and-error approach using empirically derived heuristics. Since there is no reliable method to predict the optimal additive(s) or the best additive concentration for the stabilization of a given protein, it is typical for a very large number of formulation experiments to be performed, using different types and concentrations of salts, sugars, surfactants, polyols, and other inert molecules. In addition, the interaction between common solution additives and proteins is typically weak, thus requiring a high concentration to be effective [9].

It is often the case that this approach fails to yield a formulation with acceptable stability, and in this case, the protein drug must be lyophilized and reconstituted later by the doctor or patient. Aqueous formulations, however, are greatly preferred over lyophilized ones because (1) reconstitution adds one more step to the delivery of the drug and this step can be cumbersome and increases the potential for errors, (2) lyophilization is not without major challenges (*e.g.* precipitation of buffers during freezing can lead to detrimental pH changes, proteins can undergo conformational changes in the lyophilized state leading to degradation, slight imperfections in temperature and pressure control can lead to undesirable product quality, *etc.*) and (3) cutting-edge, highly advantageous delivery methods (*e.g.* delivery via slow-release implantable devices) cannot be used with lyophilized drugs. Furthermore, even if the drug can be successfully formulated in aqueous solution, the development of stable but higher concentration formulations is often a goal of drug development. The higher the concentration of the biopharmaceutical, the better the delivery methods can be. For example, if very high concentrations (100 mg/ml+) were possible, these drugs could be packaged in implantable devices for controlled release, a completely new and highly desirable method for delivery of biopharmaceuticals. At this point, however, the most significant new family of biopharmaceuticals, therapeutic antibodies, typically must be delivered intravenously. Subcutaneous injection, which requires much higher concentrations than intravenous delivery, is much less cumbersome and is highly desired over intravenous injections.

Therefore, due to the serious issues and complications associated with protein aggregation and the ineffectiveness of current stabilization methodologies, there is great interest in developing new solution additives which are very effective at deterring aggregation and are also potent at low concentrations. However, the current understanding of the mechanisms by

which commonly used additives (*e.g.* arginine hydrochloride) deter aggregation is limited and elucidating these mechanisms of action will be of great use as well [10].

1.3. Thesis Objectives

The main objective of this thesis was to ***develop and test a novel class of excipients with the aim of dramatically improving the stabilization of protein therapeutics against aggregation***. In addition to this, we also aimed to ***elucidate the mechanism by which arginine (a model excipient) and our novel excipients inhibit aggregation to further improve excipient design and biopharmaceutical stabilization***.

1.4. Organization of Thesis

The remaining six chapters in this thesis gives an overview of how we began with a simple theory developed in our group, which describes a novel class of aggregation suppressing additives called "neutral crowder", and ended with a variety of synthesized compounds that behave in such a manner and improve aggregation suppression by more than an order of magnitude. Chapter 2 (Background on Solution Additives) gives a detailed overview of currently used excipients, how additives interact with proteins, how this interaction influences stability, a description of the Neutral Crowder theory, and our initial excipient design features for creating "neutral crowder" compounds. Chapter 3 (Experimental Procedures) describes all procedures used to develop, test, and analyze the novel excipients. Chapter 4 (Elucidating the Arginine Mechanism) describes the aggregation suppression behavior of arginine, simple improvements we discovered, and what we have elucidated about its mechanism of action (which has characteristics similar to a "neutral crowder" compound). Chapter 5 (Novel Excipient Development) details the different types of novel excipients we developed and their performance at suppressing aggregation. Chapter 6 (Conclusions) gives an overview of the research conducted and summarizes the working compounds we developed. And finally, Chapter 7 (Future Work) describes much of the immediate work that will have to be accomplished before the newly developed excipients can be commercialized and utilized by the pharmaceutical industry.

Chapter 2

2. Background on Solution Additives

2.1. Aggregation Pathway Model

The mechanisms by which commonly used additives inhibit aggregation are not quite clear, though in a few cases, qualitative mechanistic models have been developed to describe how a particular additive deters aggregation. For example, during aggregation, it is generally believed that a protein molecule that is initially in the native state (N) reversibly passes through a partially unfolded intermediate (I), which is aggregation prone, before proceeding to irreversibly form aggregates with nonnative characteristics (A_n) [11]:

$$N \rightleftharpoons I \rightarrow \frac{1}{n} A_n . \tag{3.1}$$

This mechanism is possible because a protein in its native state is not static. The secondary structural elements of the protein continually undergo small movements in space, allowing for the formation of partially unfolded species. Furthermore, in this model, there are two types of additives which will deter aggregation, conformational stabilizers and association suppressors.

A conformational stabilizer is classified as any additive which shifts the folding equilibrium from the partially unfolded state toward the native conformation. It has been demonstrated that such additives (*e.g.* sucrose, glucose, polyols, *etc.*) are preferentially excluded from the protein-solvent interface and thus exhibit an unfavorable interaction with the protein. This unfavorable interaction is greater for the partially unfolded species due to a larger solvent exposed surface area and is minimized when the protein adapts a state with minimal surface area, such as its native state [12]. Thus, it is generally believed that such additives deter aggregation by decreasing the number of aggregate prone species. As a consequence, conformational stabilizers are typically added to refolding buffers (refolding being a common technique for recovering and purifying proteins that aggregated or formed inclusion bodies during fermentation) to enhance proper folding and are used during storage to stabilize the native state. However, the aggregated state also exhibits a smaller solvent exposed surface area, thus high concentrations of such additives could actually promote aggregation, often observed as a decrease in solubility [13, 14]. Moreover, during refolding, aggregation is a competing pathway under kinetic control (see Figure 2-1) and to avoid the irreversible loss of protein from

aggregation, refolding is typically done at substantially high dilutions since aggregation is a intermolecular reaction while protein folding is a unimolecular process. To refold proteins at a reasonable concentration, an association suppressing additive, which specifically inhibits association, must be added to the solution [15].

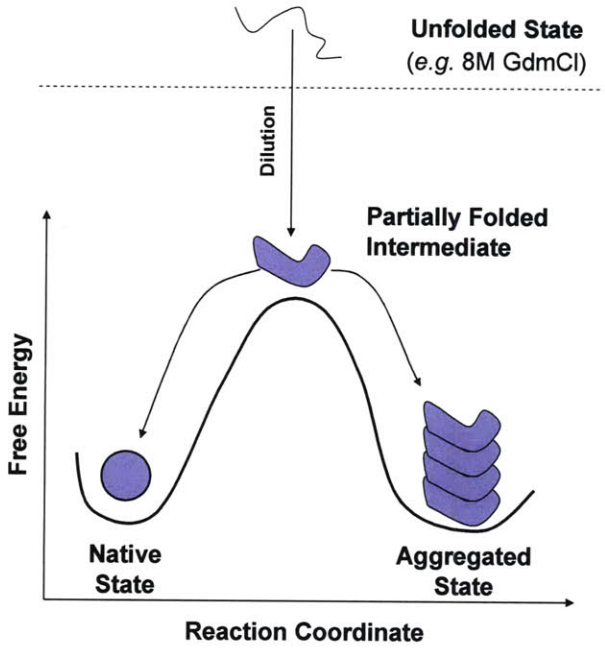

Figure 2-1: Competing pathways during refolding. To reduce the exposure of hydrophobic residues upon removal of the denaturant, the protein can either fold into the native state or aggregate into a multimeric state. Which path the protein chooses will depend on solution conditions.

One such additive is the amino acid arginine [16]. It has been well documented that arginine has very little effect on the folding equilibrium, yet it deters aggregation during refolding and storage [17-19]. Therefore, it is assumed that arginine is an association suppressant and not a conformational stabilizer. Based on the limited data available, researchers have proposed different hypotheses to explain the phenomenon [9, 15, 17, 20]. In addition, researchers have attempted to discover or develop other additives that are similar to arginine, but more effective at hindering aggregation [21-25]. Despite all of this, the mechanism by which arginine functions is still not well understood.

2.2. Commonly Used Excipients

The description above is a somewhat simple and generalized description of how excipients inhibit aggregation. Most aggregation suppressing excipients can be described in such a manner, however, some exhibit unique behavior (*e.g.* polyethylene glycol). Moreover, formulation development is usually a bit more involved, taking into consideration other factors

that influence stability (*e.g.* pH, ionic strength, buffering components, surface adsorption, solubility, other degradation pathways, *etc.*).

Determining the ideal pH is an essential first step in formulating proteins since it has a large influence on colloidal and conformational stability, solubility, and the reaction rates of other degradation pathways. Choosing an appropriate buffering component (*i.e.* one that produces favorable stability or solubility) and determining an ideal overall ionic strength are important as well. Moreover, almost all formulation recipes include a nonionic surfactant (*e.g.* polysorbate) which inhibits surface adsorption and aggregation induced by shaking and shearing. Furthermore, the most common excipients utilized for inhibiting aggregation are usually conformational stabilizers (*e.g.* sucrose, trehalose, glycerol, mannitol, sorbitol, *etc.*).

Regardless of the type of excipient used, the effect folding stabilizers, association suppressors, salts, buffering components, *etc.* have on solubility, conformational stability, and other macromolecular reactions, can be described via preferential interaction theory.

2.3. Additive-Protein Preferential Interactions

Preferential interaction theory is the thermodynamic framework developed to quantify the effect of cosolutes on protein stability. Preferential interactions, though weak in nature, will significantly influence the solubility and stability of a protein in addition to influencing protein-protein interactions [9, 26]. These interactions can be quantified through the use of parameters called preferential interaction coefficients.

2.3.1. Preferential Interaction Coefficients

2.3.1.1 Cosolute-Protein Preferential Interaction Coefficient Γ_{μ_3}

The preferential interaction coefficient of interest, Γ_{μ_3}, is defined by the following partial derivative,

$$\Gamma_{\mu_3} \equiv \left(\frac{\partial m_3}{\partial m_2} \right)_{T,P,\mu_3}, \tag{3.2}$$

which describes how the cosolute molality (m_3) must change to maintain a constant cosolute chemical potential (μ_3) when there is a change in protein concentration (m_2), with the additional constraints of constant temperature (T) and pressure (P) [27]. The subscripts used indicate solution components in Scatchard notation: water (subscript 1), the protein (subscript 2), and the cosolute (subscript 3) [28]. Alternatively, a thermodynamically equivalent definition can be

expressed by the partial derivative relating the chemical potential of the protein to that of the cosolute, at constant temperature, pressure, and protein molal concentration:

$$\Gamma_{\mu_3} \equiv -\left(\frac{\partial \mu_2}{\partial \mu_3}\right)_{T,P,m_2} . \tag{3.3}$$

These partial derivatives indicate that Γ_{μ_3} is both a measure of how the chemical potential of the protein is perturbed by the presence of the cosolute and a measure of how the cosolute concentration must change to maintain a constant chemical potential when protein is added to the solution. The later description has been interpreted using a two-domain model as the difference in the cosolute concentration between the local domain surrounding the protein and the bulk solution [29, 30]. As shown in Figure 2-2, under such interpretations, additives with a positive Γ_{μ_3} are typically described as being preferentially bound to the protein surface due to an increase in the concentration of the cosolute in the local domain and this favorable interaction, as indicated by Eq. (3.3), lowers the chemical potential of the protein. The opposite is true for additives with a negative Γ_{μ_3}, which are typically described as being preferentially excluded from the surface of the protein.

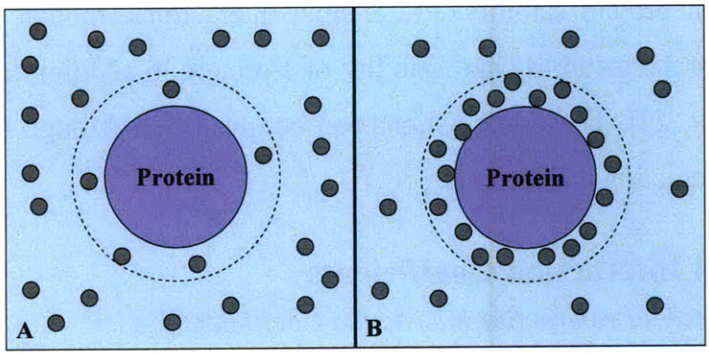

Figure 2-2: A simple diagram depicting the two extremes of preferential interactions. (A) depicts the situation in which the cosolute (grey circle) is preferentially excluded from the surface of the protein ($\Gamma_{\mu_3} < 0$) and (B) depicts the situation in which the cosolute is preferentially bound to the surface of the protein ($\Gamma_{\mu_3} > 0$).

2.3.1.2 Other Preferential Interaction Coefficients

Even though values for Γ_{μ_3} are essential in understanding how cosolute molecules interact with a protein, the parameter, however, cannot be measured directly because the chemical potential of the cosolute cannot be selectively held constant. Therefore, an approximation must be obtained. This can be accomplished by defining other preferential interaction coefficients. The first is the *isoosmolal preferential interaction coefficient*, Γ_{μ_1}, which

is defined by the partial derivative describing the change in cosolute molality (m_3) with respect to protein molality (m_2) at constant temperature, pressure, and water chemical potential (μ_1):

$$\Gamma_{\mu_1} \equiv \left(\frac{\partial m_3}{\partial m_2} \right)_{T,P,\mu_1} . \tag{3.4}$$

The constraint of constant μ_1 can be achieved via vapor pressure equilibrium and thus the parameter defined by Eq. (3.4) can be determined directly by VPO techniques, which will be described later [27].

The second preferential interaction coefficient is the *dialysis equilibrium preferential interaction coefficient*, Γ_{μ_1,μ_3}, which is assumed to be approximately equal to Γ_{μ_3} and is defined by the partial derivative describing the change in cosolute molality (m_3) with respect to protein molality (m_2) at constant T, μ_1, and μ_3:

$$\Gamma_{\mu_1,\mu_3} \equiv \left(\frac{\partial m_3}{\partial m_2} \right)_{T,\mu_1,\mu_3} \cong \Gamma_{\mu_3} . \tag{3.5}$$

The constraints of constant μ_1 and μ_3 while allowing pressure to vary are the conditions of dialysis equilibrium, given that the protein cannot diffuse across the membrane, thus the parameter defined by Eq. (3.5) can be determined directly by dialysis/densimetry techniques [31-34].

2.3.1.3 Dialysis/Densimetry Technique for Determining Γ_{μ_1,μ_3}

For several decades, the only established method for approximating Γ_{μ_3} has been through measuring Γ_{μ_1,μ_3} via the dialysis/densimetry technique. Other techniques for determining the change in cosolute concentration (*e.g.* refractive index) upon dialysis have been utilized, but the procedure is essentially the same. As described elsewhere, the change in mass and volume resulting from allowing a protein solution to reach dialysis equilibrium with the solvent (here solvent refers to the cosolute-water solution) can be attributed to a change in the mass and partial specific volume of the protein [35, 36]. By doing so, Γ_{μ_1,μ_3} can be related to the partial specific volume at infinite dilution of the protein before (ϕ_2°) and after dialysis ($\phi_2'^\circ$), along with the partial specific volume of the cosolute ($\bar{v}_3$), the density of the solvent (ρ_s°), and the molecular weights of the protein (M_2) and cosolute (M_3):

$$\Gamma_{\mu_1,\mu_3} = \frac{M_2}{M_3}\left(\frac{\phi_2^\circ - \phi_2'^\circ}{1/\rho_s^\circ - \bar{v}_3}\right). \tag{3.6}$$

The partial specific volume of the protein is taken at infinite dilution since this value represents the partial specific volume with no protein-protein interactions and can be determined by extrapolating the apparent specific volume data (ϕ), as determined through density measurements of the protein solution (ρ), to zero protein mass concentration (C_2):

$$\phi_2^\circ \equiv \lim_{C_2\to 0}\phi_2 = \lim_{C_2\to 0}\left[\frac{1}{\rho_s^\circ}\left(1-\frac{\rho-\rho_s^\circ}{C_2}\right)\right]. \tag{3.7}$$

The partial specific volume of the cosolute at a given concentration is assumed to be the same as that for a solution with no protein, thus it can be determined from changes in solvent density with respect to cosolute mass fraction (z_3):

$$\bar{v}_3 = \frac{1}{\rho_s^\circ} + (1-z_3)\left(\frac{\partial(1/\rho_s^\circ)}{\partial z_3}\right)_{T,P,z_2=0}. \tag{3.8}$$

2.3.1.4 Thermodynamic Relationship between Osmolality and Γ_{μ_3}

Recently, a new methodology utilizing vapor pressure osmometry has been introduced as a means for approximating Γ_{μ_3} [27, 37]. The advantage of this technique over dialysis/densimetry is that only a relatively small volume of sample is needed for the measurements and numerous Γ_{μ_3} data points covering a wide range of cosolute concentrations can be gathered in a single day. The only major drawback of this technique is that a very high concentration of protein is required to clearly distinguish the change in water activity from the reference solution. A vapor pressure osmometer is essentially a highly accurate dew point hygrometer. It uses the dew point temperature of a small volume of air in equilibrium with an aqueous solution to report the osmolality (*Osm*) of the solution (via the Clausius-Clapeyron equation), which is related to the activity of the water (a_1) in the solution:

$$Osm \equiv -m_1 \ln a_1. \tag{3.9}$$

To use such measurements to determine the preferential interaction coefficient, a relationship must be expressed relating the parameter to the activity of water. This is done by substituting in the relationship between chemical potential and activity into Eq. (3.3) and

substituting in the Gibbs-Duhem relationship for one of the solutes. This produces two alternative equations relating Γ_{μ_3} to the activity of water:

$$\Gamma^{I}_{\mu_3} = \frac{m_1}{m_2}\left(\frac{\partial \ln a_1}{\partial \ln a_3}\right)_{T,P,m_2} + \frac{m_3}{m_2} \tag{3.10}$$

$$\Gamma^{II}_{\mu_3} = \left[\frac{m_3}{m_1\left(\partial \ln a_1 / \partial \ln a_2\right)_{T,P,m_2} + m_2}\right]. \tag{3.11}$$

However, the change in the activity of the cosolute (Eq (3.10)) or the protein (Eq. (3.11)) cannot be measured independently, thus their values must be approximated. One such method is to assume that the change in the activity of either solute in the three component solution is approximately the same as the change in a two component solution containing only that solute and water. By doing so, the change in the activity of the protein and the cosolute can be approximated from osmolality measurements via the Gibbs-Duhem relationship,

$$\partial \ln a_3 \cong \partial \ln a_3^{\circ(2)} = -\frac{m_1}{m_3}\partial \ln a_1^{\circ(2)} \tag{3.12}$$

$$\partial \ln a_2 \cong \partial \ln a_2^{\circ(3)} = -\frac{m_1}{m_2}\partial \ln a_1^{\circ(3)}, \tag{3.13}$$

where the superscript o(i) represents a two component solution lacking component i. By applying the chain rule and Euler's reciprocity to Eqs. (3.10) and (3.11), along with substituting in these approximations, the definition of osmolality (Eq. (3.9)), and a symbol to simplify the expression,

$$\Omega_k = \left(\frac{\partial Osm}{\partial m_k}\right)_{T,P,m_{k'\neq k}}, \tag{3.14}$$

two alternative and equally valid approximations can be expressed relating Γ_{μ_3} to solution osmolality:

$$\Gamma^{I}_{\mu_3} \cong \frac{m_3}{m_2}\left[1 - \frac{\Omega_3}{\Omega_3^{\circ(2)}}\right] \tag{3.15}$$

$$\Gamma^{II}_{\mu_3} \cong \left[\frac{m_3\left(\Omega_2^{\circ(3)} - \Omega_2\right)}{m_3\Omega_3 + m_2\left(\Omega_2^{\circ(3)} - \Omega_2\right)}\right]. \tag{3.16}$$

Moreover, the second approximation can be further simplified by realizing that,

$$\Gamma_{\mu_1} = -\frac{\Omega_2}{\Omega_3}, \tag{3.17}$$

which reduces Eq (3.16) to

$$\Gamma_{\mu_3}^{II} \cong \left[\frac{m_3\left(\Omega_2^{\circ(3)} + \Gamma_{\mu_1}\Omega_3\right)}{m_3\Omega_3 + m_2\left(\Omega_2^{\circ(3)} + \Gamma_{\mu_1}\Omega_3\right)} \right]. \tag{3.18}$$

In the analysis presented here, both Eqs. (3.15) and (3.18) were utilized to compute the preferential interaction coefficient and the results presented are an average of the two computed values.

2.3.1.5 An Exact Relationship Linking Γ_{μ_1,μ_3} and Γ_{μ_3}

It is clear that the two techniques for approximating the preferential interaction coefficient measure different parameters. Therefore, it is desirable to know the relationship between the two parameters so that a comparison of the results from the two techniques can be made. As derived elsewhere, an exact thermodynamic relationship between Γ_{μ_1,μ_3} and Γ_{μ_3} can be expressed utilizing the preferential interaction coefficients Γ_{μ_1} and Γ_{μ_3}, along with the partial molar volume of water ($\overline{V}_1$) and the cosolute ($\overline{V}_3$) [38]:

$$\Gamma_{\mu_1,\mu_3} = \frac{\overline{V}_3\Gamma_{\mu_1}\left(m_2\Gamma_{\mu_3} - m_3\right) - m_1\overline{V}_1\Gamma_{\mu_3}}{\overline{V}_3\left(m_2\Gamma_{\mu_3} - m_3\right) - m_1\overline{V}_1} \tag{3.19}$$

2.3.2. Perturbation of Free Energy

For a solute to have an effect on a reaction or macromolecular process (*e.g.* unfolding, precipitation, *etc.*), a change must occur in the interactions of that component with the entity (*i.e.* the protein) undergoing the reaction. This interaction can be expressed through the change in the chemical potential of the protein when transferred from water to the cosolute system, defined as the standard free energy of transfer ($\Delta\mu_2^{tr}$) and is related to the preferential interaction coefficient as follows [39]:

$$\Delta\mu_2^{tr} = \int_0^{m_3}\left(\frac{\partial\mu_2}{\partial m_3}\right)_{m_2} dm_3 = -\int_0^{m_3}\Gamma_{\mu_3}\left(\frac{\partial\mu_3}{\partial m_3}\right)_{m_2} dm_3 \cong -RT\int_0^{m_3}\left(\frac{\Gamma_{\mu_3}}{m_3}\right)dm_3 \cong -RT\Gamma_{\mu_3} \tag{3.20}$$

Eq. (3.20) reveals that for protein-cosolute systems with positive values of Γ_{μ_3}, the interaction between macromolecule and the cosolute causes the chemical potential of the protein

to decrease as a result of introducing the cosolute to the protein solution. The opposite is true for negative values of Γ_{μ_3} (*i.e.* the interaction causes the chemical potential to increases). If during a macromolecular process (*e.g.* denaturation), the interaction between a macromolecule and a cosolute changes during the course of the reaction (in the given example of denaturation, the macromolecule surface area changes during the course of the reaction, which influences the interaction with cosolute molecules), the free energy of transfer will differ between the reactant and product, which means the difference in the standard Gibbs free energy between the product and the reactant will change as a result of adding the cosolute to the solution. This of course will influence the reaction equilibrium, which can be expressed mathematically by the Wyman Linkage Relation [39].

2.3.3. Wyman Linkage Function: Conformational Stability and Solubility

As described above, Γ_{μ_3} is a measure of how the cosolute perturbs the free energy of the protein and such interactions will influence macromolecular interactions if the interaction differs between the reactant (R) and product (P) states. According to the Wyman linkage relation, this relationship can be expressed by,

$$-\left(\frac{\partial \Delta G^{\circ}}{\partial \mu_3}\right)=\left(\frac{\partial \ln K}{\partial \ln a_3}\right)=\Gamma_{\mu_3}^{P}-\Gamma_{\mu_3}^{R}, \tag{3.21}$$

where ΔG° is the standard Gibbs free energy change for the reaction and K is the equilibrium constant. Thus, additives which have a greater affinity for the product state over the reactant state, will shift the reaction equilibrium toward the product state (the opposite being true for additives which have a lower affinity for the product state) [40]. Two classic macromolecular reactions of importance here are protein folding and protein precipitation. For example, additives which bind more strongly to the unfolded state will tend to unfold proteins. Such additives also tend to increase solubility, thus protein binding additives both denature proteins and dissolve precipitates because the unfolded and dissolved states allow for more binding of the additive.

It should be noted that preferential interaction coefficient measurements only indicate how a cosolute interacts with a protein in its native state. A direct correlation between preferential interaction measurements and solubility can be made though and is absolute because the precipitated and dissolved states only differ by the amount of surface area exposed to the solvent, thus the nature of the interaction with a cosolute does not differ between the two states.

However, the physical properties of the unfolded state are significantly different from the physical properties of the native state and therefore, the nature of the interaction with a cosolute may differ between the two states. For example, the interaction between a cosolute and a protein may be influenced by amino acid specific, binding site specific, electrostatic, hydrogen bonding and/or hydrophobic interactions, which are all dependant on the conformation of the protein. However, this is often not the case since most cosolutes interact with proteins in a nonspecific manner (*e.g.* steric exclusion, surface tension perturbation, *etc.*), thus only the amount of surface area exposed contributes to the difference in such interactions. As a result, a correlation between conformational stability and preferential interaction measurements can only be generalized. Therefore, other thermodynamic techniques which involve thermal unfolding (*e.g.* DSC, CD Spec., *etc.*) are required to confirm how a cosolute influences conformational stability.

2.4. Association Suppression

2.4.1. Baynes-Trout Neutral Crowder Theory

The work of a prior group member, Brian Baynes, has suggested the possibility of a new class of additives which should deter protein-protein association, thus acting on the second-order process of protein aggregation [41]. The inspiration for this new class of excipients derives from a theory developed in the group which describes how additives affect the rate of protein association reactions. This idea, called "gap effect theory", demonstrates that it is possible for an additive to exert a purely kinetic effect on protein association reactions [10]. To illustrate, a free energy reaction coordinate diagram of the "gap effect" is shown in Figure 2-3. In the diagram, large additives exhibiting the desired behavior (black circles) and water (grey circles) can both solvate protein molecules (P) equally well, either when in the dissociated (P+P) or associated (P_2) state, and as a result, there is no thermodynamic effect on which state the protein chooses to be in. However, at intermediate separation (center), there will be a separation distance where the gap between the protein molecules excludes the large additives for steric reasons but still allows water to solvate the gap. This results in a net preferential exclusion of the additive and a selective free energy increase in the encounter complex (*i.e.* the activation energy for the association reaction). Theoretical simulations show that this effect can greatly slow protein association reactions. Specifically, it is possible for a relatively large additive, which does not substantially affect the free energy of isolated protein molecules, to significantly increase the free energy

barrier for protein association. We call such additives "neutral crowders" because in theory, they do not affect the end states of the reaction or the free energy of unfolding, and are, hence, "neutral" to protein stability but if they are of significant size relative to water and are present in sufficient concentration, they "crowd" out protein-protein interactions.

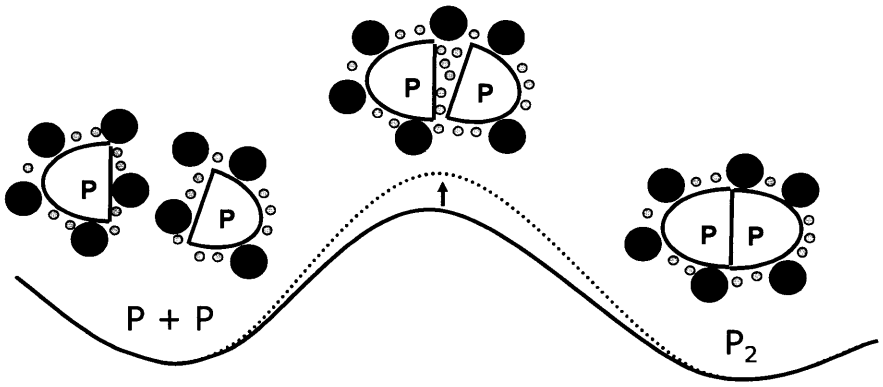

Figure 2-3: The gap effect. The entropically unfavorable exclusion of "neutral crowder" molecules raises the activation energy for protein association even though the free energy of the dissociated and associated states is not perturbed.

Figure 2-4 shows how the size and preferential interaction of an additive influences the rate of protein aggregation, as predicted by the "gap effect theory" developed by Baynes and Trout. The figure reveals that if a relatively large neutral crowder molecule (radius > 8 Å) can be synthesized and used as an additive to a protein solution, it has the potential to suppress protein association rates to a significantly greater extent than any other additive developed to date. In Figure 2-4, the relative association rate constant (k_a/k_{a0}), which compares the rate in the presence of an additive to the rate in its absence, is plotted as a function of additive radius and additive-protein preferential interaction coefficient. The preferential interaction coefficient, Γ_{XP} (which is the same parameter as Γ_{μ_3}), is the computed preferential interaction coefficient for the interaction with a spherical model protein (radius = 20 Å). These results indicate that to suppress association, an additive should be large and as attractive to proteins as possible because the gap effect is seen to increase greatly with increasing additive size (*e.g.* an 8 Å additive might be able to suppress the rate of association by 2-3 orders of magnitude).

The Baynes and Trout gap effect theory shows that, depending on the size of the additive and its energy of interaction with proteins relative to water, additives may accelerate, decelerate, or have no effect on the rate of aggregation. Specifically, for design purposes, a large additive which has an attractive interaction with proteins will slow aggregation the most. Unfortunately,

as the size of an additive is increased, the excluded volume between the additive and protein also increases, leading to a net increase in repulsion between the additive and protein [42]. Thus, merely increasing the size of an additive is not an effective strategy to deter aggregation. While larger size is desirable to emphasize the gap effect, an additive that is too large will exert a repulsive force that will prevent it from interacting "neutrally" with proteins.

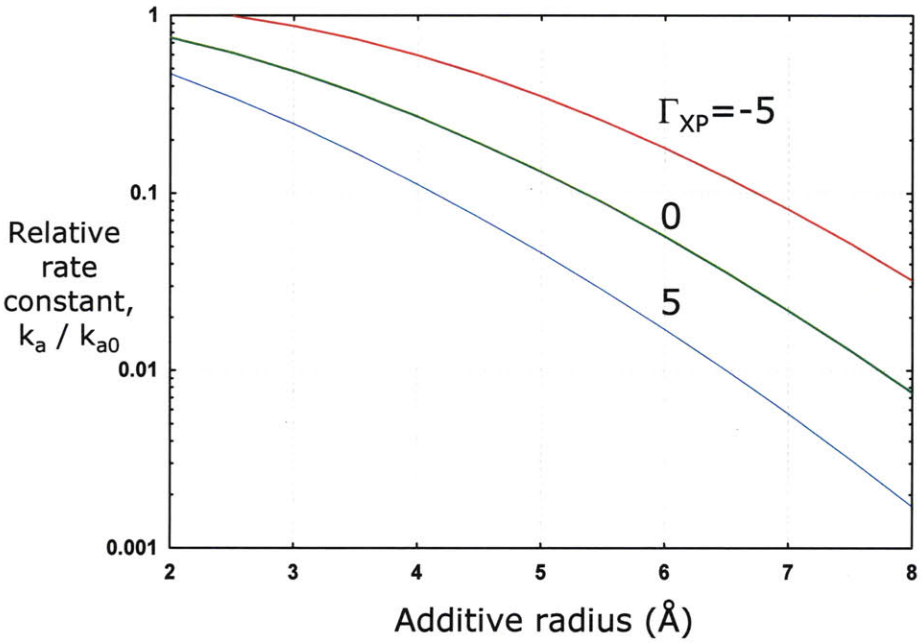

Figure 2-4: Association rate depression with increasing additive size, as predicted by gap effect theory (reproduced from [10]).

The key to designing a large, anti-aggregation additive will be to increase (moderately) the additive's size, while also making additive-protein interactions more favorable. The strategy for making interactions with protein molecules more favorable is to functionalize the additive's surface with guanidinium groups, which are known to bind to proteins. It is logical to assume that as the number of guanidinium groups on the surface of an additive is increased, the net interaction energy between the additive and proteins will become more attractive. However, if attractive forces between the new additives and proteins are too great, the additives will destabilize the protein structure and lead to unfolding or partial unfolding. Such action is similar to that of chaotropic denaturants such as guanidinium hydrochloride and urea. Thus, one must seek to balance the attractive interactions due to the surface groups with repulsive forces that result from excluded volume effects to produce an additive that acts net-neutrally with a protein. This is the key idea behind developing neutral crowders.

2.4.2. Arginine: A Putative Neutral Crowder

Gap effect theory suggests that large solution additives which do not affect the free energy of isolated protein molecules can slow down the kinetics of protein-protein association reactions. Such molecules, called "neutral crowders," are rare because as additive size is increased, the preferential interaction coefficient and hence protein transfer free energy falls off in proportion to the additive radius raised to the third power, due to an excluded volume effect. However, the aggregation suppressing behavior of arginine hydrochloride, ArgHCl, a naturally occurring compound (see Figure 2-5), seems to be consistent with the theorized "neutral crowder" effect.

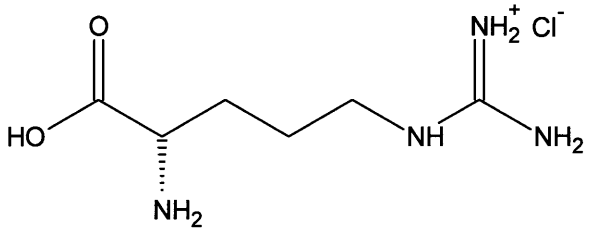

Figure 2-5: Structure of Arginine Hydrochloride (ArgHCl).

We assume that arginine hydrochloride is a "neutral crowder" because it has been shown by us and others that ArgHCl:

- Stabilizes proteins against aggregation but does not stabilize them against unfolding, as demonstrated by no improvement in a protein's melting temperature upon addition of ArgHCl to the solution [18].
- Reduces attractive protein-protein interactions as indicated by a shift from negative to a positive osmotic second virial coefficient, B_{22}, values, as indicated by light scattering experiments [43].
- Decreases the rate of association of globular proteins, as measured by surface plasmon resonance [15].
- Decreases the rate of association of unfolded and partially unfolded intermediates which are on the folding pathway during refolding, as measured by native protein activity and size-exclusion chromatography [15].

All of this observed behavior is consistent with the "neutral crowder" theory, hence our interest in fully characterizing the behavior of arginine and elucidating its mechanism of action. However, ArgHCl is only moderately effective at inhibiting aggregation, thus there is a necessity to develop more effective additives.

2.5. *Design and Testing Approach*

The "neutral crowder" theory and the observed behavior for arginine suggests that effective association suppressing excipients can be created by synthesizing a relatively large compound that has functional groups on the surface that interact favorably with protein molecules, producing a compound that is neither strongly bound nor strongly excluded from the surface of a protein due to a balance of exclusion and attraction. The background on solution additives presented above suggests that preferential interaction measurements and thermal unfolding analysis will be of use in the development of such excipients and thus, they were utilized to elucidate the mechanism by which the novel excipients influenced aggregation.

In our study, guanidinium chloride was chosen as the protein binding moiety since this functional group is present on arginine and it is known to bind to proteins. Such compounds were produced by modifying commercially available amine compounds (*e.g.* di- and tri-amines (see Figure 2-6 and Figure 2-7, respectively) and polyamidoamine (PAMAM) dendrimers (see Figure 2-8)). This in addition to synthesizing arginine peptides (see Figure 2-9). Moreover, in addition to preferential interaction and thermal unfolding analysis, the synthesized excipients were tested by incubating, at elevated temperatures, a solution containing a model protein in the presence of the novel excipients and monitoring monomer loss kinetics via size exclusion HPLC.

Tests showed that the first batch of excipients synthesized were unsuccessful at inhibiting aggregation. In fact, most accelerated aggregation. However, further analysis of the arginine mechanism revealed that ion-ion interactions are extremely important when it comes to suppression aggregation and careful analysis must be made when choosing the counterion for the guanidinium functional group. This led to the production of excipients that not only inhibit aggregation, but do so with abilities nearly an order of a magnitude great than arginine or any other commonly used excipient.

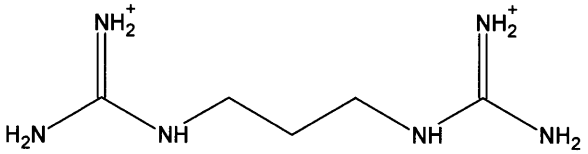

Figure 2-6: An example of a diguanidinium compound.

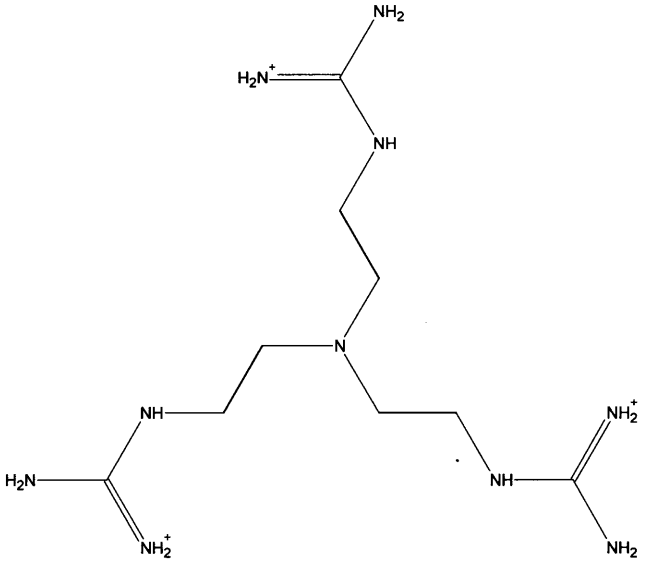

Figure 2-7: An example of a triguanidinium compound.

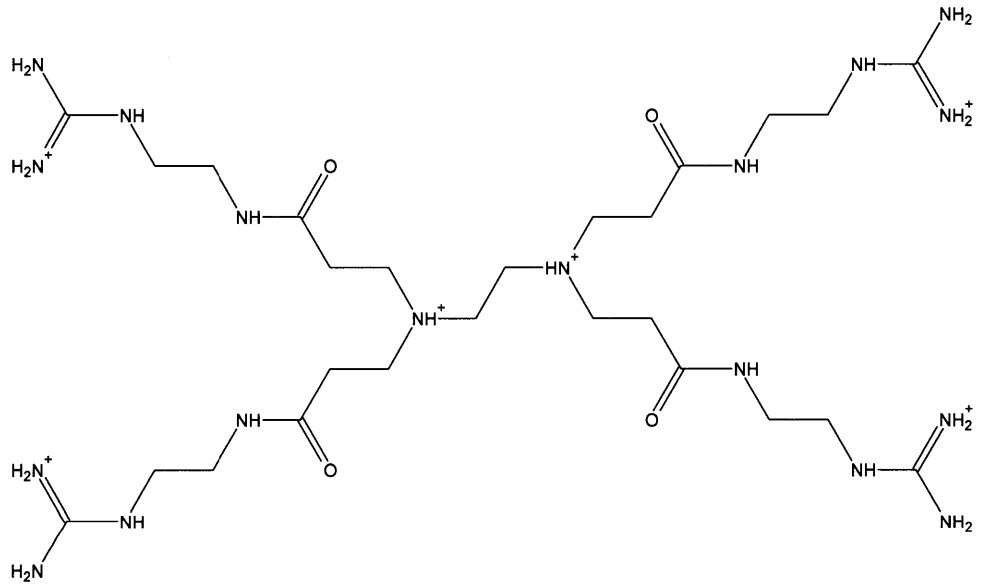

Figure 2-8: Structure of a Generation 0 PAMAM dendrimer with the surface modified to guanidinium.

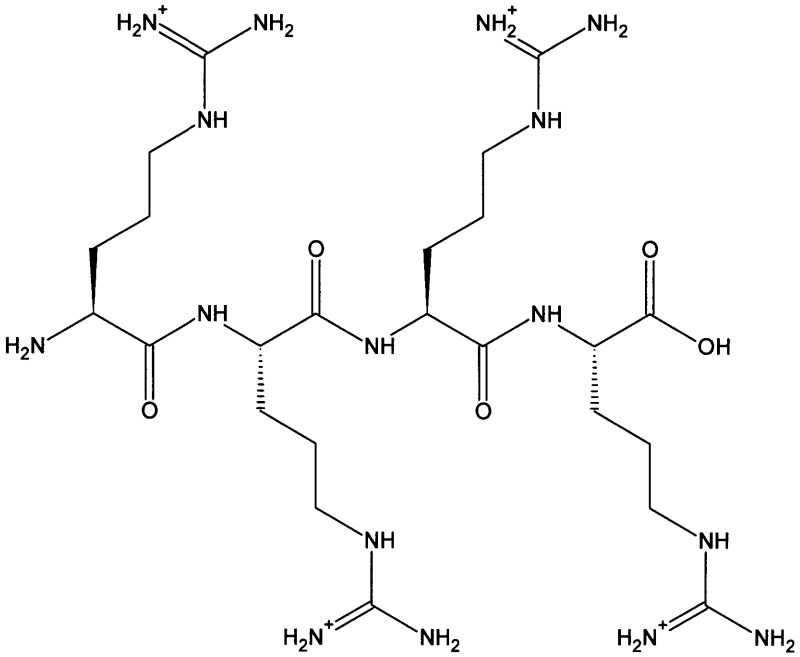

Figure 2-9: An example of an arginine peptide for use as an excipient.

Chapter 3

3. Experimental Procedures

One of the initial goals set at the beginning of this project was to establish a set of procedures and assays that would allow for a proper assessment of the effectiveness of the synthesized additives, not to mention allowing for the mechanism of action to be elucidated. To fully understand the results presented in this document, it is essential to be aware of how the data was collected. Therefore, this chapter is devoted to detailing all of the procedures utilized in order to provide clarity to the data presented and to eliminate any need for repetition in explaining the results. For a proper comparison of the data collected, these procedures remained nearly constant throughout the study and the conditions stated here are the conditions utilized for all the data presented, unless otherwise stated.

3.1. Proteins and Reagents

Bovine serum albumin (BSA), A3059, hen egg white lysozyme (Lys), L6876, bovine α-chymotrypsinogen A type II (aCgn), C4879, and jack bean concanavalin A (Con A), C2010, were obtained from Sigma-Aldrich (St. Louis, MO) as essentially salt free lyophilized powders. All other reagents used were also obtained from Sigma-Aldrich in the highest available grade. The concentration of BSA, lysozyme, aCgn, and Con A were determined by absorbance using extinction coefficients of 0.658 $mL*mg^{-1}cm^{-1}$ at 278 nm, 2.74 $mL*mg^{-1}cm^{-1}$ at 281 nm, 1.97 $mL*mg^{-1}cm^{-1}$ at 282 nm, and 1.37 $mL*mg^{-1}cm^{-1}$ at 280 nm respectively [31]. No corrections were made for cosolute concentrations or light scattering since the concentrations were typically low due to the protein solutions being diluted with pure water for the purpose of producing a sample with an optical density of about 1.0.

3.2. Chemical Synthesis & Purification

3.2.1. General Procedure

3.2.1.1 Synthesis

The general procedure developed for synthesizing the novel excipients involved converting commercially available amines into guanidine compounds through the use of a guanylating agent. The synthesis of guanidine compounds has received much attention lately

because of the biochemical and pharmaceutical importance of the guanidino moiety and a variety of guanylating agents are commercially available. For more info, Katritzky and Rogovoy have reviewed the recent developments in guanylating agents [44]. Of the many reagents available for different guanidinylation purposes, the most applicable reagents are either a thiopseudourea or a pyrazole carboxamidine because of our desire to synthesize compounds with unsubstituted guanidines on the outer surface of the compound. The later is a more potent guanylating agent but the former allows for an easier purification procedure since the leaving group is a volatile thiol that can simply be evaporated from the solution, on top of being less expensive. Most reagents come with protecting groups on the nitrogens and a *tert*-butoxycarbonyl (BOC) protecting group was chosen (see Figure 3-1) since it allowed for the reaction and purification to be carried out in an organic solvent (an aqueous phase synthesis was found to be not ideal due to limited reagent availability and purification options) and is easily removed with a strong acid, such as hydrogen chloride, forming a guanidinium salt, the desired state of the guanidino group.

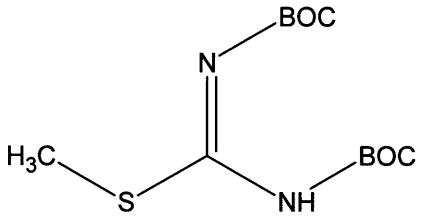

Figure 3-1: 1,3-Bis(*tert*-butoxycarbonyl)-2-methyl-2-thiopseudourea, the guanylating agent utilized in the synthesis of the excipients.

The procedure developed and presented here is a modification of the method presented by Gers and coworkers [45]. The typical procedure involved adding an excess of the guanidinylation reagent (usually a 1.25-2 fold excess) to a solution containing the amine of interest dissolved in either dichloromethane (DCM) or dimethylformamide (DMF). All reactions were conducted in a fume hood since the methanethiol produced not only has a very pungent odor but can be toxic in high doses. Reaction vessels were left open to the environment to allow the methanethiol to evaporate from the solution, allowing for a more facile reaction. DCM was found to be advantageous in some cases since it is highly volatile and thus easily removed through evaporation prior to purification. However, the substitution reaction involved proceeds somewhat slowly in DCM and solutions tend to dry out if reaction times are too long, thus it was only used for the smaller excipients, which reacted more quickly due to fewer substitution sites per molecule. DMF is better suited for facilitating the substitution reaction (*i.e.* reactions

proceeded more quickly) and for longer reaction times since it has a lower vapor pressure. However, due to this low vapor pressure, the solvent is somewhat difficult to remove from the solution using a standard laboratory vacuum pump. Therefore, DMF was only utilized when synthesizing larger excipients.

Solutions were stirred for 8-72 hours with the reaction monitored using thin layered chromatography (TLC). The TLC plates contained a fluorescent dye, allowing for spots containing organic compounds to be detected with a UV lamp (254 nm). The TLC mobile phase contained a mixture of polar and nonpolar solvents in a ratio that gave the best separation of the reaction system of interest. A typical mixture contained hexane:ethyl acetate:methanol in a ratio of 1.5:1.25:1. Unreacted amine compounds typically had no mobility due to their polar character, while the guanylating agent had high mobility due to its nonpolar character. The product (*i.e.* the excipient of interest with BOC protecting groups) usually had a low to moderate mobility. When all of the amine was consumed by the reaction, the solvent was evaporated under reduced pressure at 60°C using a rotavap. When DMF was used, only enough solvent was evaporated to reduce the volume to a reasonable level for the purification step because extended periods of high temperature and high concentrations of the dendrimer compounds caused the compounds to degrade and crosslink, producing an insoluble gel.

3.2.1.2 Purification

Several purification procedures were explored during this project. Initially, a polymer bound version of the guanylating agent was tested since no purification step was required but such products are expensive and were found to be inefficient at converting all of the substitution sites to guanidine. A liquid-liquid extraction (LLE) technique was explored but the best techniques were found to be solid-phase extraction (SPE) for small guanidinium compounds and precipitation for the dendrimers.

The small amine compounds used as starting reagents are inexpensive and available in large amounts. The products from these starting compounds are fairly nonpolar and hard to separate from the guanylating agent. It is much easier separating the product from just the polar amine compound, however, an excess of the guanylating agent is required for the reaction to go to completion because all of the compounds utilized contain multiple substitution sites. To resolve this issue, one option is to add the small amine compounds dropwise, with each drop added at a rate that allows the previous addition of amine to be completely converted, and

stopping when all of the guanylating agent is consumed. Another option is to react a known amount of the amine compound with a slight excess of the guanylating agent to consume all of the amine, then add a large excess of the amine once the reaction is completed to quickly consume the excess guanylating agent, thus creating a solution containing a known amount of the product and a small amount of unreacted and partially reacted amine. Neither procedure is advantageous in regards to the complexity of the final solution or the conversion rate for a given amount of the amine compound. However, the later procedure was utilized for simplicity in regards to the amount of monitoring required and the variability in reaction times (*i.e.* the rate at which drops are added must be in accordance to the rate at which the reaction proceeds for maximum conversion, which at the beginning of this project was unknown and varies depending on the amine starting compound). After the addition of additional amine, the solvent was evaporated at atmospheric pressure at 60°C and the residue was dissolved in hexane. The solution was then passed through a SPE column packed with silica to separate the product from unreacted or partially reacted amine compounds. The unreacted and partially reacted amines were retained on the column and the product passed right through (as determined by TLC). After washing the column with hexane, the solvent was evaporated from the collected solution under reduced pressure at 60°C.

As for the dendrimer compounds, the same procedure could not be utilized due to the high cost of the dendrimers. Also, SPE was found to be ineffective since the dendrimers had a high affinity for the silica and they could not be eluted, even with a highly polar solvent. Furthermore, the dendrimer products were found to be insoluble in hexane due to their polar character. Therefore, after much of the DMF was evaporated, the residue was dissolved in a minimal amount of diethyl ether and hexane was added to precipitate the product. The guanylating reagent was found to be soluble in hexane, thus after discarding the supernatant, washing the precipitates with hexane, redissolving the precipitates in diethyl ether and repeating the procedure several times, the product became quite pure (as determined by TLC).

3.2.1.3 Deprotection

The deprotection step removing the BOC protecting group is quite simple. After purifying the products and evaporating the solvent, the residues were dissolved in a minimal amount of DCM. Then an excess of 2M HCl dissolved in diethyl ether (sometimes 4M HCl dissolved in dioxane) was added to the solution. The vials were placed in a fume hood with the

cap left open to the environment to allow the carbon dioxide formed to escape. The final product slowly precipitated as the BOC protecting groups were removed. Occasionally, the vials were agitated or stirred with a Teflon coated spatula. After 24 hours, the precipitates were inspected. The deprotection was assessed to have been completed when the formation of carbon dioxide bubbles had ceased for at least two hours and the precipitates had a hard consistency that was easily ground into a powder (rather than having an amorphous consistency similar to gum). Following the deprotection step, the precipitates were spun down in a centrifuge, the supernatant discarded, and diethyl ether added to wash the precipitates. This was repeated at least 8 times (4 washes with diethyl ether and 4 washes with acetone). The precipitates were dried under reduced pressure at 60°C briefly and then dried under reduced pressure at room temperature for 24-48 hours. The structure of the products was analyzed using mass and NMR spectrometry. The purity was analyzed using NMR spectrometry by identifying the impurities and computing the amount of each relative to the product.

3.2.2. Ion Exchange

Ion exchange was utilized to prepare alternate salt forms of the synthesized excipients. An excess of the sodium salt of interest was used to load the anion of interest onto a column containing Amberlite IRA 400 anion exchange resin. After rinsing the column with HPLC grade water, a solution containing one of the synthesized excipients or L-arginine hydrochloride was passed through the column, in an amount less than half of the capacity of the column and at the recommended flowrate to ensure complete ion exchange. Additional water was then passed through the column to elute all of the excipient.

3.2.3. Type Specific Procedures

3.2.3.1 Arginine Salts

To help elucidate the arginine mechanism, a variety of arginine salts were produced. Most of the arginine salt forms of interest are not available commercially and thus they had to be prepared from L-arginine or L-arginine hydrochloride. Most of the salts (phosphate, sulfate, citrate, acetate, fluoride, bromide, and iodide) were prepared by titrating an L-Arginine solution with the appropriate acid solution. Arginine thiocyanate (ArgHSCN) had to be prepared via ion exchange, as described above, since thiocyanic acid is an unstable compound that is not commercially available. All arginine salt solutions were completely dried for storage and

reconstituted at the desired concentration prior to use. Most of the water was evaporated under reduced pressure at 60°C using a rotavap. The crystals that formed were further dried for at least 48 hours over a calcium sulfate desiccant under reduced pressure. Prior to use, the amount of residual water was determined by gently heating a small sample (500 mg) at around 120°C until the mass remained constant. This procedure verified that the phosphate and citrate salts were monohydrates.

3.2.3.2 Modified Amine Compounds

To summarize the developed synthesis technique, commercially available di- and tri-amine compounds were guanylated with an excess of the guanylating agent in DCM and a large excess of the amine was added once the reaction was completed to completely consume the remaining guanylating agent. The product and unreacted amine were separated using SPE with hexane as the mobile phase. The BOC protecting group was removed from the purified product with concentrated HCl dissolved in either diethyl ether or dioxane. The precipitated product was washed with diethyl ether and acetone and then dried at an elevated temperature and reduced pressure.

3.2.3.3 PAMAM Dendrimers

To summarize the developed synthesis technique, commercially available PAMAM dendrimers (generations 0-2) were guanylated with an excess of the guanylating agent in DMF. After the reaction was completed, most of the DMF was evaporated at 60°C under reduced pressure and the excess guanylating agent was removed by dissolving the residue in a minimal amount of diethyl ether and precipitating the product with hexane, repeating the procedure until the product was pure (as indicated by TLC). The BOC protecting group was removed from the purified product with concentrated HCl dissolved in either diethyl ether or dioxane. The precipitated product was washed with diethyl ether and acetone and then dried at an elevated temperature and reduced pressure. Alternate salt forms were prepared via ion exchange, with excess water evaporated at 60°C under reduced pressure and dried to completeness via lyophilization.

3.2.3.4 Arginine Peptides

The arginine peptides were not synthesized in house. Collaboration with another research group, which specializes in peptide research, was initiated prior to the beginning of this project. They were tasked with preparing the peptides. However, their contribution was quite poor. It

took them nearly a year to figure out which HPLC column to use to purify the peptides. Even after they determined the proper procedure, they could only provide a few milligrams of peptide (which was barely enough for a single experiment) every 4-6 months. The purity of the peptides they produced were questionable and the lyophilization technique they used to convert the peptides from a TFA to chloride salt was likely ineffective. Due to their poor performance, customized peptides were later purchased from GenScript® Corporation, with each batch size in an excess of 1 gram and purity greater than 98%. LC-MS data was provided to validate the purity and structure of the peptides. Alternate salt forms, including the chloride salt form, were prepared via ion exchange, with excess water evaporated at 60°C under reduced pressure and dried to completeness via lyophilization.

3.3. Structural Analysis

3.3.1. Nuclear Magnetic Resonance (NMR) Spectrometry

The structure and purity of the synthesized excipients were analyzed via NMR spectrometry. The Bruker Avance 400 NMR spectrometer in the Department of Chemistry Instrumentation Facility (DCIF) was utilized for this purpose. 1D ^{1}H (400 MHz) and 1D ^{13}C (101 MHz) spectrums were produced for official validation of the structure. 2D ^{1}H-^{1}H COSY and 2D ^{1}H-^{13}C HSQC techniques were utilized to help deconvolute the 1D spectrums. If any ambiguity persisted, results were compared to predictions made using ChemBioDraw Ultra 12.0. The guanylating agent and amine precursors were analyzed as well (which also included the amine in an ammonium chloride salt form for comparison to the guanidinium salt spectrum) to identify any unreacted precursors. The amine starting compounds and the BOC protected intermediates were each dissolved in deuterated chloroform ($CDCl_3$) and the guanidinium salt products and ammonium salt precursors were dissolved in deuterated dimethyl sulfoxide (DMSO-d6), all at a concentration around 25 mg/mL. ^{1}H and ^{13}C spectrums were calibrated using residual protio-solvent. The total integrated area of the ^{1}H spectrums were set to the total number of hydrogens present on the target molecule. No unreacted amine or guanylating agent were detected, thus the only impurities were residual solvent, which were identified by comparing chemical shifts with literature values. The residual solvent peaks were integrated separately from the product peaks to determine moles of solvent per mole of product, thus allowing a calculation of mass percentage. (See Appendix A for NMR data for each excipient analyzed).

3.3.2. Mass Spectrometry (MS)

Several ionization sources are available at MIT for mass spectrometry analysis, but electrospray ionization (ESI) was deemed best suited to analyze the synthesized compounds since (1) the technique allows for the analysis of samples with a molecular weight greater than 1000 daltons and (2) of the techniques available for ionizing large compounds, ESI produces the least amount of fragmentation without the use of a matrix, which can interfere with the analysis of ions with small mass-to-charge ratios. The DCIF houses a Bruker Daltonics APEXIV 4.7 Tesla Fourier Transform Ion Cyclotron Resonance Mass Spectrometer (FT-ICR-MS) equipped with an electrospray ionization source. Only positive ions were analyzed (with samples processed by the facility's mass spectrometry specialist) and the data was deemed satisfactory when the major peaks corresponded to a $[M+nH]^{n+}$ value, typical of ESI-MS. (See Appendix A for MS results for each excipient analyzed).

3.4. Model Proteins

The most critical aspect of this project was assessing the effectiveness of the novel excipients at inhibiting aggregation. Protein aggregation is a complex process that is completely dependent on the physical and chemical properties of the protein(s) involved. For example, some proteins are so unstable that nearly all of the dissolved protein will aggregate within a few hours at room temperature and at a relatively low concentration, while some proteins require near boiling temperatures to cause them to aggregate, even at high concentrations. Furthermore, the effect additives have on aggregation is not completely universal, thus any new additive produced must be tested on all proteins of interest to verify effectiveness. Therefore, a thorough assessment of the effectiveness of any new excipient would be an extremely difficult and tedious task. To allow an adequate amount of research and development to be conducted during this project, testing had to be restricted to one or a few model proteins that are aggregation prone and representative of pharmaceutical proteins.

Obviously, performing tests using actual pharmaceutical proteins would be ideal. However, such proteins are very expensive and obtaining an adequate amount would not be feasible without some sort of collaboration with the manufacturer. Producing a pharmaceutically relevant protein in-house would be less expensive, but such work is very time consuming and only a relatively small amount can be produced and purified in a lab setting. A tremendous amount of protein was required for all of the testing and stability analysis required in this study.

Therefore, the model protein was restricted to an inexpensive, commercially available protein that can be obtained in large quantities but also has characteristics similar to pharmaceutically relevant proteins.

3.4.1. Literature Review and Early Analysis

Many articles that deal with protein aggregation (which is also a topic of interest in other fields, such as food science, neurological diseases, diagnostics, *etc.*) have used ordinary proteins commonly found in most laboratories (*e.g.* BSA, lysozyme, beta lactoglobulin, hemoglobin, ovalbumin, ribonuclease A, carbonic anhydrase *etc.*) for the very same reasons mentioned above [15, 46-51]. Other articles have utilized more pharmaceutically relevant proteins (*e.g.* insulin, human factor VIII, immunoglobulin, human growth hormone, GCSF, human interferon gamma, *etc.*), but in most cases, the source of the protein was provided by a pharmaceutical company. Moreover, in some cases, the aggregation of the particular protein is not of a pressing concern, other cases showed the aggregation mechanism to be quite complex and for some, the protein seemed to not be ideally suited for heat induced aggregation studies [52-57].

Early work conducted in this study began with BSA, lysozyme, and carbonic anhydrase. However, those proteins did not make ideal model proteins because the aggregation kinetics of BSA were hard to analyze, lysozyme was found to show little aggregation propensity unless heated to fairly high temperatures, and carbonic anhydrase was a bit more expensive and did not respond well to high ionic strength. However, after reviewing the literature again, newly published articles showed that α-Chymotrypsinogen A would make an ideal model protein.

3.4.2. Bovine α-Chymotrypsinogen A

α-Chymotrypsinogen A (aCgn) was selected as a model protein for aggregation because it is a well characterized protein that has been used extensively as a model protein in aggregation studies due to it having aggregation characteristics similar to therapeutically relevant proteins [58-60]. More importantly, the native state is monomeric, with simple unfolding thermodynamics and kinetics that forms nonnative aggregates on a reasonable time scale at mildly elevated temperatures.

aCgn is a globular single-domain protein with a molecular mass of 25.7 kDa and an isoelectric point (pI) of 9.2 (thus it has a positive charge when formulated at physiological or low pH). The protein is the inactive precursor of the serine protease chymotrypsin. Thermal stability

studies show the protein to be most stable against unfolding in a pH range of 3 to 6 (see Figure 3-2) [61].

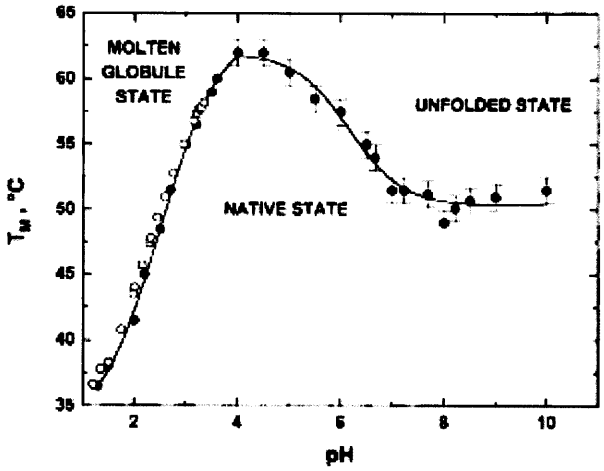

Figure 3-2: Denaturation midpoint temperature (T_m) of aCgn versus pH, demonstrating optimal conformational stability at around pH 4 (figure taken from [61]).

Other studies show that the protein has the greatest colloidal stability (as determined by osmotic second virial coefficient values) at low pH and low ionic strength [62]. However, at low pH, the protein is conformationally unstable from being in such a highly charge state and would be highly aggregation prone if it were not for strong repulsive electrostatic interactions. Furthermore, at low pH, increasing the salt content has a detrimental effect on colloidal stability due to the screening of repulsive interactions. At a higher pH, the charged state of the protein is much lower and is prone to associating due to a lack of electrostatic repulsion. Increasing the ionic strength screens attractive charge-dipole and dipole-dipole interactions. Thus at high ionic strength, the pH of the solution has little influence on colloidal stability, similar to other proteins.

These results also show that at around pH 5 (the point at which the dominate interaction shifts from repulsive electrostatic interactions to attractive dipole interactions), the protein is least sensitive to ionic strength, thus making this the ideal pH for this study. This is not only because the protein exhibits the greatest conformational stability at that pH, ionic strength effects would not contribute to the aggregation of the protein. This would allow for a clearer assessment of the aggregation suppression mechanism. These results are almost identical to those for chymotrypsin, which were obtained by a separate research group, indicating that these results are reliable [63].

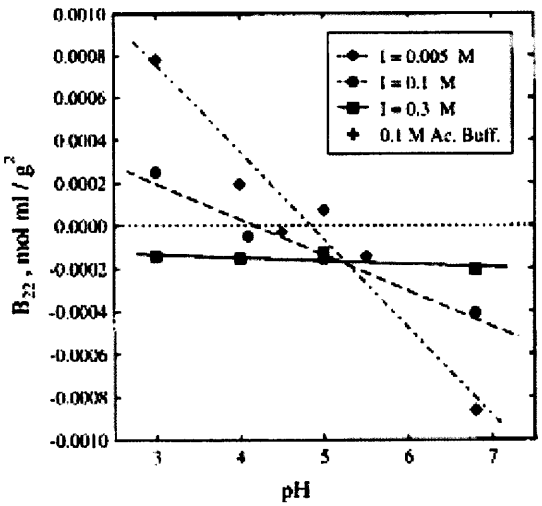

Figure 3-3: Osmotic 2nd Virial Coefficient (B_{22}) values for aCgn (at varying ionic strength) versus pH, demonstrating that aCgn is least sensitive to ionic strength at around pH 5 (figure taken from [62]).

3.4.3. Enzymatic Inhibition and Solution Preparation

It should be noted that the sample of aCgn obtained contains trace amounts of either chymotrypsin or trypsin, which exhibits proteolytic activity when the pH is greater than 4.5. When incubated at elevated temperatures, the digestion rate of aCgn is comparable to the rate of aggregation, leading to a considerable amount of small molecular weight peptides forming over the course of an experiment and likewise, a considerable amount of the protein lost to digestion rather than aggregation, making it impossible to determine how an additive influences aggregation. Lowering the pH was not an option because at low pH, the protein is highly sensitive to the presence of ions. As described above, a pH of about 5 is optimal for when using this protein. To verify this, sodium chloride and potassium chloride were used as cosolutes, up to a concentration of 1M, at a pH of 5. Neither salt exhibited any influence on the rate of aggregation under these conditions. However, this was not the case for when the pH was lowered to below 4.5. At low pH, all types of salt, including ArgHCl, sped up aggregation, with all types increasing the rate by nearly the same extent. This indicated that the loss of colloidal stability resulting from electrostatic shielding dominated any aggregation inhibition the salt compounds may have exhibited.

Therefore, to keep the protein from being digested, a stock of aCgn was first treated with phenylmethylsulfonyl fluoride (PMSF), a serine protease inhibitor that irreversibly inhibits proteolytic activity. 35 mg of PMSF was first dissolved in 1 mL of anhydrous ethanol. This solution was then added in 250 μL increments to a stirred 50 mL stock solution of aCgn (10

mg/mL, 100 mM sodium phosphate pH 7) every hour until PMSF reached its solubility limit, usually after three additions. The higher the pH of the protein solution, the faster the inhibition reaction proceeds. Likewise, the inhibitor degrades more quickly at high pH due to a reaction with water, thus a pH of 7 was chosen for an optimal balance of inhibition rate and inhibitor stability. Moreover, adding PMSF in increments was utilized because the inhibitor, though effective at low concentrations, still rapidly degrades in aqueous solutions at pH 7 [64]. Three hours after the first addition of PMSF, the solution was passed through a 0.1 micron filter (Millipore) to remove PMSF precipitates, then placed under dialysis (Spectra/Por 3500 MWCO) for 48 hours at 4°C against three changes of 20 mM sodium citrate pH 5 buffer (700 mL each). Following the dialysis, the solution was first centrifuged briefly to remove any aggregates that precipitated during dialysis and then concentrated using a centrifugal filter (Amicon Ultra 10,000 MWCO) to a final concentration of 20 mg/mL.

This treatment reduced the digestion rate to an insignificant level and all aCgn samples produced in this manner had nearly identical aggregation rates. All aCgn solutions were frozen, rather than lyophilized, in 0.3 mL aliquots, and used within 6 months. Storing the protein in this manner did not seem to have any detrimental effect on the quality of the protein (*i.e.* the amount of aggregates in the initial solution) nor the rate of aggregation when used in the study. Solutions for the aggregation study were prepared by first preparing a solution of the cosolute of interest in a 20 mM sodium citrate pH 5 buffer at a concentration double that of interest (adjusting the pH to 5 if needed with the appropriate acid or an 8M sodium hydroxide solution), then mixing 0.3 mL of this solution with 0.3 mL of a thawed aCgn solution. This solution was then divided into twelve aliquots, each 50 µL in volume, and incubated at an elevated temperature, as described below.

3.5. Accelerated Aggregation Assay

The rate of aggregation for most proteins is actually fairly slow. At refrigerated or room temperatures, it may take days, weeks, or maybe even months before significant levels of aggregates to form, if at all. One of the most obvious signs that aggregation has occurred is when the solution has turned turbid from the precipitation of high molecular weight oligomers. In some cases, the extent of aggregation can be verified by monitoring the amount of visible light obscured. However, this is not the most robust means of monitoring aggregation, because often, the aggregates will settle to the bottom of the vial or vessel. Furthermore, in many cases, the

aggregates remain soluble, thus giving no visible indication that aggregation has occurred, in addition to giving no viable means to remove the aggregates from the solution prior to injection. However, if aggregates do form, it does not take much for the solution to be deemed unacceptable. A general rule of thumb is that the maximum amount of protein in the aggregated state cannot exceed 5% of the total amount of protein. With this restriction, the shelf-life of many products is relatively short (*i.e.* a few weeks to several months) which is far short of the 18 to 24 month shelf-life typically desired. Even with a shelf-life of a few days to several weeks, this rate is far to slow to allow for adequate testing and analysis of the novel excipients. If experiments were conducted at room temperature, only a few experiments could be tested over the course of a four year project. Therefore, the rate of aggregation had to be accelerated by stressing the protein solution, preferably by enough to allow several experiments to be conducted per week.

There are several ways one could destabilize a protein in order to accelerate aggregation [5]. However, only a few are frequently used during aggregation studies (*e.g.* elevated temperatures, extreme pH, shaking and shearing, freezing and thawing, refolding, addition of a destabilizing chemical, *etc.*) because these stresses often occur during production and storage. Elevating the temperature is the simplest method and most relevant to this study since it is the most representative of formulation conditions. At slightly elevated temperatures, the main effect is accelerated association because of more frequent collisions between protein molecules. However, as the temperature is increased further, the conformation of the protein becomes less compact and ultimately unfolds. This leads to non-Arrhenius behavior, which makes shelf-life prediction from accelerated conditions extremely difficult, if not impossible within any reasonable level of certainty, unless a model is developed from low temperature data. In addition, it is possible for the mechanism of aggregation at elevated temperatures to differ significantly from normal storage condition and therefore, methods developed for inhibiting aggregation at elevated temperatures might not be applicable at lower temperatures.

However, the theory developed to guide this research suggests that the excipients developed will specifically inhibit association and thus they will inhibit aggregation regardless of the conformational state of the protein. Furthermore, the temperature selected was chosen to be more than 10°C lower than the onset of unfolding, making the amount of unfolded protein quite low. However, at such temperatures, the rate limiting step might shift from partial unfolding to

association due to a rapid increase in the rate at which partially unfolded species develop relative to the increases in the rate of association.

On a side note, it was initially hypothesized that the "neutral crowder" excipients would only be effective for when association is the rate limiting step in aggregation. However, upon further analysis, it is clear that if the excipients slow association to a rate lower than the rate of partial unfolding (which occurs prior to association), they will be effective for the case when partial unfolding is normally the rate limiting step, thus making them universally applicable.

3.5.1. Elevated Temperatures

The aggregation of aCgn was accelerated by incubating samples at an elevated temperature. A temperature of 52.5°C was found to be optimal, in that it was well below the onset of unfolding but accelerated aggregation enough to allow an aggregation experiment to be completed in less than a day. This particular temperature value was reached upon by analyzing aCgn aggregation over a range of temperature values, each separated by a 2.5°C increment. In addition to all aggregation experiments being conducted at that temperature, all samples contained 10 mg/mL aCgn, held at pH 5.0 using a 20 mM sodium citrate buffer. 50 µL aliquots of each protein-cosolute mixture were placed in 0.2 mL PCR tubes and then incubated in a Bio-Rad MyCycler thermal cycler with a timer initiated when the thermal cycler reached the desired temperature. Each series contained 12 samples and up to 8 different mixtures were tested at the same time during any given experiment. Samples were removed periodically and immediately placed in an ice bath to quench the aggregation reaction. The length of each experiment was planned so that the extent of reaction was such that at least 20% of the protein aggregated, making changes in the rate of aggregation pronounced. After at least five minutes in the ice bath, the samples were briefly spun in a minicentrifuge to remove precipitated aggregates and then transferred to autosampler vials for HPLC analysis.

3.5.2. Size Exclusion HPLC (SE-HPLC)

15 µL of each aCgn sample were injected into an Agilent 1200 series HPLC, equipped with a Zorbax GF-250 (4.6x250 mm, 4 micron) size exclusion column and a UV-Vis detector, during a sequence. The HPLC mobile phase contained 25 mM citric acid monohydrate, 25 mM sodium acetate, and 200 mM sodium chloride with the pH adjusted to 4.0 using sodium hydroxide. The flowrate was set at 1 mL/min and samples were monitored at 280 nm.

Most of the aggregates that formed either precipitated or did not pass through the column, thus only the monomer concentration was monitored with time. Peak area at each time point was compared to the peak area for an unstressed sample to create a monomer loss profile normalized with respect to the initial concentration. Each profile was compared to a profile for a sample containing no cosolute to determine the change in aggregation rate upon addition of the cosolute. For a control, each aggregation experimental setup had one protein solution that contained no cosolute to verify that the baseline aggregation rate remained constant amongst all experiments. For all aCgn concentrations studied, the aggregation exhibited 2nd order kinetics, thus rate constants were determined by fitting each data set to a 2nd order rate law.

3.5.3. Cosolute Osmotic Virial Coefficient Measurements

Vapor pressure osmometry was utilized to evaluate intermolecular interactions in aqueous solutions of the arginine salts. The most frequently used approach to describe interactions in electrolyte solutions is the ion interaction or virial coefficient model developed by Pitzer and co-workers [65, 66]. The model accounts for nonideal trends in the osmotic coefficient, ϕ, which is defined as

$$\phi = -\frac{\ln a_w}{\nu m M_w} = \frac{Osm}{\nu m}, \tag{4.1}$$

where a_w is the activity of water, M_w is the molar mass of water, m is the solute molality, and ν is the stoichiometric coefficient of the solute. For a single MX electrolyte solution, the Pitzer model can be expressed as

$$\phi - 1 = |z_M z_X| f^{\phi} + \frac{2\nu_M \nu_X}{\nu} B^{\phi} m + \frac{2(\nu_M \nu_X)^{3/2}}{\nu} C^{\phi} m^2, \tag{4.2}$$

where ν_M and ν_X are the number of M and X ions in the formula and z_M and z_X are their respective charges in electronic units; also $\nu = \nu_M + \nu_X$. The other parameters have the form

$$f^{\phi} = -A_{\phi}\left[\frac{I^{1/2}}{1 + bI^{1/2}}\right] \tag{4.3}$$

$$B^{\phi} = \beta^{(0)} + \beta^{(1)} e^{-\alpha I^{1/2}}, \tag{4.4}$$

where I is the ionic strength and A_{ϕ} is the Debye-Hückel coefficient for the osmotic coefficient, both in the molal concentration scale. The Wescor 5520 vapor pressure osmometer utilized in this study has no temperature control, but A_{ϕ} is a relatively weak function of temperature, thus

the value for water at the standard temperature of 25°C was utilized, which is A_ϕ = 0.392 $(kg/mol)^{1/2}$. The parameters α and b can be floated during analysis but they are typically fixed at 2.0 $(kg/mol)^{1/2}$ and 1.2 $(kg/mol)^{1/2}$, respectively, for all electrolytes [66]. A previous analysis of ArgHCl utilized these parameter values and as a result, the same was done in this study [67]. The parameters $\beta^{(0)}$ and $\beta^{(1)}$ describe the ionic strength dependence of the second osmotic virial coefficient, B^ϕ, which describes short range binary interactions. C^ϕ is the third virial coefficient, which describes ternary interactions and is typically small and only important at high concentrations.

In this analysis, the excipients were dissolved in HPLC grade water, with no buffering component or pH adjustment. A series of samples were prepared gravimetrically and the density of each solution was determined using an Anton Paar DMA 4500 density meter, accurate to within $1x10^{-5}$ g/mL, to determine both the molar concentration of each solution and the partial molar volume. $\beta^{(0)}$, $\beta^{(1)}$ and C^ϕ were determined by fitting the osmometry data to Eq. (4.2) using the nonlinear fitting program Igor Pro.

3.6. Preferential Interaction Determination

3.6.1. Dialysis/Densimetry

Dialysis/Density measurements were carried out according to a previously described procedure [32]. Lyophilized protein was dried in a desiccator at room temperature for 2 days under reduced pressure to remove trace amounts of water. Five samples of the dried protein were dissolved at a concentration ranging from 5 to 25 mg/mL in 4 mL of buffer solutions containing the cosolute at the desired concentration and pH. For the constant molality measurements, 2 mL of each sample were set aside, tightly sealed, and left at room temperature overnight prior to the densimetry measurements. In the constant chemical potential experiments, the remaining 2 mL of each protein sample were transferred to a dialysis bag and dialyzed at room temperature for 48 hours against two changes of solvent (400-500 mL each) prior to densimetry measurements. The density of each solution was measured using an Anton Paar DMA 4500 density meter, accurate to within 0.00001 g/mL. After the density measurements, samples of each protein solution were diluted gravimetrically with water to a final optical density of about 1.0 and the concentrations were determined spectrophotometrically with a PerkinElmer Lambda 35 UV/Vis spectrometer.

The partial specific volume of the protein from each experiment was determined using Eq. (3.7) and these values were used to compute the preferential interaction coefficient via Eq. (3.6).

3.6.2. Vapor Pressure Osmometry (VPO)

3.6.2.1 Preparation of Solutions for VPO Measurements

Osmolality measurements were carried out according to a previously described procedure [27]. For improved accuracy, all samples were prepared gravimetrically using a Mettler Toledo balance with a precision of 0.1 mg. Partial molar volume values of the cosolutes (evaluated using Eq. (3.8)) and density values of the solutions were used to determine concentration. Even though a three component system was utilized for the derivation of Eqs. (3.15) and (3.18), it was found that a pH buffering salt was necessary for the measurements, mainly to stabilize the protein solution and to control the pH of the guanidineHCl and arginineHCl solutions. Since the concentration of the buffering salt is constant for every sample, not to mention low, it has no contribution to the computed slopes used in Eqs. (3.15) and (3.18). Therefore, the molal concentration of the buffer can be lumped together with that of the water and the system can be treated as a three component system (*i.e.* solvent, protein, and additive).

Stock solutions of both BSA and lysozyme and stock solutions of the cosolutes were each prepared in a 40 mM sodium phosphate pH 6 buffer, while aCgn measurements were conducted in 20 mM sodium citrate buffer, either at pH 3.5 or pH 5.0. Stock solutions of the cosolutes were used so that a series of samples could be easily produced with varying solute concentrations by dilution with water and buffer. The protein stock solutions were further purified by placing the solutions under dialysis for 48 hours at room temperature against two changes of buffer (1 L each). To obtain the best possible results from the VPO measurements, the protein solutions were concentrated to within 100-200 mg/mL (depending on the solubility of the protein) using Amicon Ultra centrifugal filter tubes. To determine the concentration, samples of each stock solution were diluted gravimetrically with water to a final optical density of about 1.0 and the concentration determined spectrophotometrically with a PerkinElmer Lambda 35 UV/Vis spectrometer. It should be noted that the protein stock solutions were stored at 4°C and used with in one week of their preparation.

Eqs. (3.15) and (3.18) require that osmolality measurements for the three component solution at various cosolute concentrations be taken at a constant protein molal concentration. However, it is difficult to prepare such solutions and therefore, samples with a constant molar

volume were prepared instead. To correct for this, a relationship between constant protein molal concentration (m_2) and constant protein molar concentration ([2]), derived elsewhere, was used to obtain values for the desired parameter [27]:

$$\Omega_3 = \frac{\left(\partial Osm/\partial m_3\right)_{[2]}}{1-\Gamma_{\mu_1}\left(\frac{[2]\bar{V}_3}{1-[2]\bar{V}_2}\right)}. \tag{4.5}$$

A Wescor Vapro® vapor pressure osmometer (model 5520) was used for all osmometry measurements. The instrument was calibrated using 0.1, 0.29, and 1 mol/kg osmolality standards provided by Wescor. Due to slight drifts in the calibration points, the instrument was recalibrated after every four measurements and the thermocouple was cleaned daily using concentrated ammonium hydroxide followed by a rinse with HPLC grade water. Even though the instrument is designed to measure osmolality readings as high as 3.2 mol/kg, the accuracy of measurements above the last calibration point of 1 mol/kg is significantly impaired from extrapolating the calibration curve and increased system noise at high solute concentrations. It was discovered during the course of such measurements that this error significantly influenced the accuracy of the results. Therefore, cosolute concentration was limited so that all measurements had an osmolality at or below 1.2 mol/kg. The data sets for each cosolute included at least 16 individually made solutions (both for the sets with and without the protein). Each sample contained approximately 80 µL of solution (40 µL of diluted cosolute stock solution and 40 µL of the protein stock solution). A minimum of three measurements of each solution were made. If the standard deviation for a particular solution was greater than 2 mmol/kg, more measurements were taken or another sample was prepared. For a given solute, data representing the dependence of osmolality on solute molality were fitted by a quadratic equation, with all fitting parameters floated. Data were weighted by the standard deviation of the individual osmolality measurements for the same solution using the nonlinear fitting program Igor Pro. Using a cubic function or other fitting curves did not improve the fitting for any of the data.

3.6.2.2 Determination of the Isoosmolal Preferential Interaction Coefficient (Γ_{μ_1})

Though the isoosmolal preferential interaction coefficient Γ_{μ_1} can be determined directly from osmolality data (via Eq. (3.17)), it is difficult to obtain an accurate measure of the relationship between osmolality and protein concentration for every cosolute concentration. This

is mainly due to the fact that osmolality does not change significantly over the available range of protein concentration. However, the parameter can be approximated by comparing the molal concentration of the cosolute in the protein solution (m_3) to the molal concentration of the cosolute in a two component solution at the same osmolality (*i.e.* the same water chemical potential) (m_3^{Δ}):

$$\Gamma_{\mu_1} = \left(\frac{\partial m_3}{\partial m_2}\right)_{T,P,\mu_1} \cong \frac{\left(m_3 - m_3^{\Delta}\right)}{m_2}. \tag{4.6}$$

By using such a procedure, the preferential interaction coefficient Γ_{μ_3} can be determined using both approximation methods from just three series of samples (*i.e.* a series containing all three components at various cosolute concentrations, a series of just protein and buffer at various protein concentrations, and a series of just the cosolute and buffer at various cosolute concentrations). For more information, see Table B-1 in Appendix B for experimentally determined partial molar volumes and osmolality dependence on solute molality for two-component solutions.

3.7. Folding Thermodynamics/Structure Analysis

3.7.1. Differential Scanning Calorimetry (DSC)

The thermodynamic stability of aCgn in the presence of the excipients was determined by DSC (Microcal VP-Differential Scanning Calorimeter located in the Biophysical Instrumentation Facility (BIF) at MIT). Each reading began with a minimum of three buffer-buffer up and down scans (in this case, the buffer also contains the cosolute of interest) to establish a reproducible thermal history followed by a single protein-buffer up scan. aCgn was analyzed at a concentration of 1 mg/mL in a 20 mM sodium citrate pH 5 buffer containing the cosolute of interest and a scan rate of 90°C/hour. The data was analyzed in the MicroCal Origin® plotting software. In the analysis, the last buffer-buffer reference scan was first subtracted from the data, which was then normalized with respect to the protein concentration, cell volume, and scan rate. The baseline (which was progressed through the peak) was then subtracted from the plot and the denaturation midpoint temperature, T_m, was taken as the temperature at the peak maximum of the unfolding event. For each excipient, three concentrations (including zero concentration) were tested and T_m values with respect to cosolute concentration were fitted to a linear trend.

3.7.2. Circular Dichroism (CD) Spectrometry

There are additional methods available for determining the transition point of unfolding. An alternate method utilized was through the use of a circular dichroism (CD) spectrometer equipped with a temperature control system. The amount of circularly polarized UV radiation a protein absorbs will depend on its secondary and tertiary structure. At 220 nm, the ellipticity (which represents the difference in the amount of left and right circularly polarized light absorbed) for alpha helix and beta sheet structures are distinctive from a random coil. Therefore, if the circular dichroism of a protein is monitored at 220 nm as the temperature is increased, the ratio of unfolded to folded protein and the denaturation midpoint temperature can easily determined from the change in the ellipticity, assuming the ellipticity reaches a constant value and that this value represents a completely unfolded protein. The BIF possesses an Aviv Model 202 Circular Dichroism Spectrometer equipped with a thermostat that can scan from -10° to 110°C.

The advantage of CD spectrometry is that the procedure is fairly straightforward and allows for rapid measurements (as apposed to DSC, which can take hours or days for a single measurement). The procedure is often used to determine the effect cosolutes have on the structure of a protein as well as the unfolding temperature but the procedure is not applicable if the cosolute absorbs UV radiation, which is the case for guanidinium compounds. This is a major downside for this project. Similar techniques (such as fluorescence spectrometry) would also have a similar problem because of the absorbance of UV radiation, thus the technique was found to only be useful for analyzing the structure of the protein without any cosolute. This did give a profile of the amount of protein unfolded versus temperature, which was found useful in shelf-life prediction and determining an optimal incubation temperature.

Chapter 4

4. Elucidating the Arginine Mechanism

4.1. Arginine Hydrochloride (ArgHCl)

As mentioned before, one excipient that has gained much attention lately is arginine hydrochloride (ArgHCl) [68]. Its ability to greatly improve the refolding yield of proteins was discovered, nearly by accident, two decades ago, and since then it has been the topic of many research articles, mainly due to its effectiveness and unique behavior in regards to suppressing aggregation [15, 17, 69-76]. Arakawa and coworkers have written many research articles and reviews regarding the abilities and applications of ArgHCl, including its use during protein production and purification and as a solution additive for other macromolecules [9, 16, 71, 77-82]. Their collective work provides useful insight into the mechanism of action and highlights all of the observed behavior gathered thus far for ArgHCl. However, a clear mechanistic understanding has yet to be determined. Other researchers have conducted additional studies in an attempt to resolve this issue, but their results still lack a clear mechanistic understanding as well and often, the mystery is deepened by the discovery of more unique behavior [17, 20, 72-74].

As described in Chapter 2, what is known about ArgHCl is that it seems to inhibit protein-protein interactions with little influence on the conformational stability of the protein, thus slowing association rather than stabilizing the native structure [15, 18, 70]. Some researchers have attempted to explain this behavior, but no one theory has been generally accepted because most theories cannot explain all of the observed behavior, not to mention that there is often a lack of experimental evidence confirming the proposed mechanism of action [68]. A theory first proposed in our group (see Chapter 2), suggests that arginine is neutral in regards to preferential interactions, leading to little influence on the folding equilibrium but eliciting an entropic penalty for when two protein molecules aggregate due to the exclusion of arginine molecules from the gap formed between two associating protein molecules [15]. The preferential interaction coefficient values shown below seemed to, for the most part, support this theory [70]. However, at high concentrations, the preferential interaction of arginine goes from being neutral or slightly attracted to being highly excluded, and thus, we have to admit that this

model only describes a simple theoretical molecule, referred to as a "neutral crowder", and cannot explain all of the complex behavior ArgHCl exhibits, though the "neutral crowder" effect may play a large role [68, 76, 83]. To elucidate the mechanism of arginine, we began by performing a preferential interaction study, followed by molecular dynamic simulations, and an in depth stability study. Various arginine salts were studied as well, demonstrating that ion-ion interactions may be a contributing factor in the aggregation suppression mechanism.

4.1.1. Comparison between Γ_{μ_3}, Γ_{μ_1,μ_3}, and Literature Values

To evaluate the reliability of the results determined from VPO measurements, which is a relatively new technique, a comparison was made with data found in the literature (most of which was determined by dialysis/densimetry measurements, a well established technique). Table B-2 through Table B-4 (see Appendix B) summarize the preferential interaction coefficients over the concentration range detailed in the last column of Table B-1, for BSA, lysozyme, and aCgn, respectively, in combination with five common cosolutes (arginine hydrochloride, guanidinium chloride, glucose, glycerol, and urea). Theoretically, the preferential interaction coefficients Γ_{μ_3} and Γ_{μ_1,μ_3} are independent of protein concentration, which has been demonstrated elsewhere, thus there was no need to repeat the measurements at other protein concentrations [27, 84]. To make the comparison with literature values, Γ_{μ_1,μ_3} was computed using Eq. (3.19). All of the results shown in Table B-2 through Table B-4 support the assumption that Γ_{μ_1,μ_3} is approximately equal to Γ_{μ_3} because there is no significant difference between Γ_{μ_3} and Γ_{μ_1,μ_3} values for any protein-cosolute system tested. Therefore, a direct comparison can be made between Γ_{μ_3} values measured by VPO and Γ_{μ_1,μ_3} values measured by dialysis/densimetry.

In the literature, BSA is the most common model protein used in preferential interaction coefficient studies and therefore, there is more data available for comparison, which makes it the ideal protein to evaluate the reliability of the VPO methodology. For guanidinium chloride, glucose, glycerol, and urea, the trends between Γ_{μ_3} values and cosolute concentration for BSA computed from VPO measurements are nearly identical to those reported in the literature. This gives confidence that VPO is a reliable methodology for determining the preferential interaction coefficient. However, for BSA, the results for ArgHCl differ significantly from the literature values, as well as for lysozyme, thus making it unclear which values were most representative of

the preferential interaction coefficient. Thus a repeat of the dialysis/densimetry measurements was conducted to resolve this dilemma.

4.1.2. Comparison of VPO Results to Dialysis/Densimetry Measurements

The preferential interaction coefficient Γ_{μ_1,μ_3} for arginine was determined by dialysis/densimetry measurement techniques at two different arginine concentrations for both BSA and lysozyme (see Table B-5). For comparison, these results were plotted against the trends computed from VPO measurements in addition to the literature values, as depicted in Figure 4-1. It is clear that the results determined by dialysis/densimetry measurements conducted in our lab confirm the VPO results and bring into question the reliability of the literature values. For BSA (Figure 4-1a), both of the values measured by dialysis/densimetry follow the VPO trend, in particular, the preferential interaction coefficient for ArgHCl is close to zero at low concentrations and gradually becomes increasingly negative. This differs from the literature values, which depict the preferential interaction coefficient for arginine as being significantly greater than zero at low to moderate concentrations, and then decreasing and becoming increasingly negative at high concentrations.

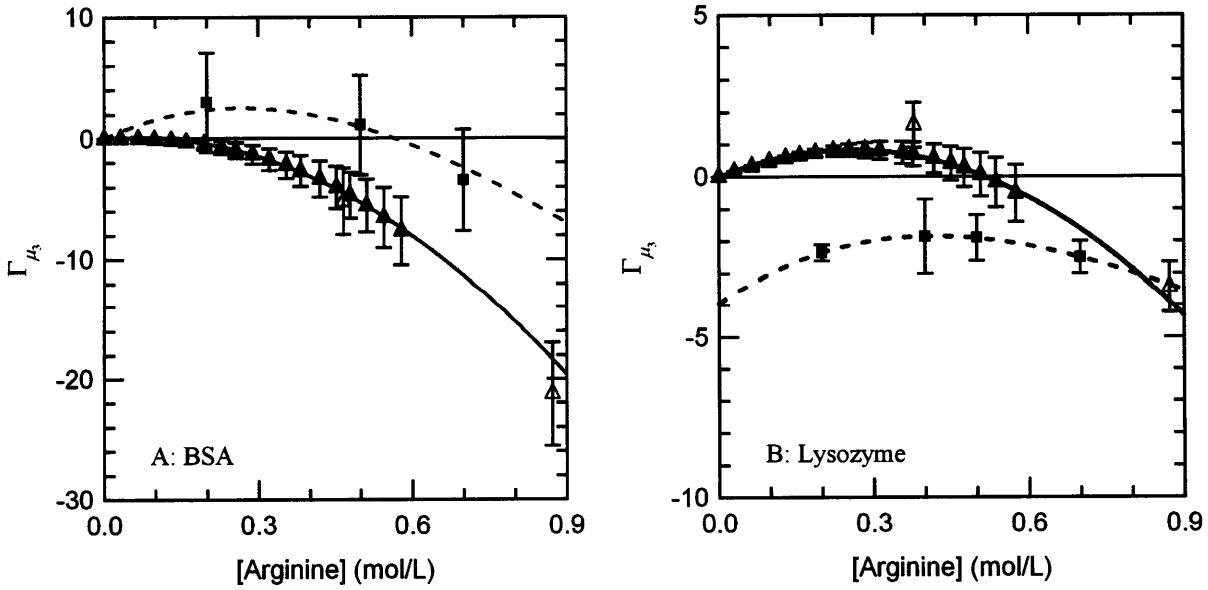

Figure 4-1: Comparison of Arginine Γ_{μ_3} values, obtained from VPO measurements (▲), dialysis/densimetry measurements (△), and the literature (taken from [85]) (■) for (A) BSA and (B) lysozyme.

For lysozyme (Figure 4-1b), the values determined by dialysis/densimetry confirm that the preferential interaction coefficient for ArgHCl is slightly positive at low concentrations and

decreases and becomes increasingly negative at high concentrations. This differs significantly from the literature values, which depict the preferential interaction coefficient for arginine being negative and somewhat constant over the entire concentration range tested. In addition the literature values seem to indicate the value for the preferential interaction coefficient is nonzero at zero arginine concentration, which is an unrealistic result.

Though the new results for BSA do not differ significantly from the literature, they do differ greatly for lysozyme. The significance of these new results is that they indicate a common trend in the interaction between arginine and protein molecules, as opposed to conflicting trends previously reported [85]. That is, arginine seems to be neither strongly bound nor attracted to the surface of the protein and becomes excluded only at high concentrations. This has never been observed before and is a result that characterizes arginine as unique among solutions additives studied.

4.1.3. Interpretation of Preferential Interaction Coefficient Values

These new results for the preferential interaction of arginine, as measured by VPO and dialysis/densimetry, differ significantly from literature values [85]. This alters the interpretation of how arginine interacts with protein molecules. Before, based on the literature values for BSA, many assumed that arginine was a bound additive (except in the case of lysozyme, where it was thought that the positively charged protein repelled the positively charged guanidinium moiety on the arginine molecule). The VPO results seem to indicate something different, and in the case for lysozyme, the opposite trend is observed. For BSA, in the concentration range of 0-0.3 M, Γ_{μ_3} is approximately zero and gradually becomes negative as the concentration increases. As for lysozyme, Γ_{μ_3} values seem to indicate that arginine is slightly bound for concentrations less than 0.5 M and then gradually becomes excluded as the concentration increases. However, for both proteins, Γ_{μ_3} values seem to indicate that arginine is neither strongly bound nor excluded from the surface of the protein for arginine concentrations less than 0.5 M but becomes excluded at higher concentrations.

This trend is further exemplified when the preferential interaction of arginine is compared to denaturing and stabilizing cosolutes. In Figure 4-2, a comparison has been made with urea and glycerol for all three proteins tested and it shows that preferential interaction coefficient values for arginine typically fall in between values for urea and glycerol.

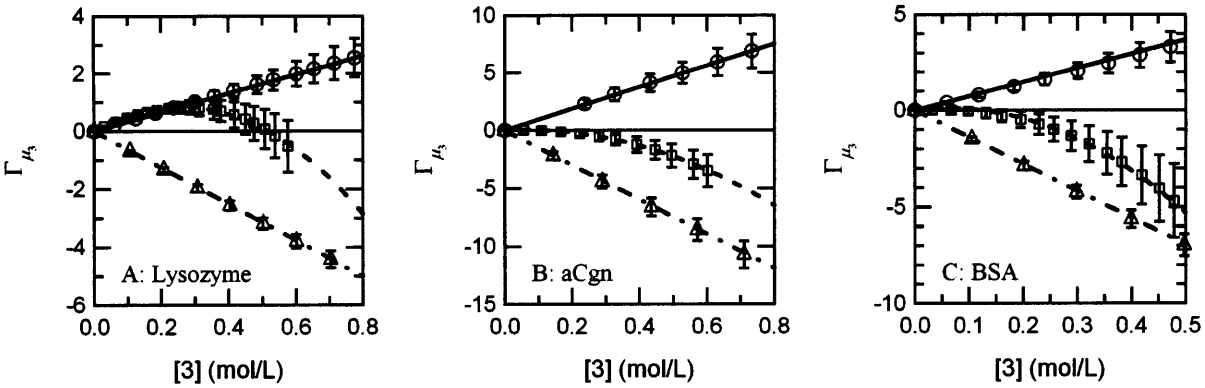

Figure 4-2: Comparison of arginine preferential interaction coefficient values (□) to that for urea (○) (a preferentially bound and denaturing cosolute) and glycerol (△) (a preferentially excluded and stabilizing cosolute) for all three proteins tested; (A) lysozyme, (B) aCgn, and (C) BSA.

Our preferential interaction measurements seem to indicate that arginine has little interaction with proteins. However, these measurements indicate only how a cosolute interacts with a protein in its native state, but assuming that the trend holds for the unfolded state, arginine should have no effect on the unfolding equilibrium because it does not significantly perturb the free energy of either state. However, some results seem to indicate that arginine has a very slight destabilizing effect at high concentrations for some proteins (including lysozyme) [86]. This effect, though, is very minor when compared to guanidinium chloride (1 M arginine HCl decreases the melting temperature by about 1 °C for lysozyme versus a decrease of about 13-17 °C for guanidinium chloride at a similar concentration). Those results seem to indicate that arginine has a very slight affinity for the unfolded or partially unfolded state but not enough to have a significant effect on the folding equilibrium.

4.1.4. Relationship of Arginine Preferential Interaction with Concentration

An interesting trend can be observed when the arginine preferential interaction coefficients for all three proteins tested are plotted together. Figure 4-3 reveals that not only is there a unique trend in the preferential interaction coefficient for arginine with concentration but there is also a trend with protein size as well. Γ_{μ_3} values for α-Chymotrypsinogen (a 25.7 kDa protein) fall in between the values for lysozyme (a 14.3 kDa protein) and BSA (a 66.4 kDa protein). Since there is no clear trend with the charge of the protein (*i.e.* the positively charged arginine seems to be most attracted to the positively charged proteins), these results seem to

indicate that as the protein size increases, arginine molecules become increasingly excluded from the protein surface.

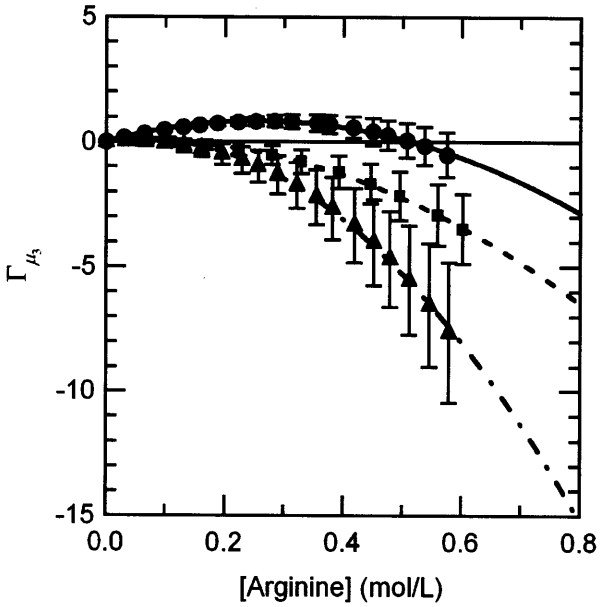

Figure 4-3: Comparison of arginine preferential interaction coefficient values for all three proteins tested demonstrating the relationship with arginine concentration and protein size; lysozyme, 14.3 kDa (●), aCgn, 25.7 kDa (■), and BSA, 66.4 kDa (▲).

The reason for this behavior is not quite clear from this data. It has been shown that arginine interacts with amino acid residues in almost the same fashion and to almost the same extent as guanidinium, with both compounds showing a strong affinity for most residues, especially aromatic residues [9]. This behavior may explain why arginine shows a slight affinity for proteins and a slightly stronger affinity for unfolded proteins, since aromatic residues are typically buried in a folded protein. However, surface tension also increases with increasing arginine concentration, much more so than with guanidinium [9]. Such behavior tends to repel additives from the protein surface and enhance the binding of water, since the protein surface constitutes a solvent interface at which there must be interfacial tension. This, in addition to the fact that arginine is a much larger molecule than guanidinium (*i.e.* the solvation layer surrounding a protein molecule that is inaccessible to the additive but accessible to water increases as the size of the additive increases) further increases the relative repulsion of arginine with respect to protein size. Thus, the attraction the guanidino residue of arginine has for the protein surface plus the repulsion generated by the greater surface tension increment and the greater size may account for why arginine is neither strongly attracted nor strongly excluded

from the protein at low to moderate concentrations. In other words, arginine experiences a balance of attraction and repulsion from the protein surface. Also, the fact that arginine becomes increasingly excluded as the size of the protein increases and is not strongly attracted to the unfolded state, indicates that the repulsion effects are more dominant than the attractive effects since guanidinium tends to be more attracted to proteins as the size increases (*i.e.* there is more surface area to bind to as the size of the protein increases).

However, the attraction that guanidinium residues have for a protein and the repulsion generated by surface tension increments and volume exclusion tend to be linear with cosolute concentration at low concentrations. Therefore, if the affinity and repulsion experienced by arginine is additive, then the preferential interaction coefficient should display a linear trend with respect to arginine concentration. However, this is clearly not the trend shown in Figure 4-3. If arginine has a similar affinity for the protein as guanidinium but at the same time, such interactions are hindered by surface tension and steric exclusion effects, then it might be possible for the protein surface to become saturated with arginine at a low concentration if the protein does not unfold to allow for more binding. If this happens, any additional arginine added to the solution upon saturation would be merely portioned away from the local domain around the protein.

The phenomenon exhibited in Figure 4-3 was furthered studied within the group through the use of molecular dynamic simulations. The work of Diwakar Shukla shows that the varying level of interaction arginine exhibits with different proteins is due to the number of interacting groups present on the surface of the protein, not with the size of the protein. Lysozyme has more amino acids present on its surface which favorably interact with arginine, thus the higher level of interaction [83]. Furthermore, the hypothesis of arginine saturating the surface was found to be only partially true. Only at high concentrations (>1.5 m) does the surface become saturated and therefore, saturation cannot explain the trend observed at moderate concentrations. MD simulations revealed that arginine tends to form clusters in solution, even at moderate concentrations, which tend to reduce preferential binding [76]. A more in depth explanation is provided below.

4.2. Arginine and the Hofmeister Series

4.2.1. Background on Ion-Ion Interactions

When reviewing the literature regarding arginine and protein aggregation, it becomes apparent that little attention has been devoted toward other salt forms besides chloride. The chloride form has been studied so exclusively that ArgHCl is often referred to simply as arginine [68]. Given how much is known about the influence various anions have on protein stability, a study of other salt forms seemed necessary. The Hofmeister Series,

$$F^- > SO_4^{2-} > HPO_4^{2-} > Acetate^- > Citrate^- > Cl^- > NO_3^- > Br^- > I^- > SCN^-$$

$$NH_4^+ > K^+ > Na^+ > Cs^+ > Li^+ > Mg^{2+} > Ca^{2+} > Ba^{2+} > Gdm^+$$

first published in 1888, ranks common ions with respect to their ability for precipitating proteins (greatest precipitation ability (lowest protein solubility) on the left, lowest precipitation ability (highest protein solubility) on the right), which in turn is related to the thermodynamic stability of the native state (with a trend opposite of protein solubility) [87]. The series has been studied extensively since the time of Hofmeister and serves as the foundation for selecting salts that will influence protein solubility, crystallization, denaturation, and aggregation, though the mechanism for this behavior is not well understood [88-91]. It has been clearly demonstrated that the choice of anion has a more pronounced effect on these properties than the choice of cation. Furthermore, chloride is often placed in the middle of the series, thus other arginine salt forms are likely to exhibit a drastically different behavior, both positively and negatively, in regards to protein stability [91, 92].

From an extensive literature search, the only other form of arginine studied in this context has been $ArgH(SO_4)_{1/2}$, though the extent of those studies is somewhat limited [93, 94]. Recently though, the combination of arginine hydrochloride with sodium glutamate was studied by two different research groups and has received much attention due to a seemingly synergetic effect when the two solutes are combined in an equimolar ratio [3, 95, 96]. Moreover, in a recently published review, Lange and Rudolph (Rudolph being one of the researchers who discovered the arginine effect) postulated, based on existing information, on how other arginine salt forms would behave as excipients and also commented on the refolding behavior of the sulfate and phosphate salt forms, stating that those two forms of arginine perform poorly as refolding excipients even though sulfate and phosphate often provide conformational stabilization [68].

However, no citations or experimental results were given to support those findings. Even though sulfate and phosphate may reduce the refolding abilities of arginine, we have found that they enhance its aggregation suppression ability when a model protein is aggregated under accelerated conditions. Moreover, switching to a thiocyanate salt induces rapid aggregation under similar conditions rather than enhancing aggregation suppression, opposite of what we and Lange and Rudolph hypothesized.

When attempting to explain the arginine mechanism, many researchers start by addressing the effect of the guanidinium moiety [86, 94]. Indeed, the guanidinium functional group must play an important role, because not only is arginine the only amino acid with a guanidinium moiety, no other amino acid exhibits aggregation suppression characteristics quite like arginine, though other amino acids, particularly proline, have been used as stabilizers [47, 97, 98]. However, without a clear understanding at the molecular level, not only of the interaction of arginine with protein molecules, but also of how arginine behaves in solution, the arginine mechanism will remain a mystery.

Recently though, two key revelations, in relation to the solution behavior of guanidinium and arginine, may help to resolve this perplexing question. The first comes from a series of articles published by Mason and coworkers, in which they studied the solution behavior of various guanidinium salts [99-104]. They initially aimed to explain the denaturing ability of guanidinium but ultimately revealed that guanidinium can form a strong attractive interaction with various anions of similar size that also have hydrogen bond accepting capabilities (*e.g.* sulfate, carbonate, etc.), which likely explains why some guanidinium salts (*e.g.* guanidinium sulfate) stabilize proteins rather than denature them because the two ions form clusters in solution rather than bind to the protein surface [103]. The hydrogen bonding interactions and cluster formations they demonstrated reveal a lot about the behavior of guanidinium. However, some of this behavior was revealed long before the work of Mason and coworkers. Bonner, using osmotic coefficient data and Raman spectroscopy, gave valuable insight into the hydrogen bond interactions between guanidinium and halide ions over three decades ago [105, 106]. Bonner showed that guanidinium and fluoride form a strong hydrogen bond, possibly even a partial H-F bond, in solution and this interaction is weaker for other halides in the order of $Cl^- > Br^- > I^-$, due to ion size and charge density. At the time though, these results were used to speculate on the hydrogen bond formation between guanidinium and proteins. More importantly, Mason and

coworkers also revealed that the structure of the poorly hydrated guanidinium molecule allows the positively charged ions to associate, with the planar molecules stacked on top of each other [100]. These results contradict long held beliefs about how certain cosolutes influence the structure of water (solutes have long been labeled either a chaotrope or a kosmotrope because of this water structure viewpoint) and how the perturbed water structure influences protein stability. [91]. However, their results are compelling and have strong implications for explaining the arginine mechanism, not only with how arginine will interact with various counterions, but how an arginine molecule interacts with other arginine molecules (and glutamate), given that guanidinium should have a strong interaction with the carboxylate moiety found on both molecules.

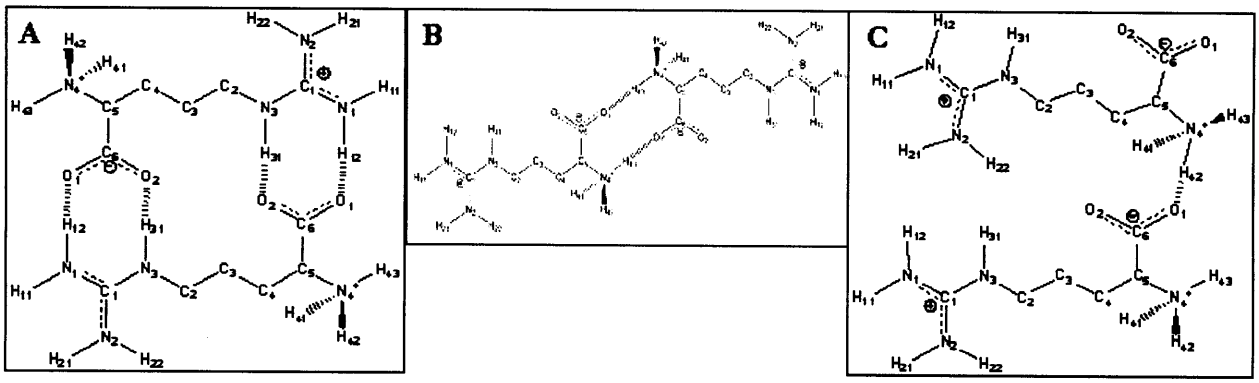

Figure 4-4: Arginine dimer formation. (A) Head-to-Tail orientation in which the guanidinium moiety hydrogen bonds with the carboxylate moiety of a nearby arginine molecule, (B) Tail-to-Tail orientation in which the carboxylate moiety hydrogen bonds with the amine moiety of a nearby arginine molecule, and (C) Head-to-Head orientation in which the faces of the guanidinium moiety stack on top of each other and the carboxylate moiety of one molecule hydrogen bonds with the amine moiety of the other molecule. Figure and results produced by Diwakar Shukla from MD simulations.

This self association of arginine (see Figure 4-4), the second revelation mentioned, has previously been revealed and has been studied extensively in our lab through molecular dynamics (MD) simulations [72, 76, 83, 107]. These simulations show that the guanidinium functional group does indeed interact with the surface of a protein molecule in a fashion similar to free guanidinium; however, this functional group also forms attractive interactions with other arginine molecules, particularly with the carboxylate moiety. This self association of arginine molecules likely prevents arginine from binding too strongly to the surface of protein molecules, preventing it from denaturing the protein but at the same time preventing protein-protein interactions. Support for this claim comes from the observation by us and others that if the carboxylate moiety is not present, the resulting molecule denatures proteins much like

guanidinium [68]. This self association also explains the unique trend observed for preferential interaction coefficient measurements (*i.e.* slightly bound at low concentrations but highly excluded at high concentrations) and explains why the arginine effect diminishes at high concentrations [70]. Therefore, we were motivated to explore other arginine salt forms, not only with the aim of discovering a better performing excipient, but to further our understanding of this mechanism and how guanidinium-anion interactions may contribute to the observed behavior.

4.2.2. Aggregation Suppression

The aggregation suppression exhibited by various arginine salt forms was tested and an example of the data collected during an accelerated aggregation experiment is shown in Figure 4-5a. In the figure, the aCgn monomer loss profiles for four different arginine salt solutions are shown (with all arginine salts at a concentration of 150 mM) in addition to the reference profile (*i.e.* no cosolute). It can be seen that ArgH(H_2PO_4), ArgH(Citrate)$_{1/2}$, and ArgHCl reduce the rate of aggregation, while ArgHSCN increases it. The rate of monomer loss in the presence of ArgHCl at 150 mM is approximately twice as slow as the reference monomer loss rate. However, at the same concentration, ArgH(Citrate)$_{1/2}$ and ArgH(H_2PO_4) slow the rate by a factor of approximately 3 and 8, respectively. On the other hand, ArgHSCN at 150 mM increases the rate by a factor of 6. This, however, only represents a fraction of the data collected, which is best represented in Figure 4-5b. The figure shows the relative rate constant for monomer loss (*i.e.* the rate constant relative to the rate constant for no cosolute) over a wide range of cosolute concentrations. It should be noted that the profile for ArgH(H_2PO_4) is limited due to its poor solubility (approximately 0.45 M at room temperature).

The results shown in Figure 4-5b demonstrate that phosphate, sulfate, and citrate significantly improve the aggregation suppression ability of arginine when compared to chloride. More importantly, at low concentrations (less than 150 mM), the aggregation suppression ability of ArgH(H_2PO_4) is significant. At a concentration of 150 mM, the rate of aCgn aggregation in the presence of ArgH(H_2PO_4) is 2.5 times slower than when in the presence of either ArgH(SO_4)$_{1/2}$ or ArgH(Citrate)$_{1/2}$. However, the aggregation suppression ability of ArgH(H_2PO_4) levels off significantly as the concentration is increased beyond 200 mM and the aggregation suppression ability of ArgH(SO_4)$_{1/2}$ and ArgH(Citrate)$_{1/2}$ likely exceeds it at concentrations above 300 mM. As will be discussed below, this trend seen for ArgH(H_2PO_4) is likely the result of the strong interaction between the guanidinium moiety of arginine and the phosphate ion.

These interactions seem to cause arginine-phosphate clusters to form at higher concentrations, leading to high preferential exclusion and a strong thermodynamic stabilization but also limited cosolute solubility and reduced association suppression.

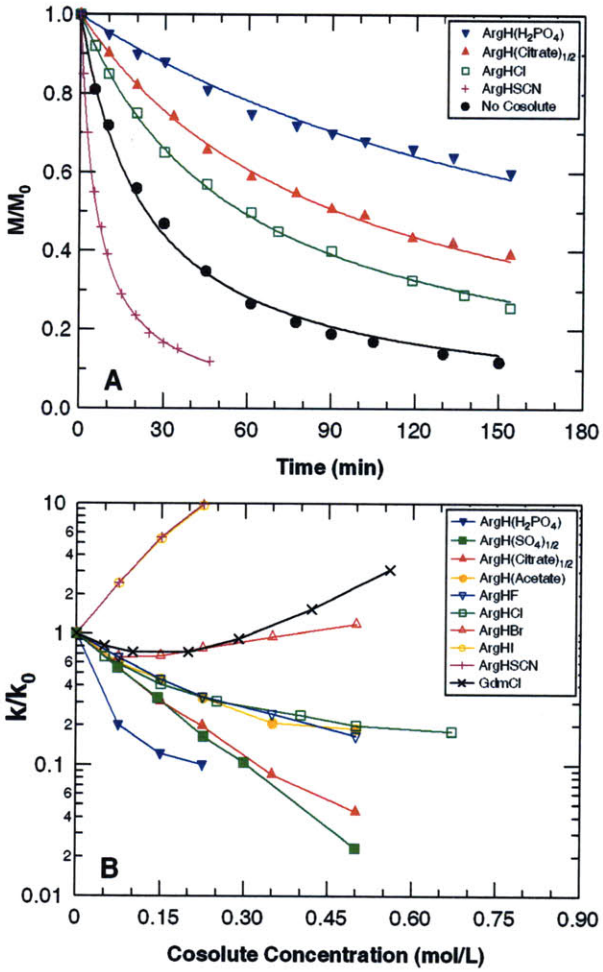

Figure 4-5: The influence of arginine salts on aCgn monomer loss due to aggregation at 52.5°C. All solutions contained 10 mg/mL aCgn and were prepared in a 20 mM sodium citrate pH 5 buffer. (A) Monomer concentration, M, normalized with respect to the initial monomer concentration, M_0, versus time from a single experiment, with plots fitted to a second order rate law. The concentration of each arginine salt shown was 150 mM. (B) aCgn monomer loss rate constant in the presence of a cosolute, k, relative to no cosolute, k_0, versus cosolute concentration, with lines drawn through the plots to aid the eye.

As for the remaining arginine salts to the left of chloride in the Hofmeister Series, the relative rate constant for fluoride and acetate are essentially the same as that for chloride. At concentrations below 100 mM, the aggregation suppression ability of these three salt forms are comparable to sulfate and citrate. However, as the concentration increases, the aggregation suppression ability of sulfate and citrate continue to improve, having a relative rate constant value between 0.02-0.04 at the highest concentration tested (0.5 M), while the relative rate

constant for fluoride, acetate, and chloride levels off, having a near constant value of about 0.2 for concentrations above 0.4 M. ArgHF seems to level off the least and has a slightly better relative rate constant value at a concentration of 0.5 M, which will likely improve slightly as the concentration is increased further, however, such concentrations are considered quite high and would rarely ever be used.

As for the arginine salts to the right of chloride in the Hofmeister Series, relative rate constant values for bromide remain fairly close to unity for all concentrations tested, while iodide and thiocyanate have nearly identical profiles, which show a large increase in the rate of aggregation, up to 10 times faster at a concentration as low as 200 mM. The relative rate constant for ArgHBr dips down to around 0.7 at 150 mM before reversing and exceeding a value of 1 at higher concentrations. This indicates that the slight denaturing effect of bromide balances out the aggregation suppression effect of arginine, while the strong denaturing effect of iodide and thiocyanate completely overpower the effect of arginine. The large size and low charge density of these ions lead to them being poorly hydrated and therefore, in thermodynamic terms, they tend to partition to the surface of a protein rather than being solvated in the bulk solution, even more so if hydrophobic residues are exposed [99]. This favorable preferential interaction leads to protein denaturation, with iodide and thiocyanate having a stronger preferential interaction than bromide due to their smaller charge density. In fact, iodide and thiocyanate are so destabilizing that ArgHI and ArgHSCN are stronger denaturants than guanidinium chloride (GdmCl), as represented by relative rate constant values (see Figure 4-5) and the denaturation midpoint temperatures (see discussion below). The relative rate constant for GdmCl initially decreases when it is added to the solution (similar to ArgHBr), reaching a minimum of about 0.6 at 200 mM and then increases there after, reaching a value of 9 at 0.7 M, the highest concentration tested. This shows that at low concentrations, GdmCl has an ability to reduce aggregation by inhibiting protein-protein interactions, similar to arginine, but at high concentration, the denaturing effect of this compound dominates and causes aggregation to proceed more rapidly. However, ArgHI and ArgHSCN show no stabilizing effect from the presence of arginine. They instead increase the rate of aggregation at all concentrations tested, doubling the rate at a concentration as low as 50 mM. All of this exemplifies the pronounced effect the choice of anion has on protein stability and shows that ArgH(H_2PO_4), ArgH$(SO_4)_{1/2}$ and ArgH$(Citrate)_{1/2}$ might be more useful than ArgHCl for preventing protein aggregation.

4.2.3. Conformation Stability

The thermodynamic stability of aCgn in the presence of the arginine salts was assessed by determining the change in the denaturation midpoint temperature, T_m, with respect to cosolute concentration. Table B-6 shows the denaturation midpoint temperature increment, $dT_m/d[3]$, defined as the slope of T_m with respect to the molar concentration of the cosolute (component 3), over the concentration range tested, as determined by DSC. This parameter allows for a comparison to be made amongst the different cosolutes since it describes the rate at which T_m changes per amount of cosolute added. Over this short initial concentration range, T_m exhibited a linear trend, giving constant values for $dT_m/d[3]$. The values obtained exhibit the pronounced effect the anions have on the thermodynamic stability of the protein and they correspond directly to ordering of the relative rate constants obtained during the aggregation study.

However, as has been mentioned above, conformational stabilization is not the only thing to consider when assessing how a cosolute influences aggregation. From the simple model provided in Chapter 2, it is clear that there are two types of additives which will deter aggregation: conformational stabilizers and association suppressors. As shown in Table B-6, $ArgH(H_2PO_4)$, $ArgH(SO_4)_{1/2}$ and $ArgH(Citrate)_{1/2}$ increase the melting temperature of the protein at a rate of 6.3, 5.4, and 3.3 $^oC*M^{-1}$, respectively. Clearly, these salts are able to stabilize the native structure of the protein and inhibit the formation of partially unfolded species. This thermodynamic stabilization is on top of the association suppression effect of arginine, leading to the significant aggregation suppression shown in Figure 4-5. However, ArgH(Acetate), ArgHF, and ArgHCl exhibit no thermodynamic stabilization, which is consistent with previous reports for ArgHCl, thus these salts inhibit aggregation solely through an association suppression mechanism. ArgHBr, on the other hand, has a denaturing effect on aCgn, lowering the melting temperature at a rate of 7.7^oC*M^{-1}. This reduction in conformational stability, which increases the number of reactive species in solution, seems to nearly balance out the association suppression effect of arginine, as exhibited by the relative rate constant having a value near unity over a wide concentration range. This reduction in thermodynamic stability is only slightly greater than that for GdmCl, which might explain why they have near identical profiles at low concentrations. However, the association suppression effect of arginine seems to be greater than that for guanidinium, thus, at high concentrations, the rate of aCgn aggregation is higher in the presence of GdmCl than in the presence of ArgHBr. As for the other arginine salts, ArgHI and

ArgHSCN lower the melting temperature significantly, both with a rate greater than 20 $^{o}C*M^{-1}$. This denaturing effect overpowers the association suppression effect of arginine and as a result, these salts increase the rate of aggregation at all concentrations. These results, along with aggregation suppression data, support the claim that $ArgH^{+}$ acts as an association suppressor and shows that the choice of anion will either enhance or counteract this aggregation suppression by influencing the stability of the native structure.

4.2.4. Ion-Ion Interactions

As mentioned above, ion-ion interactions may be contributing to the observed behavior of the arginine salts. To quantify this behavior, osmotic 2nd virial coefficient values were determined for aqueous solutions of the arginine salts via VPO and MD simulations. Experimental osmotic data were analyzed using the Pitzer model, while MD simulations were analyzed using the McMillan-Mayer model. Pitzer ion interaction parameters for each of the salts are shown in Table B-6. The second virial coefficient parameters obtained for ArgHCl closely matches values reported in the literature. However, the value for the third virial coefficient differs. This is likely due to the narrow and low concentration range used in this study. The third virial coefficient is typically small and is only important at high concentrations. Moreover, much of the time, it can be excluded from the model with little error. Therefore, there is greater uncertainty in the third virial coefficient values reported here and they are not likely to accurately represent trends at high concentrations. However, the second virial coefficient parameters reported are likely to be a fair representation of behavior in the concentration region of interest.

Since the magnitude of like and unlike ion interactions is ionic strength dependent, it is generally held that only cation-anion binary interactions make contributions to the value of $\beta^{(1)}$ (see Eq. (4.4)). Both like and unlike ion interactions make contributions to $\beta^{(0)}$, making it difficult to interpret this parameter unless radial distribution functions are available for each ion pair. The trends observed for $\beta^{(1)}$ are in accordance with the discussion above. The values for the phosphate, sulfate, and citrate salts are large and negative, indicating a strong interaction between arginine and those anions. The value for ArgHCl is negative but smaller in magnitude, indicating a weak attractive interaction. The values for the other halide salts show repulsive interactions for those salt forms with the strongest repulsion for thiocyanate and iodide and the weakest for bromide. These trends are similar to those for guanidinium halide salts, which show

a weak repulsive interaction with fluoride, a moderate repulsive interaction with bromide, and a strong repulsive interaction with iodide [67].

During MD simulations performed by Diwakar Shukla, significant ion pairing was observed in the aqueous arginine salt solutions, which supports the above experimental results. Ion-pairs in thiocyanate, chloride and acetate solutions are observed to be randomly distributed throughout the solution while citrate, phosphate and sulfate show a marked tendency to form hydrogen bonded clusters. Snapshots of the MD simulation boxes of arginine salts are shown in Figure 4-6. It can be seen that the structures formed by sulfate, citrate and phosphate salts are worm-like chains of sizes comparable to the box-size. Clustering observed in chloride, thiocyanate and acetate solutions is mainly due to Arg-Arg interaction, as reported in our recent work on aqueous arginine hydrochloride solutions [76, 108]. Similar ionic clusters were also observed in aqueous guanidinium salt solutions [100-102]. Brady and coworkers have shown that clustering in Gdm salts is not dependent on the choice of the water model and is not a simulation artifact. Due to the presence of Gdm group as a side-chain in arginine, the clustering behavior of arginine salts is expected to be similar to Gdm salts. The only difference between Gdm and arginine salts is the presence of additional charged groups and a carbon chain in arginine. These additional charged groups (N-terminal amino group and C-terminal carboxylate group) seem to further enhance the clustering in solution.

Radial distribution functions (Figure 4-7) between pairs of ions help to further characterize the solution structure. Arg-Arg pairing, as illustrated by the RDF between the Gdm carbon atoms in the arginine molecule (Figure 4-7a), shows a tendency of the Gdm side chain to stack with each other or form ion-pairs separated by the anion acting as a bridge between two Gdm side chains. The presence of an anion does not significantly affect the Arg-Arg pairing except in the case of citrate, where the large size of the anion crowds out the formation of Arg pairs. The pairing between cation and anion is strongest for the sulfate ion due to the exceptionally strong interaction with the Gdm group (Figure 4-7b). Acetate also has a strong interaction with arginine; however, the interactions for thiocyanate and chloride are weak and almost nonexistent. The pairing between anions (Figure 4-7c) is significantly weak, as compared to other ion pairs in solution. Phosphate ion pairing exhibits a peak around 0.5 nm due to the presence of hydrogen bonding acceptors and donors in the same molecular anion. Sulfate also shows a peak around 0.6-0.7 nm but this is not due to direct interaction between sulfate ions;

rather, the peak is due to the presence of multiple sulfate anions associating with adjacent or the same arginine cation.

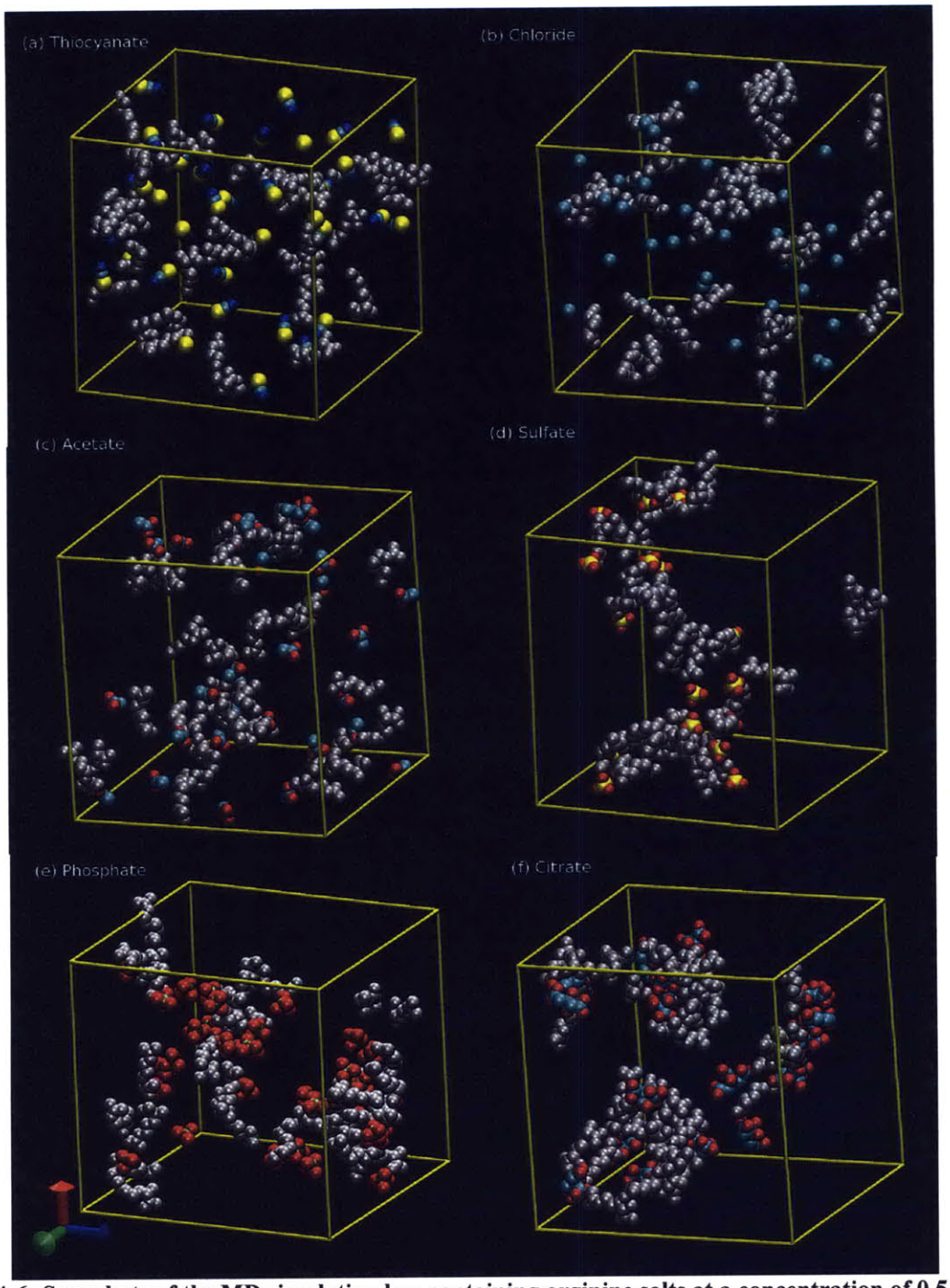

Figure 4-6: Snapshots of the MD simulation box containing arginine salts at a concentration of 0.5 mol/kg. To improve the clarity of the image, water molecules are not shown and only heavy atoms (all atoms excluding hydrogen) in the arginine molecules and counter-ions are shown. The following color code is used to represent atoms: C (cyan), O (red), N (blue), S (yellow), Cl (light blue), and P (brown). Arginine molecules are shown in silver. Figure and results produced by Diwakar Shukla.

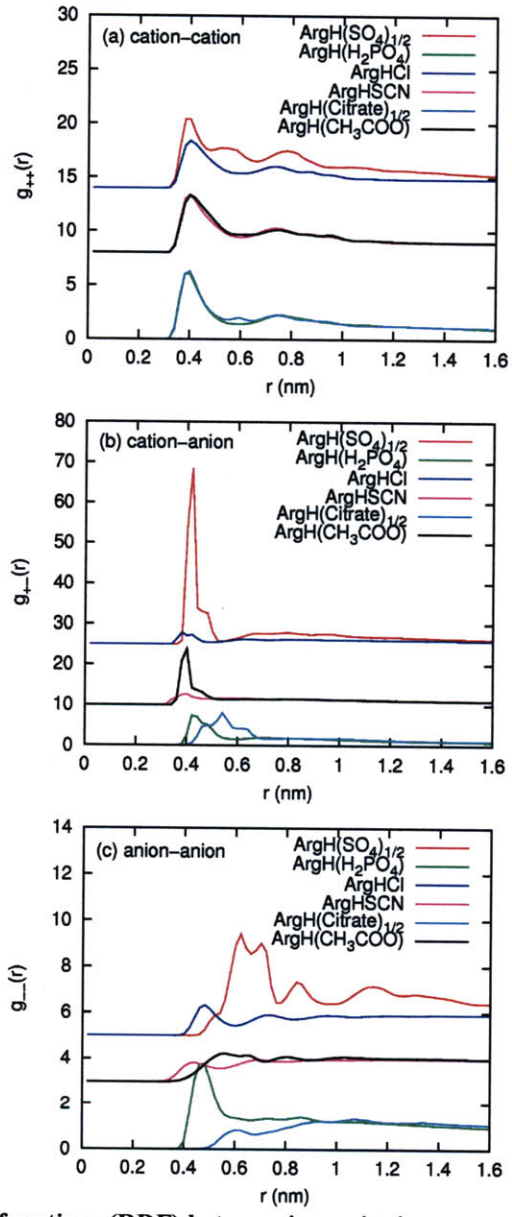

Figure 4-7: Radial distribution functions (RDF) between ion-pairs in aqueous arginine salt solutions. Cation-Cation RDF is the RDF between guanidinium carbon atoms of arginine. For the counter ions, the atoms used as centers for estimating the RDF's are: Sulfate – Sulfur atom, Phosphate – Phosphorus atom, Citrate - Central carbon atom, Thiocyanate – Nitrogen atom and Acetate – Carboxylate carbon atom. Figure and results produced by Diwakar Shukla.

The individual contribution of specific ion-pairs to the overall clustering in the solution can be estimated by calculating the osmotic second virial coefficient between ion-pairs in solution from MD simulation results. B_{22} values for ion-pairs in each salt solution are reported in Table 4-1, which show a similar trend to the experimentally obtained Pitzer coefficient data. It can be seen that the contribution of cation-cation interaction is similar in magnitude to cation-anion interaction. Σ_{ij} denotes the sum of interactions between all ion pairs. Sulfate, phosphate

and citrate form a group with strong overall interactions whereas chloride, thiocyanate and acetate form a group with weak overall interactions. The difference between the collective structures of these two groups of anions apparently results from the differences in the hydrogen bonding of the anions to the arginine cation.

Table 4-1: McMillan-Mayer Second Virial Coefficient, B_{22} (L/mol), values for ion pairs in aqueous arginine salt solutions. Results produced by Diwakar Shukla.

Arg Salt	C (mol/L)	B_{22} ++	B_{22} +-	B_{22} --	Σ_{ij}
$(SO_4)_{1/2}$	0.49	-6.58	-4.63	-4.89	-20.73
$(Citrate)_{1/2}$	0.48	-2.13	-3.24	-0.35	-8.96
(H_2PO_4)	0.48	-3.20	-2.00	-1.60	-8.80
(Acetate)	0.48	-1.80	-1.25	-0.2	-4.50
Cl	0.48	-0.99	-0.90	0.45	-2.35
SCN	0.48	-1.40	-0.90	0.40	-2.80

The numbers of hydrogen bonds between different ion-pairs in solution, as determined from MD simulations, are reported in Table 4-2. It can be seen that the highest number of hydrogen bonds between arginine molecules are formed in chloride and thiocyanate solutions due to the minimal interaction of these ions with the cation. The loss of hydrogen bonds between arginine molecules in acetate, sulfate, phosphate and citrate solutions is compensated by the formation of large number of cation-anion hydrogen bonds. Acetate also interacts strongly with the arginine, forming hydrogen bonds with the Gdm group as shown in Figure 4-8a, but due to the limited number of hydrogen bond acceptors, the acetate anion cannot act as a bridge between multiple cations. On the other hand, citrate, sulfate and phosphate can form bridged structures, as shown in Figure 4-8b, c, and d, respectively. The number of hydrogen bonds between arginine ions (with an anion acting as a bridge) is high for sulfate, phosphate and citrate as compared to acetate, chloride and thiocyanate, which have limited to negligible capacity to form such bridged structures.

These observations also explain why sulfate, phosphate and citrate form large hydrogen bonded clusters in solution as compared to acetate, chloride and thiocyanate. These results are in accordance to our previous reports, in which we showed that ArgHCl likely forms Arg-Arg clusters in solution at high concentration, which leads to a higher preferential exclusion. Moreover, according to the "Gap Effect" theory developed in our group, this clustering reduces the ability of individual ArgHCl molecules to inhibit protein-protein association, however, this is

likely compensated for by the association suppression ability of the large clusters [10, 15, 72, 76, 83].

Table 4-2: Number of hydrogen bonds between different ions in aqueous arginine salt solutions. Results produced by Diwakar Shukla.

Arg Salt	C (mol/L)	Cation-Cation	Cation-Anion	Anion-Anion	Anion-Arg-Anion
$(SO_4)_{1/2}$	0.49	42.7	51.3	0	67.3
$(Citrate)_{1/2}$	0.48	35.8	31.7	0.4	28.2
(H_2PO_4)	0.48	40.6	30.1	11.7	19.5
(Acetate)	0.48	29.3	30.6	0	8.3
Cl	0.48	49.4	6.3	0	0
SCN	0.48	38.2	4.0	0	0

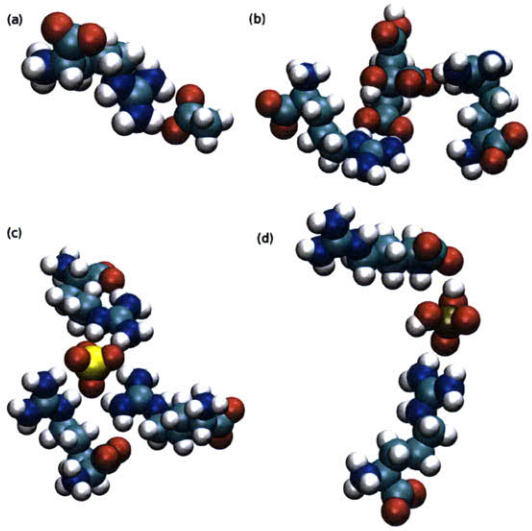

Figure 4-8: Hydrogen bonding interaction between arginine and (a) acetate, (b) citrate, (c) sulfate and (d) phosphate anions. It can be seen that sulfate, phosphate and citrate can interact with multiple arginine molecules forming large hydrogen-bonded structures. The following color code is used to represent atoms: C (cyan), O (red), N (blue), S (yellow), and P (brown). Results produced by Diwakar Shukla.

4.2.5. Preferential Interactions

To gain insight into how the arginine salts inhibit protein-protein interactions, the preferential interaction coefficient, Γ_{μ_3}, at various concentrations was determined, both experimentally via VPO measurements and computationally via MD simulations. The experimental results are depicted in Figure 4-9, which shows Γ_{μ_3} values at various cosolute concentrations, and Table B-6, which summarizes the polynomial fit and uncertainty of the experimental data. It should be noted that the preferential interaction coefficient could not be obtained via VPO for the acetate and fluoride salts due to these anions being in equilibrium with their respective conjugate acid forms at pH 5. This is because both of these acid forms are

volatile and this volatility (even though it is slight due to the low concentration of the conjugate acid) interfered with the VPO measurements.

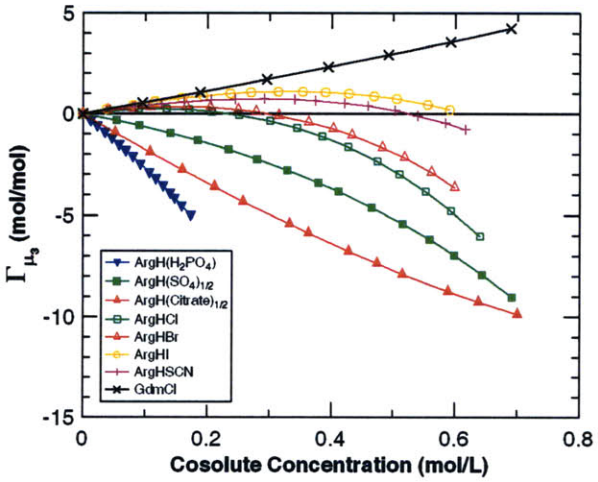

Figure 4-9: Preferential Interaction Coefficient, Γ_{μ_3}, values for the interaction between arginine salts (and guanidinium chloride for comparison) and aCgn, as determined from VPO measurements. Error bars left off for clarity and curves drawn through the plots to aid the eye (see Table B-VI for more detail).

Regardless, the experimental results for the other salts show a clear difference between the anions to the left and to the right of chloride in the Hofmeister Series. The chloride, bromide, iodide, and thiocyanate salts all have preferential interaction coefficient values that start out positive at low concentrations, with an initial slope almost identical to GdmCl, whereas the initial values for the citrate, sulfate, and phosphate salts are negative. The preferential interaction data for GdmCl is depicted in Figure 4-9 and Table B-6 as well for comparison, which exhibits a linear trend with respect to concentration.

The initial slope for the chloride, bromide, iodide, and thiocyanate salts show that at low concentrations, when there is likely little ion-ion interactions, the salts interact favorably with the protein surface, with an attractive interaction similar in strength as that for GdmCl. However, as the concentration increases, the preferential interaction coefficient values for all of these arginine salts begin to decrease and ultimately become negative, whereas the trend for GdmCl continues to increase linearly, indicating that at high concentrations, these arginine salts exhibit a net repulsive interaction with the native protein structure, opposite of what is observed for GdmCl. Furthermore, the preferential interaction coefficient values for thiocyanate and iodide salts exhibited the largest values at all concentrations, followed by bromide and chloride, indicating that the thiocyanate and iodide ions cause those salts to have a stronger attractive interaction with the protein surface, followed by bromide, which causes only a slightly stronger interaction.

This nonlinear trend is in accordance with previously reported values for ArgHCl and aCgn, with values differing slightly due to a difference in pH [70]. In our previous report, the surprising behavior exhibited by ArgHCl was initially thought to be the result of the protein surface quickly becoming saturated with the large molecule. MD simulations later revealed that this hypothesis was likely true but only at much higher concentrations (>1.5 m) and thus another mechanism contributed to the nonlinear trend observed at lower concentrations [76, 83]. As mentioned before, studies of aqueous ArgHCl solutions revealed that arginine has a strong propensity to form clusters in solution via hydrogen bonding between the guanidinium and carboxylate moieties, which was found to be stronger than the hydrogen bond formed with water, and as a result, it reduces the interaction with the protein surface. This reduced interaction is due, in part, to the larger net size of the clusters (steric exclusion) but in large part, is due to a favorable interaction with other arginine molecules in the bulk solution, which is similar or greater in strength as the interaction with the protein surface. This self interaction is concentration dependent, with more clustering obviously occurring at higher concentrations, and is observed for all arginine salts, thus leading to the nonlinear trends shown in Figure 4-9.

Even though for most cosolutes there is a direct relationship between the preferential interaction coefficient and the folding stability of the protein, there are, however, some exceptions [26]. The mechanism for each exception varies but all are the result of preferential interactions being repulsive in one conformational state and attractive in another. For compounds containing guanidinium and halide anions, exceptions can arise due to the exposure of hydrophobic residues upon unfolding. Guanidinium, the larger halides, and thiocyanate have a strong preference for hydrophobic residues which causes them to interact more strongly with the unfolded state and as a result, they denature proteins [99]. That is, if they are not inhibited from interacting with the protein surface, as in the case for arginine. The preferential interaction coefficient values obtained experimentally and MD simulations seem to indicate that arginine is inhibited from interacting strongly with the protein (either with the native state or the unfolded state) even though it has a guanidinium moiety, but the halide anions and thiocyanate are free to interact with the protein. This favorable interaction is in order of $SCN^- \approx I^- > Br^- > Cl^-$, as indicated by T_m and Γ_{μ_3} results. The interaction between F^- and the protein is likely similar in strength as that for Cl^-, which is weak at best, but as explained above, Γ_{μ_3} could not be determined for ArgHF.

The magnitude of the interaction with the protein for these salts not only arises from a difference in the protein-anion interaction but also from a difference in the guanidinium-anion interaction, as discussed above, with thiocyanate and iodide exhibiting a weaker interaction with guanidinium. Thus these ions are less likely to interact with arginine in the bulk solution. Fluoride and chloride not only exhibit a weaker interaction with the protein surface, but they also interact more favorably with the guanidinium functional group on arginine. Therefore, as a result of Arg-Arg interactions and weak Arg-anion interactions, Γ_{μ_3} values exhibit a nonlinear trend for all of the salts discussed and due to the variation in Arg-anion and protein-anion interactions, the magnitude of Γ_{μ_3} and the shift in T_m varies amongst the salts. ArgHCl and ArgHF are neither naturants nor denaturants because there is little difference between the interaction with the native and unfolded states. When a protein unfolds, free $ArgH^+$ will likely interact more strongly with the exposed hydrophobic residues but the clusters should become more excluded, counteracting this interaction. The interaction of the fluoride and chloride ions with the unfolded state likely does not change much because of the weak interaction these ions have for hydrophobic residues or interfacial surfaces. As for ArgHBr, ArgHI, and ArgHSCN, the interaction with the unfolded state is stronger due to an increase in surface area and the stronger interaction the anions exhibit for hydrophobic residues, leading to those salts denaturing proteins.

The salts to the left of chloride in the Hofmeister series exhibit similar interactions but of differing magnitude, resulting in Γ_{μ_3} trends distinct from the other salts. The phosphate, sulfate, and citrate salts have Γ_{μ_3} values that start out negative and remain negative, with values similar to other highly excluded solutes (e.g. glycerol), indicating that the salts have a strong repulsive interaction with the protein at all concentrations [70]. $ArgH(SO_4)_{1/2}$ is the least excluded of these salts, causing the ordering of Γ_{μ_3} data to differ slightly from the ordering of T_m data. Also, the profile for citrate is almost linear for most of the concentration range, likely from the large citrate ion interfering with Arg-Arg clustering and being highly excluded. It is well documented that salts with these anions typically stabilize the folded state, as is the case for the data shown here. Both this and the preferential exclusion can be contributed, in part, to the nonspecific repulsive interaction of the anions with protein molecules that result from the ions being well hydrated and preferring the bulk solution. But as discussed above, phosphate, sulfate and citrate have been

shown to exhibit a strong interaction with guanidinium, which further reduces the interaction with the protein.

Theoretical preferential interaction coefficient values computed from the MD simulation are reported in Table 4-3. The estimated preferential interaction values match well with the trend in the experimental values. The experimental and computational estimates for sulfate, phosphate (if linearly extrapolated to 0.5m) and citrate indicate large, negative preferential interaction coefficient values, while thiocyanate exhibits a large, positive preferential interaction coefficient value. The experimental and computational estimates do not match exactly, which is likely due to the inability of the force field to capture the interactions of these ions with the protein surface. Force fields are optimized for interaction of ions with water molecules. However, as shown above, for arginine salts, interactions with water is only one of the key interactions. In order to obtain an accurate estimate of preferential interaction coefficient values, force fields for electrolytes should also be benchmarked to reproduce bulk solution properties and interactions between individual ions [109, 110]. For the purposes of understanding the interesting experimental trends observed in this study, the current force fields provide reasonable estimates of the trends in the data.

On the basis of the preferential interaction coefficient of individual ions, it can be concluded that the number of cations on the protein surface is related to the number of anions. Thiocyanate ions have large positive preferential interaction coefficient values but it does not interact strongly with arginine cations. Sulfate ions have a negative preferential interaction coefficient but it interacts strongly with arginine. Therefore, there are two driving forces that push and pull arginine towards and away from the protein surface. In order to conserve the local charge near the protein surface, more arginine molecules are drawn towards the protein surface if there are a high number of anions near the protein surface. If the interactions between ions in the bulk solution are not as strong as compared to the interaction between the ion and the protein surface, the ions accumulate near the protein surface. Therefore, sulfate, phosphate and citrate salts have a large negative Γ_{μ_3} because of their exclusion from the protein surface and the strong interaction with arginine molecules, which pulls arginine away from the surface. The opposite can be said for thiocyanate, acetate and chloride salts.

Table 4-3: Theoretical preferential interaction coefficient values for α-Chymotrypsinogen A in aqueous arginine salt solutions. MD results produced by Diwakar Shukla.

Arg Salt	C (mol/L)	Γ^{VPO}	Γ^{MD}	Γ^{+}	Γ^{-}	Protein-Arg h-bonds
$(SO_4)_{1/2}$	0.5	-5.2±1.7	-4.5	-4	-3	15.3
$(Citrate)_{1/2}$	0.5	-7.7±1.5	-4.5	-4	-3	16.1
(H_2PO_4)	0.2	-5.7±0.4	-3	-2	-4	10
(Acetate)	0.5	N/A	0	1	-1	19.7
Cl	0.5	-2.6±0.3	-2.5	-1	-4	18.9
SCN	0.48	0.2±1.2	4	4	4	26.5

The error bars on the preferential coefficient values are on the order of ±1

This interaction between arginine and the anion further reduces the preferential interaction of the arginine salt with the surface of the protein and further limits interactions with the unfolded state, thus enhancing thermodynamic stabilization. This cation-anion interaction does not necessarily reduce Arg-Arg interactions though, because the anion bridges together two arginine molecules, similar to guanidinium-sulfate clusters discovered by Mason and coworkers, causing the nonlinear trend observed [101]. Therefore, the ordering of the Γ_{μ_3} and T_m data is not only ranked by how the individual anions interact with protein molecules but also, with how those anions interact with the guanidinium functional group on the arginine molecule.

The trends discussed here are related to the impact these salts have on protein aggregation. As discussed in the previous section, both changes in the folding equilibrium of the protein and changes in protein-protein interactions must be taken into consideration. The halide and thiocyanate salts exhibit an attractive interaction with the protein at low concentrations. This, according to the theory developed by Baynes and Trout, should reduce protein-protein interactions, much more so for arginine than for guanidinium due to the larger size of arginine [10, 15]. However, as discussed before, bromide, iodide, and thiocyanate also denature the protein, which counteracts this effect. This is not the case for chloride and fluoride, thus ArgHCl and ArgHF reduce aggregation solely through a reduction of protein-protein interactions. However, at high concentrations, the ArgHCl shows a net exclusion (which is likely the case for ArgHF and ArgH(Acetate) as well but to a lesser extent) and the aggregation suppression ability of ArgHCl plateaus because adding more ArgHCl at that point does not contribute any to protein-protein interaction suppression due to the additional arginine molecules being partitioned to the bulk solution. Even though ArgHCl shows a net exclusion at high concentrations, there are enough arginine molecules in the local domain of the protein for it to significantly reduce

protein-protein interactions, thus lowering the relative rate constant to about 0.2. This indicates that the effect from the arginine molecule should be significant for the highly excluded salt forms as well, though the effect of arginine reducing protein-protein interactions might be reduced some. Furthermore, aggregation suppression is further enhanced for the excluded salts via the stabilizing effect the citrate, sulfate, and phosphate salt forms have on the folding equilibrium, leading to the relative rate constant being further reduced beyond that for ArgHCl. Moreover, another contributing factor may be an increased diffusional barrier. The cluster formation induced by the phosphate, sulfate, and citrate ions likely increases the viscosity of the solution in addition to forming a network of bridged ions around the protein, thus inhibiting the movement of protein molecules. This mechanism has not been considered before and should be further investigated since it will likely be a key feature of novel excipients.

4.3. Conclusions on the Arginine Mechanism

For arginine salt solutions, we have shown that the interaction between ions and the interaction of the anion with the protein surface influences the preferential interaction coefficient value, which is directly related to the conformational stability of the native state. Attractive ion-ion interactions in solution lead to the formation of clusters that are larger in size as compared to the individual ions. This effective size enhancement of the molecules around the protein increases the "Gap effect" and inhibits protein-protein association. Furthermore, clustering should lead to an increase in the viscosity of the solution, which should lower the diffusion of proteins in solution. Therefore, ion-ion interactions affect the rate of aggregation via three mechanisms: 1) enhancing the native state conformational stability, 2) increasing the barrier for protein-protein association, and 3) increasing the viscosity of the solution medium. These results not only help to elucidate the arginine mechanism, they also have large implications for interpreting the mechanism behind the Hofmeister Series, in that, changes in the water structure did not seem to come into play. Rather, ion-ion interactions and other ion specific behavior seem to dominate the behavior of each arginine salt.

Chapter 5

5. Novel Excipient Development

As described in previous chapters, protein aggregation is a major problem in the biotech industry, which not only affects the development and production of therapeutics but also affects the use of diagnostic proteins. As a consequence, there is interest in improved methods of stabilizing protein based products, which has motivated us to synthesize and test a novel class of solution additives with aggregation suppression abilities significantly greater than currently used excipients. In an attempt to accomplish such a goal, we set out to produce so called "neutral crowder" additives. It was theorized that by balancing the attraction between a protein and guanidinium with the exclusion generated by volume exclusion effects observed for large molecules, a new class of additives could be created that would inhibit protein-protein association with little influence on the conformational stability of the protein. As described below, we created new molecules by modifying the surface of large molecules through the addition of multiple guanidinium functional groups. We later showed that with the proper choice of counterion, the synthesized molecules behaved as aggregation suppressors. For our model proteins, the new compounds fulfilled the main goal of the project, in that they lower the rate of aggregation by an order of magnitude more than commonly used excipients. We feel that this is a positive result in the right direction and hopefully, the compounds will perform equally well for other proteins.

5.1. Structure of Synthesized Compounds

5.1.1. Modified Amine Compounds

The first series of compounds synthesized were simple diguanidinium compounds. We synthesized such compounds so that different design features could be explored and assessed (*i.e.* size and number of guanidinium functional groups) to aid in the development of larger, more complex structures. Figure 5-1 depicts the general structure for one type of simple compound (*i.e.* diguanidinium compounds), which are numbered as 1n, where 1 represents a single hydrocarbon chain with n number of carbons. A variety of diguanidinium compounds were synthesized, ranging from the smallest size possible (*i.e.* a 2 carbon chain labeled 102) to the largest soluble compound, about 12 carbon units, which was labeled 112 (see Figure 5-2).

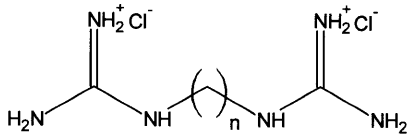

Figure 5-1: General structure for 1n series of guanidinium chloride compounds

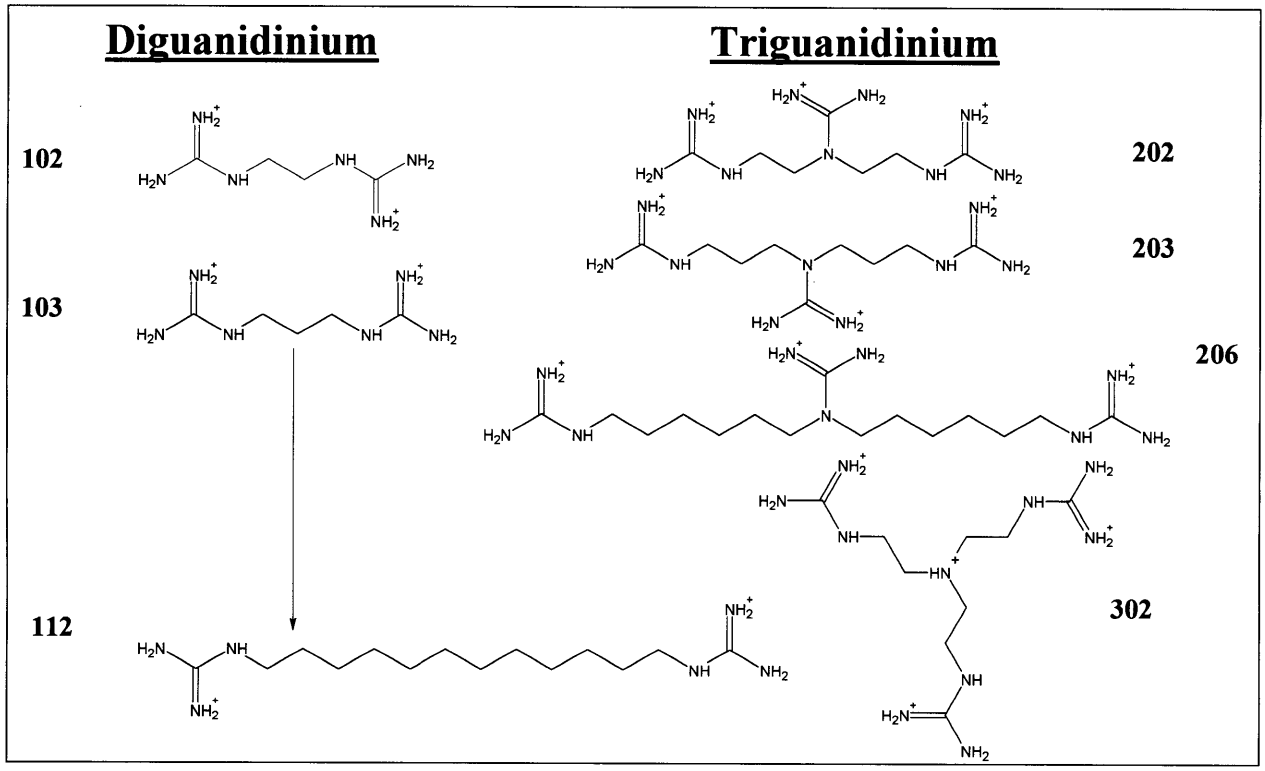

Figure 5-2: Structures and reference number of all the Di- and Triguanidinium compounds synthesized.

Larger guanidinium compounds were synthesized, with either 2 (2n series) or 3 (3n series) hydrocarbon chains attached to a central amine functional group (see Figure 5-2). Both series produced a compound with three guanidinium functional groups, thus they are referred to as triguanidiniums. The variety of commercially available amines is limited, especially for the triamines, thus the type of compounds that could be produced in this manner is somewhat limited unless the amine compounds are synthesized from scratch. Figure 5-2 represents nearly all of the compounds that could be produced from commercially available amines. There were a few other compounds that were tested, mainly amine compounds that contained hydroxide functional groups, but the use of those compounds was limited due to difficulty in purifying the product.

5.1.2. PAMAM Dendrimers

The most promising starting materials to work with are dendrimers, specifically polyamidoamine (PAMAM) dendrimers (which is probably the most well known type of

dendrimer). These molecules are repeatedly branched (each branch contains an amide functional group, which gives the branches rigidity) and are roughly spherical in shape, not to mention large but still soluble in aqueous solutions. The properties of dendrimers are dominated by the functional groups on the surface of the molecule. In the case of PAMAM dendrimers, the surface functional groups are amines, which can easily be converted to guanidinium groups via the methodology described in Chapter 3. Dendrimers are classified by generation (*i.e.* the number of repeated branching cycles that are performed during syntheses). Each subsequent branching nearly doubles the size and molecular weight of the molecule (see Table 5-1), thus making them an ideal starting material to produce excipients with a radius greater than 8 Å (which was a benchmark size set at the beginning of the project).

Table 5-1: Size and molecular weight of PAMAM Dendrimers

Generation	Core Carbons	MW Daltons	Diameter Å	Surface Groups
0	2	517	15	4
1	2	1430	22	8
2	2	3256	29	16
3	2	6909	36	32

It should be noted that there is some ambiguity in the literature in how generation naming is applied. The naming shown here follows how the manufacturer (Dendritech®) names each generation, which is to refer to the starting amine compound (usually ethylene diamine) as the core and to refer to the first dendrimer structure produced as Generation 0 (some sources refer to this structure as Generation 1 and refer to the core as Generation 0).

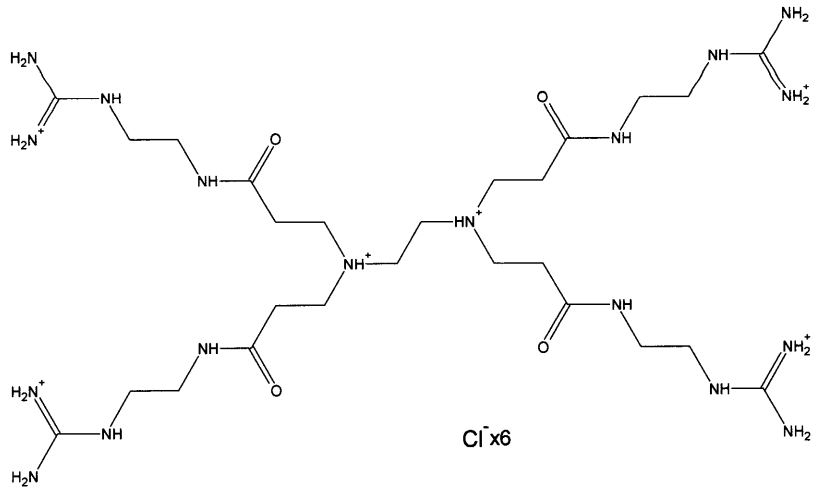

Figure 5-3: Generation 0 polyamidoamine (PAMAM) dendrimer, with an ethylene diamine core and guanidinium chloride surface groups (G0-Cl)

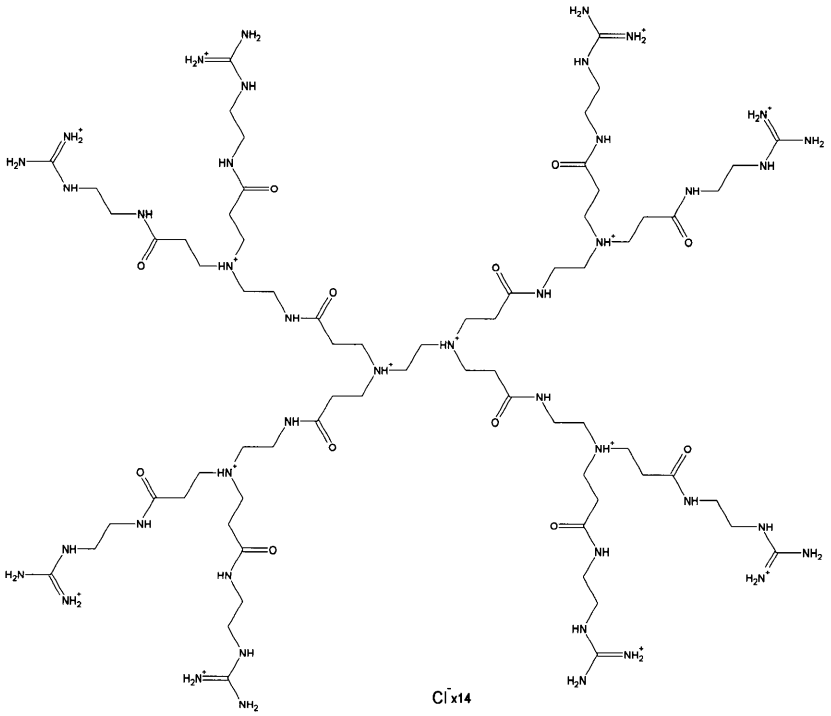

Figure 5-4: Generation 1 polyamidoamine (PAMAM) dendrimer, with an ethylene diamine core and guanidinium chloride surface groups (G1-Cl)

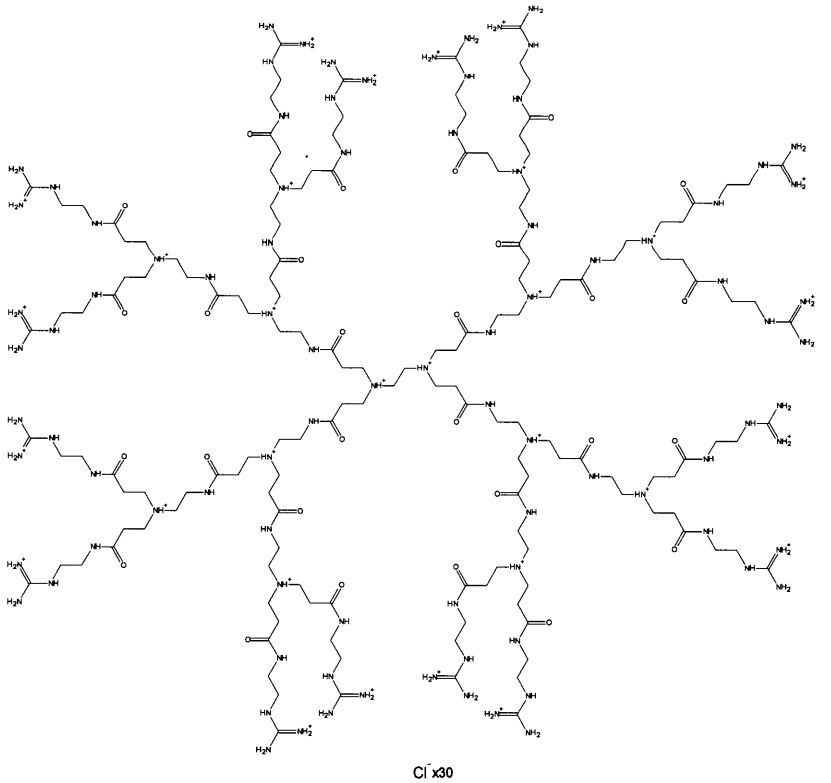

Figure 5-5: Generation 2 polyamidoamine (PAMAM) dendrimer, with an ethylene diamine core and guanidinium chloride surface groups (G2-Cl)

Figure 5-3 through Figure 5-5 depict the surface modified dendrimers structures produced (initially all structures were chloride salts). As shown in the figures and later confirmed by NMR (see Appendix A), the internal branch points are protonated as well, thus the additional counterions in excess of the number of surface groups. The core of the dendrimer can vary but there is no variation in the branch structure. A core size of 2, 4, 6, or 12 carbons are commonly available from Dendritech®. In addition, Starburst® offers a PAMAM dendrimer with ammonia as the core structure (these dendrimers have 3 initial branches and are commonly referred to as Starburst® polymers). Only dendrimers with an ethylene diamine core were utilized in this study (see Table 5-2 for the molecular weight of the protonated excipients and the entire molecular weight of various salt forms).

Table 5-2: Molecular weights of guanidinium modified PAMAM Dendrimers (including total molecular weight for various salt forms).

Generation	Core Carbons	Surface	MW (Protonated)	Total Molecular Weight of Salt		
				Cl^-	SO_4^{2-}	$H_2PO_4^-$
0	2	$GdnH^+$	690.9	903.6	979.1	1272.8
1	2	$GdnH^+$	1780.3	2276.6	2452.7	3138.1
2	2	$GdnH^+$	3959.0	5022.6	5400	6868.7

5.1.3. Peptides

Another promising class of molecules is arginine peptides. These molecules are simple to produce and were expected to have enhanced aggregation suppression properties over ordinary arginine. The peptides produced were synthesized either by our collaborators or purchased from Genscript. Most of the results presented depict data collected using the GenScript® peptides as excipients. Figure 5-6 depicts the structures of arginine peptides ranging from a dimer to a tetramer, though peptides up to a heptamers were synthesized and tested.

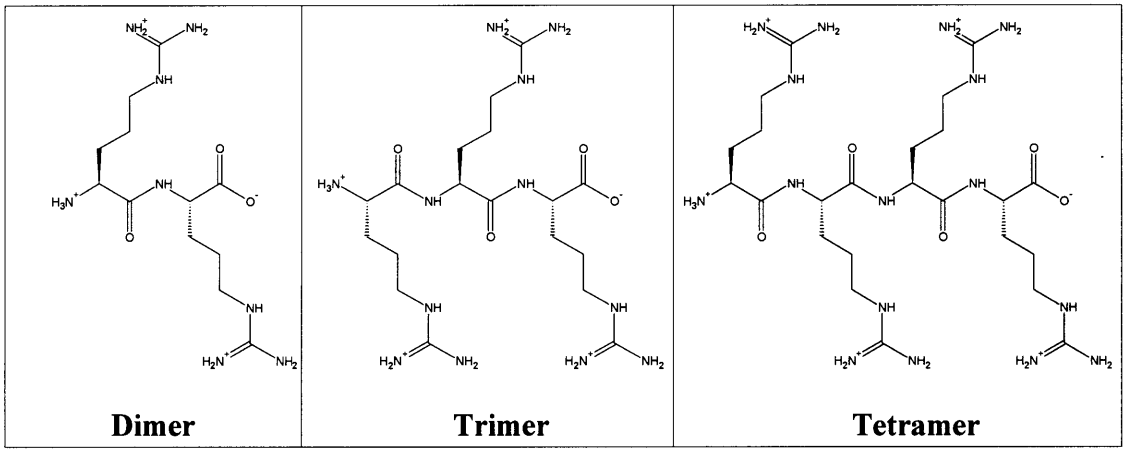

Figure 5-6: L-Arginine peptides.

Table 5-3: Peptides obtained from Genscript® Corporation.

Peptide	MW	Molecular Weight of Salt	
	Daltons	HCl	H_2SO_4
RR	330.4	403.3	428.5
RS	261.3	297.7	310.3
RL	287.4	323.8	336.4
RQ	302.3	338.8	351.4
RE*	303.3	--	--
RRR	486.6	596.0	633.7
RER*	459.5	496.0	508.5
ERE**	432.4	--	--
RRRR	642.8	788.6	838.9
RREE*	588.6	--	--
RERE*	588.6	--	--

*These peptides are amphoteric and therefore, should exist as a zwitterion without the need of a counterion
**This peptide requires the presence of a cation to neutralize the charge.

Other peptides besides all arginine residues were synthesized and tested. To gather a clear picture of the effects from changing a single residue, dimers were created with one arginine (R) residue and the second residue either serine (S), leucine (L), glutamine (Q), or glutamic acid (E). To mimic and enhance the behavior of arginine-glutamate mixtures, trimers and tetramers were produced containing different arginine and glutamic acid sequences (RER, ERE, RREE, and RERE). See Table 5-3 for a complete list of the peptides obtained from GenScript® and see Appendix A regarding purity and structural data. Once it was discovered that the aggregation suppression behavior was strongly controlled by the counterion, studying different sequences was ended so adequate work towards studying the effect of different counterions could be conducted. Furthermore, the synthesis and purification of other peptide sequences did not turn out well since the peptides containing arginine and glutamic acid could not be dried via lyophilization. The cake structure would always collapse, leaving behind a viscous liquid that could not be dried any further.

5.2. Aggregation Suppression Behavior

5.2.1. Chloride Salt Form

As mentioned previously, all excipients were initially prepared as chloride salts. A chloride salt was chosen because most excipients of this nature that are salts come in the chloride form (*e.g.* NaCl, guanidinium chloride, Arginine Hydrochloride, *etc.*) in addition to the fact that the chloride ion has a negligible preferential interaction and anhydrous HCl is one of the few

strong acids available dissolved in an organic solvent, the other being trifluoroacetic acid. Therefore, making a chloride salt from a BOC protected guanidine is only one of a handful of options without having to perform an ion exchange. It was somewhat apparent that the counterion might have an impact on preferential interactions; however, the initial strategy for adjusting the preferential interaction was by adjusting the size of the excipient and the number of surface functional groups.

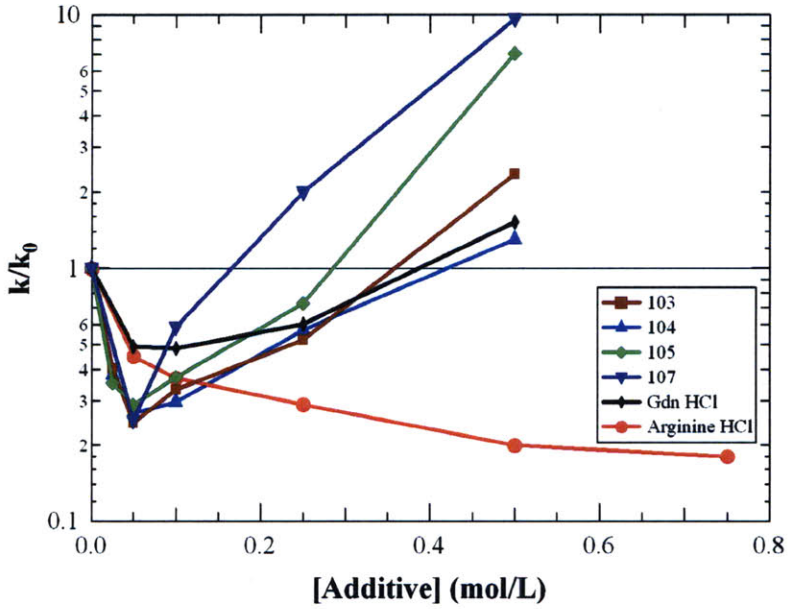

Figure 5-7: The influence of diguanidinium compounds (in the form of chloride salts) on aCgn monomer loss due to aggregation at 52.5°C. All solutions contained 10 mg/mL aCgn and were prepared in a 20 mM sodium citrate pH 5 buffer. Figure depicts the Relative Rate Constant (as determined by changes in the half life of aCgn monomer loss) versus additive concentration, with lines drawn through the plots to aid the eye.

In short, all excipient forms initially produced failed to inhibit aggregation. In fact, all induced aggregation, speeding up the process at moderate to high concentrations. Figure 5-7 depicts the relative rate constant for aCgn aggregation in the presence of diguanidinium compounds. All of the diguanidinium compounds shown are chloride salts and exhibit a behavior very similar to guanidinium chloride. That is, the compounds inhibit aggregation at low concentrations but speed up the process at concentrations above 0.4 M. The compounds are more effective than guanidinium chloride or arginine hydrochloride at low concentrations (~0.05 M), but that window of aggregation suppression is narrow and not ideal. Arginine hydrochloride, on the other hand, continues to inhibit aggregation as its concentration is increased, until ultimately reaching a plateau. Furthermore, the larger the diguanidinium compound, the more destabilizing the effect. The 109 and 112 compounds are not shown because they sped up aggregation so

rapidly that an observable amount of aggregation occurred at room temperature prior to incubation.

The arginine peptides exhibit an almost identical behavior. Figure 5-8 depicts the relative rate constant for aCgn aggregation in the presence of arginine peptides (chloride salt forms). The figure looks less dramatic because the concentration range is shorter; however, the crossover point from the concentration range in which aggregation is suppressed to the concentration range in which aggregation is induced is nearly identical, roughly 0.3-0.4 M. The arginine peptides also outperform arginine at a concentration of 0.05 M (similar to the diguanidinium compounds and to almost the exact same extent). Moreover, the larger the arginine peptide, the more it destabilizes the protein.

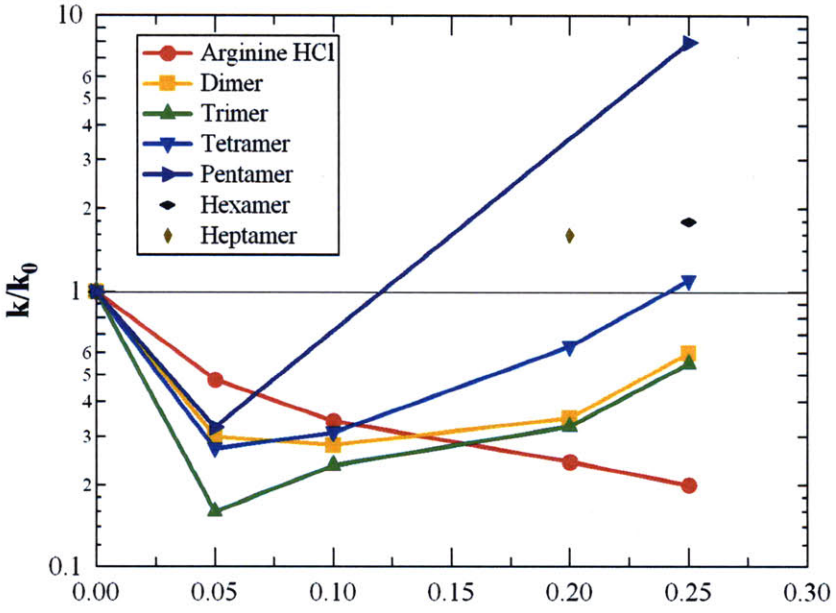

Figure 5-8: The influence of arginine peptides (in the form of chloride salts) on aCgn monomer loss due to aggregation at 52.5°C. All solutions contained 10 mg/mL aCgn and were prepared in a 20 mM sodium citrate pH 5 buffer. Figure depicts the Relative Rate Constant (as determined by changes in the half life of aCgn monomer loss) versus additive concentration, with lines drawn through the plots to aid the eye.

The guanidinium modified dendrimers (chloride salt forms) also exhibit this behavior (see Figure 5-12 in the next section for some of the data collected). Furthermore, all of these excipients lower the unfolding temperature of the protein, as determined by DSC, which gives an indication of how they are speeding up aggregation. Even if they do inhibit association, they are also shifting the unfolding equilibrium to the unfolded state, counteracting any association suppression. All of this behavior is opposite of what was expected, making it clear that the

mechanism involved in the preferential interaction is more complex than previously thought. It was initially thought that the large size of the compounds would counteract the preferential interaction of the guanidinium functional group due to steric exclusion, but in reality, the large size seems to enhance the preferential interaction. The problem with the initial theory was that the design plan was too simple and assumed that such molecules would be rigid spheres instead of soft, flexible molecules (which is what was synthesized). Moreover, the initial theory assumed that guanidinium was attracted to protein surfaces in a nonspecific manner, which is not the case since guanidinium has a stronger preference for buried hydrophobic groups. Figure 5-9 gives a simple overview of how soft, flexible structures influence preferential interactions and why monomeric arginine does not exhibit this behavior. In essence, instead of inhibiting interactions, the large size and flexible nature allows for multiple interactions that bind cooperatively, more so to the unfolded state. In other words, when one functional group interacts with the surface, it seems reasonable to assume that it makes it much more likely that a second functional group will interact due to the proximity between the second functional group and the surface of the protein brought on by the initial interaction. And once 2 functional groups interact, the process likely cascades until all possible functional groups are interacting with the protein. Moreover, for larger compounds, it seems more likely for other functional groups to find a place to bind on the surface (smaller molecules might not have the reach to bind to another spot on the surface and large molecules likely diffuse away from the protein at a slower rate). Therefore, large, soft molecules with many guanidinium functional groups bind to the surface of proteins (or the unfolded state) with an affinity stronger than anticipated.

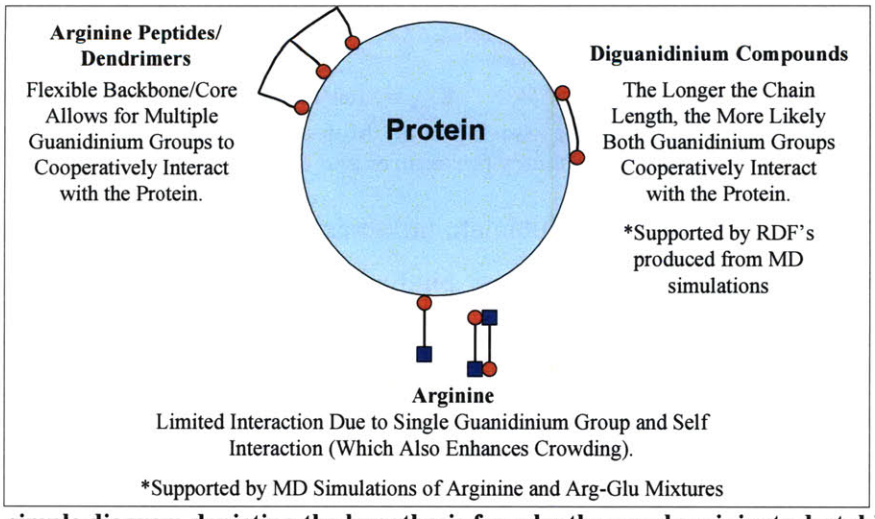

Figure 5-9: A simple diagram depicting the hypothesis for why the novel excipients destabilize proteins.

MD simulations seem to support these conclusions. The preferred orientation of diguanidinium compounds was found to be parallel with the protein surface, indicating that they lay flat on the surface with both functional groups bound. Experimental evidence for this hypothesis was recently published [111]. New molecules with guanidinium surface groups (arginine peptides and polyether dendrimers) were found to have a very strong interaction with proteins (in this case BSA) and were deemed "molecular glues" by the authors since they strongly induced protein association. The interaction was strong enough to be determined by isothermal titration calorimetry (ITC), indicating association rather than preferential binding. Moreover, the interaction became stronger as the size was increased (*i.e.* larger dendrimer) and the flexibility improved (a triethylene glycol linker was introduced into the peptide structure, which increased the extent of association). These results seem to support the hypothesis of cooperative binding.

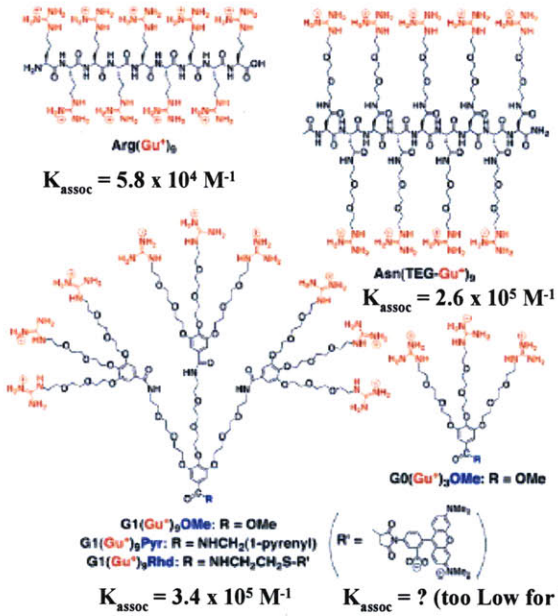

Figure 5-10: Association constants for the association of "Molecular Glues" with BSA, as determined by isothermal titration calorimetry (structures and data obtained from [111]).

Our early work with arginine-glutamate mixtures indicated that the attractive interaction between the two molecules leads to reduced binding to protein molecules, which reduces the conformational destabilization and enhances aggregation suppression by creating large clusters that behave as "neutral crowders". Attempts were made to mimic this behavior but it quickly became clear that a better or more universal method would be to exchange the chloride

counterion with an ion that favorably interacts with the guanidinium functional group (see Chapter 4 in regards to how the counterion influences the behavior of arginine).

5.2.2. Guanidinium-Carboxylate Mixtures

The structure of the diguanidinium compounds is quite similar to arginine and it is apparent that if one wants to add another compound to the solution that will interact with the diguanidinium compound in a fashion similar to the interaction between arginine and glutamate, the best candidate would be a dicarboxylate (see Figure 5-11). Fortunately, dicarboxylic acids are commercially available and inexpensive. Initial aggregation studies on smaller compounds showed that the best results came from a 50/50 mixture of a diguanidinium (chloride salt) and a dicarboxylate (sodium salt) both with identical carbon chain lengths. Such mixtures exhibited enhanced aggregation suppression (similar in ability to ArgHCl) when compared to either component by itself. However, such results were limited to the smaller compounds because mixtures containing diguanidinium and dicarboxylate compounds with a chain length greater than 3 carbons were nearly insoluble, even though either component was completely soluble by itself at all concentrations tested (between 0.5-1.0 M). This insolubility gives strong support to the hypothesis that the two molecules favorably interact in solution and from these results, it seems that the attractive interaction is quite strong. The hydrogen bonding and electrostatic interaction between the two end groups, in addition to the hydrophobic interaction between the two core hydrocarbon chains, likely induces a strong attractive interaction between the two molecules, causing the mixture to precipitate out of solution.

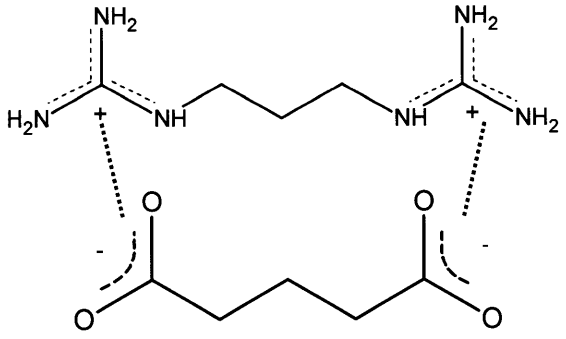

Figure 5-11: An example of the attractive interaction between diguanidinium and dicarboxylate molecules.

To make the situation worse, dicarboxylate (sodium salt) compounds with a chain length greater than 7 carbons were found to be insoluble as well, limiting the options for this approach. An attempt was made to test mixtures in which the two components had differing sized

hydrocarbon chains. The solubility of the mixtures containing larger compounds was improved slightly by this approach and such mixtures did exhibit an ability to suppress aggregation. However, the approach was limited as well, not too mention that it could not be implemented for other excipient types, such as peptides and dendrimers. Furthermore, it became apparent that the choice of counterion would produce the same effect (see Chapter 4); therefore efforts were directed toward studying the effect of exchanging the counterion. On a side note, the solubility of such compounds could possibly be improved by utilizing a more hydrophilic core structure (such as polyethylene glycol) but such compounds would have to remain small. This is due to the fact that for two compounds to associate in the manner desired, the structure would have to be linear and the number of surface groups would be limited to two or three guanidiniums.

5.2.3. Other Salt Forms

From the literature review presented in Chapter 4 and the results obtained from the diguanidinium-dicarboxylate mixtures, it became apparent that if the chloride ion was exchanged with phosphate, sulfate, or acetate, conformational destabilization would be reduced and aggregation suppression would be enhanced as a result. The work conducted by Mason and coworkers showed that guanidinium and sulfate interact so strongly that they form large clusters in solution. As a result, it was initially feared that the solubility of the excipients would be reduced by the formation of a network of compounds bridged together by sulfate or phosphate ions. Initial tests showed this to be true for the diguanidinium compounds (in addition, arginine phosphate has a low solubility as well, indicating a strong interaction). All diguanidinium compounds, except 103, precipitated completely when Na_2SO_4 was added to the solution, indicating a strong interaction and cluster formation. However, when mixed with Na_2HPO_4, the smaller diguanidinium compounds (102-104) precipitated right away while 105 turned slightly cloudy and 106 and larger compounds remained clear. When mixed with NaH_2PO_4, compounds 102-103 precipitated while the rest remained clear, indicating a stronger interaction with HPO_4^{2-}, as predicted.

These results of lower solubility only seem to be limited to the diguanidinium compounds, likely due to the hydrophobic attraction between the molecules and relative ease of forming a crystalline structure due to the simple structure of the compound. The solubility of the dendrimers is not reduced by either sulfate or phosphate and the solubility of the arginine peptides is not reduced by sulfate either. However, the solubility of the peptides in the presence

of phosphate was not tested due to the low solubility of arginine phosphate and the limited supply of arginine peptides. The performance of the dendrimer and peptide excipients in the form of other salt types is significantly different from the chloride salt form. Exchanging the chloride ion with acetate moderately improves the performance of the generation 0 dendrimer (see Figure 5-12). The acetate salt form has a profile similar to the chloride salt, in that at low concentrations it inhibits aggregation but as the concentration increases, the trend reverses and ultimately speeds up aggregation at high concentrations. The only difference is that the relative rate constant is lower than the chloride salt form for all concentrations tested.

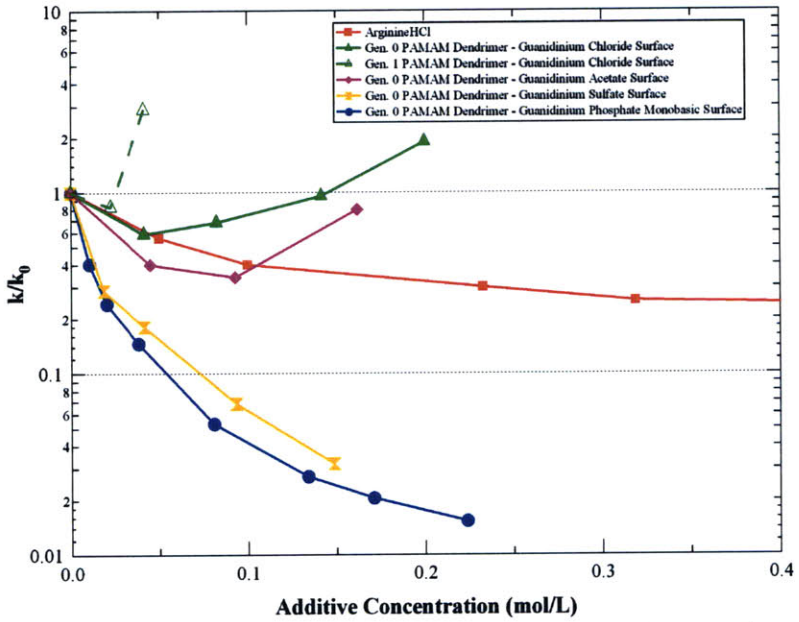

Figure 5-12: The influence of guanidinium modified PAMAM dendrimer salts on aCgn monomer loss due to aggregation at 52.5°C. All solutions contained 10 mg/mL aCgn and were prepared in a 20 mM sodium citrate pH 5 buffer. Figure depicts the Relative Rate Constant (as determined by the amount of time for 20% monomer loss) versus additive concentration, with lines drawn through the plots to aid the eye.

When the chloride ion is exchanged with either SO_4^{2-} or $H_2PO_4^-$, the performance of the dendrimers is significantly improved. As shown in Figure 5-12, at low concentrations, both the sulfate and phosphate salt forms lower the rate of aggregation to a greater extent than either the chloride or acetate salt forms. Furthermore, as the concentration increases, this trend continues but begins to plateau at a concentration near 0.2 M. At such a concentration, the relative rate constant is approximately 0.02, which is an order of magnitude better than the relative rate constant for arginine. This trend continues for larger dendrimer forms, with the larger dendrimers outperforming the generation 0 dendrimer, as expected. As shown in Figure 5-13, the relative rate constant profiles for the generation 1 and generation 2 phosphate dendrimers are similar to

the profile for the generation 0 dendrimer, but the relative rate constant is lower for all concentrations, with the generation 2 dendrimer exhibiting the lower relative rate constant profile.

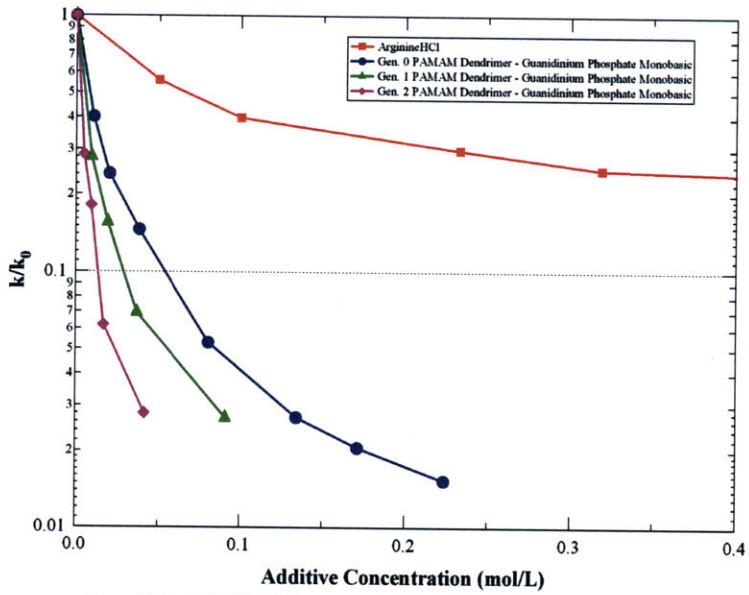

Figure 5-13: The influence of modified PAMAM dendrimers (guanidinium phosphate monobasic surface) on aCgn monomer loss due to aggregation at 52.5°C. All solutions contained 10 mg/mL aCgn and were prepared in a 20 mM sodium citrate pH 5 buffer. Figure depicts the Relative Rate Constant (as determined by the amount of time for 20% monomer loss) versus additive concentration, with lines drawn through the plots to aid the eye.

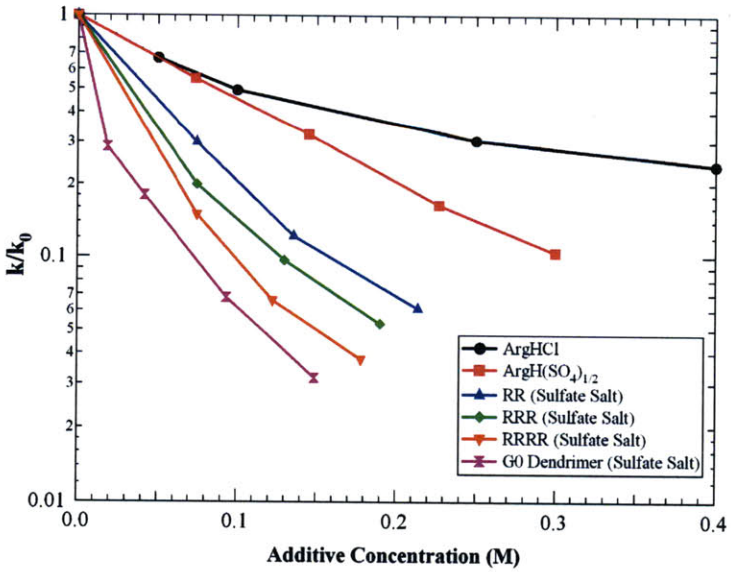

Figure 5-14: The influence of arginine peptides (in the form of sulfate salts) on aCgn monomer loss due to aggregation at 52.5°C. All solutions contained 10 mg/mL aCgn and were prepared in a 20 mM sodium citrate pH 5 buffer. Figure depicts the Relative Rate Constant (as determined by the amount of time for 20% monomer loss) versus additive concentration, with lines drawn through the plots to aid the eye.

The arginine peptides exhibit a very similar behavior when the chloride ion is replaced with sulfate. Figure 5-14 shows the improvement in aggregation suppression the sulfate ion provides over the chloride ion for monomeric arginine. Each additional arginine residue added improves the aggregation suppression, which is opposite of the behavior exhibited by the chloride salt forms. The largest peptide tested in this manner was a tetramer, which has a relative rate constant profile close to that of the generation 0 dendrimer (sulfate salt). If this trend continues, the larger peptides will likely exceed the performance of the dendrimers.

Even though the larger dendrimers and peptides perform better at inhibiting aggregation, they might not be optimal for use as a formulation excipient due to tonicity constraints (and possibly viscosity and toxicological constraints as well). Ideally, protein formulations must be isotonic (*i.e.* same osmotic pressure as blood) to minimize pain upon injection. Electrolyte solutions have a higher osmotic pressure than non-electrolyte solutions due to the presence of dissociated species. The larger excipients have many counterions per molecule, resulting in a very high osmotic pressure. Figure 5-15 compares all of the excipients tested at isotonic concentrations. The figure shows the Shelf Life Extension Factor (*i.e.* the factor by which the shelf life of the protein is extended when in the presence of that excipient) for aCgn aggregation. When compared to sucrose and ArgHCl, all of the excipients developed show improvement in the shelf life of the protein. However, for both the dendrimers and peptides, the isotonic concentration for the larger forms is quite low. So low, in fact, that the performance at such concentrations is less than the performance of the smaller excipients (that is when the smaller excipients are formulated at isotonic concentrations as well).

As shown in the figure below, the most ideal excipients for formulations seem to be arginine dimers and generation 0 dendrimers, while the other excipients may be useful during production and purification steps. A possible way to improve upon the results presented here might be to introduce other residues into the peptide structure which will increase the size but at the same time not change the tonicity. Such a design plan was initiated but not followed through with because the peptides could not be lyophilized successfully. Alternate ways of preparing such excipients would have to be explored. Furthermore, using a larger core dendrimer or synthesizing the dendrimers from scratch (possibly using solid phase synthesis techniques) could allow for larger dendrimers with the same number of surface groups, thus improving aggregation suppression without increasing the tonicity.

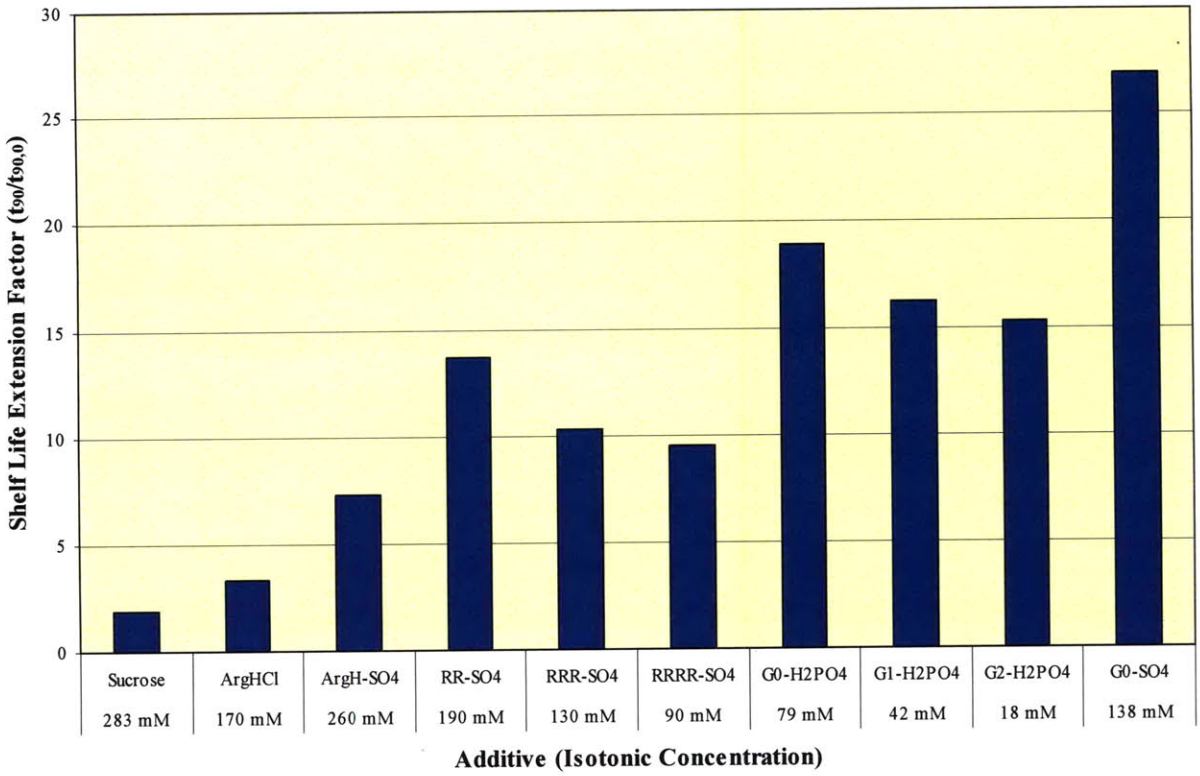

Figure 5-15: Shelf Life Extension Factor for bovine α-Chymotrypsinogen aggregation (10 mg/mL, T = 52.5°C, Buffer = 20 mM Na-Citrate, pH 5) at isotonic concentrations for the novel excipients and commonly used excipients.

5.3. Other Aggregation Conditions

5.3.1. Lower Temperatures

All of the results presented thus far have been for experiments conducted at elevated temperatures. It was predicted that "neutral crowders" would work at all temperatures, but to prove that, lower temperature experiments were conducted. Figure 5-16 depicts aCgn monomer loss with time when incubated at 37°C in the presence of G0 dendrimers formulated at isotonic concentrations. The relative profiles are almost identical to higher temperature data. With no excipient present, 5% of the monomer is lost due to aggregation in about a half a day. When in the presence of an isotonic solution of ArgHCl, a 5% loss occurs after about two days (roughly 4 times longer). However, when in the presence of an isotonic solution of G0-H_2PO_4 (the generation 0, guanidinium modified PAMAM dendrimer with a phosphate, monobasic counterion), this is extended to about 9 days (18 times longer) and when in the presence of an isotonic solution of G0-SO_4, a 5% loss occurs after 25 days (50 times longer).

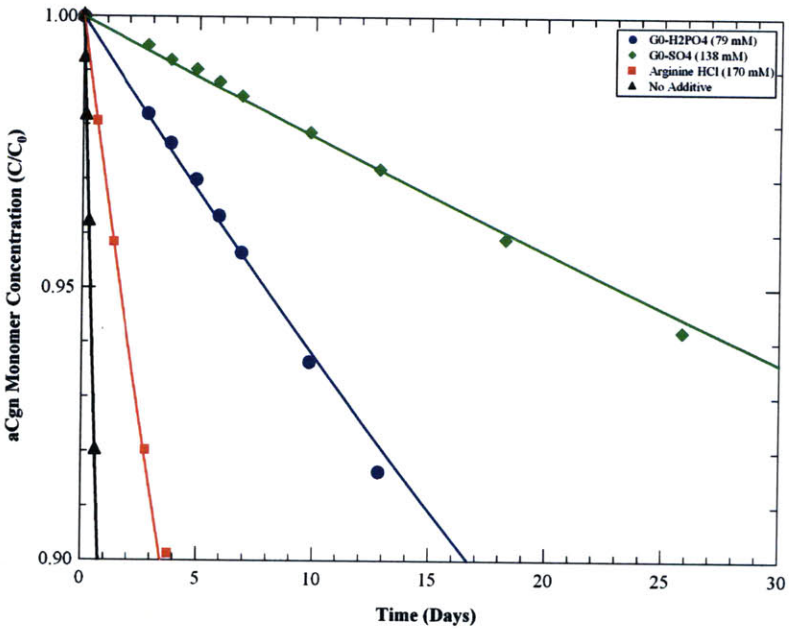

Figure 5-16: The influence of G0 dendrimer salts on aCgn monomer loss due to aggregation at 37°C. All solutions contained 10 mg/mL aCgn and were prepared in a 20 mM sodium citrate pH 5 buffer. Figure depicts monomer concentration, C, normalized with respect to the initial monomer concentration, C_0, versus time, with plots fitted to a second order rate law.

The above figure only represents data collected at one protein concentration and at one temperature. Table 5-4 depicts k/k_0 values for a variety of temperatures and two different protein concentrations. The excipients perform best at low temperatures. However, they also perform best at lower protein concentrations. Both of these results are not surprising given that protein solutions are least aggregation prone at low temperatures and at low concentrations. However, at higher protein concentrations, the aggregation rate reduction is still quite significant and the novel excipients still outperform ArgHCl by an order of magnitude.

Table 5-4: Relative rate of bovine α-Chymotrypsinogen (aCgn) aggregation at various temperatures and protein concentrations in the presence of the novel excipients at isotonic concentrations.

	k/k_0				
	10 mg/mL aCgn			40 mg/mL aCgn	
	37 °C	45 °C	52.5 °C	37 °C	45 °C
G0-H_2PO_4 (79 mM)	0.049	0.061	0.053	0.059	0.066
G0-SO_4 (138 mM)	0.02	0.025	0.037	0.0271	0.018
ArgHCl (170 mM)	0.237	0.27	0.303	0.217	0.36

Table 5-5 gives an overview of the average shelf life of aCgn at different temperatures under three different formulation scenarios. It should be noted that the 25°C data are predicted values using an Arrhenius Plot of low temperature data (37°C - 42.5°C). With no additive, the

shelf life (*i.e.* the amount of time required for a 5% loss of the monomer) of aCgn ranges from 2 minutes (52.5°C) to only 3.4 days (room temperature). When in the presence of ArgHCl, these time periods are quadrupled, allowing for a 12 day shelf life at room temperature. But when in the presences of the G0-SO_4 dendrimer (k/k_0 = 0.025), the shelf life is increased by a factor of 40, allowing for a shelf life of 5 months at room temperature (which exemplifies the necessity for performing experiments at elevated temperatures due to this long time frame for a single experiment). At refrigerated conditions, this could be extended to a much longer period of time, possibly allowing for a protein that is only stable for a few days to be stored for the time frame desired (18-24 months).

These results validate the use of high temperature data to analyze excipients. Though, it is almost impossible to obtain an accurate prediction of shelf life values from high temperature data due to non-Arrhenius behavior, it does seem that the relative rate constant values remain near constant over a wide range of temperatures for these excipients. Therefore, if one knows the shelf life of a protein at low temperature, they can predict the extended shelf life resulting from the addition of an excipient using high temperature data. These results also show that the excipients will be of use as formulation excipients. However, additional considerations must be taken into account (*e.g.* testing on other proteins, toxicological effects, rates of other degradation pathways, *etc.*).

Table 5-5: Time period for a 5% loss of α-Chymotrypsinogen monomer due to aggregation at various temperatures and rate reductions (10 mg/mL, Buffer = 20 mM Na-Citrate, pH 5).

t_{95}: 5% α-Chymotrypsinogen A Monomer Loss						
T (°C)	No Additive		Arginine k/k_0 = 0.25		Dendrimer k/k_0 = 0.025	
52.5	2	Minutes	8	Minutes	1.3	Hours
45	2.1	Hours	8.4	Hours	3.5	Days
37	8.6	Hours	1.4	Days	14	Days
25*	3.4	Days	12	Days	5	Months

10 mg/mL aCgn, 20 mM Sodium Citrate, pH 5
*Predicted Value (Arrhenius Plot of Low Temperature Data)

5.3.2. Other Proteins

The main focus throughout this project has been to develop excipients which will stabilize therapeutic proteins. However, there is a great need to stabilize proteins for diagnostic and clinical uses, as well as for research purposes. There are many types of proteins used for

such applications, however, one type that is commonly used are lectins, which are a class of proteins that bind to sugars. Lectins are important in medicine, medical research, biotech research, biomanufacturing, and recently have been the focus for therapeutic use. But more importantly, they are commonly used as a diagnostic agent [112]. The most commonly used lectin is Concanavalin A (Con A), which has seen wide spread use in lectin affinity chromatography, which is commonly used to purify glycoproteins.

At a pH greater than 7, the predominant form of Con A is a tetramer, but below pH 7 it is a dimer and is often utilized in this pH range. Furthermore, at a pH below 7, the protein is unstable, showing a significant amount of aggregation and precipitation at room temperature within just a few hours. Figure 5-17 shows the Shelf Life Extension Factor for the novel excipients (as well as several commonly used excipients). Con A was incubated at 37°C at a pH of 6.5 and the excipients were at or below isotonic concentrations. From these results, it seems as though Con A is sensitive to ionic strength at such a pH because sodium chloride, sodium phosphate, and arginine hydrochloride all speed up aggregation. Furthermore, the phosphate form of the novel excipients shows only a marginal stabilization. The G0 phosphate dendrimer exhibits a Shelf Life Extension Factor of 1.5, which is about the same as that for sucrose and glycerol. However, the sulfate forms exhibit much improved aggregation suppression. At isotonic concentrations, the G0 sulfate dendrimer extends the shelf life by a factor of 6, while the G1 sulfate dendrimer extends it by 16, showing that for Con A aggregation, the low isotonic concentrations are not a factor. In addition, it shows the importance of testing all of the salt forms due to the drastic difference between the phosphate and sulfate forms (this behavior is also observed for other sodium salts as well).

The most surprising result is the behavior of the arginine peptides. At isotonic concentrations, the arginine trimer (sulfate salt) extends the shelf life by a factor of 20 while the arginine tetramer (sulfate salt) extends the shelf life by a factor of 33. This seems to indicate that larger peptides (as well as larger dendrimers) will perform even better and all of the excipients will likely be of use at very low concentrations. These are the kind of results desired and hopefully, such behavior will be observed for other proteins.

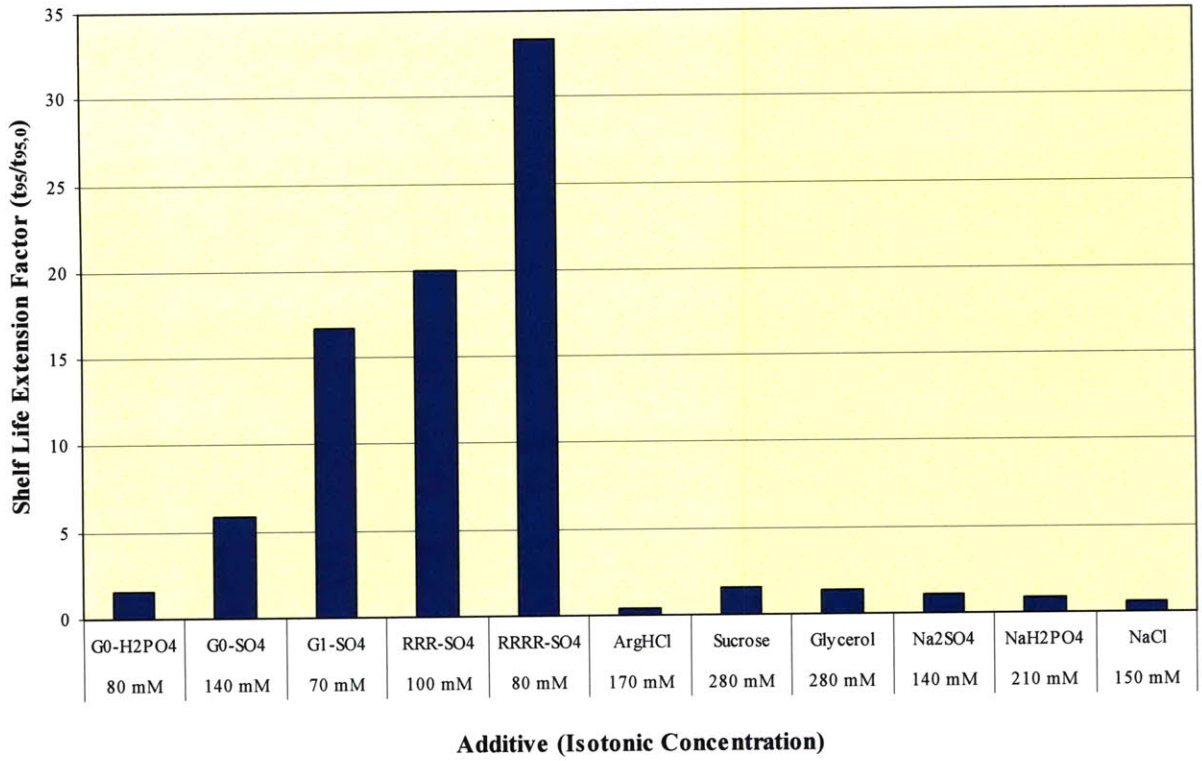

Figure 5-17: Shelf Life Extension Factor for Concanavalin A (7.5 mg/mL, T = 37°C, Buffer = 40 mM Na-Phosphate, pH 6.5) at isotonic concentrations for the novel excipients and commonly used excipients.

5.4. Thermodynamic Analysis

5.4.1. Preferential Interaction Coefficient

The preferential interaction coefficient for the novel excipients were analyzed to give insight into how they interact with proteins and influence aggregation. Figure 5-18 depicts the interaction between the diguanidinium compounds (chloride salt forms) and aCgn. Figure 5-18a shows Γ_{μ_3} values versus concentration. The error bars have been left off for clarity and lines have been drawn through the data to aid the eye. The data confirms the hypothesis that there is a strong interaction between the excipients and the protein. There is also a clear trend between the magnitude of the attractive interaction and the size of the diguanidinium compound, which is best represented in Figure 5-18b. In that figure, Γ_{μ_3} values at a concentration of 0.5 M are shown versus carbon chain length. The Γ_{μ_3} value for guanidinium chloride represents the value at a carbon chain length of 0. Initially, when the size of the molecule is increased, the attraction decreases, most likely due to larger volume exclusion. However, as the size continues to increase, the attractive interaction increases. As mentioned before, the best hypothesis for this

phenomenon is that there is likely a greater ability for cooperative binding in addition to increased hydrophobic interactions.

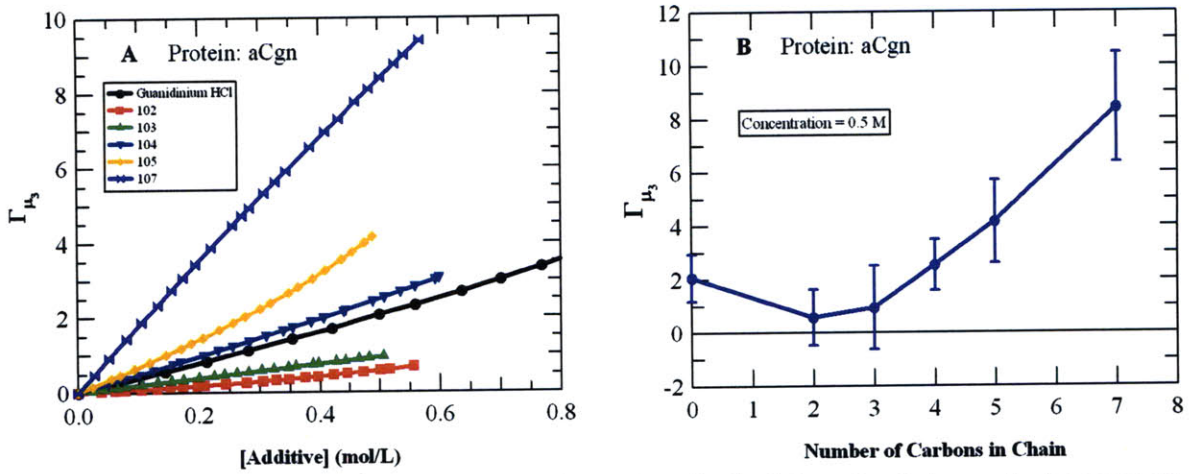

Figure 5-18: Preferential interaction coefficient values for the interaction between the 1n series of guanidinium compounds and aCgn (pH 5). (A) Preferential interaction coefficient values versus additive concentration, (B) Preferential interaction coefficient values versus additive size at a concentration of 0.5 M.

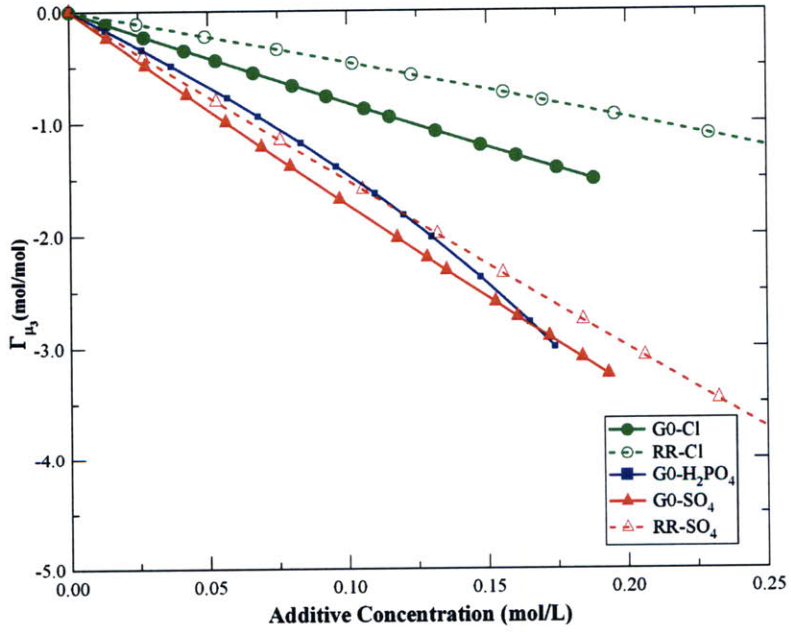

Figure 5-19: Preferential interaction coefficient values for the interaction of the G0 dendrimer salts and arginine sulfate dimers with aCgn (pH 5). The error bars have been left off for clarity and lines have been drawn through the data to aid the eye.

Figure 5-19 shows the preferential interaction coefficient values for the G0 dendrimer compounds and the arginine dimer peptides. Only the preferential interaction coefficient of the arginine dimer was measured. Larger peptides were not analyzed due to a lack of a sufficient quantity of peptide for such measurements but also due to the near identical behavior amongst

the peptides and near identical behavior to the diguanidinium compounds (that is in regards to aggregation suppression). Therefore, it was assumed that the peptides would also exhibit near identical preferential interaction behavior, thus it was deemed unnecessary to spend the time and resources measuring the parameter for all of the peptides.

Contrary to earlier predictions, the chloride form for both types of excipients exhibits a negative preferential interaction for all concentrations for both types of compounds, indicating that it is excluded from the protein surface. This is likely due to the large size inhibiting the compounds from fully solvating the local domain. However, as will be discussed below, this exclusion does not prevent the molecule from binding to the unfolded state and denaturing the protein, nor does it mean that the cosolute molecules in the local domain are not forming an attractive interaction with the protein surface. This result only indicates that the concentration of cosolute in the local domain is less than the concentration in the bulk solution. There is likely an attractive interaction formed between the local cosolute molecules and the protein. The other salt forms (phosphate and sulfate) can be considered naturants, even though they exhibit near identical preferential interaction coefficient values as the chloride form, though those salt forms are slightly more excluded. As discussed in the section on arginine salts (Chapter 4), the phosphate and sulfate ions form attractive interactions with the guanidinium functional group, something the chloride ion does not do, which likely inhibits the compounds from interacting with the unfolded state.

The key feature to note is that the preferential interaction coefficient for the sulfate and chloride dendrimer and peptide excipients exhibits a linear trend with respect to concentration. However, the phosphate dendrimer exhibits a slightly nonlinear trend. The arginine salts exhibit a nonlinear trend due to Arg-Arg interactions. It is clear that such interactions are not possible for the peptides, due to the carboxylate functional group being part of the peptide backbone, and are not possible for the dendrimers either due to the spherical structure of the polymer compounds, hence the linear trend. However, this does not exclude the possibility of salt bridges forming when the counterion is sulfate or phosphate, as observed for the arginine salts. Such conclusion in regards to cluster formation and salt bridges for the novel excipients is supported by MD simulation results. Preliminary MD results show no clustering but they show salt bridges for the sulfate and phosphate forms but not the chloride forms.

5.4.2. DSC

The denaturing midpoint temperature (T_m) of aCgn in the presence of the dendrimer and peptide compounds was determined via differential scanning calorimetry (DSC). Figure 5-20 depicts the unfolding transition excess heat capacity versus temperature. The figure shows the difference in the denaturing midpoint temperature (*i.e.* the temperature at the peak of the curve) of aCgn when in the presence of the different G0 dendrimers. Table 5-6 gives the denaturation midpoint temperature increments for the different dendrimer and peptide salts, in addition to ArgHCl, sucrose, and guanidinium chloride for comparison. The G0 dendrimer chloride salt lowers the melting temperature of aCgn at a rate nearly double that of guanidinium chloride, which indicates that the compound forms a stronger attraction to the unfolded state and is a much more potent denaturant. The arginine dimer (chloride salt) also lowers the melting temperature, but to a much lower extent. This is not surprising given that the compound only has two guanidinium functional groups and stabilizes proteins over a fairly broad concentration range. But at a concentration of 1 M, the dimer will lower the melting temperature by nearly 1°C, which is significant and is likely the reason why the compound speeds up aggregation at such concentrations.

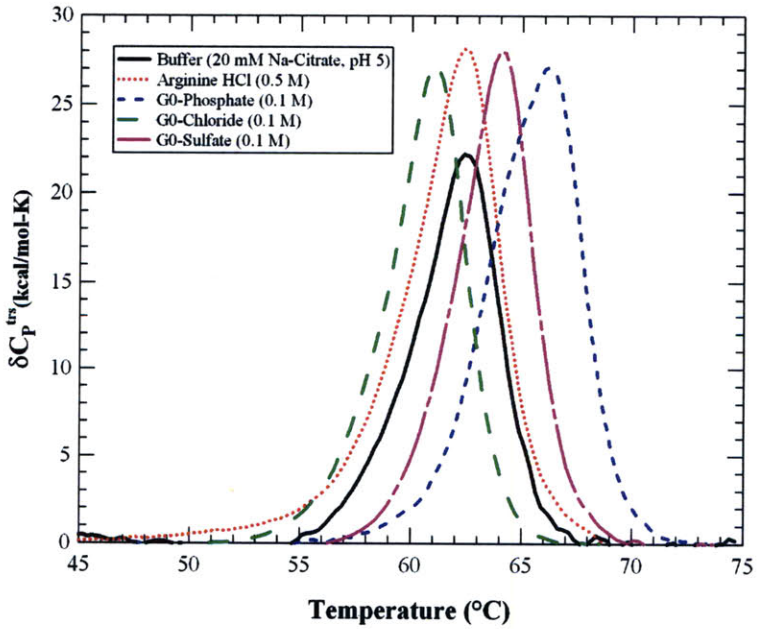

Figure 5-20: DSC scans of aCgn in the presence of G0 dendrimer salts.

Most importantly, the results show that the phosphate and sulfate G0 dendrimers and the sulfate peptides raise the unfolding temperature significantly. The G0 dendrimer phosphate salt

raises the unfolding temperature at a rate of 37.4 °C/M, the G0 dendrimer sulfate salt raises it at a rate of 15.2 °C/M, and the arginine dimer sulfate salt raises it at a rate of 10.4 °C/M (which is more than that for monomeric $ArgH(SO_4)_{1/2}$). All values are significantly higher than that for sucrose (0.7 °C/M), which is a commonly used conformational stabilizer, thus indicating that the phosphate and sulfate salts of the novel excipients are potent naturants. However, the trends in the aggregation behavior make it seem that these salt forms also inhibit association as well.

Table 5-6: Denaturing midpoint temperature increments for aCgn in the presence of G0 dendrimers, arginine dimers, and other excipients.

Additive	Max Conc.	$dT_m/d[3]$
	mol/L	°C/M
G0-Cl	0.1	-13.9
G0-H2PO4	0.1	37.4
G0-SO4	0.1	15.2
RR-Cl	0.25	-0.8
RR-SO4	0.25	10.4
ArgHCl	0.5	0.0
Sucrose	0.5	0.7
GdnHCl	0.2	-6.6

5.5. Mechanistic Inquiry

The hypothesis developed for the mechanism behind the arginine salts studied seems to extend, in part, to the dendrimers and peptide excipients. Phosphate, sulfate, and possibly citrate form attractive interactions with the guanidinium functional groups through hydrogen and electrostatic bonding. These interactions should cause the solutes to form salt bridges and inhibit them from binding to the unfolded state. The large size of the dendrimer and peptide excipients should cause them to be potent "neutral crowders", even though they are somewhat excluded from the surface of the protein. This is relative to the exclusion an inert molecule the same size as the compound should exhibit. If the excipients had no attraction to the proteins surface, their preferential interaction coefficient values would show them to be more than 4 times as excluded, making it reasonable to refer to them as "near-neutral" rather than "neutral" or "excluded". Conformational stabilization further enhances the aggregation suppression (as opposed to the denaturing effect of the chloride salt, which counteracts any association suppression that is observed at low concentrations). Moreover, the large size of the excipients and the bridged network of ions should also increase the viscosity of the solution and inhibit the motion of the protein molecules, thus causing an increased diffusional barrier to association.

MD simulations will have to be conducted to give further molecular level insight into the mechanism of action. However, the importance of the guanidinium functional group can easily be demonstrated and is illustrated below. Figure 5-21 compares the aggregation suppression ability of dendrimers with a guanidinium surface to dendrimers with the original ammonium surface (of course, dendrimers of the same size containing the same counterion were compared). A PAMAM dendrimer with an ammonium chloride surface has a similar profile to that of the guanidinium chloride surface, however, the relative rate constant for the guanidinium chloride dendrimer is significantly lower than the ammonium counterpart for all concentrations, indicating a greater ability for suppressing aggregation. In fact, at low concentrations, the relative rate constant for the ammonium chloride dendrimer is not much less than unity, indicating almost no ability for inhibiting aggregation. Furthermore, when the counterion is switched to phosphate, the ammonium dendrimer switches from an excipient that only marginally stabilizes proteins at low concentrations to one that inhibits aggregations at all concentrations, like its guanidinium counterpart (the trend plateaus at higher concentrations as well). However, the ability of the ammonium phosphate dendrimer to inhibit aggregation is an order of magnitude less than the guanidinium phosphate dendrimer. Moreover, the same behavior is observed when arginine is compared to lysine, the closest amino acid analogue to arginine. All of these results indicate that the guanidinium functional group plays an important role and is an essential feature in the mechanism of action.

These results support the hypothesis presented above in that the improved aggregation suppression is a result of the guanidinium functional group forming a strong interaction with the phosphate, sulfate, and citrate counterions. This can be said because even though the ammonium functional group can form hydrogen bonds, it cannot form hydrogen bonds with such counterions in the same manner as guanidinium because of its smaller size and fewer hydrogen bond donating sites. As a result, the interaction with the counterions is much weaker for ammonium and salt bridge formation is most likely nonexistent. The reason for the guanidinium chloride dendrimer being a better aggregation suppressor than the ammonium chloride dendrimer at low concentrations is most likely due to the stronger attractive interaction the guanidinium dendrimer forms with proteins. Guanidinium chloride by itself inhibits aggregation at low concentrations due to it inhibiting protein-protein interactions, which is a result of its interaction with proteins, while ammonium chloride does not exhibit such behavior because it cannot form as strong of an

interaction with the protein surface. Likewise, this same mechanism should extend to the dendrimers.

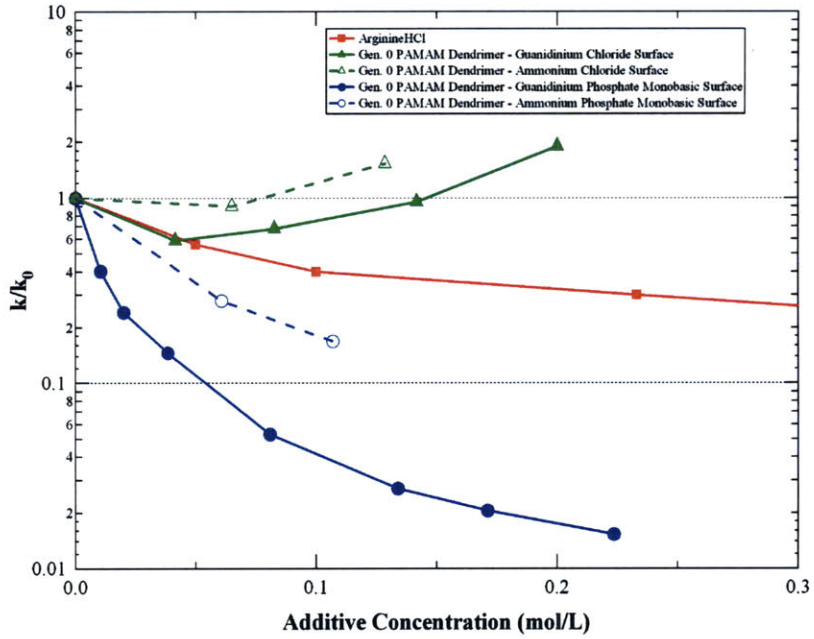

Figure 5-21: The impact the guanidinium surface has on aCgn monomer loss due to aggregation at 52.5°C. The figure compares dendrimers with a guanidinium surface to a dendrimer with an ammonium surface (same size and counterion). All solutions contained 10 mg/mL aCgn and were prepared in a 20 mM sodium citrate pH 5 buffer. Figure depicts the Relative Rate Constant (as determined by the amount of time for 20% monomer loss) versus additive concentration, with lines drawn through the plots to aid the eye.

A similar inquiry must be made into how much of the observed effect is the result of the dendrimer and how much is the result of the counterion. This question should be asked because sulfate and phosphate salts are common protein stabilizers and as shown in the figure above, the ammonium phosphate dendrimer inhibits aggregation as well. Therefore, the performance of the dendrimer phosphate and sulfate salts were compared to sodium phosphate and sodium sulfate (see Figure 5-22). One thing to note is that guanidinium dendrimers have multiple counterions, therefore, for a fair comparison, the performance of the dendrimers should be compared to the sodium analogues at a concentration 6 times that of the G0 phosphate dendrimer and 3 times that of the G0 sulfate dendrimer. When the relative rate constant for the G0 guanidinium phosphate dendrimer at 0.1 M is compared to sodium phosphate at 0.6 M, it is obvious that the dendrimer has a significantly lower rate constant (0.04 for the dendrimer and only 0.1 for sodium phosphate). A similar result is observed for the sulfate salts (a relative rate constant of 0.03 for the sulfate dendrimer at a concentration of 0.15 M and a value of 0.08 for sodium sulfate at a concentration of 0.45 M).

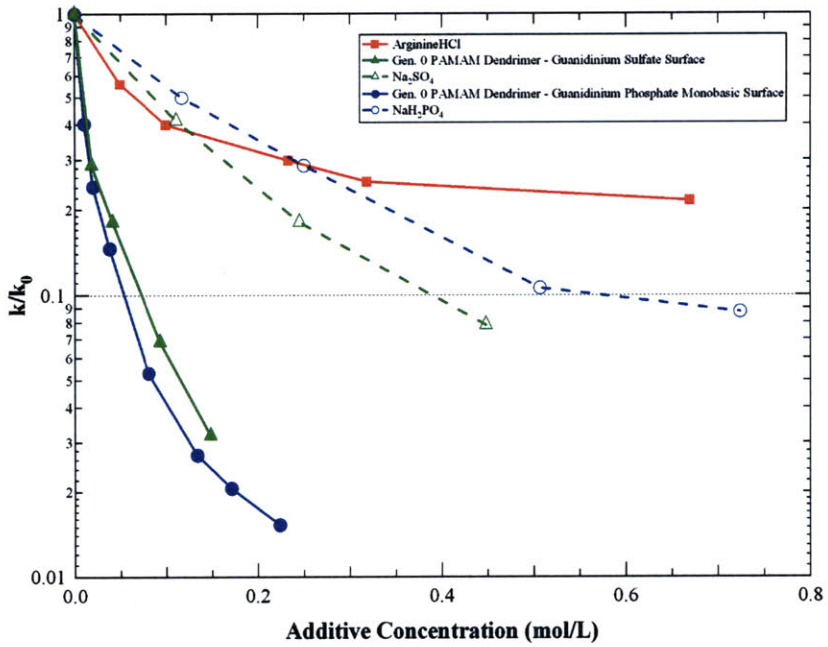

Figure 5-22: The impact the counterion has on aCgn monomer loss due to aggregation at 52.5°C. The figure compares dendrimer salts with the sodium salts containing the same anion. All solutions contained 10 mg/mL aCgn and were prepared in a 20 mM sodium citrate pH 5 buffer. Figure depicts the Relative Rate Constant (as determined by the amount of time for 20% monomer loss) versus additive concentration, with lines drawn through the plots to aid the eye.

For both types of salts, the guanidinium dendrimer version lowers the relative rate constant by a factor of 2.5 when compared at the same anion concentration, which indicates the impact the large molecule has on aggregation suppression. Contrary to this, the G0 ammonium phosphate dendrimer at 0.1 M only has a relative rate constant of 0.2. This indicates that the ammonium dendrimer has a detrimental effect on protein stability when compared to sodium, further exemplifying the role of the guanidinium functional group. Furthermore, it should be noted that sodium sulfate significantly reduces the solubility of the protein at concentrations above 0.5 M, showing that the dendrimer salts can inhibit aggregation without adversely affecting the solubility of the protein.

Chapter 6

6. Conclusions

6.1. Completion of Thesis Goals

As stated in Chapter 1, the main objectives of this thesis project were to:

1. Develop and test a novel class of excipients with the aim of dramatically improving the stabilization of protein therapeutics against aggregation.
2. Elucidate the mechanism by which arginine and the novel excipients inhibit aggregation.

As outlined in Chapter 4, a detailed study was conducted that has resulted in more insight into the mechanism by which arginine inhibits aggregation, in addition to showcasing additional arginine salts with improved aggregation suppression abilities. Furthermore, Chapter 5 details the development of a novel class of solutions additives and demonstrates their aggregation suppression abilities. Furthermore, the chapter includes thermodynamic data and a mechanistic inquiry, giving insight into the mechanism by which the novel excipients inhibit aggregation. The next step from here will involve implementing the developed excipients into pharmaceutical formulations and protein production steps.

6.2. Design Features Explored

In summary, three types of excipients were produced, modified amine compounds, modified PAMAM dendrimers, and arginine peptides. The modified amine compounds consisted of diguanidinium compounds and triguanidinium compounds with either two or three alkyl chains. Generation 0 through generation 2 PAMAM dendrimers were modified, each with an ethylene diamine core. The peptides consisted of small oligomers ranging in size from a dimer to a heptamer, most consisting of all arginine residues but some containing other residues such as glutamic acid and serine. The diguanidinium compounds showed that having a hydrophobic core is detrimental to the solubility of the solute. Moreover, the diguanidinium compounds demonstrated a counterintuitive behavior in regards to the larger compounds exhibiting a stronger rather than weaker interaction with proteins, which causes them to be more destabilizing. The dendrimers and peptides also exhibited this counterintuitive behavior.

It was discovered, though, that the key behind producing a stabilizing compounds lay, in part, with the choice of counterion to the guanidinium functional group. As described above, placing guanidinium functional groups on the surface of a large molecule to counter the strong repulsive interaction resulting from a large volume exclusion effect did not produce the results desired. The attractive interaction between the compounds and the protein were stronger than expected. However, exchanging the chloride ion with a counterion with a strong hydrogen bond accepting characteristic reduced this interaction due to a favorable interaction between the guanidinium functional group and the counterion. A study of different arginine counterions indicated that such counterions lead to cosolute clustering, which is likely the key to why the counterions enhance aggregation suppression. This mechanism likely extends to the dendrimer and peptide excipients.

6.3. Novel Excipients Produced

In summary, there are three types of novel excipients produced:

1. **Arginine Salts** (*e.g.* arginine phosphate, arginine sulfate, arginine citrate, *etc.*)
2. **Guanidinium Modified PAMAM Dendrimers** (with either a phosphate, sulfate, or citrate counterion)
3. **Arginine Peptides** (with either a phosphate, sulfate, or citrate counterion)

6.4. Elucidation of Aggregation Suppression Mechanisms

Early results seemed to indicate that arginine hydrochloride may be a naturally occurring "neutral crowder" compound. However, a more detailed investigation showed that the mechanism was more complex since arginine exhibited an unusual preferential interaction behavior. Furthermore, arginine peptides were shown to be destabilizing, indicating that the arginine mechanism was not simply a balance of attraction and repulsion with the protein surface. Rather, the behavior of arginine in solution is as important as how it interacts with the protein surface. MD simulations indicated that this behavior was likely due to attractive interactions between arginine molecules, which cause arginine molecules to form clusters in solution, something the arginine peptides could not do. Furthermore, an investigation into other counterions showed that attractive cation-anion interactions in solution lead to the formation of even larger and more extensive clusters. This effective size enhancement of the molecules around the protein inhibits arginine from binding too strongly with the protein surface and

increases the "Gap effect", which enhances the inhibition of protein-protein association. Furthermore, cluster formation likely leads to an increase in the viscosity of the solution, thus lowering the diffusion of proteins in solution. Therefore, arginine affects the rate of aggregation via three mechanisms: 1) enhancing the native state conformational stability (only for phosphate, sulfate, and citrate salts), 2) increasing the barrier for protein-protein association (both for individual arginine molecules and arginine clusters) and 3) increasing the viscosity of the solution medium.

The novel excipients should also exhibit a similar mechanism due to their similarity with arginine. The dendrimers and peptides cannot self associate like individual arginine molecules due to their large and complex structure. Therefore, they tend to interact strongly with the surface of the protein, rather than with other cosolute molecules in the bulk solution. However, when the counterion is phosphate, sulfate, or citrate, the novel excipients should form salt bridges in solution due to the interaction between guanidinium and those counterions. This clustering reduces the interaction with the protein surface, eliminating and reversing the conformation destabilization induced by the compounds, and enhancing protein association inhibition.

In conclusion, neither arginine nor the novel excipients are simple "neutral crowder" compounds. A simple "neutral crowder" compound likely cannot be created by balancing attraction with volume exclusion due to an unexpected cooperative interaction with the protein surface.

6.5. Significance of Results

The arginine salts and novel excipients exhibit an aggregation suppression ability that is greater than commonly used excipients. At isotonic concentrations, the arginine salts allow for a shelf life that is more than double that for solutions containing ArgHCl. Moreover, the peptides can quadruple the shelf life when compared to ArgHCl and the dendrimers can extend the shelf life even further, by a factor of 8. With aggregation suppression an order of a magnitude greater than commonly used excipients, the novel excipients should be useful for reducing the aggregation of diagnostic and therapeutic proteins in many different applications, such as in formulations, production, purification, clinical use, and research. Furthermore, the mechanistic inquiry into the mechanism of action of arginine and the novel excipients reveals a lot about how

to further improve the excipients developed, possibly even allowing for the creation of other excipient types.

Chapter 7

7. Future Work

7.1. Other Considerations

The main objective during this study was to develop an excipient which inhibits the aggregation of a model protein under accelerated conditions. This and only this was considered due to the large amount of work involved in synthesizing and assessing a novel class of excipients. Even though this objective was met, there are a lot of other factors that need to be considered before the excipients can be utilized. These include but are not limited to determining:

1. Whether or not the excipients are toxic and determining if they can be included in a formulation if found to be nontoxic.
2. Whether or not the novel excipient inhibit the aggregation of other proteins (therapeutic and diagnostic proteins in particular).
    a. Do the excipients work universally?
    b. If not, which type of proteins do they work best on?
    c. What is the mechanism for this if that is the case?
    d. Can the formulation conditions be tuned for them to work universally?
3. Whether or not they truly inhibit aggregation at room or refrigerated temperatures.
4. How the excipients influence the maximum concentration of a protein. For example:
    a. Do they diminish it because of their size?
    b. Will they help to increase the maximum concentration due to slower aggregation?
5. Whether or not the excipients cause the viscosity to be too high for them to be utilized.
6. The optimal pH range and formulation conditions, in addition to compatibility with other commonly used excipients.
7. Whether or not the novel excipients inhibit the aggregation caused by other stresses (*e.g.* shaking and shearing, freezing and thawing, dialysis, ultrafiltration, refolding, *etc.*).

8. How well they inhibit aggregation during protein production steps.
9. How the excipients influence other degradation pathways (*e.g.* deamidation, oxidation, *etc.*).
10. Whether or not the purity of the excipients needs to be higher (*e.g.* the impurities might be toxic or detrimental to the long term stability of proteins).
11. Whether or not the excipients themselves remain stable over time.
12. How easily the excipients can be removed from a solution if found to be not applicable for formulations but are utilized elsewhere during production and purification.

7.2. Improved Excipients

Now that a group of working excipients has been developed and the mechanism by which they inhibit aggregation somewhat elucidated, it is clear that there are many other variations that can be explored to further improve aggregation suppression performance.

7.2.1. Higher Purity

The purity of the dendrimer excipients is somewhat low (<80%) due to residual solvents. This remains to be the case even when the dendrimers are dried over a longer period of time. It was apparent after the ion exchange experiments that simply dissolving the chloride form of the dendrimer excipients in water and evaporating away the water via lyophilization that much of the residual solvent is removed and this resulted in an improved performance of the chloride form. Therefore, it is probably best to purify the dendrimer excipients via a more robust method (*e.g.* HPLC, dialysis, *etc.*) to obtain the excipients in the highest purity possible and to eliminate any undesirable impurities. This will likely improve the performance of the sulfate and phosphate form of the dendrimer excipients (even though they should already be more pure than the original chloride form they were produced from).

7.2.2. Synthesizing PAMAM Dendrimers from Scratch

As discussed earlier, one major drawback of the excipients developed is their high tonicity resulting from many counterions per molecule. It might be beneficial to make the excipients larger while keeping the number of functional groups on the surface the same. One could simply utilize the commercially available PAMAM dendrimers with larger cores, however, that could be detrimental to the smaller dendrimers due to an increase in the hydrophobic

character of the core structure, which is more exposed to the solvent in the smaller generations. Moreover, the smaller generations are likely to be the most optimal due to a higher isotonic concentration and lower toxicity, making it even less desirable to use the dendrimers with a larger core.

Even though they might not give beneficial results, the larger core dendrimers should still be explored to gain more insight into the most optimal structure. However, if one wants to make the dendrimers larger by making the branches longer, the dendrimers will have to be synthesized and purified from scratch. The commercially available PAMAM dendrimers have alkyl chains two carbons long on either side of the amide functional group because the reactants utilized to make the branches (methyl acrylate and ethylene diamine) are used specifically for their volatility (*i.e.* after the synthesis step, the reactants are easily removed via simple evaporation). To make longer branches, reactants with longer alkyl chains will have to be utilized. However, such reagents will become less and less volatile, thus a different synthesis/purification method will have to be utilized. It might be possible to synthesize the dendrimers via a solid phase synthesis technique, growing the dendrimers on a polymer substrate and washing away the reactants after each step rather than evaporating them. After the synthesis, the dendrimers can be cleaved from the addition of an acid. Furthermore, the surface of the dendrimers can be modified prior to cleavage, making the synthesis and purification less cumbersome on top of producing a product with a much higher purity without the use of a purification step. However, the synthesis of larger dendrimers might be troublesome due to an increased likelihood of crosslinking nearby branches, which is another reason why only smaller reactants are utilized.

7.2.3. Incorporating Other Amino Acids

The arginine peptides are a much more attractive group of excipients to work with in regards to making larger excipients with the same number of functional groups because the method to do so is quite simple. All one has to do is introduce other residues into the sequence which are hydrophilic but at the same time not electrolytic (*e.g.* serine). The structures that can be produced are essentially limitless, thus many different design features can be explored, allowing for the behavior of the excipient to be fine tuned. It might even be possible to tune the excipient on a case by case basis to obtain the maximum aggregation suppression possible.

7.2.4. Surface Modified to Urea

The other protein binding functional group that can possibly be utilized in the same manner as guanidinium is urea. Therefore it might be possible to modify the surface of the excipients to urea (including using lysine peptides as a starting material). However, urea has no counterion, thus the behavior of such compounds cannot be influenced by introducing a different counterion. But this might not be necessary because urea does not bind to proteins as strongly as guanidinium. If a urea modified excipient is found to inhibit aggregation, it will be beneficial because it can be utilized at higher concentrations due to a much lower tonicity resulting from a lack of counterions (that is if such compounds are soluble, which might be an issue if urea is the surface group). Furthermore, they will likely be less toxic since they will not have a positively charged surface. Large molecules with a positively charged surface have been shown to be toxic to cell lines and detrimental to red blood cells. Neutral and negatively charged compounds do not exhibit this behavior [113].

7.3. Applications

7.3.1. Formulation

Throughout this project, the main goal has been to develop new excipients for use in formulations. However, as outlined above, it will not be a straightforward or a quick process for the novel excipients to be implemented as formulation additives. There are many things to consider, with the toxicological effects at the top of the list. If the excipients pass all of the considerations outlined above, they will likely be of use in therapeutic and diagnostic formulations.

7.3.2. Production & Purification

A more immediate and possibly greater use for the excipients might be as additives during production and purification. Several of the steps involved during the production and purification of proteins are detrimental to the structure of the protein and can cause rapid aggregation. This aggregation can greatly reduce the yield of a costly product and places a huge burden on downstream purification steps. One example of this is Protein A affinity chromatography, which is often the first purification step in the production of antibodies. Antibodies bind to the column when dissolved at a moderate pH and are eluted by lowering the pH to around 2-3. This extreme pH is detrimental to the protein structure and causes significant

aggregation [114]. The novel excipients inhibit protein-protein interactions and might be of use eluting the antibodies at a higher and less detrimental pH, similar to the effect arginine has on Protein A elution. Furthermore, they will help to reduce the aggregation that occurs during elution and the virus hold step. Other production and purifications steps may include fermentation, refolding, affinity chromatography, *etc.* Furthermore, ionic exchange and size exclusion chromatography steps are often utilized during the polishing steps, which can be utilized to remove the excipients from the solution if they are not desired or cannot be used in the final formulation recipe.

7.4. Toxicological Study

As mentioned before, large molecules with a positively charged surface have been shown to be toxic to cell lines and detrimental to red blood cells. Therefore, the guanidinium modified compounds might be too toxic to be used as a formulation additive. Furthermore, small peptides can cause allergic reactions, making an immunogenic and toxicological study absolutely essential. Even though PAMAM dendrimers have received a tremendous amount of attention in regards to therapeutic uses, only a relatively small amount of research has been devoted to determining their toxicological effects, likely due to the high cost of such studies [113]. The research that has been conducted has shown an *in vitro* toxic effect that is generation dependent. The larger PAMAM dendrimers are more toxic, while the smaller generations seem promising for therapeutic use. *in vivo* studies with mice show no detrimental effect in regards to the life span or behavior of the mice, except for the very large generation 7 PAMAM dendrimer. Furthermore, the dendrimers showed no immunological effects, which is a positive sign. Similar and more rigorous studies will have to be conducted for the modified dendrimers before anyone can begin to implement the excipients into formulations.

Cited References

1. Leader, B., Q.J. Baca, and D.E. Golan, *Protein therapeutics: a summary and pharmacological classification.* Nature Reviews, 2008. **7**: p. 21-39.

2. Wang, W., *Instability, stabilization, and formulation of liquid protein pharmaceuticals.* International Journal of Pharmaceutics, 1999. **185**: p. 129-188.

3. Frokjaer, S. and D.E. Otzen, *Protein drug stability: A formulation challenge.* Nature Reviews Drug Discovery, 2005. **4**(4): p. 298-306.

4. Randolph, T.W. and J.F. Carpenter, *Engineering challenges of protein formulations.* Aiche Journal, 2007. **53**(8): p. 1902-1907.

5. Wang, W., *Protein aggregation and its inhibition in biopharmaceutics.* International Journal of Pharmaceutics, 2005. **289**: p. 1-30.

6. Carpenter, J.F., et al., *Inhibition of Stress-Induced Aggregation of Protein Therapeutics.* Methods in Enzymology, 1999. **309**: p. 236-255.

7. Wang, W., et al., *Antibody structure, instability, and formulation.* Journal of Pharmaceutical Sciences, 2007. **96**(1): p. 1-26.

8. Cleland, J.L., M.F. Powell, and S.J. Shire, *The Development of Stable Protein Formulations - a Close Look at Protein Aggregation, Deamidation, and Oxidation.* Critical Reviews in Therapeutic Drug Carrier Systems, 1993. **10**(4): p. 307-377.

9. Arakawa, T., et al., *Suppression of protein interactions by arginine: A proposed mechanism of the arginine effects.* Biophysical Chemistry, 2007. **127**: p. 1-8.

10. Baynes, B.M. and B.L. Trout, *Rational design of solution additives for the prevention of protein aggregation.* Biophysical Journal, 2004. **87**(3): p. 1631-1639.

11. Lumry, R. and H. Eyring, *Conformation changes of proteins.* Journal of Physical Chemistry, 1954. **58**(2): p. 110-120.

12. Timasheff, S.N., *Control of protein stability and reactions by weakly interacting cosolvents: The simplicity of the complicated.* Advances in Protein Chemistry, 1998. **51**: p. 355-432.

13. Frokjaer, S. and D.E. Otzen, *Protein Drug Stability: A Formulation Challenge.* Nat. Rev. Drug Discovery, 2005. **4**: p. 298-306.

14. Minton, A.P., *Influence of excluded volume upon macromolecular structure and associations in 'crowded" media.* Current Opinion in Biotechnology, 1997. **8**(1): p. 65-69.

15. Baynes, B.M., D.I.C. Wang, and B.L. Trout, *Role of arginine in the stabilization of proteins against aggregation.* Biochemistry, 2005. **44**(12): p. 4919-4925.

16. Tsumoto, K., et al., *Role of Arginine in Protein Refolding, Solubilization, and Purification.* Biotechnology Progress, 2004. **20**(5): p. 1301-1308.

17. K., R.C.R., et al., *L-Arginine increases the solubility of unfolded species of hen egg white lysozyme.* Protein Science, 2005. **14**(4): p. 929-935.

18. Arakawa, T. and K. Tsumoto, *The effects of arginine on refolding of aggregated proteins: not facilitate refolding, but suppress aggregation.* Biochemical and Biophysical Research Communications, 2003. **304**(1): p. 148-152.

19. Shiraki, K., et al., *Biophysical Effect of Amino Acids on the Prevention of Protein Aggregation.* Journal of Biochemistry, 2002. **132**(4): p. 591-595.

20. Liu, Y.-d., et al., *A newly proposed mechanism for arginine-assisted protein refolding -- not inhibiting soluble oligomers although promoting a correct structure.* Protein Expression and Purification, 2007. **51**: p. 235-242.

21. Kudou, M., et al., *Prevention of thermal inactivation and aggregation of lysozyme by polyamines.* European Journal of Biochemistry, 2003. **270**: p. 4547-4554.

22. Shiraki, K., et al., *Arginine ethylester prevents thermal inactivation and aggregation of lysozyme.* European Journal of Biochemistry, 2004. **271**: p. 3242-3247.

23. Shiraki, K., et al., *Amino acid esters prevent thermal inactivation and aggregation of lysozyme.* Biotechnology Progress, 2005. **21**(2): p. 640-643.

24. Okanojo, M., et al., *Diamines Prevent Thermal Aggregation and Inactivation of Lysozyme.* Journal of Bioscience and Bioengineering, 2005. **100**(5): p. 556-561.

25. Hamada, H. and K. Shiraki, *L-Argininamide improves the refolding more effectively than L-arginine.* Journal of Biotechnology, 2007. **130**: p. 153-160.

26. Timasheff, S.N., *The Control of Protein Stability and Association by Weak Interactions With Water: How Do Solvents Affect These Processes?* Annual Review of Biophysics and Biomolecular Structure, 1993. **22**: p. 67-97.

27. Courtenay, E.S., et al., *Vapor Pressure Osmometry Studies of Osmolyte-Protein Interactions: Implications for the Action of Osmoprotectants in Vivo and for the Interpretation of "Osmotic Stress" Experiments in Vitro.* Biochemistry, 2000. **39**(15): p. 4455-4471.

28. Scatchard, G., *Physical chemistry of protein solutions. I. Derivation of the equations for the osmotic pressure.* Journal of The American Chemical Society, 1946. **68**: p. 2315-2319.

29. Baynes, B.M. and B.L. Trout, *Proteins in mixed solvents: A molecular-level perspective.* Journal of Physical Chemistry B, 2003. **107**(50): p. 14058-14067.

30. Record, M.T. and C.F. Anderson, *Interpretation of Preferential Interaction Coefficients of Nonelectrolytes and of Electrolyte Ions in Terms of a Two-Domain Model.* Biophysical Journal, 1995. **68**: p. 786-794.

31. Gekko, K. and S.N. Timasheff, *Mechanism of Protein Stabilization by Glycerol: Preferential Hydration in Glycerol-Water Mixtures.* Biochemistry, 1981. **20**: p. 4667-4676.

32. Arakawa, T. and S.N. Timasheff, *Stabilization of Protein Structure by Sugars.* Biochemistry, 1982. **21**: p. 6536-6544.

33. Arakawa, T. and S.N. Timasheff, *Protein Stabilization and Destabilization by Guanidinium Salts.* Biochemistry, 1984. **23**: p. 5924-5929.

34. Timasheff, S.N. and G. Xie, *Preferential interactions of urea with lysozyme and their linkage to protein denaturation.* Biophysical Chemistry, 2003. **105**: p. 421-448.

35. Casassa, E.F. and H. Eisenberg, *Partial Specific Volumes of Refractive Index Increments In Multicomponent Systems.* Journal of Physical Chemistry, 1961. **65**: p. 427-433.

36. Casassa, E.F. and H. Eisenberg, *Thermodynamic analysis of multicomponent solutions.* Advances in Protein Chemistry. Vol. 19. 1964, New York: Academic Press. 287-395.

37. Zhang, W., et al., *Thermodynamic Characterization of Interactions of Native Bovine Serum Albumin with Highly Excluded (Glycine Betaine) and Moderately Accumulated (Urea) Solutes by a Novel Application of Vapor Pressure Osmometry.* Biochemistry, 1996. **35**(32): p. 10506-10516.

38. Anderson, C.F., et al., *Generalized derivation of an exact relationship linking different coefficients that characterize thermodynamic effects of preferential interactions.* Biophysical Chemistry, 2002. **101-102**: p. 497-511.

39. Timasheff, S.N., *Preferential Interactions of Water and Cosolvents with Proteins*, in *Protein-Solvent Interactions*, R.B. Gregory, Editor. 1995, Marcel Dekker, Inc.: New York. p. 445-482.

40. Wyman Jr., J., *Linked functions and reciprocal effects in hemoglobin: a second look.* Advances in Protein Chemistry. Vol. 19. 1964, New York: Academic Press. 223-286.

41. Roberts, C.J., R.T. Darrington, and M.B. Whitley, *Irreversible aggregation of recombinant bovine Granulocyte-Colony Stimulating Factor (bG-CSF) and implications for predicting protein shelf life.* Journal of Pharmaceutical Sciences, 2003. **92**(5): p. 1095-1111.

42. Arakawa, T. and S.N. Timasheff, *Mechanism of Poly(Ethylene Glycol) Interaction with Proteins.* Biochemistry, 1985. **24**(24): p. 6756-6762.

43. Tsumoto, K., *Role of arginine in protein refolding, solubilization, and purification.* Biotechnology progress, 2004. **20**(5): p. 1301-1308.

44. Katritzky, A.R. and B.V. Rogovoy, *Recent developments in guanylating agents.* Arkivoc, 2005. **4**: p. 49-87.

45. Gers, T., et al., *Reagents for efficient conversion of amines to protected guanidines.* Synthesis-Stuttgart, 2004(1): p. 37-42.

46. Vaiana, S.M., et al., *Irreversible Formation of Intermediate BSA Oligomers Requires and Induces Conformational Changes.* PROTEINS: Structure, Function, and Bioinformatics, 2004. **55**(1053-1062).

47. Samuel, D., et al., *Proline inhibits aggregation during protein refolding.* Protein Science, 2000. **9**(2): p. 344-352.

48. Panick, G., R. Malessa, and R. Winter, *Differences between the pressure- and temperature-induced denaturation and aggregation of beta-lactoglobulin A, B, and AB monitored by FT-IR spectroscopy and small-angle X-ray scattering.* Biochemistry, 1999. **38**(20): p. 6512-6519.

49. Yan, Y.B., et al., *Protein thermal aggregation involves distinct regions: Sequential events in the heat-induced unfolding and aggregation of hemoglobin.* Biophysical Journal, 2004. **86**(3): p. 1682-1690.

50. Weijers, M., et al., *Heat-induced denaturation and aggregation of ovalbumin at neutral pH described by irreversible first-order kinetics.* Protein Science, 2006. **12**: p. 2693-2703.

51. Gotte, G., F. Vottariello, and M. Libonati, *Thermal aggregation of ribonuclease A - A contribution to the understanding of the role of 3D domain swapping in protein aggregation.* Journal of Biological Chemistry, 2003. **278**(12): p. 10763-10769.

52. Sluzky, V., et al., *Kinetics Of Insulin Aggregation In Aqueous-Solutions Upon Agitation In The Presence Of Hydrophobic Surfaces.* Proceedings of The National Academy of Sciences of The United States Of America, 1991. **88**(21): p. 9377-9381.

53. Grillo, A.O., et al., *Conformational origin of the aggregation of recombinant human factor VIII.* Biochemistry, 2001. **40**(2): p. 586-595.

54. Vermeer, A.W.P. and W. Norde, *The thermal stability of immunoglobulin: Unfolding and aggregation of a multi-domain protein.* Biophysical Journal, 2000. **78**(1): p. 394-404.

55. Otzen, D.E., et al., *Structural basis for cyclodextrins' suppression of human growth hormone aggregation.* Protein Science, 2002. **11**(7): p. 1779-1787.

56. Chi, E.Y., et al., *Roles of conformational stability and colloidal stability in the aggregation of recombinant human granulocyte colony-stimulating factor.* Protein Science, 2003. **12**(5): p. 903-913.

57. Webb, S.D., et al., *A new mechanism for decreasing aggregation of recombinant human interferon-gamma by a surfactant: Slowed dissolution of lyophilized formulations in a solution containing 0.03% polysorbate 20.* Journal of Pharmaceutical Sciences, 2002. **91**(2): p. 543-558.

58. Andrews, J.M. and C.J. Roberts, *Non-Native Aggregation of α-Chymotrypsinogen Occurs through Nucleation and Growth with Competing Nucleus Sizes and Negative Activation Energies.* Biochemistry, 2007. **46**: p. 7558-7571.

59. Andrews, J.M., W.F. Weiss, and C.J. Roberts, *Nucleation, growth, and activation energies for seeded and unseeded aggregation of alpha-chymotrypsinogen A.* Biochemistry, 2008. **47**(8): p. 2397-2403.

60. Li, Y., B.A. Ogunnaike, and C.I. Roberts, *Multi-Variate Approach to Global Protein Aggregation Behavior and Kinetics: Effects of pH, NaCl, and Temperature for alpha-Chymotrypsinogen A.* Journal Of Pharmaceutical Sciences. **99**(2): p. 645-662.

61. Chalikian, T.V., et al., *The Native and the Heat-induced Denatured States of α-Chymotrypsinogen A: Thermodynamic and Spectroscopic Studies.* Journal of Molecular Biology, 1997. **274**: p. 237-252.

62. Velev, O.D., E.W. Kaler, and A.M. Lenhoff, *Protein interactions in solution characterized by light and neutron scattering: Comparison of lysozyme and chymotrypsinogen.* Biophysical Journal, 1998. **75**(6): p. 2682-2697.

63. Coen, C.J., H.W. Blanch, and J.M. Prausnitz, *Salting-Out Of Aqueous Proteins - Phase-Equilibria And Intermolecular Potentials.* Aiche Journal, 1995. **41**(4): p. 996-1004.

64. Gold, A.M. and D. Fahrney, *Sulfonyl Fluorides As Inhibitors Of Esterases.2. Formation + Reactions Of Phenylmethanesulfonyl Alpha-Chymotrypsin.* Biochemistry, 1964. **3**(6): p. 783-&.

65. Pitzer, K.S., *Thermodynamics Of Electrolytes.1. Theoretical Basis And General Equations.* Journal of Physical Chemistry, 1973. **77**(2): p. 268-277.

66. Pitzer, K.S. and G. Mayorga, *Thermodynamics Of Electrolytes.2. Activity And Osmotic Coefficients For Strong Electrolytes With One Or Both Ions Univalent.* Journal of Physical Chemistry, 1973. **77**(19): p. 2300-2308.

67. Das, B., *Pitzer ion interaction parameters of single aqueous electrolytes at 25 degrees C.* Journal of Solution Chemistry, 2004. **33**(1): p. 33-45.

68. Lange, C. and R. Rudolph, *Suppression of Protein Aggregation by L-Arginine.* Current Pharmaceutical Biotechnology, 2009. **10**(4): p. 408-414.

69. Buchner, J. and R. Rudolph, *Renaturation, Purification and Characterization of Recombinant Fab-Fragments Produced in Escherichia-Coli.* Bio-Technology, 1991. **9**(2): p. 157-162.

70. Schneider, C.P. and B.L. Trout, *Investigation of Cosolute-Protein Preferential Interaction Coefficients: New Insight into the Mechanism by Which Arginine Inhibits Aggregation.* Journal of Physical Chemistry B, 2009. **113**(7): p. 2050-2058.

71. Arakawa, T., et al., *The effects of arginine on protein binding and elution in hydrophobic interaction and ion-exchange chromatography.* Protein Expression and Purification, 2007. **54**(1): p. 110-116.

72. Das, U., et al., *Inhibition of Protein Aggregation: Supramolecular Assemblies of Arginine Hold the Key.* PLoS One, 2007. **2**(11): p. e1176.

73. Srinivas, V., et al., *Structural perturbation and enhancement of the chaperone-like activity of alpha-crystallin by arginine hydrochloride.* Protein Science, 2003. **12**(6): p. 1262-1270.

74. Ghosh, R., S. Sharma, and K. Chattopadhyay, *Effect of Arginine on Protein Aggregation Studied by Fluorescence Correlation Spectroscopy and Other Biophysical Methods.* Biochemistry, 2009. **48**(5): p. 1135-1143.

75. Shukla, D., C. Shinde, and B.L. Trout, *Molecular Computations of Preferential Interaction Coefficients of Proteins.* Journal of Physical Chemistry B, 2009. **113**(37): p. 12546-12554.

76. Shukla, D. and B.L. Trout, *Interactions of Arginine with Proteins and the Mechanism by Which Arginine Inhibits Aggregation.* Journal of Physical Chemistry B, 2010. **114**(42): p. 13426-13438.

77. Nakakido, M., et al., *To be Excluded or to Bind, that is the Question: Arginine Effects on Proteins.* Current Pharmaceutical Biotechnology, 2009. **10**(4): p. 415-420.

78. Arakawa, T., Y. Kita, and A.H. Koyama, *Solubility enhancement of gluten and organic compounds by arginine.* International Journal of Pharmaceutics, 2008. **355**(1-2): p. 220-223.

79. Hirano, A., et al., *The solubility of nucleobases in aqueous arginine solutions.* Archives of Biochemistry And Biophysics, 2010. **497**(1-2): p. 90-96.

80. Ohtake, S., T. Arakawa, and A.H. Koyama, *Arginine as a Synergistic Virucidal Agent.* Molecules, 2010. **15**(3): p. 1408-1424.

81. Arakawa, T., et al., *Stabilizing and destabilizing effects of arginine on deoxyribonucleic acid.* International Journal of Biological Macromolecules, 2010. **46**(2): p. 217-222.

82. Abe, R., et al., *Immobilized metal affinity chromatography in the presence of arginine.* Biochemical and Biophysical Research Communications, 2009. **381**(3): p. 306-310.

83. Shukla, D. and B.L. Trout, *Preferential Interaction Coefficients of Proteins in Aqueous Arginine Solutions and Its Molecular Origins.* Journal of Physical Chemistry B, (Submitted).

84. Courtenay, E.S., M.W. Capp, and J. M. Thomas Record, *Thermodynamics of interactions of urea and guanidinium salts with protein surface: Relationship between solute effects on protein processes and changes in water-accessible surface area.* Protein Science, 2001. **10**: p. 2485-2497.

85. Kita, Y., et al., *Contribution of the Surface Free Energy Perturbation to Protein-Solvent Interactions.* Biochemistry, 1994. **33**: p. 15178-15189.

86. Ishibashi, M., et al., *Is arginine a protein-denaturant?* Protein Expression and Purification, 2005. **42**: p. 1-6.

87. Hofmeister, F., *Zur Lehre von der Wirkung der Salze.* Arch. Exp. Pathol. Pharmakol, 1888. **24**: p. 247-260.

88. Baldwin, R.L., *How Hofmeister Ion Interactions Affect Protein Stability.* Biophysical Journal, 1996. **71**: p. 2056-2063.

89. Cacace, M.G., E.M. Landau, and J.J. Ramsden, *The Hofmeister series: salt and solvent effects on interfacial phenomena.* Quarterly Reviews of Biophysics, 1997. **30**(3): p. 241-277.

90. Curtis, R.A., et al., *Protein-protein interactions in concentrated electrolyte solutions - Hofmeister-series effects.* Biotechnology and Bioengineering, 2002. **79**(4): p. 367-380.

91. Zhang, Y.J. and P.S. Cremer, *Interactions between macromolecules and ions: the Hofmeister series.* Current Opinion in Chemical Biology, 2006. **10**(6): p. 658-663.

92. Tadeo, X., M. Pons, and O. Millet, *Influence of the Hofmeister anions on protein stability as studied by thermal denaturation and chemical shift perturbation.* Biochemistry, 2007. **46**(3): p. 917-923.

93. Lodderstedt, G., et al., *Hofmeister salts and potential therapeutic compounds accelerate in vitro fibril formation of the N-terminal domain of PABPN1 containing a disease-causing alanine extension.* Biochemistry, 2008. **47**(7): p. 2181-2189.

94. Maity, H., C. Karkaria, and J. Davagnino, *Effects of pH and Arginine on the Solubility and Stability of a Therapeutic Protein (Fibroblast Growth Factor 20): Relationship between Solubility and Stability.* Current Pharmaceutical Biotechnology, 2009. **10**(6): p. 609-625.

95. Golovanov, A.P., et al., *A Simple Method for Improving Protein Solubility and Long-Term Stability.* Journal of The American Chemical Society, 2004. **126**: p. 8933-8939.

96. Valente, J.J., et al., *Second Virial Coefficient Studies of Cosolvent-Induced Protein Self-Interaction.* Biophysical Journal, 2005. **89**(6): p. 4211-4218.

97. Ignatova, Z. and L.M. Gierasch, *Inhibition of protein aggregation in vitro and in vivo by a natural osmoprotectant.* Proceedings of The National Academy of Sciences of The United States of America, 2006. **103**(36): p. 13357-13361.

98. Jorgensen, L., et al., *Recent trends in stabilising peptides and proteins in pharmaceutical formulation - considerations in the choice of excipients.* Expert Opinion on Drug Delivery, 2009. **6**(11): p. 1219-1230.

99. Mason, P.E., et al., *The hydration structure of guanidinium and thiocyanate ions: Implications for protein stability in aqueous solution.* Proceedings of The National Academy of Sciences of The United States Of America, 2003. **100**(8): p. 4557-4561.

100. Mason, P.E., et al., *The structure of aqueous guanidinium chloride solutions.* Journal of The American Chemical Society, 2004. **126**(37): p. 11462-11470.

101. Mason, P.E., et al., *Nanometer-Scale Ion Aggregates in Aqueous Electrolyte Solutions: Guanidinium Sulfate and Guanidinium Thiocyanate.* Journal of Physical Chemistry B, 2005. **109**(50): p. 24185-24196.

102. Mason, P.E., et al., *Nanometer-scale ion aggregates in aqueous electrolyte solutions: Guanidinium carbonate.* Journal of Physical Chemistry B, 2006. **110**(27): p. 13477-13483.

103. Dempsey, C.E., et al., *The Reversal by Sulfate of the Denaturant Activity of Guanidinium.* Journal of The American Chemical Society, 2007. **129**(51): p. 15895-15902.

104. Mason, P.E., et al., *Specificity of Ion-Protein Interactions: Complementary and Competitive Effects of Tetrapropylammonium, Guanidinium, Sulfate, and Chloride Ions.* Journal of Physical Chemistry B, 2009. **113**(10): p. 3227-3234.

105. Bonner, O.D., *Osmotic And Activity Coefficients Of Some Guanidinium Salts At 298.15 K.* Journal Of Chemical Thermodynamics, 1976. **8**(12): p. 1167-1172.

106. Bonner, O.D., *Hydrogen-Bonding In Guanidinium Fluoride.* Journal of Physical Chemistry, 1977. **81**(24): p. 2247-2249.

107. Schlund, S., C. Schmuck, and B. Engels, *How important is molecular rigidity for the complex stability of artificial host-guest systems? A theoretical study on self-assembly of gas-phase arginine.* Chemistry-A European Journal, 2007. **13**(23): p. 6644-6653.

108. Vondrasek, J., et al., *The Molecular Origin of Like-Charge Arginine-Arginine Pairing in Water.* Journal of Physical Chemistry B, 2009. **113**(27): p. 9041-9045.

109. Kalcher, I. and J. Dzubiella, *Structure-thermodynamics relation of electrolyte solutions.* Journal of Chemical Physics, 2009. **130**(13).

110. Fyta, M., et al., *Ionic force field optimization based on single-ion and ion-pair solvation properties.* Journal of Chemical Physics, 2010. **132**(2).

111. Okuro, K., et al., *Molecular Glues Carrying Multiple Guanidinium Ion Pendants via an Oligoether Spacer: Stabilization of Microtubules against Depolymerization.* Journal of The American Chemical Society, 2009. **131**(5): p. 1626-1627.

112. Mislovicova, D., et al., *Lectinomics I. Relevance of exogenous plant lectins in biomedical diagnostics.* Biologia, 2009. **64**(1): p. 1-19.

113. Duncan, R. and L. Izzo, *Dendrimer biocompatibility and toxicity.* Advanced Drug Delivery Reviews, 2005. **57**(15): p. 2215-2237.

114. Gottschalk, U., ed. *Process Scale Purification of Antibodies*. 2009, John Wiley & Sons, Inc.: Hoboken, NJ.

115. Gekko, K. and T. Morikawa, *Preferential Hydration of Bovine Serum Albumin in Polyhydric Alcohol-Water Mixtures.* The Journal of Biochemistry, 1981. **90**: p. 39-50.

Appendix A – NMR & MS Data

Modified Di- and Tri-Amine Compounds

NMR and MS data were not collected for any of the modified amine compounds. There was high certainty in the complete conversion to guanidinium for these compounds and in the purity of the resulting products due to the simplicity of the reaction and the effective separation via TLC. Due to the cost and labor required to collect the structural data, it was deemed prudent to test the synthesized compounds first and only collect structural data if results were to be published. The excipients synthesized from the amine compounds were found to be destabilizing due to a strong preferential interaction and the conversion to other guanidinium salts to reduce this interaction failed due to low solubility. Thus, further development with such compounds was not explored due to this and the success of the dendrimer and peptide compounds.

BOC Protected Guanidine Modified PAMAM Dendrimers

Generation 0 (G0-BOC)

$[^{a}CH_2{}^{a}CH_2]$-
$(^{b}N[^{c}CH_2{}^{d}CH_2{}^{e}CO^{f}NH^{g}CH_2{}^{h}CH_2{}^{i}NH^{j}C(^{k}NH^{l}COO^{m}C(^{n}CH_3)_3)(^{k'}N^{l'}COO^{m'}C(^{n'}CH_3)_3)]_2)_2$

Purity (NMR): 91.4 %

^{1}H-NMR (400 MHz, $CDCl_3$): δ ppm 11.40 (4H, s, k), 8.51 (4H, t, J=5.5 Hz, i), 7.94 (4H, t, J=4.7 Hz, f), 3.50 (8H, q, J=5.7 Hz, h), 3.34 (8H, q, J=5.6 Hz, g), 2.61 (8H, t, J=5.5 Hz, c), 2.35 (4H, s, a), 2.29 (8H, t, J=5.6 Hz, d) , 1.433 (36H, s, n'), 1.439 (36H, s, n)

^{13}C-NMR (101 MHz, $CDCl_3$): δ ppm 173.3 (e), 163.4 (j), 157.0 (l'), 153.1 (l), 83.4 (m), 79.5 (m'), 52.2 (a), 50.5 (c), 40.6 (h), 39.8 (g), 34.3 (d), 28.2 (n'), 28.5 (n)

MS (ESI^+): m/z calculated for $C_{66}H_{120}N_{18}O_{20}$ $[M+H]^+$: 1485.90; found: 1485.91 (100%): $[M+2H-COOC(CH_3)_3]^+$: 1385.85; found: 1385.86 (43%), $[M+Na]^+$: 1507.88; found: 1507.89 (38%).

Generation 1 (G1-BOC)

$[^{a}CH_2{}^{a}CH_2]$-
$[^{b1}N(^{c1}CH_2{}^{d1}CH_2{}^{e1}CO^{f1}NH^{g1}CH_2{}^{h1}CH_2{}^{b2}N[^{c2}CH_2{}^{d2}CH_2{}^{e2}CO^{f2}NH^{g2}CH_2{}^{h2}CH_2{}^{i}NH^{j}C(^{k}NH^{l}COO^{m}C(^{n}CH_3)_3)(^{k'}N^{l'}COO^{m'}C(^{n'}CH_3)_3)]_2)_2]_2$

Purity (NMR): 90.6 %

^{1}H-NMR (400 MHz, $CDCl_3$): δ ppm 11.37 (8H, s, k), 8.51 (8H, t, J=5.5 Hz, i), 7.95 (8H, t, J=5.3 Hz, f2), 7.69 (4H, t, J=5.4 Hz, f1), 3.49 (16H, q, J=5.6 Hz, h2), 3.33 (16H, q, J=5.4 Hz, g2), 3.18 (8H, q, J=5.7 Hz, g1), 2.65-2.72 (24H, br t, c1+c2), 2.49 (8H, t, J=5.7 Hz, h1), 2.45 (4H, s, a), 2.28-2.31 (24H, br t, d1+d2) , 1.43 (72H, s, n'), 1.42 (72H, s, n)

^{13}C-NMR (101 MHz, $CDCl_3$): δ ppm 172.8-173.0 (e1+e2), 163.3 (j), 157.0 (l'), 153.1 (l), 83.4 (m), 79.4 (m'), 52.3-52.5 (a+h1), 50.4-50.5 (c1+c2), 40.6 (h2), 40.0 (g2), 37.6 (g1), 34.1-34.2 (d1+d2), 28.2 (n'), 28.5 (n)

MS (ESI$^+$): m/z calculated for $C_{150}H_{272}N_{42}O_{44}$ $[M+2H]^{2+}$: 1684.02; found: 1684.54 (100%), $[M+3H-COOC(CH_3)_3]^{3+}$: 1633.99; found: 1634.52 (97%), $[M+Na]^+$: 1695.51; found: 1695.53 (53%).

Generation 2 (G2-BOC)

$[^{a}CH_2{}^{a}CH_2]$-(^{b1}N[$^{c1}CH_2{}^{d1}CH_2{}^{e1}CO^{f1}NH^{g1}CH_2{}^{h1}CH_2{}^{b2}$N($^{c2}CH_2{}^{d2}CH_2{}^{e2}CO^{f2}NH^{g2}CH_2{}^{h2}CH_2$ ^{b3}N[$^{c3}CH_2{}^{d3}CH_2{}^{e3}CO^{f3}NH^{g3}CH_2{}^{h3}CH_2$ $^{i}NH^{j}C(^{k}NH^{l}COO^{m}C(^{n}CH_3)_3)(^{k'}N^{l'}COO^{m'}C(^{n'}CH_3)_3)]_2)_2]_2)_2$

Purity (NMR): 94.9 %

^{1}H-NMR (400 MHz, $CDCl_3$): δ ppm 11.37 (16H, s, k), 8.51 (16H, t, poorly resolved, i), 7.96 (16H, t, poorly resolved, f3), 7.82 (4H, t, poorly resolved, f1), 7.64 (8H, t, poorly resolved, f2), 3.49 (32H, q, J=4.9 Hz, h3), 3.33 (32H, q, J=4.3 Hz, g3), 3.18-3.21 (24H, br t, g1+g2), 2.65-2.72 (56H, br t, c1+c2+c3), 2.45-2.49 (28H, br m, a+h1+h2), 2.25-2.32 (56H, br t, d1+d2+d3) , 1.424 (144H, s, n'), 1.418 (144H, s, n)

^{13}C-NMR (101 MHz, $CDCl_3$): δ ppm 172.9-173.1 (e1+e2+e3), 163.3 (j), 157.0 (l'), 153.0 (l), 83.3 (m), 79.4 (m'), 52.4-52.6 (h1), 50.3-50.4 (c1+c2+c3), 40.6 (h3), 40.0 (g3), 37.5-37.6 (g1+g2), 34.1-34.2 (d1+d2+d3), 28.2 (n'), 28.5 (n)

MS (ESI$^+$): m/z calculated for $C_{318}H_{576}N_{90}O_{92}$ $[M+20H]^{20+}$: 357.42; found: 357.22 (100%)

Guanidine HCl Modified PAMAM Dendrimers

Generation 0 (G0-Gdn)

$[^{a}CH_2{}^{a}CH_2]$-(^{b}NH[$^{c}CH_2{}^{d}CH_2{}^{e}CO^{f}NH^{g}CH_2{}^{h}CH_2{}^{i}NH^{j}C(^{k}NH_2)_2]_2)_2$

Purity (NMR): 77.0 %
Yield: 81.8 %

^{1}H-NMR (400 MHz, DMSO-d_6): δ ppm 11.06 (2H, br s, b), 8.50 (4H, t, J=5.1 Hz, i), 7.89 (4H, t, J=5.2 Hz, f), 7.34 (16H, br s, k), 3.72 (4H, br s, a), 3.43 (8H, t, J=6.3 Hz, c), 3.20-3.24 (16H, br m, g+h), 2.77 (8H, t, J=6.9 Hz, d)

^{13}C-NMR (101 MHz, DMSO-d_6): δ ppm 169.3 (e), 157.3 (j), 49.2 (c), 46.4 (a), 40.1 (g), 38.0 (h), 29.3 (d)

MS (ESI$^+$): m/z calculated for $C_{26}H_{62}N_{18}O_4Cl_6$ $[M-6HCl+2H]^{2+}$: 343.24; found: 343.25 (100%); $[M-6HCl+3H]^{3+}$: 229.17; found: 229.17 (59%), $[M-6HCl+4H]^{4+}$: 172.13; found: 172.13 (16%), $[M-6HCl+H]^+$: 685.48; found: 685.51 (4%).

Generation 1 (G1-Gdn)
$[^{a}CH_2{}^{a}CH_2]$-
$[^{b1}NH(^{c1}CH_2{}^{d1}CH_2{}^{e1}CO^{f1}NH^{g1}CH_2{}^{h1}CH_2{}^{b2}NH[^{c2}CH_2{}^{d2}CH_2{}^{e2}CO^{f2}NH^{g2}CH_2{}^{h2}CH_2{}^{i}NH^{j}C(^{k}NH_2)_2]_2)_2]_2$

Purity (NMR): 70.4 %
Yield: 78.5 %

^{1}H-NMR (400 MHz, DMSO-d_6): δ ppm 11.09 (2H, br s, b1), 10.45 (4H, br s, b2), 8.67 (4H, t, poorly resolved, f1), 8.51 (8H, t, poorly resolved, i), 7.91 (8H, t, poorly resolved, f2), 7.36 (32H, br s, k), 3.73 (4H, br s, a), 3.45-3.53 (8H, br q, g1), 3.35-3.45 (24H, br t, c1+c2), 3.18-3.27 (40H, br m, h1+h2+g2), 2.71-2.85 (24H, br t, d1+d2)

^{13}C-NMR (101 MHz, DMSO-d_6): δ ppm 169.5-169.6 (e1+e2), 157.3 (j), 51.3 (h1), 48.7-48.8 (c1+c2), 46.5 (a), 40.2 (g2), 38.1 (h2), 33.7 (g1), 29.1-29.3 (d1+d2)

MS (ESI$^+$): m/z calculated for $C_{70}H_{158}N_{42}O_{12}Cl_{14}$ $[M-14HCl+6H]^{6+}$: 295.21; found: 295.21 (100%), $[M-14HCl+5H]^{5+}$: 354.05; found: 354.06 (90%), $[M-14HCl+4H]^{4+}$: 442.31; found: 442.32 (70%), $[M-14HCl+7H]^{7+}$: 253.18; found: 253.18 (62%), $[M-14HCl+3H]^{3+}$: 589.41; found: 589.43 (27%).

Generation 2 (G2-Gdn)
$[^{a}CH_2{}^{a}CH_2]$-$(^{b1}NH[^{c1}CH_2{}^{d1}CH_2{}^{e1}CO^{f1}NH^{g1}CH_2{}^{h1}CH_2{}^{b2}NH(^{c2}CH_2{}^{d2}CH_2{}^{e2}CO^{f2}NH^{g2}CH_2{}^{h2}CH_2{}^{b3}NH[^{c3}CH_2{}^{d3}CH_2{}^{e3}CO^{f3}NH^{g3}CH_2{}^{h3}CH_2{}^{i}NH^{j}C(^{k}NH_2)_2]_2)_2]_2)_2$

Purity: 77.7 %
Yield: 81.1 %

^{1}H-NMR (400 MHz, DMSO-d_6): δ ppm 11.09 (2H, br s, b1), 10.40-10.44 (12H, br s, b2+b3), 8.62-8.69 (12H, br t, f1+f2), 8.51 (16H, t, poorly resolved, i), 7.91 (16H, br t, poorly resolved, f3), 7.34 (64H, br s, k), 3.65-3.75 (20H, br m, a+h2), 3.48-3.53 (24H, br q, g1+g2), 3.35-3.48 (56H, br t, c1+c2+c3), 3.18-3.27 (72H, br m, h1+h3+g3), 2.71-2.82 (56H, br t, d1+d2+d3)

^{13}C-NMR (101 MHz, DMSO-d_6): δ ppm 169.5-169.7 (e1+e2+e3), 157.3 (j), 51.3-51.4 (h1+h2), 48.7-48.8 (c1+c2+c3), 46.6 (a), 40.2 (g3), 38.1 (h3), 33.7-33.8 (g1+g2), 29.1-29.3 (d1+d2+d3)

MS (ESI$^+$): m/z calculated for $C_{158}H_{350}N_{90}O_{28}Cl_{30}$ $[M-29HCl+11H]^{11+}$: 361.25; found: 361.26 (100%); $[M-30HCl+8H]^{8+}$: 492.11; found: 492.11 (97%), $[M-30HCl+9H]^{9+}$: 437.54; found: 437.54 (92%), $[M-30HCl+10H]^{10+}$: 393.89; found: 393.88 (47%), $[M-30HCl+7H]^{7+}$: 562.27; found: 562.27 (33%), $[M-30HCl+6H]^{6+}$: 655.81; found: 655.81 (10%).

Arginine Peptides
Arginine Dimer (RR)
HPLC Purity: 99.2%
MS (ESI$^+$): m/z calculated for RR $[M+2H]^{2+}$: 166.20; found 166.10 (100%), $[M+H]^+$: 331.39; found 331.15 (30%).

Arginine Dimer (RS)
HPLC Purity: 99.3%
MS (ESI^+): m/z calculated for RS $[M+H]^+$: 262.28; found 262.10 (100%).

Arginine Dimer (RL)
HPLC Purity: 99.1%
MS (ESI^+): m/z calculated for RL $[M+H]^+$: 288.36; found 288.15 (100%).

Arginine Dimer (RQ)
HPLC Purity: 99.7%
MS (ESI^+): m/z calculated for RQ $[M+H]^+$: 303.34; found 303.10 (100%).

Arginine Dimer (RE)
HPLC Purity: 99.1%
MS (ESI^+): m/z calculated for RQ $[M+H]^+$: 304.34; found 304.15 (100%).

Arginine Trimer (RRR)
HPLC Purity: 98.2%
MS (ESI^+): m/z calculated for RRR $[M+2H]^{2+}$: 243.79; found 244.25 (100%), $[M+3H]^{3+}$: 163.19; found 163.10 (52%), $[M+H]^+$: 487.57; found 487.30 (18%).

Arginine Trimer (RER)
HPLC Purity: 99.1%
MS (ESI^+): m/z calculated for RER $[M+2H]^{2+}$: 230.75; found 230.70 (100%) , $[M+H]^+$: 460.50; found 460.25 (41%).

Arginine Trimer (ERE)
HPLC Purity: 98.2%
MS (ESI^+): m/z calculated for ERE $[M+H]^+$: 433.43; found 433.15 (100%), $[M+2H]^{2+}$: 217.22; found 217.10 (76%).

Arginine Tetramer (RRRR)
HPLC Purity: 98.1%
MS (ESI^+): m/z calculated for RRRR $[M+2H]^{2+}$: 322.38; found 322.35 (100%), $[M+3H]^{3+}$: 215.25; found 215.20 (96%), $[M+H]^+$: 643.76; found 643.35 (13%).

Arginine Tetramer (RREE)
HPLC Purity: 98.0%
MS (ESI^+): m/z calculated for RREE $[M+2H]^{2+}$: 295.31; found 295.30 (100%), $[M+H]^+$: 589.62; found 589.30 (22%).

Arginine Tetramer (RERE)
HPLC Purity: 99.4%
MS (ESI^+): m/z calculated for RERE $[M+2H]^{2+}$: 295.31; found 295.25 (100%), $[M+H]^+$: 589.62; found 589.30 (18%).

Appendix B – Preferential Interaction Data

Table B-1: Summary of solute partial molar volume and dependence of solution osmolality on solute molality for two-component solutions

cosolute (comp. 3)	partial molar volume (L/mol) × 10^2	$\Omega_3^{o(2)}$ (pH 6.0)[a]	$\Omega_3^{o(2)}$ (pH 3.5)[b]	range (m) mol/kg
argHCl	$(14.21 \pm 0.01) + (0.22 \pm 0.01)m_3$	$(1.716 \pm 0.001) - (0.674 \pm 0.004)m_3$	$(1.678 \pm 0.001) - (0.542 \pm 0.002)m_3$	$m_3<0.70$
gdnHCl	$(6.94 \pm 0.01) + (0.048 \pm 0.003)m_3$	$(1.795 \pm 0.01) - (0.598 \pm 0.002)m_3$	$(1.747 \pm 0.001) - (0.4906 \pm 0.002)m_3$	$m_3<0.76$
glucose	$(11.192 \pm 0.003) + (0.044 \pm 0.003)m_3$	1.016 ± 0.001	1.013 ± 0.001	$m_3<1.20$
glycerol	$(7.118 \pm 0.003) - (0.030 \pm 0.001)m_3$	0.9956 ± 0.0005	0.974 ± 0.001	$m_3<1.14$
urea	$(4.452 \pm 0.002) - (0.117 \pm 0.003)m_3$	$(0.9744 \pm 0.0005) - (0.0918 \pm 0.001)m_3$	$(0.916 \pm 0.001) - (0.008 \pm 0.001)m_3$	$m_3<1.22$
protein (comp. 2)	partial molar volume (L/mol)	$\Omega_2^{o(3)}$ (pH 6.0)[a]	$\Omega_2^{o(3)}$ (pH 3.5)[b]	range (m) mmol/kg
BSA	47.731 ± 0.001	3.6 ± 0.9	N/A	$1.9 \le m_2 \le 2.1$
lysozyme	10.005 ± 0.001	3.2 ± 0.9	N/A	$3.5 \le m_2 \le 3.8$
aCgn	17.910 ± 0.001	N/A	4.4 ± 0.8	$1.65 \le m_2 \le 1.75$

[a] The base buffer was 40 mM Na-Phosphate; [b] The base buffer was 20 mM Na-Citrate.

Table B-2: Summary of preferential interaction coefficients for BSA as measured by VPO and comparison to literature values

cosolute	Ω_3	Γ_{μ_1}	Γ_{μ_3}	Γ_{μ_1,μ_3}	$\Gamma_{\mu_1,\mu_3}^{Lit}$
argHCl	$(1.713 \pm 0.003) - (0.611 \pm 0.006)m_3$	$(-1.64 \pm 0.6) + (1.7 \pm 3.9)m_3 - (13.3 \pm 5.4)m_3^2$	$(1.72 \pm 0.96)[3] - (23.2 \pm 4.0)[3]^2$	$(1.43 \pm 0.96)[3] - (22.8 \pm 4.0)[3]^2$	$(12.6 \pm 4.5)[3] - (22.0 \pm 5.3)[3]^2$ [a]
gdnHCl	$(1.755 \pm 0.001) - (0.576 \pm 0.005)m_3$	$(-1.8 \pm 0.4) + (11.2 \pm 1)m_3$	$(12.4 \pm 1)[3]$	$(12.3 \pm 1)[3]$	$(13.0 \pm 1.2)[3]$ [b]
glucose	$(1.055 \pm 0.001) - (0.004 \pm 0.002)m_3$	$(-1.4 \pm 0.6) - (17.8 \pm 0.8)m_3$	$(-20.6 \pm 0.4)[3]$	$(-20.8 \pm 0.4)[3]$	$(-19.8 \pm 0.9)[3]$ [c]
glycerol	$(1.014 \pm 0.001) - (0.0012 \pm 0.001)m_3$	$(-4.1 \pm 0.8) - (12.3 \pm 1.2)m_3$	$(-14.0 \pm 0.4)[3]$	$(-14.3 \pm 0.4)[3]$	$(-12 \pm 4)[3]$ [d]
urea	$(0.963 \pm 0.002) - (0.092 \pm 0.002)m_3$	$(-5.0 \pm 0.4) + (6.2 \pm 0.6)m_3$	$(6.4 \pm 0.4)[3]$	$(6.2 \pm 0.4)[3]$	$(6 \pm 2)[3]$ [e]

It should be noted that [3] represents the molar concentration of the cosolute. The base buffer was 40 mM Na-Phosphate, pH 6.0 and BSA concentration held constant at 1.8 mM. The values represented are an average of the separate approximations computed by Eqs. (3.15) and (3.18). [a] Calculated from data in Reference [85]. [b] Calculated from data in Reference [33]. [c] Calculated from data in Reference [32]. [d] Calculated from data in Reference [115]. [e] Calculated from data in Reference [84].

Table B-3: Summary of preferential interaction coefficients for Lysozyme as measured by VPO and comparison to literature values

cosolute	Ω_3	Γ_{μ_1}	Γ_{μ_3}	Γ_{μ_1,μ_3}	$\Gamma_{\mu_1,\mu_3}^{Lit}$
argHCl	$(1.670 \pm 0.001) - (0.560 \pm 0.004)m_3$	$(-1.9 \pm 0.3) + (5.2 \pm 2.2)m_3 - (6.5 \pm 3.3)m_3^2$	$(5.8 \pm 0.4)[3] - (10.8 \pm 1.4)[3]^2$	$(5.5 \pm 0.4)[3] - (10.6 \pm 1.4)[3]^2$	$(-3.6 \pm 1.6) + (8.0 \pm 9)[3] - (9.2 \pm 11.3)[3]^2$ [a]
gdnHCl	$(1.770 \pm 0.001) - (0.584 \pm 0.002)m_3$	$(-2.0 \pm 0.3) + (2.9 \pm 0.6)m_3$	$(3.7 \pm 1)[3]$	$(3.6 \pm 1)[3]$	N/A
glucose	$(1.026 \pm 0.001) - (0.001 \pm 0.002)m_3$	$(-2.8 \pm 0.3) - (2.4 \pm 0.5)m_3$	$(-2.74 \pm 0.3)[3]$	$(-3.07 \pm 0.3)[3]$	N/A
glycerol	$(1.018 \pm 0.001) - (0.002 \pm 0.002)m_3$	$(-2.2 \pm 0.3) - (5.7 \pm 0.5)m_3$	$(-6.3 \pm 0.3)[3]$	$(-6.5 \pm 0.3)[3]$	N/A
urea	$(0.962 \pm 0.001) - (0.090 \pm 0.002)m_3$	$(-4.3 \pm 0.3) + (2.7 \pm 0.3)m_3$	$(3.4 \pm 1)[3]$	$(3.2 \pm 1)[3]$	$(6 \pm 1)[3]$ [b]

It should be noted that [3] represents the molar concentration of the cosolute. The base buffer was 40 mM Na-Phosphate, pH 6.0 and lysozyme concentration held constant at 3.5 mM. The values represented are an average of the separate approximations computed by Eqs. (3.15) and (3.18). [a] Calculated from data in Reference [85]. [b] Calculated from data in Reference [34].

Table B-4: Summary of preferential interaction coefficients for α-Chymotrypsinogen A as measured by VPO

cosolute	Ω_3	Γ_{μ_1}	Γ_{μ_3}	Γ_{μ_1,μ_3}
argHCl	$(1.676 \pm 0.001) - (0.514 \pm 0.004)m_3$	$(-2.8 \pm 0.8) + (0.2 \pm 5.8)m_3 - (6.6 \pm 8.1)m_3^2$	$(0.9 \pm 0.7)[3] - (10.5 \pm 2.5)[3]^2$	$(0.5 \pm 0.7)[3] - (10.4 \pm 2.5)[3]^2$
gdnHCl	$(1.736 \pm 0.001) - (0.488 \pm 0.002)m_3$	$(-0.8 \pm 0.6) + (4.1 \pm 1.1)m_3$	$(4.0 \pm 1)[3]$	$(3.9 \pm 1)[3]$
glucose	$(1.022 \pm 0.001) - (0.001 \pm 0.002)m_3$	$(-4.5 \pm 1.2) - (5.1 \pm 0.8)m_3$	$(-5.2 \pm 0.5)[3]$	$(-5.8 \pm 0.5)[3]$
glycerol	$(0.999 \pm 0.001) - (0.002 \pm 0.002)m_3$	$(-3.7 \pm 0.8) - (13.6 \pm 1.2)m_3$	$(-15.0 \pm 0.7)[3]$	$(-15.3 \pm 0.7)[3]$
urea	$(0.902 \pm 0.001) - (0.015 \pm 0.001)m_3$	$(-6.6 \pm 0.8) + (8.8 \pm 1.0)m_3$	$(9.4 \pm 1)[3]$	$(9.1 \pm 1)[3]$

It should be noted that [3] represents the molar concentration of the cosolute. The base buffer was 20 mM Na-Citrate, pH 3.5 and aCgn concentration held constant at 1.7 mM. The values represented are an average of the separate approximations computed by Eqs. (3.15) and (3.18).

Table B-5: Preferential interaction coefficients for proteins in arginine HCl solutions as determined by dialysis/ densimetry measurements

[arginine] (M)	ϕ_2° (mL/g)	$\phi_2'^{\circ}$ (mL/g)	Γ_{μ_1,μ_3}
	BSA, pH 6.0		
0.47	0.708 ± 0.002	0.713 ± 0.001	(-5.2 ± 2.8)
0.87	0.717 ± 0.002	0.734 ± 0.002	(-21.3 ± 4.3)
	lysozyme, pH 6.0		
0.38	0.713 ± 0.002	0.706 ± 0.002	(1.6 ± 0.7)
0.87	0.697 ± 0.002	0.710 ± 0.002	(-3.4 ± 0.8)

The base buffer was 40 mM sodium phosphate, pH 6.0.

Table B-6: Summary of VPO preferential interaction coefficient data, aCgn denaturation midpoint temperature increments, cosolute partial molar volume (PMV), and Pitzer ion interaction parameters for arginine salts and guanidinium chloride.

Salt	Γ_{μ_3}	$dT_m/d[3]$ K*L/mol	Max. [3] mol/L	PMV L/mol x 10^2	$\beta^{(0)}$ kg/mol	$\beta^{(1)}$ kg/mol	C^{ϕ} $(kg/mol)^2$	Range mol/kg	σx10^4
ArgH(H_2PO_4)	(-24.2 ± 3.4)[3] - (25.6 ± 8.8)$[3]^2$	6.3	0.2	15.50 + 1.44m	0.1885	-1.5303	-0.3143	0.2-0.4	2.4
ArgH(SO_4)$_{1/2}$	(-4.3 ± 0.5)[3] - (12.4 ± 6.2)$[3]^2$	5.4	0.5	13.50 + 0.36m	0.0279	-1.5809	-0.0454	0.7-1.3	1.9
ArgH(Citrate)$_{1/2}$	(-18.5 ± 0.8)[3] + (6.3 ± 4.6)$[3]^2$	3.3	0.5	16.73 + 0.53m	0.1146	-1.0850	-0.1123	0.6-1.2	4.0
ArgH(Acetate)	N/A	0.0	0.5	16.30 + 0.40m	N/A	N/A	N/A	N/A	N/A
ArgHF	N/A	0.0	0.5	11.83 + 0.65m	N/A	N/A	N/A	N/A	N/A
ArgHCl	(5.9 ± 0.1)[3] - (23.4 ± 4.8)$[3]^2$	0.0	0.5	14.05 + 0.29m	-0.0317	-0.3357	-0.0391	0.4-0.9	2.7
ArgHBr	(5.9 ± 0.1)[3] - (19.5 ± 4.3)$[3]^2$	-7.7	0.5	14.54 + 0.08m	-0.1532	0.0818	0.0093	0.4-0.9	1.4
ArgHI	(6.9± 0.3)[3] - (10.8± 4.0)$[3]^2$	-22.6	0.5	15.89 + 0.57m	-0.2996	0.2778	0.1215	0.4-0.9	2.4
ArgHSCN	(5.9 ± 0.2)[3] - (11.2 ± 4.3)$[3]^2$	-24.2	0.5	15.85 + 0.51m	-0.3304	0.5404	0.1299	0.4-0.9	1.7
GdmCl	(6.0 ± 1)[3]	-6.6	0.5	4.45 + 0.12m	N/A	N/A	N/A	N/A	N/A
ArgHCl (Lit.)[67]	N/A	N/A	N/A	N/A	-0.0429	-0.3878	0.0084	0.1-1.6	30.0

aCgn solutions for the Γ_{μ_3} (50 mg/mL) and T_m (1 mg/mL) data contained 20 mM sodium citrate pH 5 buffer. PMV and Pitzer Ion Interaction Parameters data were for cosolute only solutions containing no buffering component.