2018-12-01

Engineering Cell-Free Biosystems for On-Site Production and Rapid Design of Next-Generation Therapeutics

Kristen Michelle Wilding
Brigham Young University

Engineering Cell-Free Biosystems for On-Site Production and Rapid Design of Next-Generation Therapeutics

Kristen Michelle Wilding

A dissertation submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Bradley C. Bundy, Chair
Joel S. Griffitts
Thomas A. Knotts
Joshua L. Price
Matthew J. Memmott

Department of Chemical Engineering

Brigham Young University

ABSTRACT

Engineering Cell-Free Biosystems for On-Site Production and Rapid Design of Next-Generation Therapeutics

Kristen Michelle Wilding
Department of Chemical Engineering, BYU
Doctor of Philosophy

While protein therapeutics are indispensable in the treatment of a variety of diseases, including cancer, rheumatoid arthritis, and diabetes, key limitations including short half-lives, high immunogenicity, protein instability, and centralized production complicate long-term use and on-demand production. Site-specific polymer conjugation provides a method for mitigating these challenges while minimizing negative impacts on protein activity. However, the location-dependent effects of polymer conjugation are not well understood. Cell-free protein synthesis provides direct access to the synthesis environment and rapid synthesis times, enabling rapid evaluation of multiple conjugation sites on a target protein. Here, work is presented towards developing cell-free protein synthesis as a platform for both design and on-demand production of next generation polymer-protein therapeutics, including (1) eliminating endotoxin contamination in cell-free reagents for simplified therapeutic preparation, (2) improving shelf-stability of cell-free reagents via lyophilization for on-demand production, (3) coupling coarse-grain simulation with high-throughput cell-free protein synthesis to enable rapid identification of optimal polymer conjugation sites, and (4) optimizing cell-free protein synthesis for production of therapeutic proteins.

Keywords: synthetic biology, cell-free protein synthesis, endotoxin removal, lyophilization, lyoprotectant, unnatural amino acid, high-throughput screening

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my husband, Paul Wilding, for supporting me in this endeavor. I appreciate your sacrifices in either working or watching Joshua during countless early mornings, late nights, and weekends in order to allow us both to simultaneously pursue our doctorates.

I am grateful to the Department of Chemical Engineering for the facilities, instruction, and opportunities provided me during my studies. I would particularly like to thank my advisor, Dr. Bundy, for his guidance, support, and consistent confidence in my abilities and also for the time which he put in to helping me put together strong fellowship applications. I would also like to thank my committee members, Dr. Thomas Knotts, Dr. Joel Griffitts, Dr. Joshua Price, and Dr. Matthew Memmott, for their time and willingness to mentor me.

I am also thankful for my various funding sources, the Utah NASA Space Grant Consortium, the National Science Foundation Graduate Research Fellowship Program, and the BYU Simmons Center for Cancer Research, which have provided resources essential to accomplishing this work.

Finally, I would like to thank my current and former colleagues, Addison Smith, Derek Bush, Matt Schinn, Seung-Ook Yang, Mark Smith, Andrew Broadbent, Amin Salehi, Emily Long Zhao, Porter Hunt, and Mehran Soltani for their contributions, suggestions, and friendship. I would especially like to thank Addison Smith for his time and effort in helping me with the coarse-grain simulation, and Emily Long Zhao for her persistence and flexibility in helping with the lyophilization work. I would also like to extend a thank you to the many undergraduate students who contributed in these projects, with a special thanks to Joshua Wilkerson for his dependability, hard work, and insight.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

In the 1980's, medicine was changed forever with the FDA approval of the first recombinant therapeutic – human insulin – followed closely by a second, human growth hormone [1]. Approval of these therapeutics marked the first approval of human proteins produced in *E. coli* and ushered in a new age of medicine with the burgeoning field of synthetic biology [1]. The ability to produce proteins in hosts other than their native organisms has enabled safe, large-scale production of a wide variety of therapeutics such that the list of FDA-approved protein and peptide therapeutics now reaches upwards of 240 [1]. These protein therapeutics have emerged as indispensable components in the treatment of many diseases, including diabetes, arthritis, and cancer, and comprised 7 of the top 10 best-selling drugs in 2017 [2]. However, development and use of these protein drugs is accompanied by several key challenges: (1) many proteins are rapidly filtered from the body in the kidneys, resulting in short half-lives which require frequent dosing [3,4]; (2) protein therapeutics, particularly those not human in origin, often elicit immune responses, which can further decrease drug activity and potentially result in dangerous hypersensitivity reactions [4-6]; (3) proteins are highly susceptible to protease degradation, thermal stress, and aggregation, which can further limit half-lives and increase immunogenicity and requires stringent cold-chain storage [4-9]; and (4) stringent cold-chain storage requirements for both proteins and the cellular machinery used to produce them limits the accessibility to these vital therapeutics [8,9].

Cell-free protein synthesis (CFPS) offers a promising platform for addressing these challenges. CFPS presents multiple advantages over *in vivo* approaches for protein optimization and on-demand production. Central advantages include rapid synthesis and high volumetric yields [10-13], emancipation from cellular viability constraints to allow flexible reaction conditions [12-16], direct access to the reaction environment due to the absence of a cell wall [12,13,16], and broad scalability [17,18]. These attributes make CFPS a promising tool for addressing the challenges associated with protein therapeutics. First, the open protein synthesis environment of CFPS allows for simplified protein purifications for on-site production by eliminating the need for sophisticated cell lysis equipment and enabling treatment of CFPS prior to protein expression to reduce contamination [12,19]. The flexible environment independent of live cells and the directly scalable nature of CFPS reactions also makes CFPS amenable to on-demand protein synthesis by enabling lyophilization for shelf-stability and ease of transport [19]. Finally, because of the rapid protein synthesis capability, flexibility, and scalability, CFPS lends itself well to the optimization of engineered proteins for enhanced therapeutic development [20,21]. Here, we capitalized upon these unique attributes of CFPS to both improve and apply CFPS technology towards the improvement of protein therapeutics.

1.1 Project scope

The challenges inherent to protein therapeutics are complex and multi-faceted. The aim of this project is to provide a tool to enhance access to and development of protein therapeutics. Specifically, the objectives of this project are:

1. To eliminate endotoxin in CFPS systems to enable simplified purifications for on-demand protein production

2. To enhance shelf-stability of CFPS systems to improve stability during transport and storage towards distributed, on-site therapeutic production
3. To improve design of active, stable polymer-protein conjugates for next-generation therapeutic development
4. To produce therapeutic proteins at high yields

1.1.1 Endotoxin elimination

Approximately one third of protein therapeutics are produced in *E. coli*, targeting a wide variety of diseases. However, due to immune recognition of endotoxin (a lipid component in the *E. coli* cell membrane), these protein products must be extensively purified before application to avoid adverse reactions such as septic shock. Cell-free systems offer the unique ability to remove endotoxin prior to protein synthesis. Here we evaluate three strategies for removing endotoxins from *E. coli* cell lysate: Triton X-114 two-phase extraction, polylysine affinity chromatography, and extract preparation from genetically engineered, endotoxin-free ClearColi® cells. We demonstrate the first adaptation of ClearColi® cells to prepare cell-free extract with high protein synthesis capability. Pre-expression endotoxin removal from CFPS reagents could simplify downstream processing, potentially enabling on-demand production of unique protein therapeutics.

1.1.2 Cryoprotected CFPS for improved shelf-stability

In recent years, lyophilized cell-free systems have widened the application space by enabling improved stability outside of cold-chain storage. Current embodiments of the system, however, demonstrate poor stability at elevated temperatures. Lyoprotectants have long been recognized for the ability to preserve the activity of biological molecules during drying processes,

however the application of this technology to lyophilized cell-free systems has been limited. Here, we demonstrate improved cryoprotection for stable, highly active lyophilized cell-free protein synthesis capable of producing up to 10x protein yields of similar systems.

1.1.3 Site-specific PEGylated protein screening

PEGylation, or covalent attachment of polyethylene glycol (PEG) to proteins, is a well-established method for addressing many of these issues. However, the traditional approach to PEGylation generally involves targeting naturally reactive protein residues, such as lysines or the N- or C-terminus, resulting in insufficient control over conjugation site and often significantly reducing specific activity [3,22,23]. Unnatural amino acids (uAA) can be incorporated to provide a unique handle for PEGylation [20,24], but optimal site selection for uAA-incorporation and site-specific PEGylation is poorly understood. We apply CFPS towards evaluation of the locational effects of uAA-incorporation and site-specific PEGylation and develop a computationally-efficient coarse-grain simulation model as a potential tool to enhance PEGylation site screening.

1.1.4 Optimized CFPS production of cancer therapeutic

Acute lymphoblastic leukemia (ALL) is the most common childhood cancer. Successful treatment using FDA-approved asparaginase therapies are plagued by extreme hypersensitivity and rapid drug clearance. Additional reports of anti-PEG antibodies threaten the utility of currently available PEGylated therapies for the disease, including pegaspargase. We demonstrate optimized production of a second-line therapeutic for ALL, crisantaspase. This work demonstrates the first reported production of highly active crisantaspase in a cell-free system. We further propose a method for high-throughput optimization of site specific polymer conjugates of these therapeutics

using a PEG-alternative, poly(2-ethyl-2-oxazoline). This technology could improve the development of stabilized protein therapeutics.

1.2 Outline

This work focuses on developing CFPS technology towards addressing four key challenges of protein therapeutics: (1) short half-lives, (2) immunogenicity, (3) limited stability, and (4) poor accessibility due to stringent cold-chain storage requirements. The bulk of this dissertation constitutes adaptations of publications (published or in-progress) which I have lead, developed, and authored. In addition, this work frequently references, but does not include, several additional publications which I either authored or coauthored (as summarized in Table 1-1).

Table 1-1: Work Summary & Outline

Publication	Chapter	Challenge	Presented Solution
Wilding et al., 2018. *Biotechol. J.*	2	Accessibility, immunogenicity	Endotoxin elimination for simplified purification
Wilding et al., (In preparation)	3	Accessibility, stability	Stabilized CFPS lyophilization for distributed, on-demand synthesis
Wilding et al., 2018. *ACS Synth. Biol.*	4	Short half-life, immunogenicity, stability	Site-specific PEGylation optimization for improved stability and activity of PEGylated proteins.
Wilding et al., 2018. *Curr. Opin. Biotech.*	Referenced throughout	Enzyme stability, cost, and operational sustainability	Review of immobilization, encapsulation, and energy system optimization
Hunt JP, Yang SO, Wilding KM, Bundy BC, 2016. *Bioengineered*	Referenced throughout	Accessibility, stability	Review of CFPS lyophilization progress
Smith MT, Wilding KM, et al., 2014. *FEBS Letters*	Referenced throughout	Biotechnology applications	Review of CFPS applications

.

2 ENDOTOXIN-FREE E. COLI-BASED CELL-FREE PROTEIN SYNTHESIS FOR ON-DEMAND CANCER THERAPEUTIC PRODUCTION

Chapters 2 and Appendix A are adapted from a publication entitled "Endotoxin-free *E. coli*-based cell-free protein synthesis: Pre-expression endotoxin removal approaches for on-demand cancer therapeutic production" (Wilding et al. 2018). The presented work was led and conducted by myself, with contributions from my graduate student colleague Porter Hunt and undergraduate students Joshua Wilkerson, Parker Funk, Rebecca Swensen, William Carver, and M. Lance Christian, who I trained and supervised.

2.1 Introduction

The past few decades have seen a dramatic increase in the number of recombinant protein therapeutics developed to target a wide variety of diseases such as cancer, arthritis, and diabetes. Approximately one third of all these products are produced in *E. coli* [25,26]. The *E. coli* platform is attractive for therapeutic production due to its low cost and rapid growth kinetics. However, like other gram negative bacteria, *E. coli* produce high levels of lipopolysaccharide (LPS), or endotoxin, during cell growth, division, and lysis. Human exposure to endotoxin activates an immune response and can result in septic shock [25,27,28]. Therefore, therapeutics produced in *E. coli* must be extensively purified to remove endotoxin to extremely low levels before administration [29]. This additional processing and quality control adds to the cost of protein therapeutic production, with downstream processing in general accounting for 45-92% of total manufacturing costs [30].

The removal process can result in poor protein recovery, reduced protein activity, and the introduction of contaminants which often necessitates optimization for each individual product [25,27,30,31].

One potential solution is to use cell-free protein synthesis (CFPS), where the biological machinery necessary for transcription and translation is harvested from cells and activated *in vitro*. This approach to protein synthesis provides the unique potential to shift several processing steps upstream, such as cell lysis and separation. Thus, the system provides an opportunity for endotoxin removal prior to protein synthesis. This pre-expression endotoxin removal could potentially simplify or eliminate downstream endotoxin removal steps. Pre-expression endotoxin removal would be especially important for the emerging application of lyophilized CFPS technology for on-demand, deployable therapeutic synthesis, which could enable rapid production of personalized therapeutics or therapeutics in response to a pandemic threat [19,32-34]. Here, we evaluate three methods of pre-expression endotoxin removal from CFPS systems. We report that conventional methods of endotoxin removal from clarified *E. coli* cell lysate, or cell-free extract, show promise, but significant cost and engineering is necessary to make these methods viable. By contrast, optimized production of cell-extract from ClearColi® (an LPS-free genetically engineered BL21 DE3 *E. coli* strain) produced consistently high yields of protein without the need for specific endotoxin removal steps. Further, we demonstrate production of a clinically relevant FDA-approved cancer therapeutic, crisantaspase, from low endotoxin and endotoxin-free *E. coli* cell-free extract.

2.2 Materials and methods

Unless otherwise specified, materials were obtained from Sigma Aldrich (St. Louis, MO). All falcon tubes, microcentrifuge tubes, and pipette tips were endotoxin-free. Additional details are provided in Appendix A.

2.2.1 Extract preparation

Cell extracts were made using either *E. coli* BL21 Star™ (DE3) (Invitrogen, Carlsbad, CA) or ClearColi® BL21 DE3 (Lucigen, Middleton, WI). Extract was produced as described previously with either LB, 2xYT, or TB fermentation media [19,35-37].

2.2.2 CFPS pre-treatment

For pre-treatment studies, standard BL21 extract was treated using either Triton X-114 two-phase extraction or polylysine affinity chromatography. Triton X-114 extractions were performed using 1% volume Trion X-114 according to a modification of a method previously described for purified proteins [29,38,39]. Endotoxin content and protein production capacity of extract samples were assessed after 1-3 cycles. For polylysine affinity chromatography treatment, Pierce™ High Capacity Endotoxin Removal Spin Columns, 0.5 mL were obtained from ThermoFisher Scientific. Columns were equilibrated and washed as directed by the manufacturer. 1 mL samples of *E. coli* BL21 Star™ (DE3) extracts (referred to as BL21 (DE3*)) were diluted to 2 mL using endotoxin-free LAL reagent water (Lucigen, Middleton, WI), and applied to spin columns as directed by the manufacturer.

2.2.3 CFPS production of sfGFP and crisantaspase

CFPS of sfGFP was performed and yields assessed as previously described [19,40]. The crisantaspase gene with a C-terminal linker (GGSGGS) and 6-His tag was purchased and cloned into the pY71 vector (GenScript). Crisantaspase CFPS reactions were performed similarly to previously described CFPS reactions, varying the yield using 12-36 nM DNA and incubating at 30°C for 8 hours or 37°C for 3 hours [20]. Reactions were performed in 20 µL – 400 µL volumes in 2 mL microcentrifuge tubes (Bioexpress, Kaysville, UT) or 15 mL falcon tubes. Crisantaspase was purified using Ni-NTA spin columns according to the manufacturer's instructions. Soluble protein concentration and the protein concentration of purified samples was determined via TCA precipitation and scintillation counting as reported previously [20].

2.2.4 LAL testing

The endotoxin equivalent activity of the extracts, treated extracts, and purified samples was determined using the LAL Kinetic Turbidimetric Assay (Lonza, Allendale, NJ) according to the manufacturer's instructions.

2.2.5 Crisantaspase activity assay

The activity of crisantaspase was assessed using an Asparaginase Activity Assay Kit (Sigma Aldrich). Sample activity was determined by measuring the change in OD_{570} 37°C using a Synergy-MX microplate reader (BioTek, Winooski, VT).

2.3 Results and discussion

The open format of CFPS provides the opportunity to remove or eliminate endotoxin prior to protein expression, potentially reducing or eliminating downstream endotoxin removal steps to enable on-demand, deployable therapeutic synthesis. However, to our knowledge, no attempt has been made to remove endotoxin from *E. coli* extract prior to CFPS. Towards this end, we assessed the impact of endotoxin removal before CFPS expression by the following approaches: Triton X-114 phase extraction, polylysine affinity chromatography, and extract production from genetically-engineered endotoxin-free ClearColi® *E. coli* cells (Figure 2-1).

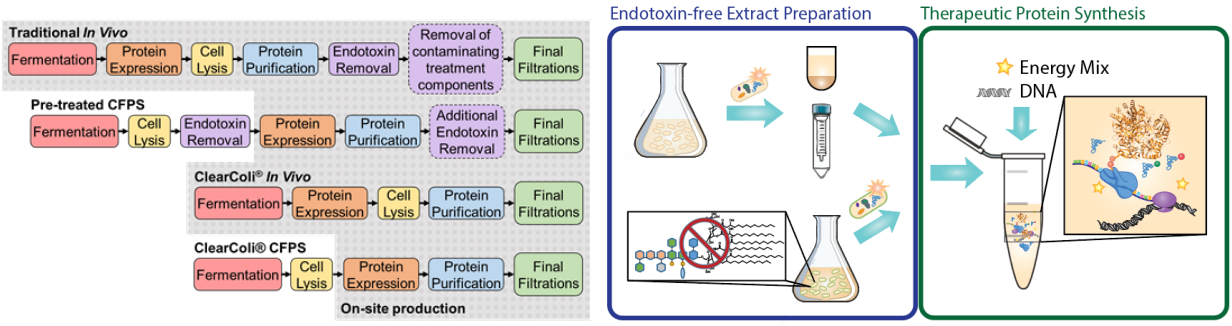


Figure 2-1: Pre-expression Endotoxin Removal from CFPS Systems Simplifies On-site Processing. (Left) Flowchart comparing workflow for the production and purification of protein therapeutics via in vivo E. coli expression, Triton X-114 or polylysine affinity pre-treated E. coli-based CFPS, in vivo ClearColi®, and ClearColi® CFPS. Dotted box represents a step which may or may not be necessary. Shading indicates steps which could be performed on-site. (Right) Schematic of endotoxin-free extract production for on-demand therapeutic synthesis.

2.3.1 Pre-expression endotoxin removal from standard BL21 (DE3*)

Among the most common methods of endotoxin removal from purified proteins are Triton-X-114 two-phase extraction and affinity chromatography [28]. While these methods have been quite successful for some purified proteins, multiple protocol repetitions and trade-offs between protein stability and removal efficiency can complicate treatments, reducing yields and protein activities and introducing additional contaminants [27,30,31,38,41]. We therefore assessed the

potential of Triton X-114 extraction and polylysine affinity chromatography as endotoxin removal methods for extract pre-treatment, testing the endotoxin content and protein synthesis capacity of extracts after one or more cycles of one of the protocols.

As shown in Figure 2-2, both treatment methods significantly, but insufficiently, reduced the endotoxin content of the treated extracts. Both methods also significantly reduced the protein synthesis capacity. Of the two, Triton X-114 extraction was the most effective approach, removing 95% endotoxin after 3 treatments while maintaining ~30% protein synthesis capacity. The losses of protein synthesis capacity are not surprising, considering the complex mixture of proteins, small molecules, and macromolecules involved in protein transcription and translation which could be lost during the extract treatment. However, the initial endotoxin content of CFPS extract (~18 million endotoxin units (EU) per mL – Figure 2-2) is orders of magnitude greater than that of protein sample that has undergone purification steps prior to endotoxin removal [25,27,29,38,42,43]. Thus ~99.999% of endotoxin would need to be removed to reach FDA approved levels of endotoxin administration (on the order of 100 EU per mL) to eliminate the risk of endotoxin toxicity and eliminate the need for endotoxin removal steps from CFPS-produced protein [25,27,29,38,42]. Further optimization of Triton X-114 extraction protocols may enable increased removal efficiencies beyond 95% and may mitigate some of the loss of CFPS components [41]. However, considering the 70% loss in CFPS activity and the significant concentration of residual endotoxin (1 million EU per mL), the current embodiments of extract treatment methods are currently ill-suited towards large-scale applications.

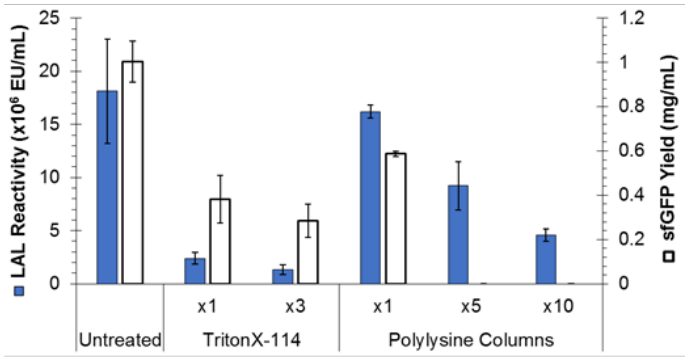


Figure 2-2: Pre-expression Extract Treatment Reduces Extract Endotoxin Content but Reduces Viability. Endotoxin content, as measured by LAL reactivity (blue bars), and sfGFP production yields (white bars) from untreated and treated standard E. coli extracts. x1, x3, x5, and x10 represents the number of extractions or column passes. Error bars represent one standard deviation, where $n \geq 2$ for LAL reactivity, $n \geq 3$ for sfGFP yield. When not visible, error bars are hidden beneath the markers.

2.3.2 Endotoxin-free ClearColi®-based extract

An alternative to pre-expression endotoxin removal for preparation of an endotoxin-free CFPS system is to prepare extract from an *E. coli* strain that is inherently endotoxin-free. Recently researchers have developed such a strain, the genetically modified BL21 DE3 *E. coli* strain ClearColi® (*ΔgutQ, ΔkdsD, ΔlpxM, ΔpagP, ΔlpxP, ΔeptA*). In the ClearColi® membrane lipid, the core oligosaccharide and two of the lipid chains from the LPS have been eliminated, rendering the lipid unrecognizable to the toll-like receptor 4 (TLR4) complex responsible for the endotoxin reaction in humans (Figure 2-3) [25,44,45]. This strain has therefore been shown in multiple studies to be "endotoxin free" and can be used to produce therapeutics without downstream endotoxin removal steps [25,26,45-47].

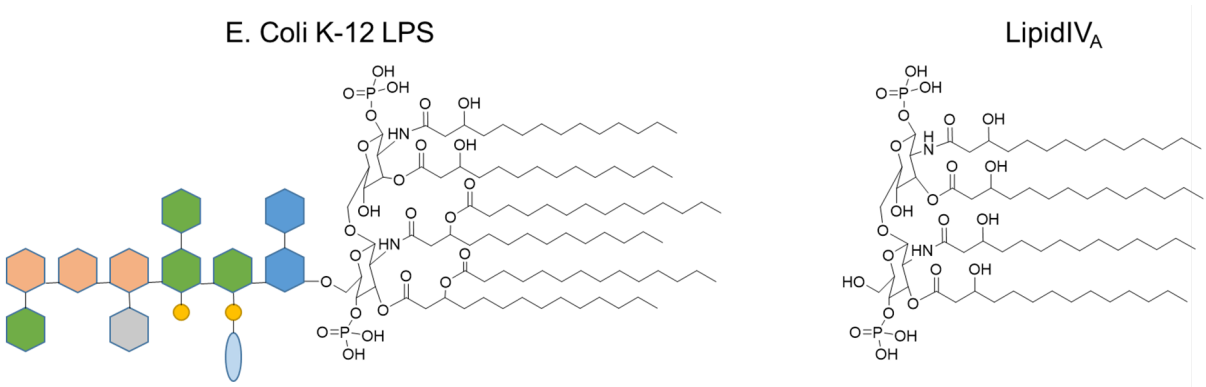


Figure 2-3: E. coli lipid A (LPS) vs ClearColi® lipid IV$_A$. Hexagons represent various sugars: grey = D-galactose, orange = D-glucose, green = l-glycerol-D-manno-heptose, blue = 3-deoxy-D-manno-oct-ulosonic acid. Yellow circle represents phosphate, light blue oval represents phosphoethanolamine.

Here we report the first use of ClearColi® to make cell-extract for CFPS and report its viability in cell-free systems. The ClearColi® extract preparation procedure differs slightly from that of standard extract due to the altered characteristics of the ClearColi® cells. ClearColi® cells are reportedly more sensitive to osmolarity, grow at approximately half the rate of regular BL21 cells, and reach lower final densities [25,48,49]. Gene expression in bacteria changes dramatically with growth rate and cell phase, making induction and harvest points important parameters when preparing extract for CFPS [50,51]. Using the lab-scale shake-flask fermentation approach with commonly used media, induction, and harvest protocols (detailed in 2.2.1 and 0), the protein production yields from ClearColi® extracts were ~20% of regular BL21 DE3 extracts (Figure 2-4). Therefore, in order to create a high-yielding ClearColi® extract, both the effect of media type on the ClearColi® growth and the effect of induction and harvest times on cell-free extract activity were investigated.

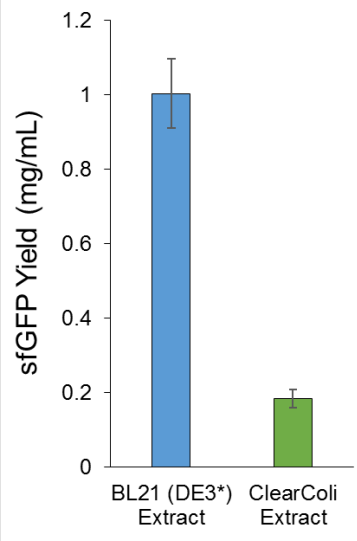
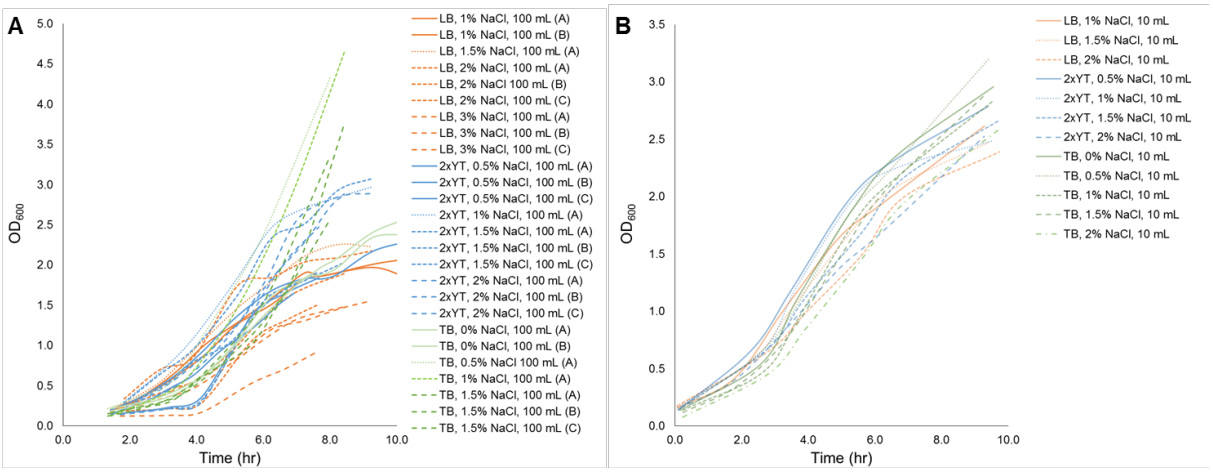


Figure 2-4: Initial ClearColi Extract Yields. CFPS protein synthesis capacity of standard BL21 extract compared to initial ClearColi® extract preparation, determined by sfGFP yields. The initial ClearColi® extract was induced with 1 mM IPTG at an OD600 ~0.53 (~1.75 hours of growth after inoculation with 100 mL intermediate growth) and harvested when the growth rate was seen to decrease, indicating the start of exit from log phase, at an OD600 of 1.47 (approximately 5.2 hours after inoculation with 100 mL intermediate growth, ~3.4 hours after induction). Error bars indicate standard deviation from $n \geq 3$.

2.3.3 Media optimization for ClearColi® extract preparation

Three different media formulations and varying levels of NaCl were tested: LB-Miller (the recommended media for ClearColi® protein preparations), 2xYT media (reported to produce high-yielding CFPS extracts), and TB media (which supports *E. coli* growth to higher cell densities) [35,48,52]. The three media formulations were evaluated at increasing NaCl levels, as enhanced ClearColi® growth has been suggested at higher osmolarity and ionic strength [52]. Thus, NaCl was added to LB, 2xYT, and TB media ranging from original concentration in the media (10 gm/L for LB-Miller, 5 gm/L for 2xYT, and 0 gm/L for TB) up to 20 gm/L. Results indicated that TB media consistently facilitated faster final growth rates and higher cell densities, and that increasing NaCl concentration beyond an additional 5 g/L did not significantly improve cell growth rates and densities (Figure 2-5). Therefore, TB media with an additional 5 g/L or 10 g/L NaCl was used to grow ClearColi® for cell-free extract preparation. While the ClearColi® growth rate with TB was

faster than other media formulations, it remained lower than BL21 (DE3*) (~0.72±0.12 hr^{-1} vs 1.48±0.25 hr^{-1} characteristic growth rate as determined by the Verhulst-Pearl equation, Figure 2-6B). A comparison of the ClearColi® growth kinetics to BL21 (DE3*) is shown in Figure 2-6A.



C

Final OD (after ~8 hrs, shake flask)

	0%	0.50%	1%	1.50%	2%
LB-Miller	--	--	1.86450	2.25600	1.83300
2xYT	--	1.77300	2.84700	2.29400	2.72700
TB	2.03550	4.32600	4.00050	2.96300	--

Final Instantaneous Growth rate (OD ~2-3, shake flask)

	0%	0.50%	1%	1.50%	2%
LB-Miller	--	--	0.03496	0.16319	0.03933
2xYT	--	0.12240	0.04230	0.08921	0.09749
TB	0.15184	0.36097	0.36705	0.38177	--

D

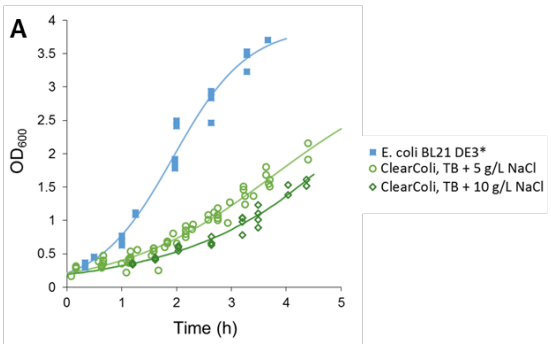
Final OD (9 hrs, falcon tube)

	0%	0.50%	1%	1.50%	2%
LB-Miller	--	--	2.61495	2.47013	2.38973
2xYT	--	2.78723	2.49015	2.66655	2.52510
TB	2.95695	3.19605	2.83823	2.88053	2.58870

Final Instantaneous growth rate (at OD 2-3, falcon tube)

	0%	0.50%	1%	1.50%	2%
LB-Miller	--	--	0.09737	0.06244	0.06306
2xYT	--	0.06568	0.03419	0.07260	0.13043
TB	0.07746	0.12154	0.09203	0.11521	0.10020

Figure 2-5: ClearColi Media Optimization. (A)&(C): Growth curves and final instantaneous growth rates from media optimization for 100 mL growths in 500 mL shake flasks, where orange curves are grown in LB media, blue curves represent growths in 2xYT media, and green curves represent growths in TB media. Note that TB media with 5 g/L or 10 g/L NaCl consistently reaches the highest OD600 during the growth period, and TB with >0 g/L NaCl has highest final instantaneous growth rates (n=1-3); (B)&(D): Growth curves and final instantaneous growth rates (average of n=2) from media optimization for 10 mL growths in 50 mL falcon tubes, where orange curves are grown in LB media, blue curves represent growths in 2xYT media, and green curves represent growths in TB media. Note that TB media again reaches the greatest OD600 and is consistently among the highest final instantaneous growth rates. Oxygen transfer in the 50 mL falcon tubes may have been limiting, considering the lack of air flow through the cap, which may explain some of the discrepancies in final OD600 and instantaneous growth rates between the 100 mL and 10 mL growths.



B

Verhulst-Pearl Equation:

$$P(t) = \frac{kP_0e^{rt}}{k + P_0(e^{rt} - 1)}$$

Cells	Media Base	[NaCl] (g/L)	k ± 95% CI	P_0 ± 95% CI	r ± 95% CI
BL21 DE3*	2xYT	5	3.90 ± 0.37	0.209 ± 0.076	1.48 ± 0.25
ClearColi	TB	5	3.34 ± 1.02	0.207 ± 0.036	0.72 ± 0.12
ClearColi	TB	10	8.00 ± 23.2	0.191 ± 0.065	0.53 ± 0.21

Figure 2-6: ClearColi vs Standard BL21 (DE3*) Growth Parameters. (A) Growth curves for standard E. coli BL21 (DE3*) in 2xYT media and ClearColi® in TB with 5 g/L and 10 g/L added NaCl. Smooth lines represent Verhuslt-Pearl growth equation fits to the data; (B) Parameters obtained by regressing growth data from Figure 2-6A to the Verhulst-Peart equation (right) for time-dependent, self-limiting population growth, where P0 is the initial population, r is the growth rate, and k is the carrying capacity. Note that the growth rate parameter, r, is significantly different between the BL21 DE3* strain and the ClearColi® strains, but the growth rate of the ClearColi® is not significantly different between the two salt concentrations.

2.3.4 Early induction/harvesting produces high-yielding ClearColi® extracts

The induction and harvesting OD_{600} was next optimized for ClearColi® extract preparation. IPTG induction is required for T7 RNA polymerase production and induction at 0.5 OD_{600} with harvest at mid to late log phase has previously been shown to be optimal [35]. The impact of induction and harvest OD_{600} on CFPS activity of ClearColi® extract was assessed (Figure 2-7A). The highest yielding ClearColi® extracts were induced just before log phase (OD_{600} ~ 0.6) and harvested in mid-log phase (OD_{600} 2-3, ~5-7 g cell/L media). Interestingly, comparable CFPS yields could be obtained when extracts were induced earlier (OD_{600} ~0.1) and harvested earlier (OD_{600} 0.6-2.5, ~2-5 g cells/L media). The sfGFP production yields of these highest performing ClearColi® extracts were comparable to standard BL21 yields, at about 80-90% of the yields of standard extract (Figure 2-7B).

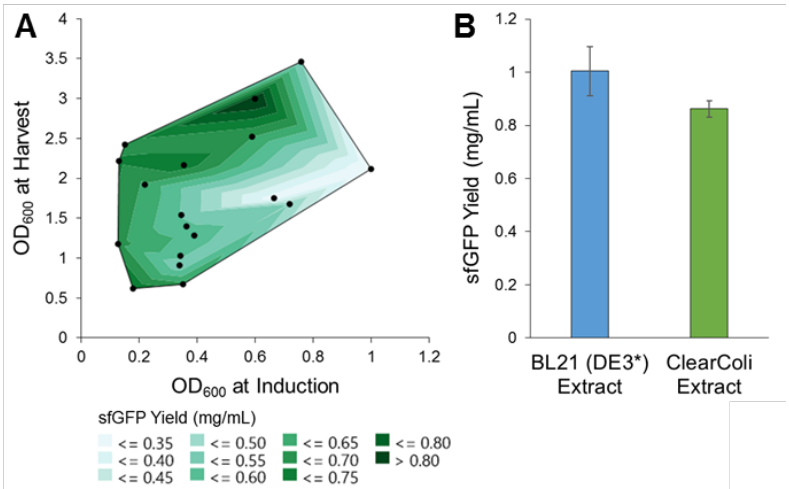


Figure 2-7: ClearColi Extract Induction/Harvest Optimization. (A) sfGFP yield from ClearColi® extracts versus induction/harvest OD600; (B) Comparison of sfGFP yield from CFPS reactions using either BL21 (DE3*) extract versus induction/harvest optimized ClearColi® extract fermented with TB media, $n \geq 3$, error bars represent one standard deviation.

2.3.5 High titers of crisantaspase are produced using reduced-endotoxin CFPS

Crisantaspase (ErA) is an asparaginase derived from *Erwinia chrysanthemi* bacteria which is FDA-approved as a secondary treatment for acute lymphocytic leukemia in patients who have developed immune responses against the primary treatment, *E. coli*-derived asparaginase [53]. To meet restrictions for intravenous administration, the administered endotoxin levels for ErA must not exceed 5 EU/kg patient/hr [42]. Therefore, reducing the initial endotoxin contamination of the sample could significantly simplify downstream purifications for this drug.

Active ErA was produced using the standard BL21 extract, Triton X-114-treated extract, and ClearColi® extract, which is the first time to our knowledge that ErA has been produced using CFPS. ErA yields for each of the three cell-free systems is shown in Figure 2-8A. ErA produced in these reactions was purified via a single Ni-NTA affinity column, and the endotoxin content of each sample was evaluated using the widely used limulus ameobocyte lysate (LAL) assay (Figure 2-8B). A single Ni-NTA spin column treatment reduced the LAL reactivity of ErA by greater than

1000-fold for ErA produced using standard BL21 DE3 extract and Triton X-114-treated extract, and by approximately 20,000-fold for ErA produced from the ClearColi® extracts (Table 2-1). The ~10-fold higher reduction of LAL reactivity from the ClearColi™ CFPS-derived ErA compared to the other extracts is consistent with previous reports that the absence of the oligosaccharide in lipid IV_A makes it easier to remove than standard LPS [25]. The residual LAL activation from the ClearColi® derived product is likely due to the previously reported non-specific nature of the LAL assay, which has been reported to react with a broader spectrum of lipids than does the human immune system [25,46].The ErA produced in the Triton X-114-treated extract had significantly reduced endotoxin concentration compared to the sample purified from standard BL21 extract (Figure 2-8B), However, the lower yields of ErA with this system suggest that, per mg protein, the current embodiment of the Triton X-114 treated extract may not produce sufficiently reduced levels of endotoxin for intravenous administration after a single affinity column purification. By contrast, ClearColi® extracts produced approximately ~70% as much ErA as standard BL21 during short reaction times while Triton X-114 extracts produced ~25%. Extended reaction times enabled comparable yields from both BL21 and ClearColi® extracts, at up to 1 mg/mL ErA, and 0.45 mg/mL from Triton X-114-treated extract (Figure 2-8A). The higher yield of the ClearColi®-derived extract, coupled with the simplicity of extract preparation and the reported activity of the lipid IV_A as an antagonist rather than an agonist for endotoxin response in humans, makes the cell-free ClearColi® system a promising approach to simplify purification requirements of deployable, on-demand therapeutic synthesis [45]. In addition, the ErA produced in the new ClearColi® extract retained the same activity of that produced in BL21 (DE3*) extract (Figure 2-8C).

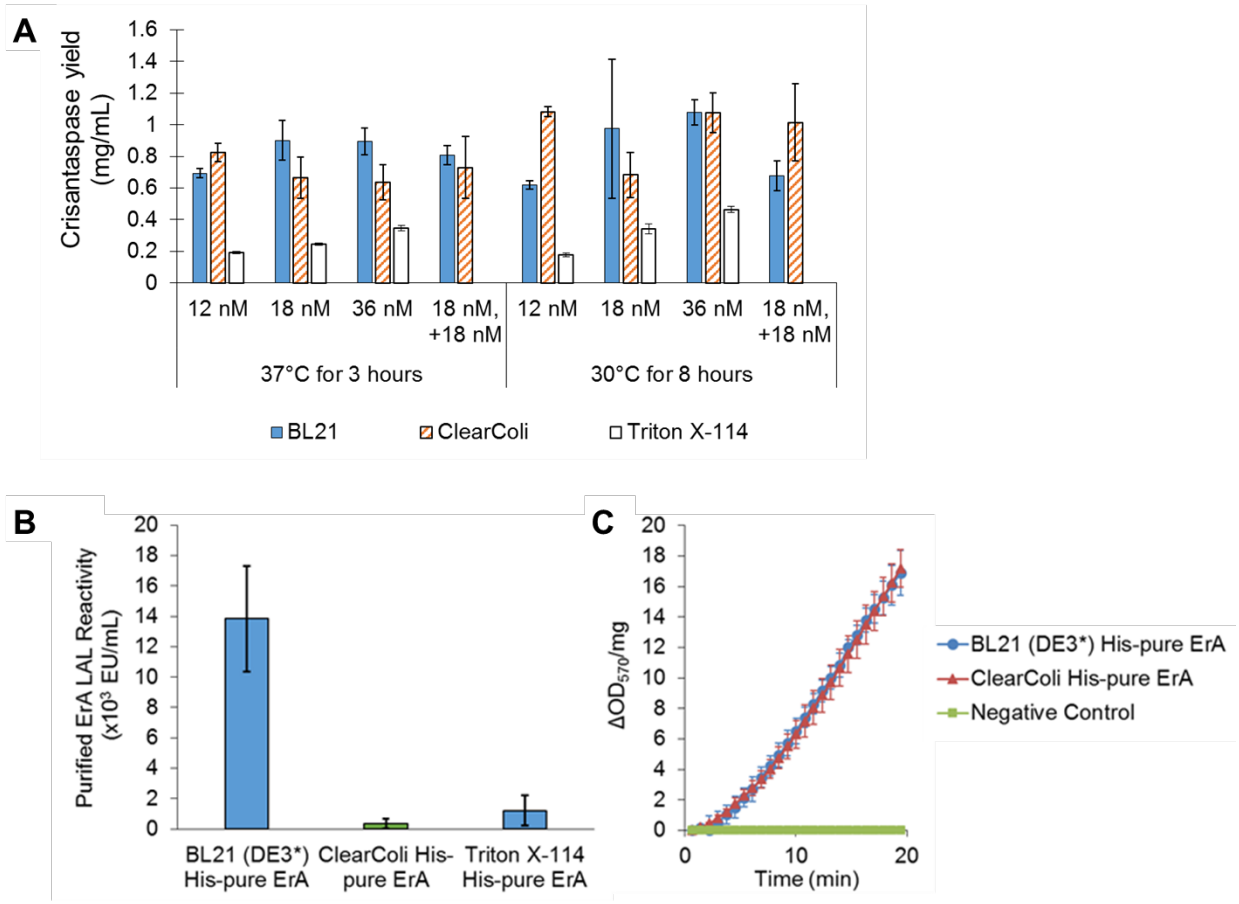


Figure 2-8: Cell-free Protein Synthesis of Active, Endotoxin-free Crisantaspase at High Yields. (A) Endotoxin content determined by LAL kinetic turbidimetric assay of ErA samples purified by His-tag purification with single pass Ni-NTA spin columns from CFPS reactions using BL21 (DE3*) extract, ClearColi® extract, or Triton X-114-treated BL21 (DE3*) extract. The remaining endotoxin levels for the ClearColi® sample are attributed to the non-specific nature of the LAL assay. Error bars represent standard deviations, n=4; (B) Specific activity of ErA samples produced in cell-free using BL21 (DE3*) extract or ClearColi® extract, normalized by protein concentration to 1 mg/mL ErA. Negative control is a buffer-only control. Error bars represent standard deviations, n = 8. (C) Increasing reaction times improves ErA yields. Increasing DNA concentrations were tested at both 37°C for 3 hours and 30°C for 8 hours. In most cases, increasing DNA did not have a significant effect on BL21 (DE3*) or ClearColi® reactions at the shorter time periods. Incremental improvements in the yield from Triton X-114 treated extract were observed at higher DNA concentrations, though yields remained significantly lower than those from the other two extract types. For BL21 (DE3*) and ClearColi®, an additional DNA condition was tested, where reactions were initiated with 18 nM DNA, and then 18 nM more DNA was added after 45 min for 3 hour reactions or 3 hours for 8 hour reactions. No significant improvement in yield was observed for these conditions. For all three extract systems, greatest ErA yields were obtained during the 8 hour reactions. Error bars represent the standard deviation for $n \geq 3$.

Table 2-1: LAL Reactivity of Purified Samples from CFPS Platforms.

	LAL Activity of Ni+ purified sample (EU/mL)		Compared to BL21	Fraction of Initial LAL Activity Remaining		
BL21(DE3)*	13849	± 1739	100.0%	3.34E-04	±	4.20E-05
ClearColi	356	± 171	2.6%	4.91E-05	±	2.36E-05
BL21(DE3)*, 3x Triton X-114	1214	± 502	8.8%	3.36E-04	±	1.39E-04

LAL Activity after one Ni+ affinity purification is significantly reduced for the samples produced via ClearColi CFPS compared to the other two methods. Furthermore, Ni+ purification reduces LAL Activity for the ClearColi sample by nearly 10-fold more than the other E. coli-based samples due to the altered LPS structure. LipidIV_A is detected by the LAL assay, but does not elicit an immune response in humans [25].

2.3.6 Conclusions

The presence of endotoxin is a limitation of *E. coli*-based CFPS systems and holds the risk of septic shock if not adequately removed during purification. Triton X-114 treatment significantly reduced but did not eliminate endotoxin content and allowed up to 45% yields. By contrast, highly active and inherently endotoxin-free extract was produced using the ClearColi® cell strain after extract preparation procedures were modified to account for the altered growth kinetics of this strain. In addition, consistent production of significant amounts of the clinically relevant therapeutic protein crisantaspase (ErA) was demonstrated using this system. The presented platform improves upon the endotoxin-free *in vivo* production enabled by the ClearColi® strain by shifting steps such as fermentation and lysis upstream to enable rapid therapeutic synthesis in a ready-to-use, open format which is more amenable to deployable, on-demand production. The presented platforms are an important step towards on-site therapeutic production, potentially allowing rapid synthesis of therapeutic proteins without the need for complex downstream purifications.

3 ANTIPLASTICIZED LYOPHILIZED CELL-FREE PROTEIN SYNTHESIS SYSTEMS FOR ENHANCED STABILITY AND ON-DEMAND THERAPEUTIC PRODUCTION

Having developed an endotoxin-free CFPS platform, the next aim of this work was to further the potential of this platform to enable rapid, on-demand therapeutic synthesis by stabilizing CFPS reagents for shipping, storage, and on-site protein production. Here, we build upon previous work by our lab, in which standard extract systems were stabilized for room temperature storage via lyophilization, by employing various lyoprotectants to enhance the stability of stored extracts and to develop a stabilized, "just-add-water" system for protein synthesis. The following work is being prepared for dissemination in a peer reviewed journal. The presented work was primarily led and conducted by myself, with contributions from my graduate student colleague Emily Long Zhao and undergraduate students Conner Earl and Rebecca Swensen who I trained and supervised.

3.1 Introduction

Protein therapeutics have emerged as some of the most promising treatments to a wide variety of diseases, including diabetes, arthritis, and cancer [12,16,33,54]. As biotechnology and understanding of disease mechanisms continue to improve, personalized protein therapeutics may become increasingly possible and impactful [8,54]. However, several aspects of the current paradigm for production and distribution preclude the on-site, on-demand synthesis essential for personalized medicine on emergency response. First, biologic production is primarily done in live

cells, necessitating specialized facilities, trained experts, and stringent quality control, which generally restricts the production to centralized facilities [8,9]. Second, limited stabilities of proteins often necessitate cold-chain storage during transportation and storage, which can limit access in remote areas without the proper distribution framework [8].

Cell-free protein synthesis has emerged as a method to circumvent many of the limitations associated with dependence on live cells for protein expression. Cell-free systems provide direct access, facile scalability, and emancipation from cell viability constraints, which can better enable small-scale or potentially automatic therapeutic production and simplified purification [12,13,34]. When coupled with the ability to produce protein directly from linear PCR products, these attributes lend the CFPS system towards rapid on-demand protein production [40,55]. However, cell-free reagents traditionally require the same cold-chain storage as other biological reagents, limiting the potential for distributed, point-of-use synthesis.

For traditional purified proteins, the cold-chain storage and stability issues are often mitigated by storing and distributing the samples in a lyophilized state. However, this strategy precludes on-demand synthesis of small batches of therapeutics, such as would be necessary for personalized therapeutics and stratified medicine [8]. Additionally, protein denaturing and aggregation during the lyophilization and storage process can reduce the activity of the lyophilized therapeutics and increase their immunogenicity [56,57]. An alternative would be to stabilize the expression system for transit and storage, enabling on-site production of fresh protein as needed (Figure 3-1). Accordingly, we recently demonstrated for the first time that lyophilization of cell-free protein synthesis reagents confers stabilizing benefits, enabling enhanced activity after prolonged storage at room-temperatures while simultaneously rendering the reagents sterile [19].

More recently, our lab and others have demonstrated the potential to lyophilize whole cell-free systems onto paper towards low-cost, on-demand biosensing and production of a variety of therapeutics [9,58,59]. While these studies have demonstrated significant progress, they have been limited to room-temperature studies, costly PURE (protein synthesis using recombinant elements) CFPS reagents, extremely small-scale volumes, and have necessitated storage under an inert nitrogen atmosphere and silica desiccant, and hence would be impractical for larger scale reactions for on-demand and stratified medicine applications [59]. Furthermore, actual GFP yields were not reported and therapeutic proteins were produced from extracts which were presumably not stored after freeze-drying. Finally, a straightforward protocol for generation of the microscale lyophilized products is not given. Other previously demonstrated lyophilized CFPS systems have similarly lacked applicability to large-scale operations [60]. Therefore, we sought to create a simple, more robust, scalable procedure for stabilizing extract and whole-cell free reactions at a variety of temperatures relevant for transportation and storage of CFPS reagents for distributed, on-demand therapeutic production.

In order to ensure activity of lyophilized CFPS reagents after transportation to and storage at remote locations, CFPS reagents must be stabilized against degradation at elevated temperatures such as those which may be encountered during unrefrigerated transport. Despite significant improvement over aqueous systems, previously reported lyophilized systems retain minimal viability after even short times at elevated temperatures [19,60]. A method commonly employed to stabilize purified proteins involves addition of sugar molecules, most commonly disaccharides such as trehalose, to the protein formulation prior to drying [57,61-65]. Addition of sugars as lyoprotectants in lyophilized protein preparations is common-place [57,61-65]. However, this stabilization has been only minimally explored towards the stabilization of whole cell-free protein

synthesis systems and was plagued by marked inhibitory effects of the selected lyoprotectant, trehalose [60]. While the exact mechanism of sugar-based stabilization for lyophilized protein preparation is not completely elucidated, two primary mechanisms have been posited: (1) water replacement or (2) vitrification [57,64,66,67]. In the first mechanisms, sugars are hypothesized to gradually replace water in hydrogen bonding at the surface of proteins during the drying process [57,64,66,67]. For this interaction-based stabilization, small sugars such as disaccharides are thought to be most beneficial, as large polymeric sugars are thought to be more rigid and less accessible to hydrogen bonding with the protein surface as compared to small disaccharides [56,57]. Conversely, polymeric sugars are thought to stabilize protein matrices via the second mechanism, vitrification. Relaxation dynamics in glasses take place over multiple time-scales, from picoseconds to milliseconds [56]. The fast relaxation dynamics, referred to as the β relaxations, are related to the vibration of molecules and diffusion and have been shown to be the most related to protein stabilization [56,57]. Slowing these relaxations can stiffen the glass and thereby stabilize the protein. In this way, addition of small amounts of small-molecules to a glass with a higher T_g, termed antiplasticization, has been shown to increase the stability of lyophilized protein/sugar-glass matrices by slowing the fast β relaxations, stiffening the material to slow protein motion despite lowering the T_g [56,63,65,67-73]. Indeed, significant improvements to the stability of lyophilized protein/sugar-glass formulations have resulted from addition of small molecules and intermediate linker molecules to sugar-glasses, particularly for polymeric sugar glasses [67,70]. However, despite the success of antiplasticization in lyophilized purified protein preparations [56,63,65,68-72], the approach has not been investigated in the context of the much more complex environment of lyophilized whole cell-free systems.

We therefore sought to apply the principles of antiplasticization towards the stabilization of lyophilized CFPS systems in order to enhance their stability for storage at elevated temperatures. It has been previously reported that residual water can act as a plasticizer [64,66,74]. Given that we were storing samples exposed to atmospheric conditions, such that some rehydration during storage was a risk, we selected only lyprotectants with high T_g so as to ensure that the overall glass would still have sufficiently high T_g in the presence of small amounts of residual water. Due to the complex and incompletely understood interactions involved in antiplasization, optimal glass-former/linker/plasticizer combinations are difficult to predict [57,70]. Therefore, we selected two antiplasticized polymeric glasses for our study – ficoll/maltitol/DMSO (FMD) and dextran/maltitol/DMSO (DMD) – which have heretofore proved successful for the stabilization of purified protein [70]. In addition, due to the reported success with trehalose as a cryprotectant [60], we also included trehalose and trehalose plasticized by DMSO in our study. Our objectives for this work are three-fold: (1) enhance the stability of CFPS systems to keep greater yield at elevated temperatures, (2) create a CFPS system which retains viability after storage at extreme temperature storage, and (3) create enhanced, truly "just-add-water" system amenable to distributed, on-demand therapeutic synthesis.

Here, we demonstrate enhanced preparation of CFPS systems for on-demand therapeutic production by (1) enhance CFPS stability at elevated temperatures by identifying favorable alternative lyoprotectants which enhance the stability of stored CFPS systems without negatively effecting productivity, including multiple antiplasticized matrices, (2) providing a simple, more industrially relevant protocol to produce lyophilized samples robust against standard atmospheric conditions, and (3) demonstrating a simple, stable, and industrially relevant truly "just-add-water"

cell-free protein synthesis system to enable facile transportation and storage of CFPS systems for on-demand therapeutic production.

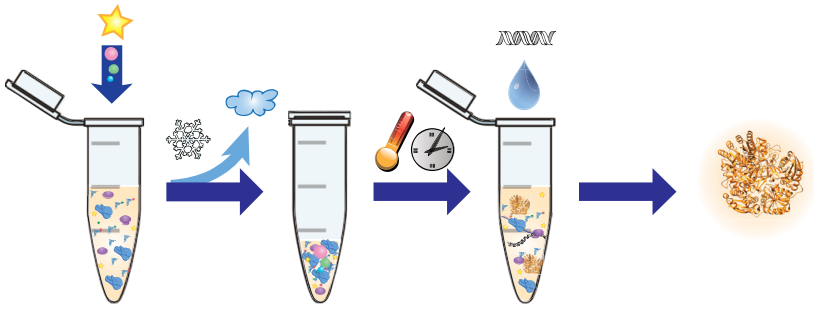


Figure 3-1: Lyophilized cell-free systems enable distributed, on-demand therapeutic protein synthesis.

3.2 Materials and methods

3.2.1 Extract preparation

Extracts were prepared from BL21 (DE3*) and ClearColi BL21 (DE3) strains as described previously [12]. Briefly, cells were grown in 2.5 L tunair flasks (IBI Scientific, Peosta, IA) and production of T7 RNA polymerase was induced via addition of 1 mM IPTG (GoldBio, St Louis, MO) during early exponential growth ($O.D._{600}$ ~ 0.1 – 0.4). Cells were then harvested in mid log phase ($O.D._{600}$ ~2-3) by centrifuging at 6,000 rpm for 10 minutes. Finally, cells were homogenized at 21,000 psig via 3 passes in an Avestin French Press, and the lysate clarified by centrifugation at >12,000 rcf.

3.2.2 Lyophilization

For samples with lyoprotectants added, protectants were added to a final weight percent of 20% of the final solution. For samples lyophilized with PANOxSP energy mix (made as previously described [75]), a 1:1 by volume mixture was used. Two lyophilization methods were initially investigated, using either a LabConco machine and drying multiple samples simultaneously inside

the inner chamber or using a FlexiDry manifold freeze drying with each sample in an individual peripheral tube. Results between the two drying methods were very similar (Figure B-1), so due to the significantly faster drying time with the FlexiDry method, this setup was used for the remainder of the experiments.

Control and cryoprotected samples were loaded into 75 mL cylindrical glass vials (Labglass), and frozen in a -40°C ethanol bath (Just-A-Tilt Shell Freezer Chiller SF-4Az, FTS Systems, Warminster, PA). Vials were then connected to the manifold of the freeze dryer (Flexi-dry MP, FTS Systems). System pressure remained ≤ 300 mTorr, with a collector temperature of -60°C. At 45 minute intervals, samples were removed briefly for weighing, and then promptly re-frozen in the shell freezer and re-attached to the manifold. After approximately 1.5 hours samples reached the minimum weight, corresponding to approximately 10% of initial extract weight. Sample weight remained constant over an additional hour of freeze drying, suggesting that samples were sufficiently dried. Remaining water content was estimated to be < 3% using Equations B-1 through B-3. Dried sample shells were then carefully ground with chemical spatulas, aliquoted by mass into 1.5 mL microcentrifuge tubes for storage. Because the tubes were not vacuum-sealed or back-filled with inert gas, samples were exposed to air during storage.

3.2.3 Rehydration ratio determination

Estimated rehydration ratios were calculated from the initial and dried sample weights according to Equation 3-1, where RHR_{est} is the estimated rehydration ratio and $wt_{initial}$ and wt_{dried} are the initial and dried weights of the sample, respectively. A DC™ protein assay (Bio-Rad, Hercules, CA) was performed on the rehydrated extracts per the manufacturer's instructions to verify that the extracts were rehydrated to the correct protein concentration compared to standard

aqueous extract. Results and the estimated rehydration ratios are included in Figure B-3 and Table B-1.

$$RHR_{est} = \frac{wt_{initial} - wt_{dried}}{wt_{dried}} \quad (3\text{-}1)$$

3.2.4 Cell-free protein synthesis

Yield of the stored extracts at each point was assessed by the production of sfGFP. These reactions were performed as described previously in triplicate [12].

3.3 Results

3.3.1 Antiplasticization improves lyophilized extract storage stability at high temperatures

We first determined the impact of the various lyoprotectant systems on storage stability of lyophilized *E. coli* cell-free extract with the goal to improve the viability of the system above that previously reported [19,60]. *E. coli* extracts were prepared as previously reported [12]. Cryprotectants were then added to the extract to a final weight of 20% of the solution for all samples except ficoll-based samples, in which additives totaled 11% by weight of the solution due to solubility constraints. The 20% or 11% solution weight which accounted for lyoprotectants was split between each additive as reported in Table 3-1, based on previously determined optimal ratios [63,65-67,70]. Samples were lyophilized over a period of 2.5 hours, with less than 3% estimated water remaining in all cases (see Equations B-1 through B-3). Dried extract pellets were then ground and aliquoted into 1.5 mL eppindorf tubes for storage. Rehydration ratios were calculated based on the water removed, and verified via DC assay analysis (Figure B-2, Figure B-3, and Table B-1). At each time point, aliquots of dried extract stored at each temperature were rehydrated with the appropriate amount of sterile ultrapure water by weight. Viability was then assessed by

measuring sfGFP yields produced from each rehydrated sample (Figure 3-2A). sfGFP yields were calculated as previously reported [12]. Results are depicted in Figure 3-2B - E.

Table 3-1: Protectant Fractions in Aqueous Extract Samples Before Lyophilization.

Sample Name	Total $x_{protectants}$	Glass-former fraction	Linker fraction	Plasticizer fraction
Control	0	--	--	--
Trehalose	0.20	1	--	--
Trehalose/DMSO	0.20	0.95	--	0.05
Dextran/Maltitol/DMSO	0.20	0.9	0.05	0.05
Ficoll/Maltitol/DMSO	0.11	0.8	0.1	0.1

3.3.1.1 Extract storage at 4°C

The lowest temperature investigated in our study was 4°C. Because storage at 4°C requires only a standard refrigerator, we reasoned that this storage temperature could reasonably be obtained for long-term storage under many circumstances. At 4°C storage, all samples retained activity very well. Over the course of 3 months, the control, DMD, and FMD systems maintained approximately 90% of the activity of fresh extract. Trehalose-protected samples retained the least activity, at approximately 25% of full extract activity throughout the duration of the test. The sample protected with DMSO-plasticized trehalose initially exhibited low activity (~4%) but appeared to recover slightly over the first few weeks, eventually performing similarly to the trehalose-protected samples after 1-2 weeks. Over long-term storage at 4°C, trehalose/DMSO protected samples performed similarly to trehalose-protected samples.

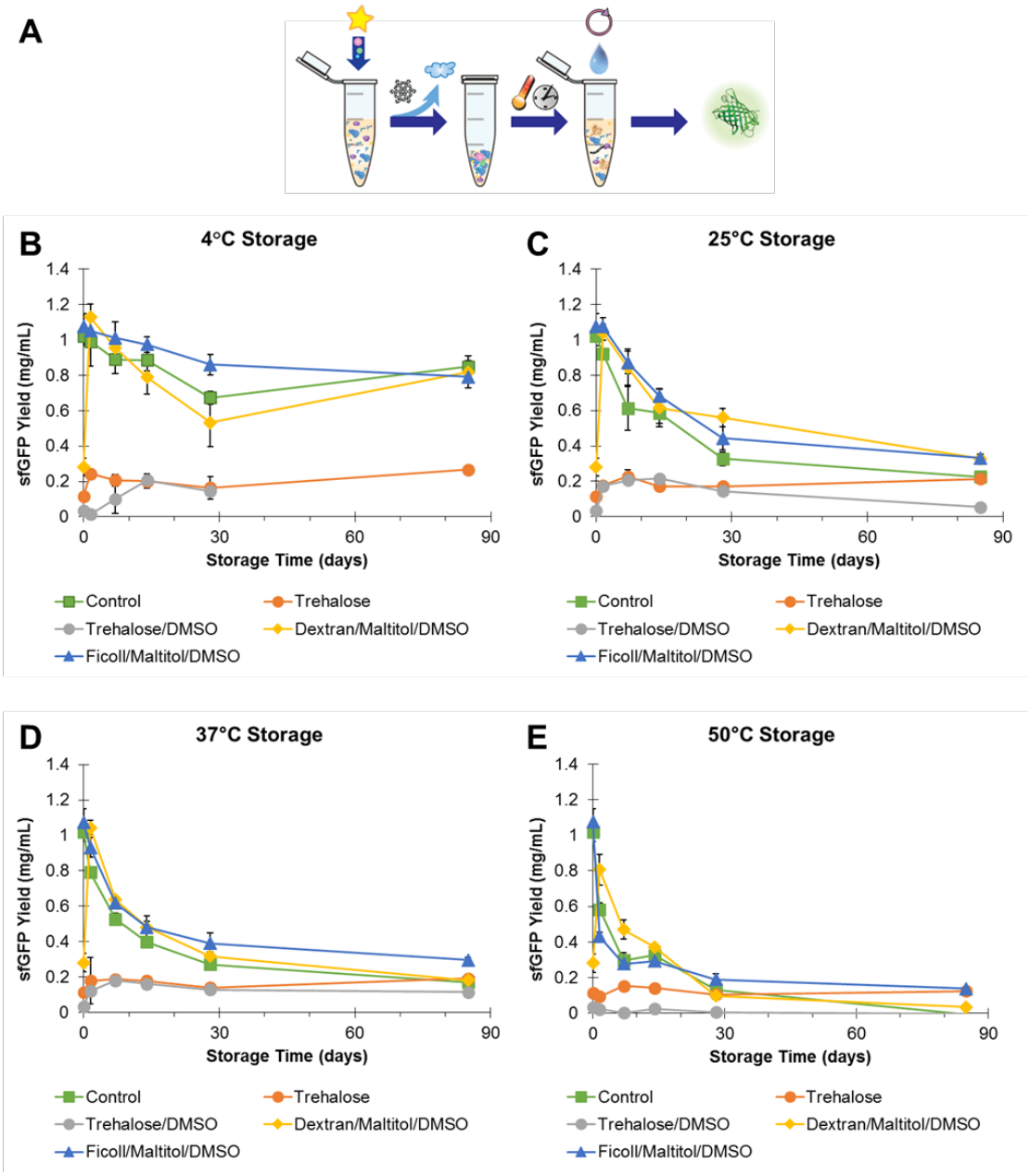


Figure 3-2: CFPS Yields for Extracts Lyophilized with and without Protectants After Storage at 4°C, 25°C, 37°C, and 50°C.

3.3.1.2 Extract storage at 25°C

During storage at a warm room temperature of about 25°C during the experiment, lyoprotectants were more beneficial than at 4°C. While control lyophilized extract maintained approximately 25% activity after 3 months, which is in good agreement with our previously

reported results [19], DMD- and FMD-protected samples each retained upwards of 36%. The polymeric protectant samples also maintained greater than 90% activity over 1 week, which is beneficial for unrefrigerated transport of these freeze-dried systems. As at 4°C, the trehalose- and trehalose/DMSO- protected samples retained the lowest activity at 25°C, remaining at approximately 20% of full activity throughout the duration of the 3-month test. Interestingly, the trehalose/DMSO-protected samples appeared to have an initial recovery period, similar to that seen at 4°C, however the recovery period appeared to be shorter than at 4°C.

3.3.1.3 Extract storage at 37°C

Storage at high temperatures such as 37°C is of interest for transport to remote locations and storage at facilities with no air conditioning or refrigeration capacity. Control and DMD- or FMD-protected samples retained over 50% activity after 1 week storage at 37°C. Control lyophilized extract maintained approximately 19% of full activity after 3 months storage under these conditions. After 3 months, trehalose-protected extract was approximately as active as the control, at about 20% of full activity, having maintained approximately constant activity throughout the 3 month period. Trehalose/DMSO-protected samples again performed very similarly to trehalose-protected samples, reducing in activity after 3 months to about 12%. DMD samples initially maintained slightly higher activity than the extract-only control, but after 3 months were approximately equivalent to the control at 20% full extract activity. Both DMD- and FMD-protected samples maintained >50% activity for nearly 2 weeks. FMD-protected samples performed the best of those tested at >30% full extract activity after 3 months, approximately a 40% improvement over the extract-only lyophilized control.

3.3.1.4 Extract storage at 50°C

Storage at 50°C represents storage under extreme conditions, nearing the maximum recorded temperatures on earth. Storage at high temperatures can also be useful in predicting long-term storage at lower temperatures. At 50°C, all samples reduced to less than 50% activity within the first week. After 3 months, the control extract was no longer viable. Trehalose samples retained about 13% of full extract activity after 3 months, again remaining approximately constant over the entire test period. Trehlaose/DMSO-protected samples performed the poorest, completely losing viability after the first week. DMD-protected samples initially performed the best, retained about 47% activity after the first week, however over the final two months the activity reduced to approximately 3%, barely out-performing the control. FMD proved the most advantageous protectant for extreme temperature storage, maintaining activity similar to the control during the first month and approximately >15% full extract activity after 3 months storage at 50°C.

3.3.2 Lyophilized whole-CFPS system retains high activity at room temperature

We next investigated the potential for lyophilizing a complete CFPS system for "just-add-water" on-demand protein synthesis. Our approach is unique compared to previous approaches, as (1) both energy mix and extract are manufactured in-house, reducing costs compared to previously reported commercial systems [35,59], (2) larger volumes and related methods are more amenable to large-scale production and protein synthesis with minimal specialized equipment compared to previously reported paper-based and micro-scale reactions [59], (3) stability is demonstrated under harsher conditions involved in atmospheric storage where samples are exposed to oxygen and water, and (4) our system demonstrates stability of the high-energy compound PEP, the basis of some of the highest-yielding CFPS energy systems [19,76,77].

Cell-free reactions were assembled as previously described [12], except that additional water and DNA were excluded (Figure 3-3A). The mixed reactions were then lyophilized and aliquoted into 1.5 mL microcentrifuge tubes for storage. Results are shown in Figure 3-3B. After 1 month storage at room temperature, the freeze-dried CFPS reactions ($CFRx_{lyo}$) retained greater than 60% of yields with fresh extract and fresh PANOxSP (~ 0.6 mg/mL). While the yield of the stored $CFRx_{lyo}$ system is initially lower than that of lyophilized extract with fresh PANOxSP ($xSTD_{lyo}$ + $PANOx_{fresh}$), after approximately 2 weeks of storage $CFRx_{lyo}$ such that after 1 month of storage $CFRx_{lyo}$ yields are approximately 2x higher. Although exact yields were not reported in a previous study using trehalose-protected CFPS [60], based on normalized data provided and the standard GFP yields reported the present yields are at least 10x higher.

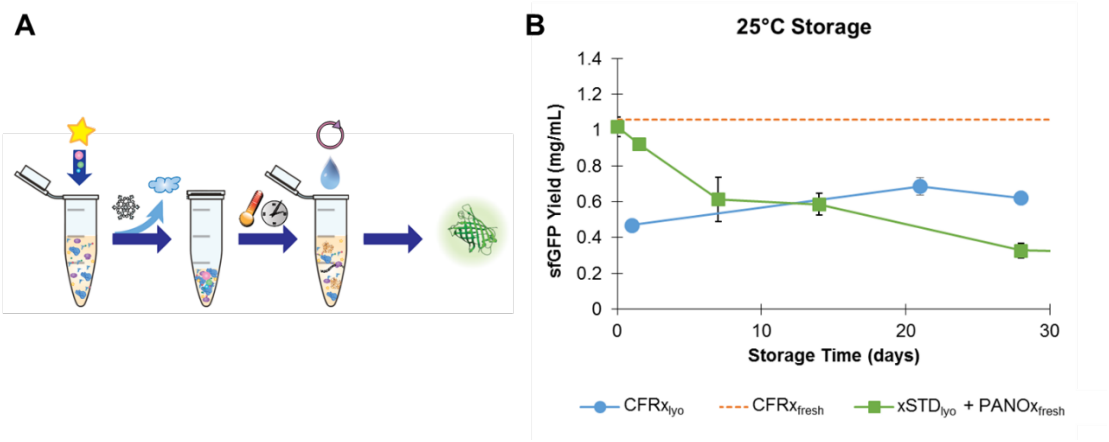


Figure 3-3: Storage of Whole-CFPS ($CFRx_{lyo}$) at 22°C Compared to Stored Lyophilized Extract with Fresh PANOx [$xSTD_{lyo}$ + $PANOx_{fresh}$. For reference the average sfGFP yield fresh extract and fresh PANOx (CFRxfresh) is also shown as a dotted line.

3.4 Discussion

We have demonstrated the feasibility of applying the technology of antiplasticization to enhance the preservation of lyophilized CFPS extract. After 12 hours of storage at 4°C, the control extract lyophilized without lyoprotectants as well as the Dextran/Maltitol/DMSO (DMD) sample and Ficoll/Maltitol/DMSO (FMD) sample had lost no activity over fresh extract samples,

indicating that the extract was robust against the lyophilization method and that these lyoprotectants were not inhibitory of CFPS at the levels added. The negligible loss in activity following lyophilization indicates that protein stability during the lyophilization process is not a major concern for lyophilized extracts. Indeed, components native to the complex extract mixture may already be filling the role of a lyoprotectant in stabilizing the protein components during drying. This hypothesis is further supported by our previous work in which sucrose, another disaccharide commonly used to stabilize proteins during drying, failed to improve the stability or activity of lyophilized extract [19]. Therefore, we hypothesize that stabilization afforded to freeze-dried extract by sugar components can be attributed largely to the second proposed mechanism of sugar-based cryoprotection: vitrification. None of the tested combinations of protectants proved beneficial at 4°C; however, due to the high retained activity it is reasonable that no perceptible benefit would be achieved with the lyoprotectants. It is also reasonable that antiplasticization had no significant benefit at 4°C, as the β relaxations of the lyophilized extract glass at this temperature may be sufficiently slow as to not benefit from increased stiffening of the matrix. At increased temperatures, there was a notable benefit to cryoprotection with the plasticized polymeric glasses. Both DMD and FMD plasticized glasses significantly stabilized lyophilized extract at room temperature, producing nearly 1.5 times the sfGFP yield compared to control lyophilized extract after 3-month storage at room temperature. Similarly, at 37°C and 50°C, the FMD-protected samples outperformed the control extract after 3 months storage and at nearly every time point tested. These results support the hypothesis that degradation of lyophilized extract is primarily due to molecular vibrations, which are increased at higher temperatures. Therefore, plasticized polymeric glasses would provide enhanced stability at temperatures sufficiently higher than the T_g of the resulting glass.

Trehalose has been used successfully in the past to stabilize multiple proteins and was recently demonstrated to improve yield retention of lyophilized CFPS systems [60,65,66,68,71-73]. Because of its small size and relatively high T_g, trehalose may potentially stabilize lyophilized protein formulations by both the water replacement and vitrification mechanisms [60,65,66,68,71-73]. However, due to the complex nature of CFPS systems, a trade-off exists between high initial CFPS yields, which are reduced with increasing trehalose concentration, and high percent activity retention [60]. At 20 wt%, a similar fraction previously reported to be stabilizing for CFPS, trehalose appears to have a pronounced negative effect on the activity of CFPS. Indeed, over every temperature and time point tested, the maximum sfGFP yield obtained with a trehalose-protected system remained less than 30% of full extract activity. Therefore, although trehalose-protected samples did consistently maintain nearly 100% of the initial sample activity for all temperatures, higher even than the retention previously reported [60], the low initial yields of this system render it inferior to other protection systems and even to control lyophilized extract at all conditions except high temperatures for extended periods of time. After extended storage at high temperatures, trehalose-protected extract performs similarly to extract protected with antiplasticized polymeric glass such as FMD. Therefore, in our study we have shown that ficoll antiplasticized with maltitol and DMSO is preferable to trehalose as a lyoprotectant for CFPS reagents.

The high yield of the $CFRx_{lyo}$ samples demonstrates the capacity of CFPS for distributed, on-demand protein production for applications such as personalized medicine, point-of-care therapeutic production, and emergency therapeutic generation such as for pandemic response. Our $CFRx_{lyo}$ systems outperform similarly reported systems and offer an atmospherically robust, more scalable approach than others previously reported [59,60]. The enhanced storage of the combined

CFPS components is in contrast to previous reports which have reported higher yields when energy mix is stored separately from CFPS extract [60], but is reasonable considering previous demonstrations that multiple amino acids, all contained in the energy mix, can act as lyoprotectants [78]. Indeed, we hypothesize that the potentially lyoprotectant activity of the energy mix on the lyophilized extracts increasingly compensates for the loss in energy mix activity over time, rendering the $CFRx_{lyo}$ system more active than $xSTD_{lyo}$ + $PANOx_{fresh}$ after extended storage at elevated temperatures. This could explain why the $CFRx_{lyo}$ system initially has lower yields than $xSTD_{lyo}$ with fresh extract, which could potentially be attributed to an initial pronounced decrease in the activity of the energy system upon storage at elevated temperatures as seen with storage of aqueous PANOx [19]. As the decline in energy mix activity slows, its lyoprotectant activity on the freeze-dried extract may compensate for the decreased energy mix activity my preserving significantly greater extract activity. Future work testing fresh PEP and other high-energy components with a stored, lyophilized extract and amino acid mixture may further elucidate the mechanism of $CFRx_{lyo}$ stabilization. Additional work may also further improve the stability of such systems through thorough optimization of lyoprotectant combinations and ratios, and by investigating the stability under an inert atmosphere.

3.5 Conclusions

We have demonstrated enhanced stability of lyophilized CFPS extracts through antiplasticization. These antiplasticized lyophilized extracts are stable over a broad temperature range, ensuring their stability against a variety of conditions potentially encountered during various transportation methods and storage conditions. Furthermore, we have demonstrated the storage stability of a one-pot "just-add-water" CFPS system under atmospheric storage conditions. Coupled with our recent demonstration of high-yield endotoxin-free extract production, the

production of a highly active, one-pot “just-add-water” CFPS system using straight forward in-house reagent preparation procedures and minimal specialized equipment is an important step towards industrially relevant production of protein expression systems for on-demand therapeutic manufacture.

4 THE LOCATIONAL IMPACT OF SITE-SPECIFIC PEGYLATION: STREAMLINED SCREENING WITH CELL-FREE PROTEIN EXPRESSION AND COARSE-GRAIN SIMULATION

PEGylation is an invaluable approach to stabilizing and mitigating the immunogenicity of protein therapeutics, however, locational impacts of PEGylation are not well understood. Therefore, for rapid development of enhanced therapeutics for on-demand applications such as stratified medicine, a rapid screening approach is needed to identify optimal PEGylation sites. Chapters 4 and Appendix C are adapted from the publication "The Locational Impact of Site-Specific PEGylation: Streamlined Screening with Cell-free Protein Expression and Coarse-grain Simulation" (Wilding et al. 2018) published in ACS Synthetic Biology in February 2018. The presented work was designed, led, and conducted by myself, with contributions from my graduate student colleagues Addison Smith and Derek Bush as well as undergraduate student Joshua Wilkerson.

4.1 Introduction

Since its invention in the 1970's, PEGylation has proved to be a valuable tool for pharmaceutical applications [79-82]. Several PEGylated therapeutics are already available for clinical use including the 15th best-selling pharameutical Neulasta [2,5,6,80]. PEGylated therapeutics are reported to have improved pharmacokinetics and reduced immunogenicity, due to

slower renal filtration and increased resistance to degradation and aggregation [83-87]. Enzyme biocatalysts could similarly benefit from PEGylation through improved stability and greater hydrodynamic radius, leading to improved recoverability and retention in matrices [5,88]. However, commercially available PEGylated proteins to date are non-specifically PEGylated, targeting multiple natural residues such as lysine or cysteine [4,89], or are produced by targeting naturally occurring, uniquely reactive sites, such as the N- or C- terminus [6,90], disulfide bonds [91], or at less prevalent natural amino acids such as cysteine which have been mutated into the protein [92,93]. Because these techniques limit the sites available for targeting, the tethering locations on the protein may be in suboptimal locations such that conjugation would hinder protein stability or important protein-protein interactions [90], both of which can dramatically reduce the protein's activity. In some proteins, application of these techniques may also require extensive mutagenesis in order to prevent undesired PEGylation where the targeted moiety occurs naturally in the protein.

Site-specific insertion of unnatural amino acids (uAA) via stop codon suppression offers the ability to conjugate a protein at potentially any site with minimal mutation [24,94,95]. The wide variety of uAA with useful side-chain chemistries adds a flexibility to uAA-based conjugation which makes it a powerful tool for producing optimally PEGylated proteins. However, the optimal site for PEG conjugation is not well understood and activities of different analogs can vary greatly [24,40]. In order to mitigate the costs associated with development of optimized PEGylated proteins, improved guidelines are necessary to inform conjugate design.

E. coli-based cell-free protein synthesis (CFPS) offers an ideal platform for rapidly and economically screening various sites for uAA incorporation [40,96] and PEGylation, providing a promising tool for both developing guidelines to inform PEGylated protein design and identifying

optimal PEGylated proteins from a pool. The flexibility of the CFPS has enabled cell-free synthesis of a wide variety of challenging proteins, including cytotoxic proteins [33], disulfide-bonded therapeutics [15], virus-like particles [21,97,98], proteins requiring chaperones [99], and antibodies [94]. Although the *E. coli*-based cell-free system is currently unable to replicate mammalian glycosylation patterns, recent advances in engineering glycosylation into the system and advances in uAA incorporation are promising developments towards this end. Here we present a cell-free based approach that allows rapid, scalable assessment of PEGylated proteins, enabling optimization of PEGylated proteins in a more time- and resource-efficient manner.

As a proof-of-concept study, we apply our system to evaluate existing guidelines for PEGylated protein design, specifically (1) PEG size and number as a factor in PEGylation efficiency [79], protein stability [83,87,89,100-102], and protein activity [85,89,103,104]; (2) PEGylation site solvent accessibility as a predictor of efficient PEGylation [105-108]; and (3) flexible loops as preferred conjugation sites to minimize stability and activity loss [82,109,110]. Finally, to further enhance the screening process, we also develop a coarse-grain molecular simulation to inform candidate site selection and demonstrate the utility of the simulation for coarse PEGylation site assessment.

4.2 Materials and methods

4.2.1 Extract preparation

Extracts were prepared using an *Escherichia coli* BL21*(DE3) pEVOL-AzF strain, a kind gift from Peter Shultz [111,112]. The extract was prepared in a manner similar to that which has been described previously [14,35,37,113,114] with a few modifications. Cells were grown at 37°C

and 280 rpm in sequential growths. Growths were started in 5 mL of 2xYT media, incubated overnight, and moved into 100 mL 2xYT. The 100 mL growth was then incubated until an O.D. of 2.0 and then added to 900 mL L 2xYT media in a 2.5 L Tunair baffled shake flask (IBI Scientific, Peosta, IA) for a final volume of 1 L. When the 1 L growth reached an O.D. between 0.5 and 0.7, cells were induced with both 1 mL of 1 M isopropyl-1-thio-b-D-galactopyranoside (IPTG) and 0.20 g arabinose. Cells were then monitored and harvested in mid-log phase (an O.D. of ~2.0 in this work) at 8,000 rpm for 30 minutes. After washing in Buffer A, cells were re-suspended in Buffer A at a ratio of 1 mL per gram and lysed in three passes through an Avestin Emulsiflex B-15 cell disruptor (Ottawa, Canada) at 21,000 psi. Lysate was centrifuged at 12,000 rcf for 10 minutes, following which supernatant was removed and incubated at 37°C for 30 minutes. Extract was then flash frozen and stored at -80°C until use.

4.2.2 Plasmid preparation

A cysteine-free T4 lysozyme (T4 Lyz) variant was obtained from Addgene (Cambridge, MA) and cloned into the pY71 plasmid, and a C-terminal strep-tag was added for purification purposes, as described previously [109]. Six variants – K16Amber, S44Amber, N53Amber, L91Amber, K135Amber, and Ins163Amber – were created using the Quikchange II mutagenesis protocol (Agilent Technologies, Santa Clara, CA) [109]. A seventh variant with two Amber codons was created also using the Quikchange II mutagenesis protocol to insert an additional amber stop between K162 and N163 on the N53Amber variant. This variant is hereafter referred to as N53AzF/Ins163AzF. Plasmids were purified for use in cell-free protein synthesis using a Qiagen Plasmid Maxi Kit (Valencia, CA).

4.2.3 Cell-Free protein synthesis

Cell-free protein synthesis was performed using the standard PANOxSP system, with a few modifications [14]. The reaction mixture was as follows, with components obtained from Sigma-Aldrich unless otherwise specified: 25% v/v *E. coli* pEVOL pAzF extract, 25% v/v 19-amino acid PANOxSP mixture (Glutamate was added separately as a salt with Mg in order to optimize the Mg content of the reaction), 18 mM $Mg(Glu)_2$, 12 nM plasmid purified with Qiagen Plasmid Maxi Kit, 5 μM C_{14} leucine (PerkinElmer), 3 mM AzF (Chem Impex International, Wood Dale, IL), and the remaining volume distilled deionized water. For synthesis of the N53AzF/Ins163AzF variant, AzF was added at 6 mM in order to improve yields. The reactions were assembled under a safe-light and incubated in darkness in order to preserve the azide group, which decays upon exposure to UV or near-UV light [14,112,115]. Reactions were performed in 15 mL falcon tubes (GeneMate) at 300 – 400 μL volumes and incubated overnight (~15 hours in this work) at 30°C. Negative control reactions were performed in 50 μL volumes in 2 mL microcentrifuge tubes (GeneMate) for the same time at 30°C. Purified synthetase was not added as yields were sufficiently high with only the synthetase provided in the prepared pEVOL extract. Total protein synthesis yields were determined using liquid scintillation as discussed previously [14], using 5% trichloroacetic acid for protein precipitation. Synthesis of full-length protein was verified by running 3 μL of the CFPS reaction on a NuPAGE 10% Bis-Tris Gel (Invitrogen, Carlsbad, CA). The gels were run according to manufacturer's instructions, at 200V for 35 minutes using MES buffer. After running, gels were stained with SimplyBlue SafeStain (Invitrogen), dried, and autoradiograms were performed using Kodak MR Autoradiogram Films with 2 day exposure time. Yields of full-length variants were determined from densitometry using ImageJ software [116] to compare the relative band intensities to that of WT lysozyme and scale the yields accordingly to the WT yields.

4.2.4 T4 lysozyme purification

T4 Lyz was purified using Strep-Tactin Spin Columns (IBA Life Sciences, Gottingen, Germany) according to manufacturers' specifications, with the following variations. To improve recovery, CFPS samples were run through the spin columns three times and columns were then washed 5 times with the provided Buffer W. T4 Lyz was then eluted according to the procedure specified for high concentration. Liquid scintillation was used to determine the concentration of the purified product. Using the total protein yields calculated from the scintillation of the CFPS reaction, the CPM/mg/mL was determined for each sample, which was then used to calculate the concentration of T4 Lyz variant in the purified samples.

4.2.5 PEGylation reactions

Conjugation reactions were performed using strain-promoted azide-alkyne cycloaddition, or SPAAC. The number of PEG equivalents was optimized to obtain maximal conjugation with the minimal allowable PEG. The optimal PEG equivalents was determined to be 20 for 20kDa PEG and 50 for 5kDa PEG. Reactions were assembled similarly to protocols which have been described previously [117,118], with the following specifications. In a PCR tube, 5 μM T4 lysozyme was combined with 20 or 50 equivalents of 20kDa or 5kDa DBCO-mPEG (Click Chemistry Tools, Scottsdale, AZ), respectively, in PBS buffer. PEG equivalents were doubled for the double amber suppression variant in order to obtain higher conjugation efficiency. The reactions were incubated at 37°C and 280 rpm for 18 hours. Although complete conjugation with SPAAC reactions has been reported with much shorter reaction times [94,118], an 18 hour reaction time was used to maximize conjugation at low-efficiency sites and thereby obtain more accurate stability and activity data for these T4 Lyz analogs. Control reactions were also assembled with WT T4 Lyz, mimicking the conditions of the SPAAC reactions for both PEG sizes and unconjugated controls. These reactions

provided a baseline to account for the effects of incubation time and unconjugated PEG on the stability and activity of the lysozyme.

4.2.6 PEGylation efficiency analysis

To determine efficiency of each PEGylation reaction, a sample of each reaction was run on a NuPAGE 10% Bis-Tris Gel at 200V for 35 minutes using MES buffer. The gels were stained with SimplyBlue SafeStain (Invitrogen), dried, and used to produce an autoradiogram. Autoradiograms were produced using Kodak MR Autoradiogram Film with a 2 day exposure. Using ImageJ, the relative intensities of the PEG-shifted bands and the un-shifted bands were calculated. The click efficiency was determined as the intensity of the PEG-shifted band divided by the sum of the intensity of the un-shifted and shifted band(s) in the gel lane. PEGylation efficiencies were calculated from the autoradiogram to eliminate error from possible interference of PEG with staining.

4.2.7 Stability analysis

4.2.7.1 Stability assay

Stability was analyzed by examining the shifts in the protein melting temperature, T_m. Melting temperatures were determined using the Protein Thermal Shift Assay (Thermo Fisher Scientific, Carlsbad, CA) and the corresponding Protein Thermal Shift Software, version 1.3. This assay uses a hydrophobic dye which fluoresces when it binds to the hydrophobic regions of the protein as they are exposed during melting. The protein is combined with the dye and gradually stepped through increasing temperatures in a real-time PCR machine, which monitors the change in florescence. The florescence curve is then used to determine the melting temperature. Melting temperatures were calculated using the derivative method to avoid skewing of a Boltzmann fit by

any remaining peak for the unconjugated protein. The derivative method identifies the T_m by computing a second-derivative to identify the inflection point of the fluorescence curve.

The melt reactions were assembled according to the manufacturer's instructions, with the following specifications: 5 μL Protein Thermal Shift Buffer, 7.5 μL PBS buffer, 5 μL click reaction, 2.5 μL Diluted Protein Thermal Shift Dye (8x) diluted with PBS buffer. Reactions were carried out in triplicate, assembled in a 96-well, semi-skirted Framestar Fast Plate (Midsci, St. Louis, MO) and covered with a MicroAmp Optical Adhesive Film (Applied Biosystems,Thermo Fisher Scientific). The assay was set up using StepOne Software v2.3, and run in a StepOnePlus™ Real-Time PCR System (Applied Biosystems). In order to obtain maximal resolution of the T_m, the assay was run using a standard ramp speed from a starting temperature of 25°C to 99°C. Analysis was done using the derivative method. Results are reported as a change from the average T_m of the corresponding WT reaction (no PEG, 5kDa PEG, or 20kDa PEG), ΔT_m. For the simulation comparison, the experimental data is reported as a difference in T_m between a PEGylated T4 Lyz variant and the corresponding unconjugated variant, $\Delta T_{m,conj}$.

4.2.7.2 Simulation-predicted T_m

Coarse-grain simulations were performed using the Go-like model and replica-exchange algorithm which has been described previously [119-122]. In this coarse-grain model, each residue is approximated as a single site, centered at the location of the C_α atom of the residue in the crystal structure. The model input files were created using the MMTSB Web Service (mmtsb.org/webservices/gomodel.html) [123,124] based on the T4 Lyz structure from the Protein Data Bank (PDB ID: 2LZM). PEG monomers were approximated as single residues with a radius of 1.48 angstroms and a center-to-center distance of 3.7 angstroms, and the DBCO connecting group was approximated as a residue with a radius of 5.18 angstroms. These distances were obtained using

a GaussView [125] model of the SPAAC product and attached PEG chain. Bond energies were also obtained from GaussView [125]. Other relevant parameters, including Lennard-Jones parameters, were chosen to resemble those of the peptide bonds within the protein. Additional details pertaining to the model form are included in Appendix C section C.1.

PEG chains were added to the protein model extending linearly away from the attached site. The PEG chain was then allowed to equilibrate to a more realistic conformation in NVE MD equilibrium simulations with constraints in place to hold the protein in the properly folded conformation while the PEG polymer equilibrated. Results of these equilibration simulations were then used as starting points for replica exchange simulations where an additional unconstrained equilibrium phase ensured system equilibration before production steps were recorded. Replica exchange simulations were done in the NVT ensemble using three Nose-Hoover thermostats, a time step of 3 fs, and a mass of 7.81338×10^{-22} kg*angstroms^2. Each simulation contained 20 x 10^6 equilibration steps and 60 x 10^6 productions steps. Sixty-six boxes, with temperature steps of 1.5K between boxes in the 21K range surrounding the expected melting temperatures (345K – 366K for 5kDa conjugates, 327K – 348K for WT and 20kDa conjugates) and steps of 3K for the remaining range, were used for all simulations with a box size of 4000 x 4000 x 4000 angstroms. The replica exchange simulations were run 10 times for all cases except for K16 and S44 PEGylation models, which were run 20 times due to slightly more scatter in the data. The results for all runs were averaged and the standard error reported.

The average relative native contacts, representing the remaining interactions pertaining to the protein secondary and tertiary structure relative to the 100% folded state, was plotted versus temperature. This plot shows how the degree of protein folding decreases over temperature as the protein unfolds, and is analogous to the fluorescence versus temperature plot obtained from the

experimental assay. To mimic the experimental analysis, the derivative of the relative native contacts was calculated numerically with respect to temperature using a central difference formula. The temperature at which the magnitude of this derivative was greatest was taken to be T_m, as it represents the point where the rate of change in the degree of protein folding is greatest, and corresponds with the point where the rate of increase in fluorescence is greatest. This point also corresponded with the maximum heat capacity value, which is expected as T_m is also often determined to be the temperature at which heat capacity is a maximum. The melting points calculated from simulation were then compared to the experimentally determined T_m.

4.2.8 Activity assay

Activity of each lysozyme construct was determined via the EnzChek Lysozyme Assay Kit (Thermo Fisher Scientific). Reactions were performed according to the manufacturer's instructions, with the following specifications and modifications. EnzChek buffer was not used, as it has been reported that T4 Lyz is less active at high ionic strengths [126]. Unpurified PEGylated lysozyme from each SPAAC reaction was added to each reaction at a final lysozyme concentration of 0.06 μM. This lysozyme concentration was selected because it did not cause the activity assay to saturate and required less dilution than other suitable concentrations, which eliminated unnecessary error. In a black 96-well untreated polystyrene plate (VWR), 1.2 μL of unpurified SPAAC reaction was added to 48.8 μL ddH2O and mixed thoroughly by pipetting. The 1 mg/mL fluorogenic substrate was diluted 20x in ddH2O, and 50 μL was added to each well. The plate was then immediately placed into a plate reader (Synergy MX, BioTek) and incubated at 37°C. Fluorescence was monitored every 1.5 minutes for 2 hours (494nm/518nm). Activity was determined from the endpoint fluorescence of each well and normalized to the endpoint activity of the corresponding WT reaction. Each sample was assayed in n=3+. Relative activity of each PEGylated lysozyme analog was calculated from the

mixture activity data, assuming additive activities of the PEGylated and unPEGylated lysozyme. WT activity data for a set of assays was averaged, and then conjugate activity (a_c) was calculated for each replicate according to Equation 4-1, where x_c and x_{uc} is the fraction of conjugated and unconjugated T4 Lyz, respectively, a_{rxn} is the activity of the reaction mixture, and a_{uc} is the activity of the unconjugated variant. The activity for di-PEGylated N53AzF/Ins163AzF (a_{c2}) was calculated as shown in Equation 4-2, where the activity of mono-PEGylated N53AzF/Ins163AzF (a_{c1}) is calculated as shown in 4-3, where $a_{N53AzF,c}$ and $a_{N53AzF,uc}$ is the activity of conjugated and unconjugated N53AzF, respectively, with analogous activities for Ins163AzF and N53AzF/Ins163AzF. Calculated activities were then averaged over all replicates and over all sets of assays.

$$a_c = 1/x_c * (a_{rxn} - x_{uc}a_{uc}) \quad (4\text{-}1)$$

$$a_{c2} = 1/x_{c2} * (a_{rxn} - a_{uc}x_{uc} - a_{c1}x_{c1}) \quad (4\text{-}2)$$

$$a_{c1} = 0.5\left(\frac{a_{N53AzF,c}}{a_{N53AzF,uc}} + \frac{a_{Ins163AzF,c}}{a_{Ins163AzF,uc}}\right) * a_{N53AzF/Ins163AzF,uc} \quad (4\text{-}3)$$

4.3 Results and discussion

4.3.1 CFPS-based screening system

In this study we first seek to demonstrate the utility of a cell-free expression system to quickly assess the impact of PEG size and conjugation site on the properties of a target protein. Using CFPS as opposed to *in vivo* methods reduced resource costs and accelerates protein preparation from weeks to days by (1) eliminating cell culturing of each protein variant, enabling protein synthesis in ~8 hours, (2) allowing microliter scale protein synthesis reactions for efficient reagent use, and (3) simplifying protein purification. Additionally, in less than 5 hours and with less than 2 μg of each protein sample, we are able to screen PEGylation efficiency and thermal

stability in parallel with protein activity. We selected T4 lysozyme (T4 Lyz, EC 3.2.1.17) as a model enzyme for this study, and at the conclusion of our study, we evaluate our results against several generic PEGylation guidelines.

4.3.1.1 CFPS of screening pool

Six sites were selected for uAA p-azido-L-phenylalanine (AzF) incorporation, spanning a range of solvent accessible surface areas (SASA) and structures. The selected sites are as follows: substitutions at sites K16, S44, N53, L91, and K135, and an amber stop insertion between residues K162 and N163 (Ins163). Locations of the selected sites along with their corresponding relative SASA values, as scored by ASA-View [127], are reported in Figure 4-1A. Plasmids were constructed to incorporate the amber stop codon, TAG, at each of these sites, as detailed in section 4.2.2.

We also constructed a double amber suppression variant, as di-PEGylation may further stabilize proteins compared to mono-PEGylation [87] [128], Di-PEGylation may also provide additional advantages for therapeutic proteins, such as reduced vacuole formation in the kidneys [129]. We constructed a double amber suppression variant with sites N53 and Ins163 (N53/Ins163), because they have high solvent accessibility and are located on opposite sides of the protein.

Using an *E. coli*-based cell-free system, AzF was incorporated as reported previously [14], resulting in the production of the following lysozyme variants: K16AzF, S44AzF, N53AzF, K135AzF, Ins163AzF, and N53AzF/Ins163AzF. As our lab and others have observed previously [24,40,130], yields of full-length lysozyme varied depending on the location of uAA incorporation. Average yields fell between 160-500 μg/mL (Figure 4-1B) – exceeding 25% of wild-type (WT) yield in all cases – and very low levels of full-length T4 Lyz expression were observed in the absence of AzF. The yield of the double-AzF-modified protein was 0.28 mg/mL, significantly

exceeding previously reported yields for double AzF-incorporation via amber suppression [104,107,110]. The ability to achieve higher yields in hours, even for double amber suppression, is a primary advantage of the CFPS expression system as a basis for a conjugate screening platform. Full-length AzF-incorporated variants were then rapidly purified using spin columns, as detailed in section 4.2.3.

Using the cell-free system enables sufficient protein yields for our screen from ≤ 400 µL reactions in just 8+ hours with minimal optimization. As such, CFPS reduces time and resource costs of the screen by eliminating the need for cell culturing of each construct. This system also simplifies purification by eliminating the lysis and clarification steps between expression and purification, allowing direct addition of the expression reaction to a spin column for rapid, small-scale purifications.

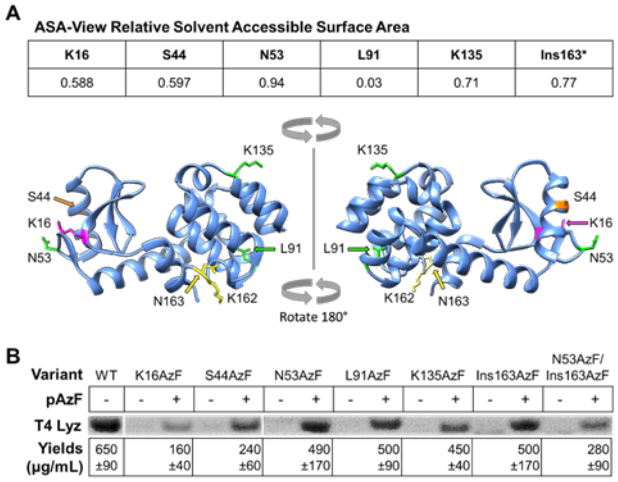


Figure 4-1: T4 Lysozyme Variants and Yields. (A) ASA-View relative solvent accessible surface areas, normalized to the average surface area of each amino acid in 30 random configurations of a Gly-X-Gly tripeptide (*SASA reported for Ins163 insertion site is an average of the ASA-View values for the SASA of K162 and N163), and Chimera image showing locations of these residues in T4 Lyz (PDB ID: 2LZM). Residues depicted in green are part of unstructured loops, while structured sites are represented in pink or orange (pink for beta sheet, orange for alpha helix). The residues surrounding the insertions site are shown in yellow, and are also part of an unstructured loop; (B) Full-length T4 Lyz CFPS yields and autoradiogram. Data shown are averages and standard deviations of n ≥ 2.

4.3.1.2 PEGylation efficiency screening

Purified T4 Lyz variants were PEGylated with 5kDa and 20kDa PEG using a strain-promoted azide-alkyne cycloaddition (SPAAC) reaction. The SPAAC reaction, illustrated in Figure 4-2A, is advantageous because it can be done rapidly and at physiological conditions without additional components such as protecting groups or copper catalysts [14,131,132]. SPAAC reactions were performed using 5 μM purified T4 Lyz variant and 20 or 50 equivalents, respectively, of 20kDa or 5kDa PEG with a dibenzocyclooctyne (DBCO) terminal group providing the strained alkyne DBCO-mPEG. The DBCO-mPEG equivalents were doubled for PEGylation of N53AzF/Ins163AzF. These conditions provided high click efficiency at most of the chosen sites.

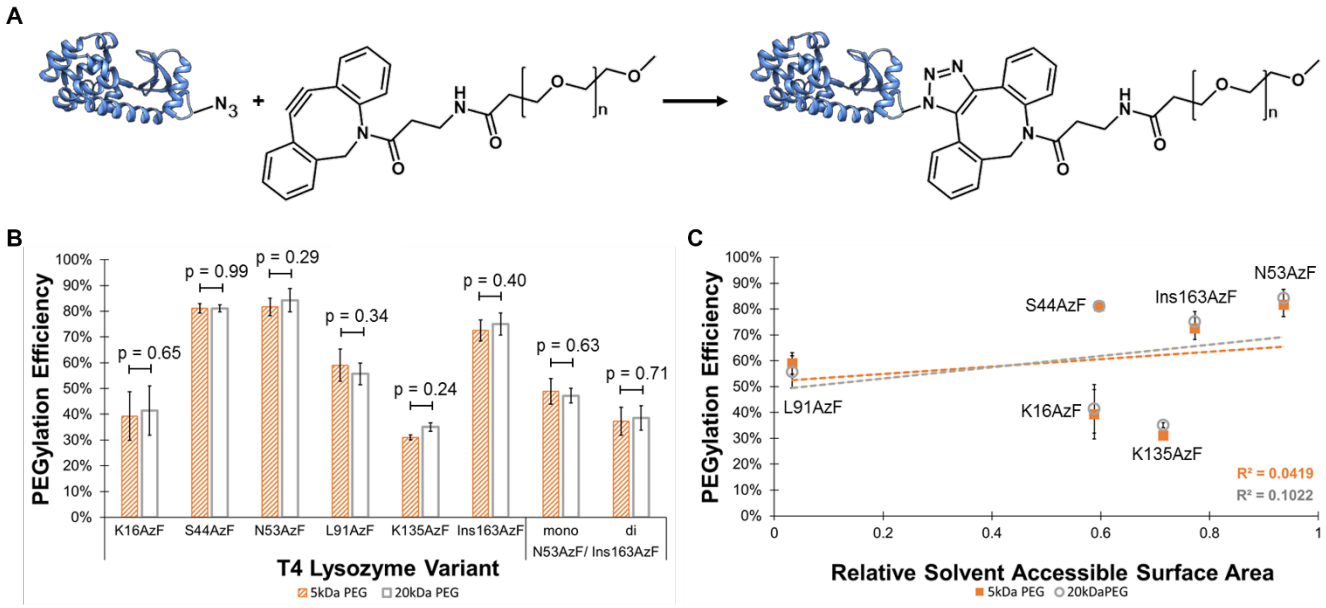


Figure 4-2: PEGylation Reaction Scheme and Efficiency. (A) Illustration of the SPAAC reaction used for PEGylation where DBCO-mPEG and AzF-substituted T4 Lysozyme are covalently conjugated; (B) PEGylation efficiency of T4 lysozyme variants ($n \geq 2$, error bars represent one standard deviation, p values are reported for each variant comparing the 5kDa to the 20kDa PEGylation efficiency and $p > 0.05$ in all cases); (C)PEGylation efficiency vs relative SASA, note that relative SASA does not appear to be linearly correlated with conjugation efficiency in this study ($n \geq 2$, error bars represent one standard deviation and are hidden under marker when not visible). Dashed lines represent a linear regression, with the orange line fit to the 5kDa PEG data and the gray line fit to the 20kDa PEG data. The corresponding R2 values are also reported.

Average PEGylation efficiency is shown in Figure 4-2B as calculated by SDS-PAGE electrophoresis, autoradiography, and densitometry (Figure 4-3). As alluded to previously, the conjugation sites selected in this study were chosen in part to evaluate PEG size and conjugation site accessibility, which are common guidelines for the design of PEGylated proteins with high conjugation efficiency [79,105-108]. The results of our screen revealed no significant impact of PEG size on click efficiency with the 5kDa and 20kDa unbranched-PEG molecules employed in our study, even for the di-PEGylated N53AzF/Ins163AzF (Figure 4-2B). This result suggests that the accessibility of the site to the larger DBCO group plays a stronger role in conjugation efficiency than the 15kDa increase in size of the flexible PEG chain. Indeed, the impact of PEG size on PEGylation efficiency may be influenced by the conjugation mechanism, and this possible relationship should be considered when optimizing PEGylation location. We also observed that while the fraction of unmodified N53AzF/Ins163AzF following SPAAC reaction is on par with that of N53AzF and Ins163AzF, the efficiency of dual PEGylation is lower than would be expected based on the conjugation efficiencies of each site individually (expected: 62%, 63%). These results suggest that steric hindrance from the first PEG chain may inhibit conjugation with the second PEG chain, despite the relatively large distance between the sites, and is an important factor when considering the advantages of multi-PEGylation verses mono-PEGylation with longer PEG. This phenomenon is expected based on observations for dual-PEGylation using natural amino acids [129].

Using our screen, we also found that higher SASA does not necessarily correlate with higher conjugation efficiency (Figure 4-2C). Using Chimera [133], we also evaluated the local surface hydrophobicity of these sites, a trait which has been suggested to improve SPAAC efficiencies via interaction with the hydrophobic DBCO group [108,134]. Again, we found no clear correlation (Figure 4-4), suggesting that neither SASA nor local surface hydrophobicity are sufficient predictors

of high-efficiency conjugation sites. The discrepancies between common design guidelines and our results as described above further demonstrate the need to rapidly screen multiple PEGylation sites in parallel. Importantly, the small-scale expression reactions and low concentration conjugation reactions described above mitigate screening costs by reducing the amount of product needed.

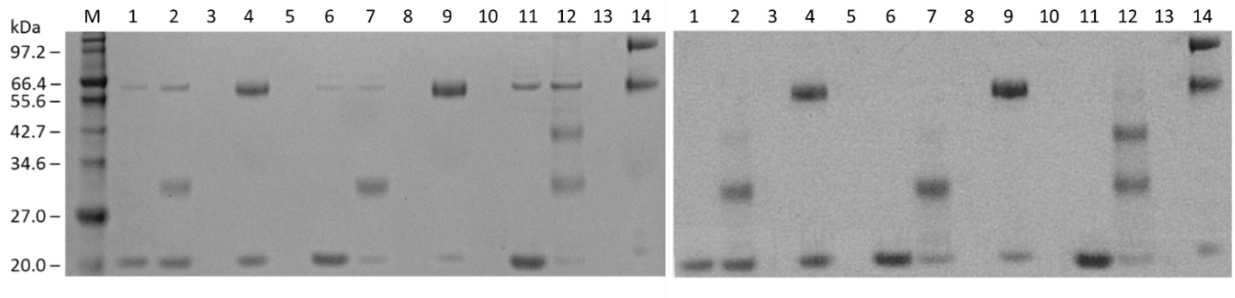


Figure 4-3: Sample SDS-PAGE and Autoradiogram of SPAAC Reactions for Click Efficiency analysis. (Left) 10% Bis-Tris gel of SPAAC samples (6 µL each) run against Unstained Protein Marker, Broad Range (2-212 kDa) (P7702, NEB). Lanes: 1-K135, 2-K135 + DBCO-PEG5kDa, 4 - K135 + DBCO-PEG20kDa, 6 - Ins163AzF, 7 – Ins163AzF+DBCO-PEG5kDa, 9 – Ins163AzF + DBCO-PEG20kDa, 11 – N53AzF/ Ins163AzF, 12 - N53AzF/ Ins163AzF + DBCO-PEG5kDa, 14 - N53AzF/ Ins163AzF + DBCO-PEG20kDa; (Right) autoradiogram of gel shown on the left, exposed for 4 days.

4.3.1.3 Stability screening

A primary motivation for PEGylating proteins is to improve both thermal stability and protease resistance [87,89,135]. Hence, stability evaluation is a key step in the conjugate screening process. Stability against protease degradation has previously been shown to correlate with protein conformational stability [87]. Thus, in our screening approach we use changes in protein melting temperature to characterize the stability of the screened proteins and corresponding conjugates. Thermal shift assays provide a higher-throughput, more cost-effective method for assessing protein melting temperature when compared to more traditional methods such as differential scanning calorimetry by allowing rapid, accurate T_m characterization with small samples and low-cost reagents. Here, we use a protein thermal shift assay to evaluate stability of our screening pool

using only 1-1.5 μg protein from each unpurified SPAAC reaction (3 replicates with ~0.3-0.5 μg of protein per replicate) in under 5 hours at less than $0.30 per screened protein or conjugate (~ $ 0.09 per replicate). The Protein Thermal Shift assay requires three orders of magnitude less protein at an order of magnitude lower protein concentration than traditional differential scanning calorimetry [136].

Using the Protein Thermal Shift assay, control T_m values for WT, WT+PEG5kDa, and WT+PEG20kDa were determined to be 61.0°C ± 0.10, 61.1°C ± 0.81, and 60.7°C ± 0.39, respectively, at a pH of 7.4, which agrees well with values previously reported for WT T4 lysozyme in literature [137-141]. The change in T_m due to incorporation of AzF or PEGylation at the incorporated AzF residue(s) relative to the corresponding control WT T_m is shown as ΔT_m in Figure 4-5A.

For all seven lysozyme variants, the magnitude of ΔT_m associated with AzF incorporation was greater than the magnitude of the change in ΔT_m as a result of PEGylation (Figure 4-5A, blue bars vs the difference between blue bars and either orange striped or white bars. This change in ΔT_m between the conjugated and unconjugated T4 Lyz variants will be referred to as $\Delta T_{m,conj}$, where $\Delta T_{m,conj} = \Delta T_{m,PEGylated} - \Delta T_{m,unconjugated}$). The observed impact on stability of even a single amino acid change demonstrates the challenge of engineering conjugation sites into proteins, as even minimal mutation can cause significant impact on protein stability. This further highlights the utility of the presented CFPS-based conjugate screening approach, as cell-free systems have been used to corporate a variety of uAAs [14], providing an opportunity to minimize the impact on stability through optimization of the choice of uAA.

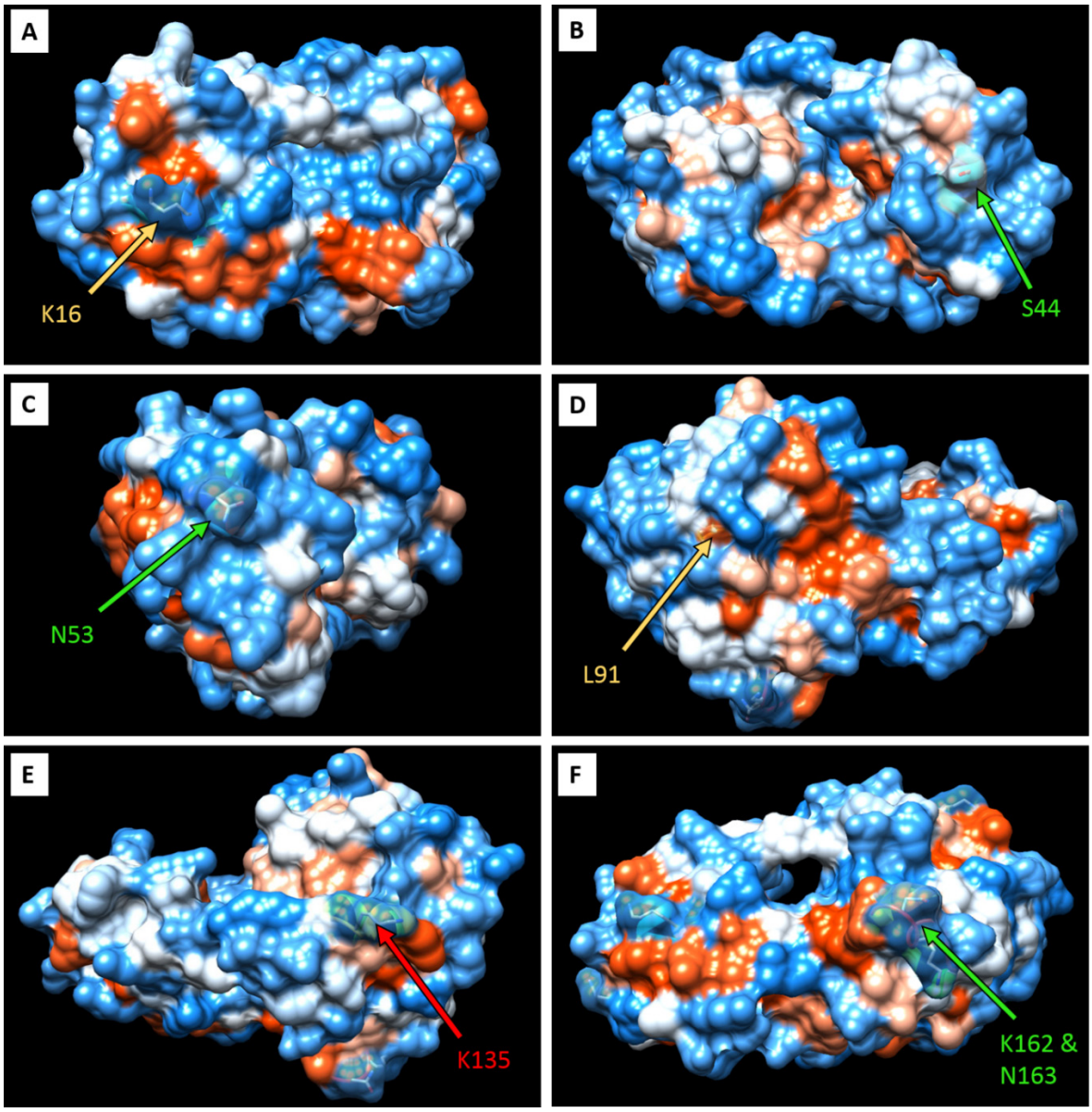


Figure 4-4: Hydrophobicity Surfaces for T4 Lyz. Blue areas are hydrophilic, white areas are neutral, and orange areas are hydrophobic. Surfaces of residues targeted are transparent, showing the side chains of the native residues underneath. Green labels/arrows indicate a conjugation site that conjugated very efficiently, yellow labels/arrows indicate a conjugation site that conjugated efficiently, and red labels/arrows indicate a site which conjugated inefficiently. (A) surface in vicinity of residue K16; (B) surface in vicinity of residue S44; (C) surface in vicinity of N53; (D) surface in vicinity of residue L91; (E) surface in vicinity of residue K135; (F) surfaces in vicinity of residues K162 and N163. Generated using Chimera [133] (PDB ID 2LZM).

PEGylation had the greatest stabilizing impact on K16AzF and Ins163AzF, but modestly improved the stability at all mono-AzF T4 Lyz variants with the exception of L91AzF. Slight

improvements in stability following double PEGylation of N53AzF/Ins163AzF were not significant, indicating that in this case stabilizing benefits did not compound as previously reported [87]. The varying impact of PEGylation on T_m for the different T4 Lyz variants was also not well predicted by SASA (Figure C-1).

In all cases, increasing PEG size from 5kDa to 20kDa did not affect the protein stability (Figure 4-5A), which agrees with many previous reports [100-102]. Studies which have found increased conjugate thermal stability with increasing PEG size have focused on smaller PEG oligomers [87,142], which could explain the difference to the present study (see Appendix C section C.2 for a more detailed discussion).

Our data also indicate that unstructured sites are not necessarily preferable to structured sites, either for uAA substitution or for PEGylation. Indeed, the site which was most tolerant of the AzF substitution was S44, which is located inside an alpha helix. By contrast, the least tolerant site, L91, is in an unstructured loop just outside an alpha helix. Similarly, the K16AzF variant was most stabilized by PEGylation, despite its location in a beta sheet. In contrast, PEGylation at an alpha helix (S44AzF) or at less structured sites was less stabilizing.

The dramatic destabilization from PEGylation at site L91 is interesting, given that this site was determined to be the optimal site for lysozyme immobilization [109,119]. These results indicate that the effects on protein stability from conjugation to a polymer differ significantly from the effects due to conjugation to a surface. A more detailed discussion of these differences is contained in Appendix C section C.2.

Overall, our results highlight the need for screening approaches such as the one presented herein due to the lack of predictive factors for PEGylation-based stabilization. The CFPS-based

screening approach provides a method for rapid, cost-effective screening of the effect of PEGylation at a variety of sites on protein thermal stability.

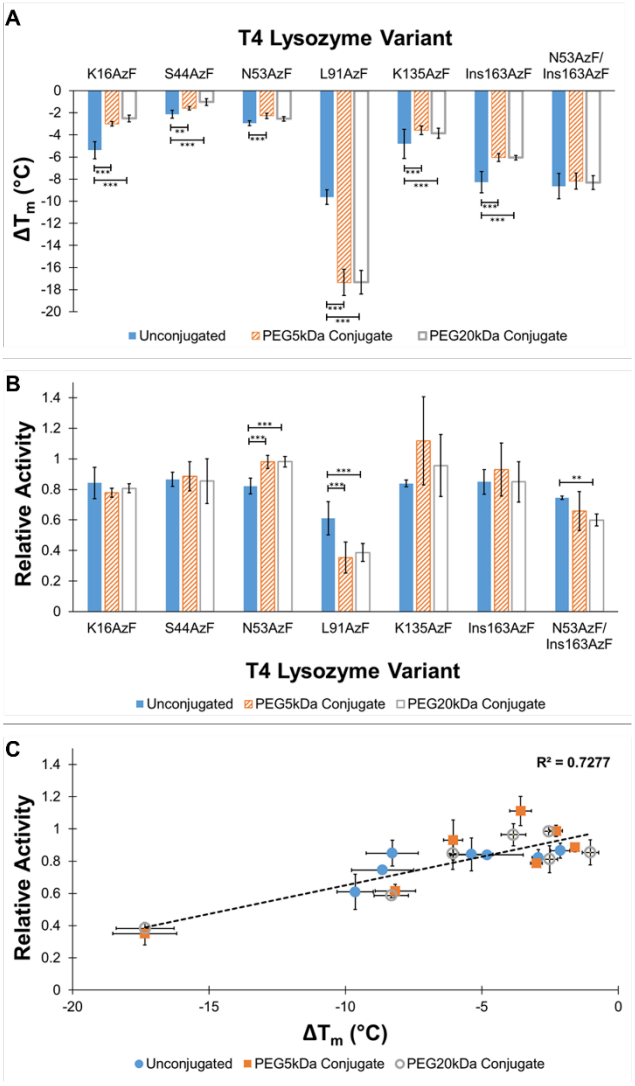


Figure 4-5: T4 Lysozyme Stability and Activity Before and After PEGylation. (A) Thermal stability of T4 Lyz variants and conjugates. ΔTm is determined as Tm – Tm,control, where Tm, control is for WT T4 Lyz; (B) Relative activity of T4 Lyz variants and conjugates; (C) Relative Activity vs ΔTm. Data from A and B plotted against each other. For A and B, data is shown as averages with one standard deviation, where n n ≥ 3, and stars indicate significant differences, where *** = $p < 0.001$ and ** = $p < 0.01$.

4.3.1.4 Activity screening

Retained protein activity is another essential metric in the design of PEGylated protein. Fortunately, a reduction in specific activity following PEGylation can be compensated for, in many

cases, by a corresponding increase in stability [79]. However, minimizing the negative effects of PEGylation on activity would require less PEGylated protein in its final application, reducing costs and negative side effects of therapeutics. As PEG size and number have been previously reported to impact the activity of PEGylated proteins, we applied our screen to evaluate these effects on T4 Lyz activity [85,89,103,104]. The activity of each T4 Lyz variant and its PEG5kDa and PEG20kDa conjugates was determined using the Enzchek Lysozyme Assay (ThermoFisher Scientific) as reported in Figure 4-5B.

The activities of unconjugated K16AzF, S44AzF, N53AzF, K135AzF, and Ins163AzF were not statistically different ($p > 0.05$). However, all are greater than the activities of L91AzF and N53AzF/Ins163AzF. These results suggest that, like stability, activity of AzF substituted variants are not well predicted by modification site secondary structure. The reduction in the activity of the unconjugated N53AzF/Ins163AzF was less than the sum of the reductions from the two individual AzF incorporations (N53AzF and Ins163AzF), suggesting that the activity losses in multi-uAA incorporations should not be considered additive. L91AzF had the lowest activity of any unconjugated T4 Lyz variant, including the double amber suppression variant, at only 61% WT activity.

Of the six mono-PEGylated lysozyme variants tested, only PEGylation of N53AzF clearly increased the activity. While unexpected, improved activity following PEGylation is not without precedent [83,143]. PEGylation at both the structured sites K16AzF and S44AzF and at the unstructured sites K135AzF and Ins163AzF did not significantly change the activity compared to the unconjugated variants. This suggests that the effect of PEGylation on activity is not well predicting by conjugation site secondary structure.

In stark contrast to the other five mono-PEGylated variants, PEGylation of L91AzF resulted in an additional loss of activity, with conjugates retaining only about 60% of the activity of unconjugated L91AzF. This decrease in activity correlates with the significant decrease in stability seen as a result of PEGylation of L91AzF. There may be a general correlation between high activity retention and SASA, given the reduction in activity following PEGylation of L91AzF, minimal impact of PEGylation on activity for K16AzF and S44AzF, and improvement in activity following PEGylation of N53AzF. However because only PEGylation of L91AzF and N53AzF significantly changed the activity, the utility of the trend in predicting sites with the highest activity retention is uncertain (Figure C-2). Still, SASA could be a useful guideline in narrowing potential PEGylation sites to exclude sites with very low SASA.

In addition, activity of diPEGylated N53AzF/Ins163AzF decreased with 20kDa PEG. As neither PEGylation of N53AzF nor Ins163AzF with 20kDa PEG had a negative effect on activity, and because PEGylation of the double amber suppression variant had no significant effect on the stability of the variant, this decrease in activity can reasonably be attributed to reduced accessibility of the active site to the macromolecular substrate (Micrococcus lysodeikticus cell walls) due to steric hindrance from the two large PEG chains.

In all cases, there was no significant difference between the activities of the 5kDa PEG conjugates vs the 20kDa PEG conjugates. PEG size has previously been reported to be negatively correlated with in vitro activity [85,89,103,144], but other studies have also shown that in some cases activity is independent of PEG size [145,146]. For example, another study in which IFN was PEGylated with 5kDa, 10kDa, and 20kDa PEG at a disulfide bridge and the activity impact was found to be size independent [146].

Literature also reports that point mutations in a protein have minimal, localized impact on the structure [140], so, given their distance from the active site, it is less likely that any of these mutations impact activity by deforming the active site. Thus, it is reasonable that any changes in activity would be related to changes in the dynamics of the protein. In analyzing the data in this study, there indeed appears to be a correlation between T4 Lyz analog stability and activity, as shown in Figure 4-5C. This relationship between stability and enzymatic activity could be used as a screening tool to eliminate variants which are likely less active based on their stability. This would reduce costs of PEGylation screening, as activity assays are more expensive than the stability assay. It is important to note that more data is necessary to confirm the existence of this trend for proteins beyond T4 Lyz, however, the presented screen could facilitate verification of this trend for other proteins. Verification of this trend would be especially important for proteins with more than 2 folding states, as destabilization of certain domains may have a smaller impact on activity than others. However, our study indicates that stability could be a useful metric by which to eliminate less active candidates in a large screen without directly testing the activity of all candidates.

4.3.2 Evaluating common PEGylation guidelines with experimental screen results

In the decades since the introduction of PEGylation, several recommendations have evolved towards the design of PEGylated proteins. A synopsis of key similarities and differences between the effects of PEGylation observed experimentally through the present screen and some common PEGylation guidelines in literature is summarized in Table 2-1.

Our observations regarding PEG size support the claim that, in general, large PEGs are more advantageous than small PEGs for conjugate optimization, as we found no significant negative impacts of PEG size on the stability or activity of T4 Lyz which might offset the improved

pharmacokinetics generally reported for conjugates with larger PEG chains [129,147-149]. In our evaluation of *in vitro* assays of a double site-specifically PEGylated conjugate, we found no motivation for double site-specific PEGylation due to the lack of additional stabilizing benefit and the negative effects on activity. However, as it is well documented that increased PEG weight improves pharmacokinetics of a therapeutic, these results suggest that there may be an optimal trade-off between specific activity and half-life which could potentially be achieved through double site-specific PEGylation. Finally, we evaluated trends in click efficiency, activity, and stability relating to the uAA incorporation/PEGylation site, finding that current design metrics were insufficiently predictive of the effects of PEGylation observed through our screen.

Overall, the experimental screen enabled the identification of S44, followed closely by N53, as the best sites for PEGylation of the sites screened due to their high PEGylation efficiency, high thermal stability, and high retention of activity. It should be noted that PEGylation had the greatest stabilizing effect on K16AzF, K135AzF, and Ins163AzF, however these sites were less tolerant of the AzF incorporation and had lower conjugation efficiencies. The locational effects of site-specific PEGylation were not well predicted by the common design guidelines evaluated in the screen. While some guidelines, such as conjugation at sites with high SASA, may have merit in some aspects of PEGylation engineering, their predictive capacity is incomplete. For example, K135 may be correctly predicted by SASA to have high activity retention, however PEGylated K135AzF would be more costly to produce commercially than similarly active S44AzF due to its low PEGylation efficiency. Hence, the screening approach presented is an attractive tool to facilitate the efficient screening of PEGylation sites in order to optimize the multiple facets of PEGylated protein design.

Table 4-1: Comparison to PEGylation Design Recommendations in Literature.

PEGylation Design Facet		PEGylation Recommendations in Literature	Observations with CFPS Screen
PEG Size		- Larger PEGs offer improved pharmacokinetics and stability relative to smaller PEGs [83,89] - ΔT_m independent of PEG size [100-102] - *In vitro* activity is inversely proportional to attached PEG mass [85,89,103,104]	- No significant difference in stability or activity of 20kDa conjugates vs 5kDa conjugates
PEG Number		- Higher modification number correlates with lower activity [85] - PEGylation at two individually stabilizing sites results in further stability improvement [87] - Single large PEG is preferable to multiple small PEGs [89]	- Double PEGylation with 20kDa PEG decreased activity slightly - Double PEGylation at two individually stabilizing sites did not significantly improve stability
Conjugation Site	**Efficiency**	- High SASA (Residues with ASA-View score > 0.4) improves conjugation efficiency[g][105-108] - Buried residues in hydrophobic pockets may conjugate with high efficiency in SPAAC reactions [108]	- SASA was not predictive of conjugation efficiency - No clear correlation to surface hydrophilicity/hydrophobicity
	Stability/ Activity	- Conjugation at unstructured loops may minimize the strain on protein structure [82,109,110]	- No clear correlation between stabilizing potential and PEGylation site secondary structure or SASA. - General trend of increased activity retention after conjugation at higher SASA may help to narrow experimental screens by eliminating sites with very low SASA

4.3.3 Enhancing CFPS-based screen with coarse-grain simulation

Considering the necessity of an experimental screen, an *in silico* tool to narrow candidate PEGylation sites would be useful in further reducing the costs of PEGylated protein design. Protein stability is a driving factor in protein PEGylation and is highly dependent on PEGylation site, therefore we sought to integrate a rapid molecular dynamics simulation as part of the screening process. Specifically, the potential of coarse grain simulation was evaluated as a pre-screen tool to

inform selection of candidate conjugation sites, thereby reducing the size of the screening pool. Coarse-grain molecular dynamics simulations have previously been used to correctly identify an optimal site for T4 Lyz immobilization [109,119]. The method employs replica exchange molecular dynamics to calculate the heat capacity (C_v) of the molecule as a function of temperature. Peaks in the C_v curve occur during structural transitions (unfolding/folding), and changes in the thermal stability can be obtained by comparing the temperatures at which the peaks occur in the C_v profiles of WT T4 Lyz and PEGylated T4 Lyz (Figure C-3 and Figure C-4). The computational efficiency of coarse-grain, molecular simulation make it especially appealing as a screening tool because many simulations can be done quickly and in parallel.

We modified the coarse-grain simulation to include a conjugated PEG chain at a specified site. The methods used are those that have been outlined previously [119-122] with a few modifications to include PEG, as detailed in the Materials and Methods. While the mechanism behind PEG-based stabilization is not well understood, there are two primary theories: PEG stabilizes proteins through (1) direct interaction with the protein surface [142,150,151] or (2) entropic interactions with the solvent [87,150]. In this study we limited the simulation to evaluate only the entropic effect. The entropy-only approach has been successfully applied to immobilized proteins [120,152], and it has been previously reported that the stabilizing effect of PEG is primarily entropic [87]. The polymer-protein interactions were thus purely repulsive. The tethering of PEG to the protein was done with a simple harmonic restraint without changing the tethering residue to AzF. Currently AzF cannot be included in the model due to insufficient experimental data for parametrization of AzF in the coarse-grain interactions. Therefore, predicted changes in melting temperature correspond to the change in melting temperature solely due to PEGylation ($\Delta T_{m,conj}$). The three-dimensional structure of T4 Lyz, needed for the model, was obtained from

PDB ID 2LZM and is expected to be a reasonable representation of the structures of AzF-substituted lysozyme as any structural perturbations resulting from such substitutions have been previously reported to be minimal and highly localized [140]. However, we excluded the variants involving an AzF insertion (Ins163AzF and N53AzF/Ins163AzF) from the simulation due to concerns that the WT structure may be a less accurate representation of the structure for these variants.

We compared the simulation-predicted $\Delta T_{m,conj}$ to the experimentally determined values in our screen, with the results shown in Figure 4. The predictions from the entropy-based PEGylated T4 Lyz simulations agree well with the results from experiment. While the simulation did not predict the exact change in melting temperature due to PEGylation, it accurately predicted the relative change in stability after PEGylation. For both the PEG5kDa and PEG20kDa, the simulation accurately predicted the most stabilizing site for conjugation to be residue 16 and the least stabilizing site to be residue 91. The simulation also accurately predicted the remaining 3 sites to be slightly less stabilizing than conjugation at residue 16 but significantly more stabilizing than conjugation at site 91. As such, the entropy-only simulation could be a powerful tool for informing design of a screening pool. Adapting the simulation to include enthalpic effects of PEG-protein interactions may further improve the predictive capacity of the simulation. Indeed, previous results have suggested that the enthalpic interaction effects of PEGylation can be destabilizing in some cases [87], which could account for the over-prediction of $\Delta T_{m,conj}$ by the entropy-only model. Interestingly, the simulation seems to provide better quantitative estimates of $\Delta T_{m,conj}$ for the 20kDa conjugates. We hypothesize that this is due to a more dominant impact of entropy with the larger polymer. Hence, incorporation of PEG-protein interaction parameters may also allow the model to predict the relative impact of PEGylation with different lengths of PEG.

Overall, our use of coarse grain simulations indicates that valuable insight can be gained about comparative effects of PEGylation from a simple, entropy-based model. This model could be incorporated as part of the screening process to create a hybrid simulation- and experimental-based screening method in which all of the sites on a target protein which are available to conjugation are first screened with the coarse grain model. The simulations in this study ran to completion in under 11 hours, potentially enabling simultaneous evaluation of all potential conjugation sites on a protein in less than 24 hours. Sites which are predicted by the simulation to be least stabilizing could then be eliminated from the subsequent experimental screen in order to conserve resources by reducing the number of candidate conjugation sites. Further improvements in the model, including parameterization of AzF and PEG-protein interactions, may enhance the capacity of the model for quantitative ΔT_m prediction.

4.4 Conclusions

Here, we have presented a screening system based on cell-free protein synthesis in order to rapidly assess the location-dependent impact of PEGylation on protein stability and function with microliter-scale reactions. Notably, the approach also enables rapid assessment of protein production yields after uAA incorporation as well as PEGylation efficiency, both of which also vary depending on the location. Thus, production yields, PEGylation efficiency, and post-PEGylation protein stability and activity can all be determined and compared relative to other sites in less than two days. The approach can be used with a wide variety of proteins, where the primary limitation is an inability to facilitate mammalian post-translational glycosylation [15,21,33,94,97-99]. In addition, we have developed a coarse-grain simulation tool for rapid *in silico* screening of protein conjugation sites and present a linear correlation between simulation and experimental results. Thus the coarse-grain simulation tools could further expedite the screening process by potentially narrowing the

experimental screening pool. Finally, we have compared our findings to currently held PEGylation guidelines and highlight the need for technology such as that presented to rapidly evaluate protein PEGylation sites, and also provide more data to help develop more accurate and refined PEGylation guidelines.

5 PRODUCTION OF THERAPEUTIC PROTEINS IN CFPS FOR PEOZYLATION OPTIMIZATION

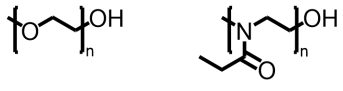
To address concerns about potential generation of anti-PEG antibodies, researchers have begun to look beyond PEG for biocompatible polymers for pharmacokinetic stabilization of protein therapeutics. Here we present preliminary work towards enabling PEOzylation optimization of multiple therapeutics in the cell-free system. The presented work was designed, led, and conducted by myself, with contributions from my graduate student colleague Emily Long Zhao.

5.1 Introduction

Acute lymphocytic leukemia (ALL) is the most common childhood cancer in the U.S. Over the past several decades, survival rates have increased from about 10% to close to 70% or more, primarily due to the discovery of bacterial asparaginases as effective treatments [153]. These enzymes deplete blood asparagine levels, selectively targeting leukemia cells which are unable to synthesize asparagine as healthy cells can. However, the bacterial origin of these asparaginases make them highly immunogenic, and short half-lives of 1.5 days or less necessitate multiple administrations weekly. As a result, between 20-40% of patients develop antibodies against the primary therapeutic, *E. coli*-derived asparaginase (aspargase), within the first few weeks of treatment, necessitating treatment with a second asparaginase derived from *Erwinia chrysanthemi* bacteria (crisantaspase) [53]. However, because crisantaspase is also highly immunogenic,

incidence of hypersensitivity or neutralizing antibodies is about 33% [53]. A high rate of relapse amongst ALL patients exacerbates the immunogenicity problem, as treating these relapsed patients with the same asparaginase they were treated with previously can cause a "vaccination effect," where the administered asparaginase is met with increased hypersensitivity and rapid clearance [53]. PEGylation of both aspargase and crisantaspase has improved treatment options by reducing their immunogenicity and dramatically increasing the half-life [53,154]. These PEGylated asparaginases, however, suffer from two key limitations: (1) PEG-associated toxicities that develop with frequent administration, such as kidney vacuole formation and anti-PEG antibodies [155-157], and (2) reduced thermal stability and specific activity [53].

The incidence of anti-PEG antibodies, however, could negatively effect the efficacy of asparaginase drugs by increaseing their immunogenicty and decreasing their half-life. A promising alternative to PEGylation is conjugation with another highly water soluble, biocompatible polymer with better renal clearance: poly(2-ethyl 2-oxazoline) (PEO) (Figure 5-1). Recent work with PEOzylation (conjugating PEO to a protein) has demonstrated a similar potential for "stealth" behavior to reduce immunogenicity of conjugated protein therapeutics, and it has been approved for consumption by the FDA [155,156]. Recent efforts to produce a PEOzylated protein therapeutic have shown high activity retention, but have been plagued by low conjugation efficiency and have been limited to native unique sites such as the N-terminus [156]. Therefore, asparaginase therapies could be improved by enabling higher conjugation efficiency, activity, and stability through conjugation to a PEO as a PEG alternative.



PEG **PEO**

Figure 5-1: PEG vs PEO structure.

Careful design of such conjugates allows improvements over the traditional benefits associated with PEGylation, namely increased half-life and reduced immunogenicity, by enhancing conjugate stability and activity through selection of an optimal conjugation site. Unnatural amino acids (uAA) provide the precise control necessary for conjugate optimization due to unique side chains which can be targeted in orthogonal conjugation reactions [20,24]. However, inadequate understanding of the effects of polymer conjugation on proteins necessitates testing many different conjugation sites, which, with traditional *in vivo* expression, is expensive and onerous [20]. Chapter 4 discusses the development of an integrated coarse-grain simulation and cell-free protein synthesis based approach to rapidly determine optimal PEGylation sites [20]. Here, work is presented towards reengineering this platform for optimization of protein therapeutics. Production of highly active crisantaspase is optimized. The proposed work would provide the basis towards the development of an optimized ALL therapeutic and could provide a framework for development of optimal PEOzylated therapeutics in the future.

5.2 Materials and methods

Materials were purchased from Sigma Aldrich unless otherwise indicated.

5.2.1 Extract preparation

Extracts were made as described previously in sections 2.2.1 and 4.2.1.

5.2.2 Cell-free protein synthesis

CFPS reaction mixtures were based on of the same procedure described in previous sections, with crisantaspase-specific modifications made during optimization. These modifications are discussed in section 5.3.

5.2.3 Protein purification

Ni-NTA spin purification was performed according to the manufacturer's instructions. For spin cation exchange chromatography, samples were diluted with buffer (buffers were based on either 25 mM sodium acetate buffer or 50 mM HEPES buffer, with NaCl added to wash and elution buffers at concentrations ranging up to 2 M) to a final volume of 400 μL and then applied to Pierce Strong Cation Exchange Spin Columns (ThermoFisher Scientific) and purified according to the manufacturer's instructions.

5.2.4 SDS-PAGE gels and autoradiography

SDS-PAGE gels and corresponding autoradiograms were performed as described in section 4.2.6.

5.2.5 Activity assays

Activity assays of crisantaspase were performed as described in section 2.2.5.

5.2.6 Stability assays

Stability assays to determine T_m were performed as described in section 4.2.7.1.

5.3 Results and discussion

The first step in developing and evaluated PEOzylated therapeutics in the high-throughput CFPS system was to re-engineering the CFPS high-throughput screening system for production and purification of these more complicated proteins. Crisantapase, a tetrameric protein presents unique challenges for high-throughput optimization, including modified CFPS conditions and restrictive purification options.

5.3.1 Crisantaspase production optimization

Crisantaspase (ErA) was initially produced and purified as described in section 2.2.3, however, it was noted that the purified ErA was only minimally active (~0.1% of reported activity). We hypothesized that this was a result of steric hindrance due to the C-terminal 6-histidine tag, as the C-terminus is located at the interface between monomers and is also implicated in enzyme activity (Figure 5-2). We therefore performed QC mutagenesis to remove the 6-histidine tag and preceding glycine/serine linker from the C-terminus of the enzyme. Activity analysis of the un-tagged ErA variant confirmed restoration of activity to levels in good agreement with literature values (Figure 5-2) [42,161-163]. In connection with the dramatically increased activity, protein synthesis yields were notably lower. We hypothesized that the decrease in yields represented an asparagine-limited synthesis reaction for the highly active untagged ErA. Therefore, we assessed the effect of adding an initial bolus or continuous additions of asparagine in an otherwise standard CFPS reaction. We found that yields could be nearly doubled by addition of an initial bolus of 15 mM asparaginase (Figure 5-2). Highly active crisantaspase has been produced in a CFPS system, enabling rapid protein synthesis towards high-throughput PEOzylation optimization.

The next step in the optimization of crisantaspase production is to identify a rapid, reliable tag-less purification strategy. This would improve the applicability of the screening process by the production of minimally-modified therapeutics and would be especially advantageous for proteins which, like crisantaspase, are negatively impacted by C-terminal affinity tags. A more widely applicable procedure would be ion exchange chromatography. We evaluated the feasibility of purifying crisantaspase using spin cation exchange columns. Example SDS-PAGE results are shown in Figure 5-3 with two different buffer systems. An apparent disadvantage of this purification method is the notably less pure protein fractions. However, using 25 mM sodium

acetate buffer, pH. 5.5, enables high recovery of ErA with minimal contaminating protein. Although the pI is expected to change slightly with each different AzF substitution, because the pH of this buffer is sufficiently lower than the expected pI of ErA, AzF-incorporated ErA should be amenable to purification under the same conditions [164].

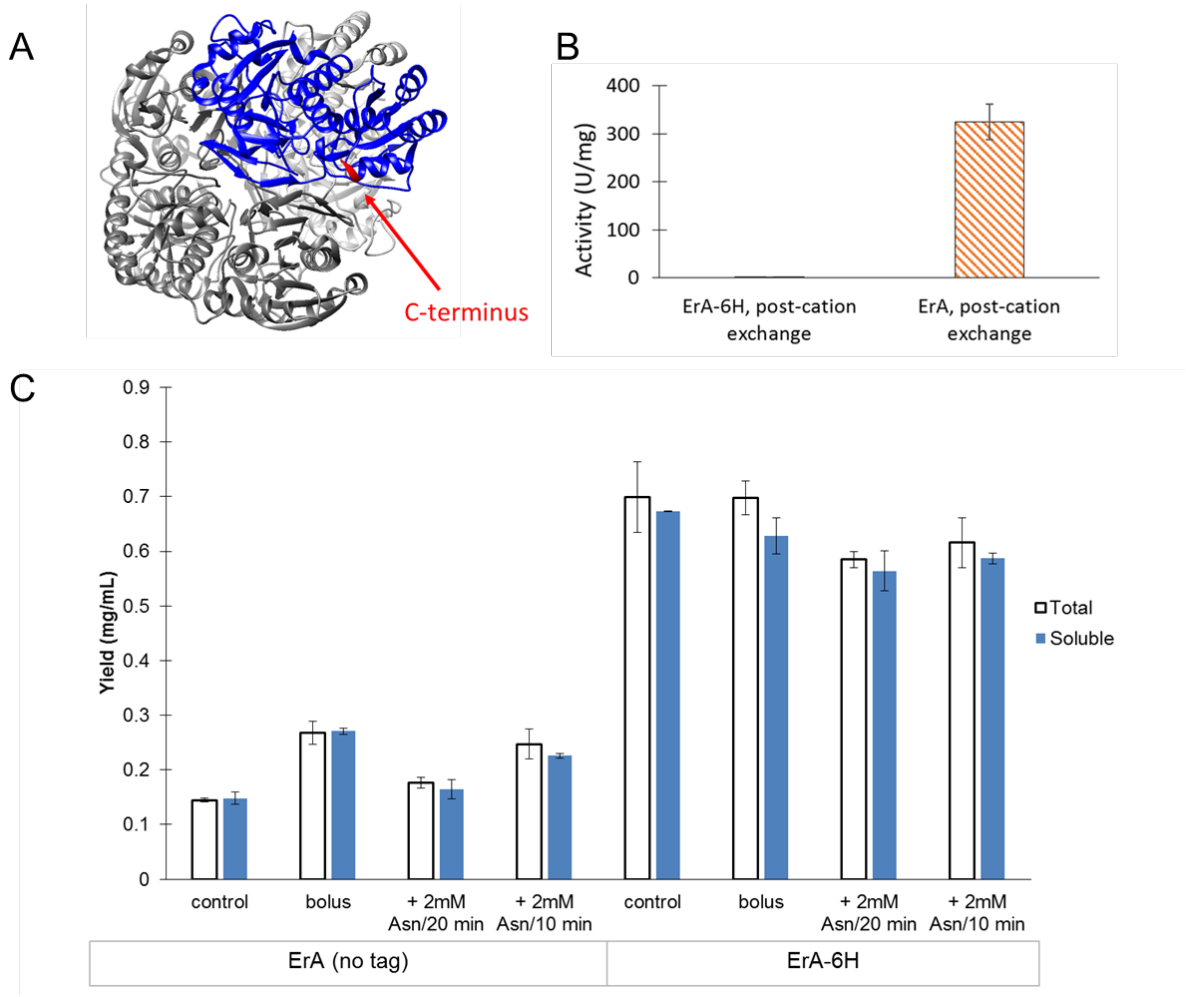


Figure 5-2: Active ErA Production Optimization. (A) Crystal structure of tetrameric ErA showing the orientation of the C-terminus (red) of one monomer (blue) at the interface between monomers; (B) Comparison of ErA activity for His-tagged and un-tagged ErA (averages with one standard deviation, $n \geq 4$); (C) CFPS optimization by asparagine supplementation (averages with one standard deviation, $n \geq 2$.

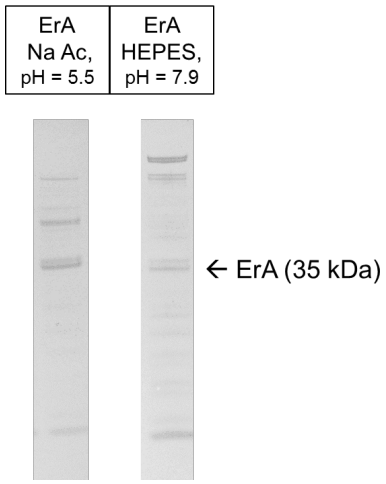


Figure 5-3: SDS-PAGE Analysis of Spin Cation Exchange-purified ErA Using Different Buffers.

5.3.2 uAA Incorporation and PEOzylation

5.3.2.1 uAA incorporation site selection

In Chapter 4, a CFPS-based high-throughput screening system and a coarse grain simulation were developed to evaluate current heuristics informing design of optimally PEGylated proteins. Current heuristics were found to be insufficiently predictive, necessitating a high-throughput screen for optimization of novel conjugates [20]. Therefore, sites for PEOzylation were selected to span a variety of sites, including locations in α helices, β sheets, and unstructured loops, moderate to high solvent accessibility, and epitope score. Structural context was evaluated using Chimera and the published protein structure for crisantaspase (PDB ID 1O7J) [133]. Solvent accessible surface area was assessed via the web-based program ASA-View, as done previously for T4 Lysozyme [20]. Because a primary concern for the *in vivo* stability of crisantaspase is its high immunogenicity, immunogenicity score was evaluated based on the B cell mapping functions of the Immune Epitope Database [167-169], which predicts 6-7 epitopes in the crisantaspase sequence, shown in Figure 2-1A. Amino acid substitutions in predicted epitope sites has been

demonstrated to reduce the immunogenic potential of some protein therapeutics by reducing their ability to bind MHC II peptide grooves [170,171]. Therefore, ncAA substitution and subsequent PEOzylation at sites in these predicted epitopes, such as the predicted immunodominant epitope between residues 283-293 [170], could simultaneously reduce the binding potential of the peptide to MHC II and also provide protection against degradation by other immune cells or proteases. Considering each of these attributes – structural context, solvent accessible surface area, and immunogenicity, 30 representative sites were selected for uAA screening. We propose to test the effect of AzF incorporation at each of these sites using the previously presented high-throughput conjugate screening platform with the CFPS modifications detailed in the previous section. This work will be continued by Ph.D. candidate Emily Long Zhao.

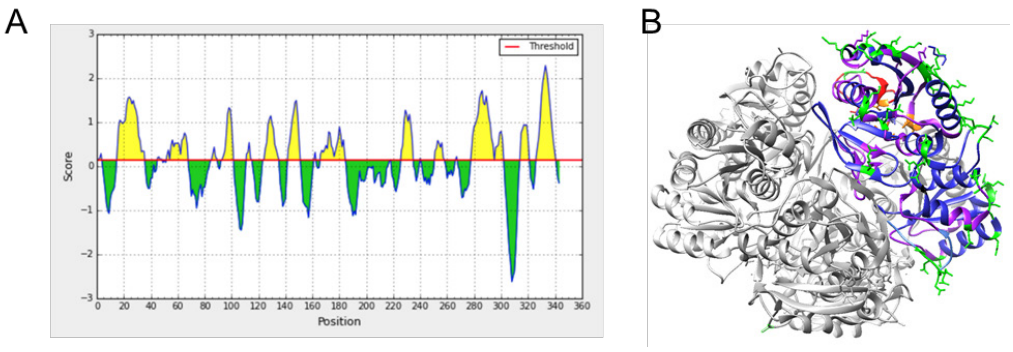


Figure 5-4: Site Selection for Crisantaspase. (A) Epitope mapping of crisantaspase using IEDB [167-169]; (B) Selected site locations on crisantaspase crystal structure (PDB ID 1O7J). The colored chain represents one monomer subunit of the tetrameric crisantaspase. Purple regions correspond to immunogenic regions identified in (A), orange and red regions are indicated in enzyme activity, and green residues indicate sites selected for PEOzylation screening.

5.3.2.2 PEOzylation scheme

Due to recent reports of possible anti-PEG antibodies, development of alternative polymers is an active area of research [156]. One such alternative polymer is poly(2-ethyl-2-oxazoline) (PEOz), which has been reported to have similar stealth properties to PEG with potentially further

decreased immunogeneicity [155,156]. PEOz was recently conjugated to two sites in G-CSF and the resulting conjugate maintained high activity [156]. We therefore propose to evaluate the location-dependent effects of PEOzylation vs PEGylation and enable rapid optimization of PEOzylated therapeutics via the adapted CFPS-based screen herein described. However, DBCO-functionalized PEOz for copper-free click chemistry is not commercially available, so we have identified a reaction scheme to functionalize PEOz with DBCO with a click reaction in order to minimize reactivity of unreacted linker or PEOz with protein (Figure 5-5). We have procured a 5kDa amine-functionalized PEOz polymer from Sigma Aldrich and a corresponding DBCO-NHS linker (Click Chemistry Tools). The DBCO-NHS linker will react with the amine group on the PEOz polymer, however, it is insoluble in water. Therefore, the reaction may be carried out in organic solvent such as methlylene chloride, which can be easily evaporated following the reaction. Water can then be added to the dried DBCO functionalization reaction to solubilize the DBCO-functionalized PEOz and unreacted PEOz, while unreacted DBCO/NHS linker remains insoluble. The amine-terminated PEOz is not expected to react with any amino acid side chains at physiological conditions, thereby insuring that PEOzylation remains site-specific.

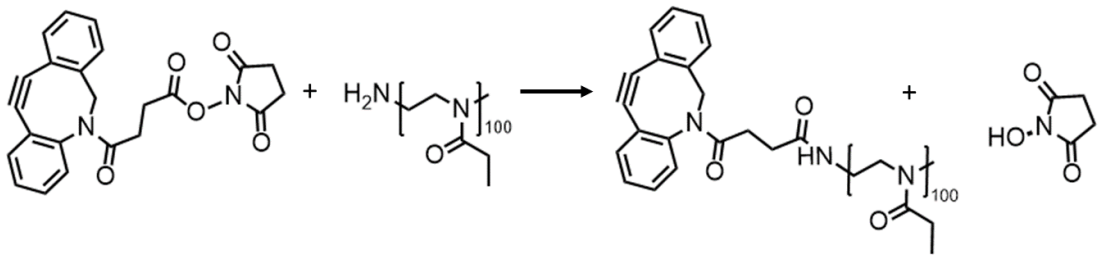


Figure 5-5: DBCO Functionalization of PEOz.

5.3.3 Stability evaluation

5.3.3.1 Crisantaspase stability

Stability of wild-type crisantaspase was evaluated using a thermal shift assay as described in 4.2.7.1. Despite some contaminating proteins in the cation-exchange purified samples, wild-type melting temperatures agreed well with previously reported values, shown in Figure 5-6, where $T_{m,\,dsc}$ represents T_m as determined by differential scanning calorimetry [172] and $T_{m,\,act}$ is the temperature at which the enzyme reportedly loses 50% activity [173] which is understandably lower than the actual T_m of the protein. Discrepancies between the determined T_m and $T_{m,\,dsc}$ may be due to the high ionic strength of the cation exchange elution buffers. Notably, there was no significant difference in the T_m between ErA with a C-terminal 6-histidine tag (ErA-6H) and untagged ErA, supporting the hypothesis that the dramatic decrease in activity of ErA-6H is the result of steric hindrance from the 6H tag as opposed to significant structural perturbation of the monomer. We also noted that the stability of the crisantaspase monomer did not differ significantly between samples purified in sodium acetate buffer or HEPES buffer.

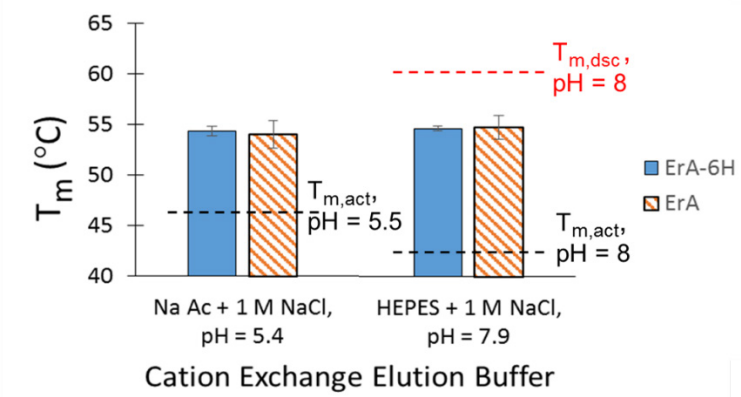


Figure 5-6: Stability of Wild-type Crisantaspase in Cation Exchange Elution Buffers.

5.3.4 Conclusions and future directions

We have demonstrated production of highly active crisantaspase in a CFPS system, as well as the highest reported yields of soluble G-CSF. Furthermore, we have adapted the purification schemes previously used in our high-throughput screening system to accommodate a wider variety of proteins and have demonstrated precise determination of T_m in these purified fractions. Finally, the framework has been established for evaluation of the locational impact of site-specific PEOzylation on two therapeutic proteins with implications in cancer treatment. This research will be further pursued by Ph.D. student Emily Long Zhao, who will begin the process of AzF incorporation and PEOzylation and will assess the reliability of linear expression templates (LETs) as a method for expedited template generation to further streamline PEOzylation evaluation. The ability to rapidly identify optimal therapeutically relevant polymer-protein conjugates could potentially expedite optimization of emerging protein therapeutics. Furthermore, the CFPS expression of these therapeutics is an important step towards on-demand production of protein therapeutics.

6 CONCLUSIONS AND FUTURE WORK

In this work, we have made several important advancements towards re-engineering cell-free protein synthesis to improve design and production of future protein therapeutics. We have improved the platform to be more amenable to distributed, on-demand protein expression by eliminating endotoxins from the system, enhancing extract storage stability, and demonstrating lyophilization of a one-pot whole cell-free protein synthesis system in a scalable format. We also designed a CFPS-based screening platform coupled with a coarse-grain simulation and used this system to evaluate the locational-dependent effects of site-specific PEGylation. Finally, we have laid the groundwork for polymer-protein conjugate optimization of therapeutic proteins in the CFPS by enabling rapid production of high yields of therapeutic proteins, including a multimeric protein and a disulfide bonded protein.

This work continues through the efforts of multiple students as shown below.

- **Chapter 2:** Further work to extend endotoxin-free cell-free expression systems is being continued by Ph.D. candidate Emily Long Zhao. We are currently extending this system to a lyophilized, shelf-stable format.
- **Chapter 3:** Further work to extend cryoprotection to lyophilized whole cell-free systems is being continued by Ph.D. candidate Emily Long Zhao.

- **Chapter 4:** Work towards improving PEGylation optimization is being continued by students in both Dr. Bundy's and Dr. Knotts's labs. Ph.D. student Emily Long Zhao will address the challenge of experimental PEGylation optimization of therapeutic proteins using therapeutics such as Onconase, crisantaspase, or G-CSF. To further the simulation work, Ph.D. student Addison Smith and undergraduate student Joshua Wilkerson are working towards parameterization of uAA such as AzF in the coarse-grain model. Joshua Wilkerson is also working to continue high-throughput synthesis and stability analysis of AzF-incorporated T4 lysozyme in order to generate experimental data to validate the coarse-grain parameters under development.
- **Chapter 5:** Continued work towards PEOzylation will be conducted by Emily Long Zhao, who will address the challenge of high-throughput uAA-incorporation of additional therapeutic proteins, such as Onconase, GM-CSF, or G-CSF and subsequent conjugation to PEOz.

Overall, the contributions of this dissertation have furthered the ability to produce next-generation therapeutics from a cost-effective on-demand *E. coli*-based cell-free protein synthesis system. My hope is this will lead to greater access to improved therapeutics.

REFERENCES

1. Usmani SS, Bedi G, Samuel JS, Singh S, Kalra S, Kumar P, Ahuja AA, Sharma M, Gautam A, Raghava GPS: **THPdb: Database of FDA-approved peptide and protein therapeutics**. *PLoS ONE* 2017, **12**:e0181748.

2. Philippidis A: **The Top 15 Best-Selling Drugs of 2017**. Genetic Engineering and Biotechnology News; 2018.

3. Cong Y, Pawlisz E, Bryant P, Balan S, Laurine E, Tommasi R, Singh R, Dubey S, Peciak K, Bird M, et al.: **Site-specific PEGylation at histidine tags**. *Bioconjug Chem* 2012, **23**:248-263.

4. Schumacher D, Hackenberger CPR: **More than add-on: chemoselective reactions for the synthesis of functional peptides and proteins**. *Current Opinion in Chemical Biology* 2014, **22**:62-69.

5. Pelegri-O'Day EM, Lin E-W, Maynard HD: **Therapeutic Protein–Polymer Conjugates: Advancing Beyond PEGylation**. *Journal of the American Chemical Society* 2014, **136**:14323-14332.

6. Maiser B, Dismer F, Hubbuch J: **Optimization of Random PEGylation Reactions by Means of High Throughput Screening**. *Biotechnology and Bioengineering* 2014, **111**:104-114.

7. Sathish J, Sethu S, Bielsky MC, de Haan L, French N, Govindappa K, Green J, Griffiths C, Holgate S, Jones D, et al.: **- Challenges and approaches for the development of safer immunomodulatory biologics Y1 - 2013**. - *Nature Reviews Drug Discovery*:- 306.

8. Ogonah OW, Polizzi KM, Bracewell DG: **Cell free protein synthesis: a viable option for stratified medicines manufacturing?** *Current Opinion in Chemical Engineering* 2017, **18**:77-83.

9. Pardee K, Slomovic S, Nguyen PQ, Lee JW, Donghia N, Burrill D, Ferrante T, McSorley FR, Furuta Y, Vernet A: **Portable, on-demand biomolecular manufacturing**. *Cell* 2016, **167**:248-259.

10. Swartz JR: **Transforming biochemical engineering with cell-free biology**. *AIChE Journal* 2012, **58**:5-13.

11. Kim T-W, Keum J-W, Oh I-S, Choi C-Y, Park C-G, Kim D-M: **Simple procedures for the construction of a robust and cost-effective cell-free protein synthesis system**. *Journal of biotechnology* 2006, **126**:554-561.

12. Wilding KM, Hunt JP, Wilkerson JW, Funk PJ, Swensen RL, Carver WC, Christian ML, Bundy BC: **Endotoxin‐free E. coli‐based cell‐free protein synthesis: Pre‐expression endotoxin removal approaches for on‐demand cancer therapeutic production**. *Biotechnology journal* 2018:1800271.

13. Smith MT, Wilding KM, Hunt JM, Bennett AM, Bundy BC: **The emerging age of cell-free synthetic biology**. *FEBS Letters* 2014, **588**:2755-2761.

14. Bundy BC, Swartz JR: **Site-Specific Incorporation of p-Propargyloxyphenylalanine in a Cell-Free Environment for Direct Protein−Protein Click Conjugation**. *Bioconjugate Chemistry* 2010, **21**:255-263.

15. Yang JH, Kanter G, Voloshin A, Levy R, Swartz JR: **Expression of active murine granulocyte-macrophage colony-stimulating factor in an Escherichia coli cell-free system**. *Biotechnology Progress* 2004, **20**:1689-1696.

16. Wilding KM, Schinn S-M, Long EA, Bundy BC: **The emerging impact of cell-free chemical biosynthesis**. *Current Opinion in Biotechnology* 2018, **53**:115-121.

17. You C, Shi T, Li Y, Han P, Zhou X, Zhang Y-HP: **An in vitro synthetic biology platform for the industrial biomanufacturing of myo-inositol from starch**. *Biotechnology and Bioengineering* 2017, **114**:1855-1864.

18. Zawada JF, Yin G, Steiner AR, Yang J, Naresh A, Roy SM, Gold DS, Heinsohn HG, Murray CJ: **Microscale to manufacturing scale‐up of cell‐free cytokine production—a new approach for shortening protein production development timelines**. *Biotechnology and bioengineering* 2011, **108**:1570-1578.

19. Smith MT, Berkheimer SD, Werner CJ, Bundy BC: **Lyophilized *Escherichia coli*-based cell-free systems for robust, high-density, long-term storage**. *BioTechniques* 2014, **56**:8.

20. Wilding KM, Smith AK, Wilkerson JW, Bush DB, Knotts TA, Bundy BC: **The Locational Impact of Site-Specific PEGylation: Streamlined Screening with Cell-Free Protein Expression and Coarse-Grain Simulation**. *ACS Synthetic Biology* 2018.

21. Smith MT, Varner CT, Bush DB, Bundy BC: **The incorporation of the A2 protein to produce novel Q β virus‐like particles using cell‐free protein synthesis**. *Biotechnology progress* 2012, **28**:549-555.

22. Liu JW, Wang Z, He JT, Wang GZ, Zhang RB, Zhao BH: **Effect of site-specific PEGylation on the fibrinolytic activity, immunogenicity, and pharmacokinetics of staphylokinase**. *Acta Biochimica Et Biophysica Sinica* 2014, **46**:782-791.

23. Hadadian S, Mirzahoseini H, Shamassebi DN, Shokrgozar MA, Bouzari S, Asgari S: **Chemoselective PEGylation of Cysteine Analogs of Human Basic Fibroblast Growth Factor (hbFGF) - Design and Expression**. *Tropical Journal of Pharmaceutical Research* 2014, **13**:1601-1607.

24. Cho H, Daniel T, Buechler YJ, Litzinger DC, Maio Z, Putnam A-MH, Kraynov VS, Sim B-C, Bussell S, Javahishvili T, et al.: **Optimized clinical performance of growth hormone with an expanded genetic code**. *Proceedings of the National Academy of Sciences* 2011, **108**:9060-9065.

25. Mamat U, Wilke K, Bramhill D, Schromm AB, Lindner B, Kohl TA, Corchero JL, Villaverde A, Schaffer L, Head SR, et al.: **Detoxifying Escherichia coli for endotoxin-free production of recombinant proteins**. *Microbial Cell Factories* 2015, **14**:57.

26. Sanchez-Garcia L, Martín L, Mangues R, Ferrer-Miralles N, Vázquez E, Villaverde A: **Recombinant pharmaceuticals from microbial cells: a 2015 update**. *Microbial cell factories* 2016, **15**:33.

27. Magalhães PO, Lopes AM, Mazzola PG, Rangel-Yagui C, Penna TC, Pessoa Jr A: **Methods of endotoxin removal from biological preparations: a review**. *J Pharm Pharm Sci* 2007, **10**:388-404.

28. Wells E, Robinson AS: **Cellular engineering for therapeutic protein production: product quality, host modification, and process improvement**. *Biotechnology journal* 2017, **12**.

29. Lopes AM, Magalhaes PO, Mazzola PG, Rangel‐Yagui CO, De Carvalho J, Penna TCV, Pessoa A: **LPS removal from an E. coli fermentation broth using aqueous two‐phase micellar system**. *Biotechnology progress* 2010, **26**:1644-1653.

30. Saraswat M, Musante L, Ravidá A, Shortt B, Byrne B, Holthofer H: **Preparative Purification of Recombinant Proteins: Current Status and Future Trends**. *BioMed Research International* 2013, **2013**:312709.

31. Anspach FB: **Endotoxin removal by affinity sorbents**. *Journal of biochemical and biophysical methods* 2001, **49**:665-681.

32. Fogarty JA, Swartz JR: **The exciting potential of modular nanoparticles for rapid development of highly effective vaccines**. *Current Opinion in Chemical Engineering* 2018, **19**:1-8.

33. Salehi ASM, Smith MT, Bennett AM, Williams JB, Pitt WG, Bundy BC: **Cell-free protein synthesis of a cytotoxic cancer therapeutic: Onconase production and a just-add-water cell-free system**. *Biotechnology Journal* 2016, **11**:274-281.

34. Boles KS, Kannan K, Gill J, Felderman M, Gouvis H, Hubby B, Kamrud KI, Venter JC, Gibson DG: **Digital-to-biological converter for on-demand production of biologics**. *Nature biotechnology* 2017, **35**:672.

35. Smith MT, Hawes AK, Shrestha P, Rainsdon JM, Wu JC, Bundy BC: **Alternative fermentation conditions for improved Escherichia coli-based cell-free protein synthesis for proteins requiring supplemental components for proper synthesis**. *Process Biochemistry* 2014, **49**:217-222.

36. Salehi ASM, Yang SO, Earl CC, Shakalli Tang MJ, Porter Hunt J, Smith MT, Wood DW, Bundy BC: **Biosensing estrogenic endocrine disruptors in human blood and urine: A RAPID cell-free protein synthesis approach**. *Toxicology and Applied Pharmacology* 2018, **345**:19-25.

37. Smith MT, Bennett AM, Hunt JM, Bundy BC: **Creating a completely "cell-free" system for protein synthesis**. *Biotechnology Progress* 2015, **31**:1716-1719.

38. Liu S, Tobias R, McClure S, Styba G, Shi Q, Jackowski G: **Removal of Endotoxin from Recombinant Protein Preparations**. *Clinical Biochemistry* 1997, **30**:455-463.

39. Aida Y, Pabst MJ: **Removal of endotoxin from protein solutions by phase separation using Triton X-114**. *Journal of immunological methods* 1990, **132**:191-195.

40. Schinn S-M, Bradley W, Groesbeck A, Wu JC, Broadbent A, Bundy BC: **Rapid in vitro screening for the location-dependent effects of unnatural amino acids on protein expression and activity**. *Biotechnology and Bioengineering* 2017, **114**:2412-2417.

41. Teodorowicz M, Perdijk O, Verhoek I, Govers C, Savelkoul HFJ, Tang Y, Wichers H, Broersen K: **Optimized Triton X-114 assisted lipopolysaccharide (LPS) removal method reveals the immunomodulatory effect of food proteins**. *PloS one* 2017, **12**:e0173778.

42. Nguyen HA, Su Y, Lavie A: **Design and Characterization of Erwinia Chrysanthemi l-Asparaginase Variants with Diminished l-Glutaminase Activity**. *Journal of Biological Chemistry* 2016, **291**:17664-17676.

43. Kaltschmidt E: **Ribosomal proteins. XIV: Isoelectric points of ribosomal proteins of E. coli as determined by two-dimensional polyacrylamide gel electrophoresis**. *Analytical Biochemistry* 1971, **43**:25-31.

44. Chng S-S, Ruiz N, Chimalakonda G, Silhavy TJ, Kahne D: **Characterization of the two-protein complex in Escherichia coli responsible for lipopolysaccharide assembly at the outer membrane**. *Proceedings of the National Academy of Sciences* 2010, **107**:5363.

45. Steimle A, Autenrieth IB, Frick J-S: **Structure and function: Lipid A modifications in commensals and pathogens**. *International Journal of Medical Microbiology* 2016, **306**:290-301.

46. Meredith TC, Aggarwal P, Mamat U, Lindner B, Woodard RW: **Redefining the Requisite Lipopolysaccharide Structure in Escherichia coli**. *ACS Chemical Biology* 2006, **1**:33-42.

47. Viranaicken W, Nativel B, Krejbich-Trotot P, Harrabi W, Bos S, El Kalamouni C, Roche M, Gadea G, Desprès P: **ClearColi BL21(DE3)-based expression of Zika virus antigens illustrates a rapid method of antibody production against emerging pathogens**. *Biochimie* 2017, **142**:179-182.

48. Rueda F, Cano-Garrido O, Mamat U, Wilke K, Seras-Franzoso J, García-Fruitós E, Villaverde A: **Production of functional inclusion bodies in endotoxin-free Escherichia coli**. *Applied Microbiology and Biotechnology* 2014, **98**:9229-9238.

49. Woodard RW, Meredith TC, Aggarwal P: **Viable non-toxic gram-negative bacteria**. Google Patents; 2012.

50. Chang DE, Smalley DJ, Conway T: **Gene expression profiling of Escherichia coli growth transitions: an expanded stringent response model**. *Molecular microbiology* 2002, **45**:289-306.

51. Klammt C, Schwarz D, Dötsch V, Bernhard F: **Cell-Free Production of Integral Membrane Proteins on a Preparative Scale**. In *In Vitro Transcription and Translation Protocols*. Edited by Grandi G. Humana Press; 2007:57-78.

52. *ClearColi® BL21(DE3)*

Electrocompetent Cells. Lucigen; 2016.

53. Chien W-W, Allas S, Rachinel N, Sahakian P, Julien M, Le Beux C, Lacroix C-E, Abribat T, Salles G: **Pharmacology, immunogenicity, and efficacy of a novel pegylated recombinant Erwinia chrysanthemi-derived L-asparaginase**. *Investigational New Drugs* 2014, **32**:795-805.

54. Lagassé HAD, Alexaki A, Simhadri VL, Katagiri NH, Jankowski W, Sauna ZE, Kimchi-Sarfaty C: **Recent advances in (therapeutic protein) drug development**. *F1000Research* 2017, **6**:113.

55. Schinn S-M, Broadbent A, Bradley WT, Bundy BC: **Protein synthesis directly from PCR: progress and applications of cell-free protein synthesis with linear DNA**. *New Biotechnology* 2016, **33**:480-487.

56. Cicerone MT, Pikal MJ, Qian KK: **Stabilization of proteins in solid form**. *Advanced Drug Delivery Reviews* 2015, **93**:14-24.

57. Tonnis WF, Mensink MA, de Jager A, van der Voort Maarschalk K, Frijlink HW, Hinrichs WLJ: **Size and Molecular Flexibility of Sugars Determine the Storage Stability of Freeze-Dried Proteins**. *Molecular Pharmaceutics* 2015, **12**:684-694.

58. Pardee K, Green AA, Takahashi MK, Braff D, Lambert G, Lee JW, Ferrante T, Ma D, Donghia N, Fan M: **Rapid, low-cost detection of Zika virus using programmable biomolecular components**. *Cell* 2016, **165**:1255-1266.

59. Pardee K, Green Alexander A, Ferrante T, Cameron DE, DaleyKeyser A, Yin P, Collins James J: **Paper-Based Synthetic Gene Networks**. *Cell* 2014, **159**:940-954.

60. Karig DK, Bessling S, Thielen P, Zhang S, Wolfe J: **Preservation of protein expression systems at elevated temperatures for portable therapeutic production**. *Journal of the Royal Society Interface* 2017, **14**:20161039.

61. Larsen BS, Skytte J, Svagan AJ, Meng-Lund H, Grohganz H, Löbmann K: **Using dextran of different molecular weights to achieve faster freeze-drying and improved storage stability of lactate dehydrogenase**. *Pharmaceutical Development and Technology* 2018:1-6.

62. Garzon-Rodriguez W, Koval RL, Chongprasert S, Krishnan S, Randolph TW, Warne NW, Carpenter JF: **Optimizing Storage Stability of Lyophilized Recombinant Human Interleukin-11 with Disaccharide/Hydroxyethyl Starch Mixtures**. *Journal of Pharmaceutical Sciences* 2004, **93**:684-696.

63. Soles C, Tsai A, Cicerone M: **Glass Dynamics and the Preservation of Proteins**. In *Misbehaving Proteins*. Springer New York; 2006:193-214.

64. Bissoyi A, Kumar A, Rizvanov AA, Nesmelov A, Gusev O, Patra PK, Bit A: **Recent advances and future direction in lyophilisation and desiccation of mesenchymal stem cells**. *Stem cells international* 2016, **2016**.

65. Obrzut J, Anopchenko A, Douglas JF, Rust BW: **Relaxation and antiplasticization measurements in trehalose–glycerol mixtures – A model formulation for protein preservation**. *Journal of Non-Crystalline Solids* 2010, **356**:777-781.

66. Lerbret A, Affouard F: **Molecular packing, hydrogen bonding, and fast dynamics in lysozyme/trehalose/glycerol and trehalose/glycerol glasses at low hydration**. *The Journal of Physical Chemistry B* 2017, **121**:9437-9451.

67. Cicerone MT, Tellington A, Trost L, Sokolov A: **Substantially Improved Stability of Biological Agents in Dried Form**. *BioProcess International* 2003.

68. Devineni D, Gonschorek C, Cicerone MT, Xu Y, Carpenter JF, Randolph TW: **Storage stability of keratinocyte growth factor-2 in lyophilized formulations: Effects of formulation physical properties and protein fraction at the solid–air interface**. *European Journal of Pharmaceutics and Biopharmaceutics* 2014, **88**:332-341.

69. Abdul‐Fattah AM, Truong‐Le V, Yee L, Nguyen L, Kalonia DS, Cicerone MT, Pikal MJ: **Drying-Induced Variations in Physico-Chemical Properties of Amorphous Pharmaceuticals and Their Impact on Stability (I): Stability of a Monoclonal Antibody***. *Journal of Pharmaceutical Sciences* 2007, **96**:1983-2008.

70. Cicerone MT, Tellington A, Trost L, Sokolov A: **Plasticized hydrophilic glasses for improved stabilization of biological agents**. Google Patents; 2006.

71. Averett D, Cicerone MT, Douglas JF, de Pablo JJ: **Fast relaxation and elasticity-related properties of trehalose-glycerol mixtures**. *Soft Matter* 2012, **8**:4936-4945.

72. Cicerone MT, Soles CL: **Fast Dynamics and Stabilization of Proteins: Binary Glasses of Trehalose and Glycerol**. *Biophysical Journal* 2004, **86**:3836-3845.

73. Anopchenko A, Psurek T, VanderHart D, Douglas JF, Obrzut J: **Dielectric study of the antiplasticization of trehalose by glycerol**. *Physical Review E* 2006, **74**:031501.

74. Cicerone MT, Tellington A, Trost L, Sokolov A: **Substantially Improved Stability**. *BioProcess International* 2003.

75. Jewett MC, Swartz JR: **Mimicking the Escherichia coli cytoplasmic environment activates long‐lived and efficient cell‐free protein synthesis**. *Biotechnology and bioengineering* 2004, **86**:19-26.

76. Albayrak C, Swartz JR: **Using E. coli-based cell-free protein synthesis to evaluate the kinetic performance of an orthogonal tRNA and aminoacyl-tRNA synthetase pair**. *Biochemical and biophysical research communications* 2013, **431**:291-295.

77. Calhoun KA, Swartz JR: **Energy systems for ATP regeneration in cell-free protein synthesis reactions**. In *In vitro transcription and translation protocols*. Springer; 2007:3-17.

78. Forney‐Stevens KM, Bogner RH, Pikal MJ: **Addition of Amino Acids to Further Stabilize Lyophilized Sucrose‐Based Protein Formulations: I. Screening of 15 Amino Acids in Two Model Proteins**. *Journal of pharmaceutical sciences* 2015.

79. Veronese FM, Pasut G: **PEGylation, successful approach to drug delivery**. *Drug Discovery Today* 2005, **10**:1451-1458.

80. Pasut G, Veronese FM: **State of the art in PEGylation: The great versatility achieved after forty years of research**. *Journal of Controlled Release* 2012, **161**:461-472.

81. Pasut G, Sergi M, Veronese FM: **Anti-cancer PEG-enzymes: 30 years old, but still a current approach**. *Advanced Drug Delivery Reviews* 2008, **60**:69-78.

82. Mei B, Pan C, Jiang H, Tjandra H, Strauss J, Chen Y, Liu T, Zhang X, Severs J, Newgren J, et al.: **Rational design of a fully active, long-acting PEGylated factor VIII for hemophilia A treatment**. *Blood* 2010, **116**:270.

83. Chiu K, Agoubi LL, Lee I, Limpar MT, Lowe JW, Goh SL: **Effects of Polymer Molecular Weight on the Size, Activity, and Stability of PEG-Functionalized Trypsin**. *Biomacromolecules* 2010, **11**:3688-3692.

84. Kontermann RE: **Strategies for extended serum half-life of protein therapeutics**. *Current Opinion in Biotechnology* 2011, **22**:868-876.

85. da Silva Freitas D, Abrahão-Neto J: **Biochemical and biophysical characterization of lysozyme modified by PEGylation**. *International Journal of Pharmaceutics* 2010, **392**:111-117.

86. Dhalluin C, Ross A, Leuthold L-A, Foser S, Gsell B, Müller F, Senn H: **Structural and Biophysical Characterization of the 40 kDa PEG−Interferon-α2a and Its Individual Positional Isomers**. *Bioconjugate Chemistry* 2005, **16**:504-517.

87. Lawrence PB, Gavrilov Y, Matthews SS, Langlois MI, Shental-Bechor D, Greenblatt HM, Pandey BK, Smith MS, Paxman R, Torgerson CD, et al.: **Criteria for Selecting PEGylation Sites on Proteins for Higher Thermodynamic and Proteolytic Stability**. *Journal of the American Chemical Society* 2014, **136**:17547-17560.

88. Illanes A, Cauerhff A, Wilson L, Castro GR: **Recent trends in biocatalysis engineering**. *Bioresource Technology* 2012, **115**:48-57.

89. Bailon P, Palleroni A, Schaffer CA, Spence CL, Fung W-J, Porter JE, Ehrlich GK, Pan W, Xu Z-X, Modi MW, et al.: **Rational Design of a Potent, Long-Lasting Form of Interferon: A 40 kDa Branched Polyethylene Glycol-Conjugated Interferon α-2a for the Treatment of Hepatitis C**. *Bioconjugate Chemistry* 2001, **12**:195-202.

90. Levine PM, Craven TW, Bonneau R, Kirshenbaum K: **Intrinsic bioconjugation for site-specific protein PEGylation at N-terminal serine**. *Chemical Communications* 2014, **50**:6909-6912.

91. Brocchini S, Godwin A, Balan S, Choi J-w, Zloh M, Shaunak S: **Disulfide bridge based PEGylation of proteins**. *Advanced Drug Delivery Reviews* 2008, **60**:3-12.

92. Rosendahl MS, Doherty DH, Smith DJ, Carlson SJ, Chlipala EA, Cox GN: **A Long-Acting, Highly Potent Interferon α-2 Conjugate Created Using Site-Specific PEGylation**. *Bioconjugate Chemistry* 2005, **16**:200-207.

93. Doherty DH, Rosendahl MS, Smith DJ, Hughes JM, Chlipala EA, Cox GN: **Site-Specific PEGylation of Engineered Cysteine Analogues of Recombinant Human Granulocyte-Macrophage Colony-Stimulating Factor**. *Bioconjugate Chemistry* 2005, **16**:1291-1298.

94. Zimmerman ES, Heibeck TH, Gill A, Li X, Murray CJ, Madlansacay MR, Tran C, Uter NT, Yin G, Rivers PJ, et al.: **Production of Site-Specific Antibody–Drug Conjugates Using Optimized Non-Natural Amino Acids in a Cell-Free Expression System**. *Bioconjugate Chemistry* 2014, **25**:351-361.

95. Tian F, Lu Y, Manibusan A, Sellers A, Tran H, Sun Y, Phuong T, Barnett R, Hehli B, Song F, et al.: **A general approach to site-specific antibody drug conjugates**. *Proceedings of the National Academy of Sciences* 2014, **111**:1766-1771.

96. Salehi ASM, Smith MT, Schinn S-M, Hunt JM, Muhlestein C, Diray-Arce J, Nielsen BL, Bundy BC: **Efficient tRNA degradation and quantification in Escherichia coli cell extract using RNase-coated magnetic beads: A key step toward codon emancipation**. *Biotechnology Progress*:n/a-n/a.

97. Bundy BC, Swartz JR: **Efficient disulfide bond formation in virus-like particles**. *Journal of Biotechnology* 2011, **154**:230-239.

98. Bundy BC, Franciszkowicz MJ, Swartz JR: **Escherichia coli-based cell-free synthesis of virus-like particles**. *Biotechnology and Bioengineering* 2008, **100**:28-37.

99. Park C-G, Kim T-W, Oh I-S, Song JK, Kim D-M: **Expression of functional Candida antarctica lipase B in a cell-free protein synthesis system derived from Escherichia coli**. *Biotechnology Progress* 2009, **25**:589-593.

100. Rodríguez-Martínez JA, Solá RJ, Castillo B, Cintrón-Colón HR, Rivera-Rivera I, Barletta G, Griebenow K: **Stabilization of α-Chymotrypsin upon PEGylation Correlates with Reduced Structural Dynamics**. *Biotechnology and bioengineering* 2008, **101**:1142-1149.

101. Plesner B, Westh P, Nielsen AD: **Biophysical characterisation of GlycoPEGylated recombinant human factor VIIa**. *International Journal of Pharmaceutics* 2011, **406**:62-68.

102. Plesner B, Fee CJ, Westh P, Nielsen AD: **Effects of PEG size on structure, function and stability of PEGylated BSA**. *European Journal of Pharmaceutics and Biopharmaceutics* 2011, **79**:399-405.

103. Zang Q, Tada S, Uzawa T, Kiga D, Yamamura M, Ito Y: **Two site genetic incorporation of varying length polyethylene glycol into the backbone of one peptide**. *Chemical Communications* 2015, **51**:14385-14388.

104. Grace MJ, Lee S, Bradshaw S, Chapman J, Spond J, Cox S, DeLorenzo M, Brassard D, Wylie D, Cannon-Carlson S, et al.: **Site of Pegylation and Polyethylene Glycol Molecule Size Attenuate Interferon-α Antiviral and Antiproliferative Activities through the JAK/STAT Signaling Pathway**. *Journal of Biological Chemistry* 2005, **280**:6327-6336.

105. Lim SI, Cho J, Kwon I: **Double clicking for site-specific coupling of multiple enzymes**. *Chemical Communications* 2015, **51**:13607-13610.

106. Qiu H, Boudanova E, Park A, Bird JJ, Honey DM, Zarazinski C, Greene B, Kingsbury JS, Boucher S, Pollock J, et al.: **Site-Specific PEGylation of Human Thyroid Stimulating Hormone to Prolong Duration of Action**. *Bioconjugate Chemistry* 2013, **24**:408-418.

107. Lim SI, Hahn YS, Kwon I: **Site-specific albumination of a therapeutic protein with multi-subunit to prolong activity in vivo**. *Journal of Controlled Release* 2015, **207**:93-100.

108. Arpino JAJ, Baldwin AJ, McGarrity AR, Tippmann EM, Jones DD: **In-Frame Amber Stop Codon Replacement Mutagenesis for the Directed Evolution of Proteins Containing Non-Canonical Amino Acids: Identification of Residues Open to Bio-Orthogonal Modification**. *PLOS ONE* 2015, **10**:e0127504.

109. Wu JCY, Hutchings CH, Lindsay MJ, Werner CJ, Bundy BC: **Enhanced Enzyme Stability Through Site-Directed Covalent Immobilization**. *Journal of Biotechnology* 2015, **193**:83-90.

110. Moatsou D, Li J, Ranji A, Pitto-Barry A, Ntai I, Jewett MC, O'Reilly RK: **Self-Assembly of Temperature-Responsive Protein–Polymer Bioconjugates**. *Bioconjugate Chemistry* 2015, **26**:1890-1899.

111. Chin JW, Santoro SW, Martin AB, King DS, Wang L, Schultz PG: **Addition of p-Azido-l-phenylalanine to the Genetic Code of Escherichia coli**. *Journal of the American Chemical Society* 2002, **124**:9026-9027.

112. Young TS, Ahmad I, Yin JA, Schultz PG: **An Enhanced System for Unnatural Amino Acid Mutagenesis in E. coli**. *Journal of Molecular Biology* 2010, **395**:361-374.

113. Smith MT, Berkheimer SD, Werner CJ, Bundy BC: **Lyophilized E*scherichia coli* -based cell-free systems for robust, high-density, long-term storageLyophilized Escherichia coli-based cell-free systems for robust, high-density, long-term storage**. *BioTechniques* 2014, **56**:8.

114. Earl CC, Smith MT, Lease RA, Bundy BC: **Polyvinylsulfonic acid: A Low-cost RNase inhibitor for enhanced RNA preservation and cell-free protein translation**. *Bioengineered* 2017:1-8.

115. Chatterjee A, Sun SB, Furman JL, Xiao H, Schultz PG: **A Versatile Platform for Single- and Multiple-Unnatural Amino Acid Mutagenesis in Escherichia coli**. *Biochemistry* 2013, **52**:1828-1837.

116. Abràmoff MD, Ram SJ: **Image Processing with ImageJ**. *Biophotonics International* 2004, **11**:7.

117. Debets MF, van Berkel SS, Dommerholt J, Dirks AJ, Rutjes FPJT, van Delft FL: **Bioconjugation with Strained Alkenes and Alkynes**. *Accounts of Chemical Research* 2011, **44**:805-815.

118. Debets MF, van Berkel SS, Schoffelen S, Rutjes FPJT, van Hest JCM, van Delft FL: **Aza-dibenzocyclooctynes for fast and efficient enzyme PEGylation via copper-free (3+2) cycloaddition**. *Chemical Communications* 2010, **46**:97-99.

119. Wei S, Knotts TA: **Effects of tethering a multistate folding protein to a surface**. *The Journal of Chemical Physics* 2011, **134**:185101.

120. Wei S, Knotts TA: **Predicting stability of alpha-helical, orthogonal-bundle proteins on surfaces**. *The Journal of Chemical Physics* 2010, **133**:115102.

121. Bush DB, Knotts TA: **Probing the effects of surface hydrophobicity and tether orientation on antibody-antigen binding**. *The Journal of Chemical Physics* 2017, **146**:155103.

122. Bush DB, Knotts TA: **Communication: Antibody stability and behavior on surfaces**. *The Journal of Chemical Physics* 2015, **143**:061101.

123. Karanicolas J, Brooks Iii CL: **Improved Gō-like Models Demonstrate the Robustness of Protein Folding Mechanisms Towards Non-native Interactions**. *Journal of Molecular Biology* 2003, **334**:309-325.

124. Karanicolas J, Brooks CL: **The origins of asymmetry in the folding transition states of protein L and protein G**. *Protein Science* 2002, **11**:2351-2361.

125. Dennington R, Keith TA, Millam JM: **GaussView, Version 5**. Semichem Inc.; 2016.

126. Jensen HB, Kleppe K: **Effect of Ionic Strength, pH, Amines and Divalent Cations on the Lytic Activity of T4 Lysozyme**. *European Journal of Biochemistry* 1972, **28**:116-122.

127. Ahmad S, Gromiha M, Fawareh H, Sarai A: **ASAView: Database and tool for solvent accessibility representation in proteins**. *BMC Bioinformatics* 2004, **5**:51-51.

128. Loong BK, Knotts TA: **Communication: Using multiple tethers to stabilize proteins on surfaces**. *The Journal of Chemical Physics* 2014, **141**:051104.

129. Xu J, Bussiere J, Yie J, Sickmier A, An P, Belouski E, Stanislaus S, Walker KW: **Polyethylene Glycol Modified FGF21 Engineered to Maximize Potency and Minimize Vacuole Formation**. *Bioconjugate Chemistry* 2013, **24**:915-925.

130. Albayrak C, Swartz JR: **Cell-free co-production of an orthogonal transfer RNA activates efficient site-specific non-natural amino acid incorporation**. *Nucleic Acids Research* 2013, **41**:5949-5963.

131. Kolb HC, Sharpless KB: **The growing impact of click chemistry on drug discovery**. *Drug Discovery Today* 2003, **8**:1128-1137.

132. Thirumurugan P, Matosiuk D, Jozwiak K: **Click Chemistry for Drug Development and Diverse Chemical–Biology Applications**. *Chemical Reviews* 2013, **113**:4905-4979.

133. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE: **UCSF Chimera—A visualization system for exploratory research and analysis**. *Journal of Computational Chemistry* 2004, **25**:1605-1612.

134. Reddington SC, Tippmann EM, Dafydd Jones D: **Residue choice defines efficiency and influence of bioorthogonal protein modification via genetically encoded strain promoted Click chemistry**. *Chemical Communications* 2012, **48**:8419-8421.

135. Lee BK, Kwon JS, Kim HJ, Yamamoto S, Lee EK: **Solid-Phase PEGylation of Recombinant Interferon α-2a for Site-Specific Modification: Process Performance, Characterization, and in Vitro Bioactivity**. *Bioconjugate Chemistry* 2007, **18**:1728-1734.

136. Johnson CM: **Differential scanning calorimetry as a tool for protein folding and stability**. *Archives of Biochemistry and Biophysics* 2013, **531**:100-109.

137. Matsumura M, Becktel WJ, Levitt M, Matthews BW: **Stabilization of phage T4 lysozyme by engineered disulfide bonds**. *Proceedings of the National Academy of Sciences of the United States of America* 1989, **86**:6562-6566.

138. Gray TM, Matthews BW: **Structural analysis of the temperature-sensitive mutant of bacteriophage T4 lysozyme, glycine 156----aspartic acid**. *Journal of Biological Chemistry* 1987, **262**:16858-16864.

139. Wetzel R, Perry LJ, Baase WA, Becktel WJ: **Disulfide bonds and thermal stability in T4 lysozyme**. *Proceedings of the National Academy of Sciences of the United States of America* 1988, **85**:401-405.

140. Baase WA, Liu L, Tronrud DE, Matthews BW: **Lessons from the lysozyme of phage T4**. *Protein Science : A Publication of the Protein Society* 2010, **19**:631-641.

141. Mooers BHM, Baase WA, Wray JW, Matthews BW: **Contributions of all 20 amino acids at site 96 to the stability and structure of T4 lysozyme**. *Protein Science : A Publication of the Protein Society* 2009, **18**:871-880.

142. Pandey BK, Smith MS, Torgerson C, Lawrence PB, Matthews SS, Watkins E, Groves ML, Prigozhin MB, Price JL: **Impact of Site-Specific PEGylation on the Conformational Stability and Folding Rate of the Pin WW Domain Depends Strongly on PEG Oligomer Length**. *Bioconjugate Chemistry* 2013, **24**:796-802.

143. Hartley AM, Zaki AJ, McGarrity AR, Robert-Ansart C, Moskalenko AV, Jones GF, Craciun MF, Russo S, Elliott M, Macdonald JE, et al.: **Functional modulation and directed assembly of an enzyme through designed non-natural post-translation modification**. *Chemical Science* 2015, **6**:3712-3717.

144. Harris JM, Martin NE, Modi M: **Pegylation**. *Clinical Pharmacokinetics* 2001, **40**:539-551.

145. Deiters A, Cropp TA, Summerer D, Mukherji M, Schultz PG: **Site-specific PEGylation of proteins containing unnatural amino acids**. *Bioorganic & Medicinal Chemistry Letters* 2004, **14**:5743-5745.

146. Balan S, Choi J-w, Godwin A, Teo I, Laborde CM, Heidelberger S, Zloh M, Shaunak S, Brocchini S: **Site-Specific PEGylation of Protein Disulfide Bonds Using a Three-Carbon Bridge**. *Bioconjugate Chemistry* 2007, **18**:61-76.

147. Kim TH, Swierczewska M, Oh Y, Kim A, Jo DG, Park JH, Byun Y, Sadegh-Nasseri S, Pomper MG, Lee KC, et al.: **Mix to Validate: A Facile, Reversible PEGylation for Fast Screening of Potential Therapeutic Proteins In Vivo**. *Angewandte Chemie International Edition* 2013, **52**:6880-6884.

148. Cong Y, Pawlisz E, Bryant P, Balan S, Laurine E, Tommasi R, Singh R, Dubey S, Peciak K, Bird M, et al.: **Site-Specific PEGylation at Histidine Tags**. *Bioconjugate Chemistry* 2012, **23**:248-263.

149. Nairn NW, Shanebeck KD, Wang A, Graddis TJ, VanBrunt MP, Thornton KC, Grabstein K: **Development of Copper-Catalyzed Azide–Alkyne Cycloaddition for Increased in Vivo Efficacy of Interferon β-1b by Site-Specific PEGylation**. *Bioconjugate Chemistry* 2012, **23**:2087-2097.

150. Chao S-H, Matthews SS, Paxman R, Aksimentiev A, Gruebele M, Price JL: **Two Structural Scenarios for Protein Stabilization by PEG**. *The Journal of Physical Chemistry B* 2014, **118**:8388-8395.

151. Yang C, Lu D, Liu Z: **How PEGylation Enhances the Stability and Potency of Insulin: A Molecular Dynamics Simulation**. *Biochemistry* 2011, **50**:2585-2593.

152. Knotts TA, Rathore N, de Pablo JJ: **An Entropic Perspective of Protein Stability on Surfaces**. *Biophysical Journal* 2008, **94**:4473-4483.

153. **SEER Cancer Stat Facts: Acute Lymphocytic Leukemia**. National Cancer Institute.

154. Allas S, Sahakian P, Fichtner I, Abribat T: **Pharmacokinetics and Pharmacodynamics in Mice of a Pegylated Recombinant Erwinia Chrysanthemi-Derived L-Asparaginase**. *Blood* 2009, **114**:2033.

155. Li J, Zhou Y, Li C, Wang D, Gao Y, Zhang C, Zhao L, Li Y, Liu Y, Li X: **Poly(2-ethyl-2-oxazoline)–Doxorubicin Conjugate-Based Dual Endosomal pH-Sensitive Micelles with Enhanced Antitumor Efficacy**. *Bioconjugate Chemistry* 2015, **26**:110-119.

156. Mero A, Fang Z, Pasut G, Veronese FM, Viegas TX: **Selective conjugation of poly(2-ethyl 2-oxazoline) to granulocyte colony stimulating factor**. *Journal of Controlled Release* 2012, **159**:353-361.

157. Rau RE, Dreyer Z, Choi MR, Liang W, Skowronski R, Allamneni KP, Devidas M, Raetz EA, Adamson PC, Blaney SM, et al.: **Outcome of pediatric patients with acute lymphoblastic leukemia/lymphoblastic lymphoma with hypersensitivity to pegaspargase treated with PEGylated Erwinia asparaginase, pegcrisantaspase: A report from the Children's Oncology Group**. *Pediatric Blood & Cancer* 2018, **65**:e26873-n/a.

158. Basnett J, Xie V, Cisterne A, Bradstock K, Bendall L: **Regulation of the bone marrow microenvironment by G-CSF: Effects of G-CSF on acute lymphoblastic leukaemia**. *PLoS ONE* 2017, **12**:e0188042.

159. Mehta HM, Malandra M, Corey SJ: **G-CSF and GM-CSF in Neutropenia**. *Journal of immunology (Baltimore, Md. : 1950)* 2015, **195**:1341-1349.

160. Heinzelman P, Schoborg JA, Jewett MC: **pH responsive granulocyte colony-stimulating factor variants with implications for treating Alzheimer's disease and other central nervous system disorders**. *Protein Engineering, Design and Selection* 2015, **28**:481-489.

161. Nguyen HA, Su Y, Zhang JY, Antanasijevic A, Caffrey M, Schalk AM, Liu L, Rondelli D, Oh A, Mahmud DL, et al.: **A Novel l-Asparaginase with low l-Glutaminase Coactivity Is Highly Efficacious against Both T- and B-cell Acute Lymphoblastic Leukemias In Vivo**. *Cancer Research* 2018, **78**:1549.

162. Pourhossein M, Korbekandi H: **Cloning, expression, purification and characterisation of Erwinia carotovora L-asparaginase in Escherichia coli**. *Advanced Biomedical Research* 2014, **3**:82.

163. Krasotkina J, Borisova AA, Gervaziev YV, Sokolov NN: **One‐step purification and kinetic properties of the recombinant l‐asparaginase from Erwinia carotovora**. *Biotechnology and Applied Biochemistry* 2004, **39**:215-221.

164. Gervais D, King D: **Capillary isoelectric focusing of a difficult-to-denature tetrameric enzyme using alkylurea–urea mixtures**. *Analytical Biochemistry* 2014, **465**:90-95.

165. Hill CP, Osslund TD, Eisenberg D: **The structure of granulocyte-colony-stimulating factor and its relationship to other growth factors**. *Proceedings of the National Academy of Sciences* 1993, **90**:5167.

166. Rosenblum G, Cooperman BS: **Engine out of the Chassis: Cell-Free Protein Synthesis and its Uses**. *FEBS letters* 2014, **588**:261-268.

167. Larsen JEP, Lund O, Nielsen M: **Improved method for predicting linear B-cell epitopes**. *Immunome Research* 2006, **2**:2-2.

168. Ponomarenko JV, Bourne PE: **Antibody-protein interactions: benchmark datasets and prediction tools evaluation**. *BMC Structural Biology* 2007, **7**:64-64.

169. Haste Andersen P, Nielsen M, Lund O: **Prediction of residues in discontinuous B-cell epitopes using protein 3D structures**. *Protein Science : A Publication of the Protein Society* 2006, **15**:2558-2567.

170. Moola ZB, Scawen MD, Atkinson T, Nicholls DJ: **Erwinia chrysanthemi L-asparaginase: epitope mapping and production of antigenically modified enzymes**. *Biochemical Journal* 1994, **302**:921-927.

171. Cantor JR, Yoo TH, Dixit A, Iverson BL, Forsthuber TG, Georgiou G: **Therapeutic enzyme deimmunization by combinatorial T-cell epitope removal using neutral drift**. *Proceedings of the National Academy of Sciences of the United States of America* 2011, **108**:1272-1277.

172. Gervais D, Hayzen J, Orphanou C, McEntee A, Hallam C, Brehm R: **Understanding the process-induced formation of minor conformational variants of Erwinia chrysanthemi l-asparaginase**. *Enzyme and microbial technology* 2017, **98**:26-33.

173. Kotzia GA, Labrou NE: **L-Asparaginase from Erwinia chrysanthemi 3937: cloning, expression and characterization**. *Journal of biotechnology* 2007, **127**:657-669.

174. Shrestha P, Holland TM, Bundy BC: **Streamlined extract preparation for Escherichia coli-based cell-free protein synthesis by sonication or bead vortex mixing**. *BioTechniques* 2012, **53**:163-174.

APPENDIX A. SUPPLEMENTARY MATERIALS FOR CHAPTER 2

A.1 Detailed supplementary methods

A.1.1 Extract preparation

Cell extracts were made using either the *E. coli* BL21 Star™ (DE3) strain (BL21 (DE3)*) (Invitrogen, Carlsbad, CA) or ClearColi® BL21 DE3 chemically competent cells (Lucigen, Middleton, WI). Each extract was grown using sterile technique without the addition of antibiotic. Standard BL21 (DE3)* extract was produced as described previously [35]. Briefly, overnight cultures of 5 mL of LB were inoculated into 100 mL of 2xYT and incubated at 37°C and 280 rpm until an OD_{600} of 2. The 100 mL culture was then inoculated into a 2.5 L shake flask containing 900 mL of 2xYT. T7 RNA Polymerase expression was induced with 1 mM IPTG at an OD_{600} 0.5-0.7, and cells were harvested in mid-long phase (around OD_{600} 2.5 – 3.5 in this work). Cells were then harvested, homogenized using an Avestin Emulsiflex-B15 homogenizer, and prepared as previously described [174]. ClearColi® extracts were prepared similarly, with the following modifications. First, TB media with 5 g/L or 10 g/L NaCl was used in place of 2xYT to support higher cell densities of ClearColi® growth. Second, a 100 mL culture was inoculated directly from a frozen glycerol stock of the ClearColi® BL21 DE3 cells and the culture was allowed to grow overnight, reaching an OD of 2-3. Cells were induced at a variety of OD_{600}, varying from 0.1 –

0.7, and harvested at a variety of OD_{600} ranging from 0.6 – 3.5. Cells were harvested, lysed, and otherwise prepared according to the same procedure as the standard BL21 (DE3)* cells.

A.1.2 Triton X-114 two-phase extractions

Triton X-114 (Sigma Aldrich) was added to 0.5-1 mL BL21 (DE3)* extracts in a pyrogen-free microcentrifuge tube (VWR) to a final concentration of 1% by volume. Samples were then incubated at 4°C for 30 minutes with end-over-end spinning, following which they were transferred to a 37°C water bath for 10 minutes. After these two incubations, the samples were centrifuged at > 16,000 g for 10 minutes at 25°C. Centrifuged samples presented three phases: a lower micelle-rich phase, a middle aqueous phase, and an upper oily phase (Figure A-1). The upper two phases were extracted by pipetting, and the procedure was repeated with the extracted portion 2 additional times. When yields of extracts treated using this method were found to be low, the method was modified to decrease the processing time. This later method was based off of a protocol previously reported, with a few modifications [39]. Triton X-114 was added to extract samples (0.5-1 mL) to a final concentration of 1% by volume. Samples were then vigorously vortexed and incubated on ice for 5 minutes. Samples were then vortexed again and incubated in a 37°C water bath for 5 minutes, followed by centrifugation at >16,000 g for 7 minutes at 25°C. This procedure resulted in the sample layer formations described for the initial protocol, and the top two phases were likewise removed by pipetting. Yields and endotoxin content were assessed after 1 or 3 cycles of this phase extraction. Unless otherwise specified, Triton X-114 phase extraction results in this paper refer to treatment using this shortened protocol.

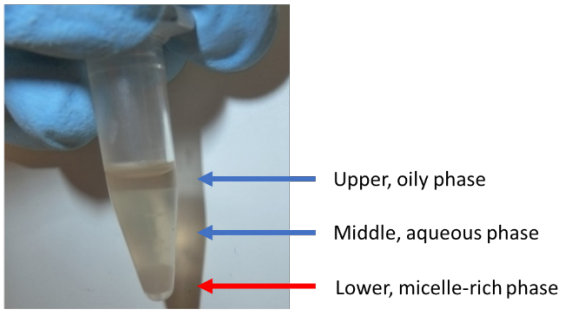


Figure A-1: Triton X-114 2-Phase Extraction with E. coli Cell Extract. Phases formed during Triton X-114 extraction with cell-free extract. The upper two layers are extracted, leaving the endotoxin-rich lower phase.

A.1.3 Polylysine affinity chromatography

Pierce™ High Capacity Endotoxin Removal Spin Columns, 0.5 mL were obtained from ThermoFisher Scientfic. Columns were equilibrated and washed as directed by the manufacturer. 1 mL samples of BL21 (DE3)* extracts were diluted to 2 mL using endotoxin-free LAL reagent water (Lucigen, Middleton, WI). The 2 mL sample was loaded onto the equilibrated, washed polylysine spin column and incubated end-over-end at 4°C for 1 hour. Columns were then centrifuged at 500 g for 1 minute to elute sample, and the eluent was collected in a 15 mL falcon tube. The collected sample was then loaded into a fresh polylysine spin column and the procedure was repeated. Samples were tested after 1, 5, and 10 column passes.

A.1.4 Media optimization

Media optimizations were performed in two sets. For the first set, 5 mL LB-Miller media was inoculated from a frozen glycerol stock of ClearColi® BL21 DE3 Electrocompetent cells and grown overnight. The 5 mL growths were then each used to inoculate 100 mL of the test media in a 500 mL shake flask. The shake flask was incubated at 37°C and 280 rpm, with OD_{600} measurements taken using a Synergy-MX every 30-60 minutes. Media tested included LB-Miller

with 10-20 g/L total NaCl, 2xYT with 5-20 g/L total NaCl, and TB with 0-15 g/L total NaCl. For the second set, LB-Miller with 10-20 g/L total NaCl, 2xYT with 5-20 g/L total NaCl, and TB with 0-20 g/L total NaCL were each tested. A 10 mL culture of each medium was inoculated to an initial OD_{600} of <0.05 in 50 mL conical centrifuge tubes and incubated at 37 °C and 280 rpm. OD_{600} measurements were taken in duplicate every 60-90 minutes for 10 hours.

A.1.5 LAL testing

Multiple dilutions of each sample ($1:10^3 – 1:10^7$) were prepared using endotoxin-free water (Lonza, Basel, Switzerland) to ensure measurement within the range of the standard curve of the assay. Samples were loaded into a clear, pyrogen-free 96-well plate and incubated for 10 minutes at 37°C. LAL reagent was then added to the plate using a multichannel pipette, and the absorbance at 340 nm was monitored using a Synergy-MX microplate reader (BioTek, Winooski, VT). Each dilution was assayed in duplicate. The time for the absorbance to increase by 0.03 was recorded and compared to a standard curve prepared with supplied LPS standard in order to determine the LAL reactivity of the sample in endotoxin equivalent units. This assay is known to not differentiate between the *E. coli* lipid IV_A and true endotoxin, which results in endotoxin readings for ClearColi® samples [25,46]. It has been well documented that the activity of these same samples using cell-based assays reveals the inactivity of this residual lipid IV_A in immune recognition [25,46].

A.2 Supplementary results and discussion

A.2.1 Pre-expression endotoxin removal via Triton X-114 two-phase extraction

Triton X-114 extraction has been in use for decades and is one of the most common endotoxin removal strategies [28]. Traditionally, Triton X-114 is added to a protein sample and incubated at both 4°C and 37°C. Upon centrifugation two phases form due to excluded volume interactions, with the upper aqueous phase being enriched in protein and the lower micelle-rich phase being enriched in LPS [27,38]. Extraction of the aqueous layer then allows separation from the LPS. This technique has been applied with great success to a variety of purified proteins, with >98% endotoxin removal reported [27,38]. However, multiple repetitions of the procedure are typically required to sufficiently remove endotoxin, which can further reduce protein recovery yield and activity [27,38]. Additional purification steps may also be required to remove residual surfactant [27,41]. However, if Triton X-114 was used to pre-treat CFPS extract, residual Triton X-114 would likely be removed during protein purification. It may even help reduce endotoxin in the purified product, as inclusion of Triton X-114 in wash steps of affinity chromatography has been reported to enhance separation from endotoxin [27].

Triton X-114 two phase extraction was thus used for pre-expression endotoxin removal from *E. coli* extract. Two methods were initially compared for Triton X-114 extraction, one with significantly shorter incubation periods that the other, as described in the Methods. The protocol employing shorter incubation times removed endotoxin with a similar efficiency as the longer protocol, yet produced approximately double the yields in protein expression (Figure A-2). The protocol with shorter 5 min incubation periods was therefore selected for the bulk of this work. Endotoxin removal and yield retention were assessed after performing the phase extraction one or

three times, with the results reported in Figure 2-2. While there was significant variability in the performance of the Triton X-114 treated extracts, the majority of endotoxin extraction and the majority of the loss of extract activity occurred after the first extraction, with extracts maintaining ~40% of its protein production activity and ~13% of the original endotoxin. After three Triton X-114 extractions, the endotoxin content was reduced to ~5% while the extract retained ~30% of its protein production yield. It should be noted that Triton X-114 two-phase extraction and affinity based endotoxin removal have been designed for final stage treatment of a protein therapeutic that has already been significantly purified, whereas untreated cell-free extract contained approximately 18 million endotoxin units (EU) per milliliter–at least 100 times more than the initial endotoxin content of untreated samples previously reported [25,27,29,38]. Therefore, while 95% endotoxin removal is consistent with previous reports for Triton X-114 extractions on purified proteins, due to the high initial endotoxin content of standard CFPS extract, even 95% endotoxin removal results in significant levels of residual endotoxin (1 MEU/mL). The loss of extract activity is not unexpected, due to the many small molecules and macromolecules involved in protein transcription and translation which could be lost during the phase extraction. Further optimization of Triton X-114 extraction protocols may mitigate some of these losses [41]. However, considering the loss in activity and the significant concentration of residual endotoxin, additional methods for endotoxin removal were explored.

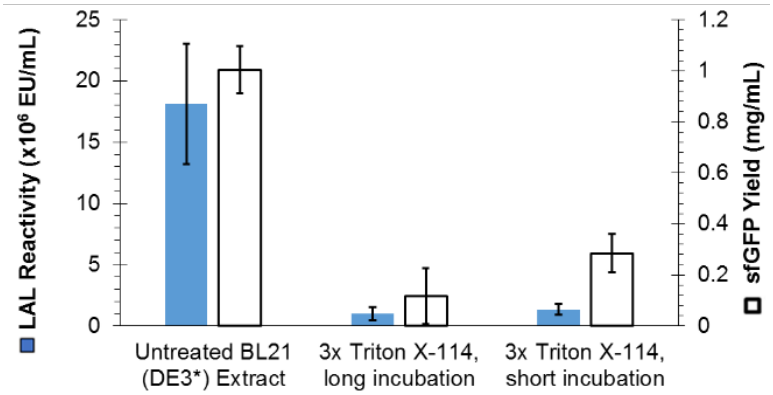


Figure A-2: Comparison of Triton X-114 Extraction Methods. LAL reactivity and sfGFP yield of untreated extract and extracts treated with Triton X-114 using the long or short incubation times described in the methods. Shorter incubation times remove approximately the same amount of endotoxin but preserve greater activity. Error bars represent standard deviation from n = 4.

A.2.2 Pre-expression endotoxin removal via polylysine affinity chromatography

Another method of endotoxin removal employs affinity adsorption to selectively remove endotoxin from protein samples. Many successful embodiments of this strategy have exploited various chemical characteristics of LPS, including its hydrophobicity and negative charge [31]. However, affinity adsorption also presents several challenges. For example, endotoxin adsorption is reportedly best when performed at a pH near the pI of the desired protein in order to maximize endotoxin removal and protein recovery; the neutral charge of the protein at this pH is desirable because positively charged proteins will interact more strongly with endotoxin and negatively charged proteins will bind to the column along with the endotoxin [31]. This pH guideline is not always ideal, however, as the target protein may be less stable at this pH. A common endotoxin adsorbent is polmyxin B, an antibiotic with high affinity for the negatively charged LPS. However, polymyxin B treatment requires long contact times and can introduce polymyxin B contamination which must be removed before administration [31]. As an alternative, high capacity polylysine columns similarly interact with the negatively charged LPS but require only 1 hour contact times.

Polylysine treatment has also been reported to have improved protein recoveries compared to polymyxin B treatment [31]. For these reasons, endotoxin removal and cell extract viability after treatment with polylysine affinity columns were assessed.

Despite the resin's high reported binding capacity of 2 MEU/mL, extract samples were diluted 1x with water in order to meet the recommended 4:1 sample to resin ratio without overwhelming the column. Still, due to the high endotoxin content of untreated extract, multiple treatments were necessary. Endotoxin content and protein synthesis capability of cell-free extracts was evaluated after 1, 5, or 10 column passes (Figure 2-2). Although the columns removed a significant amount of endotoxin, they were less effective than the Triton X-114 treatment. A single column pass reduced the endotoxin content of the treated extract by ~2 MEU which represented a decrease of only ~10%, though the extract did retain nearly 60% of its original protein synthesis activity. Additional column treatments further reduced the endotoxin content, but extract viability was lost after 5 column passes. The dramatic effect of the polylysine treatment on extract viability is likely due to the column adsorbing essential CFPS components. For example, at a neutral pH where the extracts were treated, tRNA, rRNA, and multiple protein components of the *E. coli* ribosome are expected to be negatively charged [43]. While the columns effectively removed endotoxin and further optimization of extract dilution or pH could improve nucleic acid and protein recovery, the current negative effects on extract protein synthesis activity necessitate further engineering before such a process is economical for bulk extract preparations.

APPENDIX B. SUPPLEMENTARY MATERIALS FOR CHAPTER 3

B.1 Estimated percentage water loss during lyophilization

$$\%H_2O\ lost = \frac{\Delta m_{total}}{m_{H2O}} \quad \text{(B-1)}$$

$$m_{H2O} = \rho_{H2O} \times vol_{ext} \quad \text{(B-2)}$$

$$\rho_{H2O} = \rho_{Extract} - \rho_{protein} \quad \text{(B-3)}$$

B.2 Comparison of lyophilization methods

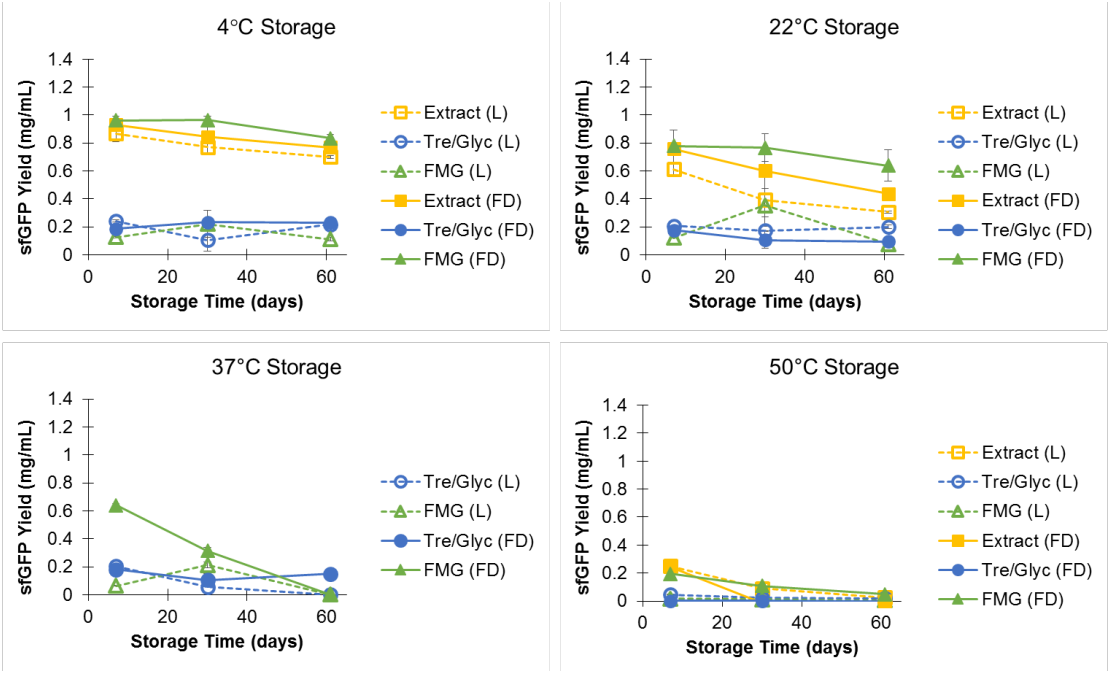


Figure B-1: Comparison of Lyophilization Methods. Samples were dried in either (1) 50 mL falcon tubes in a LabConco freeze dryer (hollow data points), or (2) 75 mL glass vials on a FlexiDry manifold freeze dryer (filled data points). In all three formulations tested, the samples dried in the manifold freeze dryer (FD) performed approximately as well as or slightly better than samples dried in the LabConco freeze-dryer. Because the manifold freeze drying also allowed for faster drying and, in some cases, visibly more complete drying, the manifold freeze dryer was used for the bulk of this work.

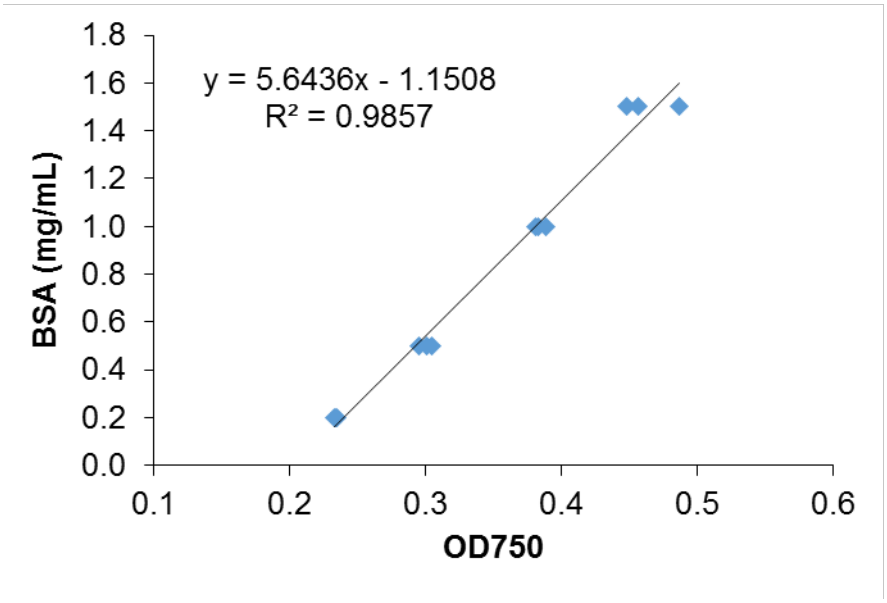


Figure B-2: Sample Standard Curve for Protein Concentration Determination using DC Protein Assay. Bovine serum albumin (BSA) was used as a protein standard, and a new standard was run for each DC assay.

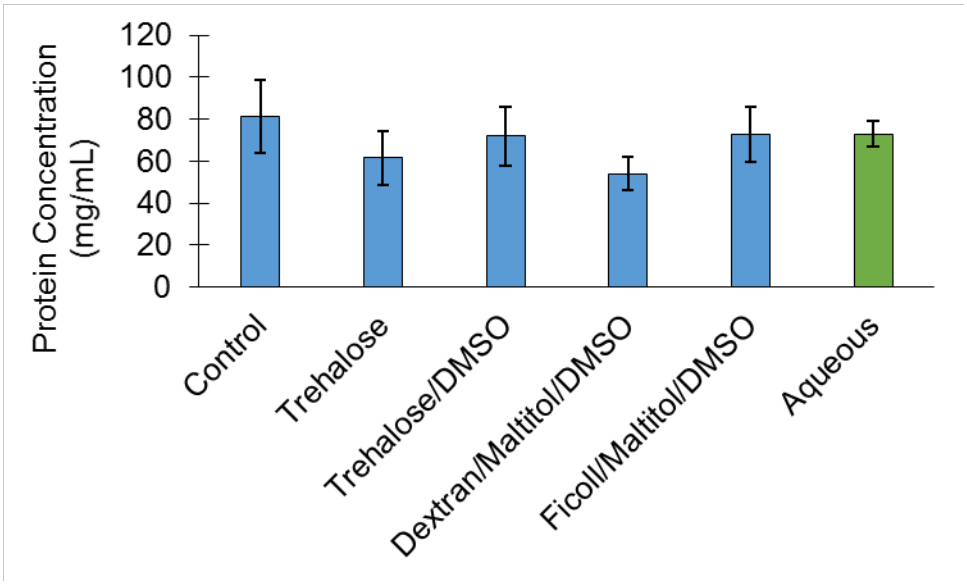


Figure B-3: DC Protein Assay Results for Rehydrated Lyophilized Extract Using Estimated Rehydration Ratios. Error represents standard deviation, n = 6 except for Trehalose/DMSO where n = 2.

Table B-1: Rehydration Ratios for Lyophilized Samples

Sample	RHR (uL H2O/mg dried sample)	Protein Concentration (mg/mL)
Aqueous	--	73 ± 6
Control	8.5	81 ± 17
Trehalose	2.57	61 ± 13
Trehalose/DMSO	2.57	71 ± 14
Dextran/Maltitol/DMSO	2.57	54 ± 8
Ficoll/Maltitol/DMSO	3.5	72 ± 17

APPENDIX C: SUPPLEMENTARY MATERIALS FOR CHAPTER 4

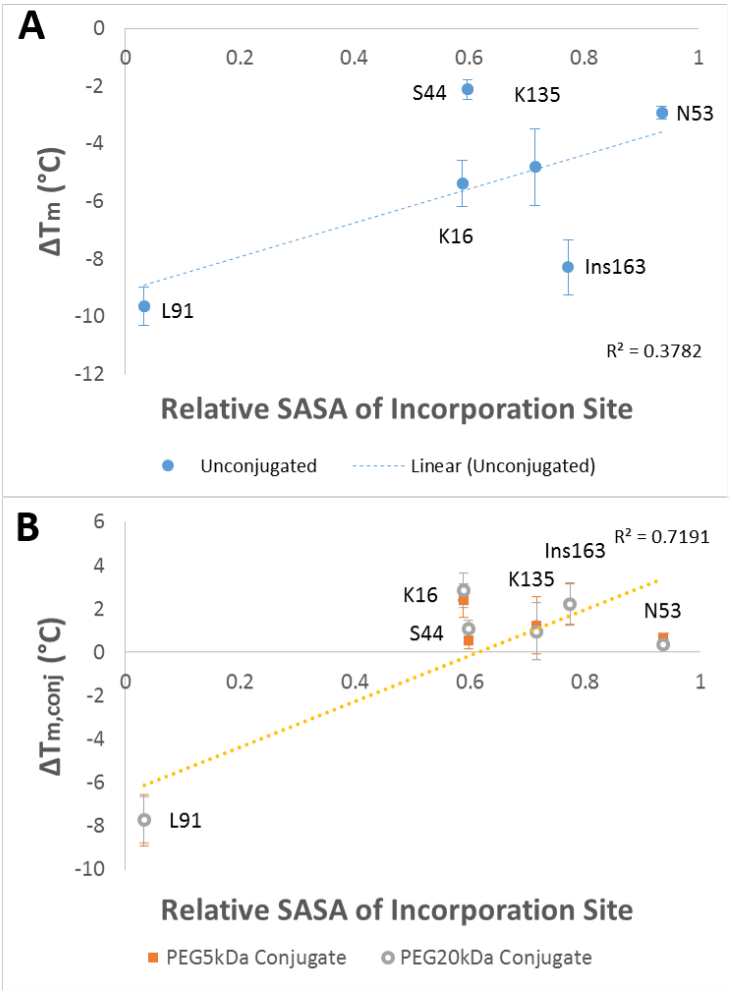


Figure C-1: Evaluating Contributing Factors for Prediction of Conjugate Stability. Relative SASA values from ASA-View[127] for each incorporation site. For the insertion Ins163, an average SASA of the two adjacent resides (162 and 163) is used for the evaluation. A: ΔT_m vs relative SASA of incorporation site of unconjugated variants. Note that while there appears to be a weak correlation, it is not predictive of Ins163 or S44, and would not have predicted that S44 would be least destabilized by AzF incorporation; B: $\Delta T_{m,conj}$ vs relative SASA of incorporation site. Again, there appears to be a weak correlation - mostly created by the low SASA and $\Delta T_{m,conj}$ of site L91 – which fails to predict the most stabilized sites, K16 and Ins163.

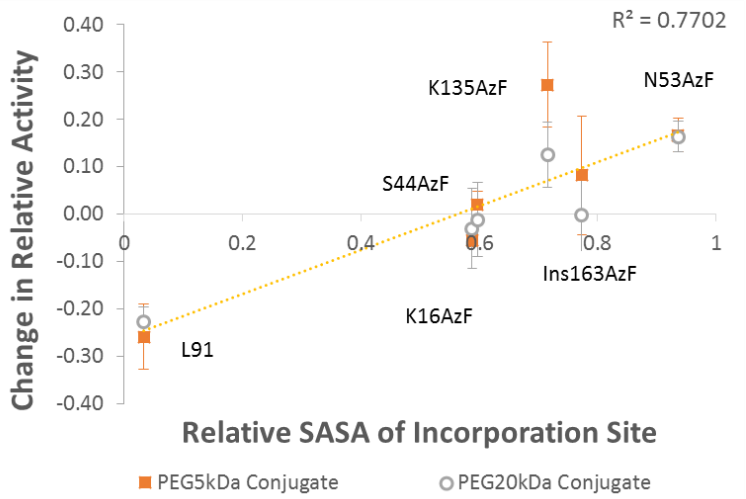


Figure C-2: Change in Activity due to PEGylation vs SASA of Incorporation Site. Change in relative activity, the activity of a lysozyme variant normalized to the activity of the corresponding WT (+ PEG) reactions, following PEGylation of each mono-AzF-incorporated Lyz shows a general positive trend with relative SASA of the conjugation site. However, the change in relative activity as a result of PEGylation was only significant for L91AzF and N53AzF, diminishing the utility of the trend. Nevertheless, the data indicates that SASA may provide general guidance in relation to which sites may be poor candidates for conjugation but may not be useful for predicting optimal conjugation sites.

C.1 Coarse-grain simulation

C.1.1 Additional model details

In the Gō-like model, the tertiary structure of the protein is defined by native contacts. Residues which hydrogen bond in the native structure or have one non-hydrogen atom within 4.5 angstroms of a non-hydrogen atom of another residue form a native contact. The energy of this interaction, V_{ij}, is defined as according to Equation C-1, where r_{ij} is the instantaneous distance between the C_α of the two residues i and j, ε_{ij} is the interaction energy between residues i and j, and σ_{ij} is the equilibrium distance between the C_α atoms of residues i and j in the crystal structure. Here, the ε_{ij} is the same for each residue and is equal to 378. Solvent effects are included implicitly, as an energy barrier for the water two residues must displace in order to come into close proximity.

$$V_{ij} = \epsilon_{ij}\left[13\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - 18\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{10} + 4\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \quad \text{(C-1)}$$

C.1.2 Heat capacity and native contacts results

Coarse-grain replica exchange simulations were run to determine the melting temperature for lysozyme PEGylated at each of the substitution sites. The change from WT melting temperature was compared to the experimental results, as reported in the main text. The plots below show the heat capacity and native contacts traces for WT T4 Lyz and T4 Lyz PEGylated at each of the substitutions with both PEG sizes. The point of maximum heat capacity is generally taken to be the protein melting temperature, which agreed well with the point of maximum change in native contacts. For each PEG size, the positions and magnitudes of the first two peaks appear to be relatively constant between the five variants, although both peaks are stabilized for all three variants in the 5kDa PEG conjugate group. Conjugation to site 91 appears to affect primarily the largest peak, which also corresponds to the greatest change in native contacts and therefore the T_m, causing it to merge with the second peak. These smaller two peaks are not reflected in the fluorescence data from the Protein Thermal Shift Assay. However, the general behavior of the major denaturation peak agrees well with the experimental data – that $\Delta T_{m,conj}$ for variants PEGylated at site 91 is significantly more destabilizing than any other site, K16 is most stabilizing, and other sites are similar to K16 but slightly less stabilizing (Figure C-3 and Figure C-4).

C.2 Supplemental PEGylation discussion

C.2.1 Possible explanations for absence of PEG size-dependent effects

One possible explanation for the independence from PEG size can be provided by examining studies which have found conformational stability to be dependent on PEG size. Two

studies have used smaller PEG oligomers, comparing the stability effects between conjugates of varying sizes of small PEG oligomers and between the small oligomers and a larger 2kDa PEG [87,142]. In one such study, the mechanism behind the size-dependent stabilization was found to be that longer PEG chains decreased the rate of protein unfolding in addition to increasing the rate of folding, while small PEGs only increased the rate of folding [142]. As the PEG chains used in this study are significantly larger than the PEG oligomers examined in the aforementioned studies, it is possible that the 5kDa PEG is sufficiently large to decrease the rate of protein unfolding to a point that the additional length of the 20kDa PEG has a negligible effect on the unfolding rate.

C.2.2 Difference between immobilization-based and PEGylation-based stabilization

The dramatic destabilization from PEGylation at site L91 is interesting, given that this site was determined to be the optimal site for lysozyme immobilization [109]. These results indicate that the effects on protein stability from conjugation to a polymer differ significantly from the effects due to conjugation to a surface. One possible explanation for this difference is that the flexible PEG chain does not constrain the protein in the same way that a rigid surface does. Immobilization was previously shown to stabilize T4 Lyz by eliminating an unfolded intermediate present in the WT folding mechanism [119]; apparently, the flexible PEG chain does not lead to the same elimination. To verify this, we performed molecular simulations and found that the folding of the protein occurs through a 4-state process in both the WT and PEGylated T4 Lyz variants meaning no intermediate was eliminated as was observed for immobilization [119] (See Figure C-4). The dominant effect of PEG at L91 may be interference with correct folding rather than prevention of unfolding. In a study with small PEG oligomers attached to accessible residues, Lawrence et al. reported that PEGs larger than 2 kDa stabilized protein folding and prevented protein unfolding [87], an effect similar to that observed with protein immobilization [119]. This

appears to be the mechanism by which PEGylation at the other sites stabilize T4 Lyz with respect to the unconjugated mutant, as the folding/unfolding pathway does not appear to be altered by the PEG (See Figure 4-5 and Figure C-3, and Figure C-4). However, PEGylation at L91 does appear to alter the shapes and distribution of the heat capacity peaks, suggesting that PEGylation of L91 may impact the folding/unfolding pathway in such a way as to destabilize the protein and downshift T_m.

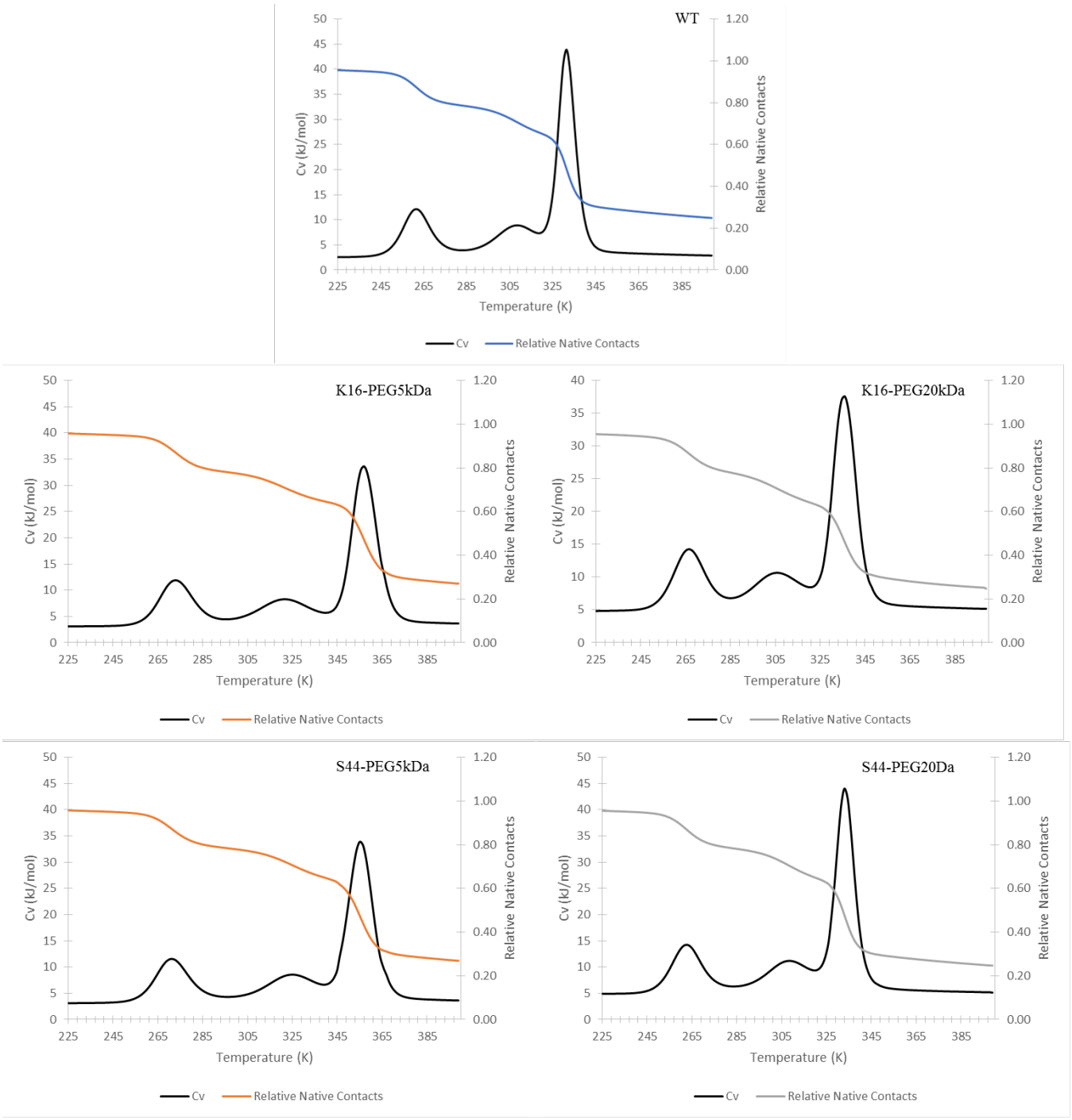


Figure C-3: Heat Capacity and Native Contacts Curves for WT, K16, and S44 from Simulation.

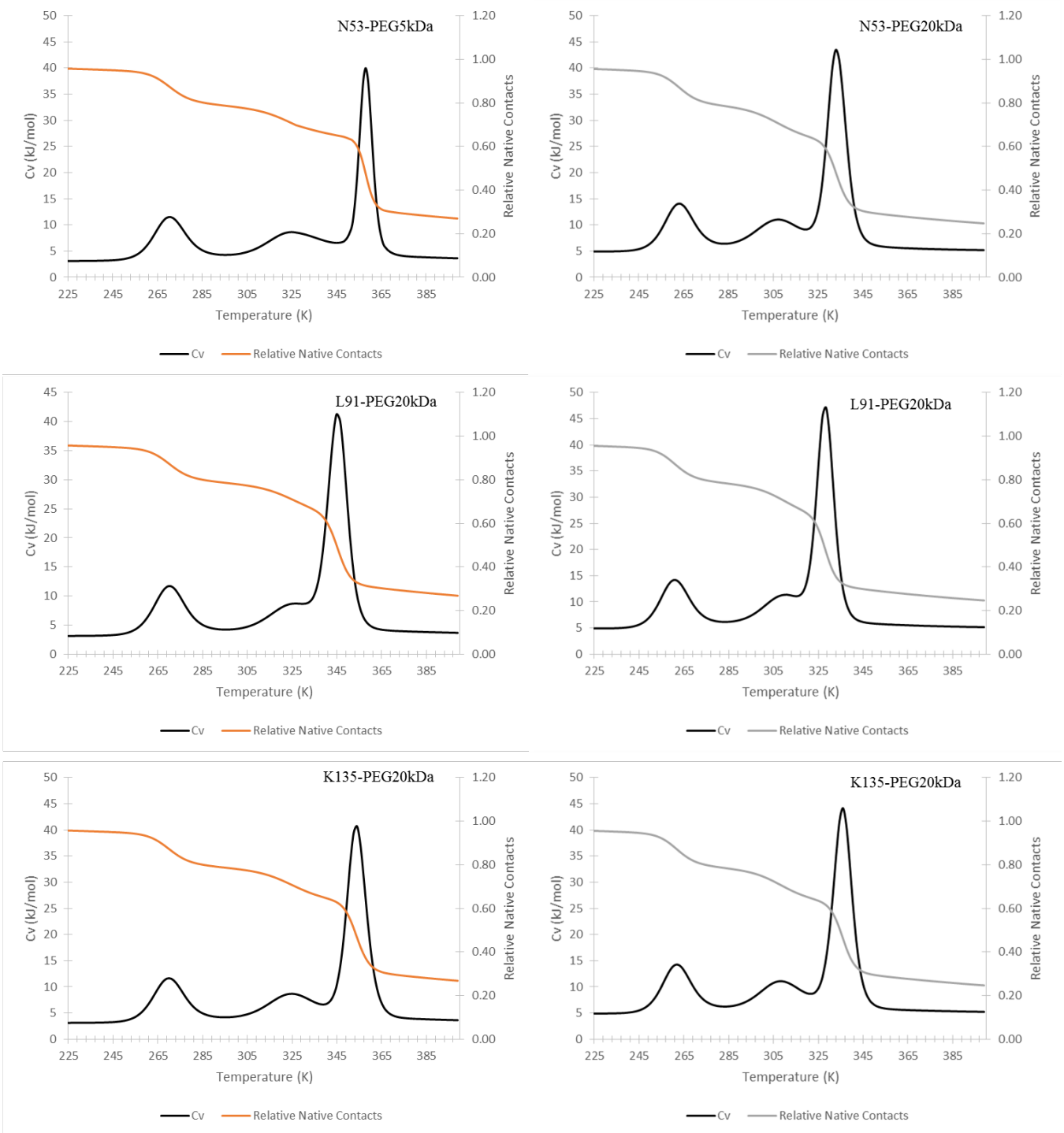


Figure C-4: Heat Capacity and Native Contacts Curves for N53, L91, and K135 from Simulation. Note that the heat capacity curves for lysozyme PEGylated at site 91 do not show elimination of any of the three peaks from the WT curve, but do show the highest temperature peak shifting down to begin merging with the intermediate temperature peak. This supports the hypothesis that excluded volume interactions due to PEGylation at this site in some way alters the folding and unfolding kinetics of T4 lysozyme, though no intermediates seem to be eliminat