Investigation of Curli-Specific Proteins CsgD and CsgA Reveals a Potential Target for Neurodegenerative Therapeutics and Provides Structural Insights into the Bacterial Amyloid CsgA

Jude Kinkead

December, 2022

Director of Thesis: John Cavanagh, Ph.D.

Major Department: Department of Biochemistry and Molecular Biology

ABSTRACT

A growing body of evidence implicates curli-containing biofilms in the onset of Alzheimer's and Parkinson's diseases. Curli-containing biofilms are composed primarily of two bacterial amyloid proteins, CsgA and CsgB. These proteins are prone to misfolding and possess common secondary structural features that allow for template-driven polymerization resulting in insoluble protein deposits. Deposition of endogenous amyloids αSyn and Aβ42 are a hallmark characteristic of Parkinson's and Alzheimer's respectively.

We are host to an array of microorganisms, which are able to directly interact with their hosts through the gut-brain axis. Curli-containing biofilms produced by enteric gut bacteria have been shown to promote motor impairment in mice, as well promote motor impairment and αSyn aggregation in aged Fischer 344 rats and *C. elegans*. Soluble oligomers of CsgA may be able to serve as a template for rapid polymerization of αSyn through a process known as cross-seeding. Curli fibrils are recognized by TLR2 and may also contribute to neuroinflammation.

Expression of curli fibrils is dependent upon a number of transcription factors, as well as the DnaK-DnaJ-GrpE chaperone system. Following translation of the *csgDEFG*-encoded proteins, the response regulator, CsgD, upregulates the transcription of the *csgBAC* operon.

CsgA and CsgB are kept soluble within the cell by CsgC and DnaK prior to periplasmic secretion through SecYEG machinery. During and following extracellular secretion, CsgA and CsgB are postulated to begin adopting the β-sheet rich structure similar to the mature fibril.

A class of compounds have shown specificity for response regulators and have demonstrated efficacy in both biofilm inhibition and dispersal. The 2-aminoimidazole compounds contain a 2-aminoimidazole pharmacore linked to various substituent groups. Precise tuning of the substituent group can provide for drastically increased specificity to a response regulator. In this work, a variety of approaches were implemented to isolate both *E. coli* CsgD and *S.*Typhimurium CsgD. Following purification, these proteins were screened utilizing a thermal shift assay to investigate their propensity for ligand binding by the 2-aminoimidazole compounds. *E. coli* CsgD, as well as *S.* Typhimurium CsgD suggest ligand binding in samples treated with AGL-869.

Diagnostic efforts for these diseases have largely focused on identifying stable oligomeric intermediates of endogenous amyloids. As a result of their propensity for aggregation, a typical approach utilized for the isolation of soluble, monomeric amyloids is the application of high concentrations of denaturants like guanidinium hydrochloride. While this is a standard approach, its application to an intrinsically disordered protein like CsgA may prohibit the purified isolate from inhabiting its native range of conformations - limiting investigations of stable oligomeric species. In order to combat this, we co-expressed the DnaK-DnaJ-GrpE in order to efficiently solubilize and isolate recombinant *E. coli* CsgA under native conditions. Natively-purified CsgA was shown to exhibit secondary structural characteristics that diverge from established literature. Characterization of natively-purified CsgA and initial comparisons to CsgA purified using high concentrations of denaturant are also included.

Investigation of Curli-Specific Proteins CsgD and CsgA Reveals a Potential Target for Neurodegenerative Therapeutics and Provides Structural Insights into the Bacterial Amyloid CsgA

A Thesis

Presented to the Faculty of the Department of Biochemistry and Molecular Biology

East Carolina University

In Partial Fulfillment of the Requirements for the Degree

Master of Science in Biomedical Sciences

By

Jude Kinkead

December, 2022

Director of Thesis: John, Cavanagh Ph.D.
Thesis Committee Members:
Tonya Zeczycki, Ph.D.
Morgan Milton, Ph.D.

ACKNOWLEDGEMENTS

I thank the Zeczycki and Cavanagh Labs for welcoming me into their labs, as well as for providing mentorship and numerous contributions to this research. I thank my committee members Dr. John Cavanagh, Dr. Tonya Zeczycki, and Dr. Morgan Milton for their aid in producing this work. I thank my parents, family, and girlfriend for their constant support and encouragement. I would also like to thank the Wooten family and the Wooten Family Initiative for Brain Health Research for funding this study.

TABLE OF CONTENTS

LIST OF TABLES v

LIST OF FIGURES vi

LIST OF ABBREVIATIONS ix

Section 1: Background 1

Section 2: E. coli CsgD Purification and 2-AI Binding Studies 11

Introduction 11

Methods 15

Data Analysis 21

Results 23

Discussion 43

Section 3: Background 48

Introduction 48

Methods 53

Data Analysis 55

Results 56

Discussion 64

Section 4: Background 67

Introduction 67

Methods 71

Data Analysis 76

Results 78

Discussion 97

REFERENCES 103

LIST OF TABLES

1. Primers Used for Generation of SUMO-CsgD Fusion Construct..... 22

2. Primers Used for Generation of dsDNA Fragments Containing CsgD Box 55

LIST OF FIGURES

1. Structural Breakdown of *E. coli* CsgA Sequence 5

2. NarL Receiver Domain Depicting $(\beta\alpha)_5$ Motif 7

3. LuxR C-terminal HTH DNA-binding Domain.....8

4. Expression Testing of *Ec* CsgD in BL21(DE3).....24

5. Expression Testing of *Ec* CsgD in Alternative Expression Strains 25

6. Ni^{2+} Purification of *Ec* CsgD Grown in HMS174(DE3) 26

7. Thermal Shift Assay of *Ec* CsgD Grown in HMS174(DE3) 27

8. Sarkosyl-Mediated Ni^{2+} Purification of *Ec* CsgD 28

9. Thermal Shift Assay of Sarkosyl-Purified *Ec* CsgD 29

10. Sarkosyl-Mediated Ni^{2+} Purification of *Ec* CsgD for BioBead Treatment 30

11. Thermal Shift Assay of BioBead-Treated *Ec* CsgD 31

12. Expression testing of SUMO-CsgD 32

13. Ni^{2+} Purification of SUMO-CsgD 33

14. Expression testing of SUMO-CsgD and *Ec* CsgD co-expressions with DnaK-DnaJ-GrpE Chaperone System 35

15. Ni^{2+} Purification of *Ec* CsgD co-expressed with DnaK-DnaJ-GrpE 36

16. Thermal Shift Assay of *Ec* CsgD co-expressed with DnaK-DnaJ-GrpE 37

17. Ni^{2+} Purification of SUMO-CsgD co-expressed with DnaK-DnaJ-GrpE 38

18. Expression testing of *Ec* CsgDN 39

19. Ni^{2+} Purification of *Ec* CsgDN 40

20. 2-AI Screening of *Ec* CsgDN 42

21. CsgD Sequence Alignments 49

22. *S*T CsgD N-terminal Domain Dimer 50

23. *S*T CsgD N-terminal Domain Organization.....51

24. *S*T CsgD N-terminal Domain Substitution Sites 51

25. Ni^{2+} Purification of *S*T CsgD.....56

26. Ni^{2+} Purification of *S*T CsgDN.....57

27. 2-AI Screening of *S*T CsgDN 59

28. 2-AI Screening of *S*T CsgD 59

29. Thermal Shift Assay for Detection of DNA-Binding in *S*T CsgD 60

30. Electrophoretic Mobility Shift Assay for Detection of DNA-Binding in *S*T CsgD .61

31. Crystallization Screening of *S*T CsgD.....63

32. AlphaFold Predicted Structure of *E. coli* CsgA.....68

33. Expression Testing of *E. coli* CsgB.....78

34. Expression Testing of *E. coli* CsgB co-expressed with DnaK-DnaJ-GrpE Chaperone System.....80

35. Ni^{2+} Purification of *E. coli* CsgB.....81

36. Expression Testing of *E. coli* CsgA co-expressed with DnaK-DnaJ-GrpE Chaperone System.....82

37. Ni^{2+} Purification of *E. coli* CsgA.....83

38. 6x His Staining, SDS-PAGE, and Native-PAGE of Natively-Purified *E. coli* CsgA 85

39. Circular Dichroism of Natively-Purified *E. coli* CsgA.....86

40. JC-1 Assay of Natively-Purified *E. coli* CsgA 87

41. Size Exclusion and Anion Exchange Attempts to Purify Monomeric *E. coli* CsgA .89

42. Chemically Induced Unfolding of Natively-Purified *E. coli* CsgA.....91

43. SDS-PAGE of On-Column Denatured *E. coli* CsgA.....93

44. Circular Dichroism of On-Column Denatured *E. coli* CsgA.....94

45. SDS-PAGE of Denaturant-Purified *E. coli* CsgA.....95

46. Native-PAGE of Denaturant-Purified *E. coli* CsgA.....96

LIST OF ABBREVIATIONS

AD Alzheimer's Disease 1

PD Parkinson's Disease 1

Aβ42 amyloid-beta 42 peptide 1

αSyn alpha-synuclein 1

APP Amyloid Precursor Protein 1

GABA γ-aminobutyric acid 3

c-di-GMP bis-(3′–5′)-cyclic dimeric guanosine monophosphate 6

rdar Red, Dry, and Rough 6

HSP70 Heat-Shock Protein 70 6

HTH Helix-Turn-Helix 7

AHL N-acyl homoserine lactone 8

2-AI 2-aminoimidazole 11

NMR Nuclear Magnetic Resonance 12

IPTG Isopropyl β-d-1-thiogalactopyranoside 12

Sarkosyl Nuclear Magnetic Resonance 12

SUMO Small Ubiquitin-like Modifier 13

ATP Adenosine-triphosphate 13

LB Luria-Broth 15

BME β-mercaptoethanol 17

$MgCl_2$ Magnesium Chloride 17

GuHCl, GdnHCl Guanidinium Hydrochloride 44

EDTA Ethylenediamine tetraacetic acid 54

EtBr Ethidium Bromide 54

Section 1. BACKGROUND

1.1 **Amyloidosis in Alzheimer's Disease (AD) and Parkinson's Disease (PD)**

Alzheimer's and Parkinson's disorders combine to affect more than 7 million Americans, with an associated healthcare cost exceeding $350 billion per year. Both the prevalence of these disorders and their economic impact are expected to increase in the coming decades. The healthcare cost of Alzheimer's disease alone is expected $1 trillion by the year 2050. These diseases are characterized by neurodegenerative and motor defects that typically present late in adulthood – one in three seniors will die afflicted with a dementia-related disorder.[1,2] A hallmark characteristic of both disorders is the deposition of amyloid proteins within localized regions of the brain. Classical amyloid nomenclature defines these proteins as those capable of demonstrating in vivo tissue deposition, as well as Congo red binding and green birefringence upon inspection by polarization microscopy. [3] Amyloid monomers are typically intrinsically-disordered or contain intrinsically-disordered regions. [4] The gain in flexibility afforded by existing in a partially unfolded state allows a single protein to participate in multiple processes as well as interact with their targets over large surface areas, however this functional gain is marred by its propensity for misfolding and aggregation. The amyloid-beta 42 peptide (Aβ42) is implicated in Alzheimer's disorder, while alpha-synuclein (αSyn) protein is evidenced in Parkinson's. [5,6]

1.2 **Amyloid Misfolding and Cytotoxicity**

Evidence supports the functionality of αSyn and Aβ42's precursor protein, Amyloid Precursor Protein (APP), within the cell, participating in regulation of the synaptic vesicle cycle

at the nerve terminal and in neural growth and repair respectively. [7,8] The transition from functional, soluble protein or peptide to insoluble fibril is described by a three state kinetic model sensitive to temperature, pH, and protein concentration. [9] This process begins with a slow lag phase during which soluble monomeric and multimeric species existing in the partially unfolded state begin to misfold, aggregate, and nucleate fibril development. The lag phase is followed by a drastic change in production of the insoluble fibril, which is believed to arise from cooperative binding and relatively unidirectional polymer development. Following the growth phase, fibrillation reaches equilibrium resulting in a sigmoidal curve for the development of amyloid fibrils. [9] In addition to the association of amyloid-containing neuritic inclusions with disease pathology and progression, familial mutations coinciding with increased neurodegenerative pathologies show greater propensity for aggregation in vitro. [6] This serves to identify the amyloid proteins as a pathologic agent in the progression of diseases like AD and PD. Since these proteins adopt multiple identities from soluble monomers to highly aggregated, insoluble fibrils, identification of the species and structures associated with disease pathology is critical to elucidating the mechanism of disease progression. For Aβ42, there is an inverse relationship between the molecular weight of multimers and their toxicity once a specific threshold is passed – implicating the soluble oligomers as the toxic species. [10] Toxicity of transient, soluble oligomers is supported by evidence in other amyloid disease systems. [11] Cytotoxicity through interactions with the plasma membrane may be increased for oligomers which exhibit greater disorder, diffusibility, number of reactive ends, and exposure of hydrophobic residues. [10]

1.3 **Alternative Sources for Amyloid Exposure**

In addition to endogenously-coded elements, the human body is host to an array of micro-organisms. For purposes of contextualizing the tremendous abundance of organisms within the human body, the number of bacteria inhabiting the gut is equal to the number of cells in the human body. [12] The diversity and abundance of these organisms fluctuates and varies in response to both intrinsic and extrinsic effects on the host including diet, drug usage, stress, and pregnancy. [13] These microorganisms are known to play critical roles in the regulation of mood, pain, metabolism, and both active and adaptive immunity, interacting with the host through its own metabolites and proteins. Microbial dysbiosis is clearly evidenced in certain cancers and systemic autoimmune disorders, as well as metabolic disorders and inflammatory-bowel disease. [13] In addition, evidence of interaction exists between intestinal-microbial communities and extra-intestinal organ communities in modulating tissue-specific immune responses of the skin, lungs, and liver, as well as in cancer and the central nervous system. The ability of some bacteria to produce γ-aminobutyric acid (GABA) and regulate GABA receptors in the brain is evidence of their ability to directly modulate neurological activity. [14] The vagus nerve may provide the avenue by which these interactions are able to occur, providing a physical linkage between the brainstem and gastrointestinal tract and allowing for bidirectional communication. [15]

1.4 **Bacterial Amyloids in AD and PD Progression**

Evidence implicating involvement of the gastrointestinal tract in Parkinson's has existed since James Parkinson's observations describing shaking palsy. [16] Later observation of patients that had undergone a vagotomy revealed a reduced incidence of Parkinson's disorder, highlighting

the potential role of the vagus nerve as a vehicle for disease pathology. [17] In patients, genome-wide screens identified positive association between prevalence of bacterial amyloids CsgA and CsgB and the observance of PD pathology. [18] In addition, analysis of fecal microbiomes in PD patients revealed elevated presence of Enterobacteriaceae. [19] Enterobacteriaceae are known to express extracellular curli fibrils composed primarily of the bacterial amyloids CsgA and CsgB and serve as a major structural component of enteric biofilms. Inability to correctly coordinate the curli fibrillation process results in a single layer biofilm as opposed to an elaborate and three-dimensional architecture seen in curli-containing biofilms. [20] Curli has been shown to promote αSyn aggregation and enhance motor impairment in aged Fischer 344 rats and *Caenorhabditis elegans (C. elegans)*. [21] Curli components have also been shown to colocalize with αSyn in *C. elegans* and human neuroblastoma cells, as well as contribute to neuronal death in vivo. [18] Curli's pathologic effect on αSyn aggregation is believed to stem from a well-documented process known as cross-seeding. [18] Cross-seeding refers to misfolded oligomeric and protofibril amyloids of one species serving as a physiological template for multimerization of lag-phase amyloids of another, allowing for the transition to rapid growth phase; bypassing the thermodynamic limits of lag-phase and resulting in a heterogeneous fibril. [22]

1.5 **Characteristics of the bacterial amyloid CsgA**

E. coli CsgA is a 151 residue protein with three defined regions; a secretory-targeting sequence, the N-terminal 22 amino acids of the mature protein, and an amyloid core consisting of five 19-22 AA long imperfect repeats with conserved serine, glutamine, glycine, asparagine, glutamine arrangement. [23]

SecYEG	M	K	L	L	K	V	A	A	I	A	A	I	V	F	S	G	S	A	L	A							
N22	G	V	V	P	Q	Y	G	G	G	G	N	H	G	G	G	G	N	N	S	G	P	N					
R1	S	E	L	N	I	Y	Q	Y	G	G	G	N	S	A	L	A	L	Q	T	D	A	R	N				
R2	S	D	L	T	I	T	Q	H	G	G	G	N	G	A	D	V	G	Q	G	S	D	D					
R3	S	S	I	D	L	T	Q	R	G	F	G	N	S	A	T	L	D	Q	W	N	G	K	N				
R4	S	E	M	T	V	K	Q	F	G	G	G	N	G	A	A	V	D	Q	T	A	S	N					
R5	S	S	V	N	V	T	Q	V	G	F	G	N	N	A	T	A	H	Q	Y								

Figure 1: Structural breakdown of *E. coli* CsgA sequence. The conserved arrangement of serine, glutamine, glycine, asparagine, and glutamine within repeats R1-R5 are highlighted in red.

At the cell-surface, the CsgA fibrillation process is nucleated by an amyloidogenic counterpart CsgB. [24] The repeating units R1, R3, and R5 have all demonstrated amyloidogenicity in vitro, with mutations in R1 and R5 exhibiting reduced or eliminated production of thin aggregative fimbrae in vivo. [24, 25] These studies have revealed the necessity of CsgB-mediated interactions with R1 and R5 of mature CsgA to induce heteronucleation. In addition it appears that key aspartate and glycine residues in R2, R3, and R4 serve to suppress the inherent aggregative potential of CsgA in vivo, reducing the incidence of fibril mislocalization. [26] Fibril localization is mediated by electrostatic interactions between surface lipids and lipoproteins that allow for amyloid retention at the cell-surface. [27] Current models predict that monomeric CsgA is excreted as an unstructured, monomeric coil that proceeds through β-sheet rich oligomers, resulting in a predicted β-helix or β-solenoid type fibril. [23, 28, 29] Polymerization of αSynuclein by CsgA cross-seeding is evidenced both in vivo and in vitro. [18, 23] Recent work has identified a stable homodimeric species of CsgA as well as an αSyn-CsgA complex that result in accelerated aggregation of αSyn. [30] It is important to consider that amyloid oligomeric intermediates are the species believed to exert cytotoxicity. In addition, for purposes

of soluble, monomeric isolation of CsgA from the cell, the established practice is to employ high concentrations of denaturant that allow for solubilization of membrane components and isolation of CsgA without aggregation. [31]

1.6 **Coordination of Biofilm Development**

Curli expression is a highly coordinated and regulated process dependent on transcription of both *csgDEFG* and *csgBAC* operons. [42] Biofilm development to confer physical and antibiotic resistance to its bacterial inhabitants is affected by numerous factors including pH, temperature, osmolarity, oxygen levels, specific ions, and hydrodynamics. [33] Transcription of the *csgBAC* operon encoding the structural components CsgB and CsgA is dependent upon both the transcription and activation of the master regulator CsgD. Control of *csgD* transcription is multi-faceted and affected by cell growth and environmental stimuli, as well as documented regulation by the secondary messenger bis-(3′–5′)-cyclic dimeric guanosine monophosphate (c-di-GMP) and at least five transcription factors. [33, 34] In addition to promoter acting elements and reliance on CsgD expression, development of the wild-type red, dry, and rough (rdar) biofilm phenotype is reliant on the DnaK chaperone system. [36] The DnaK chaperone system is the *E. coli* homolog of the eukaryotic Heat Shock Protein 70 (HSP70) system and interacts with transcriptional regulators of *csgD*, as well as both CsgD and CsgA.

1.7 **CsgD Transcriptional Regulation**

CsgD is the master regulator of biofilm formation in Enterobacteriaceae like *E. coli* and *S.* Typhimurium, orchestrating the transition from planktonic to biofilm growth states. [37] CsgD is a putative LuxR-like response regulator. [38] The canonical LuxR/FixJ family of response regulators possess an N-terminal receiver domain composed of five βα motifs that responds to cellular signals, as well as a helix-turn-helix (HTH) C-terminal DNA-binding domain connected by a disordered linker-region. [38]

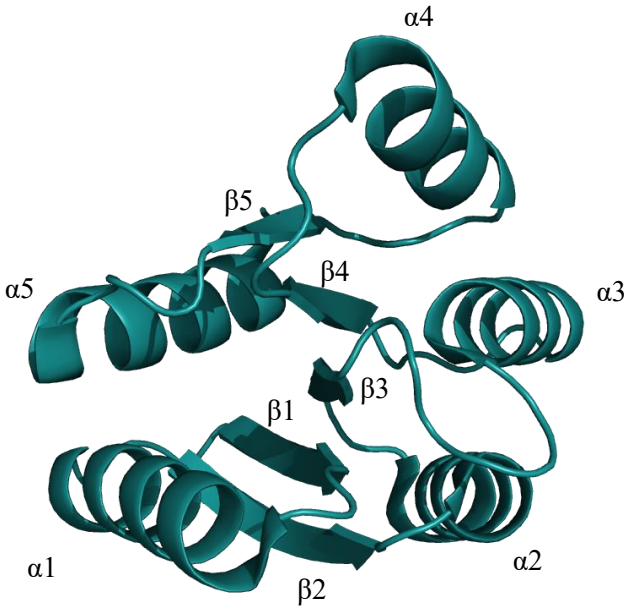

Figure 2: N-terminal receiver domain of NarL from *Mycobacterium tuberculosis* displaying the canonical $(\beta\alpha)_5$ motif PDB ID (3EUL).

In the classical quorum-sensing LuxR protein, in the absence of its N-terminal activator, N-acyl homoserine lactone (AHL), the N-terminal domain orients to obscure the DNA-binding activity of its C-terminal domain; sufficient concentrations of AHL trigger binding to the N-terminal domain leading to conformational change, multimerization, and DNA-binding through its C-terminal HTH domain. [39]

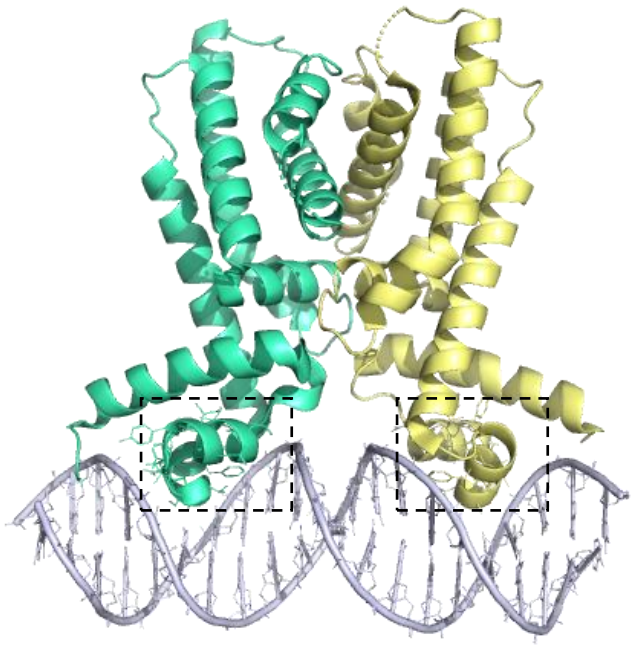

Figure 3: The C-terminal domain of LuxR bound to DNA. The helix-turn-helix DNA-binding motif that putatively mediates DNA-binding in *E. coli* and *S.* Typhimurium CsgD is highlighted within the boxed regions. PDB ID (7AMT).

In contrast to the LuxR system, CsgD is believed to be under the control of a two-component system, however its cognate histidine kinase has yet to be identified. [38] Understanding the basis for activation in a response regulator is key for determining its mechanism of DNA-binding, which dictates the potential avenues for inhibition. Studies of *S.* Typhimurium CsgD mutants featuring amino acid substitution at the conserved site of phosphorylation (D59) indicate

that CsgD exerts its DNA-binding activities in the unphosphorylated form throughout biofilm development. [40] Interestingly, the structure of *E. coli* CsgD receiver domain shows structural similarity to the receiver domain of *Vibrio Cholerae*, whose sequential activation, dimerization, and DNA-binding activity are stimulated by the secondary messenger c-di-GMP. [41] Crystal structure of the *S.* Typhimurium N-terminal receiver domain of CsgD as well as Congo red assays of point mutations, which assesses biofilm formation and features in vivo, reveal two dimerization interfaces, one N-terminal and one C-terminal, both of which appear to be crucial to CsgD function.[38] DNase I footprinting in addition to electrophoretic mobility shift assays unveiled a *S.* Typhimurium CsgD consensus target sequence 5′-AAAAGNGNNAAAWW and identified direct regulation of both positive and negative biofilm determinants. [40, 42] Among those identified was the flagella-related protein FliF, the diguanylate cyclase AdrA, and the csgBA operon encoding curli structural components CsgA and CsgB. The *E. coli* form of CsgD shows an altered consensus targets sequence, 5′-CGGGKGAGNKA, with promoter recognition regions that overlap in part with the *S.* Tyhpimurium CsgD promoter recognition regions. [42] The differences in CsgD recognition box may arise from the differences in primary sequence, with the two sequences differing in 17 amino acids within the N-terminus at 92% identity and 95% similarity.

1.8 **Curli Expression and Chaperone-Dependence**

The action of several chaperone-class proteins are critical for the efficient production and localization of curli oligomers and fibrils. [36, 43] It has been shown that CsgC is a selective

and highly effective chaperone protein that functions to keep CsgA soluble within the cytoplasm, while CsgE and CsgF exert chaperone-like activities. [45] CsgE forms a periplasmically-localized cap on the pore protein CsgG and directs CsgA to the pore, while CsgF is extracellularly located and tethers the pore to surface-assembled CsgB. In vitro experiments have revealed potent inhibition of CsgA amyloid formation by the molecular chaperones DnaK, DnaJ, and Spy proteins. [45] Interestingly, DnaK had an appreciable effect on the lag time of CsgA polymerization yet did not alter the time from rapid growth to plateau. [45] The HSP70 system also plays a multifaceted role in coordinating biofilm expression. [46] DnaK mediates *csg* gene regulation through modulation of the transcriptional regulator RpoS, participates in the folding of nascent CsgD, and binds to the targeting signal sequence of CsgA and CsgB to stabilize the translocation-competent forms of each amyloid. [46]

Section 2. *E. coli* CsgD Purification and 2-AI Binding Studies

Introduction

The increased prevalence of Enterobacteriaceae in individuals afflicted with neurodegenerative disorders like Parkinson's and Alzheimer's Disorders, as well as in vitro and in vivo evidence demonstrating cross-seeding activity between bacterially-sourced biofilm components and endogenous amyloid proteins has implicated curli-containing biofilms in the onset and progression of neurodegenerative disorders. [18, 21, 30] Curli production in enteric bacteria like *S.* Typhimurium and *E. coli* is orchestrated by the master regulator CsgD – a putative LuxR-like response regulator unique to prokaryotes. [37, 38] CsgD is believed to participate in a two-component system with its C-terminal DNA-binding activity under regulation by interactions between its N-terminal receiver domain and a membrane-associated histidine kinase; however, the cognate histidine kinase has yet to be identified. [38]

CsgD's identity as a bacterial response regulator makes it an ideal target for host-inert neurodegenerative therapeutics. Previous work by Cavanagh et al. identified a compound produced by marine sponges in the family *Agelasidae* known as Ageliferin, which contains a bio-active 2-aminoimidazole (2-AI) pharmacore. [47] Small molecule derivatives containing the 2-AI motif have been demonstrated to inhibit bacterial biofilm production in planktonic-state bacterium, as well as disperse established biofilms across a number of species. [47] Pull-down and binding assays, as well as molecular docking models have provided a molecular mechanism for 2-AI binding and evidence for specificity to response regulators – with the 2-AI derivatives preferentially making contacts with both the N and C-termini across the inter-domain interface. [48] The specificity of these compounds for response regulators invited investigation of CsgD-2-

AI interactions and repurposing of these previously anti-biofilm compounds for neurodegenerative therapeutics.

Previous work with the response regulators BfmR and QseB have indicated that tailoring of the structural motif attached to the 2-AI pharmacore may promote their interaction with the inter-domain interface and improve specificity. [49] Due to the interaction of 2-AI molecules with both the N- and C-termini, full length structural determination is critical for the design of highly-specific 2-AI molecules for inhibitory purposes. Soluble, recombinant protein amenable to concentration greatly aids these investigations – since nuclear magnetic resonance (NMR) studies and crystallographic techniques typically require high concentrations.

Initial purification efforts began with a growth protocol optimized for the response-regulator BfmR. This approach involves a standard growth temperature of 37 °C, followed by temperature reduction and induction of protein expression with Isopropyl β-d-1-thiogalactopyranoside (IPTG). These efforts were followed by adjustments to the growth protocol, including depression of growth temperature to thermodynamically restrict bacterial growth and promote correct-protein folding by reduction of bacterial stress. Additionally, multiple *E. coli* expression lines including ArcticExpress(DE3), RIPL(DE3), BL21(DE3), and HMS174(DE3) were investigated for their ability to mitigate inclusion body formation.

In an alternative approach to solubilize *E. coli* CsgD, denoted *Ec* CsgD henceforth, localized in bacterial inclusion bodies following induction, a mild-solubilization with the anionic surfactant N-lauroyl-sarcosinate (sarkosyl) was employed. Since solubilization involves working with detergents above their critical micelle concentration, a modified purification protocol

involving a column-purification step with a polystyrene-derived resin for detergent removal was also pursued.

In efforts to improve expression-level and solubility, a small ubiquitin-like modifier (SUMO) fusion construct was developed and subjected to purification. The fusion of the SUMO tag to the recombinant protein has been evidenced to drastically improve soluble expression level and folding of the recombinant protein through conferred stability and nucleation by the SUMO tag. [50] Primer design for the amplification of the *Ec* CsgD gene insert included 5′and 3′ BsaI restriction sites complementary to the overhangs in the pSUMO vector following digestion by BsaI. This allows for insertion of the *Ec* CsgD gene insert immediately following the final codon of the N-terminal 6 x histidine containing SUMO tag.

Evidence implicating DnaK as the primary chaperone responsible for promoting correct-folding of *Ec* CsgD in vivo invited growth supplementation with the pKJE7 plasmid encoding heat-shock chaperone proteins DnaK, DnaJ, and GrpE to both *Ec* CsgD and SUMO-CsgD recombinant expression. [36] Adenosine-triphosphate (ATP)-bound DnaK encounters nascent peptide, as well peptides being actively transcribed on the ribosome with low substrate affinity, stimulation by its J-domain protein DnaJ promotes ATP hydrolysis and conformational change that increases affinity for peptide substrate. [51] The nucleotide exchange factor GrpE facilitates ADP-ATP exchange and return to the low substrate affinity conformation, releasing correctly folded protein. [51] Removal of the chaperone-complex was achieved by incubation of clarified lysate with magnesium chloride, ATP, and soluble, denatured *E. coli* proteins prior to affinity chromatography.

Difficulties in obtaining, soluble, correctly-folded *Ec* CsgD prompted redirection of efforts towards the individual domains. A protocol described for purification of *S.* Typhimurium full length CsgD and the N-terminal domain was applied to the N-terminal domain of *E. coli* CsgD. Following purification, the N-terminal domain was screened with a selection of 2-AI molecules.

Each purification approach utilized immobilized nickel affinity chromatography in a one-step purification method. Binding of 2-AI molecules to the N-terminal domain, as well as qualitative evaluation of protein folding post-purification were evaluated through a thermal shift assay. The goal of this study was to investigate 2-AI binding to the *E. coli* response regulator CsgD and determine whether this response regulator is a potential target for neurodegenerative therapeutics.

Methods

Plasmid Maintenance

Plasmids encoding DNA for the protein of interest were transformed into either *E. coli* K-12 derivative Top10s or NovaBlue maintenance strains by electroporation and heat shock respectively. Single colonies were selected from antibiotic-containing Luria Broth (LB) agar plates and grown at 37°C, 225 rpm to an OD_{600} of 1.0-1.2. Plasmid isolation was achieved with the usage of QIAprep Spin (QIAGEN) and ZymoPure (Zymo Research) plasmid isolation reagents following manufacturer's centrifugal protocol.

E. coli CsgD Full Length, N-terminus, C-terminus Growth and Expression

pET 28a plasmids containing 6 x histidine tagged, *E. coli* master regulator protein *Ec* CsgD full length, C-terminus, and N-terminus were obtained from GenScript (Piscataway, NJ). Plasmids were transformed into chemically competent *E. coli* BL21(DE3) host cells for expression. In some cases, as indicated, alternative expression systems were used. Cells from an overnight liquid culture were used to inoculate 1 L cultures of LB medium supplemented with kanamyacin (50 mg/mL) to a starting OD_{600} of 0.02. Cultures were grown to an OD_{600} of 0.6-0.8 at 37°C, 160 rpm and induced with 1 mM isopropylthio-β-galactoside. Following induction, cells were incubated at 18°C, 120 rpm and harvested at 6,000 rpm. Cell paste was stored at -80°C.

Generation of SUMO-CsgD DNA construct

Primers as detailed in Table 1 corresponding to the N-terminus and C-terminus of the pET28a CsgD plasmid insert with engineered BsaI restriction sites were purchased from Integrated DNA Technologies (IDT). Reaction mixtures containing dNTPs, primers, CsgD-

containing pET28a, and commercially purchased LongAmp (New England BioLabs) were thermocycled on a T-100 Thermo-cycler (BioRad) according to the LongAmp PCR protocol provided by New England BioLabs. Verification of insert amplification was performed via DNA agarose gel electrophoresis. Removal of reaction components excepting the amplified insert was performed via QIAspin Prep PCR Cleanup (QIAGEN) according to manufacturer's centrifugal protocol. pSUMO vector containing a BsaI restriction site immediately following the upstream 6 x histidine-SUMO tag, as well as BsaI-Hfv2 (New England BioLabs) was incubated for 15 minutes at 37°C. The digestion products were analyzed by agarose gel electrophoresis and isolated via QIAquick (QIAGEN) gel-extraction reagents and centrifugal protocol. Digested pSUMO vector and amplified CsgD insert were incubated overnight with DNA Ligase (New England BioLabs). Ligation products were heat inactivated at 65°C, transformed into NovaBlue maintenance strain and plated on kanamycin (50 mg/mL) containing LB-agar plates. Single colonies were isolated and sequenced by Eton Bioscience (Research Triangle, NC).

Initial Purification of Full Length E. coli CsgD

Cell paste containing overexpressed *Ec* CsgD was resuspended in 25 mM Tris pH 8.0, 400 mM NaCl and lysed via sonication on ice with 20 s pulse intervals at 40% amplitude, followed by 59 s rest for a total of 30 minutes then spun down at 12,000 x g. Supernatant of the clarified lysate was loaded onto a glass column packed with Ni-NTA (Invitrogen) resin equilibrated with the resuspension buffer. The column was washed with additional resuspension buffer followed by 25 mM Tris pH 8.0, 1 M NaCl, prior to elution with 25 mM Tris pH 8.0, 400

mM NaCl, 250 mM Imidazole in a linear gradient elution. Purification results were analyzed by SDS-PAGE.

Sarkosyl Purification of E. coli CsgD

Resuspended cell pellet in 50 mM Tris pH 7.5, 300 mM NaCl, 5 mM Magnesium chloride ($MgCl_2$), 5 mM 2-mercaptoethanol (BME), and 5% w/v sarkosyl and was incubated overnight with stirring at 4°C. Following incubation overnight, solution was diluted to 200 mL with 25 mM Tris pH 8.0, 300 mM NaCl and supplemented to 1% TritonX-100 and 10 mM CHAPS. The solution was sonicated on ice with 20 s pulse intervals at 40% amplitude, followed by 59 s rest for a total of 30 minutes then spun down at 12,000 x g. The supernatant of the clarified lysate was loaded onto a glass column packed with Ni-NTA (Invitrogen) resin equilibrated with diluted resuspension buffer. The column was washed with 25 mM Tris pH 8.0, 300 mM NaCl, 10 mM imidazole, 0.5% sarkosyl followed by 25 mM Tris pH 8.0, 1M NaCl, 10 mM imidazole, 0.25% sarkosyl. SUMO-CsgD was then eluted with 25 mM Tris pH 8.0, 300 mM NaCl, 250 mM imidazole in a linear gradient elution. For complete removal of detergent molecules in specified instances, purification eluants were flowed over a glass column packed with BioBeads-Sm2 (BioRad) polystyrene resin. Flow through was pooled and concentrated via centrifugation in a Centricon 10K MWCO centrifugal filter (Fisher Scientific) and analyzed via thermal shift assay.

SUMO-CsgD Purification

Cell paste was resuspended in 25 mM Tris pH 8.0, 200 mM NaCl prior to sonication on ice with 20 s pulse intervals at 40% amplitude, followed by 59 s rest for a total of 30 minutes then spun down at 12,000 x g. The supernatant was loaded onto an equilibrated Ni-NTA (Invitrogen) column. The column was washed with 25 mM Tris pH 8.0, 200 mM NaCl followed by 25 mM Tris pH 8.0, 1 M NaCl, followed by another wash with 25 mM Tris pH 8.0, 200 mM NaCl. The sample was eluted with 25 mM Tris pH 8.0, 250 mM imidazole, 200 mM NaCl in a linear gradient elution. Purification results were analyzed by SDS-PAGE.

Preparation of Soluble Denatured E. coli Proteins

LB-media was inoculated with *E. coli* BL21(DE3) containing no antibiotic resistant plasmid and grown to an OD_{600} of 1.0-1.2. The growth contents were subjected centrifugation at 6,000 rpm in an Avanti JXN-26 (Beckman Coulter) centrifuge to separate media from the bacteria. The supernatant was disposed of and the bacterial cell paste was resuspended in 20 mM HEPES pH 8.0, 150 mM NaCl. And sonicated on ice with 20 s pulse intervals at 40% amplitude, followed by 59 s rest for a total of 30 minutes. The whole-cell lysate was clarified by centrifugation at 12,000 x g. The supernatant of the clarified-lysate was diluted to ~ 2 mg/mL with 20 mM HEPES pH 8.0, 150 mM NaCl and placed in a water bath with stirring at 65 °C. Following clouding of the supernatant, heating was continued for an additional 10 minutes prior to centrifugation of the lysate to remove insoluble precipitates. The supernatant was retained and concentration estimated using a NanoDrop One^{C} (Fischer Scientific). The supernatant was diluted with 20 mM HEPES pH 8.0, 150 mM NaCl to 1 mg/mL and stored at -20 °C.

CsgD, KJE7 Purification

Cell paste was resuspended in 50 mL 50 mM sodium phosphate pH 7.9, 300 mM NaCl and sonicated with 20 s pulse intervals at 40% amplitude, followed by 59 s rest for a total of 30 minutes then spun down at 12,000 x g.The supernatant of the clarified lysate was incubated with 0.1 mg/mL denatured *E. coli* proteins, 5 mM ATP, 12.5 mM $MgCl_2$ for an hour with stirring prior to loading onto a glass column packed with Ni-NTA (Invitrogen) equilibrated with the resuspension buffer. The column was washed with 50 mM sodium phosphate pH 7.9, 300 mM NaCl followed by 50 mM sodium phosphate pH 7.9, 1M NaCl, followed by another wash with 50 mM sodium phosphate pH 7.9, 300 mM NaCl. The sample was eluted with 50 mM sodium phosphate pH 7.9, 300 mM NaCl, 250 mM imidazole in a linear gradient elution.

Purification of CsgDN

Ec CsgDN was purified according to a protocol adapted from Wen et al. [38] Bacterial cell paste containing overexpressed *Ec* CsgDN was resuspended in 50 mM Tris-HCl pH 8.0, 300 mM NaCl, and 10 mM $MgCl_2$ lysed by sonication on ice with 20 s pulse intervals at 40% amplitude, followed by 59 s rest for a total of 30 minutes then spun down at 12,000 x g. The supernatant of the clarified lysate was applied to a column packed with Ni-NTA (Invitrogen) resin equilibrated with the resuspension buffer. The sample was eluted with 50 mM Tris-HCl pH 8.0, 300 mM NaCl, 10 mM $MgCl_2$, 250 mM Imidazole in a linear gradient elution.

SDS-PAGE

SDS-PAGE analysis was performed on tris-glycine 14% acrylamide gels. Both 10 and 15-well combs were used with 20 µL and 10 µL loads respectively. Electrophoresis was

performed using a BioRad Power Pac 1000 at 40 mA per gel until dye front began to reached the bottom of the gel. Gels were stained with coommassie and destained with 30% methanol and 10% acetic acid. Gels were imaged on a GelDoc EZ Imager (BioRad).

Expression Test

At both pre-induction and harvest, 10 mL samples of bacterial growths were taken and centrifuged at 2,500 x g for 15 minutes. The supernatant was discarded and the pellet was resuspended in 1 mL of appropriate lysis buffer. Samples were sonicated on ice at 10% amplitude for four rounds of 15 s bursts. Following sonication, the "total" sample was taken. The lysate was clarified by centrifugation at 10,000 x g for 10 minutes. A sample of the supernatant was taken to represent the "soluble" fraction while the remainder of the supernatant was discarded. The remaining pellet was resuspended in 1 mL of appropriate lysis buffer and clarified by centrifugation at 10,000 x g for 10 minutes. The supernatant was discarded and the remaining pellet was resuspended in 1 mL of appropriate lysis buffer and a sample was taken to represent the "insoluble" fraction. The samples were analyzed by SDS-PAGE.

Data Analysis

Thermal Shift Assay Analysis

Raw spectra imported from QuantStudio 3 (Fisher Scientific) were imported to Protein Thermal Shift™ (Fisher Scientific). Regions of analysis, determined from the temperature-axis, were defined to maximize baseline-stability. The results were analyzed in the Protein Thermal Shift™ software by the Boltzmann-derived T_m method. This method provides for determination of the temperature at which the protein is in the 50% unfolded state via measurement of the fluorescent probe Sypro-Orange which binds hydrophobic regions of the unfolding protein. The Boltzmann method for determination of T_m utilizes the linear correlation between the fluorescence intensity and the fraction of unfolded protein as shown in equation 1.

(eq. 1) $$\%U(T) = F_{min} + \frac{F_{max} - F_{min}}{1 + e^{\frac{Tm - T}{a}}}$$

Where a is a parameter determined by the steepness of the unfolding response, T_m is the temperature at which 50% unfolding is achieved, F_{max} and F_{min} are the maximum and minimum in relative fluorescence for the unfolding curve. Datasets were gathered in triplicate for each reaction condition. Datasets containing outliers and those with distorted unfolding curves were excluded from T_m determination.

Table 1: Primers containing BsaI-Hfv2 restriction sites (red) for generation of SUMO-CsgD fusion product

CsgD-SUMO forward primer	5'-CCGGTCTCTAGGTATGTTTAATGAAGTACATTCAATACACGG
CsgD-SUMO reverse primer	5'-CCGGTCTCTCTAGTTATTAACGGCGCAGATTATCATTAGCCC

Results

E. coli CsgD Growth and Expression Testing

Initial efforts to produce 6 x histidine tagged *E. coli* CsgD utilized a growth protocol optimized for another bacterial response regulator BfmR. In brief, cells were grown at 37 °C to an OD_{600} of 0.6-0.8 prior to induction with 1 mM IPTG and temperature reduction to 18 °C. Throughout the post-induction incubation, aliquots of the growths were taken to assess recombinant protein expression level, solubility, and in vivo viability. Expression testing of these samples is designed to assess the potential yield of a standard nickel affinity purification and so mimics protein isolation from sonicated whole-cell extract. Samples corresponding to the contents of the whole cell lysate, the clarified supernatant, and the insoluble cellular remnants and inclusions were analyzed via SDS-PAGE. In addition to the BfmR-optimized growth protocol, adjustments to growth temperature in the effort to reduce bacterial stress throughout the post-induction window, as well as alternative expression systems. Analysis of growth conditions through expression tests are shown in figures 4 and 5. Of the approaches taken, the most common result was poor soluble expression and robust localization to inclusion bodies. However, expression in *E. coli* HMS174(DE3) noticeably improved the expression of soluble *Ec* CsgD.

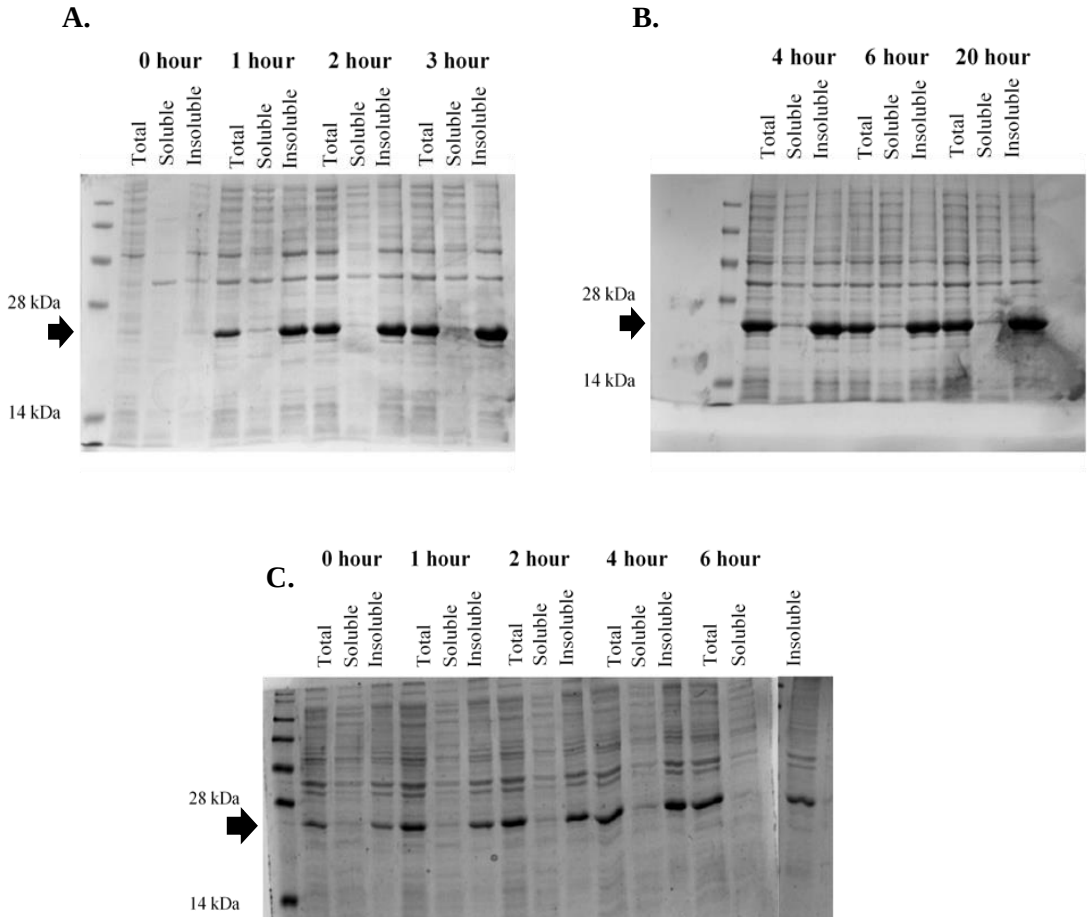

Figure 4: SDS-PAGE analysis of Expression Tests. Initial Expression testing of *Ec* CsgD is shown in A-B, while expression testing of *Ec* CsgD grown at 20 °C is shown in panel C. The protein of interest is indicated with an arrow.

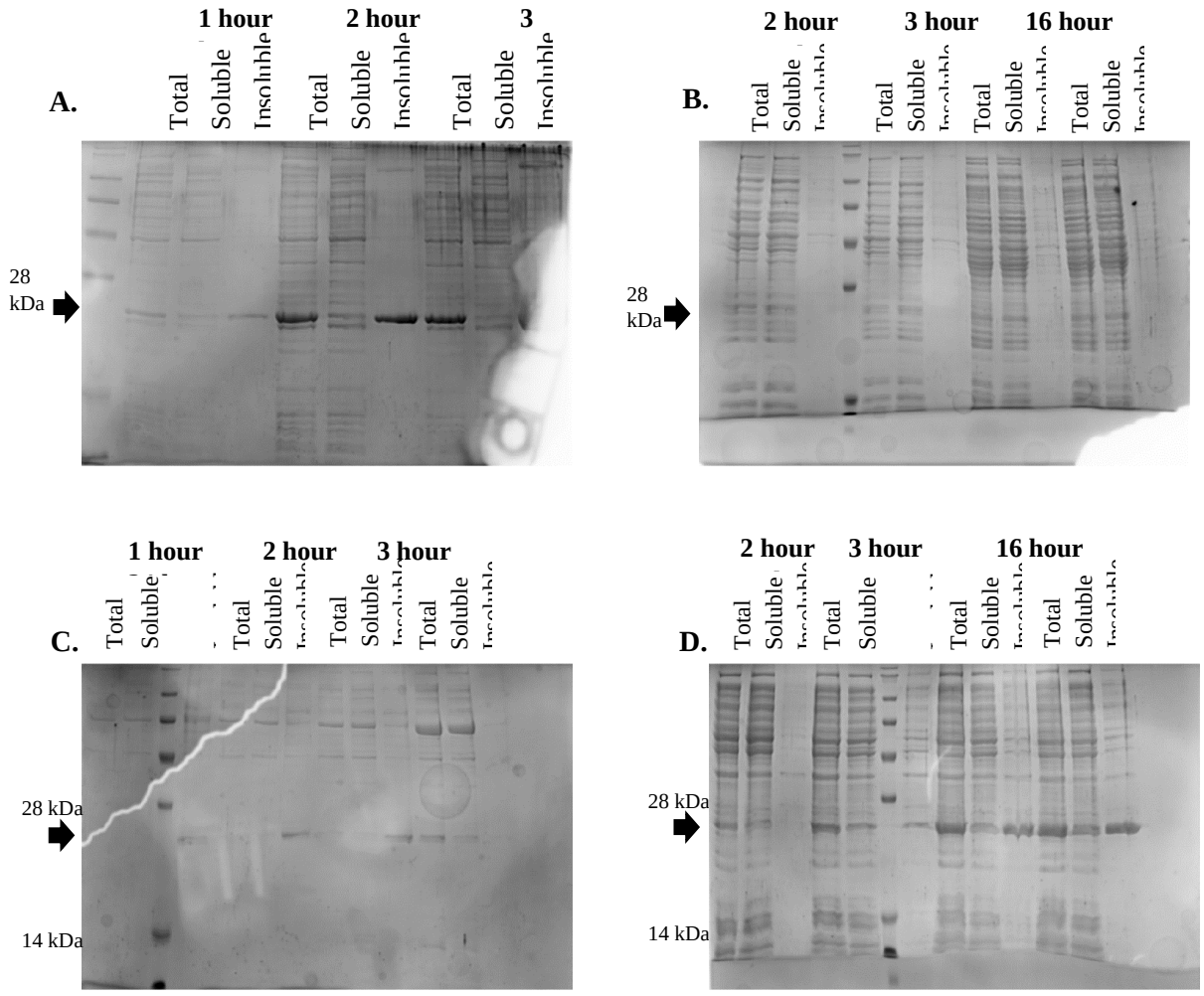

Figure 5: SDS-PAGE of expression test results in alternative *E. coli* expression strains is shown above. Panel A shows expression in RIPL(DE3). Panel B shows expression in BL21(DE3). Panel C shows expression in ArcticExpress(DE3). Panel D shows expression in HMS174(DE3). The presence and location of *Ec* CsgD is indicated with an arrow.

Standard Nickel Affinity Chromatography of E. coli CsgD grown in HMS174(DE3)

To isolate CsgD expressed within the soluble, cytoplasmic fraction, cell paste was resuspended and subjected to sonication. Following centrifugation, the supernatant was loaded to a nickel-affinity column packed with Ni-NTA resin. The sample was subjected to successive wash steps prior to elution with imidazole. Results of the purification were analyzed by SDS-PAGE. Results are shown in figure 6. Isolation of soluble CsgD was achieved at low concentrations, however the majority of CsgD remained localized to insoluble inclusion bodies. Loss of *Ec* CsgD in the load indicates some level of aggregation or obstruction of the 6x His tag.

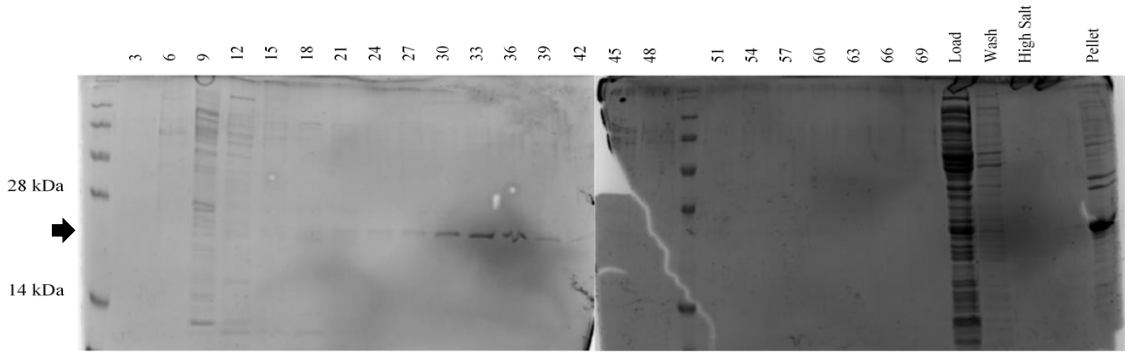

Figure 6: Purification results were analyzed by SDS-PAGE on tris-glycine polyacrylamide gels. Samples from each of the indicated fractions, as well as samples from the load flow through as well as the wash phases are shown above. Presence of the protein of interest is indicated with an arrow.

Thermal Shift Assay of E. coli CsgD grown in HMS174(DE3)

To determine thermostability and perform qualitative assessment of protein folding, *Ec* CsgD purified from *E. coli* HMS174(DE3) was analyzed by thermally-induced unfolding in the presence of 8x Sypro-Orange. Fluorescence spectra of technical triplicates are shown in figure 7. In contrast to the sigmoidal unfolding curve typical of folded proteins, the result below indicated unfolded protein. This result, in addition to the low expression post-induction and low yield of the purification process, encouraged us to pursue alternative methods of expression and purification.

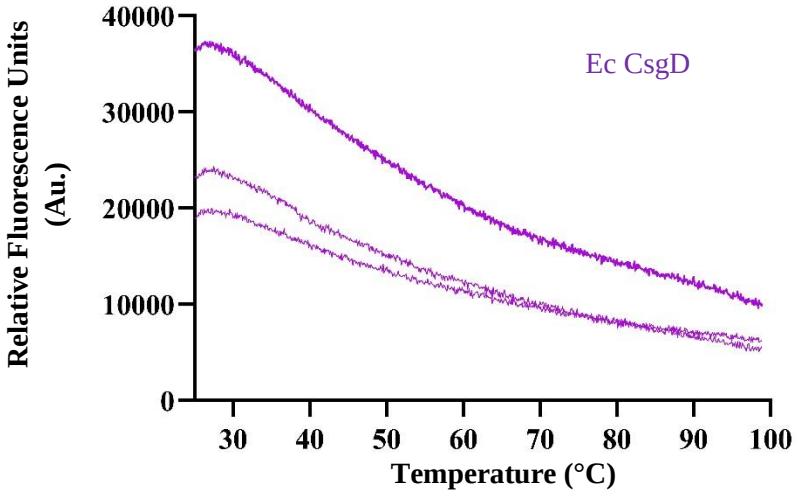

Figure 7: *Ec* CsgD and Sypro-Orange buffered in 25 mM Tris pH 8.0, 400 mM NaCl were subjected to a thermal ramp from 25 – 98 °C. Fluorescence responses for each of the replicate runs are shown above.

Sarkosyl Solubilization of E. coli CsgD from Inclusion Bodies

In an alternative purification approach that capitalizes on robust insoluble expression, the desired bacterial cell paste was resuspended in a lysis buffer containing 5% sarkosyl and incubated overnight to solubilize inclusion bodies. Following solubilization, the solution was subjected to standard nickel affinity chromatography and sarkosyl was removed through centrifugal buffer exchange. SDS-PAGE of purification fractions are shown in figure 8. While a large portion of *Ec* CsgD remained localized in inclusion bodies, as well as protein loss in both the load and wash phases, isolation of soluble *Ec* CsgD improved drastically.

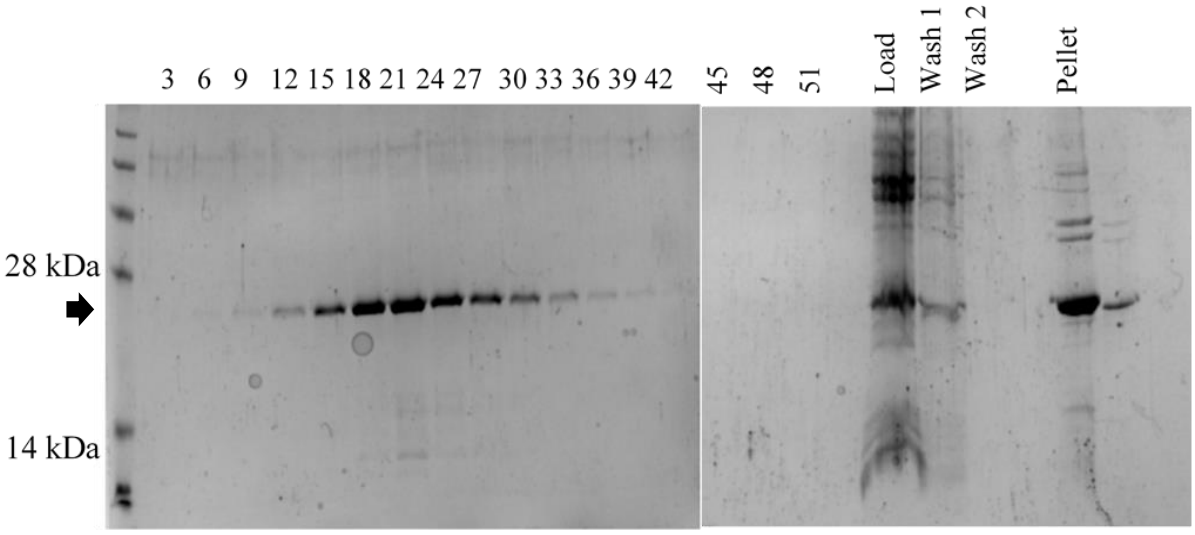

Figure 8: Results of the sarkosyl-mediated purification of *Ec* CsgD were analyzed on a denaturing tris-glycine polyacrylamide gel. Samples from each of the indicated elution fractions, as well as samples from the load flow through, wash flow throughs, and the insoluble cell remnants were taken and shown above. Presence of protein of interest is indicated with an arrow.

Thermal Shift Assay of Sarkosyl-Purified E. coli CsgD

Thermal unfolding of sarkosyl-purified CsgD in the presence of 8x Sypro-Orange dye allowed for determination of thermostability and qualitative assessment of folding in the purified protein via fluorescence-measurement. Fluorescence spectra of technical triplicates are shown in figure 9. While the purification results indicate improved yield, thermal shift analysis of sarkosyl-purified *Ec* CsgD indicates lack of folding, making the isolate inapplicable for structural studies.

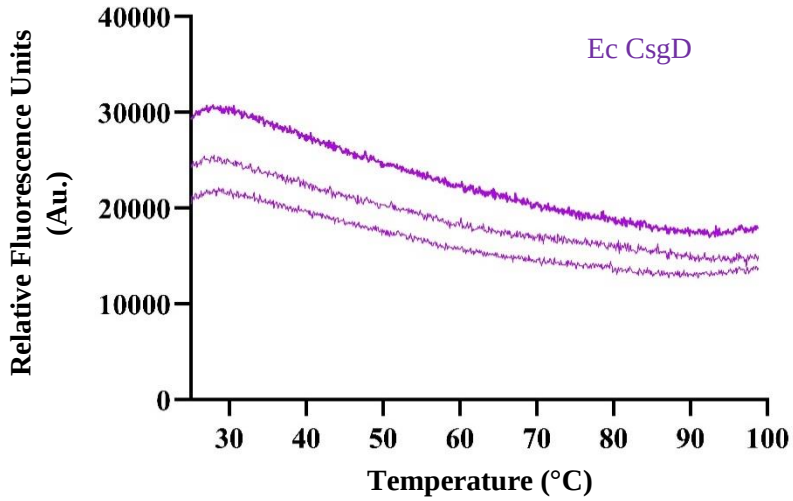

Figure 9: Sarkosyl-purified *Ec* CsgD and Sypro-Orange buffered in 25 mM Tris pH 8.0, 300 mM NaCl were subjected to a thermal ramp from 25 – 98 °C. Fluorescence responses for each of the replicate runs are shown above.

Bio-Beads Polystyrene Resin for Removal of Detergents

Oddly, bubbling of samples was noticed in eluants from the sarkosyl-based purification method, even in samples determined protein-deficient by SDS-PAGE. To ensure that residual detergent was effectively removed, eluants from sarkosyl-aided nickel affinity chromatography were passed over a column containing Bio-Beads polystyrene resin. The polystyrene resin adsorbs highly hydrophobic components including the acyl chains of detergents. Purification results were analyzed by SDS-PAGE and shown in figure 10. Results echo previous sarsokyl-based purifications of *Ec* CsgD with improved yield and some protein loss during the load and wash phases. Fractions containing *Ec* CsgD with high purity were pooled and concentrated prior to BioBead-SM2 treatment.

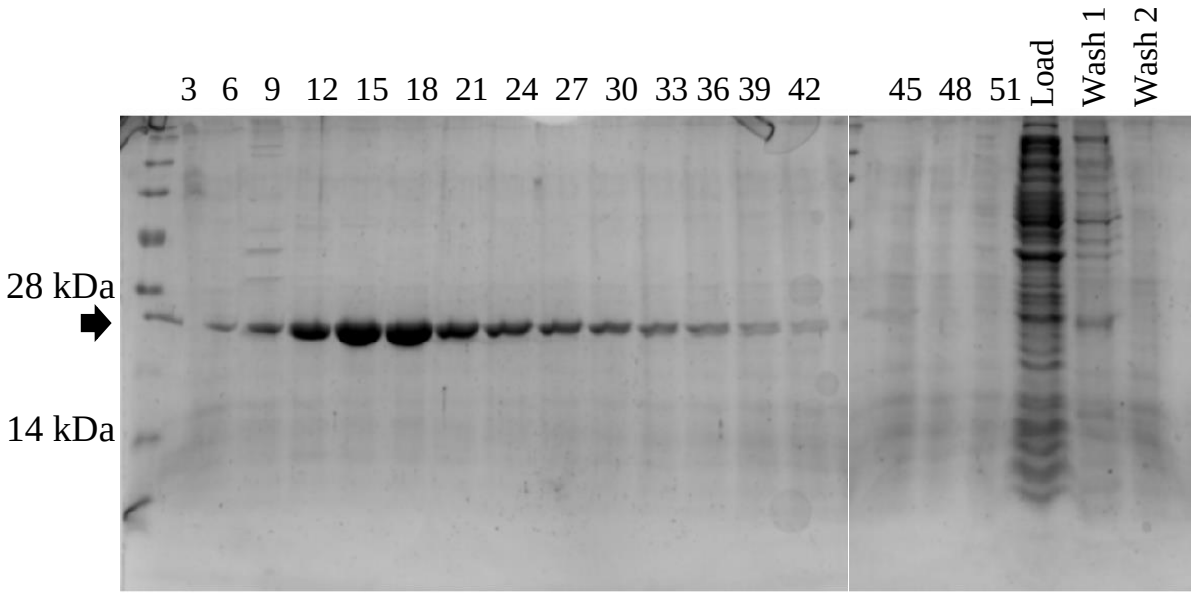

Figure 10: Treatment with BioBeads-SM2 necessitated repetition of the sarkosyl-mediated purification protocol prior to removal of residual detergents. SDS-PAGE on a tris-glycine polyacrylamide gel allowed for determination of the protein content in the indicated elution fractions, as well as the load and wash flow throughs. Protein of interest is indicated with an arrow.

Thermal Shift Assay of Bio-Bead Treated E. coli CsgD

Thermostability and folding of purified *E. coli* CsgD in the absence of residual detergents was analyzed by thermally-induced unfolding in the presence of 8x Sypro-Orange. Fluorescence spectra of technical triplicates are shown below in figure 11. As opposed to the continuously decreasing unfolding curves seen previously, BioBead-SM2 treated *Ec* CsgD exhibits a plateau at the beginning of the temperature ramp. However, these results were not indicative of a physiologically-relevant isolate and prompted alternative approaches.

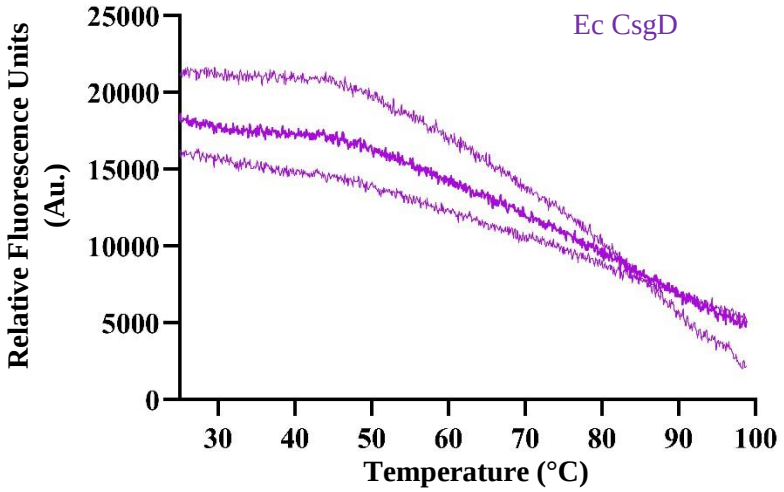

Figure 11: *Ec* CsgD treated by exposure to BioBeads-SM2 resin for removal of residual detergents, as well as Sypro-Orange were buffered in 25 mM Tris pH 8.0, 300 mM NaCl. Fluorescence responses of each replicate run were recorded across the temperature ramp from 25 – 98 °C and shown above.

Expression Test of SUMO-CsgD grown in BL21(DE3)

Yield and localization of the fusion construct were obtained via expression testing of IPTG-induced cultures and analyzed by SDS-PAGE. Results are shown in figure 12. The addition of the SUMO fusion partner to *Ec* CsgD resulted in robust, soluble expression of the fusion construct.

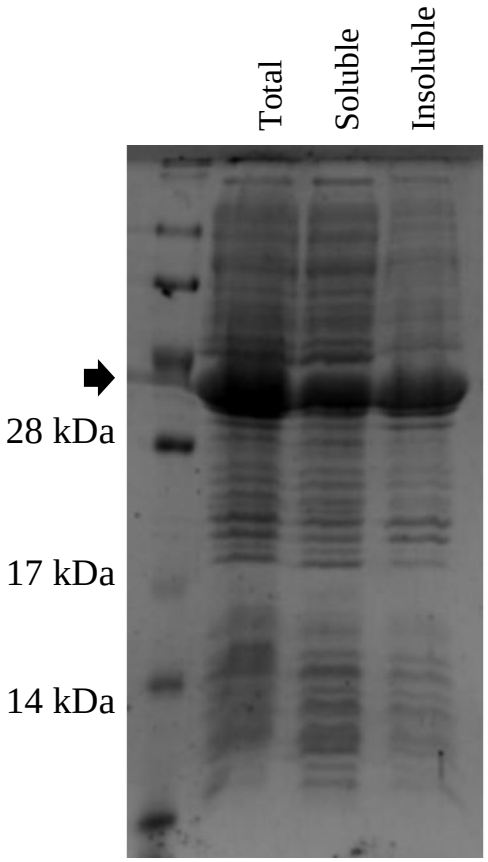

Figure 12: Results of the expression test were analyzed by SDS-PAGE on a tris-glycine polyacrylamide gel. Fractions corresponding to the whole-cell lysate (Total), the cytoplasmic fraction (Soluble), and inclusion bodies and cellular debris (Insoluble) are shown above. SUMO-CsgD is indicated with an arrow.

Purification of SUMO-CsgD Fusion Product

SUMO-CsgD was purified by a nickel affinity chromatography method previously utilized for initial *Ec* CsgD purification from HMS174(DE3) and results of the purification were analyzed by SDS-PAGE. Purification results are shown in figure 13. Despite robust expression, yield of the purification was considerably lower than expected. In similar fashion to previous approaches, loss of the protein of interest was noted in both the load flow through and wash flow through.

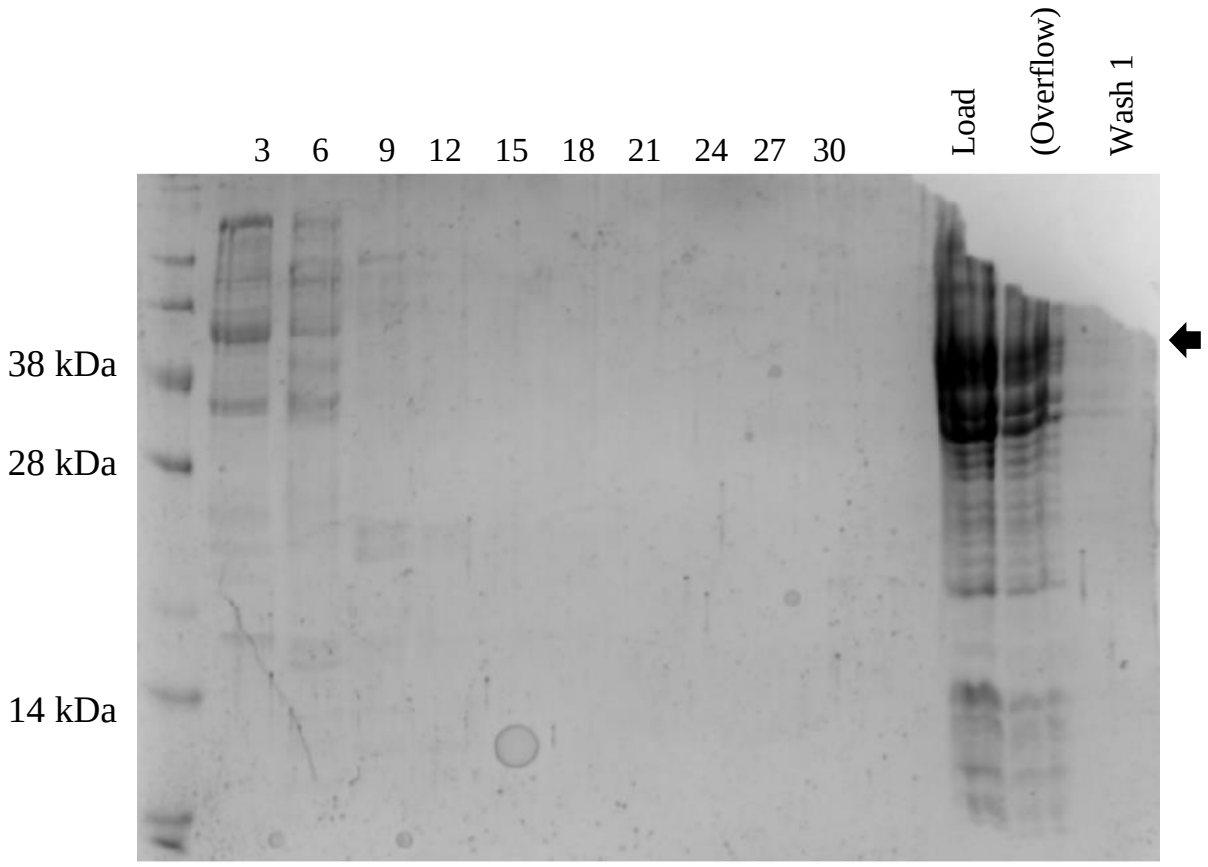

Figure 13: Attempted purification of SUMO-CsgD. Purification results were analyzed by SDS-PAGE on a tris-glycine polyacrylamide gel. Samples from the elution fractions indicated above, as well as the load and wash flow throughs are shown above. Location of the protein of interest is indicated with an arrow.

Growth Supplementation with DnaK, DnaJ, and GrpE

Given the poor yield of SUMO-CsgD purification, as well as difficulties in obtaining the correct protease, attention was turned to cellular chaperones for improvement of soluble *Ec* CsgD expression. Dual expression with molecular chaperones and its effect on the yield of soluble *Ec* CsgD was evaluated via expression testing. SDS-PAGE results of the expression test are shown in figure 14. Even in the growth supplemented with 3% EtOH, an alternative method for DnaK induction, upregulation of chaperone proteins was noted as well as moderate improvement in the expression of soluble CsgD.

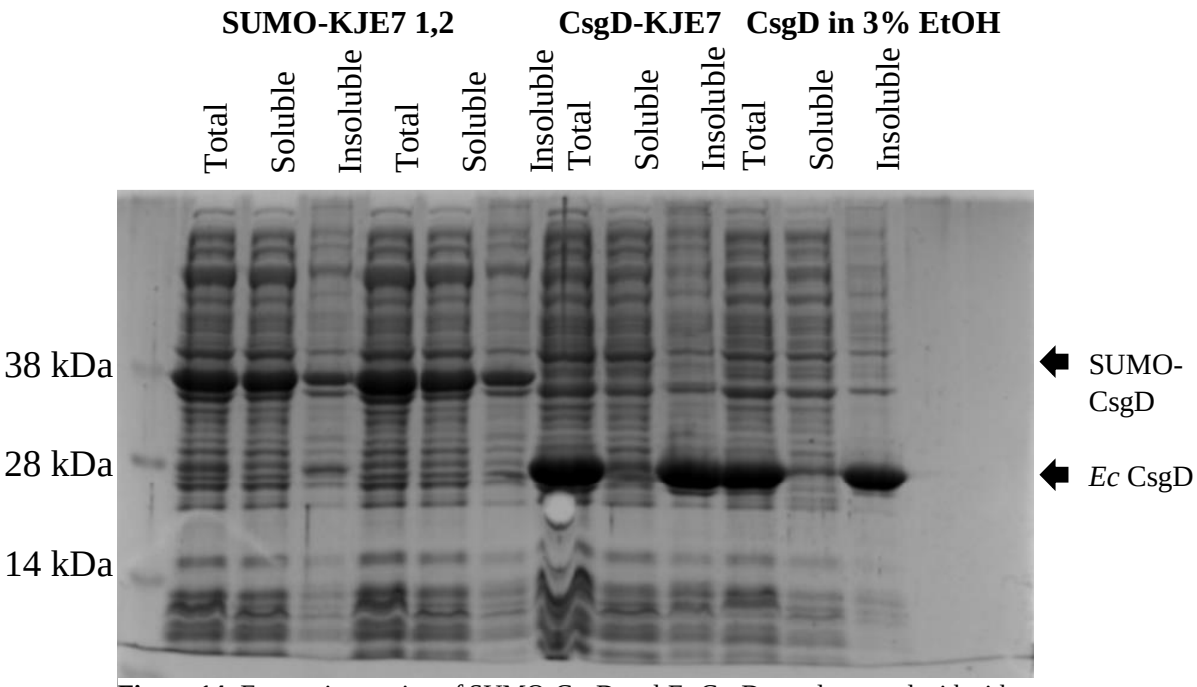

Figure 14: Expression testing of SUMO-CsgD and *Ec* CsgD supplemented with either pKJE7 plasmid or 3% ethanol were analyzed by SDS-PAGE on a tris-glycine polyacrylamide gel. The fractions corresponding to the whole-cell lysate (load), cytoplasmic fraction (soluble), and inclusion bodies and cellular debris (insoluble) are shown for a duplicate growth of SUMO-CsgD dual expression, as well as for *Ec* CsgD dual expression, and *Ec* CsgD supplemented with ethanol. Putative locations for SUMO-CsgD and *Ec* CsgD are indicated with arrows.

Purification of Chaperone Supplemented E. coli CsgD

Cell paste containing overexpressed Ec CsgD, as well as DnaK, DnaJ, and GrpE was resuspended in tris-buffered lysis solution. Following sonication, the supernatant was incubated with soluble, denatured *E. coli* proteins in the presence of $MgCl_2$ and ATP. This provides alternative substrate for chaperone complexes and promotes release of folded *Ec* CsgD. Following incubation, the sample was subjected to multiple wash steps in a nickel-affinity purification protocol. The results of the purification were analyzed by SDS-PAGE and are shown in figure 15. Purification fractions show residual contaminants in early fractions, as well as low concentration of protein of interest. Protein loss is observed in both the load and wash flow through. In addition, protein was noted to elute lower in the linear gradient than for *Ec* CsgD alone.

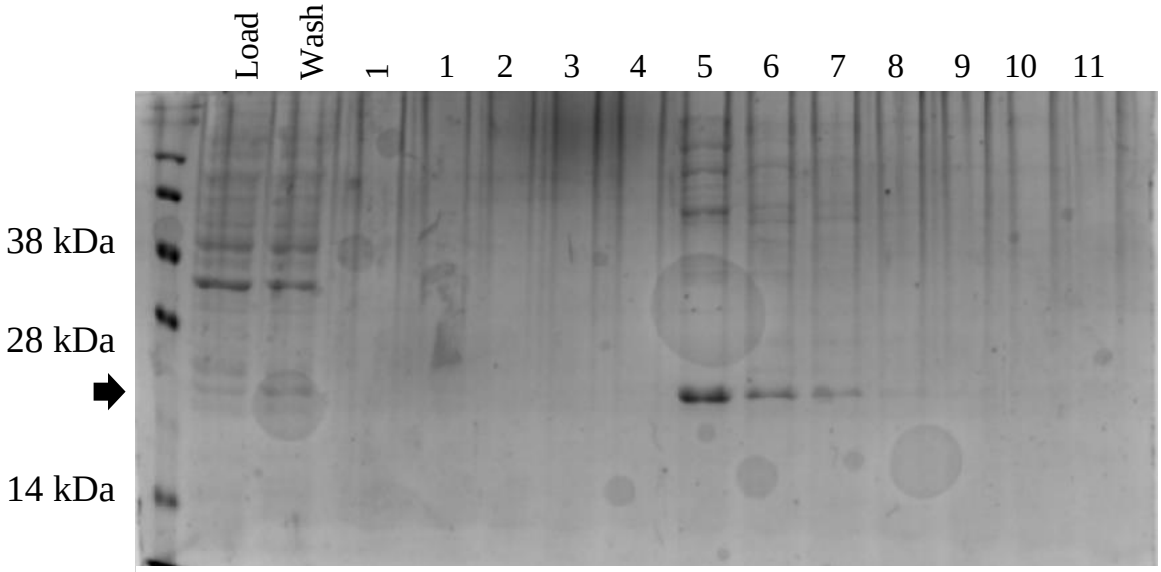

Figure 15: Purification results of *Ec* CsgD supplemented with pKJE7 were analyzed by SDS-PAGE on a tris-glycine polyacrylamide gel. Samples corresponding to the indicated elution fractions, as well as the load and wash flow throughs are shown above. The location of *Ec* CsgD is indicated with an arrow.

Thermal Shift Assay of CsgD supplemented with DnaK, DnaJ, and GrpE

Following purification and bulk removal of molecular chaperones, fractions containing *Ec* CsgD with minimal contamination were incubated with 8x sypro-orange dye and analyzed via thermal unfolding along the indicated temperature range. The spectra for technical triplicates, are shown below in figure 16. Results present continuously decreasing unfolding curves as seen for previous purification approaches, excepting the BioBead-SM2 treated isolate. Despite the appearance of what resembles a peak between 60-70 °C, this may be the product of residual contaminant and overall unfolding curves indicate protein isolate that is not physiologically relevant.

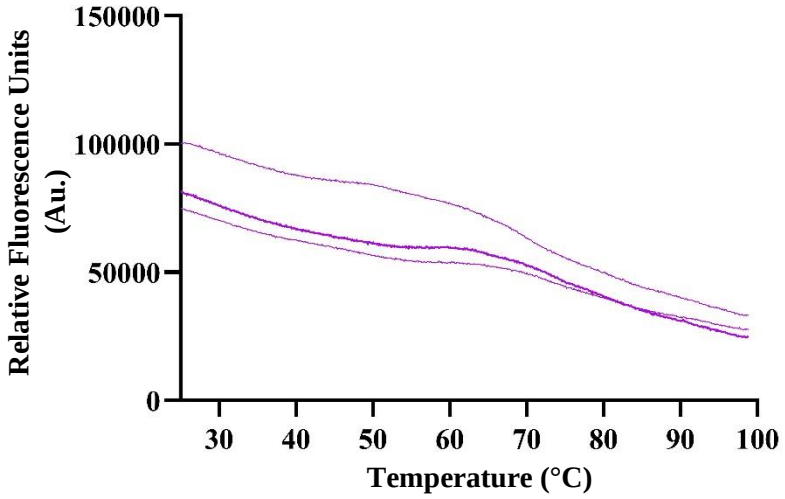

Figure 16: *Ec* CsgD treated with exposure to BioBead-SM2 resin, as well as Sypro-Orange were buffered in 25 mM Tris pH 8.0, 300 mM NaCl. Fluorescence responses over the temperature ramp 25 – 98 °C for each replicate are shown above.

Purification of Chaperone Supplemented SUMO-CsgD

In an effort to maximize the solubility of recombinant protein and promote purification yield, supplementation with DnaK, DnaJ, and GrpE was also applied to the fusion construct SUMO-CsgD. The same purification approach utilized for the dual expression of *E. coli* CsgD was applied to the fusion construct. Purification results were analyzed by SDS-PAGE and results are shown in figure 17. Results indicated incomplete chaperone removal and show protein loss in the load and wash, with little improvement to soluble yield of the protein of interest.

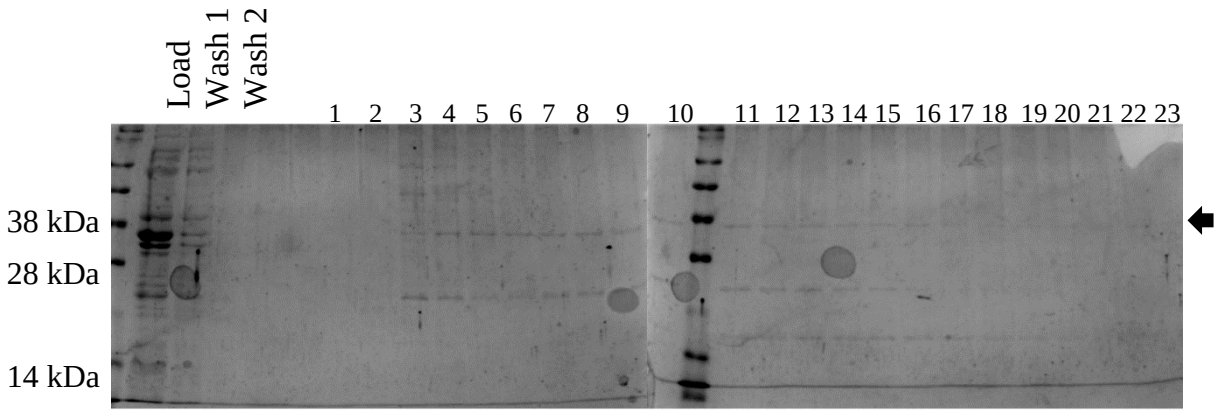

Figure 17: Purification results of SUMO-CsgD supplemented with pKJE7 were analyzed by SDS-PAGE on tris-glycine polyacrylamide gels. Samples from the indicated elution fractions, as well as the load and wash phase flow throughs are recorded above. The location of SUMO-CsgD is indicated with an arrow.

Expression Testing E. coli CsgDN

Following difficulties in expressing full length *Ec* CsgD, the solubility and expression level of the N-terminal domain, CsgDN, were evaluated through expression testing and SDS-PAGE analysis is shown in figure 18. Results indicate that the N-terminal domain exhibits greater propensity for soluble expression than the full-length protein.

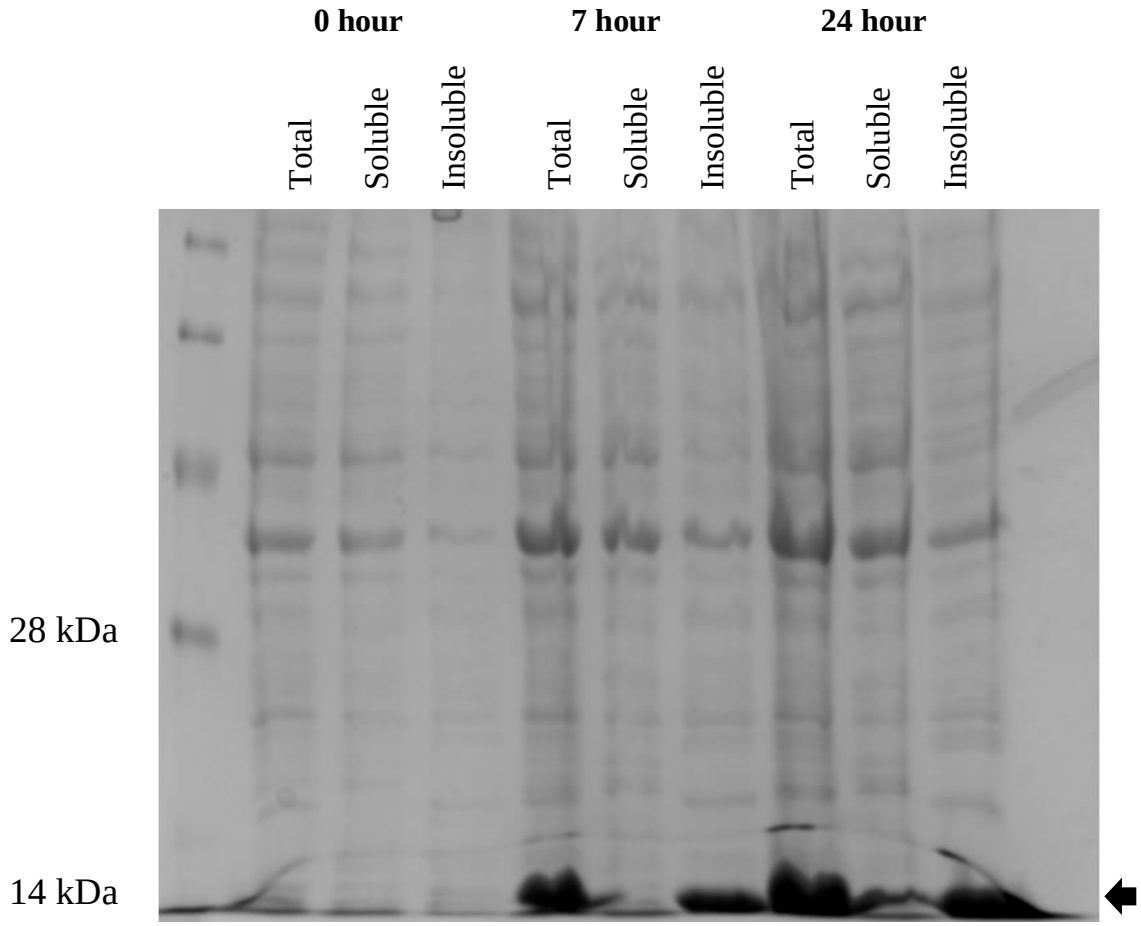

Figure 18: Expression test results were analyzed by SDS-PAGE on a tris-glycine polyacrylamide gel. Fractions corresponding to the whole-cell lysate (total), the cytoplasmic fraction (soluble), and the cellular debris and inclusion bodies (insoluble) are shown above. The location of *Ec* CsgDN is indicated with an arrow.

Purification of CsgDN

Standard nickel-affinity chromatography was conducted following resuspension of cell paste containing overexpressed CsgDN and sonication. SDS-PAGE analysis of purification results are shown in figure 19. Results indicate successful isolation of CsgDN with high purity, albeit at low concentrations for structural determination.

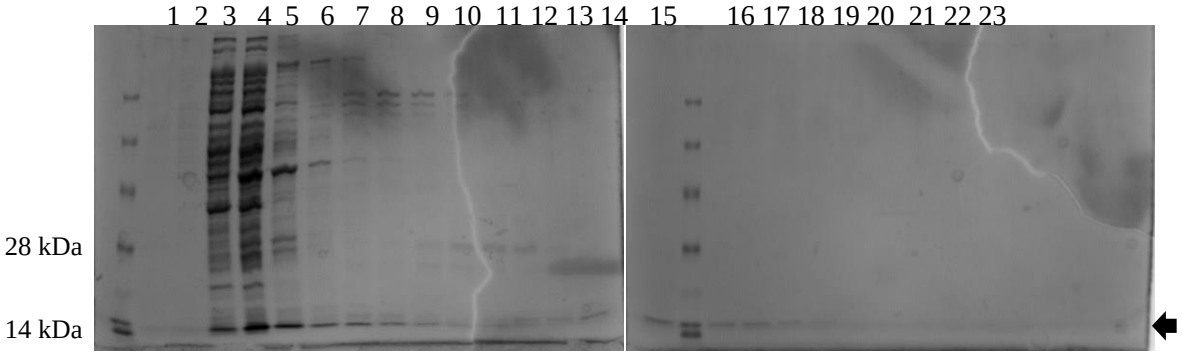

Figure 19: Purification results were analyzed by SDS-PAGE on a tris-glycine polyacrylamide gel. Samples were taken from the elution fractions indicated above.

Thermal Shift Assay and 2-AI Screen of E. coli CsgDN

Thermostability of the N-terminal receiver domain of *E. coli* CsgD was probed via thermally-induced unfolding in the presence of 8x Sypro-Orange dye. To screen the ability of 2-aminoimidazole (2-AI) compounds to bind the receiver domain, 2x CsgDN reaction mixes were supplemented to the indicated concentration of 2-AI molecule. The binding effect was interpreted through shifts in the T_m. Fluorescence spectra of technical replicates runs along with analysis of ΔT_m for each compound tested are shown in figure 20. Of the 2-AI molecules screened, comparison of *Ec* CsgDN + AGL-869 to the DMSO control was indicative of 2-AI binding. Unfolding curves for samples supplemented with AGL-600 and AGL-726 contained outliers, however remaining unfolding curves were not indicative of 2-AI binding. For both AGL-833 and AGL-949 supplemented samples, baseline and curve distortion were observed – making determination of T_m unreliable for these treated groups.

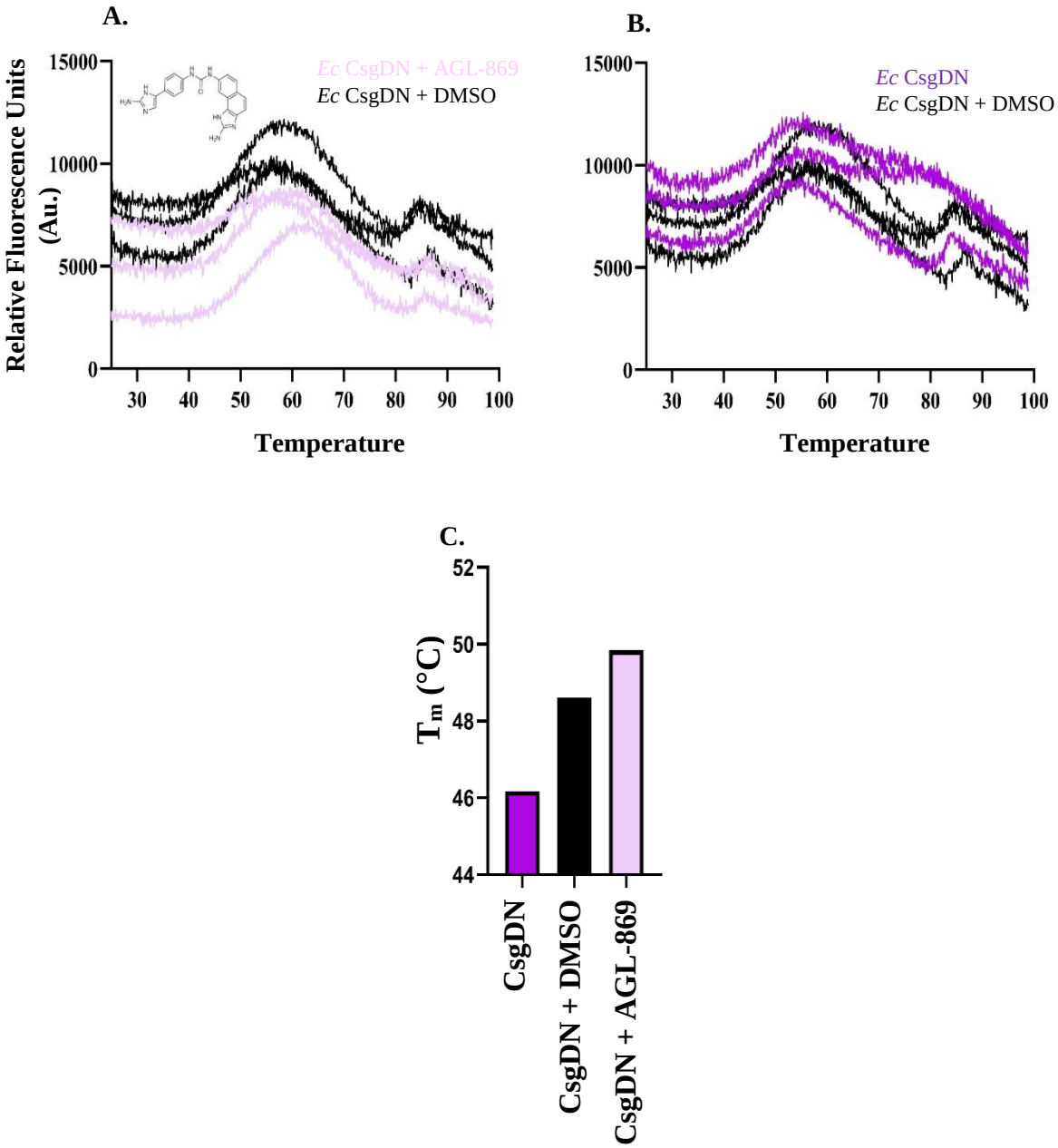

Figure 20: Thermal shift analysis of *Ec* CsgDN (5 µM) buffered in 25 mM Tris pH 8.0, 300 mM NaCl, 10% DMSO in the presence of 50 µM of 2-AI molecule and 8x Sypro-Orange. Fluorescence responses comparing unfolding curves for *Ec* CsgDN + DMSO and *Ec* CsgDN + AGL-833 are shown in panel A. Panel B shows unfolding curves of replicate runs for *Ec* CsgDN alone and in the presence of DMSO. Panel C reports the shifts in T_m determined via Boltzmann fitting in the Protein Thermal Shift Software.

Discussion

Traditional recombinant protein growth and expression methods optimized for BfmR, as well as adjustments to induction-length and temperature all resulted in robust *E. coli* CsgD expression localized to bacterial inclusion bodies. (Figure 4) Expression testing was pursued in multiple *E. coli* expression systems including RIPL(DE3), ArcticExpress(DE3), and HMS174(DE3), each designed to improve a specific aspect of the transformation to expression process. RIPL(DE3) are designed to improve incorporation and enhance translation of proteins rich in arginine, isoleucine, proline, and leucine, while ArcticExpres(DE3) express chaperonins Cpn60 and Cpn10, as well as grow at low temperatures – improving solubility and folding. The lack of improvement in soluble CsgD expression in RIPL(DE3) and ArcticExpress(DE3) imply that difficulties in obtaining soluble CsgD do not stem from insufficient bacterial tRNA or from lack of Cpn60 and Cpn10 chaperonin activity. (Figure 5) In the case of HMS174(DE3), which are designed to provide high translation efficiency and promote protection of the DE3 prophage, soluble expression of full length CsgD was moderately improved, with the majority of expressed protein residing in the soluble, cytoplasmic fraction. (Figure 5) This effect may stem from the improved translation efficiency and decreased the level of bacterial-stress induced by the transformation process. Purification via nickel affinity chromatography allows for coordination of the histidine residues within the 6 x His tag. The immobilized nickel ion coordinates the imidazole side chains of two histidine residues at the 1 position forming a stable six-membered coordination complex – retaining the sample within the resin of the column. Following elution with imidazole (Figure 6), the molecular analog of histidine side-chains, thermal shift analysis of

the purified protein revealed a thermal unfolding curve that was continuously decreasing and reached maximum relative fluorescence at the outset of the temperature ramp. (Figure 7) For a folded protein, sypro fluorescence is typically low and constant early in the temperature ramp followed by a positive sigmoidal fluorescence response to the linear increase in temperature. As the protein unfolds, sypro is able to access hydrophobic pockets within the protein, reducing the incidence of nonradiative excited-state decay and fluorescence quenching by water molecules and resulting in an increase in fluorescence intensity. This indicates that at the outset, sypro was able to freely access hydrophobic regions of the unstructured protein. The continuous decrease in fluorescence intensity is likely a result of entropic effects at higher temperatures. This may indicate that cellular levels of endogenous chaperones did not meet the requirements of the overexpressed *Ec* CsgD, or that the isolate was not amenable to concentration via centrifugation.

The abundant expression of full-length *Ec* CsgD in bacterial inclusion bodies provided another avenue by which we could obtain soluble protein. Many methods exist to solubilize inclusion bodies, including treatment with high concentrations of Guanidinium Hydrochloride (GuHCl, GdnHCl), Urea, or SDS which typically result in complete protein unfolding. To avoid this, we utilized sarkosyl, which exhibits milder unfolding effects than SDS but through a similar mechanism - beginning with interactions between its anionic head group and the basic side chains of the protein prior to disruption of the hydrophobic core through its acyl tail. [52, 53] Soaking of bacterial cell paste containing overexpressed CsgD followed by purification via nickel affinity chromatography resulted in high yield purification of soluble CsgD. (Figure 8) However, probing of protein quality via thermal shift assay revealed a similar unfolding curve to that seen for *Ec* CsgD purified from HMS174(DE3). (Figure 9) CsgD may have been localized to inclusion bodies in an unfolded state, or is unable to refold correctly following removal of

sarkosyl. Interestingly, sarkosyl-purified protein eluants were noted to froth atypically during pipette transfer. This effect was noted even in purification fractions determined to be protein-deficient by SDS-PAGE and implicated residual detergent as the abnormality's source. To mitigate potential presence of residual detergents, purification eluants were run over a gravity column packed with polystyrene resin capable of adsorbing the acyl tails of the anionic surfactants. Following treatment, protein quality was probed via a thermal shift assay revealing a similar result. (Figure 11) However, the appearance of a plateau at the beginning of the temperature ramp may indicate the appearance of some localized structures; however their stability is compromised at low temperatures and not indicative of a physiologically-relevant isolate. These results imply the need for in vivo chaperone activity, however the possibilities that *Ec* CsgD was localized to inclusion bodies in an unfolded state as well as the possibility that *Ec* CsgD is not amenable to concentration via centrifugation remain. These approaches indicated the need for both improved folding and expression-level of the protein of interest.

Affinity and fusion tags including maltose-binding-protein, glutathoione-s-transferase, and small ubiquitin-like modifier (SUMO), have been shown to exhibit intrinsic chaperone-like activity. [54] The fusion tags typically express robustly and rapidly adopt their natively-folded structure, this confers stability and solubility to its fusion partner through interaction between hydrophobic patches of the tag and protein of interest - reducing self-aggregation. We chose to generate an N-terminally 6x histidine tagged SUMO-CsgD fusion construct. The SUMO (Smt3) protein sourced from yeast confers several advantages including enhanced expression, solubility, and folding. [54] Additionally, SUMO protease, the protease domain of Ulp1, recognizes the tertiary structure of SUMO and mitigates cleavage internal to the protein of interest. [50] Expression of SUMO-CsgD was achieved in high amount and solubility. (Figure 12) However,

protein loss during loading of the sample to Ni-NTA resin would indicate disruption of the histidine tag responsible for coordinating immobilized nickel ions. (Figure 13) In addition, the available Ulp1 smt3 protease domain expressed at an unexpected molecular weight and lacked protease activity. These factors encouraged redirection of efforts to the individual domain constructs, as well as growth supplementation with molecular chaperones DnaK, DnaJ, and GrpE.

Growth supplementation with DnaK, DnaJ, and GrpE afforded moderate increase in the soluble, cytoplasmic fraction of overexpressed CsgD and SUMO-CsgD. (Figure 14) Purification of both SUMO-CsgD and CsgD dual expressions were achieved with limited success. (Figure 15, 17) Residual presence of high molecular weight bands in the case of *Ec* CsgD dual expressions and low molecular weight bands, one of which possibly corresponding to GrpE, in the case of SUMO-CsgD dual expressions indicated the need for purification refinement - which could potentially be achieved through increases in wash volumes. Interestingly, both SUMO-CsgD and CsgD eluted at lower imidazole concentrations. This potentially implicates action of chaperone machinery at the N-terminus, disrupting coordination of nickel ions by their 6x His tags. The deficit in purification yield associated with this approach makes structural investigation through NMR and x-ray crystallography challenging.

Of the individual domains, the N-terminal domain of CsgD (CsgDN) showed promise for soluble expression following growth and expression according to a protocol previously described by Wen et al. [38](Figure 18) Following purification by nickel affinity chromatography, protein quality was probed utilizing a thermal shift assay. (Figure 19) Thermal unfolding of CsgDN revealed a typical sigmoidal fluorescence response indicative of a structured protein. This result prompted screening of 2-AI compounds to the natively purified N-terminal domain. (Figure 20)

Inspection of unfolding curves revealed many alterations to the shape and intensities of the unfolding responses. For the cases in which non-sigmoidal responses were observed, as for AGL-833 and AGL-949, it is possible that cylic, polyaromatic substituents are responsible for alterations to the unfolding response by directly interacting with sypro-orange dye. The data presented does not provide definitive evidence of 2-AI binding to *E. coli* CsgDN, however the *E. coli* CsgD system shows promise as a target for repurposed 2-AI molecules. The apparent shift in unfolding response in the presence of AGL-869 was indicative of 2-AI binding and may indicate that interactions with the N and C-terminal domains of full length *Ec* CsgD occur through a combination of polar and nonpolar interactions with the 3H-naphtho[1,2-d]imidazol-2-amine substituent, providing some guidance in selection of 2-AI molecules for future screening. In addition, SUMO-CsgD expression characteristics and newly obtained, active smt3 protease make it an attractive approach and refinement of the purification and cleavage process may allow for purification of recombinant *E. coli* CsgD at high concentration.

Section 3. *S.* Typhimurium CsgD Purification and Identification of 2-AI binding

Introduction

Following purification attempts for full length *Ec* CsgD, efforts were redirected to *S.* Typhimurium CsgD for which a crystal structure of the N-terminal domain exists (PDB ID: 5XP0), as well as evidence of direct binding to defined promoter sequences and characterization of its in vitro phosphorylation response. [38, 40] *S.* Typhimurium CsgD, denoted as *S*T CsgD, (uniprot ID: O54294) and *Ec* CsgD (uniprot ID: P52106) share 92% sequence identity, with complete conservation of the C-terminal domain. The *S.* Typhimurium N-terminal domain, denoted *S*T CsgDN, presents several abnormalities, lacking the canonical LuxR/FixJ α2 helix, shown in figure 2, as well as structural similarity between its α5 helix and the α6 helix of Vibrio cholerae master regulator VpsT. [38] The conserved aspartate D59 can be phosphorylated, a process that requires association of a divalent cation, typically magnesium, in the aspartate pocket. [55] Previous work determining the effect of divalent cations on phosphoryl transfer in Spo0F has indicated that the magnesium located in the loop between α1 and β1 is likely the magnesium responsible for stabilizing the phosphoryl-group transfer. [55] Additionally, the magnesium associated with the α4/ β5/ α5 interface is thought to be a product of solution conditions, however some evidence in Spo0F indicated that copper binding at this site affected conformational change in the α4-β4 loop. [55] Correct topology of the α4-β4 loop has been indicated as a prerequisite for sensor histidine kinase interactions and signal transduction. [55] Mapping of substituted sites onto the 5XP0 crystal structure reveal substitutions in the core β-sheets, as well as residues responsible for coordinating the two magnesium ions. This implies that functional differences between the CsgD homologs could arise from differences in

phosphorylation mechanism, differences in signal transduction, as well as arrangement of the hydrophobic interior.

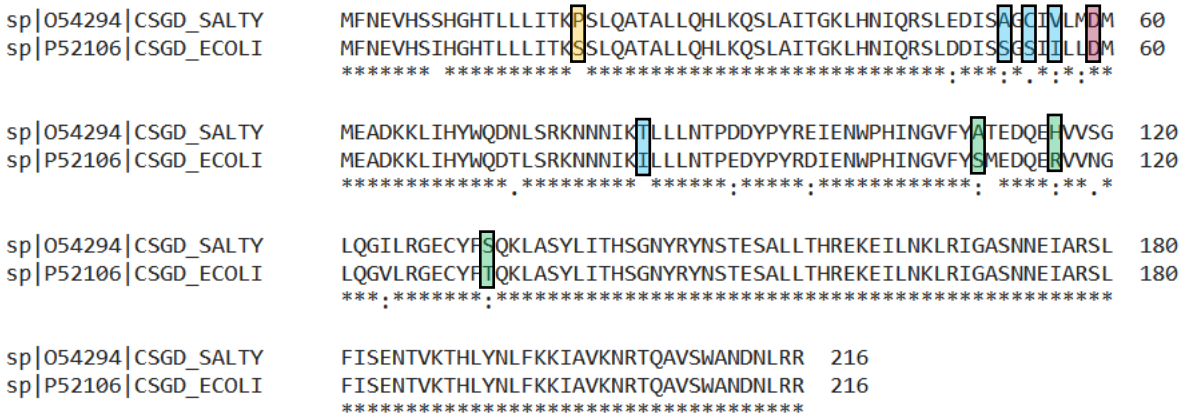

Figure 21: Sequence alignment of ST CsgD (CsgD_SALTY) and *Ec* CsgD (CsgD_ECOLI) are shown above. Substituted sites lack a * denotion underneath the residue alignment. The conserved site of phosphorylation is highlighted in red (D59), while residue substitutions internal to the hydrophobic core are highlighted in blue. Residues highlighted in green are substitutions implicated in coordination of magnesium in the α4/β5/α5 interface, while the residue highlighted in yellow indicates a substitution in a residue involved with coordination of magnesium to the loop following α1.

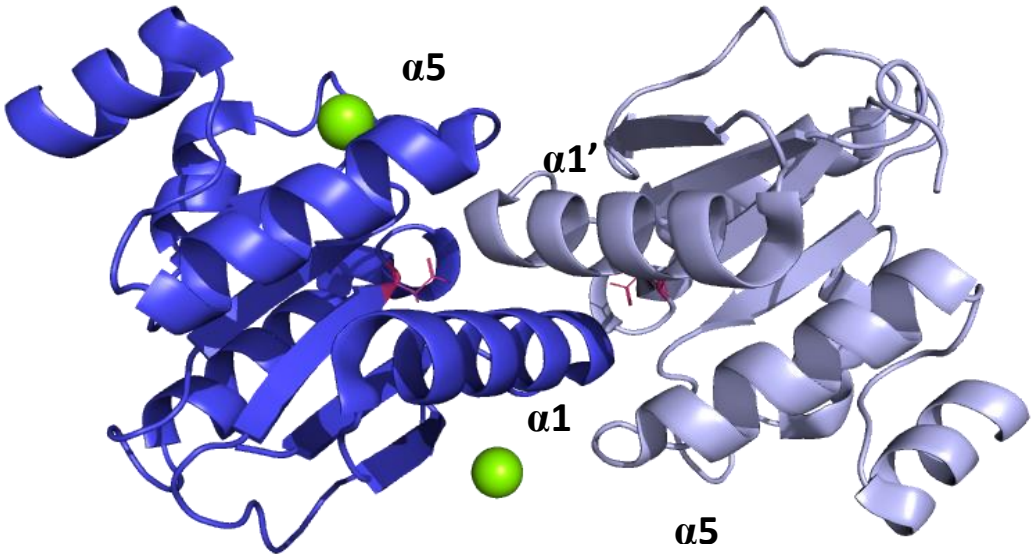

Figure 22: The dimerized N-terminal domain of ST CsgD (PDB ID: 5XP0) with individual domains denoted in blue and lilac. Magnesium ions are shown in lime-green. The dimerization interface mediated by the α1 and α5 helices of each domain is shown above. The conserved site of phosphorylation is shown in red.

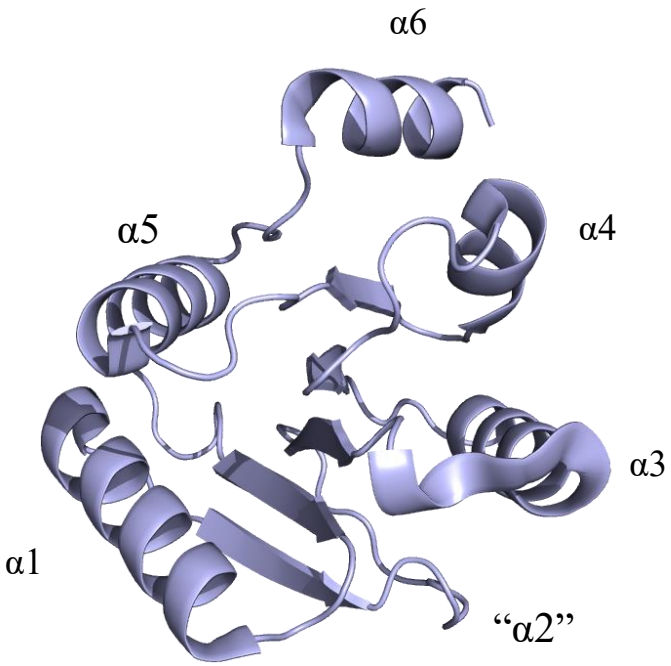

Figure 23: The N-terminal receiver domain of ST CsgD (PDB ID: 5XP0), reveals atypical characteristics. In addition to presentation of α2 as a disordered loop, the appearance of α6 is also atypical. The arrangement of α-helices, each connected by a β-sheet is shown above.

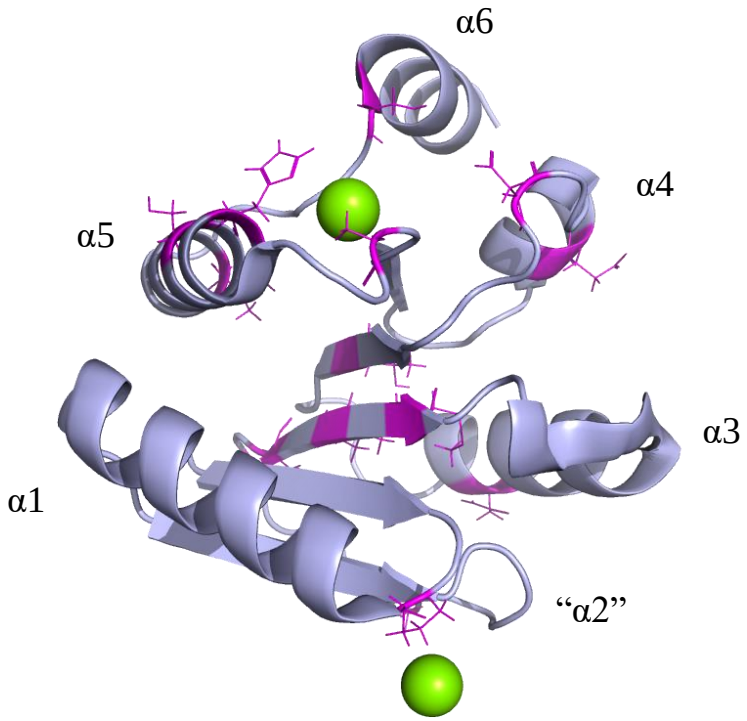

Figure 24: Substituted sites of the ST CsgD receiver domain (PDB ID: 5XP0) indicated in magenta are localized to the core β-sheets, as well as areas responsible for magnesium (lime-green) coordination.

Investigations into 2-AI binding to *S.* Typhimurium CsgD began with growth and expression of full-length *S*T CsgD, as well as the N-terminal domain crystalized by Wen et al. [38] Purification of both full length *S*T CsgD denoted as CsgD_mFL and the 5XP0 construct denoted as CsgD144 followed a corrected protocol obtained from Wen et al. Identification of 2-AI binding was qualitatively assessed through ΔT_m values obtained via thermal shift assay. Thermal unfolding curves were fit to Boltzmann distributions to extrapolate ΔT_m values. In addition to identifying *S*T CsgD as a viable target of 2-AI small molecules, structural determination of full-length *S*T CsgD remains crucial. Preliminary efforts to obtain crystallographic structural data are also included.

Methods

Plasmid Maintenance

Plasmid maintenance was performed as described in Section 2.

S. Typhimurium CsgD Constructs Growth and Expression

DNA corresponding to *S*T CsgD (*S*T CsgD and *S*T CsgDN) constructs were inserted into pET28a vectors and transformed into chemically competent *E. coli* BL21(DE3) expression strain. Cells from an overnight culture were used to inoculate 1 L of LB medium supplemented with kanamycin (50 mg/mL) to a starting OD_{600} of 0.02. Cultures were grown to an OD_{600} of 0.6-0.8 at 37°C, 200 rpm and induced with 1 mM IPTG. Following induction, cells were incubated for 8 hours at 28°C, 200 rpm and harvested at 6,000 rpm. Cell waste paste was stored at -80°C.

*S*T *CsgD Purification*

Cell paste from 1 L of bacterial growth was resuspended in 50 mL of 25 mM Tris pH 8.0, 150 mM NaCl, 5% glycerol and sonicated on ice. The lysate was clarified by centrifugation at 15,000 rpm, 4°C for 15 minutes. The supernatant was loaded onto a glass column containing 10 mL of equilibrated Ni-NTA (Invitrogen) resin. The column was washed with 500 mL of 25 mM Tris pH 8.0, 400 mM NaCl, 5% glycerol followed by 100 mL of 25 mM Tris pH 8.0, 1 M NaCl, 5% glycerol, and another 100 mL of 25 mM Tris pH 8.0, 150 mM NaCl, 5% glycerol. *S*T CsgD was eluted with a linear salt-bridge gradient 0-100% 25 mM Tris pH 8.0, 400 mM NaCl, 500 mM Imidazole, 5% glycerol. Following elution, CsgD-containing fractions were pooled and

dialyzed into 25 mM Tris pH 8.0, 400 mM NaCl, 5% glycerol, 1 mM β-mercaptoethanol (BME), 10 mM $MgCl_2$, and 0.1 mM Ethylenediamine tetraacetic acid (EDTA) for usage in subsequent analysis.

Electrophoretic Mobility Shift Assay (EMSA)

Reaction mixes buffered in 25 mM Tris pH 8.0, 400 mM mM NaCl, 5% glycerol, 1 mM β-mercaptoethanol (BME), 10 mM $MgCl_2$, and 0.1 mM EDTA containing 4 µM DNA and specified concentration of CsgD_mFL were incubated at 30°C for 15 minutes prior to loading onto a pre-chilled, 8% Tris-Borate-EDTA polyacrylamide (TBE) gel. Samples were run for 45–75 minutes at 80 V prior to soaking in Ethidium Bromide (EtBr) for 5 minutes. Following imaging of EtBr staining, gels were stained with coommassie-dye and destained with 30% methanol, 10% acetic acid prior to secondary imaging.

Hanging-Drop Vapor Diffusion Crystallization

Purified *S*T CsgD was concentrated using a 10K MWCO Amicon Ultra centrifugal filter (Sigma Aldrich) to a concentration of 5 mg/mL. 500 µL of crystallization condition (Hampton Research) was added to each well in a 24-well cell culture plate. Borosilicate coverslips were spotted with 2 uL drops of *S*T CsgD and each drop diluted with 2 uL of well-condition. Coverslips were inverted and sealed above each well and allowed to incubate at room temperature.

Thermal Shift Assay

As described in Section 2.

Gel Methods

As described in Section 2.

Table 2: *Primers used for synthesis of dsDNA fragments containing the S*T *CsgD box (red)*

csgB	Forward	GTG GGT TTT AAT ACT TTG GTA TGA ACT AAA AAA GAA AAA TAC AAC GCG CGG GTG AGT TAT TAA AAA TAT TTC CGC
	Reverse	GCG GAA ATA TTT TTA ATA ACT CAC CCG CGC GTT GTA TTT TTC TTT TTT AGT TCA TAC CAA AGT ATT AAA ACC CAC
yccT	Forward	CAG CTA ATA CGA ATG AAT GAG A
	Reverse	TCT CAT TCA TTC GTA TTA GCT G

Data Analysis

As described in Section 2.

Results

Purification of S. Typhimurium full-length CsgD and CsgD_144

Both *S*T CsgD and *S*T CsgDN were purified by nickel affinity chromatography. Results of each purification were analyzed by SDS-PAGE and the results are shown below in figure 25 and 26. In contrast to *Ec* CsgD, *S*T CsgD and *S*T CsgDN, both express readily in the soluble, cytoplasmic fraction and are purified at high concentrations following affinity chromatography.

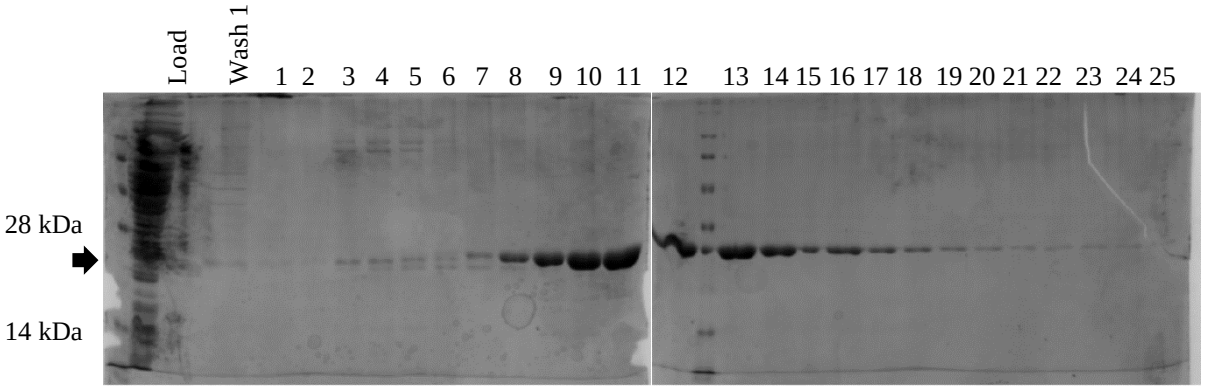

Figure 25: Purification results for *S*T CsgD were analyzed by SDS-PAGE on a tris-glycine polyacrylamide gel. Samples from the indicated elution fractions, as well as the load and wash flow throughs were recorded. The location of *S*T CsgD is indicated with an arrow.

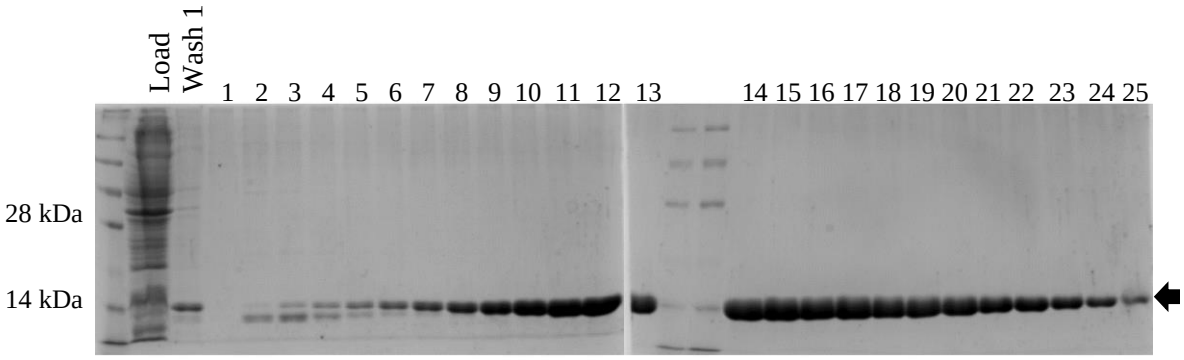

Figure 26: Purification results for *S*T CsgDN were analyzed by SDS-PAGE on tris-glycine polyacrylamide gels. Samples corresponding to the indicated elution fractions, as well as the load and wash flow throughs were recorded. The location of *S*T CsgDN is indicated with an arrow.

Identification of 2-AI Binding in full-length CsgD and CsgD_144

To assess ligand binding of 2-AI compounds to either *S*T CsgD or *S*T CsgDN identified through a thermal shift assay. Protein mixtures containing indicated protein and ligand concentrations were subjected to thermally-induced unfolding in the presence of 8x Sypro-Orange dye. Fluorescence spectra of technical replicates and ΔT_m values are reported in figures 27 and 28. Thermostabilizing shifts in ΔT_m, indicative of 2-AI binding, were observed for both samples treated with AGL-869 and AGL-853 in *S*T CsgD and *S*T CsgDN. For both proteins, treatment with AGL-949 and AGL-833 resulted in altered baseline and distorted unfolding curves, while AGL-503, AGL-600, and AGL-726 treatment were not indicative of 2-AI binding.

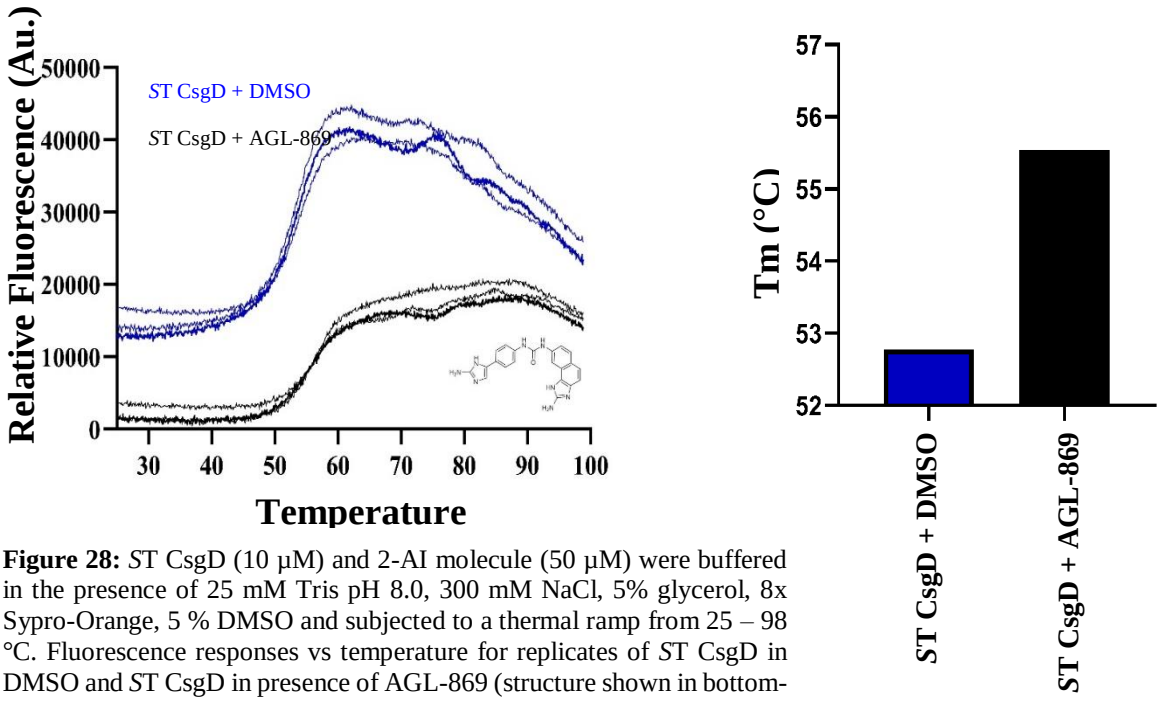

Figure 28: *S*T CsgD (10 µM) and 2-AI molecule (50 µM) were buffered in the presence of 25 mM Tris pH 8.0, 300 mM NaCl, 5% glycerol, 8x Sypro-Orange, 5 % DMSO and subjected to a thermal ramp from 25 – 98 °C. Fluorescence responses vs temperature for replicates of *S*T CsgD in DMSO and *S*T CsgD in presence of AGL-869 (structure shown in bottom-left) are shown on the left. T_m values determined from Boltzmann fits are reported on the right.

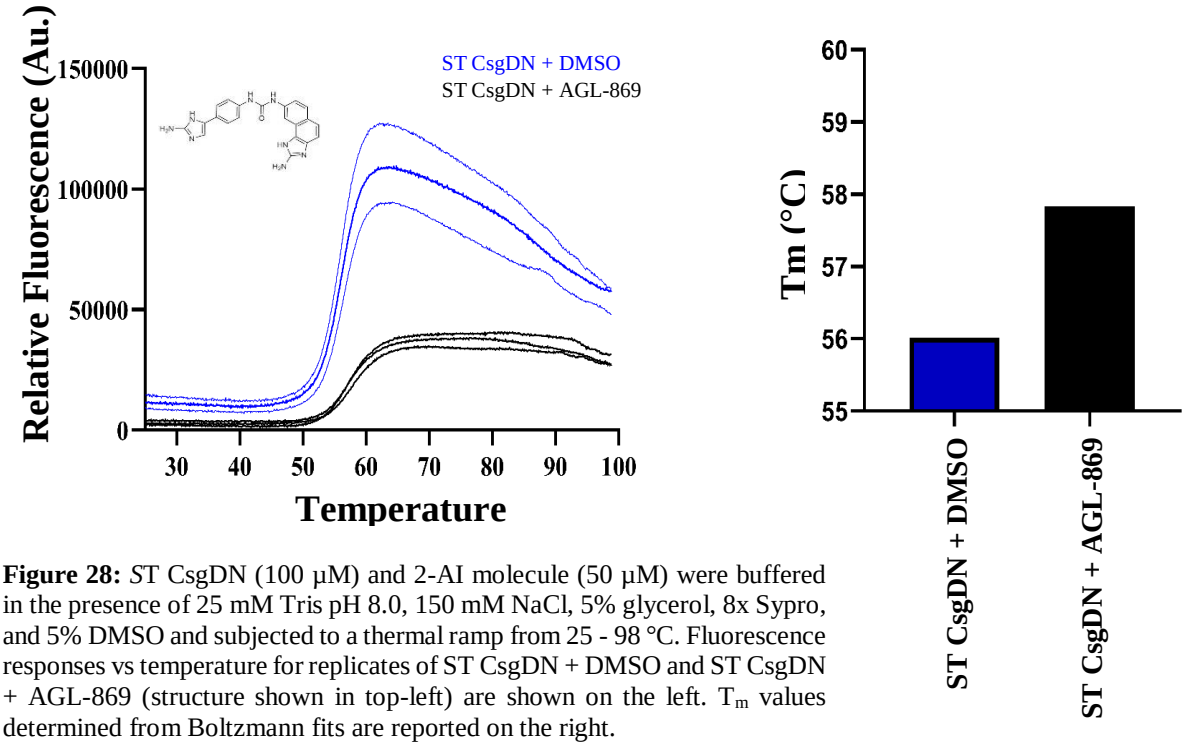

Figure 28: *S*T CsgDN (100 µM) and 2-AI molecule (50 µM) were buffered in the presence of 25 mM Tris pH 8.0, 150 mM NaCl, 5% glycerol, 8x Sypro, and 5% DMSO and subjected to a thermal ramp from 25 - 98 °C. Fluorescence responses vs temperature for replicates of ST CsgDN + DMSO and ST CsgDN + AGL-869 (structure shown in top-left) are shown on the left. T_m values determined from Boltzmann fits are reported on the right.

Thermal Shift Assay of ST CsgD in the Presence of DNA

To investigate DNA substrate-binding by *S*T CsgD, protein at the indicated concentration was incubated with dsDNA containing the consensus target sequence upstream of the *fliF* gene in *S.* Typhimurium, as well as 8x Sypro-Orange. Fluoresence spectra and reports of ΔT_m values are given in figure 29. A destabilizing effect was noted for the effect on T_m. Both the shape of the unfolding response and the intensity of the curve differ for the DNA-containing samples. Notably, following the rapid increase in fluorescence intensity, *S*T CsgD alone exhibits concave-up characteristics over a short temperature range, while this feature is absent from the DNA-containing samples.

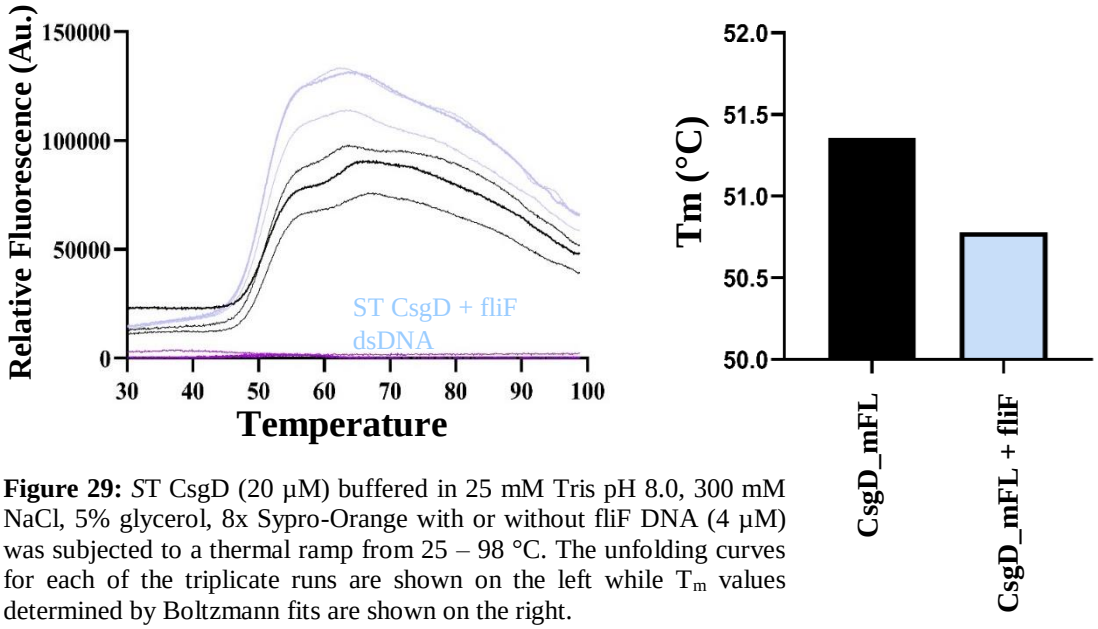

Figure 29: *S*T CsgD (20 µM) buffered in 25 mM Tris pH 8.0, 300 mM NaCl, 5% glycerol, 8x Sypro-Orange with or without fliF DNA (4 µM) was subjected to a thermal ramp from 25 – 98 °C. The unfolding curves for each of the triplicate runs are shown on the left while T_m values determined by Boltzmann fits are shown on the right.

Electrophoretic Mobility Shift Assay

Purified *S*T CsgD was incubated with 4 μM DNA for in-gel detection of DNA-binding. The DNA present corresponded to dsDNA fragments containing the consensus target sequences for *S*T CsgD in various genes it is known to regulate. Incubated samples were loaded onto an 8% TBE acrylamide gel and subjected to gel electrophoresis until the dye front neared the end of the gel. Imaging of EtBr staining is shown below in figure 30. Results showed no difference between the electrophoretic mobility of the reaction mixtures and DNA controls.

CsgD	+	-	+	+	-	+
yccT	+	+	-	-	-	-
csgB	-	-	-	+	+	-

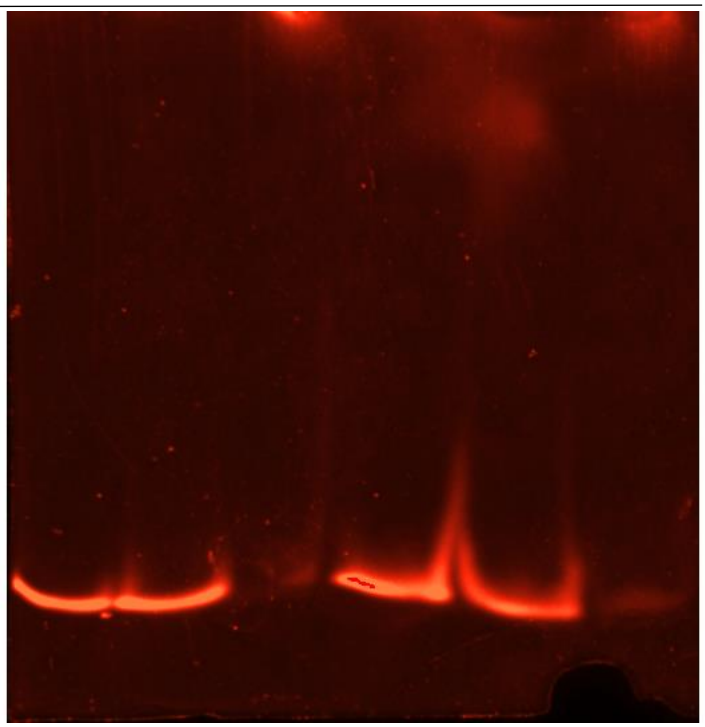

Figure 30: 20 μM ST CsgD was incubated with 4 μM of promoter fragments for indicated genes prior to gel electrophoresis on an 8% TBE. Reaction mixtures, as well as both DNA and Protein controls for each promoter fragment are visualized by EtBr staining and shown above.

*Crystallization of S*T *CsgD*

Structural determination of *S*T CsgD began with screening of various crystal conditions provided by Hampton Research (Aliso Viejo, CA). Images of crystals formed in the presence of 0.1 M Tris pH 8.5, 0.2 M Ammonium Phosphate Monobasic, and 50% v/v (+/-)-2-Methyl-2,4-pentanediol are shown below in figure 31.

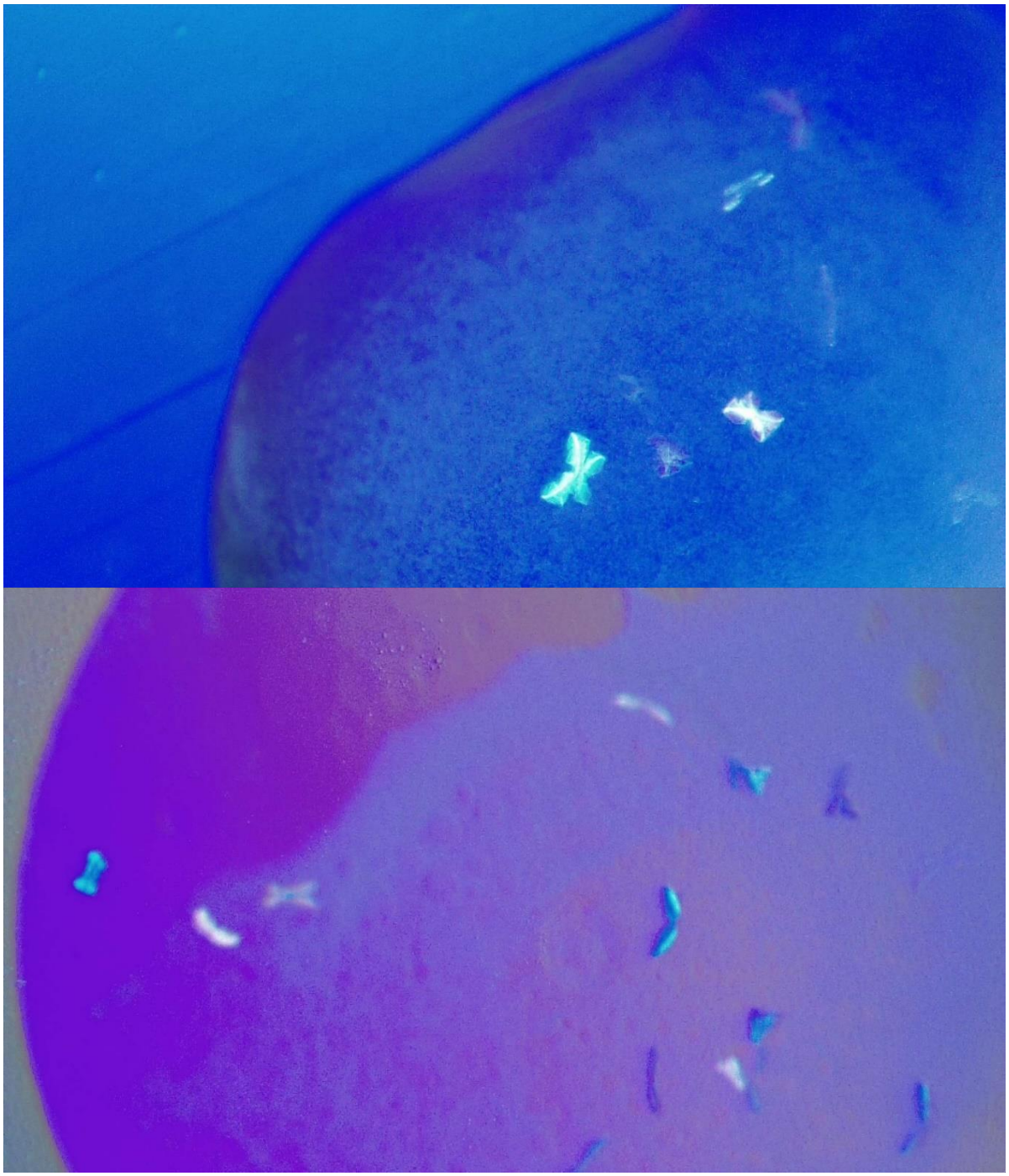

Figure 31: Crystals formed from the crystallization condition 0.1 M Tris pH 8.5, 0.2 M Ammonium Phosphate Monobasic, and 50% v/v (+/-)-2-Methyl-2,4-pentanediol. A negative control containing only the original buffer and crystallization condition was also performed – in which phase separation, but not crystal formation, was observed.

Discussion

In stark contrast to the *E. coli* homolog, *S*T CsgD, and *S*T CsgDN expressed readily and with high solubility in E. coli Bl21(DE3) cells following a growth and expression protocol obtained from Wen et. al. [38] (Figure 25, 26) One possibility is that the *S*T CsgD homolog is nonfunctional within the background of an *E. coli* expression system, while *Ec* CsgD may express functionally and be localized to inclusion bodies as a result of its disruptive effect on growth state and cellular signaling. Both *S*T CsgD and *S*T CsgDN displayed sigmoidal thermal unfolding curves in the presence of 8x sypro-orange dye, allowing for qualitative description of the purified protein as a structured isolate. Investigation of 2-AI binding to *S*T CsgD and *S*T CsgDN through thermal shift assay reveal a number of shifts in T_m of treated groups indicative of 2-AI binding. (Figure 27, 28) For both *S*T CsgD and *S*T CsgDN, AGL-853 and AGL-869 treatment resulted in statistically significant shifts in T_m compared to the DMSO control group. Interestingly, AGL-869 was also identified as a potential ligand for *Ec* CsgD – this common feature may arise from shared secondary structures at the inter-domain interface. Each of these shifts were thermostabilizing in nature and indicative of 2-AI binding to the respective isolates. 2-AI binding to *S*T CsgDN indicated preference for molecules containing an aromatic substituent with polar substitutions, as well as linkage of the aromatic substituent to a carbonyl-carbon via a secondary amine. Specificity of the full length *S*T CsgD indicated preference for polyaromatic substituents bearing similar polar substitutions. In addition, studies revealed that molecules lacking linkage of the substituent via a secondary amine showed reduced shifts in T_m. This indicates that intelligent design of inhibitory 2-AI molecules will rely on specific tuning of both polar and hydrophobic components attached to the 2-AI pharmacore.

Efforts to identify DNA-binding activity in *S*T CsgD began with in-gel detection of DNA-binding through electrophoretic mobility shift assay. dsDNA fragments corresponding to *S*T CsgD target sequences identified by Ogasawara et. al. were incubated with purified *S*T CsgD and analyzed by native-PAGE on an 8% TBE gel. (Figure 30) However, DNA-binding activity was not seen. This may be due to its high theoretical pI at 8.99, calculated using the ProtParam tool by Expasy, which would cause the protein to run in the opposite direction as the DNA during gel electrophoresis. Difficulties in gel-entry by CsgD were evidenced during multiple attempts, for which coommassie staining of the gel revealed smearing in the wells or absence of coommassie following destaining. Additionally, presence of the N-terminal histidine tag throughout the DNA-binding assay may interfere with the dimerization interface provided by the α1 and α5 helices of the N-terminal domain that have been shown critical for dimerization and DNA-binding activity. [38]

Further attempts to identify DNA-binding were inspired by Vietor et al. who utilized a thermal shift assay for detection of DNA-binding by the NF-κB family of dimeric transcription factors. [56] Visual comparison between unfolding curves for *S*T CsgD alone and in the presence of DNA reveal a destabilizing effect on T_m for the DNA-treated sample. (Figure 29) This effect was coupled to an increase in the magnitude of relative fluorescence. Shift in the T_m may be indicative of DNA-binding activity, however further investigations should be performed prior to making definitive statements about its DNA-binding activity. One possibility is that DNA-binding shifts population equilibrium in favor of the dimeric form, a structural arrangement that may be more amenable to thermally induced unfolding. High Performance Liquid Chromatography by size exclusion, as well as repeat of the assay at saturating concentrations of DNA would aid in understanding this phenomenon.

While existing structures and fold-recognition of the C-terminal domain serve to identify *S*T CsgD as a NarL/FixJ family protein with a LuxR-like HTH DNA-binding domain, efforts to glean near-atomic level structural determinations of full-length *S*T CsgD have shown that CsgD is amenable to concentration and was successfully concentrated to 5 mg/mL without precipitating from solution. While the crystallized product shown in Figure 31 was determined to be a component of the crystallization condition, additional crystallization screens and alterations to previously examined conditions may provide an avenue for full length structural determination for purposes of designing specific and inhibitory 2-AI molecules.

Section 4. Native Purification Provides Structural Insights into the Bacterial Amyloid CsgA

Introduction

CsgA is a bacterial amyloid that serves as the major structural component of curli-containing biofilms produced by enteric bacteria. [23] It has been shown to colocalize with αSyn both in vitro and in vivo. [18, 23] CsgA's ability to cross-seed endogenous amyloids like αSyn, as well as nucleating fibrillation has identified CsgA as a key player in neurodegenerative onset and progression.

Amyloid cytotoxicity is believed to stem from the soluble oligomers as opposed to the insoluble fibrils. [10] The toxicity of the transient soluble oligomers is postulated to stem from pore-like oligomeric intermediates exerting aberrant functions at the surface of the plasma membrane. [10, 11]

Investigations into *E. coli* CsgA have described three protein domains, an N-terminal sequence that targets it for periplasmic translocation by SecYEG machinery, as well as an N-terminal 22 amino acids prior to five imperfect repeats believed to serve as the amyloid core. [57] The monomeric unit of CsgA is believed to initially be secreted to the extracellular space in an unstructured form with repeat regions quickly adopting a β-sheet-turn-β-sheet motif. [23] The β-sheet rich structure is postulated to enhance amyloid fibrillation by providing an amyloid-template to soluble monomers and multimers of endogenous amyloids. [22] While this provides insight into amyloid fibril cross-seeding, the transient oligomers remain the central focus for understanding cytotoxicity in neurodegenerative disease pathologies. Understanding the

structures and features associated with monomeric and oligomeric conformational ensembles of CsgA is crucial for gaining a mechanistic framework for the acceleration of neurodegenerative onset.

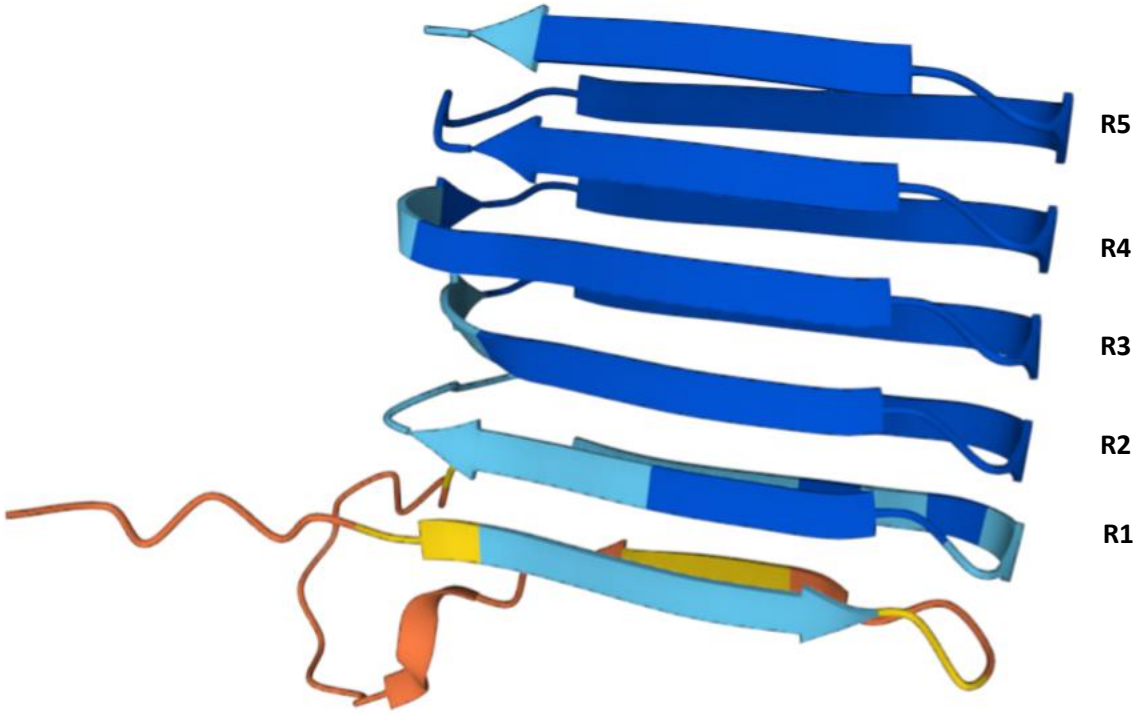

Figure 32: An AlphaFold structural prediction of *E. coli* CsgA. The predicted structure depicts the confidence with orange being the lowest and deep blue being the highest.

One of the key aspects of working with amyloid proteins is their propensity for oligomerization and aggregation. Within the cell, CsgC and DnaK work to keep CsgA in a soluble, translocation-competent state throughout transport to the extracellular matrix. [36, 43] Typical approaches to mitigate this effect include the application of chemical denaturants like GuHCl through either on-column denaturation to solubilize CsgA isolated from the cytoplasm, as well as soaking of bacterial cell paste in a lysis buffer containing GuHCl to solubilize CsgA

from inclusion bodies. [23, 58] Denaturant purifications of amyloid proteins aid in promoting protein amenability to experimentation, as well as increasing yield, however the application to a highly dynamic, intrinsically-disordered protein known to interact with chaperone-protein complexes may prohibit the purified recombinant protein from sampling its complete range of conformational identities. This would restrict ability to investigate all physiologically-relevant complexes that contribute to neurodegenerative outcomes.

Initial purification approaches began with the closely-related and structurally similar minor curli subunit CsgB. Since both CsgA and CsgB are directed towards the periplasm, initial growth efforts followed a periplasmic growth and expression protocol optimized for αSyn. In addition, post-induction incubation at 37 °C may aid in the production of the stress-related proteins CsgB and CsgB by exacerbating cellular stress.

Initial purification efforts were followed by a modified growth and purification protocol involving growth supplementation with the pKJE7 vector containing genes for heat-shock chaperone proteins DnaK, DnaJ, and GrpE. Natively-purified CsgA and CsgB were obtained via incubation with magnesium chloride ($MgCl_2$), ATP, and soluble, denatured *E. coli* proteins throughout a standard immobilized nickel affinity chromatography with thorough washing.

Natively-purified CsgA was verified through His-staining of SDS-PAGE for in-gel detection of the 6x histidine tagged CsgA. Concentration-dependent effects and response to the anionic surfactant SDS were characterized through SDS-PAGE and Native-PAGE, while thermodynamic parameters and understanding of its unfolding transition were investigated through chemically-induced unfolding studies. Additionally, structural characteristics of the

natively purified conformational ensemble were investigated through far-UV circular dichroism for qualitative description of secondary structural motifs.

Finally, a comparative approach was taken to investigate differences between the structural and conformational features of natively-purified CsgA and both denaturant-treated CsgA and denaturant-purified CsgA.

Methods

Plasmid Maintenance

Plasmid maintenance was performed as described in Section 2.

CsgA/CsgB Growth and Expression

DNA fragment for *E. coli* CsgA/CsgB was inserted into pET28a vector and transformed via electroporation along with L-arabinose inducible pKJE7 plasmid encoding chaperone proteins DnaK, DnaJ, and GrpE into either *E. coli* BL21(DE3) or HMS174(DE3) expression strain. Cells from an overnight culture were used to inoculate 1L of LB medium supplemented with kanamycin (35 mg/mL), Chloramphenicol (20 mg/mL), and L-arabinose (0.1 mg/mL) to a target OD600 of 0.003. Cells were grown at 37°C, 225 rpm to an OD600 of 0.6-0.8. Protein expression was induced with 1 mM isopropylthio-β-galactoside for 48 hours at 30°C, 225 rpm. Following harvest at 6,000 rpm, supernatant was decanted. Cell paste was flash frozen in liquid nitrogen and stored at -80°C.

CsgA Native Purification

Cell paste containing overexpressed CsgA was thawed and resuspended in 200 mL of 20 mM HEPES pH 8.0 supplemented with 5mM ATP, 12.5 mM MgCl2, and 0.1 mg/mL denatured *E. coli* proteins. Following sonication on ice, cell paste was clarified by centrifugation at 15,000 rpm, 4°C for 15 minutes. The supernatant was equilibrated with 10 mL Hi-Bind Ni QR agarose resin (BioVision) and batch bound with stirring at 4°C for 90 minutes. The sample was then

loaded onto a glass column fitted to an Econo Gradient pump (BioRad) at 4°C. The resin was washed with 500 mL of 20 mM HEPES pH 8.0. Following wash, CsgA was eluted with 20 mM HEPES pH 8.0, 150 mM imidazole in a linear gradient elution. CsgA-containing fractions, identified by SDS-PAGE, were pooled and dialyzed into 20 mM HEPES pH 8.0, 2 mM EDTA. The pooled eluant was then dialyzed into 10 mM Potassium Phosphate pH 7.2 and concentrated using absorbent. Concentrated CsgA was supplemented to 10% glycerol prior to flash freezing in liquid nitrogen and storage at -80°C.

Gel Methods

SDS-PAGE analysis was performed on 12% tris-glycine polyacrylamide gels. Both 10 and 15-well combs were used with 20 µL and 10 µL loads respectively. Electrophoresis was performed using a BioRad PowerPac HC at 150 V for 75 minutes. Gels were stained with coommassie and destained with 30% methanol and 10% acetic acid. Gels were imaged on a GelDoc EZ imager (BioRad).

Native-PAGE was performed on both 9% and 12% tris-glycine acrylamide gels. 10 well combs were used with 20 µL loads. Electrophoresis was performed using a BioRAD PowerPac HC at 150 V for 5 hours on ice. Gels were stained with coommassie and destained with 30% methanol, 10% acetic acid. Gels were imaged on a GelDoc EZ imager (BioRad).

His-Stain SDS-PAGE

Following SDS-PAGE, unstained gels were fixed in 30% methanol, 10% acetic acid for 30 minutes. Gels were thoroughly washed with excess water prior to addition of 50 mL of His-Tag Stain (Thermo Fisher) for 15 minutes with shaking. Gels were then washed thoroughly with excess water and shaking prior to addition of 50 mL of His-Tag Stain Developer 15 minutes with shaking. Following stain development, gels were washed with excess water and imaged with the EtBr setting on a GelDoc EZ Imager (BioRad).

JC-1 Assay

Reaction mixtures containing 5% DMSO, 1µMJC-1, 20 mM HEPES pH 8.0, and indicated concentration of natively purified His-CsgA were prepared and incubated in a 96-well, clear bottom plate. The samples were incubated at 37°C with shaking for 10 minutes in a SpectraMax iD5 plate reader. Using excitation at 498 nm, emission spectra were recorded from 510 to 680 nm. The cycle was repeated for 280 minutes. Spectra were baseline-corrected prior to plotting and secondary analysis.

Circular Dichroism

CsgA, both natively purified and refolded, were buffer exchanged into 10 mM Potassium phosphate pH 7.2. Following dilution to 5µMCsgA, spectra were recorded from 180 – 300 nm using a Chirascan V100 UV-CD (Applied Photophysics). Spectra were recorded at room temperature, with a 0.1 cm path length, at a scan speed of 50 nm/min with a 0.5 nm step. Each

scan was repeated four times, averaged, and buffer-subtracted prior to conversion to molar elipcities.

CsgA Denaturant Purification

CsgA was purified using GdnHCl based on a protocol described by Chapman et al. [23] Cell paste was thawed and resuspended in 100 mL of 10 mM Potassium phosphate pH 7.2. Following sonication on ice, cell paste was clarified by centrifugation at 10,000 x g for 15 minutes. The supernatant was incubated with 10 mL of Ni-NTA (Invitrogen) resin for 45 minutes at 4°C. The sample was loaded onto a glass column and washed with 500 mL 10 mM Potassium phosphate pH 7.2. The column was then washed with 100 mL 10 mM potassium phosphate pH 7.2, 8M GdnHCl, and eluted with 100 mL of 50 mM potassium phosphate pH 2.0, 8M GdnHCl. CsgA-containing fractions were desalted using a PD10 (GE Healthcare) column. Following desalting, CsgA was buffer exchanged into 10 mM potassium phosphate pH 7.2 and analyzed by circular dichroism, Native-PAGE, and SDS-PAGE.

Chemically Induced Unfolding Studies

Purified CsgA was buffer exchanged into 10 mM potassium phosphate pH 7.2 and its concentration determined by BCA. Following buffer exchange, samples containing 10 mM Potassium phosphate pH 7.2, 5μM CsgA, and varying denaturant concentrations spanning 0-6M or 0-8M were incubated overnight with rotation. A Shimadzu RF-5301 Fluorimeter was used to

investigate excitations at both 280 nm and 297 nm with emission spectra recorded from 300-650 nm. Spectra were then normalized and baseline-corrected prior to secondary analysis.

Expression Test

Performed as described in Section 2.

Data Analysis

JC-1 Assay

Raw fluorescence emission spectra recorded from 300-650 nm were exported from the SpectraMaxID5 (Molecular Devices) instrument. The raw spectra were baseline-corrected using fractional difference. $\frac{Value-Baseline}{Baseline}$, where the baseline was determined from the average of the first 3 three values and last 3 values. To mitigate baseline distortion as a result of excitation overlap with the emission spectra, points below 518 nm were excluded from the analysis. Following baseline-correction, plots were generated relating fractional fluorescence intensity and time at specified emission wavelengths.

Circular Dichroism

Background-subtracted spectra were exported from the Chirascan (Applied Photophysics) instrument. The data was exported to Pro-Data Viewer (Applied Photophysics). Runs corresponding to the buffer background were averaged and defined as baseline. Runs corresponding to protein-sample were also averaged and baseline-subtracted prior to conversion from mdeg to molar elipcities.

Chemically-Induced Unfolding

Raw spectra were exported from the Shimadzu RF-5301 instrument into Prism (GraphPad Software). The raw spectra were first normalized, defining upper and lower limits as

the largest and smallest values in the dataset respectively. The normalized data was then baseline-corrected utilizing fractional difference $\frac{Value-Baseline}{Baseline}$ prior to generation of full-spectra. Secondary analysis to determine m-values and ΔG_{UN} was performed via plotting of fractional fluorescence intensity vs the concentration of denaturant followed by nonlinear least squares fit involving six parameters as shown in equation 2.

(eq. 2)

$$\frac{\left\{\left(\gamma_F^0+m_F[denaturant]\right)+\left(\gamma_U^0+m_U[denaturant]\right)e^{\left\{\frac{(\Delta G(water)-m[denaturant])}{RT}\right\}}\right\}}{1+e^{\frac{\Delta G(water)-m[denaturant]}{RT}}} = \gamma_{obs}$$

Where γ_F and γ_U are intercepts of the linear fit during the folded and unfolded regimes of the unfolding curve and m_U and m_F are the slopes. γ_{obs} refers to the observed value of the parameter utilized to measure unfolding.

Results

E. coli CsgB Growth and Expression

Initial growth procedures were performed in accordance with a protocol optimized for αSyn. In brief, transformed expression strain cells – either BL21(DE3) or HMS174(DE3) – were grown and induced with IPTG at 37°C. Growth results were interpreted via an expression test. SDS-PAGE of the expression test is shown in figure 33. CsgB expression was difficult to distinguish from the background of the whole cell-lysate.

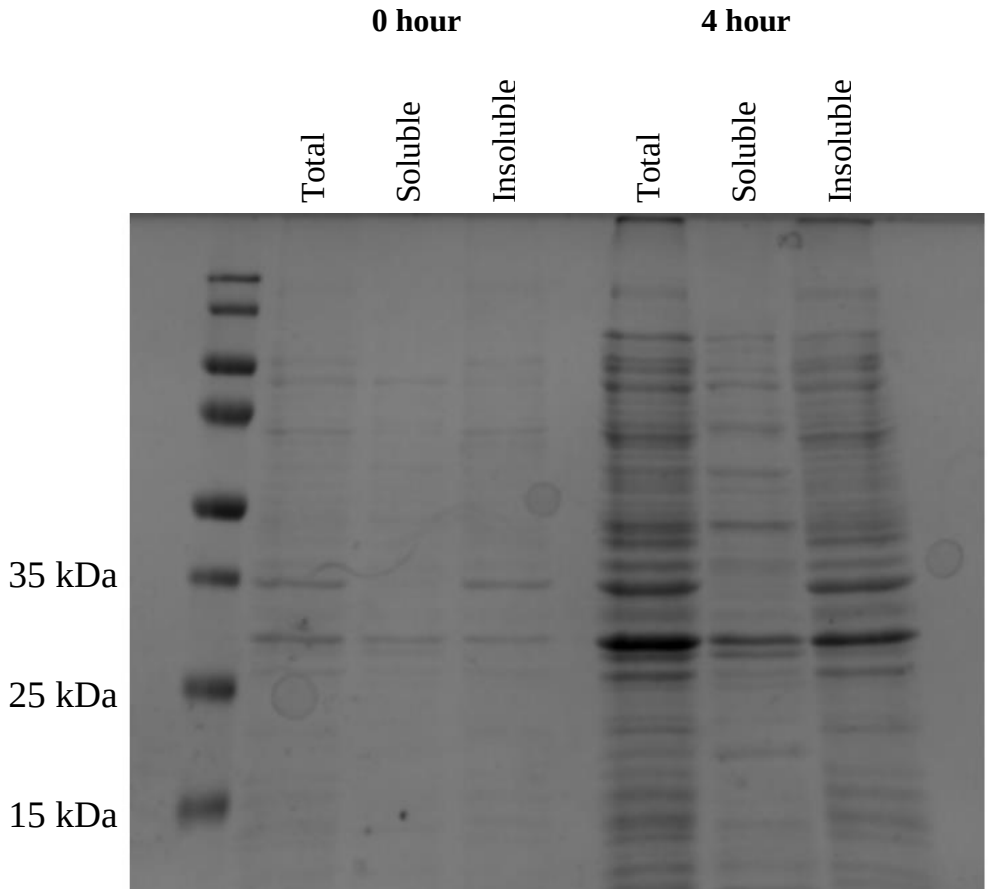

Figure 33: Expression test results were analyzed by SDS-PAGE on tris-glycine polyacrylamide gels. The fractions corresponding to the whole-cell lysate (total), cytoplasmic fraction (soluble), and cellular debris and inclusion bodies (insoluble) were recorded over time.

E. coli CsgB/pKJE7 Dual Growth and Expression

The pKJE7 plasmid containing DnaK, DnaJ, and GrpE is arabinose inducible and was induced at the outset of the growth. Cells were allowed to grow to appropriate turbidity prior to induction of CsgB with IPTG. The results of each growth were interpreted through expression testing. SDS-PAGE of an expression test is shown in figure 34. CsgB expression was not obvious and difficult to distinguish from the background of the whole-cell lysate.

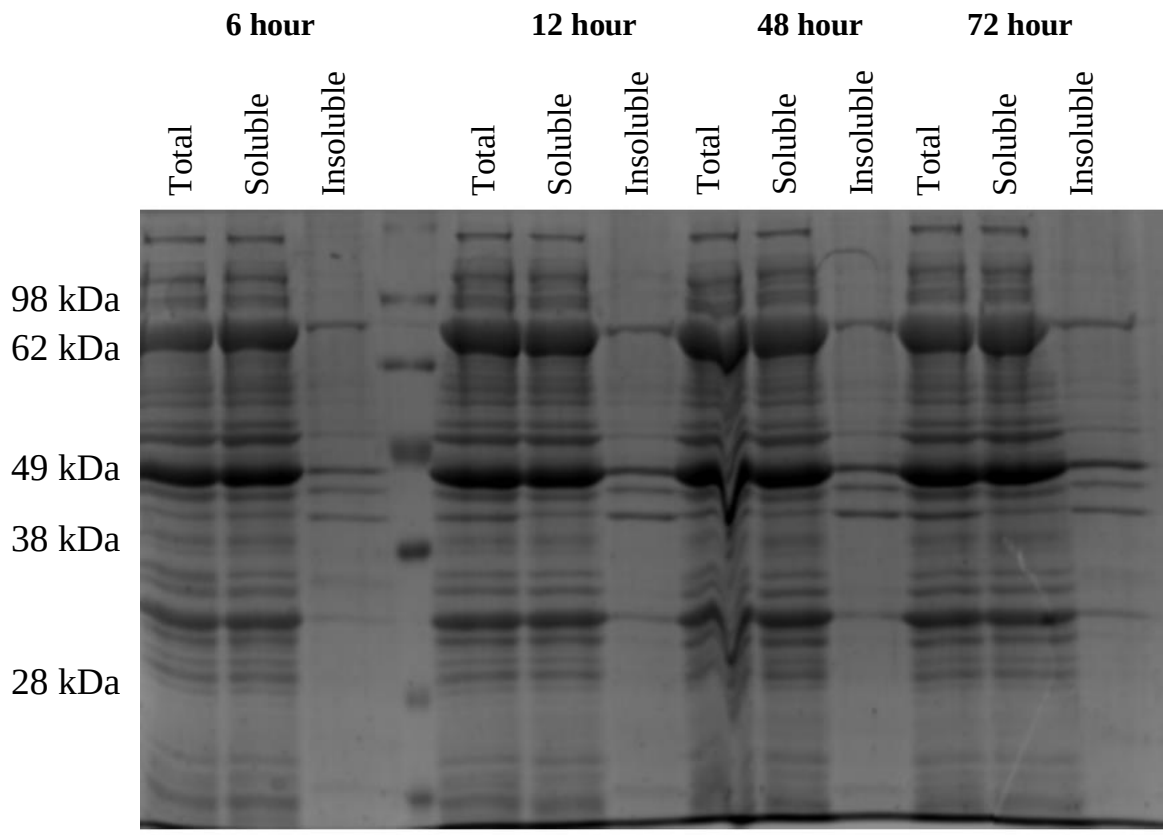

Figure 34: Expression test results were analyzed by SDS-PAGE on tris-glycine polyacrylamide gels. The fractions corresponding to the whole-cell lysate (total), cytoplasmic fraction (soluble), and cellular debris and inclusion bodies (insoluble) were recorded over time. Bands corresponding to overexpressed protein are thought to either be KJE7-encoded proteins or CsgB, however distinction of the two possibilities is inconclusive.

E. coli CsgB Ni^{2+} Affinity Chromatography for Removal of Chaperone Proteins

Cell paste containing overexpressed *E. coli* CsgB and pKJE7 encoded proteins were resuspended in a HEPES-buffered solution. Following sonication and clarification of lysate via centrifugation, the supernatant was incubated with Magnesium chloride ($MgCl_2$), ATP, and soluble and denatured *E. coli* proteins to allow for DnaK to release CsgA and bind to alternative substrate. Incubation was followed by standard nickel affinity chromatography at 4°C. Sample

was eluted from the column with imidazole and purification result was analyzed by SDS-PAGE shown in figure 35. Results indicate residual chaperone contamination, however allow for putative identification of CsgB. Which appears to run slightly higher than the expected 17 kDa for the His-tagged recombinant protein.

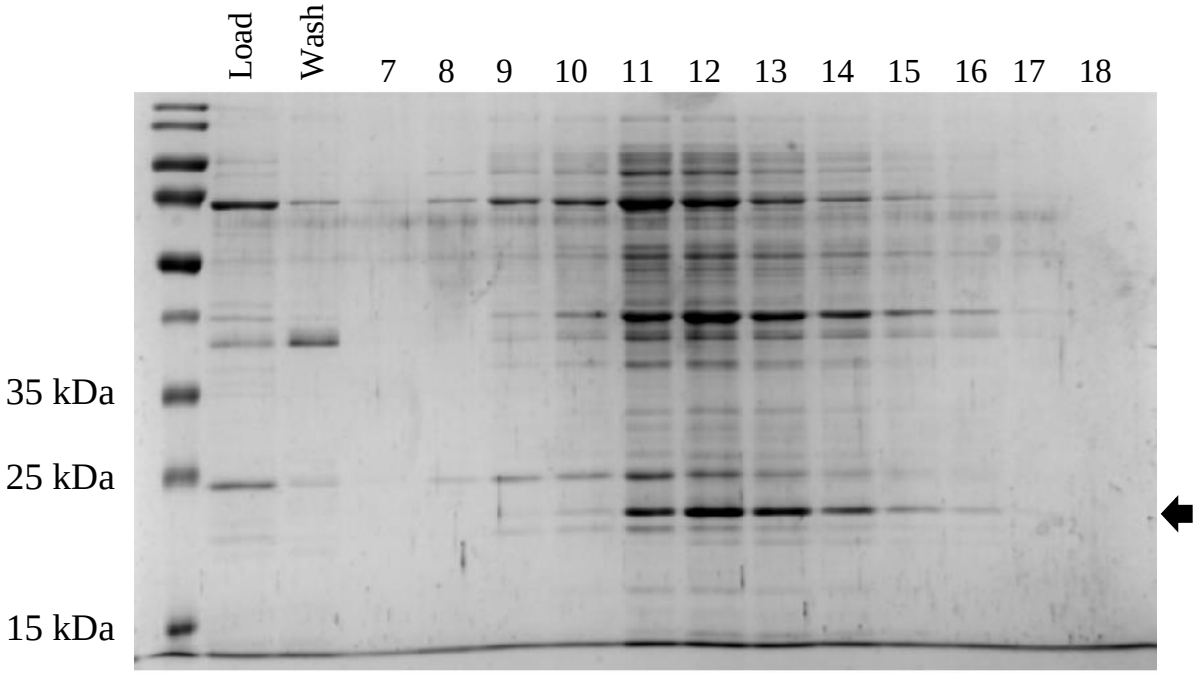

Figure 35: Purification results of CsgB supplemented with pKJE7 were analyzed by SDS-PAGE on a tris-glycine polyacrylamide gel. Fractions corresponding to samples with the highest A280s, as well as samples of the load and wash flow throughs were recorded. Putative location of monomeric CsgB is identified with an arrow.

E. coli CsgA/pKJE7 Growth and Expression

In a similar manner to CsgB, a dual-expression protocol to supplement the growth with DnaK, DnaJ, and GrpE was also implemented for expression of *E. coli* CsgA. Growth results were analyzed by expression testing. SDS-PAGE displaying growth result is shown below in figure 36.

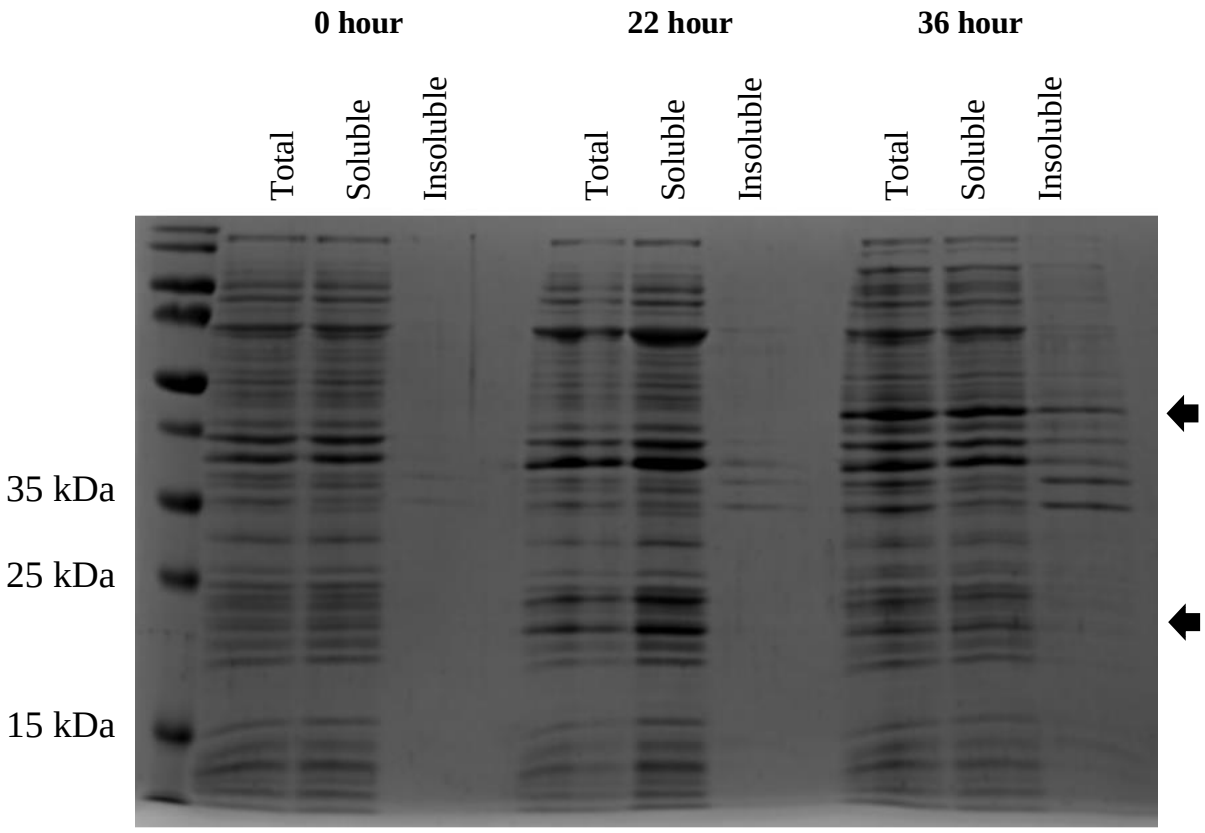

Figure 36: Results of expression testing were analyzed by SDS-PAGE on a tris-glycine polyacrylamide gel. Fractions corresponding to the whole-cell lysate (total), the cytoplasmic fraction (soluble), and cellular debris and inclusion bodies (insoluble) were recorded at indicated time points. Putative locations for the presence of monomeric and oligomeric CsgA are identified with arrows.

E. coli CsgA Ni^{2+} Affinity Chromatography for Removal of Chaperone Proteins

Following guidance from previous purifications in CsgB, cell paste containing overexpressed CsgA was resuspended in a HEPES-buffered solution containing $MgCl_2$, ATP, and soluble, denatured *E. coli* proteins. The resuspension was subjected to sonication and centrifugation. The supernatant was incubated with equilibrated Ni-HiBind resin to allow batch-binding of CsgA to the resin. Purification result was analyzed by SDS-PAGE and is shown in figure 37.

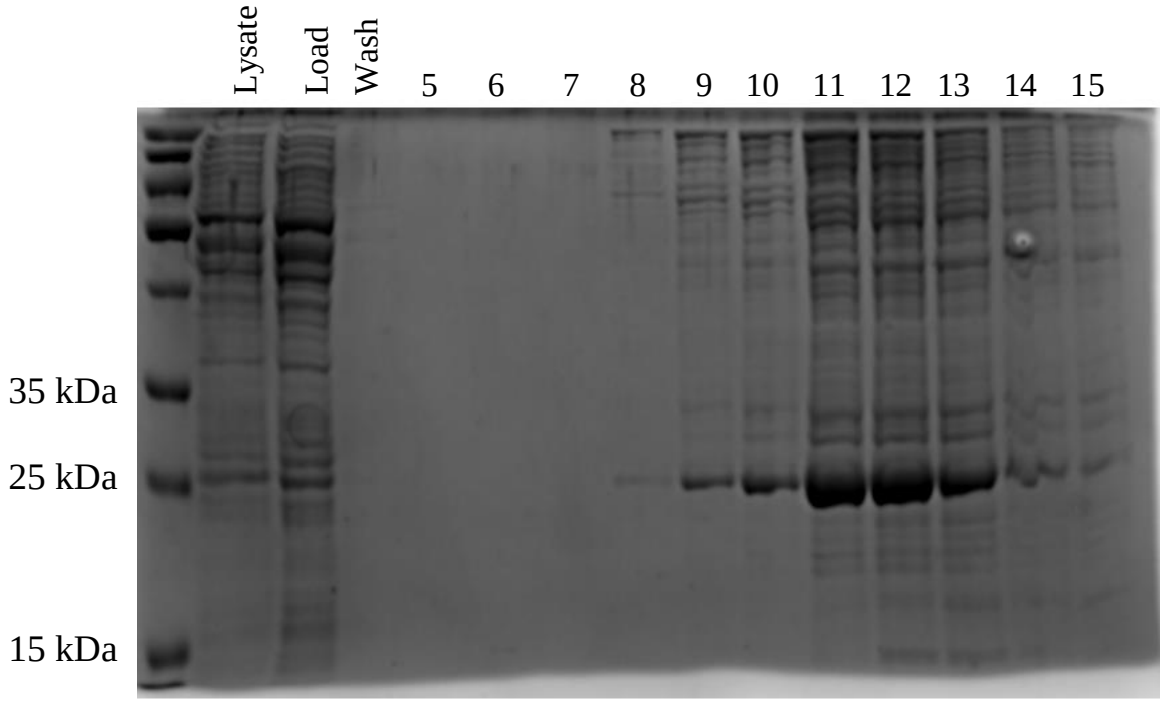

Figure 37: Purification results for CsgA supplemented with DnaK, DnaJ, and GrpE were analyzed by SDS-PAGE on a tris-glycine polyacrylamide gel. Samples were taken from the indicated elution fractions, as well as from the whole-cell lysate (lysate) and the load and wash flow throughs were recorded.

Native-PAGE, SDS-PAGE, and His-Staining of Natively Purified CsgA

Natively purified CsgA was concentrated and protein concentration estimated by Bicinchoninic acid assay (BCA), which provides colorimetric determination of protein concentration by measuring the reduction of Cu^{2+} to Cu^{+}. Dilutions of concentrated CsgA were analyzed by Native-PAGE, SDS-PAGE, and His-staining to verify purification results, understand the effect of SDS on the overall population of purified CsgA, as well as investigate concentration-dependent effects. Results are shown in figure 38. Results indicate SDS-induced oligomerization and verifies the identity of the multiple bands seen as His-tagged CsgA.

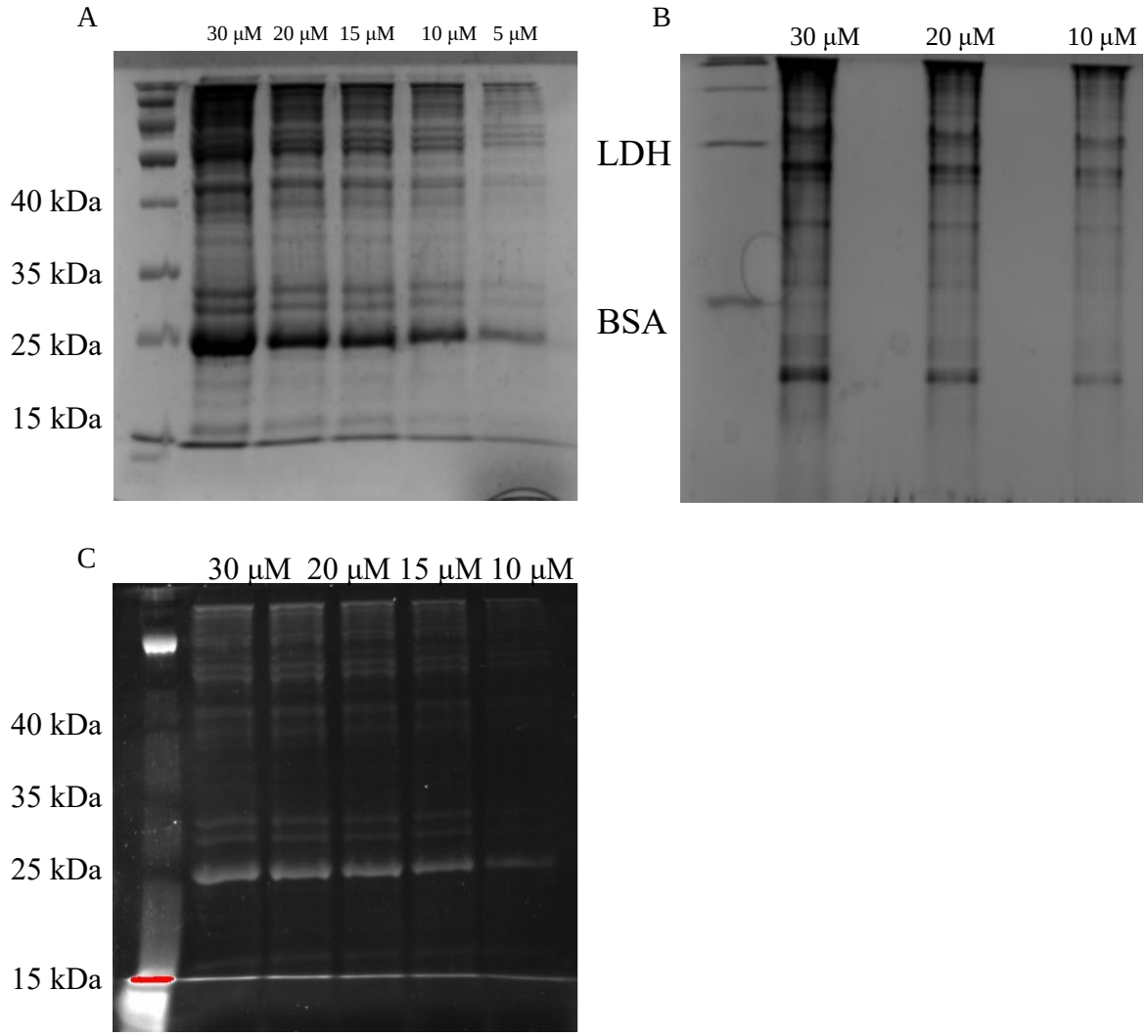

Figure 38: Natively-Purified CsgA was concentrated in 3.5 kDa MWCO dialysis tubing packed with absorbent. The purified product was analyzed by SDS-PAGE and Native-PAGE on tris-glycine polyacrylamide gels (Panels A-B). Panel C shows a duplicate gel of the one seen in Panel A, however (C) was stained with His-stain rather than coommassie and was imaged using the EtBr setting on a GelDoc Imager.

Circular Dichroism of Natively Purified E. coli CsgA (frozen vs unfrozen)

Natively purified CsgA was diluted to the indicated concentration and analyzed using a Chirascan (Applied Photophysics). Spectra were recorded from 180 nm to 350 nm in order to observe changes in the secondary structure of the protein through monitoring of protein backbone components. Spectra were background and buffer subtracted prior to conversion to molar elipcities (ME). The average ME of repeat reads, n=4, for natively purified CsgA prior to and post-freezing were recorded to evaluate structural pertubations arising from the freezing process. Spectra are shown below in figure 39. Results indicate glycerol addition did not perturb the average secondary structural characteristics of the population, as well as indicate α-helical properties.

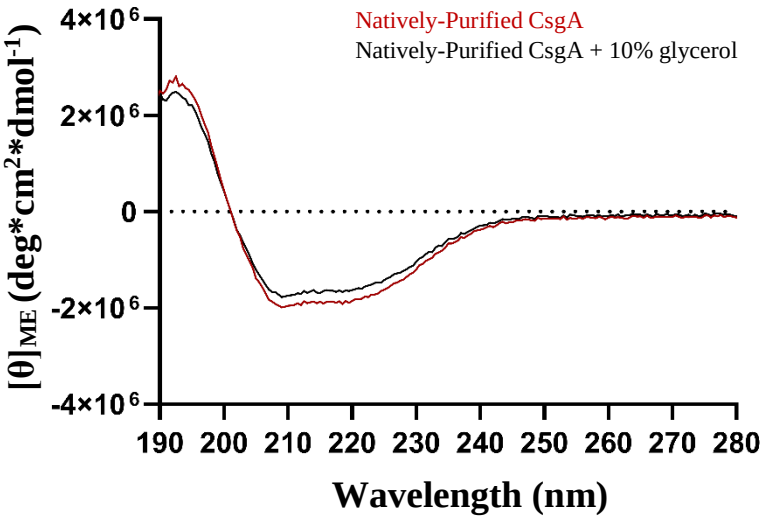

Figure 39: Natively-purified CsgA and flash-frozen natively-purified CsgA supplemented to 10% glycerol were buffer exchanged into 10 mM potassium phosphate pH 7.2 and diluted to 10 µM. Spectra were recorded from 180-360 nm at a 0.5 nm step-size. Repeat reads, n = 4, were taken for the buffer as well as each of the protein samples. Spectra were background and buffer subtracted prior to averaging and conversion to molar elipcities.

JC-1 Assay of E. coli CsgA

In the effort to observe oligomer formation in the population of CsgA, as well as assess for shared binding characteristics between αSyn and CsgA, natively purified *E. coli* CsgA was incubated at the indicated concentration of JC-1 and fluorescence emission spectra were recorded over time. Spectra were normalized and baseline corrected prior to plotting of fractional fluorescence intensity vs time at each peak. Recording of secondary analysis are presented in figure 40. Results indicate dose-dependent interaction of pre-formed oligomers with JC-1 aggregates, while oligomer formation as seen previously for αSyn was absent. [58]

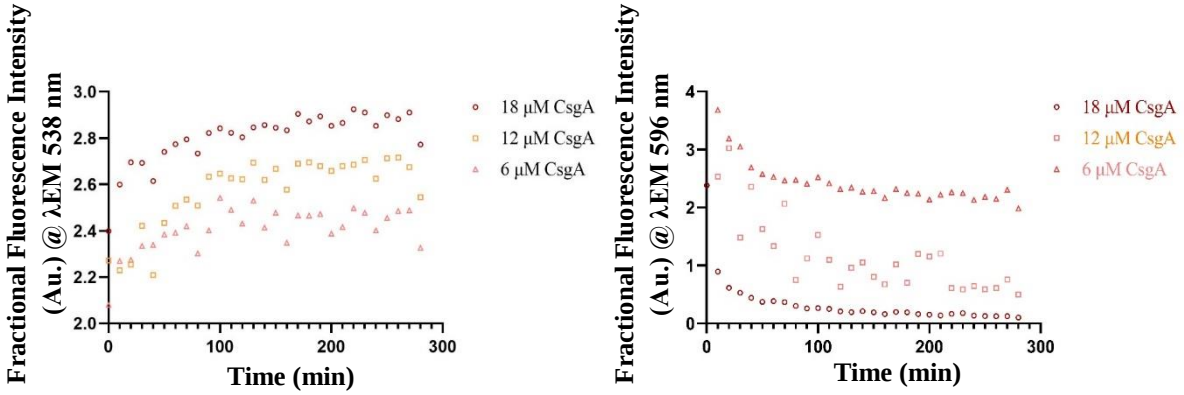

Figure 40: Reaction conditions included natively-purified CsgA at the indicated concentration, as well as 10 µM JC-1, 5% DMSO, in 20 mM HEPES pH 7.2. Each well of a 96-well Greiner black/clear-bottom was subjected to successive cycles of shaking at 37 °C for 10 minutes, followed by measurement of fluorescent emission spectra. Raw plots were baseline-corrected prior to secondary analysis shown above. The left shows the increase in fractional fluorescence intensity at 538 nm, while the right shows the decrease in fractional fluorescence intensity at 596 nm.

Size Exclusion Chromatography and Anion Exchange for Isolation of Monomeric E. coli CsgA

Efforts to isolate monomeric CsgA following native purification invited the usage of size exclusion chromatography or anion exchange chromatography. Following the native purification of *E. coli* CsgA as described above, eluants were subjected to either Fast Protein Liquid Chromatography with a sephacryl size exclusion column or HighQ anion exchange chromatography. Elution profiles for both purification approaches were recorded via SDS-PAGE and are shown in figure 41. Results indicate unsuccessful isolation of monomeric CsgA.

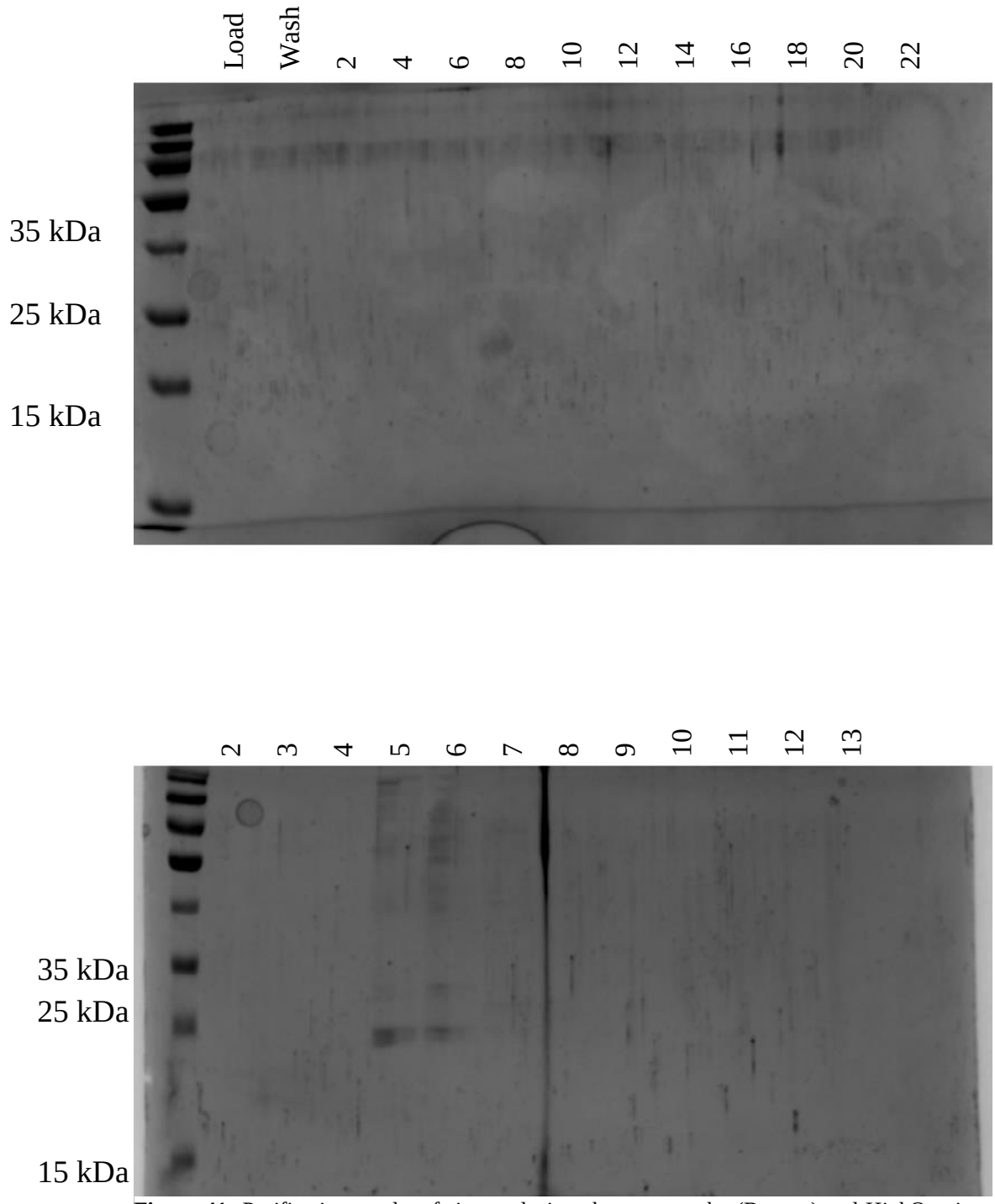

Figure 41: Purification results of size-exclusion chromatography (Bottom) and HighQ anion-exchange chromatography (Top) were analyzed by SDS-PAGE on a tris-glycine polyacrylamide gel. Fractions corresponding to the indicated elution fractions are shown for both approaches, while load and wash flow throughs were also recorded for HighQ anion exchange.

Chemically Induced Unfolding of E. coli CsgA

Natively purified *E. coli* CsgA was analyzed through chemically-induced unfolding mediated either by GuHCl or Urea to yield thermodynamic properties of the protein mixture, as well as investigate conformational changes in the protein through monitoring of localized unfolding. Purified *E. coli* CsgA was incubated with the indicated concentration of denaturant and allowed to reach equilibrium overnight. Solutions were excited at both 280 nm and 297 nm with spectra recorded from 300-650 nm. Spectra were normalized and baseline-corrected prior to analysis of changes in fractional fluorescence intensity at a specific wavelength, as well as shifts in maximum λEM. Spectra and secondary analysis are shown in figure 42.

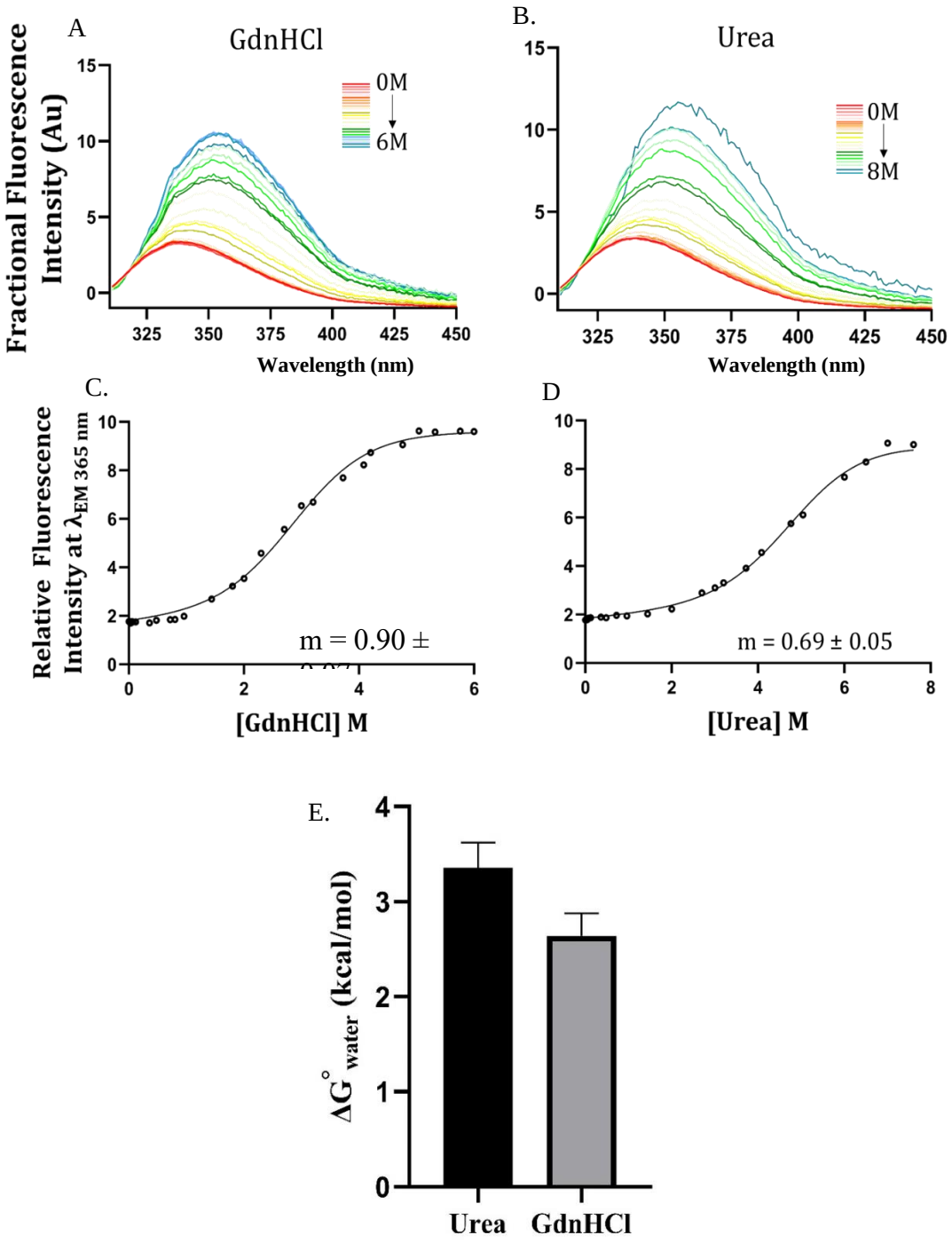

Figure 42: Chemically-induced unfolding was achieved with the usage of GuHCl (GdnHCl), or Urea. Raw emission spectra following excitation at 297 nm were normalized and buffer subtracted prior to plotting of fractional fluorescence intensity vs wavelength (A-B). The relative fluorescence intensities at 365 nm were both recorded vs concentration. (C-D) M-values and ΔG_{UN} were extrapolated using nonlinear regression and ΔG_{UN} values for each denaturant are shown in E. Error bars shown correspond to the upper limit of the 95% confidence interval.

On-Column Denaturation of E. coli CsgA

To investigate the potential effect of denaturant application to the conformational and oligomeric heterogeneity of *E. coli* CsgA, a modified protocol adapted from Chapman et al. was employed. [23] In brief, purified CsgA was batch bound to Ni-NTA and subjected to both a non-denaturant wash followed by washing with 8 M GuHCl and elution by pH depression. Purification results were analyzed by SDS-PAGE are shown below in figure 43. Results indicate a difference in the band profile of natively-purified CsgA present in the load flow through compared to the contents of the GuHCl-treated Elution.

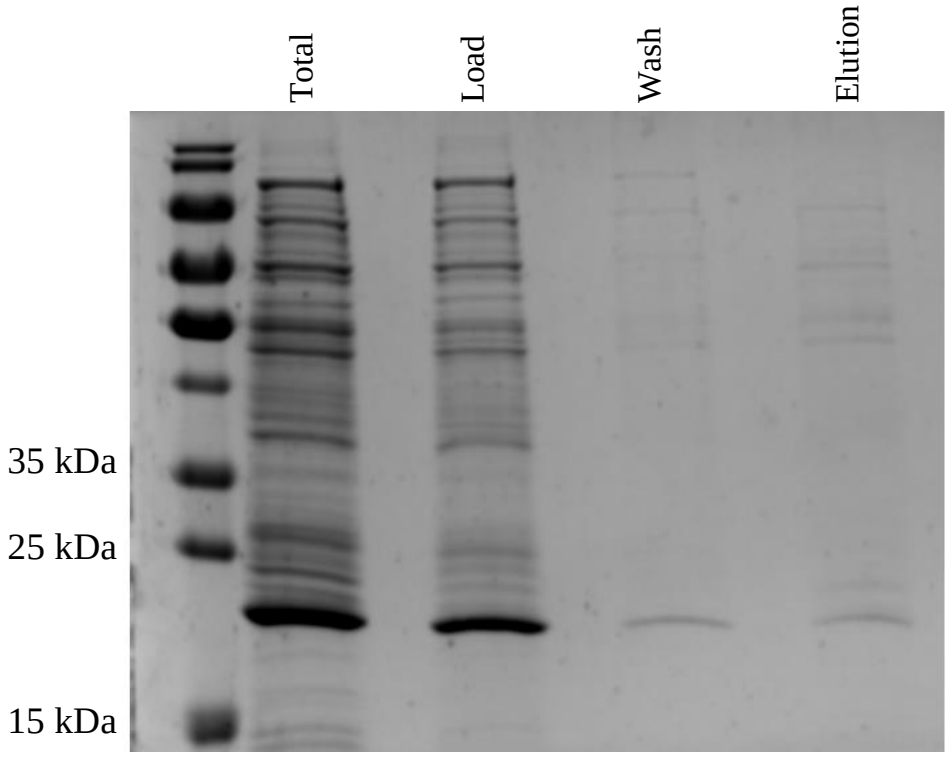

Figure 43: Results of the on-column denaturation were analyzed by SDS-PAGE on a tris-glycine polyacrylamide gel. Samples containing natively-purified CsgA untreated by GuHCl are shown in the whole-cell lysate (total) and the load and wash flow throughs. The sample corresponding to GuHCl treated CsgA is recorded within the elution lane.

Circular Dichroism of On-column Denatured E. coli CsgA

To qualitatively describe the effect of denaturant on secondary structure of the conformational ensemble, CsgA purified utilizing the on-column denaturation protocol was desalted using a PD10 column and buffer exchanged to ensure removal of denaturant. The on-column denatured CsgA was diluted to the indicated concentration and analyzed through circular dichroism using a Chirascan (Applied Photophysics). Spectra were background and buffer

subtracted prior to conversion to molar elipcities. Molar elipcities are reported in figure 44. Results indicate changes in the average of the secondary structure of the two populations.

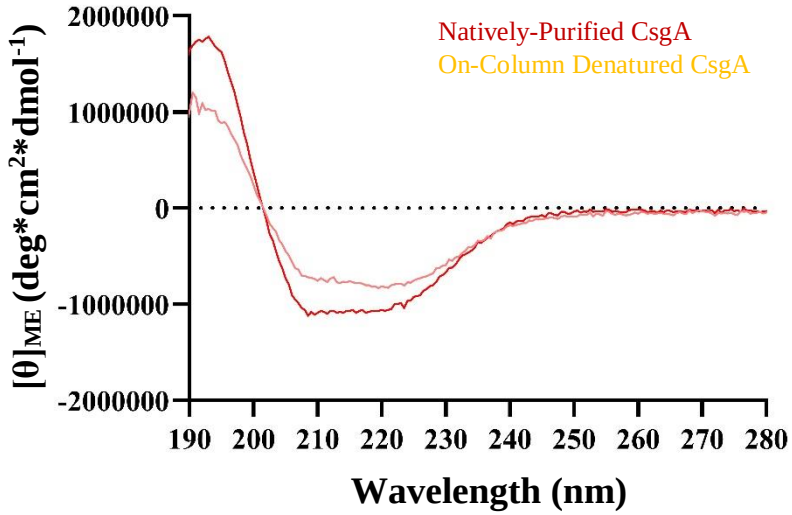

Figure 44: Cirular dichroism comparing the two purification approaches. Protein concentration (5 µM) was determined by BCA. Both natively-purified CsgA and on-column denatured CsgA were buffered in 10 mM potassium phosphate pH 7.2 . Spectra were recorded from 180 to 360 nm and read repeats, n = 4, were buffer and background subtracted prior to averaging and conversion to molar elipcities.

Comparison between Natively Purified and Denaturant Purified E. coli CsgA by SDS and Native-PAGE

Investigation of the effect of denaturant on the conformational heterogeneity was also investigated through a protocol designed to mimic solubilization of CsgA from inclusion bodies. While the on-column denaturation protocol required starting from natively purified CsgA, this protocol was designed to isolate CsgA directly from the cell lysate. Remaining steps of the

purification were in accordance with the modified protocol described by Chapman et al. [23] Comparison of denaturant-purified CsgA in the elution fraction to natively purified CsgA in the load fraction was conducted via SDS-PAGE and Native-PAGE. Results are recorded in figures 45 and 46. Results indicate a visually distinct band-profile when compared to natively-purified CsgA.

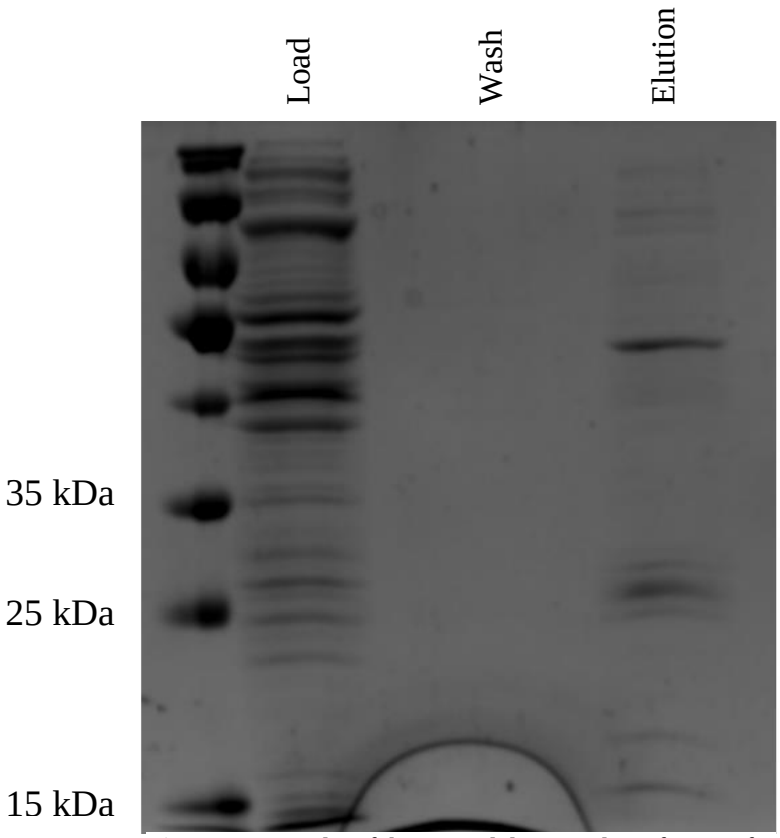

Figure 45: Results of the GuHCl denatured purification for extraction of CsgA from whole-cell lysate is shown above. The elution lane corresponds to a CsgA-containing fraction in the elution profile.

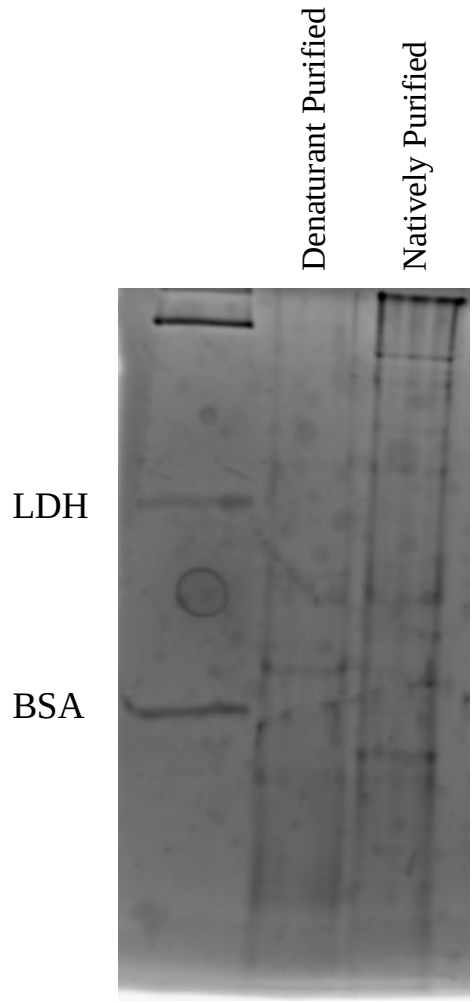

Figure 46: Denaturant-purified CsgA corresponding to a CsgA-containing fraction of the elution profile and natively-purified CsgA were analyzed by BCA assay to determine protein concentration prior to Native-PAGE analysis on a tris-acrylamide polyacrylamide gel.

Discussion

Initial efforts to express the structurally similar minor curli subunit CsgB, which possesses amyloidogenic activity and is believed to nucleate the fibrillation of CsgA, followed an approach taken for periplasmic expression of αSyn. [24] (Figure 33) This growth protocol includes induction with IPTG at 37 °C to contribute to cellular stress, an environment typically conducive to production of stress-response associated CsgA and CsgB. Multiple attempts and expression testing in HMS174(DE3) showed little improvement in detecting expression of CsgB, which may stem from its propensity to aggregate within the cell in the absence of CsgC – making oligomeric intermediates difficult to distinguish from the background of whole-cell-lysate.

To mitigate the aggregative propensity of CsgB and CsgA within the cell, bacterial growths were supplemented with the pKJE7 plasmid containing heat shock chaperone proteins DnaK, DnaJ, and GrpE. For both CsgA and CsgB, this approach provided an avenue for putative identification of CsgA from whole-cell-lysate. (Figure 34, 36)

Initial purification efforts were applied to CsgB, with clarified supernatant from whole-cell lysates incubated with magnesium chloride, ATP, and denatured *E. coli* proteins prior to isolation of CsgB via nickel affinity chromatography. Results of this purification indicated presence of CsgB in the eluant, however apparent chaperone contamination prompted refinement of the purification protocol. (Figure 35)

Refined purification attempts were applied to the major curli subunit CsgA. In the refined protocol, bacterial cell paste was resuspended in a lysis buffer containing magnesium chloride, ATP, and denatured *E. coli* proteins to increase the duration of exposure and enhance chaperone

removal. Additionally, column washes were extended to ensure an A280 less than 0.003 prior to elution. Results of this purification approach yielded detection of multiple identities through SDS-PAGE analysis. (Figure 37)

Results of the purification were verified through in-gel detection of 6x Histidine tagged proteins. This in-gel detection stain provides a method for specific detection of 6x Histidine tags following exposure to UV light. [60] (Figure 38) Interestingly, his-staining of gel electrophoresed through an SDS-PAGE gel identified the numerous protein bands visualized with coomasie staining as the only 6x histidine tagged protein, CsgA. This necessitates SDS-resistance of CsgA and its multimeric intermediates, which may be achieved through secondary-structural and multimerization interfaces with features resistant to unfolding. In addition, comparison between native-PAGE and SDS-PAGE of purified CsgA reveals an increase in the band number for protein analyzed by SDS-PAGE. This serves as potential evidence for SDS-induced oligomerization of CsgA. (Figure 38) This is not an unexpected effect as CsgA is known to polymerize at the extracellular membrane interface and SDS has previously been implemented as a lipid membrane mimetic. [27]

JC-1 is an aggregation-state dependent fluorescence probe that has exhibited selective activity towards αSyn. JC-1 exhibits differential fluoresence λEM depending on its aggregation state, sharing common excitation at 490 nm. It is hypothesized JC-1 binds to monomeric αSyn without perturbing its secondary structure, αSyn-JC-1 complexes are believed to encourage template-dependent multimerization of JC-1, resulting in the red-shifted fluorescence of JC-1 in its aggregated form. It has also been postulated that fibril interactions with monomeric JC-1 were marked by red-shift and increase in fluorescence intensity at ~538 nm. [58] JC-1 assays for monitoring of oligomer formation revealed increased prevalence of the fluorescence peak at 538

nm and a decrease in the fluorescence peak at 596 nm. The intermediate peak previously noted to correspond to oligomers of αSyn was not visualized. [58] (Figure 40) JC-1 interactions are protein-specific and predicted to be electrostatically-driven. It is postulated that interactions between monomeric amyloid subunits with monomers of JC-1 are responsible for the appearance of the intermediate fluorescent peak. Possibilities for the lack of this observation include potentially low concentrations of monomeric CsgA within the mixed sample, as well as requiring longer observation intervals. The decrease in fluorescence intensity at 596 nm and increase in fluorescence intensity at 538 nm could be attributed to interactions between pre-formed oligomers and J-aggregates in a dose-dependent manner. (Figure 40)

Secondary structural characteristics were probed via far-uv circular dichroism for analysis of the protein backbone characteristics. In contrast to previous documentation of denaturant-purified CsgA, which describes an unstructured monomer that proceeds through β-sheet rich oligomeric intermediates, it was noted that the average of the conformational ensemble of natively-purified CsgA possessed α-helical characteristics. Specifically, negative peaks at 208 nm and 222 nm, as well as a positive peak at 193 nm. [61] (Figure 39)

α-helical characteristics of the average of the conformational ensemble and the contrast this highlights between denaturant and natively purified CsgA prompted further characterization of the properties and behavior of the natively purified protein prior to direct comparison between the two purification approaches. To evaluate thermodynamic parameters like ΔG_{UN} and investigate localized unfolding transitions, natively purified CsgA was analyzed via chemically induced unfolding in the presence of either GuHCl or urea. (Figure 42) Nonlinear regression of fractional fluorescence intensity vs concentration of denaturant allowed for determination of ΔG_{UN}, which were calculated to be 3.38 kcal/mol and 2.64 kcal/mol for urea and GuHCl

respectively. M-values were also extrapolated via nonlinear regression of the same plot, resulting in m values of 0.69 and 0.90 respectively. (Figure 42) The differences in ΔG_{UN} can be attributed to the differing unfolding mechanisms between the two denaturants. GuHCl is believed to unfold proteins primarily through electrostatic interactions, while urea unfolds via typically-hydrophobic interactions. While there is a low thermodynamic threshold for unfolding in both GuHCl and urea, the higher energetic requirement and denaturant concentration for urea induced unfolding may stem from stable secondary structural elements mediated by electrostatic interactions. Both the low m-values and low values for ΔG_{UN} are characteristic of an intrinsically-disordered protein that undergoes little change in solvent accessible surface area throughout the unfolding process.

To directly compare the purification approaches, an on-column denaturation protocol adapted from Chapman et al. was implemented. [23] Results of the on-column denaturation protocol revealed decreased band number when compared to the natively purified protein present in the load flow through when analyzed via SDS-PAGE, as well as exhibiting differences in the average molar elipticity of its conformational ensemble through far-uv circular dichroism with loss of the 208 nm negative peak characteristic of α-helices. (Figure 43, 44) Efforts to mimic protocols for solubilization of CsgA from inclusion bodies were also performed, with SDS-PAGE and native-PAGE analysis supporting previous results. (Figure 45, 46)

This investigation into structural and behavioral features of the major curli subunit CsgA points to the necessity for native-purification of this protein to gain insight into the many structural and stoichiometric variations of this amyloid protein. We present evidence for the restriction of CsgA's native conformational variation in protein treated with denaturants like GuHCl and Urea.

Conclusions and Future Directions

Amyloidosis has long been evidenced in the neurodegenerative pathologies of AD and PD. In addition, recent evidence has implicated production of curli amyloid fibrils by enteric gut bacteria and subsequent cross-seeding between bacterial and endogenous amyloids in the onset and progression of these disorders. The work presented herein serves to identify potential targets for repurposed, existing anti-biofilm compounds to inhibit the production of these extracellular aggregative fimbrae, as well as identify an area of standard practice in working with CsgA that may restrict future investigations.

The response regulator CsgD serves as an attractive target for preventative and therapeutic treatments since the two-component systems and structures associated with them are features unique to prokaryotic populations – giving rise to the potential for host-inert therapeutics. Evidence indicates the ability of both *Ec* CsgD and *S*T CsgD in binding the 2-AI small molecules, however follow-up investigations with the full length response regulator *Ec* CsgD are critical. Evidence for 2-AI binding in *S*T CsgD was indicative of a preference for structural substituents containing a balance of polar and nonpolar motifs. In addition, the SUMO-CsgD approach and newly obtained smt3 protease for purification of *Ec* CsgD shows promise.

In regards to CsgA, the application of high concentrations of denaturant have become standard practice for the isolation of monomeric CsgA from either the cytoplasmic, soluble fraction of from insoluble inclusion bodies. However, when comparing secondary structural characteristics, SDS-PAGE, and Native-PAGE profiles, we were surprised to note alterations to the average secondary structure between the two purification approaches, as well as reduced heterogeneity of the sample in GuHCl purified CsgA.

Identification of potential targets for neurodegenerative therapeutics encourages further investigation to elucidate the full-length structures of both *Ec* CsgD and *S*T CsgD, as well as determination of their activation and DNA-binding mechanism for improved specificity of 2-AI small molecules. The discrepancies in purification approaches provide basis for continued investigation into the interaction of natively-purified CsgA and αSyn in the possibility that structures contained within the conformational ensemble of natively-purified CsgA are those responsible for cross-seeding and neuroinflammation seen in AD and PD pathologies. This work could also be extended broadly to investigate the effect of denaturant application to amyloid proteins studied in other human disease systems.

References

(1) Alzheimer's Association. 2022 Alzheimer's Disease Facts and Figures. **2022**.

(2) Marras, C.; Beck, J. C.; Bower, J. H.; Roberts, E.; Ritz, B.; Ross, G. W.; Abbott, R. D.; Savica, R.; Eeden, S. K. Van Den. Prevalence of Parkinson ' s Disease across North America. *npj Park. Dis.* **2020**, No. May 2018, 1–7. https://doi.org/10.1038/s41531-018-0058-0.

(3) Sipe, J. D.; Benson, M. D.; Buxbaum, J. N.; Ikeda, S.; Saraiva, M. J. M.; Westermark, P. Amyloid Fibril Protein Nomenclature : 2010 Recommendations from the Nomenclature Committee of the International Society of Amyloidosis. **2010**, *6129*. https://doi.org/10.3109/13506129.2010.526812.

(4) Avni, A.; Swasthi, H. M.; Majumdar, A.; Mukhopadhyay, S. *Intrinsically Disordered Proteins in the Formation of Functional Amyloids from Bacteria to Humans*, 1st ed.; Elsevier Inc., 2019; Vol. 166. https://doi.org/10.1016/bs.pmbts.2019.05.005.

(5) Stefanis, L. A -Synuclein in Parkinson ' s Disease. *Cold Spring Harb Perspect Med* **2012**.

(6) Article, R. The Amyloid- β Pathway in Alzheimer ' s Disease. *Mol. Psychiatry* **2021**, No. February. https://doi.org/10.1038/s41380-021-01249-0.

(7) Goodsell, D. Molecule of the Month: Amyloid-beta Precursor Protein. https://doi.org/10.2210/rcsb_pdb/mom_2006_7.

(8) Bendor, J.; Logan, T.; Edwards, R. H. The Function of α -Synuclein. *Neuron* **2014**, *79* (6). https://doi.org/10.1016/j.neuron.2013.09.004.The.

(9) Lee, C.; Nayak, A.; Sethuraman, A.; Belfort, G.; Mcrae, G. J. A Three-Stage Kinetic Model of Amyloid Fibrillation. *Biophys. J.* **2007**, *92* (10), 3448–3458. https://doi.org/10.1529/biophysj.106.098608.

(10) Verma, M.; Vats, A.; Taneja, V. Toxic Species in Amyloid Disorders : Oligomers or Mature Fibrils. *Ann. Indian Acad. Neurol.* **2015**, *18* (2). https://doi.org/10.4103/0972-2327.144284.

(11) Sengupta, U.; Nilson, A. N.; Kayed, R. EBioMedicine The Role of Amyloid- β Oligomers in Toxicity , Propagation , and Immunotherapy. *EBioMedicine* **2016**, *6*, 42–49. https://doi.org/10.1016/j.ebiom.2016.03.035.

(12) Sender, R.; Fuchs, S.; Milo, R. Commentary Are We Really Vastly Outnumbered ? Revisiting the Ratio of Bacterial to Host Cells in Humans. *Cell* **2016**, *164* (3), 337–340. https://doi.org/10.1016/j.cell.2016.01.013.

(13) Mohajeri, M. H. The Role of the Microbiome for Human Health : From Basic Science to Clinical Applications. *Eur. J. Nutr.* **2018**, *57* (1), 1–14. https://doi.org/10.1007/s00394-018-1703-4.

(14) Zheng, D.; Liwinski, T.; Elinav, E. Interaction between Microbiota and Immunity in Health and Disease. *Cell Res.* **2020**, No. April. https://doi.org/10.1038/s41422-020-0332-7.

(15) Breit, S.; Kupferberg, A.; Rogler, G.; Hasler, G.; Hasler, G. Vagus Nerve as Modulator of the Brain–Gut Axis in Psychiatric and Inflammatory Disorders. *Front. Psychiatry* **2018**, *9* (March). https://doi.org/10.3389/fpsyt.2018.00044.

(16) Parkinson, J. An Essay on the Shaking Palsy. *J Neuropsychiatry Clin Neurosci* **2002**, 223–236.

(17) Svensson, E.; Thomsen, R. W.; Djurhuus, J. C.; Pedersen, L.; Borghammer, P.; Sørensen, H. T. Vagotomy and Subsequent Risk of Parkinson ' s Disease. *Ann. Neurol.* **2015**, 522–529. https://doi.org/10.1002/ana.24448.

(18) Wang, C.; Yin, C.; Ma, F.; Zheng, C. Genome-Wide Screen Identifies Curli Amyloid Fibril as a Bacterial Component Promoting Host Neurodegeneration. *PNAS* **2021**, 1–12. https://doi.org/10.1073/pnas.2106504118.

(19) Scheperjans, F.; Aho, V.; Pereira, P. A. B.; Koskinen, K.; Paulin, L. Gut Microbiota Are Related to Parkinson' s Disease and Clinical Phenotype. *Mov. Disord.* **2015**, *30* (3), 9–12. https://doi.org/10.1002/mds.26069.

(20) Kikuchi, T.; Mizunoe, Y.; Takade, A.; Naito, S.; Yoshida, S. Curli Fibers Are Required for Development of Biofilm Architecture in Escherichia Coli K-12 and Enhance Bacterial Adherence to Human Uroepithelial Cells. *Microbiol. Immunol.* **2005**, *49* (9), 875–884.

(21) Chen, S. G.; Stribinskis, V.; Rane, M. J.; Demuth, D. R.; Gozal, E.; Roberts, A. M.; Jagadapillai, R.; Liu, R.; Choe, K.; Shivakumar, B.; Son, F.; Jin, S.; Kerber, R.; Adame, A.; Masliah, E.; Friedland, R. P. Exposure to the Functional Bacterial Amyloid Protein Curli Enhances Alpha-Synuclein Aggregation in Aged Fischer 344 Rats and Caenorhabditis Elegans. *Sci. Rep.* **2016**, *6* (October), 1–10. https://doi.org/10.1038/srep34477.

(22) Morales, R.; Moreno-gonzalez, I.; Soto, C. Cross-Seeding of Misfolded Proteins : Implications for Etiology and Pathogenesis of Protein Misfolding Diseases. *PLOS Pathog.* **2013**, *9* (9), 1–4. https://doi.org/10.1371/journal.ppat.1003537.

(23) Wang, X.; Smith, D. R.; Jones, J. W.; Chapman, M. R. In Vitro Polymerization of a Functional Escherichia Coli Amyloid Protein. *J. Biol. Chem.* **2007**, *282* (6), 3713–3719. https://doi.org/10.1074/jbc.M609228200.

(24) Hammer, N. D.; Schmidt, J. C.; Chapman, M. R. The Curli Nucleator Protein , CsgB , Contains an Amyloidogenic Domain That Directs CsgA Polymerization. *PNAS* **2007**. https://doi.org/10.1073/pnas.0703310104.

(25) Wang, X.; Hammer, N. D.; Chapman, M. R. The Molecular Basis of Functional Bacterial Amyloid Polymerization and Nucleation. *J. Biol. Chem.* **2008**, *283* (31), 21530–21539. https://doi.org/10.1074/jbc.M800466200.

(26) Wang, X.; Zhou, Y.; Ren, J.; Hammer, N. D.; Chapman, M. R. Gatekeeper Residues in the Major Curlin Subunit Modulate Bacterial Amyloid Fiber Biogenesis. *PNAS* **2010**, *107* (1), 163–168. https://doi.org/10.1073/pnas.0908714107.

(27) Swasthi, H. M.; Mukhopadhyay, S. Electrostatic Lipid – Protein Interactions Sequester the Curli Amyloid Fold on the Lipopolysaccharide Membrane Surface. *J. Biol. Chem.* **2017**, *292* (48), 19861–19872. https://doi.org/10.1074/jbc.M117.815522.

(28) Perov, S.; Lidor, O.; Salinas, N.; Nimrod, G.; Tayeb-, E. Structural Insights into Curli CsgA Cross- β Fibril Architecture Inspire Repurposing of Anti-Amyloid Compounds as Anti-Biofilm Agents. *PLOS Pathog.* **2019**, 1–31.

(29) Collinson, S. K.; Parker, J. M. R.; Hodges, R. S.; Kay, W. W. Structural Predictions of AgfA , the Insoluble Fimbrial Subunit of Salmonella Thin Aggregative Fimbriae. *J. Mol. Biol.* **1999**.

(30) Bhoite, S. S.; Han, Y.; Ruotolo, B. T.; Chapman, M. R. Mechanistic Insights into Accelerated α -Synuclein Aggregation Mediated by Human Microbiome-Associated Functional Amyloids. *J. Biol. Chem.* **2022**, *298* (7), 102088. https://doi.org/10.1016/j.jbc.2022.102088.

(31) Zhou, Y.; Smith, D. R.; Hufnagel, D. A.; Chapman, M. R. Experimental Manipulation of the Microbial Functional Amyloid Called Curli. *Methods Mol. Biol.* **2013**, No. 14, 1–21. https://doi.org/10.1007/978-1-62703-245-2.

(32) Hammar, M.; Arnqvist, A.; Bian, Z.; Olsdn, A.; Normark, S. Expression of Two Csg Operons Is Required for Production of Fibronectin- and Congo Red-Binding Curli Polymers in Escherichia Coil K-12. *Mol. Microbiol.* **1995**, *16*, 661–670.

(33) C. C. Goller and T. Romeo. Environmental Influences on Biofilm Development. *Curr. Top. Microbiol. Immunol.* **2008**, *322*, 37–66.

(34) Ahmad, I.; Cimdins, A.; Beske, T.; Römling, U. Detailed Analysis of C-Di-GMP Mediated Regulation of CsgD Expression in Salmonella Typhimurium. **2017**, 1–12. https://doi.org/10.1186/s12866-017-0934-5.

(35) Ogasawara, H.; Yamada, K.; Kori, A.; Yamamoto, K.; Ishihama, A. Regulation of the Escherichia Coli CsgD Promoter : Interplay between Five Transcription Factors. **2010**, 2470–2483. https://doi.org/10.1099/mic.0.039131-0.

(36) Sugimoto, S.; Yamanaka, K.; Niwa, T.; Terasawa, Y.; Kinjo, Y.; Mizunoe, Y.; Ogura, T. Hierarchical Model for the Role of J-Domain Proteins in Distinct Cellular Functions. *J. Mol. Biol.* **2021**, *433* (3), 166750. https://doi.org/10.1016/j.jmb.2020.166750.

(37) Brombacher, E.; Baratto, A.; Dorel, C.; Landini, P. Gene Expression Regulation by the Curli Activator CsgD Protein : Modulation of Cellulose Biosynthesis and Control of Negative Determinants for Microbial Adhesion. *J. Bacteriol.* **2006**, *188* (6), 2027–2037. https://doi.org/10.1128/JB.188.6.2027.

(38) Wen, Y.; Ouyang, Z.; Devreese, B.; He, W.; Shao, Y.; Lu, W.; Zheng, F. Crystal Structure of Master Biofilm Regulator CsgD Regulatory Domain Reveals an Atypical Receiver Domain. *Protein Sci.* **2017**, *26*, 2073–2082. https://doi.org/10.1002/pro.3245.

(39) Lau, Y. Y.; How, K. Y.; Yin, W.; Chan, K. Functional Characterization of Quorum Sensing LuxR-Type Transcriptional Regulator , EasR in Enterobacter Asburiae. *PeerJ* **2020**, 1–21. https://doi.org/10.7717/peerj.10068.

(40) Zakikhany, K.; Harrington, C. R.; Nimtz, M.; Hinton, J. C. D.; Römling, U. Unphosphorylated CsgD Controls Biofilm Formation in Salmonella Enterica Serovar Typhimurium. *Mol. Microbiol.* **2010**, *77* (July), 771–786. https://doi.org/10.1111/j.1365-2958.2010.07247.x.

(41) Manuscript, A. Vibrio Cholerae VpsT Regulates Matrix Production and Motility by Directly Sensing Cyclic Di-GMP. *Science (80-.).* **2011**. https://doi.org/10.1126/science.1181185.Vibrio.

(42) Ogasawara, H.; Yamamoto, K.; Ishihama, A. Role of the Biofilm Master Regulator CsgD in Cross-Regulation between Biofilm Formation and Flagellar Synthesis □ †. *J. Bacteriol.* **2011**, *193* (10), 2587–2597. https://doi.org/10.1128/JB.01468-10.

(43) Evans, M. L.; Chorell, E.; Almqvist, F.; Chapman, M. R.; Go, A.; Li, F.; Koch, M.; Sefer, L.; Evans, M. L.; Chorell, E.; Taylor, J. D.; Matthews, S. J.; Wittung-stafshede, P.; Almqvist, F.; Chapman, M. R. The Bacterial Curli System Possesses a Potent and Selective Inhibitor of Amyloid Formation Article The Bacterial Curli System Possesses a Potent and Selective Inhibitor of Amyloid Formation. *Mol. Cell* **2015**, 445–455. https://doi.org/10.1016/j.molcel.2014.12.025.

(44) Evans, M. L.; Chapman, M. R. Curli Biogenesis: Order out of Disorder. *BBA - Mol. Cell Res.* **2014**, *1843* (8), 1551–1558. https://doi.org/10.1016/j.bbamcr.2013.09.010.

(45) Evans, M. L.; Schmidt, J. C.; Ilbert, M.; Doyle, S. M.; Quan, S.; Bardwell, J. C. A.; Jakob, U.; Wickner, S.; Chapman, M. R. E. Coli Chaperones DnaK, Hsp33 and Spy Inhibit

Bacterial Functional Amyloid Assembly. *Prion* **2011**, *5* (4), 323–334. https://doi.org/10.4161/pri.5.4.18555.

(46) Sugimoto, S.; Arita-Morioka, K. ichi; Terao, A.; Yamanaka, K.; Ogura, T.; Mizunoe, Y. Multitasking of Hsp70 Chaperone in the Biogenesis of Bacterial Functional Amyloids. *Commun. Biol.* **2018**, *1* (1), 2–15. https://doi.org/10.1038/s42003-018-0056-0.

(47) Richards, J. J.; Huigens III, R. W.; Ballard, T. E.; Basso, A.; Cavanagh, J.; Melander, C. Inhibition and Dispersion of Proteobacterial Biofilms. *Chem. Commun.* **2008**, No. 14.

(48) Thompson, R. J.; Bobay, B. G.; Stowe, S. D.; Olson, A. L.; Su, Z.; Actis, L. A.; Melander, C.; Cavanagh, J. Identification of BfmR, a Response Regulator Involved in Biofilm Development, as a Target for a 2-Aminoimidazole-Based Anti-Biofilm Agent. *Biochemistry* **2013**, *51* (49), 9776–9778. https://doi.org/10.1021/bi3015289.Identification.

(49) Milton, M. E.; Minrovic, B. M.; Harris, D. L.; Kang, B.; Jung, D.; Lewis, C. P.; Thompson, R. J.; Melander, R. J.; Zeng, D.; Melander, C.; Cavanagh, J. Re-Sensitizing Multidrug Resistant Bacteria to Antibiotics by Targeting Bacterial Response Regulators: Characterization and Comparison of Interactions between 2-Aminoimidazoles and the Response Regulators BfmR from Acinetobacter Baumannii and QseB from Francisella Spp. *Front. Mol. Biosci.* **2018**, *5* (FEB), 1–12. https://doi.org/10.3389/fmolb.2018.00015.

(50) LifeSensors. SUMOstar ® Gene Fusion Technology lifesensors.com/wp-content/uploads/2019/09/1100K_1100A_1101__1106_Ecoli_SUMOstar_Manual.pdf.

(51) Mayer, M. P.; Bukau, B. Hsp70 Chaperones: Cellular Functions and Molecular Mechanism. *Cell. Mol. Life Sci.* **2005**, *62* (6), 670–684. https://doi.org/10.1007/s00018-004-4464-6.

(52) Schlager, B.; Straessle, A.; Hafen, E. Use of Anionic Denaturing Detergents to Purify Insoluble Proteins after Overexpression. *BMC Biotechnol.* **2012**.

(53) Tao, H.; Liu, W.; Simmons, B. N.; Harris, H. K.; Cox, C.; Massiah, M. A. Purifying Natively Folded Proteins from Inclusion Bodies Using Sarkosyl, Triton X-100, and CHAPS. *Benchmarks* **2010**, 61–63. https://doi.org/10.2144/000113304.

(54) Costa, S.; Almeida, A.; Castro, A.; Domingues, L. Fusion Tags for Protein Solubility , Purification , and Immunogenicity in Escherichia Coli: The Novel Fh8 System. *Front. Microbiol.* **2014**, *5* (February), 1–20. https://doi.org/10.3389/fmicb.2014.00063.

(55) Kojetin, D. J.; Thompson, R. J.; Benson, L. M.; Naylor, S.; Waterman, J.; Davies, K. G.; Opperman, C. H.; Stephenson, K.; Hoch, J. A.; Cavanagh, J. Structural Analysis of Divalent Metals Binding to the Bacillus Subtilis Response Regulator Spo0F : The Possibility for in Vitro Metalloregulation in the Initiation of Sporulation. *BioMetals* **2005**, 449–466. https://doi.org/10.1007/s10534-005-4303-8.

(56) Leitner, P. D.; Vietor, I.; Huber, L. A.; Valovka, T. Fluorescent Thermal Shift - Based Method for Detection of NF - KB Binding to Double - Stranded DNA. *Sci. Rep.* **2021**, No. 0123456789, 1–10. https://doi.org/10.1038/s41598-021-81743-1.

(57) Yan, Z.; Yin, M.; Chen, J.; Li, X. Assembly and Substrate Recognition of Curli Biogenesis System. *Nat. Commun.* No. 2020, 1–10. https://doi.org/10.1038/s41467-019-14145-7.

(58) Sewell, L.; Stylianou, F.; Xu, Y.; Taylor, J.; Sefer, L.; Matthews, S. NMR Insights into the Pre-Amyloid Ensemble and Secretion Targeting of the Curli Subunit CsgA. *Sci. Rep.* **2020**, *10* (1), 1–10. https://doi.org/10.1038/s41598-020-64135-9.

(59) Lee, J.; Lee, I.; Choe, Y.; Kang, S.; Kim, H. Y.; Gai, W.; Hahn, J.; Paik, S. R. Real-Time Analysis of Amyloid Fibril Formation of α -Synuclein Using a Fibrillation-State-Specific Fluorescent Probe of JC-1. *Biochem. J.* **2009**, *323*, 311–323. https://doi.org/10.1042/BJ20081572.

(60) ThermoFisher. Pierce ™ 6xHis Protein Tag Stain Reagent Set www.thermofisher.com/document-connect/document-connect.html?url=https://assets.thermofisher.com/TFS-Assets%2FLSG%2Fmanuals%2FMAN0011397_Pierce_6xHis_Protein_Tag_Stain_Reag_UG.pdf.

(61) Greenfield, N. J. Using Circular Dichroism Spectra to Estimate Protein Secondary Structure. *Nat. Protoc.* **2009**, *1* (6), 2876–2890. https://doi.org/10.1038/nprot.2006.202.Using.