THE DETERMINANTS OF BACTERIAL AMYLOID NUCLEATION AND POLYMERIZATION

By

Xuan Wang

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Molecular Cellular and Developmental Biology)
in The University of Michigan
2008

Doctoral Committee:

Assistant Professor Matthew R. Chapman, Chair
Professor James Bardwell
Professor Robert A. Bender
Professor Ari Gafni

To my wife, Yue

Acknowledgements

I would like to thank my mentor Dr. Matthew Chapman for his help, guidance and contributions to all of the work presented here. He has been extremely supportive during my whole graduate education, especially at my rough beginning. I have benefited a lot from the scientific environments he has built. I sincerely appreciate all the wise advice he gave me about my career and personal life. He has been a great mentor and role model for me. The parties he threw and his tasty grilled chicken wings will be part of my best memory in Ann Arbor.

I would like to thank the past and present lab members. They have contributed greatly to my science life and personal life in Ann Arbor. Without their support and friendship, my Ph.D life will be much less colorful. Dr. Michelle Barnhart and Elisabeth Ashman Epstein provided a lot of help at my beginning in Chapman lab. Especially, I sincerely appreciate Elisabeth to teach me English tongue twisters and her patience and help to improve all my manuscripts. I would like to thank my great lab bay mate, Neal Hammer, to provide numerous rides for me in the past four years. I also would like to thank him for introducing tailgating and many slang words to me. He has brought a lot of fun to my life. I would like to thank Daniel Smith for his significant contribution to my *in vitro* work presented here. It is always fun to have "team-building activities" with him. I would like to thank Ryan Frisch for his help on my projects. After four and a half years I still cannot figure out where he finds so many useful technology tips. Because of his

scientific skill with yeast his homebrewed beer tastes always good.

I would like to thank many contributors for their help during my graduate work. I would like to thank Daniel Smith and Jonathan Jones for their contributions to Chapter 2. Daniel Smith improved the CsgA purification strategy, developed the denaturing protocol for CsgA and investigated the size of the A11 antibody interacting species. Jonathan initially constructed the kinetic analysis of CsgA polymerization using the plate reader and discovered the self-seeding event of CsgA. The work presented in Chapter 2 was derived from Wang *et al*., 2007, J Biol Chem *282*, 3713-3719.

I would like to thank Neal Hammer for his contribution to Chapter 3. He provided the important reagent. The work presented in Chapter 3 was derived from Wang *et al.*, 2008, J Biol Chem *283*, 21530-21539.

I would like to thank Elisabeth Epstein, Neal Hammer and Ryan Frisch for their helpful comments for Chapter 4. The work presented in Chapter 4 was derived from Wang *et al.*, 2008, J Mol Biol *380*, 570-580.

I would like to thank Juan-Jie Ren, Yizhou Zhou and Neal Hammer for their contributions to Chapter 5. All of them contributed their work to make the gatekeeper story more comprehensive. Juan-Jie Ren and Yizhou Zhou performed some mutagenesis analysis presented in Chapter 5. Neal Hammer designed a new purification strategy and made the expression constructs used in Chapter 5. I also would like to thank Wei Dai for making constructs used in Chapter 6.

I would like to thank my committee members, Dr. Robert A. Bender, Dr. James

Bardwell and Dr. Ari Gafni for their support and valuable suggestions in all committee meetings.

I would also like to thank the Rackham graduate school to provide the Predoctoral Fellowship for the last year of my graduate education.

I would to like to thank my parents Jiarong Wang and Li Yu for their immeasurable support and valuable advice throughout my life. I am also indebted to my brother Ke Wang for his advice and inspiration. I would like to thank my old friend Dr. Yan Liu for his brilliant advice and warm encouragement. He is like my brother and I always can trust him and count on him. I would like to thank my friend Dr. Guoping Ren for providing many interesting suggestions to my projects. The discussion with him is always inspiring.

Finally and the most importantly, I would like to thank my wife Yue Cheng for her constant support and unconditional love.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AD, Alzheimer's disease

CD, circular dichroism

CR, Congo red

FA, formic acid

GdnHCl, guanidine hydrochloride

HFIP, hexafluoroisopropanol

KPi, potassium phosphate buffer

MTP, *M. tuberculosis* pili

R, repeating unit

TEM, transmission electron microscopy

TFA, trifluoroacetic acid

ThT, thioflavin T

Chapter 1

General Introduction[1]

Amyloid Biology

Amyloids are distinct proteinaceous fibers

Amyloids are distinct proteinaceous fibers traditionally associated with many human ailments such as neurodegenerative diseases and prion-based encephalopathies. Amyloid was first termed by Rudolph Vichow in 1854 to describe the abnormal substance in brain tissue stained positively by iodine (reviewed by Sipe and Cohen, 2000). Amyloid was derived from the Latin amylum and the Greek amylon since the chemical nature of amyloid was thought to be starch-like because it stained with iodine (Sipe and Cohen, 2000). Later, amyloid was demonstrated to be protein aggregates with ordered structures (Bonar et al., 1969; Glenner et al., 1969; Glenner et al., 1971; Glenner et al., 1971; Kim et al., 1969; Shirahama and Cohen, 1967). Amyloids could be stained with Congo red, displaying an apple-green birefringence when viewed with polarized light (Figure 1.1A) (Divry P, 1927). Besides their distinct tinctorial properties, amyloid deposits are ordered structures with distinctive filaments as detected by electron microscopy (EM) (Figure 1.1B). These fibers are 4-10 nm wide and long unbranched and found in various amyloid deposits from different amyloidotic tissues. X-ray diffraction of amyloid powders showed that these ordered proteinaceous structures have cross-beta

[1] Some part of this chapter is derived from Hammer *et al*., 2008, J Alzheimers Dis *13*, 407-419.

sheet structure, in which beta strands are oriented perpendicular to fibril axis (Sunde et al., 1997). Due to the aggregative properties of amyloidogenic proteins, the amyloid structure at atomic level has been difficult to solve. Solid-state NMR and X-ray microcrystallography have been successfully used to elucidate the amyloid structure of Aβ40 (Figure 1.1C), the fragments of yeast prions Sup35p (Figure 1.1D) and HET-s (Figure 1.1E) (Nelson et al., 2005; Petkova et al., 2002; Wasmer et al., 2008). In these amyloid structures at atomic resolution, β strands form either β helices by paring two β sheets or β solenoid composed of three β sheets perpendicular to their fibril axis (Figure 1.1C, 1.1D and 1.1E).

Amyloid diseases

Amyloid-like fibril aggregates are associated with at least 10 neurodegenerative diseases including Alzheimer's disease, Parkinson's disease and Huntington's disease (Chiti and Dobson, 2006). Amyloid fibers are also found in certain tissues proposed to be associated with more than 20 nonneuropathic amyloidoses in humans (Chiti and Dobson, 2006). The accumulation of similar amyloid aggregates in these various diseases suggests they may share some common etiologies

To better illustrate amyloid pathology and biogenesis, I will detail the biochemical and pathological characteristics of Aβ, arguably the most studied of all amyloids. The clinical and neuropathological characteristics of Alzheimer's disease (AD) were first reported by Alois Alzheimer (Alzheimer, 1907). Later, EM analysis identified long, unbranched 4-10 nanometer wide fibers to be the core component of these abnormal brain lesions (Kidd, 1963; Terry et al., 1964). The Aβ polypeptide was purified from AD associated plaques and was determined to be the major protein component of amyloid

plaques (Glenner and Wong, 1984; Wischik et al., 1988). Aβ is formed when the amyloid precursor protein (APP) is sequentially cleaved by β- and γ-secretases (Haass, 2004). It is proposed that APP plays important physiological roles in cell adhesion, neurite outgrowth, synaptogenesis and synapse remodeling (Zheng and Koo, 2006), however, the function of the Aβ polypeptide is currently unknown. There are two major cleavage products, Aβ40 and Aβ42 (Hartmann et al., 1997). The primary sequences of Aβ40 and Aβ42 only differ in that Aβ42 has 2 additional C-terminal residues, Ile^{41} and Ala^{42}. Mutations in presenilins, a central component of γ-secretase, account for most cases of familial AD. These mutations increase the production of Aβ42 in both transfected cells and transgenic mice (Citron et al., 1997).

Several lines of evidence link APP and misfolded Aβ to AD. Firstly, genetic evidence links chromosome 21 and familial AD, including missense mutations in the APP gene (Cheng et al., 1988; Patterson et al., 1988; St George-Hyslop et al., 1987). (Goate et al., 1991). Furthermore, AD like neuropathologies were observed in Down's syndrome (trisomy of chromosome 21) (Masters et al., 1985; Robakis et al., 1987). It is plausible that increased APP expression and high Aβ level induces cerebrovascular amyloidosis in Down's syndrome. In 2006, Rovelet-Lecrux et al showed that duplication of the APP locus on chromosome 21 in five families with autosomal dominant early-onset Alzheimer disease resulting in Aβ accumulation and early-onset Alzheimer disease (Rovelet-Lecrux et al., 2006). Secondly, various experimental *in vivo* models suggest a critical role of Aβ in disease development. Overexpression of a carboxyl-terminal fragment of APP in pheochromocytoma (PC12) and fibroblast (NIH 3T3) cell lines induced degenerative changes. Also overexpression of full-length human APP cDNA

caused degenerative changes to post-mitotic neurons (Yankner et al., 1989; Yoshikawa et al., 1992). Two different transgenic mouse models that express high level of human mutant APP developed many pathological hallmarks of AD (Games et al., 1995; Hsiao et al., 1996). Thirdly, there is an overwhelming amount of evidence showing *in vitro* chemically synthesized Aβ peptides are potent neurotoxins (Hartley et al., 1999; Lambert et al., 1998; Pike et al., 1991). Collectively, these findings point toward Aβ as the mitigating factor in AD development.

However, the molecular mechanism behind Aβ misfolding and AD development remain unclear. Hardy and Selkoe proposed the "amyloid cascade hypothesis" in which the central event in AD development is an imbalance between Aβ production and clearance leading to the accumulation and aggregation of Aβ (Hardy and Selkoe, 2002). The biochemical species of Aβ that induces pathogenesis is still a hotly debated topic. Experimental evidence suggests two major models: (1) the final amyloid fiber product causes neuronal damage or (2) neuronal cells are exposed to a toxic intermediate formed as Aβ polymerizes into the amyloid fiber.

Amyloid laden plaques are often found in post-mortem AD brains, which led to the suggestion that mature insoluble fiber aggregates of the causative agent for AD. However, statistical analyses have found only a weak correlation between the number of amyloid aggregates and the severity of AD (Lue et al., 1999; Terry et al., 1991). In addition, high molecular weight Aβ42 aggregates do not correlate with toxicity in the *Caenorhabiditis elegans* AD model (Cohen et al., 2006). The formation of high molecular weight protein aggregates may be a mechanism to protect cells from cytoxicity by sequestering the toxic intermediates present formed during Aβ misregulation (Arrasate et

al., 2004; Slow et al., 2005).

A wide range of nonfibrillar Aβ forms including dimer, trimer, oligomer, spherical aggregates and protofibrils have been reported to be cytotoxic and support this idea (Cleary et al., 2005; Hartley et al., 1999; Hoshi et al., 2003; Lambert et al., 1998; Lesne et al., 2006; Nilsberth et al., 2001; Roher et al., 1996; Townsend et al., 2006; Walsh et al., 2002). Different Aβ forms, including the small diffusible Aβ oligomer, high molecular weight oligomer, and fibers effect cortical neurons differently (Deshpande et al., 2006). Collectively, these data suggest nonfibrillar intermediates trigger neuropathologies. Therefore, the development of neurodegeneration could be induced by a complicated combinatory effect of several toxic Aβ conformers.

Despite evidence that prefibrillar aggregates may be the causative agents AD toxicity, many researchers have reported that mature Aβ fibers to be toxic to cultured neuronal cells (Busciglio et al., 1992; Deshpande et al., 2006; Hartley et al., 1999; Petkova et al., 2005; Tsai et al., 2004). How can this apparently conflicting data be reconciled? Recent study has demonstrated that the amyloid fiber is not a static structure. For instance, amyloid fibers formed from an SH3 domain showed very dynamic properties, in which molecules can be recycled by a dissociation and re-association mechanism within the fibril population (Carulla et al., 2005). Therefore, amyloid fibrils could provide a reservoir for toxic soluble oligomers, which could trigger the pathology (Haass and Selkoe, 2007). Under different experimental conditions amyloid fibrils may have different potentials to liberate soluble oligomers, which may be cytotoxic to the cultured neurons.

Nevertheless, the ability to effectively develop therapeutics that will deter

amyloid disease development is hindered due to the lack of experimental evidence defining a toxic species. Moreover, the molecular mechanism behind the initial misfolding events that convert soluble amyloid proteins into an amyloid fiber *in vivo* has not been forthcoming. Perhaps exploring systems where amyloid formation occurs as a natural functional process will provide answers to these questions.

Functional Amyloids

Although amyloid formation is traditionally associated with protein misfolding human disease, recent studies have identified functional amyloid fibers in bacteria, fungi, and mammals (Fowler et al., 2007; Hammer et al., 2008). Functional amyloid fibers are not the products of protein misfolding but those of protein folding and assembly, and they contribute toward the physiological well-being of the cell.

Bacterial functional amyloids

Amyloid fibers have distinct physical strength that is comparable by weight to steel (Smith et al., 2006). Amyloid fibers are also chemically stable and insoluble at some extent even in common denaturants such as sodium dodecyl sulfate (SDS), urea and guanidine hydrochloride (GdnHCl) (Masters et al., 1985). Bacteria can reside in relatively harsh environments that lack nutrients, or that are chemically challenging. Recent studies have shown that amyloid-like structures are abundant in natural biofilms (Larsen et al., 2007). It is plausible that bacteria use these strong fibers as extracellular organelles to interact with their environments to survive from the difficult environments (Barnhart et al., 2006; Epstein and Chapman, 2008).

The best understood bacterially produced amyloid fiber is curli. Curli fibers share

all the hallmarks of amyloids including staining thioflavin T and Congo red (Chapman et al., 2002). Curli fibers bound to Congo red also display apple-green birefringence under cross-polarized light (Lundmark et al., 2005). In addition, curli fibers are protease-resistant and SDS insoluble (Collinson et al., 1991). Curli fibers are 4-10 nm wide fibrous proteinaceous structure composed of curlin subunits, CsgA and CsgB (*csg* stands for curli specific gene) (Barnhart and Chapman, 2006). Curli fibers can mediate interactions between individual bacteria, between bacteria and host tissues, and even to inert surfaces that often resist bacterial colonization, like Teflon and stainless steel (Bokranz et al., 2005; Gophna et al., 2001; Gophna et al., 2002; Ryu and Beuchat, 2005; Uhlich et al., 2002; Zogaj et al., 2003). Curli induce a potent host inflammatory response, initiate binding to host cells, and increase the ability of the bacteria to persist within the environment and the host (Barnhart and Chapman, 2006).

The chaplins are another well-known example of bacterial extracellular amyloid structure produced by the Gram-positive bacterium *Streptomyces coelicolor*. *S. coelicolor* are soil bacteria and produce aerial hyphae growing into the air that are critical for their spore formation. In the absence of the chaplins, development of aerial hyphea is impaired (Elliot et al., 2003). It was proposed that the amyloid properties of chaplins may reduce the surface tension at the water/air interface by its hydrophobic nature to free hyphae to grow up into the air (Claessen et al., 2003; Claessen et al., 2004). Chaplins isolated from the *S. coelicolor* cell wall can form thioflavin T-positive, β-sheet-rich fibers (Claessen et al., 2003). The assembly of chaplins is a coordinated process. Chaplins in *S. coelicolor* are encoded by *chpA-H* and ChpE and ChpH are secreted and assembled into a network of amyloid fibers at the air-water interface (Claessen et al., 2003). ChpC, ChpE and

ChpH were shown to be most important for assembling chaplins (Di Berardo et al., 2008)

Since bacterial amyloid structures are abundant in the native environment (Larsen et al., 2007), the number of bacterially produced amyloids described is increasing. Two recent discoveries showed that extracellular structures of *Mycobacterium tuberculosis* and bacterial endospore from *Bacillus* and *Clostridium* have amyloid characteristics (Alteri et al., 2007; Plomp et al., 2007). *M. tuberculosis* pili (MTP) shows that they are non-branching fibers similar to curli and other amyloid fibers as detected by electron microscopy (Alteri et al., 2007). Isolated MTP also showed biochemical properties similar to amyloid fibers such as SDS-insolublity and Congo red binding (Alteri et al., 2007). Bacteria like *Bacillus* and *Clostridium* can produce an endospore when the nutrients are limited. Atomic force microscopic analysis showed that the spore coat of *Bacillus atrophaeus* is composed of fibers similar to amyloid (Plomp et al., 2007). It is plausible that the amyloid properties make the endospore an extremely stable structure highly resistant to pH extremes, heat and radiation as a protective coat.

Besides using amyloid fibers to perform physiological tasks, bacteria also use toxic amyloid intermediates to defend themselves. *Klebsiella pneumoniae* produces Microcin E492 (MccE492), assembled into oligomeric pores to induce the death of neighboring enteric bacteria (Destoumieux-Garzon et al., 2003). Like other oligomeric amyloid intermediates (Quist et al., 2005), MccE492 has been shown to form voltage independent ion channels in planar lipid bilayers (Lagos et al., 1993). MccE492 loses its cytotoxic properties when it polymerizes into amyloid fibers (Bieler et al., 2005). The cytotoxicity of MccE492 was proposed to be correlated to its pre-amyloid oligomric state (Bieler et al., 2005). It is possible that the toxic MccE492 species share some common

structural features and toxic mechanism with disease-associated amyloid intermediates. Another interesting example that bacteria used the toxic properties of amyloid is hairpin produced by some plant pathogens like *Xanthomonas* and *Psedomonas* species (Oh et al., 2007). Hairpins are secreted virulence factors that induce the hypersensitive response in plants (Kim et al., 2003). The hypersensitive response is a plant defense mechanism that inhibits intracellular pathogen growth by eliciting plant cell death akin to apoptosis in animal cells (Pennell and Lamb, 1997). It was recently shown that HpaG, a harpin produced by *Xanthomonas*, can polymerize into amyloid-like fibers (Oh et al., 2007). In contrast to MccE492, injection of HpaG protofibrils and mature fiber into plant cells induced cell death (Oh et al., 2007). The competence to induce hypersensitive response appears to be correlated to amyloidogenic tendency of Hairpin proteins (Oh et al., 2007). However, the exact cytotoxicity mechanism induced by hairpins remains elusive.

Fungal functional amyloids

Amyloid-like structures were discovered in various fungal species. Actually much of our knowledge concerning prion infectivity comes from a lot of excellent work using fungal prion as a model system (Shorter and Lindquist, 2005). Prion proteins were initially identified in the context of the transmissible spongiform encephalopathies (TSEs), the fatal neurodegenerative disorders including Creutzfeldt-Jakob disease and kuru in humans and scrapie and mad cow disease in animals (Prusiner et al., 1998). A prion is defined as an infectious protein (Prusiner et al., 1998). Disease-associated prions with amyloid properties could convert the soluble prion with a native conformation into disease-associated conformation (Prusiner et al., 1998). Prion proteins have a high propensity to assemble into self-propagating amyloid fibers under physiological

conditions both *in vivo* and *in vitro* (Glover et al., 1997; Prusiner et al., 1983; Prusiner et al., 1998; Taylor et al., 1999). Just like amyloid formation, preformed fibers composed of prion proteins could function as "seeds" to promote the conformation conversion of prion proteins in the native fold (Glover et al., 1997). This seeded catalysis of polymerization is proposed to be critical for prion infectivity (Prusiner et al., 1998). The definitive evidence to show prion infectivity is to induce prion-based phenotype by purely recombinant proteins with amyloid properties. The failure to accomplish this with recombinant animal prions and wild-type animal raised the doubts about the prion hypothesis. However, the first definitive evidence for "protein only" hypothesis was achieved by Maddelein *et al* (Maddelein et al., 2002). They successfully infected wild-type filamentous fungus *Podospora anserine* with *in vitro* preformed amyloid fibers from *P. anseria* prion protein HET-s, suggesting amyloid fibers alone can harbor prion infectivity (Maddelein et al., 2002).

The cellular function of HET-s in filamentous fungus *P. anserine* is associated with heterokaryon incompatibility (Coustou-Linares et al., 2001). Cells fusions could occur between adjacent individual cells of the filamentous fungus *P. anserine* and result in multinucleated cells called heterokaryon (Dales et al., 1983). The *het-s* locus of *P. anserine* has two alleles, *het-s* and *het-S* which encode HET-s and HET-S, respectively. HET-S is soluble whereas HET-s could be soluble or insoluble aggregates in the cytosol (Coustou et al., 1997). When hyphae that contain HET-s aggregates fuse with cells containing HET-S, the heterokaryon occurs and results in cell death to prevent further fusion (Coustou et al., 1997). Even though the biological significance of heterokaryon incompatibility is not completely understood, it is suggested that prion state of HET-s

play a beneficial role potentially in reducing the transmission of viruses or other infectious agents (Shorter and Lindquist, 2005). The nonamyloidogenic mutants of HET-s also lost its infectivity, suggesting the infectivity of HET-s aggregates is due to its amyloid properties (Ritter et al., 2005).

Yeast prion proteins such as Sup35p, Ure2p and Rnq1p have been extensively studied. They all have Gln/Asn rich domain essential for their amyloidogenic properties. Sup35p, Ure2p and Rnq1p can form amyloid-like aggregates in the cytosol, leading to a certain phenotype heritable by daughter cells in a non-Mendelian manner (Ross et al., 2005; Shorter and Lindquist, 2005). Sup35p is a termination factor in its native conformation (Liebman and Cavenagh, 1980). The aggregation of Sup35p leads to the increased read through of the stop codon and alters translation-termination fidelity (King et al., 1997). This prion phenotype [PSI^+] was thought to be beneficial for yeast because yeast populations with [PSI^+] have selective growth advantage under some conditions and it can confer both phenotypic and evolutionary plasticity (True et al., 2004; True and Lindquist, 2000).

In a good nitrogen source, native Ure2p regulates nitrogen catabolism by binding to the positive transcriptional regulator Gln3 and repressing the genetic network to use poor nitrogen sources (Blinder et al., 1996). When it forms amyloid-like aggregates, it loses its repressing activity (Masison and Wickner, 1995; Wickner, 1994). The potential benefits of Ure2P amyloid state remains unclear. It was reported that yeasts proliferate more rapidly without Ure2p activity compared to wild-type strain under the salt stress (Crespo et al., 2001). In addition, a recent study showed that Ure2p missing the amyloidogenic Gln/Asn rich domain had substantially reduced stability and activity

(Shewmaker et al., 2007).

Rnq1p was identified as an epigenetic modifier to induce the prion phenotype [*PSI*$^{+}$] in yeasts (Sondheimer and Lindquist, 2000). The cellular function of Rnq1 remains poorly understood. However, Rnq1p in amyloid state is commonly found in natural yeast isolates, suggesting its potential beneficial roles (Shorter and Lindquist, 2005). It is proposed that Rnq1p in amyloid state facilitate protein-folding transition including regulating [*PSI*$^{+}$] phenotype (Shorter and Lindquist, 2005).

Amyloid-like structures are also invoked by *Candida albicans*, the most common fungal opportunistic pathogen, as cell surface-expressed adhesins (Otoo et al., 2008; Rauceo et al., 2004). *Candida albicans* has broad specificity in adhering to host tissues, and adherence is followed by biofilm formation and colonization (Blankenship and Mitchell, 2006). Among many adhesins produced by *C. albicans,* the agglutinin-like (Als) proteins have been implicated in pathogenesis and biofilm formation (Chandra et al., 2001). Als adhesins were shown to have amyloid-like properties and these amyloid-like properties of Als proteins mediate multicellular aggregation reminiscence of bacterial extracellular amyloid-like matrix (Otoo et al., 2008; Rauceo et al., 2004).

Animal functional amyloids

Functional amyloid-like structures are also present in various animals (Fowler et al., 2007). A functional amyloid composed of the protein Pmel17 was reported to function as a template and to accelerate the biosynthesis of melanin polymers, which protects against UV and oxidative damage in various mammalian species such as mouse (Fowler et al., 2006). Pmel17 is synthesized as a transmembrane protein transiently localized in endoplasmic reticulum membrane. It is transported to Golgi and then to melanosome

organelles, specialized membrane enclosed vesicles where the melanin is synthesized (Raposo et al., 2001). In a post-Golgi compartment, the luminal fragment (Mα) was cleaved from its transmembrane domain (Mβ) and the degradation of Mβ free Mα to initiate amyloid polymerization process (Fowler et al., 2006). Mα has remarkably high amyloidogenic tendency as shown *in vitro* (Fowler et al., 2006). The Mα fibers bind the melanin precursors and facilitate the synthesis of melanin (Fowler et al., 2006). This suggests that amyloid is a conserved protein quaternary structure produced from bacteria to mammals contributing to normal cell physiology. Moreover, cells have evolved the mechanism to control amyloid propagation. The proteolysis of Pmel17 is critical to initiate amyloid formation, which is tightly coordinated and controlled.

Amyloid was also proposed to be the molecular basis of certain long-term memory formation (Si et al., 2003). Cytoplasmic polyadenylation element binding proteins (CPEBs) are highly conserved RNA-binding proteins localized at neuronal synapses that promote the elongation of the polyadenine tail of messenger RNA (Hake and Richter, 1994). In *Aplysia* (sea slug), a neuron-specific isoform of CPEB was reported to be important for long-term synaptic facilitation, but not for early expression (Si et al., 2003). This isoform of CPEB is conserved in *Drosophila*, mouse and human (Si et al., 2003). Interestingly, this CPEB has an N-terminal Gln/Asn rich domain very similar to the prion domains of yeast prions (Si et al., 2003). CPEB protein at full length has amyloidogenic properties and can induce prion-based phenotype in a yeast model (Si et al., 2003). It is intriguing to investigate whether the amyloid state of CPEB linked to its activity. The CPEB at the prion state is more active than it at the non-prion state, suggesting the prion property promotes CPEB biological function (Si et al., 2003).

However, the existence of CPEB at the prion state in *Aplysia* remains undescribed. It is plausible that the high tolerant amyloid properties could constitute the molecular basis of long-term memory due to its physical and chemical endurance.

The discovery of functional amyloids suggests that amyloid may be an ancient and conserved quaternary structure. Functional amyloids provide a unique niche to understand how nature modulates amyloid propagation and prevents cytotoxicity associated with amyloidogenesis.

Amyloid Polymerization Mechanism

Amyloid formation is proposed to be a common property of most proteins by Dobson and his colleagues (Chiti and Dobson, 2006). The denatured globular proteins such as muscle myoglobin can be converted into amyloid fibrils *in vitro* (Fandrich et al., 2001). Polypeptide SH3 domain which is not associated with diseases can be converted into amyloid fibrils and early nonfibrillar intermediates are inherently highly cytotoxic, suggesting a common mechanism for amyloid proteotoxicity (Bucciantini et al., 2002). Even though their primary sequences are very different, all amyloid proteins including disease-associated or not disease-associated assemble into structurally similar fibrils. During early *in vitro* aggregation, they share a structurally similar transient species which can be recognized by a conformational-specific antibody A11 (Figure 1.2) (Kayed et al., 2003). A11 reactive species of different disease-associated amyloid proteins are cytotoxic and this toxicity can be inhibited by A11, suggesting a common pathogenesis mechanism for different amyloid diseases (Kayed et al., 2003).

In vitro many amyloid proteins polymerize into ordered fibers follows a triphasic

process (Figure 1.2A black line). The polymerization is proposed to follow nucleation-dependent kinetics like protein crystallization, microtubule assembly, flagellum assembly, and virus coat assembly (Jarrett and Lansbury, 1993; Lomakin et al., 1996). Before mature fiber aggregates are detectable, there is a time period where the polymerization of amyloidogenic polypeptides appears to be stagnant. However, during this lag phase trace amounts of dimer, trimer, and eventually, nucleus (oligomer) are formed (Figure 1.2B) (Podlisny et al., 1995; Podlisny et al., 1998; Walsh et al., 1999). Nucleus formation is the rate-limiting step of fibril assembly. Once a nucleus has formed, monomer addition to the growing fiber becomes thermodynamically favorable and occurs quickly (Figure 1.2B) (Jarrett and Lansbury, 1993). As with any dynamic polymerization process where different folding intermediates are present at any one time, the monomer, oligomer, protofibrils (short fibrillar aggregates) and fibrils have been observed using different techniques including atomic force microscopy (Roher et al., 1996; Walsh et al., 1999).

Nucleation is the rate-limiting step of *in vitro* amyloid polymerization and is potentially critical to control amyloidogenesis *in vivo*. The nucleation might represent the latent phase of the development of various amyloid diseases (Selkoe, 2003). Most of amyloidoses and neurodegenerative diseases occur in middle or late life (Selkoe, 2003). Presumably, the factors promoting nucleation or shortening the latent lag time of amyloidogenesis will enhance the disease development. Since amyloid polymerization is similar to the process of crystallization, Jarrett and Lansbury proposed “seeding” may play an important role occur in amyloid assembly like crystallization (Jarrett and Lansbury, 1993). In facts, amyloid can promote its own polymerization, where soluble subunits can be efficiently incorporated into the preformed fiber (Figure 1.2A red line).

This template-directed polymerization or seeding potentially drives prion infectivity and the development of noninfectious amyloid diseases (Harper and Lansbury, 1997; Prusiner et al., 1998; Shorter and Lindquist, 2005). For instance, serum amyloid A amyloidosis (AA amyloidosis) in different animal models were enhanced by providing different preformed amyloid fibrils (Lundmark et al., 2002; Lundmark et al., 2005; Sigurdsson et al., 2002; Zhang et al., 2008). Even though nucleation and seeding is very important, the nature of this template-driven polymerization remains elusive. In this thesis, I used a functional amyloid propagation system to extensively elucidate this process.

Sequence Determinants of Amyloid Formation

Anfinsen proposed that the structure of a protein is determined by its specific amino acid sequence (Anfinsen and Scheraga, 1975; Haber and Anfinsen, 1961). Interestingly, amyloid formation seems an inherent property of a polypeptide chain independent of the amino acid sequence (Chiti and Dobson, 2006). Amyloid structure was reported to be dictated by main chain interactions (Fandrich and Dobson, 2002). Amyloidogenic proteins do not necessarily share any similarity at amino acid level even though the assembled fibers share the common cross-beta structure (Chiti and Dobson, 2006). However, there were many reports that the mutation or modification of the certain side chains changed protein aggregation and therefore changed the pathological effects (Chamberlain et al., 2001; Games et al., 1995; Hsiao et al., 1996; Van Nostrand et al., 2001). It is plausible that amyloidogenic proteins use diverse side-chain interactions to facilitate polymerization. For instance, hydrophobic stretch of Aβ42 instead of the specific sequence is sufficient for its aggregation (Kim and Hecht, 2005). It was proposed

that the interaction between aromatic residues in IAPP fragment is essential for its aggregation *in vitro* (Azriel and Gazit, 2001). Yeast prion proteins Sup35p, Ure2p and Rnq1p all have Gln/Asn rich domains that are essential for prion propagation (Ross et al., 2005; Shorter and Lindquist, 2005). Gln/Asn rich domains were proposed to have a high propensity to form self-propagating amyloid fibrils (Michelitsch and Weissman, 2000). Moreover, the specific sequences in the Gln/Asn rich domain of Sup35p govern its self-association and species-specific seeding activity (Tessier and Lindquist, 2007). Gln repeats are involved in the development of Huntington's and Kennedy's diseases (Scherzinger et al., 1997). The CAG repeat expansion of Huntingtin gene results in the increase of the length of poly-Gln stretch and the self-aggregation tendency *in vitro* (Chen et al., 2002). When CAG expansion is above the threshold, typically 37 Gln residues, huntingtin forms aggregates and induces the neuron death (Chen et al., 2002). This poly-Gln codon expansion causes amyloidogenic pathology has been validated in *Caenorhabditis elegans* (Faber et al., 1999), zebrafish (Schiffer et al., 2007) and mouse model (Reddy et al., 1998). It is plausible that poly-Gln can facilitate the polypeptide to form amyloid, therefore inducing degenerative pathology.

Dobson and Chiti *et al*. proposed that simple intrinsic physicochemical properties, such as hydrophobicity, secondary structure propensity and charge of unfolded polypeptides were correlated to aggregation rate (Chiti and Dobson, 2006; Chiti et al., 2003). However, this prediction of amyloidogenecity only based on global amino acid composition is limited because it did not consider the positional effect. In fact, the strong positional effects for proline substitution in Aβ and the point mutation for amyloidogenic peptide STVIIE were observed (Lopez de la Paz and Serrano, 2004; Williams et al., 2004;

Williams et al., 2006). My data presented in Chapter 4 demonstrates that CsgA protein has critical residues that drive self-aggregation and assembly into curli fibers. Secondly, this prediction algorithm did not consider the cellular factors influencing amyloidogenesis. For instance, Hsp104 plays an important role in yeast prion Sup35p polymerization (Shorter and Lindquist, 2004) and CsgB in *E. coli* initiates CsgA polymerization *in vivo* (Hammar et al., 1996). A protein that has high intrinsic aggregation properties does not necessarily efficiently assemble into fibers *in vivo* if its sequence disfavors the interactions with other important factors in amyloidogenesis. One interesting example like this was shown in Chapter 3.

Many amyloidogenic proteins including synuclein, yeast prions, HET-s, human prion, bacterial amyloid protein CsgA, CsgB *etc.* have imperfect repeated sequences. These repeats constitute the protease-resistant core and were shown to play important role in fiber assembly (Ross et al., 2005). For instance, the N-terminal prion-determining domain of Sup35p has five imperfect oligopeptide repeats, and certain deletions of the repeats resulted in the defectiveness in propagation of Sup35p fibrils (Liu and Lindquist, 1999). Moreover, *in vitro*, repeat expansion peptides (with two extra repeats) were shown to be more amyloidogenic than wild-type peptides (Liu and Lindquist, 1999). It is plausible that intramolecular interactions within repeated sequence facilitate amyloid formation.

There have been several prediction algorithms reported to successfully identify amyloidogenic regions (Fernandez-Escamilla et al., 2004; Tartaglia and Vendruscolo, 2008). These aggregation-prone sequences are very abundant in globular proteins (Linding et al., 2004). Cellular molecular chaperons and quality control mechanisms are

employed to prevent uncontrolled protein aggregation. Besides these mechanisms, it has been proposed that proteins have evolved intrinsic strategies to prevent aggregation using the arrangement of structural and sequence elements (Richardson and Richardson, 2002). For instance, some proteins have conserved residues that are important for preventing misfolding or aggregation. It has been postulated that charged residues and other β breaker residues such as Pro could function as 'gatekeeper' residues to prevent amyloid formation because their low tendency to be incorporated into cross-beta amyloid conformation (Monsellier and Chiti, 2007; Richardson and Richardson, 2002). Gatekeeper residues are commonly found at the edge β strand (the peripheral β strand) in the native β sheet conformations to prevent protein aggregation (Richardson and Richardson, 2002). The existence of gatekeeper residues could change folding kinetic partitioning to prevent nonproductive pathways including aggregation or other misfolding pathway (Figure 1.3A). It is also possible that the existence of gatekeeper residues alter the free energy state of the folded products. Mutation of gatekeeper residues resulted in more efficient amyloidogenesis and therefore early onset of amyloid pathology (Figure 1.3B) (Chiti et al., 2003; Orpiszewski and Benson, 1999; Otzen et al., 2000). For example, the mutation of Asp at position 23 in Aβ is linked to some familial Alzheimer's disease and this mutation D23N enhances fibrillogenesis *in vitro* (Grabowski et al., 2001; Van Nostrand et al., 2001).

Collectively, amyloid structure appears to be a main-chain dominated structure, but side chains influence both positively and negatively on amyloidogenesis *in vivo* and *in vitro*. What side chains exactly contribute to amyloid polymerization, especially *in vivo* amyloidogenesis, remains poorly understood. In this thesis I aim to elucidate the roles of

important side-chains in curli assembly.

Curli Assembly Mechanism

Our lab studies the biogenesis of bacterially produced amyloid fibers, curli. Curli are extracellular amyloids that contribute in many ways to cell survival and maintenance. Curli are important components of extracellular matrix of biofilms produced by enteric bacteria such as *E. coli* and *Salmonella* spp (Barnhart and Chapman, 2006). Due to the amyloid properties of curli fibers, the colonies of curli-producing *E. coli* stain red when grown on Congo red indicator plates, which provides a convenient assay to monitor curli assembly *in vivo*. Curli assembly in *E. coli* requires genes encoded by the *csgBA* and *csgDEFG* operons (Figure 1.4). CsgD is a transcriptional regulator in the FixJ/UhpA family and activates the transcription of *csgBA* operon under specific conditions (Barnhart et al., 2006). The expression of *csgBA* absolutely requires CsgD and there is no other known positive regulator for the *csgBA* operon (Barnhart et al., 2006). The regulation of curli gene expression is complex and is influenced by many environmental cues. Most laboratory strains of *E. coli* and most *Salmonella* strains, curli expression is best at temperatures below 30 °C (Arnqvist et al., 1992). Interestingly, many clinic strains of *E. coli* including sepsis isolates, can express curli at 37 °C and some mutations in the *csgD* promoter result in strains that produce curli regardless of temperature (Romling et al., 1998; Uhlich et al., 2001). In addition, the low salt condition and nutrient limitation (nitrogen, phosphate and iron) stimulate curli gene expression (Gerstel and Romling, 2001; Romling et al., 1998). Oxygen tension also plays a role in curli expression, with microaerophilic conditions resulting in maximal *csgD* transcription (Gerstel and Romling, 2001). There are many transcriptional regulators modulating the *csgDEFG* transcription.

OmpR/EnvZ, RpoS, Crl, MlrA and IHF positively activate the transcription of *csgDEFG*. CpxA/R and Rcs negatively influence the transcription of *csgDEFG* (Barnhart et al., 2006).

CsgE, CsgF, CsgG, CsgA and CsgB expressed from *csgBA* and *csgDEFG* operons have N-terminal Sec-signal sequence which directs these proteins across the inner membrane into periplasmic space through SecYEG machinery (Figure 1.4). CsgG is an outer membrane lipoprotein and is proposed to be the major component of curli-specific secretion machinery. In the absence of CsgG, curli fiber subunits CsgA and CsgB are not stable (Loferer et al., 1997). Overexpression of both CsgG and CsgA resulted in secretion of CsgA into the extracellular milieu (Chapman et al., 2002). CsgG interacts with CsgE and CsgF at the outer membrane (Robinson et al., 2006). Purified CsgG forms barrel-structured oligomers and overexpression of CsgG increases outer membrane permeability (Robinson et al., 2006). All these data suggest CsgG may function as a curli-specific secretion machine and coordinate the efficient curli assembly.

CsgE and CsgF are indispensable for curli assembly (Chapman et al., 2002). Their exact functions are beginning to be elucidated. In the absence of CsgE, the protein levels of CsgA and CsgB significantly are decreased and curli formation is dramatically disrupted (Chapman et al., 2002). A *csgF* mutant is defective in the nucleation of CsgB and many fibers produced by *csgF* cells were found not associated with cells (Chapman et al., 2002; Hammer et al., 2007).

The amino acid sequences of CsgA and CsgB share approximately 30% identity and 50% similarity. Their primary structure can be divided into three parts, a Sec signal sequence, an N-terminal 22 or 23 residues and five imperfect repeating units (R1, R2, R3,

R4 and R5) (Figure 1.5A and 1.5B). The Sec signal sequences are cleaved when they pass through inner membrane (Collinson et al., 1999; White et al., 2001). Therefore, the mature CsgA and CsgB only contain N-terminal stretches and five repeating units. The N-terminal stretches are sensitive to proteinase K treatment, whereas the five repeating units are resistant to proteinase K treatment indicating that they comprise the amyloid cores (Collinson et al., 1999; White et al., 2001). The N-terminal stretch of CsgA was reported to target CsgA to CsgG secretion machinery while the role of the N-terminal stretch of CsgB remains elusive (Robinson et al., 2006).

The repeating units of CsgA contain a consensus sequence (S-X_5-Q-X-G-X_2-N-X-A-X_3-Q) and each repeating unit is predicted to form a β strand-loop-β strand motif, thereby five repeats comprising a cross-β structure with two β sheets (Figure 1.5C) (Collinson et al., 1999). Similar to CsgA, the first four repeating units of CsgB have a consensus sequence (X_6-Q-X-G-X_2-N-X-A-X_3-Q) and the five repeating units are predicted to form a cross-β structure (White et al., 2001). The R5 of CsgB is different from other repeats in that it does not have conserved Gln and Gly residues (Figure 1.5B). Although CsgA and CsgB share similar amino acid sequences and predicted structural features, the theoretic isoelectric points (pI) of mature CsgA and CsgB are 4.51 and 9.30, respectively, because CsgB has significantly more cationic residues such as Arg and Lys than CsgA, especially in R5 (Figure 1.5B). Unlike CsgA, CsgB is associated to outer membrane and tethered on the cell surface (Hammar et al., 1996; Loferer et al., 1997). The deletion of the last repeat of CsgB resulted in a molecule disassociated from cells (Hammer et al., 2007). The majority component of outer membrane of Gram-negative bacteria is lipopolysaccharides (LPS) and it was reported that some LPS-binding

sequences are rich in cationic residues (Ferguson et al., 2000). It is possible that this distinct enrichment of cationic residues especially at the C-terminus of CsgB contributes to its affinity to LPS at the outer membrane.

The polymerization of CsgA *in vivo* requires the nucleator CsgB. CsgA is secreted to the milieu and remains soluble in the absence of CsgB (Hammar et al., 1996). CsgA secreted from *csgB* cells can interact with CsgB presented on the *csgA* cell surface and assemble into fibers (Hammar et al., 1996). This successful interbacterial complementation builds the basis of our current curli assembly model (Figure 1.4). CsgB is proposed to prime the polymerization of CsgA, termed nucleation-precipitation process (Bian and Normark, 1997; Hammar et al., 1996). The molecular details of this mechanism are beginning to be elucidated. We proposed that there are two key events that facilitate CsgA polymerization on the cell surface. Firstly, CsgB functions as a template to initiate and accelerate secreted CsgA polymerization (Figure 1.4). Since most amyloid formation is a self-propagation process, we hypothesize that preformed CsgA fibers also can function as a template to direct CsgA polymerization (Figure 1.4). Elucidation of these two critical events is one goal of this thesis. There are several fundamental questions I am trying to address using curli as a model system to study highly evolved amyloidogenic machine.

1) The polymerization of most disease-associated amyloids is a triphasic process and follows self-propagating mechanism. Is the polymerization of bacterial amyloid protein such as CsgA similar to disease-associated amyloids? If amyloid formation is an inherent property of polypeptide chains, is the folding or polymerization process conserved?

2) What are the sequence determinants for CsgA nucleation and polymerization? Is the

five-fold internal sequence homology of CsgA critical to its ability to polymerize into an amyloid fiber?

3) Amyloid structures are dictated by backbone interactions. The high conservation of polar residues such as Ser, Gln and Asn of CsgA suggests their potentially importance for CsgA polymerization. How do these side-chains influence bacterial amyloid formation?

4) How does *E .coli* control the toxicity usually associated with amyloid propagation? How does *E.coli* harness the polymerization of CsgA using the evolved machinery? What is the intrinsic property of CsgA to render the control by cells?

Figures

Figure 1.1. Amyloids are distinct proteinaceous fibers.

(A) Amyloids formed in brains associated with serum A amyloidosis in the presence of Congo red under polarized light (Zhang et al., 2008). (B) Negative-stain EM micrographs of Aβ40 amyloid fibers. (C) The structural model of Aβ40 resolved by solid-state NMR (Petkova et al., 2002). (D) The structure of peptide GNNQQNY derived from Sup35p and resolved by X-ray microcrystallography (Nelson et al., 2005). (E) The structural model of amyloid fibril of the HET-s prion domain in Het-s resolved by solid-state NMR (Wasmer et al., 2008).

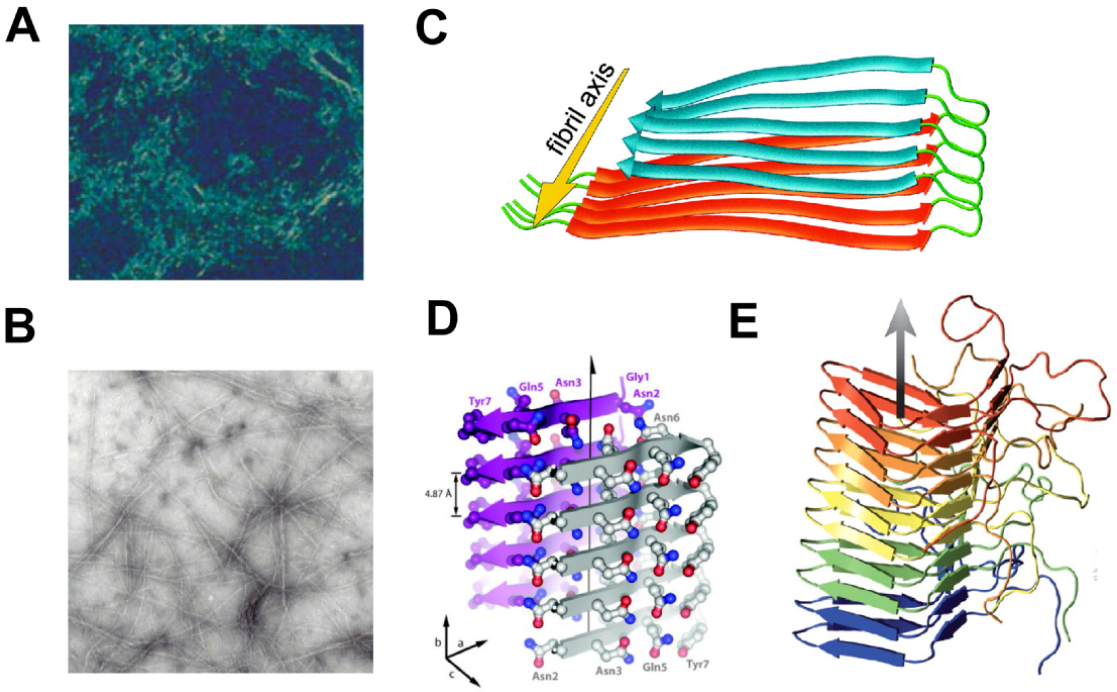


Figure 1.2. *In vitro* polymerization mechanism of amyloid proteins.
(A) A graphic representation of amyloid fiber formation showing a characteristic lag phase in the absence of seeds (black). By contrast, preformed amyloid fibers can act as seeds to eliminate the lag phase (red). (B) Schematic of amyloid fiber polymerization. Soluble precursor proteins assemble into a common intermediate prior to polymerizing into a stable amyloid fiber. The formation of the intermediate is proposed to be the rate-limiting step of amyloidogenesis.

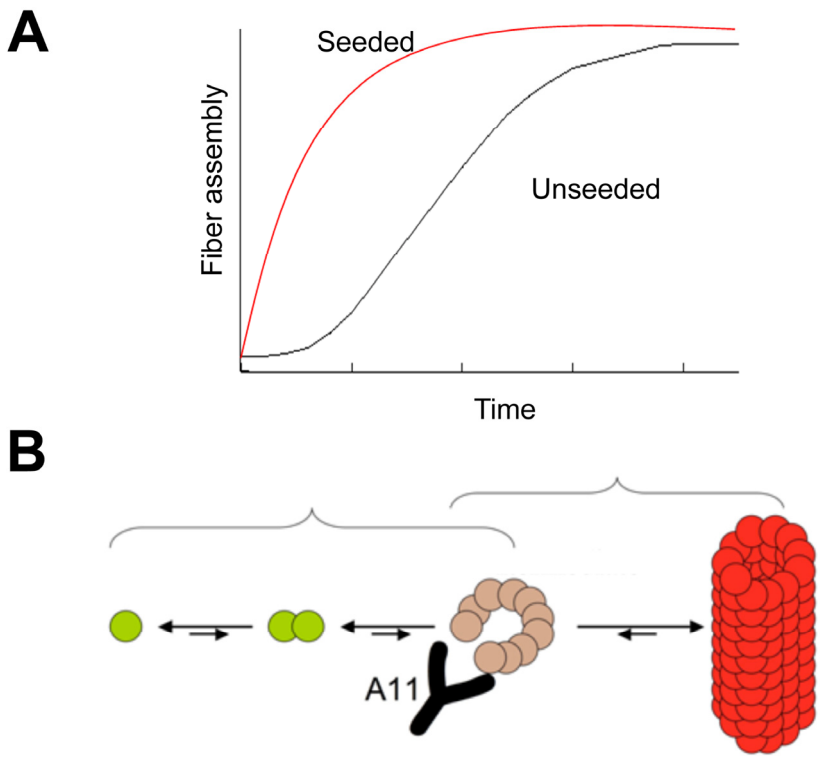


Figure 1.3. The gatekeeper residues influence the aggregation of amyloid proteins. (A) Schematic of the influence of gatekeeper residues on amyloid aggregation. The gatekeeper residues prevent the association of aggregative domains and thereby decrease the nonproductive folding intermediates. (B) In the absence of gatekeeper residues, amyloid aggregation is favored and aggregative domains form ordered fiber aggregates.

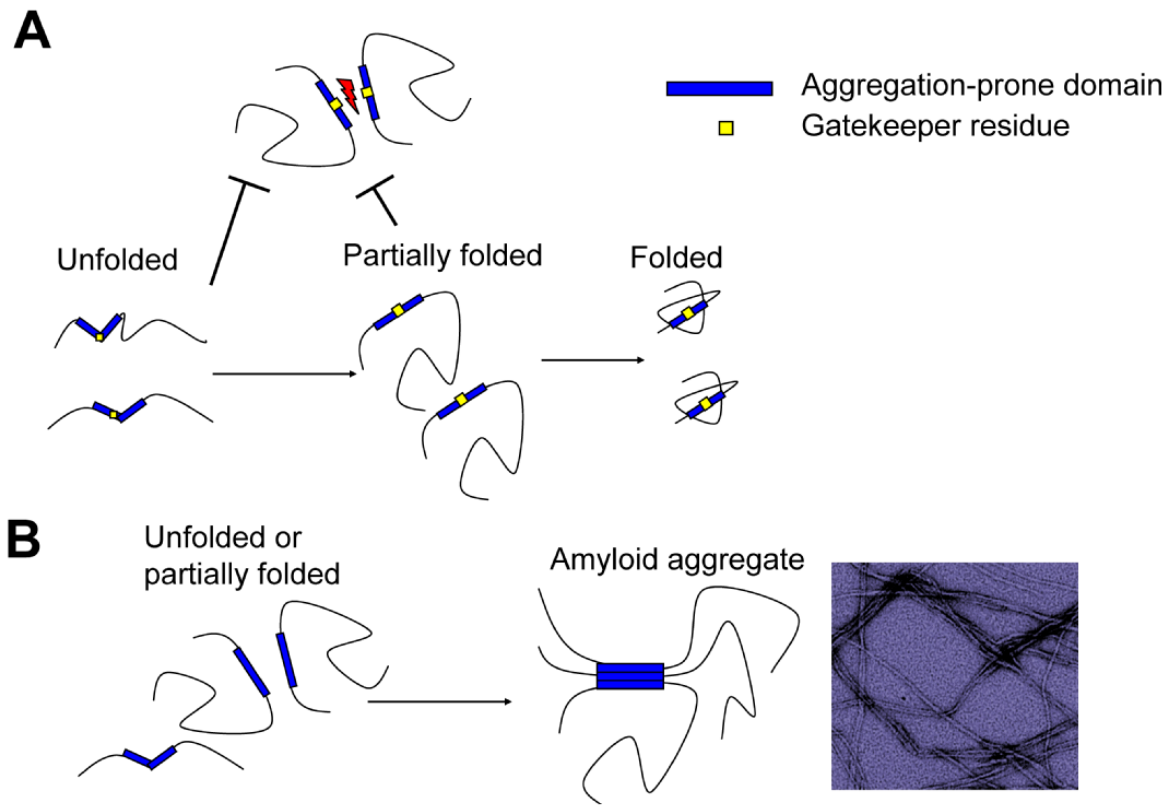


Figure 1.4. Model of curli regulation and assembly.
Except CsgD, All other Csg proteins have Sec-dependent signal sequences. CsgE and CsgF are required for efficient curli assembly. The major subunit protein (CsgA) and the nucleator (CsgB) are secreted to the cell surface in a CsgG dependent fashion. CsgB associates with the outer membrane where it nucleates soluble, unstructured CsgA into a highly ordered amyloid fiber as indicated by (1). Preformed CsgA fibers also promote soluble CsgA assembly as indicated by (2).

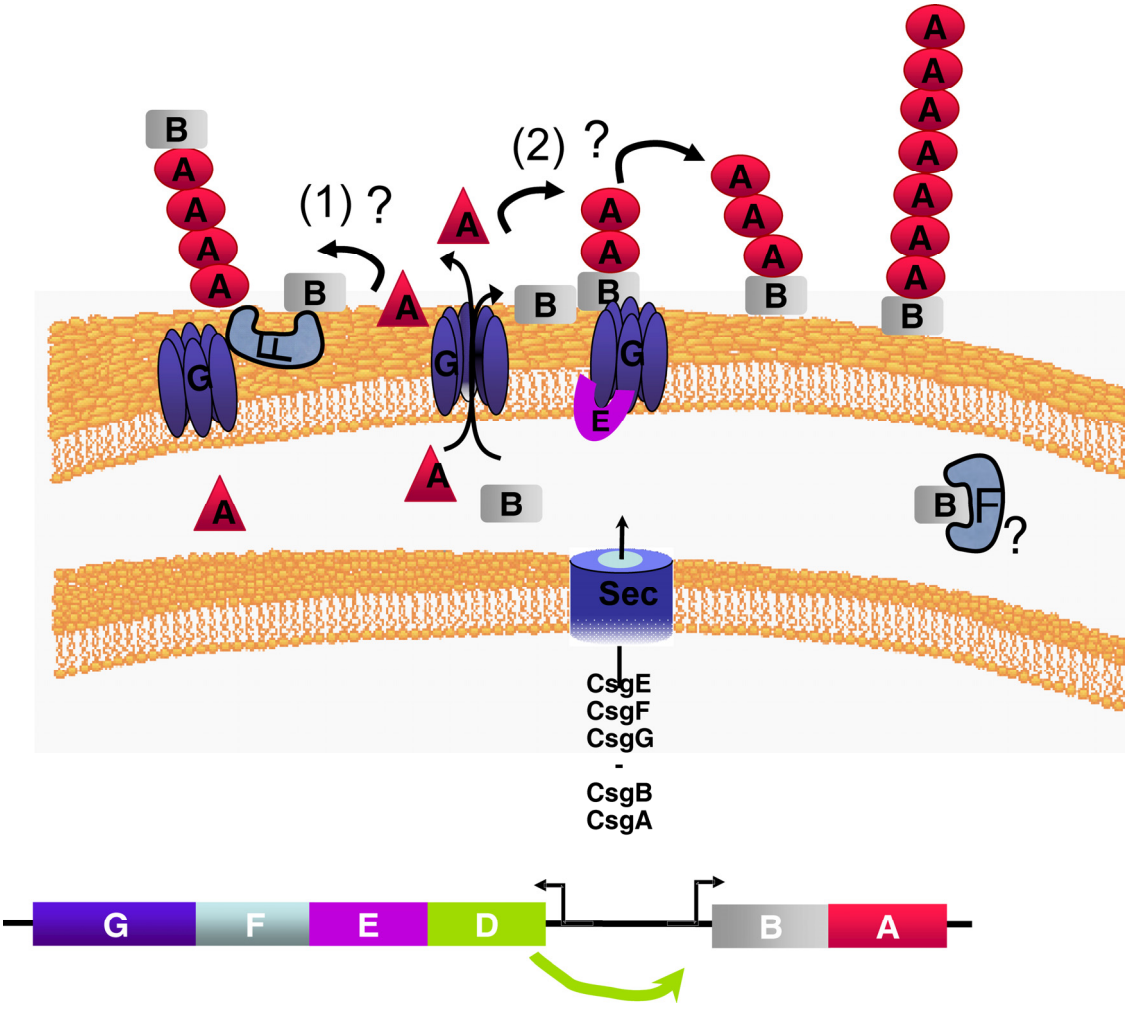
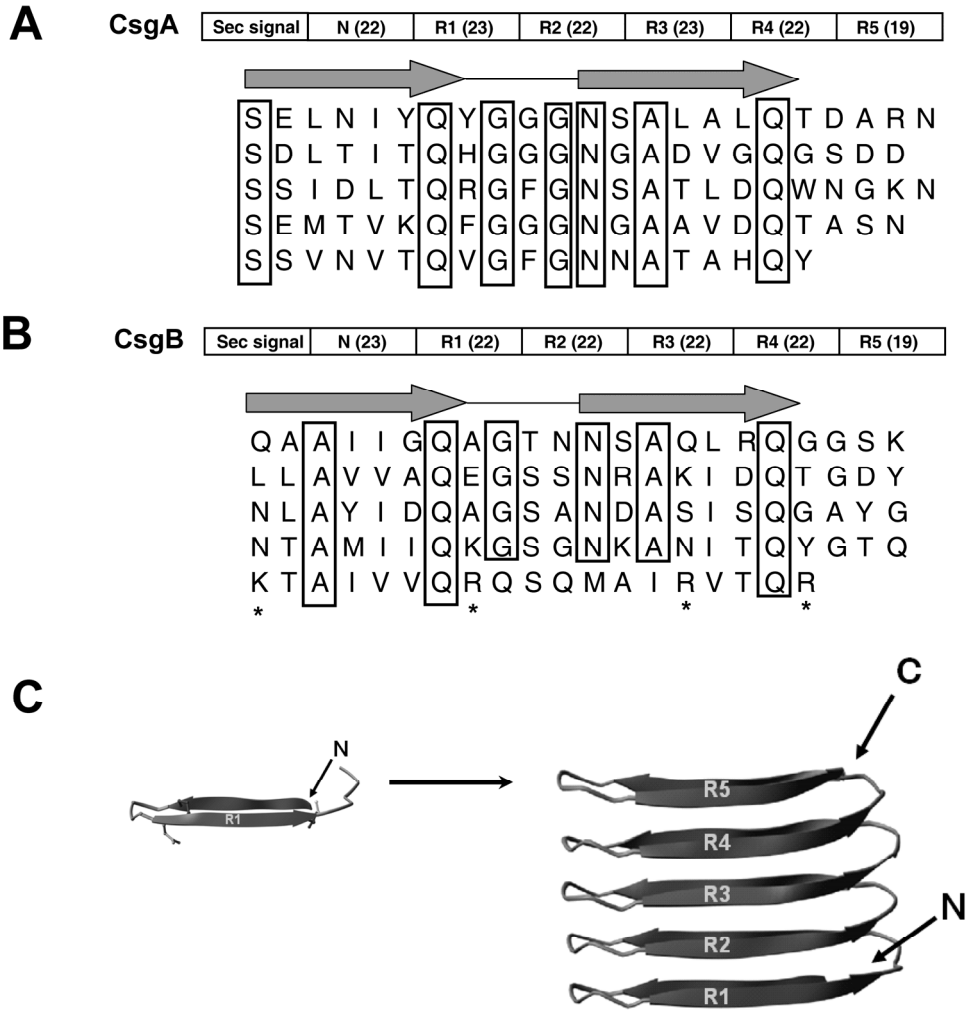


Figure 1.5. The amino acid sequence and predicted structure of CsgA and CsgB. (A) Schematic of CsgA including an N-terminal Sec signal sequence and the N-terminal 22 residues that precede the five repeating units. The imperfect repeating units in CsgA (R1–R5) are predicted to form β strand-loop- β strand structures. Amino acids comprising the β strand are located below the arrows. Boxed letters represent amino acids conserved throughout the repeating units. (B) Schematic of CsgB amino acid sequence. Boxed letters represent amino acids conserved throughout the repeating units. Positively charged amino acids in the fifth repeating unit of CsgB are marked with asterisks. (C) The predicted structural model of CsgA shows that the repeating units form β strand-loop- β strand motifs comprising cross β structure. The structure was constructed by Dr. Craig Smith based on previous work from the Kay lab (Collinson et al., 1999).



Reference

Alteri, C. J., Xicohtencatl-Cortes, J., Hess, S., Caballero-Olin, G., Giron, J. A., and Friedman, R. L. (2007). Mycobacterium tuberculosis produces pili during human infection. Proc Natl Acad Sci U S A *104*, 5145-5150.

Alzheimer, C. (1907). Über eine eigenartige Erkrankung der Hirnrinde. Allg Zeitschr Psychiatr Psychiatr-Gerichtl Med *64*, 146.

Anfinsen, C. B., and Scheraga, H. A. (1975). Experimental and theoretical aspects of protein folding. Adv Protein Chem *29*, 205-300.

Arnqvist, A., Olsen, A., Pfeifer, J., Russell, D. G., and Normark, S. (1992). The Crl protein activates cryptic genes for curli formation and fibronectin binding in *Escherichia coli* HB101. Mol Microbiol *6*, 2443-2452.

Arrasate, M., Mitra, S., Schweitzer, E. S., Segal, M. R., and Finkbeiner, S. (2004). Inclusion body formation reduces levels of mutant huntingtin and the risk of neuronal death. Nature *431*, 805-810.

Azriel, R., and Gazit, E. (2001). Analysis of the minimal amyloid-forming fragment of the islet amyloid polypeptide. An experimental support for the key role of the phenylalanine residue in amyloid formation. J Biol Chem *276*, 34156-34161.

Barnhart, M. M., and Chapman, M. R. (2006). Curli biogenesis and function. Annu Rev Microbiol *60*, 131-147.

Barnhart, M. M., Lynem, J., and Chapman, M. R. (2006). GlcNAc-6P levels modulate the expression of Curli fibers by *Escherichia coli*. J Bacteriol *188*, 5212-5219.

Bian, Z., and Normark, S. (1997). Nucleator function of CsgB for the assembly of adhesive surface organelles in *Escherichia coli*. Embo J *16*, 5827-5836.

Bieler, S., Estrada, L., Lagos, R., Baeza, M., Castilla, J., and Soto, C. (2005). Amyloid formation modulates the biological activity of a bacterial protein. J Biol Chem *280*,

26880-26885.

Blankenship, J. R., and Mitchell, A. P. (2006). How to build a biofilm: a fungal perspective. Curr Opin Microbiol *9*, 588-594.

Blinder, D., Coschigano, P. W., and Magasanik, B. (1996). Interaction of the GATA factor Gln3p with the nitrogen regulator Ure2p in Saccharomyces cerevisiae. J Bacteriol *178*, 4734-4736.

Bokranz, W., Wang, X., Tschape, H., and Romling, U. (2005). Expression of cellulose and curli fimbriae by *Escherichia coli* isolated from the gastrointestinal tract. J Med Microbiol *54*, 1171-1182.

Bonar, L., Cohen, A. S., and Skinner, M. M. (1969). Characterization of the amyloid fibril as a cross-beta protein. Proc Soc Exp Biol Med *131*, 1373-1375.

Bucciantini, M., Giannoni, E., Chiti, F., Baroni, F., Formigli, L., Zurdo, J., Taddei, N., Ramponi, G., Dobson, C. M., and Stefani, M. (2002). Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. Nature *416*, 507-511.

Busciglio, J., Lorenzo, A., and Yankner, B. A. (1992). Methodological variables in the assessment of beta amyloid neurotoxicity. Neurobiol Aging *13*, 609-612.

Carulla, N., Caddy, G. L., Hall, D. R., Zurdo, J., Gairi, M., Feliz, M., Giralt, E., Robinson, C. V., and Dobson, C. M. (2005). Molecular recycling within amyloid fibrils. Nature *436*, 554-558.

Chamberlain, A. K., Receveur, V., Spencer, A., Redfield, C., and Dobson, C. M. (2001). Characterization of the structure and dynamics of amyloidogenic variants of human lysozyme by NMR spectroscopy. Protein Sci *10*, 2525-2530.

Chandra, J., Kuhn, D. M., Mukherjee, P. K., Hoyer, L. L., McCormick, T., and Ghannoum, M. A. (2001). Biofilm formation by the fungal pathogen Candida albicans: development, architecture, and drug resistance. J Bacteriol *183*, 5385-5394.

Chapman, M. R., Robinson, L. S., Pinkner, J. S., Roth, R., Heuser, J., Hammar, M., Normark, S., and Hultgren, S. J. (2002). Role of *Escherichia coli* curli operons in directing amyloid fiber formation. Science *295*, 851-855.

Chen, S., Ferrone, F. A., and Wetzel, R. (2002). Huntington's disease age-of-onset linked to polyglutamine aggregation nucleation. Proc Natl Acad Sci U S A *99*, 11884-11889.

Cheng, S. V., Nadeau, J. H., Tanzi, R. E., Watkins, P. C., Jagadesh, J., Taylor, B. A., Haines, J. L., Sacchi, N., and Gusella, J. F. (1988). Comparative mapping of DNA markers from the familial Alzheimer disease and Down syndrome regions of human chromosome 21 to mouse chromosomes 16 and 17. Proc Natl Acad Sci U S A *85*, 6032-6036.

Chiti, F., and Dobson, C. M. (2006). Protein misfolding, functional amyloid, and human disease. Annu Rev Biochem *75*, 333-366.

Chiti, F., Stefani, M., Taddei, N., Ramponi, G., and Dobson, C. M. (2003). Rationalization of the effects of mutations on peptide and protein aggregation rates. Nature *424*, 805-808.

Citron, M., Westaway, D., Xia, W., Carlson, G., Diehl, T., Levesque, G., Johnson-Wood, K., Lee, M., Seubert, P., Davis, A.*, et al.* (1997). Mutant presenilins of Alzheimer's disease increase production of 42-residue amyloid beta-protein in both transfected cells and transgenic mice. Nat Med *3*, 67-72.

Claessen, D., Rink, R., de Jong, W., Siebring, J., de Vreugd, P., Boersma, F. G., Dijkhuizen, L., and Wosten, H. A. (2003). A novel class of secreted hydrophobic proteins is involved in aerial hyphae formation in Streptomyces coelicolor by forming amyloid-like fibrils. Genes Dev *17*, 1714-1726.

Claessen, D., Stokroos, I., Deelstra, H. J., Penninga, N. A., Bormann, C., Salas, J. A., Dijkhuizen, L., and Wosten, H. A. (2004). The formation of the rodlet layer of streptomycetes is the result of the interplay between rodlins and chaplins. Mol Microbiol *53*, 433-443.

Cleary, J. P., Walsh, D. M., Hofmeister, J. J., Shankar, G. M., Kuskowski, M. A., Selkoe,

D. J., and Ashe, K. H. (2005). Natural oligomers of the amyloid-beta protein specifically disrupt cognitive function. Nat Neurosci *8*, 79-84.

Cohen, E., Bieschke, J., Perciavalle, R. M., Kelly, J. W., and Dillin, A. (2006). Opposing activities protect against age-onset proteotoxicity. Science *313*, 1604-1610.

Collinson, S. K., Emody, L., Muller, K. H., Trust, T. J., and Kay, W. W. (1991). Purification and characterization of thin, aggregative fimbriae from Salmonella enteritidis. J Bacteriol *173*, 4773-4781.

Collinson, S. K., Parker, J. M., Hodges, R. S., and Kay, W. W. (1999). Structural predictions of AgfA, the insoluble fimbrial subunit of Salmonella thin aggregative fimbriae. J Mol Biol *290*, 741-756.

Coustou-Linares, V., Maddelein, M. L., Begueret, J., and Saupe, S. J. (2001). In vivo aggregation of the HET-s prion protein of the fungus Podospora anserina. Mol Microbiol *42*, 1325-1335.

Coustou, V., Deleu, C., Saupe, S., and Begueret, J. (1997). The protein product of the het-s heterokaryon incompatibility gene of the fungus Podospora anserina behaves as a prion analog. Proc Natl Acad Sci U S A *94*, 9773-9778.

Crespo, J. L., Daicho, K., Ushimaru, T., and Hall, M. N. (2001). The GATA transcription factors GLN3 and GAT1 link TOR to salt stress in Saccharomyces cerevisiae. J Biol Chem *276*, 34441-34444.

Dales, R. B., Moorhouse, J., and Croft, J. H. (1983). The location and analysis of two heterokaryon incompatibility (het) loci in strains of Aspergillus nidulans. J Gen Microbiol *129*, 3637-3642.

Deshpande, A., Mina, E., Glabe, C., and Busciglio, J. (2006). Different conformations of amyloid beta induce neurotoxicity by distinct mechanisms in human cortical neurons. J Neurosci *26*, 6011-6018.

Destoumieux-Garzon, D., Thomas, X., Santamaria, M., Goulard, C., Barthelemy, M.,

Boscher, B., Bessin, Y., Molle, G., Pons, A. M., Letellier, L.*, et al.* (2003). Microcin E492 antibacterial activity: evidence for a TonB-dependent inner membrane permeabilization on *Escherichia coli*. Mol Microbiol *49*, 1031-1041.

Di Berardo, C., Capstick, D. S., Bibb, M. J., Findlay, K. C., Buttner, M. J., and Elliot, M. A. (2008). Function and redundancy of the chaplin cell surface proteins in aerial hypha formation, rodlet assembly, and viability in Streptomyces coelicolor. J Bacteriol *190*, 5879-5889.

Divry P, F. M. (1927). Sur les propriétés optiques de l'amyloïde. CR Société de Biologie (Paris) *97*, 1808-1810.

Elliot, M. A., Karoonuthaisiri, N., Huang, J., Bibb, M. J., Cohen, S. N., Kao, C. M., and Buttner, M. J. (2003). The chaplins: a family of hydrophobic cell-surface proteins involved in aerial mycelium formation in Streptomyces coelicolor. Genes Dev *17*, 1727-1740.

Epstein, E. A., and Chapman, M. R. (2008). Polymerizing the fibre between bacteria and host cells: the biogenesis of functional amyloid fibres. Cell Microbiol *10*, 1413-1420.

Faber, P. W., Alter, J. R., MacDonald, M. E., and Hart, A. C. (1999). Polyglutamine-mediated dysfunction and apoptotic death of a Caenorhabditis elegans sensory neuron. Proc Natl Acad Sci U S A *96*, 179-184.

Fandrich, M., and Dobson, C. M. (2002). The behaviour of polyamino acids reveals an inverse side chain effect in amyloid structure formation. Embo J *21*, 5682-5690.

Fandrich, M., Fletcher, M. A., and Dobson, C. M. (2001). Amyloid fibrils from muscle myoglobin. Nature *410*, 165-166.

Ferguson, A. D., Welte, W., Hofmann, E., Lindner, B., Holst, O., Coulton, J. W., and Diederichs, K. (2000). A conserved structural motif for lipopolysaccharide recognition by procaryotic and eucaryotic proteins. Structure *8*, 585-592.

Fernandez-Escamilla, A. M., Rousseau, F., Schymkowitz, J., and Serrano, L. (2004).

Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. Nat Biotechnol *22*, 1302-1306.

Fowler, D. M., Koulov, A. V., Alory-Jost, C., Marks, M. S., Balch, W. E., and Kelly, J. W. (2006). Functional amyloid formation within mammalian tissue. PLoS Biol *4*, e6.

Fowler, D. M., Koulov, A. V., Balch, W. E., and Kelly, J. W. (2007). Functional amyloid--from bacteria to humans. Trends Biochem Sci *32*, 217-224.

Games, D., Adams, D., Alessandrini, R., Barbour, R., Berthelette, P., Blackwell, C., Carr, T., Clemens, J., Donaldson, T., Gillespie, F., and et al. (1995). Alzheimer-type neuropathology in transgenic mice overexpressing V717F beta-amyloid precursor protein. Nature *373*, 523-527.

Gerstel, U., and Romling, U. (2001). Oxygen tension and nutrient starvation are major signals that regulate agfD promoter activity and expression of the multicellular morphotype in Salmonella typhimurium. Environ Microbiol *3*, 638-648.

Glenner, G. G., Cuatrecasas, P., Isersky, C., Bladen, H. A., and Eanes, E. D. (1969). Physical and chemical properties of amyloid fibers. II. Isolation of a unique protein constituting the major component from human splenic amyloid fibril concentrates. J Histochem Cytochem *17*, 769-780.

Glenner, G. G., Ein, D., Eanes, E. D., Bladen, H. A., Terry, W., and Page, D. L. (1971). Creation of "amyloid" fibrils from Bence Jones proteins in vitro. Science *174*, 712-714.

Glenner, G. G., Page, D., Isersky, C., Harada, M., Cuatrecasas, P., Eanes, E. D., DeLellis, R. A., Bladen, H. A., and Keiser, H. R. (1971). Murine amyloid fibril protein: isolation, purification and characterization. J Histochem Cytochem *19*, 16-28.

Glenner, G. G., and Wong, C. W. (1984). Alzheimer's disease: initial report of the purification and characterization of a novel cerebrovascular amyloid protein. Biochem Biophys Res Commun *120*, 885-890.

Glover, J. R., Kowal, A. S., Schirmer, E. C., Patino, M. M., Liu, J. J., and Lindquist, S.

(1997). Self-seeded fibers formed by Sup35, the protein determinant of [PSI+], a heritable prion-like factor of S. cerevisiae. Cell *89*, 811-819.

Goate, A., Chartier-Harlin, M. C., Mullan, M., Brown, J., Crawford, F., Fidani, L., Giuffra, L., Haynes, A., Irving, N., James, L., and et al. (1991). Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. Nature *349*, 704-706.

Gophna, U., Barlev, M., Seijffers, R., Oelschlager, T. A., Hacker, J., and Ron, E. Z. (2001). Curli fibers mediate internalization of *Escherichia coli* by eukaryotic cells. Infect Immun *69*, 2659-2665.

Gophna, U., Oelschlaeger, T. A., Hacker, J., and Ron, E. Z. (2002). Role of fibronectin in curli-mediated internalization. FEMS Microbiol Lett *212*, 55-58.

Grabowski, T. J., Cho, H. S., Vonsattel, J. P., Rebeck, G. W., and Greenberg, S. M. (2001). Novel amyloid precursor protein mutation in an Iowa family with dementia and severe cerebral amyloid angiopathy. Ann Neurol *49*, 697-705.

Haass, C. (2004). Take five--BACE and the gamma-secretase quartet conduct Alzheimer's amyloid beta-peptide generation. Embo J *23*, 483-488.

Haass, C., and Selkoe, D. J. (2007). Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid beta-peptide. Nat Rev Mol Cell Biol *8*, 101-112.

Haber, E., and Anfinsen, C. B. (1961). Regeneration of enzyme activity by air oxidation of reduced subtilisin-modified ribonuclease. J Biol Chem *236*, 422-424.

Hake, L. E., and Richter, J. D. (1994). CPEB is a specificity factor that mediates cytoplasmic polyadenylation during Xenopus oocyte maturation. Cell *79*, 617-627.

Hammar, M., Bian, Z., and Normark, S. (1996). Nucleator-dependent intercellular assembly of adhesive curli organelles in *Escherichia coli*. Proc Natl Acad Sci U S A *93*, 6562-6566.

Hammer, N. D., Schmidt, J. C., and Chapman, M. R. (2007). The curli nucleator protein, CsgB, contains an amyloidogenic domain that directs CsgA polymerization. Proc Natl Acad Sci U S A *104*, 12494-12499.

Hammer, N. D., Wang, X., McGuffie, B. A., and Chapman, M. R. (2008). Amyloids: friend or foe? J Alzheimers Dis *13*, 407-419.

Hardy, J., and Selkoe, D. J. (2002). The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. Science *297*, 353-356.

Harper, J. D., and Lansbury, P. T., Jr. (1997). Models of amyloid seeding in Alzheimer's disease and scrapie: mechanistic truths and physiological consequences of the time-dependent solubility of amyloid proteins. Annu Rev Biochem *66*, 385-407.

Hartley, D. M., Walsh, D. M., Ye, C. P., Diehl, T., Vasquez, S., Vassilev, P. M., Teplow, D. B., and Selkoe, D. J. (1999). Protofibrillar intermediates of amyloid beta-protein induce acute electrophysiological changes and progressive neurotoxicity in cortical neurons. J Neurosci *19*, 8876-8884.

Hartmann, T., Bieger, S. C., Bruhl, B., Tienari, P. J., Ida, N., Allsop, D., Roberts, G. W., Masters, C. L., Dotti, C. G., Unsicker, K., and Beyreuther, K. (1997). Distinct sites of intracellular production for Alzheimer's disease A beta40/42 amyloid peptides. Nat Med *3*, 1016-1020.

Hoshi, M., Sato, M., Matsumoto, S., Noguchi, A., Yasutake, K., Yoshida, N., and Sato, K. (2003). Spherical aggregates of beta-amyloid (amylospheroid) show high neurotoxicity and activate tau protein kinase I/glycogen synthase kinase-3beta. Proc Natl Acad Sci U S A *100*, 6370-6375.

Hsiao, K., Chapman, P., Nilsen, S., Eckman, C., Harigaya, Y., Younkin, S., Yang, F., and Cole, G. (1996). Correlative memory deficits, Abeta elevation, and amyloid plaques in transgenic mice. Science *274*, 99-102.

Jarrett, J. T., and Lansbury, P. T., Jr. (1993). Seeding "one-dimensional crystallization" of amyloid: a pathogenic mechanism in Alzheimer's disease and scrapie? Cell *73*, 1055-

1058.

Kayed, R., Head, E., Thompson, J. L., McIntire, T. M., Milton, S. C., Cotman, C. W., and Glabe, C. G. (2003). Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis. Science *300*, 486-489.

Kidd, M. (1963). Paired helical filaments in electron microscopy of Alzheimer's disease. Nature *197*, 192-193.

Kim, I. C., Franzblau, C., Shirahama, T., and Cohen, A. S. (1969). The effect of papain, pronase, Nagarse and trypsin on isolated amyloid fibrils. Biochim Biophys Acta *181*, 465-467.

Kim, J. G., Park, B. K., Yoo, C. H., Jeon, E., Oh, J., and Hwang, I. (2003). Characterization of the Xanthomonas axonopodis pv. glycines Hrp pathogenicity island. J Bacteriol *185*, 3155-3166.

Kim, W., and Hecht, M. H. (2005). Sequence determinants of enhanced amyloidogenicity of Alzheimer A{beta}42 peptide relative to A{beta}40. J Biol Chem *280*, 35069-35076.

King, C. Y., Tittmann, P., Gross, H., Gebert, R., Aebi, M., and Wuthrich, K. (1997). Prion-inducing domain 2-114 of yeast Sup35 protein transforms in vitro into amyloid-like filaments. Proc Natl Acad Sci U S A *94*, 6618-6622.

Lagos, R., Wilkens, M., Vergara, C., Cecchi, X., and Monasterio, O. (1993). Microcin E492 forms ion channels in phospholipid bilayer membrane. FEBS Lett *321*, 145-148.

Lambert, M. P., Barlow, A. K., Chromy, B. A., Edwards, C., Freed, R., Liosatos, M., Morgan, T. E., Rozovsky, I., Trommer, B., Viola, K. L.*, et al.* (1998). Diffusible, nonfibrillar ligands derived from Abeta1-42 are potent central nervous system neurotoxins. Proc Natl Acad Sci U S A *95*, 6448-6453.

Larsen, P., Nielsen, J. L., Dueholm, M. S., Wetzel, R., Otzen, D., and Nielsen, P. H. (2007). Amyloid adhesins are abundant in natural biofilms. Environ Microbiol *9*, 3077-3090.

Lesne, S., Koh, M. T., Kotilinek, L., Kayed, R., Glabe, C. G., Yang, A., Gallagher, M., and Ashe, K. H. (2006). A specific amyloid-beta protein assembly in the brain impairs memory. Nature *440*, 352-357.

Liebman, S. W., and Cavenagh, M. (1980). An antisuppressor that acts on omnipotent suppressors in yeast. Genetics *95*, 49-61.

Linding, R., Schymkowitz, J., Rousseau, F., Diella, F., and Serrano, L. (2004). A comparative study of the relationship between protein structure and beta-aggregation in globular and intrinsically disordered proteins. J Mol Biol *342*, 345-353.

Liu, J. J., and Lindquist, S. (1999). Oligopeptide-repeat expansions modulate 'protein-only' inheritance in yeast. Nature *400*, 573-576.

Loferer, H., Hammar, M., and Normark, S. (1997). Availability of the fibre subunit CsgA and the nucleator protein CsgB during assembly of fibronectin-binding curli is limited by the intracellular concentration of the novel lipoprotein CsgG. Mol Microbiol *26*, 11-23.

Lomakin, A., Chung, D. S., Benedek, G. B., Kirschner, D. A., and Teplow, D. B. (1996). On the nucleation and growth of amyloid beta-protein fibrils: detection of nuclei and quantitation of rate constants. Proc Natl Acad Sci U S A *93*, 1125-1129.

Lopez de la Paz, M., and Serrano, L. (2004). Sequence determinants of amyloid fibril formation. Proc Natl Acad Sci U S A *101*, 87-92.

Lue, L. F., Kuo, Y. M., Roher, A. E., Brachova, L., Shen, Y., Sue, L., Beach, T., Kurth, J. H., Rydel, R. E., and Rogers, J. (1999). Soluble amyloid beta peptide concentration as a predictor of synaptic change in Alzheimer's disease. Am J Pathol *155*, 853-862.

Lundmark, K., Westermark, G. T., Nystrom, S., Murphy, C. L., Solomon, A., and Westermark, P. (2002). Transmissibility of systemic amyloidosis by a prion-like mechanism. Proc Natl Acad Sci U S A *99*, 6979-6984.

Lundmark, K., Westermark, G. T., Olsen, A., and Westermark, P. (2005). Protein fibrils in nature can enhance amyloid protein A amyloidosis in mice: Cross-seeding as a disease

mechanism. Proc Natl Acad Sci U S A *102*, 6098-6102.

Maddelein, M. L., Dos Reis, S., Duvezin-Caubet, S., Coulary-Salin, B., and Saupe, S. J. (2002). Amyloid aggregates of the HET-s prion protein are infectious. Proc Natl Acad Sci U S A *99*, 7402-7407.

Masison, D. C., and Wickner, R. B. (1995). Prion-inducing domain of yeast Ure2p and protease resistance of Ure2p in prion-containing cells. Science *270*, 93-95.

Masters, C. L., Simms, G., Weinman, N. A., Multhaup, G., McDonald, B. L., and Beyreuther, K. (1985). Amyloid plaque core protein in Alzheimer disease and Down syndrome. Proc Natl Acad Sci U S A *82*, 4245-4249.

Michelitsch, M. D., and Weissman, J. S. (2000). A census of glutamine/asparagine-rich regions: implications for their conserved function and the prediction of novel prions. Proc Natl Acad Sci U S A *97*, 11910-11915.

Monsellier, E., and Chiti, F. (2007). Prevention of amyloid-like aggregation as a driving force of protein evolution. EMBO Rep *8*, 737-742.

Nelson, R., Sawaya, M. R., Balbirnie, M., Madsen, A. O., Riekel, C., Grothe, R., and Eisenberg, D. (2005). Structure of the cross-beta spine of amyloid-like fibrils. Nature *435*, 773-778.

Nilsberth, C., Westlind-Danielsson, A., Eckman, C. B., Condron, M. M., Axelman, K., Forsell, C., Stenh, C., Luthman, J., Teplow, D. B., Younkin, S. G*., et al.* (2001). The 'Arctic' APP mutation (E693G) causes Alzheimer's disease by enhanced Abeta protofibril formation. Nat Neurosci *4*, 887-893.

Oh, J., Kim, J. G., Jeon, E., Yoo, C. H., Moon, J. S., Rhee, S., and Hwang, I. (2007). Amyloidogenesis of type III-dependent harpins from plant pathogenic bacteria. J Biol Chem *282*, 13601-13609.

Orpiszewski, J., and Benson, M. D. (1999). Induction of beta-sheet structure in amyloidogenic peptides by neutralization of aspartate: a model for amyloid nucleation. J

Mol Biol *289*, 413-428.

Otoo, H. N., Lee, K. G., Qiu, W., and Lipke, P. N. (2008). Candida albicans Als adhesins have conserved amyloid-forming sequences. Eukaryot Cell *7*, 776-782.

Otzen, D. E., Kristensen, O., and Oliveberg, M. (2000). Designed protein tetramer zipped together with a hydrophobic Alzheimer homology: a structural clue to amyloid assembly. Proc Natl Acad Sci U S A *97*, 9907-9912.

Patterson, D., Gardiner, K., Kao, F. T., Tanzi, R., Watkins, P., and Gusella, J. F. (1988). Mapping of the gene encoding the beta-amyloid precursor protein and its relationship to the Down syndrome region of chromosome 21. Proc Natl Acad Sci U S A *85*, 8266-8270.

Pennell, R. I., and Lamb, C. (1997). Programmed Cell Death in Plants. Plant Cell *9*, 1157-1168.

Petkova, A. T., Ishii, Y., Balbach, J. J., Antzutkin, O. N., Leapman, R. D., Delaglio, F., and Tycko, R. (2002). A structural model for Alzheimer's beta -amyloid fibrils based on experimental constraints from solid state NMR. Proc Natl Acad Sci U S A *99*, 16742-16747.

Petkova, A. T., Leapman, R. D., Guo, Z., Yau, W. M., Mattson, M. P., and Tycko, R. (2005). Self-propagating, molecular-level polymorphism in Alzheimer's beta-amyloid fibrils. Science *307*, 262-265.

Pike, C. J., Walencewicz, A. J., Glabe, C. G., and Cotman, C. W. (1991). In vitro aging of beta-amyloid protein causes peptide aggregation and neurotoxicity. Brain Res *563*, 311-314.

Plomp, M., Leighton, T. J., Wheeler, K. E., Hill, H. D., and Malkin, A. J. (2007). In vitro high-resolution structural dynamics of single germinating bacterial spores. Proc Natl Acad Sci U S A *104*, 9644-9649.

Podlisny, M. B., Ostaszewski, B. L., Squazzo, S. L., Koo, E. H., Rydell, R. E., Teplow, D. B., and Selkoe, D. J. (1995). Aggregation of secreted amyloid beta-protein into sodium

dodecyl sulfate-stable oligomers in cell culture. J Biol Chem *270*, 9564-9570.

Podlisny, M. B., Walsh, D. M., Amarante, P., Ostaszewski, B. L., Stimson, E. R., Maggio, J. E., Teplow, D. B., and Selkoe, D. J. (1998). Oligomerization of endogenous and synthetic amyloid beta-protein at nanomolar levels in cell culture and stabilization of monomer by Congo red. Biochemistry *37*, 3602-3611.

Prusiner, S. B., McKinley, M. P., Bowman, K. A., Bolton, D. C., Bendheim, P. E., Groth, D. F., and Glenner, G. G. (1983). Scrapie prions aggregate to form amyloid-like birefringent rods. Cell *35*, 349-358.

Prusiner, S. B., Scott, M. R., DeArmond, S. J., and Cohen, F. E. (1998). Prion protein biology. Cell *93*, 337-348.

Quist, A., Doudevski, I., Lin, H., Azimova, R., Ng, D., Frangione, B., Kagan, B., Ghiso, J., and Lal, R. (2005). Amyloid ion channels: a common structural link for protein-misfolding disease. Proc Natl Acad Sci U S A *102*, 10427-10432.

Raposo, G., Tenza, D., Murphy, D. M., Berson, J. F., and Marks, M. S. (2001). Distinct protein sorting and localization to premelanosomes, melanosomes, and lysosomes in pigmented melanocytic cells. J Cell Biol *152*, 809-824.

Rauceo, J. M., Gaur, N. K., Lee, K. G., Edwards, J. E., Klotz, S. A., and Lipke, P. N. (2004). Global cell surface conformational shift mediated by a Candida albicans adhesin. Infect Immun *72*, 4948-4955.

Reddy, P. H., Williams, M., Charles, V., Garrett, L., Pike-Buchanan, L., Whetsell, W. O., Jr., Miller, G., and Tagle, D. A. (1998). Behavioural abnormalities and selective neuronal loss in HD transgenic mice expressing mutated full-length HD cDNA. Nat Genet *20*, 198-202.

Richardson, J. S., and Richardson, D. C. (2002). Natural beta-sheet proteins use negative design to avoid edge-to-edge aggregation. Proc Natl Acad Sci U S A *99*, 2754-2759.

Ritter, C., Maddelein, M. L., Siemer, A. B., Luhrs, T., Ernst, M., Meier, B. H., Saupe, S.

J., and Riek, R. (2005). Correlation of structural elements and infectivity of the HET-s prion. Nature *435*, 844-848.

Robakis, N. K., Wisniewski, H. M., Jenkins, E. C., Devine-Gage, E. A., Houck, G. E., Yao, X. L., Ramakrishna, N., Wolfe, G., Silverman, W. P., and Brown, W. T. (1987). Chromosome 21q21 sublocalisation of gene encoding beta-amyloid peptide in cerebral vessels and neuritic (senile) plaques of people with Alzheimer disease and Down syndrome. Lancet *1*, 384-385.

Robinson, L. S., Ashman, E. M., Hultgren, S. J., and Chapman, M. R. (2006). Secretion of curli fibre subunits is mediated by the outer membrane-localized CsgG protein. Mol Microbiol *59*, 870-881.

Roher, A. E., Chaney, M. O., Kuo, Y. M., Webster, S. D., Stine, W. B., Haverkamp, L. J., Woods, A. S., Cotter, R. J., Tuohy, J. M., Krafft, G. A.*, et al.* (1996). Morphology and toxicity of Abeta-(1-42) dimer derived from neuritic and vascular amyloid deposits of Alzheimer's disease. J Biol Chem *271*, 20631-20635.

Romling, U., Sierralta, W. D., Eriksson, K., and Normark, S. (1998). Multicellular and aggregative behaviour of Salmonella typhimurium strains is controlled by mutations in the agfD promoter. Mol Microbiol *28*, 249-264.

Ross, E. D., Minton, A., and Wickner, R. B. (2005). Prion domains: sequences, structures and interactions. Nat Cell Biol *7*, 1039-1044.

Rovelet-Lecrux, A., Hannequin, D., Raux, G., Le Meur, N., Laquerriere, A., Vital, A., Dumanchin, C., Feuillette, S., Brice, A., Vercelletto, M.*, et al.* (2006). APP locus duplication causes autosomal dominant early-onset Alzheimer disease with cerebral amyloid angiopathy. Nat Genet *38*, 24-26.

Ryu, J. H., and Beuchat, L. R. (2005). Biofilm formation by *Escherichia coli* O157:H7 on stainless steel: effect of exopolysaccharide and Curli production on its resistance to chlorine. Appl Environ Microbiol *71*, 247-254.

Scherzinger, E., Lurz, R., Turmaine, M., Mangiarini, L., Hollenbach, B., Hasenbank, R.,

Bates, G. P., Davies, S. W., Lehrach, H., and Wanker, E. E. (1997). Huntingtin-encoded polyglutamine expansions form amyloid-like protein aggregates in vitro and in vivo. Cell *90*, 549-558.

Schiffer, N. W., Broadley, S. A., Hirschberger, T., Tavan, P., Kretzschmar, H. A., Giese, A., Haass, C., Hartl, F. U., and Schmid, B. (2007). Identification of anti-prion compounds as efficient inhibitors of polyglutamine protein aggregation in a zebrafish model. J Biol Chem *282*, 9195-9203.

Selkoe, D. J. (2003). Folding proteins in fatal ways. Nature *426*, 900-904.

Shewmaker, F., Mull, L., Nakayashiki, T., Masison, D. C., and Wickner, R. B. (2007). Ure2p function is enhanced by its prion domain in Saccharomyces cerevisiae. Genetics *176*, 1557-1565.

Shirahama, T., and Cohen, A. S. (1967). Reconstitution of amyloid fibrils from alkaline extracts. J Cell Biol *35*, 459-464.

Shorter, J., and Lindquist, S. (2004). Hsp104 catalyzes formation and elimination of self-replicating Sup35 prion conformers. Science *304*, 1793-1797.

Shorter, J., and Lindquist, S. (2005). Prions as adaptive conduits of memory and inheritance. Nat Rev Genet *6*, 435-450.

Si, K., Giustetto, M., Etkin, A., Hsu, R., Janisiewicz, A. M., Miniaci, M. C., Kim, J. H., Zhu, H., and Kandel, E. R. (2003). A neuronal isoform of CPEB regulates local protein synthesis and stabilizes synapse-specific long-term facilitation in aplysia. Cell *115*, 893-904.

Si, K., Lindquist, S., and Kandel, E. R. (2003). A neuronal isoform of the aplysia CPEB has prion-like properties. Cell *115*, 879-891.

Sigurdsson, E. M., Wisniewski, T., and Frangione, B. (2002). Infectivity of amyloid diseases. Trends Mol Med *8*, 411-413.

Sipe, J. D., and Cohen, A. S. (2000). Review: history of the amyloid fibril. J Struct Biol *130*, 88-98.

Slow, E. J., Graham, R. K., Osmand, A. P., Devon, R. S., Lu, G., Deng, Y., Pearson, J., Vaid, K., Bissada, N., Wetzel, R.*, et al.* (2005). Absence of behavioral abnormalities and neurodegeneration in vivo despite widespread neuronal huntingtin inclusions. Proc Natl Acad Sci U S A *102*, 11402-11407.

Smith, J. F., Knowles, T. P., Dobson, C. M., Macphee, C. E., and Welland, M. E. (2006). Characterization of the nanoscale properties of individual amyloid fibrils. Proc Natl Acad Sci U S A *103*, 15806-15811.

Sondheimer, N., and Lindquist, S. (2000). Rnq1: an epigenetic modifier of protein function in yeast. Mol Cell *5*, 163-172.

St George-Hyslop, P. H., Tanzi, R. E., Polinsky, R. J., Haines, J. L., Nee, L., Watkins, P. C., Myers, R. H., Feldman, R. G., Pollen, D., Drachman, D., and et al. (1987). The genetic defect causing familial Alzheimer's disease maps on chromosome 21. Science *235*, 885-890.

Sunde, M., Serpell, L. C., Bartlam, M., Fraser, P. E., Pepys, M. B., and Blake, C. C. (1997). Common core structure of amyloid fibrils by synchrotron X-ray diffraction. J Mol Biol *273*, 729-739.

Tartaglia, G. G., and Vendruscolo, M. (2008). The Zyggregator method for predicting protein aggregation propensities. Chem Soc Rev *37*, 1395-1401.

Taylor, K. L., Cheng, N., Williams, R. W., Steven, A. C., and Wickner, R. B. (1999). Prion domain initiation of amyloid formation in vitro from native Ure2p. Science *283*, 1339-1343.

Terry, R. D., Gonatas, N. K., and Weiss, M. (1964). Ultrastructural Studies In Alzheimer's Presenile Dementia. Am J Pathol *44*, 269-297.

Terry, R. D., Masliah, E., Salmon, D. P., Butters, N., DeTeresa, R., Hill, R., Hansen, L. A.,

and Katzman, R. (1991). Physical basis of cognitive alterations in Alzheimer's disease: synapse loss is the major correlate of cognitive impairment. Ann Neurol *30*, 572-580.

Tessier, P. M., and Lindquist, S. (2007). Prion recognition elements govern nucleation, strain specificity and species barriers. Nature *447*, 556-561.

Townsend, M., Shankar, G. M., Mehta, T., Walsh, D. M., and Selkoe, D. J. (2006). Effects of secreted oligomers of amyloid beta-protein on hippocampal synaptic plasticity: a potent role for trimers. J Physiol *572*, 477-492.

True, H. L., Berlin, I., and Lindquist, S. L. (2004). Epigenetic regulation of translation reveals hidden genetic variation to produce complex traits. Nature *431*, 184-187.

True, H. L., and Lindquist, S. L. (2000). A yeast prion provides a mechanism for genetic variation and phenotypic diversity. Nature *407*, 477-483.

Tsai, J., Grutzendler, J., Duff, K., and Gan, W. B. (2004). Fibrillar amyloid deposition leads to local synaptic abnormalities and breakage of neuronal branches. Nat Neurosci *7*, 1181-1183.

Uhlich, G. A., Keen, J. E., and Elder, R. O. (2001). Mutations in the csgD promoter associated with variations in curli expression in certain strains of *Escherichia coli* O157:H7. Appl Environ Microbiol *67*, 2367-2370.

Uhlich, G. A., Keen, J. E., and Elder, R. O. (2002). Variations in the csgD promoter of *Escherichia coli* O157:H7 associated with increased virulence in mice and increased invasion of HEp-2 cells. Infect Immun *70*, 395-399.

Van Nostrand, W. E., Melchor, J. P., Cho, H. S., Greenberg, S. M., and Rebeck, G. W. (2001). Pathogenic effects of D23N Iowa mutant amyloid beta -protein. J Biol Chem *276*, 32860-32866.

Walsh, D. M., Hartley, D. M., Kusumoto, Y., Fezoui, Y., Condron, M. M., Lomakin, A., Benedek, G. B., Selkoe, D. J., and Teplow, D. B. (1999). Amyloid beta-protein fibrillogenesis. Structure and biological activity of protofibrillar intermediates. J Biol

Chem *274*, 25945-25952.

Walsh, D. M., Klyubin, I., Fadeeva, J. V., Cullen, W. K., Anwyl, R., Wolfe, M. S., Rowan, M. J., and Selkoe, D. J. (2002). Naturally secreted oligomers of amyloid beta protein potently inhibit hippocampal long-term potentiation in vivo. Nature *416*, 535-539.

Wasmer, C., Lange, A., Van Melckebeke, H., Siemer, A. B., Riek, R., and Meier, B. H. (2008). Amyloid fibrils of the HET-s(218-289) prion form a beta solenoid with a triangular hydrophobic core. Science *319*, 1523-1526.

White, A. P., Collinson, S. K., Banser, P. A., Gibson, D. L., Paetzel, M., Strynadka, N. C., and Kay, W. W. (2001). Structure and characterization of AgfB from Salmonella enteritidis thin aggregative fimbriae. J Mol Biol *311*, 735-749.

Wickner, R. B. (1994). [URE3] as an altered URE2 protein: evidence for a prion analog in Saccharomyces cerevisiae. Science *264*, 566-569.

Williams, A. D., Portelius, E., Kheterpal, I., Guo, J. T., Cook, K. D., Xu, Y., and Wetzel, R. (2004). Mapping abeta amyloid fibril secondary structure using scanning proline mutagenesis. J Mol Biol *335*, 833-842.

Williams, A. D., Shivaprasad, S., and Wetzel, R. (2006). Alanine scanning mutagenesis of Abeta(1-40) amyloid fibril stability. J Mol Biol *357*, 1283-1294.

Wischik, C. M., Novak, M., Edwards, P. C., Klug, A., Tichelaar, W., and Crowther, R. A. (1988). Structural characterization of the core of the paired helical filament of Alzheimer disease. Proc Natl Acad Sci U S A *85*, 4884-4888.

Yankner, B. A., Dawes, L. R., Fisher, S., Villa-Komaroff, L., Oster-Granite, M. L., and Neve, R. L. (1989). Neurotoxicity of a fragment of the amyloid precursor associated with Alzheimer's disease. Science *245*, 417-420.

Yoshikawa, K., Aizawa, T., and Hayashi, Y. (1992). Degeneration in vitro of post-mitotic neurons overexpressing the Alzheimer amyloid protein precursor. Nature *359*, 64-67.

Zhang, B., Une, Y., Fu, X., Yan, J., Ge, F., Yao, J., Sawashita, J., Mori, M., Tomozawa, H., Kametani, F., and Higuchi, K. (2008). Fecal transmission of AA amyloidosis in the cheetah contributes to high incidence of disease. Proc Natl Acad Sci U S A *105*, 7263-7268.

Zheng, H., and Koo, E. H. (2006). The amyloid precursor protein: beyond amyloid. Mol Neurodegener *1*, 5.

Zogaj, X., Bokranz, W., Nimtz, M., and Romling, U. (2003). Production of cellulose and curli fimbriae by members of the family Enterobacteriaceae isolated from the human gastrointestinal tract. Infect Immun *71*, 4151-4158.

Chapter 2

Investigation of CsgA *in vitro* Polymerization[2]

Abstract

Amyloid formation is characterized by the conversion of soluble proteins into biochemically and structurally distinct fibers. Although amyloid formation is traditionally associated with diseases like Alzheimer's disease, a number of biologically functional amyloids have recently been described. Curli are amyloid fibers produced by *Escherichia coli* that contribute to biofilm formation and other important physiological processes. We characterized the polymerization properties of the major curli subunit protein CsgA. CsgA polymerizes into an amyloid fiber in a sigmoidal kinetic fashion with a distinct lag, growth and stationary phase. Adding sonicated preformed CsgA fibers to the polymerization reaction can significantly shorten the duration of the lag phase. I also demonstrated that the conversion of soluble CsgA into an insoluble fiber involves the transient formation of an intermediate similar to that characterized for several disease-associated amyloids. The CsgA core amyloid domain can be divided into 5 repeating units that share sequence and structural hallmarks. I showed that peptides representing three of these repeating units are amyloidogenic *in vitro*. Although the defining characteristics of CsgA polymerization appear conserved with disease-associated

[2] A version of this chapter has been published as Wang *et al*., 2007, J Biol Chem *282*, 3713-3719.

amyloids, these proteins evolved in diverse systems and for different purposes. Therefore, amyloidogenesis appears to be an innate protein folding pathway that can be capitalized on to fulfill normal physiological tasks.

Introduction

Amyloid formation is the hallmark of medically related disorders such as Alzheimer's disease, Huntington's disease, Parkinson's disease, and the transmissible spongiform encephalopathies (Chiti and Dobson, 2006). The root of these diseases is the uncontrolled conversion of seemingly unrelated soluble proteins into biochemically and structurally related fibers known as amyloids. Despite their diversity in size and amino acid content, all amyloid proteins assemble into 4-12nm wide fibers that are β-sheet rich and exhibit conserved tinctorial properties. Soluble pre-amyloid species also share common pore-like epitopes and these globular species may induce cytotoxicity (Hardy and Selkoe, 2002; Kayed et al., 2003; Quist et al., 2005).

Numerous studies have revealed that amyloidogenic proteins are mostly unstructured or contain mixtures of β-sheets and α-helices in their native state, but when polymerized into fibers they invariably adopt a characteristic cross β-sheet structure (Barrow and Zagorski, 1991; Barrow et al., 1992; Hurshman et al., 2004). This cross β-sheet structure is common to all amyloids and is characterized by β-strands that orient perpendicular to the fiber axis. *In vitro,* disease-associated amyloids polymerize into fibers with nucleation-dependent kinetics with characteristic lag, growth and stationary

phase. The lag phase is proposed to contain folding intermediates that are key to the toxicity associated with certain amyloidogenic proteins (Lashuel et al., 2002; Lesne et al., 2006). During the lag phase, amyloidogenic proteins adopt a transient folding species that disrupts membrane integrity (Bucciantini et al., 2002; Kayed et al., 2003; Lashuel et al., 2002). Loss of membrane integrity is proposed to underlie the cell death and disease associated with many amyloids (Bucciantini et al., 2002; Kayed et al., 2003). A conformational-specific antibody has been generated that recognizes a transient intermediate formed during amyloidogenesis of several disease-associated proteins (Kayed et al., 2003).

A new class of amyloids has recently been found that play important physiological roles for the cell. These so-called 'functional' amyloids are found in bacteria (Bieler et al., 2005; Chapman et al., 2002; Elliot et al., 2003), fungi (Coustou-Linares et al., 2001; True and Lindquist, 2000) and mammals (Fowler et al., 2005). The first example of a functional amyloid in bacteria was curli (Chapman et al., 2002). Curli compose part of the complex extracellular matrix that is required for biofilm formation (Austin et al., 1998; Zogaj et al., 2003; Zogaj et al., 2001), host cell adhesion (Johansson et al., 2001) and invasion (Gophna et al., 2001; Gophna et al., 2002), and they are proposed to be important stimulants of the host inflammatory response (Bian et al., 2000; Bian et al., 2001; Tukel et al., 2005). An intriguing question is whether these functional amyloid proteins polymerize in a manner similar to disease-associated amyloids.

Curli formation is the result of an elegant biosynthetic pathway directed by the

Csg proteins in *E. coli.* The major curli subunit, CsgA, can be secreted to the cell surface as a soluble, unstructured protein (Barnhart and Chapman, 2006; Chapman et al., 2002). CsgA is efficiently nucleated into an insoluble amyloid fiber in the presence of the outer membrane-associated protein, CsgB (27). After nucleation, the fibers are predicted to grow by subsequent CsgA addition to the amyloid fiber's tip (Barnhart and Chapman, 2006).

Both CsgA and CsgB display a remarkable five fold internal symmetry characterized by conserved polar residues. These five repeating units consist of 19-24 amino acids and align along serine, glutamine and asparagine residues (Barnhart and Chapman, 2006; Collinson et al., 1999). Each repeating unit is predicted to form a strand-loop-strand motif that closely resembles the cross β-spine structure described for many disease-associated amyloids (Collinson et al., 1999; Nelson and Eisenberg, 2006; Nelson et al., 2005).

Here we characterized the folding of purified CsgA and showed that its polymerization is similar to that of disease-associated amyloids. CsgA polymerization involves a transient structurally conserved intermediate that implies a common polymerization pathway between functional and disease-associated amyloids. We found that the conserved folding intermediate for CsgA is a monomer or low molecular weight multimer. We demonstrated that at least three of five repeating units of CsgA are amyloidogenic. These results suggest that covalent linkage of multiple amyloidogenic units facilitates efficient fiber formation.

Experimental Procedures

CsgA Purification

CsgG and CsgA-his were overexpressed in LSR12 (C600::Δ*csgBA* and Δ*csgDEFG*) as previously described (Chapman et al., 2002). Following centrifugation for 15 minutes at 10,000g, the supernatant was clarified by filtration though a 0.22 μm PES bottle-top filter (Corning, Acton, MA). Filtrates containing CsgA were passed over a HIS-Select™ HF NiNTA (Sigma Aldrich, Atlanta, GA) column, washed with 10 volumes of 10 mM potassium phosphate buffer (KPi) pH 7.2, and eluted with 10 mM KPi 100 mM imidazole pH 7.2. CsgA-containing fractions were combined and passed through a 0.02 μm Anotop 10 filter (Whatman, Maidstone, England). The N-terminal portion of purified CsgA-his was sequenced by mass spectrometry and found to be identical to the predicted CsgA amino acid sequence. A modified protocol using guanidine hydrochloride (GdnHCl) was employed to fully denature CsgA-his. Following the first wash, the column was equilibrated with 5 volumes 10 mM KPi 8 M GdnHCl pH 7.2 and eluted with 50 mM KPi 8 M GdnHCl pH 2. Sephadex G25 was used for desalting/buffer exchange. To create CsgA-his seeds, 2-week-old fibers were sonicated using a Sonic Dismembrator (Fisher Model 100, Fisher, Pittsburg, PA) for three 15-second bursts on ice. Where indicated, CsgA samples were filtered through a prewashed Amicon Ultra-4 (Millipore, Bedford, MA) centrifugal filter devices. Samples were centrifuged at 4,000g for 2 minutes and the retentate and filtrate fractions were collected. A plasmid encoding CsgA-his can complement *ΔcsgA* cells *in vivo* and purified CsgA-his polymerizes into an

amyloid fiber with similar kinetics as wild-type CsgA (Chapman et al., 2002). Therefore, CsgA-his is referred to as 'CsgA' throughout this chapter.

Thioflavin T (ThT) assay

Following desalting to remove imidazole or GdnHCl, CsgA was incubated at room temperature. At different time intervals, CsgA samples were mixed with 20 μM ThT and fluorescence was measured using a Spectramax M2 plate reader (Molecular Devices, Sunnyvale, CA) set to 438nm excitation and 495nm emission with a 475nm cutoff. Alternatively, samples amended with 25 μM ThT were assayed directly in the Spectramax M2 plate reader every 10 minutes after shaking for 5 seconds.

Circular Dichroism (CD) Spectroscopy

CsgA samples (10 μM CsgA in 50 mM KPi, pH 7.2) were assayed in a Jasco J-810 spectropolarimeter from 190 to 250nm in a quartz cell with 1-mm path length at 25 °C.

Blot assay with A11 and anti-CsgA antibody

CsgA samples were dripped onto 0.2μm Transblot Nitrocellulose membranes (Bio-Rad, Hercules, CA) as described (Kayed et al., 2003) and allowed to dry for 5 minutes. The membrane was blocked in 5% milk in TBS-T (0.01% Tween-20) for at least 1 hour. The dot blots were washed 3 times in TBS-T before and after incubating with 1:10000 dilutions of A11 primary antibody (BioSource, Camarillo, California) and goat anti-rabbit-HRP (Sigma-Aldrich) in 3%BSA/TBS-T. The blots were developed using the SuperSignal® West Dura system (Pierce, Rockford, IL). Blots were stripped and

reprobed with a 1:10000 dilution of rabbit anti-CsgA antibody (Barnhart et al., 2006).

Electron microscopy

Philips CM12 Scanning transmission electron microscope was used to visualize the fiber aggregates. Samples (10 μl) were placed on formvar coated copper grids (Ernest F. Fullam, Inc, Latham, NY) for 2 min, washed with deionized water, and negatively stained with 2% uranyl acetate for 90 seconds.

Peptide preparation

Peptides were chemically synthesized by Proteintech Group Inc., Chicago. Purity was greater than 90% by HPLC and size was confirmed by mass spectroscopy. To equilibrate the pH of each sample and to remove any potential seed in the peptide preparations, the peptides were denatured using a modified protocol described previously (Chen and Wetzel, 2001). Briefly, peptides were dissolved to 0.5 mg/ml in TFA/HFIP (1:1 v/v) and sonicated for 10 min. The suspensions were incubated at room temperature until they visibly cleared. The solvent was then removed by vacuum. Peptides were then dissolved in cold 2 mM HCl and centrifuged at 100,000 X g in a TLA-55 (Beckman, Fullerton, CA) for 1 hour at 4°C. The supernatants were equilibrated to 50mM K-Pi pH7.2 by 200mM K-Pi pH7.2 on ice. When the samples were shifted to room temperature, the polymerization was measured by ThT assay.

Results

CsgA polymerization kinetics

To determine the polymerization kinetics of CsgA, an *in vitro* polymerization assay was developed. The transition of freshly purified, soluble CsgA to amyloid fibers was monitored using thioflavin T (ThT), an amyloid-specific dye commonly used to assay amyloid formation (LeVine, 1993; LeVine, 1999). The ThT fluorescence of CsgA samples followed a sigmoidal curve with distinguishable lag, growth and stationary phases (Figure 2.1A). The length of the lag phase was concentration independent when CsgA was incubated at concentrations above 4 μM (Figure 2.1A). Concentration-independent lag phases have been reported for other amyloidogenic proteins (Lomakin et al., 1997; Rhoades and Gafni, 2003). ThT fluorescence signal did not appreciably change after 8 hours, remaining at approximately the same level for over 30 days (data not shown).

Circular dichroism spectroscopy and transmission electron microscopy (TEM) were used to measure the structural changes that occurred during CsgA amyloidogenesis. Circular dichroism spectrum indicated that immediately after purification CsgA was largely unstructured (Figure 2.1B). However, CsgA adopted a β-sheet-rich structure after 2 days of incubation at room temperature (Figure 2.1B). Immediately after purification there was no apparent fiber formation or aggregation by TEM (Figure 2.1C). Two hours after purification regular, unbranched fibers were readily observed (Figure 2.1D). Dense fiber aggregates were also observed 7 days post purification (Figure 2.1E).

The A11 antibody recognizes a transient CsgA folding species

The polymerization of eukaryotic amyloids involves the formation of an

intermediate folding species proposed to cause amyloid-associated toxicity to host cells (Hardy and Selkoe, 2002; Kayed et al., 2003). The A11 antibody recognizes an Aβ transient intermediate (Kayed et al., 2003). Remarkably, this antibody also recognizes a transient intermediate formed by IAPP, poly Q, PrP, and Sup35p, among others (Kayed et al., 2003; Shorter and Lindquist, 2004). The A11 antibody recognizes only a transient intermediate species, not soluble monomers or mature amyloid fibers derived from these proteins.

The A11 antibody was used to determine if CsgA shared a common polymerization intermediate with eukaryotic amyloids. We found that immediately after purification CsgA was recognized by the A11 antibody (Figure 2.2A). As fiber formation proceeded, evidenced by increased ThT fluorescence and the appearance of fiber aggregates by TEM, the A11 antibody lost its affinity for CsgA (Figure 2.2A). A polyclonal antibody generated against CsgA recognized purified CsgA independent of its polymerization status (Figure 2.2A).

The observation that the A11 antibody recognized CsgA suggested that CsgA polymerization intermediates might be structurally similar to those formed by disease-associated amyloid proteins. It also suggested that, immediately after purification, CsgA had already begun its transition to an amyloid fiber. To prevent CsgA from folding during purification, the CsgA-containing fractions were amended with 8 M GdnHCl. Under these strongly denaturing conditions, the A11 antibody did not recognize CsgA; however, denatured CsgA was strongly recognized by the CsgA antibody (Figure 2.2B).

Immediately after GdnHCl removal with a desalting column, CsgA was recognized by the A11 antibody (Figure 2. 2B).

To determine the minimum size of the CsgA transient intermediate, freshly purified protein was passed through Amicon Ultra centrifugal membranes with different molecular weight cutoffs. The retentate and filtrate were probed with the A11 antibody (Figure 2.2C). The A11 antibody recognized a species in the filtrate of the 30kD membrane, suggesting that the smallest reactive species of CsgA is 30 kD or less. Because CsgA-his has a predicted molecular mass of 13.9 kD, the species recognized by the A11 antibody is likely either a monomer or dimer.

CsgA fibers can catalyze self-polymerization

The approximately sigmoidal ThT fluorescence curve suggests that CsgA polymerizes in a nucleation-dependent mechanism. Therefore, the growing fiber would be expected to direct the polymerization of new CsgA molecules. We tested the hypothesis that preformed CsgA fibers could catalyze CsgA polymerization. Addition of 2.5% (w/w) sonicated CsgA fibers to freshly purified, soluble CsgA resulted in a significant reduction of the lag phase (Figure 2.3A). Coincident with the dramatically shorter lag phase in seeded reactions, CsgA was recognized by the A11 antibody for a significantly shorter time compared to unseeded reactions (Figure 2.3B).

CsgA is composed of several amyloid-forming units

The observation that CsgA was recognized by the A11 antibody immediately after passing through a 30 kDa cutoff filter (Figure 2.2C) was unexpected since the A11

antibody is thought to recognize an oligomeric form of amyloidogenic proteins (Kayed et al., 2003; Shorter and Lindquist, 2004; Lesne et al., 2006). The number of molecules present in the oligomeric state recognized by A11 varies among amyloidogenic proteins, and Aβ oligomers have been estimated to be larger than tetramers (Kayed et al., 2003; Lesne et al., 2006). However, CsgA is recognized by A11 as a monomer or at most a dimer as estimated by cutoff filtration. It is possible that a single CsgA molecule includes multiple amyloidogenic domains that collectively contribute to its interaction with the A11 antibody. The primary sequence of CsgA can be divided into three parts: the Sec-dependent signal sequence, the N-terminal 22 amino acids of the mature protein, and a repeat domain that contains five 19-22 amino acid repeating units (Figure 2.4A). The five repeating units form a protease resistant structure that is proposed to be the amyloid core of CsgA (Barnhart and Chapman, 2006; Collinson et al., 1999). Each repeat has four conserved polar amino acids: serine, glutamine, asparagine and glutamine (Figure 2.4A). The regular arrangement of glutamine and asparagine residues also occurs in CsgA homologs from different *Enterobacteriaceae* (Wang and Chapman unpublished observation).

We hypothesized that these repeating units might represent single amyloid forming units. Peptides corresponding to each repeating unit were chemically synthesized and tested for their ability to form amyloid fibers. Two independently derived preparations of each peptide were assayed. Peptides corresponding to repeating unit 1, 3 or 5 (R1, R3 or R5) produced a ThT-positive signal and self-assembled into fibers as

evidenced by TEM when incubated at 0.2 mg/ml (Figure 2.4B-E). Neither R2 nor R4 showed evidence of amyloidogenesis when resuspended at a concentration of 0.2 mg/ml, although fibers were observed by TEM when R2 or R4 were incubated at 2 mg/ml (Figure 2.4B and data not shown). The morphology of R1 fibers was similar to those formed by purified CsgA, being generally longer than 1000nm (compare Figure 2.4C to Figure 2.1D and 2.1E). R3 fibers were consistently shorter (ranging from 200nm-1000nm) than those formed by CsgA (compare Figure 2.4D to Figure 2.1D and 2.1E). R5 fibers appeared more rigid and aggregated than CsgA fibers (Figure 2.4E). The morphologies of the fibers did not appreciably change over the course of a ten-day incubation. This analysis suggests that CsgA contains at least 3 highly amyloidogenic domains, R1, R3 and R5, that likely drive fiber formation *in vivo.*

Discussion

Amyloid formation is traditionally associated with uncontrolled protein misfolding and aggregation that results in many systemic and neurodegenerative disorders (Chiti and Dobson, 2006; Shorter and Lindquist, 2005). However, there are a growing number of functional amyloids that suggest amyloidogenesis is also a general tenet of normal cellular physiology. In fact, amyloid formation may be a common property of most proteins (Fandrich et al., 2001; Guijarro et al., 1998).

The work presented in this chapter, as well as that published previously, demonstrates that both disease-associated and functional amyloids share a common

amyloid formation pathway (Glover et al., 1997). CsgA polymerizes with nucleation-dependent kinetics and fiber formation is ameliorated by the addition of pre-formed CsgA fibers. We also found that CsgA polymerization involves the formation of a transient species similar to that produced by other amyloidogenic proteins such as Aβ, synuclein, IAPP, insulin, lysozyme and polyglutamine (Kayed et al., 2003).

The transient species that the A11 antibody recognizes during CsgA polymerization is a monomer or low-molecular weight multimer (Figure 2.2C). It was reported that the A11-recognized species of Aβ and Sup35p were probably large molecular weight oligomers (Kayed et al., 2003; Shorter and Lindquist, 2004). Unlike Aβ and Sup35p (Kayed et al., 2003; Shorter and Lindquist, 2004), CsgA was immediately recognized by the A11 antibody upon removal of strong denaturants like GdnHCl or after its passage through a 30 kD Amicon filter. At least two hypotheses can be proposed to explain CsgA's ability to be recognized by A11 immediately after denaturation or passage through a 30 kD cutoff filter. First, CsgA may adopt an oligomeric conformation so quickly that our ability to measure this transition is lost in the time that it takes to immobilize CsgA on the blotting paper. Another possibility is that the CsgA species recognized by A11 is not an oligomer, but a monomer that contains multiple amyloidogenic units. It is possible that CsgA monomer or dimer rapidly folds into a conformation recognized by A11. Nevertheless, these two hypotheses are not mutually exclusive, and there may be other plausible interpretations.

Nonetheless, CsgA contains multiple amyloidogenic domains that may contribute

to its ability to efficiently transition from a soluble protein to an amyloid fiber. Many studies have led to the proposal that amyloid fibers themselves are not toxic to cells, instead toxicity is proposed to be caused by transient folding intermediates (Demuro et al., 2005; Glabe and Kayed, 2006; Hardy and Selkoe, 2002). Therefore, one mechanism that might be used by functional amyloids to prevent toxicity is to minimize the duration of toxic folding intermediates. This is apparently how Pmel17, an extremely rapidly forming functional amyloid found in mammalian cells, is able to assemble within the cell without eliciting a toxicity cascade (Fowler et al., 2005).

CsgA has a striking primary sequence arrangement (Figure 2.4A). The five repeats of CsgA are very similar and share greater than 30% sequence identity. We showed that each repeating unit is potentially a single amyloid domain and that R1, R3 and R5 are the highly amyloidogenic *in vitro* (Figure 2.4B-E). The covalent linkage of multiple amyloid domains may facilitate amyloid fiber formation by increasing the number of amyloidogenic building blocks. This also results in rapid formation of the intermediate recognized by the A11 antibody. Other amyloidogenic proteins contain repeat sequences that have been postulated to facilitate fiber formation (Ross et al., 2005; Wright et al., 2005). For instance, the N-terminal prion-determining domain of Sup35p has five imperfect oligopeptide repeats and certain deletions of the repeats are defective in propagation of Sup35p fibrils. Moreover, *in vitro*, repeat-expansion peptides (with 2 extra repeats) were shown to be more amyloidogenic than wild-type peptides (Liu and Lindquist, 1999). Previous work suggested the most amyloidogenic domains of CsgA

were contained in the hexapeptide GHGGGN and QFGGGN, which are present in R2 and R4 respectively (Cherny et al., 2005). However, our analysis suggests that R1, R3 and R5 contain the more highly amyloidogenic sequences. A thorough mutagenesis study in later chapters defined the residues that contribute to the highly amyloidogenic nature and interaction specificity of CsgA.

The amyloidogenic peptides R1 and R5 contain sequences that contribute significantly to the ability of CsgA to bind human proteins such as fibronectin, plasminogen, tissue plasminogen activator, and β2-microglobulin (Olsen et al., 2002). This correlation suggests amyloidogenecity of CsgA may be directly linked to these biological activities. In fact, work by Gebbink *et al* suggested that curli contribute to colonization in animal hosts by activating host proteases that are involved in haemostasis (Gebbink et al., 2005).

Curli can also enhance amyloid protein A amyloidosis in mice (Lundmark et al., 2005). It is proposed that cross-seeding may play a role in the development of amyloid diseases (Kisilevsky, 2000; Lundmark et al., 2005). The *in vitro* system that we have established here provides an ideal vehicle to test the specificity of curli seeding with other amyloids. Understanding how functional amyloid proteins interact with other host proteins may lead to new ideas about cellular physiology and the processes that promote the toxicity associated with many amyloids.

Most amyloids are known to self-propagate in a process called seeding. In prion diseases such as Bovine Spongiform Encephalopathy, seeding underlies protein

infectivity (Prusiner, 1998). Amyloid self-propagation is also critical to disease development in the non-transmissible amyloid diseases(Kisilevsky, 2000; Lansbury, 1999). Our demonstration of CsgA seeding suggests that functional amyloids also utilize a controlled self-propagation process to fulfill their biological function. *In vivo* CsgA polymerization is nucleated by the outer membrane associated protein CsgB, which shares nearly 49% sequence similarity with CsgA (Hammar et al., 1996). One proposed model of nucleation is that CsgB provides an amyloid-like template that initiates CsgA polymerization (Barnhart and Chapman, 2006; Hammar et al., 1996). The growing fiber tip could then act as a template to direct subsequent CsgA polymerization.

Proteins that are not predicted to form stable globular folds may be prone to aggregation and amyloid formation, and indeed most functional amyloid proteins have natively disordered segments (Nelson and Eisenberg, 2006; Uversky et al., 2000). Consistent with this, some proteins have been shown to form amyloid fibers only after the native, globular fold has been compromised by chemical denaturants or by mutations (Fandrich et al., 2001; Guijarro et al., 1998; Booth et al., 1997). Circular dichroism studies presented here suggest that CsgA is unstructured after secretion (Figure 2.1B). Mature CsgA is also predicted to be natively unfolded (data not shown) by the Uversky and Galzitskaya algorithms (Galzitskaya and Garbuzynskiy, 2006; Uversky et al., 2000). The natively unfolded segments of CsgA may facilitate amyloidogenesis indirectly by preventing formation of stable globular structures that would be less likely to aggregate and precipitate into a fiber. Alternatively, the unfolded regions of CsgA may actively

direct amyloidogenesis by presenting specific aggregative surfaces to neighboring molecules. In this context ‘natively unfolded’ would be a transition state during the formation of a stably folded amyloid fiber. Importantly, in the case of functional amyloids the amyloid fiber would not be the product of protein misfolding, but that of protein folding. Certainly, the growing number of functional amyloids suggests that amyloid is an evolutionarily conserved structure. The selective processes that have been employed by functional amyloids to limit cellular toxicity provide a unique context from which to investigate disease-associated amyloidogenesis.

Figures

Figure 2.1. *In vitro* polymerization of CsgA measured by ThT fluorescence, CD and TEM.

(A) The fluorescence of freshly purified CsgA mixed with 25 μM ThT was measured in 10-minute intervals at 495nm after excitation at 438nm. (B) Circular dichroism analysis of 10 μM CsgA immediately after purification, 2 days post-purification, 7 days post-purification and 15 days post-purification. CsgA was incubated at room temperature without shaking after purification. (C-E) Transmission Electron Microscopy (TEM) micrographs of 30 μM CsgA after incubation at room temperature for the indicated times. (Scale bar: 500nm)

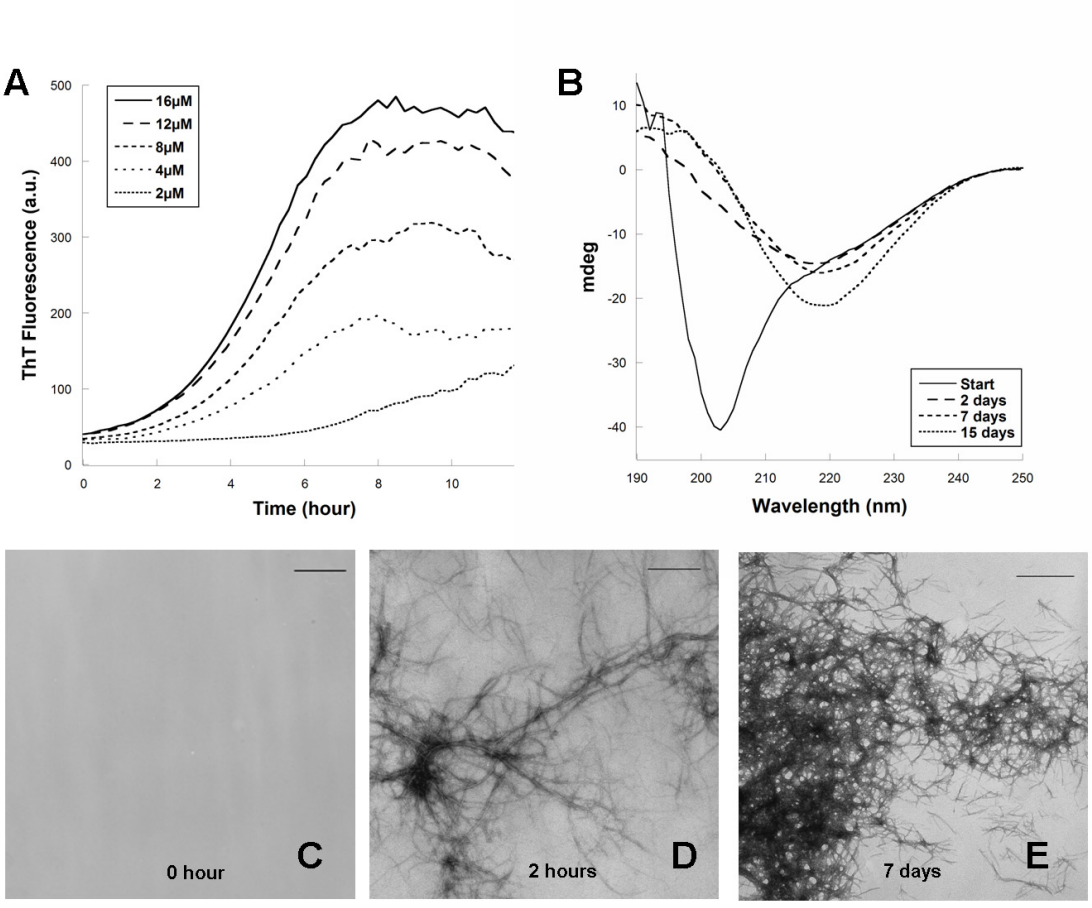


Figure 2.2. Detection of transient conserved intermediate species during CsgA polymerization.

(A) ThT fluorescence (top) and immunoblotting (bottom) of 80 μM CsgA incubated for the indicated time post-purification. At the indicated times, samples were removed, ThT was added to a final concentration of 20 μM, and fluorescence was measured. Samples were blotted onto nitrocellulose membrane and probed with the A11 antibody, and after stripping, with the anti-CsgA antibody. (B) CsgA denatured with 8 M GdnHCl was blotted onto nitrocellulose and probed with the A11 and anti-CsgA antibodies (left). GdnHCl was removed using a Sephadex G25 column (final buffer: 50 mM KPi pH 7.2) and then immediately blotted onto nitrocellulose and probed with the A11 and anti-CsgA antibodies (right). (C) Amicon ultra filters were used to separate CsgA solutions prior to probing with the A11 and anti-CsgA antibodies. The molecular weight cutoff of the filters is indicated. Retentates and filtrates were immediately blotted onto nitrocellulose and probed with the A11 antibody, and after stripping, with the anti-CsgA antibody.

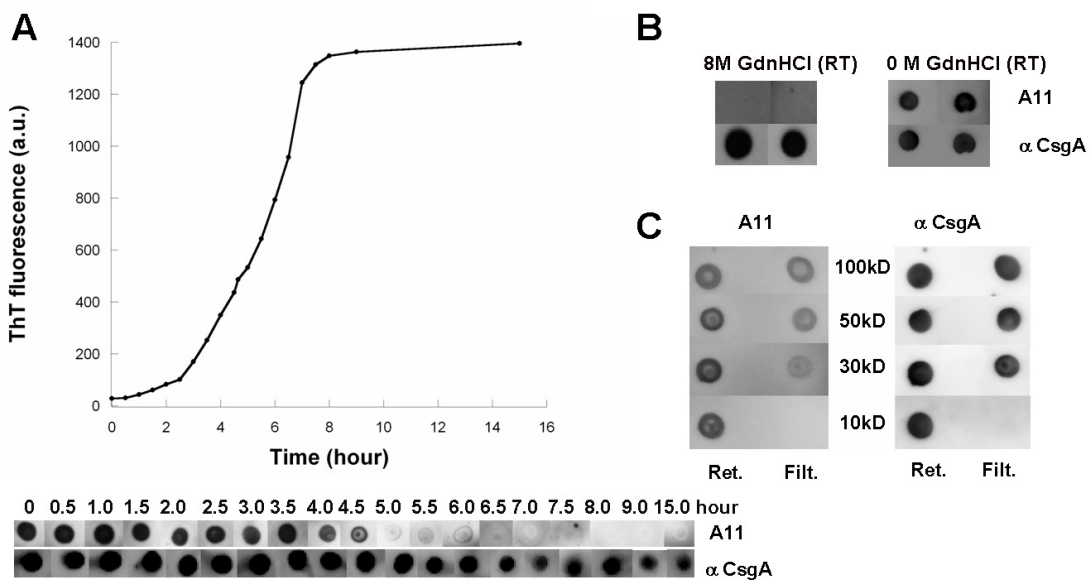


Figure 2.3. CsgA fibers can catalyze self-polymerization.
(A) CsgA (40 μM) fluorescence in the absence (solid line) or presence of 2.5% by weight of sonicated CsgA fibers (dashed line). Samples were incubated at room temperature, collected at the indicated times and amended with 20 μM ThT prior to excitation at 438nm and measurement at 495nm. (B) Samples were removed at the indicated times and immediately blotted to nitrocellulose. Blots were probed with the A11 antibody and after stripping with an anti-CsgA antibody.

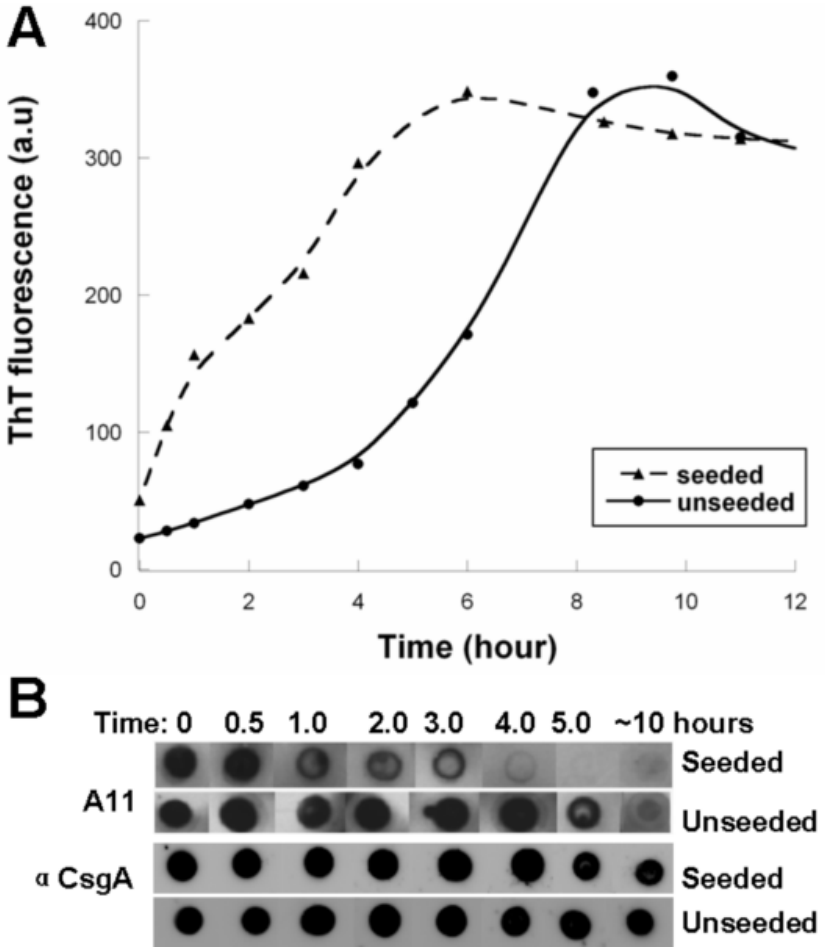
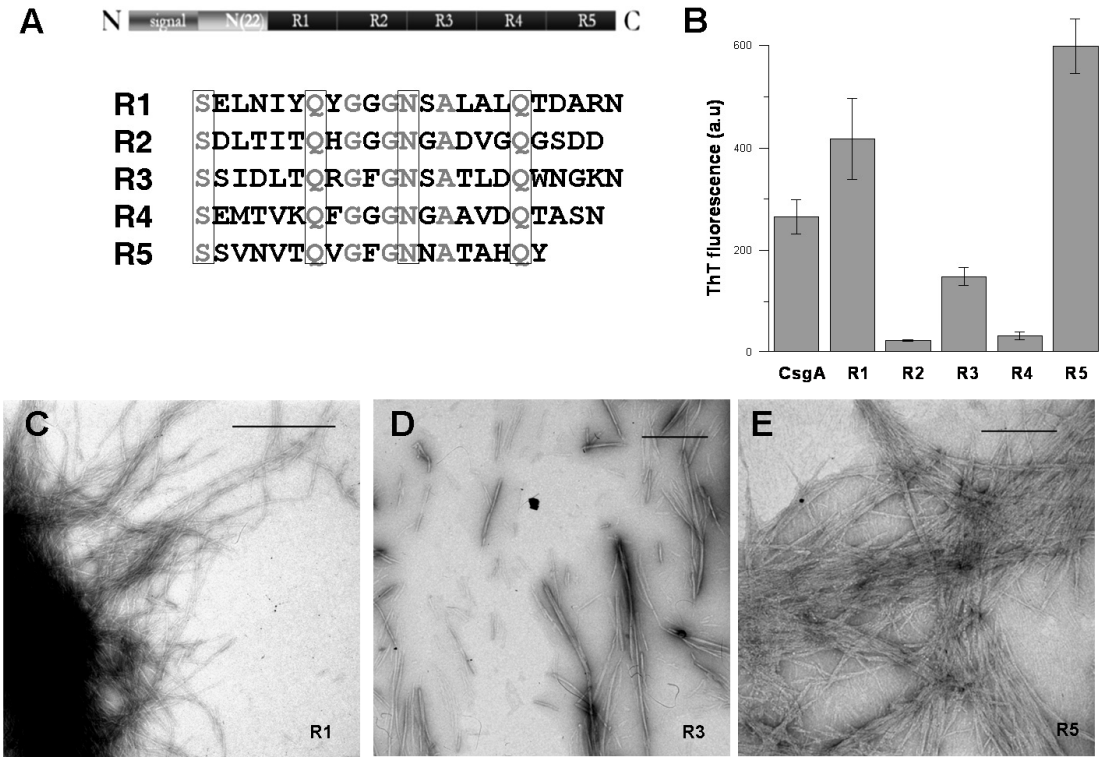


Figure 2.4. Three CsgA intramolecular peptide repeats can assemble into amyloid fibers.

(A) Alignment of internally conserved residues. CsgA primary sequence shows the repeated consensus sequences. The identical amino acid residues of five repeats are in gray color and the conserved polar amino acid residues are enclosed in 4 boxes.

(B) Oligopeptides of R1, R2, R3, R4 and R5 at 0.2 mg/ml in KPi were incubated at room temperature for 5 days before ThT fluorescence measurements were taken. Error bar indicates the standard error mean for at least three separate experiments. CsgA fibers were diluted to 0.2 mg/ml and assayed for ThT fluorescence. (C-E) Transmission Electron Microscopy (TEM). 0.5 mg/ml of R1, R3 and R5 in pH7.2 KPi were incubated at room temperature for 5 days. Samples of different peptide solutions were directly applied on formvar-coated grids and visualized with negative staining electron microscopy. Scale bar is equal to 500nm.



References

Austin, J. W., Sanders, G., Kay, W. W., and Collinson, S. K. (1998). Thin aggregative fimbriae enhance *Salmonella enteritidis* biofilm formation. FEMS Microbiol Lett *162*, 295-301.

Barnhart, M. M., and Chapman, M. R. (2006). Curli biogenesis and function. Annu Rev Microbiol *60*, 131-147.

Barnhart, M. M., Lynem, J., and Chapman, M. R. (2006). GlcNAc-6P levels modulate the expression of curli fibers by *Escherichia coli*. J Bacteriol *188*, 5212-5219.

Barrow, C. J., and Zagorski, M. G. (1991). Solution structures of beta peptide and its constituent fragments: relation to amyloid deposition. Science *253*, 179-182.

Barrow, C. J., Yasuda, A., Kenny, P. T., and Zagorski, M. G. (1992). Solution conformations and aggregational properties of synthetic amyloid beta-peptides of Alzheimer's disease. Analysis of circular dichroism spectra. J Mol Biol *225*, 1075-1093.

Bian, Z., Brauner, A., Li, Y., and Normark, S. (2000). Expression of and cytokine activation by *Escherichia coli* curli fibers in human sepsis. J Infect Dis *181*, 602-612.

Bian, Z., Yan, Z. Q., Hansson, G. K., Thoren, P., and Normark, S. (2001). Activation of inducible nitric oxide synthase/nitric oxide by curli fibers leads to a fall in blood pressure during systemic *Escherichia coli* infection in mice. J Infect Dis *183*, 612-619.

Bieler, S., Estrada, L., Lagos, R., Baeza, M., Castilla, J., and Soto, C. (2005). Amyloid formation modulates the biological activity of a bacterial protein. J Biol Chem *280*,

26880-26885.

Booth, D. R., Sunde, M., Bellotti, V., Robinson, C. V., Hutchinson, W. L., Fraser, P. E., Hawkins, P. N., Dobson, C. M., Radford, S. E., Blake, C. C., and Pepys, M. B. (1997). Instability, unfolding and aggregation of human lysozyme variants underlying amyloid fibrillogenesis. Nature *385*, 787-793.

Bucciantini, M., Giannoni, E., Chiti, F., Baroni, F., Formigli, L., Zurdo, J., Taddei, N., Ramponi, G., Dobson, C. M., and Stefani, M. (2002). Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. Nature *416*, 507-511.

Chapman, M. R., Robinson, L. S., Pinkner, J. S., Roth, R., Heuser, J., Hammar, M., Normark, S., and Hultgren, S. J. (2002). Role of *Escherichia coli* curli operons in directing amyloid fiber formation. Science *295*, 851-855.

Chen, S., and Wetzel, R. (2001). Solubilization and disaggregation of polyglutamine peptides. Protein Sci *10*, 887-891.

Cherny, I., Rockah, L., Levy-Nissenbaum, O., Gophna, U., Ron, E. Z., and Gazit, E. (2005). The formation of *Escherichia coli* curli amyloid fibrils is mediated by prion-like peptide repeats. J Mol Biol *352*, 245-252.

Chiti, F., and Dobson, C. M. (2006). Protein misfolding, functional amyloid, and human disease. Annu Rev Biochem *75*, 333-366.

Collinson, S. K., Parker, J. M., Hodges, R. S., and Kay, W. W. (1999). Structural predictions of AgfA, the insoluble fimbrial subunit of *Salmonella* thin aggregative

fimbriae. J Mol Biol *290*, 741-756.

Coustou-Linares, V., Maddelein, M. L., Begueret, J., and Saupe, S. J. (2001). In vivo aggregation of the HET-s prion protein of the fungus Podospora anserina. Mol Microbiol *42*, 1325-1335.

Demuro, A., Mina, E., Kayed, R., Milton, S. C., Parker, I., and Glabe, C. G. (2005). Calcium dysregulation and membrane disruption as a ubiquitous neurotoxic mechanism of soluble amyloid oligomers. J Biol Chem *280*, 17294-17300.

Elliot, M. A., Karoonuthaisiri, N., Huang, J., Bibb, M. J., Cohen, S. N., Kao, C. M., and Buttner, M. J. (2003). The chaplins: a family of hydrophobic cell-surface proteins involved in aerial mycelium formation in Streptomyces coelicolor. Genes Dev *17*, 1727-1740.

Fandrich, M., Fletcher, M. A., and Dobson, C. M. (2001). Amyloid fibrils from muscle myoglobin. Nature *410*, 165-166.

Fowler, D. M., Koulov, A. V., Alory-Jost, C., Marks, M. S., Balch, W. E., and Kelly, J. W. (2005). Functional Amyloid Formation within Mammalian Tissue. PLoS Biol *4*, e6.
Galzitskaya, O. V., and Garbuzynskiy, S. O. (2006). Entropy capacity determines protein folding. Proteins *63*, 144-154.

Gebbink, M. F., Claessen, D., Bouma, B., Dijkhuizen, L., and Wosten, H. A. (2005). Amyloids--a functional coat for microorganisms. Nat Rev Microbiol *3*, 333-341.
Glabe, C. G., and Kayed, R. (2006). Common structure and toxic function of amyloid oligomers implies a common mechanism of pathogenesis. Neurology *66*, S74-8.

Glover, J. R., Kowal, A. S., Schirmer, E. C., Patino, M. M., Liu, J. J., and Lindquist, S. (1997). Self-seeded fibers formed by Sup35, the protein determinant of [PSI+], a heritable prion-like factor of S. cerevisiae. Cell *89*, 811-819.

Gophna, U., Barlev, M., Seijffers, R., Oelschlager, T. A., Hacker, J., and Ron, E. Z. (2001). Curli fibers mediate internalization of *Escherichia coli* by eukaryotic cells. Infect Immun *69*, 2659-2665.

Gophna, U., Oelschlaeger, T. A., Hacker, J., and Ron, E. Z. (2002). Role of fibronectin in curli-mediated internalization. FEMS Microbiol Lett *212*, 55-58.

Guijarro, J. I., Sunde, M., Jones, J. A., Campbell, I. D., and Dobson, C. M. (1998). Amyloid fibril formation by an SH3 domain. Proc Natl Acad Sci U S A *95*, 4224-4228.
Hammar, M., Bian, Z., and Normark, S. (1996). Nucleator-dependent intercellular assembly of adhesive curli organelles in *Escherichia coli*. Proc Natl Acad Sci U S A *93*, 6562-6566.

Hardy, J., and Selkoe, D. J. (2002). The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. Science *297*, 353-356.

Hurshman, A. R., White, J. T., Powers, E. T., and Kelly, J. W. (2004). Transthyretin aggregation under partially denaturing conditions is a downhill polymerization. Biochemistry *43*, 7365-7381.

Johansson, C., Nilsson, T., Olsen, A., and Wick, M. J. (2001). The influence of curli, a MHC-I-binding bacterial surface structure, on macrophage-T cell interactions. FEMS

Immunol Med Microbiol *30*, 21-29.

Kayed, R., Head, E., Thompson, J. L., McIntire, T. M., Milton, S. C., Cotman, C. W., and Glabe, C. G. (2003). Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis. Science *300*, 486-489.

Kisilevsky, R. (2000). Review: amyloidogenesis-unquestioned answers and unanswered questions. J Struct Biol *130*, 99-108.

Lansbury, P. T. J. (1999). Evolution of amyloid: what normal protein folding may tell us about fibrillogenesis and disease. Proc Natl Acad Sci U S A *96*, 3342-3344.

Lashuel, H. A., Hartley, D., Petre, B. M., Walz, T., and Lansbury, P. T., Jr. (2002). Neurodegenerative disease: amyloid pores from pathogenic mutations. Nature *418*, 291.

Lesne, S., Koh, M. T., Kotilinek, L., Kayed, R., Glabe, C. G., Yang, A., Gallagher, M., and Ashe, K. H. (2006). A specific amyloid-beta protein assembly in the brain impairs memory. Nature *440*, 352-357.

LeVine, H. r. (1993). Thioflavine T interaction with synthetic Alzheimer's disease beta-amyloid peptides: detection of amyloid aggregation in solution. Protein Sci *2*, 404-410.

LeVine, H. r. (1999). Quantification of beta-sheet amyloid fibril structures with thioflavin T. Methods Enzymol *309*, 274-284.

Liu, J. J., and Lindquist, S. (1999). Oligopeptide-repeat expansions modulate

'protein-only' inheritance in yeast. Nature *400*, 573-576.

Lomakin, A., Teplow, D. B., Kirschner, D. A., and Benedek, G. B. (1997). Kinetic theory of fibrillogenesis of amyloid beta-protein. Proc Natl Acad Sci U S A *94*, 7942-7947.

Lundmark, K., Westermark, G. T., Olsen, A., and Westermark, P. (2005). Protein fibrils in nature can enhance amyloid protein A amyloidosis in mice: Cross-seeding as a disease mechanism. Proc Natl Acad Sci U S A *102*, 6098-6102.

Nelson, R., and Eisenberg, D. (2006). Recent atomic models of amyloid fibril structure. Curr Opin Struct Biol *16*, 260-265.

Nelson, R., Sawaya, M. R., Balbirnie, M., Madsen, A. O., Riekel, C., Grothe, R., and Eisenberg, D. (2005). Structure of the cross-beta spine of amyloid-like fibrils. Nature *435*, 773-778.

Olsen, A., Herwald, H., Wikstrom, M., Persson, K., Mattsson, E., and Bjorck, L. (2002). Identification of two protein-binding and functional regions of curli, a surface organelle and virulence determinant of *Escherichia coli*. J Biol Chem *277*, 34568-34572.

Prusiner, S. B. (1998). The prion diseases. Brain Pathol *8*, 499-513.

Quist, A., Doudevski, I., Lin, H., Azimova, R., Ng, D., Frangione, B., Kagan, B., Ghiso, J., and Lal, R. (2005). Amyloid ion channels: a common structural link for protein-misfolding disease. Proc Natl Acad Sci U S A *102*, 10427-10432.

Rhoades, E., and Gafni, A. (2003). Micelle formation by a fragment of human islet

amyloid polypeptide. Biophys J *84*, 3480-3487.

Ross, E. D., Minton, A., and Wickner, R. B. (2005). Prion domains: sequences, structures and interactions. Nat Cell Biol *7*, 1039-1044.

Shorter, J., and Lindquist, S. (2005). Prions as adaptive conduits of memory and inheritance. Nat Rev Genet *6*, 435-450.

Shorter, J., and Lindquist, S. (2004). Hsp104 catalyzes formation and elimination of self-replicating Sup35 prion conformers. Science *304*, 1793-1797.

True, H. L., and Lindquist, S. L. (2000). A yeast prion provides a mechanism for genetic variation and phenotypic diversity. Nature *407*, 477-483.

Tukel, C., Raffatellu, M., Humphries, A. D., Wilson, R. P., Andrews-Polymenis, H. L., Gull, T., Figueiredo, J. F., Wong, M. H., Michelsen, K. S., Akcelik, M., Adams, L. G., and Baumler, A. J. (2005). CsgA is a pathogen-associated molecular pattern of Salmonella enterica serotype Typhimurium that is recognized by Toll-like receptor 2. Mol Microbiol *58*, 289-304.

Uversky, V. N., Gillespie, J. R., and Fink, A. L. (2000). Why are "natively unfolded" proteins unstructured under physiologic conditions? Proteins *41*, 415-427.

Wright, C. F., Teichmann, S. A., Clarke, J., and Dobson, C. M. (2005). The importance of sequence diversity in the aggregation and evolution of proteins. Nature *438*, 878-881.

Zogaj, X., Bokranz, W., Nimtz, M., and Romling, U. (2003). Production of cellulose and

curli fimbriae by members of the family Enterobacteriaceae isolated from the human gastrointestinal tract. Infect Immun *71*, 4151-4158.

Zogaj, X., Nimtz, M., Rohde, M., Bokranz, W., and Romling, U. (2001). The multicellular morphotypes of *Salmonella typhimurium* and *Escherichia coli* produce cellulose as the second component of the extracellular matrix. Mol Microbiol *39*, 1452-1463.

Chapter 3

N- and C-terminal Repeating Units Govern CsgA Nucleation and Seeding Responsiveness[3]

Abstract

Amyloid fibers are filamentous proteinaceous structures commonly associated with mammalian neurodegenerative diseases. Nucleation is the rate-limiting step of amyloid propagation and its nature remains poorly understood. *E. coli* assembles functional amyloid fibers called curli on the cell surface using an evolved biogenesis machine. *In vivo*, amyloidogenesis of the major curli subunit protein, CsgA, is dependent on the minor curli subunit protein, CsgB. In this chapter, I directly demonstrated that $CsgB^+$ cells efficiently nucleated purified soluble CsgA into amyloid fibers on the cell surface. CsgA contains five imperfect repeating units that fulfill specific roles in directing amyloid formation. Deletion analysis revealed that the N- and C-terminal most repeating units were required for *in vivo* amyloid formation. I found that CsgA nucleation specificity is encoded by the N- and C-terminal most repeating units using a blend of genetic, biochemical and electron microscopic analyses. In addition, I found that C-terminal most repeat was most aggregation-prone and dramatically contributed to CsgA polymerization *in vitro*. This work defines the elegant molecular signatures of bacterial amyloid nucleation and polymerization, thereby revealing how nature directs amyloid formation to occur at the correct time and location.

[3] A version of this chapter has been published as Wang *et al.*, 2008, J Biol Chem *283*, 21530-21539.

Introduction

Amyloid formation is most readily recognized as the underlying cause of neurodegenerative diseases and prion-based encephalopathies (Chiti and Dobson, 2006). At the root of these diseases is the uncontrolled conversion of soluble proteins into an insoluble fiber known as amyloid. Amyloid fibers are 4-10 nm wide, unbranched filaments possessing a characteristic cross-β sheet structure, and specific tinctorial properties when stained with Congo red and thioflavin T (ThT) (Chiti and Dobson, 2006). The disease-associated amyloid forming proteins have little similarity at amino acid level, although the resulting fibers are biochemically and structurally similar (Chiti and Dobson, 2006). Although the pathology of amyloid diseases is incompletely understood, there is growing evidence that soluble folding intermediates are key to cytotoxicity and disease development (Bucciantini et al., 2002; Hartley et al., 1999; Lambert et al., 1998; Lesne et al., 2006; Roher et al., 1996; Walsh et al., 2002).

An increasing number of examples suggest that amyloid fibers, or amyloidogenic intermediates, can be utilized to facilitate a particular physiological task (Fowler et al., 2007). These 'functional amyloids' have been found in many organisms, including bacteria, fungi and mammals (Fowler et al., 2007). The functional amyloids provide a unique perspective on amyloidogenesis, since it is assumed that the cell has evolved mechanisms to control and propagate functional amyloid formation so that the cytotoxicity normally associated with amyloid formation is minimized. A compelling example of a functional amyloid is curli, a bacterially produced extracellular fiber that is required for biofilm formation and other community behaviors (Barnhart and Chapman, 2006; Gerstel and Romling, 2003). Curli fibers exhibit the biochemical and structural

properties of amyloids (Chapman et al., 2002). *E. coli* possesses at least six proteins, encoded by the *csgBA* and *csgDEFG* (*csg*, curli specific genes) operons, that are dedicated to curli biogenesis. This highly regulated assembly machine ensures that the curli subunits interact at the correct time and location.

In *E. coli*, the polymerization of the major curli fiber subunit protein CsgA into an amyloid fiber depends on the minor curli subunit protein, CsgB (Hammar et al., 1996). The mature CsgA protein has a predicted molecular weight of 13.1 kDa, and at least two distinct functional domains, an N-terminal outer membrane secretion domain (Robinson et al., 2006) and an amyloid core domain. The N-terminal domain is 22 amino acids long (residues 21 to 42 as shown in Figure 3.1A, as the first 20 amino acids are cleaved during translocation by the Sec inner membrane secretion complex) (Collinson et al., 1999). The amyloid core domain (residues 43 to 151) contains five imperfect repeating units, each 19-23 amino acids (Figure 3.1A). The five repeating units form a protease-resistant structure (Collinson et al., 1999) that is proposed to be the amyloid core of CsgA (Wang et al., 2007). These repeats are distinguished by the consensus sequence Ser-X5-Gln-X4-Asn-X5-Gln linked by 4 or 5 residues, which is found in all known CsgA homologs from *Enterobacteriaceae* (Wang et al., 2007). Each repeat is predicted to form a β strand-loop-strand motif and five repeated motifs compose cross-β structure (Collinson et al., 1999). *In vitro*, chemically synthesized oligopeptide repeating unit 1 (R1), R3 and R5 can efficiently assemble into amyloid-like fibers, suggesting that the individual repeats are amyloidogenic as presented in Chapter 2 (Wang et al., 2007). It is plausible that these five repeating units are critical for CsgA polymerization into an amyloid fiber *in vivo*. But how these repeat sequences direct CsgA polymerization remain

elusive.

Most amyloids have the ability to promote their own polymerization in a process called seeding where preformed fibers provide templates for fiber elongation. This process potentially drives prion infectivity and the development of noninfectious amyloid diseases (Harper and Lansbury, 1997; Prusiner et al., 1998; Shorter and Lindquist, 2005). Likewise, CsgA polymerization is seeded by preformed CsgA amyloid fibers (Wang et al., 2007). *In vivo* the nucleator protein CsgB initiates the polymerization of CsgA (Hammar et al., 1996). Because CsgB presents an amyloid-like template to CsgA, it was proposed that fiber-mediated self-seeding and CsgB-mediated heteronucleation are mechanistically similar processes (Hammer et al., 2007). As seeding/nucleation underlies the limiting step of amyloid propagation, understanding the nature of this mechanism will shed light on how to control amyloid formation. Here, I use a powerful blend of genetics, biochemistry and electron microscopy analysis to elucidate the sequence determinants and specificity of curli nucleation and polymerization. Although five repeats of CsgA share high sequence similarity, I found that only N- and C- terminal repeats are indispensable and responsive to CsgB-mediated heteronucleation and CsgA seeding. This work provides unique insights into the *in vivo* nucleated amyloid polymerization.

Experimental Procedures

Bacterial Growth

To induce curli production, bacteria were grown on YESCA plates (1 g yeast extract, 10 g Casamino Acids and 20 g agar per liter) at 26°C for 48 hours (Chapman et al., 2002). When needed, antibiotics were added to plates at the following concentrations:

kanamycin 50 μg/ml, chloramphenicol 25 μg/ml, or ampicillin 100 μg/ml.

Strains and Plasmids

Strains and plasmids used in this study are listed in Table 3.1. Primer sequences used in this study are listed in Table 3.2. Plasmids containing repeat deletions were constructed by site-specific mutagenesis using overlapping PCR extension. The PCR products contained the relevant mutations and NcoI/BamHI sites at the 5'/3' ends. PCR products were cloned downstream of the *csgBA* promoter into NcoI/BamHI sites of control vector pLR2 (Robinson et al., 2006). The *csgA* strain LSR10 (MC4100 Δ*csgA*) and expression vector pMC3 were generated previously (Chapman et al., 2002). To express and purify CsgA mutant proteins, PCR amplified mutant sequences including sequence encoding 6 histidine residues at C-terminus were subcloned into pMC3 to replace sequence encoding CsgA-his. The truncated CsgA proteins and the plasmids encoding them were named to reflect the particular deletion. For example, CsgA lacking R1 was named ΔR1, and the plasmid encoding ΔR1 was named pΔR1.

Western Blot Analysis

Whole-cell and plug western blot analysis of CsgA were described previously (Chapman et al., 2002). Briefly, for whole-cell western analysis, bacteria were grown on YESCA plates at 26°C for 48 hours. Cells were scraped off the plates and normalized by optical density at 600 nm. Cell pellets (including cell associated protein aggregates) were resuspended in 2X SDS loading buffer either with or without prior formic acid (FA) treatment. Wild-type curli fibers require brief FA treatment prior to SDS PAGE in order to depolymerize CsgA monomers from the fiber (Collinson et al., 1991). For plug western analysis, cultures were normalized by optical density at 600 nm and spotted on the

YESCA plate. After growth at 26°C for 48 hours, 8 mm circular plugs that included cells and the underlying agar (plugs) were collected and resuspended in 2X SDS loading buffer either with or without prior FA treatment. Samples were electrophoresed on a 15% sodium dodecyl sulfate (SDS)-polyacrylamide gel and blotted onto polyvinylidene difluoride membrane using standard techniques. Western blots were probed by anti-CsgA polyclonal antibody that was raised in rabbits against purified CsgA (Proteintech, Chicago, IL) and was used at a dilution of 1:10,000. The secondary antibody was anti-rabbit antibodies conjugated to horseradish peroxidase (Sigma, St. Louis, MO) and was used at a dilution of 1:7,000. The blots were developed using the Pierce super signal detection system as previously described (Hammer et al., 2007).

Transmission Electron Microscopy

A Philips CM10 Transmission Electron Microscope was used to visualize the curliated cells and protein fiber aggregates. Samples (10 μl) were placed on Formvar-coated copper grids (Ernest F. Fullam, Inc., Latham, NY) for 2 min, washed with deionized water, and negatively stained with 2% uranyl acetate for 90 sec.

Purification of Mutant Proteins

C-terminal hexahistidine tagged CsgA or CsgA mutant proteins were overexpressed along with CsgG by induction with 0.25 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) in LSR12 (C600 Δ*csgBAC* and Δ*csgDEFG*). In this strain, CsgA or CsgA mutant proteins were secreted to the medium and purified as previously described in Chapter 2 (Wang et al., 2007). Cell-free medium containing CsgA or CsgA mutant proteins was passed through a HIS-select^{TM} HF nickel-nitrilotriacetic acid column (Sigma, St. Louis, MO). His-tagged CsgA or CsgA mutant proteins were

eluted with 100 mM imidazole in 10 mM potassium phosphate (pH 7.2). Sephadex G25 (balanced in 50 mM potassium phosphate, pH7.2) was used to remove the imidazole in the CsgA-containing fractions. Immediately after removal of imidazole, freshly purified proteins in 50 mM potassium phosphate (pH7.2) were soluble and unstructured as reported previously (Wang et al., 2007). CsgA and CsgA mutant proteins were purified to homogeneity by this approach as evidenced by single monomer band on SDS-PAGE stained by Coomassie blue. Plasmids encoding CsgA-His (Chapman et al., 2002) or ΔR2-His can complement a *csgA* mutant *in vivo* (data not shown), suggesting that His tag does not disrupt amyloidogenecity of CsgA and its mutants.

***In vitro* polymerization assay**

For ThT assays, freshly purified CsgA or CsgA mutant proteins were passed through a 0.02-μm Anotop 10 filter (Whatman, Maidstone, UK) before being loaded on 96-well opaque plate. ThT was added to a concentration of 20 μM. Fluorescence was measured every 10 min after shaking 5 sec by a Spectramax M2 plate reader (Molecular Devices, Sunnyvale, CA) set to 438 nm excitation and 495 nm emission with a 475 nm cutoff. ThT fluorescence was normalized by $(F_i-F_0)/(F_{max}-F_0)$. F_i was the ThT intensity (fluorescence arbitrary unit) of samples and F_0 was the ThT background intensity. For the measurements of polymerization at different concentrations, F_{max} was the maximum ThT intensity of samples at the highest concentration. For seeding reactions, F_{max} was the maximum ThT intensity of the reaction in the absence of seeds. CsgA polymerization can be described as a triphasic process with a lag, growth and stationary phase (Wang et al., 2007). We used the time intervals of lag phase (T_0) and fiber growth phase (T_c) to describe the efficiency of polymerization. T_0 and T_c were previously used to compare the

polymerization of yeast prion Sup35p and its mutants (DePace et al., 1998). Here, the T_0 was obtained by measuring the time interval between the starting point (0 hour) and the intersection point of the extrapolation line from the lag phase and extrapolation line of the major portion of the growth phase. The length of the growth phase (T_c) was calculated by subtracting T_0 from the time at the beginning of the stationary phase, which includes the major portion of growth phase.

Peptide preparation

Peptides were chemically synthesized by Proteintech Group Inc., Chicago, IL. The peptides were dissolved in hexafluoroisopropanol (HFIP)/trifluoroacetic acid (TFA) (1:1 v/v) as previously described in Chapter 2 (Wang et al., 2007). Alternatively, peptides were dissolved in 8.0 M guanidine hydrochloride (GdnHCl) buffered by 50 mM potassium phosphate buffer at pH 7.2 to fully denature the sample. After at least 1 hour incubation, GdnHCl was quickly removed by a Sephadex G10 column that was balanced in 50 mM potassium phosphate, pH7.2 at room temperature. Peptides in potassium phosphate buffer (50 mM, pH7.2) were passed through a 0.02-μm Anotop 10 filter (Whatman, Maidstone, UK) and their polymerization was measured using ThT fluorescence as described for CsgA and mutant proteins. Similar results were obtained using either denaturing protocol and the HFIP/TFA denaturing protocol was used, unless indicated otherwise.

***In vitro* seeding**

Mature fibers (2-week-old fibers) were sonicated using a Fisher Model 100 sonic dismembrator (Fisher, Pittsburgh, PA) for three 15-sec bursts on ice to produce seeds. Seeds were added to freshly purified or prepared samples immediately before the start of

ThT florescence assay.

Overlay assay

10 μl freshly prepared proteins or peptides were dripped on a lawn of the $CsgB^+$ (*csgA* mutant strain LSR10) and $CsgB^-$ (*csgAB* mutant strain LSR13) cells which were grown on YESCA plates at 26°C for 48 hours. Samples were incubated on cells for 10 min at room temperature unless otherwise specified, stained with 0.5 mg/ml Congo red solution (50 mM potassium phosphate, pH 7.2) for 5 min, and then washed with potassium phosphate buffer. Cells incubated with protein or peptide solution were collected for TEM analysis.

Results

R1 and R5 are critical for CsgA polymerization *in vivo*

The amyloid core domain of CsgA (residues 43 to 151) contains five imperfect repeating units (Figure 3.1A). Using a series of *csgA* alleles that contained precise in-frame deletions, the role of each repeating unit was determined. The ability of *csgA* alleles to complement a non-amyloidogenic *csgA* mutant strain was assessed by western blotting and transmission electron microscopy (TEM). A *csgA*$^-$/pCsgA strain produced curli fibers that were microscopically similar to those assembled by wild-type strain MC4100 as shown by TEM (Figure 3.2). Wild type curli are SDS insoluble and require brief treatment with formic acid (FA) to liberate CsgA monomers from the fiber (Collinson et al., 1991). Similarly, the CsgA produced by *csgA*$^-$/pCsgA was SDS insoluble and did not migrate into an SDS PAGE gel without FA pretreatment (Figure 3.1B, lanes 1 and 2). A *csgA* mutant transformed with pΔR2 produced curli fibers that

were similar to those produced by *csgA*⁻/pCsgA, suggesting that R2 is dispensable for curli formation (Figure 3.1B, lanes 7-8, Figure 3.1D and Figure 3.2). Fibers assembled by *csgA*⁻/pΔR3 and *csgA*⁻/pΔR4 appeared shorter than wild-type fibers (Figure 3.1D and Figure 3.2). Also, ΔR3 and ΔR4 fibers were less SDS resistant and migrated into an SDS-PAGE gel without FA treatment (Figure 3.1B, lanes 9, 10, 11 and 12). Deletion of R1 or R5 resulted in the most defective *csgA* alleles. Nearly no fibers were produced by *csgA*⁻/pΔR1 or *csgA*⁻/pΔR5 when analyzed by TEM (Figure 3.1D). Furthermore, ΔR1 and ΔR5 proteins were nearly undetectable in whole cell lysates scraped off YESCA plates (Figure 3.1B, lanes 5 and 6; lanes 13 and 14). To test the possibility that ΔR1 and ΔR5 were secreted away from the cell as soluble proteins, cells and the underlying agar were collected and analyzed by western blotting. Both ΔR1 and ΔR5 were readily detected in samples that included the underlying agar, demonstrating that, like wild-type CsgA, ΔR1 and ΔR5 were stable and secreted to the cell surface (Figure 3.1C, lanes 5, 6, 7 and 8). However, unlike CsgA, ΔR1 and ΔR5 were not assembled into an SDS-resistant fiber after secretion (Figure 3.1B and 3.1D). I also asked if ΔR1 and ΔR5 could form fiber aggregates if cells were incubated for 100 hours instead of our standard incubation time of 48 hours. A small number of fibril aggregates were observed in *csgA*⁻/pΔR1 samples incubated for 100 hours, although the fibers were rarely detected and they did not associate with cells (Figure 3.3). In contrast to ΔR1, ΔR5 did not form any detectable fiber aggregates, even after 100 hours of incubation (Figure 3.3).

We noted that ΔR1 and ΔR3 migrated more slowly than other mutant proteins on SDS-PAGE even though they are all of similar predicted molecular weights. Purified homogeneous proteins ΔR1 and ΔR3 also migrated more slowly than other repeat

deletion mutant proteins, suggesting the different migration rates resulted from the intrinsic properties of mutant proteins. The nature of these aberrant mobilities was not investigated further in this study.

R5 is critical for *in vitro* self-polymerization of CsgA

At least two possibilities existed as to why ΔR1 and ΔR5 did not assemble into amyloid fibers *in vivo*. The first possibility is that ΔR1 and/or ΔR5 were missing CsgB-responsive domains that prevented them from participating in nucleation. Alternatively, ΔR1 and ΔR5 might inefficiently participate in intermolecular interactions with themselves. To test the ability of mutant proteins to self-assemble into an amyloid fiber, I utilized an *in vitro* polymerization assay that is CsgB-independent as presented in Chapter 2 (Wang et al., 2007). ΔR1, ΔR2, ΔR3, ΔR4 and ΔR5 were purified and their polymerization was compared to wild-type CsgA using a ThT assay described in our previous work (Wang et al., 2007). *In vitro* wild-type CsgA polymerization is characterized by three distinct phases: lag, growth and stationary as shown in Chapter 2 (Wang et al., 2007). When comparing the polymerization of CsgA mutant proteins I utilized two simple parameters that were previously used to describe the polymerization of yeast prion Sup35p (DePace et al., 1998). The first kinetic parameter was the time period preceding rapid fiber growth, called lag phase, or T_0. The second parameter was the time period encompassing fiber growth phase (T_c) (DePace et al., 1998). Interestingly, when the samples were incubated at room temperature with intermittent shaking, T_0 and T_c for ΔR1, ΔR2, ΔR3 and ΔR4 were similar to those of wild-type CsgA (Figure 3.4A and 4B). However, both T_0 and T_c for ΔR5 were dramatically increased (Figure 3.4A and 4B). TEM analysis confirmed that the ThT positive aggregates formed by CsgA mutant

proteins were ordered fibers (Figure 3.5). Furthermore, the T_0 values of CsgA, ΔR1, ΔR2, ΔR3 and ΔR4 were concentration independent above 10 μM, while the T_0 value for ΔR5 was inversely correlated to concentration up to 200 μM (Figure 3.4A and data not shown). When the proteins were incubated without intermittent shaking, ΔR5 polymerization was even more defective relative to wild-type CsgA. At a concentration of 50 μM T_0 and T_c values of ΔR5 polymerization were two orders of magnitude greater than those of CsgA (Figure 3.6). ΔR5 incubated at 10 μM without shaking did not assemble into fibers within 800 hours (Figure 3.6). These results suggest that R5 plays a critical role in CsgA self-polymerization. Peptides corresponding to each repeating unit were synthesized (Wang et al., 2007) and their polymerization kinetics was determined by ThT fluorescence (Figure 3.4C and 3.4D). The strong denaturing reagents, 8.0 M GdnHCl or HFIP/TFA (1:1 v/v), were employed to remove possible seeds that would unpredictably influence polymerization kinetics. Chemically synthesized R1, R3 and R5 peptides efficiently assembled into ThT-positive fibers *in vitro* as shown in Chapter 2 (Wang et al., 2007). Like CsgA, R1 and R3 peptides polymerized into ThT-positive fibers in a triphasic fashion that included a discernable lag, growth and stationary phase (Figure 3.4C). At a concentration of 0.6 mg/ml (~240 μM), T_0 values of R1 and R3 were approximately 7.5 and 15 hours, respectively (Figure 3.4C). Unlike R1 and R3, the R5 polymerized with no apparent lag phase, becoming ThT-positive immediately after removal of the denaturant (Figure 3.4C). Even when R5 was incubated at relatively low concentrations (15 μM), there was no detectable lag phase (Figure 3.4D). The T_c value of R5 was also shorter than that of R1 and R3 (Figure 3.4C). Under the same conditions, R2 and R4 were never observed to polymerize as measured by ThT fluorescence and TEM (Figure 3.4C and

data not shown). Collectively, these results demonstrate that R5 is the most aggregation-prone of the repeating unit peptides.

CsgA efficiently responds to CsgB nucleation to assemble into amyloid fibers

Although CsgA can self-assemble into amyloid fibers *in vitro*, CsgA polymerization *in vivo* is dependent on the CsgB nucleator. Direct evidence that soluble CsgA can be converted to an amyloid fiber by wild-type CsgB is missing. To test whether the cell surface-exposed CsgB can nucleate the polymerization of purified CsgA, I developed an overlay assay using $CsgB^+$ or $CsgB^-$ cells and freshly purified soluble CsgA protein (Figure 3.7A). Different concentrations of CsgA were overlaid on $CsgB^+$ (*csgA*) or $CsgB^-$ (*csgAB*) cells and CsgA polymerization was detected using the amyloid specific dye Congo red and TEM (Figure 3.7B-E). CsgA overlaid on $CsgB^+$ cells efficiently assembled into amyloid-like fibers as shown by Congo red staining and TEM (Figure 3.7B and 3.7D), while CsgA overlaid on $CsgB^-$ cells did not induce Congo red binding and did not assemble into fibers shown by TEM (Figure 3.7C and 3.7D). CsgA fibers formed on $CsgB^+$ cells in the overlay assay were similar to wild-type curli fibers formed by *csgA*$^-$/pCsgA (Compare Figure 3.7B and 3.1D). The CsgB-mediated polymerization of purified CsgA occurred in as little as 6 minutes (1 min incubation and 5 min Congo red staining) (Figure 3.7E).

Mapping the CsgB-responsive sequences in CsgA

The CsgB-responsive repeating units in CsgA were mapped using the overlay assay and peptides corresponding to each repeating unit. 2.0 mg/ml of freshly prepared R1, R2, R3, R4 or R5 peptides was overlaid onto $CsgB^+$ and $CsgB^-$ cells. When either R1 or R5 was overlaid on $CsgB^+$ cells, amyloid-like fibers were detected by Congo red

binding and TEM (Figure 3.8A and 3.8B). When either R1 or R5 was overlaid on $CsgB^-$ cells for the same incubation time, no Congo red binding was observed and no fibers were detected by TEM (Figure 3.8A and data not shown). The R1 and R5 fibers formed on CsgB-presenting cells were shorter and thinner than fibers assembled by wild-type CsgA polymerization on the surface of the cell (Compare Figure 3.8B, 3.7B, 3.1D and 3.2). When 2.0 or 4.0 mg/ml R2, R3 and R4 were overlaid on $CsgB^+$ or $CsgB^-$ cells, no fibers were formed as evidenced by the lack of Congo red binding and by TEM analysis. This suggested that R2, R3 and R4 were unable to interact with CsgB to promote fiber assembly (Figure 3.8A and data not shown). Therefore, we concluded that R1 and R5 were the CsgB heteronucleation-responsive domains of CsgA.

To further test the notion that R1 and R5 were the only CsgB-responsive domains of CsgA, a *csgA* allele without both R1 and R5 (ΔR1&5) was engineered and its CsgB responsiveness was measured. ΔR1&5 was unable to complement a *csgA* mutant for curli biogenesis (data not shown). Unlike CsgA, ΔR1&5 did not respond to CsgB nucleation on $CsgB^+$ cells measured with the overlay assay even after extended incubation (Figure 3.8C). *In vitro*, ΔR1&5 polymerized into ThT-positive fibers, with T_0 and T_C values comparable to those of ΔR5 (compare Figure 3.4A and 3.8D). Fiber aggregates formed by ΔR1&5 *in vitro* were short and protofibril-like, similar to fibers assembled by peptide R3 by TEM (data not shown). A nucleation-competent CsgB truncation mutant ($CsgB_{trunc}$) was purified as previously described (Hammer et al., 2007). $CsgB_{trunc}$ has been demonstrated to seed wild-type CsgA (Hammer et al., 2007). Unlike CsgA, the polymerization of ΔR1&5 was not stimulated by $CsgB_{trunc}$ *in vitro* (Figure 3.8D). Collectively, these results demonstrate that a *csgA* allele missing R1 and R5 is not

responsive to CsgB-mediated heteronucleation *in vivo* and *in vitro*.

R1 and R5 are CsgA seeding responsive domains

CsgA polymerization is a self-propagating process in which CsgA seeds can completely eliminate the lag phase as presented in Chapter 2 (Wang et al., 2007). However, the determinants of CsgA seeding specificity are poorly understood. Our results demonstrate R1 and R5 are CsgB-responsive domains. To test whether R1 and R5 also play an important role in growth of CsgA fibers, I tested CsgA seeding on ΔR1, ΔR5 and ΔR1&5 *in vitro*. CsgA seeds eliminated the lag phase of ΔR1 and ΔR5 polymerization (Figure 3.9A and 3.9B), while ΔR1&5 polymerization was not promoted by CsgA seeding (Figure 3.9C). These findings demonstrated that both R1 and R5 are involved in the response to CsgA seeding and specificity is encoded in R1 and R5 sequences.

In support of the hypothesis that only R1 and R5 are CsgA seeding-responsive domains, each repeating unit peptide was incubated with wild-type CsgA fiber seeds. The lag phase of R1 polymerization was completely eliminated when 2% preformed CsgA fibers were added to the reaction (Figure 3.9D). In order to test the responsiveness of R5 to CsgA seeding, I needed to slow the aggregation of R5 so that a discernable lag phase could be observed. In the presence of 1.0 M GdnHCl, 0.2 mg/ml R5 remained unpolymerized for over 40 hours (Figure 3.9E). CsgA seeds eliminated the lag phase of R5 even under these mildly denaturing conditions (Figure 3.9E). However, CsgA seeds did not accelerate R3 polymerization (Figure 3.9F). CsgA seeds were also unable to stimulate the polymerization of R2 and R4 peptides, as both R2 and R4 remained unpolymerized for 48 hours detected by ThT assay and TEM (data not shown).

Discussion

Amyloid formation has been associated with neurodegenerative diseases for over a century, yet very little is known about the mechanism of nucleated polymerization *in vivo*. Curli provide a sophisticated genetic and molecular tool set with which to explore amyloid nucleation and polymerization. I have described the molecular determinants of a highly evolved and elegant amyloid assembly pathway. CsgB-mediated heteronucleation is requisite for *in vivo* curli fiber assembly. I extensively investigated the CsgA sequence determinants for nucleation and polymerization. I found that the N- and C- terminal repeats (R1 and R5) are indispensable for curli biogenesis *in vivo*. R5 is the most aggregation-prone repeating unit and it contributes dramatically to CsgA polymerization *in vitro*. R1 and R5 play critical roles in CsgA seeding and CsgB heteronucleation response.

Molecular details of curli assembly

Curli biogenesis occurs via a nucleation-precipitation mechanism, where surface-exposed CsgB nucleates secreted CsgA into an insoluble fiber (Hammar et al., 1996; Hammer et al., 2007). I found that R1 and R5 are indispensable for *in vivo* curli assembly (Figure 3.1). I have provided direct evidence that surface-exposed CsgB promotes the transition of CsgA from a soluble protein to an insoluble amyloid-like fiber (Figure 3.7). Both R1 and R5 of CsgA were demonstrated to respond to CsgB heteronucleation (Figure 3.8). Similarly, preformed CsgA fibers (in the absence of CsgB) were able to recruit soluble CsgA into the fiber in a process called seeding (Wang et al., 2007). Seeding was also dependent on R1 and R5 (Figure 3.9). Based on these results I propose a promiscuous nucleation model to describe the molecular details of curli

assembly (Figure 3.10). CsgA is not amyloidogenic within or outside the cell until it encounters outer membrane-localized CsgB. This temporal and spatial control might prevent the proteotoxicity of amyloid formation. After CsgA is secreted from cells, either R1 or R5 of CsgA interacts with CsgB or the growing fiber tip, facilitating CsgA polymerization (Figure 3.10). Collectively, the curli fiber assembly machinery on the cell surface is composed of multiple nucleation templates (CsgB and CsgA fiber tip) and multiple responsive domains of fiber subunit CsgA (R1 and R5), which makes nucleation polymerization an efficient process *in vivo*.

Microbes assemble many proteinaceous fibers on their cell surface. One challenge to assembling extracellular fibers is prevention of subunit diffusion to the extracellular milieu. Therefore, the assembly of curli, pili and flagella must be regulated so that efficient subunit incorporation into the fiber occurs. P and type-I pilus assembly is dependent on the chaperone/usher pathway, where periplasmic chaperone proteins deliver the pilus subunits to the outer membrane-localized usher protein. On the periplasmic side of the usher, pilus subunits are added to the base of the growing pilus. Because pilus subunits never leave the periplasmic space before incorporation into the pilus, there is no chance for diffusion of unassembled subunits to the extracellular space (Sauer et al., 2004). Flagella are another common extracellular fiber assembled by Gram-negative bacteria. Soluble flagellin is secreted through the growing fiber until it is bound by the cap protein and polymerized onto the distal end (Yonekura et al., 2002). In contrast to flagella and pili assembly, the major curli subunit CsgA is secreted from the cell as a soluble protein, which subsequently polymerizes into a fiber. To ensure that secreted subunits are efficiently incorporated into an amyloid fiber, curli assembly is dependent on

the CsgB nucleator and multiple CsgA responsive domains. Nucleation, the stalwart of curli formation, is also important for disease-associated amyloid formation. Understanding the molecular signatures of nucleation will lead to new insights into disease-associated protein aggregation.

Distinction between *in vivo* and *in vitro* bacterial amyloid formation

Even though CsgA is amyloidogenic *in vitro*, it remains soluble in the extracellular milieu in the absence of nucleator CsgB *in vivo* (Hammar et al., 1996). Like a *csgB* mutant, *csgA* alleles missing one or both CsgB-responsive domains (R1 and R5) are unable to efficiently polymerize *in vivo*. For example, ΔR1 is completely defective in curli assembly *in vivo*, even though ΔR1 polymerizes with similar kinetics as wild-type CsgA *in vitro* (Figure 3.1 and Figure 3.4). ΔR1 was mostly incapable of forming fiber aggregates *in vivo,* even if cells were incubated for over 100 hours instead of the standard 48 hours. A small amount of ΔR1 short fibril aggregates not associated with cells were observed by TEM, while ΔR5 was never observed to form fiber aggregates by TEM (Figure 3.3). These observations were consistent with the *in vitro* polymerization results: ΔR1 polymerization was much more efficient than ΔR5 polymerization (Figure 3.4A and 4B). This suggests that within natural cellular environments the amyloidogenecity of CsgA is tightly controlled by the CsgB heteronucleator. Therefore, these intermolecular interactions between CsgA and CsgB initiate bacterial amyloid formation at the correct place and time.

Nucleation specificity of amyloid formation

Nucleation/seeding is the rate-limiting step of amyloid formation (Jarrett and Lansbury, 1993). The mechanism and specificity of nucleation is poorly understood.

Nucleation specificity is arguably the best understood in the yeast prion Sup35p. One small region (amino acids 9-39) located in the N-terminal prion domain of Sup35p governs self-recognition and species-specific seeding activity (Tessier and Lindquist, 2007). The residues located in this region are critical for prion propagation (DePace et al., 1998). Here, I demonstrated that bacterial amyloid protein CsgA also contains two small regions (R1 and R5) that control nucleation response. It may be a general rule that only small portions (nucleation sites) of amyloid proteins govern the nucleation response and specificity. The residues contained in the nucleation site determine the nucleation response and specificity. Mutagenesis analyses on Sup35p (DePace et al., 1998) and CsgA support this notion. The nucleation sites of Sup35p and CsgA are located within the protease-resistant amyloid core of the proteins, and peptides containing the nucleation sites of Sup35p and CsgA are sufficient to form amyloid fibers *in vitro* (Glover et al., 1997; Wang et al., 2007).

CsgA homologues from different bacteria share high sequence similarity suggesting cross-species nucleation might occur. *Salmonella typhimurium* CsgA and CsgB homologues are 74.8% and 82.1% identical to *E.coli* CsgA and CsgB, respectively. *Salmonella typhimurium csgA* and *csgB* genes can complement *E.coli csgA* and *csgB* mutations in terms of fiber assembly (Romling et al., 1998). Because curli subunits are secreted into the extracellular milieu prior to incorporation into the fiber, complex bacterial communities may have members that secrete CsgA-like molecules and others that produce CsgB-like nucleators. It is plausible that nucleation promiscuity facilitates interspecies fiber formation in complex communities, allowing curli to act as a structural signal that links different species together.

Figures and Tables

Figure 3.1. R1 and R5 are critical for CsgA *in vivo* polymerization into an amyloid fiber. (A) Schematic of CsgA including an N-terminal Sec signal peptide and the N-terminal 22 residues that precede the five repeating units. (B) Whole-cell western blots of *csgA* mutant strains harboring the following plasmids: pCsgA (lanes 1 and 2), vector control (VC) (lanes 3 and 4), pΔR1 (lanes 5 and 6), pΔR2 (lanes 7 and 8), pΔR3 (lanes 9 and 10), pΔR4 (lanes 11 and 12) and pΔR5 (lanes 13 and 14). Samples were treated with (+) or without (-) FA. The blot was probed with anti-CsgA antibody. (C) Western blot of whole-cells and underlying agar (agar plugs) from *csgA* strains containing constructs pCsgA (lanes 1 and 2), vector control (lanes 3 and 4), pΔR1 (lanes 5 and 6) and pΔR5 (lanes 7 and 8) grown 48 hours at 26°C on YESCA plates. Samples were treated with (+) or without (-) FA. The blots were probed with anti-CsgA antibody. (D) Negative-stain EM micrographs of *csgA* mutant cells containing the indicated plasmids. Cells were grown on YESCA plates for 48 hrs at 26°C prior to staining with uranyl acetate. Scale bars are equal to 500 nm.

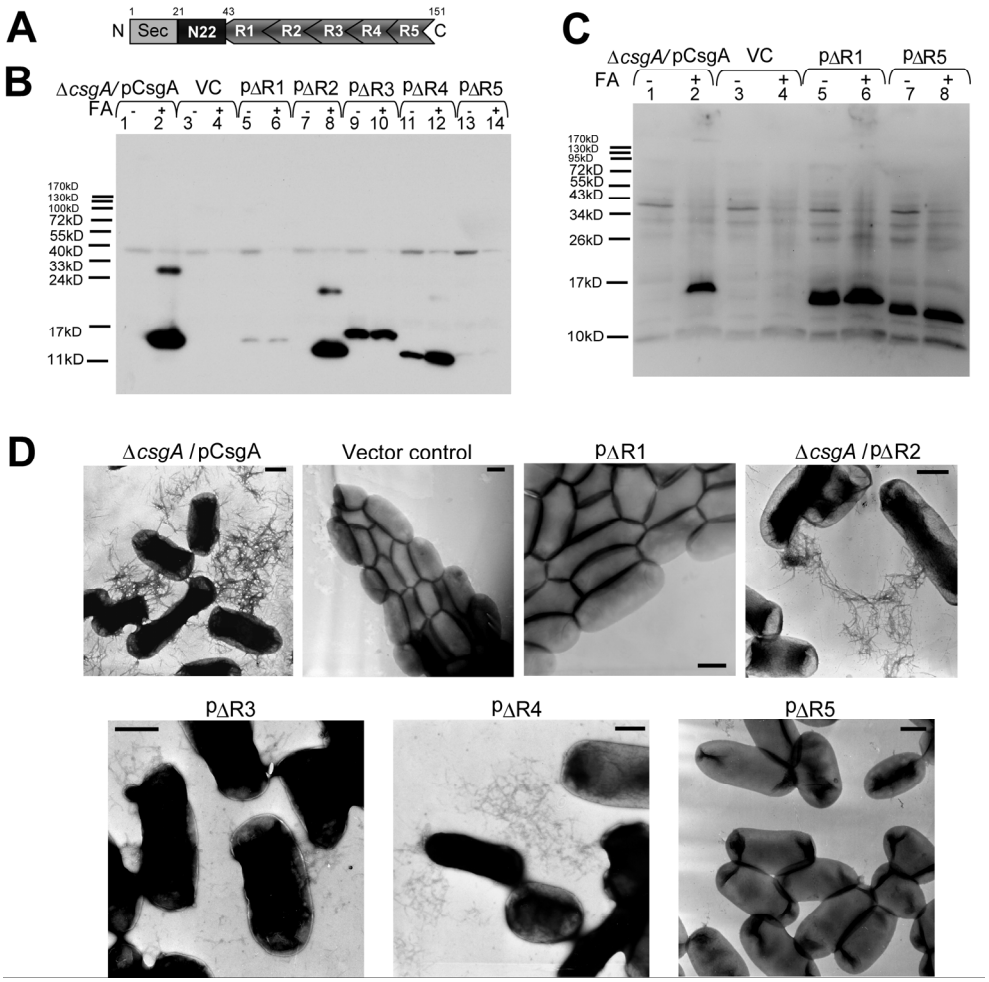


Figure 3.2. Negative-stain EM micrographs of CsgA, ΔR2, ΔR3 and ΔR4 fibers assembled *in vivo* at high magnification.
Negative-stain EM micrographs of wild-type strain MC4100 and *csgA* mutant cells containing the indicated plasmids. Cells were grown on YESCA plates for 48 hrs at 26°C prior to staining with uranyl acetate. Scale bars are equal to 200 nm.

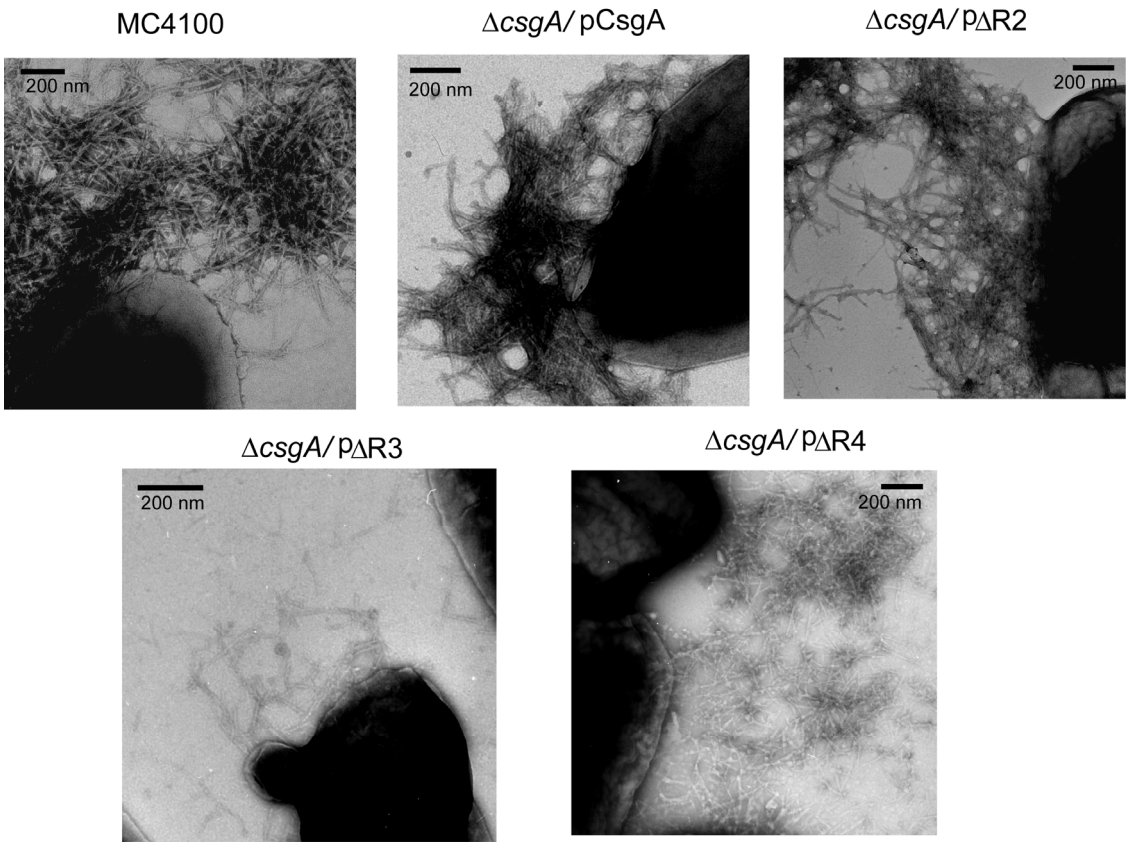


Figure 3.3. Negative-stain EM micrographs of *csgA* cells expressing CsgA, ΔR1 or ΔR5 for a long-term growth.
Negative-stain EM micrographs of *csgA* mutant cells containing the indicated plasmids. Cells were grown on YESCA plates for 100 hrs at 26°C prior to staining with uranyl acetate. Scale bars are equal to 500 nm.

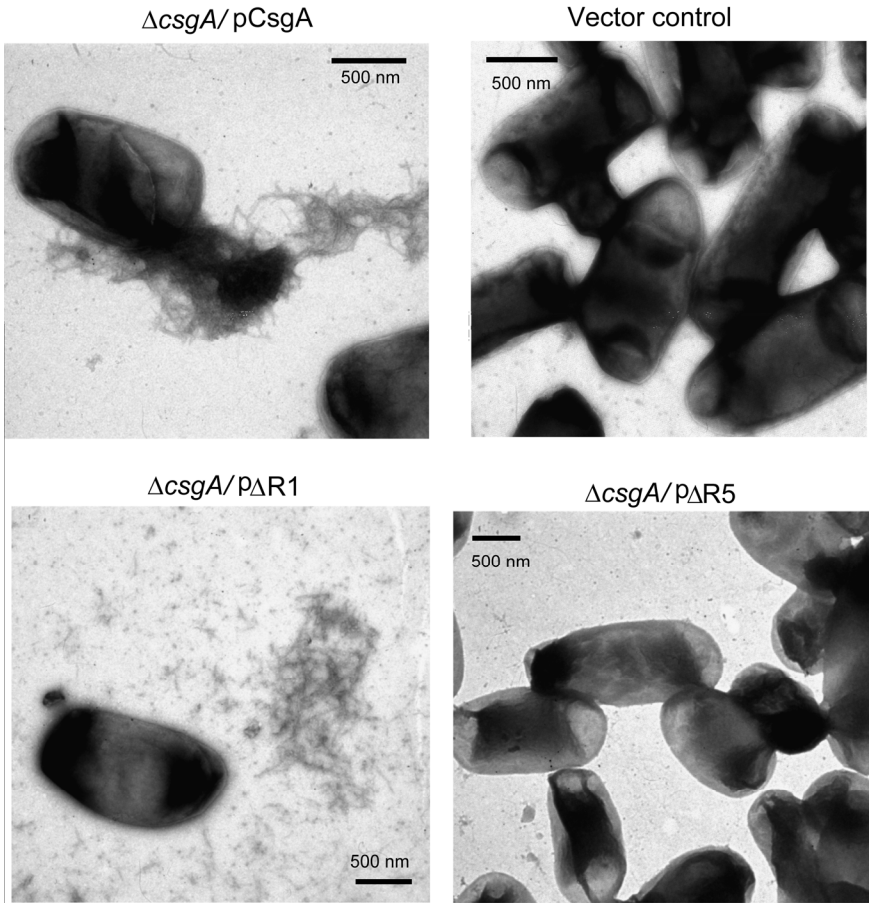


Figure 3.4. *In vitro* polymerization of CsgA, mutant CsgA proteins and peptides. (A) The normalized fluorescence intensity of freshly purified CsgA and repeat deletion analogues was measured at 495 nm after excitation at 438 nm in the presence of 20 μM ThT. Samples were shaken for 5 seconds prior to fluorescence measurements in 10 minute intervals. The protein used in the assay is indicated at the top of the graph and the protein concentration is above the relevant data points. The X-axis of each graph spans from 0 hr to 10 hrs, except for ΔR5 which spans from 0 hr to 35 hrs. (B) Lag time (T_0) of freshly purified CsgA and CsgA analogues incubated at 10 μM. Data were expressed as the mean ± SEM of three independent experiments. (C) The polymerization of 0.6 mg/ml (~240 μM) chemically synthesized peptides R1, R2, R3, R4 and R5 was measured by ThT fluorescence. (D) The polymerization of 15 μM and 30 μM R5 peptide was measured by ThT fluorescence.

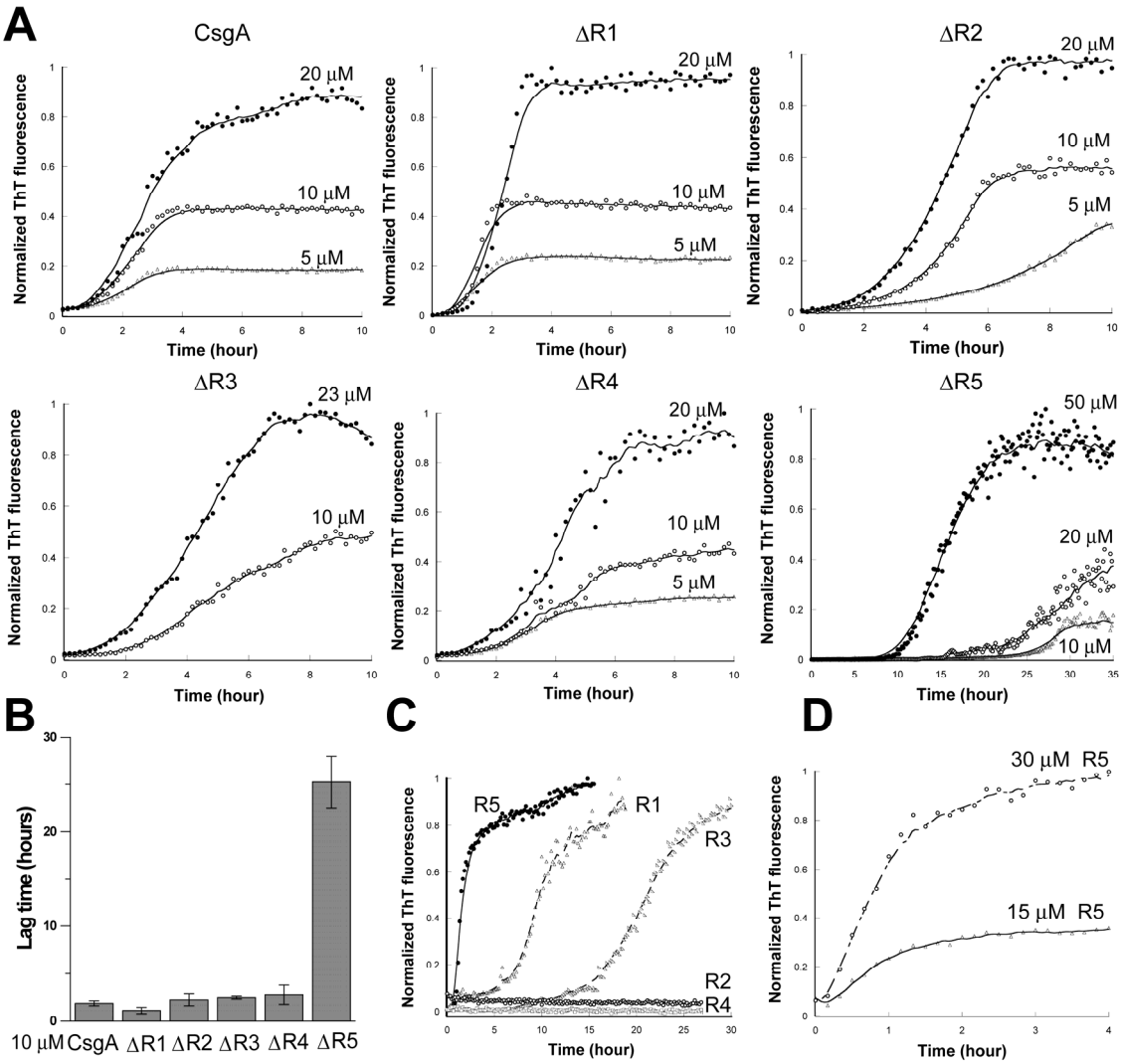


Figure 3.5. EM Micrographs of *in vitro* polymerized fibers from CsgA and mutant proteins.
(A-F) Negative-stain EM micrographs of *in vitro* polymerized fibers of CsgA (A), ΔR1(B), ΔR2 (C), ΔR3 (D), ΔR4 (E) and ΔR5 (F). Scale bars are equal to 500 nm.

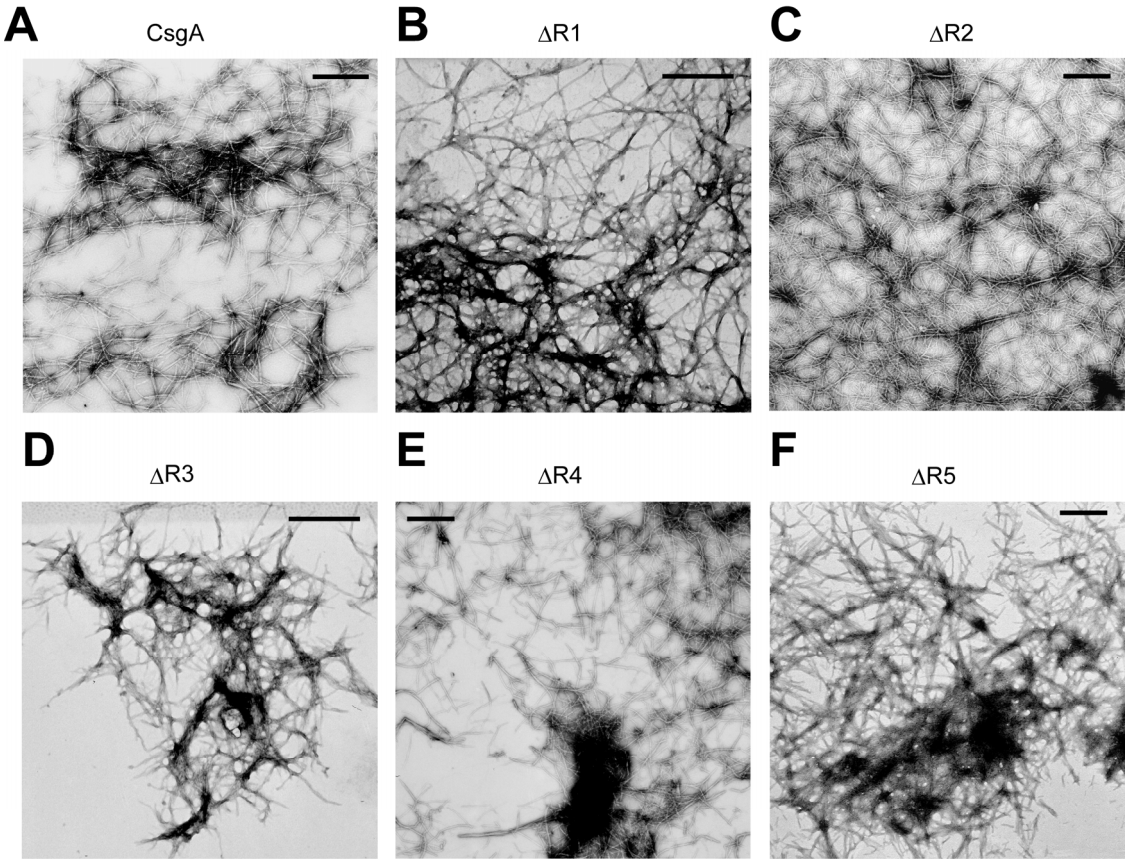


Figure 3.6. *In vitro* polymerization of CsgA mutant proteins under the quiescent condition.

13 μM CsgA (A) and ΔR5 at various concentrations (B) were incubated at room temperature without agitation in the presence of 0.02% NaN_3. At the indicated time points, samples were withdrawn, ThT was added at a concentration of 20 μM and fluorescence was measured at 495 nm after excitation at 438 nm by a Spectramax M2 plate reader (Molecular Devices, Sunnyvale, CA). ThT fluorescence was normalized as described in the Materials and Methods.

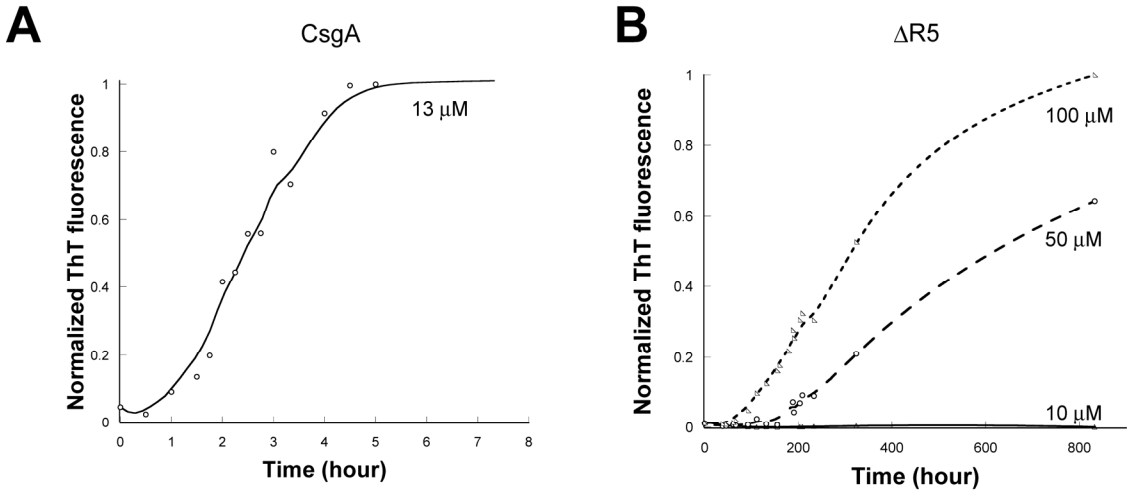


Figure 3.7. Purified CsgA is efficiently nucleated when overlaid on CsgB-expressing cells.
(A) A schematic presentation of the overlay assay in which freshly purified and soluble CsgA is dripped onto cells expressing the nucleator protein, CsgB. In the presence of CsgB, CsgA (shown as a circle) undergoes a conformational change and polymerizes into a fiber (shown as a chevron). In the absence of CsgB, CsgA remains soluble and does not assemble into an amyloid fiber. (B-C) Negative-stain EM micrographs of $CsgA^{-}B^{+}$ (B) and $CsgA^{-}B^{-}$ (C) cells grown on YESCA plates for 48 hours at 26°C that were overlaid with freshly purified 40 μM CsgA. Scale bars are equal to 500 nm. (D) Congo red staining of $CsgA^{-}B^{+}$ and $CsgA^{-}B^{-}$ cells after being overlaid with different concentrations of soluble CsgA. (E) 72 μM CsgA was overlaid on $CsgA^{-}B^{+}$ and $CsgA^{-}B^{-}$ cells and incubated for the indicated time intervals before staining for 5 minutes with Congo red solution.

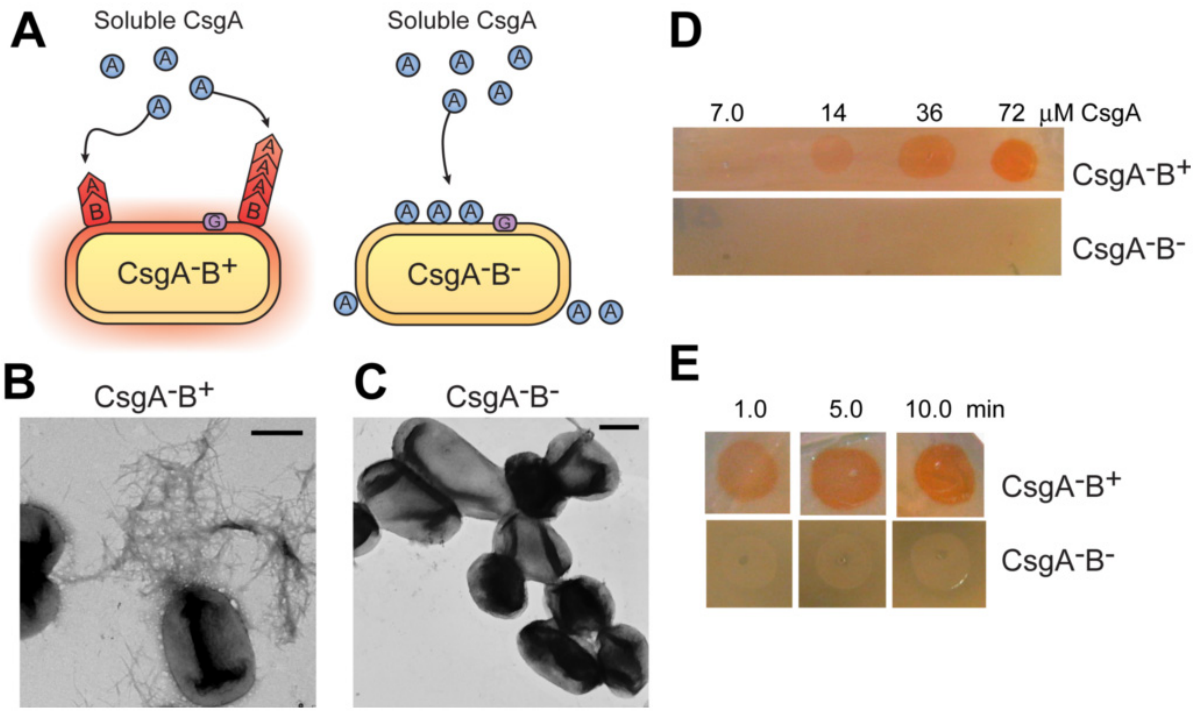


Figure 3.8. R1 and R5 are responsive to CsgB heteronucleation.
(A) 2.0 mg/ml of freshly prepared peptide solution was overlaid on CsgA$^-$B$^+$ and CsgA$^-$B$^-$ cells. After a 5-minute incubation, cells were stained with 0.5 mg/ml Congo red solution for 5 min and washed once by 50 mM potassium phosphate buffer.
(B) Negative-stain EM micrographs of CsgA$^-$B$^+$ cells overlaid with freshly prepared R1 (left) and R5 (right). Conditions are the same as in (A). Scale bars are equal to 200 nm.
(C) Freshly purified 105 μM ΔR1&5 was overlaid onto CsgA$^-$B$^+$ and CsgA$^-$B$^-$ cells. ΔR1&5 was incubated on cells for 10 hours before staining with Congo red. (D) 3% and 6% $CsgB_{trunc}$ by weight was added to freshly purified 50 μM ΔR1&5. The polymerization was measured by ThT fluorescence.

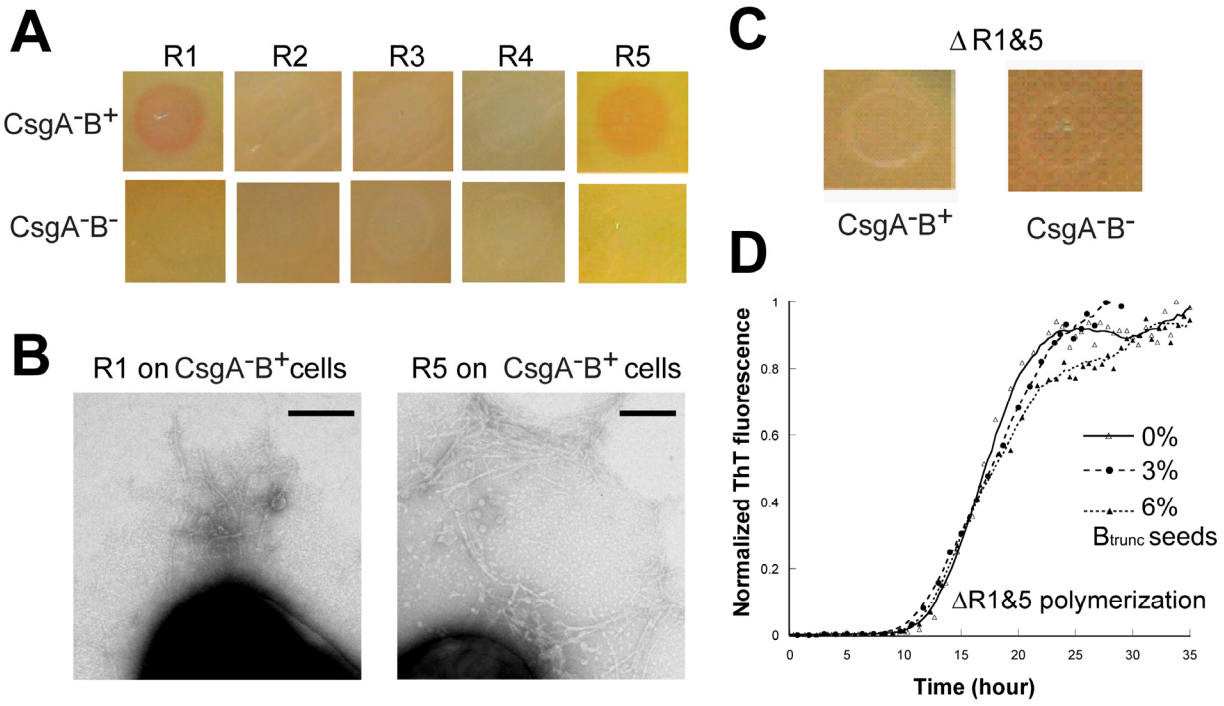


Figure 3.9. R1 and R5 are responsive to CsgA seeding.
(A)-(C) Indicated percentages of sonicated CsgA fibers were added to freshly prepared 17 μM ΔR1 (A), 50 μM ΔR5 (B) and 50 μM ΔR1&5 (C). The polymerization was measure by ThT fluorescence. (D) 2% sonicated CsgA fibers were added to freshly prepared 1.0 mg/ml R1 peptide solution. The polymerization was measured by ThT fluorescence. (E) R5 was solubilized in HFIP/TFA. Upon removal of HFIP/TFA, R5 was resuspended in 1.0 M GdnHCl buffered by 50 mM potassium phosphate buffer at pH 7.2. 4% and 8% sonicated CsgA fibers were added to 0.2 mg/ml R5 solution in the presence of 1.0 M GdnHCl, mixed with ThT at 20 μM and fluorescence was measured. (F) Indicated percentages of sonicated CsgA fibers were added to freshly prepared 1.0 mg/ml R3 peptide solution. The polymerization was measured by ThT fluorescence.

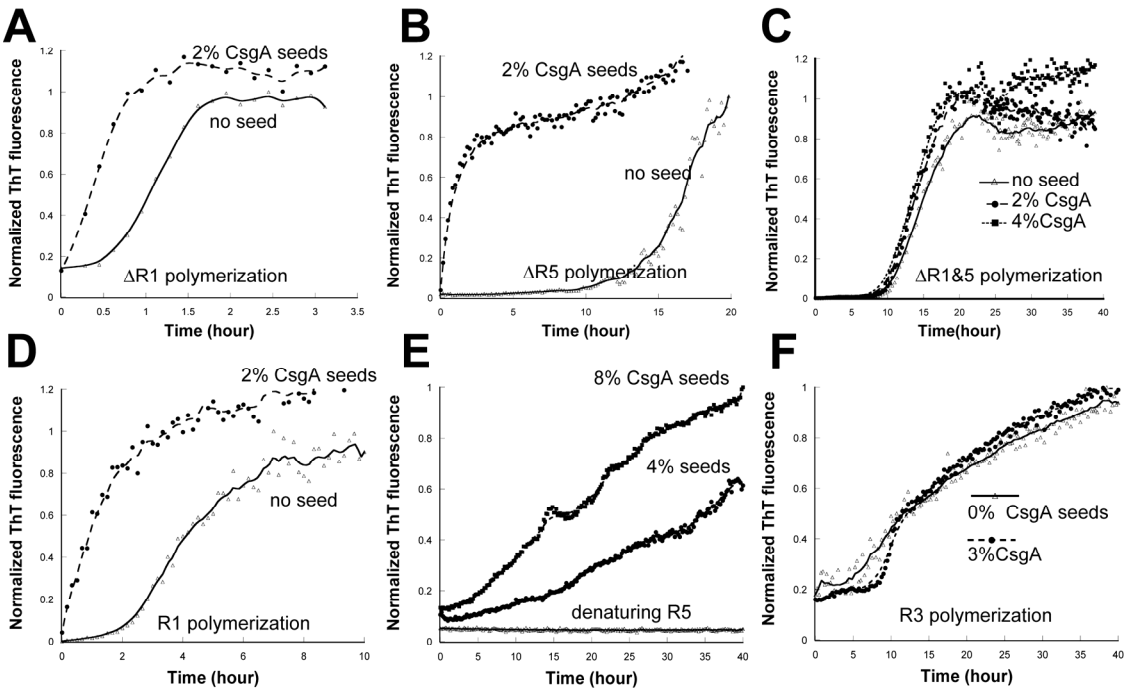


Figure 3.10. The promiscuous nucleation model of CsgA *in vivo* polymerization. Curli assembly is governed by surface-localized CsgB (heteronucleation) and by the growing fiber tip (homonucleation). After secretion from the periplasm to extracellular space, R1 and R5 of CsgA can interact with CsgB or fiber tips, initiating curli formation.

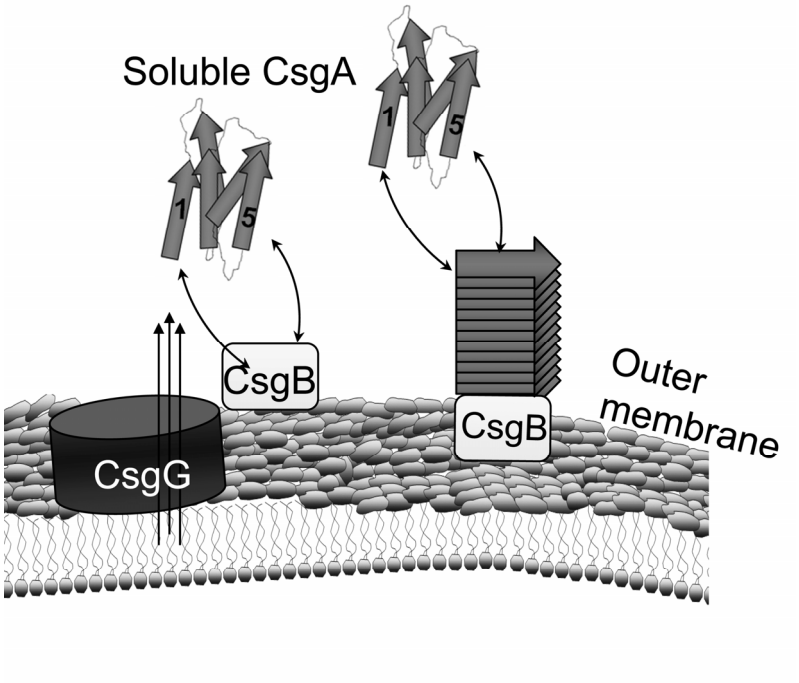


Table 3.1. Strains and plasmids used in Chapter 3[a]

Strains or Plasmids	Relevant characteristics	References	Primer used
Strains			
csgA (LSR10)	MC4100 Δ*csgA*	(Chapman et al., 2002)	
csgBA (LSR13)	MC4100 Δ*csgBA*	(Hammer et al., 2007)	
LSR12	C600 Δ*csgBA and*Δ*csgDEFG*	(Chapman et al., 2002)	
Plasmids			
pCsgA (pLR5)	*csgA* sequence in pLR2	Hultgren lab	
pLR2	Control vector containing csg*BA* promoter	(Robinson et al., 2006)	
pΔR1	*csgA* without R1(S^{43} to N^{65}) in pLR2	This Study	FpLR5, RpLR5, pΔR1 P1, P2
pΔR2	*csgA* without R2 (S^{66} to D^{87}) in pLR2	This study	FpLR5, RpLR5, pΔR2 P1, P2
pΔR3	*csgA* without R3 (S^{88} to N^{110}) in pLR2	This study	FpLR5, RpLR5, pΔR3 P1, P2
pΔR4	*csgA* without R4 (S^{111} to N^{132}) in pLR2	This study	FpLR5, RpLR5, pΔR4 P1, P2
pΔR5	*csgA* without R5 (S^{133} to Y^{151}) in pLR2	This study	FpLR5, pΔR5 P1
pΔR1&5	*csgA* without R1 and R5 in pLR2	This study	FpLR5, pΔR5 P1[b]

[a] Vectors used for expression and purification are not listed. These vectors were constructed by insertion of PCR amplified mutant *csgA* sequences with C-terminal hexahistidine tag into NdeI/EcoR1 sites in pMC3 (Chapman et al., 2002) replacing sequence encoding CsgA-his.

[b] pΔR1 was used as template for PCR to make pΔR1&5.

Table 3.2. Sequence of primers used in Chapter 3

Primer Name	Sequence
FpLR5[a]	5' CATGCCATGGCGAAACTTTTAAAAGTAGC 3'
RpLR5[b]	5' CGGGATCCTGTATTAGTACTGAT 3'
PΔR1 P1[c]	5' AATAGTCAAGTCAGAATTTGGGCCGCTATT 3'
PΔR1 P2[d]	5' AATAGCGGCCCAAATTCTGACTTGACTATT 3'
PΔR2 P1	5' GATCGATTGAGCTGTTACGGGCATCA 3'
PΔR2 P2	5' TGATGCCCGTAACAGCTCAATCGATC 3'
PΔR3 P1	5' CCGTCATTTCAGAGTCATCTGAGCCCT 3'
PΔR3 P2	5' AGGGCTCAGATGACTCTGAAATGACGG 3'
PΔR4 P1	5' CGTTGACGGAGGAATTTTTGCCGTTC 3'
PΔR4 P2	5' GAACGGCAAAAATTCCTCCGTCAACG 3'
PΔR5 P1	5' CGGGATCCTGTATTAGTTAGATGCAG 3'

[a] FpLR5 is paired to noncoding strand immediately upstream of the start codon of *csgA* in pLR5.
[b] RpLR5 is paired to coding strand immediately downstream of the stop codon of *csgA* in pLR5.
[C] The primers with odd number such as pΔR1 P1 are paired to the coding strand of *csgA* template.
[d] The primers with even number such as pΔR1 P2 are paired to the noncoding strand of *csgA* template.

References

Barnhart, M. M., and Chapman, M. R. (2006). Curli biogenesis and function. Annu Rev Microbiol *60*, 131-147.

Bucciantini, M., Giannoni, E., Chiti, F., Baroni, F., Formigli, L., Zurdo, J., Taddei, N., Ramponi, G., Dobson, C. M., and Stefani, M. (2002). Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. Nature *416*, 507-511.

Chapman, M. R., Robinson, L. S., Pinkner, J. S., Roth, R., Heuser, J., Hammar, M., Normark, S., and Hultgren, S. J. (2002). Role of *Escherichia coli* curli operons in directing amyloid fiber formation. Science *295*, 851-855.

Chiti, F., and Dobson, C. M. (2006). Protein misfolding, functional amyloid, and human disease. Annu Rev Biochem *75*, 333-366.

Collinson, S. K., Emody, L., Muller, K. H., Trust, T. J., and Kay, W. W. (1991). Purification and characterization of thin, aggregative fimbriae from Salmonella enteritidis. J Bacteriol *173*, 4773-4781.

Collinson, S. K., Parker, J. M., Hodges, R. S., and Kay, W. W. (1999). Structural predictions of AgfA, the insoluble fimbrial subunit of Salmonella thin aggregative fimbriae. J Mol Biol *290*, 741-756.

DePace, A. H., Santoso, A., Hillner, P., and Weissman, J. S. (1998). A critical role for amino-terminal glutamine/asparagine repeats in the formation and propagation of a yeast prion. Cell *93*, 1241-1252.

Fowler, D. M., Koulov, A. V., Balch, W. E., and Kelly, J. W. (2007). Functional amyloid--from bacteria to humans. Trends Biochem Sci *32*, 217-224.

Gerstel, U., and Romling, U. (2003). The csgD promoter, a control unit for biofilm

formation in Salmonella typhimurium. Res Microbiol *154*, 659-667.

Glover, J. R., Kowal, A. S., Schirmer, E. C., Patino, M. M., Liu, J. J., and Lindquist, S. (1997). Self-seeded fibers formed by Sup35, the protein determinant of [PSI+], a heritable prion-like factor of S. cerevisiae. Cell *89*, 811-819.

Hammar, M., Bian, Z., and Normark, S. (1996). Nucleator-dependent intercellular assembly of adhesive curli organelles in *Escherichia coli*. Proc Natl Acad Sci U S A *93*, 6562-6566.

Hammer, N. D., Schmidt, J. C., and Chapman, M. R. (2007). The curli nucleator protein, CsgB, contains an amyloidogenic domain that directs CsgA polymerization. Proc Natl Acad Sci U S A.

Harper, J. D., and Lansbury, P. T., Jr. (1997). Models of amyloid seeding in Alzheimer's disease and scrapie: mechanistic truths and physiological consequences of the time-dependent solubility of amyloid proteins. Annu Rev Biochem *66*, 385-407.

Hartley, D. M., Walsh, D. M., Ye, C. P., Diehl, T., Vasquez, S., Vassilev, P. M., Teplow, D. B., and Selkoe, D. J. (1999). Protofibrillar intermediates of amyloid beta-protein induce acute electrophysiological changes and progressive neurotoxicity in cortical neurons. J Neurosci *19*, 8876-8884.

Jarrett, J. T., and Lansbury, P. T., Jr. (1993). Seeding "one-dimensional crystallization" of amyloid: a pathogenic mechanism in Alzheimer's disease and scrapie? Cell *73*, 1055-1058.

Lambert, M. P., Barlow, A. K., Chromy, B. A., Edwards, C., Freed, R., Liosatos, M., Morgan, T. E., Rozovsky, I., Trommer, B., Viola, K. L.*, et al.* (1998). Diffusible, nonfibrillar ligands derived from Abeta1-42 are potent central nervous system neurotoxins. Proc Natl Acad Sci U S A *95*, 6448-6453.

Lesne, S., Koh, M. T., Kotilinek, L., Kayed, R., Glabe, C. G., Yang, A., Gallagher, M., and Ashe, K. H. (2006). A specific amyloid-beta protein assembly in the brain impairs memory. Nature *440*, 352-357.

Prusiner, S. B., Scott, M. R., DeArmond, S. J., and Cohen, F. E. (1998). Prion protein biology. Cell *93*, 337-348.

Robinson, L. S., Ashman, E. M., Hultgren, S. J., and Chapman, M. R. (2006). Secretion of curli fibre subunits is mediated by the outer membrane-localized CsgG protein. Mol Microbiol *59*, 870-881.

Roher, A. E., Chaney, M. O., Kuo, Y. M., Webster, S. D., Stine, W. B., Haverkamp, L. J., Woods, A. S., Cotter, R. J., Tuohy, J. M., Krafft, G. A*., et al.* (1996). Morphology and toxicity of Abeta-(1-42) dimer derived from neuritic and vascular amyloid deposits of Alzheimer's disease. J Biol Chem *271*, 20631-20635.

Romling, U., Bian, Z., Hammar, M., Sierralta, W. D., and Normark, S. (1998). Curli fibers are highly conserved between Salmonella typhimurium and *Escherichia coli* with respect to operon structure and regulation. J Bacteriol *180*, 722-731.

Sauer, F. G., Remaut, H., Hultgren, S. J., and Waksman, G. (2004). Fiber assembly by the chaperone-usher pathway. Biochim Biophys Acta *1694*, 259-267.

Shorter, J., and Lindquist, S. (2005). Prions as adaptive conduits of memory and inheritance. Nat Rev Genet *6*, 435-450.

Tessier, P. M., and Lindquist, S. (2007). Prion recognition elements govern nucleation, strain specificity and species barriers. Nature *447*, 556-561.

Walsh, D. M., Klyubin, I., Fadeeva, J. V., Cullen, W. K., Anwyl, R., Wolfe, M. S., Rowan, M. J., and Selkoe, D. J. (2002). Naturally secreted oligomers of amyloid beta protein

potently inhibit hippocampal long-term potentiation in vivo. Nature *416*, 535-539.

Wang, X., Smith, D. R., Jones, J. W., and Chapman, M. R. (2007). In vitro polymerization of a functional *Escherichia coli* amyloid protein. J Biol Chem *282*, 3713-3719.

Yonekura, K., Maki-Yonekura, S., and Namba, K. (2002). Growth mechanism of the bacterial flagellar filament. Res Microbiol *153*, 191-197.

Chapter 4

The Critical Roles of the Conserved Glutamine and Asparagine Residues in CsgA Nucleation and Polymerization[4]

Abstract

Amyloids are proteinaceous fibers commonly associated with neurodegenerative diseases and prion-based encephalopathies. Many different polypeptides can form amyloid fibers, leading to the suggestion that amyloid is a primitive main-chain-dominated structure. A growing body of evidence suggests that amino acid side chains dramatically influence amyloid formation. The specific role fulfilled by side chains in amyloid formation, especially *in vivo*, remains poorly understood. In this chapter, I determined the role of internally conserved polar and aromatic residues in promoting amyloidogenesis of the functional amyloid protein, CsgA. CsgA is the major protein component of curli fibers assembled by enteric bacteria such as *Escherichia coli* and *Salmonella* spp. *In vivo* CsgA polymerization into an amyloid fiber requires the CsgB nucleator protein. The CsgA amyloid core region is composed of five repeating units, defined by regularly spaced Ser, Gln and Asn residues. The results of a comprehensive alanine scan mutagenesis screen showed that Gln and Asn residues at positions 49, 54, 139 and 144 were critical for curli assembly. Alanine substitution of Q49 or N144 impeded the ability of CsgA to respond to CsgB-mediated heteronucleation, and the ability of CsgA to self-polymerize *in vitro*. However, CsgA proteins harboring these

[4] A version of this chapter has been published as Wang *et al.*, 2008, J Mol Biol *380*, 570-580.

mutations were still seeded by preformed wild-type CsgA fibers *in vitro*. This suggests that CsgA-fibril-mediated seeding and CsgB-mediated heteronucleation have distinguishable mechanisms. Remarkably, Gln residues at positions 49 and 139 could not be replaced by Asn residues without interfering with curli assembly, suggesting that the side chain requirements were especially stringent at these positions. This analysis demonstrates that bacterial amyloid formation is driven by specific side chain contacts and provides a clear illustration of the essential roles of specific side chains in promoting amyloid formation.

Introduction

Amyloid formation is readily associated with neurodegenerative diseases and prion-based encephalopathies (Chiti and Dobson, 2006). Amyloid fibers are 4-10 nm wide, unbranched proteinaceous filaments (Chiti and Dobson, 2006). Amyloid fibers possess a characteristic cross-β sheet quaternary structure, in which β strands are perpendicular to fibril axis.[1] These fiber structures have specific tinctorial properties, binding to dyes such as Congo red and thioflavin T (ThT) (Chiti and Dobson, 2006). Amyloid toxicity is complex, but a growing body of work suggests that pre-fiber aggregates are cytotoxic, while mature fibers are relatively benign (Hardy and Selkoe, 2002). Therefore, understanding the folding cascades involved in amyloid formation is necessary for developing new therapies to target these processes.

A newly described class of 'functional' amyloids suggest that amyloid formation can be an integral part of normal cellular physiology (Fowler et al., 2007; Hammer et al., 2008). Functional amyloids provide a unique perspective on amyloidogenesis because the

cell must have coevolved mechanisms to prevent the toxicity normally associated with amyloid formation. One of the best understood functional amyloids is curli, a bacterially produced extracellular fiber required for biofilm formation and other community behaviors (Barnhart and Chapman, 2006). In *E. coli*, the polymerization of the major curli fiber subunit protein CsgA into an amyloid fiber is dependent on the minor curli subunit protein, CsgB (Hammar et al., 1996). CsgA remains soluble until it encounters outer membrane-localized CsgB (Barnhart and Chapman, 2006). CsgB has also been demonstrated to have amyloid-forming properties and apparently serves as a template for CsgA polymerization (Hammer et al., 2007).

We previously showed that, like many other amyloids, preformed CsgA fibers could seed soluble CsgA polymerization *in vitro* as presented in Chapter 2 (Wang et al., 2007). Therefore, we proposed that the growing CsgA fiber on the cell surface could serves as a template promoting soluble CsgA polymerization in a process akin to seeding (Wang et al., 2007). The molecular details of CsgA fiber-mediated seeding and CsgB-mediated heteronucleation remain poorly described. Because nucleation underlies the rate-limiting step of amyloid propagation, understanding the nature of this mechanism will shed light on how cells control amyloid formation.

The primary sequence of CsgA can be divided into three functional domains: an N-terminal Sec signal sequence (cleaved after translocation into the periplasmic space), an N-terminal 22 amino acid segment (N22) that directs CsgA secretion across the outer membrane (Robinson et al., 2006), and an amyloid core region (residues 43 to 151) that contains five imperfect repeating units, each 19-23 amino acids in length (Figure 4.1A) (Barnhart and Chapman, 2006). The five repeating units form a protease-resistant

structure (Collinson et al., 1999), which is proposed to be the amyloid core of CsgA (Wang et al., 2007). These repeats are distinguished by the consensus sequence Ser-X5-Gln-X4-Asn-X5-Gln and are linked by 4 or 5 residues (Collinson et al., 1999). These Ser, Gln and Asn residues are conserved among CsgA homologs of many enteric bacteria (Wang et al., 2007). This high degree of amino acid sequence conservation suggests that these residues may play an important role in curli assembly.

Many polypeptides, if not all, can assemble into an amyloid fiber *in vitro* even though amyloid-forming proteins do not necessarily share amino acid similarities (Chiti and Dobson, 2006). Therefore, it has been proposed that amyloid formation is an inherent property of polypeptide main chains (Chiti and Dobson, 2006). However, specific residues likely play a role in promoting both disease-associated and functional amyloid formation. Yeast prion protein Sup35p has a Gln/Asn rich domain at N-terminus that has been implicated in prion propagation (DePace et al., 1998; Liu and Lindquist, 1999; Osherovich et al., 2004). Moreover, the specific sequences in this Gln/Asn rich domain govern self-recognition and species-specific seeding activity (Tessier and Lindquist, 2007). Aromatic residues in the islet amyloid polypeptide fragment positively contribute to its polymerization into amyloid fibers *in vitro* (Gazit, 2002). Structural analysis of Aβ 40 and Aβ42 revealed that two β sheets form a parallel β-sandwich stabilized by specific intermolecular side chain contacts and changes of these side chains resulted in defective fiber assembly (Luhrs et al., 2005; Petkova et al., 2002). Therefore, it is clear that side chains can influence amyloid polymerization, but the contribution of side chains to *in vivo* amyloidogenesis and the exact roles of amino acid side chain contacts remain poorly understood. Here, I performed a comprehensive mutagenesis study on CsgA and

identified the residues that promote CsgA amyloidogenesis. I showed that CsgA amyloidogenesis is driven by the side chain contacts of four Gln and Asn residues in N- and C-terminal repeats. These Gln and Asn residues play essential roles in the response to CsgB-mediated heteronucleation and the initiation of efficient self-assembly *in vitro*. Strikingly, these four Gln and Asn residues are not required for CsgA seeding suggesting CsgA seeding and CsgB-mediated heteronucleation have distinct mechanisms.

Experimental Procedures

Bacterial Growth

To induce curli production, bacteria were grown on YESCA plates (1.0 g yeast extract, 10 g casamino acids and 20 g agar per liter) at 26°C for 48 h (Chapman et al., 2002). Antibiotics were added to plates at the following concentrations: kanamycin 50 μg/ml, chloramphenicol 25 μg/ml, or ampicillin 100 μg/ml. Curli production was monitored by using Congo red-YESCA (CR-YESCA) plate (Chapman et al., 2002).

Strains and Plasmids

Strains LSR10 (MC4100::*csgA*), LSR12 (C600::*csgBAC* and *csgDEFG*), LSR13 (MC4100::*csgBA*) and MHR261 (MC4100::*csgB*) and plasmids pLR5 (encoding CsgA), pLR2 (vector control) and pMC1(encoding CsgG) were previously constructed (Chapman et al., 2002; Hammar et al., 1996; Hammer et al., 2007; Robinson et al., 2006). Plasmids containing CsgA mutations were constructed by site-specific mutagenesis using standard overlapping PCR extension. PCR products containing relevant mutations and NcoI/BamHI restriction endonuclease sites at 5'/3' ends were cloned downstream of the *csgBA* promoter into NcoI/BamHI sites of vector pLR2. The expression vector pMC3

was generated previously (Chapman et al., 2002). To express and purify CsgA mutant proteins, PCR amplified mutant sequences including C-terminal hexahistidine tag were cloned into pMC3, replacing the sequence encoding CsgA-his. Sequences of constructs were verified by DNA sequencing.

Western analysis

The immunoblotting of CsgA in whole-cells and plugs was performed as previously described (Chapman et al., 2002). Briefly, bacterial cells grown on YESCA plate for 48 h were scraped off and normalized by optical density at 600 nm. Cell pellets were resuspended in 2X SDS loading buffer either with or without prior formic acid (FA) treatment as previously described (Chapman et al., 2002). Alternatively, 8 mm circular plugs including cells and underlying agar were resuspended in 2X SDS loading buffer either with or without prior FA treatment. Samples were electrophoresed on 15% sodium dodecyl sulfate (SDS)-polyacrylamide and transferred onto polyvinylidene difluoride membrane using standard techniques. Western blots were probed and developed as previously described (Hammer et al., 2007).

Transmission Electron Microscopy

A Philips CM10 Transmission Electron Microscope was used to visualize the cell samples and protein fiber aggregates. Samples (10 μl) were placed on Formvar-coated copper grids (Ernest F. Fullam, Inc., Latham, NY) for 2 min, washed with deionized water, and negatively stained with 2% uranyl acetate for 90 sec.

***In vitro* polymerization and nucleation assay**

CsgA or CsgA mutant proteins and CsgG were co-overexpressed by induction with isopropyl β-D-1-thiogalactopyranoside (IPTG) in LSR12 (C600::*csgBAC* and

csgDEFG) strain. CsgA or CsgA mutant proteins were secreted to the supernatant and purified as previously described (Wang et al., 2007). After removal of imidazole by a desalting column, purified proteins were passed through a 30 kD cutoff filter (Amicon® Ultra, MA) to remove possible aggregates or seeds that might alter polymerization kinetics. Purified homogeneous proteins were loaded on 96-well opaque plate and ThT was added to a concentration of 20 μM. ThT fluorescence was measured every 10 min after shaking 5 sec by a Spectramax M2 plate reader (Molecular Devices, Sunnyvale, CA) set to 438 nm excitation and 495 nm emission with a 475 nm cutoff. Alternatively, samples were kept at room temperature in the presence of 0.02% NaN_3. At the indicated times, aliquots were removed, ThT was added at a concentration of 20 μM and fluorescence intensity was measured as described above. For seeding reactions, CsgA fiber seeds, prepared as previously described in Chapter 2 (Wang et al., 2007), were added to freshly purified CsgA mutant proteins immediately before the start of ThT fluorescence assay.

Overlay assay

10 μl freshly purified proteins were spotted onto a lawn of *csgA* (LSR10) or *csgBA* (LSR13) cells grown on YESCA plates at 26°C for 48 h. Samples were incubated on cells for 10 min at room temperature, stained with a 0.5 mg/ml Congo red in 50 mM potassium phosphate (pH 7.2) for 5 min, then washed with potassium phosphate buffer and photographed.

Interbacterial complementation

Cells that secrete CsgA or CsgA mutant proteins were streaked adjacently to *csgA* cells on a CR-YESCA plate. The plates were photographed after 48 h growth at 26°C.

Results

Ala scan mutagenesis of internally conserved polar residues

The amyloid core of CsgA is composed of five repeating units, defined by internally conserved and regularly spaced Ser, Gln and Asn residues conserved among many enteric bacteria (Figure 4.1A) (Wang et al., 2007). We performed an Ala scan mutagenesis on these 20 polar residues to test their importance in directing bacterial amyloid formation. Amyloid formation of each CsgA mutant was initially assessed by growing bacteria on plates amended with Congo red, as cells expressing wild-type CsgA will stain a deep red color on this medium (Figure 4.1B). Only two mutations, Q49A and N144A, significantly reduced Congo red binding relative to wild-type CsgA (Figure 4.1B). Transmission electron microscopy (TEM) was used to validate the Cong red binding phenotypes. *csgA* cells (LSR10) transformed with pLR5 (encoding CsgA) produced curli fibers that were indistinguishable from those assembled by wild-type strain MC4100 by TEM. Cells expressing $CsgA^{Q49A}$ or $CsgA^{N144A}$ assembled fewer fibers than cells expressing wild-type CsgA observed by TEM (Figure 4.1C).

CsgA polymerization into an amyloid fiber *in vivo* can also be monitored by its ability to migrate as a monomer on SDS PAGE gels after dissociation by a strong acid, formic acid (FA) (Chapman et al., 2002). For example, CsgA produced by wild-type cells is whole cell-associated and SDS insoluble (Collinson et al., 1991). Brief treatment with FA liberates CsgA monomers from curli fibers produced by wild-type strain MC4100(Hammar et al., 1996). Similar to the wild-type strain, CsgA produced by *csgA*/pLR5 was SDS insoluble and required brief pretreatment with FA to migrate into the gel (Figure 4.2A, lanes 1 and 2). All of the Ser mutants (S43A, S66A, S88A, S111A

and S133A) were also SDS insoluble and associated with the whole cell fraction (Figure 4.2A, lanes 5, 6, 13, 14, 21, 22, 29, 30, 37 and 38). However, 13 of 15 Ala substitution mutants of Gln/Asn residues (Q49A, N54A, Q60A, Q72A, N77A, Q94A, N99A, Q105A, Q117A, N122A, Q128A, Q139A and N144A) showed different levels of SDS solubility by whole-cell western analysis, suggesting these polar residues help stabilize the amyloid structure (Figure 4.2A). $CsgA^{Q49A}$ and $CsgA^{N144A}$ were unable to complement Congo red binding to a *csgA* mutant (Figure 4.1B), and very little of these mutant proteins could be recovered from whole cell lysates scraped off YESCA plates (Figure 4.2A, lanes 7, 8, 41 and 42).

To test the possibility that $CsgA^{Q49A}$ and $CsgA^{N144A}$ were secreted away from the cell as soluble proteins, cells and the underlying agar were collected and analyzed by western blotting. In these samples, called 'plugs', both $CsgA^{Q49A}$ and $CsgA^{N144A}$ were readily detected and SDS soluble, demonstrating that $CsgA^{Q49A}$ and $CsgA^{N144A}$ were stable, secreted to the cell surface and unpolymerized (Figure 4.2B, lanes 2, 3, 11 and 12). $CsgA^{N54A}$ and $CsgA^{Q139A}$ were also significantly different from other mutants in the whole cell SDS solubility assays. $CsgA^{N54A}$ was completely SDS soluble (Figure 4.2A, lanes 9 and 10) and $CsgA^{Q139A}$ was not predominately cell associated (Figure 4.2A, lanes 39 and 40). $CsgA^{N54A}$ and $CsgA^{Q139A}$ were SDS soluble detected by western analysis of cells and the underlying agar (Figure 4.2B, lanes 4, 5, 8 and 9), suggesting $CsgA^{N54A}$ and $CsgA^{Q139A}$ were not assembled into wild-type like fibers *in vivo*. Collectively, Q49A and N144A were the most defective mutants in curli formation of the 20 mutants examined. In addition, the N54A and Q139A mutants were also significantly defective in curli assembly as measured by western analysis.

Ala scan mutagenesis of the aromatic residues

It was reported that aromatic residues may play an important role during amyloidogenesis (Azriel and Gazit, 2001; Gazit, 2002). We tested the contribution of aromatic residues in the CsgA amyloid core region by Ala scan mutagenesis. Except for $CsgA^{Y151A}$, Ala substitutions in the aromatic residues resulted in proteins that were phenotypically identical to wild-type CsgA as detected by Congo red binding and whole-cell western analysis (Figure 4.3A and 4.3B). $CsgA^{Y151A}$ was undetectable by western analysis of cells and the underlying agar, indicating that Y151A was unstable (Figure 4.3C, lanes 5 and 6). When Tyr^{151} was changed to Phe, Trp, Lys or Asp, curli formation was restored as evidenced by Congo red binding and whole-cell western analysis, suggesting Tyr residue at position 151 is not absolutely required (Figure 4.3D and data not shown).

$CsgA^{Q49A}$ and $CsgA^{N144A}$ are defective in self-polymerization *in vitro*

To further characterize the most defective CsgA mutant proteins, we purified $CsgA^{Q49A}$ and $CsgA^{N144A}$ and compared their polymerization to wild-type CsgA using the ThT assay described in Chapter 2 (Wang et al., 2007). Wild-type CsgA assembles into an amyloid fiber *in vitro* at concentrations above 2.0 μM in the absence of CsgB (Wang et al., 2007). Two parameters were used to compare the polymerization kinetics of CsgA and its mutant analogues as presented in Chapter 3. The first kinetic parameter was the time period preceding rapid fiber growth, called lag phase or T_0. The second parameter was the time period encompassing the fiber growth phase from initiation of rapid polymerization to its completion, called conversion time (T_c) (DePace et al., 1998). At a concentration of 40 μM, the T_0 of $CsgA^{Q49A}$ was similar to that of CsgA, while the T_c was

much greater than that of CsgA (Figure 4.4A). $CsgA^{N144A}$ polymerization had much greater T_0 and T_c than those of CsgA, suggesting the amido group of Asn at position 144 is critical for aggregation (Figure 4.4A). After 120 hrs, both $CsgA^{Q49A}$ and $CsgA^{N144A}$ had assembled into amyloid fibers with similar fiber morphology to wild-type CsgA fibers (Figure 4.4B, 4.4C and 4.4D).

$CsgA^{Q49A}$ and $CsgA^{N144A}$ are defective in heteronucleation response

Even though $CsgA^{Q49A}$ and $CsgA^{N144A}$ were defective in self-polymerization, in the presence of wild-type CsgA seeds they polymerized with efficiency similar to wild-type CsgA (data not shown). To test of the ability of $CsgA^{Q49A}$ and $CsgA^{N144A}$ to respond to CsgB-mediated heteronucleation, two different approaches were employed. The first was an overlay assay using freshly purified CsgA or CsgA mutant proteins and cells expressing the CsgB nucleator protein (Wang et al., 2008). In a CsgB-dependent manner, soluble wild-type CsgA was converted into an amyloid fiber within 1 minute of the overlay as evidenced by Congo red staining, or by TEM as shown in Chapter 3 (Wang et al., 2008). Various concentrations of CsgA, $CsgA^{Q49A}$ and $CsgA^{N144A}$ were overlaid on *csgA* cells ($CsgB^{+}$). At a concentration of 10 μM or higher, CsgA polymerized into an amyloid fiber in a CsgB-dependent fashion detected by Congo red staining (Figure 4.5A). However, 10 μM $CsgA^{Q49A}$ and $CsgA^{N144A}$ did not polymerize into amyloid-like fibers under the same conditions (Figure 4.5A). When the concentration of $CsgA^{Q49A}$ and $CsgA^{N144A}$ was increased to 40 μM, a Congo red-binding structure on the surface of CsgB-expressing cells was detected (Figure 4.5A).

Curli fibers are proposed to assemble after CsgA is secreted to the extracellular space. This can be illustrated by interbacterial complementation where a *csgA* mutant

strain ($CsgB^+$) is grown in close proximity to a *csgB* mutant strain ($CsgA^+$). CsgA molecules secreted by *csgB* donor cells polymerize into amyloid fibers on the *csgA* acceptor cells, as shown in Figure 4.5B (Hammar et al., 1996). We performed an interbacterial complementation assay to test CsgB-mediated heteronucleation responsiveness of CsgA, $CsgA^{Q49A}$ and $CsgA^{N144A}$ *in vivo*. CsgA secreted from *csgBA*/pLR5 (donor) responded to the heteronucleation of CsgB on the *csgA* cells (acceptor), and formed an amyloid fiber on the surface of *csgA* cells as demonstrated by Congo red (Figure 4.5B) (Hammar et al., 1996). There was reduced Congo red binding on the surface of *csgA* cells when *csgBA* cells containing plasmid encoding $CsgA^{Q49A}$ or $CsgA^{N144A}$ were grown adjacently to *csgA* cells (Figure 4.5B). The protein levels of CsgA, $CsgA^{Q49A}$ and $CsgA^{N144A}$ produced in the *csgBA* genetic background were similar as detected by western analysis of cells and the underlying agar (data not shown). These data collectively demonstrate that $CsgA^{Q49A}$ and $CsgA^{N144A}$ were defective in responding to the heterogeneous nucleator CsgB.

Ala substitutions of Q49, N54, Q139 and N144 result in the dramatic loss of heteronucleation response to CsgB

In order to abolish the four critical Gln/Asn residues identified by our Ala scan mutagenesis, we constructed the mutant Q49A/N54A/Q139A/N144A. *In vivo* $CsgA^{Q49A/N54A/Q139A/N144A}$ ($CsgA^{slowgo}$) completely lost the ability to assemble into amyloid fibers as detected by Congo red binding, TEM and whole-cell western analysis (data not shown). We tested the competence of $CsgA^{slowgo}$ to respond to CsgB-mediated heteronucleation by overlay assay. Purified homogeneous $CsgA^{slowgo}$ could not be promoted to form amyloid fibers by cell-associated CsgB as evidenced by the lack of

Congo red binding in the overlay assay even at a concentration of 100 μM (Figure 4.6A). There was no Congo red binding on the surface of *csgA* cells when *csgBA* cells containing plasmid encoding $CsgA^{slowgo}$ were grown adjacently to *csgA* cells shown in Figure 4.6B, suggesting that $CsgA^{slowgo}$ completely lost the response to CsgB-mediated heteronucleation in the interbacterial complementation. The protein level of $CsgA^{slowgo}$ was similar to wild-type CsgA in this genetic background (data not shown). Collectively, these results demonstrate that $CsgA^{slowgo}$ remains unpolymerized even in the presence of the heteronucleator CsgB, suggesting these four polar residues are required in CsgB-mediated heteronucleation.

Q49, N54, Q139 and N144 are required for efficient self-polymerization and $CsgB_{trunc}$ cross-seeding not for CsgA seeding

$CsgA^{slowgo}$ was purified and its polymerization kinetics was analyzed *in vitro*. The T_0 and T_c of $CsgA^{slowgo}$ were much higher than CsgA (Figure 4.7A), suggesting that Q49, N54, Q139 and N144 play critical roles in self-polymerization *in vitro*. Self-assembled $CsgA^{slowgo}$ fibers were similar to CsgA fibers shown by TEM (Figure 4.7B). Interestingly, wild-type CsgA seeds efficiently eliminated the lag phase of the polymerization of $CsgA^{slowgo}$ as detected by ThT assay (Figure 4.7C). In the presence of CsgA seeds, the polymerization rate $CsgA^{slowgo}$ is very similar to wild-type CsgA as shown in Chapter 2 (Wang et al., 2007) and ThT signal reached stationary phase within 5 hours (Figure 4.7C). Polymerized $CsgA^{slowgo}$ fibers promoted by wild-type CsgA were similar to CsgA fibers shown by TEM (Figure 4.7D). A nucleation-competent CsgB truncation mutant (called $CsgB_{trunc}$) was purified as previously described (Hammer et al., 2007). $CsgB_{trunc}$ has been demonstrated to efficiently seed the polymerization of wild-type CsgA (Hammer et al.,

2007). Unlike CsgA seeds, $CsgB_{trunc}$ seeds did not efficiently promote the polymerization of $CsgA^{slowgo}$ under the same condition shown in Figure 4.7C. Self-assembled $CsgA^{slowgo}$ fibers were able to seed wild-type CsgA polymerization, suggesting $CsgA^{slowgo}$ fibers are similar to wild-type CsgA fibers in terms of seeding specificity (data not shown).

Conservative Q/N substitutions

It was proposed that *in vitro* amyloid formation is strongly influenced by simple physicochemical factors of polypeptides such as hydrophobicity, charge and secondary structure propensity. (Chiti and Dobson, 2006; Chiti et al., 2003) Ala substitutions at positions 49, 54, 139 and 144 decreased CsgA polymerization *in vivo*. To test the stringency of the amino acid present at these critical positions, I constructed four mutants (Q49N, N54Q, Q139N and N144Q) that contained relatively conservative amino acid changes. Gln and Asn residues have similar physicochemical features, including the same amido group. Surprisingly, even though the side chain of Asn is only one carbon distance shorter than that of Gln, $CsgA^{Q49N}$ had dramatically reduced Congo red binding and was almost undetectable in lysates of whole cells scraped off YESCA plate (Figure 4.8A and 4.8B, lanes 7 and 8). $CsgA^{Q49N}$ was stable and secreted into the agar as detected by western analysis of cells and the underlying agar (Figure 4.8C, lanes 7 and 8). In addition, whole-cell western analysis showed that Q139N (Figure 4.8B, lanes 15 and 16) was as defective as Q139A (Figure 4.2A, lanes 39 and 40). The sequence requirements at positions 49 and 139 appear extremely stringent during fiber assembly since either Ala or Asn was not tolerated at these positions. The Congo red binding of $CsgA^{N144Q}$ and $CsgA^{N54Q}$ was similar to that of wild-type CsgA (Figure 4.8A). In addition, $CsgA^{N144Q}$ and $CsgA^{N54Q}$ were predominately SDS insoluble as shown by whole-cell western

analysis (Figure 4.8B, lanes 11, 12, 13 and 14). The SDS insolubility of $CsgA^{N144Q}$ and $CsgA^{N54Q}$ was similar to wild-type CsgA as detected by western analysis of cells and the underlying agar (Figure 4.8C, lanes 11, 12, 15 and 16). We observed wild-type like fibers assembled by the *csgA* strain that contained plasmid encoding $CsgA^{N144Q}$ by TEM, while relatively few fibers were observed on the *csgA* strain expressing $CsgA^{Q49N}$ (Figure 4.8D). Collectively, *in vivo* Gln residues at positions 49 and 139 could not be replaced by Asn, while Asn residues at position 54 and 144 could be replaced by Gln and function like wild-type CsgA. *In vitro* $CsgA^{Q49N}$ was as defective as $CsgA^{Q49A}$ in both heteronucleation responsiveness and self-polymerization. $CsgA^{Q49N}$ polymerization had greater T_c than that of CsgA at the same concentration measured by ThT assay (Figure 4.8E). $CsgA^{Q49N}$ did not respond to CsgB nucleation at 10 μM as detected by overlay assay (Figure 4.8E insert). It has been proposed that amino acid content can be an important factor in promoting amyloid formation (Ross et al., 2004; Ross et al., 2005). To test whether N144Q can suppress the polymerization defect of Q49N, we constructed the mutant Q49N/N144Q. Like Q49N, Q49N/N144Q had significantly reduced Congo red binding relative to wild-type CsgA (Figure 4.8F). Collectively, these findings suggest that CsgB-mediated heteronucleation and CsgA self-polymerization require spatially constrained contacts between certain amino acid side chains.

Discussion

Curli function as part of the extracellular matrix produced by many Gram-negative enteric bacteria. Curli assembly is a precisely coordinated process that is highly amenable to study because of the sophisticated genetic and biochemical tools afforded by

E. coli. Here, I elucidated the CsgA sequence determinants that drive amyloid formation. I found that Gln and Asn residues at the N- and C- terminal repeats are critical for curli assembly. These Gln and Asn residues are necessary for CsgA to respond to CsgB-mediated heteronucleation and to initiate efficient CsgA self-assembly *in vitro*.

The contribution of amino acid side chains to functional amyloid formation

Unlike amorphous aggregates formed by nonspecific hydrophobic interactions, amyloid fibers are highly ordered proteinaceous polymers that are stabilized by specific interactions. The side chains can clearly influence amyloidogenesis (Chiti and Dobson, 2006). But how specific sequences influence amyloid formation *in vivo* remains poorly understood. We demonstrated that four internally conserved Gln and Asn residues were required to respond to CsgB-mediated heteronucleation. When these four residues are changed to Ala, the resulting protein, $CsgA^{slowgo}$, cannot efficiently initiate self-assembly *in vitro* as indicated by an extremely long lag phase before polymerization (Figure 4.7A). $CsgA^{slowgo}$ also cannot efficiently respond to CsgB-mediated heteronucleation (Figure 4.6). These results suggest that the specific side chain interactions are the driving forces to efficiently initiate bacterial amyloid formation *in vivo* and *in vitro*. Interestingly, we found that these Gln/Asn residues required for CsgB-mediated heteronucleation are not necessary for CsgA fiber-mediated seeding, suggesting that CsgA seeding and CsgB-mediated heteronucleation may have somewhat distinct mechanisms (Figure 4.7C). $CsgA^{slowgo}$ is completely defective in amyloid formation *in vivo*, although it can inefficiently assemble into amyloid fibers *in vitro* (Figure 4.7A and 4.7B). This distinction between *in vitro* and *in vivo* behaviors suggests that *in vivo* amyloid formation has highly stringent requirements that limit initiation of amyloid formation to CsgB-

mediated heteronucleation. For example, CsgA remains soluble in the extracellular milieu in the absence of nucleator CsgB even though CsgA is amyloidogenic *in vitro* (Wang et al., 2007). This indicates that CsgA may diffuse into the extracellular milieu if it cannot participate in efficient nucleation. Curli also provides an elegant example how amyloidogenecity can be temporally and spatially controlled. It is possible that an evolved functional amyloid such as curli capitalizes on specific side-chain interactions to tightly regulate and control amyloidogenecity.

The diverse sequence determinants of amyloid formation

The sequence-specific determinants of amyloid formation have been most studied *in vitro*. Clearly, different amyloid proteins employ various interactions such as hydrogen bonding, hydrophobic and electrostatic interactions to promote intra- and inter-molecular associations. For example, the hydrophobic stretches in Aβ42, rather than specific side chains, are sufficient to promote aggregation (Kim and Hecht, 2006). Yeast prion proteins Sup35p, Ure2p and Rnq1p all have Gln/Asn rich domains that are essential for prion propagation *in vivo* and *in vitro* (DePace et al., 1998; Michelitsch and Weissman, 2000; Sondheimer and Lindquist, 2000; Taylor et al., 1999). In Huntington's disease, polyglutamine peptides play an essential role in molecular etiology. In addition, the length of polyglutamine peptides is positively correlated to *in vitro* aggregation (Chen et al., 2002). Gln/Asn rich domains are proposed to have a high potential to be amyloidogenic because they can form intramolecular hydrogen bonds (Michelitsch and Weissman, 2000). Here, I showed that the internally conserved Gln and Asn residues at the N- and C- terminal repeats are critical for CsgA fiber formation. The initiation of

CsgA polymerization both in *vitro* and *in vivo* appears to be dependent on hydrogen bonds formed by these Gln and Asn residues.

Moreover, I showed that 13 of 15 Ala substitution mutants of Gln/Asn residues (Q49A, N54A, Q60A, Q72A, N77A, Q94A, N99A, Q105A, Q117A, N122A, Q128A, Q139A and N144A) showed somewhat more sensitive to SDS relative to wild-type CsgA (Figure 4.2A), suggesting these polar residues help stabilize the amyloid structure. Interestingly, the polar amides of these internally conserved Gln and Asn residues were proposed to form hydrogen bonds with backbones to stabilize its fibril structure in a CsgA predicted structural model. (Collinson et al., 1999) It is possible that these critical Gln/Asn residues not only play an important role to initiate fiber assembly but also stabilize the fiber structure when fibers are formed.

It was proposed that interactions among aromatic residues by π-stacking possibly play a critical role in self-assembly of some amyloid fibrils such as islet amyloid polypeptide (Gazit, 2002). To assess the role of aromatic residues in guiding *in vivo* CsgA amyloidogenesis, I individually replaced each aromatic residue in the amyloid core of CsgA, revealing that aromatic residues do not significantly contribute to curli assembly (Figure 4.3).

Bacterial amyloid formation is influenced by the spatial arrangement of Q/N residues

It has been suggested that amyloid formation is determined by simple physicochemical factors of polypeptides such as hydrophobicity, charge and secondary structure propensity (Chiti et al., 2003). Furthermore, the amyloidogenecity of yeast prions Ure2p and Sup35p appears to be encoded by a particular amino acid composition,

rather than specific sequences (Ross et al., 2004; Ross et al., 2005). In contrast, my work demonstrates that CsgA amyloidogenesis, especially *in vivo*, is highly dependent on its specific sequence rather than on its overall amino acid composition or the physicochemical properties of polypeptide chains. The Gln residue at positions 49 and 139 of CsgA could not be replaced by Asn residues without interfering with curli assembly (Figure 4.8), even though Asn shares many physicochemical features with Gln. $CsgA^{Q49N\text{-}N144Q}$ and wild-type CsgA share an identical composition of amino acids, but $CsgA^{Q49N\text{-}N144Q}$ is dramatically defective in curli assembly (Figure 4.8F). Additionally, although $CsgA^{Q49A}$, $CsgA^{Q60A}$, $CsgA^{Q94A}$, $CsgA^{Q105A}$, $CsgA^{Q117A}$, $CsgA^{Q128A}$, $CsgA^{Q150A}$ have the identical composition of amino acids, $CsgA^{Q49A}$ is dramatically defective than other Ala mutants in fiber assembly (Figure 4.1B and 4.2A). We previously showed that N- and C-terminal CsgA repeats drive curli assembly in Chapter 3 (Wang et al., 2008). Our current study clearly suggests the position of Gln or Asn residues plays an equally important role for amyloid assembly.

Globular protein folding depends on specific side chain contacts. Whether amyloid formation, especially *in vivo,* depends on specific side chain contacts is just beginning to be tested. My work demonstrates that the initiation of CsgA amyloid formation requires specific amino acid side chains. Specific side chain contacts are likely to facilitate the spatial and temporal control over the polymerization of functional amyloids.

Figures

Figure 4.1. Effect of Ala substitutions of conserved Ser, Gln and Asn of CsgA on curli assembly.

(A) Schematic of CsgA including an N-terminal Sec signal sequence and the N-terminal 22 residues that precede the five repeating units. The regularly-spaced Ser, Gln and Asn residues are indicated in red color and the aromatic residues are indicated in blue color. The position of highlighted residues is indicated above the residue. (B) CR-YESCA plate with *csgA* cells transformed with the plasmids encoding CsgA with Ala substitution after 48 h of growth at 26°C. *csgA*/pLR5 and *csgA*/pLR2 were used as the positive control (PC) and vector control (VC), respectively. The specific Ala substitution is indicated. (C) Negative-stain TEM micrographs of *csgA* cells containing vector control (pLR2) or plasmids encoding wild-type CsgA (pLR5), $CsgA^{Q49A}$ and $CsgA^{N144A}$. Scale bars are equal to 500 nm.

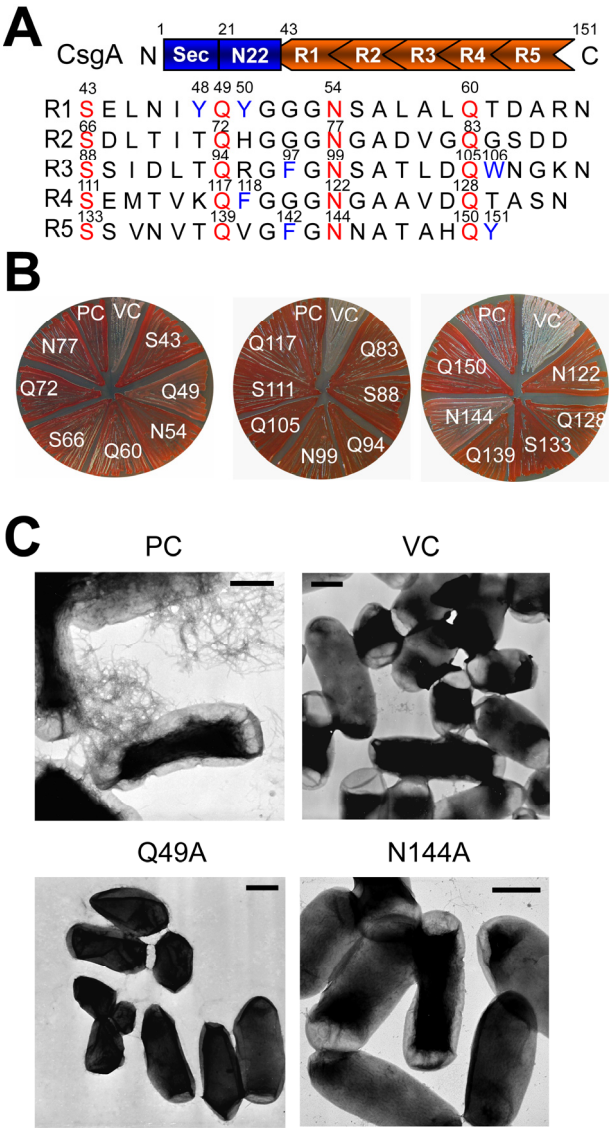


Figure 4.2. Western analysis of CsgA mutants with Ala substitutions of internally conserved Ser, Gln and Asn.
(A) Western blots of whole-cell lysates from *csgA* cells containing vector control or plasmids encoding wild-type CsgA or indicated Ala substituted mutants. (B) Western blot of whole-cells and underlying agar (agar plugs) from *csgA* strains containing plasmids encoding CsgA, $CsgA^{Q49A}$, $CsgA^{N54A}$, $CsgA^{Q139A}$ or $CsgA^{N144A}$. Samples were treated with (+) or without (-) FA. The blots were probed with anti-CsgA antibody.

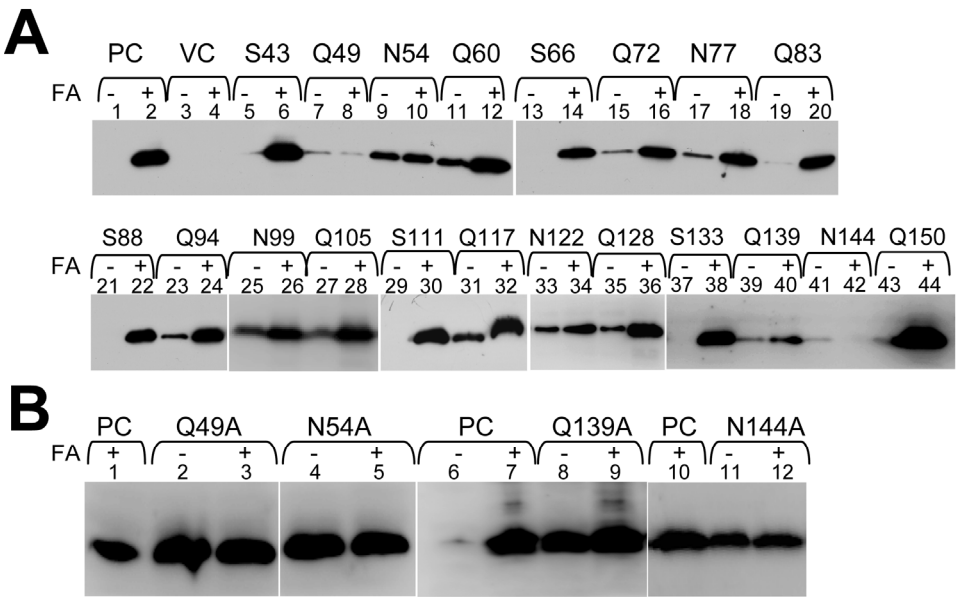


Figure 4.3. Ala scan mutagenesis of aromatic residues in CsgA amyloid core regions. (A) CR-YESCA plate with *csgA* cells transformed with the plasmids encoding CsgA, $CsgA^{Y48A}$, $CsgA^{Y50A}$, $CsgA^{F97A}$, $CsgA^{W106A}$, $CsgA^{F118A}$, $CsgA^{F142A}$ or $CsgA^{Y151A}$. (B) Western blots of whole-cell lysates from *csgB* cells and *csgA* cells containing plasmids encoding wild-type CsgA or indicated Ala mutants. (C) Western blot of whole-cells and underlying agar (agar plugs) from *csgB* cells and *csgA* strains containing plasmids encoding CsgA or $CsgA^{Y151A}$. (D) The CR-YESCA plate with *csgA* cells transformed with the plasmids encoding CsgA, $CsgA^{Y151F}$, $CsgA^{Y151W}$, $CsgA^{F151D}$ or $CsgA^{Y151K}$.

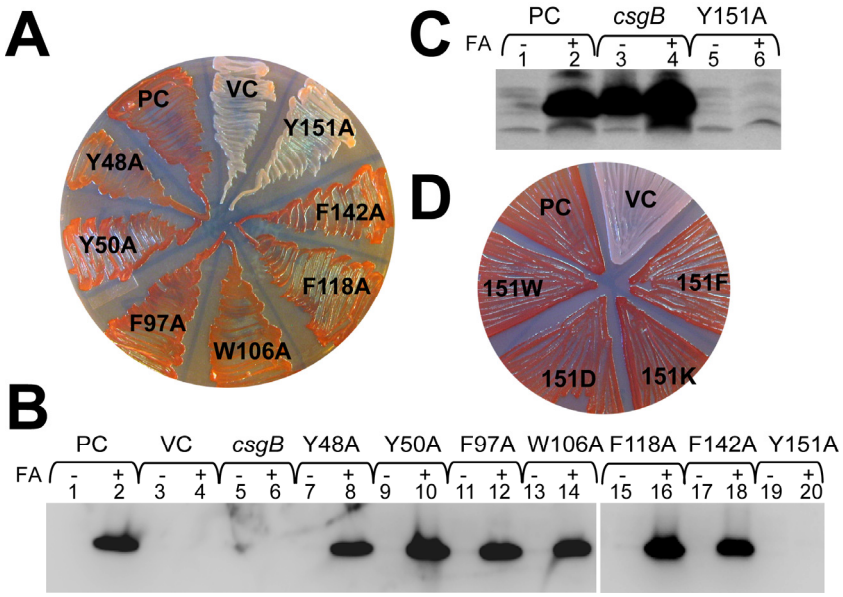


Figure 4.4. *In vitro* self-polymerization of CsgA^{Q49A} and CsgA^{N144A} are defective. (A) The fluorescence of 40 μM CsgA (solid circle), CsgA^{Q49A} (open circle) and CsgA^{N144A} (open square) mixed with 20 μM ThT was measured in 10-minute intervals at 495nm after excitation at 438nm. *a.u.*, arbitrary units. (B-D) Negative-stain EM micrographs of *in vitro* self-polymerized fibers of CsgA (B), CsgA^{Q49A} (C) and CsgA^{N144A} (D). Scale bars are equal to 200 nm.

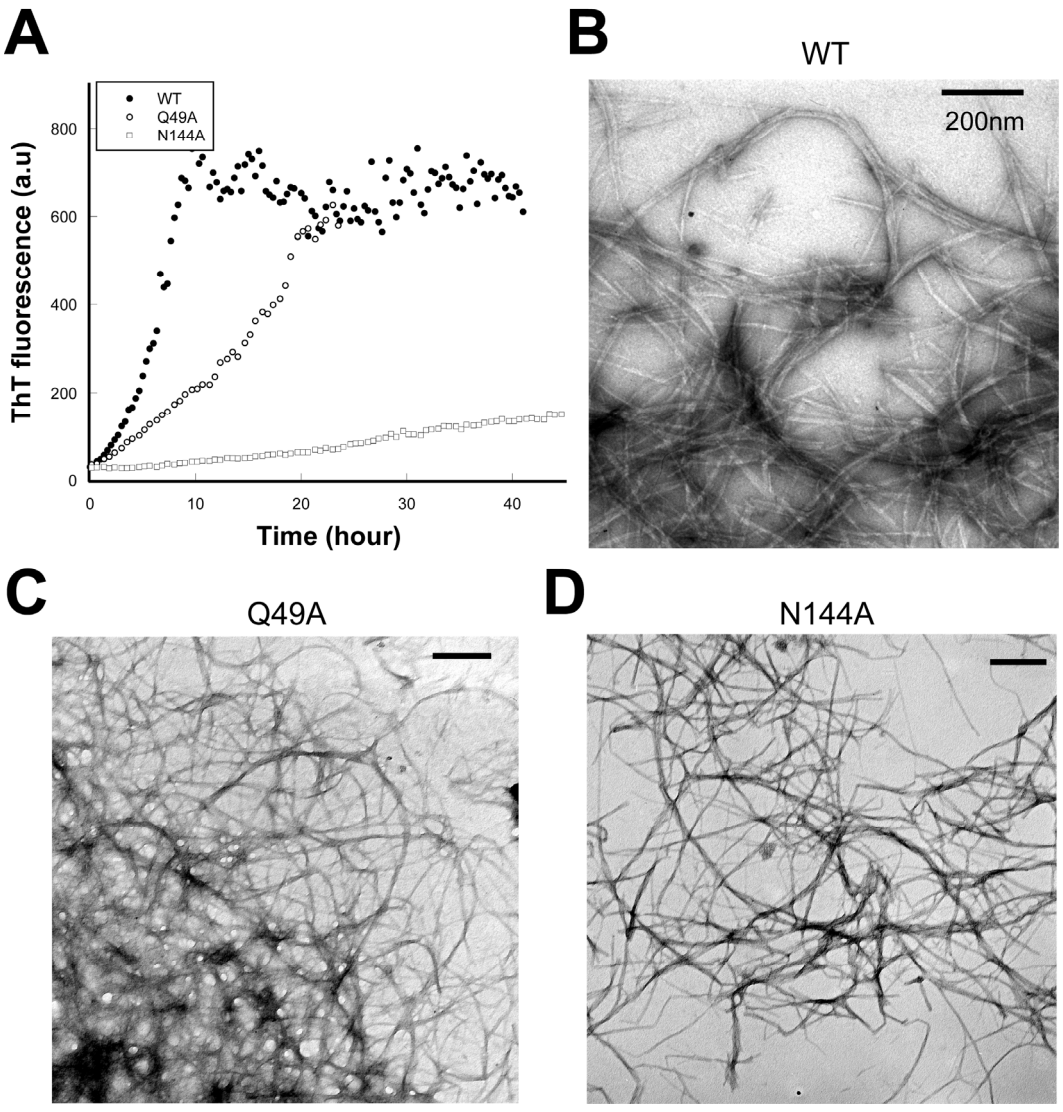


Figure 4.5. CsgA^{Q49A} and CsgA^{N144A} are defective in response to CsgB-mediated heteronucleation.
(A) A schematic presentation of the overlay assay in which soluble CsgA is spotted onto cells (*csgA*) expressing the nucleator protein, CsgB, where CsgA polymerization is catalyzed by CsgB (top). Freshly purified CsgA, CsgA^{Q49A} or CsgA^{N144A} at the indicated concentrations was spotted onto on *csgA* cells. After 10 min incubation, cells were stained with 0.5 mg/ml Congo red solution and washed by potassium phosphate buffer (bottom). (B) A schematic presentation of the interbacterial complementation assay in which soluble CsgA molecules secreted from *csgB* cells interact with CsgB on the surface of *csgA* cells to form fibers (top). *csgBA* cells containing plasmids encoding CsgA, CsgA^{Q49A} or CsgA^{N144A} were grown adjacently to *csgA* cells on a CR-YESCA plate for 48 h at 26°C (bottom).

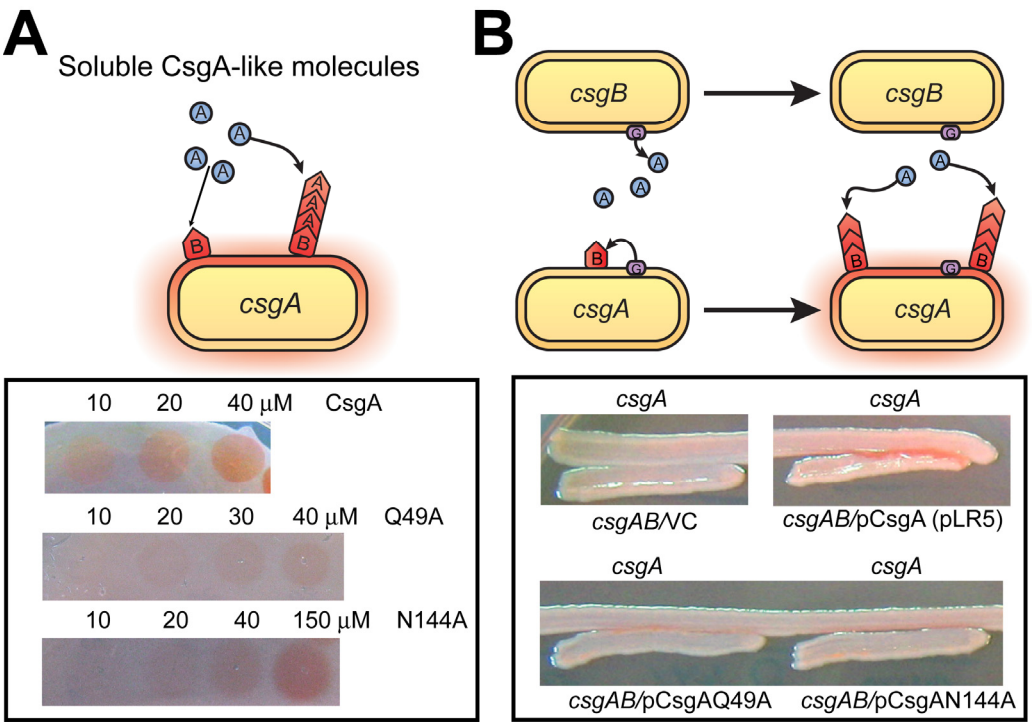


Figure 4.6. The heteronucleation responsiveness of CsgA^{slowgo} (CsgA$^{Q49A/N54A/Q139A/N144A}$).
(A) Freshly purified homogeneous CsgA^{slowgo} at the indicated concentrations was spotted onto *csgA* or *csgBA* cells. After 10 min incubation, cells were stained with 0.5 mg/ml Congo red solution and washed by potassium phosphate buffer. (B) *csgBA* cells containing plasmids encoding CsgA or CsgA^{slowgo} were grown adjacently to *csgA* cells on a CR-YESCA plate for 48 h at 26°C.

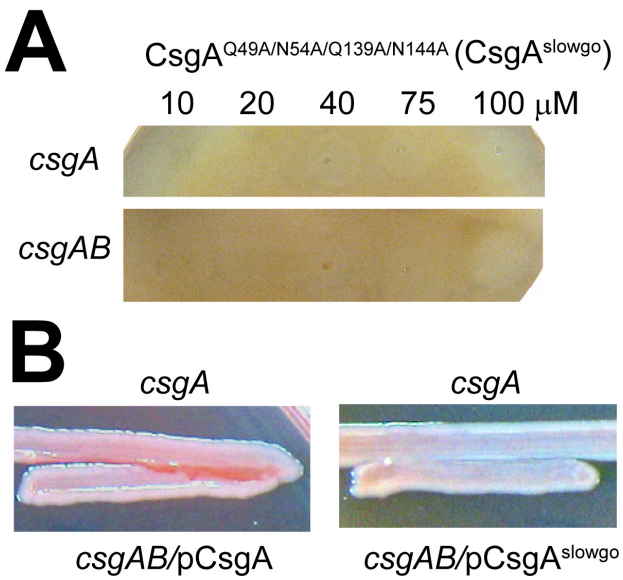


Figure 4.7. *In vitro* self-polymerization and seeding responsiveness of CsgA^{slowgo}. (A) ThT fluorescence intensity of 40 μM CsgA^{slowgo} incubated for the indicated time intervals. *a.u.*, arbitrary units. (B) Negative stain TEM micrograph of *in vitro* self-polymerized aggregates of CsgA^{slowgo}. Scale bars are equal to 200 nm. (C) 5% CsgA seeds, CsgB$_{trunc}$ seeds by weight or buffer vehicle were added to 60 μM CsgA^{slowgo} and the polymerization was measured by ThT fluorescence assay. The fluorescence intensity of nucleated CsgA^{slowgo} samples were shown at the indicated time intervals. The contribution to fluorescence intensity by seeds was subtracted. (D) Negative-stain TEM micrograph of *in vitro* polymerized aggregates of CsgA^{slowgo} promoted by CsgA. Scale bars are equal to 200 nm.

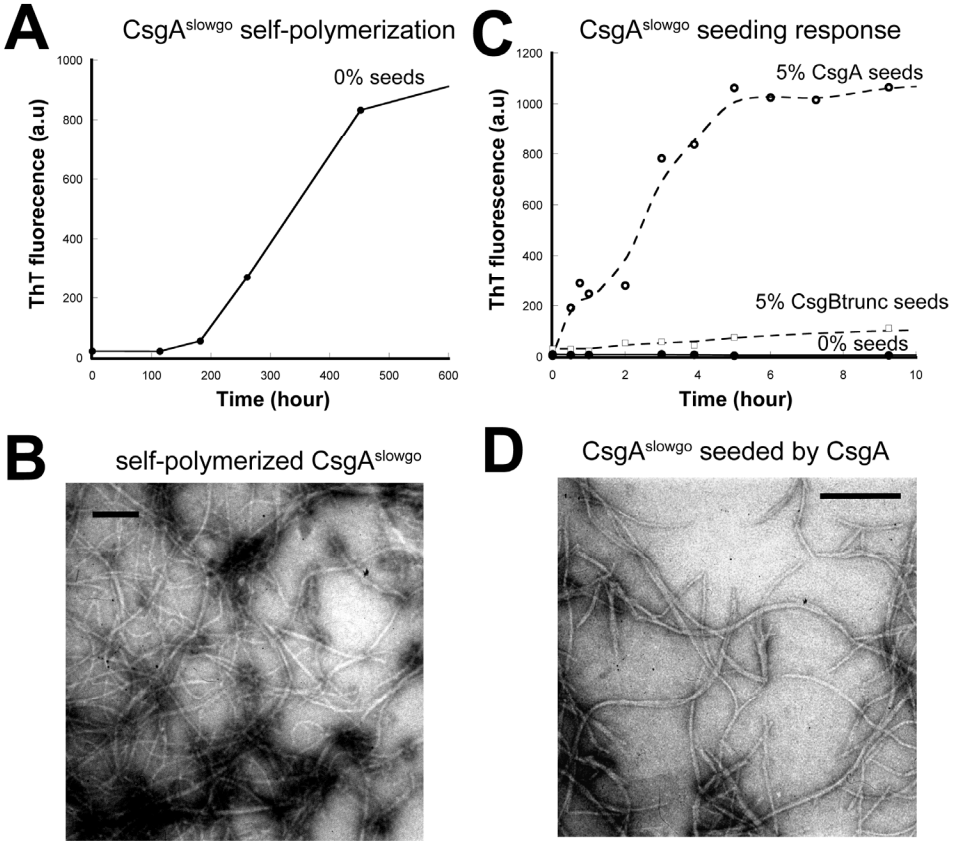
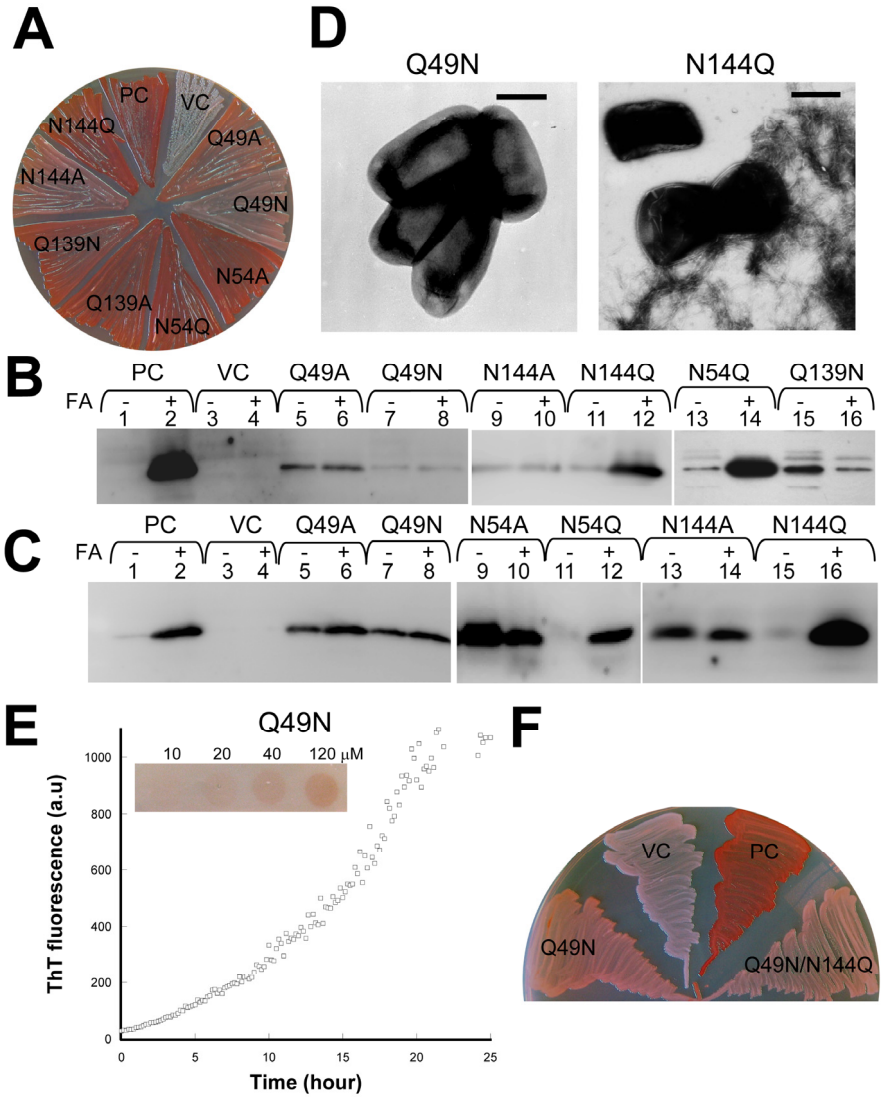


Figure 4.8. Conservative Q/N substitutions at positions 49, 54, 139 and 144. (A) CR-YESCA plate with *csgA* transformed with the plasmids encoding CsgA, $CsgA^{Q49A}$, $CsgA^{Q49N}$, $CsgA^{N54A}$, $CsgA^{N54Q}$, $CsgA^{Q139A}$, $CsgA^{Q139N}$, $CsgA^{N144A}$ or $CsgA^{N144Q}$. (B) Western blots of whole-cell lysates from *csgA* cells containing plasmids encoding wild-type CsgA, $CsgA^{Q49A}$, $CsgA^{Q49N}$, $CsgA^{N54Q}$, $CsgA^{Q139N}$, $CsgA^{N144A}$ or $CsgA^{N144Q}$. (C) Western blot of whole-cells and underlying agar (agar plugs) from *csgA* strains containing plasmids encoding CsgA, $CsgA^{Q49A}$, $CsgA^{Q49N}$, $CsgA^{N54A}$, $CsgA^{N54Q}$, $CsgA^{N144A}$ or $CsgA^{N144Q}$. Samples were treated with (+) or without (-) FA. The blots were probed with anti-CsgA antibody. (D) Negative-stain TEM micrographs of *csgA* cells containing plasmids encoding $CsgA^{Q49N}$ and $CsgA^{N144Q}$. Scale bars are equal to 500 nm. (E) The polymerization of 56 μM $CsgA^{Q49N}$ was measured by ThT fluorescence. Congo red staining of *csgA* cells overlaid with $CsgA^{Q49N}$ at various concentrations is shown in the graph inset. (F) CR-YESCA plate of a *csgA* strain transformed with the control vector pLR2 or the plasmids encoding CsgA, $CsgA^{Q49N}$ or $CsgA^{Q49N/N144Q}$.



References

Azriel, R., and Gazit, E. (2001). Analysis of the minimal amyloid-forming fragment of the islet amyloid polypeptide. An experimental support for the key role of the phenylalanine residue in amyloid formation. J Biol Chem *276*, 34156-34161.

Barnhart, M. M., and Chapman, M. R. (2006). Curli biogenesis and function. Annu Rev Microbiol *60*, 131-147.

Chapman, M. R., Robinson, L. S., Pinkner, J. S., Roth, R., Heuser, J., Hammar, M., Normark, S., and Hultgren, S. J. (2002). Role of *Escherichia coli* curli operons in directing amyloid fiber formation. Science *295*, 851-855.

Chen, S., Ferrone, F. A., and Wetzel, R. (2002). Huntington's disease age-of-onset linked to polyglutamine aggregation nucleation. Proc Natl Acad Sci U S A *99*, 11884-11889.

Chiti, F., and Dobson, C. M. (2006). Protein misfolding, functional amyloid, and human disease. Annu Rev Biochem *75*, 333-366.

Chiti, F., Stefani, M., Taddei, N., Ramponi, G., and Dobson, C. M. (2003). Rationalization of the effects of mutations on peptide and protein aggregation rates. Nature *424*, 805-808.

Collinson, S. K., Emody, L., Muller, K. H., Trust, T. J., and Kay, W. W. (1991). Purification and characterization of thin, aggregative fimbriae from Salmonella enteritidis. J Bacteriol *173*, 4773-4781.

Collinson, S. K., Parker, J. M., Hodges, R. S., and Kay, W. W. (1999). Structural predictions of AgfA, the insoluble fimbrial subunit of Salmonella thin aggregative fimbriae. J Mol Biol *290*, 741-756.

DePace, A. H., Santoso, A., Hillner, P., and Weissman, J. S. (1998). A critical role for amino-terminal glutamine/asparagine repeats in the formation and propagation of a yeast prion. Cell *93*, 1241-1252.

Fowler, D. M., Koulov, A. V., Balch, W. E., and Kelly, J. W. (2007). Functional amyloid--from bacteria to humans. Trends Biochem Sci *32*, 217-224.

Gazit, E. (2002). A possible role for pi-stacking in the self-assembly of amyloid fibrils. Faseb J *16*, 77-83.

Hammar, M., Bian, Z., and Normark, S. (1996). Nucleator-dependent intercellular assembly of adhesive curli organelles in *Escherichia coli*. Proc Natl Acad Sci U S A *93*, 6562-6566.

Hammer, N. D., Schmidt, J. C., and Chapman, M. R. (2007). The curli nucleator protein, CsgB, contains an amyloidogenic domain that directs CsgA polymerization. Proc Natl Acad Sci U S A. *104*, 12494-12499.

Hammer, N. D., Wang, X., McGuffie, B. A., and Chapman, M. R. (2008). Amyloids: Friend or Foe? J Alzheimers Dis (in press).

Hardy, J., and Selkoe, D. J. (2002). The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. Science *297*, 353-356.

Kim, W., and Hecht, M. H. (2006). Generic hydrophobic residues are sufficient to promote aggregation of the Alzheimer's Abeta42 peptide. Proc Natl Acad Sci U S A *103*, 15824-15829.

Liu, J. J., and Lindquist, S. (1999). Oligopeptide-repeat expansions modulate 'protein-only' inheritance in yeast. Nature *400*, 573-576.

Luhrs, T., Ritter, C., Adrian, M., Riek-Loher, D., Bohrmann, B., Dobeli, H., Schubert, D., and Riek, R. (2005). 3D structure of Alzheimer's amyloid-beta(1-42) fibrils. Proc Natl Acad Sci U S A *102*, 17342-17347.

Michelitsch, M. D., and Weissman, J. S. (2000). A census of glutamine/asparagine-rich regions: implications for their conserved function and the prediction of novel prions. Proc Natl Acad Sci U S A *97*, 11910-11915.

Osherovich, L. Z., Cox, B. S., Tuite, M. F., and Weissman, J. S. (2004). Dissection and design of yeast prions. PLoS Biol *2*, E86.

Petkova, A. T., Ishii, Y., Balbach, J. J., Antzutkin, O. N., Leapman, R. D., Delaglio, F., and Tycko, R. (2002). A structural model for Alzheimer's beta -amyloid fibrils based on experimental constraints from solid state NMR. Proc Natl Acad Sci U S A *99*, 16742-16747.

Robinson, L. S., Ashman, E. M., Hultgren, S. J., and Chapman, M. R. (2006). Secretion of curli fibre subunits is mediated by the outer membrane-localized CsgG protein. Mol Microbiol *59*, 870-881.

Ross, E. D., Baxa, U., and Wickner, R. B. (2004). Scrambled prion domains form prions and amyloid. Mol Cell Biol *24*, 7206-7213.

Ross, E. D., Edskes, H. K., Terry, M. J., and Wickner, R. B. (2005). Primary sequence independence for prion formation. Proc Natl Acad Sci U S A *102*, 12825-12830.

Sondheimer, N., and Lindquist, S. (2000). Rnq1: an epigenetic modifier of protein function in yeast. Mol Cell *5*, 163-172.

Taylor, K. L., Cheng, N., Williams, R. W., Steven, A. C., and Wickner, R. B. (1999). Prion domain initiation of amyloid formation *in vitro* from native Ure2p. Science *283*, 1339-1343.

Tessier, P. M., and Lindquist, S. (2007). Prion recognition elements govern nucleation, strain specificity and species barriers. Nature *447*, 556-561.

Wang, X., Hammer, N. D., and Chapman, M. R. (2008). The molecular basis of bacterial amyloid polymerization and nucleation. J Biol Chem, In presee

Wang, X., Smith, D. R., Jones, J. W., and Chapman, M. R. (2007). *In vitro* polymerization of a functional *Escherichia coli* amyloid protein. J Biol Chem *282*, 3713-3719.

Chapter 5

Gatekeeper Residues Modulate Bacterial Amyloidogenesis

Abstract

Amyloid fibers are filamentous protein structures commonly associated with neurodegenerative diseases. Unlike disease-associated amyloids that are the products of protein misfolding, *Escherichia coli* assembles membrane anchored extracellular functional amyloid fibers called curli using an evolved biogenesis machine. Functional amyloids provide a template for understanding controlled amyloid formation. Curli fibers are composed of two proteins, CsgA and CsgB. *In vivo*, the polymerization of the major curli subunit protein, CsgA, is dependent on CsgB-mediated nucleation. The amyloid core of CsgA is composed of 5 imperfect repeats (R1 to R5). I previously showed that R1 and R5 govern the ability of CsgA to respond to CsgB nucleation and self-seeding. I further characterized the determinants of this specificity in this chapter and found certain aspartate and glycine residues in R2, R3 and R4 function as "gatekeepers" of polymerization by inhibiting their intrinsic aggregation propensities and responsiveness to nucleation. Mutations of the gatekeeper residues in R2, R3 and R4 to the corresponding amino acids found in either R1 or R5 rendered these repeats similar to R1 and R5. Mutations of all gatekeeper residues of CsgA resulted in uncontrolled amyloid propagation even in the absence of the nucleation machinery. In addition, the ectopic induction of CsgA with mutated gatekeeper residues was cytotoxic, suggesting that miss-regulation of CsgA polymerization is detrimental to the cell. This is the first

example how nature uses the gatekeeper residues to exquisitely control functional amyloid propagation.

Introduction

Amyloids are ordered proteinaceous filaments, which are associated with many mammalian neurodegenerative diseases and prion-based encephalopathies (Chiti and Dobson, 2006). Amyloid fibers have distinct biochemical and biophysical properties such as great stability and specific tinctorial properties when bound to Congo red and thioflavin T (Chiti and Dobson, 2006). The molecular mechanism of neurodegenerative disease development induced by amyloid propagation remains elusive, partially due to their erratic and uncontrolled nature. Recently a growing body of research has shown that amyloids can also be integral part of physiology found in different organisms including bacteria, fungi and mammals (Fowler et al., 2007; Hammer et al., 2008). How nature coordinates functional amyloid propagation and reduces the associated cytotoxicity is poorly understood.

Curli, a bacterially produced amyloid, is an important component of extracellular matrix and is involved in bacterial community behaviors (Barnhart and Chapman, 2006). Due to the amyloid properties of curli fibers, the colonies of curli-producing *E. coli* stain red when grown on Congo red indicator plates, which provides a convenient assay to monitor curli assembly *in vivo*. In *E. coli* at least six proteins are dedicated to assembling curli fibers at correct time and space. Curli fibers are composed of a major subunit CsgA and a minor subunit CsgB. CsgA remains unpolymerized until it encounters the surface-tethered nucleator CsgB, which initiates CsgA polymerization (Hammar et al.,

1996). CsgD is a transcriptional activator for the *csgBA* operon (Barnhart and Chapman, 2006). CsgG, CsgE and CsgF are non-fiber structural accessory proteins involved in subunit stability, secretion of the fiber subunits and modulation of fiber assembly. CsgG is proposed to be the curli secretion apparatus that directs the exportation of CsgA and CsgB across the outer membrane (Robinson et al., 2006). CsgE and CsgF interact with CsgG at the outer membrane (Robinson et al., 2006). CsgF is required for efficient CsgB-mediated nucleation and CsgE is critical for CsgA and CsgB stability (Chapman et al., 2002; Robinson et al., 2006).

The CsgA primary amino acid sequence comprises a Sec signal peptide (positions 1-20), 22 residues (positions 21-42) dedicated to interacting with CsgG secretion apparatus and an amyloid core region composed of five imperfect repeats (positions 43-151), each 19-23 amino acids (Figure 5.1A). Sec signal sequence is cleaved when CsgA reaches the periplasmic space (Collinson et al., 1999). The repeats (R1, R2, R3, R4 and R5) share at least 30% identity at the amino acid level distinguished by the consensus sequence Ser-X_5-Gln-X_4-Asn-X_5-Gln. I previously showed that R1 and R5 direct CsgA to respond to CsgB-mediated nucleation and are critical for fiber elongation (Wang et al., 2008). Moreover, internally conserved Gln and Asn residues located in R1 and R5 are required for CsgB nucleation and curli assembly (Wang and Chapman, 2008). Because all repeats have these critical Gln and Asn residues, the specific determinants for why R1 and R5 are distinct from other repeats still remain elusive. Here, we elucidate these determinants and found certain Asp and Gly residues in R2, R3 and R4 function as gatekeeper residues inhibiting their amyloidogenic properties.

Gatekeeper residues are defined as residues in many globular proteins to cap the

aggregation-prone sequences and thereby promote the proteins to fold into their native conformations (Monsellier and Chiti, 2007; Otzen et al., 2000; Richardson and Richardson, 2002; Rousseau et al., 2006). Herein, I explore the role of gatekeeper residues in functional amyloidogenesis. Instead of simply opposing aggregation, I found that gatekeeper residues actually play a positive role in harnessing the aggregation of CsgA. A CsgA mutant lacking all gatekeeper residues efficiently polymerized into amyloid fibers independently of the other curli assembly proteins CsgB and CsgF. In addition, the induced expression of CsgA without gatekeeper residues dramatically decreased the viability of *E. coli* cells.

Experimental Procedures

Bacterial growth

Bacteria were grown at 26 °C on YESCA plates to induce curli production (Chapman et al., 2002). When needed, antibiotics were added to the plates in the following concentrations, kanamycin 50 μg/ml, chloramphenicol 25 μg/ml, or ampicillin 100 μg/ml. Curli production was monitored using Congo red-YESCA plates (Chapman et al., 2002).

Bacterial strains and plasmids

Strains and plasmids used in this study are listed in Table 6.1. Primer sequences used in this study are listed in Table 6.2. To study CsgA mutants for *in vivo* polymerization, PCR products containing CsgA mutations and NcoI and BamHI restriction endonuclease sites at 5' and 3' ends were obtained using standard overlapping PCR extension. PCR products were cloned into NcoI/BamHI sites of vector pLR2 that is

downstream of *csgBA* promoter. To purify CsgA and CsgA* from cell pellets, constructs pNH11 and pNH12 encoding His_6-tagged CsgA and CsgA* without periplasmic Sec signal sequence were generated by standard PCR extension. The PCR product containing NcoI and BamHI sites at 5' and 3' ends was cloned NcoI/BamHI sites into pET11d downstream to T7 promoter. The empty vector pNH3, pMC3 and pCsgAslowgoTRC containing C-terminal His_6-tagged full-length CsgA or $CsgA^{slowgo}$ sequence under the control of TRIC promoter were generated previously (Wang and Chapman, 2008; Wang et al., 2008). pCsgA*TRC was constructed by replacing His_6-tagged CsgA with His_6-tagged CsgA* in pMC3.

Purification of CsgA* and CsgA

C-terminal His_6 tagged CsgA* and CsgA without the Sec periplasmic signal sequence were purified from cell pellets. NEB C2566 cells harboring a pET11d vector encoding His-tagged CsgA (pNH11) or CsgA* (pNH12) were grown to OD_{600} 0.9 in LB broth containing 100 μg/ml ampicillin. The expression of CsgA or CsgA* was induced with 0.5 mM IPTG and induction proceeded at 37 °C for an hour. Cells were collected by centrifugation and the pellets were stored at -80 °C. The cells were resuspended and lysed by extraction solution (8 M GdnHCl, 50 mM K_2HPO_4/KH_2PO_4, pH 7.2). A total 50 ml of extraction solution was used for cell pellets from a 500 ml culture. The lysate was incubated at 4 °C with magnetic stirring for 48 hrs. The insoluble portion of the lysate was removed by centrifuging at 10,000 g and the supernatant was incubated with 1.2 ml HIS-select™ HF nickel-nitrilotriacetic acid resin (NiNTA) (Sigma, St. Louis, MO) for 1 hour at room temperature. NiNTA beads that bound His-tagged CsgA or CsgA* were loaded on the column and lysis supernatant passed through the column. GdnHCl in

NiNTA was washed away by 12 ml potassium phosphate buffer (50 mM K_2HPO_4/KH_2PO_4, pH 7.2). CsgA or CsgA* was eluted from the column by 0.5 M imidazole 50 mM potassium phosphate buffer pH7.2. Purified proteins were passed through a 30 kDa cutoff filter (Amicon® Ultra, MA) to remove possible aggregates and seeds that might alter the polymerization kinetics. His tagged CsgA and CsgA* showed no difference to CsgA and CsgA* without a His tag in curli assembly *in vivo* (data not shown). Therefore, CsgA*-His and CsgA-His were referred to as "CsgA*" and "CsgA" for *in vitro* polymerization study throughout this report.

Preparation of peptides

Peptides R3, R4 and derivative peptides $R3^{D4N/D17L}$ and $R4^{G13N/D17H}$ were synthesized by Proteintech Group Inc., Chicago, IL. Peptides were prepared as in Chapter 2 and 3.

Thioflavin T assay

Thioflavin T assays were performed and normalized as previously described (Wang et al., 2008). Seeds were prepared as previously described in Chapter 2.

Transmission electron microscopy

Philips CM12 transmission electron microscope was used to visualize the cells and fiber aggregates. Samples were placed on formvar coated copper grids (Ernest F. Fullam, Inc, Latham, NY) for 2 min, washed with deionized water twice, and stained with 2% uranyl acetate for 90 seconds.

SDS-PAGE and western blotting

Bacteria whole cell lysates or cells and underlying agar (plugs) were prepared and probed for CsgA using previously described methods (Chapman et al., 2002).

Overlay assay

Purified proteins were spotted on the *csgA* cells and stained with 0.5 mg/ml Congo red solution as previously described in Chapter 3 and 4.

Spotting assay for cell viability

Freshly transformed cells with plasmids pMC3, empty vector pNH3, pCsgAslowgoTRC or pCsgA*TRC were scraped off the plates and normalized by absorbance at 600 nm. Cells with serial dilutions were spotted on LB chloramphenicol plates with various concentrations of IPTG to induce expression of CsgA, CsgA^{slowgo} or CsgA*. After overnight growth, colony number was manually counted and colony forming units per ml cell suspension under different IPTG concentrations were calculated.

Results

Seeding specificity is encoded in R1 and R5

I previously showed that R1 and R5 are the domains responsible for both CsgA-CsgB interaction, defined as nucleaton, and CsgA-CsgA interaction, defined as seeding (Wang et al., 2008). In support of the hypothesis that only R1 and R5 are the key seeding elements of CsgA and are important for fiber elongation, 8% (w/w) each repeating unit peptide was prepared as seeds and added to freshly purified CsgA and the polymerization was measured by ThT fluorescence over time as previously described in Chapter 3 (Wang et al., 2007). Before being used as seeds, R1, R3 and R5 peptides were polymerized into fibers as described (Wang et al., 2007). The lag phase of CsgA polymerization was completely eliminated when preformed R1 or R5 fibers were added

to the reaction (Figure 5.1B), while R2, R3 and R4 did not accelerate CsgA polymerization (Figure 5.1B). 1% or 2% R1 or R5 seeds were sufficient to accelerate CsgA polymerization as detected by ThT assay (data not shown). A nucleation-competent CsgB truncation mutant was purified as previously described (Hammer et al., 2007), and its nucleation effect on peptides were tested. Only peptide R1 and R5 polymerization was accelerated in the presence of 3% (w/w) $CsgB_{trunc}$ (Figure 5.1C and Figure 5.2). We also tested the seeding specificity among the repeating unit peptides. R1 and R5 seeded both each other and themselves, while other repeats were deficient for both seeding and nucleation properties (Figure 5.1C and Figure 5.3). These results further support the idea that R1 and R5 govern the seeding and nucleation properties in CsgA.

R1 and R5 are interchangeable *in vivo*

R1 and R5 are somewhat functionally redundant: both seed CsgA polymerization and constitute nucleation responsive domains to CsgB and CsgA. To test the hypothesis that R1 and R5 are functionally redundant, the residues of R1 of CsgA (R12345) were replaced by the corresponding residues from R5 to form the mutant construct with the following repeating unit order: R52345 (Table 6.1). Despite the fact that the R52345 molecule does not contain R1, when expressed from the plasmid, pR52345, it was able to complement a csgA mutant and assembles into fibers on the cell surface (Figure 5.1D). Similarly, a CsgA analogue in which R5 was replaced by R1 (R12341) was also able to assemble fibers on the cell surface (Figure 5.1D). Wild-type curli fibers are SDS insoluble and formic acid (FA) treatment is required to dissociate the monomers from the fibers (Collinson et al., 1991). Both R52345 and R12341 formed SDS insoluble fibers similar to those formed by wild-type CsgA (Figure 5.1E lanes 1, 2, 5, 6, 7 and 8),

suggesting that R1 and R5 are functionally redundant *in vivo*. Just like CsgA, the *in vivo* polymerization of R52345 and R12341 is dependent on CsgB (data not shown). R3 is also an amyloidogenic domain in CsgA and the sequence similarity (70%) and identity (48%) between R3 and R5 are higher than those between R1 and R5 (52% similarity and 39% identity). We asked whether R3 can substitute for R1 or R5 *in vivo*. R32345 was somewhat defective *in vivo* as fewer fibers were observed by TEM from csgA/pR32345 compared to *csgA*/pR12341 or csgA/pR52345 (Figure 5.1D). Furthermore R32345 remained SDS soluble and did not require FA treatment to migrate into the gel (Figure 5.1E lanes 9 and 10). The phenotype of R12343 was even more dramatic. R12343 was not detected by western blotting of the cell pellets but instead was only detected when both cells and the underlying agar were collected, suggesting R12343 was stably secreted and completely defective in curli assembly (Figure 5.1E lanes 11 and 12). Nearly no fibers were produced by csgA/pR12343 when analyzed by TEM (Figure 5.1D). Therefore, R3 was not able to substitute for R5 at the C-terminus, and was partially capable of replacing R1 at the N-terminus, suggesting the nucleation/seeding response and specificity is encoded only by R1 and R5 and not by R3 even though R5 is more similar to R3 than R1 at the amino acid level.

Identification of gatekeeper residues of R3 that inhibit amyloidogenic properties

My results indicate that common residues found in both R1 and R5 possibly are involved in the observed nucleation response and seeding specificity. Alternatively, R3 should contain specific differences compared to R1 and R5 that make it less amyloidogenic than R1 and R5. Guided by this perspective, we compared R3 to R1/R5, looking for residues in R3 that were distinctive from both R1 and R5. Four residues were

targeted as candidates: Asp^{136}, Arg^{140}, Asp^{149} and Trp^{151} in the molecule R12343 (Figure 5.4A). If these residues hinder the ability of R3 to successfully replace R5, then conversion of R12343 to wild-type CsgA in curli assembly could be accomplished by changing these residues to those found in either R1 or R5. Constructs pXW56 ($R12343^{D136N}$), pXW57 ($R12343^{R140Y}$), pXW58 ($R12343^{R140V}$), pXW59 ($R12343^{D149L}$), pXW60 ($R12343^{D149H}$), pXW61 ($R12343^{W151T}$) and pXW62 ($R12343^{W151Y}$) were engineered to test this hypothesis (Table 6.1) (Figure 5.4A). Strikingly, $R12343^{D136N}$, $R12343^{D149L}$and $R12343^{D149H}$ in *csgA* cells assembled into amyloid-like fibers shown by TEM and Congo red binding (Figure 5.4B and data not shown), while other mutant proteins such as $R12343^{R140Y}$, $R12343^{R140V}$, $R12343^{W151T}$ and 12343^{W151Y} were as defective as R12343 in fiber formation *in vivo* as measured by TEM and Congo red binding (data not shown). There were also more SDS insoluble $R12343^{D136N}$, $R12343^{D149L}$ and $R12343^{D149H}$ than SDS soluble forms detected by whole-cell western analysis, while other mutants such as $R12343^{R140Y}$, $R12343^{R140V}$, $R12343^{W151T}$ and 12343^{W151Y} remained similar to R12343 (Figure 5.4C). *In vitro* polymerization of $R12343^{D149H}$ was dramatically faster than that of R12343 as measured by previously described ThT assay (Figure 5.4D) (Wang et al., 2007).

An overlay assay was previously constructed in Chapter 3 to test the response of CsgA and the CsgA mutant proteins to CsgB-mediated nucleation (Wang et al., 2008). In a CsgB-dependent manner, soluble freshly purified CsgA was converted into amyloid fibers on the *csgA* cells expressing CsgB as evidenced by TEM or Congo red staining (Wang et al., 2008). When $R12343^{D149H}$ and R12343 at 10 μM were spotted on *csgA* cells, $R12343^{D149H}$ assembled into fibers more readily than R12343 as indicated by stronger

Congo red staining (Figure 5.4D, inset). The Asn^{136} and His^{149} residues of CsgA are not conserved among CsgA homologues from different enteric bacteria (Figure 5.4E) and the His^{149} residue of $R12343^{D149H}$ can be replaced by Arg, Lys or Glu without interfering with curli assembly (data not shown). These results suggest Asn^{136} and His^{149} residues are not essential for fiber formation while Asp residues at positions 136 and 149 in the molecule R12343 specifically impede amyloidogenic properties of R3. Therefore, we designated Asp residues at positions 91 and 104 in R3 gatekeeper residues because they disrupt amyloidogenic properties of R3. The R3 derivative peptide with these Asp residues changed to R1-like residues (D4N/D17L) polymerized much faster than R3 peptide, supporting the idea that these Asp residues inhibit the intrinsic aggregation of R3 (Figure 5.5).

Identification of gatekeeper residues in R2 and R4

Peptides R2 and R4 have poor amyloidogenic properties and cannot self-assemble into amyloid fibers efficiently *in vitro* (Wang et al., 2007). It is possible that R2 and R4 also contain gatekeeper residues as in R3 that disrupt their amyloidogenic properties. Both R1 and R5 are required for curli assembly and deletion of either R1 or R5 or both result in nearly complete loss of fiber assembly *in vivo* (Wang et al., 2008). R2345 (ΔR1) and R1234 (ΔR5) were defective in curli assembly presumably because R2 and R4 cannot function like R1 or R5. If the potential gatekeeper residues of R2 or R4 were mutated, R2345 or R1234 could be capable of assembling amyloid fibers. Guided by this perspective, we compared R4 to R1/R5, looking for residues in R4 that were different from both R1 and R5. Four residues were targeted as candidates: Thr^{114}, Arg^{116}, Gly^{123} and Asp^{127} in the molecule R1234 (Figure 5.6A). When Gly^{123} or Asp^{127} or both were

mutated to the corresponding residues from R5 (Asn and His, respectively), the *in vivo* amyloidogenesis of R1234 was significantly improved as detected by Congo red binding and TEM, suggesting $Gly1^{23}$ and Asp^{127} are gatekeeper residues in R4 (Figure 5.6). Similarly, after Gly^{55}, Asp^{57} and Gly^{59} were mutated to the corresponding residues from R1 (Ser, Leu and Leu, respectively), mutated R2345 could efficiently assemble into fibers as detected by Congo red binding and TEM (Figure 5.7). To test whether these gatekeeper residues impede self-polymerization of R2 or R4, we tested self-polymerization of the R4 derivative peptide with mutations of gatekeeper residues ($R4^{G13N/D17H}$). When the gatekeeper residues were mutated to the corresponding residues from R5, $R4^{G13N/D17H}$ polymerized into amyloid fibers much more efficiently compared to original R4 peptide as measured by ThT fluorescence and TEM (Figure 5.8 and data not shown).

Gatekeeper residues are not tolerated in R5

The five repeating units of CsgA have the consensus sequence Ser-X_5-Gln-X_4-Asn-X_5-Gln. Seven gatekeeper residues (Gly^{78}, Asp^{80}, Gly^{82}, Asp^{91}, Asp^{104}, Gly^{123} and Asp^{127} for *E. coli* CsgA) found in R2, R3 and R4 were located in the 5 residues stretches between Ser and Gln or Asn and Gln (Figure 5.4E). R1 and R5 of CsgA homologues from different enteric bacteria do not have any Asp or Gly residues in the same regions (Figure 5.4E). This observation led us to hypothesize that the existence of gatekeeper residues discriminates R2, R3 and R4 from R1 and R5. We introduced Asp residues into R5 at positions 136 and 149 of wild-type CsgA and tested the amyloid formation properties of this mutant ($CsgA^{N136D/H149D}$) *in vivo*. These positions corresponded to where Asp residues are found in R3. Like R12343, $CsgA^{N136D/H149D}$ did not assemble into amyloid fibers *in vivo* as detected by whole-cell western analysis and

TEM (Figure 5.9).

CsgA lacking all gatekeeper residues polymerized in the absence of CsgB and CsgF *in vivo*

All the CsgA homologues from enteric bacteria have gatekeeper residues in R2, R3 and R4, suggesting the gatekeeper residues play an important role in CsgA assembly. To gain some insight about the roles of gatekeeper residues in bacterial amyloid propagation, we constructed a CsgA mutant with all the gatekeeper residues changed to the corresponding residues from either R1 or R5 ($CsgA^{G78S/D80L/G82L/D91N/D104L/G123N/D127H}$), called CsgA* (Figure 5.10A) and tested its behavior in curli assembly and self-polymerization. CsgB and CsgF are required for efficient curli assembly (Chapman et al., 2002; Hammar et al., 1996). CsgB tethered on the cell surface functions as a nucleator for CsgA polymerization and deletion of CsgF disrupts CsgB-mediated nucleation (Chapman et al., 2002; Hammar et al., 1996). Surprisingly, in the absence of CsgB and CsgF, CsgA*-expressing cells contained a deep Congo red staining pattern suggesting fiber had been assembled (Figure 5.10B). Western blotting of *csgFBA*/pCsgA cells and the underlying agar showed that CsgA remained SDS soluble in the absence of CsgB and CsgF (Figure 5.10C lanes 3 and 4), while CsgA* from *csgFBA*/pCsgA* cells was SDS insoluble (Figure 5.10C lanes 5 and 6). Many of the CsgA* fibers formed in *csgFBA* genetic background were mislocalized and were not tethered with cells as detected by TEM, while no fibers were produced by *csgFBA*/pCsgA cells (Figure 5.10D). In both *csgBA* and *csgBF* genetic backgrounds, CsgA* also formed SDS insoluble fibers and bound Congo red (Figure 5.10B, Figure 5.11 and data not shown). After cells were scraped off the plate, the agar plate stained with Congo red suggesting fibers composed of

CsgA* were mislocalized (data not shown). Like CsgA, the stability of CsgA* was still dependent on CsgG and CsgE. In the absence of CsgG or CsgE, CsgA* was undetectable by western blotting of cells and underlying agar and could not bind Congo red (Figure 5.10B and Figure 5.11). Collectively, secretion of CsgA* is dependent on CsgG and CsgE. However, in contrast to CsgA, the polymerization of CsgA* *in vivo* does not require CsgB and CsgF. It is plausible that these gatekeeper residues render CsgA fiber assembly under the control of CsgB and CsgF.

CsgA* polymerized faster than CsgA *in vitro* and induced expression caused *E. coli* to be less viable

To understand what biochemical properties distinguish CsgA* from CsgA, we tried to purify CsgA and CsgA* and compared their polymerization kinetics *in vitro* using a previously described method (Wang et al., 2007). Wild-type CsgA was found to be secreted to the medium by co-overexpressing CsgA and CsgG (Wang et al., 2007). Overexpression of full length CsgA* resulted in cell lysis and no CsgA* was secreted to medium or associated with cells, which suggested that CsgA* is potentially cytotoxic. We took an alternative approach to purify mature CsgA and CsgA*. The expression vector containing CsgA or CsgA* lacking the Sec-dependent periplasmic signal sequence was transformed into the expression strain NEB C2566. Protein expression was induced by IPTG and was purified from cell pellets resuspended in guanidine hydrochloride (see Materials and Methods). Freshly purified CsgA polymerization follows a triphasic process with a lag, a rapid growth and stationary phase similarly as previously described (Wang et al., 2007). The lag phase for CsgA at a concentration higher than 5 μM is around ~2 hrs as detected by the ThT assay (Figure 5.12A) (Wang et al., 2007). In

contrast to CsgA, CsgA* started polymerization without a clear lag phase and the total time interval for its ThT signal to reach stationary phase is much shorter than that of CsgA (Figure 5.12A). Like CsgA, CsgA* self-assembled aggregates were ordered amyloid-like fibers as evidenced by TEM (data not shown). To carefully study the potential cytotoxicity of CsgA-like molecules, we used spotting assay to measure how cell viability was influenced by the induced expression of full-length CsgA, CsgA* or $CsgA^{slowgo}$ ($CsgA^{Q49A/N54A/Q139A/N144A}$). $CsgA^{slowgo}$ was previously reported to be extremely defective in curli assembly and its self-polymerization has a lag time two orders of magnitude greater than that of CsgA polymerization (Wang and Chapman, 2008). The viability of *csgA* cells did not change significantly when the expression of $CsgA^{slowgo}$ was induced (Figure 5.12B). At relative high IPTG concentration (>50 μM), expression of CsgA resulted in less viable cells compared to either $CsgA^{slowgo}$ or empty vector control (Figure 5.12B). The induced expression of CsgA* with IPTG at 25 μM caused a marked loss of cell viability compared to wild-type CsgA (Figure 5.12B and Figure 5.13) and colony forming units of cells expressing CsgA* were around three orders of magnitude lower than those of cells expressing CsgA at IPTG concentration higher than 25 μM (Figure 5.12B). Expression of CsgA* resulted in decreased viability of various *E. coli* K12 cells such as MC4100, C600 and BW25113 (data not shown).

Discussion

I previously showed that two repeating units of CsgA (the N- and C-terminal repeats) govern the nucleation/seeding response and are critical for fiber assembly (Wang et al., 2008). Data presented here identified the specific residues that inhibit

amyloidogenic properties of the middle repeats of CsgA. We call these residues gatekeeper residues. The gatekeeper residues of CsgA appear to reduce cytotoxicity associated with amyloidogenesis by controlling the polymerization kinetics of CsgA. When these gatekeepers are mutated CsgA polymerization is accelerated *in vitro* and occurs *in vivo* without the need for the nucleator protein, CsgB.

During the disease-associated amyloidogenesis, the rate-limiting step is nucleus formation, the initial development of small, metastable oligomers of a protein (Jarrett and Lansbury, 1993; Selkoe, 2003). For curli assembly, this step is proposed to be overcome in *E. coli via* a template mediated mechanism where the nucleator protein, CsgB, adopts a β-sheet rich secondary structure and propagates the β-sheet rich amyloid fold onto CsgA molecules when it comes in contact with at the cell surface (Hammer et al., 2007). The nucleation of CsgA by CsgB requires the participation of specific conserved Gln and Asn residues in R1 and R5 of CsgA (Wang and Chapman, 2008). The middle repeats, R2, R3 and R4, also have Gln and Asn residues at these conserved positions, yet they cannot respond to CsgB nucleation (Wang et al., 2008). Further, it seems that the nucleation/seeding specificity is not simply correlated to sequence similarity or identity because similarity and identity between R3 and R5 are higher than those between R1 and R5. Using rationally designed mutagenesis, we identified specific Asp and Gly residues required for unresponsiveness to CsgB nucleation and defectiveness of self-assembly (Figure 5.4-8). The changes of these residues render R2, R3 and R4 both efficient self-polymerization and responsiveness to CsgB-mediated nucleation. The intrinsic aggregation properties seem correlated to nucleation response, suggesting the self-polymerization and nucleated polymerization share common features.

Results from several different studies suggest that Pro, Asp and Gly are among the residues least likely to be incorporated into a β-sheet structure (Kallberg et al., 2001; Street and Mayo, 1999). In general, the residues that potently disrupt β-strand formation, such as Pro and Asp, inhibit amyloid formation (Chiti et al., 2003; Soto et al., 1998). It was postulated that charged residues and other β breaker residues in globular proteins could function as "gatekeeper" residues to prevent amyloid formation (Monsellier and Chiti, 2007; Otzen et al., 2000; Thirumalai et al., 2003). Some gatekeeper residues are found at the edge β strand (the peripheral β strand) in the native β sheet conformations to prevent protein aggregation (Richardson and Richardson, 2002). Since CsgA and CsgB are dedicated to form amyloid fibers efficiently, it is not surprising that there is not a single Pro residue in the five repeats of CsgA or CsgB from *E. coli*. What was not expected was the finding of somewhat conserved gatekeeper residues located in the middle repeats of CsgA homologues from different bacteria (Figure 5.4). Gatekeeper residues are not tolerated in end repeats of CsgA which was predicted to be the edge β strands (Figure 5.9) (Collinson et al., 1999). In contrast to gatekeepers residues in the edge β strand of globular proteins that appear to simply oppose aggregation (Richardson and Richardson, 2002), the gatekeeper residues of wild-type CsgA located in the middle repeats of CsgA do not completely inhibit CsgA polymerization, since CsgA still can self-assemble into amyloid fibers (Wang et al., 2007). The self-polymerization of a CsgA mutant lacking all gatekeeper residues (CsgA*) is much faster than CsgA and there is almost no lag time before its rapid polymerization (Fig. 4A). This strong aggregation propensity of CsgA* resulted in its efficient polymerization in the extracellular environment in the absence of CsgB and CsgF (Figure 5.10 and 5.11). Many of the

CsgA* fibers assembled in the absence of CsgB were not associated with cells and thereby presumably may not help cells to attach the surface. It is plausible that gatekeeper residues of CsgA modulate CsgA polymerization efficiency so that its polymerization and localization are under the control of CsgB and CsgF.

It has been previously reported that the efficiency of amyloid formation correlates to cytotoxicity and disease development in various models (Hardy and Selkoe, 2002; Luheshi et al., 2007; Selkoe, 2003; Van Nostrand et al., 2001). Here, CsgA derivatives with distinct polymerization properties showed similar pattern (Figure 5.12B). The induced expression of rapidly-polymerizing CsgA* caused cells less viable compared to wild-type CsgA (Figure 5.12-13). One possibility is that the existence of gatekeeper residues in CsgA decreases the load of aggregated proteins for cellular molecular chaperones and dedicated quality control machinery. Alternatively, the toxic properties of CsgA* may be distinct from wild-type CsgA. Functional amyloid propagation is a balance to maximize polymerization efficiency while minimizing potential cytotoxicity associated with amyloidogenesis. The work presented in this chapter demonstrates that the absence of gatekeeper residues results in the uncontrolled aggregative properties and increased cytotoxicity. Therefore, the primary sequence of wild-type CsgA has been shaped by evolution to fit the requirements of both efficient polymerization and control by dedicated curli assembly factors.

Figures and Tables

Figure 5.1. Seeding/nucleation specificity of CsgA was determined by R1 and R5. (A) The schematic of CsgA primary sequence. (B) 8% indicated seeds were added to CsgA polymerization. (C) Summary of CsgA and peptides seeding/nucleation specificity. Data previously reported were indicated by "*" (Hammer et al., 2007; Wang et al., 2008; Wang et al., 2007). (D) Negative-stain EM micrographs of *csgA* mutant strain containing plasmids pR12341, pR52345, pR32345 or pR12343. Scale bars are equal to 500 nm. (E) Western blots of whole-cell lysates (top panel) or plugs (whole cells and underlying agar) (bottom panel) from *csgA* mutant cells containing plasmids pCsgA (lanes 1 and 2), Vector control (lanes 3 and 4), pR12341 (lanes 5 and 6), pR52345 (lanes 7 and 8), pR32345 (lanes 9 and 10) or pR12343 (lanes 11 and 12) grown for 48 hours at 26°C on YESCA plates. Cells were treated with (+) or without (-) FA prior to electrophoresis as indicated. The blots were probed with anti-CsgA antibody.

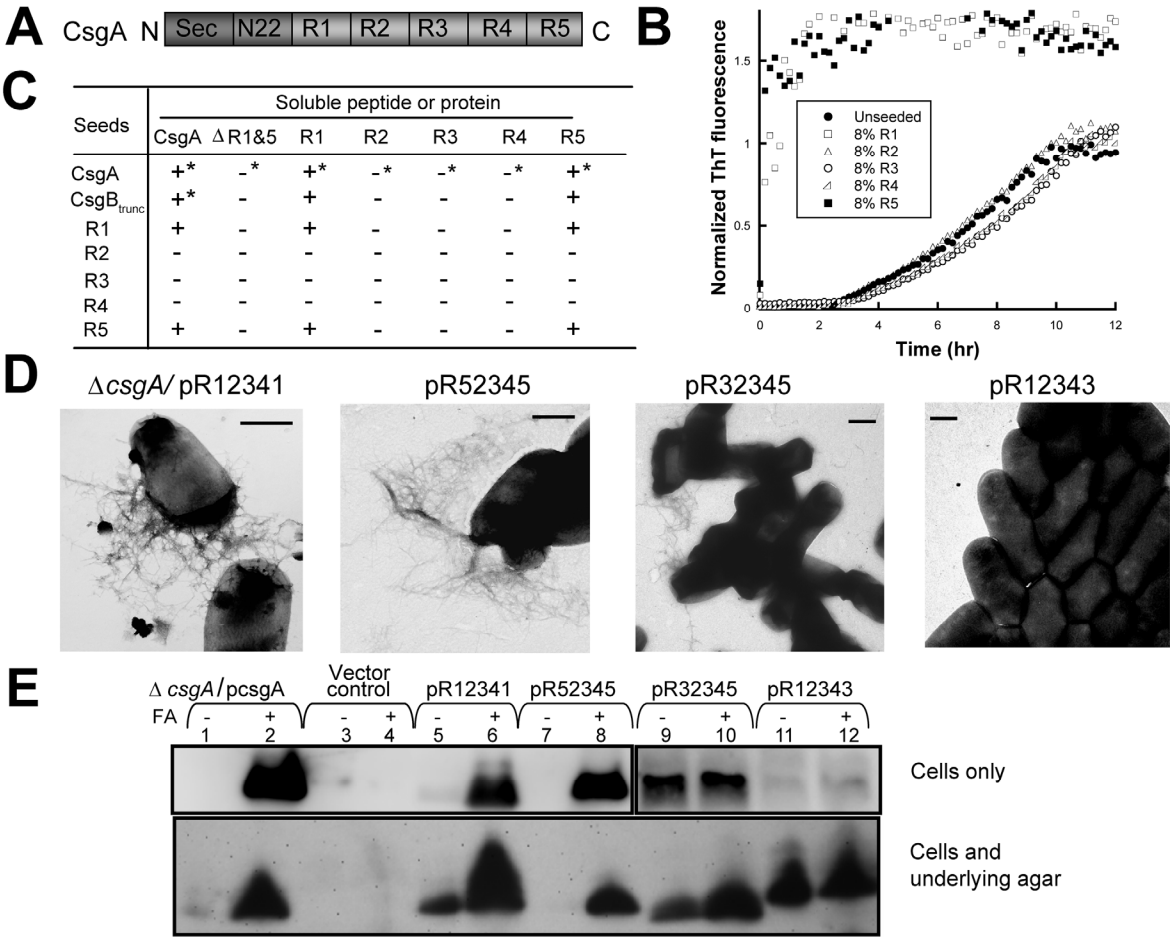


Figure 5.2. $CsgB_{trunc}$ seeds R1 and R5 polymerization *in vitro*. (A) 3% by weight sonicated $CsgB_{trunc}$ fibers or buffer were added to freshly prepared 0.5 mg/ml peptide R1 and mixed with ThT before fluorescence was measured at 438 nm after excitation at 495 nm. (B) 3% by weight sonicated $CsgB_{trunc}$ fibers were added to 0.2 mg/ml R5 in 1.0 M GdnHCl buffered by 50mM potassium phosphate at pH7.2. Samples were mixed with ThT and fluorescence was measured at 438 nm after excitation at 495 nm.

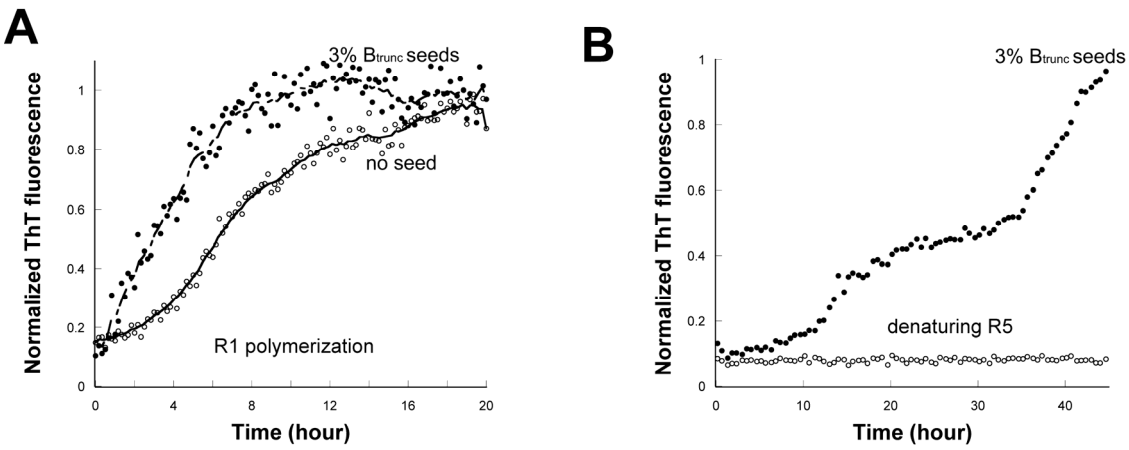


Figure 5.3. Seeding specificity of R1 and R5. (A)-(C) Indicated percentages by weight of R1, R3, R5 sonicated fibers and buffer vehicle were added to freshly prepared 1.0 mg/ml oligopeptide R1 (A), 1.0 mg/ml μM oligopeptide R3 (B) and 17 μM (0.2 mg/ml) ΔR1 which contains repeats R2, R3, R4 and R5 (C). The polymerization was measured by ThT fluorescence.

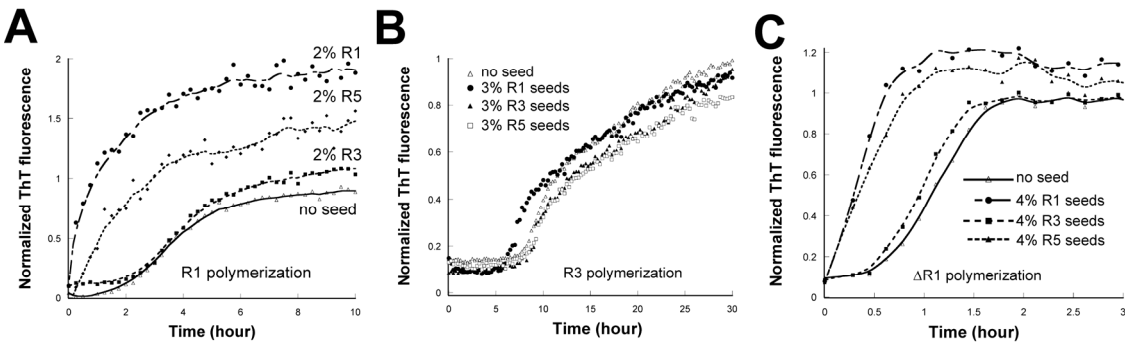


Figure 5.4. Mutations of Asp residues of C-terminal R3 in R12343 rendered amyloidogenic properties. (A) Differences between R3 and R1/R5 amino acid sequence are indicated with boxes. Constructs to mutate C-terminal R3 of R12343 are indicated in the bottom panel. (B) Negative-stain EM micrographs of *csgA* cells containing plasmids pXW56 (R12343^{D136N}), pXW59 (R12343^{D149L}) and pXW60 (R12343^{D149H}). Cells were grown 48 hours at 26°C on YESCA plates. Scale bars are equal to 500 nm. (C) Western blots of whole-cell lysates from *csgA* cells containing constructs pR12343 (lane 1 and 2; lane 17 and 18), pXW56 (R12343^{D136N} lane 15 and 16), pXW57 (R12343^{R140Y} lane 3 and 4), pXW58 (R12343^{R140V} lane 5 and 6), pXW59 (R12343^{D149L} lane 7 and 8), pXW60 (R12343^{D149H} lane 9 and 10), pXW61(R12343^{W151T} lane 11 and 12) and pXW62 (R12343^{W151Y} lane 13 and 14) grown 48 hours at 26°C on YESCA plates. Samples were treated with (+) or without (-) FA as indicated. The blots were probed with anti-CsgA antibody. (D) The polymerization of 20 μM R12343^{D149H} and R12343 measured by ThT fluorescence. Congo red staining of *csgA* (CsgB^{+}) cells overlaid with10 μM and 20 μM R12343^{D149H} and R12343 was shown in graph inset. (E) The alignment of five repeating units of CsgA from different enteric bacteria. The consensus sequences Ser-X_5-Gln and Asn-X_5-Gln indicated in yellow background were predicted to form β-strands (Collinson et al., 1999). The identified gatekeeper residues of CsgA from *E. coli* are indicated with blue boxes. The putative gatekeeper residues of CsgA from other enteric bacteria are indicated with red boxes.

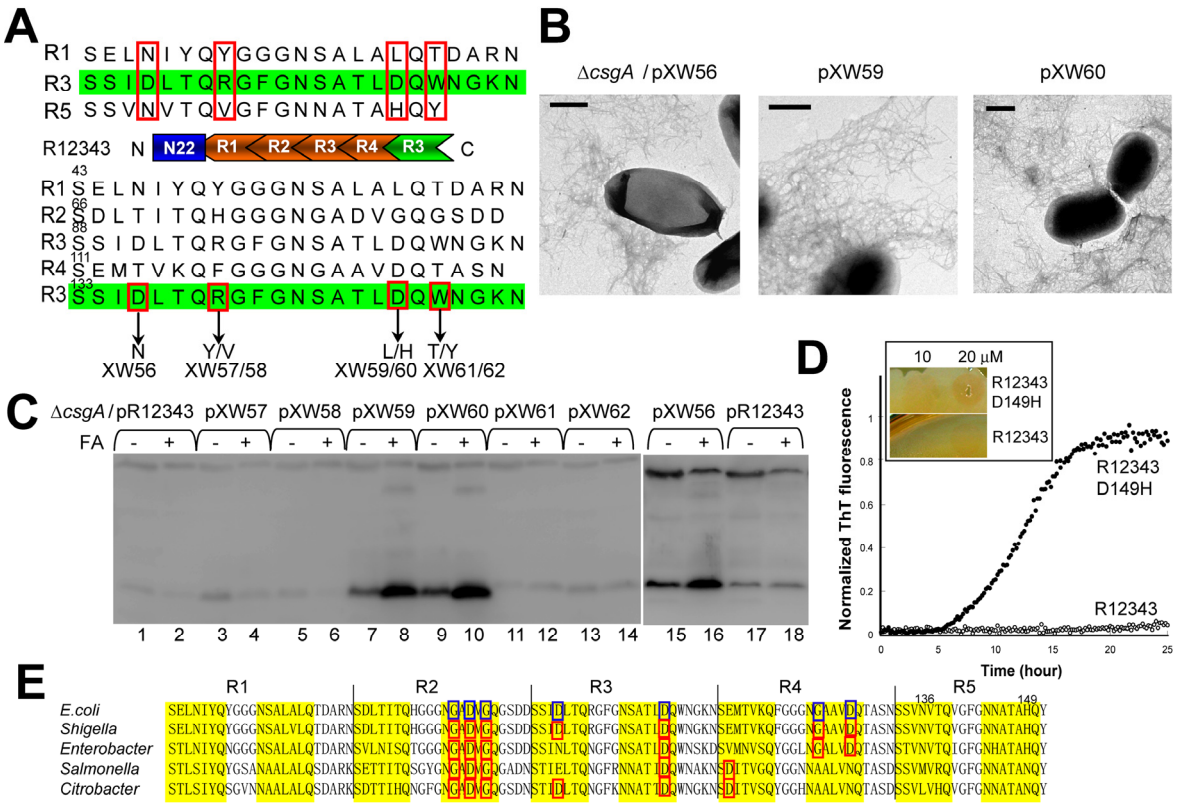


Figure 5.5. The polymerization of peptide R3 and R3$^{D4N/D17L}$. The polymerization of 0.5 mg/ml chemically synthesized peptides R3 and R3 $^{D4N/D17L}$ was measured by ThT fluorescence. The changes of amino acid sequence of derivative peptide R3$^{D4N/D17L}$ from R3 were indicated on the top of the graph.

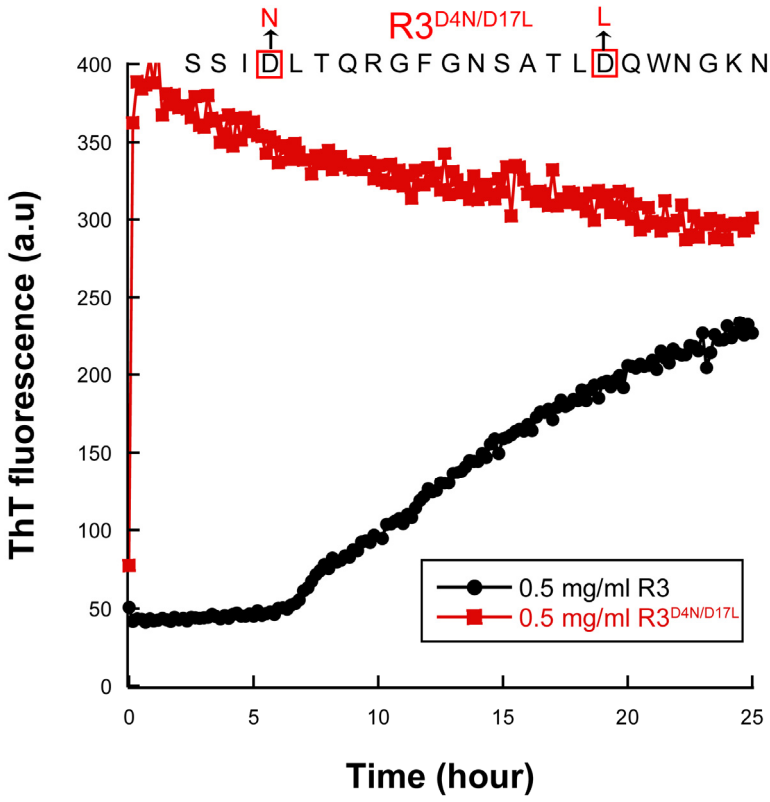


Figure 5.6. Identification of gatekeeper residues of R4. Mutations of Asp or Gly residues of C-terminal R4 in R1234 can restore amyloidogenic properties. (A) Differences between R4 and R1/R5 amino acid sequence are indicated with boxes. Constructs to mutate R4 of R1234 are indicated in the bottom panel. (B) Congo red-YESCA plate with csgA cells transformed with the control vector or plasmids encoding CsgA (pCsgA), R1234, R1234^{T114N}, R1234^{K116T}, R1234^{G123N}, R1234^{D127H} or R1234$^{G123N/D127H}$. (C)-(D) Negative-stain EM micrographs of *csgA* cells containing plasmids pR1234 (C) or pR1234$^{G123N/D127H}$ (D). Cells were grown 48 hours at 26°C on YESCA plates. Scale bars are equal to 500 nm.

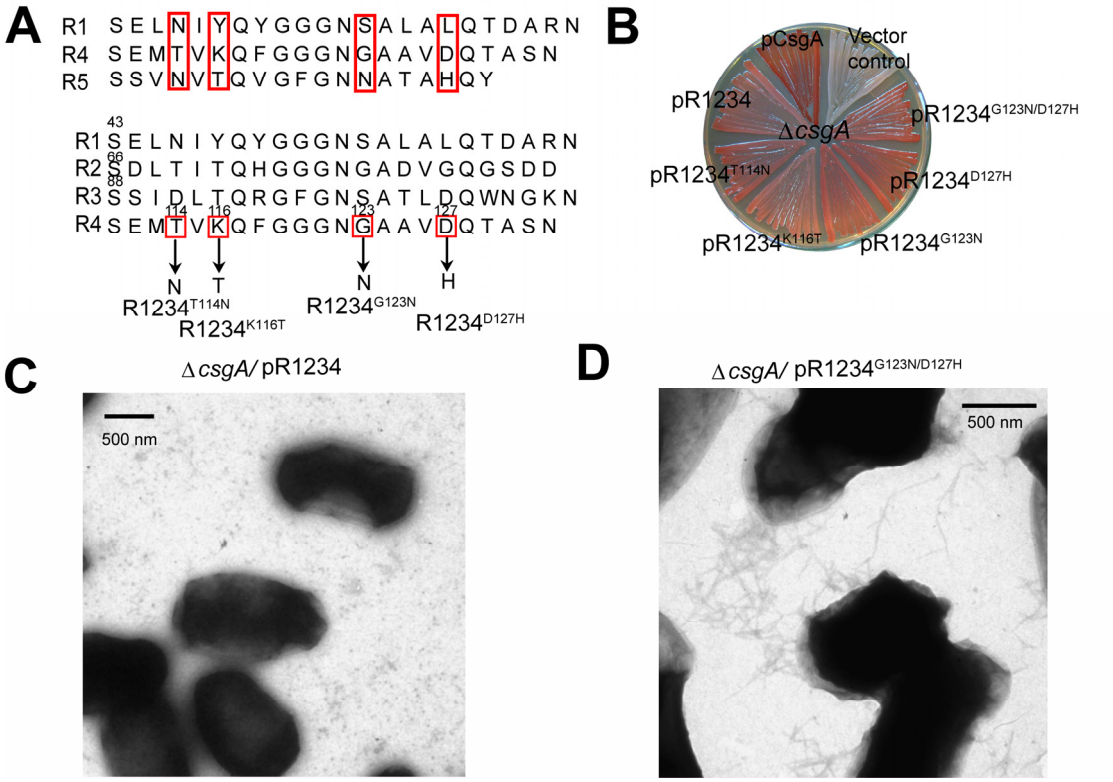


Figure 5.7. Identification of gatekeeper residues of R2. Mutations of Asp and Gly residues of N-terminal R2 in R2345 can restore amyloidogenic properties. (A) Differences between R2 and R1/R5 amino acid sequence are indicated with boxes. Constructs to mutate R2 of R2345 are indicated in the bottom panel. (B) Congo red YESCA plate with csgA cells transformed with the control vector or plasmids encoding CsgA, pR2345, pR2345^{D44E} or pR2345$^{G55S/D57L/G59L}$. (C)-(D) Negative-stain EM micrographs of *csgA* cells containing plasmids pR2345 (C) or pR2345$^{G55S/D57L/G59L}$ (D). Cells were grown 48 hours at 26°C on YESCA plates. Scale bars are equal to 500 nm.

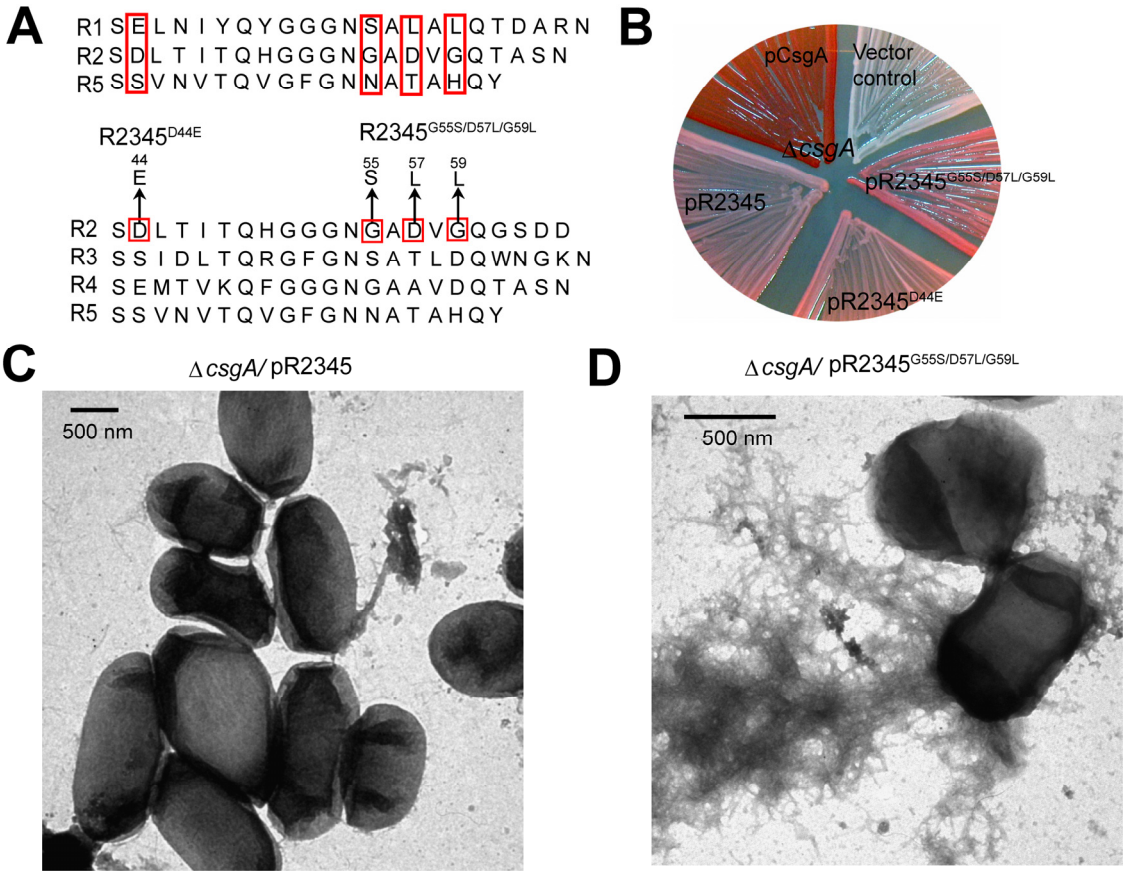


Figure 5.8. The polymerization of peptide R4 and R4$^{G13N/D17H}$. The polymerization of 0.2 mg/ml chemically synthesized peptides R4 and R4$^{G13N/D17H}$ was measured by ThT fluorescence. The changes of amino acid sequence of derivative peptide R4 $^{G13N/D17H}$ from R4 were indicated on the top of the graph.

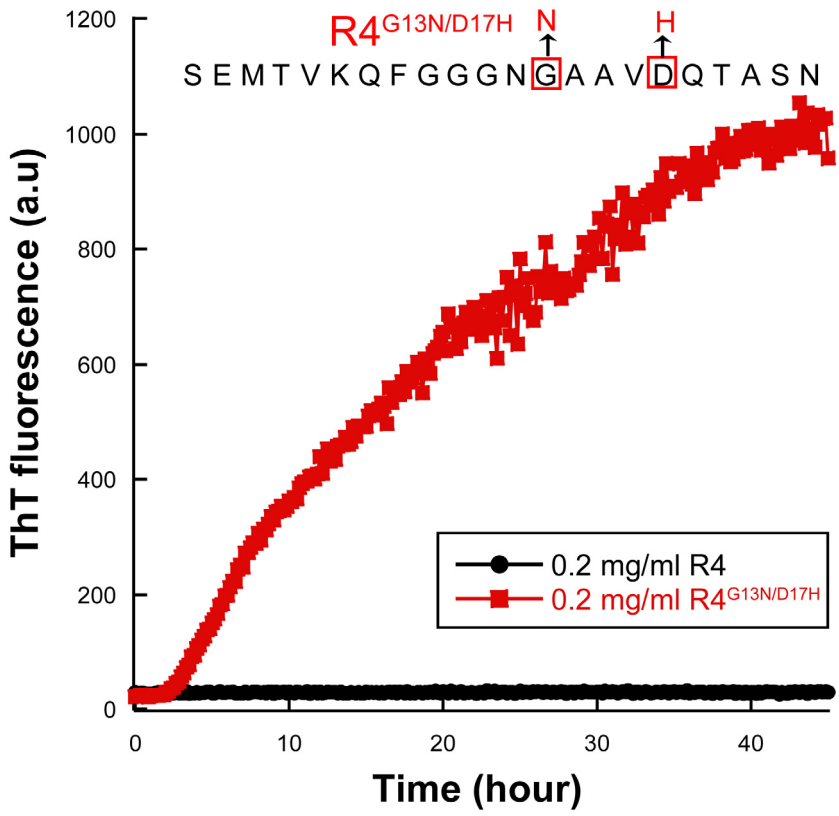


Figure 5.9. Gatekeeper residues can disrupt R5 amyloidogenic properties. (A) Western blots of whole-cell lysates from *csgA* cells containing constructs pCsgA (lane 1 and 2), Control vector (lane 3 and 4) and pCsgA$^{N136D/H149D}$ (lane 5 and 6) grown 48 hours at 26°C on YESCA plates. Samples were treated with (+) or without (-) FA as indicated. The blots were probed with anti-CsgA antibody. (B)-(C) Negative-stain EM micrographs of *csgA* cells containing plasmids pCsgA (B) or pCsgA$^{N136D/H149D}$ (C). Cells were grown 48 hours at 26°C on YESCA plates. Scale bars are equal to 1 μm.

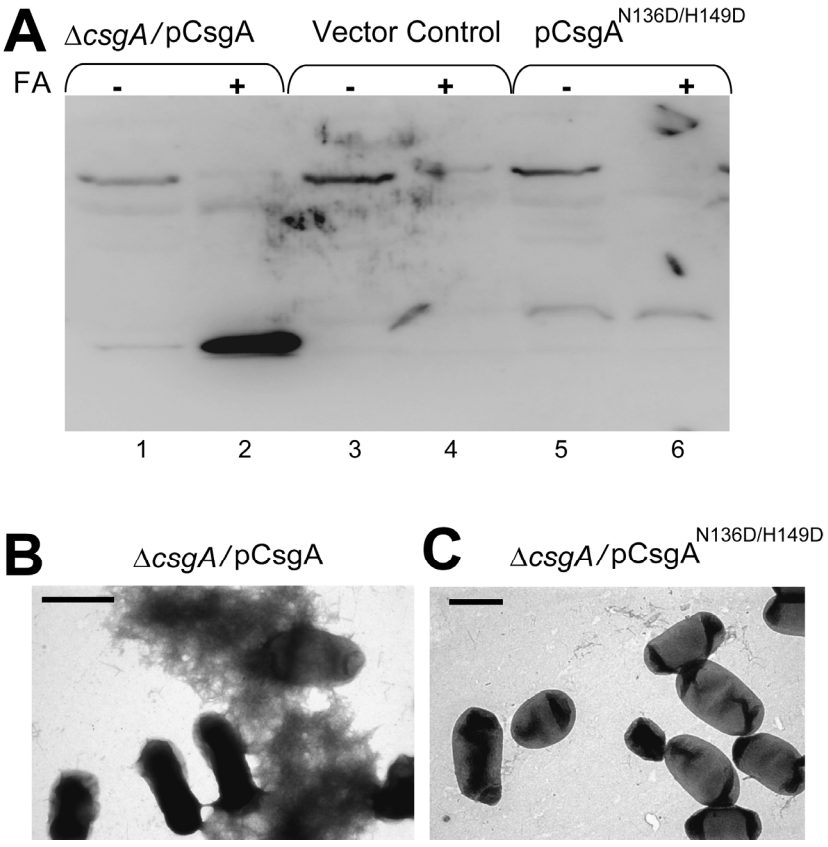


Figure 5.10. *In vivo* polymerization of CsgA* is independent of nucleator CsgB and CsgF. (A) Seven gatekeeper residues were changed to R1/R5-like residues (indicated above the CsgA sequence) in CsgA*. (B) CR-YESCA plate with *csgA*, *csgFBA*, *csgG*, *csgFB*, *csgE* cells transformed with the plasmids encoding CsgA or CsgA*. Cells were grown 48 hours at 26°C on YESCA plates. (C) Western blot of whole cells and underlying agar (agar plugs) from *csgFBA* cells containing control vector (lanes 1 and 2) or pCsgA (lanes 3 and 4) or pCsgA* (lanes 5 and 6). Samples were treated with (+) or without (-) FA as indicated. The blots were probed with anti-CsgA antibody. (D) Negative-stain EM micrographs of *csgFBA* cells containing plasmids pCsgA or pCsgA*. Scale bars equal 500 nm.

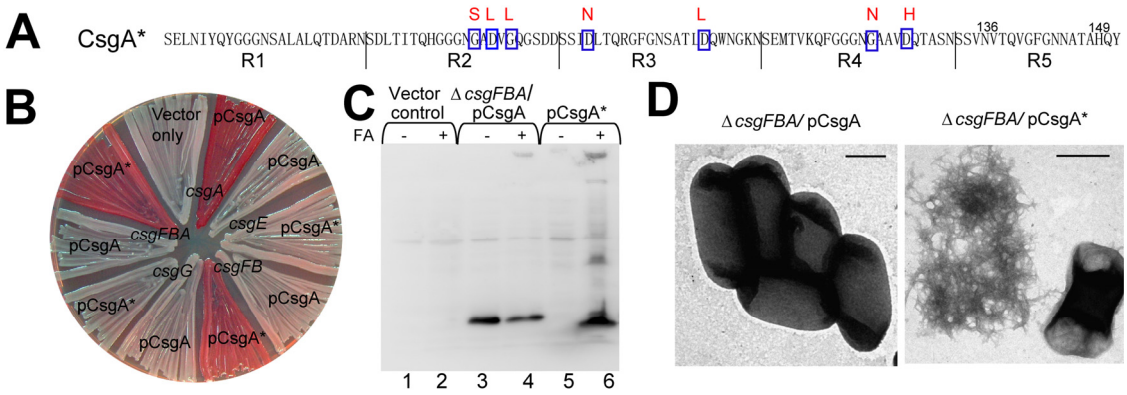


Figure 5.11. Western analysis of CsgA and CsgA* in different curli specific gene deletion backgrounds. Western blot of whole cells and underlying agar (agar plugs) from *csgA csgBA*, *csgG* and *csgE* cells containing pCsgA or pCsgA*. The blots were probed with anti-CsgA antibody. Lanes 1 and 2 contain *csgA*/pCsgA cells and lanes 3 and 4 contains *csgA*/pCsgA*; lanes 5 and 6 contain *csgBA*/pCsgA cells and lanes 7 and 8 contain *csgBA*/pCsgA* cells; lanes 9 and 10 contain *csgG*/pCsgA* and lanes 11 and 12 contain *csgE*/pCsgA*. Samples were treated with (+) or without (-) FA as indicated. The blots were probed with anti-CsgA antibody.

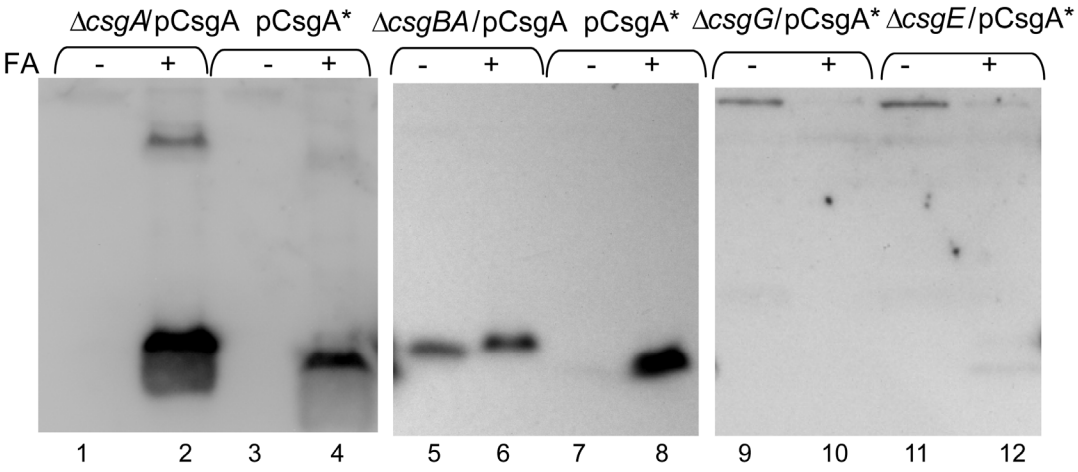


Figure 5.12. *In vitro* polymerization of CsgA* is faster than CsgA and induced expression of CsgA* cause cells less viable. (A) The polymerization of CsgA and CsgA* at 7 μM was measured by ThT fluorescence. (B) Normalized cell suspensions of *csgA* containing empty vector or plasmids encoding $CsgA^{slowgo}$, CsgA or CsgA* were spotted on LB chloramphenicol plates in the presence of IPTG at various concentrations. Colony forming units of serial dilutions of cell suspensions were manually counted.

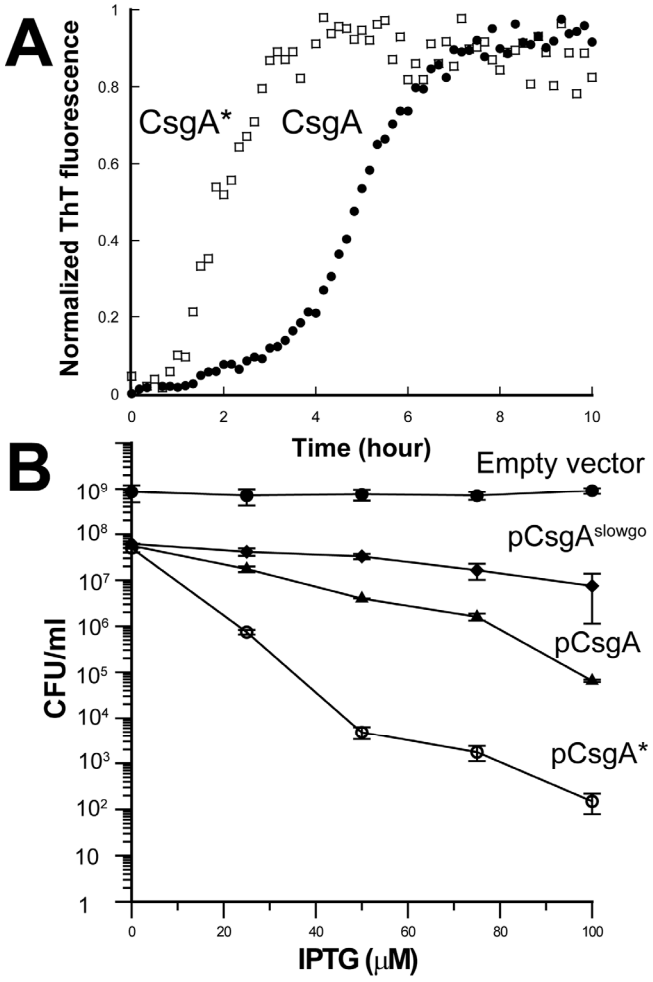


Figure 5.13. Spotting assay of cells with the induced expression of CsgA, CsgA^{slowgo} or CsgA*. LSR10 cells freshly transformed with pNH3, pCsgAslowgoTRC, pMC3 or pCsgA*TRC were scraped off from the plates and normalized by absorbance at 600 nm. Cell suspensions with serial dilutions were spotted on the LB plates supplemented with 25 μg/ml chloramphenicol and 0 μM or 25 μM IPTG.

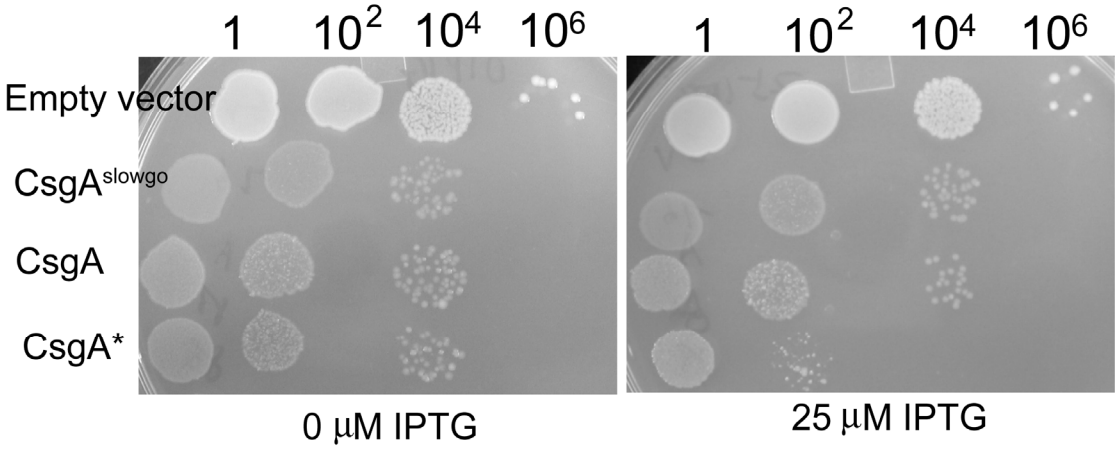


Table 5.1 Strains and plasmids used in Chapter 5

Strains or Plasmids	Relevant characteristics	References
Strains		
csgA (LSR10)	MC4100 Δ*csgA*	(Chapman et al., 2002)
csg*BA* (LSR13)	MC4100 Δ*csgBA*	(Hammer et al., 2007)
csgG (LSR1)	MC4100 *csgG*::Tn105	(Robinson et al., 2006)
csgE (MHR480)	MC4100 Δ*csgE*	(Chapman et al., 2002)
csgFB (MHR442)	MC4100 Δ*csgFB*	(Hammer et al., 2007)
csgFBA (NDH108)	*csgF* Δ*csgBA*	(Hammer et al., 2007)
NEB #C2566	T7 Express Competent *E. coli* (High Efficiency)	NEB Inc.
LSR12	C600 *csgBA and csgDEFG*	Chapman et al., 2002
Plasmids		
pCsgA (pLR5)	*csgA* sequence in pLR2	(Wang et al., 2008)
pLR2	Control vector containing *csgBA* promoter	(Robinson et al., 2006)
pR12341	CsgA repeat 5 replaced by 1 in pLR2	This study
pR52345	CsgA repeat 1 (residues 43-61) replaced by repeat 5 in pLR2	This study
pR32345	CsgA repeat 1 replaced by 3 in pLR2	This study
pR12343	CsgA repeat 5 replaced by 3 in pLR2	This study
pXW56	D136N in R12343 in pLR2	This study
pXW57	R140Y in R12343 in pLR2	This study
pXW58	R140V in R12343 in pLR2	This study
pXW59	D149L in R12343 in pLR2	This study
pXW60	D149H in R12343 in pLR2	This study
pXW61	W151T in R12343 in pLR2	This study
pXW62	W151Y in R12343 in pLR2	This study
pCsgA* (pXW86)	G78S/D80L/G82L/D91N/D104L /G123N/D127H in CsgA in pLR2	This study
pR1234 (pΔR5)	*csgA* without R5 (S^{133} to Y^{151}) in pLR2	(Wang et al., 2007)
pR1234T114N (pYZ1)	T114N in R1234 in pLR2	This study
pR1234K116T (pYZ2)	K116T in R1234 in pLR2	This study
pR1234G123N (pYZ3)	G123N in R1234 in pLR2	This study
pR1234D127H (pYZ4)	D127H in R1234 in pLR2	This study
pR1234G123N/D127H (pYZ12)	G123N and D127H in R1234 in pLR2	This study
pR2345 (pΔR1)	*csgA* without R1 (S^{43} to S^{65}) in pLR2	(Wang et al., 2007)
pR2345D44E (pYZ9)	D44E in R2345 in pLR2	This study
pR2345G55S/D57L/G59L (pYZ10)	G55S/D57L/G59L in R2345 in pLR2	This study
pCsgAN136D/H149D (pXW81)	N136D/H149D in CsgA in pLR2	This study

pMC3	Sequence encoding His-tagged *csgA* cloned into the IPTG inducible plasmid pHL3	(Chapman et al., 2002)
pCsgAslowgoTRC	Sequence encoding His-tagged $CsgA^{Q49A/N54A/Q139A/N144A}$ cloned into pHL3	(Wang et al., 2008)
pNH3	Empty expression vector for pHL3	(Hammer et al., 2007)
pCsgA*TRC	Sequence encoding His-tagged $CsgA^{G78S/D80L/G82L/D91N/D104L/G123N/D127H}$ cloned into pHL3	This study
pNH11	Sequence encoding CsgA without Sec-signal peptide (residues 1-20) cloned into pET11d under T7 promoter	This study
pNH12	Sequence encoding CsgA* without Sec-signal peptide (residues 1-20) cloned into pET11d under T7 promoter	This study

Table 5.2 Sequence of primers used Chapter 5

Primer Name	Sequence
FpLR5[a]	5’ CATGCCATGGCGAAACTTTTAAAAGTAGC 3’
RpLR5[b]	5’ CGGGATCCTGTATTAGTACTGAT 3’
pR12341P1	5’ GTTCAGCTCAGAGTTAGATGCAGTCTG 3’
pR12341P2	5’ CAGACTGCATCTAACTCTGAGCTGAAC 3’
pR12341P3	5’ CGGGATCCTTAGTTACGGGCATCAG 3’
pR52345 P1	5’ CGTTGACGGAGGAATTTGGGCCGCTATT 3’
pR52345 P2	5’ AATAGCGGCCCAAATTCCTCCGTCAACG 3’
pR52345 P3	5’ AGTCAAGTCAGAGTTACGGGCATCGTACTGATGAGCG 3’
pR52345 P4	5’ CGCTCATCAGTACGATGCCCGTAACTCTGACTTGACT 3’
pR32345 P1	5’ GGGTAATAGTCAAGTCAGAATTTTTGCCGTTCCACTGATC 3’
pR32345 P2	5’ GATCAGTGGAACGGCAAAAATTCTGACTTGACTATTACCC 3’
pR12343 P1	5’ CAGATCGATTGAGCTGTTAGATGCAGTCTG 3’
pR12343 P2	5' CAGACTGCATCTAACAGCTCAATCGATCTG 3'
pR12343 P3	5’ CGGGATCCTTAATTTTTGCCGTTCCA 3’
pXW56 P2	5’ CGGGATCCTTAATTTTTGCCGTTCCACTGATCAAGAGTAGCG CTGTTACCGAAGCCACGTTGGGTCAGATTGATTGAGCTGTTAGA 3’
pXW57 P2	5’ CGGGATCCTTAATTTTTGCCGTTCCACTGATCAAGA GTAGCGCTGTTACCGAAGCCGTATTGGGTCAGATCG 3’
pXW58 P2	5’ CGGGATCCTTAATTTTTGCCGTTCCACTGATCAAGA GTAGCGCTGTTACCGAAGCCAACTTGGGTCAGATCG3’
pXW59 P2	5’ CGGGATCCTTAATTTTTGCCGTTCCACTGATGAAGAGTAGCGCTG 3’
pXW60 P2	5’ CGGGATCCTTAATTTTTGCCGTTCCACTGAAGAAGAGTAGCGCTG 3’
pXW61 P2	5’ CGGGATCCTTAATTTTTGCCGTTGGTCTGATCAAGAGTAG 3’
pXW62 P2	5’ CGGGATCCTTAATTTTTGCCGTTGTACTGATCAAGAGTAG 3’
pCsgA* P1	5’ GCCACGTTGGGTCAGGTTGATTGAGCTGTCATCTGAGCCC TGCAGAACCAGTGCAGAATTACCGCCGCCATG 3’
pCsgA* P2	5’ CATGGCGGCGGTAATTCTGCACTGGTTCTGCAGGGCTCAGAT GACAGCTCAATCAACCTGACCCAACGTGGC 3’
pCsgA* P3	5’ GTTAGATGCAGTCTGATGAACTGCAGCGTTGTTGCCACCACCGAAC 3’
pCsgA* P4	5’ GTTCGGTGGTGGCAACAACGCTGCAGTTCATCAGACTGCATCTAAC 3’
pXW81 P1	5’ CTGAGTCACGTCGACGGAGGAG 3’
pXW81 P2	5’ CTCCTCCGTCGACGTGACTCAG 3’
pYZ1 P1	5’ CGGGATCCTTAGTTAGATGCAGTCTGGTCAACTGCA GCACCGTTGCCACCACCGAACTGTTTAACGTTCATTTCAGA 3’
pYZ2 P1	5’CGGGATCCTTAGTTAGATGCAGTCTGGTCAACTGCAGCAC CGTTGCCACCACCGAACTGAGTAACCGTCATTTCAG3’
pYZ3 P1	5’ CGGGATCCTTAGTTAGATGCAGTCTGGTCA ACTGCAGCGTTGTTGCCACCACCG 3’
pYZ4 P1	5’ CGGGATCCTTAGTTAGATGCAGTCTGATGAACTGCAGCACCG 3’
pYZ12 P1	5’ CGGGATCCTTAGTTAGATGCAGTCTGATGAACTGCAGCG TTGTTGCCACCACCG 3’
pYZ9 P1	5’ GGTAATAGTCAACTCAGAATTTGGGCCG 3’
pYZ9 P2	5’ CGGCCCAAATTCTGAGTTGACTATTACC 3’

Reference

Barnhart, M. M., and Chapman, M. R. (2006). Curli biogenesis and function. Annu Rev Microbiol *60*, 131-147.

Chapman, M. R., Robinson, L. S., Pinkner, J. S., Roth, R., Heuser, J., Hammar, M., Normark, S., and Hultgren, S. J. (2002). Role of *Escherichia coli* curli operons in directing amyloid fiber formation. Science *295*, 851-855.

Chiti, F., and Dobson, C. M. (2006). Protein misfolding, functional amyloid, and human disease. Annu Rev Biochem *75*, 333-366.

Chiti, F., Stefani, M., Taddei, N., Ramponi, G., and Dobson, C. M. (2003). Rationalization of the effects of mutations on peptide and protein aggregation rates. Nature *424*, 805-808.

Collinson, S. K., Emody, L., Muller, K. H., Trust, T. J., and Kay, W. W. (1991). Purification and characterization of thin, aggregative fimbriae from Salmonella enteritidis. J Bacteriol *173*, 4773-4781.

Collinson, S. K., Parker, J. M., Hodges, R. S., and Kay, W. W. (1999). Structural predictions of AgfA, the insoluble fimbrial subunit of Salmonella thin aggregative fimbriae. J Mol Biol *290*, 741-756.

Fowler, D. M., Koulov, A. V., Balch, W. E., and Kelly, J. W. (2007). Functional amyloid--from bacteria to humans. Trends Biochem Sci *32*, 217-224.

Hammar, M., Bian, Z., and Normark, S. (1996). Nucleator-dependent intercellular assembly of adhesive curli organelles in *Escherichia coli*. Proc Natl Acad Sci U S A *93*, 6562-6566.

Hammer, N. D., Schmidt, J. C., and Chapman, M. R. (2007). The curli nucleator protein, CsgB, contains an amyloidogenic domain that directs CsgA polymerization. Proc Natl Acad Sci U S A *104*, 12494-12499.

Hammer, N. D., Wang, X., McGuffie, B. A., and Chapman, M. R. (2008). Amyloids:

friend or foe? J Alzheimers Dis *13*, 407-419.

Hardy, J., and Selkoe, D. J. (2002). The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. Science *297*, 353-356.

Jarrett, J. T., and Lansbury, P. T., Jr. (1993). Seeding "one-dimensional crystallization" of amyloid: a pathogenic mechanism in Alzheimer's disease and scrapie? Cell *73*, 1055-1058.

Kallberg, Y., Gustafsson, M., Persson, B., Thyberg, J., and Johansson, J. (2001). Prediction of amyloid fibril-forming proteins. J Biol Chem *276*, 12945-12950.

Luheshi, L. M., Tartaglia, G. G., Brorsson, A. C., Pawar, A. P., Watson, I. E., Chiti, F., Vendruscolo, M., Lomas, D. A., Dobson, C. M., and Crowther, D. C. (2007). Systematic *in vivo* analysis of the intrinsic determinants of amyloid Beta pathogenicity. PLoS Biol *5*, e290.

Monsellier, E., and Chiti, F. (2007). Prevention of amyloid-like aggregation as a driving force of protein evolution. EMBO Rep *8*, 737-742.

Otzen, D. E., Kristensen, O., and Oliveberg, M. (2000). Designed protein tetramer zipped together with a hydrophobic Alzheimer homology: a structural clue to amyloid assembly. Proc Natl Acad Sci U S A *97*, 9907-9912.

Richardson, J. S., and Richardson, D. C. (2002). Natural beta-sheet proteins use negative design to avoid edge-to-edge aggregation. Proc Natl Acad Sci U S A *99*, 2754-2759.

Robinson, L. S., Ashman, E. M., Hultgren, S. J., and Chapman, M. R. (2006). Secretion of curli fibre subunits is mediated by the outer membrane-localized CsgG protein. Mol Microbiol *59*, 870-881.

Rousseau, F., Schymkowitz, J., and Serrano, L. (2006). Protein aggregation and amyloidosis: confusion of the kinds? Curr Opin Struct Biol *16*, 118-126.

Selkoe, D. J. (2003). Folding proteins in fatal ways. Nature *426*, 900-904.

Soto, C., Sigurdsson, E. M., Morelli, L., Kumar, R. A., Castano, E. M., and Frangione, B. (1998). Beta-sheet breaker peptides inhibit fibrillogenesis in a rat brain model of amyloidosis: implications for Alzheimer's therapy. Nat Med *4*, 822-826.

Street, A. G., and Mayo, S. L. (1999). Intrinsic beta-sheet propensities result from van der Waals interactions between side chains and the local backbone. Proc Natl Acad Sci U S A *96*, 9074-9076.

Thirumalai, D., Klimov, D. K., and Dima, R. I. (2003). Emerging ideas on the molecular basis of protein and peptide aggregation. Curr Opin Struct Biol *13*, 146-159.

Van Nostrand, W. E., Melchor, J. P., Cho, H. S., Greenberg, S. M., and Rebeck, G. W. (2001). Pathogenic effects of D23N Iowa mutant amyloid beta -protein. J Biol Chem *276*, 32860-32866.

Wang, X., and Chapman, M. R. (2008). Sequence determinants of bacterial amyloid formation. J Mol Biol *380*, 570-580.

Wang, X., Hammer, N. D., and Chapman, M. R. (2008). The molecular basis of functional bacterial amyloid polymerization and nucleation. J Biol Chem *283*, 21530-21539.

Wang, X., Smith, D. R., Jones, J. W., and Chapman, M. R. (2007). *In vitro* polymerization of a functional *Escherichia coli* amyloid protein. J Biol Chem *282*, 3713-3719.

Chapter 6

Conclusions and Future Directions

Conclusions

Amyloid propagation underlies diverse mammalian ailments such as Alzheimer's disease, systemic amyloidosis, bovine spongiform encephalopathy (mad cow disease) and Creutzfeldt-Jacob disease (Chiti and Dobson, 2006). The product of protein misfolding, disease-associated amyloid formation can be erratic and apparently uncontrolled. Certainly, the physiological and physical constraints on disease-associated amyloidogenesis are poorly described and understood. This is due, in part, to the sporadic nature of *in vivo* amyloid model systems. However, certain organisms have developed the ability to faithfully direct amyloid formation temporally and spatially (Fowler et al., 2007; Hammer et al., 2008; Shorter and Lindquist, 2005). These "functional amyloids" are not toxic to the organism that produces them, but instead fulfill a variety of important physiological roles (Fowler et al., 2007; Hammer et al., 2008; Shorter and Lindquist, 2005). Functional amyloids have been described in bacteria, fungi and mammals (Fowler et al., 2007; Hammer et al., 2008). Because functional amyloid formation can be a tightly controlled process, these biogenesis systems represent a unique platform from which to better understand amyloidogenesis and the cellular toxicity associated with it.

The functional amyloid studied in this thesis is curli, bacterially produced extracellular fibers required for biofilm and other community behaviors (Barnhart and

Chapman, 2006). Curli fibers are composed of a major subunit, CsgA, and minor subunit, CsgB. Curli assembly is initiated by the CsgB nucleator protein, which is proposed to provide an amyloid template to CsgA on the cell surface (Hammer et al., 2007). Curli formation in *E. coli* requires a specific secretion machinery that directs curli subunits to the cell surface and prevents their polymerization inside the cell (Barnhart and Chapman, 2006; Chapman et al., 2002; Robinson et al., 2006). This thesis work has been aimed at understanding the molecular details of CsgA polymerization and nucleation.

Bacterial Amyloid Formation is a conserved folding pathway

The final product of bacterial amyloid curli has been previously demonstrated to be biochemically similar to disease-associated amyloids (Chapman et al., 2002). Part of my thesis work was to determine how the polymerization of functional amyloids compared with that of disease-associated amyloids. The polymerization processes of many disease-associated amyloids share common features. Disease-associated amyloid propagation *in vitro* follows a nucleation-dependent process in which nucleus formation is the rate-limiting step (Jarrett and Lansbury, 1993). A common intermediate of many disease-associated amyloids during the lag phase of polymerization can be detected by a conformational specific antibody, A11 (Kayed et al., 2003). CsgA *in vitro* polymerization was shown to have similar properties to those of disease-associated amyloids in Chapter 2. CsgA polymerization is a triphasic process as measured by ThT assay and preformed CsgA fibers can eliminate the lag phase. The lag phase of CsgA polymerization is independent of its concentration when CsgA concentration is higher than 4 μM,

reminiscent of *in vitro* polymerization of IAPP. Moreover, A11 antibody can recognize the transient form during CsgA lag phase. All these results suggest that not only the final amyloid fibers are conserved, but *in vitro* amyloidogenesis is also conserved between functional and disease-associated amyloids. This supports the notion that amyloidogenesis is an inherent property of polypeptides and assembly of these fibers is also common process. The self-seeding of CsgA also suggests that self-propagation is not only the critical step of disease-associated amyloidogenesis but also may play an important role in fiber assembly.

Bacterial amyloid nucleation was controlled by N- and C-terminal repeats of CsgA

At the cell surface, CsgA is secreted as a soluble, unstructured protein that is converted to an amyloid fiber in a CsgB-dependent fashion (Hammar et al., 1996). *In vitro*, certain CsgB truncation mutants readily assemble into amyloid, which suggested that CsgB might initiate *in vivo* CsgA amyloid formation in a mechanism akin to seeding (Hammer et al., 2007). To directly demonstrate that cell-associated wild-type CsgB accelerates CsgA polymerization, I performed overlay assays using purified CsgA protein and proved that CsgB on the cell surface could promote CsgA polymerization. Preformed CsgA amyloids also promote CsgA polymerization as shown in Chapter 2, demonstrating CsgA responds to self-seeding and CsgB nucleation. To explore the mechanistic details of CsgA polymerization into an amyloid, I extensively analyzed the sequence determinants for curli assembly using a powerful blend of genetic and biochemical approaches.

The CsgA amino acid sequence can be divided into three domains, an N-terminal

Sec signal sequence (cleaved after secretion through the inner membrane), an N-terminal 22 residues involved in secretion through the outer membrane and five imperfect repeating units (R1, R2, R3, R4 and R5) that comprise the protease-resistant amyloid core. I showed that peptides representing R1, R3 and R5 are amyloidogenic *in vitro* in Chapter 2. I showed Peptides corresponding to R1 and R5 responsive to CsgA seeding and CsgB nucleation in Chapter 3. Moreover, deletion of R1, R5, or both R1 and R5 resulted in a molecule that no longer polymerized *in vivo*. A mutant CsgA molecule missing R1 and R5 (ΔR1&R5) was completely incapable of being nucleated by CsgB or being seeded by preformed CsgA fibers. From this analysis, I proposed a simple model of curli assembly in Chapter 3. After secretion across the outer membrane, the R1 and R5 regions of CsgA would interact with either membrane-associated CsgB, or CsgA fiber tips, which template a conformational change in CsgA so that it now adopts an amyloid form (Figure 3.10). The redundancy of seeding/nucleation responsive domains would make curli assembly a highly efficient process.

Remarkably stringent requirements of specific side-chain interactions for curli assembly

Many proteins, if not all, can assemble into amyloid-like fibers *in vitro*, which has lead to the suggestion that amyloid formation is an inherent property of polypeptides (Chiti and Dobson, 2006). Clearly, though, the chemical nature of particular side chains in the polypeptide can influence fiber assembly (Chiti and Dobson, 2006). However, the specific role of side-chain interactions during amyloid formation remains poorly

described. I addressed this question in Chapter 4 by dissecting the internally conserved residues in the CsgA repeating units which are predicted to form cross-beta structure (Collinson et al., 1999). The consensus sequence (Ser-X_5-Gln-X_4-Asn-X_5-Gln) is found in each CsgA imperfect repeating unit. The high degree of the conservation of these polar residues suggests their potential important role in curli assembly. To explore the specific roles of side chains in curli assembly, I analyzed the contribution of conserved polar residues such as Ser, Gln and Asn residues and all the aromatic residues by Ala substitutions. I identified four critical residues, Gln and Asn residues at positions 49, 54, 139 and 144 (Figure 4.1). These results also support the importance of N- and C-terminal repeats in curli assembly as shown in Chapter 3. Ala substitutions of these four residues resulted in a complete loss to CsgB nucleation and an extremely long lag time before fiber formation is initiated, suggesting these four polar side chains contribute to CsgB nucleation response and self-polymerization. Surprisingly, Gln residues at position 49 and 139 cannot even be replaced by Asn residues without interfering with curli assembly. $CsgA^{Q49N}$ is defective in CsgB nucleation response and self-assembly compared to wild-type CsgA. My findings suggest interaction of CsgA to itself and to CsgB is tightly controlled by very specific side-chain contacts and these important side chains are located in R1 and R5 of CsgA. In addition, I found wild-type CsgA fibers still was able to seed the CsgA mutant with all four critical polar side chains abolished, suggesting CsgA fiber seeding and CsgB nucleation have distinguishable mechanisms. It is plausible that fiber mediated seeding is not dependent on polar side chains, whereas the CsgB nucleation

requires the participation of polar side chains. All aromatic residues of CsgA do not positively contribute curli assembly. Maintenance and formation of native globular structures depends on specific side-chain contacts. My work in Chapter 4 demonstrates that the formation of amyloid structure of CsgA also depends on specific side-chain interactions. These side-chain interactions between CsgA-CsgA and CsgA-CsgB allow bacterial amyloid propagation to occur at the correct time and correct location.

Gatekeeper residues of CsgA exquisitely modulate its polymerization

R1 and R5 constitute the nucleation and seeding responsive regions of CsgA. Although the sequence similarity and identity between of R3 and R5 are higher than those between R1 and R5, R3 cannot replace R1 or R5 without interfering with curli assembly as shown in Chapter 5. This suggests that the nucleation/seeding specificity is not simply correlated to sequence similarity or identity. The critical Gln and Asn residues are also present in all five repeats, indicating this specificity is not determined by internally conserved Gln and Asn residues. The molecular determinants of this specificity will be an important step to understand how nature evolved this elegant amyloid propagation system. I extensively investigated the sequence determinants of CsgA for nucleation specificity and self-polymerization in Chapter 5. I found that there are gatekeeper residues in R2, R3 and R4. If these gatekeeper residues were mutated, R2, R3 and R4 became similar as R1 and R5 in terms of intrinsic amyloidogenic properties and nucleation responsiveness. Moreover, I found that gatekeeper residues are not present in R1 and R5 and not tolerated in R5. The presence of gatekeeper residues distinguishes

repeats R2, R3 and R4 from amyloidogenic repeats R1 and R5. Since CsgA is dedicated for amyloid assembly, it is surprising to find the presence of gatekeeper residues in CsgA. I proposed that these gatekeeper residues potentially provide an exquisite way to modulate CsgA amyloidogenesis *in vivo*. Without these gatekeeper residues, CsgA polymerizes into fibers in extacellular space independent of CsgB and CsgF due to its extremely aggregative properties. In addition, the induced expression of CsgA molecule without gatekeeper residues has more significant toxicity than the expression of wild-type CsgA. It seems that the primary sequence of CsgA is shaped by evolution to fit the requirements of both efficient polymerization and prevention for its toxicicty associated with amyloidogenesis.

Future Directions

In this thesis I extensively explore the determinants for CsgA polymerization and nucleation. One of my most exciting discoveries is that nature uses gatekeeper residues to exquisitely modulate functional amyloid propagation and reduce the cytotoxicity. It seems the curli assembly is a balance between polymerization efficiency and control by curli assembling factors. Some interesting questions remain to be answered in the near future.

Without gatekeeper residues, CsgA mutant, CsgA*, was shown to polymerize into amyloid fibers faster than wild-type CsgA. It is worth comparing the kinetic details of CsgA and CsgA* to understand how CsgA* overcomes the nucleation step common for

most amyloid propagation. Is CsgA* polymerization a downhill reaction instead of a nucleation dependent mechanism? By studying the kinetic details of CsgA* polymerization, we will gain some insights of the molecular details of lag phase, which is the rate-limiting step for amyloid propagation. How the efficient polymerization of CsgA* causes *E. coli* less viable is also an intriguing question. CsgA* may stuck in the periplasmic space or interfere with essential cellular process. We can test whether CsgA* exist in the periplasm extraction after induction by western blotting. It is also possible that there are CsgA-specific chaperones or degradation machines that cannot cope with CsgA* due to its uncontrolled polymerization. Chaperones or proteins that help cells to overcome cytotoxicity caused by expression of CsgA* can be identified by a genetic screen to look for more viable *E. coli* cells coexpressing two plasmids containing the sequences of a *E. coli* genomic library and CsgA* sequence, respectively. It is possible to identify the detoxifying machines for amyloid propagation in *E. coli*.

The gatekeeper residue appears to oppose aggregation in some globular proteins (Monsellier and Chiti, 2007; Richardson and Richardson, 2002). The gatekeeper residues in the middle region of CsgA modulate functional amyloid propagation. It is possible that other functional amyloids also have gatekeeper residues to control their propagation. The gatekeeper residues such as Gly and Asp of CsgA are located in the five-residues stretches between Ser and Gln or between Asn and Gln in the consensus sequence Ser-X_5-Gln-X_4-Asn-X_5-Gln. CsgB has a similar repeated primary structure with a consensus sequence X_6-Gln-X_4-Asn-X_5-Gln. Interestingly, CsgB also has Gly and Asp

residues located in the six-residues stretches before the conserved Gln and five-residues stretches between Asn and Gln. Whether these putative gatekeeper residues play a role in the nucleation function of CsgB is worth investigation. It is also exciting to perform an extensive survey to identity the gatekeeper residues in other functional amyloid proteins and their potential roles in regulating amyloidogensis.

Even though my mutagenesis work elucidates the sequence determinants of CsgA nucleation and polymerization, the exact the arrangement of β strands in CsgA fibrils remains elusive. My data showed that the fibers formed by different CsgA mutants have different SDS solubility suggesting the different structural details. The X-ray diffraction and atomic force microscopy were successfully used to reveal the structural features of different amyloid fibers and strength (Serpell et al., 1999; Smith et al., 2006). We can take these approaches to further elucidate the structural differences and their influence to nucleation response and fiber strength.

ThT was convenient to trace the polymerization of CsgA and its mutants. But ThT binding sites on amyloids are unclear and ThT assay did not provide useful information in the early polymerization process. To better understand CsgA polymerization and its interplay with other curli assembling factors such as CsgB, CsgF and CsgE, more sensitive biophysical assays should be constructed such as light scattering and fluorescence correlation spectroscopy. To achieve this goal, I constructed a series of CsgA mutants with Cys residues inserted at various locations that can be labeled by thio-reactive fluorescence dyes. I found $CsgA^{Q150C}$ is wild-type like *in vivo* and it

self-assembles into amyloid fibers *in vitro* with less efficiency (data not shown). $CsgA^{Q150C}$ is able to respond CsgB nucleation and CsgA seeding (data not shown). This mutant is a useful candidate to start studying the early event of CsgA assembly and the interaction with other interaction partners by specifically labeling cysteine residue at position 150.

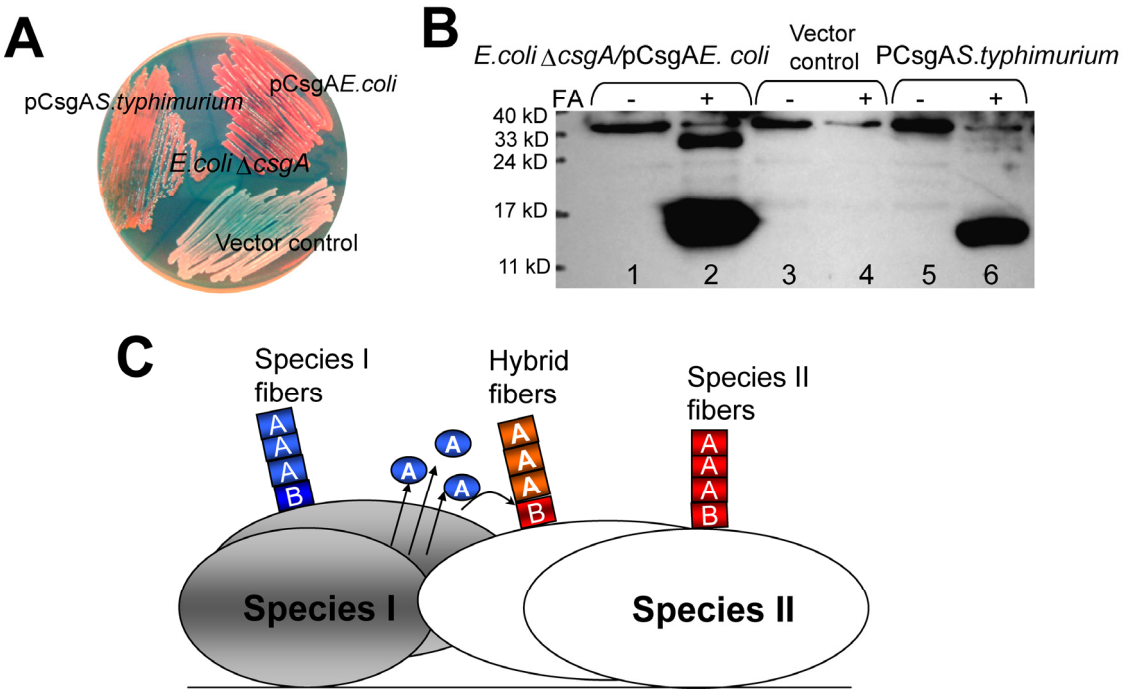
It was reported that seeding effects are strongly influenced by sequence similarity between seeds and soluble amyloid proteins (Chien et al., 2003; Krebs et al., 2004; O'Nuallain et al., 2004). For yeast prions, cross-species transmission was determined by amino acid sequence and conformations of prion proteins (Tanaka et al., 2005). Prion conversion, including mammal and yeast prions, requires a very high level of identity of the interacting protein sequences (Chen et al., 2007; Prusiner, 1998). In contrast to prion proteins, R1 and R5 constitute the critical seeding/nucleation regions of CsgA, although they share only 39% sequence identity between them. In addition, CsgB shares less than 30% sequence identity to CsgA, and it can nucleate CsgA polymerization. The somewhat relaxed requirement of sequence identity for cross-seeding and nucleation in curli biogenesis may have important physiological consequences. It was reported that amyloid-like structures are very abundant in natural biofilms comprising different bacterial species (Larsen et al., 2007). Since the nucleated propagation occurs on the cell surface, bacterial cross-species nucleated amyloidogenesis might happen. *Salmonella typhimurium csgA* and *csgB* genes can complement *Escherichia coli csgA* and *csgB* mutations in terms of positive Congo red binding, indicating amyloid-like structure

formed on the cell surface (Figure 6.1A) (Romling et al., 1998). Like *E. coli csgA*, *S. typhimurium* CsgA formed SDS-insoluble, fibrous structures in *E. coli* Δ*csgA* cells with the plasmid encoding *S. typhimurium* CsgA as detected by western analysis and electron microscopy (Figure 6.1B and data not shown). Presumably, *E. coli* CsgB converts soluble *S. typhimurium* CsgA to the amyloid conformation. This cross-species nucleation promiscuity supports the possibility that the CsgA-like molecules of one bacterial species could interact with the CsgB-like molecules of another bacterial species, facilitating the building of complicated matrixes covering bacterial communities (Figure 6.1C). The study of cross-species nucleation in bacteria could be interesting.

Cross seeding between bacterial amyloids and disease-associated amyloids may have pathological relevance. It has been shown that curli can bind a lot of host cell proteins including MHC-I molecules such as β_2-microglobulin (Olsen et al., 1998). β_2-microglobulin has been implicated in hemodialysis-related amyloidosis. Can curli promote the polymerization β_2-microglobulin since they can physically interact with each other? It has been shown that natural fibrils such as curli could enhance the AA amyloidosis in mice model (Lundmark et al., 2005). Moreover, it was reported that the isolated *E. coli* associated with sudden infant death syndrome were curliated and the sera from the dead infant has certain amount of CsgA by immunoblotting (Goldwater and Bettelheim, 2002). The possible pathological role of bacterial amyloids remains unanswered.

Figures and Tables

Figure 6.1. *Salmonella typhimurium* CsgA responds to *Escherichia coli* CsgB in *E.coli* Δ*csgA* cells and assembles into curli.
(A) The YESCA plates supplemented with Congo red dye with *E.coli* Δ*csgA* cells transformed with vector control or plasmids encoding *E.coli* CsgA or *S. typhimurium* CsgA. Cells were grown for 48 hrs at 26 °C. (B) The western blot of *E.coli* Δ*csgA* cells transformed with vector control (lanes 3 and 4) or plasmids encoding *E. coli* CsgA (lanes 1 and 2) or *S. typhimurium* CsgA (lanes 5 and 6). Samples were treated with (+) or without (-) formic acid (FA) to depolymerize the polymers. The blot was probed with anti-CsgA antibody. (C) Schematic shows cross nucleation and polymerization occurring between different bacterial species in the same environment. CsgA molecules secreted from species I (indicated in blue) interact with CsgB of Species II (indicated in red), undergo a conformation change and assemble into hybrid fibers (indicated in orange). Species I and II can assemble their own fibers (indicated in blue and red, respectively).



Reference

Barnhart, M. M., and Chapman, M. R. (2006). Curli biogenesis and function. Annu Rev Microbiol *60*, 131-147.

Chapman, M. R., Robinson, L. S., Pinkner, J. S., Roth, R., Heuser, J., Hammar, M., Normark, S., and Hultgren, S. J. (2002). Role of Escherichia coli curli operons in directing amyloid fiber formation. Science *295*, 851-855.

Chen, B., Newnam, G. P., and Chernoff, Y. O. (2007). Prion species barrier between the closely related yeast proteins is detected despite coaggregation. Proc Natl Acad Sci U S A *104*, 2791-2796.

Chien, P., DePace, A. H., Collins, S. R., and Weissman, J. S. (2003). Generation of prion transmission barriers by mutational control of amyloid conformations. Nature *424*, 948-951.

Chiti, F., and Dobson, C. M. (2006). Protein misfolding, functional amyloid, and human disease. Annu Rev Biochem *75*, 333-366.

Collinson, S. K., Parker, J. M., Hodges, R. S., and Kay, W. W. (1999). Structural predictions of AgfA, the insoluble fimbrial subunit of Salmonella thin aggregative fimbriae. J Mol Biol *290*, 741-756.

Fowler, D. M., Koulov, A. V., Balch, W. E., and Kelly, J. W. (2007). Functional amyloid--from bacteria to humans. Trends Biochem Sci *32*, 217-224.

Goldwater, P. N., and Bettelheim, K. A. (2002). Curliated Escherichia coli, soluble curlin and the sudden infant death syndrome (SIDS). J Med Microbiol *51*, 1009-1012.

Hammar, M., Bian, Z., and Normark, S. (1996). Nucleator-dependent intercellular assembly of adhesive curli organelles in Escherichia coli. Proc Natl Acad Sci U S A *93*,

6562-6566.

Hammer, N. D., Schmidt, J. C., and Chapman, M. R. (2007). The curli nucleator protein, CsgB, contains an amyloidogenic domain that directs CsgA polymerization. Proc Natl Acad Sci U S A *104*, 12494-12499.

Hammer, N. D., Wang, X., McGuffie, B. A., and Chapman, M. R. (2008). Amyloids: friend or foe? J Alzheimers Dis *13*, 407-419.

Jarrett, J. T., and Lansbury, P. T., Jr. (1993). Seeding "one-dimensional crystallization" of amyloid: a pathogenic mechanism in Alzheimer's disease and scrapie? Cell *73*, 1055-1058.

Kayed, R., Head, E., Thompson, J. L., McIntire, T. M., Milton, S. C., Cotman, C. W., and Glabe, C. G. (2003). Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis. Science *300*, 486-489.

Krebs, M. R., Morozova-Roche, L. A., Daniel, K., Robinson, C. V., and Dobson, C. M. (2004). Observation of sequence specificity in the seeding of protein amyloid fibrils. Protein Sci *13*, 1933-1938.

Larsen, P., Nielsen, J. L., Dueholm, M. S., Wetzel, R., Otzen, D., and Nielsen, P. H. (2007). Amyloid adhesins are abundant in natural biofilms. Environ Microbiol *9*, 3077-3090.

Lundmark, K., Westermark, G. T., Olsen, A., and Westermark, P. (2005). Protein fibrils in nature can enhance amyloid protein A amyloidosis in mice: Cross-seeding as a disease mechanism. Proc Natl Acad Sci U S A *102*, 6098-6102.

Monsellier, E., and Chiti, F. (2007). Prevention of amyloid-like aggregation as a driving force of protein evolution. EMBO Rep *8*, 737-742.

O'Nuallain, B., Williams, A. D., Westermark, P., and Wetzel, R. (2004). Seeding specificity in amyloid growth induced by heterologous fibrils. J Biol Chem *279*, 17490-17499.

Olsen, A., Wick, M. J., Morgelin, M., and Bjorck, L. (1998). Curli, fibrous surface proteins of Escherichia coli, interact with major histocompatibility complex class I molecules. Infect Immun *66*, 944-949.

Prusiner, S. B. (1998). Prions. Proc Natl Acad Sci U S A *95*, 13363-13383.

Richardson, J. S., and Richardson, D. C. (2002). Natural beta-sheet proteins use negative design to avoid edge-to-edge aggregation. Proc Natl Acad Sci U S A *99*, 2754-2759.

Robinson, L. S., Ashman, E. M., Hultgren, S. J., and Chapman, M. R. (2006). Secretion of curli fibre subunits is mediated by the outer membrane-localized CsgG protein. Mol Microbiol *59*, 870-881.

Romling, U., Bian, Z., Hammar, M., Sierralta, W. D., and Normark, S. (1998). Curli fibers are highly conserved between Salmonella typhimurium and Escherichia coli with respect to operon structure and regulation. J Bacteriol *180*, 722-731.

Serpell, L. C., Fraser, P. E., and Sunde, M. (1999). X-ray fiber diffraction of amyloid fibrils. Methods Enzymol *309*, 526-536.

Shorter, J., and Lindquist, S. (2005). Prions as adaptive conduits of memory and inheritance. Nat Rev Genet *6*, 435-450.

Smith, J. F., Knowles, T. P., Dobson, C. M., Macphee, C. E., and Welland, M. E. (2006). Characterization of the nanoscale properties of individual amyloid fibrils. Proc Natl Acad Sci U S A *103*, 15806-15811.

Tanaka, M., Chien, P., Yonekura, K., and Weissman, J. S. (2005). Mechanism of

cross-species prion transmission: an infectious conformation compatible with two highly divergent yeast prion proteins. Cell *121*, 49-62.