Biophysical Characterization of Naturally Occurring Bacterial Amyloids

Dueholm, Morten Simonsen

Publication date: 2009

Document Version
Publisher's PDF, also known as Version of record

Link to publication from Aalborg University

Citation for published version (APA):
BIOPHYSICAL CHARACTERIZATION OF NATURALLY OCCURRING BACTERIAL AMYLOIDS

Morten Simonsen Dueholm

Department of Biotechnology, Chemistry and Environmental Engineering
Section of Biotechnology
Aalborg University
Ph.D. Dissertation, 2009
Preface

The focus of this thesis is discovery and biophysical characterization of functional bacterial amyloids (FuBA). The work is presented in four papers, all included in this thesis.

Paper 1:


Paper 2:


Paper 3:


Paper 4:


The work was carried out in the laboratories of Professor Daniel E. Otzen (Department of Molecular Biology, Interdisciplinary Nanoscience Centre, Aarhus University, Denmark, previously Department of Biotechnology, Chemistry and Environmental Engineering, Aalborg University, Denmark) and the laboratories of Professor Per Halkjær Nielsen (Department of Biotechnology, Chemistry and Environmental Engineering, Aalborg University, Denmark). The majority of the work in paper 1 was performed by PhD, Poul Larsen (Previously Department of Biotechnology, Chemistry and Environmental Engineering, Aalborg University, Denmark) while the majority of the work in paper 2 was performed by MSc, Peter L. Jensen (Previously Department of Biotechnology, Chemistry and Environmental Engineering, Aalborg University, Denmark). Electron microscopy was carried out at the Institute of Medical Microbiology and Immunology, Aarhus University, in collaboration with Professor Gunna Christiansen. MS/MS sequencing was done in the laboratories of Professor Jan J. Enghild (Department of Molecular Biology, University of Aarhus, Denmark) by associate professor Steen V. Petersen. Full genome sequencing was performed by associate professor Kåre L. Nielsen and Ph.D student Mads Sønderkær (Department of Biotechnology, Chemistry and Environmental Engineering, Aalborg University, Denmark). X-ray fiber diffraction spectra were obtained in collaboration with PhD student, Kim L. Hein (Department of Molecular Biology University of Aarhus, Denmark).

The PhD project was funded by pre-doctoral grants from Aalborg University, the Danish National Research Foundation and the BioNET research network, funded by the Villum Kann Rasmussen Foundation.
I would like to acknowledge all present and former members of the protein biophysics group for their support and the many constructive conversations. You have all helped me take this project to a higher level. Especially I would like to thank my supervisor Professor Daniel E. Otzen and co-supervisor Per H. Nielsen for always showing true interest and enthusiasm about my work. I would also dedicate a special thank to Peter A. Christensen, Søren B. Nielsen, Lise Giehm, Martin Jeppesen, Brian Vad, and Rajiv V. Basaiawmoit for three great years together. I would also like to thank Steen V. Petersen, Gunna Christiansen, Kåre L. Nielsen, Mads Sønderkær, and Kim L. Hein for wonderful collaborations. Finally, I would like to thank my family, friends and last but not least my girlfriend for always being supportive.
Resumé


Idet mulighederne for at undersøge amyloiddannelsen in vivo er meget begrænsede, er de fleste undersøgelser foretaget in vitro. Vores viden om amyloiddannelsen er derfor begrænset til disse simplificerede systemer, som på mange punkter ikke imiterer omgivelserne, hvor sygdomsrelaterede amyloider dannes.

Et voksende antal undersøgelser har i midlertid vist at den amyloide struktur ikke kun er associeret med sygdom, men også udfører gavnlige funktioner i naturen. Et eksempel er curli fimbriaer fra *Escherichia coli* og *Salmonella* stammer. Disse amyloider findes på overfladen af bakterieceller, hvor de formidler celle til celle/vært interaktioner. De er ydermere involveret i dannelsen af biofilm.

De bakterielle amyloider er i kontrast til de sygdoms relaterede amyloider optimeret med hensyn til strukturel stabilitet. Viden om deres egenskaber kan derfor bidrage til den grundlæggende forståelse af fibrilleringsprocessen. Udover at komme med svar på hvorfor amyloide strukturer dannes og medfører sygdom, kan denne viden også være et værdifuldt værktøj i udviklingen af strategier til at forhindre eller kontrollerer uhensigtsmæssig biofilm dannelse.

Summary

Approximately 20 different human diseases are known to be associated with the deposition of protein fibrils, known as amyloid fibrils, resulting in tissue damage and degradation. The amyloid diseases are very common and comprise devastating diseases like Parkinson’s and Alzheimer’s disease. Detailed understanding of the mechanisms by which amyloids are formed from native proteins and how they cause disease are therefore of both social and economic importance.

As the possibilities for studying amyloid formation in vivo are very limited, most studies on the fibrillation process have been carried out in vitro. This limits our knowledge of the amyloid formation to simplified systems that in many regards do not mimic the natural settings in which the disease related amyloids are formed.

Interestingly, a growing number of studies have now shown that the amyloid structure is not only associated with disease but is actually being used beneficially in nature. One example is curli fimbriae from Escherichia coli and Salmonella sp. These bacterial amyloids are found on the surface of the bacterial cells, where they mediate cell to cell/host interactions. They are furthermore involved in the formation of biofilms.

Functional bacterial amyloids are in contrast to amyloids involved in human diseases evolutionarily optimized for structural stability. Knowledge of their properties might therefore provide fundamental insight into the fibrillation process. Besides providing answers to why amyloid structures form and cause disease, this insight may help in the development of strategies to prevent or control fimbriae formation and thus potential control unwanted biofilm production.

In this study we were able to show that the expression of amyloid structures is a common feature of many bacteria living in biofilms. A detailed investigation of pure culture mycolata systems showed that amyloids in many species may constitute a hitherto overlooked integral part of the spore and the cellular envelope. We were furthermore able to isolate and characterize a functional amyloid from a Pseudomonas strain and identify the operons responsible for its expression. This operon was conserved in many other Pseudomonas strains including the opportunistic pathogen Pseudomonas aeruginosa and represented a novel genetic organization for amyloid expression. Finally we performed at detailed study on the effect of environmental conditions on the fibrillation of CsgA, the major subunit of the curli fimbriae. It was found that CsgA was extremely robust to changing environmental conditions. This is consistent with evolutionarily optimized fibrillation properties.
Abbreviations

Ordered alphabetically

Aβ: Amyloid β-peptide of Alzheimer’s disease
AFM: Atomic force microscopy
ALF: Amyloid-like fimbriae
BCA: Bicinchoninic acid assay
BTACD: Circular dichroism
CF: Cystic fibrosis
CLSM: Confocal laser scanning microscopy
CR: Congo red
CWS: Cell-wall sorting signal
DAPI: 4’,6-diamidino-phenylindole
EM: Electron microscopy
FA: Formic acid
FISH: Fluorescence in situ hybridization
FTIR: Fourier-transform infrared spectroscopy
FuBA: Functional bacterial amyloids
GdmCl: Guanidinium chloride
hIAPP: Human islet amyloid polypeptide
HR: Hypersensitive response
IgG: Immunoglobulin G
MS/MS: Tandem mass spectrometry
MTP: Mycobacterium tuberculosis pili
SEM: Scanning electron microscopy
PFA: Paraformaldehyde
PrP: Prion protein
PrP<sup>Sc</sup>: Prion protein in the endogenous cellular form
PrP<sup>Sc</sup>: Prion protein in the misfolded, disease associated form
SDS: Sodium dodecylsulphate
SDS-PAGE: Sodium dodecylsulphate polyacrylamide gel electrophoresis
Spp. or Sp.: Species
ssNMR: Solid state nuclear magnetic resonance
STEM: Scanning transmission electron microscopy
Tafi: Thin aggregative fimbriae
TEM: Transmission electron microscopy
TFA: Trifluoroacetic acid
TFE: 2,2,2-trifluoroethanole
ThT: Thioflavín T
TSE: Transmissible spongiform encephalopathy
WWTP: Wastewater treatment plant
# Table of Contents

Preface ................................................................................................................................. 1
Resumé ................................................................................................................................. 3
Summary .................................................................................................................................. 5
Abbreviations ........................................................................................................................ 7

1. Introduction ..................................................................................................................... 11
   1.1 Amyloids ....................................................................................................................... 11
      1.1.1 Discovery and Medical Implications ......................................................................... 11
      1.1.2 Structure of the Amyloid ......................................................................................... 11
      1.1.3 Amyloid Detection .................................................................................................. 12
      1.1.4 The Folding Funnel and Amyloid Formation ......................................................... 13
   1.2 Functional Amyloid ..................................................................................................... 16
      1.2.1 Curli and Thin Aggregative Fimbria ........................................................................ 16
      1.2.2 Mycobacterium tuberculosis Pili ............................................................................ 18
      1.2.3 Harpins .................................................................................................................... 19
      1.2.4 Chaplins .................................................................................................................. 20
      1.2.5 Hydrophobins ......................................................................................................... 22
      1.2.6 Yeast Prions Provide Non-Mendelian Traits ......................................................... 23
      1.2.7 Silkmoth Chorion .................................................................................................... 24
      1.2.8 Pmel17 .................................................................................................................... 26
   1.3 Aim of this Project ....................................................................................................... 27

2. Papers .............................................................................................................................. 29
   2.1 Paper 1 ........................................................................................................................ 29
      2.1.1 Motivation ............................................................................................................... 30
      2.1.2 Outcome .................................................................................................................. 30
      2.1.3 The Paper ............................................................................................................... 31
   2.2 Paper 2 ........................................................................................................................ 45
      2.2.1 Motivation ............................................................................................................... 46
      2.2.2 Outcome .................................................................................................................. 46
      2.2.3 The Paper ............................................................................................................... 47
   2.3 Paper 3 ........................................................................................................................ 57
      2.3.1 Motivation ............................................................................................................... 58
      2.3.2 Outcome .................................................................................................................. 58
      2.3.3 The Paper ............................................................................................................... 59
   2.4 Paper 4 ........................................................................................................................ 73
      2.4.1 Motivation ............................................................................................................... 74
      2.4.2 Outcome .................................................................................................................. 74
      2.4.3 The Paper ............................................................................................................... 75
3. Discussion and Perspective ........................................................................................................... 87
   3.1 Prevalence of Functional Amyloids ....................................................................................... 87
   3.2 Localization and Function of Functional Bacterial Amyloid ............................................... 87
   3.3 Functional Bacterial Amyloids in *Pseudomonas* ............................................................... 88
   3.4 Repeated Sequences and Conserved Gln/Asn Residues ..................................................... 88
   3.5 Functional Amyloid Precursor are highly Amyloidogenic ............................................... 89
   3.6 Molecular Structure of a Functional Amyloid ...................................................................... 89
   3.7 Environmental Degradation of Functional Amyloids .......................................................... 89
   3.8 Working with Functional Amyloid ....................................................................................... 90

4. Conclusions .................................................................................................................................. 93

5. References ................................................................................................................................... 95
1. Introduction

1.1 Amyloids
Amyloids are insoluble, proteinaceous aggregates sharing a well-defined fibrous cross-β quaternary structure in which the amyloid polypeptides fold in a regular manner on itself so that adjacent chain segments are laterally arranged perpendicular to the fibril axis (Figure 1) [1-3].

![Image of amyloid structures](image_url)

**Figure 1: Ribbon representation of amyloid structures formed by Aβ40.** (A) Structure a protofilament with twofold symmetry, based on solid state nuclear magnetic resonance (ssNMR), with constrains from electron microscopy (EM), X-ray fiber diffraction, electron paramagnetic resonance (EPR) and biochemical techniques. Reprinted from [4] with permission from Elsevier. (B) Structure of a protofilament with threefold symmetry, based on ssNMR, with constrains from scanning transmission electron microscopy (STEM) and specific secondary, tertiary, and quaternary structural data from ssNMR. Reprinted from [5] with permission from the National Academy of Sciences.

1.1.1 Discovery and Medical Implications
The term amyloid was first introduced in connection with disease in 1854 by Rudolph Virchow, who discovered that cerebral corpora amylacea that had an abnormal macroscopic appearance stained pale blue when treated with iodide and that the color subsequently turned purple when the sample was treated with sulfuric acid. He concluded that the structures underlying the apparent macroscopic abnormalities were cellulose or starch and termed them amyloids, derived from the Latin amyllum and the Greek amylon [6, 7]. Five years later Carl Friedreich and August Kekule demonstrated that major component of the amyloid was proteinaceous and not cellulose or starch as earlier suggested by Virchow. Despite these findings the amyloid term survived [8].

Since then more than 20 different human diseases have been linked to the formation of amyloids, including Parkinson’s and Alzheimer’s disease, type II diabetes, and the transmissible spongiform encephalopathies (TSE) such as Creutzfeld-Jacob disease [9]. The high prevalence of the amyloid diseases signifies that research into the field of amyloid formation is both of social and economic importance. For each disease a biochemically unique protein makes up the majority of the amyloid deposit [9]. Interestingly, these proteins do not share any secondary or tertiary structure natively. It is therefore suggested that the amyloid protein fold is supported by backbone interactions and thus represent a generic conformation which can be adopted by all proteins given the right conditions. This has been substantiated by the fact that many non-amyloidogenic proteins can form amyloid-like fibrils in vitro [10].

1.1.2 Structure of the Amyloid
When examined by electron microscopy (EM) or atomic force microscopy (AFM) amyloids are found to be made of straight fibrils composed of twined filaments. The fibrils are of indeterminate length
and have a width in the order of 7-20nm, varying with the fibril morphology (Figure 2). A given protein is not restricted to one single fibril morphology. An in vitro fibrillation study of the 29 residue peptide hormone glucagon showed that at least four different fibrils morphologies can be formed by changing the fibrillation conditions, such as pH, concentration of salts, and temperature. Some of the fibril morphologies show a periodic pitch, whereas others seem straight (Figure 2) [11]. Another study carried out on hIAPP20-29, a fragment of the human islet amyloid polypeptide amylin, using solid state NMR (ssNMR) showed fibril polymorphism even in a single fibrillation batch [12].

![Figure 2: Amyloid fibril morphologies.](image)

The cross-β quaternary structure of amyloids was first observed and described in 1969 using X-ray fiber diffraction [14]. The amyloid fold is characterized by an intense meridional reflection at 4.7-4.8Å, which result from the mean separation of the hydrogen-bonded β-strands that are arranged perpendicular to the fiber axis in the cross-β fold, and a weaker equatorial reflection at approximately 10Å, which originates from the spacing of the β-sheets [1].

Fourier transform infrared spectroscopy (FTIR) has also proven useful as a biophysical tool for the identification of amyloid structures. A comparison of the FTIR spectra of various globular β-sheet proteins and amyloids showed an interesting feature in the amide I band. For the globular proteins the amide I maximum for β-sheet was found from 1630-1643cm⁻¹ whereas it was shifted to 1611-1630cm⁻¹ for the amyloid fibrils [15]. The shift of the β-strand absorbance maximum to lower wavenumbers for amyloids results from the very strong hydrogen bonds formed in the cross-β protein fold [16].

1.1.3 Amyloid Detection
Since amyloids were first identified using iodide, there has been a remarkable development in amyloid specific dyes. Metachromatic dyes, such as methyl violet, were the first to replace the iodide sulfuric acid test. These dyes allowed the identification of the extracellular nature amyloids deposits [7].

The aniline dye Congo red (Figure 3A) was introduced for amyloid staining in 1922 by H. Bennhold, who discovered that when Congo red (CR) was injected intravenously into patients with amyloid, the dye disappeared from the plasma and accumulated in the amyloid deposits [7]. The characteristic apple-green birefringence of amyloids stained with CR under polarized light was first described in 1927 by Divry and Florkin [7]. This phenomenon is thought to arise from linear orientation of CR molecules along the fibril, but the exact nature of the binding is not well understood [17-20]. CR
binding specificity is relatively high, but whereas the characteristic birefringence appears to be specific for the continuous β-sheet structure of amyloids, CR also binds all-α, α+β, all-β and β-helical proteins as well as extended β-sheets of the short-chain polyamines polyglycine and poly-L-serine although with low affinity [21-23].

![Figure 3: Molecular structure of common amyloid dyes. (A) Congo red (B) Thioflavin T](image)

Thioflavin T (ThT) (Figure 3B) is a fluorescent benzoethiazole dye, which have been widely used as a probe for amyloid formation in vitro, because it does not interfere appreciably with the formation of fibril structure [24]. ThT has a higher sensitivity than CR but a more profound non-specificity [25]. ThT binds poorly to natively folded proteins and non-fibrillar structures formed along the fibrillation pathway although non-amyloid conformations with ThT-affinity have been described [26-28]. The affinity of the charged ThT molecule towards fibrils varies greatly depending on pH. Differential binding is also observed depending on the properties of the fibrillating polypeptide making direct quantitative comparison between fibril species challenging [11, 26]. Several fibril species that bind ThT poorly or do not bind the dye in detectable amounts have been identified including fibrillar structures of poly-L-serine or poly-L-lysine, fibrils formed of truncated amylin peptide and ThT-negative, CR-positive fibrils of Aβ1-40 [29, 30]. When ThT binds to amyloids a new peak, not seen for the free dye, appears at 450nm in the excitation spectrum. When this peak is excited fluorescence emission can be measured at 482nm [31]. There have been proposed a number of different mechanisms for the binding of ThT to amyloid fibrils. One of the more plausible models states that ThT binds in hydrophobic channels running parallel to the fibril axis [32, 33]. The change in the spectral properties of ThT results from a chiral bias imposed by the interaction with the amyloid fibril, and has been interpreted as an indication that the molecule adopts a twisted conformation upon fibril binding [34, 35].

In 2002 O’Nnuaillain and Wetzel introduced conformational antibodies (WO1 and WO2) against a common conformation epitope in amyloid fibrils with little apparent dependence on amino acid side chain information [36]. The antibodies were raised against fibrils of Aβ, but were found also to bind amyloids of other protein, while no affinity was found for the monomeric precursors. Furthermore, the antibodies showed no affinity for non-amyloid biological aggregates, such as collagen, gelatin, and elastin [36]. These properties show that the conformational antibodies represent highly specific probes for amyloid detection.

**1.1.4 The Folding Funnel and Amyloid Formation**

Many proteins need to adopt a unique protein structure, termed the native structure, in order to be functional active. In vitro folding studies have shown that small globular proteins such as RNase can fold spontaneous into the native structure in absence of any catalytic biomolecules [37, 38]. This lead to the assumption that the native state represented a global free energy minimum state [37]. However, the numbers of possible conformations of the polypeptide backbone is so large that if all
conformations should be screened, a protein would not be able to fold within the lifetime of the universe. This is called the Levinthal paradox [37, 39, 40].

A solution to this problem was presented by Leopold et al. who introduced the folding funnel (Figure 4A) [37]. The folding funnel principle follows as few guide lines. (i) Proteins fold from a random state by collapsing and reconfiguring, (ii) reconfiguration occurs diffusively and follows a drift from high energy towards low energy conformations, and (iii) reconfiguration occurs between conformations that are structurally similar. The result of the latter is that global conversion is prohibited after the collapse and local configuration alone is considered [37]. The folding funnel presented by Leopold et al. has one major drawback; it does not consider intermolecular contacts, and thus only apply for very dilute protein solutions. In order to make the folding funnel theory more applicable Clark presented a reshaped folding funnel with an additional deep energetic minima, which reflected off-pathway interchain interactions (Figure 4B) [38]. With this modification the folding funnel extended to describe the effect of chaperons, protein aggregation and amyloid formation.

![Figure 4: Cartoon representation of the folding funnel. (A) The folding funnel as described by Leopold et al. [37]. (B) The folding funnel modified to include intermolecular interactions as described by Clark [38]. The following abbreviations are used: M=molten Globule, N=native state, Agg=amorphous aggregates, Amy=amyloid fibrils, U=unfolded structures, I=partly folded intermediates. Stable conformations are highlighted in dark color.](image)

In vitro the fibrillation of most soluble proteins into amyloid fibrils follows a nucleation dependent mechanism containing a lag phase, a growth phase, and a stationary phase (Figure 5A+B). The lag phase represents the formation of fibrillation prone nuclei. The nuclei formation is followed by an elongation reaction, where fibrils form by incorporation of additional precursors to the nuclei. This yields the growth phase. The stationary phase is reached when the precursors are depleted [41-44]. For many native folded proteins the formation of fibrillation prone nuclei requires that at least a fraction of the proteins is partly unfolded, but not so well solvated that they do not engage in noncovalent intermolecular interactions [45, 46]. A very stable native fold precludes sufficient population of the amyloidogenic state and results in the absence of significant fibril nucleation and elongation [47]. The 101 residue S6 from *Thermus thermophilus*, for example requires elevated temperature, low pH, and the presence of moderate concentrations of salt in order to fibrillate. At these conditions some regions of the protein have been shown to be partially unfolded [45]. For Aβ a low concentration of TFE (20%) converts the predominantly unstructured monomers into partially ordered, quasi-stable conformers. This conversion results in a temporal decrease in the lag phase for fibril formation and a significant increase in the rate of fibril elongation [47]. Both studies indicate the requirement of a semi folded precursor for nucleation to be efficient.
The nature of the fibrillation nuclei is still debated, but it is generally considered as structured oligomers, termed protofibrils [9, 42]. Both the Aβ₁₋₄₀ and Aβ₁₋₄₂ have been shown to exist as soluble oligomers in rapid equilibrium with the corresponding monomeric forms. These oligomers appear to be composed of 2-4 and 5-6 molecules for Aβ₁₋₄₀ and Aβ₁₋₄₂, respectively. Circular dichroism (CD) measurements have suggested that the secondary structure of these oligomers is relatively disordered [48]. Similar oligomers have also been identified in equilibrium with the monomeric forms of the NM region of the yeast prion Sup35p, yeast phosphoglycerate kinase, and the SH3 domain from the bovine phosphatidylinositol 3’ kinase [49-51]. These small oligomers are hypothesized to associate with each other to form the structured protofibrils, which in turn function as fibrillation prone nuclei [49, 51, 52]. Protofibrils can be visualized by AFM or TEM [53-55]. Here they appear to be single or chained spherical beads with a diameter of 2-5nm. Protofibrils can also appear as annular structures, which is apparently formed by circularization of the beaded chains [53-55].

Recent research has indicated, that for at least some neurodegenerative diseases, small oligomers or protofibrils, rather than amyloid fibrils, are the real pathogenic species. For Alzheimer’s disease the severity of the cognitive impairment correlates with the presence of small oligomers, rather than with the amyloid burden [56-58]. This is furthermore substantiated by the fact that the progressive “Arctic” mutation (E693G) of the Aβ precursor protein, associated with a hereditary early-onset manifestation of Alzheimer’s disease, has been found in vitro to boost protofibril formation, but not fibril formation [59]. Parkinson’s disease is associated with fibrillation and deposition of the protein α-synuclein in the dopaminergic neurons of the substantia nigra [60]. The dopaminergic neurons that survive the disease show no quantifiable difference in viability, whether or not they contain fibrillar deposits [61, 62]. Mutations related to early-onset forms of Parkinsonism give rise to neuronal degeneration in the absence of accumulated fibrillar deposits [63]. These results both hint that prefibrillar intermediates, rather than the mature fibrils, represent the cytotoxic species. Prefibrillar species of many other proteins have also been shown to be toxic to neuronal cells under conditions where the native and fibrillar forms of the proteins are not [64-67].

A characteristic of the fibrillation process is the ability of preformed seeds (fragmented mature fibrils) to nucleate the fibrillation, thus removing or minimizing the lag phase (Figure 5A) [11, 42, 68]. In general the preformed seeds need to be of the same protein, however cross seeding is seen for
some proteins with high sequence similarity. A study of seeded lysozyme fibrillation showed that for efficient cross seeding proteins with a sequence identity of more than 36% is needed [68]. Similar results were obtained in a study on the fibrillation of hetero-dimerized human immunoglobulin domains. When the linked domains had a more than 70% sequence identity, they were highly prone to co-aggregation, whereas no interaction was seen when the sequence identity was below 30-40% [69].

1.2 Functional Amyloid

The amyloid fold is not restricted to protein misfolding and disease. The last decade has provided evidence that many living systems use the generic amyloid fold for functional purposes. These living systems stretch from simple bacteria to humans [70-73]. In the following some of the important systems for functional amyloid will be discussed.

1.2.1 Curli and Thin Aggregative Fimbria

Fimbriae are proteinaceous, filamentous appendages exposed on bacterial surfaces [74, 75]. These structures are mainly composed of a single protein species called the fimbrin, which upon polymerization forms filaments of 2-7nm in diameter, that extent 1-2µm from the bacterial surface [75, 76]. For Gram-negative bacteria, for which fimbriae have been extensively studied, the molecular weight of the fimbrins are in the range of 14 to 30kDa [76]. The fimbriae enable the bacteria to make specific recognition and adhesion to diverse target molecules such as mammalian host tissue components, eukaryotic and prokaryotic cell surface elements or abiotic surfaces [74, 77, 78]. These factors suggest that fimbriae are important actors in the formation of biofilms, which are defined as sessile bacterial communities of cells that live attached to each other and/or inanimate surfaces [79].

In 1989 a unique and highly aggregative type of fimbria was purified from a natural isolate of E. coli [80]. This type of fimbria was morphologically and biochemically distinct from other surface appendages of E. coli and it was composed of subunits, that differed from all previous known fimbrins. The fimbriae were named curli due to their coiled appearance, as judged by EM (Figure 6A) [80]. Homologous fimbriae were later discovered in Salmonella spp. [74, 76]. These fimbria was named thin, aggregative fimbriae (Tafi) [76]. In 2002 Chapman et al. showed that curli fimbriae share the general properties of amyloids, i.e. a rich β-sheet secondary structure, as seen by CD, binding of Congo red, resulting in the characteristic green birefringence and amyloid-like binding of ThT [81]. The curli and Tafi are extremely resistant to mechanically and chemically denaturation, and do not depolymerize in 5M NaOH, 8M urea or boiling in 2% SDS [76, 82]. This property can be linked to the high degree of hydrogen bonds in the amyloid fold stabilizing the structure. However, curli and Tafi can be depolymerized using 90% formic acid [76, 82]. The effect of formic acid can be described by two effects: (i) Formic acid is a less polar solvent than water and thus a better solvent for the apolar parts of the protein and (ii) formic acid introduce alternative hydrogen bond. The combined result is a disruption of the amyloid structure [83].
Expression of curli and Tafi are carried out by genes that are arranged and regulated in a similar fashion [74]. The genes for curli expression are clustered in the csgBAC and csgDEFG operons, which encodes the curli subunits and regulate their transcription and transport, respectively (Figure 6B) [77, 84-86]. In Salmonella the corresponding operons have been termed agfBAC and agfDEFG [74, 87]. Due to the high similarities between the curli and the Tafi systems, we will consider only the curli system in the following. Additional information obtained from studies on Tafi will be introduced to shed further light on the system.

The csg genes are responsible for the expression of the respective Csg proteins. CsgA is the major subunit of the curli fimbriae. It contains an N-terminal Sec signal sequence, which is cleaved off during translocation to the periplasm. Furthermore, it contains five imperfect repeats with highly conserved glutamine and asparagine residues that are important for the amyloid formation (Figure 6C) [88, 89]. CsgB is a homolog to CsgA which also contains the repeated units and shares 30% amino acid sequence identity (Figure 6C) [84, 88]. CsgB is translocated to the cell surface where it acts as a nucleator protein in the formation of curli (Figure 6B). It has been suggested that the nucleation is a result of conformational changes of soluble CsgA upon interaction with CsgB [84, 90]. This has been substantiated by molecular modeling of AgfB, which shows that AgfB can form a β-helix structure compatible with that predicted for the fibril form of AgfA [91]. It is not known whether CsgB is linked directly to the cell membrane or to the assembly complex by protein-protein interactions. An additional function has been indicated for the CsgB homolog in the Tafi system. AgfB subunits can be found in the length of the native Tafi at a stoichiometry of approximately one AgfB molecules per 20 AgfA molecules. Here they seem to be involved in the formation of fibril branch-points [91]. No transcripts have been detected for csgC and there is no reported role for CsgC in curli biogenesis [92]. However, mutation studies have indicated that AgfC is localized in the periplasm and that it is important for AgfA extracellular assembly, facilitating Tafi synthesis [87]. CsgD is a regulator of the LuxR superfamily that positively controls the production of curli by transcriptional activation of the csgBA operons [93]. CsgG is an outer membrane-located lipoprotein, which is highly resistant
to protease digestion both in vivo and in vitro [90]. It has been indicated that CsgG forms an outer membrane channel that stabilize CsgA and CsgB during curli assembly by mediating their translocation across the outer membrane [94]. Furthermore, it has been shown that CsgG participates in a complex with CsgE and CsgF [94]. Knock-out experiments have shown that CsgE and CsgF stabilize curli, and they have therefore been suggested to be involved in the extracellular nucleation of CsgA [81]. The strict regulation of the curli operons and the separation of the nucleation and seeding process into CsgB and CsgA (in fibrillar form), respectively, ensures that the amyloid fibril formation happen at the right time and place. Production of potential cytotoxic intermediates therefore does not endanger the cell. A detailed understanding of the molecular basis for curli formation might therefore provide new insight on how to reduce the toxicity of disease related amyloids [73].

Two different methods have been employed to isolate curli and Tafi fimbriae. The first was developed for purification of Tafi fimbriae [76]. In this method bacteria are grown on CFA agar plates, suspended in buffer and homogenized by sonication. Contaminating cell macromolecules are enzymatically degraded or dissolved in detergents and removed by washing steps. Contaminating proteins are removed by preparative SDS-PAGE. The latter is possible because boiling in SDS-PAGE loading buffer do not depolymerize the insoluble Tafi, which can be collect from the top of the SDS-PAGE gel after electrophoresis [76]. The second method was developed for purification of curli fimbriae [81]. In this method bacteria are grown on YESCA agar plates and dissolved in buffer. Curli are then sheared of the bacteria using an Omni-mixer homogenizer, and subsequently isolated from other cell components and soluble proteins by differential centrifugation [81]. If high purity is needed, a preparative SDS-PAGE step, similar to that for the Tafi purification, is included [81].

1.2.2 Mycobacterium tuberculosis Pili

*Mycobacterium tuberculosis* is responsible for nearly 3 million human deaths annually, and is thereby the number one bacterial cause of human mortality. Even though it is widely held that *Mycobacteria* do not produce pili, Alteri *et al.* has shown that *M. tuberculosis* produce pili when grown for in vitro for 2-3 weeks on certain solid mycobacteriological medium [95]. The *M. tuberculosis* pili (MTP) are 2-3nm wide and are morphologically very similar to curli and Tafi fimbriae (Figure 7). MTP was shown not to dissociate when boiled in SDS-PAGE loading buffer, a biochemical property that is also characteristic for the curli and Tafi amyloid fibrils [76, 95]. The MTP was furthermore found to bind laminin, but not fibronectin or type IV collagen [95]. Finally, the *M. tuberculosis* was indirectly shown to produce MTP in at least 60% of human infection using an immunofluorescence microscopy experiment with affinity purified IgG antibodies from patients with active tuberculosis. IgG antibodies purified from healthy donors was used as negative control [95]. These results suggest that MTP is directly involved in the pathogenesis of *M. tuberculosis*. 
Despite, the many similarities between MTP and curli and Tafi fimbriae, no obvious primary sequence homology were seen between them. The only parallel is the occurrence of a large fraction of glycine and proline residues, which account for the hydrophobicity that is a common property of pilins [95]. The open reading frame that encodes MTP is not organized in an operon or clustered with genes that are commonly associated with pili biogenesis. This is in sharp contrast to the well organized genetic structure for curli and Tafi production [92, 95].

Although an amyloid structure was hinted for the MTP, no direct biophysical data was provided to support this suggestion. This hypothesis may be verified using a combination of FTIR, CD, ThT binding, and X-ray fiber diffraction studies.

The purification of MTP can be done in the following way. Bacteria are grown on 7H11 agar plates lacking OADC for two weeks at 37°C before being suspended in PBS buffer. MTP are sheared of the cell surface by vigorous vortexing with glass beads. Cell debris are removed by gentle centrifugation steps. Contaminating lipids are removed by extractions with equal volumes of a 2:1 chloroform:methanol mixture. The MTP is finally pellet by ultracentrifugation [95]. The MTP enriched preparation show no bands corresponding to fimbrin monomers when examined by SDS-PAGE. However, when acid hydrolysates of the MTP preparation are examined by liquid-chromatography MS/MS the fimbrin can be identified [95]. It might be possible to depolymerize the MTP using concentrated formic acid of TFA as have previously been done for Tafi and curli [76, 81].

1.2.3 Harpins

Harpins are heat stable, glycine-rich, type III-secreted proteins expressed by plant pathogenic bacteria, which cause a hypersensitive response (HR) when infiltrated into the intercellular space of tobacco leaves [96, 97]. HpaG is a harpin produced by *Xanthomonas axonopodia* pv. *glycines*. It is composed of two predicted α-helix domains (motif 1 and 3) situated in the N- and C-terminal regions, and a Gln and Gly repeated sequence (motif 2) located between the two α-helices (Figure 8A). The motif 2 region is homologous to the prion-forming domain of the yeast prion protein Rnq1p [98].

---

**Figure 7: Mycobacterium tuberculosis Pili (MTP) visualized by TEM.** (A) MTP produced by the clinical isolate CDC1551. (B) Purified MTP. (Magnification: ×45,000). Reprinted from [95] with permission from the National Academy of Sciences.
Recombinant HpaG produced as a his-tagged protein (His$_6$-HpaG) is purified as a tetramer and exhibit bead- and annular-like spherical forms similar to spherical oligomers formed by A$_B$ protein when examined by TEM (Figure 8B). It is furthermore able to induce HR in tobacco leaves (Figure 8D). The mutant His$_{56}$-HpaG(L50P) is in contrast purified as a monomer and do not form oligomers (Figure 8C) [96]. His$_{56}$-HpaG(L50P) is not able to induce HR (Figure 15D). This indicates that the ability to induce HR is linked to the formation of active oligomers. After incubation in 20mM Tris-HCl, 10mM NaCl at pH 8.0 for 4-8 days at 27°C His$_{56}$-HpaG forms elongated protofibrils, which mature into fibrils after 10 days. Fibril formation was also tested under plant apoplast-like conditions. At these conditions mature fibrils are formed already after 3 days. The mature fibrils are able to induce HR to the same extent as the tetramer [96].

To determine if the fibrillar structure of HpaG is amyloid, CR binding was examined. It was found that HpaG fibrils bind CR in an amyloid specific manner, causing a red shift in the absorbance spectrum with maximum difference at 541nm compared to free CR. CR stained fibrils also produce green-yellow birefringence when examined with polarized light. Secondary structure analysis of the fibrils by CD showed a classical signature of a β-sheet protein, compatible with an amyloid conformation [96]. Put together these findings show that the harpins do indeed form amyloid fibrils. Other harpins were also shown to form fibrils at similar conditions, indicating that fibril formation is a common feature of harpins [96].

Harpins represent an interesting example of how bacteria can benefit from the cytotoxic nature of the amyloids or protofibrils. A better understanding of the involvement of amyloids in the interaction of plant with plant pathogenic bacteria might provide new insight into the nature of amyloidogenic proteins as well as into the mechanism of cell death due to these proteins.

1.2.4 Chaplins
The life cycle of filamentous bacteria such as *Streptomyces coelicolor* resembles that of filamentous fungi. After formation of feeding submerged mycelium, aerial hyphae is formed which eventually septate into chains of spores. The spores are dispersed by wind or insects to give rise to new vegetative mycelium [100, 101]. When the submerged hyphae grow into the air their surface properties is dramatically changed. The surface of the submerged hyphae is hydrophilic, whereas that of the aerial hyphae are hydrophobic. The hydrophobicity correlates with the formation of a rodlet layer which consists of a mosaic of 8-10nm wide rods (Figure 9). The rodlet layer resembles that formed by class I hydrophobins on aerial hyphae of filamentous fungi (See Chapter 1.2.5).
However, the proteins which make up rodlet layer in the bacteria are evolutionary distinct from the hydrophobins [100, 102, 103].

Figure 9: Ultrastructure of the rodlet layer on spores of wild-type Streptomyces coelicolor. Image was obtained by using scanning electron microscopy. Reprinted from [99] with permission from American Society for Microbiology.

Two classes of protein have been shown to be involved in the formation of the rodlet layer. The rodlin proteins RdlA and RdlB are responsible for organization of the rodlet layer. Mutants lacking either rdlA or rdlB result in the absence of the rodlet layer. Instead a layer of parallel 4-6nm fibers is seen. These fibers are composed of chaplin proteins and are believed to have amyloid structure as judged by ThT binding, CD, and the ability of these proteins to self assemble in vitro [100, 101]. The chaplin family consist of eight members, ChpA-H, which all have a conserved hydrophobic domain called the chaplin domain (Figure 10). The chaplins contain three regularly spaced GN residues, a motif that is also seen in major curli subunit, CsgA. The chaplins can be grouped in two according to their size and structure. The long chaplins, ChpA-C, contains two chaplins domains along with an C-terminal cell-wall sorting signal (CWS) containing a LAXTG motif followed by a hydrophobic region and a charged tail [104]. The CWS signal result in covalent attachment of the long chaplins to the cell-wall. The short chaplins, ChpD-H, only contain one chaplin domain [100]. It is hypothesized that the long chaplins, which are bound to the cell surface, act as a fibrillation prone nuclei for the fibrillation of the short chaplins in a manner similar to that of CsgB mediated CsgA fibrillation in the curli system [88, 100].

<table>
<thead>
<tr>
<th>Long Chaplins</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
<th>70</th>
<th>80</th>
<th>90</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>ChpA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ChpB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ChpC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ChpD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ChpE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ChpF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ChpG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ChpH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Short Chaplins</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
<th>70</th>
<th>80</th>
<th>90</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>ChpD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ChpE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ChpF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ChpG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ChpH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 10: Amino acid sequence of long and short chaplins from Streptomyces coelicolor. Chaplin domains are highlighted in light gray and conserved GN motifs within these are marked by asterisks at the bottom. The cell-wall sorting signals of the long chaplins are highlighted in dark gray. The LAXTG signal (LAXTG), the hydrophobic region (H) and the charged tail (C) are marked at the bottom.
Short secreted chaplins have been purified in the following way. Bacteria are grown on cellophane discs on the surface of MS medium for four days. The resulting biomass is harvested, suspended in HEPES buffer and sonicated. The sample is centrifuged and the pellet, containing chaplins, is collected and boiled in 2% SDS. The sample is then centrifuged and the supernatant containing SDS soluble proteins discarded [103]. The chaplin enriched pellet is lyophilized, dissolved in concentrated TFA and lyophilized again. The sample can then be dissolved SDS-PAGE loading buffer and subjected to SDS-PAGE [100, 103].

1.2.5 Hydrophobins

Hydrophobins constitute a class of low molecular weight proteins expressed by filamentous fungi. They are secreted as soluble proteins, but when they interact with an interface (e.g. medium-air or cell-wall-air) they structurally rearrange and polymerize to form amphipathic membranes [105]. The expression of hydrophobins has multiple roles in fungal growth and development. Hydrophobins act as natural surfactants and reduce the surface tension of the growth medium and thereby allow fungi to breach the water-air interface and produce aerial structures such as aerial hyphae [106, 107]. Spores, which mature at the tip of the air exposed hyphae, are also coated by a layer of hydrophobins that renders their surface hydrophobic. This makes the spores resistant to wetting and facilitates their effective dispersal in air [102]. Coating of hyphae with hydrophobins also allows the fungus to attach to hydrophobic surfaces such as Teflon [108]. Fungi also use hydrophobins in the interaction with living organism, both as pathogens and mutualists [102].

The hydrophobins can be divided into two classes according to solubility characteristics and hydrophathy patterns [106]. Class I hydrophobins generate highly insoluble assemblies, which resists boiling in 2% SDS and has only been found to be depolymerized using TFA or formic acid [109, 110]. Class II forms assemblies that are less stable. These assemblies can be dissolved in ethanol or 2% SDS or by applying pressure or cooling [105, 106]. The hydrophobic side of the interface membrane created between growth medium and air by class I hydrophobins has been shown to contain of a mosaic of rodlets by AFM and EM [111]. The rodlets have a diameter of 5-15nm, depending on the hydrophobin, and lengths of >100nm [111]. The class I hydrophobin rodlets show many amyloid characteristics. They have a high degree of β-sheet structure, show an amyloid pattern by X-ray fiber diffraction with reflections at 4.8 and 10-12Å, and are resistant to proteases [106, 112, 113]. They have also been reported to bind ThT and CR in an amyloid specific manner [114-116]. No rodlets have been observed for class II hydrophobins.

The primary structure of the hydrophobins are characterized by a conserved pattern of eight cysteine residues that form four disulphide bonds, but otherwise they share little amino acid sequence homology (Figure 11). They do however show very similar hydrophobicity plots [102]. The many disulphide bonds help the hydrophobins to stay soluble in solution, and therefore represent a way to control unwanted aggregation and fibrillation [114].
**Class I hydrophobins**

<table>
<thead>
<tr>
<th>SC4</th>
<th>HYD4</th>
<th>HFBII</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGSG-PVQ---CINNTTT---VANMQ---KGQSLDGG---LLQYVV---GPTQGLVIGNP---ISVGG---LTVQQ---TA-QVCD---HVTGQ---LNN---VQG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HYD4</td>
<td>CPTQQ---ASSPRTVAIL---LLQYV---LADDNLVIGGP---LTVGG---GQSGQ---AQVCTQGNTRYF---LNN---IGQ</td>
<td></td>
</tr>
<tr>
<td>HFBII</td>
<td>CPTQQ---ASSPRTVAIL---LLQYV---LADDNLVIGGP---LTVGG---GQSGQ---AQVCTQGNTRYF---LNN---IGQ</td>
<td></td>
</tr>
</tbody>
</table>

**Class II hydrophobins**

<table>
<thead>
<tr>
<th>HFBII</th>
<th>HCF1</th>
<th>PRI2</th>
</tr>
</thead>
<tbody>
<tr>
<td>HFBII</td>
<td>HCF1</td>
<td>PRI2</td>
</tr>
</tbody>
</table>

**Figure 11:** Amino acid sequence of class I and II hydrophobins. Only amino acids between the first and last Cys residues are shown due to high sequence variations at the termini. The conserved Cys residues are highlighted in gray background with the conserved disulfide bonding pattern indicated with brackets. The relatively low degree of sequence conservation is apparent. The abbreviations used are: SC4, *S. commune* (accession P16933); SC4, *S. commune* (accession P16934); EAS, *N. crassa* (accession Q04571); MPG1, *M. grisea* (accession P52751); HCF1, *C. fulvum* (accession Q00367); ABH1, *A. bisporus* (accession P49072); PRI2, *A. aegerite* (accession Q9Y8F0); RODA, *A. fumigatus* (accession P41746); HFBII, *T. reesei* (accession P52754); HFBII, *T. reesei* (accession P79073); CU, *O. ulmi* (accession Q65135); CRP, *C. parasitica* (accession P52753); HCF6, *C. fulvum* (accession Q9C2X0); MPG, *M. grisea* (accession O94196); HYD4, *G. maniliformis* (accession Q6YF29) and SRH1, *T. harzianum* (accession P79072). Reprinted from [102] with permission from Elsevier.

Due to their special properties numbers of applications have been proposed for the hydrophobins. Biocompatibility plays a major role in the field of medical implants, as insufficient biocompatibility will result in the rejection of the implant. The ability of hydrophobins to alter the hydrophathy of surfaces can improve cell adhesion to implants and thus improve the biocompatibility [117]. Enzyme immobilization on biosensors often results in protein denaturation [105]. Hydrophobins coating can be used to alter the hydrophathy of biosensors and thus prevent denaturation of immobilized protein. This has been shown to work for enzymes that are covalently attached as a fusion protein to the hydrophobin coating and for non-covalently attached enzymes [105]. Hydrophobins have furthermore received attention in the cosmetic and food industry due to their stable emulsifying properties [105].

Hydrophobins can be purified based on their ability to self assemble at air-medium interfaces. The medium from five day old standing cultures is cleared by centrifugation. The medium containing soluble hydrophobins is mixed vigorously with air, resulting in aggregation of the hydrophobins, which are then collected by centrifugation. The pellet is washed with water and lyophilized. The lyophilized material is dissolved in concentrated TFA, lyophilized, resuspended in tris buffered 60% ethanol (pH 8.0) and dialyses against water. The hydrophobins are finally precipitated by bobbling nitrogen through the solution. If the sample contains lipid contaminations these can be extracted with an equal volume of 2:1 chloroform:methanol [110].

### 1.2.6 Yeast Prions Provide Non-Mendelian Traits

Prions protein were originally discovered in relation to the transmissible spongiform encephalopathies [70]. These diseases may be genetic, sporadic, or infectious [118]. However, a common trait is that they all involve modification of the prion protein (PrP), a constituent of normal mammalian cells. The normal, cellular PrP, denoted PrP<sup>c</sup>, is converted into PrP<sup>Sc</sup> through a process,
whereby a portion of its α-helical and coiled structure is refolded into an amyloid β-sheet [70, 118]. PrPSc can cause PrPSc to become amyloid, thus propagation itself [70].

The prion hypothesis has been extended with the finding that several non-Mendelian traits in fungi are due to heritable changes in protein conformation. These may in some cases be beneficial [119]. Ure2p and Sup35p represent two potentially functional prions in the yeast *Saccharomyces cerevisina* [9, 70, 73]. In its soluble form, Ure2p regulates the nitrogen catabolism by repressing the Gln3p transcription factor that controls expression of a series of genes involved in the uptake of poor nitrogen sources [119]. When Ure2p is aggregated it loses its ability to bind Gln3p, which in turn is constitutively activated. The resulting yeast cells [URE3] can grow on media that, for example lack uracil but contain its precursor ureidosuccinate [119]. Sup35p is a protein involved in the termination of mRNA translation. Aggregation of Sup35p prevents translation termination and results in stop-codon read-through and protein C-terminal elongation [120-122]. This result in yeast cells [PSI+] which have altered phenotypes, presumably due to the creation of an altered proteome [121, 122]. The natural occurrence of [URE3] and [PSI+] strains is low. Therefore it has been suggested that the corresponding phenotypes are generally deleterious [123]. However, they can still be advantageous under particular environmental conditions [9].

In filamentous fungi vegetative cell fusions occur within and between individual cells. These fusion events lead to cytoplasmic mixing and the production of vegetative heterokaryons (i.e., cells containing different nuclear types) [124]. In *Podospora anserina* heterokaryon formation is regulated by the prion HET-s [70, 125]. The het-s locus has two alleles, het-s and het-S. HET-S is the soluble protein product, whereas HET-s has the ability to convert into an aggregated prion state. When fusion occur between a cell with aggregated HET-s and one with soluble HET-S, a heterokaryon incompatibility reaction takes place, which kills the heterokaryon and creates a barrier between the two colonies to prevent further fusion. However, fusion can take place between similar cells and between those in which HET-s is soluble and those expressing HET-S [70, 73, 125].

### 1.2.7 Silkmoth Chorion

Chorion is the major component of the eggshell of many insects such as for the silkmoth. It is highly proteinaceous (ca. 95%) and make up the outer layer of egg shell. Due to its extraordinary mechanical and physiological properties, the chorion acts as protection for the oocyte and developing embryo toward environmental stresses such as temperature variations, mechanical pressure, proteases, bacteria, and viruses [126-128]. The silkmoth chorion proteins can be classified into two major classes, A and B. Proteins of both classes are organized into three domains. The central domain is conserved within both classes and it contains characteristic hexapeptide tandem repeats with regularly spaced glycine residues (Figure 12). The flanking N- and C-terminal domains show more variation, but do also contain characteristic tandem repeats. These repeats are different from the ones in the central domain [128]. The variable terminal regions endow the proteins with specific functional and structural attributes [129].
When thin transverse sections of silkmoth chorion are examined by TEM, a lamellar ultrastructure of packed fibrils is seen (Figure 13). A combination of X-ray fiber diffraction, FTIR, and Raman spectroscopy have shown that anti-parallel β-sheets are the dominant secondary structure in the chorion proteins [126]. These results suggest that the chorion is composed of amyloid fibers. This is further supported by studies made on a 51-residue peptide analogue of the entire central domain of a class A chorion (cA) and a 18-residue peptide of the central domain of a class B chorion. Using TEM, CR birefringence, X-ray fiber diffraction and FTIR, both of these peptides were found to form amyloid fibrils in vitro [127, 130].

The first step in the fibrillogenesis of the chorion peptide cA is the formation of nuclei of liquid crystalline nature. These liquid-crystalline nuclei then collapse or deteriorate to form amyloid fibrils (Figure 14) [127].

---

**Figure 12:** Alignment of the central domains of chorion proteins from the A and B classes of *Mombyx mori*. Asterisks at the bottom mark the regular spaces glycine residues.

---

**Figure 13:** Ultrastructure of the silkmoth *Antheraea polyphemus* chorion. The image was obtained by TEM and shows an oblique section through the proteinaceous chorion. Arrays of ca. 10nm protein fibrils (F, arrowheads) are seen in each lamella (L). Arrows pointing toward the outer (O) and the inner (I), closest to the oocyte, surface of chorion is shown on the left side. Reprinted from [126] with permission from Elsevier.

---

**Figure 14:** Transformation from spherulites to immature amyloid-like fibrils. A suspension of 6.5mg/mL cA peptide in a 1:1 methanol:water mixture was examined for a period of one week using TEM. It is clearly seen that spherulite numbers decrease gradually, with a concomitant increase of straight amyloid-like fibrils. Reprinted from [127] with permission from Elsevier.
The molecular switch from spherulites to amyloid fibrils has been suggested to be a conformational transition from a left-handed parallel β-helix to an anti-parallel β-pleated sheet (Figure 15A). In the latter, the regular spaced glycine residues occupy the second position in the β-turns (Figure 15B). The mature amyloid-like fibrils are observed as uniform, unbranched, straight or slightly curved 9nm wide fibrils with indeterminate length. Each fibril is composed of two 3-4nm protofilaments that form a double helix [126, 127].

Figure 15: Proposed model for molecular switching of cA peptide from spherulites to amyloid-like fibrils. (A) In the spherulites the cA peptide is folded in a left-handed parallel β-helix conformation (left). The transformation into amyloid-fibrils is accompanied with a structural reorganization into an anti-parallel twisted β-sheet conformation (right). (C) Sequence representation of the anti-parallel twisted β-sheet model proposed for the cA peptide. Sequence should be read continuously, beginning at the bottom. Regularly spaced glycines occupying the second position in the β-turns are black-boxed. Tentative II’ β-turns alternate with four-residue β-strands. Reprinted from [126] with permission from Elsevier.

As chorion is used in by many insects and also by fish, this type of functional amyloid represents a valuable model system for the study of amyloid formation. Information from the chorion destruction mechanism used by the embryo upon hatching might furthermore provide medical relevant information on how to avoid disease related amyloid formation [126].

1.2.8 Pmel17

Figure 16: Melanin biogenesis. Pmel17 is synthesized as a transmembrane protein in the endoplasmic reticulum (ER) (1); it is trafficked first to the Golgi (2) and finally to melanosome organelles. Proprotein convertase (PC) cleavage in a post-Golgi compartment (3) creates a luminal fragment, Mα, and a transmembrane fragment, Mβ. Mα remains disulfide-bonded to Mβ to prohibit aberrant Mα amyloidogenesis (3). Degradation of Mβ releases Mα from the membrane (4), enabling Mα to form amyloid fibers within the melanosome organelle (5). Mα amyloid fibers orchestrate the synthesis of melanin from tyrosine-derived reactive indolequinones and protect the melanocyte from melanin-associated toxicity (6). Adapted from [70] with permission from Elsevier.

Melanocytes and retinal epithelium are specialized cell types responsible for the biogenesis of melanin, a tyrosine based polymer that function as a chemical defense against pathogens, toxic small molecules, and UV-radiation [131]. The mammalian biogenesis of melanin takes place within specialized membrane enclosed vesicles termed melanosomes (Figure 16) [73, 131]. The maturation of melanosomes and polymerization of melanin are dependent on insoluble fibers composed of the
80kDa lumenal fragment of Pmel17 called Mα [131]. These fibers show the biochemical hallmarks of amyloids, including binding of thioflavin S and CR resulting in amyloid specific fluorescence, amyloid-like FTIR, CD and X-ray fiber diffraction spectra, and the ability of the monomers to self assemble into amyloid fibrils [131]. The Mα fibrils are believed to kinetically enhance the formation of mature melanin, presumably by acting as a template for the polymerization of reactive melanin precursors [73, 131]. Several mechanisms are employed by the cells to minimize the exposure to toxic intermediates formed during the Mα fibrillation process. (A) Fibrillation is spatially restricted to the melanosome and folding intermediates are thus separated from the cytoplasm. (B) Liberation and fibrillation of Mα is initiated by the proteolysis of Pmel17, a trans-membrane protein unable to self-assemble. (C) The reaction kinetics/thermodynamics favor the formation of stable mature amyloid fiber and limits the time window open for intermediates [70, 73, 131].

The Pmel17 functional amyloid system shows for the first time that amyloid fold is not only associated with disease, but actually play an important role in normal mammalian cell physiology. Studies of Mα may contribute with important information of the pathological basis of protein misfolding diseases.

1.3 Aim of this Project
It is well established that amyloid formation is a keystone in the development of a variety of devastating human diseases. Despite the considerable effort put into elucidating the nature of amyloid toxicity and the initiation of amyloid formation, much still remains to be learned.

Functional amyloids systems contain highly amyloidogenic proteins which are under the strict control of sophisticated regulation to avoid the pit falls of amyloid toxicity. Research on functional amyloids systems may therefore provide valuable information on the molecular mechanisms of amyloid toxicity and initiation of amyloid formation.

The aim of this project is a thorough biophysical characterization of functional bacterial amyloid in vitro and in vivo. These studies may provide knowledge on the use of functional amyloids in natural bacterial systems and on the mechanisms of amyloid formation and toxicity.
2. Papers

2.1 Paper 1

With reference to Ministerial Order no. 114 of 8 March 2002 regarding the PhD Degree § 14, article 3, statements from each author about the PhD student’s part in the shared work must be included in case the thesis is based on already published articles.

**Co-author statement concerning the contribution by Morten S. Dueholm**

- M. S. D. performed the optimization of growth conditions and biophysical characterization of the two control strains (SM2257 and SM2258)
- M. S. D. isolated 4 out of 6 of the pure cultures described in table 3 on Congo red indicator plates and performed the identification by 16s rRNA sequencing.
- M. S. D. participated in design and evaluation of the experimental work.
- M. S. D. participated in preparation of the manuscript.

Poul Larsen 14/09-2009
Jeppe L. Nielsen 14/09-2009
Morten S. Dueholm 14/09-2009
Ronald Wetzel 10/09-2009
Daniel E. Otzen 15/09-2009
Per H. Nielsen 14/09-2009

Signature Date
2.1.1 Motivation

Although the existence of FuBA have been known since 2002, only a few FuBA systems have been described and little is known of the abundance of FuBA in natural systems. However, the fact that most polypeptides have the propensity to form amyloid in vitro, suggests that the amyloid fold could be a common structure in nature [132]. The work of paper 1 was carried out with the intention of developing a method to investigate the presence on amyloid structures in natural biofilms. This method should than be used to estimate the abundance of FuBA in environmental biofilms from different habitats and also link FuBA expression to various bacterial phyla.

2.1.2 Outcome

The first result of paper 1 was the development of a method for in situ detection and quantification of bacteria expressing FuBA in natural biofilms using confocal laser scanning microscopy (CLSM). This method relied on the use of the amyloid specific fluorescent dye ThT or the amyloid specific conformational antibodies WO1 and WO2 in combination with either DAPI, which stains all bacteria, or oligonucleotide probes by FISH, which stains specific bacterial phyla. A curli expressing and a curli deficient E. coli K-12 strain were used for the method development. Cross straining with different amyloid specific probes was used to confirm the selectivity of the probes. The amyloid specific antibodies proved to be very selective toward amyloid structures. ThT stained more bacteria and was therefore suggested to be less selective. However, the fact that ThT stains more bacteria could also be a result of a better permeability of ThT in the biofilms compared to the antibodies, due to the considerable smaller size.

The newly developed tool for in situ examination of FuBA in environmental biofilms was used to quantify the fraction of biofilm containing FuBA in biofilms isolated from multiple habitats. It was found that approximately 5-40% of the biofilm fraction contained FuBA depending on habitat. The highest abundance was found in a biofilm isolated from a drinking water reservoir and the lowest in biofilms from waste water treatment plant (WWTP) digesters.

The combination of FISH and amyloid probes allowed us to link FuBA expression with specific bacterial phyla. It was found that FuBA was expressed by several phyla including Proteobacteria, Bacteriodetes, Chloroflexi, and most likely Actinobacteria. Both amyloid positive and negative cells were found within these phyla.

This paper has indicated that FuBA are very common in natural biofilms. They may therefore play an important part in the architecture of these bacterial communities. The fact that FuBA was expressed in several bacterial phyla indicates that the amyloid structure has been selected many times during evolution and may thus represent a fundamental structural motif.
Amyloid adhesins are abundant in natural biofilms

Poul Larsen,1 Jeppe Lund Nielsen,1 Morten Simonsen Dueholm,2 Ronald Wetzel,3 Daniel Otzen2 and Per Halkjær Nielsen1* 
Department of Biotechnology, Chemistry and Environmental Engineering, 1Section of Environmental Engineering, 2Section of Biotechnology, Aalborg University, Søhngaardsholmsvej 57, DK-9000, Aalborg, Denmark. 3Department of Structural Biology, Pittsburgh Institute for Neurodegenerative Diseases, University of Pittsburgh School of Medicine, Pittsburgh, PA 15260, USA.

Summary

Surface-associated amyloid fibrils have been described by bacteria in the family Enterobacteriaceae, but it is unknown to what extent amyloid adhesins are present in natural biofilms. In this study, amyloid adhesins were specifically stained with Thioflavin T and two conformationally specific antibodies targeting amyloid fibrils. These three independent detection methods were each combined with fluorescence in situ hybridization using fluorescently labelled oligonucleotide probes in order to link phenotype with identity. Escherichia coli mutants with and without amyloid adhesins (curli) served as controls. In biofilms from four different natural habitats, bacteria producing extracellular amyloid adhesins were identified within several phyla: Proteobacteria (Alpha-, Beta-, Gamma- and Deltaproteobacteria), Bacteriodetes, Chloroflexi and Actinobacteria, and most likely also in other phyla. Quantification of the microorganisms producing amyloid adhesions showed that they constituted at least 5–40% of all prokaryotes present in the biofilms, depending on the habitat. Particularly in drinking water biofilms, a high number of amyloid-positive bacteria were identified. Production of amyloids was confirmed by environmental isolates belonging to the Gammaproteobacteria, Bacteriodetes, Firmicutes and Actinobacteria. The new approach is a very useful tool for further culture-independent studies in mixed microbial communities, where the abundance and diversity of bacteria expressing amyloid adhesins seems much greater than hitherto anticipated.

Introduction

The great majority of bacteria in different habitats grow in sessile communities known as biofilms. These are matrix-enclosed bacterial populations adherent to each other and/or to surfaces or interfaces. In the medical field, bacteria in biofilms may colonize mucosal surfaces as well as invasive devices and implants, causing serious infections (Costerton et al., 1995; Silverstein et al., 2006). In technical systems, biofilms may cause biofouling, biodeterioration and biocorrosion (Gibson et al., 1999; Kjellerup et al., 2005), but they also serve a useful purpose in water and wastewater treatment and in several biotechnological processes. Thus, microbial biofilms can be of great benefit for humans, but they can also be a nuisance, causing maintenance expenses or life-threatening diseases. Therefore, it is extremely important to understand the structure and function of biofilm communities as well as the mechanisms regulating biofilm formation and processes.

The bacterial surface structures important for adhesion, and later development of mature biofilm, are primarily proteinaceous adhesions covering a broad group of different fimbriae and non-fimbriated structures (Klemm and Schembri, 2000; Barnhart and Chapman, 2006; Latasa et al., 2006). One such important surface structure is the curli fibril produced by Escherichia coli (Olsen et al., 1989). The length of curli can vary from about 0.1 to 10 μm, with a width of 4–12 nm. Curli consist mainly of the csgA protein, which is assembled in a fibrillar tertiary structure on the outer bacterial membrane (Hammar et al., 1995). Proteins folded as β-sheet-rich fibrils are also known as amyloid fibrils, and they are folded as β-sheets stacked perpendicular to the fibril axis (Chapman et al., 2002). This is a general property of fibrillated proteins regardless of the structure or composition of the original protein, and such amyloids are known in neurodegenerative diseases such as Alzheimer’s and Parkinson’s diseases in humans and formation of prions in other mammals (Pan et al., 1993; Goedert and Spillan- tini, 2006). Most amyloids are highly insoluble in water and can be resistant to thermal and chemical denaturants, particularly prions and curli. Shortly after the discovery of curli on E. coli, presence of adhesions with similar amyloid-like structure was also described for other members of the family Enterobacteriaceae in the Gammaproteobacteria,
such as Salmonella typhimurium, Citrobacter spp. and Enterobacter sakazaki (Collinson et al., 1993; Zogaj et al., 2003). Curli fibrils are described as having different roles in pathogenesis, owing mainly to their specific binding to fibronectin, laminin, plasminogen and human contact phase proteins. This may allow these pathogens to colonize and invade host tissues. Furthermore, exposure to exogenous amyloid fibrils may increase the risk of amyloid diseases in at-risk individuals (Solomon et al., 2007).

Other types of amyloid-like adhesions have also been described on the surface of certain bacteria and fungi (Gebbink et al., 2005). Gram-positive Streptomyces (S. coelicolor) secrete hydrophobic proteins forming amyloid-like fibrils called chaplins, which are believed to be important for the formation of aerial hyphae at the water–air interface (Claessen et al., 2003). Hydrophobins are secreted by fungi and may be an important class of hydrophobic proteins promoting adsorption to surfaces, also often important at the water–air interface (Wosten, 2001). Furthermore, transmission electron microscopy (TEM) studies have shown presence of fibrillated structures with a diameter of 4–6 nm in river biofilms and microbial aggregates (flocs) from activated sludge wastewater treatment plants (Liss et al., 1996). The exact composition of these structures is unknown, but these observations suggest that amyloid surface structures may be more widespread among many microorganisms than hitherto anticipated, but it has not yet been investigated.

Detection of bacterial amyloid adhesins in pure culture studies is usually based on the binding of the dye Congo red to the β-sheet of the proteins. In this way, bacteria that produce amyloid-like adhesins are easily identified by red colonies when growing on agar plates containing Congo red (Collinson et al., 1993). This method was validated by Hammar and colleagues (1995), who found complete correlation between Congo red, autoaggregation assay and TEM for detection of the amyloid curli on positive and negative E. coli strains. Amyloid fibrils can also be stained by other methods, but, to our knowledge, these have not been used to detect bacterial amyloid adhesins. Besides Congo red, also Sirius red can be used with bright-field or polarized light microscopy (Westermark et al., 1999) and also TEM showed presence of long amyloid fibrils on prokaryotes producing amyloid structures in natural biofilms were high in several environmental habitats.

Results

Thioflavin T staining and antibody labelling of amyloid adhesins

Two E. coli mutants and aerial hyphae from S. coelicolor were used to test various methods for detection of amyloid adhesins on a single-cell level with the use of a confocal laser scanning microscope (CLSM). The results for using Congo red staining, ThT staining and fluorescently labelled antibodies are presented in Table 1. The mutant producing curli (OmpR 234 Tc6) and hyphae from S. coelicolor stained positive with all methods applied, and also TEM showed presence of long amyloid fibrils on

<table>
<thead>
<tr>
<th></th>
<th>Congo red CFA agar plates</th>
<th>Congo red fluorescence</th>
<th>ThT</th>
<th>WO1</th>
<th>WO2</th>
</tr>
</thead>
<tbody>
<tr>
<td>E. coli (OmpR 234 Tc6)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Spores from S. coelicolor</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>E. coli (CsgDAB::Kan)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

+ indicated presence of amyloids; – indicate absence of amyloids.

Two positive controls [E. coli (OmpR 234 Tc6)] and spores from S. Coelicolor and one negative control [E. coli (CsgDAB::Kan)] were applied.

Table 1. Validation of the detection methods of amyloids on pure cultures.
the E. coli mutant. Congo red fluorescence gave a strong signal but bleached very fast. ThT staining gave a bright fluorescence, and the fluorescence emission spectrum was identical to the spectrum reported to be induced by binding of other amyloid proteins to ThT (Levine, 1993). Also the two types of antibodies (WO1 and WO2) bound to the cells tested, providing a clear fluorescence. In all cases, the E. coli mutant unable to produce amyloid adhesins (CsgDAB::Kan) was negative with the two stains and the two antibodies, showing that these did not target other cell surface components (Table 1). An example of antibody binding (WO1) to the two mutants is shown in Fig. 1m and n.

The different methods were also tested on undefined mixed cultures from biofilms. Congo red staining, ThT staining and antibodies all resulted in a clear fluorescent signal, with variable intensities of microcolonies of bacteria with different morphologies (Fig. 1). However, again Congo red bleached very fast. ThT and both antibodies emitted clear, stable fluorescent signals from certain microcolonies in the biofilms, while other microcolonies and cells remained unstained (Fig. 1). This indicated presence of amyloid adhesins on the surface of certain, but not all bacterial cells or microcolonies in the biofilm. In order to investigate whether the same microorganisms stained with Congo red and ThT, a number of tests were made with double staining of the same biofilms. As shown in Fig. 1a–d, both ThT and Congo red stained the same microcolonies, indicating that they stained the same structures in the biofilms. In contrast, when fluorescent antibodies were tested together with the ThT staining, not all ThT-positive microcolonies were positive with the antibodies. This indicates that ThT and Congo red also may stain non-amyloid structures, that some amyloids had another structure that failed to bind to the antibodies, or that some amyloids were masked by other surface components in such a way that the antibodies did not bind.

Cells stained with ThT and antibodies could also be counterstained with 4′,6-Diamidino-2-Phenylindole (DAPI) that stains all microorganisms in a biofilm sample. The best combined signal from staining of cells and staining of amyloid adhesins was achieved using relocation (see Experimental procedures). This enabled quantification of the biofilm fraction positive with ThT or antibodies relative to the total amount of cells using digital image analysis. Two examples of this combination are shown in Fig. 2e–h.

Combination of ThT and antibodies with fluorescence in situ hybridization

The identity of microorganisms producing amyloid adhesins in biofilms was assessed by combining ThT or antibodies with fluorescence in situ hybridization (FISH). Figure 2a–d shows an example of a ThT staining of a wet fresh sample. After capture of images from a number of ThT-positive bacteria, the sample was dried, fixed, hybridized with gene probes, digital images of FISH signals captured from the same bacteria as were stained with ThT, and a digital overlay of the ThT-positive and the FISH-positive microorganisms was produced. Also a combination with antibody-FISH was developed. By using a range of oligonucleotide probes targeting different phylogenetic groups, it was possible to identify antibody-positive cells within the biofilms. Due to steric interactions between the gene probes and the gelatine that was applied for blocking of unspecific sites in the matrix, the signal quality from the gene probes was variable in some samples. However, whenever possible, the antibody labelling was preferred due to the potential unspecificity of the ThT staining as indicated above.

Quantification and identification of ThT and antibody-positive microorganisms in natural biofilms

The fraction of microorganisms expressing amyloid adhesins in natural biofilms from seven different habitats was quantified by using the combination of antibodies and DAPI. Around 5% to almost 45% of all microorganisms were positive with the antibodies, highest in the drinking water biofilm and lowest in the digesters (Fig. 3). The positive fraction of bacteria in biofilms from Limfjorden, from seawater and from the freshwater lake might be slightly underestimated due to a relatively large fraction of algae, which also stains positive with the DAPI stain. In biofilms from the drinking water reservoir, the freshwater lake and the activated sludge, the antibodies were often bound to apparent monospecies microcolonies (Fig. 2e–p). In the biofilms from other habitats, the antibodies were bound more irregularly, which reflected that the general 3D structure of these biofilms was less organized. The positive signal from the antibodies was usually limited to the surface and the outer regions of dense microcolonies (Fig. 1f and g), while the positive ThT signal on the same microcolonies was homogeneously distributed throughout the colonies, indicating that the penetration of antibodies was limited. Besides a direct binding of antibodies to the cell surfaces, part of the loosely attached extracellular polymeric substances often also stained with both ThT and antibodies. Interestingly, most types of biofilms contained different filamentous bacteria, which stained positive with the antibodies (Fig. 2q–t).

An identification of bacteria staining positive with ThT, WO1 and WO2 in the biofilms was conducted by a screening with 10 different phyla or group-specific gene probes (Table 2). A number of probe-defined groups were...
Fig. 1. a–d. Simultaneous staining of microcolonies in biofilm with ThT and Congo red.
a. Transmission light.
b. Same sample stained with ThT and (c) Congo red. Image (d) is an overlay of (a–c).
c–f. Simultaneous staining with ThT and WO1.
d. Transmission light.
e. Same sample stained with ThT and (g) WO1. Image (h) is an overlay of (e–g).
f–h. Simultaneous staining with ThT and WO2.
g. Transmission light.
h. Same sample stained with ThT and (k) WO2. Image (l) is an overlay of (i–k).
i–l. Simultaneous staining with ThT and WO2.

m–n. WO1 staining of the E. coli-negative control (m) and WO1 staining of the E. coli-positive control (n). White bar represents 20 μm.
### Table 2. Production of amyloid adhesins by probe-defined groups of bacteria in biofilm sampled from different habitats.

| Probe defined groups          | Activated sludge | Biofilm from freshwater lake | Biofilm from Limfjorden | Biofilm from drinking water reservoir
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Abundance</td>
<td>ThT</td>
<td>WO1</td>
<td>WO2</td>
</tr>
<tr>
<td>Alpha proteobacteria (ALF968)</td>
<td>x</td>
<td>ND</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Betaproteobacteria (BET42a)</td>
<td>xx</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Gammaproteobacteria (GAM42a)</td>
<td>x</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Delta proteobacteria (SRB385 + SRB385Db)</td>
<td>x</td>
<td>ND</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Cytophagales (CF319a + b)</td>
<td>x</td>
<td>ND</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Chloroflexi (CFXmix)</td>
<td>x</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Firmicutes (LGC354mix)</td>
<td>x</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Actinobacteria (HGC69a)</td>
<td>x</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Planctomycetes (Pla46)</td>
<td>NP</td>
<td>x</td>
<td>ND</td>
<td>–</td>
</tr>
<tr>
<td>Archaea (ARCH 915)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

**Legend:**
- **Abundance:** NP, the probe-defined group constituted < 1% of all Bacteria, Archaea and could therefore not be evaluated properly; x, 1–30%; xx, 30–60%; xxx, 60–80%.

*a. It was not possible to apply the combination of antibody and any of the tested oligonucleotide probes in the biofilm from the drinking reservoir because of interactions between gelatine and oligonucleotide probe. The amyloids were detected by ThT-FISH or antibody-FISH. Ten phyla or group-specific probes were applied to all biofilm samples. Only groups exhibiting more than 1% of all Bacteria (EUBmix) were investigated. ND: the combination of FISH and antibody was not possible with the oligonucleotide probe in this habitat. For ThT results, ND means that the bacteria targeted with the oligonucleotide probe was absent in the relocated fields of view. +, Some of the FISH-positive bacteria were positive with ThT or antibodies; –, None of the FISH-positive bacteria were positive with ThT or antibodies.

**Note:** The table provides a summary of the production of amyloid adhesins by different probe-defined groups of bacteria in biofilms sampled from various habitats. The abundances are indicated using a scale from 1 to 3 and ND for not detected.
not found in high number and were not further investigated. These are indicated as not present (NP). WO1 and WO2 labelling was in some cases hampered when the gelatine-blocking agent prevented penetration of gene probes, as mentioned above; these are indicated as ND. The ThT or antibody-positive bacteria were primarily found among the phyla *Proteobacteria*, *Bacteriodetes* (*Cytophagales*), and *Chloroflexi* (Table 2). Gram-positive bacteria (which include *Streptomyces*) were only present in very low numbers in all habitats investigated except the wastewater treatment plant sample. Here, both positive and negative microcolonies of *Actinobacteria* were found (Fig. 2a–d). The four habitats were also screened for *Archaea*, without finding enough cells from this group to establish conclusive results. In all samples examined, positive microorganisms were found among the most abundant phylogenetic groups, but not all bacteria in each probe-defined group were positive, most likely because there were several species present (different morphotypes could be observed), or because not all bacteria of a certain species expressed the amyloid adhesins. A large number of *Planctomycetes* were found in the freshwater lake biofilm. These were tested negative with the antibody staining.

In all freshwater biofilms, bacteria affiliating to the *Betaproteobacteria* were abundant. In activated sludge, a large proportion of these were amyloid-positive (Table 2). Furthermore, some so far unidentified microcolonies in the sludge with characteristic cocci morphology (1–5% of the total number of *Bacteria* hybridizing with the EUBmix) showed high fluorescence intensity with ThT. The lake biofilm was dominated by *Betaproteobacteria* as well, and also *Deltaproteobacteria* were abundant. Both groups were ThT-positive, often with different morphologies indicating several species, but only *Betaproteobacteria* were positive with WO1 and WO2. The drinking water reservoir also consisted of many *Alpha-, Beta-, and Gammaproteobacteria*, having many ThT-positive

![Fig. 2](image-url)

**Fig. 2.** a–d. Simultaneous labelling of microcolonies in biofilm with ThT and oligonucleotide probes by FISH in activated sludge.

- a. All bacteria labelled with EUBmix (blue).
- b. Actinobacteria positive with probe HGC69a (red).
- c. Same sample stained with ThT (green) and (d) is an overlay of (a–c).
- e–h. Simultaneous labelling of microcolonies in biofilm with DAPI and antibody (WO1).
- i. Sludge and (g) freshwater lake biofilm stained with DAPI (blue).
- j. The corresponding fields of view stained with WO1 (green).
- k. Same sample stained with WO1 (green) and (l) is an overlay of (i–k).
- m–r. Overlays of WO1 and oligonucleotide probes. Blue is the EUBmix oligonucleotide probe, red is a group specific oligonucleotide probe, and green is antibody WO1.
- t. Overlay of DAPI (green) autofluorescence from cyanobacteria (red) and WO1 (green) in brackish water biofilm from Limfjorden.
- u. Overlay of DAPI and WO1 on activated sludge. White bar represents 20 μm.

![Fig. 3](image-url)

**Fig. 3.** Fraction of biofilm positive with the antibodies WO1 and WO2 in different biofilms. It was quantified by dual staining with antibodies WO1 or WO2 and DAPI.
bacteria. This was, however, not possible to confirm with antibodies due to the above-mentioned blocking problems. However, as the fraction of antibody-positive bacteria in the drinking water reservoir constituted about 45% of total DAPI-positive microorganisms, the antibodies must have targeted different species outside Gammaproteobacteria. In the brackish water biofilm from Limfjorden, Cytophagales hybridizing with probe CFB319a + b was the most abundant group, and many bacteria were positive with ThT and antibodies. Some of these were attached to the sheath presumably from cyanobacteria (Fig. 2s). Alpha- and Gammaproteobacteria were also found in relatively high numbers. Among the Gammaproteobacteria there were indications of ThT-positive bacteria, but the quality of the re-locations was not sufficient to draw clear conclusions. Alphaproteobacteria from Limfjorden were not found to be ThT-positive. Interestingly, some filamentous microorganisms in the activated sludge sample stained positive with the antibodies only at the septa dividing the individual cells within the filaments (Fig. 2t).

Isolation of microorganisms producing amyloid adhesins

A number of isolates producing amyloid adhesins were generated from biofilms sampled in four different environments. The biofilms were homogenized, cultured on Congo red agar plates, and in most cases Congo red-positive colonies were numerous. Some of these were further tested positive with ThT and antibodies (Table 3), supporting that these isolates produced amyloid adhesins. All isolates were identified by analysis of their 16S rRNA gene sequences (Table 3). The results support the in situ results that expression of amyloid adhesins is widely distributed among different Gram-negative and Gram-positive phylogenetic groups growing in different habitats.

Table 3. Bacteria producing amyloid adhesins isolated from different biofilms.

<table>
<thead>
<tr>
<th>Closest relative</th>
<th>Habitat</th>
<th>% identity to closest relative</th>
<th>Congo red CFA agar plates</th>
<th>ThT</th>
<th>WO1</th>
<th>WO2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aeromonas hydrophila</td>
<td>Activated sludge</td>
<td>99</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Class Gammaproteobacteria</td>
<td>Drinking water biofilm</td>
<td>98</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Pseudomonas fluorescens</td>
<td>Drinking water biofilm</td>
<td>96</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Class Gammaproteobacteria</td>
<td>Drinking water biofilm</td>
<td>99</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Chryseobacterium sp.</td>
<td>Drinking water biofilm</td>
<td>96</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Phylum Bacteroidetes</td>
<td>Drinking water biofilm</td>
<td>99</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Pseudomonas fluorescens</td>
<td>Drinking water biofilm</td>
<td>96</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Bacillus licheniformis</td>
<td>Thermophilic digester sludge</td>
<td>99</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Thermophilic digester sludge</td>
<td>Thermophilic digester sludge</td>
<td>99</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

The presence of amyloid adhesins was confirmed on Congo red agar plates, and with ThT and antibody staining.

Discussion

Bacteria producing amyloid adhesins were found to be present in all biofilms tested. The high abundance in most habitats shows that amyloid adhesins are important components of the extracellular matrix of most biofilms, although the exact function is still unknown. Interestingly, the phylogenetic affiliation of bacteria producing amyloid adhesins was found to be distributed within a broad range of gene probe-defined phylogenetic groups and not only Enterobacteriaceae and Gram-positive Streptomyces. Comparing the genes in E. coli and S. coelicolor encoding for the amyloidal structures (results not shown) showed that the genes are very different, which, together with our results, suggests that these functional amyloids during evolution have been independently developed in many microorganisms and that they are of general importance for biofilm formation and structure.

Detection of amyloid adhesins

The reliability of the staining and immunohistochemical methods applied is fundamental for the in situ assessment of presence, quantity and phylogenetic relations of bacteria producing amyloid adhesins. The stains usually applied in the medical field (Congo red and ThT), as well as amyloid-specific antibodies (WO1 and WO2), all proved suitable for in situ detection of curli on E. coli and spores from S. coelicolor using CLSM. These tests clearly showed that the methods all selectively stained curli on the positive control while the negative control remained unstained.

ThT staining was further validated by double staining with Congo red on a natural biofilm. ThT resulted in a characteristic shift in fluorescence excitation/emission maxima upon specific binding to fibrils with amyloid-like structure (Levine, 1993). ThT produced a stable signal
compared with Congo red, which bleached very fast, as also described by Westermark and colleagues (1999). Double staining with ThT and either WO1 or WO2 showed that ThT stained a larger fraction of the bacteria than the antibodies, whereas bacteria staining positive with antibodies in all cases were positive with ThT. The specificity of ThT is not as good as that of the antibodies. Other organic molecules, such as DNA and cellulose, can also be stained by ThT (Raj and Ramaraj, 2001; Ilanchelian and Ramaraj, 2004). Both compounds may be present in the extracellular matrix of biofilms (Matthysse et al., 1981; Jahn et al., 1999), and particularly cellulose can be produced in combination with amyloid fibrils in certain Salmonella and E. coli species (Zogaj et al., 2001). The larger fraction of ThT-positive cells was also partly a consequence of ThT being a much smaller molecule than the antibodies. The positive signal from the antibodies was usually limited to the surface and the outer regions of dense microcolonies, while the positive ThT signal on the same microcolonies was homogeneously distributed throughout the colonies, strongly indicating that the penetration of antibodies was limited. Both antibodies have been tested on several proteinaceous aggregates (O’Nuallain and Wetzel, 2002), and because only proteins folded as amyloid fibrils are stained, these antibodies seem very specific for labelling of amyloid-like fibrils. Furthermore, several types of amyloid fibrils produced from different precursor proteins all resulted in binding of antibodies (O’Nuallain and Wetzel, 2002). Both antibodies gave comparable results in pure culture tests and in the in situ studies, so in future studies it is sufficient to apply only one of the antibodies. The 1 h PFA fixation after labelling with WO1 or WO2 resulted in a very strong binding, allowing the combination with DAPI or FISH without any loss of antibody signal. Combination of ThT and FISH was only possible via a tedious relocation of the two signals, so our general recommendation for future studies is the use of one of the antibodies in combination with FISH.

This new approach with ThT or antibody labelling, in combinations with gene probes for identification of the amyloid-producing microorganisms in natural biofilms, expands our possibility of investigating structure and function of complex microbial communities with cultivation-independent methods. This toolbox already includes methods for the analysis of, e.g. substrate uptake, exo-enzymatic activity and surface properties (Zita and Hermansson, 1997; Klokeoe and Geesey, 1999; Lee et al., 1999; Neu and Swerhone, 2001; Xia et al., 2007).

**Quantity and diversity of bacteria producing amyloid adhesins**

The quantification of amyloid-positive bacteria in seven different habitats revealed that they constituted a significant part, 5% to 40%, of all DAPI-positive bacteria. The data support TEM studies performed by Liss and colleagues (1996) and Leppard (1986), showing that ‘fibrillated’ structures in a size range similar to amyloid adhesins were abundant in natural biofilms. Interestingly, the tendency to produce amyloid adhesins seemed more pronounced in the oligotrophic environment such as the drinking water reservoir. This indicates that the genes responsible for production of amyloid adhesins were expressed to a greater extent under these conditions, which matches the observation that starvation triggers production of curli in *S. typhimurium* strains (Gerstel and Romling, 2001). The fraction of bacteria producing amyloid adhesins in each digester sludge sample was only about 5%, but more investigations are needed to explain these differences.

The screening of four different habitats for bacteria producing amyloid adhesins showed that these were distributed among several phyla. Mainly bacteria belonging to the Enterobacteriaceae (Zogaj et al., 2003) in the Gammaproteobacteria, such as *E. coli* and *S. typhimurium*, are known to produce amyloid adhesins, but this study has shown that bacteria belonging to other groups of Proteobacteria, Bacteroidetes, Chloroflexi, and most likely Actinobacteria and others also form these structures in natural systems. The use of broad gene probes did not allow a complete identification of the microorganisms, and each probe most likely covered the presence of many species in the biofilms investigated. This was supported by the observation of several morphotypes in each gene probe-defined group, and presence of bacteria, which was both positive and negative with ThT, WO1 and WO2. The abundance of Archaea in the four habitats investigated was very low, so the production of amyloid adhesins from this group remains to be investigated in more detail.

The identity of the isolates producing amyloids was also divers, and they belonged to the Gammaproteobacteria, Bacteriodetes, Firmicutes and Actinobacteria. The amyloids of *Pseudomonas fluorescens* were purified and further characterized, confirming the characteristics of amyloids (M.S. Dueholm, P. Larsen, J.L. Nielsen, P.H. Nielson, and D. Otzen, unpublished). These results support the in situ study that the ability to produce amyloid adhesins seems universal throughout many phyla. The exact function of the amyloids still needs to be better understood. For *E. coli*, it has been shown that the structures are important for initial adhesion (Prigent-Combaret et al., 2001) and protection against chlorine (Ryu and Beuchat, 2005). However, in biofilms investigated in our study most microorganisms were present in microcolonies, and thus represent more mature biofilms. The function is unknown, but they may be of importance for the cohesive properties of the biofilms. An interesting observation was also that several filamentous bacteria from...
different phyla stained positive with ThT. WO1 and WO2. In activated sludge, filamentous bacteria belonging to Chloroflexi expressed amyloids, and it will be very interesting to understand their function, perhaps related to hydrophobicity or strength of the filaments. More studies on the diversity, biophysical properties and genetic regulation are needed from a range of different species in order to understand more about evolution and function of these hitherto undescribed functional amyloids. Because most amyloids are non-covalent assemblies, in principle they should exist in equilibrium with their soluble, monomeric components. Some amyloids exhibit such equilibria (O’Nuallain et al., 2005), but most amyloid studies have very low, difficult-to-measure, critical concentrations in aqueous buffer. Presumably amyloids that have evolved to provide some function have been tuned by evolution for a certain insolubility and robust resistance to denaturation. The degradation of amyloid adhesins in nature is another interesting issue, as they remain insoluble under all but the harshest conditions (80–90% formic acid or trifluoroacetic acid). Nevertheless, purified Csga curli are completely degraded by proteases in vitro (A. Yde, D.E. Otzen and P.H. Nielsen, unpub. data). To our knowledge, the only related data available are from degradation of prions, a related amyloid, which can be insoluble under all but the harshest conditions (80–90%).

The results are also interesting in relation to human health. The amyloid-producing microbes constitute a rich potential source of environmental amyloids, which become more predominant at higher purities of water, such as drinking water. In the work by Lundmark and colleagues (2005), they showed that amyloid fibrils added to the drinking water of a mouse model for amyloidosis substantially decreased the lag time to onset of ‘disease’, suggesting that the added amyloid had almost a prion-like ability to get through the gut and act as a seed for amyloid growth and disease induction. This would raise the possibility that one variable in classical amyloid diseases is exposure to environmental amyloids – and might amyloids produced in biofilms be such a source?

**Experimental procedures**

**Biofilm samples**

Biofilm samples were collected from seven different habitats. One biofilm sample was grown on glass slides submerged in a non-chlorinated drinking water reservoir for 2 months. One biofilm sample was taken from a glass surface submerged in a freshwater lake close to Aabybro, Denmark. One biofilm sample was taken from the surface of a stone in the Limfjorden, a brackish water area close to Aalborg, Denmark. One biofilm sample was grown on the inside of a polycarbonate pipe submerged in seawater in Hou Harbour, Denmark.

The biofilms were scraped off the surfaces and suspended in sterile, filtered (0.2 μm polycarbonate filter) water from the location. Another biofilm sample (activated sludge flocs) was collected from the aeration tank at Aabybro wastewater treatment plant, Denmark, which is an advanced plant performing biological nitrogen removal. Two samples were taken from digesters, one from a thermophilic anaerobic digester at Aalborg East wastewater treatment plant and the other from a mesophilic anaerobic digester at Hjørring wastewater treatment plant. All experiments were initiated within 2 h after sampling.

**Organisms and growth conditions**

An E. coli mutant (OmpR 234 Tc$^6$) producing amyloid adhesins (Prigent-Combaret et al., 2001) was used as a positive control, and the negative control was an E. coli mutant (CsgDAB::Kan) unable to produce amyloid adhesins (Vidal et al., 1998). The two mutants were grown in M63 medium or on colonization-factor antigen (CFA) agar plates. Briefly, 1 l of M63 media is a mixture of: 3 g KH$_2$PO$_4$, 7 g K$_2$HPO$_4$, 2 g (NH$_4$)$_2$SO$_4$, 0.5 ml of 1 mg l$^{-1}$ FeSO$_4$, 2 ml of 0.5 M MgSO$_4$, 20 ml of 1.5 M glyceral, 20 ml of 5 mg ml$^{-1}$ thiamine, and 2 g glucose and 957 ml distilled water. The CFA agar plates contain 1% casamino acids, 0.15% yeast extract, 0.005% MgSO$_4$, 0.0005% MnCl$_2$ and 2% agar (Evans et al., 1977). pH was adjusted to 7.4 with NaOH. All incubations were performed at 26°C.

As a further positive control the aerial hyphae of S. coelicolor (DSMZ 40233) was used. Streptomyces coelicolor was grown on CFA agar plates. The spores contain amyloid-like fibers called chaplins (Claessen et al., 2003).

Six pure cultures producing amyloid adhesins were isolated from different natural biofilms. The biofilms were homogenized, grown on Congo red CFA agar plates and incubated at different temperatures, both identical with and different from the temperatures from the natural habitat (approximately 8–45°C). The isolates were chosen based on their ability to bind Congo red as described below. After further tests confirming the production of amyloid adhesins, the isolates were grown in liquid M63 or CFA media for later 16S rRNA gene sequencing and phylogenetic analysis as detailed below.

**DNA extraction, PCR amplification, DNA sequencing and phylogenetic analysis**

Total DNA was extracted from the pure cultures with a FastDNA® spin kit for soil and the FastPrep® instrument (Bio 101; Qbiogene, Carlsbad, CA). 16S rRNA gene fragments were amplified by PCR using bacterial primer sets 8F (Brosius et al., 1981) and 1390R (Zheng et al., 1996). The PCR products purified with a NucleoSpin Extract II purification system (Macherey-Nagel) ligated into a pCRII-TOPO vector with a TOPO TA cloning Kit according to the manufacturer’s instructions (Invitrogen, Groningen, the Netherlands). Plasmids were extracted and purified (FastPlasmid mini-preps, Eppendorf). Nucleotide sequencing was performed by Macrogen (South Korea). All sequences were checked for chimeric artifacts by using a CHECK_CHIMERA program from
the Ribosomal Database Project (Maidak et al., 1997). Nearly complete sequences of the 16S rRNA gene of each isolate were manually aligned and analysed by the ARB software package (http://www.arb-home.de).

Nucleotide sequence accession numbers

The sequences from this study have been deposited in the NCBI GenBank database under accession numbers EF566913-16 and EU010382-83.

Detection of amyloid adhesions on isolates

Amyloid adhesions expressed by isolates from various habitats were initially detected on CFA agar containing 20 μg ml⁻¹ Congo red and 10 μg ml⁻¹ Coomassie brilliant blue G (Hammar et al., 1995). Congo red binding colonies were identified by a red colour, whereas other colonies were white.

The presence of amyloid adhesions on the isolates was confirmed with ThT staining on liquid culture, followed by fluorescence spectroscopy making emission spectra in the interval of 460–550 nm and excitation at 450 nm. A stock solution of ThT dissolved in 96% ethanol was prepared, and the concentration was assigned by absorbance at 412 nm with \( E_{412\text{nm}} = 36 \text{ 000 } M^{-1} \text{ cm}^{-1} \) (Johnson et al., 2003). The working concentration of ThT was tested in the range of 0–80 μM on the E. coli mutants (diluted to OD₆₀₀ = 0.150), and the optimal ThT concentration for fluorescence spectrosopy was determined to 7.5 μM. When ThT binds to fibrils, the excitation maximum of 385 nm of the free dye changes to 450 nm, and the emission changes from 445 to 482 nm (Levine, 1993; Chapman et al., 2002).

Two conformationally specific antibodies, WO1 and WO2 (O’Nuallain and Wetzel, 2002), were also used to assess presence of amyloid adhesions. Both monoclonal antibodies developed in mouse spleen cells target amyloid fibril structures without staining the monomers (O’Nuallain and Wetzel, 2002). The bacteria were centrifuged at 9500 × g for 5 min, removal of supernatant, re-suspension in 400 μl PBS containing 0.1% Triton X-100 and 30 s vortexing. After the final washing step, the bacteria were re-suspended in 400 μl PBS and stored at 4°C until microscopy.

In situ detection of amyloid adhesins in natural biofilms

Staining of amyloid adhesins in natural biofilms was performed by Congo red, ThT, and with the antibodies WO1 and WO2, primarily in combination with either DAPI or FISH. Staining of amyloid adhesins with Congo red was performed according to the alkaline Congo red method (Puchtler et al., 1985; Westermark et al., 1999). Confocal laser scanning microscopy was used for detection (Puchtler and Sweat, 1965; Linke, 2000; Salt and Basdemir, 2003). Thioflavin T staining of the natural biofilms was performed by adding ThT to a final concentration of 1.5 μM. The working concentration was lowered compared with the analysis in pure culture described above in order to minimize background signal when the slides were analysed on CLSM. Antibody staining was performed as described above using gelatine-coated cover slides to prevent the biofilms from being washed off.

Staining of total bacteria with 4′,6-Diamidino-2-Phenyindole

4′,6-Diamidino-2-Phenyindole (DAPI) (Molecular Probes) was used to quantify the fraction of bacteria producing amyloid adhesins in the biofilms. 4′,6-Diamidino-2-Phenyindole staining was performed according to Fréulund and colleagues (1996).

Fluorescence in situ hybridization

The identity of bacteria in biofilms was determined with FISH and a range of oligonucleotide probes: EUB338mix [mixture of EUB338I (Amann et al., 1990), EUB338II, EUB338III (Daums et al., 1999)] to target all Bacteria, ALF968 for Alphaproteobacteria (Neef, 1997), BET42A (GAM42A as competitor) for Betaproteobacteria (Manz et al., 1992), GAM42A (BET42A as competitor) for Gammaproteobacteria (Manz and colleagues, 1992), SRB385 and SRB385Db for Deltaproteobacteria (Amann et al., 1990; Rabus et al., 1999), CF319a + b for Cytophagales in Bacteroidetes (Manz et al., 1996), CFXmix [mixture of GNSB941 (Gich et al., 2001) and CFX1223 (Björnsson et al., 2002)] for Chloroflexi, LGC354mix (mixture of LGC354A, LGC354B and LGC354C) (Meier et al., 1999) for Firmicutes, HG69a for Actinobacteria (Roller et al., 1994), PLA46 for Planctomycetes (Neef et al., 1998) and ARCH915 for Archaea (Stahl and Amann, 1991). NONEUB (Walther et al., 1993) was applied as a negative control in all hybridizations. Further details on the gene probes are available in probeBase (Löytönen et al., 2003). The gene probes were labelled with 5(6)-carboxyfluorescein-N-hydroxy-succinimide ester (FLUO5) or with sulfoindocyanine dye (Cy3) from Thermo Haibyd, Germany.

Fluorescence in situ hybridization was performed on biofilm samples alone or in combination with ThT or
antibodies. When combined with ThT, the cover glasses, which were stained and analysed with ThT, were dried and fixed for 1.5 h with 4% paraformaldehyde (PFA) for fixation of Gram-negative cells and 50% cold ethanol for fixation of Gram-positive cells. After fixation, the cover glasses were dehydrated for 2 min in increasing concentrations of ethanol (50%, 80% and 96%). The hybridization procedure was performed as described previously by Amann (1995).

Combination of FISH and antibodies was performed by applying the protocol described above on samples fixed in 4% PFA. After antibody labelling, the samples were fixed again in 1% PFA for 1 h to increase the binding strength between antibody and epitope. Finally, the samples were immobilized on glass slides, dehydrated and hybridized with oligonucleotide probes as described above.

Microscopy

Thioflavin T and Congo red were detected using a Zeiss LSM510 CLSM equipped with a Meta-filter and a 63× oil objective. Thioflavin T bound to fibrils was excited using an Argo laser (458 nm), and the emission signal was detected in the range of 473–494 nm. Congo red excitation wavelength was 543 with a HeNe laser. The Congo red emission wavelength was recorded in the interval of 580–591 nm. To maintain uniform light exposure in all samples analysed, fixed values of detector gain, amplifier offset/gain and pinhole were applied. The background signal was recorded in an unstained sample and on stained spots without cell material.

The combination of ThT and FISH was carried out sequentially, as the ThT signal was undetectable after the FISH procedure. First, an initial recording and localization of several ThT-positive microcolonies on the slide was performed, followed by fixation and the FISH procedure. Subsequently, the ThT-positive cells were relocated using the automatic stage control on the CLSM and any overlay between the gene probes, and the ThT stain was recorded. For samples consisting of multiple layers of cells, a stack of images was recorded for ThT-positive colonies in the z-axis. A projection of these images was used to compare with a similar projection of the FISH images.

The combination of antibody and DAPI was also performed sequentially first by recording images of 10 fields of view with the antibodies, DAPI staining, relocation and recording images of biomass in the original 10 fields of view.

Quantification of antibody-positive cells was performed with digital image analysis using the ImageJ software (available at http://rsb.info.nih.gov/ij; developed by Wayne Rasband, National Institutes of Health, Bethesda, MD). The quantified result is presented as biovolume of antibody-positive biofilm area relative to the area positive with the DAPI stain in very thin samples. The standard error was calculated on the different measured area fractions.

Acknowledgements

We are grateful to Gunna Christiansen, Århus University, for performing the TEM on the E. coli mutants and Avijs A. Hansen for help with the phylogenetic analysis. This study was in part funded by the Danish Technical Research Council in the Framework program ‘Activity and Diversity in Complex Microbial Systems’ and in part by NIH R01 NS46356 to RW.

References


© 2007 The Authors

Journal compilation © 2007 Society for Applied Microbiology and Blackwell Publishing Ltd, Environmental Microbiology


2.2 Paper 2


With reference to Ministerial Order no. 114 of 8 March 2002 regarding the PhD Degree § 14, article 3, statements from each author about the PhD student’s part in the shared work must be included in case the thesis is based on already published articles.

**Co-author statement concerning the contribution by Morten S. Dueholm**

- M. S. D. acted as a practical supervisor in the aspects of functional amyloid purification and biophysical characterization.
- M. S. D. purified the FuBA from *G. obscurus* (figure 5C) and performed the biophysical characterization of the purified material (figure 6C).
- M. S. D. was deeply involved in design and evaluation of the experimental work.
- M. S. D. participated in preparation of the manuscript.

Peter B. Jordal

Morten S. Dueholm

Poul Larsen

Steen V. Petersen

Jan J. Enghild

Gunna Christiansen

Peter Højrup

Per H. Nielsen

Daniel E. Otzen

Signature

Date
2.2.1 Motivation
The work of paper 1 had shown that FuBA are abundant in environmental biofilm, but the complexity of the systems did not allow us to determine the detailed localization and function of these structures. In order to deal with these questions, we examined pure culture systems. The Gram-positive bacteria group mycolata was selected for the investigation because this group comprises a number of genera with disease-causing species, including the severely pathogenic organisms Corynebacterium diphtheriae and Mycobacterium tuberculosis. The latter is the leading cause of death due to a single infectious agent globally [133]. The mycolata group has furthermore great environmental and economical impact, since several species (e.g., Gordonia spp.) may lead to unwanted foaming in wastewater treatment plants (WWTP) [134-136]. Finally, M. tuberculosis have been shown to use long entangled pili (MTP), which are very reminiscent of the amyloid-like curli fimbriae expressed by E. coli and Salmonella, indicating that FuBA is expressed in this bacteria group.

2.2.2 Outcome
The expression of FuBA was investigated for 14 mycolata species as well as six other distantly related Gram-positive bacteria using the method developed in paper 1. It was found that 12 of the mycolata species expressed FuBA as did the six other Gram-positive bacteria. This supports the conclusion of paper 1, which states that FuBA expression in a common trait for many bacteria. In few cases a harsh saponification procedure was required to remove lipid layers surrounding the bacteria and make the FuBA accessible to the WO2 antibody.

A detailed investigation of the localization of the FuBA suggested that these structures have multiple roles, sometimes even for a single bacteria strain. (A) Some bacteria expressed FuBA as extracellular fibrils. The appearance of these fibrils showed high variability with organisms. For C. glutamicum FuBA was found only in cellular microcolonies and for G. obsures they were found as a massive extracellular matrix with few or no cells attached. M. avium revealed a peculiar velvet-like WO2-positive extracellular matrix in the vicinity of cells. T. spumae exposed large cellular colonies containing fibrils more than 50μm long. (B) The need of a saponification to make the FuBA of some bacteria accessible suggests that FuBA could also be situated in the cell envelope. The presence of fibrillar material after saponification was verified by TEM, which showed the presence of fibrils that were 9nm wide and predominantly close to or integrated into the cell wall. (C) FuBA was finally found to cover the surface of spores from all spore producing bacteria tested.

FuBA was isolated and characterized biophysically from two bacteria in order to verify the histological data. The first bacteria produced encapsulated FuBA that requires saponification for exposure. The second produced extracellular FuBA. The amyloid-like nature of the FuBA was confirmed using TEM, CD, FTIR and ThT binding experiments.
Widespread Abundance of Functional Bacterial Amyloid in Mycolata and Other Gram-Positive Bacteria

Peter Bruun Jordal,1,2 Morten Simonsen Dueholm,1,2 Poul Larsen,2 Steen Vang Petersen,1 Jan Johannes Enghild,1 Gunna Christiansen,3 Peter Højrup,4 Per Halkjær Nielsen,2,4 and Daniel Erik Otzen1,2,*

Interdisciplinary Nanoscience Center (iNANO), Department of Molecular Biology, University of Aarhus, Gustav Wieds Vej 10C, DK-8000 Aarhus C, Denmark; Department of Biotechnology, Chemistry and Environmental Engineering, Aalborg University, Sohngardsholmvej 49, DK-8000 Aalborg, Denmark; Department of Medical Microbiology and Immunology, Bartholin Building, University of Aarhus, DK-8000 Aarhus C, Denmark; and Department of Biotechnology, Chemistry and Environmental Engineering, Aalborg University, Sohngardsholmvej 49, DK-9000 Aalborg, Denmark.

Received 3 September 2008/Accepted 16 April 2009

Until recently, extracellular functional bacterial amyloid (FuBA) has been detected and characterized in only a few bacterial species, including Escherichia coli, Salmonella, and the gram-positive organism Streptomyces coelicolor. Here we probed gram-positive bacteria with conformationally specific antibodies and revealed the existence of FuBA in 12 of 14 examined mycolata species, as well as six other distantly related species examined belonging to the phyla Actinobacteria and Firmicutes. Most of the bacteria produced extracellular fimbriae, sometimes copious amounts of them, and in two cases large extracellular fibrils were also produced. In three cases, FuBA was revealed only after extensive removal of extracellular material by saponification, indicating that there is integrated attachment within the cellular envelope. Spores of the genera Streptomyces, Bacillus, and Nocardia were all coated with amyloids. FuBA was purified from Gordonia amarae (from the cell envelope) and Geodermatophilus obscurus, and they had the morphology, tinctorial properties, and β-rich structure typical of amyloid. The presence of approximately 9-nm-wide amyloids in the cell envelope of G. amarae was visualized by transmission electron microscopy analysis. We conclude that amyloid is widespread among gram-positive bacteria and may in many species constitute a hitherto overlooked integral part of the spore and the cellular envelope.

The gram-positive bacterial group mycolata (mycolic acid-containing actinomycetes) comprises a number of genera with disease-causing species, including the severely pathogenic organisms Corynebacterium diphtheriae and Mycobacterium tuberculosis. The latter species is the leading cause of death due to a single infectious agent globally (17). Furthermore, mycolata have great environmental and economical impact, since several species (e.g., Gordonia spp.) may lead to unwanted foaming in wastewater treatment plants (10, 27, 43). M. tuberculosis was recently shown to use long entangled pili (MTP) to adhere to endothelium, eventually invading and infecting human and animal tissue (1). MTP’s morphology and tinctorial properties are very similar to those of the amyloid-like curli fibrils found in Escherichia coli and Salmonella species (7), although it has not been determined whether they contain the characteristic cross-β structure with β-strands perpendicular to the long fibril axis (44). In higher organisms, amyloid occurs mainly as an aberrant product of protein misfolding in, e.g., neurodegeneration and systemic amyloidosis, but bacteria are adept at turning amyloid to good use. In addition to the two bacteria mentioned above, functional bacterial amyloid (FuBA) has also been reported for streptomycetes (8) and xanthomonads (35). These examples are only the tip of the iceberg. Our recent in situ studies using WO2 antibodies specific for the amyloid conformation (36) in conjunction with 16S rRNA-targeted oligonucleotide probes for identification of the microbes revealed that amyloid-like adhesins are widespread in many phyla in environmental biofilms (29). In view of the occurrence of potential amyloid-like fibrils in one species belonging to the mycolata and the observed link between infection by a mycolata genus (Nocardia) and neurodegenerative Parkinson’s disease (13, 25, 26, 47), we have investigated this group of bacteria more closely for the presence of amyloid. Here we show that 12 of 14 different species of mycolata, as well as 6 of 6 other gram-positive bacteria, harbor amyloid. Furthermore, in some cases the amyloid can be visualized only after harsh saponification procedures which remove surrounding lipid molecules, indicating that the amyloid is deeply embedded in the cell envelope. Thus, amyloid may play a hitherto unappreciated central role in the composition of the bacterial envelope in many gram-positive bacteria.

MATERIALS AND METHODS

Organisms and growth conditions. All bacteria were grown in liquid shaken cultures (120 rpm, 28°C). E. coli SM2257 curl-deficient mutant (39) and E. coli SM2258 with upregulated curl production (48) were grown in liquid M63 minimal media (29). The following gram-positive organisms were obtained from Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH.

4101
Saponification. Fifty-milliliter 3-week-old bacterial cultures in stationary phase were centrifuged (10,000 × g, 10 min), the growth medium of each culture was removed, and the pellet was resuspended in 0.5% (wt/vol) KCl dissolved in absolute ethanol. Saponification was performed in an incubator (80°C, 200 rpm) using Teflon-sealed Greiner tubes placed in an upright position (3, 22). After 4 days, the remaining bacteria were centrifuged and resuspended in phosphate-buffered saline (PBS).

Antibody labeling and fluorescence microscopy of amyloid. A WO2 labeling (30) immunofluorescence protocol optimized for use with bacterial samples was performed as described previously (29). DAPI (4′,6-diamidino-2-phenylindole) staining was used to compare the relative positions of bacteria to the positions of bound WO2. DAPI staining was performed in the dark for 15 min on slides with air-dried bacteria using 2% (vol/vol) DAPI in PBS as described previously (10).

Purification of FuBA. Twenty milliliters of a dense 5-day (28°C, 120 rpm) culture of G. amarae, G. glutamicum, or G. obscurus in exponential phase was used to inoculate 1 liter of M63 medium. After 3 weeks of growth (28°C, 120 rpm), stationary-phase bacteria and extracellular matrix were harvested by centrifugation (10,000 × g, 30 min). The pellet was resuspended in PBS and sonicated (B. Braun Labsonic 1000L rod sonicator) on ice at medium intensity (30 min, 10°C, 30 min, Sorvall T640.1 rotor), and subjected to saponification at 80°C, 200 rpm, 6 h (Table 1). This is a very good indication that saponification resulted in a decrease (e.g., N. asteroides) or complete loss (e.g., M. avium) of WO2 binding (see below), suggesting that the purified material consists of lipids, polysaccharides, and proteins (15), and these molecules could block access of antibodies (but not of the small molecule ThT) to FuBA present either in the capsule of G. amarae or on the surface of the bacteria. A saponification step to remove lipids was therefore performed with G. amarae prior to WO2 labeling. G. amarae saponified at 80°C bound WO2 to a high degree (Fig. 1B), indicating that removal of lipids by saponification exposes FuBA present in the G. amarae cell envelope. Our immunochromatography data were supported by the observation of fibrillar structures by TEM (Fig. 2A to D).

When saponification was performed at higher temperatures, including 37°C (Fig. 2B) and 60°C (Fig. 2C), the bacteria were gradually dissolved and fibril-like structures were observed. After saponification at 80°C, nearly all bacteria had disintegrated, leaving material with a remarkable 9-nm-wide fibril-like morphology as determined by TEM (Fig. 2D). Furthermore, TEM analysis revealed that the purified FuBA from G. amarae had a fibrous morphology (see below) and was able to bind WO2 (see below), suggesting that the purified material was indeed the WO2 binding substances embedded in the capsule of G. amarae.

When immunolabeling was performed with saponified C. flavescens, there was a similar increase in WO2 binding compared to the binding with untreated C. flavescens. In other cases, saponification resulted in a decrease (e.g., G. glutamicum) or complete loss (e.g., N. asteroides) of WO2 binding (Table 1). This is a very good indication that saponification (which includes highly alkaline conditions, which are known to dissolve many types of protein aggregates) by itself is very
unlikely to cause formation of amyloid. Altogether, 12 of 14 mycolata species belonging to seven different families bound WO2.

**Production of FuBA in nonmycolata cultures.** The abundance of FuBA in mycolata prompted us to examine closely related actinomycetes. *S. coelicolor* is well known for assembly of chaplins into FuBA that confer hydrophobicity to submerged hyphae, allowing hyphae to grow into the air and form spores (8). This phenomenon has previously been reported to occur only in minimal media or liquid standing cultures, as also shown by the lack of WO2 labeling of *S. coelicolor* cultivated at 120 rpm in rich media (Fig. 3B). However, when the same *S. coelicolor* culture was saponified, strong WO2 binding was observed (Fig. 3D), indicating that *S. coelicolor* contains encapsulated amyloid, like *G. amarae*. Spores from *S. coelicolor* showed strong WO2 binding without saponification. The other nonmycolata actinomycetes *S. cinnabarium*, *G. obscurus*, and *A. acidiphila* were all able to bind WO2, as were the distantly related Firmicutes species *E. aquimarinus* and *B. mycoides* (Table 1). The fact that 18 of 20 gram-positive organisms examined, belonging to a wide array of species, produce FuBA indicates that FuBA is remarkably widespread among gram-positive organisms.

**Morphology of the extracellular FuBA.** The appearance of extracellular fibrils binding WO2 was highly variable among the organisms. The cells were visualized by staining with the nucleic acid-binding dye DAPI. *C. glutamicum* produced FuBA only in cellular microcolonies (Fig. 4A and B), whereas *G. obscurus* also produced a massive extracellular matrix with few or no cells attached (Fig. 4C and D). WO2 labeling of *M. avium* revealed a peculiar velvet-like WO2-positive extracellular matrix in the vicinity of cells (Fig. 4E and F), while WO2 labeling

### TABLE 1. Amyloid prevalence among gram-positive organisms: binding of WO2 to untreated and saponified cells based on immunofluorescence data

<table>
<thead>
<tr>
<th>Species</th>
<th>WO2 binding to untreated cells&lt;sup&gt;a&lt;/sup&gt;</th>
<th>WO2 binding to saponified cells&lt;sup&gt;a&lt;/sup&gt;</th>
<th>WO2 binding locations (untreated)&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Corynebacterineae: Corynebacterium</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. flavescens</td>
<td>+</td>
<td>+++</td>
<td>Within a few large microcolonies</td>
</tr>
<tr>
<td>C. glutamicum</td>
<td>+++</td>
<td>+</td>
<td>Within all microcolonies and on single cells</td>
</tr>
<tr>
<td><em>Corynebacterineae: Dietzia</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D. maris</td>
<td>+++</td>
<td>+</td>
<td>Within all microcolonies and on single cells</td>
</tr>
<tr>
<td>D. papillomatosis</td>
<td>+++</td>
<td>ND</td>
<td>Within all microcolonies and on single cells</td>
</tr>
<tr>
<td><em>Corynebacterineae: Gordonia</em></td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>G. amarae</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G. hydrophobaica</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Corynebacterineae: Nocardiella</em></td>
<td>+</td>
<td>-</td>
<td>On mycelia and within all microcolonies, spores and aerial hyphae positive</td>
</tr>
<tr>
<td>N. asteroides&lt;sup&gt;c&lt;/sup&gt;</td>
<td>+++</td>
<td>ND</td>
<td>Within all microcolonies and on single cells</td>
</tr>
<tr>
<td>N. polyresistens&lt;sup&gt;c&lt;/sup&gt;</td>
<td>+++</td>
<td>ND</td>
<td>Within all microcolonies and on single cells</td>
</tr>
<tr>
<td><em>Corynebacterineae: Mycobacterium</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M. avium</td>
<td>+</td>
<td>ND</td>
<td>Within velvet-like aggregates</td>
</tr>
<tr>
<td>M. pidei</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td><em>Corynebacterineae: Tsukamurella</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T. paurometabola</td>
<td>+</td>
<td>-</td>
<td>Cell microcolonies and large fibrous aggregates</td>
</tr>
<tr>
<td>T. spumae</td>
<td>+++</td>
<td>ND</td>
<td>Cell microcolonies and large fibrous aggregates</td>
</tr>
<tr>
<td><em>Corynebacterineae: Williamsia</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W. maris</td>
<td>+++</td>
<td>ND</td>
<td>Within all cell microcolonies and on single cells</td>
</tr>
<tr>
<td>W. muralis</td>
<td>+++</td>
<td>-</td>
<td>Primarily within large cell microcolonies</td>
</tr>
<tr>
<td><em>Non-Corynebacterineae in Actinobacteria</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S. coelicolor</td>
<td>-</td>
<td>+++</td>
<td>Only spores positive</td>
</tr>
<tr>
<td>S. cinnabarium</td>
<td>+++</td>
<td>ND</td>
<td>Within large mycelia and on single cells</td>
</tr>
<tr>
<td>G. obscurus</td>
<td>+++</td>
<td>ND</td>
<td>Within cells and huge extracellular matrices</td>
</tr>
<tr>
<td>A. acidiphila</td>
<td>++</td>
<td>ND</td>
<td>Primarily within large cell microcolonies</td>
</tr>
<tr>
<td><em>Firmicutes</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E. aquimarinus</td>
<td>+</td>
<td>-</td>
<td>On large cellular microcolonies</td>
</tr>
<tr>
<td>B. mycoides&lt;sup&gt;e&lt;/sup&gt;</td>
<td>+</td>
<td>+++</td>
<td>Within cell microcolonies and on large fibrous aggregates, spores positive</td>
</tr>
</tbody>
</table>

<sup>a</sup>-no WO2 binding; +, low level of WO2 binding, and only few microcolonies of cells were positive; ++, intermediate level of WO2 binding, and there was an obvious distinction between labeled highly positive cell microcolonies and unlabeled single cells; ++++, strong WO2 binding (i.e., no WO2-negative cells were observed); ND, not done.

<sup>b</sup>WO2 DAPI counterstaining allowed assessment of the position of FuBA relative to the cells.

<sup>c</sup>*N. asteroides*, *N. polyresistens*, and *B. mycoides* were grown in liquid standing cultures, and massive sporulation was observed; therefore, submerged mycelia and cells, aerial mycelia and cells, and spores were present when the cultures were labeled with WO2.
of *T. spumae* exposed large cellular colonies containing fibrils more than 50 μm long (Fig. 4G to H). Immunolabeling of the two *Nocardiopsis* spp. and *B. mycoides* showed that both the aerial hyphae and spores bound WO2. Figures 4I and J show spore formation by *B. mycoides* and the positive WO2 signal.

**Biophysical characterization of purified FuBA.** In order to examine the amyloid-like properties of FuBA, a purification protocol was developed. Two bacteria with different FuBA morphologies were chosen for detailed analysis: (i) *G. amarae*, which produces encapsulated FuBA that requires saponification for exposure, and (ii) *G. obscurus*, a nonmycolata species with remarkable production of extracellular FuBA, which is accessible to WO2 antibodies without saponification.

The optimized protocol was based on principles for purification of fungal hydrophobins (50) and chaplins from *S. coelicolor* (8) along with our own experiences from purification of amyloids from *Escherichia* and *Pseudomonas* species (M. Dueholm, P. Nielsen, and D. Otzen, unpublished results). Briefly, FuBA was liberated from bacteria using prolonged sonication on ice, and contaminating substances were removed using hot 2% SDS treatment followed by preparative SDS-PAGE. TEM analysis revealed that the purified FuBA from *G. amarae* had a fibrous morphology (Fig. 5A); furthermore, purified FuBA bound WO2 (Fig. 5B), suggesting that the purified material was indeed the WO2 binding substances embedded in the capsule of *G. amarae*. The TEM analysis, however, also revealed that other residual cell wall components were part of the purified FuBA. In particular, minor parts of the *G. amarae* capsule with extracellular material (Fig. 2A) were found interspersed between some of the fibers. Treatment with lysozyme

![Image 2](https://example.com/image2.png)

**FIG. 2.** Saponification of *G. amarae* at increasing temperatures reveals gradual liberation of fibril-like substances: TEM micrographs with 1% phosphotungstic acid staining of (A) nonsaponified *G. amarae*, (B) bacteria saponified for 4 days at 37°C, (C) bacteria saponified for 4 days at 60°C, and (D) bacteria saponified for 4 days at 80°C. Bars in panels A, B, and C represent 0.5 μm; the bar in panel D represents 0.1 μm. The arrows indicate the positions of (A) a dense extracellular matrix and (B to D) fibrillar material.
to remove potential peptidoglycan in the sample was attempted, but this did not lead to noticeable removal of impurities (data not shown).

Nevertheless, purified FuBA bound WO2 to a high degree (Fig. 5B), suggesting that a high proportion of the purified material was indeed the WO2 binding substances embedded in the capsule of *G. amarae*. SDS-PAGE was performed in order to examine the composition of FuBA from the two species. However, when purified FuBA was loaded, the material was not able to enter the pores of the gel due to FuBA’s large size (Fig. 5C); this factor had been used as a purification principle for the preparative SDS-PAGE system. Nevertheless, extraction with 100% TFA (50) resulted in faint protein bands at ~110 kDa for *G. amarae* and at ~55 kDa for *G. obscurus*, indicating the presence of protein components in FuBA. Due to the proteins’ low solubility, we have not been able to determine the molecular masses of these bands more accurately by mass spectrometry. However, we noted that for both species the molecular masses correspond to those of monomers and dimers, just as we observed for the *E. coli* FuBA CsgA (M. Dueholm et al., unpublished results) and similar to what has been reported for rodlin and chaplin(s) (8).

The amyloid-diagnostic dye ThT was used to substantiate these indications of the presence of amyloid-like material. A remarkably great (63-fold) increase in ThT emission was observed when purified *G. amarae* FuBA was added to ThT (Fig. 6A, inset). The emission maximum was close to 482 nm, which is characteristic of amyloid (31). Purified FuBA from *G. obscurus* was black and concealed the ThT signal (as well as the CD spectrum), probably due to interference. However, 50-fold dilution of the sample resulted in a ThT fluorescence signal that was four times greater than the background value (data not shown). The CD spectrum of purified FuBA from *G. amarae* had a single minimum at ~220 nm (Fig. 6A), indicating a β-sheet secondary structure in good agreement with the expectations for the cross-β amyloid fibrils. Finally, the FTIR spectra for purified FuBA from *G. amarae* (Fig. 6B) and *G. obscurus* (Fig. 6C) both contained a strong peak in the range from 1,620 to 1,630 cm⁻¹ characteristic of amyloid-like material (32, 51). We unsuccessfully attempted to sequence SDS-PAGE bands of the purified FuBA using trypsin digestion or chemical cleavage after Met, Trp, or Cys coupled with mass spectrometry or Edman degradation. This failure may reflect the small amounts of protein available and/or an unusual amino acid composition.

**DISCUSSION**

**Amyloid production in gram-positive bacteria.** Previously, FuBA has been detected and characterized only in the gram-negative bacterial taxa *E. coli*, *Salmonella*, and *Xanthomonas* and the gram-positive genus *Streptomyces* (7, 8, 36). However, our recent in situ immunolabeling experiments with WO2, Congo red, and ThT in natural bacterial habitats and wastewater treatment plants have revealed the widespread presence of FuBA in several other bacterial phyla (29). The report of FuBA in 18 mycolata and other gram-positive bacteria in this study substantiates these observations and shows that the ability to produce amyloid is much more common in bacteria than previously expected. The WO2 antibody has been shown to bind only to amyloid and not to other kinds of protein aggregates (36), and our previous work documented that WO2 binds to curli-producing *E. coli* but not to *E. coli* strains in which curli production has been knocked out (29, 30). In addition, we were able to purify FuBA to such an extent that only two bands (possibly a monomer and a dimer) were revealed by SDS-
FIG. 4. FuBA occurs in various species-specific shapes and sizes. (A, C, E, and G) Bright-field images. (B, D, F, and H) WO2 labeling (green) and DAPI counterstaining (blue). (A and B) C. glutamicum. FuBA is present around all cells. (C and D) G. obscurus. FuBA occurs in large extracellular aggregates. The arrows indicate extracellular material with a high level of amyloid but low cell density. Bars = 10 µm. (E and F) M. avium. Velvet-like substances strongly bind WO2. (G and H) T. spumae. Long (>50 µm) WO2 binding fibrils (arrows) are present. Bars = 10 µm. (I and J) Binding of WO2 to B. mycoides cells and spores. (I) Phase-contrast image of cells and spores (rings). (J) Fluorescence image of the same field, showing a high level of binding of WO2 to both cells and spores. Bars = 10 µm.
PAGE. We also observed a small amount of impurities (Fig. 2A) which may correspond to other cell wall components arising from the simple nature of our purification protocol and the inability of these materials to penetrate the SDS gel. However, the purified material showed ThT binding, had a fibrous morphology, and contained β-sheet-rich structures according to CD and FTIR. Therefore, we believe that it is unlikely that WO2 binds to a nonamyloid component in the 18 gram-positive strains that were positive for WO2. Only 2 of the 20 species tested did not produce FuBA under the growth conditions used. It is well known that amyloids are not expressed by all strains of *E. coli* and *Salmonella* (4, 39) and also that the growth conditions are important, particularly nutrient stress, which can promote curli production, adhesion, and biofilm formation (4, 39). Hence, it is possible that the two FuBA-negative species could produce FuBA under other growth conditions. The growth conditions in this study were optimized for *G. amarae* (medium, temperature, age of culture) and used for all species. The evidence that supports the widespread presence of FuBA suggests that this type of fibril is a multifunctional tool used by several bacterial species for purposes as diverse as formation of part of the cellular envelope, coating of spores, and as extracellular fibrils.

**Morphology of extracellular amyloid.** We do not have any evidence that FuBA is a prerequisite for flocculation, but we suggest that FuBA affects the properties of the resulting biofilm, although other experiments will have to elucidate this. Thus, gram-positive FuBA may act in a fashion similar to that of the extracellular curli fibrils from *E. coli* and *Salmonella*, which are known to facilitate adhesion and bacterial survival by initiating biofilm formation (4, 39). Several mycolata, including *G. amarae* (27), have been implicated in biofilm formation and operational problems in wastewater treatment plants (2). There have been several publications describing extracellular mycolata fibrils, but until very recently these fibrils have not been linked to amyloid (9, 15, 23, 34). Alteri et al. (1) showed that *M. tuberculosis* produces curli-like fimbriae that have crucial roles in infection. In the present work, FuBA was detected in the closely related pathogenic organisms *M. avium* and *N. asteroides*, and it therefore seems compelling that the MTP of *M. tuberculosis* are amyloid. Understanding the basis of adherence is the first step in combating tubercu-
Losos, especially in light of the extremely resistant M. tuberculosis strains found recently (19).

We observed a striking amount of extracellular FuBA produced by G. obscurus, M. avium, and T. spuamae. FuBA production at this level has not been described previously for bacteria. The observation of extracellular fibrils that are more than 50 μm long is very interesting and suggests that these fibrils play a central role in the development of the three-dimensional architecture of biofilms. The identity of these structures is not known, but as this study shows, they may at least in part consist of amyloids, perhaps in combination with other extracellular polymers. The same type of loosely attached material has been described for the pathogen Mycobacterium lepraemurium and the obligate pathogen Mycobacterium leprae (16). Whether the large structures observed for T. spuamae are similar to the so-called “honeycombs” recently observed in biofilms of Staphylococcus epidermidis (42) remains to be investigated.

Amyloid in the cell envelope. The cell wall in many mycolata is thought to consist of an outer layer consisting of mycolic acids, lipids, proteins, and polysaccharides and an inner electron-dense cell wall core consisting of peptidoglycan and arabinogalactan (40, 45). Our results show that some species of mycolata contain amyloids in the cell envelope not accessible to WO2, and this strongly indicates a previously uncharacterized function of FuBA. However, analysis of some nonmycolata and even bacteria belonging to the distant phylum Firmicutes revealed cell envelope amyloids. Thus, the presence of amyloids seems to be a more universal property of many gram-positive bacteria, so more detailed studies are needed to reveal the exact nature of these amyloids in the cell envelope.

Fibril-like structures were visible in G. amarae samples after saponification at different temperatures, indicating that the amyloids were not produced during the treatment. TEM images revealed that FuBA fibrils that were 9 nm wide were predominantly close to or integrated into the cell wall, whereas a minor fraction was distant from the cells. Several authors have used freeze fracture electron microscopy to show that the outer capsule surrounding intraphagosomal M. avium and M. lepraemurium consists of a multilaminar structure (41, 46). Each lamella of the M. avium coat is made up of parallel straight fibrils that are 5 nm wide. This structure is very similar to that of amyloid hydrophobins on the surface of Schizopyllum commune. Perhaps FuBA in the capsule could be partially responsible for the amazing survival of specific pathogenic mycolata species inside macrophages. This would be analogous to the silk moth chorion, where a lamellar ultrastructure of packed amyloid fibrils protects the developing embryo against temperature variations, mechanical pressure, proteases, bacteria, etc. (21). If a layer of lamellar amyloid is also present in the G. amarae envelope, this could explain this bacterium’s remarkable resistance to permeabilization and disruption (27).

Very interestingly, the results of this study also show that all bacterial species that formed spores under the conditions tested produced spores coated with amyloids. Spores covered by amyloids have been described for various fungi, where they are known as hydrophobins (20). This coating facilitates the dispersal of the spores by wind and enhances their attachment to surfaces and possibly also their pathogenic properties (20).
Some amyloids (hydrophobins) seem to have an important role in helping fungal conidia avoid clearance by neutrophils and macrophages in the early stages of infection. The presence of amyloid-like structures on spores of Bacillus atrophaeus has also recently been observed in detailed atomic force microscopy studies (38). Spores from B. mycoides are very hydrophobic, as verified by atomic force microscopy force measurements (6), and this could be due to the presence of amyloids. Biofilm formation by Bacillus cereus also takes place primarily at the air-liquid surface, where the bacteria sporulate and may be dispersed (49). Our results indicate that spores produced by many spore-forming gram-positive bacteria also are covered by amyloids, which promote wind dispersal, surface attachment, and pathogenicity, and this may also explain the extreme resistance of the spores to environmental stresses.

S. coelicolor is assumed to produce amyloid only in connection with formation of aerial hyphae and spore formation in liquid standing or solid medium (8), but the saponification and immunofluorescence analysis revealed that the amyloid was an integrated part of the cell wall also in nonsporulating cells. Saponification is known to remove substances (especially lipids) from the outer layers of the mycolata capsule (3) and could thus also remove embedding molecules from S. coelicolor, making amyloid accessible for WO2 binding. This is in agreement with a recent atomic force microscopy study (11), which showed that fibrous material is present on the surface of S. coelicolor before the onset of aerial hypha formation. Thus, the amyloid may be formed in the cell envelope prior to the formation of aerial fungal hyphae not only in S. coelicolor but also in other sporulating species, such as Nocardia and Bacillus species.

ACKNOWLEDGMENTS

We thank Ronald Wetzel for generously providing the WO2 antibodies, Torben Lund Skovhus (Danish Technological Institute, Aarhus, Denmark) for providing the necessary class II facilities for cultivation of pathogenic bacteria, Søren Nielsen and Zhila Nikrozi for initial help with electron microscopy, and Sarita Singh and Dominik Dominiak for spor staining. D.E.O. and M.S.D. acknowledge support from the Villum Kann Rasmussen Foundation for operating costs, as well as a predoctoral stipend to M.S.D. through the research network BioNET. D.E.O., S.V.P., and J.J.E. were supported by the Danish Research Foundation and Aalborg University.

REFERENCES

2.3 Paper 3


With reference to Ministerial Order no. 114 of 8 March 2002 regarding the PhD Degree § 14, article 3, statements from each author about the PhD student’s part in the shared work must be included in case the thesis is based on already published articles.

Co-author statement concerning the contribution by Morten S. Dueholm

- M. S. D designed the research supervised by P. H. N. and D. E. O.
- M. S. D. performed the majority of the experimental work, including:
  - Isolation of Pseudomonas strain UK4 on Congo red indicator plates.
  - Development and optimization of the purification protocol for Pseudomonas FuBA.
  - Biophysical characterization of the FuBA (FTIR, CD, ThT fluorescence, fibrillation assay, and preparation of samples for TEM (G. C. collected the images)).
  - Preparation and optimization of samples for MS/MS and N-terminal sequencing (S. V. P. performed the data acquisition).
  - Identification of fap homologs in other Pseudomonas spp.
- M. S. D. performed all data analysis except assembly of the Pseudomonas genome which was done by M. S. Identification of the fap operon was done in collaboration with M. S.
- M. S. D. wrote the paper supervised by P. H. N. and D. E. O.

Morten S. Dueholm 14/09-2009
Steen V. Petersen 11/09-2009
Mads Sønderkær 14/09-2009
Poul Larsen 14/09-2009
Gunna Christiansen 11/09-2009
Jan J. Enghild 10/09-2009
Jeppe L. Nielsen 14/09-2009
Kåre L. Nielsen 14/09-2009
Per H. Nielsen 14/09-2009
Daniel E. Otzen 15/09-2009

Signature Date
2.3.1 Motivation
The work of paper 3 was originally intended as screening study for FuBA in pure culture systems of bacteria isolated from natural biofilms. One of the first bacteria we identified was a *Pseudomonas* strain which expressed amyloid-like fimbriae very similar to the curli fimbriae. Such fimbriae have previously only been described for bacteria within the Enterobacteriaceae [74]. It was therefore decided to investigate these FuBA in detail. As no homologues to the curli subunits were found in the genome of sequenced *Psuedomonas* strains, it was hypothesized that the *Pseudomonas* fimbriae could represent a novel FuBA system.

2.3.2 Outcome
The first step of characterizing the amyloid-like fimbriae from the *Pseudomonas* strain was to develop a FuBA purification protocol. We achieved this based on a principle previously used for the purification of thin aggregative fimbriae from *Salmonella enteritidis* [76]. The method involves making all other cell macromolecules soluble either through the use of detergents or by enzymatic degradation, and thereafter removing them by washing steps. Contaminating proteins were removed by preparative SDS-PAGE taking advantage of the ability of the FuBA to avoid depolymerization in boiling SDS-PAGE loading buffer [76].

A biophysical characterization of the purified fimbriae was performed in order to verify their amyloid-like structure. TEM showed that the purified material was fibrillar and reminiscent of curli fimbriae isolated from *E. coli*. The amyloid-like secondary structure was confirmed by CD and FTIR. The spectra were very similar to those for curli fimbriae. Finally, the *Pseudomonas* fimbriae were shown to bind ThT in an amyloid specific manner.

To gain further insight into the *Pseudomonas* FuBA system, we tried to identify the major fimbriae subunit with MS/MS and N-terminal. This investigation allowed the identification of 4 fragments of 12-37 amino acids. When these sequenced were used to search for homologues in established databases no candidates could be found. We therefore performed a full genomic sequencing of the bacteria. This enabled us to identify the full gene sequence and furthermore the operon structure in which it was situated. The protein shared no sequence homology with CsgA, the major curli subunit of *E. coli*. However, one striking similarity was found between the two. Both sequences contained repeated units with highly conserved glutamine and asparagines residues.

The full gene sequence of the *Pseudomonas* fimbriae FuBA operon was analyzed with the blastn program and the *Pseudomonas* genome database (http://www.pseudomonas.com). The operon was found to be conserved within many *Pseudomonas* strains including the opportunistic pathogen *Pseudomonas aeruginosa*. 
NOVEL CLASS OF FUNCTIONAL AMYLOIDS DISCOVERED IN PSEUDOMONAS

Morten S. Duholm, Steen V. Petersen, Mads Sønderkær, Poul Larsen, Gunna Christiansen, Jan J. Enghild, Jeppe L. Nielsen, Kaare L. Nielsen, Per H. Nielsen, Daniel E. Otzen.

Centre for Insoluble Protein Structures, Interdisciplinary Nanoscience Center (iNANO), Department of Molecular Biology University of Aarhus (iNANO), 8000 Aarhus C, Denmark; Department of Biotechnology, Chemistry, and Environmental Engineering, Aalborg University, 9000 Aalborg, Denmark; Department of Medical Microbiology, University of Aarhus, 8000 Aarhus C, Denmark.


The authors declare no conflict of interest

Abstract: A unique type of fimbriae found among certain biofilm forming bacteria in the family Enterobacteriaceae shares many characteristics of the protein amyloids associated with neurodegenerative diseases. We report the discovery of a novel functional amyloid expressed by a Pseudomonas strain of the P. fluorescens group. The amyloid protein was purified and the amyloid-like structure verified. Partial sequencing by MS/MS combined with full genomic sequencing of the Pseudomonas strain allowed identification of the full gene which we term fapC. The fimbrin contains a thrice repeated motif with the consensus sequence NNAX4SX2NX2GNX2GXNXAAGXGNQQXN, where X is a non-aromatic residue. This motif differs from those previously found in curli fimbrins and prion proteins. The lack of aromatic residues in the repeat shows that aromatic side chains are not needed for efficient amyloid formation. In contrast, Asn and Gln residues seem to play a major role in amyloid formation as these are highly conserved in curli, prion proteins, and the newly identified fimbrin. fapC is conserved in many Pseudomonas strain including the opportunistic pathogen P. aeruginosa and is situated in a conserved gene cluster containing 6 genes, of which one codes for a homolog to FapC. This cluster may represent a new genetic organization for the production of functional bacterial amyloid.

INTRODUCTION

The ability of proteins to form organized deposits of thin needle-like structures, called amyloid fibrils, has been proposed to be a generic protein property [1]. Although such deposits are mostly associated with neurodegenerative diseases such as Alzheimer’s and Parkinson’s [1-4], amyloid formation actually plays important, beneficial roles in nature. An example is the production of a morphologically and biochemically distinct type of fimbriae by certain Escherichia coli strains. These fimbrina are called curli due to their coiled appearance as judged by electron microscopy [5]. Similar fimbrinae have also been found in certain Salmonella strains referred to as thin aggregative fimbrinae (Tafi) [6]. Curli and Tafi have been shown to be involved in colonization of abiotic surfaces, biofilm formation, and internalization of bacteria into eukaryotic cells [7, 8]. They also mediate binding of a variety of extracellular matrix and serum proteins, such as soluble fibronectin, laminin and plasminogen activator protein [9]. The amyloid-like nature of the curli has been demonstrated by their ability to bind the dyes Congo red and Thioflavin T (ThT), which undergoes the same spectral changes as when they bind to disease-related amyloids. This was corroborated by a high content of β-sheet structure [10]. Furthermore, purified curli fimbrins have been shown to self assemble into fibrils similar to those produced in vivo when incubated for one week at 4°C at neutral pH [10]. Amyloid formation by full length proteins at neutral pH is usually only seen for highly amyloidogenic disease-related proteins, such
as α-synuclein as well as mutants of lysozyme, gelsolin, and transthyretin [11-14]. The ability of curli fimbrins to self assemble in vitro at neutral pH, therefore, provides additional evidence for the amyloid-like nature of curli.

Until now the structure and expression of amyloid-like fimbriae (ALF) have only been described in detail for E. coli and Salmonella, although they have also been shown to be produced by some Citrobacter strains [15, 16]. Furthermore, it has been suggested that they might be found in other Enterobacteriaceae, such as Shigella and Enterobacter based on hybridization experiments using DNA-fragments targeting *agfA*, the structural gene for Tafi [17, 18]. Recent studies in our group, based on confocal laser scanning microscopy combined with amyloid specific dyes or amyloid specific conformational antibodies [19], have shown that a wide range of microorganisms in natural biofilms produce amyloid-like surface structures depending on the habitat tested. This indicates that there might be a large array of phylogenetically distant bacteria expressing amyloid-like structures.

The intricate system required for curli and Tafi production, which involves numerous proteins from multiple operons [9], strongly implies that the ALF systems have been highly selected and refined by evolution. It is therefore suggested that the fimbrins have been fine-tuned with respect to fibrillation. If ALF expression systems like the ones for E. coli and Salmonella are also found for other bacteria, ALF might provide a valuable library of evolutionarily distinct fimbrins, which all have been selected due to their ability to form amyloid-like structures. Given that there are protein sequences or motifs which enhance amyloid formation, these are very likely to be found as conserved or repeated sequences in the primary structure of such fimbrins. Information on such motifs may provide valuable insight into the mechanisms of amyloid formation.

Here we show that a *Pseudomonas* strain produces ALF. The ALF was purified and the amyloid-like structure was confirmed. The sequence of the fimbrin was identified by combined MS/MS, N-terminal, and full genome sequencing. The fimbrin gene was found to be conserved in many *Pseudomonas* strain and is situated in a conserved but hitherto uncharacterized gene cluster containing 6 genes.

**RESULTS**

**Bacteria Isolation and Characterization**

In an earlier study we used confocal laser scanning microscopy combined with thioflavin T and the amyloid conformational specific antibodies WO1 and WO2 to identify bacterial strains which produce amyloid-like structures. Strains were isolated from a biofilm grown on glass slides submerged in a non-chlorinated drinking water reservoir [19]. One of these strains was chosen for further investigation.

The 16S rRNA gene sequence of the bacteria (GenBank accession # NZ_ACOQ01000017) revealed a 99.8% match with the published sequences for strain *Pseudomonas gessardii* CIP 105469 [20]. We name this strain *Pseudomonas* strain UK4.

TEM was used to verify that *Pseudomonas strain* UK4 produces fimbriae. The cells are surrounded by a massive extracellular matrix which contain embedded fimbriae (Fig. 1A) which are similar to, but less distinct than, curli produced by *E. coli* (Fig. 1B).

**Purification of Amyloid-like Fimbriae**

Purification of ALF was carried out using a modified protocol developed for isolation of Tafi from *S. enteritidis* [6]. This method relies on cell disruption by freeze-thaw cycles, enzymatic degradation of cell wall component and nucleic acid, and removal of contaminating proteins by preparative SDS-PAGE. The latter was possible due to the unusual stability of the ALF, which includes the ability to withstand disruption in boiling SDS-PAGE sample buffer even with β-mercaptoethanol [6]. In order to investigate the fimbrins by SDS-PAGE, the isolated ALF were dissolved in 90% formic acid and lyophilized. This treatment dissociates ALF into fimbrins [6].

SDS-PAGE showed that all contaminating proteins were removed during the purification (Fig. 2). When the purified ALF was treated with formic acid, a major band at 25kDa and a minor at 50kDa appeared on the SDS-PAGE. These bands corresponded in size to a fimbrin monomer and dimer, respectively, analogous to the formation of dimers by the Tafi and curli fimbrins [10, 21]. However, two different protein species could not be ruled out at this stage.

**Verification of the Amyloid-like Structure**

The fibrillar nature of the purified ALF was also verified by TEM (Fig. 3A), highlighting a resemblance to curli purified from *E. coli* (Fig. 3B).

The purified ALF was then investigated by synchrotron radiation circular dichroism (SRCD) and FTIR, which in a complementary fashion provide information about the secondary structure of the protein. The Far-UV SRCD spectrum of the purified ALF was very similar to that of curli and showed predominant β-structure with the characteristic single minima at 217nm and a maxima at 195nm (Fig. 4A) [22, 23]. The FTIR spectrum of the purified ALF (Fig. 4A) was also very similar to that of curli and showed two very intensive peaks, one at 1625cm⁻¹ and another at 1660cm⁻¹. The peak at 1625cm⁻¹ is characteristic of β-pleated structures maintained by
very strong hydrogen bonds, whereas the peak at 1660 cm\(^{-1}\) is proposed to result from a splitting of the main \(\beta\)-sheet component, indicative of an anti-parallel arrangement of the \(\beta\)-strands [23, 24]. A minor peak at approximately 1650 cm\(^{-1}\) was also observed, indicating small amount of random conformations [25]. Thus SRCD and FTIR both confirm the presence of extensive \(\beta\)-sheet structure as expected for an amyloid.

Binding of ThT to the purified ALF was determined by fluorescence spectroscopy. The ThT emission spectrum (Fig. 2C) was comparable to that of curli. The strong fluorescence emission, with a maximum at 485 nm when excited at 450 nm, showed that ThT binds specifically to the purified ALF, indicating amyloid-like structure of the purified ALF [26, 27]. The negative control, native lysozyme, did not show any binding of ThT.

The fibrillation kinetics of formic acid treated *Pseudomonas* ALF was assayed by ThT fluorescence at room temperature (Fig. 2D). From the increase in ThT fluorescence, it was found that samples of formic acid treated ALF starts to fibrillate immediately after resuspension.

**Primary Structure of the ALF Monomer**

Using a combination of MS/MS sequencing on the purified *Pseudomonas* ALF fimbrin and full genomic sequencing of the bacterium, we were able to obtain the full sequence of the ALF precursor protein, which we call FapC, after functional amyloid in *Pseudomonas* (GenBank accession # ZP_04596554) (Supplementary Fig. 1). N-terminal sequencing showed that FapC contains a 24 amino acid signal sequence which is cleaved off during the translocation to the outer membrane (Fig. 5A). Furthermore, FapC contains a thrice repeated motif with the consensus sequence NNAX\(_4\)SX\(_2\)NX\(_2\)GNX\(_2\)GXNX\(_2\)AGXGNQQXN (Fig. 5B). The three copies of this motif are separated by linkers of 35 and 43 residues, respectively.

**ALF Genes are Conserved in Pseudomonas**

The sequence of FapC was blasted against the *Pseudomonas* genome database version 2.3 in order to identify orthologs in other *Pseudomonas* strains. Orthologs were found in many strains (Supplementary Table. 1). An investigation of the protein sequences of the orthologs showed that they all contain a thrice repeated motif similar to that in FapC (Supplementary Fig. 2). These repeated units are connected by linker regions which vary both in amino acid composition and length (Supplementary Table 2).

**ALF Genes are Located in a Conserved Gene Cluster**

Analysis of the genes flanking FapC and its orthologs showed that it was part of an uncharacterized gene cluster consisting of six tightly packed genes, which we name *fapA-F* (GenBank accession # NZ_ACOQ01000001) (Supplementary Table 1). The gene cluster also contained a highly conserved promoter region (Supplementary Table 1), which was found to contain putative -35 and -10 elements using the EPROM tool for promoter prediction. The presence of ribosomal binding sites (Shine-Dalgarno sequences) was identified within the gene cluster (Fig. 5C). These results showed that the gene cluster very likely represents an operon.

One of the genes (*fapB*) was found to code for a homolog to FapC and sequence analysis showed that this protein also contains the thrice repeated motif. Attempts to assign the function of the proteins encoded by the genes using ortholog identification by blast searches, subcellular localization predictions (Supplementary Table 3) or sequence-based structure prediction did not provide further insight.

**DISCUSSION**

It is now known that amyloid formation is not only associated with disease but also plays important, beneficial roles in nature. Functional amyloids have been found in connection with the chorion of the silkmoth, spider silk, the coating of fungal hyphae, the melanosome biogenesis, bacterial inclusion bodies, and not least as ALF [28-32]. In this study we examined the ALF produced by *Pseudomonas* strain UK4, which has earlier been shown to bind ThT and the amyloid specific conformational antibodies WO1 and WO2 [19, 33]. *Pseudomonas* strain UK4 was surrounded by a thick extracellular matrix which contains embedded fimbriae similar to, but less distinct, than those of curli-producing *E. coli*. This shows parallels to *S. enteric*, which in the presence of cellulose, produces diffuse Tafi where no fine detail or individual fibers can be resolved [34].

Using a purification protocol modified from Collinson et al. (1991) [6], we were able to purify ALF from *Pseudomonas* strain UK4. The purified ALF resembled *E. coli* curli morphologically and showed the expected \(\beta\)-sheet signatures according to SRCD [35] and FTIR [23, 24, 36] as well as the ability to bind ThT and self-assemble after dissolution in formic acid. Taken together this is a clear indication of amyloid structure.

By a combination of MS/MS, N-terminal, and full genome sequencing we were able to obtain the full gene sequence of the ALF fimbrin FapC, whose most remarkable feature is a thrice repeated unit with the consensus sequence NNAX\(_4\)SX\(_2\)NX\(_2\)GNX\(_2\)GXNX\(_2\)AGXGNQQXN, where X is a non-aromatic residue.
The existence of repeated units in functional amyloid monomers has also been seen for curli and Tafi fimbrins, spider silk proteins as well as prion proteins from both human and yeast [37, 38]. We note that the repeats found in FapC are significantly different from those found in the curli and Tafi fimbrins. The lack of aromatic residues in the repeats clearly shows that such residues are not needed for efficient fibrillation. This is supported by Wang et al. (2008), who used substitution mutations on aromatic residues in the curli fimbrin (CsgA) to show that these were not required for efficient polymerization in vitro [39]. In contrast, the presence of many conserved Asn/Gln residues in FapC and Tafi and curli fimbrins as well as in prion and spider silk proteins indicates that interaction of these polar residues have a key role in the fibrillation and stability of functional amyloids. This is again supported by data from Wang et al. (2008), who showed that single substitution mutations of Asn and Gln in the repeats of CsgA have major impact on the efficiency of fibrillation in vitro [39]. Note that hydrophobins and chaplins, the functional amyloids involved in controlling the hydrophobicity of hyphae in fungi and actinobacterial streptomycetes, do not contain repeated units or extensive use of Asn and Gln [40, 41]. This indicates multiple ways of stabilizing functional amyloids.

We were able to identify orthologs to FapC in many other Pseudomonas strains, including the pathogenic P. aeruginosa. This indicates that the expression of ALF is common for many Pseudomonas strains. Tafi and curli fimbriae have been shown to be important for the development of Salmonella and E. coli biofilms [42, 43]. It is therefore very likely that ALF may have that same function in Pseudomonas biofilms. Note that even though the repeated units are well conserved among the orthologs, the linker regions show high a degree of variation. A strikingly large variation is seen in the length of linker region 2, which varies from 43aa in FapC to 276aa in PSEEN2941, the FapC ortholog in Pseudomonas entomophila L48. The variations in the linker regions could be employed to control the physiochemical properties of the ALF or display species-specific features on the fimbriae.

An examination of the genes flanking fapC showed that it was part of a gene cluster consisting of 6 genes. This gene cluster is conserved in all Pseudomonas strains containing FapC orthologs. The organization of the gene cluster strongly implies that it represents a novel operon for functional amyloid production. fapB codes for a FapC homolog, which also contain the thrice repeated unit. FapB may therefore represent a nucleator protein analogous to CsgB and AgfB of the curli and Tafi systems respectively [10, 34]. The usage of nucleator proteins might therefore be a common trait in the control of functional amyloid formation.

### MATERIALS AND METHODS

The E. coli MG1665 mutant SM2258 [44], used as a positive control of ALF production, was kindly provided by Professor Søren Molin (BioCentrum-DTU, Denmark). For all experiments, bacteria was grown on colonization factor antigen (CFA) agar plates at 26°C [45].

#### Identification of Bacterial Strain

The full 16S rRNA gene was identified in the Pseudomonas strain UK4 genomic data by blast searches using the 16S rRNA sequence of P. fluorescences Pf-5 (Pseudomonas Genome database version 2.3: PFL_0119). Phylogenetic analysis was carried out as described by Thomsen et al. [46].

#### Transmission Electron Microscopy

TEM was performed on both intact bacterial cells and on purified ALF. Bacterial cells were grown on CFA agar plates and a colony was picked up and suspended in 50μL of phosphate buffered saline pH 7.4. 10μL of bacterial suspension was placed on top of a carbon coated, glow discharged nickel grid for 30sec. For the purified ALF 10-20μL of a 4mg/mL solution was applied to the nickel grid for 30sec. The grids were washed on one drop of glass distilled water and stained with 3 drops of 1% phosphotungstic acid pH 6.9 and blotted dry. Electron microscopy was done using a JEOL 1010 TEM at 60 kV. Images were taken using a Sony XCD-SX900 camera. For size determination, a standard grid-size nickel plate (2160lines/mm) was used [47].

#### Purification of Amyloid-like Fimbriae

ALF were isolated by a protocol modified from Collinson et al. [6]. Bacteria were grown on CFA agar plates (48hr, 26°C), harvested from 20 plates and suspended in 30mL of 10mM Tris-HCl, pH 8.0 (Tris buffer), containing 0.1mg/mL RNase A (bovine pancreas; Fluka) and 0.1mg/mL DNase I (bovine pancreas; Sigma). Cell suspensions were homogenized using a tissue grinder. MgCl₂, lysozyme (Sigma) and triton X-100 (Fluka) were added to 1mM, 1mg/mL and 0.1% (v/v), respectively. Cells were broken by three freeze-thaw cycles using a dry ice/ethanol bath and a water bath at 37°C. Suspensions were incubated (150rpm, 37°C,12hr), after which they were adjusted to 1% SDS and incubated for further 30min. Insoluble material containing the ALF was collected by centrifugation (38.000rpm, 30min, 4°C), resuspended in 10mL Tris buffer, heated to 95°C for 15min to melt contaminating agar, and centrifuged to collect insoluble material. This step was preformed twice. The insoluble material was resuspended in 10mL of Tris buffer containing 0.1mg/mL RNase A and 0.1mg/mL DNase I and incubated for 20min.
Lysozyme was added to 1mg/mL and the samples were incubated for 12hr, after which they were adjusted to 1% SDS and incubated for further 30min. Insoluble material was collected by centrifugation, washed with 10mL Tris buffer, centrifuged, and suspended in 4mL SDS-PAGE sample buffer. The samples were boiled for 15min, loaded onto a preparative 12% polyacrylamide gel, and subjected to electrophoresis at 20mA for 6hr. Material that did not enter the gel was washed twice with 10mL Tris buffer and resuspended in 3mL Tris buffer. The purified ALF was briefly sonicated to disrupt large aggregates. All centrifugations were carried out in a Sorvall Discovery 90 ultracentrifuge equipped with a T647.5 rotor. Protein concentrations were estimated by UV-absorbance.

**Electrophoresis**

SDS-PAGE was preformed according to the method of Laemmeli with modifications by Ames [49, 50]. Samples were mixed with an equal amount of 2x SDS-PAGE sample buffer (30% glycerol, 0.1M dithiothreitol (DTT), 1.2% SDS, 18µg/mL bromophenol blue, 150mM Tris-HCl, pH 6.8) and loaded onto a 12% (w/v) polyacrylamide gel poured with a 5% (w/v) stacking gel. Electrophoresis was carried out at a constant voltage of 140V. ALF required an additional pretreatment so that the ALF monomers would enter the gel. 50µL of purified fimbria was mixed with 450µL of 98-100% formic acid and frozen in liquid nitrogen. The frozen samples were lyophilized in a Christ Beta 1-8K lyophilizer. The samples were then resuspended in 3mL Tris buffer. The samples were boiled for 15min, loaded onto a preparative 12% polyacrylamide gel, and subjected to electrophoresis. Samples were not boiled prior to electrophoresis. Proteins were visualized by staining with 0.25% (w/v) Coomassie brilliant blue G-250 (Sigma) in 45% (v/v) ethanol, 10% (v/v) acetic acid. Destaining was carried out with 45% (v/v) ethanol, 10% (v/v) acetic acid.

**Thioflavin T Fluorescence**

ThT emission was measured at 25°C on a PerkinElmer LS55 Luminescence spectrophotometer. Emission spectra of 0.25mg/mL lysozyme, curli or purified ALF in 10mM Tris-HCl, pH 8.0 and 40µM ThT were recorded from 460-750nm with excitation at 450nm. A bandwidth of 5nm was used on both excitation and emission and the scan speed was 200nm/min. Five spectra were averaged to improve the signal to noise ratio.

Fibrillation of formic acid treated ALF was investigated using ThT fluorescence on a SpectraMax GeminiXS plate reader with the following settings: Excitation at 450nm, emission at 490nm, auto cut off at 475nm and a sensitivity of 10 using a low photomultiplier. Fibrillation was carried out in 96 well opaque plates with 150µL of 10mM Tris-HCl, pH 8.0 using 2mg/mL of formic acid treated ALF and 40µM ThT and was followed for 12hr with measurements every minute. 5sec of shaking was applied between each measurement.

**Circular Dichroism**

SRCD spectra were collected on beamline UV1 at ASTRID, part of the Institute for Storage Ring Facilities, the University of Aarhus, Denmark. The beamline was calibrated for wavelength, optical rotation and magnitude at the beginning of each fill of the storage ring (once a day). ALF was diluted to 0.2mg/mL in 10mM Tris-HCl, pH 8.0 and measured in 100 µm path-length Suprasil open cells (Hellma GmbH & Co KG). Three spectra were collected at 20°C and with 1 nm step size. The limit for the lowest wavelength was determined by the absorption of water, the buffer and the solute. The spectra were averaged, smoothed, and baselines subtracted.

**Fourier-transform Infrared Spectroscopy**

FTIR was carried out using a Tensor 27 (Bruker) FTIR spectrophotometer equipped with a DTGS Midinfrared detector and a Golden Gate single reflection diamond attenuated total reflectance (ATR) cell (Specac). Approximately 4µg of purified ALF was dried on the ATR crystal using dry nitrogen. Spectra were recorded from 4000-1000cm⁻¹ using a nominal resolution of 2cm⁻¹ and 64 accumulations. Fourier self-deconvolution of the spectra in the amide I region was performed using a built-in apodization function with a deconvolution factor of 2, a noise reduction factor of 0.5, and Lorentzian lineshape in the OPUS 5.5 system (Bruker). Identification of the different components of the amide I after Fourier self-deconvolution was performed by a Levenberg-Marquardt fit function in the OPUS 5.5 system.

**Mass Spectroscopy**

Bands of interest were excised from Coomassie Blue-stained SDS-PAGE gels, cut into small fragments, and washed in distilled water. The gel pieces were then incubated in acetoniitrile and subsequently equilibrated in NH₄HCO₃, as previously described by Karring et al. [51]. After lyophilization using a speed-vac centrifuge, the gel pieces were swelled in 20µL of 50mM NH₄HCO₃ containing 25µg/mL procine trypsin (sequencing-grade, Promega). After rehydration, 50mM NH₄HCO₃ was added to cover the gel pieces and the digestion was allowed to proceed for 16hr at 37°C. The generated peptides were recovered using Stage tips (C18, Proxen) as recommended by the manufacturer. To improve sequence analysis the peptides were derivatized using the Ettan CAF MALDI Sequencing kit (GE Healthcare). All steps were conducted using peptides absorbed to the C18 Stage tip. Briefly, lysine residues were converted to
homoarginine by the incubation with O-methylisourea hydrogen sulphate for 2hr at 37°C. Subsequently, the N-termini were sulfonated by incubation with the CAF reagent for 5min at room temperature. The derivatized peptides were eluted onto a target plate using 1µL matrix solution (0.4% (w/v) recrystallized α-cyano-4-hydroxycinnamic acid in 70% (v/v) acetonitrile, 0.03% (v/v) trifluoroacetic acid). The samples were analyzed by MALDI-TOF MS/MS by using a Q-Tof Ultima Global mass spectrometer (Micromass/Waters Corp.) calibrated using a mixture of PEG-200, -600, -1000, -2000 and NaI in 50% acetonitrile.

N-terminal Sequencing

Formic acid treated ALF was separated by SDS-PAGE and electroblotted to a PVDF membrane using Towbin buffer (25 mM Tris, 192 mM glycine, 20% methanol, pH 8.3) containing 0.01% SDS. Bands were visualized by Coomassie Brilliant Blue staining, excised and analysed by automated Edman degradation using a Procise 494HT amino acid sequencer (Applied Biosystems) with on-line phenylthiohydantoin analysis.

Genomic Sequencing

Library Preparation for Illumina Sequencing

Pseudomonas strain UK4 was grown in TSB-medium at 26°C for 48hr. DNA was purified using the PowerSoil™ DNA Isolation Kit (MO BIO Laboratories, Inc., Carlsbad) according to the manufacturer’s instructions. A DNA library for Illumina genomic DNA sequencing was prepared from 4μg DNA using a Genomic Sample Prep Kit (FC-102-1001, Illumina Inc., Cambridge, UK) according to the manufacturer’s instructions. The final DNA fragment library was kept at -20°C until used.

Genomic Sequencing

The flow cell was prepared using a Standard Cluster Generation Kit (FC-103-1001, Illumina Inc., Cambridge, UK) according to the manufacturer’s instructions. Sequencing was performed on a 1G Genome Analyser using a 36-Cycle Illumina Sequencing Kit (FC-104-1003, Illumina Inc., Cambridge, UK). 4pM of the library was used to achieve c. 25,000–34,000 clusters per tile.

Solexa data were assembled using VELVET 0.7.26 (Zerbino & Birney 2008), EDENA 2.1.1 (Hernandez et al. 2008) and the CLC Genomics Workbench 3.0.1 respectively.

Gene Cluster Assembly

The published sequences of P. fluorescens Pf-5 (Paulsen et al. 2005) (Refseq: NC_004129) and P. fluorescens PIO-1 (Refseq: NC_007492) was downloaded from the NCBI FTP site (ftp://ftp.ncbi.nih.gov/ genomes/Bacteria). Coding gene sequences located in the same gene cluster as the coding sequences of the ortholog to the ALF precursor protein was blasted against each of the three de novo assemblies using TBLASTX. Contigs having a significant hit (E-value < 1E-10) was aligned. Sequence coverage was calculated by reference genome assembly of reads to the aligned sequence using the CLC Genomics Workbench 3.0.1. Promoter regions was predicted using BPROM (http://linux1.softberry.com/berry.phtml)

ACKNOWLEDGEMENTS

This work is supported by a grant from the Villum Kann Rasmussen Foundation to M.S.D. and D.E.O. J.J.E. and D.E.O. are supported by the Danish Research Foundation (inSPI N). We are grateful to Dr. Soren V. Hoffmann for access to synchrotron radiation circular dichroism facilities.

REFERENCES


34. White, A.P., et al., Extracellular polysaccharides associated with thin


Figure 1: TEM performed on negatively stained cells of (A) *Pseudomonas* strain UK4 and (B) curli producing *E. coli* SM2258. ALF like material produced by *Pseudomonas* strain UK4 is indicated with an arrow.

Figure 2: SDS-PAGE of purified ALF with and without formic acid treatment. Proteins were visualized by Coomassie staining. The molecular mass markers are noted on the left (in kDa).
Figure 3: TEM of (A) ALF purified from *Pseudomonas* strain UK4 and (B) curli purified from *E. coli*.

Figure 4: Biophysical characterization of ALF purified from *Pseudomonas* strain UK4. (A) Far-UV SRCD spectra of *Pseudomonas* ALF (solid line) and curli fimbriae (dashed line). (B) FTIR spectra of *Pseudomonas* ALF (solid line) and curli fimbriae (dashed line) in dry, solid state. (C) Fluorescence emission of 40μM ThT when mixed with 0.25mg/mL native lysozyme (solid gray line), purified *Pseudomonas* ALF (solid black line) or curli fimbriae (dashed line) when excited at 450nm. (D) Fibrillation of 2mg/mL formic acid depolymerized *Pseudomonas* ALF (black line) in 10mM Tris-HCl, pH 8.0, measured by ThT fluorescence. Negative control (gray line) was in the absence of ALF.
Figure 5: Structure of FapC and genomic organization of the fap genes. (A) Schematic representation of FapC. The protein is composed of an N-terminal signal sequence (Sig) which is cleaved off during the translocation of the outer membrane, an N-terminal domain (Nterm), three repeats (R1-3) separated by two linker regions (L1-2) and a C-terminal tail (Cterm). (B) Repeated units in FapC. Conserved residues are highlighted in bold. (C) Organization of the fap operon. Promoter region containing -35, and -10 elements is indicated as a solid box, and protein coding sequences are indicated by open boxes. Arrow indicates transcription start site and solid dots indicate potential ribosomal binding sites (Shine-Dalgarno sequences).

Supplementary Figure 1: Identification of FapC. Sequences obtained by MS/MS (different hues of blue) and N-terminal sequencing (red) are shown on top of the translated fapC gene.
Supplementary Figure 2: Comparison of repeats in FapC and its orthologs in other *Pseudomonas* strains. For each protein the consensus sequence of the repeat is given, based on internal comparison of the three repeat sequences within each protein. Residues which are highly conserved among the orthologs are highlighted in color.

<table>
<thead>
<tr>
<th>Strain/Protein</th>
<th>Promoter</th>
<th>FapA</th>
<th>FapB</th>
<th>FapC</th>
<th>FapD</th>
<th>FapE</th>
<th>FapF</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>P. fluorescens</em> Pf-5 (PFL 3118-3123)</td>
<td>74%</td>
<td>40%</td>
<td>59%</td>
<td>38%</td>
<td>77%</td>
<td>62%</td>
<td>81%</td>
</tr>
<tr>
<td><em>P. fluorescens</em> PfO-1 (Pf01 2598-2593)</td>
<td>70%</td>
<td>39%</td>
<td>62%</td>
<td>36%</td>
<td>77%</td>
<td>61%</td>
<td>81%</td>
</tr>
<tr>
<td><em>P. entomophila</em> L48 (PSEEN 2939-2944)</td>
<td>69%</td>
<td>39%</td>
<td>56%</td>
<td>37%</td>
<td>72%</td>
<td>53%</td>
<td>73%</td>
</tr>
<tr>
<td><em>P. putida</em> F1 (Pput 2831-2836)</td>
<td>69%</td>
<td>38%</td>
<td>54%</td>
<td>35%</td>
<td>72%</td>
<td>53%</td>
<td>73%</td>
</tr>
<tr>
<td><em>P. putida</em> KT2440 (PP 2858-2853)</td>
<td>67%</td>
<td>38%</td>
<td>55%</td>
<td>35%</td>
<td>72%</td>
<td>53%</td>
<td>73%</td>
</tr>
<tr>
<td><em>P. putida</em> W619 (PputW619 2762-2767)</td>
<td>67%</td>
<td>40%</td>
<td>52%</td>
<td>36%</td>
<td>73%</td>
<td>54%</td>
<td>72%</td>
</tr>
<tr>
<td><em>P. aeruginosa</em> LESB5 (PALES 33671-33721)</td>
<td>67%</td>
<td>38%</td>
<td>47%</td>
<td>37%</td>
<td>63%</td>
<td>39%</td>
<td>65%</td>
</tr>
<tr>
<td><em>P. aeruginosa</em> PAO1 (PA 1956-1951)</td>
<td>68%</td>
<td>37%</td>
<td>47%</td>
<td>40%</td>
<td>63%</td>
<td>39%</td>
<td>65%</td>
</tr>
<tr>
<td><em>P. aeruginosa</em> PA14 (PA14 39220-39270)</td>
<td>68%</td>
<td>37%</td>
<td>47%</td>
<td>37%</td>
<td>63%</td>
<td>39%</td>
<td>65%</td>
</tr>
</tbody>
</table>

**Supplementary Table 1:** Protein identity between FapA-FapF and their orthologs in other *Pseudomonas* strains. 1Identity of the promoter region is given on DNA level.
<table>
<thead>
<tr>
<th>Strain/Protein</th>
<th>Signal Seq.</th>
<th>N-term</th>
<th>Repeats</th>
<th>Linker1</th>
<th>Linker2</th>
<th>C-term</th>
</tr>
</thead>
<tbody>
<tr>
<td>P. strain UK4</td>
<td>24</td>
<td>37</td>
<td>30</td>
<td>35</td>
<td>43</td>
<td>21</td>
</tr>
<tr>
<td>P. fluorescens PfO-1 (Pf01 2596)</td>
<td>24</td>
<td>54</td>
<td>35</td>
<td>32</td>
<td>107</td>
<td>7</td>
</tr>
<tr>
<td>P. entomophysila L48 (PSEEN 2941)</td>
<td>24</td>
<td>45</td>
<td>36</td>
<td>38</td>
<td>276</td>
<td>3</td>
</tr>
<tr>
<td>P. putida F1 (Pput 2833)</td>
<td>24</td>
<td>51</td>
<td>36</td>
<td>37</td>
<td>260</td>
<td>3</td>
</tr>
<tr>
<td>P. putida KT2440 (PP 2856)</td>
<td>24</td>
<td>51</td>
<td>36</td>
<td>37</td>
<td>260</td>
<td>5</td>
</tr>
<tr>
<td>P. putida W619 (PputW619 2764)</td>
<td>24</td>
<td>46</td>
<td>38</td>
<td>36</td>
<td>274</td>
<td>1</td>
</tr>
<tr>
<td>P. aeruginosa LESB5 (PALES 33691)</td>
<td>24</td>
<td>42</td>
<td>30</td>
<td>37</td>
<td>127</td>
<td>20</td>
</tr>
<tr>
<td>P. aeruginosa PAO1 (PA 1954)</td>
<td>24</td>
<td>42</td>
<td>35</td>
<td>32</td>
<td>122</td>
<td>15</td>
</tr>
<tr>
<td>P. aeruginosa PA14 (PA14 39240)</td>
<td>24</td>
<td>42</td>
<td>35</td>
<td>32</td>
<td>122</td>
<td>15</td>
</tr>
<tr>
<td>P. fluorescens Pf-5 (PFL 3120)</td>
<td>24</td>
<td>55</td>
<td>35</td>
<td>32</td>
<td>106</td>
<td>7</td>
</tr>
</tbody>
</table>

**Supplementary Table 2:** Length of different motifs in FapC and its orthologs.
<table>
<thead>
<tr>
<th>Strain/Protein</th>
<th>FapA</th>
<th>FapB</th>
<th>FapC</th>
<th>FapD</th>
<th>FapE</th>
<th>FapF</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>P. Strain UK4</em> (FapA-F)</td>
<td>EX</td>
<td>PP</td>
<td>OM</td>
<td>CP</td>
<td>EX</td>
<td>PP</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>53.10%</td>
<td>53.10%</td>
<td>53.10%</td>
<td>53.10%</td>
<td>53.10%</td>
<td>53.10%</td>
</tr>
<tr>
<td><em>P. fluorescens Pf-5</em> (PFL 3118-3123)</td>
<td>PP</td>
<td>EX</td>
<td>EX</td>
<td>IM</td>
<td>PP</td>
<td>PP</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>68.10%</td>
<td>53.10%</td>
<td>53.10%</td>
<td>53.10%</td>
<td>53.10%</td>
<td>53.10%</td>
</tr>
<tr>
<td><em>P. fluorescens PfO-1</em> (Pf01 2598-2593)</td>
<td>EX</td>
<td>EX</td>
<td>OM</td>
<td>CP</td>
<td>PP</td>
<td>PP</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>53.10%</td>
<td>53.10%</td>
<td>53.10%</td>
<td>68.10%</td>
<td>53.10%</td>
<td>53.10%</td>
</tr>
<tr>
<td><em>P. entomophila L48</em> (PSEEN 2939-2944)</td>
<td>PP</td>
<td>OM</td>
<td>OM</td>
<td>IM</td>
<td>EX</td>
<td>PP</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>53.10%</td>
<td>53.10%</td>
<td>53.10%</td>
<td>71.10%</td>
<td>71.10%</td>
<td>53.10%</td>
</tr>
<tr>
<td><em>P. putida F1</em> (Pput 2831-2836)</td>
<td>PP</td>
<td>OM</td>
<td>PP</td>
<td>IM</td>
<td>PP</td>
<td>PP</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>68.10%</td>
<td>53.10%</td>
<td>53.10%</td>
<td>71.10%</td>
<td>53.10%</td>
<td>53.10%</td>
</tr>
<tr>
<td><em>P. putida KT2440</em> (PP 2858-2853)</td>
<td>PP</td>
<td>OM</td>
<td>PP</td>
<td>IM</td>
<td>PP</td>
<td>EX</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>53.10%</td>
<td>53.10%</td>
<td>53.10%</td>
<td>68.10%</td>
<td>68.10%</td>
<td>53.10%</td>
</tr>
<tr>
<td><em>P. putida W619</em> (PputW619 2762-2767)</td>
<td>PP</td>
<td>PP</td>
<td>OM</td>
<td>IM</td>
<td>PP</td>
<td>PP</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>53.10%</td>
<td>53.10%</td>
<td>53.10%</td>
<td>68.10%</td>
<td>53.10%</td>
<td>53.10%</td>
</tr>
<tr>
<td><em>P. aeruginosa LESB5</em> (PALES 33671-33721)</td>
<td>EX</td>
<td>PP</td>
<td>EX</td>
<td>CP</td>
<td>PP</td>
<td>PP</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>68.10%</td>
<td>53.10%</td>
<td>53.10%</td>
<td>68.10%</td>
<td>71.10%</td>
<td>68.10%</td>
</tr>
<tr>
<td><em>P. aeruginosa PAO1</em> (PA 1956-1951)</td>
<td>EX</td>
<td>PP</td>
<td>EX</td>
<td>CP</td>
<td>PP</td>
<td>PP</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>68.10%</td>
<td>53.10%</td>
<td>53.10%</td>
<td>68.10%</td>
<td>71.10%</td>
<td>68.10%</td>
</tr>
<tr>
<td><em>P. aeruginosa PA14</em> (PA14 39220-39270)</td>
<td>EX</td>
<td>PP</td>
<td>EX</td>
<td>CP</td>
<td>PP</td>
<td>PP</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>53.10%</td>
<td>53.10%</td>
<td>53.10%</td>
<td>68.10%</td>
<td>71.10%</td>
<td>68.10%</td>
</tr>
<tr>
<td>Consensus</td>
<td>PP/EX</td>
<td>Not CP</td>
<td>Not CP</td>
<td>CP/IM</td>
<td>PP/EX</td>
<td>PP</td>
</tr>
<tr>
<td></td>
<td>Weak</td>
<td>Bad</td>
<td>Bad</td>
<td>Weak</td>
<td>Bad</td>
<td></td>
</tr>
</tbody>
</table>

**Supplementary Table 3:** Subcellular localization prediction of FapA-F and their homologs in other *Pseudomonas* strains using the PSLpred method (Bhasin, M., Garg, A., & Raghava, G. P. (2005) *Bioinformatics* (Oxford, England) 21, 2522-2524.). Predicted localization (CP=cytoplasm, IM=inner-membrane, OM=outer-membrane, PP=Periplasma, and EX=extra cellular), Reliability Index (1-5, higher number is better), and accuracy.
2.4 Paper 4


With reference to Ministerial Order no. 114 of 8 March 2002 regarding the PhD Degree § 14, article 3, statements from each author about the PhD student’s part in the shared work must be included in case the thesis is based on already published articles.

Co-author statement concerning the contribution by Morten S. Dueholm

- M. S. D designed the research supervised by D. E. O.
- M. S. D. performed almost all experimental work, including:
  - Development of the CsgA purification protocol.
  - Optimization and running of the fibrillation assay.
  - Biophysical characterization of the fibrils, including:
    - Far-UV CD, FTIR, ThT fluorescence
    - Sample preparation for TEM (Images were collected by G. C.)
    - Sample preparation for X-ray fiber diffraction (Spectra were collected by K. L. H).
- M. S. D. performed all data analysis supervised by D. E. O.
- M. S. D. wrote the paper supervised by D. E. O.

Morten S. Dueholm ___________________________ 14/09-2009
Kim L. Hein ___________________________ 14/09-2009
Gunna Christiansen ___________________________ 11/09-2009
Daniel E. Otzen ___________________________ 15/09-2009

Signature ___________________________ Date
2.4.1 Motivation
The effect of environmental conditions on amyloid formation has been studied for many proteins in vitro [11, 41, 137, 138]. These studies showed that changing environmental conditions often result in amyloid fibril polymorphism. Functional amyloids are in contrast to the disease related amyloid proteins selected due to their ability to form well-defined amyloid structures. We therefore hypothesized that these amyloidogenic proteins are more robust toward changing environmental conditions. To test this hypothesis we decided to examine the effect of changing environmental conditions on the fibrillation of CsgA, the major subunit of the E. coli curli.

2.4.2 Outcome
The first problem, we needed to solve, was how to prepare a stock solution of pure monomer CsgA. This was not trivial due to the highly aggregative nature of the protein. We developed a protocol in which native curli fimbriae first was purified using the method developed in paper 3. These fimbriae were then depolymerized in concentrated formic acid. Lyophilization was performed to remove the formic acid. The proteins were then dissolved in 8M urea and subjected to anion exchange chromatography. Fractions containing CsgA was collected and dialyzed against water. This resulted in aggregation of the CsgA. Another step of formic acid treatment was performed and the protein dissolved in 6M GdmCl. A gel filtration step in the presence of 6M GdmCl was performed to remove remaining aggregated material. This left a stable solution of CsgA monomers, which could be desalted on a column right before use.

We then investigated the fibrillation of CsgA at pH 3-9 and at NaCl concentrations from 0-500mM using a plate reader setup. CsgA was found to fibrillate at all conditions tested. To gain further insight into the fibrillation kinetics we fitted the fibrillation data to the Finke-Watzky minimalistic 2-step model for nucletion followed by autocatalytic surface growth [139, 140]. Using this model, we were able to determine rate constants for overall nucletion and growth. No direct correlation was found between the rate constants and the pH. Both the rate constant for nucletion and growth increased with increasing NaCl concentrations in a manner suggesting a simple electrostatic screening effect on aggregation. The effect of protein concentration and seeding with native curli seed was also investigated.

A thorough biophysical characterization comprising TEM, CD, FTIR, and X-ray fiber diffraction was performed on the fibrils to examine if the fibrils were similar. No differences could be observed, suggesting that the fibrils are similar.

Additional information on the fibrillation process was obtained by examining samples collected at various time points during the fibrillation with TEM, CD, SDS-PAGE, and ANS fluorescence. The latter probes the formation of hydrophobic material. It was found that the fibrillation was initiated by a conversion of the monomers to thin needle-like protofibrils with no apparent structure. These protofibrils later converted into hydrophobic amyloid fibrils which clumped together in larger aggregates.
ROBUST FIBRILLATION OF THE MAJOR CURLI SUBUNIT CSGA UNDER DIFFERENT SOLVENT CONDITIONS

Morten S. Dueholm*a, Kim L. Heinb, Gunna Christiansend, Daniel E. Otzena,1.

aCentre for Insoluble Protein Structures, Interdisciplinary Nanoscience Center (iNANO), Department of Molecular Biology University of Aarhus (iNANO), 8000 Aarhus C, Denmark; bDepartment of Biotechnology, Chemistry, and Environmental Engineering, Aalborg University, 9000 Aalborg, Denmark; cDepartment of Molecular Biology, University of Aarhus; dDepartment of Medical Microbiology, University of Aarhus, 8000 Aarhus C, Denmark.

Author Contributions: M.S.D. and D.E.O. designed research, M.S.D., K.L.H and G.C. performed research, M.S.D. and D.E.O. analyzed data, M.S.D. wrote the paper supervised by D.E.O.

The authors declare no conflict of interest

1To whom correspondence should be addressed: Telephone: +45 89425046; Fax: +45 86123178; E-mail: dao@inano.dk

Abstract: The amyloid fold is usually considered a result of protein misfolding. However, recently a number of studies have shown that this is only part of the story, and that the amyloid structure is also used in nature for functional purposes. CsgA is the major subunit of E. coli curli, the best studied functional amyloid. Here we show that in vitro fibrillation of CsgA is very robust to changing environmental conditions and that the fibrils formed have similar structure. This is consistent with evolutionarily optimized fibrillation properties and a clear contrast to what is seen for the disease related amyloid precursors. Furthermore we show that the formation of CsgA amyloid is preceded by the formation of thin needle-like protofibrils with no apparent secondary structure and followed by aggregation of the amyloid fibrils. The aggregative nature of the curli fibrils might be useful in curli mediated biofilms formation.

INTRODUCTION

Amyloids are commonly associated with devastating diseases such as Parkinson’s and Alzheimer’s disease, type II diabetes and transmissible spongiform encephalopathies such as Creutzfeldt-Jacob disease [1-4]. Amyloid formation is the result of misfolding followed by spontaneous polymerization of otherwise soluble protein into biochemically and structurally similar fibrils with a characteristic cross-β structure in which the β-strands are arranged perpendicular to the fiber axis [5-9]. When examined by TEM or AFM amyloid fibrils are typically found to be composed of intertwined filaments. The fibrils are of indeterminate length and have a width in the order of 7-20nm, varying with the fibril morphology [4, 6].

The fact that many structurally unrelated proteins can form amyloid fibrils indicates that the general amyloid fold is governed by protein backbone interactions, but stabilized by side chain interactions [6, 10-12]. Due to variations in side chain interactions, the amyloid structure does not necessarily represent a global energetic minimum in the folding landscape as is the case for the native structure of simple globular proteins, such as RNAase [13, 14]. Variations in the molecular structures of amyloid fibrils originating from a single protein are thought to be responsible for the existence of multiple strains of mammalian prion diseases and yeast prion phenotypes [15-20]. Furthermore they may produce variations in the toxicity or patterns of depositions of amyloid fibrils in Alzheimer’s and other amyloid diseases [21, 22].

Despite its common association with disease, the amyloid structure is also used for beneficial purposes in nature. The curli system in Escherichia coli is one of the best studied systems of functional amyloid. Curli fimbriae are part of the complex extracellular matrix that is essential for biofilm formation, host cell adhesion and invasion, and they are reported to be important stimulants of the host inflammatory response [23-29].
A highly regulated pathway involving two divergently expressed operons is required for the curli biogenesis. The csgBA operon encodes the curli subunits and csgDEFG codes for proteins involved in the regulation and transportation [26, 30-32]. The major curli subunit, CsgA, is secreted to the cell surface as a soluble unstructured protein [33, 34]. Here it interacts with the membrane bound CsgB, which nucleates the fibrillation of CsgA. After this initial step, fibrils are predicted to grow by subsequent addition of CsgA to the tip of the amyloid fiber [32, 34].

CsgA and CsgB both contain five imperfect repeats with highly conserved glutamine and asparagine residues that are important for the amyloid formation (Figure 1) [35]. Each repeating unit is predicted to form a strand-loop-strand motif that closely resembles the cross-β fold described for many disease related amyloids [8, 36]. It seems reasonable to expect that the presence of specific side chain interactions in and between curli subunits, aided by the presence of five repeat units in one protein sequence, guide the fibrillation of CsgA to a single well-defined amyloid form. We test this expectation by checking whether the CsgA amyloid structure is sensitive to changing environmental conditions in vitro. Using a broad range of biophysical tools we compare fibrils formed in vitro with native curli fibrils. Finally we investigate what happens structurally during the fibrillation processes. We conclude that the CsgA fibrils are very robust to changing environmental conditions which are consistent with evolutionarily optimized fibrillation properties.

RESULTS

Effect of pH and Ionic Strength

The ability of CsgA to form amyloid fibrils was investigated under a wide range of environmental conditions using ThT fluorescence as a probe for amyloid formation. For these experiments, it is very important to minimize the amount of CsgA that is not in the monomeric state, otherwise artifacts may arise from seeding of the aggregation process by preaggregated CsgA. Accordingly we developed a procedure in which CsgA was initially purified from E. coli curli in buffer containing 8M urea, dialyzed against water, lyophilized, resuspended in formic acid and subsequently purified by gel filtration in a buffer containing 6M GdmCl. The CsgA monomer fraction was then rapidly transferred to water on a desalting column immediately prior to the aggregation assay.

CsgA clearly showed fibrillation at all pH values tested (Figure 2A) with a sigmoidal time profile, in which a lag time of an hour or less gave rise to a rapid rise in fluorescence which typically stabilized at a plateau level after about 4 hrs. Higher end ThT values were observed at lower pH values. The effect of ionic strength on the fibrillation of CsgA was examined at pH 7 by varying the NaCl concentrations between 0 and 500 mM (Figure 2C). CsgA fibrillation was evident at all NaCl concentrations and the end ThT values increased with increasing ion strength.

The fluorescence emission intensities of the fibrillation curves were converted to concentration of fibrillated CsgA (measured in terms of fibrillated monomers), assuming proportionality between this form and the ThT fluorescence intensity and assuming that all monomeric CsgA is incorporated into fibrils. These assumptions are later shown to be true. The converted fibrillation curves were then fitted to the Finke-Watzky minimalistic 2-step model for nucleation followed by autocatalytic surface growth [37, 38]. Using this model, we are able to determine rate constants $k_1$ and $k_2$ for overall nucleation and growth, respectively (Figure 2B and 2D). No direct correlation was found between the nucleation rate and the pH. The nucleation rate was highest at pH 6 and 7 and lowest at pH 9. The growth rate was relatively constant, except for pH 3, where it was considerably higher. Both the rate constant for nucleation and growth increased with increasing NaCl concentrations. When the rate constants were plotted against the square root of the ion strength a linear relationship was found, indicating a simple electrostatic screening effect on aggregation (Insert to Fig. 2D).

Effect of Protein Concentration

The effect of protein concentration on the kinetics of fibrillation was investigated (Figure 3A). A linear relation was found between the protein concentrations and end ThT intensities (Insert to Fig. 3B). Remarkably, it was found that the rate constant for nucleation increased with increasing protein concentration, whereas the rate constant for growth decreased (Figure 3B). The combination of these two contrasting changes means that the ThT plateau is reached after around 6 hrs at all concentrations.

Effect of Seeding

The effect of native curli fibril seeds on the fibrillation kinetics (Figure 3C) was evaluated by fitting normalized fibrillation curves to a modified version of the Finke-Watzky model, in which the start concentration of protein on the fibril form was assumed not to be zero (Figure 3D) [37]. The rate constant for nucleation was found to increase linearly with increasing seed concentration, this shows that the added seed are more effective in seeding than the fibrils formed. In contrast, the rate constant for growth was fairly constant at lower seed concentration (≤2%), but increased slightly at higher seed concentrations. Even the lowest seed concentration (≤2%)
concentration (0.1%) produced a significant effect on the fibrillation, showing the stock solution contained no or very few seeds.

Fibril morphology

The morphology of fibrils formed at various pH values and NaCl concentrations was examined by TEM (Figure 4). All samples showed similar needle-like fibrils which aggregated into large clumps. Amorphous aggregates were also present in all samples.

Secondary Structure Analysis

The secondary structure of the fibrils was analyzed by CD and FTIR (Figure 5). All CD spectra showed similar strong β-sheet signatures with the characteristic minimum at ~217nm and maximum at ~197nm (Figure 5A-B). FTIR also showed very similar spectra for the fibrils with a well defined peak at 1625cm⁻¹, indicative of β-sheet in an amyloid-like conformation maintained by very strong hydrogen bonds and a minor shoulder at 1663 cm⁻¹ which indicates β-turns (Figure 5C-D) [39-41].

X-ray Fiber Diffraction

The structure of fibril formed at various pH values was examined by X-ray fiber diffraction (Figure 5E-F). For all samples a characteristic Bragg reflection at 4.7-4.8Å was seen, corresponding to the spacing between β-stands in the amyloid fold [6]. The broad reflection at 8-10Å, corresponding to inter-sheet spacing [6, 42], was predominantly seen at lower pH values, rather than neutral-alkaline conditions, and was not seen for the native curli either. Additional well-defined reflections were seen for all samples at 6.3Å, 5.3Å, and 4.2Å. Currently we do not have any assignments for these reflections. Due to the highly aggregative nature of the samples the fibrils did not align very well, and we therefore lack information on the orientation of the β-stands and β-sheets with respect to the fibril axis.

Fibrillation of CsgA followed by TEM

In order to gain additional insight into the fibrillation of CsgA, we examined samples collected during the fibrillation process by TEM (Figure 6). Right after mixing the monomeric CsgA with buffer a few short, thin filaments were seen in the TEM images. After 1hr larger fibrils were observed in co-existence with many thin filaments. At 2hr the fibrils appeared to have grown even larger at the expense of the small filaments. At 4hr the fibrils started to aggregate and form larger clumps. A few thin filaments were still present. These filaments were completely absent after 6hr. Some amorphous aggregates were seen in all samples except for the 0hr sample.
The rates for nucleation and growth both increased with the ionic strength. This increase is suspected to be due to screening of charged groups in CsgA.

Fibril morphology, X-ray fiber diffraction patterns and secondary structure analysis all indicated a similar amyloid structure. This is in agreement with our hypothesis that CsgA forms a well-defined amyloid structure due to the stabilizing effect of the repeating units and side chain interactions.

Detailed information on the CsgA fibrillation process was obtained by analyzing samples collected at various times during the fibrillation at pH 7. The samples were analyzed with TEM, ANS binding, CD, and SDS-PAGE. It was found that the fibrillation was initiated by a conversion of the monomers to thin needle-like protofibrils with no apparent structure (Figure 9). These protofibrils later converted into hydrophobic amyloid fibrils which clumped together in larger aggregates. The highly aggregative nature of the CsgA fibrils might be useful in curli mediated E. coli and Salmonella biofilm formation.

**MATERIALS AND METHODS**

The *E. coli* MG1665 mutant SM2258 [43], used for curli production, was kindly provided by Professor Søren Molin (BioCentrum-DTU, Denmark).

**Purification of CsgA**

Native curli fimbriae were purified as previous described (Paper 3). Curli was lyophilized, dissolved in 98% formic acid and lyophilized again. The depolymerized curli was then dissolved in AIEX binding buffer (8M urea, 1mM DTT, 20mM bis-tris, pH 6.5) and filtered through a 0.22µm filter. The sample was applied for ion exchange on a 5mL HiTrap Q HP column equilibrated with AIEX binding buffer using a flow of 5ml/min and eluted by stepwise increasing the concentration of NaCl (0, 100, 200, 500 and 1000mM). Fractions containing CsgA were pooled and dialyzed against deionized water for 2 days. This resulted in precipitation of CsgA. The precipitate was collected and resuspended in 98% formic acid lyophilized. The lyophilized material was suspended in GdmCl buffer (6M GdmCl, 20mM NaPO₄, pH 7.0). CsgA monomers were isolated from aggregated material and dimers by gel filtration in GdmCl buffer using a superfine 200 10/300 column (GE Healthcare). CsgA was desalted into deionized water using a 5mL HiTrap Desalting column right before the protein was used for fibrillation studies. This thorough procedure was required in order to obtain reproducible fibrillation curves in the fibrillation assay.

**Fibrillation Assay**

Gel-filtered CsgA in 6M GdmCl was desalted using a HiTrap desalting column equilibrated with deionized water. Protein concentrations were estimated by UV-absorbance of the peptide bond using the empirical formula by Waddell [44]. The protein was diluted in deionized water to 2x the final concentration and mixed with a premade 2x cocktail of buffer, NaCl or seeds, and 80µM ThT. 200µL samples was loading in a 96-well plate. Immediately afterward, the plate was transferred to a Tecan GENios Pro plate reader and ThT fluorescence was measured using, excitation at 448nm, emission at 485nm and a gain of 60. Measurements were obtained by bottom reads every 2min and 30sec of shaking (orbital, amplitude 2.5mm) was applied between the reads. Reads was integrated for 40µs. Six fibrillation curves were averaged to reduce the signal to noise. For seed preparation 5mL of a 1mg/mL suspension of native curli fibrils was sonicated using a rod sonicator at maximum effect for 30s.

**Circular Dichroism**

CD spectra from 250-190nm were collected on a Jasco J-810 spectropolarimeter using 0.2nm steps, scan speed of 50nm/min, bandwidth of 3nm and a response time of 2sec. A light path of 0.1mm was used and the temperature was kept constant at 20°C with a thermostatically controlled cell holder. The protein concentration was 0.2mg/mL and all spectra were baseline corrected with respect to buffer. To improve the signal-to-noise ratio five scans were averaged on each sample. The results were expressed as mean residue ellipticity (MRE). All samples except the ones collect during the fibrillated were sonicated briefly in order to minimize fibril precipitation.

**Fourier-transform Infrared Spectroscopy**

FTIR was carried out using a Tensor 27 (Bruker) FTIR spectrophotometer equipped with a DTGS Mid Infrared detector and a Golden Gate single reflection diamond Attenuated Total Reflectance (ATR) cell (Specac). Approximately 4µg of protein was dried on the ATR crystal using dry nitrogen. Spectra were recorded from 4000-1000cm⁻¹ using a nominal resolution of 2cm⁻¹ and 64 accumulations. Fourier self-deconvolution of the spectra in the amide I region was performed using a built-in apodization function with a deconvolution factor of 2, a noise reduction factor of 0.5, and Lorentzian lineshape in the OPUS 5.5 system (Bruker). Identification of the different components of the amide I after Fourier self-deconvolution was performed by a Levenberg-Marquardt fit function or by second derivative analysis in the OPUS 5.5 system.
Transmission Electron Microscopy

10-20µL of a fibril solution was applied to the nickel grid for 30sec. The grids were washed on one drop of glass distilled water and stained with 3 drops of 1% phosphotungstic acid pH 6.9 and blotted dry. Electron microscopy was done using a JEOL 1010 TEM at 60 kW. Images were taken using a Sony XCD-SX900 camera. For size determination, a standard grid-size nickel plate (2160 lines/mm) was used [45].

X-ray Fiber Diffraction

Fiber diffraction specimens were prepared on a stretch frame using a suspension of approximately 5mg/mL fibrillated CsgA, and dried at room temperature. Data were collected in-house using a Cu Kα rotating-anode source (FR591 Enraf Nonius, Delft, Holland; wavelength 1.5418 Å), equipped with a MAR Research 345 image plate X-ray detector (345 mm diameter). The sample to detector distance was 400 mm with exposure times of 20-60 min. The images were evaluated using Fibre Fix for Windows Version 1.3.1. Bragg distances were measured as the peak maximum from radial averaged intensity plots.

SDS-PAGE

SDS-PAGE was performed according to the method of Laemmli with modifications by Ames [46, 47]. Samples were mixed with SDS-PAGE loading buffer with DTT, boiled for 5min and analyzed on a 15% SDS-PAGE gel. The concentration of non-amyloid CsgA was estimated from the CsgA monomer band intensity using the ImageJ gel analysis tool (http://rsbweb.nih.gov/ij/).

ACKNOWLEDGEMENTS

This work is supported by a grant from the Villum Kann Rasmussen Foundation to M.S.D. and D.E.O. D.E.O. is supported by the Danish Research Foundation (inSPIN).

REFERENCES

41. Zandomeneghi, G., et al., FTIR reveals structural differences between native beta-


FIGURES AND TABLES

CsgA
R1 SELNIYQQGGSANSALALQTDARN
R2 SDLTITQGGSANDVGQGSDD-
R3 SSDLTITQGGSANDLVNGN
R4 SEMTVKQGGSGGAQVDQASN-
R5 SSVNVTQGSGNNATAHOY

CsgB
R1 QAAIIGQAGTNNSAQLRQGSK-
R2 LLAVVAQEGSSPAKIDQTGDY-
R3 NLAYIDQAGSANDASISGSAYG-
R4 NTAMIIQGSGFANITYQTQ-
R5 KTAVVVRQSQMAIRTQ

Figure 1: Repeated units in CsgA and CsgB. Conserved residues are highlighted in bold with gray background.

Figure 2: Effect of environmental conditions on the fibrillation of CsgA followed by ThT.

(A) Fibrillation of 20µM CsgA at pH 3-9. The following buffer systems were used, pH 3-5: 20mM citic acid/NaOH, pH 6: 20mM histidine/HCl, pH 7: 20mM phosphoric acid/NaOH, pH 8: 20mM Tris/HCl, pH 9: 20mM Glycine/HCl. Ionic strength was adjusted to 50mM by addition of NaCl. (B) Rate constants obtained by fitting the fibrillation data in (A) to the Finke-Watzky 2-state model. (C) Fibrillation of 20µM CsgA at pH 7 in the presence of various NaCl concentrations. (D) Rate constants obtained by fitting the fibrillation data in (C) to the Finke-Watzky 2-state model. Insert shows a linear relation between the rate constants and the square root of the NaCl concentration.
Figure 3: Effect of protein concentration and seeding on the fibrillation of CsgA followed by ThT. (A) Fibrillation of 2-40µM CsgA in 20mM sodium phosphate, pH 7. (B) Rate constants obtained by fitting the fibrillation data in (A) to the Finke-Watzky 2-state model. Insert shows a linear correlation between ThT plateau level and CsgA concentration. (C) Fibrillation of 20µM CsgA at pH 7 in the presence of various concentrations of native curli seeds. (D) Rate constants obtained by fitting the fibrillation data in (C) to a modified version of the Finke-Watzky 2-state model in which the start concentration of seeds is not assumed to be zero.

Figure 4: Morphology of the CsgA fibrils. TEM image of CsgA fibrils formed in 20mM sodium phosphate, pH 7.
Figure 5: Biophysical characterization of CsgA fibrils formed at various environmental conditions. (A and B) CD spectra of CsgA fibrils formed at various pH values (A) and at pH 7 with changing NaCl concentration (B). (C and D) FTIR spectra of CsgA fibrils formed at various pH values (C) and at pH 7 with changing NaCl concentration (D). (E) Radial averaged X-ray fiber diffraction spectra of fibrils formed at various pH values. (F) X-ray fiber diffraction pattern obtained for CsgA fibrils formed at pH 7.
Figure 6: Fibrillation of CsgA followed by TEM. Samples were collected at various time points during the fibrillation of 40µM CsgA at pH 7 and examined by TEM. The fibrillation curve measured by ThT is shown in the top.

Figure 7: Fibrillation of CsgA followed by ANS fluorescence and SDS-PAGE. (A) Fibrillation of 20µM CsgA followed by ThT and ANS fluorescence. (B) Fibrillation of 20µM CsgA followed by ThT and SDS-PAGE. SDS-PAGE data was expressed as CsgA monomer concentration calculated from the CsgA monomer band intensity using the ImageJ gel analysis tool (http://rsbweb.nih.gov/ij/).
Figure 8: Fibrillation of CsgA followed by circular dichroism. Samples were collected at various time points during the fibrillation of 20µM CsgA at pH 7 and examined by CD. The fibrillation curve measured by ThT is shown to the left and the CD spectra shown to the right.

Figure 9: CsgA fibrillation mechanism: Monomer CsgA forms thin needle-like protofibrils with no apparent well defined secondary structure. These protofibrils rearranges and forms amyloid fibrils, which later aggregates into dense clumps.
3. Discussion and Perspective

3.1 Prevalence of Functional Amyloids

In 2006, when I started my PhD, functional bacterial amyloids (FuBA) were still considered a rare phenomenon covered by two classes, the curli and Tafi fimbriae of *E. coli* and *Salmonella* sp., respectively, and the chaplins of *Streptomyces coelicolor* [81, 92, 100, 141]. This view was however soon to change. In order to study the prevalence of FuBA, a joint collaboration between the Environmental Biotechnology group, Aalborg University headed by Professor Per H. Nielsen and the Protein Biophysics group, Aalborg University headed by Daniel E. Otzen was established.

One outcome of this collaboration was a screening of FuBA in environmental biofilms using amyloid specific dyes and antibodies in concert with FISH (Paper 1 of this thesis). The work provided strong evidence for amyloid-like structures being a common phenomenon in natural biofilms. With this in mind it was hypothesized that FuBA are of great importance to floc properties in active sludge wastewater treatment plants (WWTP). To test this hypothesis, the abundance of FuBA in activated sludge flocs from different WWTP and the identity of these bacteria were investigated in a second paper [142]. It was found that expression of FuBA is also common for bacteria in active sludge. A more detailed analysis revealed that many denitrifiers and polyphosphate-accumulating organisms most likely produce FuBA, whereas nitrifiers do not. Many filamentous bacteria also expressed amyloid structures, as well as some foam-forming bacteria [142].

The work of the two papers has indicated that the expression of amyloid structures is a common feature of bacteria living in biofilms. In light of this, combined with our knowledge of the functional amyloid systems in other kingdoms, it could very well be assumed that the amyloid fold has been selected many times through evolution for a wide variety of functions. We might hence only have scratched the surface and much still remains to be discovered within the field of functional amyloids.

3.2 Localization and Function of Functional Bacterial Amyloid

The complexity of the environmental biofilms makes it difficult to study the localization and function of FuBA with respect to the individual bacteria. To gain further insight on the localization and function of the FuBA, pure culture studies were initiated. This resulted in paper 2 of this thesis. The work involved a detailed screening of mycolata spp. and other Gram-positive bacteria using TEM and fluorescence microscopy in combination with amyloid specific antibodies. Purification of FuBA was carried out for two bacteria and the amyloid nature verified using biophysical methods. It was not possible to identify the proteins involved in the amyloids by conventional MS/MS methodology due to extremely stable structure of the FuBA. However, it can be hypothesized that FuBA can play multiple roles as judged by their localization. (A) FuBA can be found as extracellular fibers that might have a central role in the development of the three-dimensional architecture of biofilms, as is the case for curli fimbriae of *E. coli* and *Salmonella* [143-145]. (B) They can be found as integral parts of the cell envelope, where they may mediate cell adhesion and provide a protective shield for the bacteria towards environmental treats. This would be analogous to the silkmoth chorion, where a lamellar ultra-structure of packed amyloid fibrils protects the developing embryo against temperature variations, mechanical stress, proteases, bacteria, etc. [126]. (C) Finally, FuBA was observed in the coat of spores. Spores covered by amyloids have also been described for various
fungi, where they are known as hydrophobins [141]. This coating facilitates the dispersal of the spores by wind and enhances their attachment to surfaces and possibly also their pathogenic properties [141]. These suggested functions remain to be tested. Isolation and sequencing of FuBA monomers from each of the three groups gives us the possibility to do mutation studies. A comparison of FuBA deletion mutants with wild type bacteria can be used to pinpoint the physiological role of the different kinds of FuBA.

3.3 Functional Bacterial Amyoids in Pseudomonas

The work of paper 3 involved a thorough characterization of FuBA from a Pseudomonas sp. We were able to purify the FuBA, verify the amyloid-like nature of it using biophysical tools, and determine the full protein sequence using MS/MS sequencing combined with full genomic sequencing. Furthermore, we were able to determine the full operon structure involved in the expression of the functional amyloid. The operon was found to be conserved in many Pseudomonas stains including the opportunistic pathogen Pseudomonas aeruginosa.

Colonization of the lungs of cystic fibrosis (CF) patients by P. aeruginosa is the principal cause of mortality in CF populations. P. aeruginosa infections generally persist despite the use of long term antibiotic therapy. This has been explained by postulating that P. aeruginosa forms an antibiotic-resistant biofilm consisting of bacterial communities embedded in an exopolysaccharide matrix [146]. The Pseudomonas FuBA are morphologically very similar to curli fimbriae. It might therefore be expected, that these FuBA like curli play an important role in biofilm formation. The newly discovered FuBA could thus be an important drug target for the treatment of P. aeruginosa infections. In order to check if FuBA are involved in pathogenesis of P. aeruginosa we are currently investigating if P. aeruginosa express FuBA during human infection using a real time PCR approach. An alternative to this strategy would be to check if FuBA isolated from P. aeruginosa are recognized by IgG antibodies contained in sera obtained from patients with P. aeruginosa infections. This method have previously been used to show that M. tuberculosis pili are expressed during human infection [95].

We still do not how the Pseudomonas FuBA system is regulated and which roles the other proteins expressed by the FuBA operon play. Mutational studies on the Pseudomonas FuBA operon could shed light on these questions. Knowledge on how FuBA systems are regulated provides fundamental information on why amyloids are formed and present ideas on how to avoid the cytotoxicity associated with fibrillation intermediated.

3.4 Repeated Sequences and Conserved Gln/Asn Residues

FapC, the major subunit of the functional amyloids of Pseudomonas discovered in paper 3, contains three repeated units with highly conserved glutamine and asparagine residues. Repeated units with conserved glutamine and asparagines residues are also seen for curli from both E. coli and Salmonella, chaplins from Streptomyces coelicolor and the harpin HpaG of Xanthomonas axonopodia pv. glycines. Similar motifs have furthermore been described for prions [147]. The repeated sequences could likely act as a motivator for the fibrillation of FuBA. It has previously been shown that the aggregation of multi-domain protein constructs of immunoglobulin is highly influenced by sequence identity of the domains, with the higher identity leading to faster fibrillations [69]. It could be suspected that the presence of multiple repeats give the protein the ability to form a fibrillation prone nucleus by itself. This eliminates the requirement of high protein concentration for fibrillation.
initiation, a valuable feature in systems where fibril precursors are excreted to the extracellular environment. Glutamine and asparagine are polar residues able to form hydrogen bonds. The presence of highly conserved glutamine and asparagines residues imply that interaction between such residues or with the protein backbone stabilize the structure of many functional amyloids. Molecular modeling of the TafI subunits, AgfA and AgfB, has indicated that these functional amyloids are stabilized by such hydrogen bonds [91, 148].

3.5 Functional Amyloid Precursor are highly Amyloidogenic
In paper 4 the effect of environmental conditions on the in vitro amyloid formation by CsgA, the major subunit of E. coli curli, was studied using a wide range of biophysical tools. CsgA was found to form amyloid-like fibrils at all conditions tested. The fibrils appeared very similar when examined by FTIR, CD, X-ray fiber diffraction and TEM. This suggests that CsgA is highly amyloidogenic and robust toward changing environmental conditions. This is in agreement with evolutionarily optimized fibrillation properties of and a clear contrast to what is seen for the disease related amyloid precursors. The highly amyloidogenic nature of functional amyloid precursors is also seen for Mα involved in melanin biogenesis, where it results in rapid amyloid formation [131]. This property might be general trait for functional amyloids used to avoid the toxicity associated with oligomeric intermediates. A well defined fibrillar structure could very well be a prerequisite for correct function, especially when the FuBA are used to mediate cell-host interactions, a role previously suggested for curli fimbriae [80, 86, 144, 149].

3.6 Molecular Structure of a Functional Amyloid
Although some studies have speculated on the molecular structure of functional amyloids, no true structural studies have been conducted. A high resolution-structure of a functional amyloid would therefore be of great scientific value. We are currently working on obtaining a high-resolution ssNMR structure of CsgA in the curli fimbriae of E. coli. This is however not a trivial task as we need copious amounts of pure labeled fimbriae and preliminary studies have shown that native curli purified using established protocols contains considerable amounts of extracellular polysaccharides. In order to avoid this pitfall we are working on producing artificial curli using non labeled curli seeds as a scaffold for the fibrillation of pure labeled CsgA, which are purified as described in paper 4. With a molecular structure of CsgA in the fibril form, we should be able to fully explain the function of the repeated sequences and the conserved residues herein. This knowledge may help us understand how fibrillation of functional amyloids is controlled.

3.7 Environmental Degradation of Functional Amyloids
Investigations into the natural degradation of functional amyloid may result in the isolation of bacteria, protein or small molecules that can be used to combat unwanted biofilms formation. This is very interesting because biofilms formation poses major problems in connection to chronic P. aeruginosa infections and contamination of medical implants [150]. We have been working on isolating bacteria able to degrade curli amyloid using radio labeled curli as a substrate. Bacteria which are able to assimilate curli can then be identified by microautoradiography combined with FISH as described by Gray et al. [151]. Positive candidates can then be used for the isolation of the proteins or small molecules involved in the degradation. Proteins and small molecules able to degrade functional amyloid may furthermore prove useful as therapeutic agent in the treatment of amyloid diseases.
3.8 Working with Functional Amyloid

Functional amyloids are due to their physiochemical properties difficult to handle using established techniques. They are not observed in SDS-PAGE without prior treatment with either concentrated formic acid or TFA. This is probably one of the major reasons why only a few numbers of functional amyloids have been purified so far. The monomers of functional amyloid are furthermore highly amyloidogenic and do consequently quickly form amyloid-like fibrils when dissolved in buffers. Purification of functional amyloid is hence commonly done by isolating the functional amyloid in its insoluble form from all other cell components, which are degraded enzymatically or dissolved in detergents. This strategy requires a detailed knowledge the source and do often result in an amyloid enriched product, rather than a pure amyloid preparation.

Many questions need to be considered when developing a purification strategy for functional amyloid. If FuBA is the target, the first question is how the bacteria should be grown. This is of great importance as the growth conditions have a major impact on the expression of FuBA [93, 152].

The second step is to decide if you want to separate the FuBA from the whole cells or you want to rupture the cells. The first is favorable as it yields a much cleaner starting point. It is however seldom a trivial task. Extracellular FuBA can sometimes be shed off the bacterial surface by vortexing the cells with glass beads as have been done for the purification of MTP [95] or by blending with and Omni-Mixer homogenizator as have been done for the purification of curli [81]. Alternatively they may be liberated from the cell surface by heating at 60°C. This has previously been done for some non amyloid fimbriae of *E. coli* [153]. If disruption of the cells is required, this can be done by methods like sonication or freeze-thaw cycles with an ethanol-dry ice mixture and a water bath at 37°C [76, 154, 155]. The last have in my studies proven to give the highest yield of curli and *Pseudomonas* FuBA. Enzymes may be added to give a better cell disruption. The most frequently used enzymes for this purpose are lysozyme, mutanolysin and lysostaphin. The latter is needed for the some *Staphylococcus* due to an altered peptidoglycan layer [156]. Addition of Triton X-100 or EDTA is often needed for efficient enzymatic lysis of Gram-negative bacteria as these compounds are able to disrupt the protective outer membrane [157].

The third step is to remove contaminants. DNA and RNA are degraded by DNAses and RNAses, respectively. Cell membranes are dissolved in detergents and non amyloid proteins are made soluble by heating in 2% SDS. Finally the insoluble FuBA are collected by centrifugation. If lipid contaminations are present these can be removed by extraction with an equal volume 2:1 chloroform:methanol as were described for the purification of MTP [95]. If contaminating protein remains these can be remove by preparative SDS-PAGE [76, 81].

A treatment with formic acid or TFA is needed to analyze the FuBA by SDS-PAGE. This is done by lyophilizing the FuBA, dissolving them in either concentrated formic acid or TFA, and finally removing the formic acid of TFA by lyophilization. The resulting monomer powder can then be dissolved in SDS-PAGE loading buffer and subjected to SDS-PAGE. The monomers can also be dissolved in denaturants such as 8M urea or 6M GdmCl. A monomer stock solution in 6M GdmCl is a good starting point for fibrillation studies. Denaturants can be removed on a desalting column just before the experiment is started as described in paper 4.

90
The insoluble nature of FuBA does also create difficulties when FuBA are examined using biophysical methods. It is often necessary to homogenize FuBA by sonication before examination by CD or use in fluorescence studies. Determination of protein concentration using UV methods may also provide problems. Assays like Bradford or BCA should therefore be considered for determination of protein concentration.
4. Conclusions

It has been suggested that studies on functional amyloid systems may provide valuable insight on fibril formation and toxicity. The goal of my PhD was to analyze functional bacterial amyloid in vitro and in vivo using a variety of biophysical tools. The following conclusions can be made according to the work presented in this thesis:

- A histological method has been developed for in situ studies of bacteria producing functional bacterial amyloid (FuBA). This method relies on CLSM using amyloid specific conformational antibodies (WO1 and WO2) in concert with either DAPI or oligonucleotide probes by FISH. Quantification of the FuBA expressing fraction of bacteria in biofilms is possible with the use of DAPI, which stains all bacteria. Identification of the bacteria phylum or genus responsible for FuBA production can done using phylum or genus specific FISH probes.

- In situ studies show that functional amyloid expression is a common trait for many bacteria living in natural biofilms, and FuBA expression is especially common for bacteria living in drinking water biofilms.

- Studies on pure culture mycolata species and other Gram-positive bacteria show that many of these are able to express FuBA. These FuBA may have multiple roles as judged by their localization:
  - They are found as extracellular fibrils. These could be structural elements in the biofilms or mediate cell-cell/host interactions, similar to the functions that have been proposed for curli and Tafi fimbriae.
  - They can be integrated in the cell envelope. Here they may provide the bacteria with protection against environmental threats, similar to the function seen for amyloids in the silk moth chorion.
  - They can coat spores. This may render the spores hydrophobic resulting in a better dispersal by air, similar to the function of the yeast hydrophobins and chaplins of Streptomyces coelicolor.

- Many Pseudomonas strain, including the opportunistic pathogen Pseudomonas aeruginosa, contain an operon coding for the expression of amyloid-like fimbriae. This operon does not resemble any previously described operon for FuBA expression.

- FuBA from Pseudomonas show many similarities with the curli amyloid fimbriae of E. coli, including similar fibril morphology, tinctorial properties, and secondary structure.

- There is no sequence homology between FapC, the major subunit of the Pseudomonas FuBA, and CsgA of the curli fimbriae. However, some common features are seen, including the presence of repeated motifs with highly conserved glutamine and asparagine residues.

- Fibrillation of the functional amyloid protein CsgA is very robust toward changing environmental conditions. Fibril formation is evident from pH3-9 and at NaCl concentrations from 0-500mM.
• The fibrils formed at various conditions look very similar when examined with TEM, CD, FTIR, and X-ray fiber diffraction. This indicates that a well-defined fibril structure stabilized side chain interaction is formed. This fibril structure is likely to represent a global energetic minimum in the folding landscape.

• The fibrillation of CsgA is *in vitro* initiated by a conversion of the monomers to thin needle-like protofibrils with no apparent structure. These protofibrils later convert into hydrophobic amyloid fibrils which clumped together in larger aggregates.
5. References


